BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF MECHANICAL ENGINEERING

FAKULTA STROJNÍHO INŽENÝRSTVÍ

INSTITUTE OF PHYSICAL ENGINEERING

ÚSTAV FYZIKÁLNÍHO INŽENÝRSTVÍ

ADVANCED SPECTRAL AND IMAGE PROCESSING OF HYPERSPECTRAL DATA

POKROČILÉ SPEKTRÁLNÍ A OBRAZOVÉ ZPRACOVÁNÍ HYPERSPEKTRÁLNÍCH DAT

DOCTORAL THESIS

DIZERTAČNÍ PRÁCE

AUTHOR — **Mgr. Ondřej Vaculík**

AUTOR PRÁCE

SUPERVISOR — **Ing. Mojmír Šerý, Ph.D.**

ŠKOLITEL

BRNO 2025

Abstract
This doctoral thesis investigates hyperspectral imaging, a non-destructive spatio-spectral analysis technique, with the aim to streamline its data analysis and enhance its utility for routine plant and algal growth assessment. The text describes the structure of hyperspectral data and the principles of hyperspectral acquisition systems. The theoretical background covers hyperspectral data analysis using image processing, chemometric methods, and machine learning algorithms. Among the methods specifically explored are neural networks, spectral unmixing, and dimensionality reduction. A significant practical contribution of this work is the development of a hyperspectral microscope workstation. The experimental results from this system include verifying its spectral resolution and characterising it using a reference colour calibration slide. The dataset from this calibration slide was also utilised in the study of the effectiveness of dimensionality reduction and spectral unmixing algorithms, which is presented in this text. Furthermore, the hyperspectral microscope was successfully utilised for the identification and quantification of marine cyanobacteria mixtures.

Abstrakt
Dizertační práce se zabývá hyperspektrálním zobrazováním, nedestruktivní technikou prostorově-spektrální analýzy, s cílem zefektivnit analýzu těchto dat a zvýšit její využitelnost pro rutinní posuzování růstu rostlin a řas. Text popisuje strukturu hyperspektrálních dat a principy systémů hyperspektrálního snímání. Teoretické základy zahrnují analýzu hyperspektrálních dat pomocí zpracování obrazu, chemometrických metod a algoritmů strojového učení. Mezi konkrétně zkoumané metody patří neuronové sítě, spektrální unmixing a redukce dimenzionality. Významným praktickým přínosem této práce je vývoj hyperspektrální mikroskopické pracovní stanice. Experimentální výsledky obsažené v této práci zahrnují ověření spektrálního rozlišení hyperspektrálního mikroskopu a jeho charakterizaci pomocí referenčního barevného kalibračního testu. Soubor dat z tohoto kalibračního testu byl rovněž využit při studiu účinnosti algoritmů pro redukci dimenzionality a spektrální unmixing. Dále byl hyperspektrální mikroskop úspěšně využit k identifikaci a kvantifikaci směsí mořských sinic.

Keywords
hyperspectral imaging, microscopy, image processing, machine learning, neural networks, spectroscopy, spectral unmixing, dimensionality reduction

Klíčová slova
hyperspektrální snímání, mikroskopie, zpracování obrazu, strojové učení, neuronové sítě, spektroskopie, spektrální unmixing, redukce dimenzionality

Bibliographic quotation
VACULÍK, Ondřej. Advanced spectral and image processing of hyperspectral data. Doctoral Thesis. Mojmír ŠERÝ (supervisor). Brno: Brno University of Technology, Faculty of Mechanical Engineering, 2025.

Bibliografická citace
VACULÍK, Ondřej. Pokročilé spektrální a obrazové zpracování hyperspektrálních dat. Dizertační práce. Mojmír ŠERÝ (školitel). Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2025.

Advanced Spectral and Image Processing of Hyperspectral Data

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracoval samostatně pod vedením Ing. Mojmíra Šerého, Ph.D. Další informace mi poskytli Mgr. Silvie Bernatová, Ph.D., Mgr. Katarína Rebrošová, Mgr. Martin Šiler, Ph.D. a Mgr. Jiří Šetlík. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....

Mgr. Ondřej Vaculík

30. června 2025

Poděkování

Děkuji svému vedoucímu Ing. Mojmírovi Šerému, Ph.D. za zajištění financování, cenné rady a věnovaný čas během mého doktorského studia. Velký dík patří také mým kolegům z Ústavu přístrojové techniky AV ČR prof. RNDr. Pavlovi Zemánkovi, Ph.D., Mgr. Silvii Bernatové, Ph.D., Mgr. Martinovi Šilerovi, PhD., Ing. Tadeáši Maňkovi, Ph.D., Ing. Tomáši Plichtovi, Ph.D., Ing. Veronice Richterové, Ing. Denise Šilhanové, Ing. Lukáši Šilhanovi a Tomáši Šilhanovi za pomoc při řešení projektů a přátelské pracovní prostředí.

Za spolupráci při vývoji hyperspektrální kamery děkuji firmě PSI (Photon Systems Instruments), spol. s r.o. Za poskytnutí vzorků hlubokomořských sinic děkuji Centru Algatech MBU AV ČR a Mgr. Jiřímu Šetlíkovi. Technická podpora ze skupiny Mikroskopie pro biomedicínu ÚPT AV ČR Brno a CF Elektronová mikroskopie a Ramanova spektroskopie, kterou poskytly Anna Havlíčková, Kateřina Mrázová a Kamila Hrubanová, byla neocenitelná při získávání EM snímků použitých v této práci. Děkuji také skupině Tinlicker za udržení potřebného tempa.

V poslední řadě bych chtěl srdečně poděkovat své rodině, blízkým a přítelkyni za trpělivost, nekonečnou podporu a zázemí, bez kterých by tato práce nemohla vzniknout.

Výzkum byl podpořen projekty MŠMT OP JAK (EH22_008/0004624), MPO (FV40455, EG19_262/0020294), TA ČR (FW06010453, FW10010248) a AV ČR (RVO:68081731).

Table of Contents

List of Figures

List of Tables

1 Introduction

Hyperspectral cameras are a new generation of optical systems that combine precision optics, spectroscopy, and image processing techniques to obtain spatiospectral information in a sufficiently short time. With the emergence of commercial hyperspectral imaging systems that were previously available mainly on Earth observation and astronomical satellites, hyperspectral imaging has moved closer to mainstream use. Currently, hyperspectral data are widely used in the fields of agriculture, medicine, environmental and conservation research, geological or marine research studies, forestry and oceanography, preclinical research, disease diagnosis, and *in vivo* imaging of various cellular and molecular processes [1, 2, 3, 4, 5].

The effective use of hyperspectral data requires an understanding of the origins and limitations of the collected data and various processing approaches to interpret the collected data. Hyperspectral images contain both spectral and spatial information; we can use spectral analysis tools to examine individual spectral responses and image-processing algorithms to extract spatial information from individual spectral band images. Nevertheless, due to the complexity and significant size of hyperspectral images, machine learning algorithms are more than suitable for many tasks related to hyperspectral data, such as dimensionality reduction, classification, and automated labelling [6, 7, 8].

This work builds upon this foundation and is structured in the following way. The Introduction provides a historical context for hyperspectral imaging, reviews the existing literature on hyperspectral microscopy, and states the objectives of the dissertation. The theoretical part explores the structure of hyperspectral data, various acquisition systems, calibration procedures, and methods for HS data analysis, including dimensionality reduction, spectral unmixing, and error measures. The Methods section details the experimental setup of the custom hyperspectral microscope workstation developed as part of this thesis at the Institute of Scientific Instruments of the Czech Academy of Sciences (ISI CAS), its software for automated acquisition and data management, and the specific measurements performed, such as spectral resolution verification, colour calibration slide measurement, and analysis of marine cyanobacteria mixtures. Finally, the conclusion summarises the key contributions, findings, and future implications of this thesis.

1.1 History of Hyperspectral Imaging

Spectral imaging, which involves the capture of images at different wavelengths, has evolved over time to become a crucial tool in various fields such as remote sensing, biology and materials science [9, 4, 10, 11]. At its core, spectral imaging can be divided into two main subgroups: multispectral (MS) and hyperspectral (HS) imaging. MS imaging uses a small number of individual wavelength bands, which are generally not adjacent and can have various bandwidths ($\Delta\lambda \geq 20\,\mathrm{nm}$). Each band usually has a descriptive title (red, near-infrared, thermal infrared, etc.). Bands are typically selected with a specific application in mind, as in the case of the instrumentation onboard the remote sensing satellites, which focusses on agriculture, land mapping, or change detection. In contrast, HS images contain a large number (100+) of narrow contiguous spectral bands that capture an accurate spectral response with a relatively high signal-to-noise ratio [12]. The structure of HS images, where each pixel contains detailed spectral information in the detected range of wavelengths, offers rich analytical information inherent in both spatial and spectral signatures. These signatures are intrinsically linked, meaning that the spatial distribution of a substance can

be assessed if its spectral signature is sufficiently distinct to be isolated [13]. HS imaging is closely related to the development of imaging spectrometers. The term "hyperspectral imaging" was first coined in a paper discussing the early results of the imaging spectrometry technique, described by the authors as "the acquisition of images in hundreds of contiguous, registered spectral bands such that for each pixel a radiance spectrum can be derived" [14]. The requirement of contiguity is the crucial difference between HS and MS data. An illustrative comparison of the detected spectral bands of a MS and HS data captured by two fictitious detectors is presented in Figure 1.1.

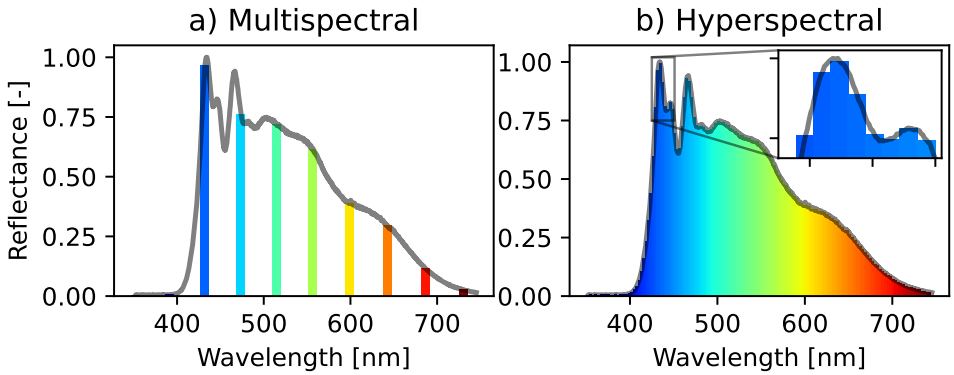


Figure 1.1: Artificial example of a) multispectral (MS) and b) hyperspectral (HS) wavelength sensitivity regions for two theoretical sensors; grey line represents a true sample reflectance, coloured bars are the measured spectral bands – a total of 10 in case of the MS and 140 bands for the HS example.

As presented in Figure 1.2, in a graph showing how selected phrases have occurred in a corpus of books over the selected years, the keyword "multispectral" gained early prominence, particularly after the 1960s, followed by a small rise of the term "spectral imaging" in the 1990s and a more recent surge of the use of the keyword "hyperspectral" from the 2010s onwards. This reflects the evolving technological advancements in imaging, with MS techniques leading due to their application in remote sensing, and HS techniques emerging as necessary technologies further advanced.

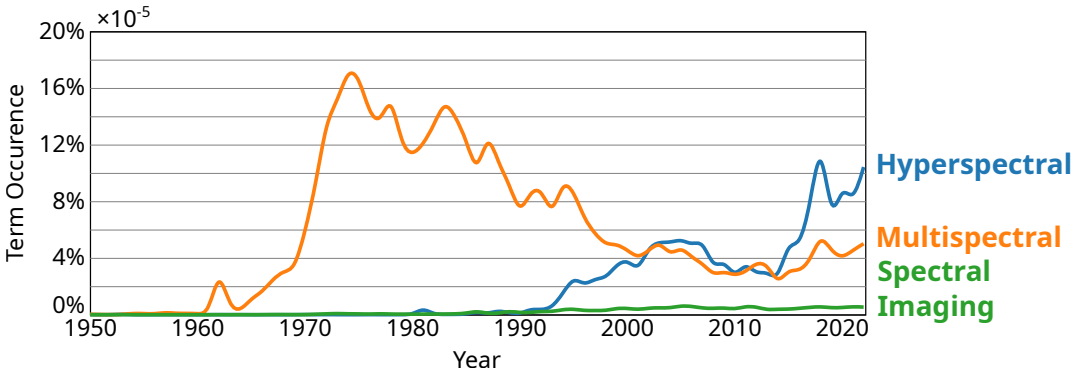


Figure 1.2: The Google N-gram plot, which shows the percentage of books in which each term appears throughout the years, indicates the evolving prominence of the three technologies over time.

The initial idea and optical design of imaging spectrometers were ahead of the sensory and data storage capabilities of the time. Therefore, the first deployment of HS systems was not made until the 1980s, when the Airborne Imaging Spectrometer (AIS) and Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) were developed by NASA's Jet Propulsion Laboratory (JPL) [12] , followed by Near-Infrared Mapping Spectrometer (NIMS) and Moderate Resolution Imaging Spectroradiometer (MODIS) instruments deployed in 1990s onboard Galileo and Aqua spacecrafts, respectively [15]. The above-mentioned HS systems have been deployed primarily for use in remote sensing.

Advancements in optics and electronics enabled the development of smaller and less expensive spectrometers, which consequently influenced the development of HS systems as well, enabling the transition from research-focused remote sensing-first platforms to other research fields and industry-oriented devices [16, 17]. The democratisation of MS and HS imaging is still an ongoing process, which has been made possible by the increasing availability of more affordable and portable solutions [18].

HS technology is being adopted by other fields such as agriculture, forest management, environmental monitoring, which take advantage of remote sensing HS data, as well as food safety and inspection, medical diagnosis, or by industries such as waste sorting, quality control, inspection, which require ground/laboratory-based HS measurements [19]. These diverse applications have collectively contributed to an increase in both the volume and complexity of HS datasets, requiring more sophisticated analytical approaches.

Alongside this growth in application domains, the increase in digital data storage capabilities and computational power has enabled significant advancement in HS data analysis methodologies. As the resolution of both the spectral and spatial components of the HS data has increased, with modern sensors capable of capturing hundreds of contiguous spectral bands at high spatial resolutions, the need for novel analytical algorithms has become increasingly evident. Traditional image processing techniques are becoming incapable of processing such high dimensionality and inherent complexity of contemporary HS datasets, therefore, the field of HS image analysis is in the stage of adoption of machine learning and neural network-based approaches [20].

To establish a comprehensive understanding of current research priorities in HS imaging and analysis, a systematic review of the scientific literature is essential to compare publication trends between the use of HS in remote sensing and microscopy. In the following section, a quantitative analysis of the disparity in research output between these two major application domains is presented.

1.2 Literature Review

The longer history of HS imaging for remote sensing is clearly reflected in the scientific literature. Table 1.1 compares the number of articles that contain the keyword "hyperspectral" and "remote sensing" in their title, abstract or keywords, with the number of articles with "hyperspectral" and "microscop*", where the star symbol stands for a wildcard symbol allowing for any letter(s) to continue, including terms such as microscope, microscopy, microscopic, and others. The search is limited to scientific articles only, with the exact form of the search strings presented in the Supplement Table S1.

The result of the search in the majority of the databases shows that there exist more articles on the topic of HS connected to remote sensing than on HS and microscopy. The exception is the PubMed database, where more articles on HS and microscopy were found, but that is probably due to the scope of the database, which focusses more on medical

Table 1.1: Number of articles found using keywords in each of the journal databases – Clarivate Web of Science (WOS), Institute of Electrical and Electronics Engineers (IEEE), Elsevier Socpus and U.S. Department of Health and Human Services PubMed.

Keywords	**WOS**	**IEEE**	**Scopus**	**PubMed**
Hyperspectral AND remote sensing	6542	1211	12488	1113
Hyperspectral AND microscop*	1391	46	1858	1132

literature rather than technical and engineering work, into which the remote sensing topic is usually categorised.

These findings indicate a knowledge gap; the connection of HS imaging and microscopy is a developing topic and was selected as an area of interest for this thesis. This knowledge gap includes both instrumentation and HS data analysis, as many of the developed algorithms and analysis approaches are focused on the remote sensing HS data and applications relevant for their users, which are not necessarily compatible or applicable for the new use cases mentioned above.

From the publications on HS microscopy, the present use cases are remarkably diverse. In biomedicine and medicine, HS microscopy has been explored for the detection and diagnosis of cancer in various tissues, by analysing the morphological and structural properties of histological samples, examining tissue surfaces and quantifying tumour characteristics [21]. HS microscopy is also applied in forensic analysis, such as document forgery and paint chip analysis [22], and medical applications such as evaluating burn wounds, cervical neoplasia, ocular diagnostics or identifying nerve fibres [4]. A notable advantage in biomedical applications is the ability of HS microscopy to accurately resolve and quantify fluorescent labels in highly auto-fluorescent tissues where standard single-band microscopy methods struggle to distinguish signals from background noise [23]. Recent developments also include automated whole-slide HS imaging microscopy for digital pathology [24]. Another key finding indicates an overlap of HS microscopy with the field of chemometrics, a discipline focused on extracting information from data in a multivariate manner, which is applicable to HS data as well [25].

In conclusion, the rapid development and application of HS microscopy have far-reaching implications for various fields. As HS instruments and techniques continue to evolve, it is essential to invest in the development of HS microscopy to enable its translation from a research tool to a practical, routine method in clinical and industrial settings.

Introduction Highlights

- HS systems were initially developed for satellite imaging and remote sensing.
- Advances in optics and electronics allowed for the development of smaller and commercial HS systems.
- HS imaging is being adopted by other scientific fields and industries.
- New algorithms for HS data analysis that can be applied outside the field of remote sensing are needed.

1.3 Objectives of the Doctoral Thesis

The aim of this doctoral thesis is to reduce the sophistication of HS data analysis and to increase the potential of this non-destructive imaging technique for routine analysis of plant and algal plant growth. The work will include the essential theoretical and practical foundation for measuring HS data and an overview of relevant spectroscopic techniques. The listed objectives of this doctoral thesis are the following:

- To create a tool for measuring, processing, and evaluating HS data, including the hardware and software necessary for acquisition, calibration, and visualisation.
- Perform a characterisation of the HS instrument using suitable calibration samples.
- Using the measured data, explore the collaboration of machine learning methods and chemometric algorithms to develop analysis algorithms that do not rely heavily on user input.
- Investigate the possibilities of reducing the dimensionality of the HS data using conventional and machine learning algorithms.

2 Theoretical Part

2.1 Structure of Hyperspectral Data

In the image processing literature [26, 27, 28], a two-dimensional matrix is used as a mathematical description of a greyscale intensity image $\mathbf{M}(x, y)$ composed of individual pixel values $p_{x,y}$ that are described with spatial location (x, y) in the image. This notation can be extended to hyperspectral imaging (HSI) using a three-dimensional matrix or tensor. In this thesis, the following notation will be used:

$$I(x, y, \lambda) \in \mathbb{R}^{W \times H \times B}, \tag{2.1}$$

where I represents the captured intensity of light (HS image), x denotes the spatial coordinate along the horizontal axis (column index), y denotes the spatial coordinate along the vertical axis (row index), and λ represents the captured wavelength, which spans the spectral range captured by the HSI system.

The individual pixel values $p_{x,y,\lambda}$ belong to a space of real numbers with the following dimensions: W is the width of the HSI (number of pixels per line, rows), H is the number of lines in the HSI (columns), and B is the number of distinct spectral bands.

Using this representation of HS data, we can specify what kind of information is processed by each analysis approach. The Methods section will also specify the relationship of this notation to the array indexing rules used in the Python programming language [29], which is the language of choice for the HS data analysis in this work. Because the HSI pixel values are arranged in an orthogonal three-dimensional space, they can be visualised as a cube; see Figure 2.1. This visualisation shows the building blocks of the HSI, the relationship between the spatial and spectral values, and also shows the possible viewpoints that can be used when examining and analysing the HSI.

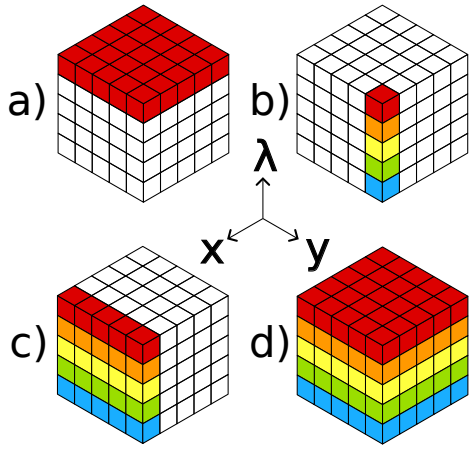


Figure 2.1: Components of a HSI a) single spectral band, b) individual spectral pixel, c) one row of pixels, an output of imaging spectrometer and d) entire HS cube.

Spatial analysis uses individual spectral bands. A single spectral band is a collection of pixels p with coordinates (x, y), belonging to a single spectral band i, defined as:

$$\mathbf{B}_i(x, y) = p_{x,y,i}, \tag{2.2}$$

where $i \in \mathbb{R}^B$. Spectral analysis examines the intensity variations between wavelengths, by studying the spectral signatures I of individual pixels at distinct locations x, y. The individual spectral pixel S is a collection of all band values located in a single point in space:

$$S_{x,y}(\lambda) = p_{x,y,\lambda}, \tag{2.3}$$

where $x, y \in \mathbb{R}^{W \times H}$. Lastly, we define a set of spectral vectors $\mathbf{S}_{2D}$, which is the entire HSI reshaped to a matrix (2D dataset) for spectral-only processing, discarding the spatial relationship between the spectral pixels:

$$\mathbf{S}_{2D} \in \mathbb{R}^{(W \times H) \times B}. \tag{2.4}$$

The following section describes how an HSI can be captured, which is tightly connected to the building blocks of HSI.

2.2 Types of Hyperspectral Acquisition Systems

One of the key factors that affect the resolution and accuracy of HSI is the type of acquisition system used. In this section, we will explore the different types of HSI acquisition systems, each with its unique characteristics of hardware solutions, advantages, and applications.

2.2.1 Spectral Filtering

A first possible approach to HSI acquisition is the use of narrow-band spectral filters combined with a monochrome camera sensor [30, 31]. A set of different narrow-band spectral filters enables the user to capture a small number of selected spectral bands $\mathbf{B}_i$, forming a multispectral image rather than HSI. A schematic representation of a single spectral band is illustrated in Figure 2.1a). However, with the use of a tunable bandpass filter (liquid crystal or Fabry-Perot), the system is capable of capturing a large number of adjacent spectral bands in a sufficiently short time [32, 33]. An example of the existing system is the HinaLea 4250 HS camera that operates in the spectral range of 400–1000 nm with the 4 nm Full Width at Half Maximum (FWHM) spectral band, capturing a total number of 300 bands thanks to the use of the front-staring Fabry-Perot spectral filter [34]. Another option is to use a liquid crystal tunable bandpass filter as a wavelength selection instrument, such as Kurios Tunable Filter, which operates in the 420–730 nm region with bandwidth ranging from 10 nm to 35 nm depending on the model [35]. The optical assembly for both systems described here is visualised in Figure 2.2.

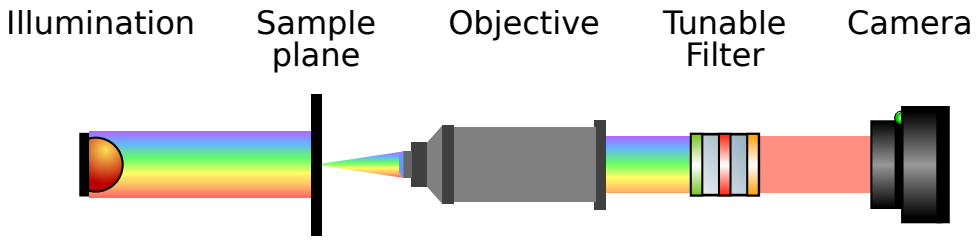


Figure 2.2: Schematic of a tunable spectral filter imaging optical assembly.

The advantage of such systems is the uniformity of the captured data in spatial domain, because the sample or camera does not move during measurement, and additionally there are less prone to spectral errors such as smile and keystone. The only significant source of distortion is the non-uniformity introduced by the optical system, filtering element, and the image sensor itself, as explored in [36]. However, spectral filter-based systems struggle with continuously capturing fast-moving objects because they require time to switch between different spectral bands. Also, the spectral resolution is lower compared to, for example, point scanning or push-broom systems, which are described in the following sections.

2.2.2 Point Scan

Another possible approach is the use of a spectrometer in conjunction with a positioning system that enables the user to capture spectral information of different locations of the observed sample. A single measured spectral pixel $S_{x,y}$ is illustrated in Figure 2.1b). This technique is used in procedures such as Raman mapping or any other spectroscopic capture performed on the area of interest within the sample. The great advantage of this approach is the unmatched spectral resolution, or even alternative spectroscopic information such as Fourier transform infrared (FTIR) imaging, or other [37]. One drawback is the time-consuming nature of the scanning and acquisition procedure, coupled with the problem of establishing an appropriate spatial sampling, as illustrated in Figure 2.3.

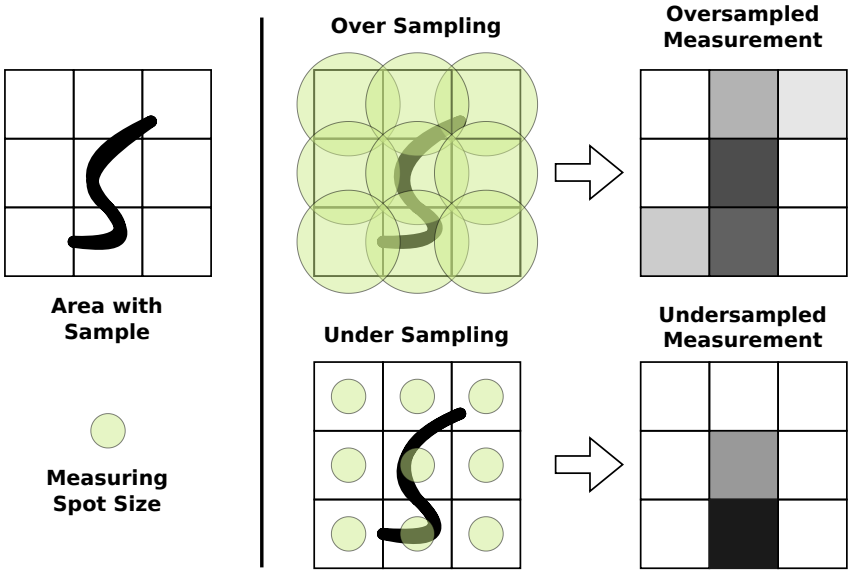


Figure 2.3: Principle of point scan measurement, with the illustration of the problem of sampling. When the measurement spot size is too large, the spatial resolution decreases and so does the signal strength. However, when the spot size is too small, there is a risk of missing certain parts of the sample. Adapted from [38].

The requirement for the point scan method is that the sample is stationary and does not move or change during the mapping procedure, which can take a long time, depending on

the required exposure time per single point spectrum measurement. In addition, the area in which the measurement can be performed is limited by the range of the motion control system. The measurement time can be reduced by using a set of two galvo mirrors or other beam steering technology to facilitate the scanning, but this greatly increases the overall cost of the system. A compromise between the spectral resolution of a point-scanner system and the spatial uniformity of a spectral filtering system is the imaging-spectrometer-based approach called pushbroom scanning, which is described in the following section.

2.2.3 Pushbroom

Advances in optical design and the availability of more complex optical elements, such as non-planar diffraction gratings [39], holographic gratings and curved prisms [40] enabled the development of imaging spectrometers, which allow the simultaneous collection of spectral information from multiple points along a line. By moving either the sample or the measurement instrument, an HSI is formed by sequential acquisition of spatiospectral frames, as illustrated in Figure 2.1c). The advantage of this approach is the reduced capture time compared to point scan while preserving the high spectral resolution. The requirement of movement perpendicular to the entrance slit of the imaging spectrometer can be satisfied "for free" by using an existing movement, such as rotation of the Earth with respect to the HS-equipped satellite (see Figure 2.4), flight movement of an aircraft or unmanned aerial vehicle system [41], or when scanning objects on a conveyor belt, which is the case for many industrial applications of HS systems [42].

