

New tools for high performance LIBS Imaging

Tomás Lopes

Mestrado em Engenharia Física

Departamento de Física e Astronomia

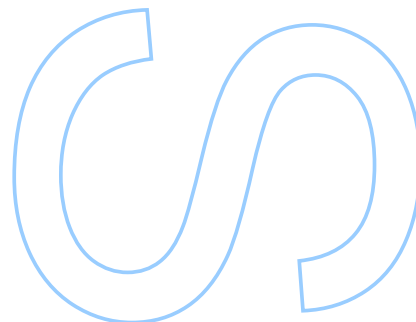
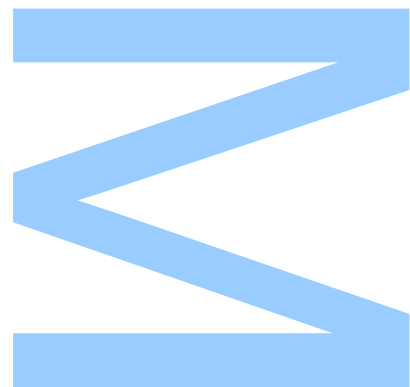
2023

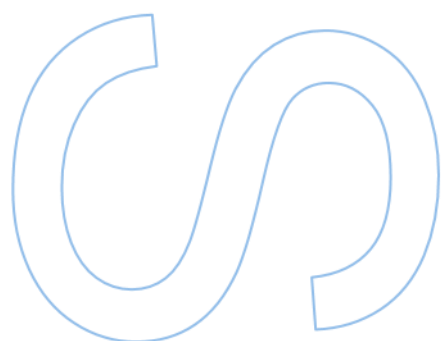
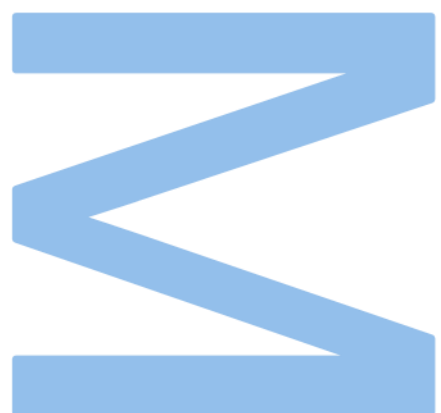
Orientador

Dr. Nuno A. Silva, Faculdade de Ciências

Coorientador

Dr. Pedro Jorge, Faculdade de Ciências





“ You’re going to make a difference. A lot of times it won’t be huge, it won’t be visible even. But it will matter just the same. ”

James Gordon - Fictional Character

Declaração de Honra

Eu, Tomás Lopes, inscrito no Mestrado em Engenharia Física da Faculdade de Ciências da Universidade do Porto declaro, nos termos do disposto na alínea a) do artigo 14.º do Código Ético de Conduta Académica da U.Porto, que o conteúdo da presente dissertação de projeto reflete as perspectivas, o trabalho de investigação e as minhas interpretações no momento da sua entrega.

Ao entregar esta dissertação de projeto, declaro, ainda, que a mesma é resultado do meu próprio trabalho de investigação e contém contributos que não foram utilizados previamente noutros trabalhos apresentados a esta ou outra instituição.

Mais declaro que todas as referências a outros autores respeitam escrupulosamente as regras da atribuição, encontrando-se devidamente citadas no corpo do texto e identificadas na secção de referências bibliográficas. Não são divulgados na presente dissertação de projeto quaisquer conteúdos cuja reprodução esteja vedada por direitos de autor.

Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Tomás Lopes

Porto, 29 de Junho de 2023

Agradecimentos

Ao terminar este ciclo de um ano de trabalho, gostava de aproveitar a ocasião para agradecer a todas as pessoas que contribuíram para esta tese. Como tudo na vida, este trabalho não teria sido possível sozinho. Em primeiro lugar gostaria de expressar a minha mais profunda gratidão aos meus orientadores. Ao Nuno Silva por ter sido acima de tudo um excelente amigo que me guiou por todo este processo, além de todas as contribuições que teve para o desenvolvimento deste trabalho e que permitiram chegar tão longe e adquirir tantos conhecimentos no processo. Agradeço também ao professor Pedro Jorge por me ter dado a oportunidade de entrar no ambiente da investigação e por toda a ajuda e preocupação que teve ao longo deste ano.

Gostaria ainda de agradecer a toda a equipa do Centro de Fotónica Aplicada do INESC que contribuíram direta e indiretamente para este trabalho. Foi um prazer desenvolver este trabalho num ambiente tão acolhedor e que realmente nos faz sentir integrados numa equipa.

Por último não podia deixar de mencionar os meus pais, amigos e a Rita, por me terem acompanhado durante todos estes anos e por todos os momentos proporcionados e que permitiram o meu desenvolvimento como indivíduo fora do contexto da faculdade. Sem eles, garantidamente este trabalho não teria sido possível.

UNIVERSIDADE DO PORTO

Abstract

Faculdade de Ciências da Universidade do Porto

Departamento de Física e Astronomia

MSc. Engineering Physics

New tools for high performance of LIBS Imaging

by [Tomás Lopes](#)

This thesis manuscript presents a research work in the field of LIBS-based Imaging, particularly the development of new and improved advanced tools to be adopted on current LIBS pipelines and that are capable of enhancing the resulting analysis provided by this technique.

Laser-induced Breakdown Spectroscopy has been an ever growing technique that has been widely adopted in various areas. As a result multiple approaches to the technique have been considered, with one of the most prominent being LIBS-based Imaging, where the analysis is extended to the spatial domain, resulting in what is referred to as a data cube, composed of two spatial dimension and one spectral dimension. To properly analyse the results of LIBS Imaging, it is of utmost importance that one has the proper knowledge and tools to explore the data cube to take advantage of all the information this technique provides. However this field is currently underexplored, with the majority of the community adopting pipelines that often result in difficulties in interpretation and visualisation of results.

To address this issues, we developed tools to better handle the complexity of these datasets. In first place we will go over an algorithm inspired by the JPEG compression algorithm, and based on the Fourier transform that facilitate the process of finding relevant emission lines for the LIBS imaging analysis, by introducing a metric that indicated how interesting each wavelength might be. Secondly we introduce a visualization pipeline that extends the LIBS analysis to the three dimensional domain, allowing for greater comprehension during the analysis, and thirdly we will go over the use of LIBS as a tandem solution with other spectral imaging techniques to enhance their individual capabilities,

testing the most often used solution of data fusion, in addition to a new approach based on the concept of knowledge distillation.

UNIVERSIDADE DO PORTO

Resumo

Faculdade de Ciências da Universidade do Porto

Departamento de Física e Astronomia

Mestrado em Engenharia Física

Novas Ferramentas para alto desempenho de LIBS Imaging

por [Tomás Lopes](#)

Esta tese apresenta um trabalho de pesquisa na área de Imaging baseado em LIBS, particularmente o desenvolvimento de novas ferramentas a ser adotadas em pipelines de LIBS atuais e que são capazes de melhorar a análise resultante fornecida por esta técnica. LIBS (espectroscopia de ruptura induzida por laser) tem sido uma técnica em constante crescimento que tem atraído um grande interesse nas mais diversas áreas. Como resultado, múltiplas abordagens à técnica foram consideradas, sendo uma das mais proeminentes LIBS Imaging, onde a análise é estendida ao domínio espacial, resultando no que é referido como um cubo de dados, composto por duas dimensões espaciais e uma dimensão espectral. Para analisar os dados resultantes de LIBS Imaging, é de extrema importância que se tenha conhecimento bem como ferramentas adequadas para explorar o cubo de dados e aproveitar todas as informações que esta técnica fornece. No entanto, este campo está atualmente subexplorado, com a maioria da comunidade a adotar pipelines que muitas vezes resultam em dificuldades na interpretação e visualização dos resultados. Para resolver estes problemas, no contexto desta tese, desenvolvemos ferramentas para melhor lidar com a complexidade destes datasets. Em primeiro lugar, vamos então analisar um algoritmo inspirado na compressão JPEG de imagens, e baseado na transformada de Fourier que facilita o processo de encontrar linhas de emissão relevantes para a análise de LIBS Imaging, introduzindo uma métrica que indica quão interessante cada comprimento de onda pode ser. Em segundo lugar, apresentamos um pipeline de visualização que estende a análise LIBS ao domínio tridimensional, permitindo uma maior compreensão durante a análise, e em terceiro lugar vamos estudar o uso de LIBS em colaboração com outras técnicas de spectral imaging numa tentativa de melhorar as suas capacidades individuais, testando a

solução mais usada de fusão de dados, além de uma nova abordagem baseada no conceito de knowledge-distillation.

Contents

Declaração de Honra	iii
Agradecimentos	v
Abstract	vii
Resumo	ix
Contents	xi
List of Figures	xiii
1 Introduction	1
1.1 A brief history of LIBS	1
1.2 Concepts of LIBS Analysis and Imaging	3
1.3 Project motivation and objectives	6
1.4 Work description	7
1.5 Major Output and Contributions	8
2 Fundamentals of Laser-Induced Breakdown Spectroscopy	11
2.1 The two Key concepts of LIBS: Laser and Plasmas	11
2.1.1 The Laser	11
2.1.2 The Plasma	14
2.2 Laser-induced plasmas	14
2.2.1 Ablation	15
2.2.2 Plasma formation	15
2.2.3 Plasma Decay and Optical Emission	16
2.3 The LIBS Technique	17
2.3.1 A typical LIBS Setup	17
2.4 LIBS Imaging	19
2.5 Concluding Remarks	20
3 LIBS Imaging Map Finder	23
3.1 Context	23
3.2 Fourier Transform	25
3.2.1 Map Finder Algorithm	25
3.3 Results and Discussion	27

3.4	Overview	30
4	Interactive three-dimensional chemical element maps with Laser-Induced Break-down Spectroscopy and photogrammetry	33
4.1	Towards interactive visualization tools for spectral imaging	33
4.2	Generating Three-dimensional Models	34
4.3	Mesh Matching Algorithm	35
4.3.1	Kabsch-Umeyama Algorithm	35
4.4	Workflow	37
4.5	Results	38
4.6	Overview	40
5	Collaborative Opportunities in Multimodal Sensor Fusion	41
5.1	The potential of Hyperspectral Imaging	41
5.2	Methodologies for Collaborative Sensing	42
5.2.1	Data Fusion	43
5.2.2	Multimodal Spectral Knowledge-Distillation	44
5.3	Results and Discussion	45
5.3.1	Data Fusion	45
5.3.2	Multimodal Spectral Knowledge-Distillation	49
5.4	Overview	51
6	Conclusion	53
	Bibliography	57

List of Figures

1.1	Schematic of the stages involved in LIBS-based Imaging analysis.	4
1.2	Some examples of analysed samples in the literature, including (A.) biological [24], (B.) industrial, (C.) heritage [28] and (D. and E.) geological samples [26, 28].	6
2.1	Schematic of the energy levels in an atom with A. and B. representing the processes of spontaneous and stimulated emission, respectively.	12
2.2	Scheme of an optical resonator based of a Fabry-Perot cavity.	13
2.3	Approximate timescales (considering a 10 ns pulse) of the processes that occur in laser-matter interactions leading to element specific emission	15
2.4	Spectral contributions on a LIBS experiment and approximate timescales for distinct phenomena	16
2.5	A. Illustration of a typical LIBS Setup. B. Experimental setups utilized during this thesis.	18
2.6	Example of a typical LIBS imaging application. A. Spectrum acquired from a rock sample where we can see some of the most prominent emission lines. B. LIBS map for the highlighted Li and Na emission lines, allowing for clear distinction of Lepidolite and Albite regions. C. Image of the sample portion being analysed. D. Mineral identification results obtained from a machine learning pipeline based of context based feature extraction and unsupervised kmeans clustering.	20
3.1	Structure of the data cube, demonstrating that the selection of emission lines isn't intensity specific as is exemplified by the oxygen emission line. . .	24
3.2	A. Black and White Image B. Representation of the image in the frequency domain (FFT) C. Post filtering representation of the frequency domain. D. Post filtering reconstruction of the image.	25
3.3	Schematic of the developed pipeline, where for each wavelength, the Fourier Transform is computed, filtering spatial frequencies in accordance to the defined feature size δ , followed by a quantization step, where the the sum of all filtered components is normalized to the sum of the total components 3.2 (here, only the DC component is filtered)	26
3.4	A. Graph of the FFT Score as a function of the wavelengths. B. Average Spectrum with two of the most prominent peaks shown C. Image of the sample being analysed. D. and E. Lithium and Sodium maps generated by the selected emission line in figure A, respectively.	29

3.5	A. Graph of the FFT Score as a function of the wavelengths, highlighting a high score line corresponding to Sodium, with the respective plot of the average spectrum. B. LIBS image generated from the highlighted emission line.	29
3.6	Graph of the time as a function of logical processes that are running the algorithm.	30
4.1	A. Process of photographing the object from various lines of sight. B. Generated 3D model of a rock sample. (Adapted from [28])	34
4.2	Graphical description of the visualization workflow. (Adapted from [28]) . .	37
4.3	Illustrative results of the potential of the approach for the generation of interactive visualization models, here applied to (A.) Heritage applications (Distribution of Cobalt on the surface of a Portuguese Tile, using the Cobalt line at $345.3nm$), (B.) non-planar geometries (distribution of silicon on a cork, using the Silicon line at $251.6nm$), and (C.) multimodal sensing (LIBS and hyperspectral data of a lepidolite-albite rock). (Adapted from [28]) . .	39
5.1	Illustration of the principle HSI. Here we can see the push broom scanning principle of this technique that scan the sample line by line, as well as the spectra that are captured at each acquisition.	42
5.2	Schematic of the adopted pipeline for sensor fusion classification	43
5.3	Schematic of the adopted pipeline for MSKD classification	45
5.4	A. Rock samples and indicative mineral zones. Spectral signature for B. LIBS and C. HSI, showcasing the typical signature of distinct mineral types.	46
5.5	Perceptual map exhibiting the average LIBS feature value and average HSI spectrum for each of the identified clusters.	47
5.6	A. Classification results obtained by applying our training pipelines to a training region. B. Classification results obtained by applying our trained models to a test sample region	48
5.7	Predictive error on training the PLS-DA algorithm using the test sample as a benchmark and LIBS labels as the ground truth.	50
5.8	Classification results from the testing set, using Multimodal Spectral Knowledge Distillation pipeline.	50
5.9	Classification results from the training set, using the trained Multimodal Spectral Knowledge Distillation model.	51

Chapter 1

Introduction

Laser-induced breakdown spectroscopy (LIBS) is a type of Optical emission spectroscopy (OES) that uses a focused laser pulse as the excitation source. The laser pulse is focused on a target surface, delivering a high energy density within a short time duration. As a consequence, a spot on the surface is ablated and a small amount of material is heated and evaporated, creating a gas plume that contains the sample atoms and ions. The gas plume then continues to absorb laser energy becoming further ionized into the state of a plasma. The optical emission of the plasma is then collected by a spectrometer and the resulting spectrum has characteristic wavelengths and intensities for each element present in the sample and can thus be used for qualitative and quantitative analysis of the target. This makes LIBS a fast, versatile, and micro-destructive technique that can be used for in situ and remote chemometric analysis of various types of sample matrices, including solids, liquids, and gases.

Still, while the potential of LIBS for a wide span of academic and technological applications is immense, the actual utilization of the technique involves the development and deployment of dedicated processing tools, algorithms, and software solutions for each specific purpose. The development of such solutions has been one of the research topics at INESC TEC in recent years and will be the main focus of this work.

1.1 A brief history of LIBS

The history of LIBS started with the invention of the laser in 1960, which provided a coherent source of optical radiation with enhanced focusing capabilities. Just two years

later, Breach and Cross [1] introduced what is now considered the precursor of LIBS, Laser-induced spark emission, harnessing the power of the laser to produce plasmas on a surface. Latter in 1963 [2], the first analytical use of laser-induced plasma for spectrochemical analysis was reported in France and is now considered the "birth" of LIBS. Still, during the first decades, LIBS stayed primarily on the academic plane, with the research focusing on the basic features of laser plasma rather than on its analytical capabilities.

The status quo started to change during the 1980s, when the analytical capabilities of LIBS became more evident, leveraging on technological advancements in lasers and detection devices with improved sensitivity, resolution, and speed. During this time, Radziemski and Cremers team at the Los Alamos National Laboratory [3] was leading the state-of-the-art, eventually coining the term LIBS for this method. Further down the road during the 1990s, LIBS made remarkable progress emerging as a versatile technique for both laboratory and field analysis, turning the century as the "future superstar" [4].

More recently, LIBS has become widely accepted as a reliable and versatile technique for elemental analysis. In particular, the technique features multiple advantages comparing to competitors such as X-Ray Fluorescence or ICP-MS, from which we can highlight:

- Ability to detect (in theory) all elements - in particular, lighter ones - with limits of detection typically in the range of a few ppm;
- Simultaneous detection of multiple elements;
- Possibility of performing a rapid analysis of a region of the sample surface, allowing LIBS-based spectral imaging, enabling spatially resolved elemental mapping with high spatial resolution and fast acquisition speeds;
- Furthermore, the substitution of a physical device (electrodes) with a laser to ignite the plasma:
 - Enables the analysis of gas, liquid, and solid samples;
 - Requires minimal sample preparation;
 - Can be employed in extreme conditions such as vacuum, underwater environments, high temperatures, and high pressures.
 - Operates at high speeds, reaching frequencies up to kHz.
 - Allows for remote (non-contact) operation, facilitating the development of portable instruments and online analysis.

- Limited and localized damage on the sample surface, with damage levels ranging from tens to hundreds of nanograms, which is particularly valuable for preserving cultural heritage or artworks;
- Provides in-depth profiling capabilities;
- Finally, LIBS can be integrated in collaboration with additional spectroscopic techniques, e.g. Raman spectroscopy, laser-induced fluorescence (LIF) or hyperspectral imaging, thus supporting interesting sensor fusion solutions for industrial applications.

These days, the current stage of LIBS technique is characterized by the increasing maturity and diversity of its applications and methodologies, with the most recent research trends now trying to push towards the improvement of the accuracy and precision of qualitative and quantitative analysis, with the development of new calibration strategies, exploring new spectral regions, coupling with other spectroscopy techniques, and much more.

