



6<sup>th</sup> INTERNATIONAL CONFERENCE ON

# Optical Characterization of Materials

MARCH 22<sup>nd</sup>–23<sup>rd</sup>, 2023  
KARLSRUHE | GERMANY

J. BEYERER | T. LÄNGLE | M. HEIZMANN (Eds.)

 **KIT** Scientific  
Publishing



Jürgen Beyerer | Thomas Längle | Michael Heizmann (eds.)

OCM 2023

6<sup>th</sup> International Conference on  
Optical Characterization of Materials

March 22<sup>nd</sup>–23<sup>rd</sup>, 2023  
Karlsruhe | Germany



OCM 2023

6<sup>th</sup> International Conference on  
Optical Characterization of Materials

March 22<sup>nd</sup> – 23<sup>rd</sup>, 2023

Karlsruhe | Germany

edited by

Jürgen Beyerer | Thomas Längle | Michael Heizmann

## Organizer

Fraunhofer Institute of Optronics,  
System Technologies and Image Exploitation IOSB  
c/o Karlsruhe Center for Material Signatures KCM  
Fraunhoferstraße 1, 76131 Karlsruhe

The proceedings are also available as an online version  
<http://dx.doi.org/10.5445/KSP/1000155014>

## Impressum



Karlsruher Institut für Technologie (KIT)  
KIT Scientific Publishing  
Straße am Forum 2  
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark  
of Karlsruhe Institute of Technology.  
Reprint using the book cover is not allowed.

[www.ksp.kit.edu](http://www.ksp.kit.edu)



*This document – excluding parts marked otherwise the cover, pictures and graphs –  
is licensed under a Creative Commons Attribution 4.0 International License  
(CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/deed.en>*



*The cover page is licensed under a Creative Commons  
Attribution-NonCommercial 4.0 International License (CC BY-ND 4.0):  
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2023 – Printed on FSC-certified paper

ISSN 2510-7240

ISBN 978-3-7315-1274-5

DOI 10.5445/KSP1000155014







## Preface

The state of the art in optical characterization of materials is advancing rapidly. New insights into the theoretical foundations of this research field have been gained and exciting practical developments have taken place, both driven by novel applications and innovative sensor technologies that are constantly emerging. The big success of the international conferences on Optical Characterization of Materials in 2013, 2015, 2017, 2019 and 2021 proves the necessity of a platform to present, discuss and evaluate the latest research results in this interdisciplinary domain. Due to that fact, the international conference on Optical Characterization of Materials (OCM) took place the sixth time in March 2023.

The OCM 2023 was organized by the Karlsruhe Center for Spectral Signatures of Materials (KCM) in cooperation with the German Chapter of the Instrumentation & Measurement Society of IEEE. The Karlsruhe Center for Spectral Signatures of Materials is an association of institutes of Karlsruhe Institute of Technology (KIT) and the business unit Inspection and Optronic Systems of the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB.

Despite the conference's young age, the organizing committee has had the pleasure to evaluate a large amount of abstracts. Based on the submissions, we selected 19 papers as posters and talks, a plenary lecture, a panel discussion and several practical demonstrations.

The present book is based on the conference held in Karlsruhe, Germany from March 22–23, 2023. The aim of this conference was to bring together leading researchers in the domain of Characterization of Materials by spectral characteristics from UV (240 nm) to IR (14  $\mu\text{m}$ ), multispectral image analysis, X-ray methods, polarimetry, and microscopy. Typical application areas for these techniques cover the fields of, e.g., food industry, recycling of waste materials, detection of contaminated materials, mining, process industry, and raw materials.

The editors would like to thank all of the authors that have contributed to these proceedings as well as the reviewers, who have in-

## Preface

vested a generous amount of their time to suggest possible improvements of the papers. The help of Lukas Roming and Jürgen Hock in the preparation of this book is greatly appreciated. Last but not least, we thank the organizing committee of the conference, led by Britta Ost, for their effort in organizing this event. The excellent technical facilities and the friendly staff of the Fraunhofer IOSB greatly contributed to the success of the meeting.

March 2023

Jürgen Beyerer  
Thomas Längle  
Michael Heizmann

**General Chairs**

|                  |           |
|------------------|-----------|
| Jürgen Beyerer   | Karlsruhe |
| Thomas Längle    | Karlsruhe |
| Michael Heizmann | Karlsruhe |

**Program Chair**

|                 |           |
|-----------------|-----------|
| Henning Schulte | Karlsruhe |
|-----------------|-----------|

**Program Committee**

|                       |              |
|-----------------------|--------------|
| Jochen Aderhold       | Braunschweig |
| Oliver Albrecht       | Dresden      |
| Johannes Anastasiadis | Karlsruhe    |
| Sebastian Bauer       | Bosten USA   |
| Andrea Büttner        | Freising     |
| Robin Gruna           | Karlsruhe    |
| Tino Hausotte         | Erlangen     |
| Andreas Herzog        | Magdeburg    |
| Thomas Hofmann        | Würzburg     |
| Olfa Kanoun           | Chemnitz     |
| Anna Kicherer         | Sieboldingen |
| Félix Salazar         | Madrid       |
| Maximilian Schambach  | Karlsruhe    |
| Heinar Schmidt        | Kulmbach     |
| Matthäus Speck        | Waldheim     |
| Hermann Wotruba       | Aachen       |
| Bernhard Zagar        | Linz         |



# Contents

Preface ..... i

## Recycling

Monitoring the sorting performance in lightweight packaging  
waste sorting plants using data of sensor-based sorters ..... 1

*S. Schlögl, G. Schmölzer, A. Weber, A. Anditsch, and A.  
Aldrian-Tischberger*

Detecting Tar Contaminated Samples in Road-rubble using  
Hyperspectral Imaging and Texture Analysis ..... 11

*P. Bäcker, G. Maier, R. Gruna, T. Längle, and J. Beyerer*

Increasing the reuse of wood in bulky waste using artificial  
intelligence and imaging in the VIS, IR, and terahertz ranges ..... 23

*L. Roming, R. Gruna, J. Aderhold, F. Schlüter, D. Cibiraite-  
Lukenskiene, D. Gundacker, F. Friederich, M. Bihler, and M.  
Heizmann*

Semi-supervised methods for CNN based classification of  
multispectral imagery ..... 37

*M. Bihler, J. Zhou, and M. Heizmann*

Regression-based Age Prediction of Plastic Waste using  
Hyperspectral Imaging ..... 51

*F. Kronenwett, P. Klingenberg, G. Maier, T. Längle, and J. Beyerer*

## Food Inspection

|   |    |
|---|----|
| Method Development for Spatially Resolved Detection of Adulterated Minced Meat . . . . .  | 65 |
| <i>E. Djaw, I. Türkmen, T. Tybussek, and T. Sauerwald</i>   |    |
| Multispectral analysis for the determination of lycopene concentration in tomatoes . . . . .  | 75 |
| <i>M. Mlynarik, G. A. Atkinson, M. L. Smith, and K. Emrith</i>  |    |
| An overview of implementing Multispectral Imaging coupled with machine learning for the assessment of microbiological quality and authenticity in foods . . . . . | 87 |
| <i>A. Lytou, L.-C. Fengou, N. Schultz, J. M. Carstensen, Y. Zhang, F. Mohareb, and G.-J. Nychas</i>   |    |
| Self-supervised Pretraining for Hyperspectral Classification of Fruit Ripeness . . . . .  | 97 |
| <i>L. A. Varga, H. Frank, and A. Zell</i>   |    |

**Sensors**

Thermographic Techniques to Explore Small-Scale Processes at Water Surfaces. . . . . 109  
*B. Jähne, L. Warmuth, R. Stewing, and K. Krall*

Sensitivity enhanced glucose sensing by return-path Mueller matrix ellipsometry . . . . . 119  
*C.-W. Chen, M. Hartrumpf, T. Längle, and J. Beyerer*

Fluorescence Imaging of Concentration Fields of Dissolved Gases at Water Interfaces . . . . . 129  
*D. Hofmann and B. Jähne*

Material Characterization using a Compact Computed Tomography Imaging Spectrometer with Super-resolution Capability . . . . . 139  
*S. Amann, M. Mel, P. Zanuttigh, T. Haist, M. Kamm, and A. Gatto*

Sulfur Dioxide Fluorescence Imaging . . . . . 149  
*B. Jähne, R. Beronova, and K. Krall*

Imaging radar systems for non-destructive material testing  
 An overview of the state of the art, the limitations and the opportunities of radar technology . . . . . 159  
*D. Nüßler, S. Leuchs, and C. Krebs*

**Algorithms**

Quick-and-Dirty Computation of Voigt Profiles, Classification of Their Shapes, and Effective Determination of the Shape Parameter. . . . . 169  
*A. Kehrein and O. Lischtschenko*





# Monitoring the sorting performance in lightweight packaging waste sorting plants using data of sensor-based sorters

Sabine Schlögl, Georg Schmölzer, Alexander Weber, Alexander Anditsch, and Alexia Aldrian-Tischberger

Montanuniversität Leoben, AVAW,  
Franz-Josef-Straße 18, 8700 Leoben

**Abstract** To achieve the necessary improvements in lightweight packaging waste sorting plants to increase the recycling rate, sensor-based material flow monitoring and plant control is the subject of current research and development. This study investigates whether data from existing sensor-based sorters could be used for this purpose. The results show that data recorded during sorting correlate strongly with ideal analysis data. Furthermore, a correlation between the data of the first sorter and the output fractions of later sorting stages could be established. The results therefore show a great potential for the use of sensor-based sorting data.

**Keywords** Monitoring, NIR, SBS, sensor-based sorting data, pixel-/object-based monitoring, lightweight packaging waste

## 1 Introduction

In 2019, 79.6 Mio. t [1] of packaging waste were created within the European Union (EU), marking the highest value recorded. To reduce the negative impact of packaging waste in general and plastic packaging in particular, a variety of new waste legislation measurements was presented throughout the last few years. One of them being the recycling rate for plastic packaging waste of 50% by 2025 [2]. This results in new requirements for lightweight packaging waste sorting plants to enable the aspired circular economy.

Many conventional sorting plants are currently operated as black boxes. Besides the manual analysis of input and output compositions, little process data is gathered and stored to enable plant control. However, the collection of such data is essential to find key aspects for optimization of both existing and newly built sorting plants. The research project “EsKorte” investigates not only the implementation of additional sensors for material flow monitoring but also the exploitability of existing, but not yet used, sensor-based sorting (SBS) data for material flow monitoring and control. Two research questions have been addressed with the presented analysis of SBS-data gathered during multi-level sorting of plastic packaging waste material using an experimental setup with a near-infrared sensor:

- (1) Is SBS-data suitable for monitoring key sorting parameters?
- (2) Is SBS-data suitable for predicting the sorting results of successive sorting steps?

## **2 Materials and Methods**

### **2.1 Materials**

The sample material was collected in a plastic packaging waste sorting plant in Austria. The samples taken in the output fractions were beverage cartons (BC), polyethylene terephthalate (PET) bottles, as well as containers made from polyethylene (PE) and polypropylene (PP). The samples included different brands, filling quantities and contents to represent the variety of plastic packaging waste. To ensure the best possible detection and sorting during the trials, the samples were manually cut into 3x3 cm pieces. This is due to the experimental setup requiring a reduced grain size. Caps and strongly curved particles were excluded from the sample material to ensure uniform particle properties. Three mixtures were created with the sample material (see Table 1). M1 represents an evenly distributed material, M2 a higher share of transparent PET-material and M3 a dominant polyolefin content. The corresponding pixel (px) and object (obj) shares differ due to the different area densities.

## Monitoring the sorting performance using SBS data

**Table 1:** Composition of sample mixtures (M1-M3) based on weighing (top) and corresponding average classified sensor data (bottom).

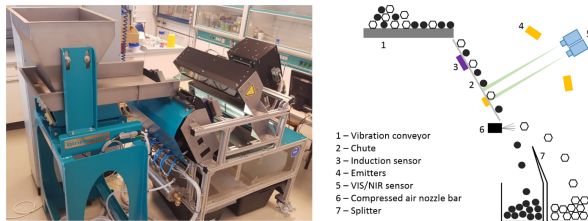
|     | M1    |  |     |  | M2     |  |     |  | M3     |  |     |  |
|-----|-------|--|-----|--|--------|--|-----|--|--------|--|-----|--|
|     | kg    |  | wt% |  | kg     |  | wt% |  | kg     |  | wt% |  |
| BC  | 0.507 |  | 25  |  | 0.4046 |  | 20  |  | 0.1925 |  | 10  |  |
| PET | 0.507 |  | 25  |  | 0.8092 |  | 40  |  | 0.1925 |  | 10  |  |
| PE  | 0.507 |  | 25  |  | 0.4046 |  | 20  |  | 0.7700 |  | 40  |  |
| PP  | 0.507 |  | 25  |  | 0.4046 |  | 20  |  | 0.7700 |  | 40  |  |

|     | px      |         | px%  |      | obj     |        | obj% |      | px     |        | px%  |     | obj |  | obj% |  |
|-----|---------|---------|------|------|---------|--------|------|------|--------|--------|------|-----|-----|--|------|--|
|     | BC      | 1073070 | 37   | 1620 | 33      | 838767 | 27   | 1232 | 23     | 403373 | 20   | 617 | 18  |  |      |  |
| PET | 1030385 | 35      | 1971 | 40   | 1644041 | 52     | 3072 | 57   | 395155 | 19     | 765  | 22  |     |  |      |  |
| PE  | 415414  | 14      | 674  | 14   | 336186  | 11     | 542  | 10   | 619754 | 30     | 1011 | 30  |     |  |      |  |
| PP  | 405092  | 14      | 654  | 13   | 336577  | 11     | 545  | 10   | 626035 | 31     | 1030 | 30  |     |  |      |  |

### 2.2 Experimental setup

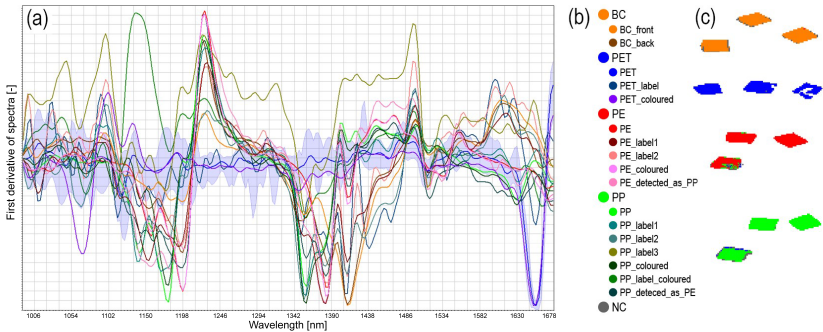
The multilevel sorting was conducted with a chute sorter (working width 500 mm, length: 455 mm) using an NIR-sensor (Model: EVK Helios-G2-NIR1 [3]). The experimental setup, including the vibration conveyor for material separation, is presented in Figure 1.



**Figure 1:** Experimental setup and associated schematic layout [4].

The detected pixels are 1.60 mm wide and have a length smaller than 1.60 mm (depending on the sliding speed). For the classification a teach-in was created in “SQALAR” [5]. To achieve the required clas-

sification close to 100% in each particle not only the pure materials, but also the mixed spectra resulting from labels on the objects were included. The settings for the differentiation of background and material (Spectrum Mean Intensity  $\leq 340$ ) were determined in an iterative process. In preliminary tests the light settings were evaluated. Lower background light caused better object localization for PET, while higher emitter light caused stronger excitation in the NIR range. The recommended default settings were altered accordingly. The reference spectra, as well as the resulting classified false color images can be seen in Figure 2.



**Figure 2:** Reference spectra for classification (a) First derivative of reference spectra (b) Created material classes with assigned spectra (c) False colour images (orange: BC, blue: PET, red: PE, green: PP, grey: Not Classified [NC]).

## 2.3 Data acquisition

Each pixel is classified based on the chosen reference spectra in the software. During the trials this classification is visualized in a livestream of false colour images on a screen. Real-time data recording is achieved by using Matlab [6] to continuously scan and analyze the false-color images on the screen. The resulting values include the total number of counted pixels per material as well as the corresponding number of objects. An object is defined as an area bigger than 70 pixels of the same colour. Objects smaller than 70 pixels are typically fault detections and therefore ignored. Further the trial time and input mass for

each sorting step is documented to calculate the throughput.

## 2.4 Experimental procedure

Each test run consists of four sorting levels (BC, PET, PE, PP), while every level includes both rougher and cleaner (see Figure 3). In a rougher, all target particles are to be sorted out, whereby the purity is low. In the cleaner, this fraction is purified by removing impurities. All input and output fractions were analysed at lower throughput to avoid overlap (Average values: rougher: 9 kg/h, cleaner: 8 kg/h, analysis: 2 kg/h). For each mixture (M1–M3) five repetitions of test runs were performed.

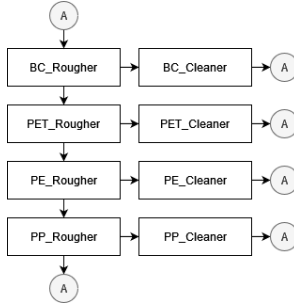


Figure 3: Flowchart of multilevel sorting; A: Analysis.