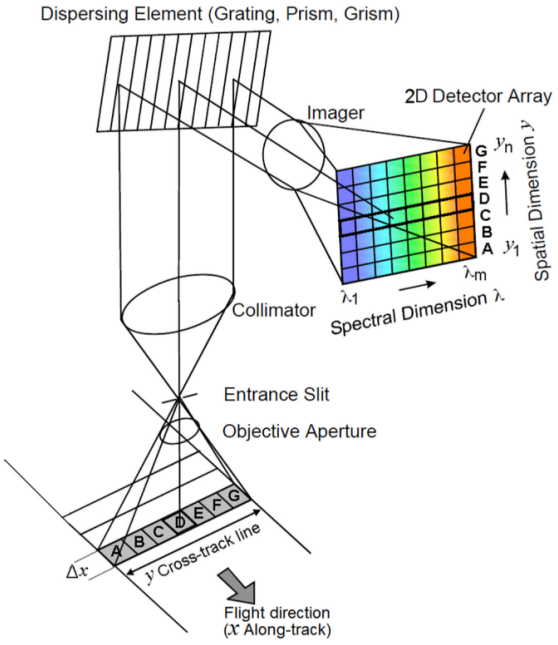


Figure 2.4: Schematic of a general dispersive elements based pushbroom HS satellite imager. Adapted from [12].

The disadvantage of the pushbroom system is the sensitivity to motion parallel to the input slit, which distorts the spatial information. This distortion mostly affects satellite or aircraft HS scanning systems and can be compensated for using motion data captured by

an inertial measurement unit during the measurement. Another disadvantage is the limited field of view, compared to the spectral filter approach, which also negatively impacts the time requirements for data collection over large areas. This limitation can be effectively addressed through the simultaneous capture of complete spatial and spectral information in a single exposure, enabling real-time HS imaging without scanning mechanisms, a capability offered by the snapshot HS system, which are discussed in the next section.

2.2.4 Snapshot

The last and most complex approach for HSI acquisition is known as snapshot, with the underlying principle of capturing the entire HSI simultaneously. This can be achieved by a clever combination of complex optics and computer image processing algorithms. There are several techniques of snapshot HSI acquisition, including the use of a coded aperture [43, 44], linear variable filter, and a micro-lens array for a single-sensor multi-aperture design [45] or even lensless solutions based on a combination of spectral array filter and diffuser [46]. The principles behind snapshot HSI acquisition are beyond the scope of this work; further information on this topic can be found in a comprehensive review presented by Hagen and Kudenov [47].

The clear advantage, as the name suggests, is the simultaneous capture of spectral and spatial data, allowing for HSI measurements of complex scenes and objects in motion (under sufficient lightning conditions to minimise the exposure time and therefore reduce motion blur). Some of the HS snapshot optical configurations are extremely compact allowing for ease of use on small unmanned aerial vehicles or handheld operation [48]. The disadvantage of these systems is the trade-off between spatial and spectral resolution, generally putting them somewhere between filter scanning and pushbroom systems in terms of spatial and spectral resolution, respectively. Snapshot systems are less common in scientific applications but are an area of active development that has great potential in the future adoption of HS imaging outside scientific and industrial settings.

Regardless of the HS image acquisition method used, the raw captured HS data need to be calibrated before any further analysis can be performed. The next chapter will examine the essential steps and techniques involved in calibrating HS data, laying the foundation for accurate and reliable analysis.

Hyperspectral Data Structure and Acquisition Highlights

- HS data form a 3D matrix $I(x, y, \lambda)$ that contains spatial and spectral information.
- At a selected wavelength value we can slice the matrix perpendicular to the spectral axis to select one spectral band $\mathbf{B}_i$.
- When fixing the two spatial coordinates, we can work with a spectral signature $S_{x,y}$ of a single point.
- HS data can be acquired using a camera with varying filter assembly by stacking images captured at different wavelengths.
- The HS cube can be formed using a spectrometer with point by point scanning or faster using a imaging spectrometer with scanning capabilities.
- Simultaneous capture of spatial and spectral data is also possible, but with a trade-off in spatial and/or spectral resolution or increased computational complexity.

2.3 Hyperspectral Data Calibration

To ensure high accuracy and low distortion of the captured HS data, the HS instrument should be properly calibrated. This includes three key steps: spectral calibration, spatial correction, and radiometric calibration.

2.3.1 Spectral Calibration

Spectral calibration of the HS system can be accomplished using a known spectral source, such as a gas mixture calibration lamp (argon, krypton, mercury-argon) or a laser source or using a monochromator. With a known spectral characteristic, given by atomic transitions or construction of the laser source, pixel-to-wavelength mapping can be performed [49, 50]. This mapping function is estimated, commonly using a linear, second-order, or third-order polynomial fitting, as indicated in Equation 2.5:

$$\hat{\lambda} = \sum_{i=0}^{n} a_i p_\lambda^i, \tag{2.5}$$

where $\hat{\lambda}$ is the estimated wavelength, n is the order of the polynomial, p_λ is the pixel index at wavelength λ, and a_i are the polynomial coefficients obtained using the fit. More advanced methods for spectral calibration do not use polynomial fit but instead use a mathematical model of the entire optical system of the spectrometer [51].

The pixel-to-wavelength map is vital for further analysis, so it is important that it is included in the HS data file, as is the case for the ENVI file format (NV5 Geospatial Solutions, Inc.) used in remote sensing, which was developed as a standardised platform for storing and manipulating geospatial images [52]. Furthermore, the spectral resolution of the HS system can be evaluated with a suitable source that emits spectral lines with a small wavelength difference, such as a low pressure sodium vapour lamp [53].

2.3.2 Spatial Calibration

Spatial calibration is important because of the spatiospectral nature of the HS image. During acquisition, the influence of optical aberrations, sensor misalignments, or mechanical inconsistencies contributes to the possibility of misregistration and distortions. Proper referencing of each pixel of the HS image to the real world is crucial to avoid distortion of spectral information [54]. In the field of remote sensing, this procedure is referred to as orthorectification. A general algorithm that can be used to align the individual spectral band is, for example, a phase correlation [55, 56]. The core of this method is based on the Fourier shift theorem, which is shown in Equations 2.6 and 2.7:

$$f_2(x, y) = f_1(x - t_x, y - t_y), \tag{2.6}$$

$$F_2(\xi, \eta) = e^{-j2\pi(\xi t_x + \eta t_y)} F_1(\xi, \eta). \tag{2.7}$$

Given two signals (f_1, f_2), related by a translation (t_x, t_y), then their Fourier transforms (F_1, F_2) in frequency space (ξ, η) have identical magnitudes, but different phase shift [55]. This phase difference can be quantified using a normalised cross power spectrum, from which by the inverse Fourier transform, the translation (t_x, t_y) can be calculated.

The optical aberrations known from spectroscopy as smile and keystone also affect HS systems. Smile aberration refers to the curvature of spectral information along the spatial

direction, and the keystone causes the spatial curvature along the spectral direction. A good optical design can minimise these aberrations; however, they can also be compensated for using correction algorithms [57].

2.3.3 Radiometric Calibration

Radiometric calibration, a concept used in remote sensing, involves determining the reflectance or transmittance of the raw captured data. In addition, it improves the signal-to-noise ratio and addresses the intensity variations caused by the optical system and sensor [3].

In a camera image, noise refers to the overall spatial and temporal fluctuations in the captured signal under consistent, uniform illumination. Noise has multiple components: dark current noise, readout noise, photon-shot noise, and fixed pattern noise [58].

Dark current noise is the unwanted signal generated in the absence of light because of thermally excited electrons in the image sensor. Thermal energy is mostly responsible for the generation of the dark current, with higher temperatures resulting in higher kinetic energy of the electrons and stronger dark current, resulting in more noise. As a consequence, the dark current is additive noise and has the nature of a random quantity with a spectral density close to white noise [59].

In the same way, readout noise is additive and equally present in every captured image but is generally independent of temperature and is rather influenced by the construction of the analogue-digital converter connected to the image sensor. It is the output of the camera sensor without illumination and at zero integration time.

Photon shot noise originates from the discrete nature of light-carrying particles, photons. Therefore, it is not equally present in each acquired image and is independent, having a different effect across series of images. In addition, it is independent of the sensor temperature.

Fixed-pattern noise is present in every image but is caused by spatial non-uniformities of the pixels originating in imperfections in the sensor construction. It is independent of the level of the signal and the temperature of the sensor.

The additive noise that consistently influences every image captured by a camera (dark current, readout noise, and fixed pattern noise) can be reduced by subtracting a dark image I_{dark}, obtained when no light reaches the detector. It is important to capture the dark frame under the same conditions as the data frames, including exposure time, gain, and sensor temperature. A temperature-controlled camera is needed for precise and consistent data collection.

Aberrations introduced by optical construction and imperfections in the light path are compensated by using a reflectance calibration reference, such as Spectralon (Labsphere) or polytetrafluoroethylene (PTFE, Teflon), which is captured as a white reference image I_{white} [60]. In case of transmission imaging, a direct image of the illumination source is obtained, including the sample tray or microscope slide glass to ensure the capture of all aberrations that are part of the raw HS image.

The reflectance $R(x, y, \lambda)$ (or transmittance T) of the HS data can be calculated using Equation 2.8:

$$R(x, y, \lambda) = \frac{I_{\text{white}}(x, y, \lambda) - I(x, y, \lambda)}{I_{\text{white}}(x, y, \lambda) - I_{\text{dark}}(x, y, \lambda)} \mathbf{C}(\lambda), \tag{2.8}$$

where I is the measured (raw) HS image, and $\mathbf{C}(\lambda)$ is a scaling vector that depends on the material used as the white calibration reference. The calibrated reflectance/transmittance data can also be converted to absorbance, using the expression in Equation 2.9:

$$A = \log_{10} \frac{1}{R}, \tag{2.9}$$

where $0 < R \leq 1$. Transformation to absorbance has been known to decrease the non-linearity in the reflectance or transmittance measurements, to some degree [61].

Once calibrated, the HS data are ready for visualisation and subsequent processing, which is discussed in the following section.

Hyperspectral Data Calibration Highlights

- Calibration is essential for obtaining precise measurements by correcting for biases and variations specific to each HS instrument.
- Various calibration techniques are employed: spatial calibration corrects for geometric distortions, spectral calibration aligns the wavelength axis with known standards, and radiometric calibration adjusts for variations in sensor response.
- The application of suitable mathematical transformations and the use of calibrated reference materials guarantees a precise measurement of the spectral properties.

2.4 Hyperspectral Data Analysis

HS image processing and analysis require significant computing power due to the size of the HS data. For instance, a single HS image cube with dimensions of $I(x, y, \lambda) = (1000, 1000, 1000)$ pixels with 16 bit precision has a size of more than 1.86 GB. This poses high requirements for both the data storage capacity and the memory size used for processing. Data compression algorithms can be used to reduce file size for storage, and dimensionality reduction techniques can be applied to HS data to reduce band redundancy before any further processing is performed [62, 63]. These methods, including Principal Component Analysis (PCA), Non-negative Matrix Factorization (NMF), and Autoencoder (AE) neural networks, have been widely used to extract relevant information from high-dimensional HS data.

Moreover, because HS images contain spectral and spatial information, we can use both spectral analysis tools to examine individual spectral responses and image processing algorithms to extract spatial information from individual spectral band images. In recent years, there has also been a development in methods that use spatial and spectral information simultaneously [64, 65, 66].

However, before performing any of the analysis steps, it is a good practise to briefly examine the HS data to ensure that they have been correctly acquired, stored, and calibrated. In addition, visual inspection of the HS image can help identify obvious issues with data quality, such as invalid spectral responses or misaligned spatial regions.

2.4.1 Spatial Analysis

Operations performed using a single band or a set of individual bands $\mathbf{B}_i$ of the HS cube are characterised as spatial analysis of HS data. The following text describes simple examples of initial evaluation that can be performed on HS data through the application of spatial analysis techniques. This category also includes algorithms from image processing that are applied to individual spectral bands, such as thresholding, segmentation, edge detection, or other algorithms depending on the objective of the analysis. Great resources on classical image processing algorithms can be found in [26, 27].

False Colour Image

The false colour image provides an initial overview of the HS data by displaying selected wavelength channels. In a false colour image, red, green, and blue values of the digital image that can be easily displayed on the screen or printed correspond to a set of selected spectral bands of the HS image. The resulting image is similar to the staining process commonly used in microscopy [3]. A single band can also be used together with a colour-mapping function, which maps the scalar values of the single-band pixels to RGB values from a defined interval to better distinguish areas with similar spectral transmittance/reflectance in the examined band. I applied this approach in my bachelor's thesis, where I used multispectral imaging for the identification of falsified banknotes [67]. This can also be used, for example, to highlight the presence of vegetation in remote sensing images by choosing the red channel to display the near-infrared reflectance values [10], as shown in Figure 2.5b).

Spatial Indices

Spectral indices represent an alternative approach to visualisation and extraction of information from HS images, by performing arithmetic operations on selected spectral bands to highlight specific information, such as vegetation cover [68], water content [69] or chlorophyll concentrations [70]. A comprehensive list of spectral indices is presented in [71]. As an example, take the Normalised Difference Vegetation Index (NDVI), defined as the difference between the near-infrared (NIR) and red reflectance divided by their sum; see Equation 2.10:

$$\mathrm{NDVI} = \frac{R_{\mathrm{NIR}} - R_{\mathrm{RED}}}{R_{\mathrm{NIR}} + R_{\mathrm{RED}}}. \tag{2.10}$$

The result of this index calculation can be seen in Figure 2.5c), where the positive values that indicate the presence of vegetation are assigned to the green channel, and negative values corresponding to water are assigned to the red channel.

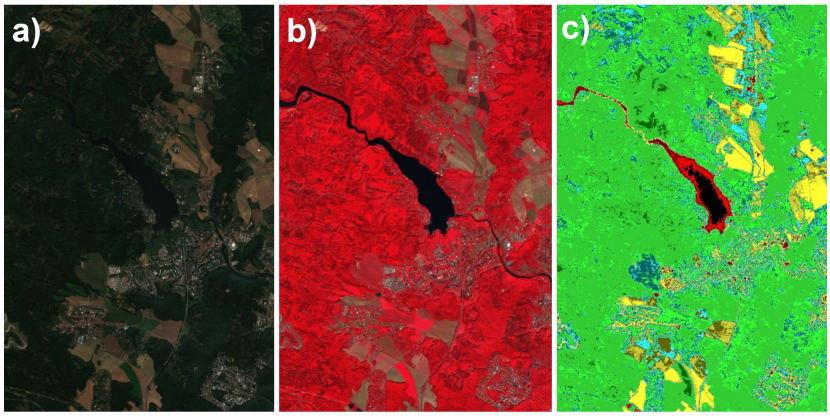


Figure 2.5: Satellite image of Brno Reservoir captured on 9 October 2022 by *Sentinel-2 L2A* in: a) natural colour (RGB), b) false colour infrared, c) Normalized Difference Vegetation Index (NDVI) [72].

To highlight specific features or to reduce the amount of data for further processing, we can select a subset of the HS data by applying image segmentation on the spectral index values or on the selected spectral band, and then working with the masked data that contain only the regions of interest.

When the HS data are refined in the spatial domain, the next logical step is to investigate along the axis perpendicular to the spatial dimensions of the HS cube, by which we mean the spectral domain.

2.4.2 Spectral Analysis

Algorithms that have an individual spectrum $S_{x,y}$ or a set of spectra $\mathbf{S}_{2\mathrm{D}}$ at their input can be categorised as spectral analysis algorithms. As a matter of fact, this domain overlaps strongly with the field of chemometrics in applications that make use of and analyse spectroscopic signals [25]. The spectral analysis algorithms can be applied to individual spectral pixels of the HS data and the results can then be visualised in the spatial domain, uncovering the information that can be hidden in the rather information dense spectral dimension. The following is a selection of algorithms used throughout this work, more on spectral analysis can be found in [61, 73].

Spectral Unmixing

Spectral unmixing can be used when a mixture of spectral signatures is present in the measured spectra. The goal of unmixing is to break down the measured spectra into pure component spectra (referred to as endmembers in remote sensing literature) and their fractional abundances. The main principle relies on the assumption that the observed spectrum is a linear (or non-linear) combination of the endmember spectra. Equation 2.11 describes a linear spectral unmixing model of n endmember spectra $e_i(\lambda)$ with coefficients (abundances) a_i and residual errors ε_i [74]:

$$S_{x',y'}(\lambda) = \sum_{i=1}^{n} a_i e_i(\lambda) + \varepsilon_i. \tag{2.11}$$

The equation mentioned above describes an unconstrained linear spectral unmixing model. In practice, however, it is more advisable to impose a set of constraints for the unmixing algorithm to achieve better results. The first possible constraint is the abundance sum-to-one constraint, defined in Equation 2.12:

$$\sum_{i=1}^{n} a_i = 1. \tag{2.12}$$

The second commonly used constraint is the abundance non-negativity constraint, stating that

$$a_i \geq 0, \quad \forall i. \tag{2.13}$$

By enforcing these two constraints, the resulting abundances can be represented as percentages, providing a direct and intuitive measure of the compositional distribution/concentration of endmembers within a HS image. This approach offers a clearer link between the image data and the reference (pure) spectra, facilitating a more straightforward interpretation of the composition of the object.

The benefits of this method include the estimation of the proportions of various materials within a pixel or sample, even when spatial resolution is limited. Drawbacks involve the necessity for precisely measured/identified endmember spectra and the possibility of errors arising from non-linear mixing or changes in endmember composition.

Without prior information on the sample, the spectral components (endmembers) can be extracted using spectral unmixing algorithms ranging from deterministic algorithms such as N-FINDR, Fast Iterative Pixel Purity Index, and machine learning-based approaches [75, 76, 77].

Spectral unmixing and mixture analysis were also the topic of my research paper titled *Rapid Identification of Pathogens in Blood Serum via Raman Tweezers in Combination with Advanced Processing Methods* [78]. This paper presents a rapid and reliable method of microbial identification in blood serum using Raman tweezers together with a processing pipeline that does not require user input, reduces possible user influence, eliminates the need for designing complicated analysis and achieves comparable (or higher) accuracy than previously used approaches. The key results of this paper are presented in the Supplementary Table S2 and Figure 2.6. The diagonal elements of the confusion matrix show the accuracy for each class (five classes – four for microbes in the serum and one for pure serum). The off-diagonal values show how the model failed in classification. The first row of the left part of Fig. 2.6 indicates that the algorithm based on FCLS spectral unmixing correctly distinguishes the spectra of blood serum (without microbes) from the spectra

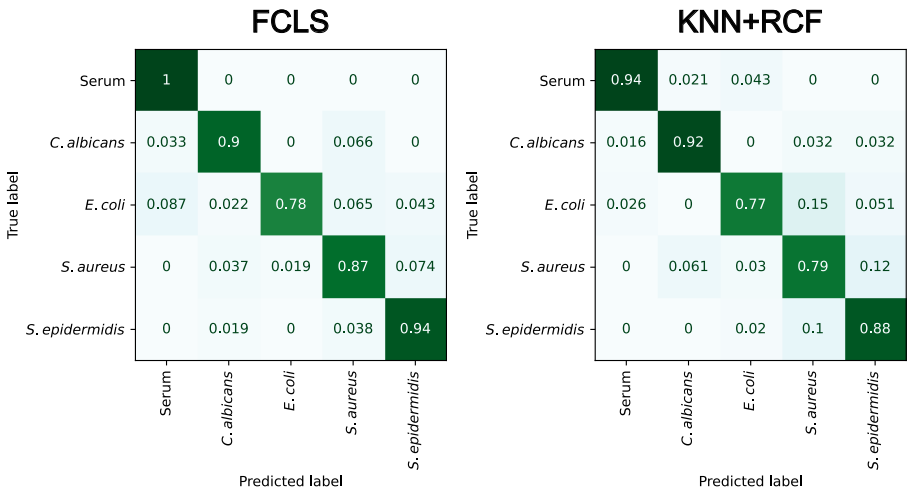


Figure 2.6: Averaged confusion matrix for identification of microbes directly in blood serum using Raman spectral data classified using the custom Fully Constrained Least Squares (FCLS) spectral unmixing based algorithm (left part) and K-Nearest Neighbors with Rolling Circle Filtering used for background correction (KNN+RCF). Data are normalized by rows (true labels) and averaged from data measured in five distinct blood sera.

of trapped cells in the blood serum. The results of the Fully Constrained Least Squares (FCLS) unmixing-based analysis indicate that the algorithm is perfectly capable of distinguishing pure serum (having no false positives, further described in Section 2.4.5), which is not the case when using the K-Nearest Neighbors with Rolling Circle Filtering background correction (KNN+RCF) algorithm (right part of Fig. 2.6). The code used for this analysis is available in [79].

The results and algorithm published in this study are also applicable to HS data, which is shown in the Methods section of this thesis. Additionally, if there is no prior information of the pure spectral components, they can be identified automatically using algorithms mentioned above or via Multivariate Curve Resolution, which is discussed in the next section.

Multivariate Curve Resolution

Multivariate Curve Resolution (MCR) is a group of techniques that originated in chemometrics with the goal of resolving a data matrix into pure component spectra and concentration profiles without prior knowledge of pure components [80, 81]. The main principle is to iteratively optimise the spectral and concentration profiles by applying suitable constraints, such as non-negativity and sum-to-one constraints discussed in the previous section. This method is used by analytical chemists, spectroscopists, and process engineers to analyse complex mixtures and reaction kinetics. Advantages include the ability to extract pure component information from complex data without requiring prior knowledge of the components. Disadvantages include the potential for ambiguity in the solution and the sensitivity to noise and initial estimates.

The core principle of MCR is described using Equation 2.14, given a dataset $\mathbf{D}$, which contains pure and mixed spectral signatures (in our case a reshaped HS data cube $\mathbf{S}_{2D}$):

$$\mathbf{D} = \mathbf{C}\mathbf{S}^{\mathbf{T}} + \mathbf{E}, \tag{2.14}$$

where the matrix $\mathbf{C}$ contains the concentrations of all components (grouping the abundance coefficients a_i from spectral unmixing) and $\mathbf{S^T}$ contains the related pure spectra (endmembers e_i). $\mathbf{E}$ is the matrix that quantifies the error or the variance that is not explained by the model (analogous to ε_i).

Non-negative Least Squares

Non-negative Least Squares (NNLS) algorithm finds the least squares solution to a system of linear equations with the constraint that the solution vector must be non-negative [82]. The main principle is to minimise the residual between the observed and predicted values while ensuring that all coefficients are positive or zero. NNLS is used in various fields, including image processing, signal processing, and chemometrics, to solve linear regression problems with non-negativity constraints. Advantages include ensuring compatibility with measured physical quantities by preventing negative values in the solution. Among the disadvantages is an increase in computational complexity compared to standard least squares and the potential for biased solutions if the true solution contains negative values, which is not the case for HS data. NNLS is used as a method for calculating the mixing coefficients a_i (abundance, concentration ratios) in spectral unmixing and MCR.

Because the HS data are composed of a larger number of spectral vectors compared to Raman spectroscopy, processing of all spectral signatures from the HS cube can be computationally demanding. To overcome this, we can reduce the dimensionality of the HS data in either the spatial and/or spectral domain.

2.4.3 Dimensionality Reduction for Hyperspectral Data Analysis

The volume and high dimensionality of the HS data introduce significant computational challenges and the statistical problem known as "curse of dimensionality" [83]. This phenomenon can lead to increased processing times, a higher likelihood of overfitting in classification models, and the need for an impractically large number of training samples. Consequently, dimensionality reduction emerges as a critical and often indispensable preprocessing step in the analysis of HS data. The primary goal is to transform the high-dimensional dataset into a lower-dimensional representation, while retaining the most significant information necessary for subsequent analysis.

This section explores several important techniques for achieving this, beginning with the straightforward approach of pixel binning. More sophisticated, data-driven methods that seek to uncover underlying structures, including the parts-based representation offered by Non-Negative Matrix Factorisation and the widely used statistical transformation of Principal Component Analysis, are examined. Lastly, a powerful non-linear approach is explored through the application of autoencoder neural networks, which have demonstrated considerable promise in learning compact and highly informative data representations.

Pixel Binning

Pixel binning is a technique employed to reduce the dimensionality of high-resolution images by aggregating pixel values into a smaller set of representative bins [84]. It is a crucial step in many image processing pipelines, particularly when dealing with large data sets or computationally intensive analysis, as it significantly reduces storage and processing requirements. This method acts using a grid of pixels, usually a 2×2 or 4×4 pixels, which are combined to form one pixel of the binned image, as shown in Figure 2.7.

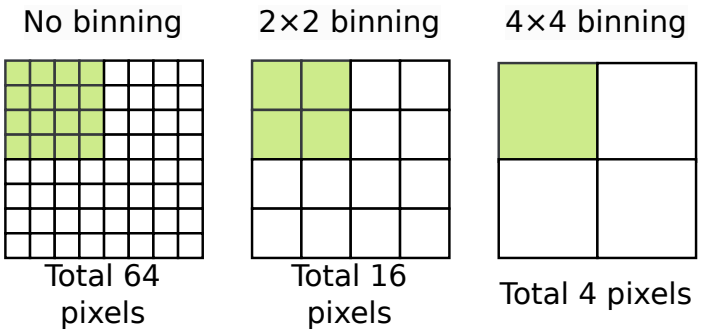


Figure 2.7: Example of image with no binning, 2×2 and 4×4 pixel binning.

The advantage of this process is the increase in the accuracy of the measured intensity by summing or averaging the values of the binning region. The disadvantage is the loss in the spatial or spectral resolution, when binning is used for HS data. However, binning can be acquired only along a selected axis, therefore, affecting only the spatial or spectral domain of the HS image.

The binning process is a form of subsampling of the data, reducing the dimensionality without any knowledge or influence of the underlying HS data it processes. The following sections describe algorithms that can be used to reduce the dimensionality of the data based on their content, by performing what in the field of machine learning is referred to as feature extraction.

Non-negative Matrix Factorisation

Non-negative Matrix Factorisation (NMF) decomposes a non-negative matrix into two non-negative matrices, representing parts-based representations of the original data [85]. The main principle is to approximate the data matrix $\mathbf{D}$ by a product of two lower-rank matrices $\mathbf{W}$ and $\mathbf{H}$, with all elements bound to be non-negative; see Equation 2.15:

$$\mathbf{D} \approx \mathbf{WH}. \tag{2.15}$$

NMF is used by data scientists, image processing researchers, and bioinformatics analysts to extract meaningful parts or features from non-negative data, such as images and spectroscopic data. Advantages include its ability to produce interpretable, parts-based representations and its suitability for data with inherent non-negativity. Disadvantages include the non-uniqueness of the solution and sensitivity to initial conditions. The use of NMF for the HS data is possible because of the inherent positivity of the measured intensities. The NMF algorithm is often related to the task of spectral unmixing, as discussed in [86, 87].

Principal Component Analysis

Principal Component Analysis (PCA) transforms a set of correlated variables into a set of linearly uncorrelated variables called principal components, which are ordered by their variance. The main principle is to find orthogonal directions (principal components) that capture the maximum variance in the data, effectively reducing the dimensionality [88].

PCA is an unsupervised learning method and is similar to clustering, as it identifies patterns without relying on prior information about samples [89]. PCA is popular in various disciplines as a tool to reduce dimensionality, visualise high-dimensional data, and identify the underlying patterns. Advantages include its ability to reduce dimensionality while preserving most of the data's variance and its computational efficiency. Disadvantages include the assumption of linearity and sensitivity to scaling and outliers. When applied to HS data, PCA is useful for identifying spectral bands with high variance across the pixel value, which may indicate a band useful for image segmentation using traditional image analysis methods. Likewise, it can be used to reduce the dimensionality of the HS image, by transforming the data along the spectral axis to the principal component space and selecting only a subset of themfor further processing [90].

Autoencoder Neural Network

Autoencoder (AE) is a type of neural network that aims to learn a compressed representation of the input data [91]. The main principle is to minimise the reconstruction error between the input and output, forcing the network to learn the essential features of the input data. The basic design of an AE network is shown in Figure 2.8.

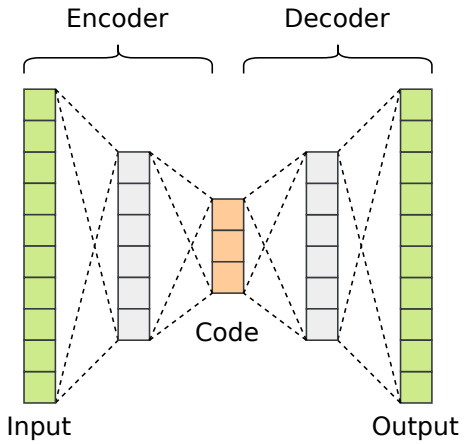


Figure 2.8: Schematic representation of AE network. The output of the encoder, the code, can be used as a dimensionally reduced representation of the input data. The decoder is used during the training and can be utilised as a reconstruction algorithm (reverse of the encoder transformation).

AEs are used in various applications, including dimensionality reduction, anomaly detection, and image denoising. Advantages include the ability to learn complex, non-linear relationships in the data and to extract meaningful features automatically. Disadvantages include the need for large datasets for training and the potential for overfitting if the network is not properly regularised. AEs can facilitate lossy compression by designing a network with desired latent space dimension (which is directly connected to the compression ratio) and training the network on the entire HS image. The building blocks of an AE network can be fully connected layers (FC-AE) or convolution layers (CNN-AE), taking the advantage of the downsampling they perform on the input data [92].

Supplementary Table S3 compares the dimensionality reduction algorithms discussed so far, listing their constraints, interpretability, linearity, and more.