1.2 Concepts of LIBS Analysis and Imaging

The typical output of a LIBS process is an emission spectrum (see Figure 1.1) with characteristic spectral lines that can be related to the spectral fingerprint of atoms, ions, and molecules. Speaking in broad terms, when talking about LIBS analysis we can divide it into two main categories: i) qualitative analysis, and ii) quantitative analysis.

Qualitative analysis is based mainly on the process of identifying the presence or absence of a constituent in the sample under analysis. Focusing on the case of atomic species for illustrative purposes, this can be done by comparing the emission lines present in the LIBS spectrum with known atomic spectra, which can be found in multiple databases such as the NIST one [5]. The presence or absence of specific emission lines thus provides information on the elemental composition of a material without the need for calibration or reference standards. Such type of analysis is often carried out using numerical algorithms for peak identification and can be also combined with further machine learning classification algorithms to distinguish between different sample types or sample regions. For example, in one of our most recent works, we have shown the possibility of identifying types of minerals when combining qualitative analysis with unsupervised clustering methodologies [6]. For such applications, the results are frequently reported as outstanding [7, 8], with

limits of detection in the range of a few ppms and even potentially reaching ppbs for LIBS systems with better detection devices (e.g. ICCD devices).

On the other hand, quantitative LIBS analysis typically refers to the process of determining the concentrations of elements or molecules in a sample, relating features of the LIBS signal (such as the intensity or the area of an emission line) to the concentration of the analyte. This can be done using multiple algorithms, from simple linear calibrations to more complicated data-driven neural network regression techniques [9]. Yet, contrary to what happens for qualitative analysis, LIBS typically underperforms for this type of analysis when dealing with more complex samples, mostly due to the intrinsic variability of the signal obtained [10]. One way to mitigate these effects is to perform calibrations with matrix-matched standards or utilize other advanced approaches such as calibration-free LIBS (CF-LIBS) [11]. Yet, both of these methodologies still do not solve the lack of control of the plasma variation with the sample type and sample surface characteristics (e.g. roughness, hardness), instrumental drifts or even environmental conditions (e.g. humidity), making reliable quantitative analysis one of the open challenges of the field.

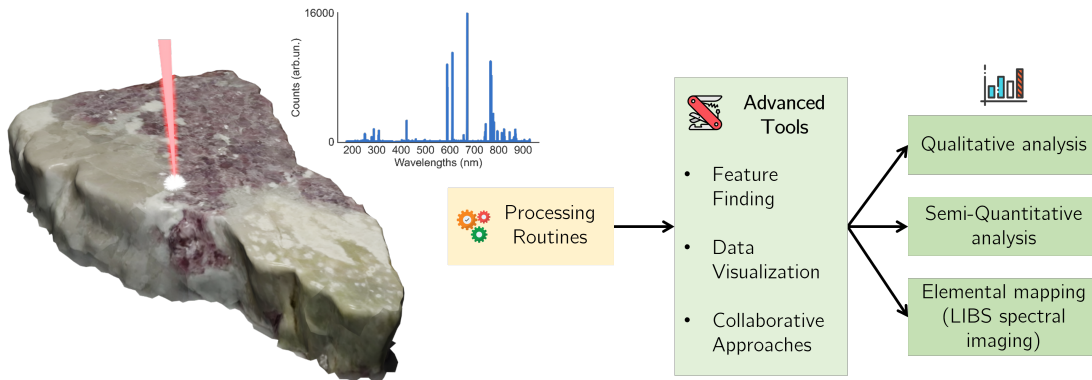


Figure 1.1: Schematic of the stages involved in LIBS-based Imaging analysis.

For the purpose of this work, we will focus mainly on qualitative analysis and in particular, on the topic of spectral imaging, widely regarded as one of the most promising concepts for technological applications for the years to come [12, 13]. LIBS-based spectral imaging or simply LIBS imaging combines localized analysis with fast acquisition rates to map the distribution of elements by spatially scanning the sample surface in a pixel-by-pixel manner, also called whiskbroom scanning. Typically, a specific emission line is selected and its value is represented in the transverse spatial directions to obtain an elemental map. Involving simple optical instrumentation, fast acquisition speeds, and multi-element capabilities with detection limits in the range of few ppms for most elements, LIBS imaging has

found an extensive and increasing range of applications in recent years, spanning a wide range of scientific areas from medicine to biology, and from geology to industry [13–16].

Some of the recent developments involving LIBS-based Imaging, address communities that greatly benefit from its capabilities. By allowing multielement detection, along with fast acquisition speeds with no constraints on sample size, several studies have been conducted and have demonstrated its ability to characterize complex samples with some examples being the analysis of a multiphase hydrothermal ore sample, with more than 30 elements identified with concentrations that range from major to trace elements [17], the analysis of biological samples, aiming to analyze trace elements in biological tissues at microscopic spatial resolutions for routine elementary investigations [18–21], and remote sensing under dangerous environments with studies conducted for the Experimental Advanced Superconducting Tokamak (EAST) magnetic fusion energy reactor, focusing on analysing the 2D distribution of Li deposition in plasma-facing materials [22] and the distributions of impurities (H, O, Ar, K, Na, and Ca) of a co-deposition layer of Li-W used in the reactor walls [23]. Another major benefit of the technique is its ability to perform depth profiling analysis, with application including multielemental 3-D images of soft tissues, either by volume reconstruction or in-depth analysis [24], to paleoclimatic research on the characterization of speleothems [25, 26] where climate proxies, such as Mg, Sr, or Ba, can be retrieved from these materials, and may indicate changes in climate or environmental conditions that occurred during the growth periods, requiring an analysis of the order of centimeters with a resolution on the scale of the annual speleothem growth of around 0.01 to 2 mm per year [27].

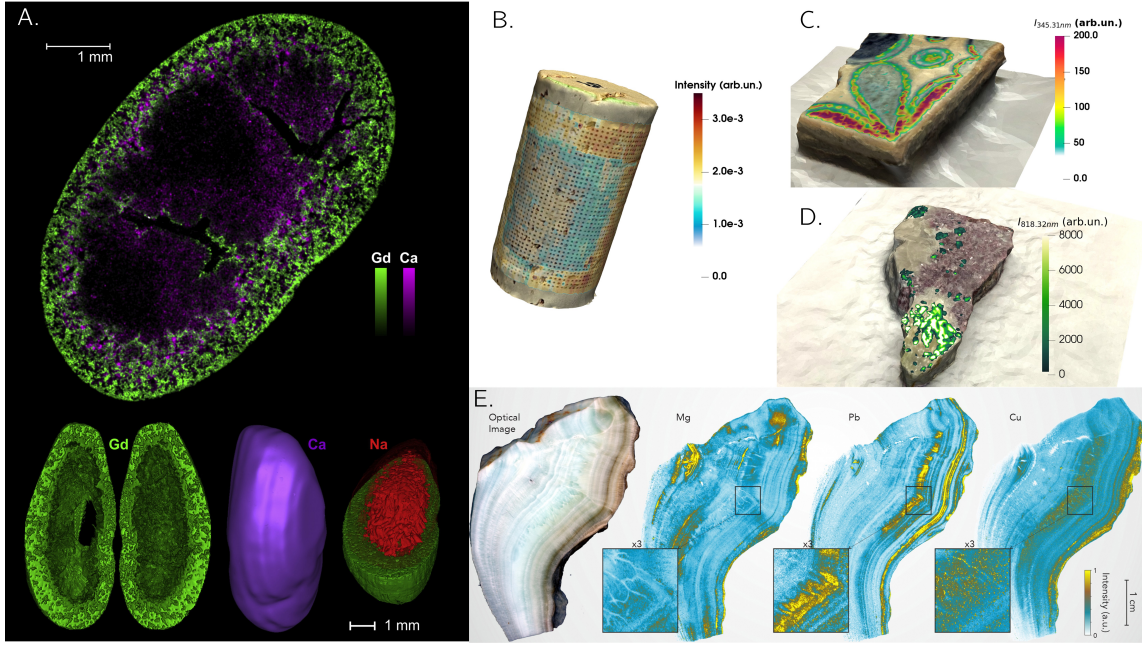


Figure 1.2: Some examples of analysed samples in the literature, including (A.) biological [24], (B.) industrial, (C.) heritage [28] and (D. and E.) geological samples [26, 28].

1.3 Project motivation and objectives

The results of LIBS imaging are only as good as its software tools and algorithms, making the development of such solutions a subject of paramount importance. In particular, these tools involve multiple operations with their own specific challenges. For example, for the pre-processing steps such as baseline removal [29, 30], peak finding, and feature detection, the most prominent challenge is to process large amounts of data, as these hyperspectral cubes ($N_x \times N_y \times N_\lambda$ points, with N_x and N_y being the spatial dimensions and N_λ the wavelength dimension) can rapidly scale to GB of information. After this, and depending on the final objective, the processing pipeline can continue with machine learning algorithms such as unsupervised clustering for classification [31], supervised classification [32], or even dimensionality reduction techniques such as Principal Component Analysis (PCA) [33]. Here the challenge normally focuses on deploying robust algorithms, capable of providing reliable, reproducible, and accurate results. Finally, data visualization also plays an important role, as they allow us to better understand the results.

Therefore, LIBS imaging has a lot of room for research, in particular in the development and deployment of software tools and strategies capable of enhancing the quality and interpretability of its results. This context constitutes the main subject of this manuscript, and its operational objectives stem from the natural progression of the research at INESC

TEC in the last year in the field of LIBS imaging, in particular:

- Develop new tools to find relevant information in the generated data of a LIBS imaging experiment, paving not only for faster processing and analysis workflows but also for future work regarding data compression strategies;
- Develop innovative solutions for interactive analysis of spectral imaging data, in particular focusing on the integration of 3D model information with 2D elemental maps, opening new perspectives for academic and technological applications;
- Deploy novel processing pipelines capable of exploring the collaboration spectral imaging sensor fusion in more effective manners, establishing a tool of promising potential for industrial applications by combining LIBS and hyperspectral imaging.

1.4 Work description

Describing the methodology to achieve these goals in an integrated manner, this manuscript is divided into 5 chapters. In addition to the introduction in the present chapter, Chapter 2 is devoted to a theoretical overview of Laser-induced Breakdown Spectroscopy, going over the core physical principles, setup and the introduction of LIBS imaging to better understand the results presented in the following chapters. In Chapter 3, we present an algorithm that uses the Fourier Transform to automatically detect features of interest that can be used for consecutive analysis, exploring its implementation with parallel computing and the benefits it has towards a real time data analysis environment. In Chapter 4, we combine 3D photogrammetry models of analysed samples with maps of the surface to create a 3D visualization environment capable of enhancing the interpretability of Spectral imaging maps by extending the analysis to the 3D domain, and lastly in Chapter 5, we will go over the use of LIBS as tandem solution alongside hyperspectral imaging (HSI), exploring the concept of collaborative sensing. The results in this chapter demonstrate the classification capabilities of two different collaborative sensing approaches, one that combines the resulting data cubes from LIBS and HSI through a mid-level fusion, and another that uses LIBS as a higher knowledge technique to train an HSI classification model that is reliable and balances good classification with fast acquisition speeds. Finally, a general conclusion will be made presenting some prospects for future works in this area.

1.5 Major Output and Contributions

In addition to the current document, this MSc dissertation has led to numerous communications and outputs in close collaboration with other members of the Spectral Imaging team at the Center for Applied Photonics at INESC TEC. These include:

Journal Publications:

- "Interactive three-dimensional chemical element maps with laser-induced breakdown spectroscopy and photogrammetry", Tomás Lopes, Pedro Rodrigues, Rafael Cavaco, Diana Capela, Miguel F.S. Ferreira, Diana Guimarães, Pedro A.S. Jorge, Nuno A. Silva, *Spectrochimica Acta Part B: Atomic Spectroscopy*, Elsevier. [28] (2023)
- "Robust and interpretable mineral identification using laser-induced breakdown spectroscopy mapping", Diana Capela, Miguel F.S. Ferreira, Alexandre Lima, Filipa Dias, Tomás Lopes, Diana Guimarães, Pedro A.S. Jorge, Nuno A. Silva, *Spectrochimica Acta Part B: Atomic Spectroscopy*, Elsevier. [6] (2023)

Conference Proceedings:

- "Multimodal approach to mineral identification: merging Laser-induced breakdown spectroscopy with Hyperspectral imaging", Tomás Lopes, Rafael Cavaco, Pedro Rodrigues, Jorge Ferreira, Diana Capela, Miguel F. S. Ferreira, Pedro A. S. Jorge and Nuno A. Silva, *Proceedings of AOP - Conference on Applications of Optics and Photonics*. (2022)
- "Integrating Laser-induced breakdown spectroscopy and photogrammetry towards 3D element mapping", Pedro Rodrigues, Tomás Lopes, Rafael Cavaco, Diana Capela, Miguel F. S. Ferreira, Pedro A. S. Jorge and Nuno A. Silva, *Proceedings of AOP - Conference on Applications of Optics and Photonics*. (2022)
- "Listening plasmas in Laser-Induced Breakdown Spectroscopy", Rafael Cavaco, Pedro Rodrigues, Tomás Lopes, Diana Capela, Miguel F. S. Ferreira, Pedro A. S. Jorge and Nuno A. Silva, *Proceedings of AOP - Conference on Applications of Optics and Photonics*. (2022)

Oral Presentations:

- "Collaborative Opportunities in Multimodal Sensor Fusion", Tomás Lopes, Diana Capela, Miguel F.S. Ferreira, Diana Guimarães, Pedro A.S. Jorge, Nuno A. Silva, IJUP2023 - Young Researcher Meeting of University of Porto. (2023)

Poster Presentations:

- "Interactive three-dimensional chemical element maps with Spectral Imaging and photogrammetry", Tomás Lopes, Pedro Rodrigues, Rafael Cavaco, Diana Capela, Miguel F.S. Ferreira, Diana Guimarães, Pedro A.S. Jorge, Nuno A. Silva, DCE2023 - Doctoral Congress in Engineering. 2023
- "Multimodal approach to mineral identification: merging Laser-induced breakdown spectroscopy with Hyperspectral imaging", Tomás Lopes, Rafael Cavaco, Pedro Rodrigues, Jorge Ferreira, Diana Capela, Miguel F. S. Ferreira, Pedro A. S. Jorge and Nuno A. Silva, PRODEF - Doctoral Program in Engineering Physics. 2023

Inventions:

- Spectral Knowledge Distillation: a disruptive approach to spectral imaging and sensor fusion for industrial applications - Under consideration for patent application;

Chapter 2

Fundamentals of Laser-Induced Breakdown Spectroscopy

This chapter is devoted to a brief overview of the theory behind Laser-Induced Breakdown Spectroscopy, a powerful tool for chemical analysis capable of providing us insight of a sample composition. Over the past few years, LIBS has gained considerable momentum in the field of chemical analysis due to its versatility compared to other techniques such as ICP-MS, XRF, or X-Ray Diffraction (XRD) leading to applications in a variety of fields, from biomedical research to heritage sciences, just to name a few.

In particular, this chapter will provide a comprehensive discussion of the inner workings of LIBS analysis, from the optical instruments utilized on a general LIBS setup to plasma formation and spectral emission. Then, we provide a brief overview of the typical LIBS Imaging analysis, attempting to provide a solid understanding of the technique to better interpret the results enclosed in the following chapters.

2.1 The two Key concepts of LIBS: Laser and Plasmas

To understand the inner-working of LIBS one can start by discussing the two fundamental concepts necessary to make it work: lasers and plasmas.

2.1.1 The Laser

LASER stands for Light Amplification by Stimulated Emission of Radiation and is arguably one of the most important technologies of the last century, branching out to an immense variety of fields. As the name implies, the mechanism behind LASER sources is

stimulated emission, a phenomenon first described by Albert Einstein in 1917, and to better understand this mechanism it is important to understand the different types of atomic emission.

Atoms emit radiation when their electrons decay from upper to lower energy levels, releasing a photon in the process. A photon is an elementary particle that carries a quantized packet of energy of a given electromagnetic field, such is the case of light. Thus, the emitted photons have a specific energy E and wavelength λ that depends on the energy difference between the upper and lower levels of the electron transition in the atom, i.e. $E = \frac{hc}{\lambda} = E_{upper} - E_{lower}$.

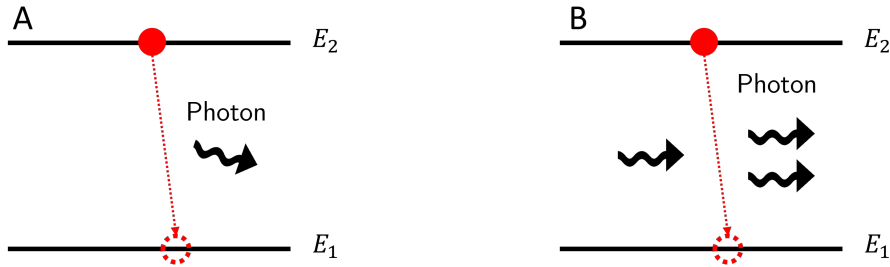


Figure 2.1: Schematic of the energy levels in an atom with A. and B. representing the processes of spontaneous and stimulated emission, respectively.

The emission of the photon can occur in a spontaneous or stimulated manner (see Figure 2.1). Spontaneous emission is the process in which an excited electron or atom decays to a lower energy level without any external influence, emitting a photon. This process occurs randomly and does not depend on the presence of other photons. The direction and phase of the emitted photon are also random and can only result in incoherent light. Stimulated emission on the other hand is the process in which an excited electron decays to a lower energy level due to the interaction with an additional photon. In this case, the incoming photon stimulates the emission of a second photon that is identical to it in frequency, phase, polarization, and direction. This is the basis of laser operation, as it allows the amplification of light by creating coherent and monochromatic beams of photons.