## 2.5 Data analysis

The data from all test runs were analysed with respect to the following parameters.  $x$  represents the number of pixels or objects.  $Yield_{Input}$  is the result in respect to the input composition, while  $Yield_{Level}$  refers to the input of the respective sorting stage.

$$(1) \text{ Coefficient of variation} = \frac{\text{standard deviation}}{\text{mean}}$$

$$(2) \text{ Yield}_{Input} = \frac{x_{i,Eject}}{x_{i,Input}}$$

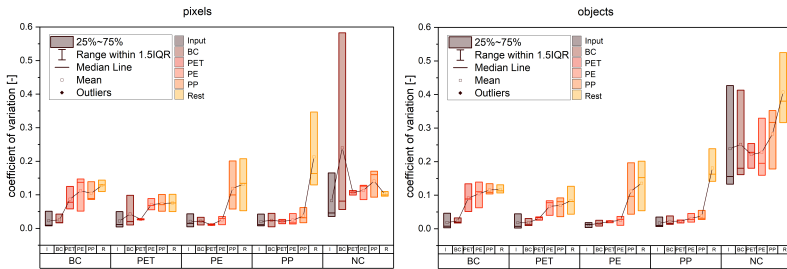
$$(3) \text{ Yield}_{Level} = \frac{x_{i,Eject}}{x_{i,Level}}$$

$$(4) \text{ Purity} = \frac{x_{i,Eject}}{x_{Eject}}$$

### 3 Results

#### 3.1 Reproducibility

Both pixel-based and object-based data was analysed to evaluate the reproducibility of data gathered from sensor-based sorters. The results show low values for the coefficient of variation (CV):  $CV_{Pixel} = 0.07$ ,  $CV_{Object} = 0.1$ . The CV values increased with each sorting level, indicating a slightly better usability of sensor data from early sorting steps (see Figure 4). The higher values for NC are noteworthy, though these are also in most cases below the critical limit of  $CV = 0.5$ . In general, the type of material class influences the CV values more than the input mixture (see Figure 5).



**Figure 4:** Coefficient of variation throughout the sorting levels. Pixel-based data (left) and object-based data (right); I: Input, R: Rest.

#### 3.2 Exploitability of sensor-based sorting data

To assess whether the SBS data of  $BC_{Rougher}$  is suitable for monitoring, a comparison was made with the input analysis data generated at optimal singulation (“ground truth”). In Figure 6 it can be seen, that the pixel data represents the ground truth slightly better than the object data. Nevertheless, the object data also shows a linear correlation and is similar to the input composition at small values.

## Monitoring the sorting performance using SBS data

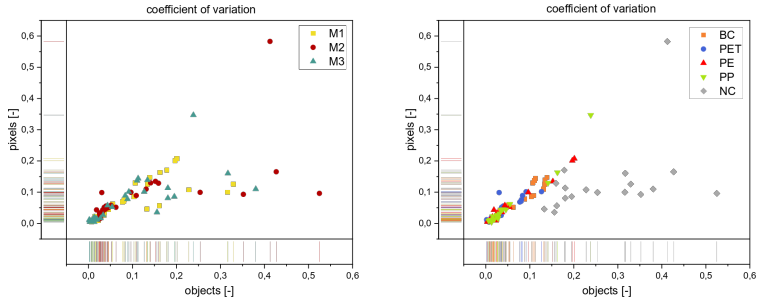


Figure 5: Influence of mixtures and materials on coefficient of variation.

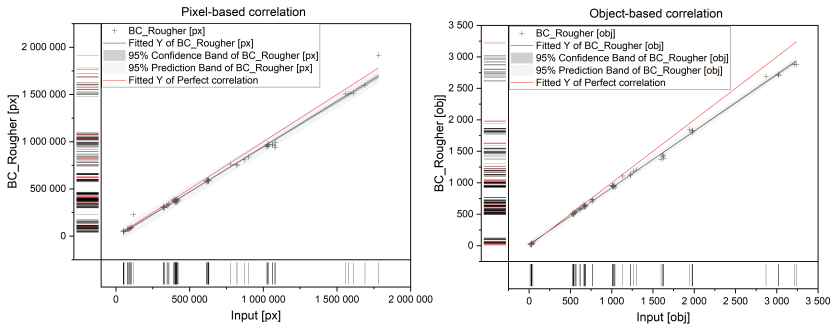
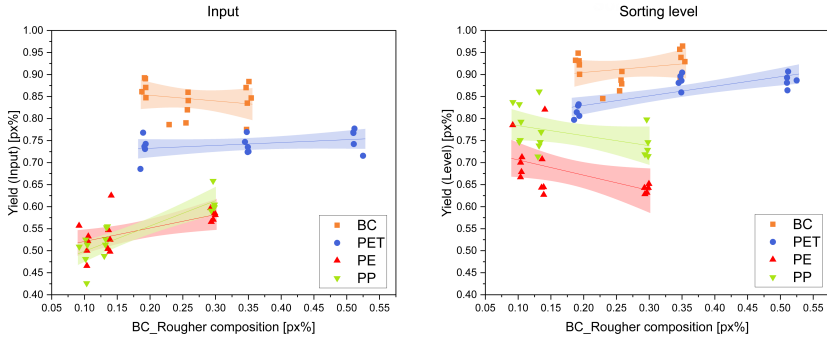


Figure 6: Comparison of Input analysis and data from SBS in  $BC_{Rougher}$ ; Pixel-based (left) and object-based (right).

### 3.3 Monitoring of Yield

To determine whether the SBS data is suitable for monitoring, the yield was assessed in relation to the input as well as in relation to the respective sorting stage (see Figure 7). There is no continuous correlation between input composition and yield but clusters depending on the sorting level were discovered. The best values are for BC, followed by PET. For  $Yield_{Input}$ , the values for PE and PP are usually around 45 – 60 px%, from which it could be deduced that the input-related yield drops sharply from the third sorting stage onwards, regardless of material. In contrast, the sorting level-related yield (Figure 7: right) shows

a clearer distinction between PE and PP. The low values of PE result from a poorer discharge behaviour, which could be observed during the tests. In general, at least a rough prediction of yield based on SBS data generated in the first sorting step appears to be possible.



**Figure 7:** Yield depending on  $BC_{Rougher}$  composition. Pixel-based values in relation to input (left) and respective sorting stages (right).

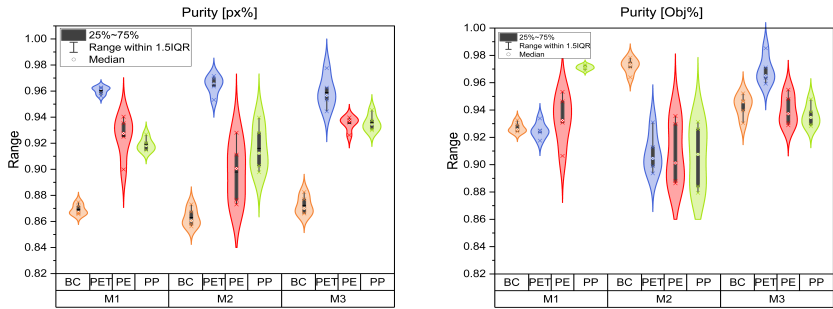
### 3.4 Monitoring of Purity

Since the purity of output fractions is a relevant criterion for recyclability, its monitoring with SBS data was further investigated. Figure 8 visualises that the composition of mixture (M1-M3) is more important than the sorting level, since there is no gradient along the sorting levels within a mixture. Lower limits and averages are higher for object-based values, which might be because pixel-based purity is degraded by misclassifications at the edges of particles.

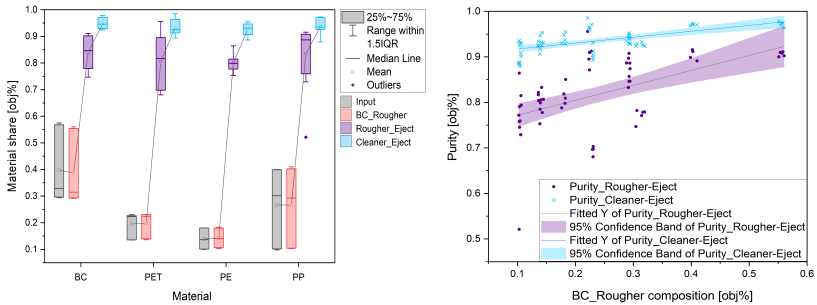
The proportion of the target fraction increases with the purification steps (see Figure 9), which is plausible since it reflects the behaviour of sorting plants. The values of the input analysis (black) and the values of  $BC_{Rougher}$  (red) are very similar, while in the eject of the rougher (purple) the purity increases strongly. The purity of the output fractions, i.e. the cleaner eject (blue), is the highest and usually has the smallest range. The correlation with  $BC_{Rougher}$  data for all output fractions has a maximum range of 10.6 percentage points. This includes results for the fourth sorting level.



## Monitoring the sorting performance using SBS data



**Figure 8:** Dependence of purity on mixtures (M1-M3) and sorting levels (BC, PET, PE, PP); left: pixel-based, right: object-based.



**Figure 9:** Increasing material shares [obj%] with increasing sorting level (left) and dependence of purity [obj%] in output fractions on  $BC_{Rougher}$  composition (right).

## 4 Conclusion

The data presented demonstrates that SBS data has high potential for material flow monitoring. The data shows a low variation with repetition and a strong correlation between the results of the optimally singulated analysis and the data recorded during sorting. Based on the data of the first sorting stage (BC), a clear distinction of the yields of the different sorting stages is possible. Furthermore, there is a clear correlation between the  $BC_{Rougher}$  data and the resulting purity of the output fractions. Based on these results, further investigations can be made to not only monitor but predict the sorting performance.

## Acknowledgements

The authors disclose receipt of the following financial support: The project “EsKorte” was funded by the Austrian Research Promotion Agency within the program “Production of the Future” under grant agreement 877341.

## References

1. Eurostat. (2023) Packaging waste by waste management operations. [Online]. Available: [https://ec.europa.eu/eurostat/databrowser/view/ENV\\_WASPAC\\_custom\\_4523594/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/ENV_WASPAC_custom_4523594/default/table?lang=en)
2. Directive (eu) 2018/852 of the european parliament and of the council of 30 may 2018 amending directive 94/62/ec on packaging and packaging waste (text with eea relevance). [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32018L0852>
3. EVK. (2023) Hyperspectral cameras for sorting and monitoring applications. [Online]. Available: <https://www.evk.biz/en/products/hyperspectral-camera/>
4. EVK. (2023) Evk sqalar - software tool for the qualitative and quantitative analysis. [Online]. Available: <https://www.evk.biz/en/products/analysis-software-tool/evk-sqalar/>
5. B. Kueppers, S. Schloegl, G. Oreski, R. Pomberger, and D. Vollprecht, “Influence of surface roughness and surface moisture of plastics on sensor-based sorting in the near infrared range. waste management and research.” *Waste Management and Research*, 2019.
6. Matlab. (2023) Mathworks. [Online]. Available: <https://de.mathworks.com/products/matlab.html>

# Detecting Tar Contaminated Samples in Road-rubble using Hyperspectral Imaging and Texture Analysis

Paul Bäcker<sup>1</sup>, Georg Maier<sup>1</sup>, Robin Gruna<sup>1</sup>, Thomas Längle<sup>1</sup>, and Jürgen Beyerer<sup>1,2</sup>

<sup>1</sup> Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstraße 1, 76131 Karlsruhe, Germany

<sup>2</sup> Vision and Fusion Laboratory (IES), Karlsruhe Institute of Technology (KIT), Haid-und-Neu-Str. 7, 76131 Karlsruhe, Germany

**Abstract** Polycyclic aromatic hydrocarbons (PAH) containing tar-mixtures pose a challenge for recycling road rubble, as the tar containing elements have to be extracted and decontaminated for recycling. In this preliminary study, tar, bitumen and minerals are discriminated using a combination of color (RGB) and Hyperspectral Short Wave Infrared (SWIR) cameras. Further, the use of an autoencoder for detecting minerals embedded inside tar- and bitumen mixtures is proposed. Features are extracted from the spectra of the SWIR camera and the texture of the RGB images. For classification, linear discriminant analysis combined with a k-nearest neighbor classification is used. First results show a reliable detection of minerals and positive signs for separability of tar and bitumen. This work is a foundation for developing a sensor-based sorting system for physical separation of tar contaminated samples in road rubble.

**Keywords** Hyperspectral Imaging, Autoencoder, Polycyclic Aromatic Hydrocarbons

## 1 Introduction

Until the 1980s, tar was primarily used as a binder for road surface construction in Germany [1]. It has since been outlawed for the construc-

tion of new roads due to its high levels of Polycyclic Aromatic Hydrocarbons (PAHs) that have been identified to be carcinogenic, mutagenic and genotoxic and can contaminate the groundwater [2]. Further, the use of recycled tar containing materials as a foundation of new road surfaces has been restricted.

Other materials present in road rubble are bitumen, which replaced tar as binder material, and minerals, which make up the biggest part of the road surface mixture (~95 wt%) and are used in the road foundation. Both of these materials are valuable for recycling, but are frequently lost as they cannot be separated from the tar containing fractions. Therefore, they are deposited at a landfill, which is increasingly expensive, or fed into a highly energy consuming tar decontamination process where they are damaged due to high temperatures altering the molecular structure of the minerals.

The mixing of tar contaminated road rubble with uncontaminated bitumen and minerals is due to different road layers and repaired road patches that appear in close proximity and are therefore mixed during demolition. Further, many uncontaminated mixtures are unnecessarily declared as tar containing, as this can be cheaper for the demolition crews than carrying out the mandated testing procedures. This testing includes taking point-samples in a certain raster and having them analyzed in a laboratory.

To acquire a rough estimate over possible PAH concentration, solvent-based paints can be sprayed onto the rubble. Such paints react with the PAHs creating a fluorescent effect that is visually observable. This method is however not sufficient for official classification, as this detection method is not accurate for all PAHs and cannot be used for dense classification and sorting of all material to limit paint usage.

As part of the InnoTeer project, the entire process from the creation of rubble at the construction site to transportation, separation and decontamination is reevaluated [3]. Fraunhofer IOSB is developing a method to efficiently separate the tar from the mixture of materials using visual inspection with the goal to develop a sensor-based sorting system.

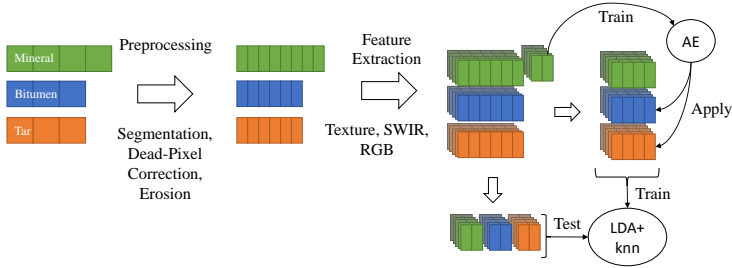
## 1.1 Related Work

Methods such as gas chromatography, high-performance liquid chromatography [4] and mass spectroscopy deliver accurate estimations of

PAH content. However, these Methods offer low throughput at a high cost and require dissolving the tested materials, rendering the methods unsuitable for recycling.

Visual methods for detecting PAHs include fluorescent spectroscopy. UV-excited fluorescence of PAH molecules in the Mid Infrared spectrum is widely used in astronomy to investigate properties of astronomical objects [5]. [6]. However, the detected PAHs are in gaseous form, which alters their fluorescence compared to PAHs in solid compounds. Quazi et al. have used fluorescent spectroscopy to detect and distinguish between different kinds of PAHs in soil samples [7]. Excitation is performed in low-wavelength regions of the visual spectrum (blue to green), detection in slightly higher wavelengths (green to red). Different excitation wavelengths have shown to excite different PAHs. In addition to detection, the varying distribution patterns of different PAHs were observed with phenanthrene forming spherical particles, whereas naphthalene forms a uniform film. The approach seems promising, however the analysis was carried out in microscopic scale and at low speeds (several seconds for a  $200 \times 200\mu\text{m}$  patch). Adaptation of this method to the macroscopic scale has to the best of our knowledge not been tried in the context of PAH detection in soil.

Li et al. use a Fourier Transform Infrared (FTIR)-Spectrometer to measure the reflectance of different PAHs in soil over a broad Mid Infrared spectrum (2500 – 16000nm) with a spectral resolution of  $4\text{cm}^{-1}$  [8]. The 35 measured samples were analyzed using a hybrid variable selection approach, that combines wavelength interval selection and wavelength point selection as preprocessing for a partial least squares regression. The method shows high accuracy, but the use of a point-measuring FTIR-Spectrometer in large throughput sensor-based sorting applications is not feasible. Jahangiri et al. have investigated differences between bitumen-based asphalts in terms of different additives using a FTIR-Spectrometer [9]. This illustrates the big variety in road surfaces which further complicates the task of separating tar from bitumen-based binder.



**Figure 1:** Data Processing pipeline. Preprocessing includes separating the samples from each capture and removing dead pixels. An autoencoder (AE) for detecting minerals embedded in tar and bitumen is trained on a subset of mineral features and applied to the training samples of tar and bitumen.

## 2 Materials and Methods

The problem of detecting tar in road rubble is posed as a classification problem between the classes tar, bitumen and minerals. Solving the problem requires data capture, preprocessing and classification. Preprocessing includes segmentation of the different samples, dead-pixel correction, feature extraction and a novel method for removing mineral patches embedded in the tar and bitumen samples. Figure 1 gives an overview of the different steps used in this work.

### 2.1 Samples

Samples for the classes tar and bitumen are both taken from the top layers of road surfaces and constitute a mixture of differently sized mineral elements and the binder (tar or bitumen). The class of minerals contains only solid pieces of minerals from the foundation layer. The sample size has been chosen to be between 16 and 32mm. Figure 2 shows examples of samples.