2.4.4 Error Measures

Error measures play a crucial role in evaluating the accuracy and performance of various analytical methods, where comparing predicted and actual values is essential. The following part briefly describes the error measures used in this work. The following equations use $\mathbf{x}$ to mark the predicted value, $\mathbf{y}$ to mark the true value, N for the number of measurements, and $\|\|$ to denote the Euclidean norm.

Root Mean Squared Error

Root Mean Squared Error (RMSE) quantifies the average magnitude of the errors, providing a measure of the overall accuracy of the prediction by calculating the square root of the mean squared difference between the predicted and observed values, as presented in Equation 2.16. It is used to evaluate the absolute accuracy of their models, with the advantage of being sensitive to large errors but the disadvantage of being influenced by outliers [93].

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\|\mathbf{y} - \mathbf{x}\|^2} \tag{2.16}$$

R-squared

R-squared (R^2) measures the proportion of variance in the dependent variable that is predictable from the independent variable(s), indicating how well the model fits the data by calculating the squared correlation coefficient. The R-squared, calculated using the expression in Equation 2.17, is used to assess the goodness of fit of the regression models, which offers the advantage of easy interpretation, but the disadvantage of not reflecting bias in the model [93].

$$R^2 = 1 - \frac{\|\mathbf{y} - \mathbf{x}\|^2}{\|\mathbf{y} - \bar{y}\|^2}, \tag{2.17}$$

where $\bar{y}$ is the mean of the true values $\mathbf{y}$.

Spectral Angle Difference

Spectral Angle Difference (SAD) error metric quantifies the spectral similarity between two spectra by measuring the angle between them, providing a measure of the spectral shape difference; see Equation 2.18. It is used by remote sensing and spectroscopy analysts to compare spectral signatures, offering the advantage of being insensitive to illumination variations but the disadvantage of not accounting for magnitude differences [94].

$$\mathrm{SAD} = \cos^{-1}\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) \tag{2.18}$$

2.4.5 Performance Metrics

When evaluating machine learning models, a common metric is needed to quantify the results. If the goal of the model is to identify contents of a spectral mixture, the values of

the individual abundances/concentrations can be interpreted as indicators of the presence or absence of the components, effectively solving a classification problem. The performance of classifier models is commonly assessed using accuracy, precision, and recall metrics, among many others [95, 92]. To calculate these metrics, we first need to establish the four categories of classifier predictions: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP and TN are correct classification results with no mismatch with respect to the underlying data and their known classes/labels. FP output, also known as Type I error, occurs when a model incorrectly indicates the presence of a condition (for example, classifying a person as sick when in fact they are not), and FN marks the opposite case (a sick person is classified as healthy). In the following equations, the abbreviations of these categories indicate the total number of such outputs of the classifier.

Accuracy A is defined as the ratio of the correct classification to the total number of classifications, as stated in Equation 2.19. Accuracy is used as a general indicator of the performance of the prediction.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.19}$$

Precision P is the proportion of all positive classifications in the model that are actually positive; see Equation 2.20. A model with high precision may miss many actual positive cases, and low precision may lead to an increase in the number of false positives.

$$P = \frac{TP}{TP + FP} \tag{2.20}$$

Recall R, also known as the true positive rate or sensitivity, quantifies the number of true positive cases accurately recognised by the model from the total of all real positive cases; see Equation 2.21. When the recall is high, the model finds all positive cases very well, and if the recall is low, there is a risk of missing some of the true positive cases.

$$R = \frac{TP}{TP + FN} \tag{2.21}$$

Precision and recall are useful metrics for working with unbalanced datasets where FP or FN errors can be costly, depending on the circumstances. In specific cases, it is advisable to optimise one of these metrics to minimise negative prediction outcomes. For example, a fire detection system tuned for high recall will cause many false alarms, which may, for example, slow down production. In contrast, a system tuned for high precision would result in fewer false alarms but carry an increased risk of missing a real fire, which could cause costly damage to the equipment.

Hyperspectral Data Analysis Highlights

- HS data are highly storage and processing intensive.
- Spatial analysis offers an initial overview of the data based on prior knowledge.
- The spectral unmixing algorithms used in chemometrics are applicable to HS data.
- To speed up and/or enhance performance, we can reduce the dimensionality of the HS data.
- Dimensionality reduction can be achieved through PCA, NMF or autoencoder-based algorithms.
- The selection of a suitable error and performance metric is important when working with spectral data.

3 Methods

3.1 Experimental Setup

During my Ph.D. studies, I have focused not only on the analytical and processing part of the HS imaging but also on the development of the HS image acquisition hardware as well. In order to have full control over the measurement and processing part of the HS imaging, I cooperated on the development of an HS microscope workstation, together with Denisa Šilhanová and Lukáš Šilhan, who designed the mechanical assembly of the workstation, and the Photon System Instruments company, who took part in the imaging spectrometer construction.

The HS microscope workstation design is presented in Figure 3.1 and the photo of the experimental setup in Figure 3.2. Its key components are described in the following sections.

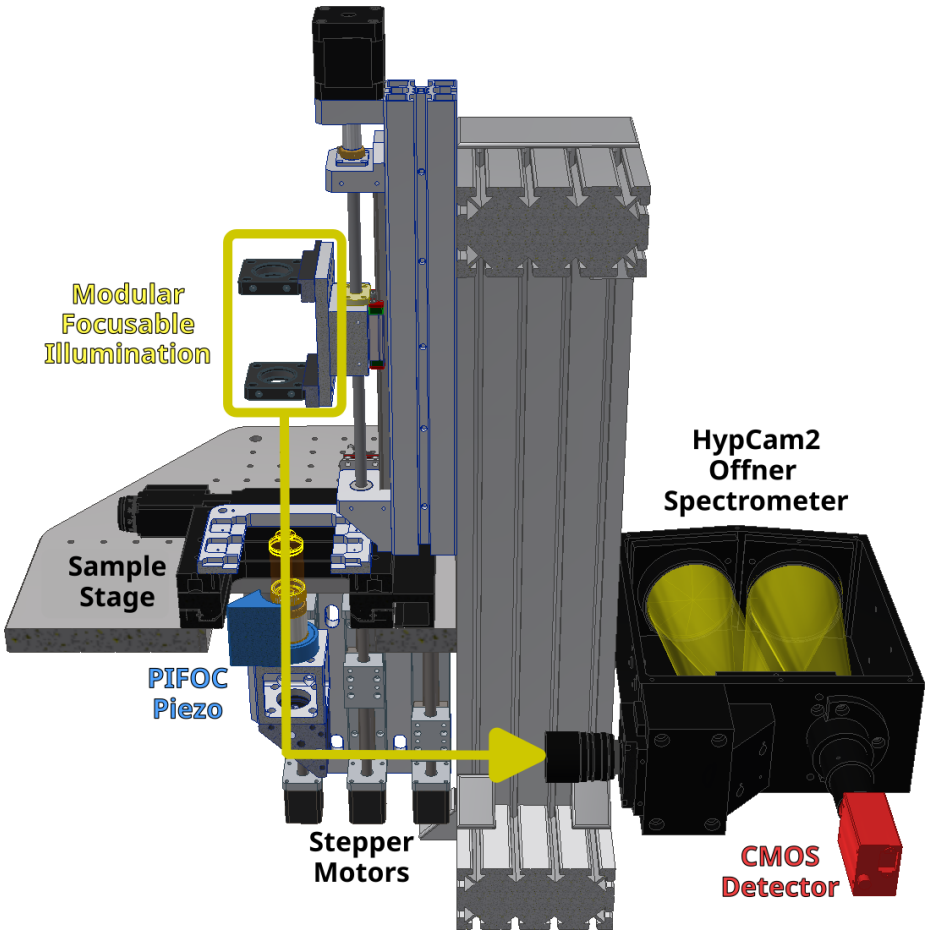


Figure 3.1: Cross section of the CAD model of HS microscope workstation, with key component description and light path denoted by yellow arrow. Image courtesy of Denisa Šilhanová and Lukáš Šilhan.

3.1.1 Illumination Unit Construction

With the aim of having optimal spectral and spatial illumination, I have selected two light sources with broadband wavelength emission and high output power. For measurement

in the visible part of the spectrum, Bridgelux BXRH-50S0601-A-74 LED is used, with a luminous flux of 675 lm and the majority of the radiant power emitted in the range of 425–675 nm [96]. For measurement in the near-infrared range, a Thorlabs SLS201LM stabilised tungsten halogen light source can be used, as it has a spectral output in the range of 360–2600 nm with peak power output at 1000 nm [97]. The configuration of the illumination unit in Figure 3.2 a) uses the Bridgelux LED with a custom designed aluminium heat sink. The LED is powered by a laboratory power supply (Manson NSP-3630), with an output controlled by a D4184 MOSFET module. The control voltage for the MOSFET module is supplied by the NI USB-6001 multifunction input/output device, which is controlled by the Python code described in Section 3.1.5. This combination of components serves as a modular LED driver, allowing us to use any LED light source and control it via software. The light source output is collimated with an aspheric lens (Thorlabs ACL2520U-B) and then focused through a microscope objective on the sample. The magnification of the focussing microscope objective is the same as the magnification of the lens used to observe the sample to ensure optimal illumination conditions. In the latest iteration, the entire illumination unit is mounted on a linear rail, which can be repositioned using a stepper motor with a precision lead screw (Misumi MSSR801-250) located at the very top of the Figure 3.1. This configuration allows for automated adjustment of the illumination to achieve optimal measurement conditions.

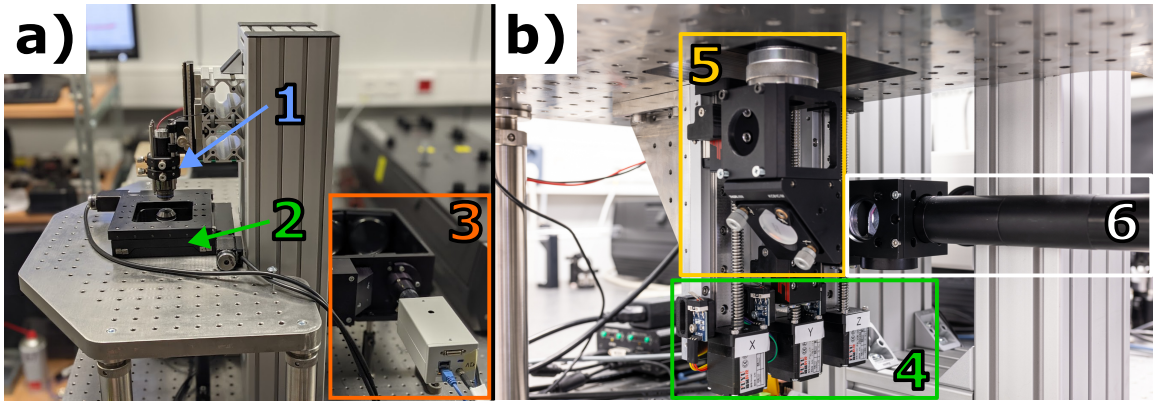


Figure 3.2: Photography of HS microscope workstation in laboratory at ISI CAS: a) top part with (1) illumination unit, (2) sample stage and (3) HS camera unit; b) Bottom part with (4) linear motorised mounts, (5) microscope objective mount with adjustable mirror, that reflects light to (6) preview camera with relay optics, or directly to HS camera, when the preview camera is moved up from the optical path.

3.1.2 Sample Stage

The observed sample is placed onto a standard 1 inch × 3 inch microscope slide, which is then secured to a holder with three side-by-side microscope slide slots. This slide holder is mounted on top of the motorised XY scanning stage (Standa 8MTF-75LS05), capable of a 75 mm × 75 mm movement range, and a step size of 2.5 µm that can operate with resolution up to 1/256 of the full step distance, the so-called microstep mode. The combination of a motorised stage and a holder capable of housing three microscope slides enables the measurement of multiple samples during a single automated measurement session, controlled using Python modules and scripts I have developed.

3.1.3 Objective Lens

The HS microscope workstation uses the standard RMS microscope objective mounting thread (W 0.8 inch × 36), enabling the use of a wide variety of microscopic lenses. The configuration at the time of writing uses the Nikon Plan Fluor 4× with a numerical aperture NA = 0.13 and a working distance of 17.1 mm. The objective lens is mounted in a high-precision piezo flexure stage (Physik Instrumente PIFOC P-721.20), allowing automated and precise focussing. The piezo flexure stage is driven using an Physik Instrumente E610.C04 piezo amplifier, which requires an analogue voltage input for position control. This is provided again by NI USB-6001 multifunction input/output device. The piezo-focussing element is connected to one of the three linear actuators (Fuyu FSK30, labelled Stepper Motors in Figure 3.1), with a travel stroke of 300 mm and a position accuracy of 0.05 mm.

Using these linear actuators, the microscope can be used in bright-field or HS capture mode. The bright field microscope mode is possible by viewing the sample through a tube lens and a separate preview camera (Basler acA1300-200uc), located at the end of the optical tube system in box 6 in Figure 3.2b). When the preview camera is moved away from the light path, HS measurements can be performed. The third actuator can be used for mounting a dichroic mirror and light source for fluorescence HS imaging. At the time of writing, the setup is optimised for transmission HS data measurement using an HS camera described in the following section.

3.1.4 Hyperspectral Camera

In the heart of the HS microscope workstation is the HypCam2 – a push-broom HS camera with an exceptional spectral resolution of 0.1 nm and a tunable wavelength range of 350–1100 nm with the measurement window bandwidth of 142 nm. HypCam2 was developed in cooperation with the Institute of Scientific Instruments of the Czech Academy of Sciences, v. v. i. and PSI (Photon Systems Instruments), spol. s r.o. company as part of the project titled *HyPerSpec – Development of Hyperspectral Camera for Biotechnology Applications and Element Analysis* (MPO TRIO FV40455) funded by the Czech Ministry of Industry and Trade. My contribution to this project was the development of a stepper motor controller, position control software, and the testing of the spectral resolution and stability of the HS system. Therefore, I am listed as one of the authors of the HypCam2 functional sample [98]. The uniqueness of the HypCam2 system can be clearly seen in the Supplementary Table S4, which provides a comparison of the HypCam2 camera with several commercially available HS systems. At the time of writing, the HypCam2 system outperforms all commercial systems in terms of spectral resolution. The HypCam2 design is based on the Offner imaging spectrometer configuration [39]; see Figure 3.3.

The trade-off for having a large spectral resolution is the reduction in measurement window bandwidth. We overcome this limitation by adding a diffraction grating positioning system that enables us to adjust the position and angle of the grating. High-precision positioning is possible thanks to the use of stepper motors connected to precision screws with 0.2 mm pitch. When the angle of the grating is changed, the wavelength range of the measurement window is shifted. The other two perpendicular axes serve as a fine-focussing adjustment. The finished assembly of the grating positioning system is presented in Figure 3.4.

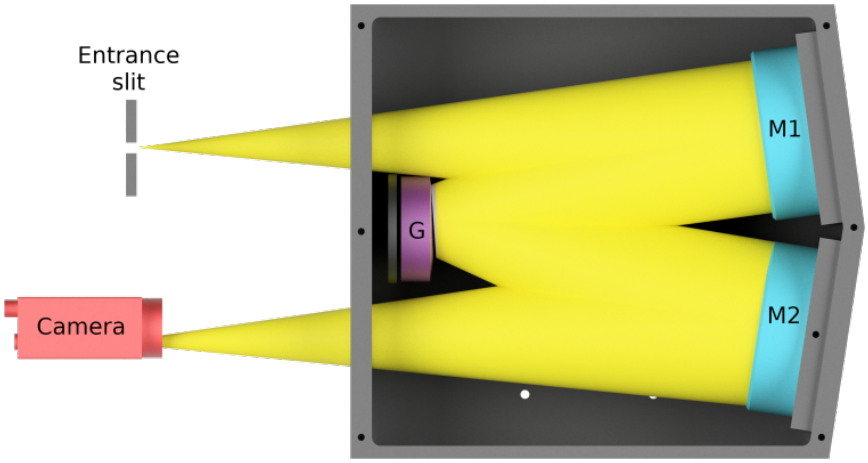


Figure 3.3: Mechanical design of the Offner type imaging spectrometer design; M1,M2 are concave mirrors, G is convex diffraction grating. The width of the entrance slit is 5 µm, the detector in use at the time of writing is a Sony IMX492 4/3" chip with 8288 × 5644 pixels, with a pixel size of 2.3 µm.

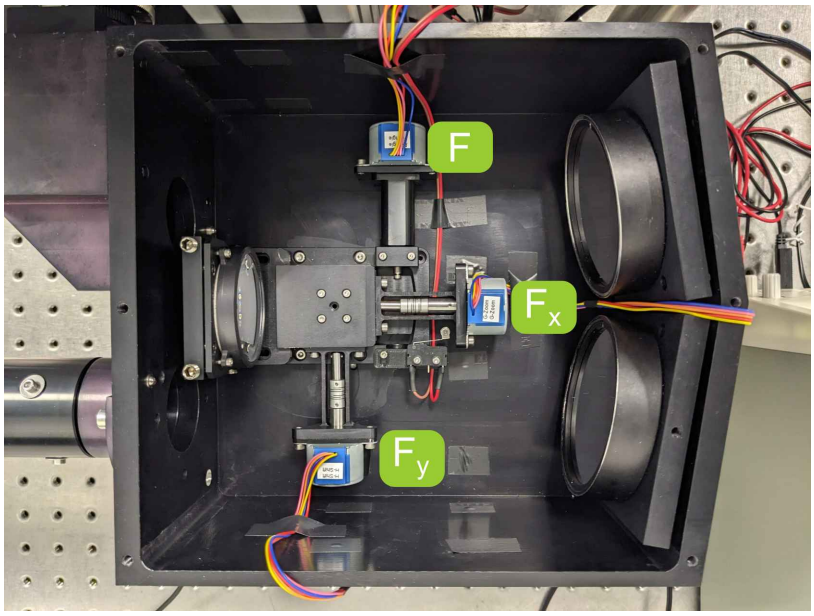


Figure 3.4: Diffraction grating positioning system. The F_x and F_y motors allow fine adjustment by displacing a two axis parallelogram placed under the grating mount, the F stepper motor controls the angle of the diffraction grating, enabling the change of measurement window bandwidth.

3.1.5 Software

To ensure repeatability and consistency of the HS acquisition, the entire process is controlled by custom software I have developed using the Python programming language. An excerpt of the measurement control script is presented in the Supplementary Listing S1. The entire codebase is available at [99]. The library used for data acquisition consists of the following modules:

- `cam_basler.py` – image acquisition using Basler cameras, developed using [100]
- `cam_psi.py` – IMX254 equipped camera communication module
- `cam_zwo.py` – IMX492 equipped camera communication module, developed using [101]
- `enviro.py` – Environmental sensor readout over USB serial
- `ni_daq.py` – Interface for the USB Multifunction I/O Device, developed using [102]
- `ramps_stepper.py` – motion control for stepper motors
- `stage.py` – XY stage control, developed using [103]
- `HS_capture.py` – main script for automated measurement
- `HS_capture_presets.py` – measurement presets for calibration and testing

The stepper motors inside the HS microscope workstation are driven using a Arduino Mega microcontroller with RepRap RAMPS 1.4 expansion board and Pololu A4988 stepper motor drivers. The Arduino board is running a custom control firmware that I have developed using the Arduino programming language. It uses serial port communication over USB with the `ramps_stepper.py` module, which is used to set individual stepper positions, as well as the speed and acceleration of movements.

A major part of the software side of working with the HS data is the storage format. The selected formats, standards, and choices behind them are presented in the following section.

3.1.6 Hyperspectral Data Collection

When communication with all components of the HS microscope station is verified and the acquisition of HS data is possible, it becomes essential to determine an appropriate storage format for both the produced data and its corresponding metadata. This was the main topic of my conference proceedings at the 14th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS) [104]. As a storage format for the captured frames, we use a Hierarchical Data Format (HDF) [105], accompanied by metadata stored in a JavaScript Object Notation (JSON) text file [106].

The HDF file format was chosen for optimised data storage and easy data sharing due to its compression capabilities and the hierarchical nature, allowing easy labelling of the calibration data and measured HS data. The dimension scales feature of the HDF file is also suitable for description of the axes and units of the HS data cube. Using the collected metadata, the measured data set satisfies the goal of an HDF file to be self-descriptive, which is an important factor when considering data reusability, one of the FAIR data principles (Findable, Accessible, Interoperable, Reusable) [107]. The structure of the captured data inside the HDF file is as follows:

- `hyperspectral` – group containing the measured data
  - `data` – raw HS data
  - `dark` – dark calibration frames
  - `white` – white (flat-field) calibration frames
  - `wavelength` – spectral reference calibration frames
- `metadata` – as JSON formatted string
- `preview` – from RGB camera

HS data are loaded using a `h5py` library stored in the NumPy array [108] with the indexing rules displayed in Listing 1:

```
import h5py

# opening the HDF file
hdf_file = h5py.File("YYYY-MM-DD-filename.h5")
# loading HS data to variable
hs_cube = hdf_file["hyperspectral"]["data"][:]
print(hs_cube.shape)
# outputs the axis sizes (bands, columns, rows)
```

Listing 1: Loading of raw HS data using the `h5py` package and printing its shape (size of the individual axes).

Consider an HS cube with the following dimensions: 2000 bands, 1200 columns, and 1000 rows. The shape of these HS data as a NumPy array is `(2000,1200,1000)`. Using the indexing rules and the implementation of the array slicing in Python, a single-point spectrum $S_{x,y}$ is equivalent to `hs_cube[:,x,y]`. A single band $\mathbf{B}_i$ can be selected using `hs_cube[:,:,i]`. A set of all spectral vectors $\mathbf{S}_{2\mathrm{D}}$ can be obtained by flattening the three-dimensional array to a lower dimension using the code presented in Listing 2.

```
bands, columns, rows = hs_cube.shape
S2D = hs_cube.reshape(bands, columns*rows)
```

Listing 2: Reshaping of HS data cube to $\mathbf{S}_{2\mathrm{D}}$.

The JSON format was selected for metadata storage due to its ease of readability by both humans and machines, its capability for validation, and interoperability due to the fact that it is an open standard file format, which is commonly used for data exchange [109]. Furthermore, the metadata we collect during our measurement can be easily transformed to the ENVI header file standard [52], which is widely used in remote sensing. This also allows for easy import of the HS data into other HS analysis software, such as Spectronon (Resonon Inc.) [110]. An example of our measured data opened in this software is shown in Figure 3.5.

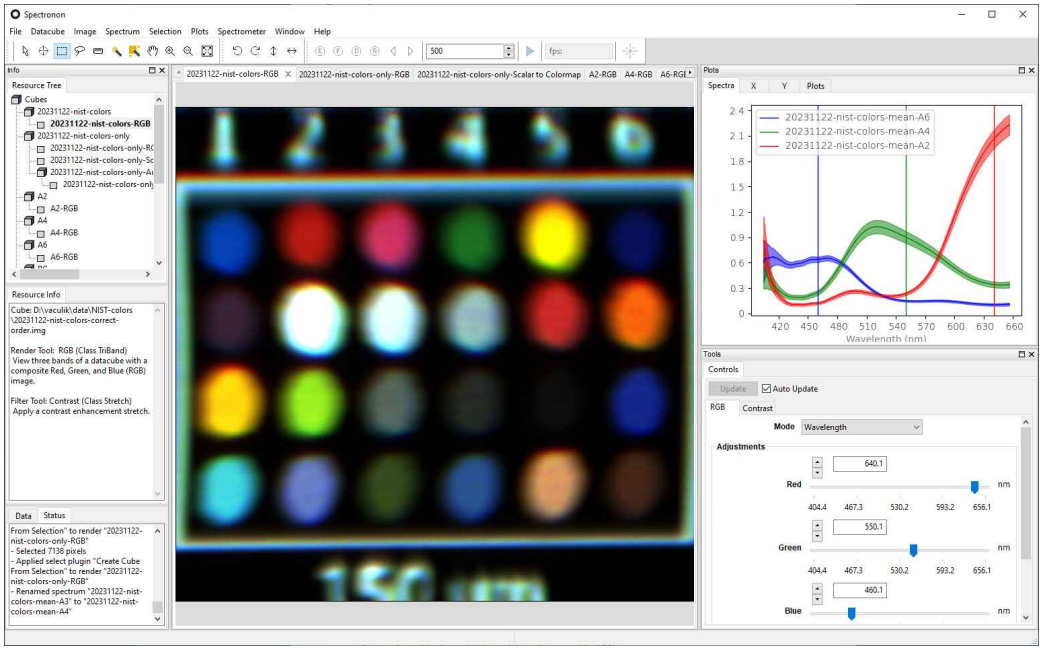


Figure 3.5: Processing of HS microscope acquired data in commercial software Spectronon (Resonon Inc.).

The structure of the metadata is defined using a JSON Schema [106], which is derived from the devices implemented in our measurement station to accurately record the measurement parameters, such as camera exposure parameters, positions of the elements of our custom automated positioning system, current system calibration, sample and author info, and more. An example of the captured metadata is presented in the Supplementary Listing S2.

In addition, environmental conditions (temperature, humidity, and pressure) are automatically recorded using a Bosch BME280 sensor connected to a microcontroller, which is used to read the measured values and relay them to the measurement software. For data management plan compliance, the JSON also contains information about the date, sample, author, project, and related funding. Finally, information on the current calibration is added, describing the pixel to wavelength mapping and spatial dimension scale.

The complete structure of the metadata is presented in Figure 3.6. Moreover, the use of well-defined structured JSON for storing the metadata allows for quick and easy catalogisation of the measured datasets using, e.g. a database system or a local index of measured datasets, which can then be uploaded to a public data repository (IEEE DataPort [111], Zenodo [112], etc.), fulfilling the requirements of findability and accessibility of the FAIR data principles. Another crucial aspect of the HS data is its size and corresponding storage demands, which are discussed in the next section.

3.1.7 Data Size Estimation

The correct assessment of the data size is a critical factor when working with HS data, as this directly influences the data storage requirements and the selection of the algorithm for the analysis of HS data. We can estimate the size of the HS image data based on the parameters of the HS measurement setup and the required sample size. We assume that the object of interest is located within an area of $22 \times 22\,\mathrm{mm}^2$, which is a standard size

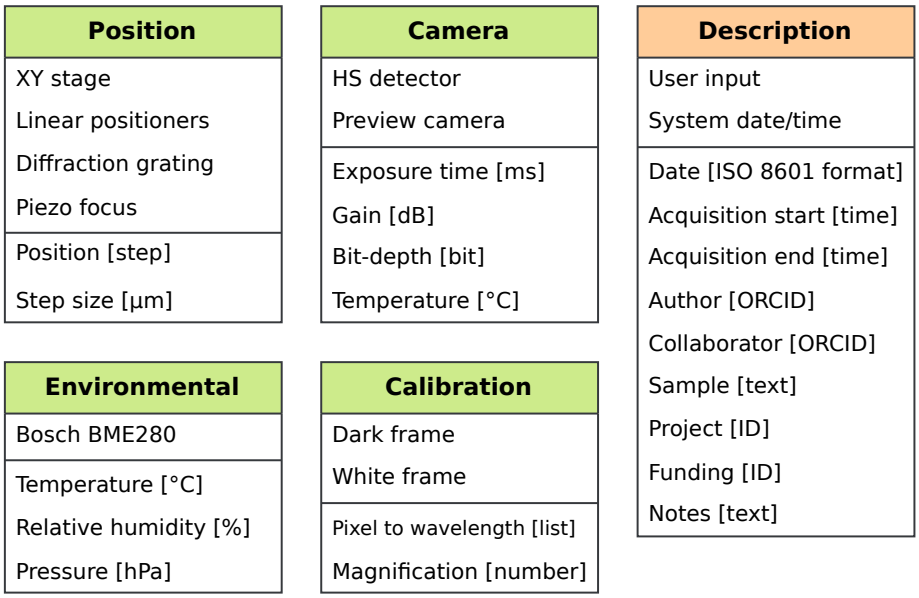


Figure 3.6: JSON metadata structure. The metadata are divided into five groups, each group contains a list of its data sources (top part of the boxes) and collected data format with corresponding unit/format in square brackets (bottom part of the boxes). The green labelled groups denote metadata that are collected automatically from the measurement hardware, the orange are collected from user input (except for date/time).

of microscope slide cover glass. When the microscope is equipped with an Olympus Plan N 10× lens, it provides us with a field of view 2 mm in width. For a complete scan of the cover glass, we need to perform 11 scans to cover the height of 22 mm. The pushbroom scanning of the HS camera requires a step size of 0.35 µm per frame when equipped with the 10× magnification lens. To cover the required area, a total number of 5714 frames is required. Given the number of pixels per frame and 14 bit resolution per pixel, the entire scan would require a total of 145 GB of storage, which is not a viable long-term approach.