In 1960 Theodore Maiman had the idea to use an optical cavity with a gain medium on its inside (a medium that can amplify the power of light), amplifying the light with each pass through the gain medium by means of stimulated emission. In stimulated emission, an incoming photon can interact with an atom or molecule in the gain medium that has excess energy, causing it to emit another photon of the same wavelength and phase as the incoming one. This way, one photon can trigger the emission of many more photons,

creating a chain reaction that amplifies the light. However, not all atoms or molecules in the gain medium are ready to emit photons. Some of them are in a lower energy state, and they can absorb photons instead of emitting them. This reduces the amplification of the light, and as such, to make a laser work, the device needs to create a situation where the number of atoms or molecules in the gain medium in the higher energy state exceeds those in the lower energy state. This is called an inversion of population and to achieve this, energy must be supplied to the gain medium, such as electricity or light in a process called pumping. This scheme can be visualized in Figure 2.2, and by placing an active medium in a mirror cavity, feedback is enabled, allowing for population inversion to occur, greatly enhancing gain, and inducing laser action.

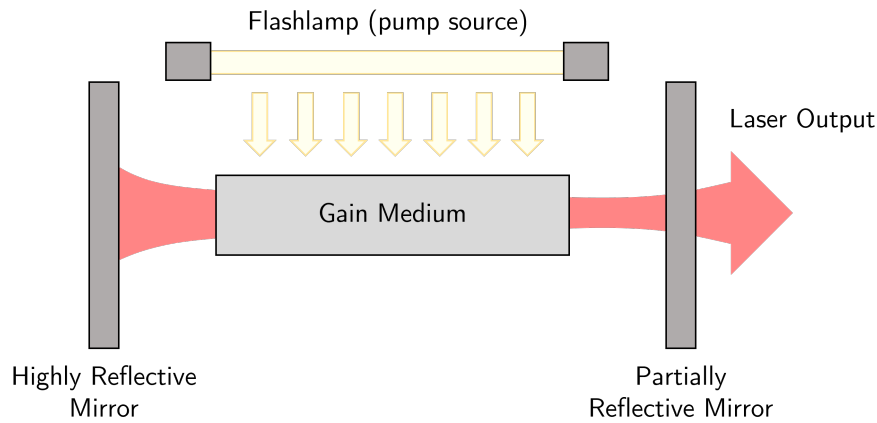


Figure 2.2: Scheme of an optical resonator based of a Fabry-Perot cavity.

Pulsed lasers are a subset of lasers that emit light in the form of short pulses rather than a continuous beam. Pulsed lasers can have various pulse durations, ranging from femtoseconds to milliseconds, and various pulse repetition rates, ranging from a few hertz to gigahertz. There are mainly two different methods for generating pulsed laser output from a laser medium, Q-switching and mode-locking. In Q-switching the optical losses inside the laser cavity are actively (using an electro-optic or acousto-optic device) or passively (using a saturable absorber) modulated, in such way that the population inversion builds up until the optical loss is suddenly reduced, triggering a giant pulse of light. In its turn, mode-locking involves locking the phases of different longitudinal modes of the laser cavity with something such as a saturable absorber or an active modulator, resulting in a train of ultrashort pulses with a fixed repetition rate determined by the cavity length. Pulsed lasers are mainly used when high intensities are required and that would otherwise be impossible

to achieve in continuous modes of operation, and therefore are fundamental in the process of laser-induced plasma formation.

2.1.2 The Plasma

The term "plasma" was initially introduced by I. Langmuir [34] in 1928 during his investigation of ionized gases in discharge tubes. Langmuir drew inspiration from the way that these discharges carried various particles such as high-velocity electrons, gas impurity molecules, and ions, and its resemblance with how blood plasma transports red and white cells, proteins, hormones, and germs. Considered the fourth state of matter, plasma is the most common state of matter in the universe as a whole, and in spite of it often being thought of as a subset of gases, the two states behave very differently. Like gases, plasmas have no fixed shape or volume, and are less dense than solids or liquids, but unlike ordinary gases, plasmas are made up of atoms in which some or all of the electrons have been stripped away and the positively charged particles, called ions, roam freely. Being made of charged particles, plasmas display distinct properties from gaseous media, like conducting electricity or generating magnetic fields. Plasmas also exhibit different properties from gases at the level of individual particles. In a gas, the particles tend to have similar speeds and energies, while in a plasma, the particles can have very different speeds and energies depending on the external fields and the degree of ionization (α),

$$\alpha = \frac{\eta_i}{\eta_i + \eta_n}, \quad (2.1)$$

where η_i and η_n are the ions and free electron densities, respectively.

In general, a plasma is formed when a gas is heated to an extremely high temperature or interacts with a strong electric or magnetic field, which causes such vigorous collisions between its atoms and molecules that electrons are ripped free in a process called ionization.

2.2 Laser-induced plasmas

Having introduced the two fundamental concepts, laser, and plasmas, we are now in a position to explore how these two come together to establish one of the fundamental pillars of the LIBS technique, the laser-induced plasmas (LIPs).

It all starts with a pulsed laser focused on a small spot on a surface sample. Having sufficient energy density (10's of GW/cm^2) - typically around a few mJ for 10ns pulse

duration - the energy is sufficient to trigger a series of events that will eventually lead to the laser-induced plasma (see Figure 2.3).

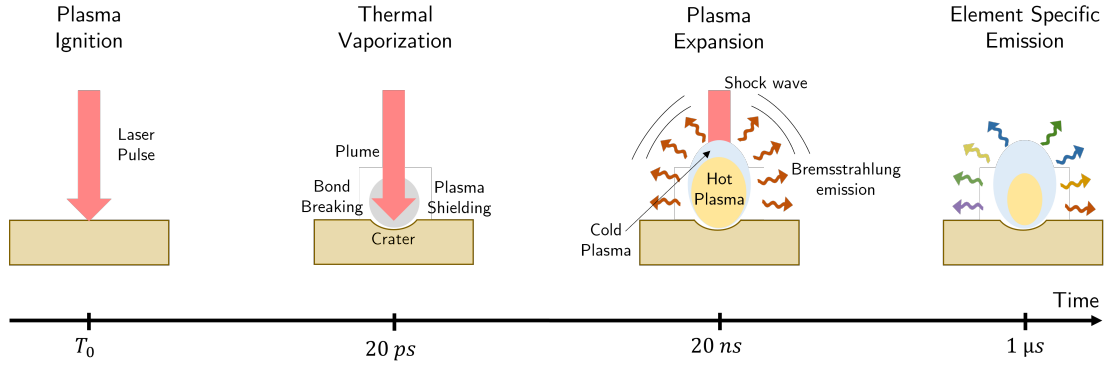


Figure 2.3: Approximate timescales (considering a 10 ns pulse) of the processes that occur in laser-matter interactions leading to element specific emission

2.2.1 Ablation

Ablation is the process of material removal and formation of a plume and is the direct consequence of two main mechanisms, photo-thermal ablation and photo-chemical ablation. Photo-thermal ablation happens when the laser energy is absorbed by the free electrons in the material and the energy is transferred to the sample matrix through electron-lattice coupling. This happens when the pulse duration is higher than the electron-lattice coupling time constant, which is of the order of picoseconds (typically what happens for a pulsed laser on the nanosecond range), effectively acting as a heat source, melting and vaporizing the material. In its turn, photo-chemical ablation occurs when the energy of the laser beam is sufficiently high to cause direct bond breaking of the material and leads to material removal through fragmentation causing in principle no temperature changes on the target surface, which is the typical scenario for femtosecond pulsed lasers. For the next sections we will focus our attention on the nanosecond regime only, which corresponds to the experimental setups we have available in the laboratory.

2.2.2 Plasma formation

Following ablation, which happens in the first microseconds, the electromagnetic field of the laser pulse eventually induces the breakdown of the gas into a plasma, leading to the subsequent generation of a shock and an acoustic wave. After this process, a plasma shielding effect continues further promoting the laser radiation absorption, heating the

plasma up to the final temperatures in the range $10^4 - 10^6 K$. The evolution of the plasma plume depends on several factors such as laser pulse characteristics, size of the focal spot, target matrix, and the surrounding atmosphere, just to name a few, and is one of the sources of variability of the signal acquired during a LIBS process.

2.2.3 Plasma Decay and Optical Emission

As soon as the laser pulse ends, the plasma begins to cool down, and de-excitation processes begin to occur. This is caused by three main mechanisms. First, we have i) recombination radiation when a free electron is captured by upper bound ionic or atomic level, accompanied by ii) Bremsstrahlung radiation, resulting from the collision and deflection of free electrons with atomic nuclei or other electrons. As the energy for these processes is not quantized, these two give rise to continuous emission that appears as a background in the optical emission spectra.

After this first period, typically around a microsecond, transitions between bound states in atoms, ions, and molecules start being dominant. These transitions result in the presence of specific atomic, ionic, and molecular emission lines in the emitted radiation spectrum. A graph demonstrating the types of emission and associated timescales can be seen in Figure 2.4.

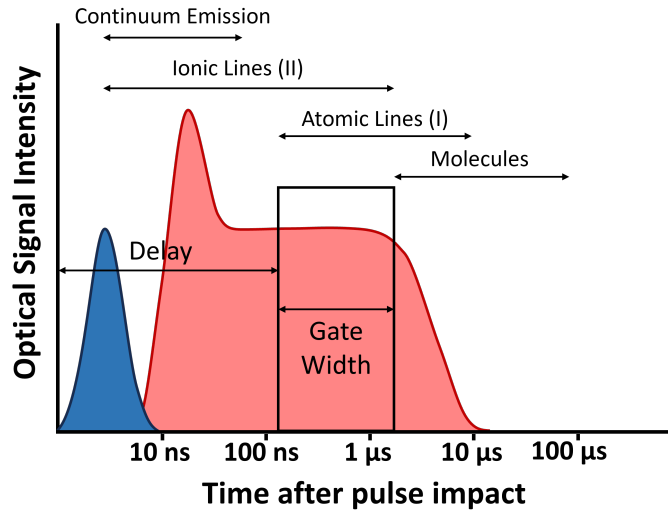


Figure 2.4: Spectral contributions on a LIBS experiment and approximate timescales for distinct phenomena

For LIBS analysis, this element-specific emission regime is the one of greater relevance. In particular, assuming local thermodynamic equilibrium conditions [35], the intensity of

the emitted lines can be approximated analytically by the Saha-Boltzmann equation

$$I_{\lambda_{ki}} = \frac{hcN g_k A_{ki}}{4\pi\lambda_{ki}Z(T)} e^{-\frac{E_k}{k_B T}} \quad (2.2)$$

where the emission line intensities $I_{\lambda_{ki}}$ are described as function of the line wavelengths λ_{ki} , total number density N , statistical weight g_k , transition probability between levels k and i A_{ki} , energy of the upper level E_k , plasma temperature T_p and the partition function $Z(T)$.

It is important to note however that experimentally the emission lines are not infinitely narrow, but have a finite width that depends on various factors and are a consequence of broadening effects. Doppler and Stark effects are the most relevant physical processes that contribute to broadening, to which we need to add an instrumental broadening that tends to dominate most lines. In particular, Doppler broadening is a direct consequence of the thermal motion of the emitting atoms or ions, which causes a Doppler shift in their emitted wavelengths producing a Gaussian line profile. The Stark broadening is caused by perturbations induced by the electric field of the electrons on the energy levels of a given transition producing a Lorentzian line profile. The resulting line profile is, therefore, a convolution of the different broadening effects, and is most often described using the Voigt profile. Typically, the broadening of the emission lines affects the resolution and accuracy of the spectral analysis but can also provide information about the physical conditions of the plasma, such as plasma density or temperature [36].

2.3 The LIBS Technique

The topics covered up to this point have enabled us to better understand the mechanism behind laser-matter interaction and how the radiation that is emitted in this process contains relevant information, that if properly harnessed, could provide useful to analyze the composition of a sample. As a whole, these summarize the working concepts of LIBS: first, ablate and generate a plasma at the sample surface using a focused pulsed laser, then collect and analyze the optical emission signal.

2.3.1 A typical LIBS Setup

The crucial components of a LIBS setup are therefore a pulsed laser acting as an excitation source to induced a plasma together with a set of lens and mirrors to steer and focus the laser pulse onto the sample. A light collection system - e.g. a collimator, a set of lenses of a parabolic mirror - collects the optical emission signal and couple it to multimode

optical fibers that are connected to a spectrometer that separates and measures spectral components of the optical signal. To this, it is common to add some translation stages, with one axis used to assure that the surface of the sample is at the focal spot (z) and two others (x and y) to allow for sample translation in the transverse directions, enabling us to capture spectra at several spots in the sample and fundamental for LIBS imaging applications. Finally, a digital delay generator synchronizes all the components and a desktop computer controls the system components and analyzes the data.

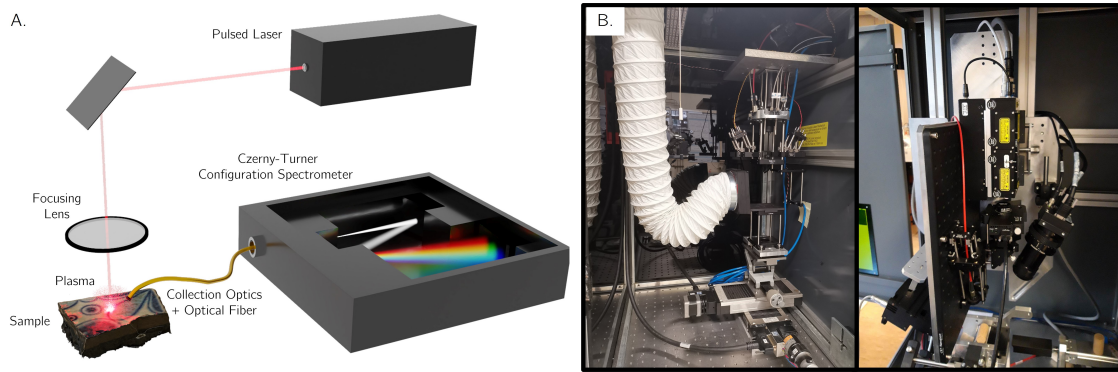


Figure 2.5: A. Illustration of a typical LIBS Setup. B. Experimental setups utilized during this thesis.

System	General Analysis	Cork Analysis
Laser	Quantel CFR 200	Quantel Viron
Pulse Energies	200 mJ	50 mJ
Fundamental Wavelength	1064 nm	1064 nm
Spectrometer	Avantes FC-UV200-1.5-SR	AvaSpec-ULS2048CL-EVO
Spectral Range	180-920 nm	226-334 nm
Repetition Rate	20 Hz	20 Hz
Sample Geometry	Planar	Cylindrical

Table 2.1: Description of the experimental setups utilized during this thesis.

Although the setups described and utilized are the most typical LIBS setup, other variants have been explored in the literature, normally designed for specific use cases. Some examples include double pulse LIBS [37] for enhanced plasma stability, telescopic LIBS [38] utilized for large distance analysis, catapulting LIBS [39] useful for trapped single particle analysis, or μ -LIBS [40], utilized for ultra-precise micrometer LIBS mapping.

2.4 LIBS Imaging

LIBS Imaging combines localized analysis with fast acquisition rates to map a sample surface through whiskbroom scanning, making it a valuable tool in numerous scientific and technological fields. Compared to other techniques like X-Ray fluorescence and Energy-dispersive X-ray spectroscopy, LIBS allows for faster multi-element detection, including lighter elements, with a simple setup and sample preparation process that increases its accessibility and versatility for use in various environments and in-situ applications. This has led to its adoption across different fields of research, such as mineralogy, metallurgy, biomedical and heritage sciences, among others, with potential applications up to the micrometer range.

In typical LIBS imaging applications, one selects a specific emission line to create a visual representation of chemical element distributions on the sample surface. This representation often consists of a map for the intensity of a given line or the ratio of chosen line intensities (as shown in Figure 1.2). Additionally, other approaches have been explored in various contexts such as depth profiling using multiple shots at the same site to create a depth-resolved spectral image.

An example of a use case of LIBS Imaging is its use for mineral classification of the rock sample of Figure 2.6C, having 4 different mineral regions Albite, Quartz, Lepidolite and Mica. Lepidolite, being the only Li-rich mineral, exhibits higher intensity of Li lines, whereas for example Albite is a Na-rich mineral that features higher Na line intensities. By merely looking at image 2.6B, where the spatial distribution of Lithium and Sodium is being shown, it is possible to make an informed decision on the location of certain mineral regions, however, some distinct minerals may have similar elemental composition, calling for the use of integrated reliable solution that are essential in LIBS-based Imaging. As one can see, in this particular case, following proper preprocessing routines, a k-means clustering algorithm was applied to locate the various minerals regions. The results obtained from such approaches often (see Figure 2.6 D.) present high degrees of accuracy and perfectly demonstrate, how capable this technique can be when proper methodologies are adopted.

In spite of its many capabilities and ever growing community of users and researchers, LIBS Imaging is still relatively new and there are still many opportunities to explore and expand on.

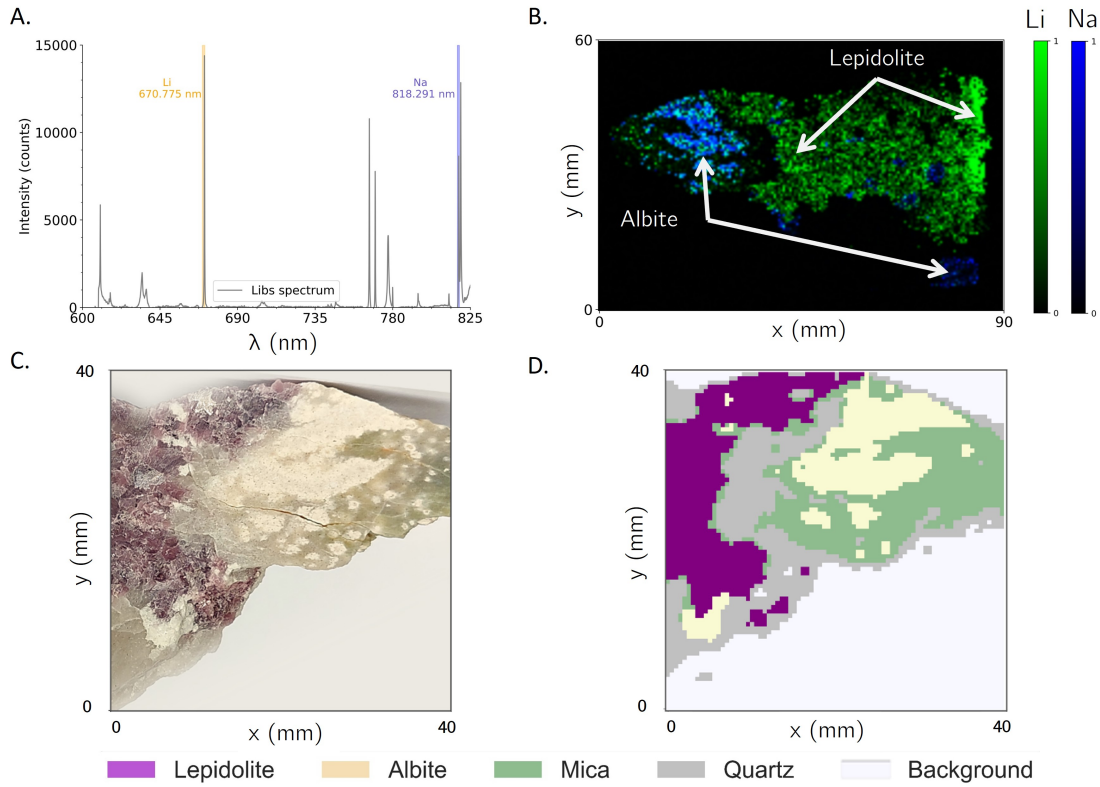


Figure 2.6: Example of a typical LIBS imaging application. A. Spectrum acquired from a rock sample where we can see some of the most prominent emission lines. B. LIBS map for the highlighted Li and Na emission lines, allowing for clear distinction of Lepidolite and Albite regions. C. Image of the sample portion being analysed. D. Mineral identification results obtained from a machine learning pipeline based of context based feature extraction and unsupervised kmeans clustering.