### 2.2 Data Acquisition Hardware

In this work, data from a hyperspectral Short Wave Infrared (SWIR) camera and a high-resolution RGB camera were combined for clas-



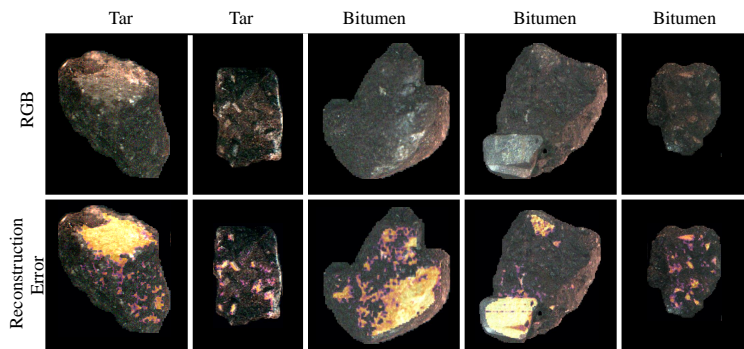
**Figure 2:** Examples for the three classes. From left to right: bitumen, tar, minerals.

sification. Both cameras are line-scanning cameras that have been mounted above the same linear stage. The linear stage carrying the samples is moving past the line-scanning cameras for image acquisition. For the hyperspectral camera, the line is illuminated using six halogen work lights. Illumination for the RGB camera is provided by two white-light LED-bars.

### 2.3 Preprocessing

As a first preprocessing step, dead-pixel correction is performed by quadratic interpolation in the spectral domain. Sample masks are automatically extracted using a binary threshold, with small artifacts being removed by morphological operation (opening) and filtering the remaining elements by size and shape.

Our goal is to be able to overlap RGB- and SWIR images (Image Registration). Therefore, the transformation between the cameras is estimated. First, the nonlinear lens distortion is calculated for each camera separately using a known calibration pattern. The resulting camera pixels are now related through a linear transformation, assuming all captured objects lie in the same plane. The main components of this transformation are a scaling factor, which is necessary because of the different resolutions and slightly different capture areas of the imaging sensors, and a translation between the cameras. These scaling and translation changes could be covered by a similarity transform (which always preserves shape). However, due to small inaccuracies in the mounting of the cameras, a more general perspective transformation is assumed (homography). The transformation matrix is estimated using a set of matching points on a calibration pattern. Using the transformation matrix, both images can be transformed into each others view.



**Figure 3:** Detection of minerals in tar and bitumen. The upper row shows unedited RGB images. The lower row shows an overlay of the RGB images and a contrast-enhanced inverse reconstruction error as computed by the autoencoder.

## 2.4 Distinguishing Surface Minerals from Tar and Bitumen

A challenge when trying to distinguish between tar, bitumen and minerals is that tar and bitumen are mixtures containing large amounts of minerals ( $\sim 95$  wt%) and much less solvent ( $\sim 5$  wt%). Although a thin layer of binder is prevalent, there are several surface patches displaying clean minerals. Figure 3 shows examples for this.

In this work, a pixelwise autoencoder was trained on a subset of samples in the minerals-class. The in- and output of the autoencoder are spectra corresponding to a single pixel. The autoencoder is structured as a multilayer perceptron network with a latent space of 32 neurons. As a preprocessing step for tar and bitumen, the autoencoder is applied to all pixels in the training set. If the reconstructed spectrum is close to the original spectrum, it is assumed that the pixel shows a mineral (see Figure 3). These pixels are disregarded for training. This results in more homogeneous training data and increases the distance between the tar and bitumen classes and the minerals. In Section 3, the effect of this measure on classification performance is discussed.



## 2.5 Feature Extraction

In this work, classification is performed both on a pixel- and an object level. For pixelwise analysis, each pixel is initially treated as a separate sample, whereas objectwise classification uses data collected for an entire sample. As pixel features, the Standard Normal Variate-normalized spectra and their derivatives are used. Object features are the object-wide means of the spectral information as well as texture information. Since texture features require multiple pixels, they are not used in the pixelwise analysis. For texture features, the frequencies in the grayscale-converted RGB image is analyzed using Discrete Fourier Transformation and Local Binary Patterns (LBP) are extracted.

## 2.6 Classification

Classification is either performed using object features, such as extracted texture features and mean spectra, or pixelwise using only the captured spectrum of each pixel. For pixelwise classification, a majority decision (MD) is added to get the desired object wide decisions. Classification is performed using Linear Discriminant Analysis (LDA), combined with a k-nearest neighbor (KNN) classifier. The LDA reduces the feature space to  $n - 1$  where  $n$  is the number of different classes. Other classifiers, such as a multilayer perceptron and a support vector machine, have also been considered, but did not perform as good.

## 3 Experimental Results

Table 1 shows the recall scores for different classification methods. For all classifications, a split of 80/20 for training- and testing data was used. The classification results were cross-validated by using 50 different training/testing splits. Classification was performed either objectwise or using a pixelwise classification with a majority decision.

The pixelwise majority decision model without an autoencoder performed best with an overall recall of 93.69%. For real life scenarios, a reliable detection of tar may be more important than the maximizing recall over all classes, since small amounts of tar can suffice to render a fraction contaminated, prohibiting the use as recycled material. Therefore for the pixelwise majority decision classifiers, robust versions were

**Table 1:** Results for different classification algorithms. Values marked with an asterisk indicate that the classes bitumen and tar were treated as a single class.

| Classification Results (Recall) |           |         |         |        |
|---------------------------------|-----------|---------|---------|--------|
| Classifier                      | Features  | Mineral | Bitumen | Tar    |
| Objectwise                      | All       | 96.65   | 86.86   | 83.57  |
| Objectwise M. vs. O.            | Texture   | 98.2    | 97.85*  | 97.85* |
| Pixel MD                        | SWIR, RGB | 99.71   | 85.60   | 94.84  |
| Pixel MD AE                     | SWIR, RGB | 100.0   | 86.23   | 91.41  |
| Pixel MD robust                 | SWIR, RGB | 98.97   | 56.34   | 100.0  |
| Pixel MD AE robust              | SWIR, RGB | 100.0   | 59.62   | 99.03  |

implemented, that assign all samples with more than 30% of pixels being classified as tar to the tar class. This achieves a perfect recall for tar samples using the pixelwise majority decision and a 99.03% recall when using the autoencoder.

The objectwise classification using both texture- and spectral features performed slightly worse overall than the pixelwise methods. However, it is more computationally which could be critical in real-world systems. For separating minerals from tar and bitumen, a single RGB camera can be sufficient to attain good separation with 98.02% of the detected minerals being true positives. This indicates the possibility of using a low-cost preselection stage using only a RGB camera to remove the minerals from the material flow.

The usage of an autoencoder for preprocessing of the training samples improves the overall classification recall for mineral and bitumen. Especially minerals can be identified consistently, as shown by the recall scores for the two models using the autoencoder. The majority decision to some degree obscures the positive effects of the autoencoder on the robustness of the detection of minerals. This improvement is observable in the overall recall over all pixels without majority decision, as shown in table 2 for pixelwise classification with- and without autoencoder. The False number of false positives in the mineral class has been halved using the autoencoder improving the recall from 98.0% to 99.19%. Recall scores for tar are slightly decreased both for the majority decision and recall over all pixels. One possible explanation for this might be that the tar samples contain a certain type of mineral that is not present in the bitumen samples. Masking out these minerals from

**Table 2:** Results for different classification algorithms on a per-pixel level.

| Pixelwise Classification Results (Recall) |           |         |         |       |
|---|-----------|---------|---------|-------|
| Classifier                                | Features  | Mineral | Bitumen | Tar   |
| Pixelwise                                 | SWIR, RGB | 98.05   | 62.27   | 71.82 |
| Pixelwise with AE                         | SWIR, RGB | 99.19   | 62.87   | 70.17 |

the training samples would therefore remove a means of detecting tar.

## 4 Conclusion and Future Work

In this work, we demonstrated that minerals, tar and bitumen can be distinguished using a combination of a hyperspectral SWIR camera and a RGB camera with overall recall scores of up to 93.69%. Using a robust majority decision, the recall for tar was further increased, resulting in mineral and bitumen fractions with high purity. The use of an autoencoder achieved mixed results, improving the detection of minerals and bitumen, but performing worse in the detection of tar. Possible reasons for this have been identified and will be investigated further.

A focus of future research is determining whether the achieved results generalize to all road rubble. Each of the used fractions in this study is taken from two different sources. Both tar- and bitumen based binders can include additives like rubber, polymer and fiber [9] to optimize for certain properties like temperature stability or noise generation. The utilized differences may be based in large parts on differences in these additives instead of strictly tar- or bitumen specific properties. Evaluation with additional test samples from multiple sources will therefore be needed to further validate the results.

The three classes used in this study do not include rocks used in the foundation layer that are in parts sprayed with a thin layer of PAH contaminated binder for adhesion with the higher road-layers. These foundation-layer rocks are challenging, as the surface contains patches of this adhesive binder as well as patches without this binder. For real-world applications, this class of samples will have to be addressed as well.

Finally, additional measurement systems like fluorescent spectroscopy and MWIR will be utilized to directly identify PAHs or other

chemical properties relating to tar or bitumen. An ideal solution to the problem will deliver estimates for the PAH concentration of each sample in addition to a classification.

## Acknowledgements

This work was supported by the Fraunhofer Internal Programs under Grant No. PREPARE 40-02829. Training samples were kindly provided by Zwisler GmbH.

## References

1. "Polycyclic aromatic hydrocarbons (PAH) [MAK Value Documentation, 2012]," in *The MAK-Collection for Occupational Health and Safety*. John Wiley & Sons, Ltd, 2013, pp. 1–216.
2. P. B. Farmer, R. Singh *et al.*, "Molecular epidemiology studies of carcinogenic environmental pollutants: Effects of polycyclic aromatic hydrocarbons (PAHs) in environmental pollution on exogenous and oxidative DNA damage," *Mutation Research/Reviews in Mutation Research*, vol. 544, no. 2, pp. 397–402, Nov. 2003.
3. "Material aus Straßen nachhaltig recyceln." [Online]. Available: <https://www.umsicht-suro.fraunhofer.de/de/presse/pressemittelungen/2022/InnoTeer.html>
4. S. Kumar, S. Negi *et al.*, "Biological and analytical techniques used for detection of polyaromatic hydrocarbons," *Environmental Science and Pollution Research*, vol. 24, no. 33, pp. 25 810–25 827, Nov. 2017.
5. J. D. T. Smith, B. Draine *et al.*, "The Mid-Infrared Spectrum of Star-forming Galaxies: Global Properties of Polycyclic Aromatic Hydrocarbon Emission," *The Astrophysical Journal*, vol. 656, no. 2, p. 770, Feb. 2007.
6. S. Foschino, O. Berné, and C. Joblin, "Learning mid-ir emission spectra of polycyclic aromatic hydrocarbon populations from observations," *Astronomy & Astrophysics*, vol. 632, p. A84, 2019.
7. F. Qazi, E. Shahsavari *et al.*, "Detection and identification of polyaromatic hydrocarbons (PAHs) contamination in soil using intrinsic fluorescence," *Environmental Pollution*, vol. 272, p. 116010, Mar. 2021.

8. M. Li, Y. Feng *et al.*, "Quantitative analysis of polycyclic aromatic hydrocarbons in soil by infrared spectroscopy combined with hybrid variable selection strategy and partial least squares," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 257, p. 119771, Aug. 2021.
9. B. Jahangiri, K. Barri *et al.*, "A molecular sensing method integrated with support vector machines to characterize asphalt mixtures," *Measurement*, vol. 179, p. 109528, Jul. 2021.



# Increasing the reuse of wood in bulky waste using artificial intelligence and imaging in the VIS, IR, and terahertz ranges

Lukas Roming<sup>1</sup>, Robin Gruna<sup>1</sup>, Jochen Aderhold<sup>2</sup>, Friedrich Schlüter<sup>2</sup>,  
Dovilė Čibiraitė-Lukenskienė<sup>3</sup>, Dominik Gundacker<sup>3</sup>, Fabian  
Friederich<sup>3</sup>, Manuel Bihler<sup>4</sup>, and Michael Heizmann<sup>4</sup>

<sup>1</sup> Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstraße 1, 76131 Karlsruhe, Germany,

<sup>2</sup> Fraunhofer Institute for Wood Research Wilhelm-Klauditz-Institut WKI, Bienroder Weg 54 E, 38108 Braunschweig, Germany,

<sup>3</sup> Fraunhofer Institute for Industrial Mathematics ITWM, Fraunhofer-Platz <sup>1</sup>, 67663 Kaiserslautern, Germany,

<sup>4</sup> Institute of Industrial Information Technology (IIIT), Karlsruhe Institute of Technology (KIT), Hertzstraße 16, 76187 Karlsruhe, Germany,

**Abstract** Bulky waste contains valuable raw materials, especially wood, which accounts for around 50% of the volume. Sorting is very time-consuming in view of the volume and variety of bulky waste and is often still done manually. Therefore, only about half of the available wood is used as a material, while the rest is burned with unsorted waste. In order to improve the material recycling of wood from bulky waste, the project ASKIVIT aims to develop a solution for the automated sorting of bulky waste. For that, a multi-sensor approach is proposed including: (i) Conventional imaging in the visible spectral range; (ii) Near-infrared hyperspectral imaging; (iii) Active heat flow thermography; (iv) Terahertz imaging. This paper presents a demonstrator used to obtain images with the aforementioned sensors. Differences between the imaging systems are discussed and promising results on common problems like painted materials or black plastic are presented. Besides that, pre-examinations show the importance of near-infrared hyperspectral imaging for the characterization of bulky waste.

**Keywords** Material characterization, waste wood, bulky waste,

## 1 Introduction

The increased use of wood is a key to achieve national and international goals in the fight against climate change and minimize the CO<sub>2</sub> footprint [1]. In this situation, the use of waste wood as a substitute for fresh wood is an interesting way to reduce the scarcity of wood. Waste wood for use as a material has meanwhile become a scarce commodity itself in Germany [2]. This is also because, according to national legislation, it can only be reused as raw material if it is free of wood preservatives and other contaminants such as PVC. The development of new sources for "clean" waste wood is therefore gaining importance. Although half of the bulky waste consists of wood, only about half of it has been used as a recycling material so far [3]. Reasons for that are the difficult separation of impurities from wood and a huge variety of materials.

Established methods for sorting bulky waste are manual picking and automatic waste sorting based on heavily shredded materials, with the cost of shredding worsening the ecological balance. A concept similar to the system proposed here was presented in [4], but for the sorting of building rubble that is not as homogeneous as bulky waste.

Thus, the project ASKIVIT (Altholzgewinnung aus Sperrmüll durch künstliche Intelligenz und Bildverarbeitung im VIS-, IR- und Terahertz-Bereich) aims at developing a solution for the automated sorting of bulky waste. The goal is to extract wood, wood-based materials, and non-ferrous metals based on a multi-sensor approach combined with artificial intelligence. Conventional RGB, near-infrared hyperspectral, and thermographic cameras, as well as a developed terahertz imaging system, are used in this work. In the first step, the different sensors are described and the fusion approach based on a convolutional neural network (CNN) is motivated. Preliminary investigations are carried out to determine the potential of near-infrared hyperspectral material characterization using machine learning. Moreover, the benefit of a multi-sensor approach is discussed and verified with sample images.



## 2 Material and methods

In this section, the different imaging systems are described and the fusion approach based on a CNN is motivated.

### 2.1 Visible imaging

Humans can characterize material from bulky waste very accurately only by its appearance in the visible spectral range. Therefore, images from conventional RGB cameras, that imitate the human eye, include highly relevant information. Furthermore, RGB cameras are available in high resolution and often by one order of magnitude more cost-effective compared to other sensors used for material characterization [5].

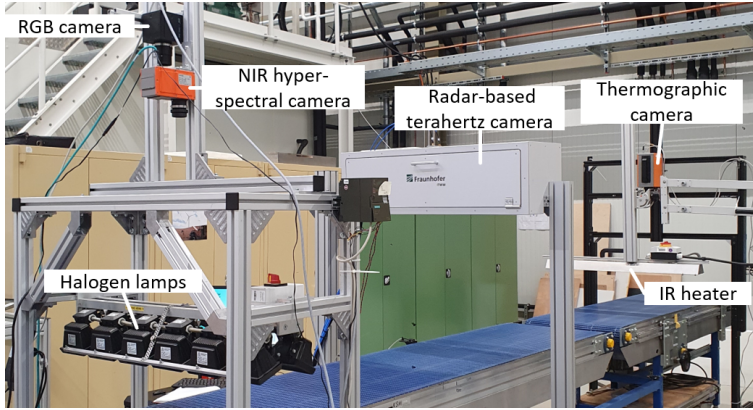
In the course of this study, a prism-based RGB line scan camera (SW-4000T-10GE) was chosen. The built-in prism of the camera splits incoming light onto three spatially separated chips, each measuring one color channel. The frame rate was set to 625 Hz. Halogen lamps were used as a light source for visible as well as near-infrared radiation. The latter was utilized for the near-infrared imaging system.

By moving the samples on a conveyor belt, images with two spatial axes were constructed using the push-broom method. The complete setup including all imaging systems presented in this paper can be seen in Figure 1.

### 2.2 Near-infrared hyperspectral imaging

Near infrared (NIR) hyperspectral imaging is another sensor principle that is used in this work to characterize bulky waste. It is particularly suitable for the detection of organic products and thus also for the identification of wood. Whereas color cameras can only view the superficial appearance, spectral information provided by NIR hyperspectral cameras shows the physical-chemical composition of the material.

As a measuring device, the camera FX17e from SPECIM is chosen. The camera collects hyperspectral images with 224 bands ranging from 900 nm to 1700 nm. The frame rate was chosen to be 104.17 Hz, such that the resolution was equal in both spatial axes of the image.