In practice, a suitable compression algorithm is used when handling and storing any type of large data. We can divide the compression algorithms into two main categories – specialised and general. The study by Dua et al. offers a review of algorithms designed for HS data compression [63]. In our HS data management, we use both of the above-mentioned types of compression. As the first step in our HS data collection, we utilise general lossless compression based on the ZIP file format, as implemented in the `h5py` library [113]. Using this approach, we can reduce the total size of the example HS image to approximately 30 GB.

Such data sizes are more manageable for archiving, but less so for loading into the personal computer memory for further analysis, considering both the size of the data and the total number of dimensions. A possible solution to this problem is the use of specialised HS compression which takes advantage of the HS data structure. Several studies have shown that the HS data structure is susceptible to high levels of redundancy, both in the spectral and spatial domains [63, 91, 19].

This redundancy further highlights the importance of the dimensionality reduction techniques described in Section 2.4, such as NMF and AE, which effectively compress the HS

data while preserving its essential information content, thus allowing more efficient processing and analysis of these complex datasets. This method of lossy compression of HS data belongs to one of the recently explored strategies that uses a convolutional autoencoder (AE-CNN) neural network, as explored in [114] and [91]. The advantage of AE-CNN lies in a lower number of parameters, relative to the standard fully connected AE network, which in turn reduces the required training time and computational complexity. The trained AE-CNN can then be used for HS data compression with significantly lower computation requirements than those needed for the training process. Another benefit of AE-CNN encoding is that the compression ratio can reach 30:1 with low reconstruction error, as shown in Table 3.1. In our HS image example above, we can reduce its size to 4.9 GB using AE-CNN encoding with the above-mentioned compression ratio. The advantage of this method is that when applied only to the spectral dimension, it does not decrease the spatial resolution of the image.

Table 3.1: A summary of compression algorithms. LZF and GZIP are compression algorithms implemented in [113]. CR stands for Compression Ratio, MSE is mean squared error. Adapted from [91].

Algorithm	**CR**	**MSE [%]**	**Lossless**
None	1:1	0	✔
LZF	2:1	0	✔
GZIP	4.8:1	0	✔
AE	20:1	5×10^{-4}	✖
	30:1	3.4×10^{-3}	✖
AE-CNN	20:1	7×10^{-5}	✖
	30:1	1.3×10^{-3}	✖

Experimental Setup Highlights

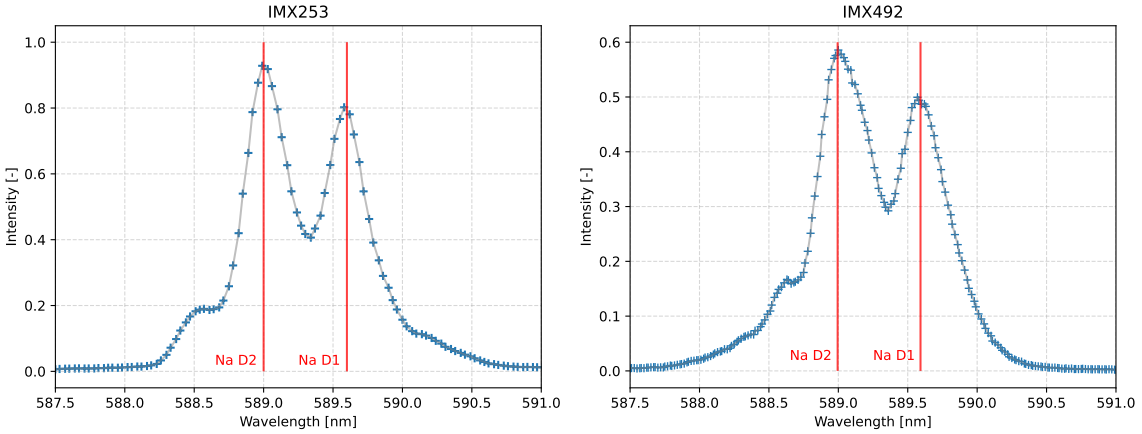
- HS microscope platform was developed at ISI CAS as part of my Ph.D. studies.
- It enables measurement with high spectral resolution across a wide band thanks to a tunable measurement range.
- The microscope supports automated measurement and data calibration using a software I have developed during my Ph.D.
- The resulting improved HS camera was developed as part of the project titled *HyPerSpec – Development of a Hyperspectral Camera for Biotechnology Applications and Element Analysis*, funded by the Czech Ministry of Industry and Trade (FV40455).

3.2 Performed Measurements

Details about the performed calibration and sample measurement, along with the corresponding analysis results, are provided in the following section. In the first measurement, we quantified the spectral resolution of the HS camera using spectral reference, along with the characterisation of the smile error. The second measurement characterised the imaging and spectral capabilities of the HS microscope by capturing a colour calibration test. The third measurement focused on the identification and unmixing of marine cyanobacteria.

3.2.1 Spectral Resolution Verification

To verify the spectral resolution of the HS system, we used a low pressure sodium vapour lamp (Helago-CZ 4035), which has strong emission of two neighbouring D-line transitions D1 and D2 at 589.592 nm and 588.995 nm respectively; see Figure 3.7.



(a) IMX253 (b) IMX492

Figure 3.7: Spectral resolution verification of the two HS camera sensors using low pressure sodium vapour lamp.

The measurement was performed using two different CMOS sensor types. The resulting parameters of these two configurations are listed in Table 3.2. The bandwidth of the configuration measurement window is indicated as B. The dispersion per pixel of the system $\Delta\lambda$ is calculated in Equation 3.1:

$$\Delta\lambda = \frac{B}{n}, \tag{3.1}$$

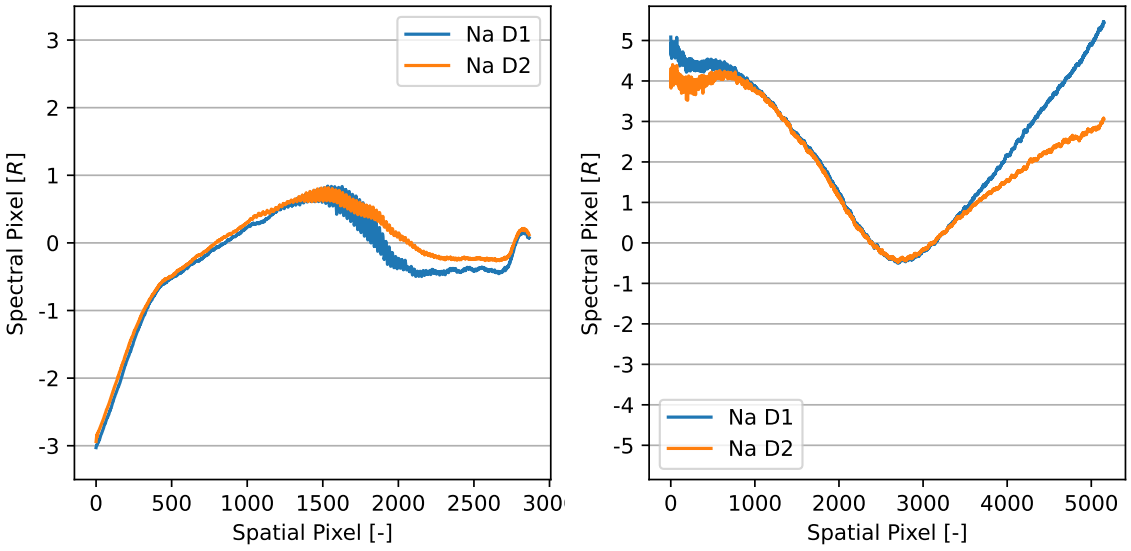
where n is the number of pixels along the spectral axis of the sensor.

For a reliable measurement of the spectra, more than one pixel value should be used [115]. In our case, three neighbouring pixels are used, and the final spectral resolution is denoted as R_λ.

We also examine the influence of smile distortion, which is common for imaging spectrometers. Smile distortion is a shift of the spectral position over the field of view of the sensor (a curvature of the spectral axis). We evaluated this by iteratively fitting a pair of Voigt profile curves to locate the centres of the D1 and D2 spectral lines for each position on the CMOS sensor. Then we plotted the central positions for both spectral lines of the fitted spectral profiles; see Figure 3.8.

Table 3.2: HS camera configuration comparison based on the used CMOS sensor.

Sensor	n [-]	λ_{min} [nm]	λ_{max} [nm]	B [nm]	$\Delta\lambda$ [nm]	R_λ [nm]
IMX253	4112	508.7	651.32	142.61	0.035	0.104
IMX492	8288	478.13	619.89	141.76	0.017	0.051



(a) IMX253

(b) IMX492

Figure 3.8: Smile distortion for the two configurations of the HS camera. The horizontal lines denote the distance in multiples of spectral resolution R_λ.

For both configurations, the smile distortion is five spectral pixels or less, which compared to the total number of pixels of the spectral axis of the sensor yields an uncertainty in the detected wavelength of 0.29% for IMX253 and 0.22% for IMX492 CMOS sensor. Smile distortion can be further suppressed by the use of pixel binning, as described in Section 2.4.3. In such a case, the spectral uncertainty would be less than the size of one pixel after binning.

3.2.2 Colour Calibration Slide

Accurate colour calibration is fundamental in various imaging applications, particularly in microscopy or HS imaging systems, to ensure reliable and reproducible results. A critical tool for achieving this accuracy is a reliable reference standard, such as the National Institute of Standards and Technology (NIST) Traceable Colour Transmission Calibration Slide (Applied Image IAM-9C-00348). This section will explore the process of acquiring and analysing HS data from the IAM-9C calibration slide, examining how techniques such as dimensionality reduction and spectral unmixing methods can be applied to this dataset, leveraging both the acquired HS images and the high-accuracy NIST-traceable reference spectral transmission data.

The IAM-9C calibration slide is designed with 24 unique colour patches of varying sizes and a large clear area, all set against a black background. The layout of the calibration slide is presented in Figure 3.9.

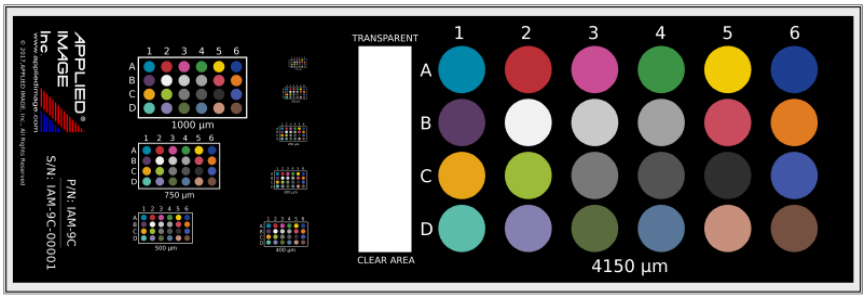


Figure 3.9: IAM-9C colour calibration microscope slide. It consists of 24 colour patches in different sizes. The slide also contains large clear area in the centre that can be conveniently used for flat-field correction.

The largest 4150 µm patches are individually calibrated for spectral transmission directly traceable to NIST. The reference transmission spectra provided in the accompanying data sheet of the IAM-9C-00348 are plotted in Figure 3.10. This spectral data also allows conversion to various colour spaces like RGB, LAB* or XYZ.

There are many reasons for employing this NIST traceable standard in HS measurement and subsequent analysis, particularly concerning the influence of dimensionality reduction (DR) on spectral unmixing. First, the reference spectra from the NIST calibration serve as a reliable reference. Secondly, these reference spectra can be used as a ground-truth database for the validation of spectral unmixing algorithms. This allows for objective evaluation of algorithm performance on a known standard. Additionally, the reference spectra are suitable for characterisation of the detector response function of the HS microscope system. Lastly, the layout and known spectral properties of the patches are valuable for evaluating spectral processing algorithms, including spectral clustering and classification, and critically, for investigating how DR affects the results of spectral unmixing algorithms on this well-defined dataset.

HS Data Capture and Preview

HS scanning data from the colour calibration sample were acquired using a Olympus Plan N 10× objective lens. The HS camera at the time of measurement was using a Sony IMX235 CMOS sensor without active cooling, and the high resolution upgrade discussed in

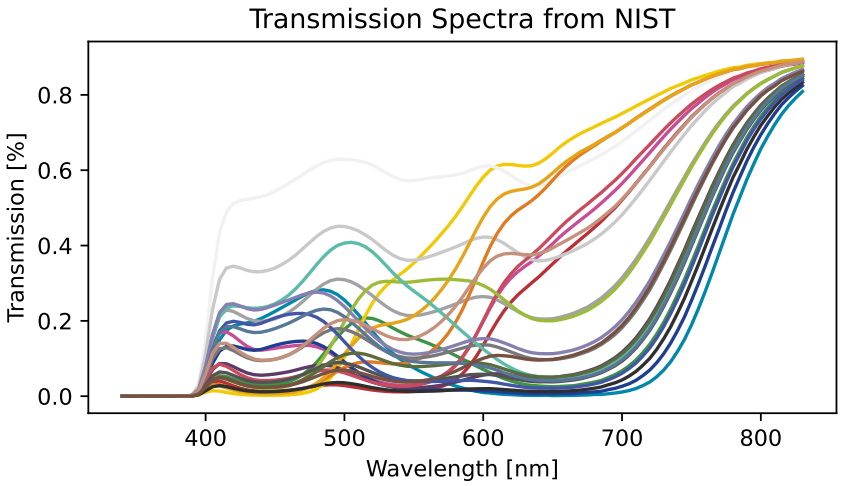


Figure 3.10: Transmission spectra measured by NIST for each of the colour patches. The colour of the line matches the colour of the patch of the calibration slide.

Section 3.2.1 was not yet installed. The corresponding metadata of the HS measurement are listed in Table 3.3.

Table 3.3: Metadata of HS measurement of NIST IAM-9C-00348 colour calibration slide.

Parameter	**Value**	**Unit**
Exposure	50	ms
Gain	5	dB
Binning	4	×
λ_{min}	404.42	nm
λ_{max}	656.07	nm
$\Delta\lambda$	0.35	nm
Camera Temperature	N/A	°C
Ambient Temperature	20	°C
Pressure	1002	hPa
Humidity	38	%
Date	2021-10-27	YYYY-MM-DD
Time	13:21:55+02:00	HH:MM:SS+offset

An overview using a RGB image composed of selected spectral bands of the HS image presented in Figure 3.11. The squares indicate the regions of interest (ROI) from which the spectral information is used. Their position was determined using the watershed algorithm [116] applied on the central wavelength of the HS image ($\lambda = 530.42\,\text{nm}$) followed by a calculation of the centroid coordinates of the identified regions, which were used as the centre of a pixel square 60×60 that defines the ROI. A comparison of the mean transmission spectral responses with the reference spectra is presented in Figure 3.12. Furthermore, all transmission spectra of the ROI of the individual coloured patches with the corresponding standard deviations are presented in the Supplementary Figure S1.

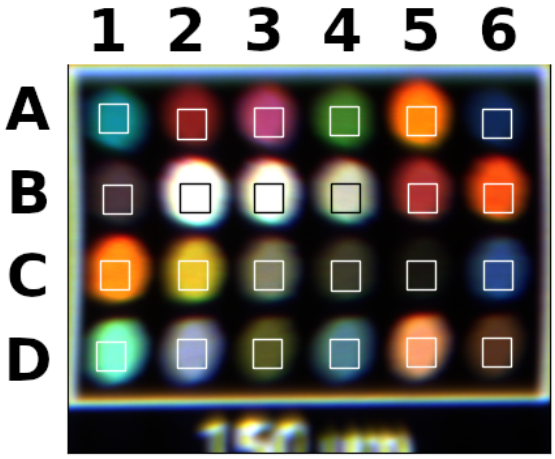


Figure 3.11: Preview of the dataset: HS image of colour calibration slide, shown using 640 nm band as red, 550 nm as green and 460 nm as blue channel. The squares indicate areas from which the spectral information is used for further analysis. The rows are labelled A, B, C, D from top to bottom. The diameter of the colour patches is 150 µm.

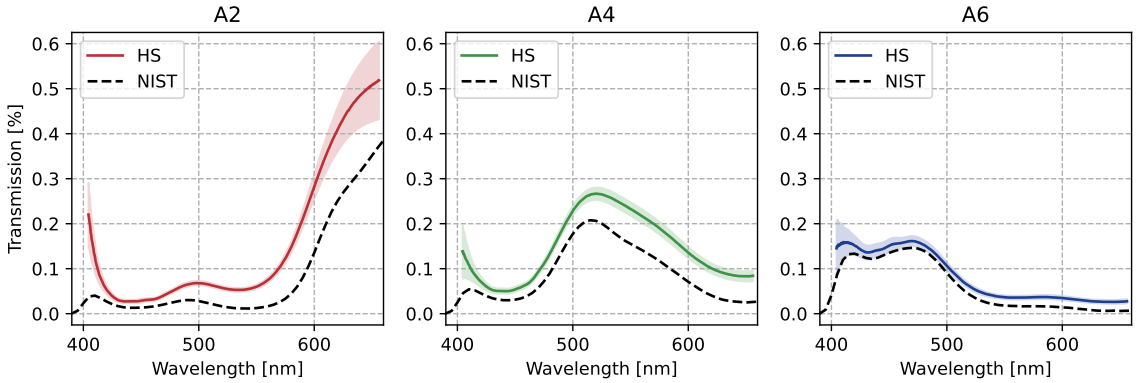


Figure 3.12: Comparison of the measured HS data and NIST reference spectra of A2 (red), A4 (green) and A6 (blue) colour patches.

Supervised Classification

To evaluate the complexity of the data, a supervised classification was performed. The goal was to train a model on the reference spectra provided by the manufacturer and test the trained model on the measured HS data, with the prediction target of RGB values from the wide gamut RGB colour space (Adobe Systems), defined for each colour patch by the calibration slide manufacturer. To ensure the compatibility of the reference spectra with the measured spectra, we interpolated the values of the reference spectra using quadratic interpolation and cropped them to match the wavelength region and sampling of the HS data. The training set consisted of NIST reference spectral signatures of patches A1 to A6 plus the spectrum of the clear (blank) part of the slide. The two test sets that were used were the rest of the colour patches of the NIST reference spectra (B1 to D6), and the spectral vectors of regions of interest in the HS data, which were introduced in Figure 3.11. The results of three supervised classification algorithms – Partial Least Squares Regres-

sion (PLSR), fully connected Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) applied to classify this problem are listed in Table 3.4 below.

Table 3.4: Comparison of supervised learning classification using three different algorithms: Partial Least Squares Regression (PLSR), fully connected Artificial Neural Network (ANN) and Convolutional Neural Network (CNN). All algorithms were trained on the first row of the reference spectrum and evaluated on both the reference spectrum (NIST) and data from the HS microscope (HS). The measure of accuracy is a R^2 metric together with Root Mean Squared Error (RMSE).

Algorithm	*NIST*		*HS*	
	R^2	RMSE	R^2	RMSE
PLSR	0.663	0.103	0.713	0.104
ANN	0.659	0.103	0.706	0.106
CNN	0.798	0.080	0.426	0.137

Upon inspection of Table 3.4, we can observe that PLSR outperforms the neural network methods on the measured HS data but performs worst on the reference dataset. The better performance of PLSR indicates that the data have a more linear dependence/trend in them. The performance of non-linear models (ANN, CNN) on the HS data achieves similar or lower accuracy. The signal-to-noise ratio in the HS data could favour simpler models, as neural networks might be overfitting to noise.

The Analysis Pipeline

The HS data analysis pipeline implements a comprehensive spectral unmixing analysis architecture for HS data, combining DR techniques with spectral unmixing algorithms, which were described in Section 2.4. The methodology consists of several key components that work in sequence. The entire pipeline is visualised using a schematic in Figure 3.13. An implementation excerpt of this pipeline is presented in the Supplementary Listing S3.

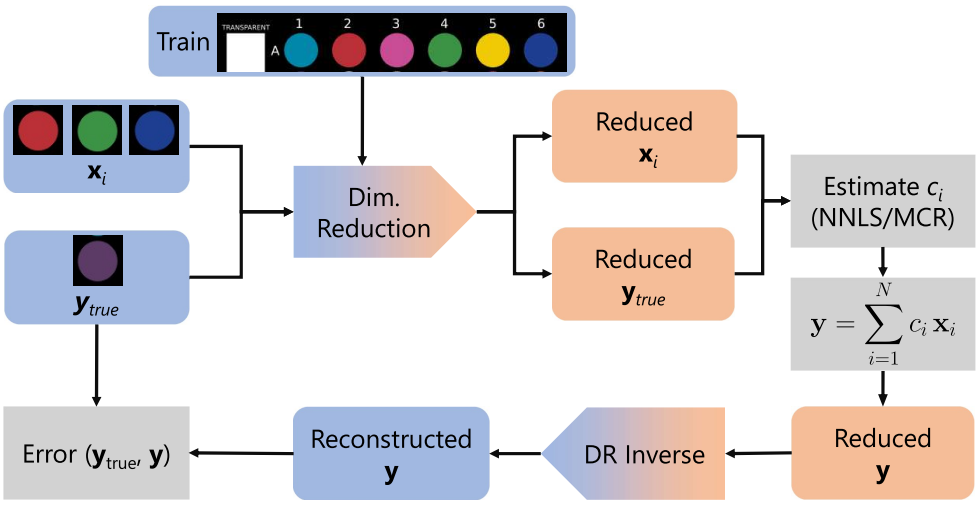


Figure 3.13: Colour calibration slide data analysis pipeline.

The step-by-step analysis procedure of the colour calibration slide HS data is as follows:

1. Data loading – HS and NIST reference spectra
2. NIST reference spectra interpolation to match the spectral resolution of HS data
3. Data calibration and normalization
4. Training of DR algorithms on NIST reference
5. DR applied to NIST and HS data – reduced datasets
6. Spectral unmixing performed on reduced datasets
7. Reconstruction of the unmixed spectra to the original dimensionality
8. Evaluation of the reconstruction error compared to the reference spectra

The analysis begins with data preparation, where reference spectral data is loaded and organised into training and testing sets. The pipeline specifically focusses on the analysis of three endmember classes $\mathbf{x_i}$ (A2, A4 and A6) that represent red, green and blue patches in the dataset. The implementation supports multiple-dimensionality reduction techniques, including PCA, NMF, and two variants of autoencoder: a fully connected architecture (AE-FC) and a convolutional neural network-based architecture (AE-CNN).

The dimensionality reduction process is implemented through a flexible pipeline that can accommodate both traditional and deep learning-based approaches. For traditional methods like PCA and NMF, the code uses scikit-learn's implementations [117], while for AE, custom models are loaded from pre-trained weights. AE models are specifically trained using the SAD loss function, described in Section 2.4.4. The network design of the AE is based on a parametric architecture, resizing the number of neurons in the hidden layers according to the dimension of the encoded data; see Listing 3.

Following dimensionality reduction, the processing continues by applying two distinct spectral unmixing algorithms: Non-negative Least Squares (NNLS) and Multivariate Curve Resolution (MCR). The NNLS implementation adapted from [119] includes robust error handling with and is used with a maximum iteration limit of 100,000 to ensure convergence. The MCR implementation, developed by NIST [120] is configured with specific parameters, including a tolerance increase factor of 10 and a maximum of 100 iterations, providing a balance between computational efficiency and solution accuracy.

The evaluation of the results of the unmixing is performed using the SAD metric. Lower SAD values indicate better preservation of spectral information. The results are systematically stored in a database, with each experiment entry containing detailed metadata including timestamp, script information, dimensionality reduction method, unmixing algorithm, number of components, and various performance metrics.

The pipeline is designed to support systematic experimentation through a grid search approach, allowing exploration of different combinations of dimensionality reduction techniques, unmixing algorithms, and dimensionality levels. This systematic approach enables a thorough investigation of the parameter space and the comparison of different methodological combinations, allowing the investigation of the trade-off between data compression and information preservation.

The pipeline also includes robust error handling and comprehensive logging of the analysis process. The code is structured in a modular fashion, with a clear separation between data loading, dimensionality reduction, unmixing, and evaluation components, making it maintainable and extensible for future research. The code for this analysis pipeline is available at [121].

```
encoder = Sequential([
        Flatten(input_shape=(input_shape,)),
        Dense(16 * encoded_shape, activation="relu"),
        Dense(8 * encoded_shape, activation="relu"),
        Dense(4 * encoded_shape, activation="relu"),
        Dense(2 * encoded_shape, activation="relu"),
        Dense(encoded_shape, activation="sigmoid"),
        Reshape((encoded_shape,))],
    name="encoder")
decoder = Sequential([
        Flatten(input_shape=(encoded_shape,)),
        Dense(2 * encoded_shape, activation="relu"),
        Dense(4 * encoded_shape, activation="relu"),
        Dense(8 * encoded_shape, activation="relu"),
        Dense(16 * encoded_shape, activation="relu"),
        Dense(input_shape, activation="sigmoid"),
        Reshape((input_shape,))],
    name="decoder")
```

Listing 3: Parametric AE implementation in Python programming language using Keras library [118]. Note the `encoded_shape` variable, used for dynamical resizing of the network architecture based on the target compressed dimension size.

Results

First, an exploration of the parameter space was performed by varying the number of components used for reconstruction of the input spectrum when using reference NIST spectra as input dataset $\mathbf{D}$, and RGB patches (A2, A4, A6) spectra as endmembers $\mathbf{S^T}$ for training the DR and unmixing algorithms. The difference between the input spectra and the reconstruction was quantified using two error functions: SAD and RMSE; see Figure 3.14.

The plot in Figure 3.14 shows that a relatively low number of dimensions is enough to reconstruct the spectrum with low error. The combination of PCA-NNLS reduces its SAD error from two to three components, but it does not show significant improvement after that. The PCA-MCR is interestingly showing worse results with three components using the SAD error, but better when evaluated using RMSE error, having similar behaviour as previous combination with SAD error. The NMF-NNLS performs best at two components and after that its performance varies when examining both of the error functions. Lastly, the NMF-MCR combination shows dramatically different performance in relation to the number of components, clearly outperforming the rest when using three components. For further processing, the number of dimensions used will be limited in the range of two to six components, due to the high variance of the results in this region and a significant compression ratio, when considering that the input dimension of the spectra is 728, which yields a 99.18 % to 99.73 % increase in storage efficiency.

Figure 3.15 presents the results that demonstrate the effectiveness of the proposed spectral unmixing pipeline in different dimensionality reduction and unmixing combinations. The analysis was performed on transmission spectral data from NIST colour patches and measured HS data separately.

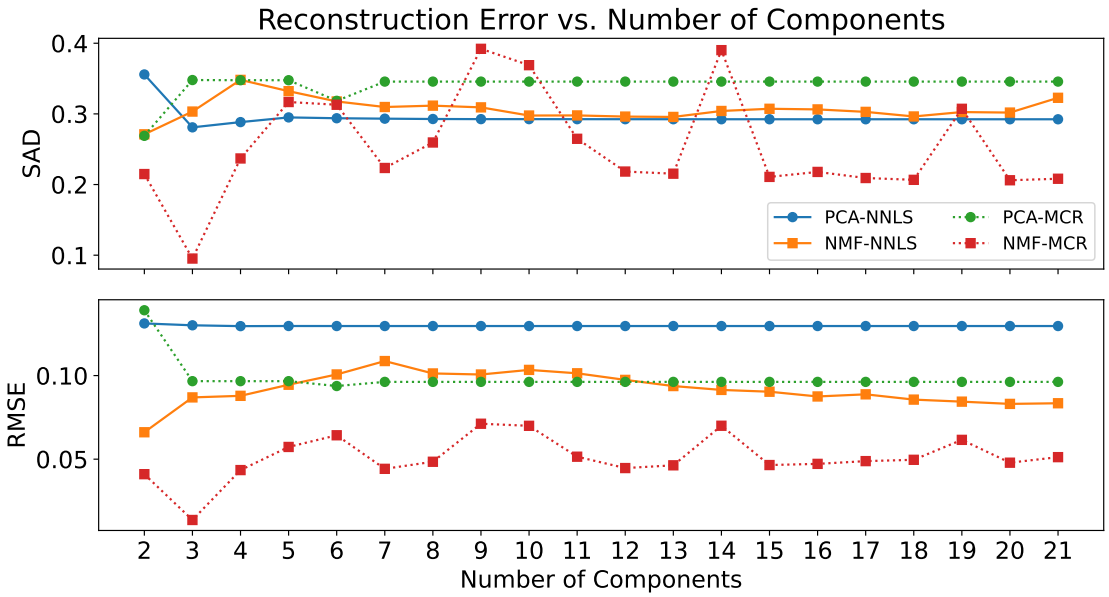
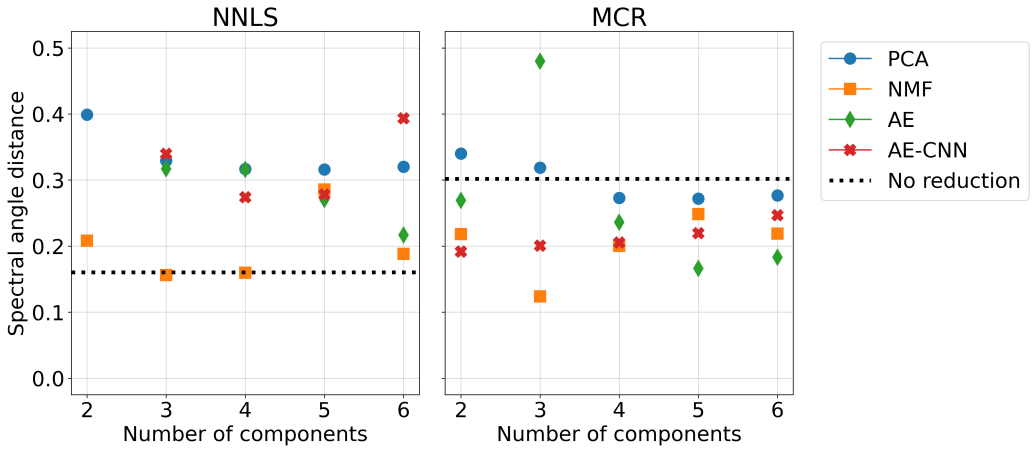


Figure 3.14: Plot of the spectrum reconstruction error versus the dimension of the reduced space (number of components). The input spectra are the reference NIST data, with RGB patches (A2, A4, A6) spectra used for training.