2.5 Concluding Remarks

Laser-induced Breakdown Spectroscopy has come a long way since its first introduction a few years after the discovery of the Laser source, with various adaptations soon following. That is the case of LIBS-based Imaging, that is currently experiencing strong developments with an increase in number of published papers in recent years, becoming a reference approach for spatially-resolved elemental characterization. All this has been enabled by the capability of the LIBS technique to reach micrometric scale resolutions with limits of detection down to ppm, accessing acquisition rates of the order of kHz [41] at ambient conditions using simple, all optical, tabletop instrumentation.

Despite the numerous advantages of LIBS-based imaging, there are still some technological barriers to overcome. In particular, the analysis of complex materials still appears to be challenging both in quantification and data processing. Then, new studies on the quantification of the technique, and the development of advanced tools becomes of utmost

importance to allow the technique to spread further and attract more applications in fields that go beyond the academic environment.

Chapter 3

LIBS Imaging Map Finder

As referred before, Laser-Induced Breakdown Spectroscopy (LIBS) has recently become an important tool for the study of the spatial distribution of chemical elements at a sample surface, a technique usually called LIBS imaging. The standard pipeline adopted for spectral LIBS imaging datasets revolves around identifying an element that is suspected to exist within the sample and generating maps based on its characteristic emission lines. In spite of this being the most common approach, maps containing relevant information are often discarded due to the lack of tools to identify relevant structures within the captured data, which can rapidly scale for a complex sample featuring hundreds to thousands of atomic, ionic, and molecular emission lines. In light of that, this chapter presents an algorithm based on the Fast Fourier Transform (FFT), capable of autonomously identifying emission lines or spectral regions containing meaningful spatial structures that can be further associated with relevant features for imaging analysis.

3.1 Context

As previously explained, in the typical LIBS Imaging process one scans the surface of a target in a point-by-point manner (whiskbroom scanning) to generate a 3D dataset, referred to as a hyperspectral or spectral data cube of size $N_x \times N_y \times N_\lambda$. The first two dimensions correspond to the two transverse spatial dimensions, while the third consists of the spectral dimension in the emission wavelengths. To construct element maps, one seeks for the presence of characteristic emission lines, integrate or slice the cube in the spectral dimension around the wavelength of interest λ_0 , and plot the resulting two dimensional data as a spatial map.

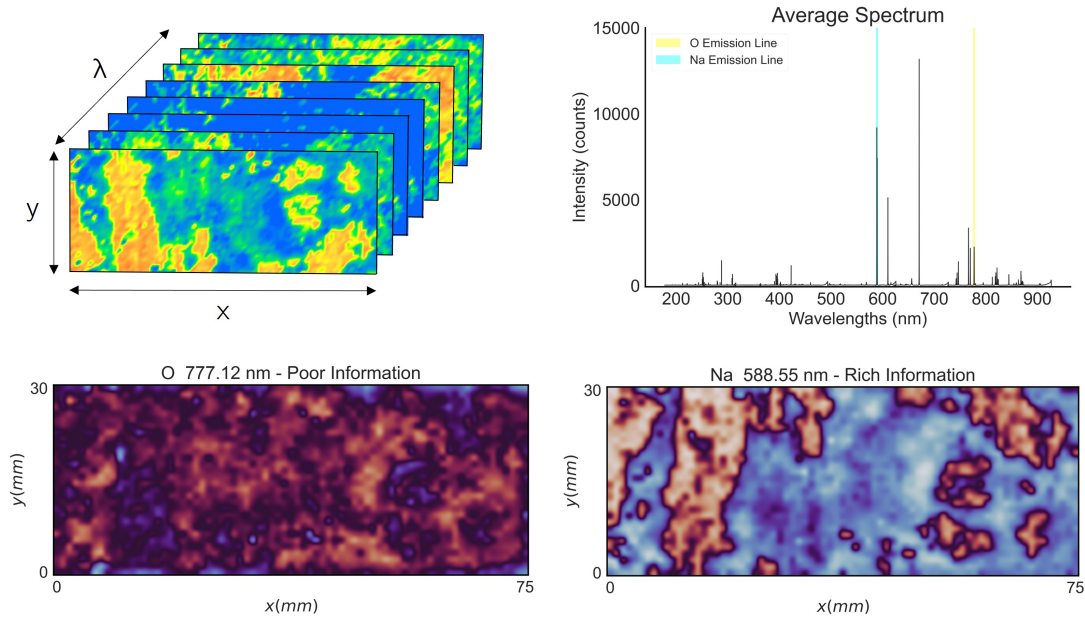


Figure 3.1: Structure of the data cube, demonstrating that the selection of emission lines isn't intensity specific as is exemplified by the oxygen emission line.

The goal is of course to find emission lines from elements that reveal underlying spatial structures in our sample and possibly be related to something we are looking for: for example, in the case of a rock, the distribution of an element at the sample surface might be related to the presence of a certain mineral [42]. This is typically done in one of two ways. The most common is to use our prior information of the sample composition and search for emission lines of elements we suspect to be present. On the other hand, if the composition of the sample is not known, one can plot the average, the minimum and the maximum spectra in the same plot, looking for lines which may contain any variation and thus possibly a spatial structure. While both these approaches have become the standard of LIBS Imaging analysis, they can be time-consuming and with the increased complexity of LIBS-generated data that can rapidly reach upwards of 100GB for a single map in advanced setups such as μ -LIBS, it becomes necessary the development of new strategies to allow the identification of relevant features for consequent analysis.

Following this reasoning and this necessity, we developed a map finder algorithm based on the Fast Fourier Transform capable of autonomously identifying without any prior knowledge, the spectral regions that may contain interesting maps for further analysis. Besides, by taking advantage of multiprocessing capabilities it may enable faster analysis of LIBS-based Imaging while identifying important features that are normally buried in the large dimensionality of the dataset.

3.2 Fourier Transform

The basis of the developed algorithm is the Fourier Transform, a mathematical tool that enables the representation of a signal in its frequency components. Its computational implementation, the Fast-Fourier Transform, is widely regarded as one of the most important algorithms of all time [43] and has a lot of applications spanning across multiple domains of science, from seismology [44] to cryptography [45], and from optics [46, 47] to medicine [48, 49]. Besides, it also plays a crucial role in image processing [50], being one of the most important ingredients of image compression algorithms (such as JPEG), making image files smaller and easing their storage and transmission.

Indeed, images can be thought of as signals that span across two dimensions, and so, by applying the Fourier Transform over both dimensions, one obtains the 2D Fourier Transform of an image, as seen in Figure 3.2B, where the spatial frequency of the components increases the further away we are from the center.

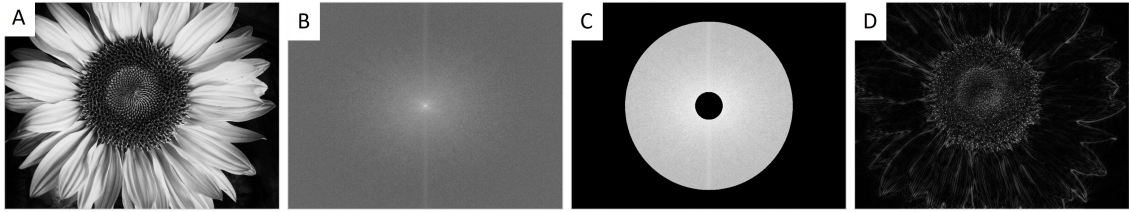


Figure 3.2: A. Black and White Image B. Representation of the image in the frequency domain (FFT) C. Post filtering representation of the frequency domain. D. Post filtering reconstruction of the image.

It is then possible to establish the relations between the dimensions in the spatial domain (δ) and its components in k-space as

$$k = \frac{2\pi}{\delta}. \quad (3.1)$$

Then, using a mask, we can filter out the k-space components that lie outside $\frac{2\pi}{\delta_{max}} < k < \frac{2\pi}{\delta_{min}}$ and be left with the components that characterize our structures within the chosen spatial range.

3.2.1 Map Finder Algorithm

Inspired by the use of FFT for image processing, our idea is then to apply the 2D Fourier Transform to the transverse directions of the data cube generated for a LIBS imaging experiment. Indeed, by establishing a spatial criteria for the structures we are looking

for, we can filter out the components of the transform which are not important - low frequencies that are related to continuous or very large structure components and high frequencies that are related to noise. This ring in the Fourier domain (see Figure 3.2C) corresponds to a bandpass filter and the information enclosed may have a higher or lower magnitude depending on whether there is, in fact, an underlying structure - i.e. a map - in the spatial dimensions we are looking for or not.

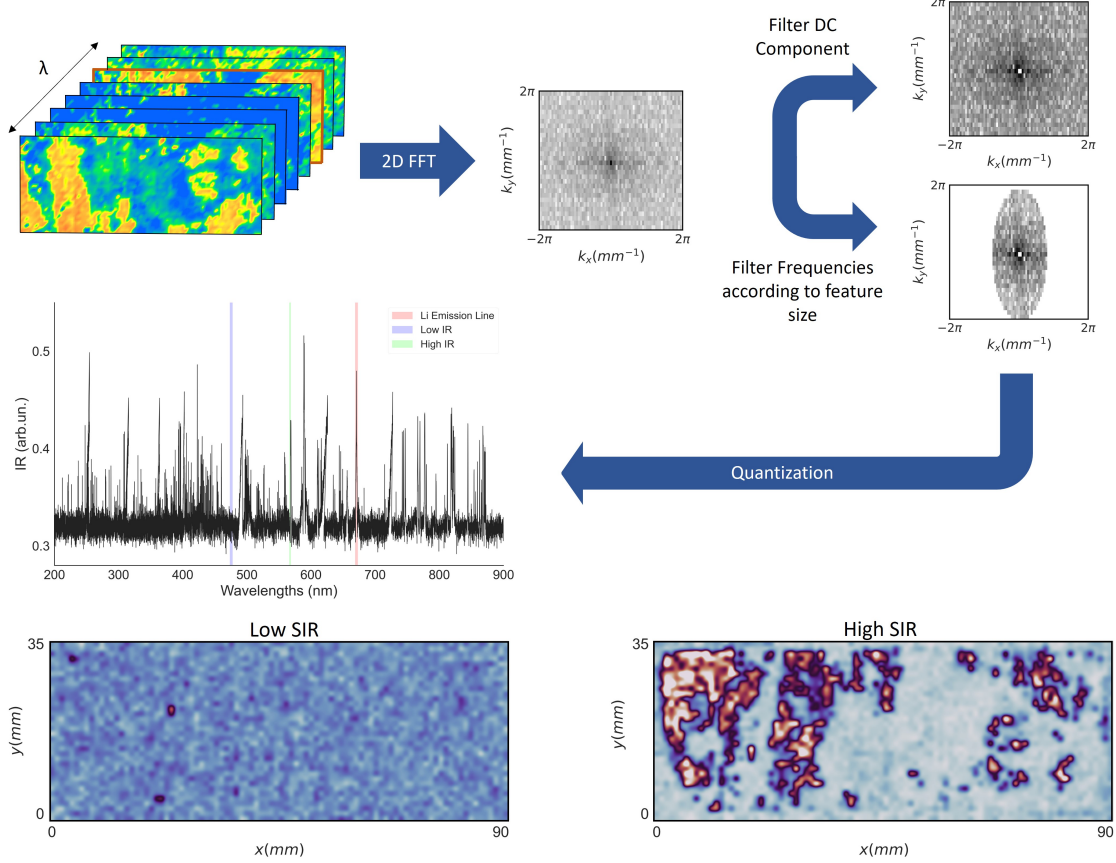


Figure 3.3: Schematic of the developed pipeline, where for each wavelength, the Fourier Transform is computed, filtering spatial frequencies in accordance to the defined feature size δ , followed by a quantization step, where the the sum of all filtered components is normalized to the sum of the total components 3.2 (here, only the DC component is filtered)

To operationalize this idea we developed and deployed an algorithm in Python using the Numpy library for mathematical operations and array manipulation. Additionally, to allow faster and scalable performances when handling big datasets, we implemented this algorithm in a manner that takes advantage of multiprocessing capabilities using the mpi4py python library, which explores the MPI framework and enables arbitrary scaling in distributed computing paradigms.

In conceptual terms, the algorithm performs as follows:

1. For each wavelength of the cube dataset, we perform the 2D Fast Fourier Transform (2D FFT) in the transverse directions xy ;
2. For each wavelength, we filter the resulting frequency components depending on the size limitations of the structures we want to identify. In particular, considering we are looking for structures with sizes between δ_{min} to δ_{max} we take the bandpass filter $k_{min} > k > k_{max}$ where $k_{min} = 2\pi/\delta_{max}$ and $k_{max} = 2\pi/\delta_{min}$;
3. For each wavelength, we compute a metric by taking the sum of the magnitude of the remaining components and normalizing it to the sum of all components before filtering, i.e.

$$SIR_{\lambda} = \frac{\iint_{Filtered} \mathcal{F}(I_{\lambda}, k_x, k_y) dk_x dk_y}{\iint_{Total} \mathcal{F}(I_{\lambda}, k_x, k_y) dk_x dk_y}, \quad (3.2)$$

where SIR stands for Spatial Information Ratio, $\mathcal{F}(I_{\lambda}, k_x, k_y)$ indicates the Fourier Transform, and I_{λ} the spatial distribution of intensities at λ .

At the end of this process we are left with a spectrum-like graph as seen in Figure 3.4A that depending on the SIR score shows if a given wavelength display structured spatial information within the target size or not. Thus, it allows to discriminate wavelengths that contain structured maps from those that only contain noise - i.e. an automatic map finding algorithm.

3.3 Results and Discussion

To test the capabilities of our algorithm we first used a rock sample constituted by Petalite, Spodumene and Albite (see Figure 3.4C), where both Petalite and Spodume are two minerals of similar composition (see Table 3.1), lithium, aluminum, silicon and oxygen, at different proportions, introducing some complexity in the analysis

Mineral	Chemical Composition
Albite	$NaAlSi_3O_8$
Petalite	$LiAlSi_4O_{10}$
Spodume	$LiAlSi_2O_6$

Table 3.1: Chemical Composition of the minerals present in the analysed rock sample.

The experimental LIBS Imaging setup (see Table 2.1) was composed of a Quantel Q-Switched Nd:YAG laser, emitting at the fundamental wavelength (1064 nm), achieving a

pulse length of 8 *ns* and 211*mJ* of maximum pulse energy. The laser pulses are steered using a single mirror and focused on the sample surface using a plano-convex lens with a 200 *mm* focal distance, creating a laser spot size at the focus point of about 300 μm . The instrumentation responsible for spectral acquisition is an 8 channel Avantes CCD spectrometer, triggered by the laser pulse, with light being guided to each channel using optical fibers and 8 individual lens collimators. To move the sample in the transverse plane, the system features two motorized sample stages that allow automatic movement in the *xy* plane, along with manual control of the *z*-axis and is used prior to the acquisition to place the sample within the Rayleigh range of the system (approximately 13 *mm*). For this particular case study, we used a laser pulse energy of 48 *mJ*, operating the spectrometers with a gate delay of 1.3 μs , and an integration time of 1.05*ms*. Additionally the resolution was set to 1 *mm/pixel* with a scanning area of 50 *mm* by 110 *mm*.

By applying the described algorithm we can see in Figure 3.4A that some wavelengths feature higher metric *SIR* values, which should indicate that the spectral region (being an emission line or not) is in fact important information in the chosen range 2*mm* to 30*mm*. This is confirmed when we select the corresponding lines with high *SIR* and examine their spatial distribution in Figure 3.4D and E, where we can clearly see the different regions that compose our sample, namely the higher intensity of the lithium emission line at 670.76 nm or the sodium line at 588.55nm. Moreover, using qualitative criteria, it is easy to see from the shape of the sodium/lithium regions that this simple LIBS imaging analysis allows us to easily differentiate the spodumene/petalite regions from the albite ones.

The relevance of the algorithm is even greater when we consider cases where there is no prior knowledge of our sample composition, and as such, informed guesses can not be made as to what emission lines could have relevant information.

Furthermore, it can also help for the case of certain emission lines with relatively low intensity, which could often be mistaken as noise and rapidly overlooked, increasing its relevance as can be seen in Figure 3.5. Here we see that an emission line corresponding to Sodium and that has a very small intensity count that could easily be overlooked if analysed by conventional methods, displays an high *SIR* score using this metric, indicating the presence of structures that can be of great importance for the remaining analysis. This is in fact confirmed by looking at the LIBS spectral map, where we see high intensities of sodium in the regions of the rock that correspond to the albite mineral, as it is expected.