**Figure 1:** Measurement setup including conventional RGB, NIR hyperspectral, terahertz, and thermography imaging.

### 2.3 Active heat flow thermography

Like the recording of RGB and NIR hyperspectral images, thermography is a camera-based sensor technology. In contrast to the first two methods mentioned, the samples in thermography do not have to be illuminated during the measurement, but are heated in advance. A detector that is sensitive in the thermal infrared range (wavelength: approx.  $3\ \mu\text{m}$  to  $14\ \mu\text{m}$ ) records the thermal radiation that the samples emit on the basis of Planck's law. The radiation intensity depends on the temperature of the samples and their emissivity. In order to be able to make statements about material parameters beyond the emissivity, the samples are heated with infrared radiators as they were transported by the conveyor belt.

The infrared camera is a Geminis 327k ML from IRCAM (Erlangen, Germany) having a dual-band HgCdTe detector (the 1st sensitivity band:  $3.7\text{--}5\ \mu\text{m}$ ; the 2nd sensitivity band:  $8\text{--}9.4\ \mu\text{m}$ ) with  $640 \times 512$  pixel. Only the 2nd band was used in order to avoid parasitic signals from direct irradiation by the infrared heater into the camera. A frame rate of 100 Hz and a 25 mm lens were used. The camera was arranged in such a way that the width of the conveyor belt filled the image along the long edge. The distance between the camera and the

heater amounted to 0.6 m.

The infrared heater consists of two Carbon Twin-Tube Emitters from Heraeus Noblelight having a length of 0.7 m and a power of 6000 W/m each. The peak wavelength of their radiation spectrum was 2  $\mu\text{m}$ . The heaters were placed about 0.28 m above the conveyor belt. Given the velocity of the conveyor belt of 0.108 m/s, the energy per area deposited in the samples is

$$E_A = \frac{6000 \frac{\text{W}}{\text{m}}}{0.108 \frac{\text{m}}{\text{s}}} = 55.56 \frac{\text{kJ}}{\text{m}^2}. \quad (1)$$

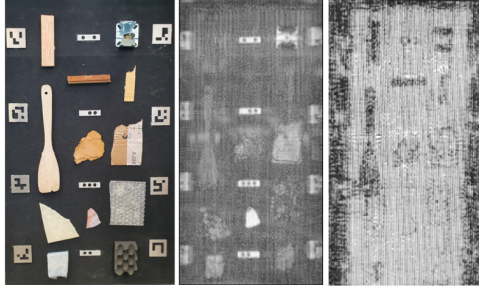
The increase in temperature on the sample surface as a result of heating by the radiant heater depends on the underlying thermophysical parameters. Therefore, structured samples can obtain a characteristic temperature pattern that allows a look underneath the sample surface.

## 2.4 Terahertz imaging

Terahertz radiation is electromagnetic radiation between far infrared and millimeter waves. Due to the capability of terahertz waves to penetrate through most of the dielectric materials, such as plastics, paper, foams, or upholstery, the differences in the refractive index may be observed in 3D [6]. Opposed to X-ray radiation, terahertz is non-ionizing. Therefore, it enables safe 3D imaging on complex structures, which are common for bulky waste.

For this application, a terahertz camera was developed as a line scan camera with 12 emitters and 12 receivers, which operate in the W-band (75–110 GHz or approx. 2.7–4 mm wavelength). A synthetic aperture radar (SAR) design for the terahertz imaging system was chosen [7]. The received signal (amplitude and phase) depends on the refractive index and spatial position of the sample structure. The aim of this system is to provide additional 3D information on overlapping and complex features of pre-crashed bulky waste.

144 effective aperture elements (12 emitter and 12 receiver combinations) are scanned for all frequencies  $Nf$  that are used to scan the scene within the W-band. The data acquisition algorithm obtains measured reference, receiver, and encoder signals. The data acquisition time as



**Figure 2:** Terahertz measurement on a sample with various materials (left), and two reconstructed terahertz images at various distances to the array to obtain reflection and shadow images, respectively.

well as the resolution depends on the number of frequency points  $N_f$  and the covered bandwidth, respectively. Each set of complex  $N_f \times 144$  data has to be reconstructed in a defined reconstruction volume in order to obtain a 3D image, which can be later observed at each reconstruction plane (referenced as a distance to the imaging array).

The used reconstruction algorithm is based on matched filter approach [8]. For the given sample shown in the photograph on the left of Figure 2, the reconstruction volume of  $80 \times 40 \times 10$  cm was chosen with corresponding  $800 \times 400 \times 50$  voxels. The reconstruction was made from 134 line scans, i.e. on average, one picture was taken every 6 mm with a speed of 0.30 m/s.

The reconstructed images show good results from the reflection of the objects (middle) as well as from the shadow image (right). The metal reflects most of the radiation, whereas shaped metals show prominent shapes due to scattering from the surfaces which are not parallel to the scanner imaging plane. A piece of a CD (as well as metallic markers) shows the strongest reflection due to conductive materials and a parallel face toward the scanner. Wood and cardboard reflect part of the radiation. The chosen rubber mat has a stripped structure, which reflects a big part of terahertz radiation giving a good contrast for shadow images of the wood. Upholstery and plastics are the most transparent in the terahertz range, therefore only tiny changes in the image can be recognized. This is important for the characterization of material composites, as terahertz radiation enables the detection

of wood and metal underneath upholstery or plastic.

The terahertz images in Chapter 3.2 were obtained using 0.108 m/s conveyor belt speed. The line scans were obtained every 6 ms. The chosen reconstruction volume was  $80 \times 55 \times 18$  cm with  $1600 \times 550 \times 80$  voxels in  $x, y, z$  directions correspondingly.

## 2.5 Sensor data fusion approach

The characterization of materials can be solved by a broad variety of classification methods, including classical and machine learning methods [9, 10]. Senecal et al. showed that using a CNN optimized for multispectral data can result in very high classification accuracy if the data set is large enough [11]. However, multispectral datasets are often very limited in size. Therefore, it is a key point in our project to enable fast data recording to capture a dataset sufficient in size. This is done by using the setup described in the previous sections. The benefit of CNN architectures is that they can use much of the spatial and all spectral information at the same time, and therefore make use of the spectral differences between the materials early. The relevant spatial and spectral features are learned by the network automatically and simultaneously, which is hard to reproduce by a classical feature design.

To combine the information of the proposed sensor modalities, a fusion technique together with a registration is necessary. In this way, the strength of each imaging system can be used to achieve a classification result better than using one technology individually. Lately, early fusion methods based on deep learning e.g. CNNs show very promising results on multispectral datasets like EuroSAT [12]. In early fusion, data from various sensors is registered and merged before classification [13].

In our project, the registration is done by using a marker-based registration approach. For the registration of RGB, NIR, and thermographic cameras, Aruco markers [14] are introduced supported by a similar marker for the Terahertz spectrum. With this marker-based approach, the image registration is robust and accurate, even if sensors show significantly different intensities on the same object. After registration, the preprocessed data from all sensors will be given into a CNN, which is currently under development. The CNN will implicitly perform an early fusion and classify the material perceived by the sensors.

### 3 Results and discussion

After describing the setup, preliminary results of NIR hyperspectral imaging will be presented. Moreover, recordings from all imaging systems will be shown and discussed.

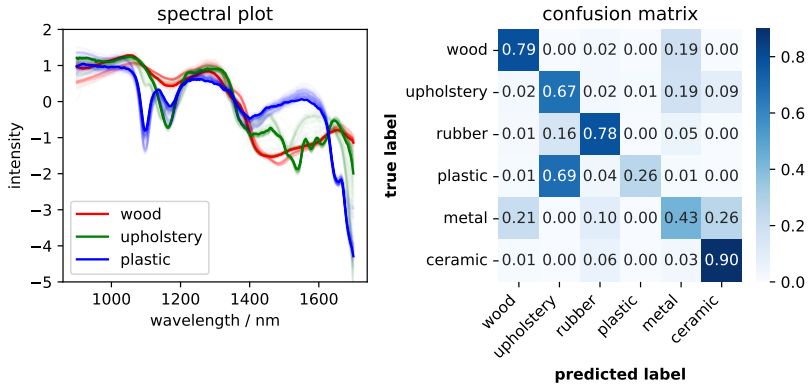
#### 3.1 Preliminary results of NIR hyperspectral imaging

Hyperspectral image analysis is state of the art for material characterization used for sorting applications. Therefore, pre-examinations have been carried out based on NIR hyperspectral data combined with a common classifier, namely partial least squares discriminant analysis (PLS-DA). The samples to be analyzed are different objects appearing in bulky waste. The objects were divided into six classes, namely wood, upholstery, rubber, plastic, metal, and ceramic. Each class can include slightly different types of material. The class wood for example included particle board, old varnished window scantlings, high-density fiberboard, and plywood.

Hyperspectral images of the samples were acquired using the FX17e camera and the setup described in section 2.2. Eight images of different sample collections were chosen for training from which  $10^5$  pixels were randomly selected. From another eight images,  $10^4$  pixels were extracted for testing. A single pixel contains 224 values, each representing the reflectance of the material at a different wavelength. As a preprocessing step, standard normal variate (SNV) correction was performed [15]. Additionally, outliers that differ more than five standard deviations from the mean have been removed from training data in order to improve the classification model.

In the spectral plot of Figure 3 the intensity over wavelengths for different materials is visualized. The intensity values can be negative due to SNV correction. Several spectra are drawn on top of each other for each class, making the variance of the data visible. It can be seen that the spectral data varies very little within each class and, by looking at the course of the spectra, the classes are visually distinguishable from each other.

The classification performance of the PLS-DA model is evaluated on test data with a confusion matrix (on the right of Figure 3). The overall accuracy on test data is 0.64. In the confusion matrix, it can be seen that



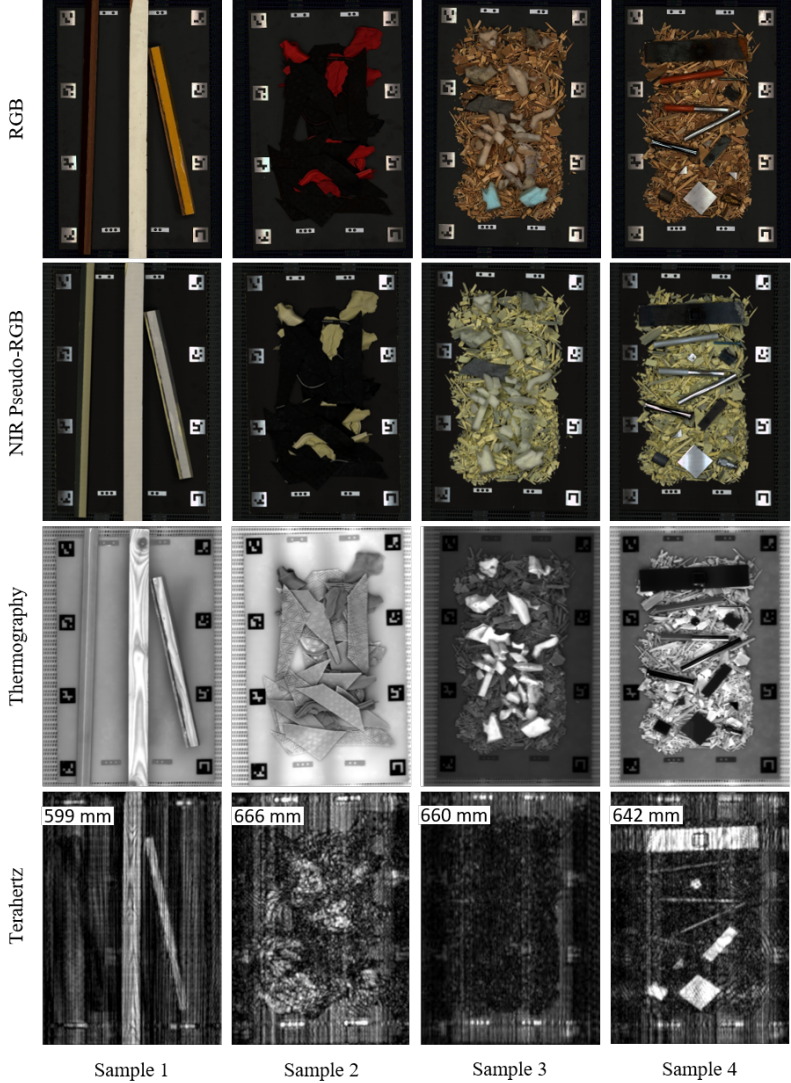
**Figure 3:** Measured spectra (left) of different materials after SNV correction and outlier removal. And confusion matrix (right) of PLS-DA classifier trained and tested on NIR hyperspectral data.

plastic is falsely classified as upholstery in most cases. A reason for that might be that the two materials are not linearly separable. However, the material wood (including particle board, varnished wood, fiberboard, and plywood) is classified correctly with a probability of 0.79. This confirms the assumption, that NIR hyperspectral imaging gives highly relevant information for detecting waste wood in bulky waste.

### 3.2 Comparison of sensor modalities

After showing the potential of hyperspectral material characterization in the near-infrared range, this section will focus on the comparison of the presented imaging systems. Therefore, four sample quantities were chosen and images were recorded using the setup shown in Figure 1. The results can be seen in Figure 4.

Sample 1 contains old varnished window scantlings, and Sample 2 are pieces of red and black rubber mats. Samples 3 and 4 are wood chips partially covered with foam and metal pieces, correspondingly. RGB and NIR hyperspectral data contain multiple channels, each representing a different wavelength. The corresponding images are in color or rather false color in the case of NIR hyperspectral data (selected



**Figure 4:** Various samples acquired by various sensor modalities. Each row shows a corresponding imaging technology from top to bottom: RGB, NIR hyperspectral, thermographic, and terahertz imaging.



wavelengths are 1100 nm, 1300 nm, and 1500 nm). In the terahertz pictures, the given number defines the visualized plane from the whole reconstruction volume by the distance of the plane to the imaging array. The distance is chosen such that the features relevant to the underlying comparison are visible. The sample carrier is approximately 680 mm away from the terahertz imaging array.

The RGB image of Sample 1 shows the paint color and surface but does not reveal the wood structure. The same applies to the NIR pseudo-RGB image, but it is less affected by the paint. The thermographic and terahertz images show the wood texture with its characteristic annual ring pattern under the paint so that this sample can be clearly identified as wood with help of thermography or terahertz. The terahertz image shows the upper plane that is 599 mm below the imaging array, which leads to a sample thickness of approx. 8 cm.

Sample 2 shows a common problem of sorting black polymers. It is not readily recycled in conventional plastic sorting facilities due to the high absorption of black pigments to radiation in NIR or visible wavelength range [16]. The red rubber chips in Figure 4 are clearly visible in the RGB image, while the black ones are hardly recognizable on the background of the black sample carrier. This also applies to the NIR pseudo-RGB image. In thermography, however, red and black rubber both have a significantly improved sensitivity and can therefore be easily distinguished from the background. The terahertz image contains information about the height of the visible mats encoded in the reconstructed volume. The image is blurred out due to the scattering of the texture of the black mats.

Samples 3 and 4 show foam and metal on wood chips, respectively. NIR pseudo-RGB images are again less influenced by the paint color of the material in comparison to the RGB images. Foam and metal are distinguishable from wood chips in almost all images. Terahertz images show strong reflection from metals, whereas wood chips absorb most of the radiation. In thermography, metal appears darker than wood because it absorbs the radiation from the radiant heater less and has a higher heat capacity and lower emissivity than wood. In contrast to that, foam appears very bright due to its low thermal capacity.

## 4 Conclusions and outlook

A novel approach for bulky waste material characterization has been presented. Different sensor modalities including visible, NIR hyperspectral, thermography, and terahertz imaging are exploited to achieve a better classification result than using a single technology individually. Regarding terahertz imaging, a synthetic aperture radar system was developed, which is specifically designed for sorting applications. The system aims to provide additional 3D information on overlapping and complex features of pre-crashed bulky waste.

All four imaging systems were brought together to build a demonstrator acquiring data using RGB, NIR, thermography, and terahertz imaging techniques in one attempt. The recorded and post-processed images showed promising results on common problems like painted materials or black plastic. The presented thermography and terahertz images reveal the wood texture with its characteristic annual ring pattern under the paint. Besides that, thermography showed good sensitivity for plastic regardless of color.

Pre-examinations on NIR hyperspectral data have shown that waste wood is distinguishable from plastic and upholstery. Furthermore, using a PLS-DA six different materials from the used set of bulky waste samples were classified with an accuracy of 0.64.

Whereas the PLS-DA estimated the class of each pixel separately, a CNN is able to make use of the spatial and spectral information at the same time. Therefore, a CNN performing a patch-wise classification on all sensor modalities will be part of future work. With an even larger dataset, the goal is to reach a high classification accuracy on a huge variety of different materials from bulky waste. With thermographic and terahertz imaging it might be even possible to look underneath overlapping material.

## Acknowledgement

The project ASKIVIT is funded by the German Federal Ministry of Food and Agriculture (BMEL) through the Fachagentur Nachwachsende Rohstoffe e. V. under the funding reference 2220HV048A.