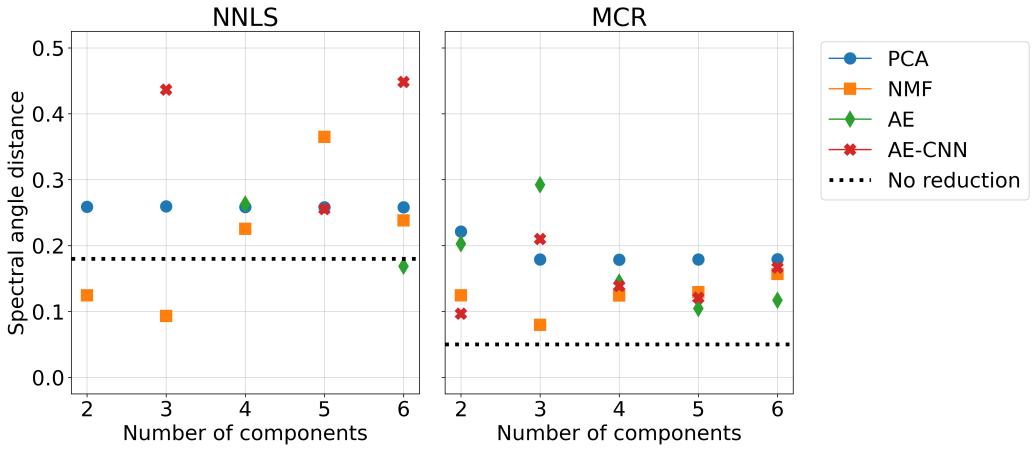
As we can observe in Figure 3.15, the SAD metrics revealed varying performance across different combinations of dimensionality reduction and unmixing methods. The results demonstrate that the choice of the dimensionality and unmixing method significantly impacts the accuracy of the spectral unmixing and the SAD reconstruction error.

Upon analysing the effect of individual DR algorithms, certain patterns become apparent. The first thing to note is that PCA performs consistently in combination with both NNLS and MCR, with the lowest SAD values for varying component counts. The more principal components are used in the unmixing, the more noise of the input data is represented, making the higher-dimensionality PCA output less useful in this case. NMF yields the best results for three components across both datasets and unmixing methods, matching or even surpassing the unmixing performance on the entire spectral data (dotted line). This may be due to the additive nature of the three photographic film layers applied to generate coloured patches, which may correlate with the decomposition process of the NMF transform. AE in most cases matches or even outperforms PCA, and generally gives better results for higher number of components, where it excels compared to all other DR approaches. The modular nature of the AE architecture may be the cause of this, since networks with higher numbers of neurons can learn to represent the input spectral information better than simpler architectures. Lastly, AE-CNN shows similar results compared to AE and NMF while using MCR for unmixing and setting the compressed dimension to four or more components. However, AE-CNN outperforms all other algorithms when using a two-component encoding dimension, indicating a capability of complex data representation with high compression ratio.

The resulting concentration identified by the unmixing pipeline can be visualised as an RGB image because it is composed of three values and also because the task is directly related to the identification of the basic colour components. Figure 3.16 shows the cropped areas of the HS image after unmixing using NMF with 3 components and MCR. The values



(a) NIST reference transmission spectra



(b) HS measured spectra

Figure 3.15: The effect of DR algorithm and components number on SAD error of original and reconstructed spectra. The DR was performed using algorithm listed in the legend, the unmixing was performed separately using NNLS and MCR, each in its own subplot. For some combinations of DR and the number of components the NNLS did not converge, hence the missing data points.

overlaying the patches are errors (differences) in the hue channel E_H, calculated by transforming the identified concentration values used as RGB colours to the hue saturation value (HSV) space [122], converted from the original RGB values, calculated using Equation 3.2:

$$E_H = 0.5 - \Big| |H_{\text{true}}(R, G, B) - H_{\text{predicted}}(c_1, c_2, c_3)| - 0.5 \Big|, \tag{3.2}$$

where $E_H \in \langle 0, 1 \rangle$ and H denotes the hue channel from HSV colour space. This error metric is based on an analogous concept as the SAD, giving no weight to the lightness or saturation of the colour, similarly as the SAD is less sensitive to offset in the two compared spectra.

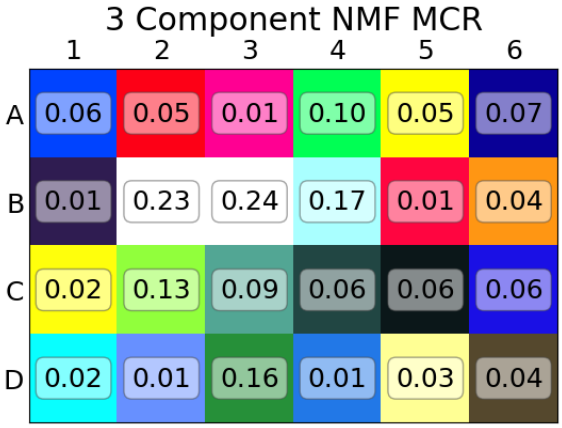


Figure 3.16: Visualization of the identified concentration c_1, c_2, c_3, identified by compressing the HS data to 3 components using NMF and unmixing using MCR, as RGB values, with the corresponding average hue error E_H for each colour patch. The axes labels represent the row and column index of the colour calibration test patches.

Colour Calibration Slide Analysis Highlights

- Reference transmission spectra can be used to train supervised classification algorithms that are applicable to the measured HS data.
- An optimal combination of DR and the unmixing algorithm can be found while keeping the SAD error low compared to no reduction in dimensionality.
- The lowest reconstruction error was achieved using the combination of NMF with 3 components and MCR for spectral unmixing.
- The best result for the highest reduction of dimensionality can be achieved using a convolutional autoencoder.
- DR combined with spectral unmixing can be used to estimate the composition of the colour patches by mixing the primary colours.

3.2.3 Marine Cyanobacteria Mixtures

Motivation

Marine cyanobacteria represent one of the main and globally abundant components of the marine ecosystem that play distinct roles in primary production and biogeochemical cycles [123]. Identification of marine cyanobacteria, shaped by varying environmental conditions such as light and nutrients, at the microscopic level is crucial to understand their ecology, physiology and subsequently to the development of remote sensing techniques [124]. Flow cytometry is one of the primary methods used to identify and differentiate cyanobacteria based on their single cell properties, such as light scattering (related to size and refractive index) and fluorescence (based on the content of cellular pigment) [125]. This unique pigment compositions can be used by HS remote sensing to monitor their distributions and infer ecosystem status and change. HS microscope imaging can be used to obtain the spectral signatures of pure cultures and defined mixtures of them needed to train and validate remote sensing algorithms to monitor their distributions globally. The following section introduces the selected samples, their preparation, and measurement with the goal of identifying their presence in mixtures and quantifying their ratios using spectral unmixing algorithms.

Samples

We were able to work with three strains of marine cyanobacteria thanks to the cooperation with Centre Algatech, Institute of Microbiology of the Czech Academy of Sciences, as part of the *PHOTOMACHINES – Photosynthetic cell redesign for high yields of therapeutic peptides* project (OP JAK, EH22_008/0004624). These strains were chosen because of their different pigment compositions.

- *Prochlorococcus marinus* strain MIT9301 (green, labelled as G)
- *Prochlorococcus marinus* strain MIT9313 (yellow, labelled as Y)
- *Synechococcus sp.* strain WH8020 (red, labelled as R)

The primary spectral differences between the two different genus, the *Prochlorococcus* and *Synechococcus* are driven by their distinct major light harvesting pigment systems: divinyl chlorophylls and an atypical antenna system in strain MIT9313 versus phycobilisomes containing phycobiliproteins and monovinyl chlorophyll in *Synechococcus*. These differences result in spectral variations, particularly in the green-yellow (490–560 nm) and red (670–680 nm) regions of their absorption and reflectance spectra [126, 127, 128]. As a result, the dominant colours of the cyanobacteria strains are red for WH8020, green for MIT9301, and yellow for MIT9313, as shown in Figure 3.17.

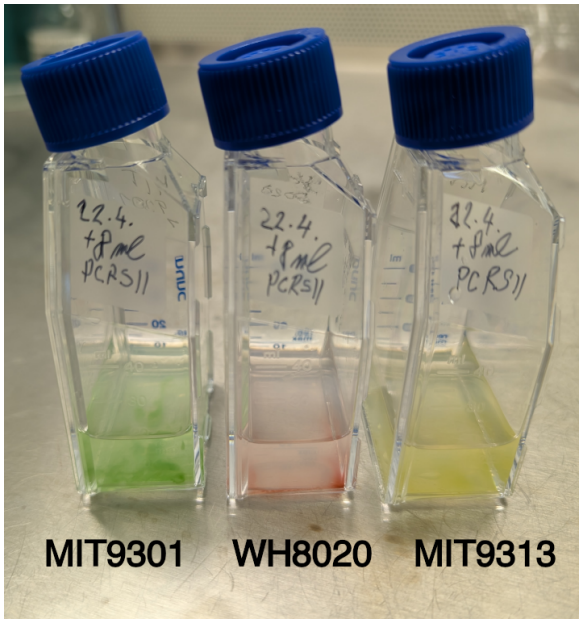


Figure 3.17: Selected marine cyanobacteria strains in cell culture flasks. Samples courtesy of Centre Algatech, Institute of Microbiology of the Czech Academy of Sciences.

To illustrate the microscopic compositional differences between the selected strains, a series of bright-field scanning transmission electron microscopy (STEM) images were acquired; see Figures 3.18, 3.19, and 3.20.

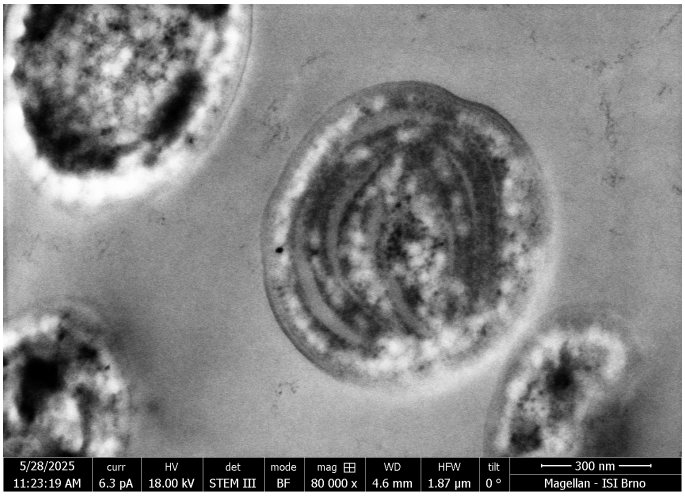


Figure 3.18: STEM image of *Prochlorococcus marinus* strain MIT9301, with clearly visible folds of thylakoid membranes containing chloroplasts, responsible for the green colour.

The samples were first centrifuged, frozen under high pressure, followed by freeze-substitution, embedded in Spurr media, sliced using an ultramicrotome (Leica Microsystems EM UC7), and images were obtained using a low voltage STEM (FEI Magellan 400/L). More details on the processing for STEM imaging can be found in [129].

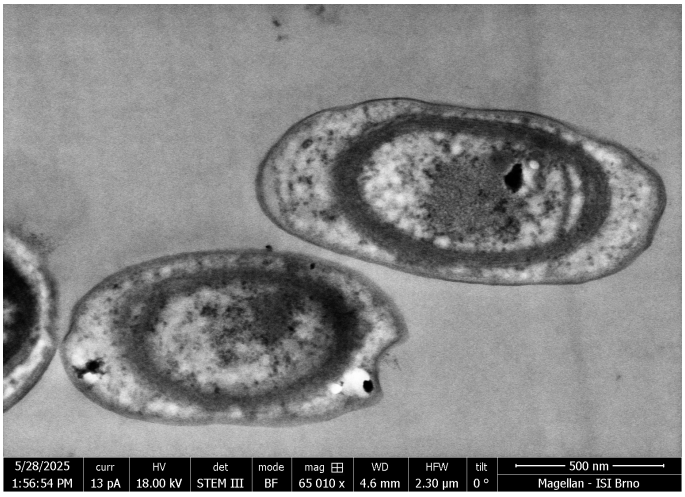


Figure 3.19: STEM image of *Prochlorococcus marinus* strain MIT9313. These cells possess three to five distinct layers of intracytoplasmic membranes, which are arranged concentrically near the cell periphery. This tight packing of the internal membranes is possible because the lack of phycobilisomes, which is also responsible for its yellow colour [130].

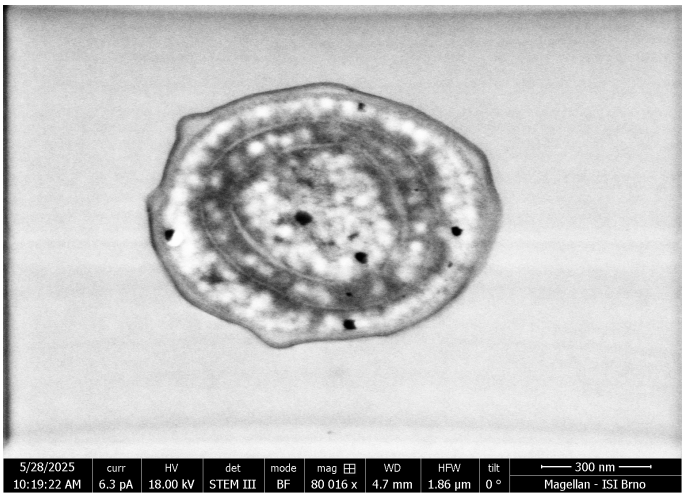


Figure 3.20: STEM image of *Synechococcus sp.* WH8020, with clearly visible rings of thylakoid membrane, which in this case contains phycoerythrin responsible for the red colour. The black dots are carboxysomes bacterial microcompartments used for carbon fixation [131].

Design of Experiment

The three different strains of cyanobacteria were sampled using a micropipette on a 96-well microplate (flat bottom, Thermo Scientific) using a scheme described in Table 3.5. The prepared microplate included pure samples of cyanobacteria strains, together with a sample of pure cultivation medium (PCR-S11 [132]) in which the cyanobacteria are dispersed. Each well was loaded with 150 µL sample liquid. The mixing ratios selected for this experiment were 1:1, 1:2 and 1:4 for the mixture of two samples, including all combinations of the ratios for all three samples. All strains were also mixed in a 1:1:1 ratio. The last mixture ratio was all cyanobacteria in a 1:2:3 ratio, for all these possible combinations; see Table 3.5.

Table 3.5: Mixing ratios of marine cyanobacteria samples with their relative positions inside the 96-well microplate. The G stands for MIT9301, R for WH8020 and Y for MIT9313. Some wells were left empty and measured as white frame reference.

Category	**Samples**					
Pure	G	R	Y	PCR-S11		
1:1 Ratio	1G:1R	1G:1Y	1R:1Y	1R:1G:1Y		
1:2 Ratio	1G:2R	1G:2Y	1R:2G	1R:2Y	1Y:2R	1Y:2G
1:4 Ratio	1G:4R	1G:4Y	1R:4G	1R:4Y	1Y:4R	1Y:4G
1:2:3 Ratio	1G:2R:3Y	1G:3R:2Y	2G:1R:3Y	2G:3R:1Y	3G:1R:2Y	3G:2R:1Y

After the preparation of the samples, the transmission spectra of the populated positions of the microplate wells were measured using the HS microscope, as well as a commercial spectrometer (Ocean Optics, Maya 2000), capturing spectral information from the centre of each well (total of 270 spectra); see Figure 3.21. The acquisition of the spectral data was performed using the HS microscope imaging optics with the only change being the replacement of the HS camera with the spectrometer collection optics.

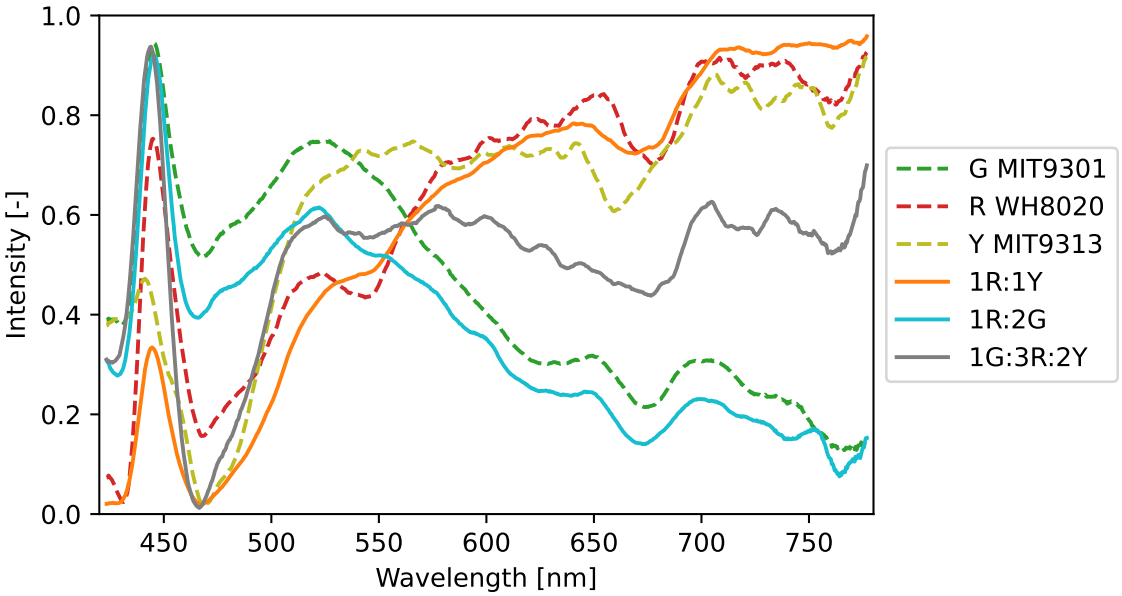


Figure 3.21: Selected transmission spectra of marine cyanobacteria and its mixtures measured using a commercial spectrometer. The measured spectra were acquired with 60 ms exposure time, calibrated using Equation 2.8, normalized to unity range, and smoothed using rolling average filter with 15 pixel window size.

HS scanning data were acquired using the HS microscope platform described in Section 3.1, in a configuration with a Nikon Plan Fluor 4× microscope objective and using a Sony IMX492 CMOS sensor. For each of the wells containing pure/mixed samples (250 frames per well, approximately 2 mm × 2 mm area around the centre of the well, total of 1 331 000 spectra). The average of the HS measured spectra for each sample is presented in the Supplementary Figure S2. The corresponding metadata of the HS measurement are listed in Table 3.6.

Table 3.6: Metadata of HS measurement of marine cyanobacteria.

Parameter	**Value**	**Unit**
Exposure	300	ms
Gain	10.8	dB
Binning	4	×
λ_{min}	477.67	nm
λ_{max}	619.85	nm
$\Delta\lambda$	0.07	nm
Camera Temperature	0.0	°C
Ambient Temperature	22	°C
Pressure	993	hPa
Humidity	40	%
Date	2025-04-08	YYYY-MM-DD
Time	10:40:47+02:00	HH:MM:SS+offset

Data Processing

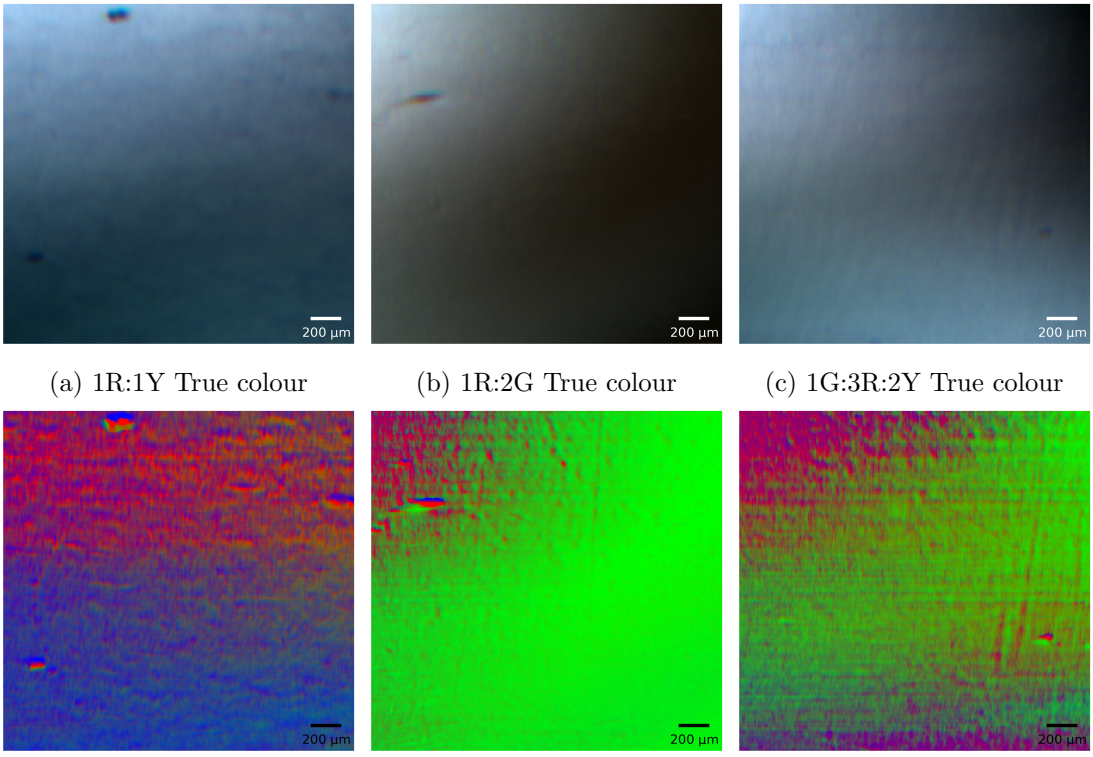
The measured HS data were calibrated using dark and white frames, as described in Section 2.3.3 and binned using a 5 pixel window in the direction perpendicular to the scanning direction to allow for a more compact visualisation of the HS images. HS images of individual wells filled with samples and then unmixed using an MCR algorithm with sum-to-one and non-negativity constraints on the estimated concentration fractions $\mathbf{C}$, as described in Section 2.4.2. The pure cyanobacteria spectra from the HS measurement were averaged and then used as input pure spectra $\mathbf{S^T}$ for the spectral unmixing algorithm in all visualisations used in the main text. To simulate the situation where the known spectra come from a spectral database that can be measured using a spectrometer with different parameters, a comparison unmixing was done using the pure spectra from the commercial spectrometer. However, these spectral data had to be cropped to the HS measurement range and upsampled (using quadratic interpolation) to match the spectral resolution of the HS data. The results of this comparison are presented in Table 3.7. With the results of the colour slide analysis described in Section 3.2.2, the marine cyanobacteria HS data were also reduced in spectral dimension using the NMF algorithm, to speed up the computation time and compare the accuracy of the result of the analysis performed on the complete HS data.

The analysis of HS marine cyanobacteria data has two goals: to detect the presence of individual cyanobacteria strains in the measured mixture and to quantify their mixing ratios. Both of these tasks are performed using the result of the spectral unmixing, the concentration matrix $\mathbf{C}$. The presence of cyanobacteria is determined using thresholding of concentration values, in this case, when the estimated concentration is higher than 5%, it is considered to be present in the mixture. From the design of the experiment, the lowest possible concentration is $16.\bar{6}\%$ (for the 1:2:3 ratio mixtures), so the threshold was set approximately as one-third of this value, leaving room for uncertainties originating from the sample preparation and intra-sample variation. The results of this identification were then compared with the true content of the measured mixtures, which enabled the calculation of the identification accuracy (the number of all identification results that were correct over the total number of samples). Identification was performed separately on both the HS data and the static spectra from the commercial spectrometer. To evaluate the

influence of DR on identification accuracy, the two data sets were also reduced using NMF to 3 components prior to unmixing. The complete results of the unmixing are presented in the Supplementary Table S5.

Results

The mixing ratio of marine cyanobacteria is determined directly from the matrix **C** and is visualised using a false colour image, where the red channel is assigned to the calculated concentration of WH8020, the green channel to MIT9301 and the blue channel to MIT9313; see Figure 3.22. These results are also visualised as separate concentration mappings for each of the marine cyanobacteria; see Figure 3.23.



(a) 1R:1Y True colour (b) 1R:2G True colour (c) 1G:3R:2Y True colour

(d) 1R:1Y False colour (e) 1R:2G False colour (f) 1G:3R:2Y False colour

Figure 3.22: Selected HS measurement of marine cyanobacteria in two different visualizations: a–c) shows true colour RGB composed from 460 nm, 550 nm and 640 nm bands from calibrated HS data, and d–e) showing false colour mapping of the concentration of the identified marine cyanobacteria, with the following colour channel assignment: R=WH8020, G=MIT9301, B=MIT9313. The concentrations were identified after NMF reduction of spectral domain to 3 components.

The samples are not easily detectable using the RGB preview images, as presented in Figure 3.22, while using the identified concentration, the presence and spatial distribution of the cyanobacteria becomes visible. To further enhance the readability of the identified concentration, it is better to plot them separately for each component, as presented in Figure 3.23.

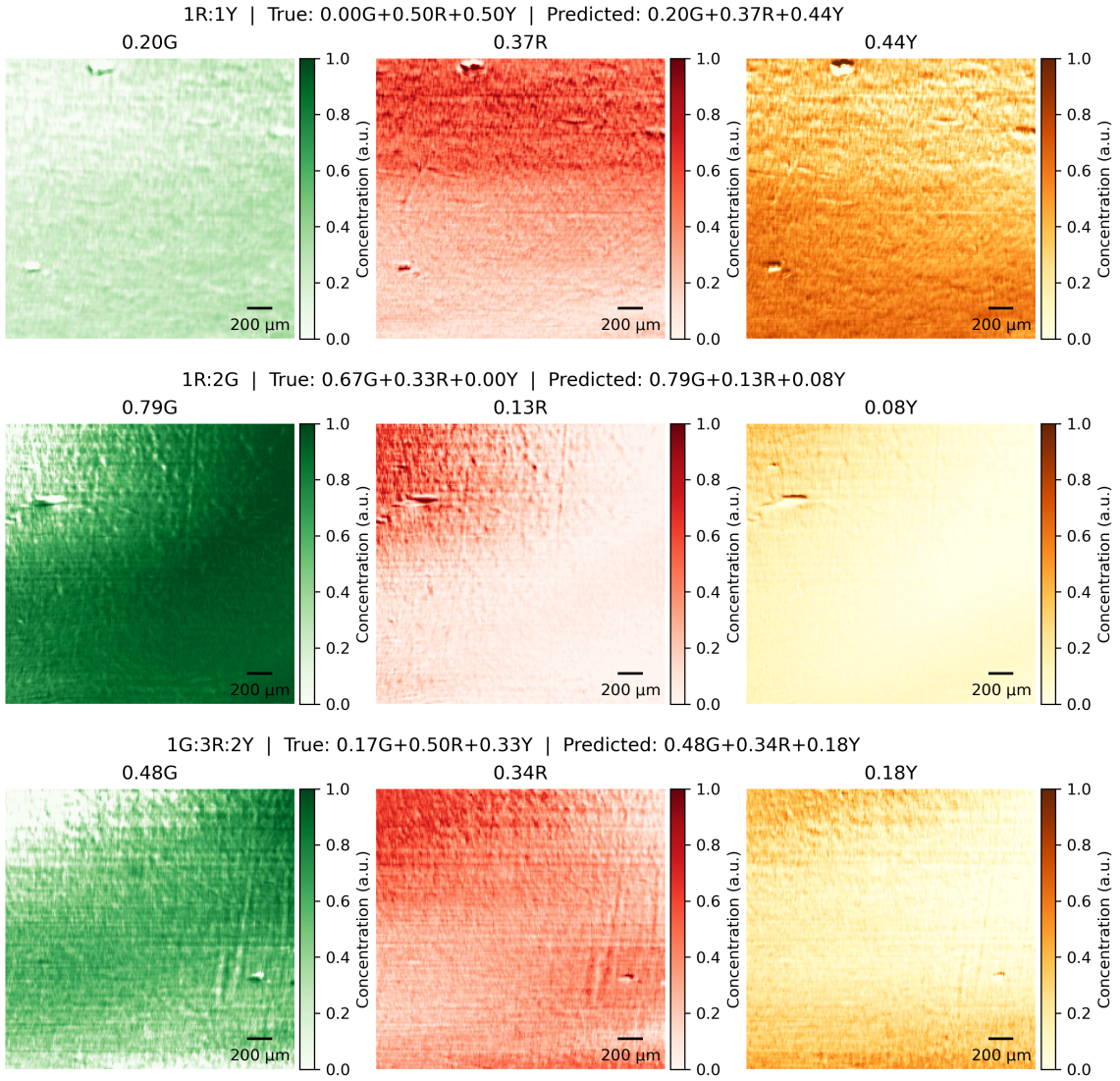


Figure 3.23: Predicted marine cyanobacteria concentrations displayed in individual channels for each of the components. The G stands for MIT9301, R for WH8020 and Y for MIT9313. The concentrations were identified after NMF reduction of spectral domain to 3 components.