Finally, and before closing the discussion, we must also comment some results on the

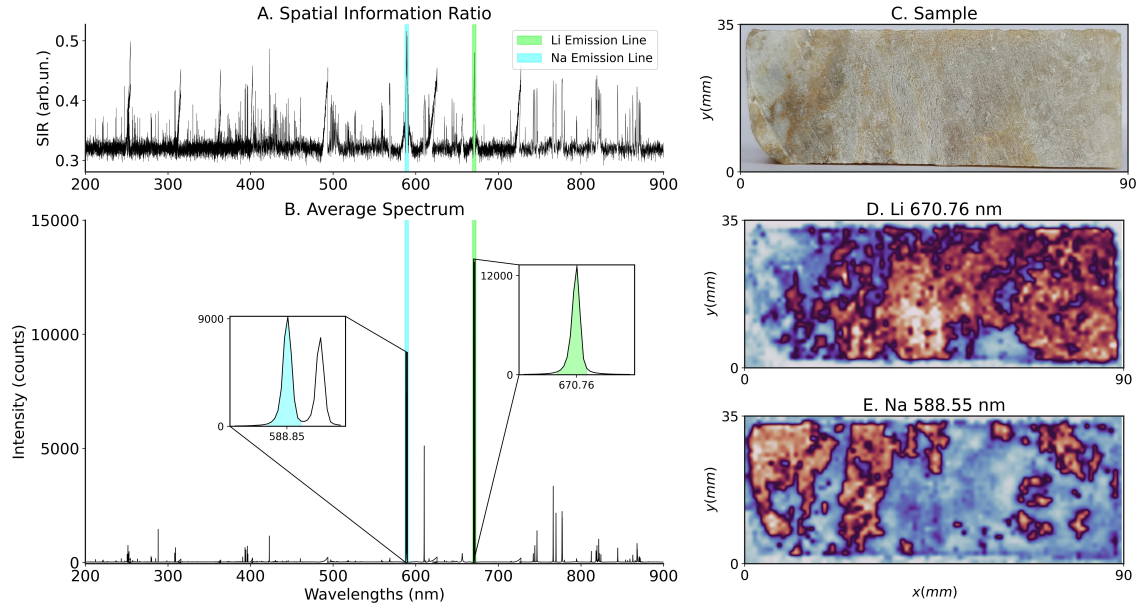


Figure 3.4: A. Graph of the FFT Score as a function of the wavelengths. B. Average Spectrum with two of the most prominent peaks shown C. Image of the sample being analysed. D. and E. Lithium and Sodium maps generated by the selected emission line in figure A, respectively.

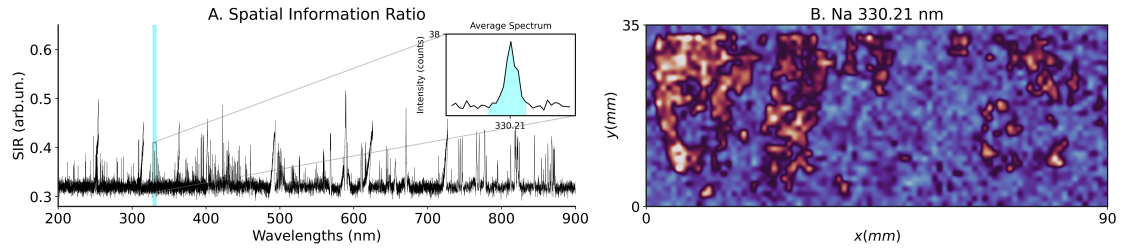


Figure 3.5: A. Graph of the FFT Score as a function of the wavelengths, highlighting a high score line corresponding to Sodium, with the respective plot of the average spectrum. B. LIBS image generated from the highlighted emission line.

computational performance of this algorithm. Indeed, since this implementation is computationally heavy, we developed and deployed an algorithm with careful performance considerations, introducing multiprocessing capabilities to enable us to run the algorithm in multiple parallel environments, from multi-core desktops to massively distributed computing infrastructures. Thus, we can effectively decrease our run-times, which can be of particular relevance when we are exploring high-resolution datasets. Our strategy was to simply divide the workload across the available workers, processing multiple wavelengths at a time, instead of processing them one by one using the MPI interface for Python. To test the effectiveness of this approach we ran the algorithm multiple times, each time varying the number of processes (N), or equivalently, the number of wavelengths that are being processed at each iteration.

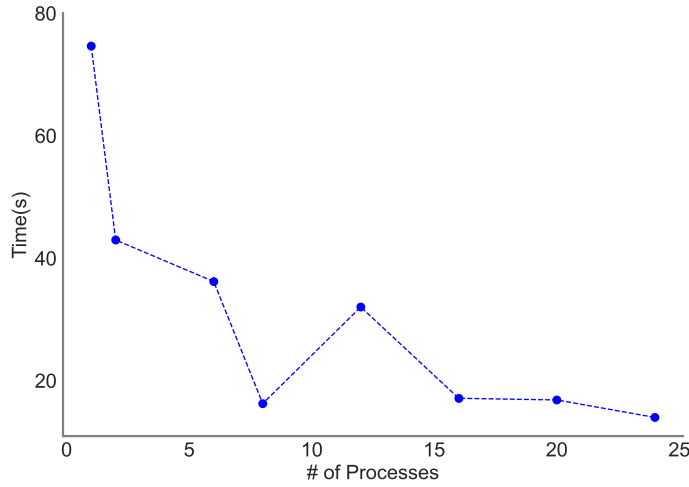


Figure 3.6: Graph of the time as a function of logical processes that are running the algorithm.

To test this we used the same dataset as of Figure 3.4 (50x110x16384 points), and run the algorithm on a desktop computer with an intel CPU i9-12900 with 16 cores and 24 threads and with 32GB of RAM at 2666 MHz. The results obtained can be seen in Figure 3.6 showing that in fact, the time to run the algorithm does exhibit the expected behavior as it decreases with the number of processes. Yet, we call the attention to the fact that for the specific case, using 8 processes can be as efficient as using 24 processes. This can have multiple explanations, from CPU throttling, to RAM limitation, etc..

Proper benchmarks would require high-performance computing (HPC) systems, yet, due to time constraints, were left out of the scope of the thesis.

3.4 Overview

In this chapter we described the development and implementation of an algorithm based on the fast fourier transform, that allows the automatic identification of emission lines and spectral regions that are of interest for LIBS Imaging analysis. By exploring a principle inspired by the JPEG compression algorithm, we have applied a bandpass filter to the frequency representation of our LIBS maps, in such a way that it allowed us to establish a metric for the information contained within a specific spatial scale range for each emission wavelength. The results further demonstrate that using this algorithm, the identification of maps of interest becomes much simpler and can be used to identify elements that have emission lines of similar intensity level to the noise, allowing us to synthetically increase our

signal-to-noise ratio. Additionally, we tested the implementation of this algorithm using a multiprocessing approach, which showed signs of being highly relevant to speed up the run times.

Given the promising results of this algorithm, future works around it could include the spectral compression of LIBS Imaging as for most application, a very small amount of the spectrum is of actual interest, as well as subsequent studies of the multiprocessing approach using HPC systems in an attempt to achieve real-time LIBS imaging analysis.

Chapter 4

Interactive three-dimensional chemical element maps with Laser-Induced Breakdown Spectroscopy and photogrammetry

After finding the relevant maps, the natural process is to display and analyze them using standard visualization tools. Yet, these are often limited to two-dimensional visualization solutions, which may be limited in terms of an interactive and contextual interpretation that can be helpful in some situations. In this chapter, we describe the development and implementation of a novel processing pipeline and workflow that makes use of dedicated photogrammetry and visualization software to easily deploy interactive three-dimensional visualization models. The results not only demonstrate the versatility of this approach but also pave for the development of novel spectral imaging diagnostic strategies that are able to deliver better qualitative analysis and insights in the three-dimensional space.

4.1 Towards interactive visualization tools for spectral imaging

While there are several two-dimensional approaches to handle spectral imaging models and images, solutions that are able to deliver interactive three-dimensional visualizations remain, to our best knowledge, underexplored in the literature. Indeed, developing a solution capable of incorporating multimodal spectral imaging data into three-dimensional models constitutes an interesting task that could lead to a plethora of new opportunities,

ranging from developing interactive teaching environments for biosciences and geology [51, 52] to providing innovative diagnostic tools for quality control of mechanical parts [53], just to name a few.

Following this reasoning, an immediate solution would be to project the maps generated by LIBS into their respective three-dimensional models. For that, it is necessary to have access to 3D models of our sample as well as develop tools capable of aligning both the generated map and the models.

4.2 Generating Three-dimensional Models

To implement a three-dimensional visualization solution, one must first create an interactive model of the sample using a three-dimensional modeling approach before merging them with a two-dimensional elemental map. For that task, many three-dimensional scanning techniques are commercially available these days with varying levels of detail and scanning precision, most of them using optical sensors for non-invasive measurement. These sensors can be divided into two broad categories: active and passive sensors. Active sensors use emitted radiation and may provide better resolution by directly gathering 3D position information, but they often require expensive hardware. Passive sensors on the other hand rely on the ambient illumination conditions and make use of acquired image data, which is then processed to infer 3D information. This constitutes the working concept of photogrammetry, which will be the main focus during this chapter.

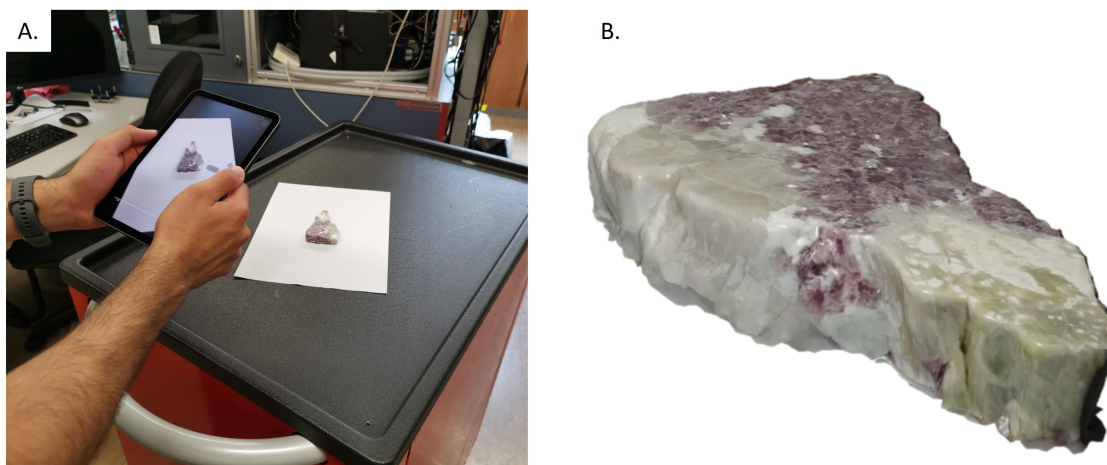


Figure 4.1: A. Process of photographing the object from various lines of sight. B. Generated 3D model of a rock sample. (Adapted from [28])

In general terms, photogrammetry is a three-dimensional coordinate measuring technique that uses multiple photographs as the fundamental medium for metrology. By using triangulation, photogrammetry analyses photographs from different locations called lines of sight, and finds overlapping points to precisely determine their relative positions. This allows the creation of 3D models of objects and/or terrain (see Figure 4.1). One of the biggest advantages of photogrammetry is that it can be performed with very basic setups such as a single digital camera, which makes it the most accessible solution to deploy an alternative for the creation of three-dimensional models starting from the centimeter range [54, 55]. Furthermore, by making use of photographs, the technique has the potential to use color information, thus creating realistic three-dimensional models of a given sample while preserving its visual features. This increases its potential for an effective merging with LIBS data.

4.3 Mesh Matching Algorithm

Having the generated 3D model and aiming to combine it with the spectral imaging data, we need to perform a crucial step of aligning both images. While being a non-trivial task, it can be done in multiple ways. Our proposed strategy is to identify coinciding features in both the three-dimensional model and the spectral image to obtain a suitable transformation that allows the alignment of both images. This requires the correct identification of three transformations: (i) Translation, (ii) Rotation, and (iii) Scaling. In short, it boils down to a point alignment problem in three-dimensional space which is commonly found in areas such as astronomy, computer vision, and molecular chemistry. One possible approach involves a least square estimation of transformation parameters between two point patterns suggested by Shingi Umeyama and commonly known today as the Kabsch-Umeyama algorithm [56].

4.3.1 Kabsch-Umeyama Algorithm

The starting point of the algorithm are the previously selected two point clouds, $\mathbf{P}$ and $\mathbf{Q}$, which we intend to align. To simplify, we can construct two matrices to store these sets of points as

$$\bar{\mathbf{P}} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix}, \quad \text{and} \quad \bar{\mathbf{Q}} = \begin{bmatrix} i_1 & j_1 & k_1 \\ i_2 & j_2 & k_2 \\ \vdots & \vdots & \vdots \\ i_n & j_n & k_n \end{bmatrix}. \quad (4.1)$$

The next step is to compute some metrics for these datasets, namely the centroid and variance given by

$$\mu_B = \frac{1}{n} \sum_{l=1}^n \mathbf{b}_l \quad (4.2)$$

$$\sigma_B^2 = \frac{1}{n} \sum_{l=1}^n \|\mathbf{b}_l - \mu_B\|^2, \quad (4.3)$$

respectively. Then, for the point clouds $\bar{\mathbf{P}}$ and $\bar{\mathbf{Q}}$, we can define the cross-covariance matrix $\bar{\mathbf{H}}$ as

$$\bar{\mathbf{H}} = \frac{1}{n} \sum_{l=1}^n (\mathbf{p}_l - \mu_{\bar{\mathbf{P}}})^T (\mathbf{q}_l - \mu_{\bar{\mathbf{Q}}}), \quad (4.4)$$

with $\mathbf{p}_l = (x_l, y_l, z_l)$ and $\mathbf{q}_l = (i_l, j_l, z_l)$ corresponding to points in the distinct meshes. This cross-covariance can then be used to compute its singular value decomposition $\bar{\mathbf{H}} = \bar{\mathbf{U}} \bar{\mathbf{D}} \bar{\mathbf{V}}^T$.

It is then possible to prove[56] that the optimal transformation parameters for the alignment of $\bar{\mathbf{P}}$ and $\bar{\mathbf{Q}}$ can be obtained by:

$$\text{Rotation: } \bar{\mathbf{R}} = \bar{\mathbf{U}} \bar{\mathbf{C}} \bar{\mathbf{V}}^T \quad (4.5)$$

$$\text{Scaling: } \bar{\mathbf{S}} = \frac{1}{\sigma_{\bar{\mathbf{P}}^2}} \text{tr}(\bar{\mathbf{D}} \bar{\mathbf{C}}) \quad (4.6)$$

$$\text{Translation: } \bar{\mathbf{T}} = \mu_{\bar{\mathbf{Q}}} - \bar{\mathbf{S}} \bar{\mathbf{R}} \mu_{\bar{\mathbf{P}}}, \quad (4.7)$$

where $\bar{\mathbf{C}}$ is defined as $\bar{\mathbf{C}} = \text{diag}(1, \dots, 1, d)$ with $d = \text{sign}(\|\mathbf{U}\| \|\mathbf{V}^T\|)$ in order to prevent reflections.

Using this procedure, an automatic algorithm was easily implemented in Python and directly used for the visualization procedure. Additionally it is important to note that the deployed algorithm is dependent on the order of the points that make up both matrices, meaning that for example point $\mathbf{p}_1$ must have a correspondence with point $\mathbf{q}_1$. In order to facilitate the feature selection process and drop this requirement, we can make this transformation for all point permutations in $\bar{\mathbf{Q}}$ and find the one with the best figure of merit (FOM), defined as the sum of the Manhattan Distances,

$$FOM = \sum_{i=1}^n |\mathbf{q}'_i - \mathbf{p}_i|, \text{ with } \mathbf{q}'_i = (\bar{\mathbf{T}} + \bar{\mathbf{S}} \bar{\mathbf{R}}) \mathbf{q}_i. \quad (4.8)$$

4.4 Workflow

Now that we have the algorithm set, we need to choose convenient software tools to implement the visualization procedure. For this work, we decided to use ParaView software [57] which provides a graphical user interface that researchers can use to explore three-dimensional datasets and further allows the visualization of big data via hardware acceleration and distributed computing approaches. In short, it simplifies tasks by allowing the user to perform a visual analysis of the dataset by creating and applying filters to the data or subsequent filter results. To allow automation of tasks, ParaView also supports the execution of processing scripts using Python which facilitates batch mode operation and will be instrumental for this work.

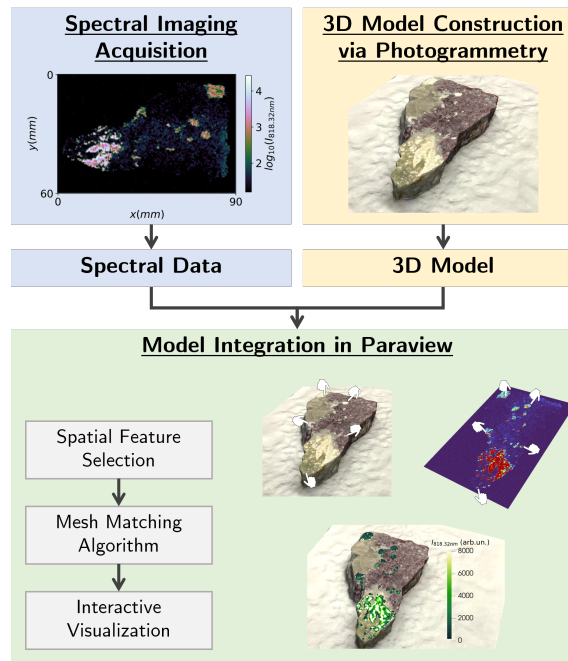


Figure 4.2: Graphical description of the visualization workflow. (Adapted from [28])

To allow interactive operation in ParaView, we need to complete three steps of a visualization workflow (see Figure 4.2). In the first step, we need to process the data obtained from the LIBS system and generate the associated chemical maps in a format that can be loaded into ParaView. ParaView is laid on top of the Visualization Toolkit VTK [58] which defines the preferable format of the input data so that it can be loaded seamlessly into the software. Therefore, after necessary preprocessing and 2D map generation, we create a two-dimensional structured grid to store the generated matrix as. All these steps are performed in Python and can be easily implemented by making use of the library to handle VTK native form. In the following step, the user scans the object using a scanning device,

with post processing being performed with proper software tools (we used a standard iPad Mini 6 together with the Polycam software) in order to create a three dimensional model. The data can then be stored, preferably in the obj format, with the accompanying texture file in the jpeg format. This method allows us to obtain not only the geometry and mesh, but also the image of the sample, resulting in a realistic three-dimensional representation of the sample.