## References

1. A. Purkus, J. Lüdtkke, G. Becher, M. Dieter, D. Jochem, R. Lehnen, M. Liesebach, H. Polley, S. Rüter, J. Schweinle *et al.*, *Evaluation der Charta für Holz 2.0: Methodische Grundlagen und Evaluationskonzept*. Thünen Report, 2019, no. 68.
2. G. Ludwig, E. Gawel, and N. Pannicke-Prochnow, *Altholz in der Kaskadennutzung—eine Bestandsaufnahme für Deutschland*. Springer, 2022.
3. B. Bilitewski, J. Wagner, and J. Reichenbach, *Bewährte Verfahren zur kommunalen Abfallbewirtschaftung; Informationssammlung über Ansätze zur nachhaltigen Gestaltung der kommunalen Abfallbewirtschaftung und dafür geeignete Technologien und Ausrüstungen*. Umweltbundesamt, Dessau-Roßlau, 2018.
4. T. J. Lukka, T. Tossavainen, J. V. Kujala, and T. Raiko, “Zenrobotics recycler-robotic sorting using machine learning,” in *Proceedings of the International Conference on Sensor-Based Sorting (SBS)*, 2014, pp. 1–8.
5. N. Kroell, K. Johnen, X. Chen, and A. Feil, “Fine metal-rich waste stream characterization based on RGB data: Comparison between feature-based and deep learning classification methods,” in *OCM 2021 - Optical Characterization of Materials : Conference Proceedings*. Ed.: J. Beyerer; T. Längle, 2021, pp. 83–97.
6. E. Castro-Camus, M. Koch, and D. M. Mittleman, “Recent advances in terahertz imaging: 1999 to 2021,” *Appl. Phys. B*, vol. 128, no. 1, Jan. 2022.
7. B. Baccouche, P. Agostini, F. Schneider, W. Sauer-Greff, R. Urbansky, and F. Friederich, “Comparison of digital beamforming algorithms for 3-d terahertz imaging with sparse multistatic line arrays,” *Advances in Radio Science*, vol. 15, pp. 283–292, 12 2017.
8. M. Pastorino, *Microwave Imaging*, ser. Wiley Series in Microwave and Optical Engineering. Hoboken, NJ: Wiley-Blackwell, march 2010.
9. R. Sharma, R. Hussung, A. Keil, F. Friederich, T. Fromenteze, M. Khalily, B. Deka, V. Fusco, and O. Yurduseven, “Coded-aperture computational millimeter-wave image classifier using convolutional neural network,” *IEEE Access*, vol. 9, pp. 119 830–119 844, 2021.
10. M. Bauer, R. Hussung, C. Matheis, H. Reichert, P. Weichenberger, J. Beck, U. Matuszczyk, J. Jonuscheit, and F. Friederich, “Fast fmcw terahertz imaging for in-process defect detection in press sleeves for the paper industry and image evaluation with a machine learning approach,” *Sensors*, vol. 21, no. 19, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/19/6569>

11. J. J. Senecal, J. W. Sheppard, and J. A. Shaw, "Efficient convolutional neural networks for multi-spectral image classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8.
12. P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 7, pp. 2217–2226, 2019.
13. K. Gadzicki, R. Khamsehashari, and C. Zetsche, "Early vs late fusion in multimodal convolutional neural networks," in *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, 2020, pp. 1–6.
14. S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.
15. Å. Rinnan, F. Van Den Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," *TrAC Trends in Analytical Chemistry*, vol. 28, no. 10, pp. 1201–1222, 2009.
16. A. Turner, "Black plastics: Linear and circular economies, hazardous additives and marine pollution," *Environment International*, vol. 117, pp. 308–318, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0160412018302125>

# Semi-supervised methods for CNN based classification of multispectral imagery

Manuel Bihler, Jiachen Zhou, and Michael Heizmann

Institute of Industrial Information Technology (IIIT), Karlsruhe Institute of Technology (KIT), Hertzstraße 16, 76187 Karlsruhe, Germany

**Abstract** Deep Convolutional neuronal networks, with their recent increase in performance, have become one of the standard techniques for RGB image classification. Due to a lack of large labeled datasets, this is not the case for multispectral image classification. To overcome this, we analyze the use of semi-supervised learning for the case of multispectral datasets. We use parameter reduction strategies to create small and efficient multispectral CNNs and combine these computationally efficient classifiers with semi-supervised learning methods. We choose the state-of-the-art semi-supervised methods MixMatch, ReMix-Match, FixMatch, and FlexMatch, to conduct experiments on the multispectral dataset EuroSAT. Additionally, we challenge this semi-supervised multispectral approach with a decreasing number of labeled images. We found that with only 15 labeled images per class, we can reach an accuracy above 80 %. If more labeled images are provided, the analyzed semi-supervised methods can even surpass basic supervised learning strategies.

**Keywords** Artificial intelligence, image processing, multispectral images, semi-supervised learning, CNN, consistency regularization, parameter reduction

## 1 Introduction

The use of deep convolutional neural networks for RGB image classification has led to a series of breakthroughs [1–4]. Extending convolutional neural networks to process multispectral imagery is becoming increasingly prevalent, especially in the field of characterization of materials, quality insurance in the food industry, or recycling of waste

materials [5]. In these fields, it is common to use multispectral (MS) data to separate materials based on their different spectral characteristics. While AI systems like CNNs show superior performance on large RGB datasets [1, 3, 4], the lack of large labeled multispectral datasets makes them difficult to employ in a multispectral setting. Compared to RGB images where there exist large publicly available datasets such as CIFAR-10 [6], and ImageNet [7], large labeled multispectral datasets are rare. In this work, we aim to improve the performance of CNNs on small unlabeled multispectral datasets by combining semi-supervised learning (SSL) methods with CNNs optimized for multispectral data (multispectral CNNs).

Semi-supervised learning provides a powerful tool to leverage unlabeled data and too largely alleviate the need for labeled data. This is particularly advantageous when collecting labeled data is expensive or time-consuming because expert knowledge or expensive machinery may be involved in the labeling process. This approach has shown impressive results in a wide variety of tasks, including facial expression recognition and natural language processing [8, 9].

To the best of our knowledge, the combination of SSL methods and multispectral CNNs is not discussed in previous work. We present a study on recently proposed state-of-the-art SSL methods in the context of classifying multispectral images. In this work, we show that modern SSL methods can be very effectively used to reduce the need for labeled data drastically. We also aim to make SLL methods more comprehensible for researchers outside the deep learning community. Therefore, in detail, we describe the methods used in the following section and then show results based on the EuroSAT dataset [10].

## 2 Semi-Supervised Methods

In image classification, semi-supervised learning (SSL) has proven to be a powerful paradigm for utilizing unlabeled data to mitigate the reliance on large labeled datasets. Compared with the results of previous SSL algorithms ( $\pi$ -Model [11], Mean teacher [12], Virtual Adversarial Training [13] and Pseudo-Label [14]), the four state-of-the-art SSL algorithms: MixMatch [15], ReMixMatch [16], FixMatch [17], and FlexMatch [18], all unify the current hybrid approaches for SSL. In this

section, we bring an overview of these four algorithms.

**1. MixMatch:** Unlike previous methods [11, 14], MixMatch introduces a single loss term unifying all three main semi-supervised approaches: entropy minimization [14, 19], consistency regularization [11, 20] and generic regularization [21, 22]. MixMatch utilizes a form of consistency regularization by using data augmentation for images. Two data augmentation methods are used subsequentially on both labeled and unlabeled images: first *random horizontal flip* and then *random crop*. Like Pseudo-Label [14], MixMatch applies multiple individual augmentations on an unlabeled image to create different instances, whose model predictions are then averaged to generate one pseudo-label for this unlabeled image. MixMatch uses a slightly changed version of the MixUp algorithm for regularization. Both labeled and unlabeled images and their corresponding labels are interpolated to generate mixed inputs and mixed labels.

**2. ReMixMatch:** To make MixMatch more data-efficient, two new techniques are introduced and directly integrated into MixMatch’s framework: distribution alignment and augmentation anchoring. Distribution alignment maximizes the mutual information between model inputs and outputs so that unlabeled data is fully utilized to improve the model’s performance. Distribution alignment encourages the marginal distribution of the model’s predictions on unlabeled data to match the marginal distribution of the ground-truth labels. Recent work found that applying stronger forms of data augmentation can significantly improve the performance of consistency regularization [23]. Augmentation anchoring is added as a replacement for the consistency regularization in MixMatch. The basic idea is to use the model’s prediction for a weakly augmented unlabeled image as the pseudo-label for many strongly augmented versions of the same image.

**3. FixMatch:** FixMatch is a significant simplification compared with MixMatch and ReMixMatch. Its simplification lies in combining only two main approaches to semi-supervised learning: consistency regularization and Pseudo-Label [14]. FixMatch first generates pseudo-labels on weakly augmented unlabeled images using their model predictions. For a given image, the pseudo-label is only retained if the model produces a high-confidence prediction. In other words, when the model assigns a probability to any class above the predefined threshold  $\tau$ , the prediction is accepted, and the model output is then converted to a

one-hot pseudo label. Then, the model’s prediction for a strongly augmented version of the same image is used to train the model against this pseudo-label.

**4. FlexMatch:** FixMatch uses a predefined constant threshold  $\tau$  for all classes to select unlabeled data that contribute to the training, thus failing to consider different learning statuses and learning difficulties of different classes. To address this issue, Curriculum Pseudo Labeling (CPL) is introduced to utilize unlabeled data according to the model’s learning status. The core of CPL is to adjust thresholds for different classes at each time step to feed the model with the fitting unlabeled data for the current learning status.

### 3 Results

In this section, we discuss our three main results. First, we present our classifier with a reduced number of parameters optimized for MS data and show the classification results on RGB and MS datasets, using supervised learning (SL). Secondly, we present the classification results using our classifier in combination with the above discussed SSL methods. Lastly, we show how the combination of MS data and SSL methods performs on datasets with a drastically decreased number of labeled images.

We use the datasets CIFAR-10 [24] and EuroSAT [10]. While CIFAR-10 is only used as a benchmarking dataset, EuroSAT is our main dataset for learning and testing the discussed strategies and methods. With 27,000 patches, EuroSAT is currently the largest labeled multispectral dataset for image patch classification. Additionally, it also contains the RGB bands, making it a perfect candidate for comparing RGB and MS learning strategies. Each multispectral image in the EuroSAT dataset consists of 13 channels, but only ten are relevant for identifying and monitoring land use classes and are used in our experiments. For the following experiments, we randomly sample 20 % and 10 % of labeled data from this dataset as validation and test sets respectively, while the remaining 18,900 labeled images are used as training data in either semi-supervised or fully supervised learning. We make sure that there is no overlap between these datasets.

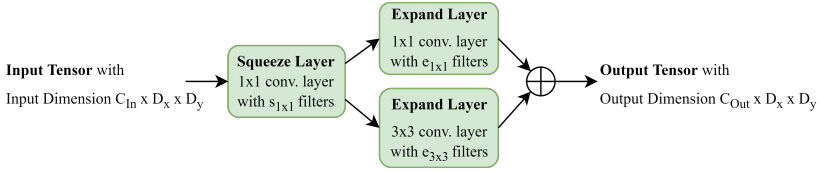


### 3.1 Parameter Reduction

The success of deep neuronal networks like ResNet [25], or Wide ResNet [26], with their thousands of layers and millions of parameters, also lies in the availability of enormous datasets like CIFAR-10. In the case of multispectral imagery, where such datasets are lacking, very deep networks would easily overfit due to the extreme number of model parameters. Additionally, applying semi-supervised algorithms with deep CNNs as backbone classifiers can consume significant computational resources, making it a very costly and time-consuming combination of methods. To tackle this problem, we develop our own classifier optimized for the case of semi-supervised learning for multispectral imagery. This classifier is based on the Wide ResNet architecture and adopts parameter-reducing strategies presented in recent work on small and efficient CNNs, such as SqueezeNet [27] and MobileNet [28].

For further modification and evaluation, we choose the following Wide ResNet structures with fewer parameters while maintaining competitive accuracy according to the results in [26]: WRN-40-04, WRN-16-08, WRN-22-08 and WRN-28-10, where the first number depicts the *depth* and the second the widening factor  $k$ .

The structure of each residual block in the Wide ResNet consists of two  $3 \times 3$  convolutional layers and hence is named  $B(3, 3)$ , where  $B$  indicates the building block and  $(3, 3)$  the list of two kernel sizes of the convolutional layers. To decrease the number of parameters further, we additionally apply the microstructure from SqueezeNet [27] in every building block. Specifically, we replace all the  $3 \times 3$  convolutional layers in each  $B(3, 3)$  building block with Fire Modules from SqueezeNet. In Figure 1 a sketch of the Fire module is depicted, and a detailed description of all variables used in the following is given in the caption. In each Fire Module, we set  $s_{1 \times 1}$  equals to  $0.125 \cdot C_{In}$ ,  $e_{1 \times 1}$  equals to  $0.75 \cdot C_{Out}$  and  $e_{3 \times 3}$  equals to  $0.25 \cdot C_{Out}$ . The number of input and output channels of each  $3 \times 3$  convolutional layer in the  $B(3, 3)$  block will be kept the same after replacement. The macro network structure of the original Wide ResNet will also be preserved. Hence, we call our network Wide ResNet with Fire Modules (WRN+FM). It closely mimics the macro-architectural design of the Wide ResNet architecture while adapting the micro-architectural elements from the SqueezeNet to reduce network parameters.



**Figure 1:** Fire Module structure as replacement for  $3 \times 3$  convolutional layer.

$C_{In}, C_{Out}$ : Number of input or output channels of the network block.

$s_{1 \times 1}$ : Number of output channels of the *Squeeze*-Layer.

$e_{1 \times 1}, e_{3 \times 3}$ : Number of output channels of the  $1 \times 1$  or  $3 \times 3$  convolutional layer in the *Expand*-Layer, where  $e_{1 \times 1} + e_{3 \times 3} = C_{Out}$ .

We evaluate the new set of classifiers on two datasets, the RGB dataset CIFAR-10, and the multispectral dataset EuroSAT. In this section, we only use fully supervised learning to be able to compare our results with other SL benchmarks. For data augmentation, we do not use heavy data augmentation as proposed in semi-supervised learning algorithms and use only horizontal flips and random crops for images. Supervised training of Wide ResNet-28-10 (without FM) consumes too much training time and computing resources; therefore, we show results from literature [26, 29]. Our experimental results are shown in Table 1.

It can be concluded from Table 1 that applying Fire Modules into the Wide ResNet structure brings benefits and also some expected downsides. With this parameter reduction strategy, the total number of network parameters can be significantly reduced, up to about 90% of the original network size. As a result, our WRN-28-10+FM consists of only 2.42 million parameters and is 15 times smaller than the original WRN-28-10. Nevertheless, it achieves a classification accuracy of 96.19% on the EuroSAT MS dataset, only 0.41% less than the benchmark network SpectrumNet. From the results on EuroSAT in Table 2, we find that WRN-28-10+FM can achieve the best validation accuracy among our four new networks.

### 3.2 Semi-supervised Methods on MS data

We conduct experiments for the four selected SSL methods on the EuroSAT dataset using our classifier WRN-28-10+FM and exhibit the re-

**Table 1:** Evaluation of different versions of Wide ResNet with and without Fire Modules on different datasets using fully supervised learning. The marked results are extracted from literature.

| Dataset                  | Classifier                | Parameter | Accuracy (%)  |
|--------------------------|---------------------------|-----------|---------------|
| CIFAR-10<br>RGB          | WRN-28-10                 | 36.49 M   | <b>95.83*</b> |
|                          | WRN-28-10+FM <sub>s</sub> | 2.40 M    | 92.51         |
|                          | WRN-22-08                 | 17.20 M   | 95.62*        |
|                          | WRN-22-08+FM <sub>s</sub> | 1.20 M    | 91.51         |
|                          | WRN-16-08                 | 11.00 M   | 95.19*        |
|                          | WRN-16-08+FM <sub>s</sub> | 0.86 M    | 90.79         |
|                          | WRN-40-04                 | 8.90 M    | 95.03*        |
|                          | WRN-40-04+FM <sub>s</sub> | 0.57 M    | 90.25         |
|                          | SpectrumNet               | 0.72 M    | 92.29         |
| EuroSAT<br>Multispectral | WRN-28-10+FM <sub>s</sub> | 2.42 M    | 96.19         |
|                          | WRN-22-08+FM <sub>s</sub> | 1.21 M    | 95.76         |
|                          | WRN-16-08+FM <sub>s</sub> | 0.87 M    | 94.89         |
|                          | WRN-40-04+FM <sub>s</sub> | 0.58 M    | 94.25         |
|                          | SpectrumNet (Benchmark)   | 0.73 M    | <b>96.60*</b> |

sults in Table 2. For semi-supervised learning, the number of labels for RGB and MS imagery is limited to 165 per class, i.e., the total number of labeled images for training is 1,650. This represents 6% of the entire dataset. The number of unlabeled images is set to 4,000 for both RGB and MS datasets to create a more realistic setting, as collecting high-dimensional MS images is more expensive and time-consuming. For comparison against supervised learning, we also conduct experiments using four different numbers of labeled images: (i) 5,650 to mimic the semi-supervised setting with the same number of samples: 4,000 unlabeled and 1,650 labeled images; (ii) 1,650 labeled images to simulate the same number of labeled images; (iii) 850 images and; (iv) 18,900 images to test the (unfair) lower and upper limit of supervised learning.

Table 2 show that all four SSL methods can still help our network achieve comparative classification accuracy, even though only limited labeled data is used. As expected, the supervised approach with the full amount of labeled images performs the best, with 96.56%. However, if the total number of labels is reduced to 5,650, the supervised method is outperformed by the semi-supervised method ReMixMatch

by 0.69%, although only 165 labeled images are used per class. One reason for this advantage of ReMixMatch lies in the utilization of strong data augmentation applied on both labeled and unlabeled images, which improves the performance of consistency regularization and helps the network achieve better robustness to noisy data. In general, MS images are expected to result in greater classification accuracy than RGB images in theory, given the additional information that is present in the spectral bands and increases the separation between classes. Except for MixMatch, all methods meet our expectations and perform better under MS conditions by 1.37% on average.