After the concentration estimation, the values of $\mathbf{C}$ were thresholded and compared to the contents of the individual mixtures. The performance metrics discussed in Section 2.4.5 were calculated based on these predictions, see Table 3.7.

Table 3.7: Comparison of the classification performance on the data from commercial spectrometer vs HS measurement and the influence of the DR in form of reduction using three component NMF. The performance metrics are calculated using equations presented in Section 2.4.5.

Dataset	DR	Runtime	Accuracy	Precision	Recall
Spectrometer	None	34 ms	0.68 ± 0.12	0.8 ± 0.1	0.8 ± 0.1
	NMF 3	74 ms	0.59 ± 0.12	0.74 ± 0.11	0.7 ± 0.1
HS	None	1692 s	0.76 ± 0.12	0.9 ± 0.1	0.8 ± 0.1
	NMF 3	57.8 s	0.76 ± 0.03	0.81 ± 0.06	0.90 ± 0.14

A key takeaway from Table 3.7 is that the use of NMF with 3 components significantly reduces the computation time (96.42% speed-up) of the unmixing in the entire HS data with little negative impact on classification performance, the accuracy remains the same, there is a decrease in precision but an increase in recall (making the model overconfident in the precisions), and for all metrics there is a reduction in their respective uncertainties. However, the same is not the case for the spectrometer data, where the DR process decreases performance. This may be due to the smaller size of the dataset compared to HS measurement (270 spectra vs. 1 331 000 spectra in HS). The negative impact of DR on runtime in the case of spectrometer data is negligible compared to the time savings of HS processing, so this is not considered a disadvantage of DR in this case.

Marine Cyanobacteria Mixtures Analysis Highlights

- Cyanobacteria play a crucial role in marine ecosystems. Their distinct pigment compositions can be used for identification by HS imaging.
- Three specific strains of marine cyanobacteria are chosen for their differing pigment compositions and resulting spectral variations.
- A 96-well microplate is prepared with pure samples and various defined mixtures of the three cyanobacteria strains, including 1:1, 1:2, 1:4, 1:1:1, and 1:2:3 ratios.
- Transmission spectra are acquired using both a commercial spectrometer and an HS microscope. The data were unmixed using an MCR algorithm with pure cyanobacteria spectra from HS images used as a reference.
- When NMF with 3 components is applied to the HS data, the computation time is significantly reduced with minimal negative impact on classification accuracy, precision, and recall.

4 Conclusion

This doctoral thesis represents a contribution to the field of hyperspectral imaging by addressing both instrumental development and advanced data analysis methodologies, with a particular focus on microscopic applications. The overarching aim of this work was to reduce the sophistication of hyperspectral data analysis and enhance the potential of this non-destructive imaging technique for routine analysis, particularly for plant and algal growth. This goal was pursued through the development of a comprehensive tool for measuring, processing, and evaluating hyperspectral data, encompassing both hardware and software, and exploring advanced analytical methods.

Key contributions of this thesis include the development of a modular hyperspectral microscope workstation at ISI CAS. This system is equipped with a hyperspectral camera, which offers exceptional spectral resolution of 0.1 nm and a tunable wavelength range of 350–1100 nm. The workstation supports automated image acquisition, which is crucial for building robust hyperspectral databases. To ensure repeatability and consistency of the HS data acquisition, custom software was developed using Python programming language, providing full control over the measurement process and enabling automated data calibration.

A standardised and efficient storage format for hyperspectral data and its metadata was established, using the Hierarchical Data Format (HDF) for optimised data storage, compression capabilities, and hierarchical structure, suitable for easy data sharing and self-description. Metadata is stored in JSON format, known for its readability and interoperability, and can be easily transformed to the ENVI header file standard, facilitating import into various HS analysis software.

The thesis also demonstrated verification of the capabilities of the hyperspectral system. The spectral resolution of the HS camera was verified using a low-pressure sodium-vapour lamp, confirming its ability to distinguish close spectral lines. The imaging and spectral capabilities of the HS microscope were characterised by capturing data from a NIST traceable colour transmission calibration slide. This dataset, coupled with NIST reference spectra, proved invaluable for validating spectral unmixing algorithms and evaluating the impact of dimensionality reduction. It demonstrated successful transfer learning/generalisation of algorithms designed based on reference spectral data to data acquired from a different instrument.

The thesis also explored the collaboration of machine learning methods and chemometric algorithms for HS data analysis. Firstly, I successfully applied the HS unmixing algorithms on Raman spectroscopy data, allowing the identification of pathogens in blood serum with an accuracy greater than 90% without the need for operator-dependent preprocessing steps. The thesis further confirmed the effectiveness of DR techniques such as non-negative matrix factorisation and autoencoder neural networks in reducing the high-dimensionality of HS data while preserving essential information. It was demonstrated that a relatively low number of dimensions is sufficient for accurate spectrum reconstruction. The designed convolutional autoencoder showed superior performance in achieving high compression ratios with low reconstruction error.

The HS measurement platform and my HS analysis algorithms were also used for proof-of-concept analysis and identification of marine cyanobacteria strains. Using pure cyanobacteria spectra as endmembers for the MCR algorithm, the unmixing process was able to identify the different types of marine cyanobacteria within mixtures with good results for balanced and complex mixing ratios, often yielding concentration ratios very close to the

sample mixing ratios. This application highlights the potential of HS microscopy to understand marine ecosystems and potentially validate remote sensing algorithms.

This doctoral thesis has successfully demonstrated the development of a sophisticated hyperspectral microscope workstation and associated software for automated data acquisition and calibration. It has also thoroughly investigated advanced processing techniques, particularly dimensionality reduction and spectral unmixing, proving their effectiveness for biological analysis, and successfully meeting the thesis objectives. The successful transfer of knowledge from reference data to actual HS measurements, and the accurate identification and quantification of marine cyanobacteria, underscore the potential of hyperspectral imaging to transition from a specialised research tool to a practical, routine method in various scientific and industrial settings.

List of Abbreviations

AE	Autoencoder
AE-CNN	Convolutional Autoencoder
AE-FC	Fully Connected Autoencoder
ANN	Artificial Neural Network
CAD	Computer Aided Design
CAS	Czech Academy of Sciences
CMOS	Complementary Metal–oxide–semiconductor
CNN	Convolution Neural Network
DR	Dimensionality Reduction
FAIR	Findable Accessible Interoperable Reusable
FCLS	Fully Constrained Least Squares
FN	False Negative
FP	False Positive
FTIR	Fourier Transform Infrared
FWHM	Full Width at Half Maximum
HDF	Hierarchical Data Format
HS	Hyperspectral
HSI	Hyperspectral Image
HSV	Hue Saturation Value
IEEE	Institute of Electrical and Electronics Engineers
IPF	Iterative Polynomial Fitting
ISI	Institute of Scientific Instruments
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbors
MCR	Multivariate Curve Resolution
MS	Multispectral
NDVI	Normalized Difference Vegetation Index
NI	National Instruments

NIR	Near-infrared
NIST	National Institute of Standards and Technology
NMF	Non-negative Matrix Factorization
NNLS	Non-negative Least Squares
PCA	Principal Component Analysis
PLSR	Partial Least Squares Regression
PSI	Photon Systems Instruments
RCF	Rolling Circle Filtering
RGB	Red Green Blue
RMSE	Root Mean Squared Error
ROI	Region(s) of Interest
SAD	Spectral Angle Distance
SNR	Signal-to-Noise Ratio
STEM	Scanning Transmission Electron Microscopy
TN	True Negative
TP	True positive
UCLS	Unconstrained Least Squares
VNIR	Visible and Near-infrared
WOS	Web of Science

References

[1] MORAIS, C. L.; BUTLER, H. J.; MCAINSH, M. R.; MARTIN, F. L. Plant Hyperspectral Imaging. In: *eLS* [online]. 1st ed. Wiley, 2019, pp. 1–12 [visited on 2023-08-18]. ISBN 978-0-470-01617-6 978-0-470-01590-2. Available from: doi:10.1002/9780470015902.a0028367.

[2] RAST, M.; PAINTER, T. H. Earth Observation Imaging Spectroscopy for Terrestrial Systems: An Overview of Its History, Techniques, and Applications of Its Missions. *Surveys in Geophysics* [online]. 2019, vol. 40, no. 3, pp. 303–331 [visited on 2023-07-06]. ISSN 1573-0956. Available from: doi:10.1007/s10712-019-09517-z.

[3] PARK, B.; LU, R. (eds.). *Hyperspectral imaging technology in food and agriculture*. New York: Springer, 2015. Food engineering series. ISBN 978-1-4939-2835-4.

[4] LU, G.; FEI, B. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*. 2014, vol. 19, no. 1, p. 10901. ISSN 1560-2281. Available from: doi:10.1117/1.JBO.19.1.010901.

[5] GOWEN, A. A.; FENG, Y.; GASTON, E.; VALDRAMIDIS, V. Recent applications of hyperspectral imaging in microbiology. *Talanta* [online]. 2015, vol. 137, pp. 43–54 [visited on 2023-08-01]. ISSN 0039-9140. Available from: doi:10.1016/j.talanta.2015.01.012.

[6] KÜSTER, J.; GROSS, W.; HEIZMANN, M.; MIDDELMANN, W. Impact of different compression rates for hyperspectral data compression based on a convolutional autoencoder [online]. 2021 [visited on 2023-07-23]. Available from: doi:10.1117/12.2598010.

[7] SHENMING, Q.; XIANG, L.; ZHIHUA, G. A new hyperspectral image classification method based on spatial-spectral features. *Scientific Reports* [online]. 2022, vol. 12, no. 1, p. 1541 [visited on 2022-06-24]. ISSN 2045-2322. Available from: doi:10.1038/s41598-022-05422-5.

[8] MANIFOLD, B.; MEN, S.; HU, R.; FU, D. A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nature Machine Intelligence* [online]. 2021, vol. 3, no. 4, pp. 306–315 [visited on 2023-07-23]. ISSN 2522-5839. Available from: doi:10.1038/s42256-021-00309-y.

[9] GOETZ, A. F. H. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment* [online]. 2009, vol. 113, S5–S16 [visited on 2024-06-26]. ISSN 0034-4257. Available from: doi:10.1016/j.rse.2007.12.014.

[10] LOWE, A.; HARRISON, N.; FRENCH, A. P. Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods* [online]. 2017, vol. 13, no. 1, p. 80 [visited on 2023-08-18]. ISSN 1746-4811. Available from: doi:10.1186/s13007-017-0233-z.

[11] BEDINI, E. The use of hyperspectral remote sensing for mineral exploration: a review. *Journal of Hyperspectral Remote Sensing* [online]. 2017, vol. 7, no. 4, pp. 189–211 [visited on 2025-05-13]. ISSN 2237-2202, ISSN 2237-2202. Available from: doi:10.29150/jhrs.v7.4.p189-211.

[12] QIAN, S.-E. Hyperspectral Satellites, Evolution, and Development History. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* [online]. 2021, vol. 14, pp. 7032–7056 [visited on 2024-07-01]. ISSN 2151-1535. Available from: doi:10.1109/JSTARS.2021.3090256.

[13] AMIGO, J. M.; BABAMORADI, H.; ELCOROARISTIZABAL, S. Hyperspectral image analysis. A tutorial. *Analytica Chimica Acta* [online]. 2015, vol. 896, pp. 34–51 [visited on 2022-05-31]. ISSN 00032670. Available from: doi:10.1016/j.aca.2015.09.030.

[14] GOETZ, A.; VANE, G.; SOLOMON, J.; ROCK, B. Imaging spectrometry for earth remote sensing. *Science*. 1985, vol. 228, no. 4704, pp. 1147–1153. Available from: doi:10.1126/science.228.4704.1147.

[15] *MODIS Web* [online]. [visited on 2024-06-26]. Available from: https://modis.gsfc.nasa.gov/about/design.php.

[16] STUART, M. B.; MCGONIGLE, A. J. S.; WILLMOTT, J. R. Hyperspectral Imaging in Environmental Monitoring: A Review of Recent Developments and Technological Advances in Compact Field Deployable Systems. *Sensors* [online]. 2019, vol. 19, no. 14, p. 3071 [visited on 2025-05-13]. ISSN 1424-8220. Available from: doi:10.3390/s19143071.

[17] TATZER, P.; WOLF, M.; PANNER, T. Industrial application for inline material sorting using hyperspectral imaging in the NIR range. *Real-Time Imaging* [online]. 2005, vol. 11, no. 2, pp. 99–107 [visited on 2025-05-13]. ISSN 1077-2014. Available from: doi:10.1016/j.rti.2005.04.003.

[18] SABIN, G. P.; SOARES, F. L. F.; DE FREITAS, D. L. D.; SILVA, H. V. d. O.; ANTUNES, C. d. M. M. O.; MOHAMED, E. A.; TEIXEIRA, C. A.; ASSIS, C.; CARDOSO, V. G. K.; VOLOCHEN, M. Chapter 5 - Hyperspectral imaging applications. In: FERNANDES, F. A. N.; RODRIGUES, S.; FILHO, E. G. A. (eds.). *Chemometrics* [online]. Elsevier, 2024, pp. 91–123 [visited on 2025-01-12]. ISBN 978-0-443-21493-6. Available from: doi:10.1016/B978-0-443-21493-6.00005-8.

[19] AHMAD, M.; SHABBIR, S.; ROY, S. K.; HONG, D.; WU, X.; YAO, J.; KHAN, A. M.; MAZZARA, M.; DISTEFANO, S.; CHANUSSOT, J. Hyperspectral Image Classification—Traditional to Deep Models: A Survey for Future Prospects. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2022, vol. 15, pp. 968–999. ISSN 2151-1535. Available from: doi:10.1109/JSTARS.2021.3133021.

[20] PAOLETTI, M. E.; HAUT, J. M.; PLAZA, J.; PLAZA, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* [online]. 2019, vol. 158, pp. 279–317 [visited on 2022-08-25]. ISSN 0924-2716. Available from: doi:10.1016/j.isprsjprs.2019.09.006.

[21] OIKNINE, Y.; AUGUST, I.; FARBER, V.; GEDALIN, D.; STERN, A. Compressive Sensing Hyperspectral Imaging by Spectral Multiplexing with Liquid Crystal. *Journal of Imaging* [online]. 2019, vol. 5, no. 1, p. 3 [visited on 2023-08-01]. ISSN 2313-433X. Available from: doi:10.3390/jimaging5010003.

[22] SILVA, C. S.; PIMENTEL, M. F.; HONORATO, R. S.; PASQUINI, C.; PRATS-MONTALBÁN, J. M.; FERRER, A. Near infrared hyperspectral imaging for forensic analysis of document forgery. *The Analyst* [online]. 2014, vol. 139, no. 20, pp. 5176–5184 [visited on 2025-05-13]. ISSN 0003-2654, ISSN 1364-5528. Available from: doi:10.1039/C4AN00961D.

[23] LEAVESLEY, S. J.; ANNAMDEVULA, N.; BONI, J.; STOCKER, S.; GRANT, K.; TROYANOVSKY, B.; RICH, T. C.; ALVAREZ, D. F. Hyperspectral imaging microscopy for identification and quantitative analysis of fluorescently-labeled cells in highly autofluorescent tissue. *Journal of Biophotonics*. 2012, vol. 5, no. 1, pp. 67–84. ISSN 1864-0648. Available from: doi:10.1002/jbio.201100066.

[24] TRAN, M. H.; GOMEZ, O.; FEI, B. An automatic whole-slide hyperspectral imaging microscope. In: *Label-free Biomedical Imaging and Sensing (LBIS) 2023* [online]. SPIE, 2023, vol. 12391, pp. 24–33 [visited on 2023-03-20]. Available from: doi:10.1117/12.2650815. Label-free Biomedical Imaging and Sensing (LBIS) 2023.

[25] AMIGO, J. M.; MARTÍ, I.; GOWEN, A. Chapter 9 - Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality. In: MARINI, F. (ed.). *Data Handling in Science and Technology* [online]. Elsevier, 2013, vol. 28, pp. 343–370 [visited on 2025-05-13]. Chemometrics in Food Chemistry. Available from: doi:10.1016/B978-0-444-59528-7.00009-0.

[26] GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. Fourth edition, global edition. New York: Pearson, 2017. ISBN 978-0-13-335672-4.

[27] DEY, S. *Hands-on Image Processing with Python: Expert Techniques for Advanced Image Analysis and Effective Interpretation of Image Data*. 1st ed. Birmingham, UK: Packt Publishing, 2018. ISBN 978-1-78934-373-1.

[28] WALT, S. van der; SCHÖNBERGER, J. L.; NUNEZ-IGLESIAS, J.; BOULOGNE, F.; WARNER, J. D.; YAGER, N.; GOUILLART, E.; YU, T. scikit-image: image processing in Python. *PeerJ* [online]. 2014, vol. 2, e453 [visited on 2023-07-20]. ISSN 2167-8359. Available from: doi:10.7717/peerj.453.

[29] VAN ROSSUM, G.; DRAKE, F. L. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN 1-4414-1269-7.

[30] FRENTRESS, Z.; YOUNG, L. C.; EDWARDS, H. D. Field Photometer with Nine-Element Filter Wheel. *Applied Optics* [online]. 1964, vol. 3, no. 2, pp. 303–308 [visited on 2025-04-03]. ISSN 2155-3165. Available from: doi:10.1364/AO.3.000303.

[31] ROSENBERGER, M.; CELESTRE, R. Smart multispectral imager for industrial applications. In: *2016 IEEE International Conference on Imaging Systems and Techniques (IST)* [online]. Chania, Greece: IEEE Press, 2016, pp. 7–12 [visited on 2025-04-03]. Available from: doi:10.1109/IST.2016.7738189.

[32] LAPRAY, P.-J.; WANG, X.; THOMAS, J.-B.; GOUTON, P. Multispectral Filter Arrays: Recent Advances and Practical Implementation. *Sensors* [online]. 2014, vol. 14, no. 11, pp. 21626–21659 [visited on 2025-04-03]. ISSN 1424-8220. Available from: doi:10.3390/s141121626.

[33] MANSHA, S.; MOITRA, P.; XU, X.; MASS, T. W. W.; VEETIL, R. M.; LIANG, X.; LI, S.-Q.; PANIAGUA-DOMÍNGUEZ, R.; KUZNETSOV, A. I. High resolution multispectral spatial light modulators based on tunable Fabry-Perot nanocavities. *Light: Science & Applications* [online]. 2022, vol. 11, no. 1, p. 141 [visited on 2025-04-03]. ISSN 2047-7538. Available from: doi:10.1038/s41377-022-00832-6.

[34] *HinaLea Model 4250* [HinaLea Imaging] [online]. 2022 [visited on 2024-11-18]. Available from: https://hinalea.ai/wp-content/uploads/2022/06/Model-4250-Intelligent-Imaging-System-2022.pdf.

[35] *Kurios® Liquid Crystal Tunable Bandpass Filters* [online]. [visited on 2025-04-03]. Available from: https://www.thorlabs.com.

[36] KOKKA, A.; PULLI, T.; HONKAVAARA, E.; MARKELIN, L.; KÄRHÄ, P.; IKONEN, E. Flat-field calibration method for hyperspectral frame cameras. *Metrologia* [online]. 2019, vol. 56, no. 5, p. 055001 [visited on 2022-05-31]. ISSN 0026-1394, ISSN 1681-7575. Available from: doi:10.1088/1681-7575/ab3261.

[37] SU, W.-H.; SUN, D.-W. Fourier Transform Infrared and Raman and Hyperspectral Imaging Techniques for Quality Determinations of Powdery Foods: A Review. *Comprehensive Reviews in Food Science and Food Safety* [online]. 2018, vol. 17, no. 1, pp. 104–122 [visited on 2023-08-01]. ISSN 1541-4337. Available from: doi:10.1111/1541-4337.12314.

[38] FOUCHER, F.; GUIMBRETIÈRE, G.; BOST, N.; WESTALL, F. Petrographical and Mineralogical Applications of Raman Mapping. In: MAAZ, K. (ed.). *Raman Spectroscopy and Applications* [online]. InTech, 2017 [visited on 2025-02-18]. ISBN 978-953-51-2907-3. Available from: doi:10.5772/65112.

[39] AERONAUTICS, N.; NASA, S. A. *Convex diffraction grating imaging spectrometer*. Inventor: M. P. CHRISP. Publ.: 1999-03-09. U.S. patent 5880834A. [Visited on 2023-07-27]. Available from: https://patents.google.com/patent/US5880834A/en.

[40] LERNER, J. M. Imaging spectrometer fundamentals for researchers in the biosciences—A tutorial. *Cytometry Part A* [online]. 2006, vol. 69A, no. 8, pp. 712–734 [visited on 2023-08-02]. ISSN 1552-4930. Available from: doi:10.1002/cyto.a.20242.

[41] GREEN, R. O.; EASTWOOD, M. L.; SARTURE, C. M.; CHRIEN, T. G.; ARONSSON, M.; CHIPPENDALE, B. J.; FAUST, J. A.; PAVRI, B. E.; CHOVIT, C. J.; SOLIS, M.; OLAH, M. R.; WILLIAMS, O. Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment* [online]. 1998, vol. 65, no. 3, pp. 227–248 [visited on 2022-05-31]. ISSN 00344257. Available from: doi:10.1016/S0034-4257(98)00064-9.

[42] QIN, J.; KIM, M.; CHAO, K.; CHAN, D.; DELWICHE, S.; CHO, B.-K. Line-Scan Hyperspectral Imaging Techniques for Food Safety and Quality Applications. *Applied Sciences* [online]. 2017, vol. 7, no. 2, p. 125 [visited on 2024-10-10]. ISSN 2076-3417. Available from: doi:10.3390/app7020125.

[43] KLEIN, L.; TOUŠ, J.; ŽÍDEK, K. Spatially encoded hyperspectral compressive microscope for ultrabroadband VIS/NIR hyperspectral imaging. *Applied Optics* [online]. 2023, vol. 62, no. 15, pp. 4030–4039 [visited on 2025-06-18]. ISSN 2155-3165. Available from: doi:10.1364/AO.484214.

[44] WAGADARIKAR, A.; JOHN, R.; WILLETT, R.; BRADY, D. Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics* [online]. 2008, vol. 47, no. 10, B44–B51 [visited on 2023-08-02]. ISSN 2155-3165. Available from: doi:10.1364/AO.47.000B44.

[45] HUBOLD, M.; MONTAG, E.; BERLICH, R.; BRUNNER, R.; BRUNNER, R.; BRÜNING, R. Multi-aperture system approach for snapshot multispectral imaging applications. *Optics Express* [online]. 2021, vol. 29, no. 5, pp. 7361–7378 [visited on 2022-08-25]. ISSN 1094-4087. Available from: doi:10.1364/OE.412655.

[46] MONAKHOVA, K.; YANNY, K.; YANNY, K.; AGGARWAL, N.; WALLER, L.; WALLER, L. Spectral DiffuserCam: lensless snapshot hyperspectral imaging with a spectral filter array. *Optica* [online]. 2020, vol. 7, no. 10, pp. 1298–1307 [visited on 2022-08-25]. ISSN 2334-2536. Available from: doi:10.1364/OPTICA.397214.

[47] HAGEN, N.; KUDENOV, M. W. Review of snapshot spectral imaging technologies. *Optical Engineering* [online]. 2013, vol. 52, no. 9, p. 090901 [visited on 2022-08-25]. ISSN 0091-3286. Available from: doi:10.1117/1.OE.52.9.090901.

[48] AASEN, H.; BURKART, A.; BOLTEN, A.; BARETH, G. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing* [online]. 2015, vol. 108, pp. 245–259 [visited on 2025-05-15]. ISSN 0924-2716. Available from: doi:10.1016/j.isprsjprs.2015.08.002.

[49] CHRIEN, T. G.; GREEN, R. O.; EASTWOOD, M. L. Accuracy of the spectral and radiometric laboratory calibration of the Airborne Visible/Infrared Imaging Spectrometer. In: *Imaging Spectroscopy of the Terrestrial Environment* [online]. SPIE, 1990, vol. 1298, pp. 37–49 [visited on 2025-04-16]. Available from: doi:10.1117/12.21334. Imaging Spectroscopy of the Terrestrial Environment.

[50] HENRIKSEN, M. B.; SIGERNES, F.; JOHANSEN, T. A. A Closer Look At A Spectrographic Wavelength Calibration. In: *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)* [online]. Rome, Italy: IEEE, 2022, pp. 1–5 [visited on 2025-04-16]. ISBN 978-1-6654-7069-8. Available from: doi:10.1109/WHISPERS56178.2022.9955104. 2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS).

[51] LIU, D.; HENNELLY, B. M. Improved Wavelength Calibration by Modeling the Spectrometer. *Applied Spectroscopy* [online]. 2022, vol. 76, no. 11, pp. 1283–1299 [visited on 2025-04-16]. ISSN 0003-7028. Available from: doi:10.1177/00037028221111796.

[52] *ENVI*. Boulder, Colorado: Exelis Visual Information Solutions, 2010. Available also from: https://www.nv5geospatialsoftware.com/Products/ENVI.

[53] KRAMIDA, A.; RALCHENKO, Y. *NIST Atomic Spectra Database, NIST Standard Reference Database 78* [online]. National Institute of Standards and Technology, 1999 [visited on 2023-07-23]. Available from: doi:10.18434/T4W30F.

[54] HENRIKSEN, M. L.; PEDERSEN, W. N.; KLARSKOV, P.; HINGE, M. One step calibration of industrial hyperspectral cameras. *Chemometrics and Intelligent Laboratory Systems* [online]. 2022, vol. 227, p. 104609 [visited on 2023-08-01]. ISSN 0169-7439. Available from: doi:10.1016/j.chemolab.2022.104609.

[55] KUGLIN, C.; HINES, D. The Phase Correlation Image Alignment Methed. In: *Proceeding of IEEE International Conference on Cybernetics and Society* [online]. New York, NY, USA: IEEE, 1975, vol. 1975, pp. 163–165 [visited on 2025-04-16]. Available from: https://scispace.com/papers/the-phase-correlation-image-alignment-methed-5gpmeuq16n.

[56] YOKOYA, N.; MIYAMURA, N.; IWASAKI, A. Detection and correction of spectral and spatial misregistrations for hyperspectral data using phase correlation method. *Applied Optics*. 2010, vol. 49, no. 24, pp. 4568–4575. ISSN 1539-4522. Available from: doi:10.1364/AO.49.004568.

[57] HARRI. *Smile and keystone* [Specim] [online]. 2021-02-15 [visited on 2023-08-01]. Available from: https://www.specim.com/smile-and-keystone/.

[58] DERENIAK, E. L.; CROWE, D. G. *Optical radiation detectors*. New York: Wiley, 1984. Wiley series in pure & applied optics. ISBN 978-0-471-89797-2.

[59] IRIE, K.; WOODHEAD, I. M.; MCKINNON, A. E.; UNSWORTH, K. Measured effects of temperature on illumination-independent camera noise. In: *2009 24th International Conference Image and Vision Computing New Zealand* [online]. 2009, pp. 249–253 [visited on 2025-05-15]. Available from: doi:10.1109/IVCNZ.2009.5378403. 2009 24th International Conference Image and Vision Computing New Zealand.

[60] SHAIKH, M. S.; JAFERZADEH, K.; THÖRNBERG, B.; CASSELGREN, J. Calibration of a Hyper-Spectral Imaging System Using a Low-Cost Reference. *Sensors (Basel, Switzerland)* [online]. 2021, vol. 21, no. 11, p. 3738 [visited on 2023-08-01]. ISSN 1424-8220. Available from: doi:10.3390/s21113738.

[61] CIURCZAK, E. W.; IGNE, B.; WORKMAN, J.; BURNS, D. A. *Handbook of Near-Infrared Analysis* [online]. 4th ed. Ed. by CIURCZAK, E. W.; IGNE, B.; WORKMAN, J.; BURNS, D. A. Fourth edition. — Boca Raton : Taylor and Francis, 2021. —: CRC Press, 2021 [visited on 2025-03-20]. ISBN 978-1-351-26988-9. Available also from: doi:10.1201/b22513.