Having both models in adequate formats natively supported by ParaView, they can now be loaded into the the visualization software where the last step is now performed. At this stage what it left to do is select matching features on both models using the Interactive Select points function of Paraview, and apply the algorithm described in section 4.3 to find the suitable set of transformation parameters that coincide the selected features in the spectral map and the three-dimensional model. Lastly we can perform the projection of the chemical map onto the three-dimensional model, resulting in an interactive visualization environment.

4.5 Results

Results of the implemented algorithm are depicted in Figure 4.3 showing the typical outcome of the combination of photogrammetry with spectral imaging techniques. Furthermore, in order to better understand the potential of the tools developed, we present three illustrative scenarios.

First, Figure 4.3A represents the chemical map for the distribution of Cobalt generated with LIBS imaging on top of a surface of a Portuguese tile. As can be seen, the algorithm performs a correct projection in this planar configuration. In this case, we can argue that the three-dimensional visualization promotes a deeper comprehension of the connections between the physical variables measured and the sample attributes, which could turn out to be a valuable resource to enhance heritage analysis, in particular by supporting the development of diagnostic tools that are able to integrate multiple sources of data in more efficient manners.

Figure 4.3B presents the chemical map for the distribution of a silicon film on top of a cork. Here we see that the algorithm is still robust when applied to cylindrical topologies, thus introducing non-trivial advantages for the analysis and diagnostics of non-planar geometries for academic and industrial purposes, allowing better identification of patterns. In this specific case, the heterogeneity of the coating alone may indicate an equipment

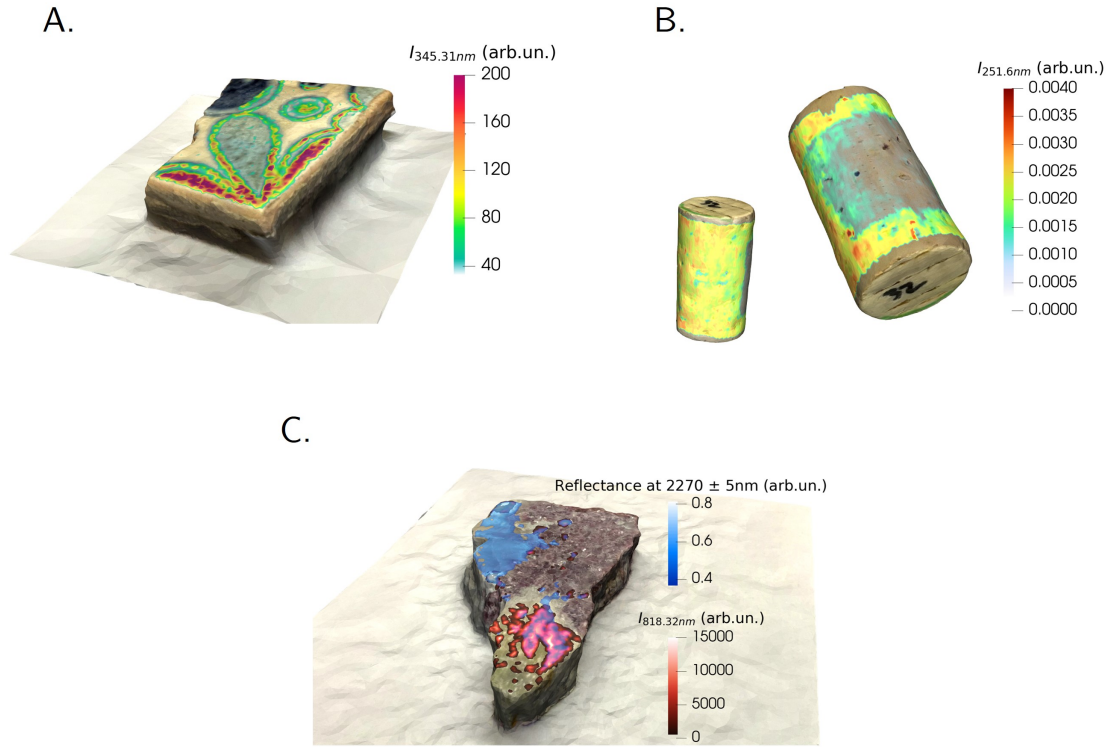


Figure 4.3: Illustrative results of the potential of the approach for the generation of interactive visualization models, here applied to (A.) Heritage applications (Distribution of Cobalt on the surface of a Portuguese Tile, using the Cobalt line at $345.3nm$), (B.) non-planar geometries (distribution of silicon on a cork, using the Silicon line at $251.6nm$), and (C.) multimodal sensing (LIBS and hyperspectral data of a lepidolite-albite rock). (Adapted from [28])

malfunction. Yet, by adding information regarding the shape and three-dimensional modeling, it is also possible to better debug the issue and identify the cause by analyzing the fabrication pipeline.

Finally, Figure 4.3C is the projection of two distinct spectral imaging techniques, LIBS and Hyperspectral Imaging for the same rock sample. This particular example shows both the spatial distribution of Potassium (K) obtained with LIBS imaging, together with the reflectance values of hyperspectral imaging in the $2270 \pm 5 nm$ band. The hyperspectral data is able to correctly distinguish the Quartz and Albite minerals (higher reflectance) from the Lepidolite ones. Then LIBS imaging can be used to perform further distinction of these regions by looking at the sodium spatial distribution (Albite is rich in Sodium). These regions are easily identified by the overlap of the spectral maps.

In our opinion, this simple example illustrates well the versatility of this approach to deal with data obtained from multimodal sensing sources, thus paving the way for new

opportunities to explore these in a collaborative manner.

4.6 Overview

In this chapter, we addressed the development of a 3D visualization pipeline, as well as the techniques necessary for the integration of two-dimensional spectral imaging data onto three-dimensional models. Using the point-cloud matching Kabsch-Umeyama algorithm we were able to deploy a solution for the projection of spectral imaging data on three-dimensional models based on the selection of multiple features using the ParaView software, a software used by researchers for interactive exploration and data analysis.

Based on the illustrative examples provided, we argue that extending the spectral analysis to the 3D spatial domain allows a deeper understanding of the relations between the various physical quantities measured and the properties of the sample surface by offering a contextual interpretation in either planar or non-planar geometries. In this way, this approach to data visualization and analysis provides non-trivial opportunities for immersive visualization and analysis, which can find applications in a wide range of subjects, including quality and manufacturing control for industry and interactive teaching in geosciences and biosciences, to name a few. Furthermore, our approach has been designed to be flexible and user-friendly, allowing for the integration of data from multiple multimodal sensing systems. Examples include maps displaying hyperspectral or fluorescence data[59, 60].

Having into consideration the benefits of this approach, future works could be done by integrating other techniques such as X-Ray computed tomography[61, 62], which may open up new opportunities for sensor fusion and broaden the interpretability of the gathered data. Its adaptability makes it simple to apply the approach to a wide range of spectroscopy applications beyond LIBS imaging. In addition, this pipeline could be a stepping stone for the creation of mixed-reality environments that could be applied for example in industrial applications for quality control, or even scientific applications in an attempt to create a virtual lab environment.

Chapter 5

Collaborative Opportunities in Multimodal Sensor Fusion

Although LIBS is a well-established technique for chemical elemental mapping, it can still lack sensitivity and specificity to the molecular structures present in the sample. In this context, the use of LIBS in collaboration with other spectroscopic techniques, such as Raman spectroscopy, can enhance its analytical capabilities and provide complementary information. In particular, one of the promising solutions is to combine LIBS with hyperspectral (NIR, SWIR) imaging (HSI). In this chapter, we start with a brief overview of the hyperspectral technique and proceed to explore two different approaches to the fusion of LIBS and HSI for the case study of the unsupervised classification of minerals in a rock sample. On one hand, a traditional data fusion approach that combines the data from each technique in an attempt to increase the amount of information and provide better classification performances. On the other hand, an innovative knowledge distillation approach that uses higher performance and knowledge of LIBS to train the HSI, hoping to enhance its overall capabilities.

5.1 The potential of Hyperspectral Imaging

In general terms, HSI is a technique based on the analysis of the reflectance spectra of a sample surface, from the visible to the infrared range. The captured data allows us to establish a relation between the presence of specific sample constituents using their optical properties at distinct wavelengths, such as absorption, emission, reflection, and scattering. These properties manifest themselves as absorption bands in the spectra and analyzing the

slope and depth of specific regions of the spectra, we may infer on the presence of some important molecules in the sample. The biggest advantage of HSI is that it can acquire data much faster than LIBS as it can operate in line-scan mode (e.g. 400 points in 5ms), with a sampling rate of up to 100KHz allowing for higher throughputs that make it suitable for industrial applications. Yet, HSI classification and distinction capabilities can often be underwhelming when compared with LIBS, in particular due to the fact that its spectra are often severely affected by noise (e.g. due to surface roughness) or convoluted with other information. In this context, there is a window of opportunity for the design of novel technological approaches to enhance the capabilities of this technique.

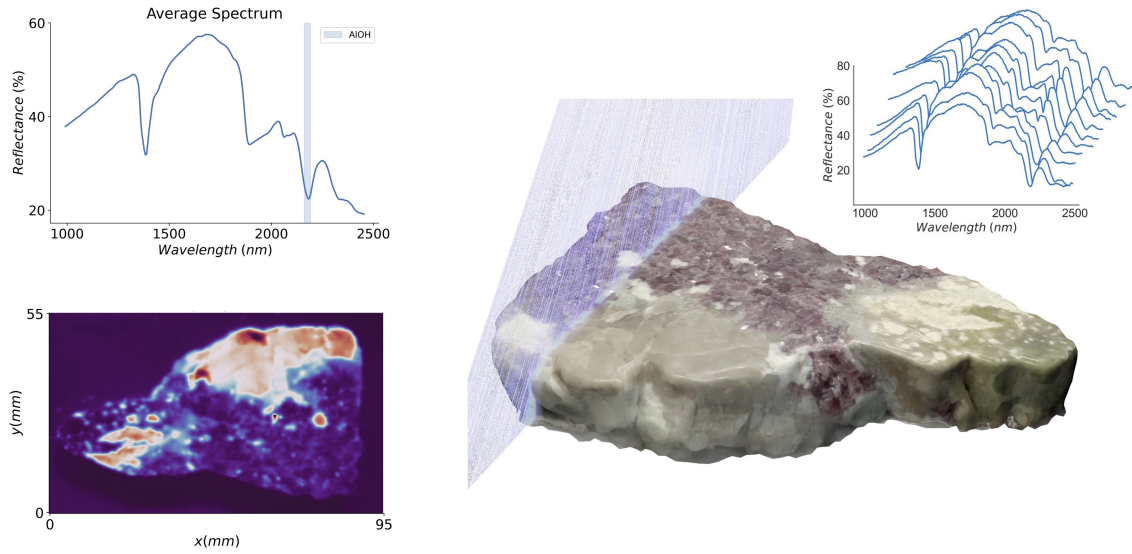


Figure 5.1: Illustration of the principle HSI. Here we can see the push broom scanning principle of this technique that scan the sample line by line, as well as the spectra that are captured at each acquisition.

5.2 Methodologies for Collaborative Sensing

To explore the opportunities of collaborative sensing for the HSI-LIBS tandem solution we will focus our attention on two distinct strategies: a traditional Data Fusion approach, and an innovative Knowledge-Distillation. Each of the techniques focuses on different goals to optimize the standalone performances of the techniques, as will be reported in the subsequent sections.

5.2.1 Data Fusion

Data fusion is the process of combining data from various modalities in an attempt to increase the information used to perform decision and obtain better output decisions. This can be done mainly in three different ways: low-level fusion, where the raw pre-processed data is combined, mid-level fusion combining the extracted features from each modality, and a high-level fusion or decision-level fusion that aggregates the outputs of multiple models using methods as majority voting and averaging. For this case in particular we intend to leverage the complementary information from each technique to effectively extend the feature space and so mid-level fusion will be mainly explored-

Our approach begins with a crucial step of aligning the spatial datasets from each modality. We select a set of matching points in each map and apply the Kabsch-Umeyama algorithm to find the optimal transformation parameters for translating, rotating, and scaling the datasets [28]. Moreover, to achieve the point-to-point match required for fusion, we use the lowest resolution of LIBS imaging as the spatial mesh and associate each point with the nearest point in the HSI imaging.

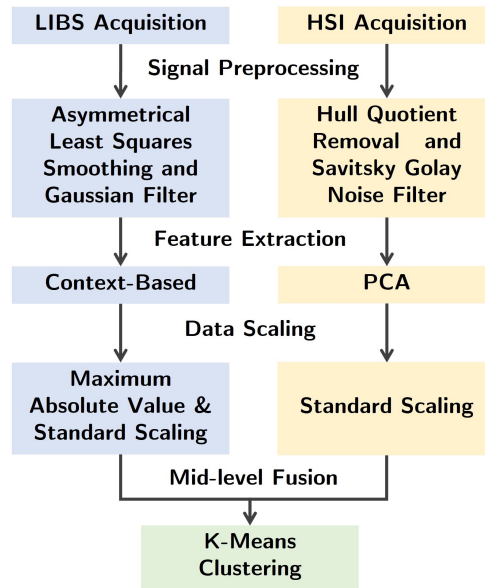


Figure 5.2: Schematic of the adopted pipeline for sensor fusion classification

In the second stage, each technique undergoes appropriate pre-processing, feature extraction, and scaling before the mid-level fusion. For LIBS imaging (see Figure ??), we remove the background from the obtained signal using a standard Asymmetrical Least Squares Smoothing algorithm [30]. This consists of continuous components from Bremsstrahlung and recombination processes that affect the emission lines non-uniformly. Additionally, we

apply a Gaussian Filter to reduce the impact of potential contaminations and edge effects. After pre-processing, we select features using a context-based approach that chooses wavelengths of interest based on our prior knowledge of possible elements in the sample (more details on the technique can be found in ref.[?]). Finally, we scale each feature to its maximum absolute value, preserving the distribution shape and preventing unequal weights from different modalities.

For HSI, we normalize the raw spectra from the camera to a white reference to obtain reflectance spectra, revealing characteristic absorption features as well as characteristic background signatures, resulting from ferrous ions, water, and carbonate absorptions that are outside the spectrometer range. While it can be argued that this background signature, also called reflectance hull, may contain additional spectral information, it also distorts the spectral absorption features, making its removal a necessary step. Given that, our HSI pre-processing pipeline begins by applying a Savitsky Golay filter to remove noise (see Figure ??), followed by hull quotient correction to remove the reflectance hull. Unlike LIBS, context-based extraction is more difficult in HSI, especially because of the convoluted spectral content. Considering this, we use a conventional PCA analysis for dimensionality reduction, selecting the first four principal components that account for an explained variance ratio of 98%. We then use the scores as the extracted features, and apply standard scaling for the same reasons discussed previously.

After extracting the features for each model, we combine them in a mid-level stage fusion (feature-level fusion) into a dataset that we use to train an unsupervised classification algorithm. We choose a conventional K-means clustering, inspired by previous results for LIBS [6]. Here we train the algorithm for a reference sample that we use to interpret the results and label the clusters (assign to mineral type) before generalizing it to unseen samples. Finally, we compare the results obtained using collaborative sensing with those for the standalone techniques, i.e., trained only with the extracted features for each technique.

5.2.2 Multimodal Spectral Knowledge-Distillation

Multimodal spectral knowledge-distillation (MSKD) is based on the concept of Cross-Modal knowledge distillation (CMKD)[63], an algorithm that uses a Teacher-Student scheme to leverage the superior performance of a modality with higher knowledge (the teacher), and transfers it to a weaker modality (the Student). However CMKD is usually applied to neural networks and depends on the supervised training of the teacher modality, which

would not be suitable for the case study at hand of unsupervised classification. Having this in mind, we developed the workflow seen in Figure 5.3 where we start with the train of a classifier using a single spectral modality in an unsupervised manner, subsequently using the output of this classifier as labels to train the weaker spectral modality with a suitable supervised learning algorithm. Using this pipeline we believe that MSKD may offer non-trivial advantages for multimodal spectral imaging, in particular because it exploits the superior performance of supervised learning [64, 65] while bypassing the laborious (and often imprecise) work of hand labeling correct pixel regions [66].

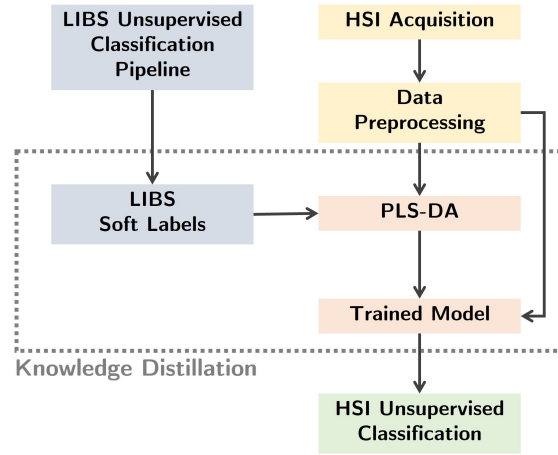


Figure 5.3: Schematic of the adopted pipeline for MSKD classification

5.3 Results and Discussion

To evaluate our collaborative sensing strategies and the proposed mineral identification pipelines, two rock samples were selected from the same mining site, both exhibiting the same mineral composition as seen in Figure 5.4.