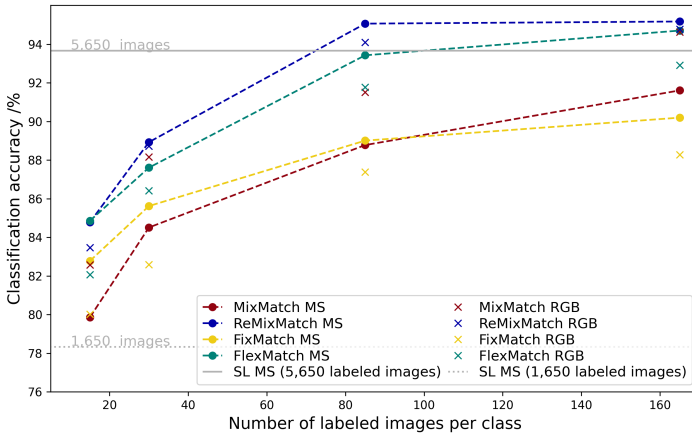
**Table 2:** Results of different semi-supervised learning methods on EuroSAT RGB and MS dataset using our WRN-28-10+FMs as classifier. Supervised learning with 850 and 18,900 images are not comparable with the SSL methods, they show the upper and lower limit of the methods for benchmarking purpose.

| Dataset        | SSL Methods                                   | Accuracy (%) |
|----------------|---|--------------|
| EuroSAT<br>RGB | MixMatch                                      | 94.64        |
|                | ReMixMatch                                    | 94.78        |
|                | FixMatch                                      | 88.28        |
|                | FlexMatch                                     | 92.91        |
| EuroSAT<br>MS  | MixMatch                                      | 91.61        |
|                | ReMixMatch                                    | <b>95.18</b> |
|                | FixMatch                                      | 90.20        |
|                | FlexMatch                                     | 94.71        |
|                | SL with 850 images                            | 68.65        |
|                | SL with 1,650 images (same number of labels)  | 78.33        |
|                | SL with 5,650 images (same number of samples) | 94.49        |
|                | SL with 18,900 images                         | 96.56        |

### 3.3 Limited number of labeled images

In this section, we drastically decrease the number of labeled images to test the limit of the discussed semi-supervised methods. The number of labeled MS images is decreased to 15, 30, 85 images per class, which represents only 0.5 %, 1% and, 3% of the entire dataset, while keeping the total number of unlabeled images the same with 4,000. This procedure is similar to other benchmarks in the literature [15–18].

The results from Figure 2 show that the classification performance of the network becomes better with an increasing number of labeled samples used in training. Among all the SSL methods, ReMixMatch consistently outperforms the other methods. FlexMatch follows ReMixMatch and proves to be the second best. The reason for this trend can be concluded as following: on the one hand, distribution alignment in ReMixMatch not only minimizes the entropy of pseudo labels for unlabeled data like all the other SSL methods do but also maximizes the mutual information between model inputs and outputs to incorporate unlabeled data for better model performance. On the other hand, a rotation loss [30] is directly included in the ReMixMatch loss term. Comparing SSL and SL for the case of 85 images per class drastically shows the power of semi-supervised learning. The SL approach with 850 images can only reach a classification accuracy of 68.65%, while the best SSL method reaches 95.07%.



**Figure 2:** Results for the four SSL methods with a limited number of labeled images. For the SSL methods, 4,000 unlabeled images are available in addition to the depicted number of labeled images. For supervised learning, a gray solid/dashed line is shown for the case of the same number of samples (5,650 images) and the same number of labeled images (1,650 images), respectively.

## 4 Conclusions and Outlook

By adjusting the macro size of the Wide ResNet architecture and changing the micro-structure according to the SqueezeNet architecture, we obtain a small and efficient network with up to 15 times fewer parameters. We show that this network can compete with other popular networks on RGB datasets and can also be effectively trained on much smaller multispectral datasets. Based on the increased computational speed, it can be combined with modern SSL methods for RGB and multispectral datasets. To the best of our knowledge, the combination of SSL methods compressed CNNs, and multispectral datasets, have not been discussed in previous work. This work proves that using 85 images per class, state-of-the-Art SSL methods reach similar or even higher accuracies than supervised learning, depending on the augmentation strategies of the supervised approach. By decreasing the number of labeled images to 15 per class, the power of semi-supervised learning becomes even more prevalent, with 84.78% compared to SL 78.33% (1,650 images). Our results show that the newest SSL method in our comparison ReMixMatch outperforms the other methods not only for RGB but also for multispectral data.

These results show that SSL can be applied to MS data, and expensive labeling can be reduced dramatically. However, more research is needed to improve the number of augmentation strategies for multispectral data. Data augmentation plays a vital role in semi-supervised learning. There are still only a few specialized data augmentations available for multispectral channels compared with RGB channels. In future work, we are interested in investigating data augmentation methods for multispectral imagery according to the characteristics of different channels. We expect that the shown methods can increase the total number of available labeled datasets, which would benefit the whole research community in the field of image classification.

## Acknowledgement

The project ASKIVIT is funded by the German Federal Ministry of Food and Agriculture (BMEL) through the Fachagentur Nachwachsende Rohstoffe e. V. under the funding reference 2220HV048A.

## References

1. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
2. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.
3. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.
4. D. Matthew Zeiler and F. Rob, "Visualizing and understanding convolutional neural networks." *ECCV*, 2014.
5. S. L. Rabano, M. K. Cabatuan, E. Sybingco, E. P. Dadios, and E. J. Calilung, "Common garbage classification using mobilenet," in *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*. IEEE, 2018, pp. 1–4.
6. A. Krizhevsky, G. Hinton *et al.*, "Learning multiple layers of features from tiny images," 2009.
7. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
8. S. Roy and A. Etemad, "Analysis of semi-supervised methods for facial expression recognition," *arXiv preprint arXiv:2208.00544*, 2022.
9. C. Liu, M. Zhang, Z. Fu, P. Hou, and Y. Li, "Flitext: A faster and lighter semi-supervised text classification with convolution networks," *arXiv preprint arXiv:2110.11869*, 2021.
10. P. Helber, B. Bischke, A. Dengel, and D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 7, pp. 2217–2226, 2019.
11. S. Laine and T. Aila, "Temporal ensembling for semi-supervised learning," *arXiv preprint arXiv:1610.02242*, 2016.
12. A. Tarvainen and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, vol. 30, 2017.

13. T. Miyato, S.-i. Maeda, M. Koyama, and S. Ishii, "Virtual adversarial training: a regularization method for supervised and semi-supervised learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 8, pp. 1979–1993, 2018.
14. D.-H. Lee *et al.*, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Workshop on challenges in representation learning, ICML*, vol. 3, no. 2, 2013, p. 896.
15. D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. A. Raffel, "Mixmatch: A holistic approach to semi-supervised learning," *Advances in neural information processing systems*, vol. 32, 2019.
16. D. Berthelot, N. Carlini, E. D. Cubuk, A. Kurakin, K. Sohn, H. Zhang, and C. Raffel, "Remixmatch: Semi-supervised learning with distribution alignment and augmentation anchoring," *arXiv preprint arXiv:1911.09785*, 2019.
17. K. Sohn, D. Berthelot, N. Carlini, Z. Zhang, H. Zhang, C. A. Raffel, E. D. Cubuk, A. Kurakin, and C.-L. Li, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," *Advances in neural information processing systems*, vol. 33, pp. 596–608, 2020.
18. B. Zhang, Y. Wang, W. Hou, H. Wu, J. Wang, M. Okumura, and T. Shinozaki, "Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling," *Advances in Neural Information Processing Systems*, vol. 34, pp. 18408–18419, 2021.
19. Y. Grandvalet and Y. Bengio, "Semi-supervised learning by entropy minimization," *Advances in neural information processing systems*, vol. 17, 2004.
20. M. Sajjadi, M. Javanmardi, and T. Tasdizen, "Regularization with stochastic transformations and perturbations for deep semi-supervised learning," *Advances in neural information processing systems*, vol. 29, 2016.
21. I. Loshchilov and F. Hutter, "Fixing weight decay regularization in adam," 2018.
22. H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization," *arXiv preprint arXiv:1710.09412*, 2017.
23. Q. Xie, Z. Dai, E. Hovy, T. Luong, and Q. Le, "Unsupervised data augmentation for consistency training," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6256–6268, 2020.
24. A. Krizhevsky, "Learning multiple layers of features from tiny images," Tech. Rep., 2009.



25. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
26. S. Zagoruyko and N. Komodakis, "Wide residual networks," *arXiv preprint arXiv:1605.07146*, 2016.
27. F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.
28. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
29. J. J. Senecal, J. W. Sheppard, and J. A. Shaw, "Efficient convolutional neural networks for multi-spectral image classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–8.
30. X. Zhai, A. Oliver, A. Kolesnikov, and L. Beyer, "S4l: Self-supervised semi-supervised learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1476–1485.



# Regression-based Age Prediction of Plastic Waste using Hyperspectral Imaging

Felix Kronenwett<sup>1</sup>, Pia Klingenberg<sup>2</sup>, Georg Maier<sup>1</sup>, Thomas Längle<sup>1</sup>,  
Elke Metzsch-Zilligen<sup>2</sup>, and Jürgen Beyerer<sup>1,3</sup>

<sup>1</sup> Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstraße 1, 76131 Karlsruhe, Germany

<sup>2</sup> Fraunhofer Institute for Structural Durability and System Reliability LBF, Bartningstraße 47, 64289 Darmstadt, Germany

<sup>3</sup> Vision and Fusion Laboratory (IES), Karlsruhe Institute of Technology (KIT), Haid-und-Neu-Str. 7, 76131 Karlsruhe, Germany

**Abstract** In order to enable high quality recycling of polypropylene (PP) plastic, additional classification and separation into the degree of degradation is necessary. In this study, different PP plastic samples were produced and degraded by multiple extrusion and thermal treatment. Using near infrared spectroscopy, the samples were examined and regression models were trained to predict the degree of aging. The models of the multiple extruded samples showed high accuracy, despite only minor spectral changes. The accuracy of the models of the thermally aged samples varied with the design of the training set due to the non-linear aging process, but showed sufficient accuracy in prediction.

**Keywords** Hyperspectral imaging, Plastic waste, Multiple Extrusion, Thermal aging, Regression, Sensor-based sorting

## 1 Introduction

With their versatile applications, plastics are indispensable for a high living standard in all areas of life, be it hygiene, lightweight construction and transport, food supply or technology [1, 2]. The plastic production worldwide amounts to 390 mio. t (2021) and in Germany alone, around 12 mio. t are consumed every year [3]. This causes massive

plastic waste streams, which are currently mainly disposed of through energy recovery in Europe and by landfill in most other regions of the world [4, 5]. However, so-called *end-of-life*-plastics are an important resource both for the plastic industry through mechanical recycling and the chemical industry through chemical recycling, yielding recycled plastic materials and platform chemicals and monomers respectively [6,7]. To underline their economical and environmental potential, plastic waste streams are referred to as *secondary raw materials* [8]. Special focus needs to be laid on the recycling of *post-consumer* secondary raw materials, which are plastics which have undergone their service-life once, as opposed to *pre-consumer*- or *post-industrial* materials, as the recycling rates of the former are very low [3,4,9].

For plastics recycling, particularly mechanical recycling, the quality of the resulting recyclate strongly depends on the characteristics of the input stream. The material homogeneity is therefore an important prerequisite for the input stream. To achieve this, the input stream is pre-processed and sorted in multiple stages, where sensor-based sorting plays a crucial role. The umbrella term *sensor-based sorting* describes a family of systems that enable the physical separation of individual particles from a material stream on the basis of information acquired by one or multiple sensors. A particular strength of the technology is its flexibility in terms of the criteria according to which sorting can be performed. This flexibility exists due to the variety of eligible sensor principles as well as the freely programmable data evaluation.

## 1.1 Contribution

During their service life, plastics undergo an aging process, inducing changes in the material's chemical and physical properties and potentially compromising its quality [10]. There are multiple factors which cause degradation effects during processing and service life such as thermo-mechanical stress during processing, causing chain scission and/or cross linking, exposure to UV-radiation, humidity, high temperatures or other weathering conditions, causing (thermo-)oxidative degradation [8, 11]. The mechanism of the oxidative degradation of polymers is referred to as autoxidation [12]. In the case of polypropylene (PP), autoxidation occurs after an induction period, accelerating the degradation exponentially [13]. Metal impurities from catalyst

residues may accelerate this process still further [14]. To counteract material degradation and to compensate a negative influence by aged polymers, stabilizers, compatibilizers and other additives are used [15]. Detailed knowledge of the degree of degradation of a secondary raw material stream is therefore highly useful for determining and adjusting the composition and concentration of the master batch in question, thereby improving the recycling of mixed materials with varying degrees of degradation.

In this study, a virgin PP homo-polymer has undergone two separate accelerated aging experiments. The first has been a recycling simulation by multiple processing and the second a service life simulation using an oven and thermo-oxidative conditions. The test specimen were injection-moulded and analyzed using NIR spectroscopy. Regression models were trained using NIR spectra to model the aging stage and predict the degree of degradation of unknown samples.

## 1.2 Related Work

Existing work has demonstrated the general suitability of NIR spectroscopy for age prediction of plastic samples. In [16], different types of plastics (virgin polymers) were investigated and regression models were trained using NIR spectra to predict the polymer degradation and a polymer quality assessment of the samples, caused by controlled, laboratory thermal aging. It showed the general suitability of NIR spectroscopy for determining polymer degradation, however accuracy depends on the type of plastic. Acrylonitrile butadiene styrene (ABS) and polyethylene terephthalate (PET) proved to be particularly suitable, while low-density polyethylene (LDPE) and PP were more difficult to evaluate. The chemical stability of polyethylene (PE) and PP was named as the cause. In [17], the investigations were extended to include the prediction of the extrusion cycles, which also showed differences in accuracy depending on the type of plastic. It was recommended to include more data in the model generation. Specifically, the prediction of the age of thermally treated PP samples was the subject of [18], with focus on the chemical modification of the polymer structure. In [19], the investigations were extended to plastic waste degraded under natural circumstances.

## 2 Materials and Methods

In the following, the production of the PP plastic samples is outlined. Subsequently, the data acquisition and the calculation of the regression models for the prediction of the aging stage are described.

### 2.1 Accelerated aging of test specimen

A PP homo-polymer (Moplen HP 500N, LyondellBasell, Rotterdam, Netherlands) in granular form was used as raw material for the accelerated aging experiments. Multiple processing was performed using a twin-screw extruder (Thermo Scientific™HAAKE™Rheomex PTW 16, Thermo Fisher, Waltham, Massachusetts, US) with a processing temperature range of 185 - 236 °C and 200 rpm. The extrusion process was repeated five times. From each extrusion cycle, a quantity was used for the preparation of test specimen (plates, 80 x 80 x 2.5 mm). Test specimen for further analysis were produced using an injection moulding system (Allrounder 320 C, Arburg, Loßburg, Germany). For the thermo-oxidative aging, test specimen were injection moulded immediately from the raw material using the above mentioned injection moulding system and conditions. The plates were placed in an aging furnace (Mettmert Universalschrank UF75, Mettmert, Büchenbach, Germany) at 150 °C and 100% ventilation. An overview can be found in Table 1.

**Table 1:** Overview of the two datasets consisting of differently aged PP samples.

|                       | Dataset A       | Dataset B                 |
|-----------------------|-----------------|---------------------------|
| Plastic type          | PP              | PP                        |
| Material              | Moplen HP 500N  | Moplen HP 500N            |
| Treatment             | extrusion       | thermal                   |
| Aging state parameter | 1, 3, 5 (times) | 10, 22, 27, 30, 34 (days) |
| Number of samples     | 3x10            | 5x10                      |

### 2.2 Data acquisition

Due to the possibility to distinguish different types of plastics, the use of hyperspectral cameras in the near-infrared (NIR) wavelength range

is widespread within the sensor-based sorting industry [20]. Based on the chemical molecules present, or specifically their functional groups, different types of plastics have individual absorption characteristics and therefore show distinct spectra in the NIR wavelength range. On an experimental level, the sensor technology has also been used to investigate different characteristics, e. g., aging states of plastic. However, the use of NIR spectra for plastic age prediction is limited due to several possible properties. Regression on the basis of NIR spectra is an inverse problem, i. e., the exact composition of the sample cannot be derived from the spectral information. One problem is the overlap of the absorption bands [21,22].

For this study, the specimen were recorded using a hyperspectral NIR line-scan camera in the wavelength range of 900 – 1700 nm. The camera model is FX17 from Specim, consisting of a spatial resolution of 640 pixels. Per pixel, 256 spectral bands were acquired, resulting in a spectral resolution of slightly more than 3 nm. Due to different reflection properties caused by surface characteristics and camera position, variations occur in the raw spectra falling through the camera apparatus and captured by the sensor. These so-called scatter effects are minimized with the help of pre-processing steps.

First, the output of the hyperspectral sensor, which can be interpreted as the spectral reflectance, was converted to absorption units  $a = \log(1/R)$ . The wavelength range was then cropped to avoid unwanted edge effects. To minimize scattering effects, the Signal Normal Variant (SNV) was applied. The mean value of each spectrum is subtracted and then divided by its standard deviation.

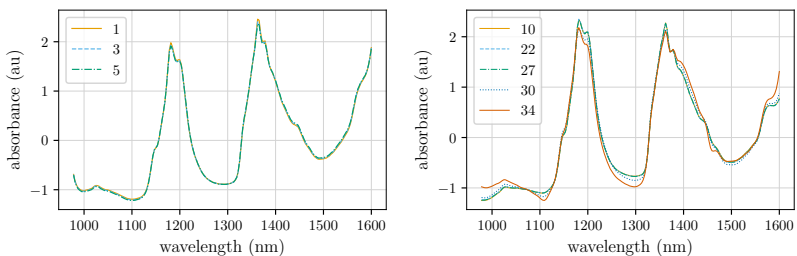
### 2.3 Evaluation of the NIR spectra of aged PP samples

For each image, the foreground pixels were segmented and an average absorption spectrum was calculated from all spectra within the sample mask. This turned out to be a relevant measure to suppress noise effects and to better highlight the small spectral changes. The mean NIR absorption spectra within a degradation stage are shown in Figure 1. Clearly visible absorption bands of the NIR spectrum are associated with  $CH_2$  and  $CH_3$  groups of the PP molecules. In the range between 1100 and 1225 nm as well as 1350 to 1450 nm, absorption bands of the second overtone region of the methylene and methyl group or the

respective combination vibrations with  $CH$  groups are located. Absorption bands of the  $CH_3$  groups are located at lower wavelengths (1195 nm, 1360 nm) compared with  $CH_2$  absorption bands (1215 nm, 1395 nm) [23]. Due to the spectral proximity, there is a strong overlap of the absorption bands.