[62] *HSZ Data Format – Scyllarus – Hyperspectral Image Processing* [online]. [visited on 2023-08-04]. Available from: https://scyllarus.data61.csiro.au/data/hsz-data-format/.

[63] DUA, Y.; KUMAR, V.; SINGH, R. S. Comprehensive review of hyperspectral image compression algorithms. *Optical Engineering* [online]. 2020, vol. 59, no. 9, p. 090902 [visited on 2023-08-04]. ISSN 0091-3286, ISSN 1560-2303. Available from: doi:10.1117/1.OE.59.9.090902.

[64] HE, Y.; TU, B.; LIU, B.; LI, J.; PLAZA, A. 3DSS-Mamba: 3D-Spectral-Spatial Mamba for Hyperspectral Image Classification. *ArXiv* [online]. 2024 [visited on 2024-05-28]. ISSN 00000000. Available from: doi:https://doi.org/10.48550/arXiv.2405.12487.

[65] YE, Z.; LI, C.; LIU, Q.; BAI, L.; FOWLER, J. E. Multiscale Spatial-Spectral Feature Extraction Network for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* [online]. 2022, pp. 1–1 [visited on 2022-06-09]. ISSN 1939-1404, ISSN 2151-1535. Available from: doi:10.1109/JSTARS.2022.3179446.

[66] ROY, S. K.; KRISHNA, G.; DUBEY, S. R.; CHAUDHURI, B. B. HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*. 2020, vol. 17, no. 2, pp. 277–281. ISSN 1558-0571. Available from: doi:10.1109/LGRS.2019.2918719.

[67] VACULÍK, O. *Využití senzoru v okrajových pásmech citlivosti*. Bakalářská práce. Olomouc: Univerzita Palackého v Olomouci, Přírodovědecká fakultaOlomouc, 2018. Available also from: https://theses.cz/id/rkg6x1/.

[68] BRAUN, M.; HEROLD, M. Mapping imperviousness using NDVI and linear spectral unmixing of ASTER data in the Cologne-Bonn region (Germany). In: *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III* [online]. SPIE, 2004, vol. 5239, pp. 274–284 [visited on 2023-07-26]. Available from: doi:10.1117/12.510978. Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III.

[69] GAO, B.-c. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* [online]. 1996, vol. 58, no. 3, pp. 257–266 [visited on 2023-07-23]. ISSN 00344257. Available from: doi:10.1016/S0034-4257(96)00067-3.

[70] HUNT, E. R.; DAUGHTRY, C. S. T.; EITEL, J. U. H.; LONG, D. S. Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index. *Agronomy Journal* [online]. 2011, vol. 103, no. 4, pp. 1090–1099 [visited on 2023-07-23]. ISSN 0002-1962, ISSN 1435-0645. Available from: doi:10.2134/agronj2010.0395.

[71] MONTERO, D.; AYBAR, C.; MAHECHA, M. D.; MARTINUZZI, F.; SÖCHTING, M.; WIENEKE, S. A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research. *Scientific Data* [online]. 2023, vol. 10, no. 1, p. 197 [visited on 2023-07-23]. ISSN 2052-4463. Available from: doi:10.1038/s41597-023-02096-0.

[72] *Sentinel Playground*. Available also from: https://apps.sentinel-hub.com/sentinel-playground.

[73] CHANG, C.-I. *Hyperspectral Data Processing: Algorithm Design and Analysis* [online]. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2013 [visited on 2023-08-18]. ISBN 978-1-118-26977-0. Available also from: doi:10.1002/9781118269787.

[74] ADAMS, J. B.; SMITH, M. O.; JOHNSON, P. E. Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site. *Journal of Geophysical Research: Solid Earth* [online]. 1986, vol. 91, pp. 8098–8112 [visited on 2023-07-26]. ISSN 2156-2202. Available from: doi:10.1029/JB091iB08p08098.

[75] WINTER, M. E. N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. In: *Imaging Spectrometry V* [online]. SPIE, 1999, vol. 3753, pp. 266–275 [visited on 2023-01-26]. Available from: doi:10.1117/12.366289. Imaging Spectrometry V.

[76] CHANG, C.-I.; PLAZA, A. A fast iterative algorithm for implementation of pixel purity index. *IEEE Geoscience and Remote Sensing Letters*. 2006, vol. 3, no. 1, pp. 63–67. ISSN 1558-0571. Available from: doi:10.1109/LGRS.2005.856701.

[77] BHATT, J. S.; JOSHI, M. V. Deep Learning in Hyperspectral Unmixing: A Review. In: *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. 2020, pp. 2189–2192. Available from: doi:10.1109/IGARSS39084.2020.9324546. IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium.

[78] VACULÍK, O.; BERNATOVÁ, S.; REBROŠOVÁ, K.; SAMEK, O.; ŠILHAN, L.; RŮŽIČKA, F.; ŠERÝ, M.; ŠILER, M.; JEŽEK, J.; ZEMÁNEK, P. Rapid Identification of Pathogens in Blood Serum Via Raman Tweezers in Combination with Advanced Processing Methods. *Biomedical Optics Express* [online]. 2023, vol. 14, no. 12, p. 6410 [visited on 2024-01-05]. ISSN 2156-7085, ISSN 2156-7085. Available from: doi:10.1364/BOE.503628.

[79] VACULÍK, O. *Raman Identification Repository* [online]. Brno, Czech Republic, 2022 [visited on 2025-05-23]. Available from: https://svn.isibrno.cz/vacuon/raman-identification.

[80] LAWTON, W. H.; SYLVESTRE, E. A. Self Modeling Curve Resolution. *Technometrics* [online]. 1971, vol. 13, no. 3, pp. 617–633 [visited on 2025-03-14]. ISSN 0040-1706. Available from: doi:10.1080/00401706.1971.10488823.

[81] GHAFFARI, M.; OMIDIKIA, N.; RUCKEBUSCH, C. Essential Spectral Pixels for Multivariate Curve Resolution of Chemical Images. *Analytical Chemistry* [online]. 2019, vol. 91, no. 17, pp. 10943–10948 [visited on 2024-11-08]. ISSN 0003-2700. Available from: doi:10.1021/acs.analchem.9b02890.

[82] BRO, R.; DE JONG, S. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics* [online]. 1997, vol. 11, no. 5, pp. 393–401 [visited on 2025-05-28]. ISSN 1099-128X. Available from: doi:10.1002/(SICI)1099-128X(199709/10)11:5<393::AID-CEM483>3.0.CO;2-L.

[83] GHAMISI, P.; PLAZA, J.; CHEN, Y.; LI, J.; PLAZA, A. J. Advanced Spectral Classifiers for Hyperspectral Images: A review. *IEEE Geoscience and Remote Sensing Magazine* [online]. 2017, vol. 5, no. 1, pp. 8–32 [visited on 2022-05-31]. ISSN 2168-6831, ISSN 2473-2397, ISSN 2373-7468. Available from: doi:10.1109/MGRS.2016.2616418.

[84] REHMAN, S. U.; KUMAR, A.; BANERJEE, A. SNR improvement for hyperspectral application using frame and pixel binning. In: XIONG, X. J.; KURIAKOSE, S. A.; KIMURA, T. (eds.) [online]. New Delhi, India, 2016, 98810Y [visited on 2022-06-20]. Available from: doi:10.1117/12.2220599. SPIE Asia-Pacific Remote Sensing.

[85] LEE, D.; SEUNG, H. S. Algorithms for Non-negative Matrix Factorization. In: *Advances in Neural Information Processing Systems* [online]. MIT Press, 2000, vol. 13 [visited on 2025-05-28]. Available from: https://proceedings.neurips.cc/paper_files/paper/2000/hash/f9d1152547c0bde01830b7e8bd60024c-Abstract.html.

[86] PAUCA, V. P.; PIPER, J.; PLEMMONS, R. J. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra and its Applications* [online]. 2006, vol. 416, no. 1, pp. 29–47 [visited on 2025-06-03]. ISSN 0024-3795. Available from: doi:10.1016/j.laa.2005.06.025.

[87] FÉVOTTE, C.; DOBIGEON, N. Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization. *IEEE Transactions on Image Processing* [online]. 2015, vol. 24, no. 12, pp. 4810–4819 [visited on 2024-10-10]. ISSN 1057-7149, ISSN 1941-0042. Available from: doi:10.1109/TIP.2015.2468177.

[88] JOLLIFFE, I. T. *Principal component analysis*. 2nd ed. New York: Springer, 2002. Springer series in statistics. ISBN 978-0-387-95442-4.

[89] LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. Principal component analysis. *Nature Methods* [online]. 2017, vol. 14, no. 7, pp. 641–642 [visited on 2023-02-08]. ISSN 1548-7105. Available from: doi:10.1038/nmeth.4346.

[90] REDDY, T.; HARIKIRAN, J. An outlook: machine learning in hyperspectral image classification and dimensionality reduction techniques. *Journal of Spectral Imaging* [online]. 2022, a1 [visited on 2022-05-31]. ISSN 2040-4565. Available from: doi:10.1255/jsi.2022.a1.

[91] MINKIN, A. The Application of Autoencoders for Hyperspectral Data Compression. In: *2021 International Conference on Information Technology and Nanotechnology (ITNT)* [online]. 2021, pp. 1–4 [visited on 2024-09-30]. Available from: doi:10.1109/ITNT52450.2021.9649318. 2021 International Conference on Information Technology and Nanotechnology (ITNT).

[92] SHUKLA, N. *Machine learning with TensorFlow* [online]. Shelter Island, NY: Manning Publications, 2018 [visited on 2023-07-17]. ISBN 978-1-61729-387-0. Available from: https://proquest.safaribooksonline.com/9781617293870.

[93] WILLARD, C. A. *Statistical Methods: An Introduction to Basic Statistical Concepts and Analysis* [online]. 2nd ed. Routledge, 2020 [visited on 2025-04-22]. ISBN 978-0-429-26103-9. Available also from: doi:10.4324/9780429261039.

[94] RASHMI, S.; ADDAMANI, S.; RAVIKIRAN, A. Spectral Angle Mapper Algorithm for Remote Sensing Image Classification. In: [online]. 2014 [visited on 2023-02-08]. Available from: https://www.semanticscholar.org/paper/Spectral-Angle-Mapper-Algorithm-for-Remote-Sensing-Rashmi-Addamani/c192da3f6560c0305926149e7f6324dab441201b.

[95] MACKAY, D. J. C. *Information theory, inference, and learning algorithms*. Cambridge, UK ; New York: Cambridge University Press, 2003. ISBN 978-0-521-64298-9.

[96] *Thrive — Bridgelux, Inc. LED Lighting* [Bridgelux Thrive] [online]. [visited on 2023-08-07]. Available from: https://www.bridgelux.com/products/thrive#specifications.

[97] *Thorlabs - SLS201L/M Stabilized Fiber-Coupled Light Source w/ Universal Power Adapter, 360 - 2600 nm, M6 Taps* [Thorlabs] [online]. [visited on 2023-08-07]. Available from: https://www.thorlabs.de.

[98] ŠERÝ, M.; ŠILHAN, L.; VACULÍK, O.; MAŇKA, T.; RATAJ, T.; TRTÍLEK, M. *Hyperspektrální kamera s vysokým rozlišením a rozsahem vlnových délek 350-1100 nm*. 2022. Available also from: https://www.isvavai.cz/riv?s=jednoduche-vyhledavani%5C&ss=detail%5C&n=0%5C&h=RIV%5C%2F68081731%5C%3A_____%5C%2F22%5C%3A00566587%5C%21RIV23-MPO-68081731.

[99] VACULÍK, O. *HyperSpec Repository*. Brno, Czech Republic, 2022. Available also from: https://svn.isibrno.cz/vacuon/hyperspec.

[100] *basler/pypylon: The official python wrapper for the pylon Camera Software Suite* [online]. [visited on 2023-08-08]. Available from: https://github.com/basler/pypylon.

[101] MARPLE, S. *Python binding for the ZWO ASI v2 library* [online]. python-zwoasi, 2025 [visited on 2025-06-30]. Available from: https://github.com/python-zwoasi/python-zwoasi.

[102] *ni/nidaqmx-python: A Python API for interacting with NI-DAQmx* [online]. [visited on 2023-08-08]. Available from: https://github.com/ni/nidaqmx-python.

[103] *libximc: libximc library* [online]. [visited on 2023-08-08]. Available from: https://libximc.xisupport.com/doc-en/.

[104] VACULÍK, O.; ŠERÝ, M.; ŠILHAN, L.; ŠILHANOVÁ, D.; ZEMÁNEK, P. Addressing Data Management for Custom Built Hyperspectral Microscopy Station. In: *2024 14th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)* [online]. Helsinki, Finland: IEEE, 2024, pp. 1–5 [visited on 2025-03-03]. ISBN 979-8-3315-1313-9. Available from: doi:10.1109/WHISPERS65427.2024.10876463. 2024 14th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS).

[105] KORANNE, S. Hierarchical Data Format 5 : HDF5. In: KORANNE, S. *Handbook of Open Source Tools* [online]. Boston, MA: Springer US, 2011, pp. 191–200 [visited on 2024-09-30]. ISBN 978-1-4419-7718-2. Available from: doi:10.1007/978-1-4419-7719-9_10.

[106] PEZOA, F.; REUTTER, J. L.; SUAREZ, F.; UGARTE, M.; VRGOČ, D. Foundations of JSON Schema. In: *Proceedings of the 25th International Conference on World Wide Web* [online]. Montréal Québec Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 263–273 [visited on 2024-09-30]. ISBN 978-1-4503-4143-1. Available from: doi:10.1145/2872427.2883029. WWW '16: 25th International World Wide Web Conference.

[107] WILKINSON, M. D.; DUMONTIER, M.; AALBERSBERG, I. J.; APPLETON, G.; AXTON, M.; BAAK, A.; BLOMBERG, N.; BOITEN, J.-W.; SILVA SANTOS, L. B. da; BOURNE, P. E.; BOUWMAN, J.; BROOKES, A. J.; CLARK, T.; CROSAS, M.; DILLO, I.; DUMON, O.; EDMUNDS, S.; EVELO, C. T.; FINKERS, R.; GONZALEZ-BELTRAN, A.; GRAY, A. J. G.; GROTH, P.; GOBLE, C.; GRETHE, J. S.; HERINGA, J.; 'T HOEN, P. A. C.; HOOFT, R.; KUHN, T.; KOK, R.; KOK, J.; LUSHER, S. J.; MARTONE, M. E.; MONS, A.; PACKER, A. L.; PERSSON, B.; ROCCA-SERRA, P.; ROOS, M.; SCHAIK, R. van; SANSONE, S.-A.; SCHULTES, E.; SENGSTAG, T.; SLATER, T.; STRAWN, G.; SWERTZ, M. A.; THOMPSON, M.; LEI, J. van der; MULLIGEN, E. van; VELTEROP, J.; WAAGMEESTER, A.; WITTENBURG, P.; WOLSTENCROFT, K.; ZHAO, J.; MONS, B. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* [online]. 2016, vol. 3, no. 1, p. 160018 [visited on 2024-09-24]. ISSN 2052-4463. Available from: doi:10.1038/sdata.2016.18.

[108] HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; RÍO, J. F. del; WIEBE, M.; PETERSON, P.; GÉRARD-MARCHANT,

P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. *Nature* [online]. 2020, vol. 585, no. 7825, pp. 357–362 [visited on 2022-06-01]. ISSN 0028-0836, ISSN 1476-4687. Available from: doi:10.1038/s41586-020-2649-2.

[109] MARRS, T. *JSON at Work: Practical Data Integration for the Web*. "O'Reilly Media, Inc.", 2017. ISBN 978-1-4919-8241-9.

[110] RESONON. *User Manual for Spectronon Version 3.5.8*. 2025. Available also from: https://docs.resonon.com/spectronon/SpectrononUserManual/SpectrononUserManual.pdf.

[111] *IEEE DataPort* [IEEE DataPort] [online]. [visited on 2024-09-26]. Available from: https://ieee-dataport.org/.

[112] EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH; OPENAIRE. *Zenodo: Research. Shared.* [online]. CERN, 2013 [visited on 2024-09-29]. Available from: doi:10.25495/7GXK-RD71.

[113] COLLETTE, A.; KLUYVER, T.; CASWELL, T. A.; TOCKNELL, J.; KIEFFER, J.; JELENAK, A.; SCOPATZ, A.; DALE, D.; CHEN; VINCENT, T.; EINHORN, M.; PAYNO; JULIAGARRIGA; SCIARELLI, P.; VALLS, V.; GHOSH, S.; PEDERSEN, U. K.; KITTISOPIKUL, M.; JAKIRKHAM; RASPAUD, M.; DANILEVSKI, C.; ABBASI, H.; READEY, J.; MÜHLBAUER, K.; PARAMONOV, A.; CHAN, L.; SCHEPPER, R. D.; SOLÉ, V. A.; JIALIN; GUEST, D. H. *h5py/h5py: 3.8.0-aarch64-wheels* [online]. Zenodo, 2023. 3.8.0-aarch64-wheels [visited on 2024-09-18]. Available from: doi:10.5281/zenodo.7568214.

[114] LA GRASSA, R.; RE, C.; CREMONESE, G.; GALLO, I. Hyperspectral Data Compression Using Fully Convolutional Autoencoder. *Remote Sensing* [online]. 2022, vol. 14, no. 10, p. 2472 [visited on 2024-09-08]. ISSN 2072-4292. Available from: doi:10.3390/rs14102472.

[115] SCHEELINE, A. How to Design a Spectrometer. *Applied Spectroscopy* [online]. 2017, vol. 71, no. 10, pp. 2237–2252 [visited on 2022-05-31]. ISSN 0003-7028, ISSN 1943-3530. Available from: doi:10.1177/0003702817720468.

[116] SOILLE, P. J.; ANSOULT, M. M. Automated basin delineation from digital elevation models using mathematical morphology. *Signal Processing* [online]. 1990, vol. 20, no. 2, pp. 171–182 [visited on 2025-04-24]. ISSN 0165-1684. Available from: doi:10.1016/0165-1684(90)90127-K.

[117] PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011, vol. 12, no. 85, pp. 2825–2830. Available from: doi:10.5555/1953048.2078195.

[118] CHOLLET, F. et al. *Keras* [https://keras.io]. 2015.

[119] THERIEN, C. *pysptools 0.15.0 documentation* [online]. 2018-10-11 [visited on 2025-06-06]. Available from: https://pysptools.sourceforge.io/.

[120] CAMP, C. H. pyMCR: A Python Library for Multivariate Curve Resolution Analysis with Alternating Regression (MCR-AR). *Journal of Research of the National Institute of Standards and Technology* [online]. 2019, vol. 124, p. 124018 [visited on 2025-06-03]. ISSN 2165-7254. Available from: doi:10.6028/jres.124.018.

[121] VACULÍK, O. *hyperspectral_ipht Repository* [online]. Brno, Czech Republic, 2024 [visited on 2025-05-23]. Available from: https://svn.isibrno.cz/vacuon/hyperspectral_ipht.

[122] JOBLOVE, G. H.; GREENBERG, D. Color spaces for computer graphics. *SIGGRAPH Comput. Graph.* [online]. 1978, vol. 12, no. 3, pp. 20–25 [visited on 2025-05-24]. ISSN 0097-8930. Available from: doi:10.1145/965139.807362.

[123] GOERICKE, R.; WELSCHMEYER, N. A. The marine prochlorophyte Prochlorococcus contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. *Deep Sea Research Part I: Oceanographic Research Papers* [online]. 1993, vol. 40, no. 11, pp. 2283–2294 [visited on 2025-06-02]. ISSN 09670637. Available from: doi:10.1016/0967-0637(93)90104-B.

[124] LEGLEITER, C. J.; KING, T. V.; CARPENTER, K. D.; HALL, N. C.; MUMFORD, A. C.; SLONECKER, T.; GRAHAM, J. L.; STENGEL, V. G.; SIMON, N.; ROSEN, B. H. Spectral mixture analysis for surveillance of harmful algal blooms (SMASH): A field-, laboratory-, and satellite-based approach to identifying cyanobacteria genera from remotely sensed data. *Remote Sensing of Environment* [online]. 2022, vol. 279, p. 113089 [visited on 2025-05-30]. ISSN 00344257. Available from: doi:10.1016/j.rse.2022.113089.

[125] PARTENSKY, F.; HESS, W. R.; VAULOT, D. Prochlorococcus, a marine photosynthetic prokaryote of global significance. *Microbiology and molecular biology reviews: MMBR*. 1999, vol. 63, no. 1, pp. 106–127. ISSN 1092-2172. Available from: doi:10.1128/MMBR.63.1.106-127.1999.

[126] WALTER, J. M.; COUTINHO, F. H.; DUTILH, B. E.; SWINGS, J.; THOMPSON, F. L.; THOMPSON, C. C. Ecogenomics and Taxonomy of Cyanobacteria Phylum. *Frontiers in Microbiology*. 2017, vol. 8, p. 2132. ISSN 1664-302X. Available from: doi:10.3389/fmicb.2017.02132.

[127] ROCAP, G.; LARIMER, F. W.; LAMERDIN, J.; MALFATTI, S.; CHAIN, P.; AHLGREN, N. A.; ARELLANO, A.; COLEMAN, M.; HAUSER, L.; HESS, W. R.; JOHNSON, Z. I.; LAND, M.; LINDELL, D.; POST, A. F.; REGALA, W.; SHAH, M.; SHAW, S. L.; STEGLICH, C.; SULLIVAN, M. B.; TING, C. S.; TOLONEN, A.; WEBB, E. A.; ZINSER, E. R.; CHISHOLM, S. W. Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. *Nature*. 2003, vol. 424, no. 6952, pp. 1042–1047. ISSN 1476-4687. Available from: doi:10.1038/nature01947.

[128] ONG, L. J.; GLAZER, A. N. Phycoerythrins of marine unicellular cyanobacteria. I. Bilin types and locations and energy transfer pathways in Synechococcus spp. phycoerythrins. *The Journal of Biological Chemistry*. 1991, vol. 266, no. 15, pp. 9515–9527. ISSN 0021-9258.

[129] HAVLÍČKOVÁ, A. *Utilization of Advanced Techniques of Electron Microscopy in Altrastructural Analysis of Microbial Cells and Microbial Biotechnology Products*. PhD thesis. Brno: Faculty of Chemistry, Brno University of Technology, 2024.

[130] TING, C. S.; HSIEH, C.; SUNDARARAMAN, S.; MANNELLA, C.; MARKO, M. Cryo-Electron Tomography Reveals the Comparative Three-Dimensional Architecture of Prochlorococcus, a Globally Important Marine Cyanobacterium. *Journal of Bacteriology* [online]. 2007, vol. 189, no. 12, pp. 4485–4493 [visited on 2025-06-11]. ISSN 0021-9193. Available from: doi:10.1128/JB.01948-06.

[131] HOHMANN-MARRIOTT, M. F. (ed.). *The Structural Basis of Biological Energy Generation*. Vol. 39 [online]. Dordrecht: Springer Netherlands, 2014 [visited on 2025-06-11]. Advances in Photosynthesis and Respiration. ISBN 978-94-017-8741-3 978-94-017-8742-0. Available from: doi:10.1007/978-94-017-8742-0.

[132] CULTURE COLLECTION, R. *PCR S11 - Red Sea medium v1* [online]. 2018 [visited on 2025-05-30]. Available from: doi:10.17504/protocols.io.sz3ef8n.

[133] PHOTONICS, H. *Headwall SIF Imaging Sensor* [online]. [visited on 2023-08-03]. Available from: https://headwallphotonics.sharefile.com/share/view/bc992cfa68934e41.

[134] INSTRUMENTS, P. S. *HySpec VNIR Camera* [Photon System Instruments] [online]. [visited on 2023-08-03]. Available from: http://www.hyperspec.org/products/vnir/.

[135] PHOTONICS, H. *MV.C VNIR Datasheet* [Headwall Photonics] [online]. 2023-01-03 [visited on 2023-08-03]. Available from: https://headwallphotonics.sharefile.com/share/view/b24696652a4a49d4.

[136] RESONON. *Pika XC2 - Hyperspectral Imaging Cameras* [online]. [visited on 2023-08-03]. Available from: https://resonon.com/Pika-XC2.

[137] HYSPEX. *HySpex VNIR-3000 N* [HySpex] [online]. [visited on 2023-08-03]. Available from: https://www.hyspex.com/hyspex-products/hyspex-classic/hyspex-vnir-3000-n/.

[138] CORP., S. O. *SOC710 Series Hyperspectral Imaging* [Surface Optics Corp.] [online]. [visited on 2023-08-03]. Available from: https://surfaceoptics.com/products/hyperspectral-imaging/soc710-portable-hyperspectral-camera/.

[139] RESONON. *Pika UV - Hyperspectral Imaging Camera* [online]. [visited on 2023-08-03]. Available from: https://resonon.com/Pika-UV.

[140] BAYSPEC. *OCI F Hyperspectral Cameras*. [N.d.]. Available also from: https://www.bayspec.com/wp-content/uploads/2022/12/New-BaySpec-Datasheet-OCI-F-Hyperspectral-Cameras.pdf.

[141] CLYDEHSI. *VNIR-HR Hyperspectral Camera 400-1000nm* [ClydeHSI] [online]. [visited on 2023-08-03]. Available from: https://www.clydehsi.com/hyperspectral-cameras/vnir-hr-hyperspectral-camera.

[142] HYSPEX. *Mjolnir UAV-based hyperspectral linescanners — SphereOptics EN* [online]. [visited on 2023-08-03]. Available from: https://sphereoptics.de/en/product/hyspex-mjolnir-uav-based-hyperspectral-linescanners/.

[143] *Hyperspectral dataset list*. [N.d.]. Available also from: https://paperswithcode.com/datasets?mod=hyperspectral-images.

[144] RESONON. *Pika L - Hyperspectral Sensors - Benchtop Systems* [online]. [visited on 2023-08-03]. Available from: https://resonon.com/Pika-L.

[145] HYSPEX. *HySpex Baldur V-1024 N* [online]. [visited on 2023-08-03]. Available from: https://www.hyspex.com/hyspex-products/hyspex-baldur/hyspex-baldur-v-1024-n/.

[146] SPECIM. *Specim FX10* [Specim] [online]. [visited on 2023-08-03]. Available from: https://www.specim.com/products/specim-fx10/.

[147] PHOTONICS, H. *MV.X Datasheet*. 2022. Available also from: https://www.headwallphotonics.com/hubfs/Headwall_MVX_Datasheet_04May2022.pdf.

[148] CLYDEHSI. *VNIR-S Hyperspectral Camera 400-1000nm* [ClydeHSI] [online]. [visited on 2023-08-03]. Available from: https://www.clydehsi.com/hyperspectral-cameras/vnir-s-hyperspectral-camera.

[149] CORNING. *microHSI 410 SHARK*. [N.d.]. Available also from: https://www.corning.com/media/worldwide/csm/documents/microHSI_410%20Sensor_and_microHSI_410_SHARK.pdf.

[150] INNO-SPEC. *RedEye* [INNO-SPEC] [online]. [visited on 2023-08-03]. Available from: https://inno-spec.de/en/redeye-en/.

[151] VACULÍK, O. */ HyperSpec · GitLab* [GitLab] [online]. 2025-04-01 [visited on 2025-05-23]. Available from: https://svn.isibrno.cz/vacuon/hyperspec.

Supplementary Material

Table S1: Boolean search strings used for the literature review.

Database	Boolean search Strings
WOS	`(AB=(hyperspectral)) AND (AB=(remote sensing))`
	`(AB=(hyperspectral)) AND AB=(microscop*)`
IEEE	`("Abstract":hyperspectral) AND ("All Metadata":remote sensing)`
	`("Abstract":hyperspectral) AND ("Abstract":microscop*)`
Scopus	`(TITLE-ABS-KEY (hyperspectral) AND TITLE-ABS-KEY (remote AND sensing)) AND (LIMIT-TO(DOCTYPE, "ar"))`
	`(TITLE-ABS-KEY (hyperspectral) AND TITLE-ABS-KEY (microscop*)) AND (LIMIT-TO(DOCTYPE, "ar")`
Pubmed	`(hyperspectral[Title/Abstract]) AND (remote sensing[Title/Abstract])`
	`(hyperspectral[Title/Abstract]) AND (microscop*[Title/Abstract])`

Table S2: Accuracy of microbe identification in blood serum (from the test set), including uncertainties calculated by the Root Mean Square Error (RMSE). Serum A-E labels serum spiked with four microbes (*C. albicans*, *E. coli*, *S. aureus*, and *S. epidermidis*). FCLS labels Fully Constrained Least Squares, NNLS labels Non-Negative Least Squares, and UCLS labels Unconstrained Least Squares, KNN+IPF labels K-Nearest Neighbors Algorithm applied to spectra with background removed by Iterative Polynomial Fitting, KNN+RCF labels K-Nearest Neighbors Algorithm applied to spectra with background removed by Rolling Circle Filter. Original result of the author, published in [78].