The two samples originate from Li-rich pregmatite vein located in the Central Iberian Zone of the Iberian Massif and from previous studies it has been shown that they are mostly composed of four mineral types, Lepidolite, Quartz, Albite, and Mica with chemical compositions that can be seen in Table 5.1

5.3.1 Data Fusion

To test the results obtained in data fusion, we start by analysing the spectral signatures of each mineral regions both in LIBS and HSI. As expected, the LIBS spectrum features variations in emission line intensities that correspond to elements presents in each mineral,

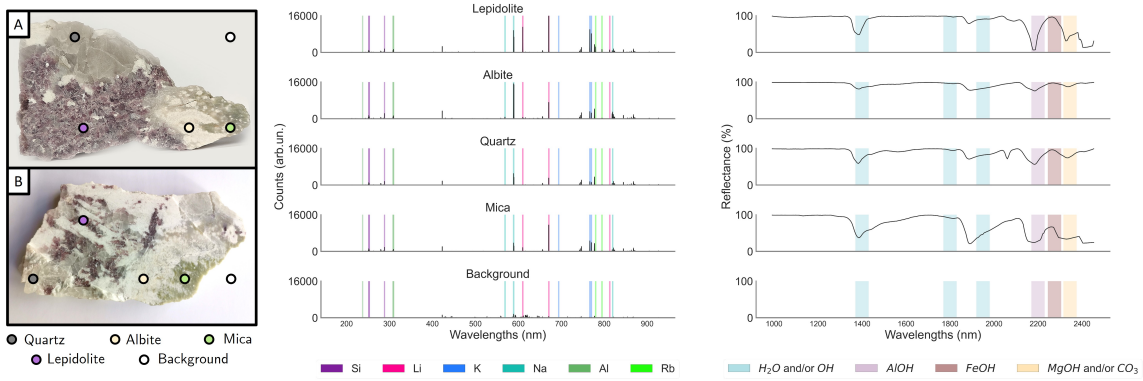


Figure 5.4: A. Rock samples and indicative mineral zones. Spectral signature for B. LIBS and C. HSI, showcasing the typical signature of distinct mineral types.

Mineral	Chemical Composition
Albite	$NaAlSi_3O_8$
Lepidolite	$KLi_2AlSi_4O_{10}F(OH)_2$
Quartz	SiO_2
Mica	$KAl_2(AlSi_3O_{10})(OH)_2$

Table 5.1: Chemical Composition of the four most relevant minerals observed in the samples.

mainly the ones previously identified in Table 5.1, and as such, we opted to extract the three emission lines for each element in a process of context-based feature extraction.

Now looking at HSI we notice that although the spectrums do significantly vary between mineral regions, they are no longer associated to specific emission lines, but rather with distinct dips and bends of the reflectance curves. This inhibits us from extracting features manually and in this context we opted to use a conventional PCA (principal component analysis) algorithm for dimensionality reduction that transforms the original feature space, to a new one of reduced dimensionality, while preserving as much variance as possible. The most important step in this method is to choose the amount of principal components to be used and in this case it is done by analysing the scree-plot of the explained variance, and setting set a threshold at 98%, which was achieved using 4 components.

Having the knowledge on how we expect each technique to provide differentiation of mineral regions, and having performed feature extraction, we can advance to deploy the mineral classification algorithm. Due to simplicity of operation and good results in literature [67, 68] we opted to train a K-means unsupervised clustering algorithm, comparing the results obtained for each standalone technique to the collaborative approach in two stages, training and testing.

For the training stage, we started by choosing a suitable training set and we opted for the region seen in Figure 5.6A, as a results of the mineral regions appearing to be better defined which is instrumental to better interpret the results qualitatively. Added to this, the dataset seems to be balanced in terms of mineral area, preventing the appearance of unwanted bias. Following the selection of the training set, K-means requires us to choose the number of regions to be used in training and in an attempt to automate the process as much as possible, we tried several approaches such as the elbow method, the silhouette score and prediction strength but the results were not conclusive, only being able to identify the amount of clusters to be within a certain range (between 4 and 6 for this particular case). We then decided to combine our prior expert analysis of the sample, confirming the presence of four major mineral regions plus the background, thus totaling five clusters.

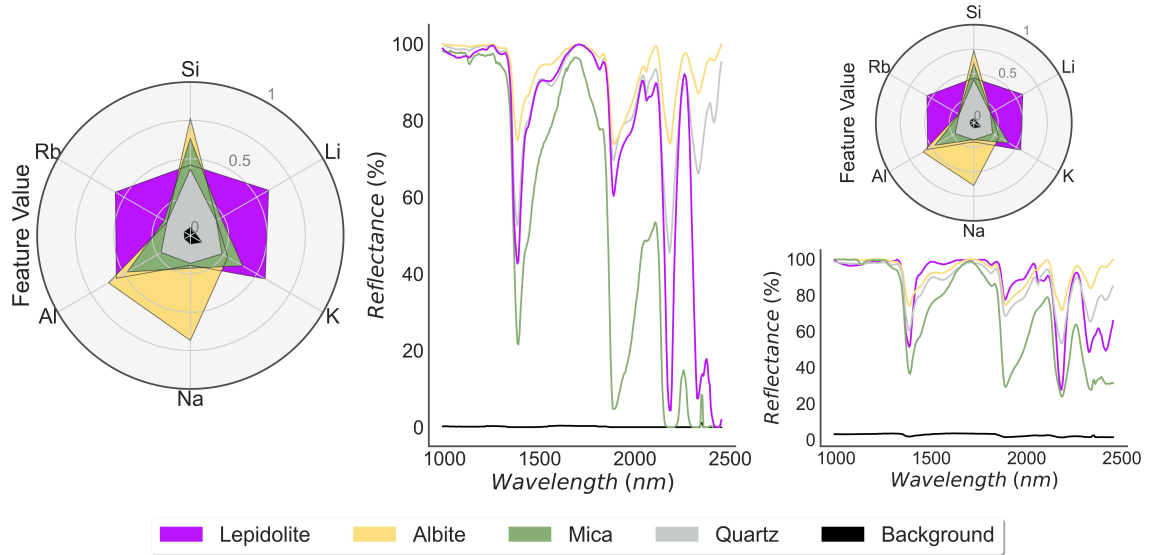


Figure 5.5: Perceptual map exhibiting the average LIBS feature value and average HSI spectrum for each of the identified clusters.

It is important to note at this stage that the identified clusters of the K-Means algorithm, only group surface zones according to their similarity of features and is not able to provide of prediction of what mineral type each region corresponds to, and a label assignment stage is then required. To achieve that, it is possible to proceed with a user interpretation of the results at the end of training. An ingenious way of doing that is to calculate the centroids in the feature space and use informative graph formats as seen in Figure 5.5 to manually assigning the clusters. In LIBS this can be done using the radar chart for the LIBS feature space, where it is easy to associate the cluster with the centroid with non-zero features in the elements of interest, according to table 5.1. For example, the

cluster with Li and Rb features shall be clearly associated with the Lepidolite mineral in this context. Alternatively, in HSI, one can still obtain the centroid and recover it onto the original feature space, obtaining a spectrum that can be compared against a reference one, in particular comparing to those in Figure 5.4. While this qualitative step can be trickier to perform than its LIBS counterpart, we expect it to be sufficient for a correct interpretation of some mineral types.

Following training we extend our analysis using a test sample to which we apply our trained model and test its generalization capabilities.

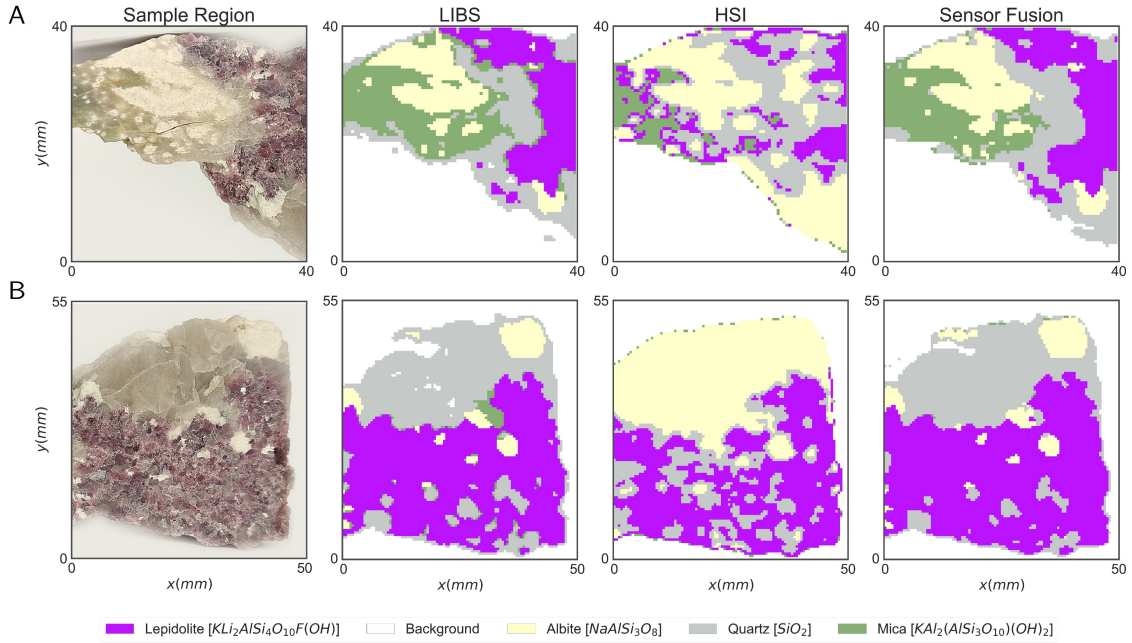


Figure 5.6: A. Classification results obtained by applying our training pipelines to a training region. B. Classification results obtained by applying our trained models to a test sample region

Starting by analysing the performances of the standalone technique as seen in Figure 5.6, it becomes evident that LIBS and HSI provide very different results. In LIBS we can see that the results are in good accordance with what is expected from an analysis of the sample figures, only noting some incorrect classification of Quartz and Mica in sample B, along with some mineral transition regions, as a result of edge effects, extensively reported in literature that are a consequence of crater size (spatial resolution), meaning that mineral boundaries often contain information of more than one mineral. In HSI on the other hand it is clear that the PCA-based method has a poor performance and isn't able to handle the complexity of hyperspectral data, as is seen by the lack of agreement of the test and training samples. Finally, the data fusion approach seems to provide a good agreement

with the sample figures, possibly overperforming the LIBS-only classification as it seems to better handle the transition regions, in addition to some previously unidentified Quartz regions now being correctly identified.

In addition to the obtained results, one must mention that the advantages of data fusion aren't limited to better classification performance but also to higher degrees of interpretability. By averaging the HSI spectra and LIBS emission lines at the cluster regions, we obtain the centroid composition (see Figure 5.5), which shows the average elemental composition using the LIBS radar chart and the average absorption spectra for each mineral type using HSI modality. This may facilitate the cluster labeling stage and enhance its interpretability, by providing more information that can be cross-checked. On another note, this approach may also compromise the versatility of these techniques, due to a more complex pipeline, while not significantly improving the results from LIBS-only, and being limited by the technique of lower resolution.

5.3.2 Multimodal Spectral Knowledge-Distillation

From the results of the previous section it has become clear that HSI doesn't perform as well as other techniques in the task of mineral identification, however this is likely not caused by poor information spectral content, but because of intermediate steps, mainly PCA. This is because PCA looks to find feature transformations that conserve variance in the original data, and while it is successful in this job, low variance components, which are frequently dropped in the process, might be superior predictors for specific assignments, often calling for a loadings analysis where each component is assigned to a specific combination of features and some are even filtered as they are related to noise and distortions present in the data. Such analysis isn't however trivial and scales with the complexity of the data, and the variance threshold ends up being the most often used metric to select the amount of components.

With this in consideration, we developed the MSKD pipeline from Figure 5.3 that uses a superior knowledge technique, in this case LIBS to train HSI. For this purpose we again opted to use the same training set as in the previous section, choosing PLS-DA as the supervised classification algorithm, as a consequence of its proven effectiveness handling hyperspectral data [69]. The choice of PLS-DA makes it so that we must identify the number of components we want to use and that is done through predictive error analysis, using the LIBS data as the ground truth, being careful to prevent overfitting. From Figure 5.7

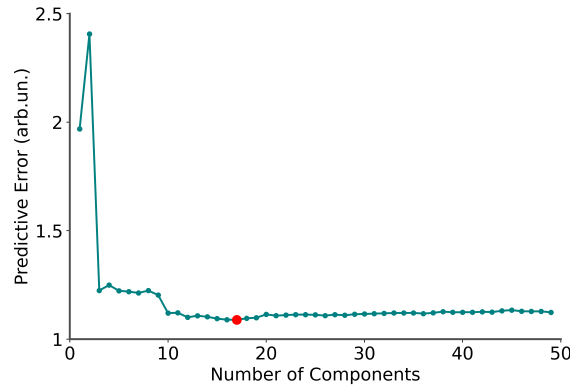


Figure 5.7: Predictive error on training the PLS-DA algorithm using the test sample as a benchmark and LIBS labels as the ground truth.

we can see that the predictive error appears to reach a minimum at around 17 components, which would up being the number of components used in our analysis.

Figure 5.8 shows the results of the training section, and although some differences can be noted comparing to the sample image, classification is on par with the one obtained from LIBS, meaning it exhibits significantly better performances than the standalone HSI approach. Based on these results, we tested the generalization performance of the model on other samples, noting that for this testing, only hyperspectral data is required, which means that from this stage forward, we are not restricted by possible limitations of another technique, and we can work at faster acquisitions speeds, using the full resolution of the hyperspectral camera.

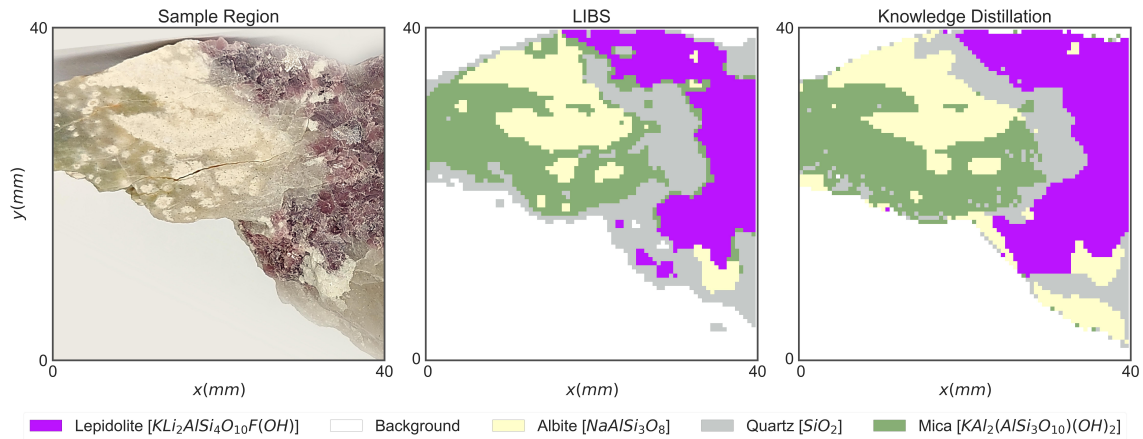


Figure 5.8: Classification results from the testing set, using Multimodal Spectral Knowledge Distillation pipeline.

The results enclosed (see Figure 5.9) in the testing sets demonstrate that the proposed method outperforms the standard PCA + KMeans pipeline, despite some minor misclassification of Quartz regions in cluttered areas of sample A and borders of sample B, showing

that in fact, this method was able to effectively transfers the LIBS knowledge to leverage the features of HSI and enhance its capabilities. Considering the versatility and higher throughput of HSI compared to LIBS, the developed model has potential for applications in various industrial and scientific domains, facilitating faster and more accurate workflows.

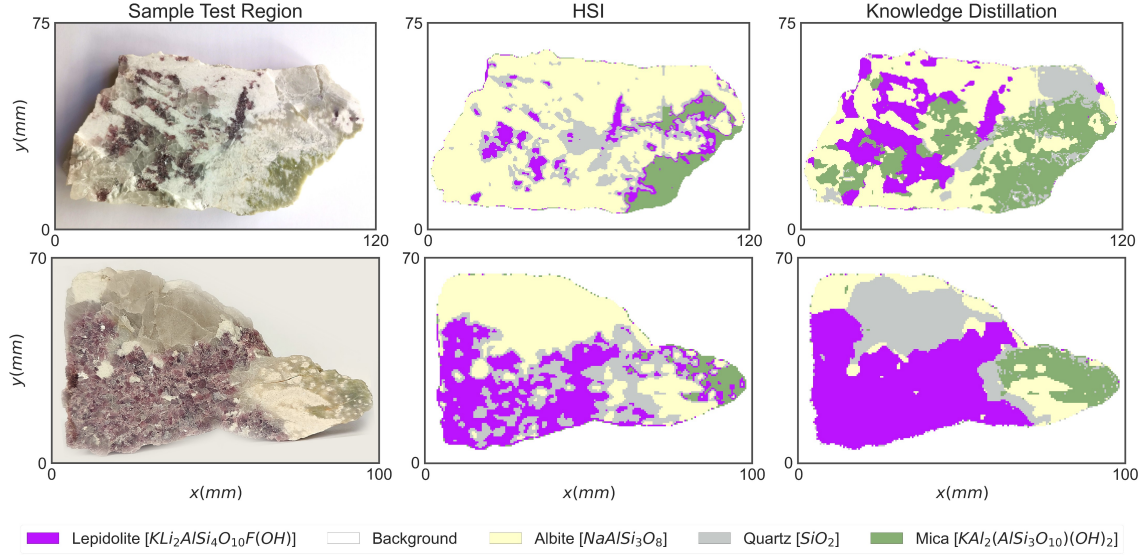


Figure 5.9: Classification results from the training set, using the trained Multimodal Spectral Knowledge Distillation model.

Given HSI's versatility and higher throughput compared to LIBS, the established learning model could be used in a variety of industrial and scientific applications speeding up workflows while granting the necessary results.

5.4 Overview

In this chapter we looked for possible synergies that arise from combining spectral imaging techniques, more specifically, LIBS and Hyperspectral imaging. For that purpose, two approaches were taken in consideration for the particular case of mineral identification: a sensor fusion approach, that combined the data of the two modalities; and a knowledge distillation approach that uses one modality to train the other one.

The results showed that when considering the standalone approaches, LIBS has a clear advantage in classification capabilities, with HSI severely lacking in comparison to it. Combining the data from both however, did slightly improve the performance when compared to LIBS and we believe that samples where LIBS struggles could highly benefit from this approach, which shall be explored in future works.

For the knowledge distillation approach, we were able to successfully train a model that takes hyperspectral data as an input which is much faster to acquire, using the LIBS technique as a supervisor during training, allowing us to obtain much better classification results of the hyperspectral technique when compared to its standalone counterpart. In future work, we plan on taking advantage of this pipeline and adopting it for industrial applications, where typically high throughput is mandatory.