When looking at the samples that have been extruded several times, a decrease in the intensity of the absorption bands associated with  $CH_2$  and  $CH_3$  can be observed. A linear relationship between spectral changes and the number of extrusion cycles can be assumed. The observations can be explained by the increasing degradation of the polymer chains per extrusion cycle.

The observation of the spectra of the thermally aged PP samples show a similar course, but clear differences are recognizable. The thermally aged samples clearly show inhomogeneous degradation behavior related to the spatial area, visible as spots on the surface. The extracted local NIR spectra of a sample therefore show different aging stages depending on the spatial pixel position. With increasing thermal age, the intensity of the  $CH_3$  and  $CH_2$  absorption bands decreases. The behavior is clearly non-linear and can rather be modeled as an exponential relationship. Furthermore, stabilizing additives prevent chain scission at the beginning of aging. Once the additives are consumed, the aging process takes its exponential course. The start of the exponential aging process therefore has an induction period.



**Figure 1:** Mean absorption spectra of multiple extruded PP samples (1-, 3- and 5-fold extruded) after SNV (left) and mean absorption spectra of thermally aged PP samples (10, 22, 27, 30, 34 days) after SNV (right).



## 2.4 Regression-based age prediction

Linear regression models were trained to predict the degree of degradation of the PP samples based on the NIR absorption spectra. For this purpose, Partial Least Squares (PLS) Regression was used. The algorithm is based on the assumption of a linear relationship  $y = Xb$  between the input data  $X$  (spectral data) and the target values  $y$  (aging time or extrusion cycles). Even though this is not the case, especially for the thermally aged samples, its application in hyperspectral data evaluation has nevertheless proved successful and showed good results even for non-linear datasets [24]. The algorithm projects the data into a space with a smaller dimension, depending on the number of latent variables (LV) defined manually beforehand. The ability to model complex relationships increases with the number of LVs, but runs the risk of overfitting. The selection of the parameter is therefore crucial. When calculating the regression model, the number of LVs must be specified. This largely determines the ability of the model to adapt to complex data. In order to obtain a highly generalizing model using only a small amount of training data, a trade-off in the training stage is necessary. To determine the number, Leave-One-Out Cross-Validation was used. In each run, one partition is used as the test set and one model is trained with the remaining partitions. A metric is calculated for each model and then averaged over the metric values to obtain an overall assessment of the suitability of the parameterization of the model. This is done for a given number of LVs, and then the number of the best, most generalized model is chosen.

### Extrusion cycle prediction model

To calculate the PLS regression for Dataset A, 10 single-extruded and 10 five-extruded samples were used for training. The remaining 10 triple-extruded samples formed the independent test set. The optimization of the numbers of LVs resulted in a number of 5, this value was later used for calculation of the PLS model.

### Thermal age prediction model

The investigations were divided into two parts, both using Dataset B. First, it was analyzed whether linear regression is suitable to model the nonlinear aging process by using only a few target values. For this purpose, the samples with aging stages 10, 27 and 34 (days) were used for training. The calculated model (Model 1) was evaluated using test data obtained from the samples with aging stages 22 and 30 (days). For the model calculation, a LV number of 8 was used after optimization.

In a second study, all 5 aging stages were used for model training. For this purpose, 5 samples per aging stage were selected for model training and 5 samples each were used for the test set. Thus, the total number of spectra used for model training was reduced compared to the first study, but included a wider range of target values. The model (Model 2) was calculated using a number of 8 LVs.

### Evaluation metrics

As a metric to evaluate the regression model, the Root Means Squared Error (RMSE) and R2 score is used. The RMSE score

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

estimates the standard deviation of the prediction of a regression model. Here,  $\hat{y}_i$  describes the prediction result and  $y_i$  the ground truth value. A distinction can be made between the RMSE of the calibration set (training) and the prediction set (test). In addition, the  $R^2$  score

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

indicates how well the independent variables are suited to explain the variance of the dependent variables, where  $n$  is the number of samples.

## 3 Experimental Results

The performance of the regression models for predicting the age of PP plastics is examined below. A distinction is made between thermal aging and aging by multiple extrusion.

### 3.1 Extrusion cycle prediction results

The performance of the model was analyzed by calculating the RMSE and  $R^2$  of the test set. Both values are depicted together with the exact structure of the training and test set in Table 2. The model achieved an RMSE of 0.367 on the independent test data of the aging stage not yet considered during training. Figure 2 shows the model-predicted values plotted against the real values. The results show a general suitability of the model for the estimation of extrusion cycles. The calculated RMSE of the training data of 0.118 shows similarity to the obtained value in the test data. In addition, the calculation of the median of the estimated aging states of the test data ( $\tilde{y}_{pred} = 3.052$ ) shows that the results scatter around the target value. The data show a linear correlation between the target value and the spectral information. Therefore, the linear PLS model can model the correlation with high accuracy using only two aging stages during training. During model calculation, it has been shown that the main focus must be on the generation of the training data and its pre-processing. Only the calculation of mean value spectra makes it possible to visualize the small change in the absorption spectrum with respect to noise influences. Thus, multiple extrusion leads only to a small change in the functional groups.

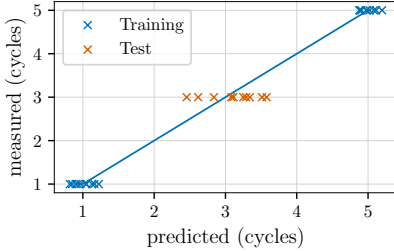
**Table 2:** Performance of the regression models on a respective independent test set for the prediction of the thermal aging stage resp. the number of extrusion cycles.

|                    | Train              | Test A             | LV | RMSE  | $R^2$ |
|--------------------|--------------------|--------------------|----|-------|-------|
| Dataset A          | 1, 5               | 3                  | 5  | 0.367 | -     |
| Dataset B, Model 1 | 10, 27, 34         | 22, 30             | 8  | 2.158 | 0.709 |
| Dataset B, Model 2 | 10, 22, 27, 30, 34 | 10, 22, 27, 30, 34 | 8  | 1.437 | 0.970 |

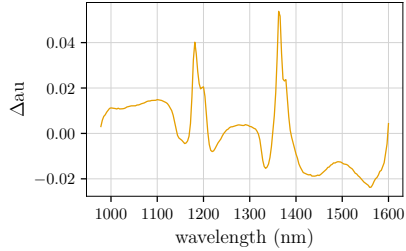
### 3.2 Thermal age prediction results

The age-prediction models of PP were assessed by calculating the RMSE and  $R^2$  of the test set. Both values are depicted together with the exact structure of the training and test set in Table 2. Figure 4 shows the model-predicted values plotted against the real values.

The evaluation of the thermally aged PP samples resulted in the calculation of two models, each based on different training data or



**Figure 2:** Results of the regression model for predicting the number of extrusion cycles. Measured versus predicted number of cycles.

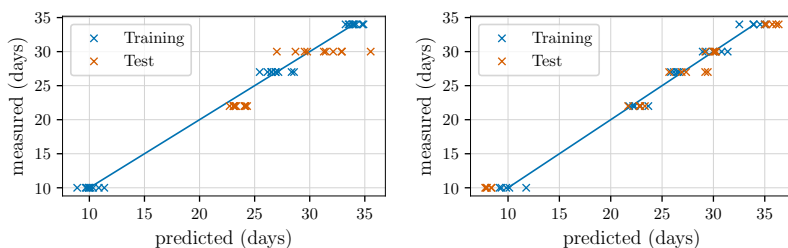


**Figure 3:** Difference of the mean NIR absorption spectra of all 1-fold and the 5-fold extruded PP samples used for model training.

different aging stages. The analysis of the spectra already showed a nonlinear course of aging. The first model, calculated from only three aging stages, achieved an RMSE of 2.158 on the test data. The scatter of the estimated aging highlights the problem of modeling the nonlinear aging process using a few target values. Prediction of the 22 days aged samples was consistently overestimated, illustrated by the median  $\tilde{y}_{pred,22} = 23.292$ . In contrast, the 30 days aged samples were only slightly overestimated on average ( $\tilde{y}_{pred,30} = 31.367$ ), but the values strongly scatter ( $\sigma_{y_{pred,30}} = 2.324$ ). The RMSE of the training data of 0.696 is also significantly lower than the RMSE of the independent test data. In addition to the nonlinear aging process, the tests also confirmed a delayed start of the aging process by admixed additives.

For the second regression model, the training set was adapted by including all 5 aging stages. The test set resulted in an RMSE of 1.437. The RMSE of the training data of 0.857 is similarly low. In addition, comparison of the medians of the test and training sets shows a uniform spread of the estimated target values around the real ones.

The comparison of both models showed that more aging stages in the training set are more important to model the nonlinear course than the absolute number of training spectra. Furthermore, it was shown that despite local differences in the aging stages within a sample, the mean spectra is suitable to represent the aging time of the entire sample.



**Figure 4:** Results of the regression models of thermally aged PP samples, measured versus predicted days. Model 1 (left) and Model 2 (right).

## 4 Conclusion and Future Work

The investigations showed the general suitability of NIR spectroscopy for the prediction of different aging and degradation stages of PP plastic. Thermally aged as well as multiple extruded PP samples were investigated. Different regression models were calculated to estimate the duration of thermal aging or the number of extrusion passes. Special attention was paid to the pre-processing and spectral averaging of the NIR spectra in order to make small spectral differences visible. The calculated regression models showed a correlation between aging condition and spectral information. The exponential progression of thermally aged samples must be modeled sufficiently well. More target values in model training greatly improves the generalizability of the model. One challenge is the inhomogeneous aging visible on the spatial area of the samples and therefore impacting the spectra, which can be investigated in further studies.

## Acknowledgment

This work was supported as a Fraunhofer LIGHTHOUSE PROJECT.

## References

1. R. Mülhaupt, "Green polymer chemistry and bio-based plastics: dreams and reality," *Macromolecular Chemistry and Physics*, vol. 214, no. 2, pp. 159–

174, 2013.

2. C. G. Schirmeister and R. Mülhaupt, "Closing the carbon loop in the circular plastics economy," *Macromolecular Rapid Communications*, vol. 43, no. 13, p. 2200247, 2022.
3. P. Europe, "Plastics - the facts 2022," <https://plasticseurope.org/knowledge-hub/plastics-the-facts-2022/>, 2022.
4. R. Geyer, J. R. Jambeck, and K. L. Law, "Production, use, and fate of all plastics ever made," *Science advances*, vol. 3, no. 7, p. e1700782, 2017.
5. R. Meys, F. Frick, S. Westhues, A. Sternberg, J. Klankermayer, and A. Bardow, "Towards a circular economy for plastic packaging wastes—the environmental potential of chemical recycling," *Resources, Conservation and Recycling*, vol. 162, p. 105010, 2020.
6. S. Al-Salem, P. Lettieri, and J. Baeyens, "Recycling and recovery routes of plastic solid waste (psw): A review," *Waste management*, vol. 29, no. 10, pp. 2625–2643, 2009.
7. S. Yin, R. Tuladhar, F. Shi, R. A. Shanks, M. Combe, and T. Collister, "Mechanical reprocessing of polyolefin waste: A review," *Polymer Engineering & Science*, vol. 55, no. 12, pp. 2899–2909, 2015.
8. F. Vilaplana and S. Karlsson, "Quality concepts for the improved use of recycled polymeric materials: a review," *Macromolecular Materials and Engineering*, vol. 293, no. 4, pp. 274–297, 2008.
9. M. K. Eriksen, A. Damgaard, A. Boldrin, and T. F. Astrup, "Quality assessment and circularity potential of recovery systems for household plastic waste," *Journal of Industrial Ecology*, vol. 23, no. 1, pp. 156–168, 2019.
10. A. Jansson, K. Möller, and T. Gevert, "Degradation of post-consumer polypropylene materials exposed to simulated recycling—mechanical properties," *Polymer Degradation and Stability*, vol. 82, no. 1, pp. 37–46, 2003.
11. E. Strömberg and S. Karlsson, "The design of a test protocol to model the degradation of polyolefins during recycling and service life," *Journal of Applied Polymer Science*, vol. 112, no. 3, pp. 1835–1844, 2009.
12. L. M. Smith, H. M. Aitken, and M. L. Coote, "The fate of the peroxy radical in autoxidation: how does polymer degradation really occur?" *Accounts of chemical research*, vol. 51, no. 9, pp. 2006–2013, 2018.
13. A. Jansson, K. Möller, and T. Hjertberg, "Chemical degradation of a polypropylene material exposed to simulated recycling," *Polymer degradation and stability*, vol. 84, no. 2, pp. 227–232, 2004.

14. G. W. Ehrenstein and S. Pongratz, *Beständigkeit von Kunststoffen*. Hanser München, 2007.
15. R. Pfaendner, H. Herbst, K. Hoffmann, and F. Sitek, "Recycling and restabilization of polymers for high quality applications. an overview," *Die Angewandte Makromolekulare Chemie: Applied Macromolecular Chemistry and Physics*, vol. 232, no. 1, pp. 193–227, 1995.
16. A. Alassali, S. Fiore, and K. Kuchta, "Assessment of plastic waste materials degradation through near infrared spectroscopy," *Waste management*, vol. 82, pp. 71–81, 2018.
17. A. Alassali, C. Picuno, T. Bébien, S. Fiore, and K. Kuchta, "Validation of near infrared spectroscopy as an age-prediction method for plastics," *Resources, Conservation and Recycling*, vol. 154, p. 104555, Mar. 2020.
18. H. Shinzawa, R. Watanabe, S. Yamane, M. Koga, H. Hagihara, and J. Mizukado, "Aging of polypropylene probed by near infrared spectroscopy," vol. 29, no. 5, pp. 259–268.
19. X. Chen, N. Kroell, T. Dietl, A. Feil, and K. Greiff, "Influence of long-term natural degradation processes on near-infrared spectra and sorting of post-consumer plastics," *Waste Management*, vol. 136, pp. 213–218, 2021.
20. D. J. da Silva and H. Wiebeck, "Current options for characterizing, sorting, and recycling polymeric waste," *Progress in Rubber, Plastics and Recycling Technology*, vol. 36, no. 4, pp. 284–303, 2020.
21. M. Vidal, A. Gowen, and J. M. Amigo, "Nir hyperspectral imaging for plastics classification," *NIR news*, vol. 23, no. 1, pp. 13–15, 2012.
22. Q. Duan and J. Li, "Classification of common household plastic wastes combining multiple methods based on near-infrared spectroscopy," *ACS ES&T Engineering*, vol. 1, no. 7, pp. 1065–1073, 2021.
23. J. Gulmine, P. Janissek, H. Heise, and L. Akcelrud, "Polyethylene characterization by ftir," *Polymer testing*, vol. 21, no. 5, pp. 557–563, 2002.
24. M. Blanco, J. Coello, H. Iturriaga, S. MasPOCH, and J. Pages, "Nir calibration in non-linear systems: different pls approaches and artificial neural networks," *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 1, pp. 75–82, 2000.





# Method Development for Spatially Resolved Detection of Adulterated Minced Meat

Ervienatasia Djaw<sup>1</sup>, Isik Türkmen<sup>1</sup>, Thorsten Tybussek<sup>1</sup>, and Tilman Sauerwald<sup>1,2</sup>

<sup>1</sup> Fraunhofer Institute of Process Engineering and Packaging IVV,  
85354 Freising, Germany

<sup>2</sup> Saarland University, Department Systems Engineering  
66123 Saarbrücken, Germany

**Abstract** This study explored the possibility of detecting different types of meat in a miniaturized patty by applying a random forest classifier on the spectral dimension followed by neighborhood majority voting on the spatial dimension to improve the random forest prediction. Hyperspectral images of patties made of 100% beef, 100% pork, and 100% horse meat were acquired with a short-wave infrared (SWIR) hyperspectral camera. The pixel-wise meat type prediction by random forest multi-class classifier was accurate to 97.5%. After the majority voting of the neighboring pixels, the prediction accuracy increased to 100%. As next, synthetic hyperspectral images of adulterated patties were generated for validating the model. The prediction accuracy of the model on the synthetic images were bigger than 98%. The findings of the proposed workflow support the development of rapid analysis tools in tandem with machine-learning to detect adulteration in minced meat.

**Keywords** Hyperspectral imaging, random forest, majority voting, food safety, adulteration, authenticity

## 1 Introduction

Meat is known for its commercial and nutritional values, yet it is prone to fraudulent and accidental adulteration which violates consumers' safety and protection [1–3]. Besides falsification of meat by other materials than the declared ingredients (e.g. beef/offal), the proportion

of ingredients or the main components (e.g. meat muscles vs fat) may deviate from the stated composition [2, 4, 5]. The DNA-based analysis is the golden standard of authenticating the meat species and their origin, but it's a time-consuming method [3].