	Spectral Mixture			**Background Removed**	
Serum	**FCLS**	**NNLS**	**UCLS**	**KNN+IPF**	**KNN+RCF**
A	0.919	0.871	0.935	0.919	0.823
B	0.897	0.897	0.845	0.845	0.655
C	0.907	0.741	0.907	0.833	0.685
D	0.927	0.927	0.891	0.909	0.818
E	0.833	0.778	0.778	0.750	0.778
Average	**0.897**	0.843	0.871	**0.851**	0.752
RMSE	0.008	0.016	0.013	0.014	0.016

Table S3: Comparison of PCA, NMF, and Autoencoder (fully connected AE-FC and Convolutional AE-CNN) algorithms.

	PCA	**NMF**	**AE-FC**	**AE-CNN**
Data Constraints	None	Non-negativity	Depends on activation	Depends on activation (non-negative)
Factorization	Orthogonal basis	Additive parts-based	Learning, via back-propagation	Back-propagation with spatial structure
Interpretability	Low	High	Medium (architecture-dependent)	Medium-High (spatial features)
Optimization Goal	Maximize variance	Minimize error under non-negativity	Minimize reconstruction loss	Same as AE-FC with spatial encoding
Linearity	Linear	Linear	Non-linear	Non-linear
Output Dimensionality	Linear projection	Linear additive	Learned latent space	Learned latent space

Table S4: Comparison of pushbroom HS systems, sorted by the spectral resolution $\Delta\lambda$, with wavelength range B_λ and number of bands N_λ.

Manufacturer	**Model**	**$\Delta\lambda$[nm]**	**B_λ[nm]**	**N_λ**	**Ref.**
PSI	HypCam2	0.1	350–1100	1370	–
Headwall	SIF	0.25	671–780	2134	[133]
PSI	VNIR	0.8	350–900	640	[134]
Headwall	MV.C	1.75	400–1000	342	[135]
Resonon	Pika XC2	1.9	400–1000	447	[136]
HySpex	VNIR-3000 N	2	400–1000	300	[137]
Surface Optics	710-E	2.31	400–1000	260	[138]
Resonon	Pika UV	2.8	330–800	255	[139]
BaySpec	OCI-F-HR	3	400–1100	240	[140]
Clyde HSI	VNIR-HR	3	400–1000	2048	[141]
HySpex	Mjolnir V-1240	3	400–1000	200	[142]
HySpex	VNIR-1800	3.26	400–1000	182	[143]
Resonon	Pika L	3.3	400–1000	281	[144]
Surface Optics	710-VP	4.69	400–1000	128	[138]
HySpex	Baldur V-1024 N	5.5	400–1000	106	[145]
Specim	FX10	5.5	400–1000	224	[146]
BaySpec	OCI-F	6	400–1100	120	[140]
Headwall	MV.X	6	400–1000	301	[147]
Clyde HSI	VNIR-S	8	400–1000	1024	[148]
Corning	microHSI 410	8	400–1000	308	[149]
INNO-SPEC	Redeye 1.7	9	950–1700	256	[150]

```
connect_hardware(linear, hs_camera, basler, stage)
hs_camera.cooler(1)
hs_camera.temperature(TEMPERATURE)
hs_camera.shutter(SHUTTER_MS * 1000)
hs_camera.gain(GAIN)
hs_camera.binning(BINNING)

if START_POSITION != {}:
    stage.move_to_coordinates(START_POSITION)
stage_start = stage.get_position()
logging.info(f"Stage arrived to start position {stage_start}")

daq.illumination(1)
daq.piezo(preset["piezo_voltage"])

with File(Path(DIRECTORY) / f"{fname}-{NAME}.h5", "w") as file:
    logging.debug("Opening HDF file")
    linear_basler(linear)
    file.create_dataset("preview", data=basler.capture_image())
    linear_hs(linear)

    file.create_group("hyperspectral")
    slit, bands = hs_camera.shape
    dset = file["hyperspectral"].create_dataset(
        "data", (bands, slit, FRAMES),  # this order generates fast HDF previews
        dtype="uint16", chunks=(16, slit, 10), compression="lzf")

    camera_cooling_wait(hs_camera, TEMPERATURE)
    for frame in range(FRAMES):
        dset[..., DIRECTION * frame] = hs_camera.one_frame().T
        stage.move_steps(2, step)
    stage_end = stage.get_position()

    stage.move_to_coordinates(WHITE_POSITION)
    white = np.array([hs_camera.one_frame().T for _ in range(CALIBRATION_FRAMES)])
    file["hyperspectral"].create_dataset("white", data=white)

    daq.illumination(0)
    dark = np.array([hs_camera.one_frame().T for _ in range(CALIBRATION_FRAMES)])
    file["hyperspectral"].create_dataset("dark", data=dark)

    file.create_group("spectral_reference")
    file["spectral_reference"].create_dataset("peak_wavelength", data=avantes_peaks[0])
    file["spectral_reference"].create_dataset("peak_intensity", data=avantes_peaks[1])
    logging.info("Wavelength calibration start")
    original_x = linear.position["X"]
    linear.move_to("X", 3500)
    daq.calibration_lamp(1)
    hs_camera.shutter(int(1e7))  # 10s
    avantes_data = np.array([hs_camera.one_frame().T for _ in range(CALIBRATION_FRAMES)])
    file["spectral_reference"].create_dataset("data", data=avantes_data)
    daq.calibration_lamp(0)
    avantes_dark = np.array([hs_camera.one_frame().T for _ in range(CALIBRATION_FRAMES)])
    file["spectral_reference"].create_dataset("dark", data=avantes_dark)
```

Listing S1: HS microscope measurement script excerpt. The code connects the measurement hardware, sets the parameters, and records the measurement and calibration data to an HDF file while automatically handling the reconfiguration needed for the capture of calibration data. For the full code, see [151].

```
{
    "description": {
        "description": "Microwell plate spectrum measurement",
        "sample": "Marine cyanobacteria mixtures",
        "acquisition_start": "2025-04-08T10:30:28+02:00",
        "acquisition_end": "2025-04-08T10:34:01+02:00"
    },
    "camera_hs": {
        "model": "ZWO ASI294MM Pro",
        "shutter": [100000, "us"],
        "gain": [108, "0.1dB"],
        "bit_depth": "12bit",
        "image_size": [5644, 8288],
        "binning": 1,
        "cooler_on": true,
        "temperature_target": [0, "degC"],
        "temperature_current": [0.4, "degC"]
    },
    "stage": {
        "axes": {
            "X": 1,
            "Y": 2
        },
        "position_start": {
            "X": [2990, 0],
            "Y": [21260, 0]
        },
        "position_end": {
            "X": [0, 0],
            "Y": [0, 0]
        },
        "step_size": [2.5, "um"],
        "microstep_mode": 256
    },
    "lens": {
        "model": "Nikon Plan Fluor",
        "magnification": 4,
        "numerical_aperture": 0.13
    },
    "enviro": {
        "temperature": {
            "value": 22,
            "unit": "C"
        },
        "humidity": {
            "value": 40,
            "unit": "%"
        },
        "pressure": {
            "value": 993,
            "unit": "hPa"
        },
        "timestamp": "2025-04-08T10:34:01+02:00"
    }
}
```


Listing S2: HS Microscope generated metadata excerpt.

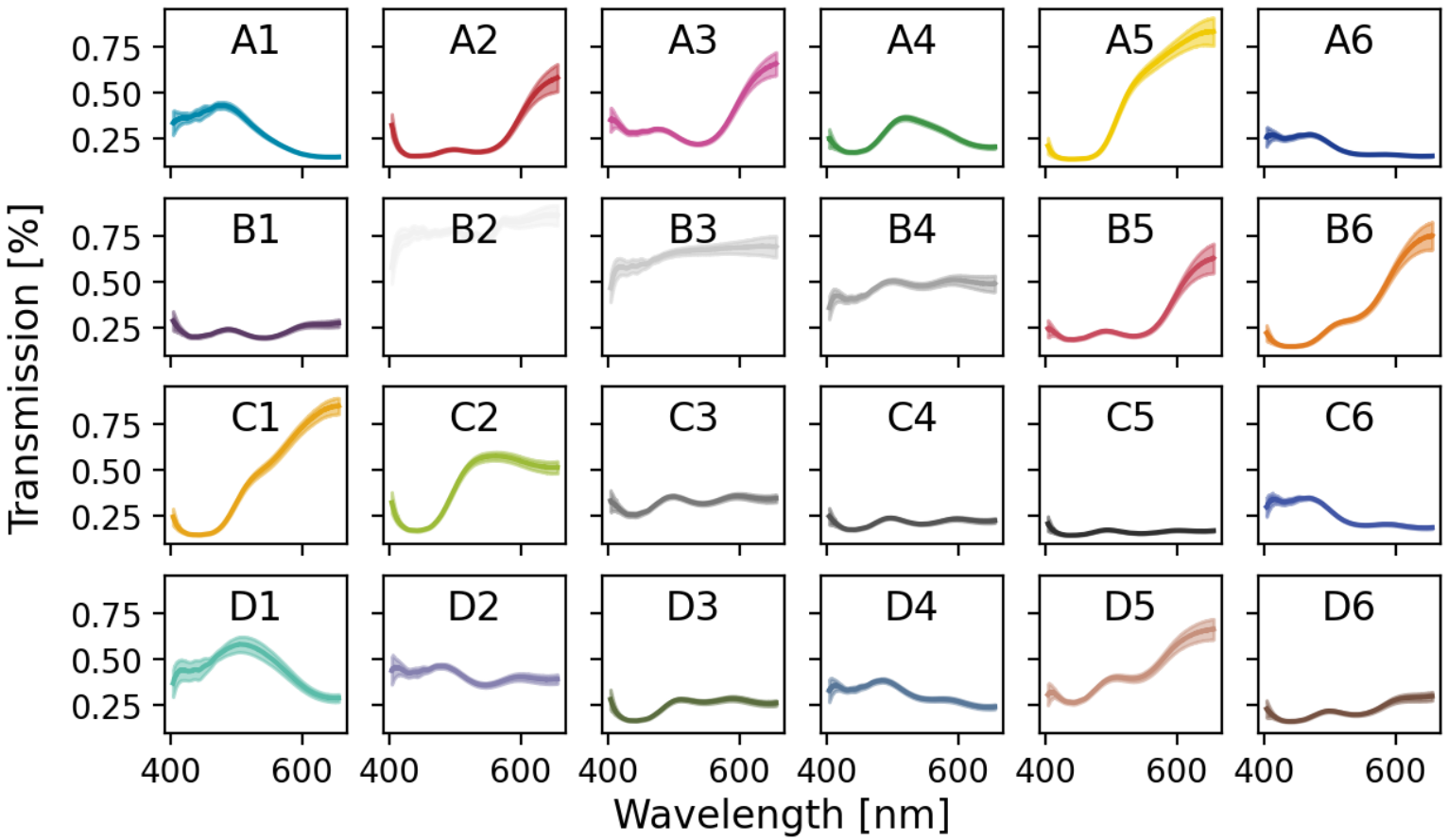


Figure S1: Average transmission spectra of individual colour patches of the NIST-IAM-9C-00348 colour calibration slide from the HS data, the shaded area represents standard deviation of the average spectra.

```
class Experiment:
def __init__(self, n_components: int = 2, reducer: str = None, unmixing: str = None):
    self.timestamp = datetime.now().strftime("%Y-%m-%d %H:%M:%S")
    self.n_components = n_components
    self.reducer_name = reducer
    self.unmixing_name = unmixing

def load_data(self, train, test, endmembers_index):
    self.train = train
    self.test = test
    self.endmembers_index = endmembers_index

def set_reducer(self):
    if self.reducer_name.upper() == "PCA":
        self.reducer = PCA(n_components=self.n_components)
        self.reducer.is_fitted_ = False
    if self.reducer_name.upper() == "NMF":
        self.reducer = NMF(n_components=self.n_components)
        self.reducer.is_fitted_ = False
    if self.reducer_name.upper() == "AE":
        self._load_autoencoder(f"autoencoder-fc-sad-{self.n_components}d.keras")
    if self.reducer_name.upper() == "AE-CNN":
        self._load_autoencoder(f"autoencoder-cnn-sad-{self.n_components}d.keras")

def _load_autoencoder(self, name_filter):
    try:
        ae_model = next(Path(MODELS_DIR/ "autoencoder").rglob(name_filter))
    except StopIteration:
        raise FileNotFoundError(f"Autoencoder model {name_filter} not found.")
    self.reducer = AutoencoderEstimator(ae_model)

def reduce(self):
    # if reducer is autoencoder
    if self.reducer.is_fitted_:
        self.train_reduced = self.reducer.transform(self.train)
        self.test_reduced = self.reducer.transform(self.test)
    else:
        self.train_reduced = self.reducer.fit_transform(self.train)
        self.test_reduced = self.reducer.transform(self.test)

def unmix(self):
    endmembers_reduced = self.train_reduced[self.endmembers_index]
    if self.unmixing_name.upper() == "NNLS":
        try:
            amap = NNLS(self.test_reduced, endmembers_reduced, maxiter=100000)
        except RuntimeError:
            print("NNLS failed to converge.")
            self.unmixing_name = "NNLS-failed"
            amap = -100 * np.ones((self.test_reduced.shape[0], len(endmembers_reduced)))
        self.test_reconstructed = np.dot(amap, endmembers_reduced)
    if self.unmixing_name.upper() == "MCR":
        mcrar = McrAR(tol_increase=10, max_iter=100)
        mcrar.fit(self.test_reduced, ST=endmembers_reduced)
        self.test_reconstructed = mcrar.D_opt_
    self.test_predicted = self.reducer.inverse_transform(self.test_reconstructed)
```

Listing S3: Python `Experiment` class excerpt, which allows to specify and combine various dimensionality reduction techniques (PCA, NMF, AE-FC, AE-CNN) with unmixing algorithms (NNLS, MCR). The class handles data loading, applying the chosen reduction and unmixing methods, and performance evaluation. For the full code, see [121].

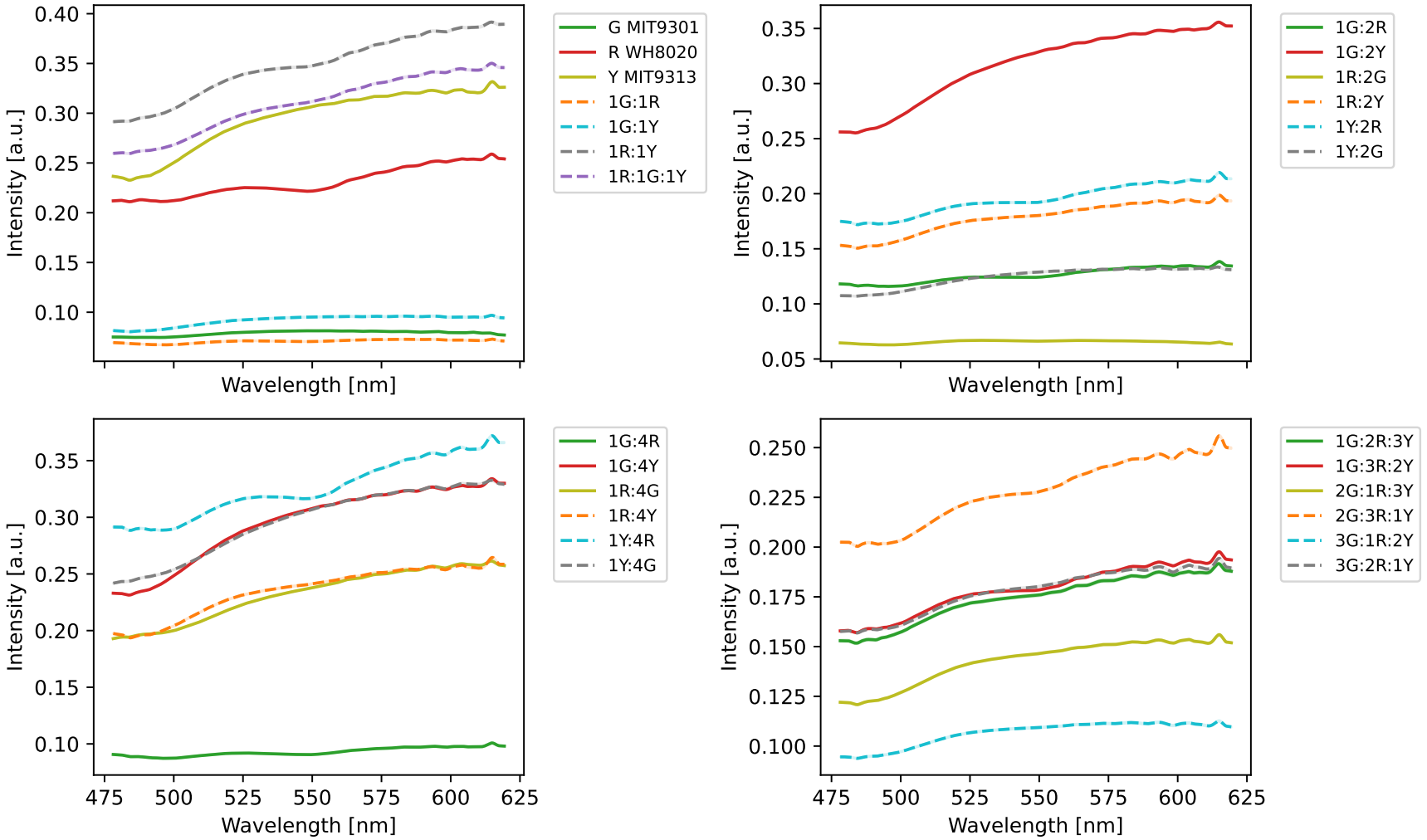


Figure S2: Average spectra of each of the marine cyanobacteria samples, measured using the HS microscope. Data were calibrated and smoothed using rolling average with a window size of 15 pixels.

Table S5: Mixing ratios and the predicted concentrations for all samples of marine cyanobacteria mixtures. The data were compressed using NMF to 3 components, before the MCR unmixing. The accuracy is calculated using the individual predictions performed on all spectral vectors from single well measurement.

Sample	**Sample Ratio**			**Predicted Ratio**			
	G	R	Y	G	R	Y	Accuracy
1G:1R	0.50	0.50	0.00	0.8229 ± 0.0018	0.1333 ± 0.0015	0.0437 ± 0.0004	0.7438 ± 0.0014
1G:1Y	0.50	0.00	0.50	0.794 ± 0.002	0.0633 ± 0.0007	0.1425 ± 0.0014	0.7363 ± 0.0015
1R:1Y	0.00	0.50	0.50	0.1981 ± 0.0007	0.3658 ± 0.0015	0.4361 ± 0.0012	0.6851 ± 0.0007
1G:2R	0.33	0.67	0.00	0.668 ± 0.002	0.2650 ± 0.0016	0.0671 ± 0.0005	0.7736 ± 0.0015
1G:2Y	0.33	0.00	0.67	0.2188 ± 0.0011	0.1255 ± 0.0012	0.6557 ± 0.0013	0.7458 ± 0.0019
1R:2G	0.67	0.33	0.00	0.7915 ± 0.0017	0.1269 ± 0.0013	0.0815 ± 0.0006	0.6441 ± 0.0015
1R:2Y	0.00	0.33	0.67	0.337 ± 0.003	0.2901 ± 0.0013	0.3734 ± 0.0013	0.7099 ± 0.0015
1Y:2R	0.00	0.67	0.33	0.2518 ± 0.0019	0.5641 ± 0.0015	0.1842 ± 0.0009	0.7240 ± 0.0017
1Y:2G	0.67	0.00	0.33	0.6853 ± 0.0019	0.0860 ± 0.0008	0.2287 ± 0.0013	0.8062 ± 0.0016
1G:4R	0.20	0.80	0.00	0.7703 ± 0.0018	0.1834 ± 0.0015	0.0463 ± 0.0004	0.8018 ± 0.0015
1G:4Y	0.20	0.00	0.80	0.2730 ± 0.0008	0.0757 ± 0.0010	0.6513 ± 0.0010	0.8525 ± 0.0015
1R:4G	0.80	0.20	0.00	0.3996 ± 0.0009	0.2145 ± 0.0007	0.3859 ± 0.0010	0.6611 ± 0.0004
1R:4Y	0.00	0.20	0.80	0.2585 ± 0.0015	0.1900 ± 0.0009	0.5515 ± 0.0013	0.6827 ± 0.0015
1Y:4R	0.00	0.80	0.20	0.1564 ± 0.0006	0.6037 ± 0.0011	0.2400 ± 0.0011	0.6574 ± 0.0011
1Y:4G	0.80	0.00	0.20	0.3483 ± 0.0010	0.1844 ± 0.0011	0.4673 ± 0.0017	0.7365 ± 0.0014
1R:1G:1Y	0.33	0.33	0.33	0.1730 ± 0.0007	0.2975 ± 0.0009	0.5295 ± 0.0009	0.9703 ± 0.0008
1G:2R:3Y	0.17	0.33	0.50	0.5562 ± 0.0017	0.2407 ± 0.0012	0.2031 ± 0.0010	0.9434 ± 0.0010
1G:3R:2Y	0.17	0.50	0.33	0.4768 ± 0.0016	0.3428 ± 0.0010	0.1804 ± 0.0010	0.9235 ± 0.0011
2G:1R:3Y	0.33	0.17	0.50	0.573 ± 0.002	0.1127 ± 0.0009	0.3148 ± 0.0015	0.8339 ± 0.0017
2G:3R:1Y	0.33	0.50	0.17	0.3139 ± 0.0014	0.4533 ± 0.0011	0.2328 ± 0.0009	0.9524 ± 0.0009
3G:1R:2Y	0.50	0.17	0.33	0.6588 ± 0.0019	0.0633 ± 0.0003	0.2778 ± 0.0017	0.7951 ± 0.0019
3G:2R:1Y	0.50	0.33	0.17	0.5909 ± 0.0019	0.2272 ± 0.0011	0.1820 ± 0.0011	0.8884 ± 0.0014

List of Publications and Results

Mgr. Ondřej Vaculík

June 2025

Publications

VACULÍK, O.; BERNATOVÁ, S.; REBROŠOVÁ, K.; SAMEK, O.; ŠILHAN, L.; RŮŽIČKA, F.; ŠERÝ, M.; ŠILER, M.; JEŽEK, J.; ZEMÁNEK, P. Rapid Identification of Pathogens in Blood Serum Via Raman Tweezers in Combination with Advanced Processing Methods. *Biomedical Optics Express*. 2023, vol. 14, no. 12, p. 6410. ISSN 2156-7085, ISSN 2156-7085. Available from: doi:10.1364/BOE.503628

- IF 2.9, Q2. Principal author, algorithm design for processing and data evaluation, the resulting implementation achieves the accuracy of pathogen identification in blood serum over 90%.

ŠILHAN, L.; ARREGI, J. A.; PLICHTA, T.; **VACULÍK, O.**; NOVOTNÝ, J.; ŠERÝ, M. Compact Vacuum Setup for Laser Induced Plasma Etching with Optical Emission Spectrum Monitoring. *Journal of Vacuum Science & Technology B*. 2025, vol. 43, no. 3, p. 034202. ISSN 2166-2746. Available from: doi:10.1116/6.0004296

- IF 1.5, Q3. Data curation, processing of the optical emission spectrum of SF_6 gas and air. Picture from this article was used on the cover of the issue *Journal of Vacuum Science & Technology B* May/June 2025.

RICHTEROVÁ, V.; GJEVIK, A.; **VACULÍK, O.**; VEJROSTA, J.; PEKAŘ, M. Impact of Collagen on the Rheological and Transport Properties of Agarose Hydrogels. *Gels*. 2025, vol. 11, no. 6, p. 396. ISSN 2310-2861. Available from: doi:10.3390/gels11060396

- IF 5, Q1. Visualisation of data of concentration profiles of eosin B and methylene blue in agarose-collagen hydrogels.

Conferences

VACULÍK, O.; ŠERÝ, M.; ŠILHAN, L.; ŠILHANOVÁ, D.; ZEMÁNEK, P. Addressing Data Management for Custom Built Hyperspectral Microscopy Station. In: *2024 14th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. Helsinki, Finland: IEEE, 2024, pp. 1–5. ISBN 979-8-3315-1313-9. Available from: doi:10.1109/WHISPERS65427.2024.10876463

VACULÍK, O. *High-Resolution Hyperspectral Microscopy for Efficient Analysis of Biological Systems* [Poster]. Jyväskylä, 2024

- This poster was voted by the audience as the best poster at 33. Jyväskylä Summer School, held on 12 August 2024 at the University of Jyväskylä, Finland.

ŠILHAN, L.; NOVOTNÝ, J.; PLICHTA, T.; JEŽEK, J.; **VACULÍK, O.**; ŠERÝ, M. Design of Setup for Laser Induced Plasma Etching. In: *2024 37th International Vacuum Nanoelectronics Conference (IVNC)*. Brno, Czech Republic: IEEE, 2024, pp. 1–2. ISBN 979-8-3503-7976-1. Available from: doi:10.1109/IVNC63480.2024.10652276

VACULÍK, O.; BERNATOVÁ, S.; ŠERÝ, M.; ŠILER, M.; SAMEK, O.; REBROŠOVÁ, K.; RŮŽIČKA, F.; ZEMÁNEK, P. Advanced Algorithms for Bacterial Analysis in Body Fluids Using Raman Tweezers. In: *Biophotonics Congress: Optics in the Life Sciences 2023 (OMA, NTM, BODA, OMP, BRAIN)*. Vancouver, British Columbia: Optica Publishing Group, 2023, AW3D.2.

ISBN 978-1-957171-21-0. Available from: doi:10.1364/OMA.2023.AW3D.2

VACULÍK, O.; ŠILHAN, L.; ŠKRABALOVÁ, D.; ŠERÝ, M.; ZEMÁNEK, P. Hyperspectral Imaging for Life Sciences. In: Lednice: Ústav přístrojové techniky AV ČR, v. v. i., 2021, vol. 61. Available also from: https://laser.isibrno.cz/. LASER61

Projects

MEYS P JAC EH2_008/0004624 – *PHOTOMACHINES – Photosynthetic cell redesign for high yields of therapeutic peptides* (2024–2028)

- Development of Raman spectroscopy measurement station control software for a robot-controlled system for high-throughput phenotyping of phototrophs (WP 1.1).

MEYS P JAC EH22_008/0004649 – *Quantum Engineering and Nanotechnology* (2024–2028)

- Development of a UI for generating Fresnel lens writing structures for UV lithography.

TA CR FW10010248 – *Research and development of innovative LiDAR technology for increasing aviation safety* (2024–2026)

- Characterisation and verification of scientific camera parameters.

TA CR FW06010453 – *Innovative projector development for Safety and Industry 4.0* (2023–2025)

- Measurement and evaluation of lens parameters for projection of security features, development of projection system with dynamically generated image.

MIT CR FV40455 – *HyPerSpec – Development of a hyperspectral camera for biotechnology applications and element analysis* (2019–2022)

- Development of HypCam2 hyperspectral camera, verification of imaging spectroscope parameters, evaluation of calibration data, automation of measurements.

MIT CR EG19_262/0020294 – *Technology for advanced optics and its industrial application* (2020–2023)

- Parametric generation of Fresnel lens structures for two-photon lithography.

Technical Outputs

ŠERÝ, M.; ŠILHAN, L.; **VACULÍK, O.**; FLEK, O.; HRABAL, J.; SNÍŽEK, J. *High Spatial Resolution Projection System with Fixed Beam Angle.* Prototype, 2024. RIV/68081731:_____/24:00603702

ŠERÝ, M.; ŠILHAN, L.; **VACULÍK, O.**; FLEK, O.; HRABAL, J.; SNÍŽEK, J. *High Spatial Resolution Projection System with Dynamically Variable Beam Angle.* Prototype, 2024. RIV/68081731:_____/24:00605258

ŠERÝ, M.; ŠILHAN, L.; **VACULÍK, O.**; FLEK, O.; HRABAL, J.; SNÍŽEK, J. *High Spatial Resolution Projection System with Dynamically Generated Image.* Functional sample, 2024. RIV/68081731:_____/24:00605262

ČERNÝ, Š.; CHLUMSKÁ, J.; KANDRA, M.; MATĚJKA, M.; STŘÍTESKÝ, S.; TĚTHAL, T.; JÁKL, P.; PLICHTA, T.; ŠERÝ, M.; ŠILHAN, L.; ŠKRABALOVÁ, D.; **VACULÍK, O.** *Prototype device for laser beam writing using 2PP.* Prototype, 2022, RIV/29137756:_____/23:N0000003

ŠERÝ, M.; ŠILHAN, L.; **VACULÍK, O.**; MAŇKA, T.; RATAJ, T.; TRTÍLEK, M. *High resolution hyperspectral camera with a wavelength range of 350-1100 nm* HypCam2. Functional sample, 2023, RIV/60646594:_____/22:N0000006