Chapter 6

Conclusion

The subjects developed over the course of this thesis have been motivated by the currently underexplored capabilities of the promising LIBS-based Imaging technique, in particular the development of new advanced tools to support the currently adopted analysis pipelines was presented, aiming for a better comprehension and quality of final results.

For that purpose, three different problems related to LIBS Imaging were studied: feature finding, data visualization, and collaborative approaches. Feature finding is the process of detecting the main informative sources providing valuable details that allow for consecutive analysis to be conducted. This step is crucial in LIBS since the information is contained within the specific emission lines of each element meaning that for the majority of cases most of the spectrum can be discarded. For this purpose, the average spectrum and dimensionality reduction techniques such as PCA are often used and have become standardized within the LIBS community but lack in regards to interpretability and often disregard important information due to the signal-to-noise ratio. With this in mind, we developed an automatic map finding technique based on the Fourier Transform, giving us a metric to easily identify emission lines with spatial structures of relevance. Additionally we tested a multiprocessing implementation, to allow for faster runtimes taking a step closer to obtaining real-time analysis capabilities.

Chapter 4 was devoted to data visualization. Here we deployed a solution for the projection of the spectral imaging data on three-dimensional models, using a pipeline that combines a point alignment solution with manual spatial feature selection to find a suitable set of transformation parameters that correctly align the spectral data to the 3D models. We then argued that this approach has the potential to enhance the overall experience of data visualization and analysis as it may allow a deeper understanding of the relationships

between the various physical quantities measured and the properties of the sample surface by providing a contextual interpretation either in planar or non-planar geometries.

Lastly we explored the potential of collaborative approaches to enhance the performance of standalone analytical techniques for sample characterization. Here we focused on two techniques: laser-induced breakdown spectroscopy (LIBS) and hyperspectral imaging (HSI), where we proposed two methods to combine them: one based on data fusion, and another based on knowledge distillation. In the data fusion method, we integrated the spectral information from both techniques in a mid-level fusion, to train a classifier. We found that this method slightly improved the accuracy, while also allowing for a great degree of interpretability of the results, however, the data fusion method was constrained by the lowest resolution and smallest throughput of the two techniques. In contrast, in the multimodal spectral knowledge distillation method, we used the high accuracy of LIBS as a teacher to guide a student model that learned from HSI data. We found that this method significantly improved the accuracy and generalization of HSI classification, while maintaining its high resolution and throughput.

Perspectives and Future Works

LIBS-based imaging is a cutting-edge technique that offers fast and reliable spectral analysis in various environments and applications, such as biology, geology, industry, and remote sensing. However, there are still many challenges and opportunities to improve and extend the capabilities of this technique, from the acquisition to the interpretation of the data. In this thesis, we have proposed novel methods and tools to address some of these challenges and to demonstrate the potential of LIBS-based imaging in different scenarios. Some possible directions for future research based on our work are:

- Map finding algorithm: the algorithm could be applied to other tasks, such as data compression, where we could exploit the sparsity of the LIBS images to reduce their size and storage requirements. Moreover, this algorithm could be optimized for high-performance computing systems, where we could leverage parallelization and distributed computing techniques to achieve real-time LIBS imaging analysis.
- Data visualization: This pipeline could be a stepping stone useful for various applications beyond the academic domain, such as education, and industry. Future work

could focus on further automating this pipeline and integrating it with virtual lab environments, where users could interact with realistic simulations of LIBS experiments and learn from them.

- **Multimodal Spectral Knowledge Distillation:** This approach could be extended to other combinations of sensors and techniques, such as Raman spectroscopy, X-ray fluorescence, or infrared spectroscopy. Furthermore, this approach could be particularly beneficial for remote sensing applications, where HSI is widely used but often suffers from low interpretability and atmospheric interferences. By using ground measurements of LIBS or other techniques as teachers, we could enhance the performance and capabilities of remote sensors, which could have a great impact on various domains, such as environmental monitoring, agriculture, or security.

Bibliography

- [1] F. Brech, Optical microemission stimulated by a ruby laser, *Appl. Spectrosc.* 16, 59 (1962).
- [2] J. Debras-Guédon and N. Liodec, De l'utilisation du faisceau d'un amplificateur a ondes lumineuses par émission induite de rayonnement (laser à rubis), comme source énergétique pour l'excitation des spectres d'émission des éléments, *CR Acad. Sci* 257, 3336 (1963).
- [3] L. J. Radziemski, From LASER to LIBS, the path of technology development, *Spectrochimica Acta Part B: Atomic Spectroscopy* 57, 1109 (2002).
- [4] J. D. Winefordner, I. B. Gornushkin, T. Correll, E. Gibb, B. W. Smith, and N. Omenetto, Comparing several atomic spectrometric methods to the super stars: special emphasis on laser induced breakdown spectrometry, LIBS, a future super star, *Journal of analytical atomic spectrometry* 19, 1061 (2004).
- [5] A. Kramida, K. Olsen, and Y. Ralchenko, Nist libs database, National Institute of Standards and Technology, US Department of Commerce (2019).
- [6] D. Capela, M. F. Ferreira, A. Lima, F. Dias, et al., Robust and interpretable mineral identification using laser-induced breakdown spectroscopy mapping, *Spectrochimica Acta Part B: Atomic Spectroscopy* , 106733 (2023).
- [7] Q. Wang, Y. Zhao, L. Liu, Z. Peng, et al., Study on cluster analysis used with laser-induced breakdown spectroscopy, *Plasma Science and Technology* 18, 647 (2016).
- [8] N. Gyftokostas, D. Stefanis, V. Kokkinos, C. Bouras, and S. Couris, Laser-induced breakdown spectroscopy coupled with machine learning as a tool for olive oil authenticity and geographic discrimination, *Scientific Reports* 11, 5360 (2021).

-
- [9] L. Narlagiri and V. R. Soma, Simultaneous quantification of Au and Ag composition from Au–Ag bi-metallic LIBS spectra combined with shallow neural network model for multi-output regression, *Applied Physics B* 127, 135 (2021).
 - [10] Y. Fu, Z. Hou, T. Li, Z. Li, and Z. Wang, Investigation of intrinsic origins of the signal uncertainty for laser-induced breakdown spectroscopy, *Spectrochimica Acta Part B: Atomic Spectroscopy* 155, 67 (2019).
 - [11] Z. Hu, D. Zhang, W. Wang, F. Chen, et al., A review of calibration-free laser-induced breakdown spectroscopy, *TrAC Trends in Analytical Chemistry* , 116618 (2022).
 - [12] L. Huang, R. Luo, X. Liu, and X. Hao, Spectral imaging with deep learning, *Light: Science & Applications* 11, 61 (2022).
 - [13] B. Busser, S. Moncayo, J.-L. Coll, L. Sancey, and V. Motto-Ros, Elemental imaging using laser-induced breakdown spectroscopy: A new and promising approach for biological and medical applications, *Coordination Chemistry Reviews* 358, 70 (2018).
 - [14] C. Fabre, Advances in Laser-Induced Breakdown Spectroscopy analysis for geology: A critical review, *Spectrochimica Acta Part B: Atomic Spectroscopy* 166, 105799 (2020).
 - [15] R. Noll, C. Fricke-Begemann, S. Connemann, C. Meinhardt, and V. Sturm, LIBS analyses for industrial applications—an overview of developments from 2014 to 2018, *Journal of Analytical Atomic Spectrometry* 33, 945 (2018).
 - [16] A. K. Pathak, R. Kumar, V. K. Singh, R. Agrawal, S. Rai, and A. K. Rai, Assessment of LIBS for spectrochemical analysis: a review, *Applied Spectroscopy Reviews* 47, 14 (2012).
 - [17] C. Fabre, D. Devismes, S. Moncayo, F. Pelascini, et al., Elemental imaging by laser-induced breakdown spectroscopy for the geological characterization of minerals, *Journal of Analytical Atomic Spectrometry* 33, 1345 (2018).
 - [18] V. Motto-Ros, L. Sancey, Q. Ma, F. Lux, et al., Mapping of native inorganic elements and injected nanoparticles in a biological organ with laser-induced plasma, *Applied Physics Letters* 101, 223702 (2012).
 - [19] L. Sancey, V. Motto-Ros, B. Busser, S. Kotb, et al., Laser spectrometry for multi-elemental imaging of biological tissues, *Scientific reports* 4, 6065 (2014).

-
- [20] L. Sancey, V. Motto-Ros, S. Kotb, X. Wang, et al., Laser-induced breakdown spectroscopy: a new approach for nanoparticle's mapping and quantification in organ tissue, *Journal of visualized experiments: JoVE* 88, e51353 (2014).
- [21] V. Motto-Ros, L. Sancey, X. Wang, Q. Ma, et al., Mapping nanoparticles injected into a biological tissue using laser-induced breakdown spectroscopy, *Spectrochimica Acta Part B: Atomic Spectroscopy* 87, 168 (2013).
- [22] R. Hai, C. Li, H. Wang, H. Ding, H. Zhuo, J. Wu, and G.-N. Luo, Characterization of Li deposition on the first wall of EAST using laser-induced breakdown spectroscopy, *Journal of Nuclear Materials* 438, S1168 (2013).
- [23] C. Li, X. Wu, C. Zhang, H. Ding, J. Hu, and G.-N. Luo, In situ chemical imaging of lithiated tungsten using laser-induced breakdown spectroscopy, *Journal of Nuclear Materials* 452, 10 (2014).
- [24] Y. Gimenez, B. Busser, F. Trichard, A. Kulesza, et al., 3D imaging of nanoparticle distribution in biological tissue by laser-induced breakdown spectroscopy, *Scientific Reports* 6, 1 (2016).
- [25] I. J. Fairchild and E. A. McMillan, Speleothems as indicators of wet and dry periods, *International Journal of Speleology* 36, 2 (2007).
- [26] V. Motto-Ros, V. Gardette, L. Sancey, M. Leprince, et al., LIBS-based imaging: Recent advances and future directions, *Spectroscopy* 35, 34 (2020).
- [27] D. Genty, A. Baker, and B. Vokal, Intra-and inter-annual growth rate of modern stalagmites, *Chemical Geology* 176, 191 (2001).
- [28] T. Lopes, P. Rodrigues, R. Cavaco, D. Capela, et al., Interactive three-dimensional chemical element maps with laser-induced breakdown spectroscopy and photogrammetry, *Spectrochimica Acta Part B: Atomic Spectroscopy* 203, 106649 (2023).
- [29] S.-J. Baek, A. Park, Y.-J. Ahn, and J. Choo, Baseline correction using asymmetrically reweighted penalized least squares smoothing, *Analyst* 140, 250 (2015).
- [30] J. Peng, S. Peng, A. Jiang, J. Wei, C. Li, and J. Tan, Asymmetric least squares for multiple spectra baseline correction, *Analytica chimica acta* 683, 63 (2010).

- [31] G. Yangmin, T. Yun, D. Yu, T. Shisong, et al., Cluster analysis of polymers using laser-induced breakdown spectroscopy with K-means, *Plasma Science and Technology* 20, 065505 (2018).
- [32] S. Moncayo, S. Manzoor, F. Navarro-Villoslada, and J. Caceres, Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy, *Chemometrics and Intelligent Laboratory Systems* 146, 354 (2015).
- [33] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D. W. Hahn, and J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, *Spectrochimica Acta Part B: Atomic Spectroscopy* 148, 65 (2018).
- [34] I. Langmuir, Oscillations in ionized gases, *Proceedings of the National Academy of Sciences* 14, 627 (1928).
- [35] J. Aguilera and C. Aragón, Multi-element Saha–Boltzmann and Boltzmann plots in laser-induced plasmas, *Spectrochimica Acta Part B: Atomic Spectroscopy* 62, 378 (2007).
- [36] H. Griem, *Spectral line broadening by plasmas* (Elsevier, 2012).
- [37] V. Babushok, F. DeLucia Jr, J. Gottfried, C. Munson, and A. Miziolek, Double pulse laser ablation and plasma: Laser induced breakdown spectroscopy signal enhancement, *Spectrochimica Acta Part B: Atomic Spectroscopy* 61, 999 (2006).
- [38] O. Samek, D. C. Beddows, J. Kaiser, S. V. Kukhlevsky, M. Liska, H. H. Telle, and A. J. Whitehouse, Application of laser-induced breakdown spectroscopy to in situ analysis of liquid samples, *Optical Engineering* 39, 2248 (2000).
- [39] M. Abdelhamid, F. J. Fortes, J. Laserna, and M. Harith, in *AIP Conference Proceedings*, Vol. 1380 (American Institute of Physics, 2011) pp. 55–59.
- [40] P. Fichet, J.-L. Lacour, D. Menut, P. Mauchien, et al., Micro LIBS technique, *Laser-Induced Breakdown Spectroscopy (LIBS): Fundamentals and Applications*, 539 (2006).
- [41] K. Rifai, F. Doucet, L. Özcan, and F. Vidal, LIBS core imaging at kHz speed: Paving the way for real-time geochemical applications, *Spectrochimica Acta Part B: Atomic Spectroscopy* 150, 43 (2018).

- [42] P. Liu, Z. Wang, and Z. Zhang, A New Quantitative Approach for Element-Mineral Determination Based on “EDS (Energy Dispersive Spectroscopy) Method”, *Geofluids* 2021, 1 (2021).
- [43] J. W. Cooley, P. A. Lewis, and P. D. Welch, The fast Fourier transform and its applications, *IEEE Transactions on Education* 12, 27 (1969).
- [44] B. Mosser, D. Mekarnia, J. Maillard, J. Gay, D. Gautier, and P. Delache, Seismological observations with a Fourier transform spectrometer-Detection of Jovian oscillations, *Astronomy and Astrophysics* 267, 604 (1993).
- [45] J. L. Massey, The discrete fourier transform in coding and cryptography, *IEEE Inform* (1998).
- [46] H. M. Ozaktas and D. Mendlovic, Fourier transforms of fractional order and their optical interpretation, *Optics Communications* 101, 163 (1993).
- [47] H. Stark, *Application of Optical Fourier Transforms* (Elsevier, 2012).
- [48] R. A. Brown, H. Zhu, and J. R. Mitchell, Distributed vector processing of a new local multiscale Fourier transform for medical imaging applications, *IEEE Transactions on medical imaging* 24, 689 (2005).
- [49] R. Sahu and S. Mordechai, Fourier transform infrared spectroscopy in cancer detection, (2005).
- [50] J. S. Lim, *Two-dimensional signal and image processing*, Englewood Cliffs (1990).
- [51] M. Wang, Q. Wang, M. Zhu, L. Sun, X. Peng, L. Liu, and J. Qu, in 2015 Optoelectronics Global Conference (OGC) (IEEE, 2015) pp. 1–3.
- [52] L. Sancey, V. Motto-Ros, B. Busser, S. Kotb, et al., Laser spectrometry for multi-elemental imaging of biological tissues, *Scientific reports* 4, 1 (2014).
- [53] I. Lopez-Quintas, M. Mateo, V. Piñon, A. Yañez, and G. Nicolas, Mapping of mechanical specimens by laser induced breakdown spectroscopy method: Application to an engine valve, *Spectrochimica Acta Part B: Atomic Spectroscopy* 74, 109 (2012).
- [54] W. Linder, *Digital photogrammetry*, Vol. 1 (Springer, 2009).

- [55] R. M. Ruiz, M. T. M. Torres, and P. S. Allegue, Comparative Analysis Between the Main 3D Scanning Techniques: Photogrammetry, Terrestrial Laser Scanner, and Structured Light Scanner in Religious Imagery: The Case of The Holy Christ of the Blood, *ACM Journal on Computing and Cultural Heritage (JOCCH)* 15, 1 (2021).
- [56] S. Umeyama, Least-squares estimation of transformation parameters between two point patterns, *IEEE Transactions on Pattern Analysis & Machine Intelligence* 13, 376 (1991).
- [57] N. Fabian, K. Moreland, D. Thompson, A. C. Bauer, et al., in *2011 IEEE Symposium on Large Data Analysis and Visualization (IEEE, 2011)* pp. 89–96.
- [58] W. J. Schroeder, L. S. Avila, and W. Hoffman, Visualizing with VTK: a tutorial, *IEEE Computer graphics and applications* 20, 20 (2000).
- [59] H. Huang, L. Liu, and M. O. Ngadi, Recent developments in hyperspectral imaging for assessment of food quality and safety, *Sensors* 14, 7248 (2014).
- [60] M. Schäferling, The art of fluorescence imaging with chemical sensors, *Angewandte Chemie International Edition* 51, 3532 (2012).
- [61] D. Prochazka, T. Zikmund, P. Pořízka, A. Břínek, et al., Joint utilization of double-pulse laser-induced breakdown spectroscopy and X-ray computed tomography for volumetric information of geological samples, *Journal of Analytical Atomic Spectrometry* 33, 1993 (2018).
- [62] J. Lavicky, M. Kolouskova, D. Prochazka, V. Rakultsev, et al., The development of dentin microstructure is controlled by the type of adjacent epithelium, *Journal of Bone and Mineral Research* 37, 323 (2022).
- [63] S. Gupta, J. Hoffman, and J. Malik, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) pp. 2827–2836.
- [64] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, A systematic review on supervised and unsupervised machine learning algorithms for data science, *Supervised and unsupervised learning for data science* , 3 (2020).
- [65] J. Delua, Supervised vs. unsupervised learning: What’s the difference, *IBM. com Blog* 12, 2021 (2021).

-
- [66] J. A. Dunnmon, A. J. Ratner, K. Saab, N. Khandwala, et al., Cross-modal data programming enables rapid medical machine learning, *Patterns* 1, 100019 (2020).
- [67] Q. Wang, Y. Zhao, L. Liu, Z. Peng, et al., Study on cluster analysis used with laser-induced breakdown spectroscopy, *Plasma Science and Technology* 18, 647 (2016).
- [68] S. A. El_Rahman, Hyperspectral image classification using unsupervised algorithms, *International Journal of Advanced Computer Science and Applications* 7 (2016).
- [69] J. M. Amigo, C. Ravn, N. B. Gallagher, and R. Bro, A comparison of a common approach to partial least squares-discriminant analysis and classical least squares in hyperspectral imaging, *International journal of pharmaceutics* 373, 179 (2009).