Most of the past studies utilized hyperspectral imaging (HSI) in the visible and near-infrared region (VNIR) (450 to 1000 nm) in tandem with chemometrics and artificial intelligence with promising outcomes. Both minced meat and meat cuts can be authenticated via these tools by examining either the whole composition or only the fatty acids profiles [4–9]. However, the spatial information was often left out due to the complexity of the data dimension, and the prediction models were often trained by averaged spectra [4, 6, 8–10]. Ropodi et al. demonstrated the application of multi-spectral imaging in the visible region using 16 spectral features with help of the support vector machine (SVM) giving 93.5% accuracy in detecting horse meat in beef minced meat. The authors also reported that the color-change during storage had a negative influence on the prediction results [6]. Jiang et al. used HSI in the VNIR range coupled with pixel-wise partial least square regression (PLSR) to quantify duck in beef minced meat. The PLSR model was trained by average spectra of patties with different levels of adulteration. Afterwards, the pixel-wise regression was applied in the spatial domain to generate adulteration heat maps [8].

This paper explored the feasibility of detecting different meat species in a patty by using a hyperspectral camera in the short-wave infrared (SWIR) region between 930 to 2500 nm in tandem with a pixel-wise random forest (RF) multi-class classifier, followed by neighborhood majority voting on every pixel across the 2D spatial dimension. The trained RF classifier aimed to classify every pixel into one of three classes as beef, horse or pork, regardless of the meat's freshness level. The neighborhood majority voting was applied subsequently on spatial dimension to improve the pixel-wise classification.

## **2 Materials and methods**

### **2.1 Meat Sample Preparation and Training Datasets**

Minced meat of 100% pork, 100% beef, and 100% horse were purchased from local butchers in Munich, Germany. A patty with ca.10 g of each

meat type was placed on a sterile Petri dish and measured on the purchase day (Day 0) and five days after the purchase day (Day 5). Between Day 0 and Day 5, meat was stored in the fridge at  $T = 6 \pm 2^\circ\text{C}$ . Patties containing different meat types were not used in this study to avoid the uncertainty in the ground truth image pixel labels of those mixtures. Instead, synthetic patties were generated to validate the model. The process of generating synthetic patty is elaborated in section 2.4.

## **2.2 SWIR hyperspectral imaging system and data acquisition**

The SWIR spectra in the region (930 - 2500 nm ) were captured using HySpex SWIR 384 SN 3197 (Norsk Elektro Optikk AS, Oslo, Norway) with a 5.45 nm sampling interval which delivers 288 data points per spectrum. The camera was equipped with 1m objective with ca. 84 cm distance between the objective and the sample's surface, resulting in an image resolution of 0.33 mm/px with 32 bit color depth. The samples on the translating stage were exposed to two halogen light sources mounted at a symmetrical angle. The reflection spectra were recorded by the push broom method at an acquisition rate of 33800  $\mu\text{s}$  per spectral line.

## **2.3 Radiometric Correction and Initial Pre-processing**

A radiometric correction was applied to all images using the software HyRad (Norsk Elektro Optikk AS, Oslo, Norway), which adjusted each spectrum based on the reflection of a white reference. The subsequent data preprocessing explained below was performed using the Python 3.9.12 programming language.

Initially the saturated spectral values of a given pixel were replaced by the nearest pixel's unsaturated spectral values or by the averaged spectrum of the surrounding unsaturated pixels [11,12]. Then the region of interest (ROI) was extracted by removing the irrelevant image sections, such as background, sampling stage, and Petri dish. The ROI extraction process utilized Gaussian blurring filter with a kernel size of (4x4) and 0.5 standard deviations on the grayscale image obtained from the first spectral feature (930 nm) followed by the automatic Otsu thresholding method to create a mask [13,14]. Finally, all spectra within

the mask were extracted and scaled using the 'Standard Scaler' function from scikit-learn python library.

## 2.4 Random Forest Classification and Dataset

A random forest (RF) multi-class classifier with 100 trees, 'entropy' as the criterion for node-splitting and 20 as the tree's maximum depth, was trained using all spectral features (288 features) in 3 cross-validations. A balanced amount of data across three meat categories were ensured in the training data set. There were 43200 data points from meat measured on Day 0 and 28800 data points from meat measured on Day 5. Not all data points were used for training; the unused data points were set aside to generate synthetic hypercubes in validation stage.

**Pixel-Wise Prediction & Majority Class Of The Neighboring Pixels.** Every pixel was classified into one of three classes (beef, horse, or pork) by the trained random forest classifier. Consecutively, each prediction result was evaluated spatially by comparing it to the majority class from its surrounding pixels (kernel size 3x3). In case of a class mismatch between the observed pixel and the majority class within the neighbors, the RF prediction probability for all classes of the observed pixel were replaced by the averaged probabilities of its surrounding pixels.

**Synthetic Patties For Validation.** Synthetic patties (50x50 px) with segmented regions in various shapes, sizes, and grey levels were generated automatically using the function „random\_shapes“ from the scikit-image python library. Every shape and the background were assigned to a particular class based on its grey level (Figure 5). Sequentially, each pixel was filled with a random spectrum belonging to the assigned class hence generating a hypercube.

## 3 Results and discussion

As seen in Figure 1, no difference can be observed by the naked eye either between the spectra of different meat types or between fresh and old. Nevertheless, the classification model in this study focused on differentiating the meat types, not the freshness level of the meat. There-

fore, the experiment aimed to generalize 100% beef patty regardless of the mixture of fresh or old beef as a beef patty.

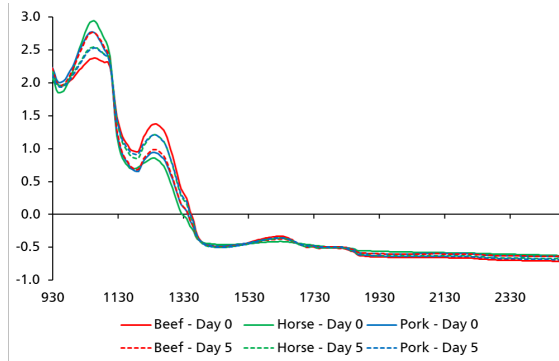


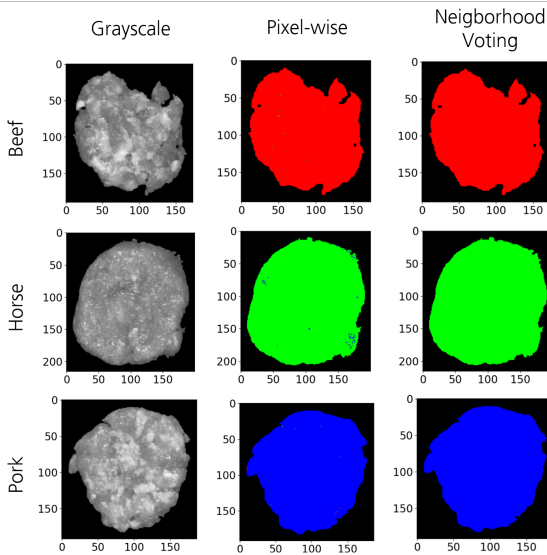
Figure 1: Average spectra of all patties.

|                 |            |       |       |       |
|-----------------|------------|-------|-------|-------|
| Predicted Class | Beef       | 23111 | 703   | 186   |
|                 | Horse      | 666   | 23167 | 167   |
|                 | Pork       | 116   | 124   | 23760 |
|                 | Beef       | Horse | Pork  |       |
|                 | True Class |       |       |       |

Figure 2: Confusion matrix from pixel-wise RF multi-class classifier.

The pixel-wise RF classification gave an accuracy of 97.3%, where 'pork' has the highest precision, recall, and f1-score values (each 99%), followed by 'horse' (each 97%) and 'beef' with 96% recall and 97% of each precision and f1-score. A closer look at the confusion matrix in Figure 2 shows a higher number of falsely predicted 'beef' as 'horse' and vice versa. The mis-classifications from pixel-wise RF classifier were more apparent to occur on single pixels than in a region (Figures 3 and 4, pixel-wise images).

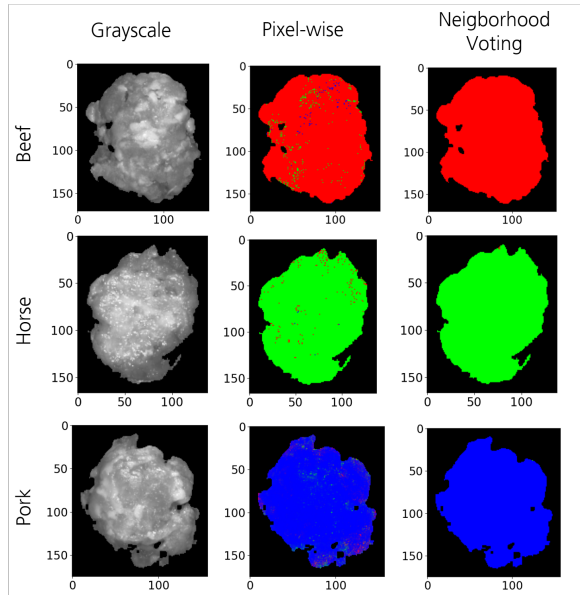
The falsely predicted pixels by pixel-wise RF classifier were corrected by comparing each pixel with its neighbors (majority voting; 3x3 kernel; see 2.4). The significant improvement can be observed in fresh (Figure 3) and five days old patties (Figure 4), comparing the images in



**Figure 3:** Fresh Meat (Day 0) Classification; Left: grayscale images at 1115 nm; Center: Pixel-wise classification results; Right: Pixel-wise classification results after neighborhood majority voting and probability values correction. Each pixel was colored based on the predicted class: red refers to 'beef', green refers to 'horse', and blue refers to 'pork'.

the middle (pixel-wise) to the images on the right (neighborhood majority voting). The success of the model is remarkably dependent on the correlation between the camera's spatial resolution, the accuracy of pixel-wise prediction, and the kernel size used in neighborhood majority voting.

The spectra at 1115, 930, and 1250 nm respectively appeared to be the most important features observed by RF. The inclusion of these features led to the biggest decrease of a tree's impurity in RF model [15]. These regions refer to the 2nd and 3rd overtone regions of C-H molecular group, except at 930 nm where O-H and C-H are overlap [16]. These findings indicate that a prediction model can be built using only these spectral features, which is to be explored further. For instance, the fat region seems to be in the highest contrast after observing the gray

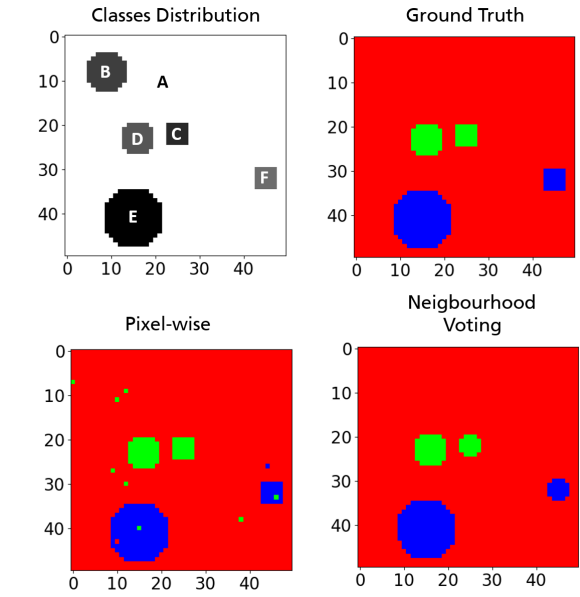


**Figure 4:** Old Meat (Day 5) Classification; Left: grayscale images at 1115 nm; Center: Pixel-wise classification results; Right: Pixel-wise classification results after neighborhood majority voting and probability values correction.

scale images at 1115 nm (see the pictures on the left in figure 3 and 4). Besides, a study by Lestari et al. demonstrated an improved prediction by using 1D FTIR on the extracted fat from meatballs in detecting rats in beef meatballs [17].

A comparison between patties from day 0 and day 5 shows that false predictions occurred more often on patties from day 5 (Figure 4, middle images) than day 0 (Figure 3, middle images), as previously stated by Ropodi et al [6]. However, in our case, this could also be due to fewer spectra collected for old patties (day 5) than fresh patties (day 0).

The validation of the complete workflow on synthetic patties showed promising results. The falsely classified pixels were mostly corrected by neighborhood majority voting. The shape of the kernel, which was square altered the shape of regions containing edges, as depicted on the synthetic patty images in figure 5.



**Figure 5:** Synthetic Patties with 90.4% beef (A + B or "Red" area), 2.8% horse (C + D or "Green" area), and 6.8% pork (E + F or "Blue" area) of which 2.8% old beef (B), 1.0% old horse (C), and 5.8% old pork (F).

Furthermore, figure 5 also validates the model's generalization, regardless of the freshness level. The old beef spectra (B) were mostly falsely predicted as horse and some of the old pork (E) were predicted as horse or beef.

## 4 Conclusions and Outlook

Random forest multi-class classification on the spectral dimension followed by neighborhood majority voting in the spatial dimension showed promising results to authenticate minced meat of different types (beef, horse, and pork). The prediction by pixel-wise RF classifier based solely on spectral dimension was accurate to 97.5%. After introducing the majority voting of the neighboring pixels in the spatial



dimension, the prediction accuracy increased to 100%.

The findings of this study can be used to develop rapid analysis tools for minced meat authentication. Furthermore, a prior image processing on the grayscale image to separate high-fat from low-fat regions may also provide an alternative approach, which is to be explored in detail as next.

## 5 Acknowledgement

This research work was supported by the Leistungszentrum Sichere intelligente Systeme (LZSiS) - Fraunhofer Society of Germany.

## References

1. S. Tähtkää, R. Majjala, H. Korkeala, and M. Nevas, "Patterns of food frauds and adulterations reported in the eu rapid alert system for food and feed and in finland," *Food Control*, vol. 47, pp. 175–184, 2015.
2. K. Robson, M. Dean, S. Brooks, S. Haughey, and C. Elliott, "A 20 year analysis of reported food fraud in the global beef supply chain," *Food Control*, vol. 116, p. 107310, 2020.
3. Y.-C. Li, S.-Y. Liu, F.-B. Meng, D.-Y. Liu, Y. Zhang, W. Wang, and J.-M. Zhang, "Comparative review and the recent progress in detection technologies of meat product adulteration," *Comprehensive Reviews in Food Science and Food Safety*, vol. 19, no. 4, pp. 2256–2296, 2020.
4. M. Kamruzzaman, M. Haque, and M. Ali, "Hyperspectral imaging technique for offal quantification in minced meat," *Journal of the Bangladesh Agricultural University*, vol. 12, no. 1, pp. 189–194, 2014.
5. H. Jiang, F. Cheng, and M. Shi, "Rapid identification and visualization of jowl meat adulteration in pork using hyperspectral imaging," *Foods*, vol. 9, no. 2, p. 154, 2020.
6. A. I. Ropodi, E. Z. Panagou, and G.-J. E. Nychas, "Multispectral imaging (msi): A promising method for the detection of minced beef adulteration with horsemeat," *Food Control*, vol. 73, pp. 57–63, 2017, special issues from the '9th EFFoST International Conference.
7. J. Mendez, L. Mendoza, J. Cruz-Tirado, R. Quevedo, and R. Siche, "Trends in application of nir and hyperspectral imaging for food authentication," *Scientia Agropecuaria*, vol. 10, no. 1, pp. 143–161, 2019.

8. H. Jiang, W. Wang, H. Zhuang, S.-C. Yoon, Y. Yang, and X. Zhao, "Hyperspectral imaging for a rapid detection and visualization of duck meat adulteration in beef," *Food Analytical Methods*, vol. 12, no. 10, pp. 2205–2215, 2019.
9. M. Kamruzzaman, Y. Makino, and S. Oshita, "Hyperspectral imaging in tandem with multivariate analysis and image processing for non-invasive detection and visualization of pork adulteration in minced beef," *Analytical Methods*, vol. 7, no. 18, pp. 7496–7502, 2015.
10. H. Jiang, Y. Yang, and M. Shi, "Chemometrics in tandem with hyperspectral imaging for detecting authentication of raw and cooked mutton rolls," *Foods*, vol. 10, no. 9, p. 2127, 2021.
11. J. Burger and P. Geladi, "Hyperspectral nir image regression part ii: dataset preprocessing diagnostics," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 20, no. 3-4, pp. 106–119, 2006.
12. Fraunhofer-Allianz-Vision, "Leitfaden zur hyperspektralen bildverarbeitung," 2019.
13. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
14. J. Yousefi, "Image binarization using otsu thresholding algorithm," *Ontario, Canada: University of Guelph*, 2011.
15. L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
16. L. Bokobza, *Origin of near-infrared absorption bands*. Wiley-VCH, 2002.
17. D. Lestari, A. Rohman, S. Syofyan, N. D. Yuliana, N. K. B. A. Bakar, and D. Hamidi, "Analysis of beef meatballs with rat meat adulteration using fourier transform infrared (ftir) spectroscopy in combination with chemometrics," *International Journal of Food Properties*, vol. 25, no. 1, pp. 1446–1457, 2022.

# Multispectral analysis for the determination of lycopene concentration in tomatoes

Marcel Mlynarik<sup>1</sup>, Gary A. Atkinson<sup>1</sup>, Melvyn L. Smith<sup>1</sup>, and Khemraj Emrith<sup>2</sup>

<sup>1</sup> Centre for Machine Vision, Bristol Robotics Laboratory, University of the West of England Bristol, BS16 1QY UK

<sup>2</sup> Mohamed bin Zayed University of Artificial Intelligence, Masdar City, Abu Dhabi UAE

**Abstract** This paper describes a novel computer vision method for the estimation of lycopene concentration in tomatoes using a multispectral imaging approach with up to 15 bands. It is shown that combining intensity measurements at wavelengths from near-infrared to ultraviolet using a neural network model achieved correlation of  $R^2=0.977$  and RMS error=4.63 mg/kg against ground truth lycopene concentration. Our results are comparable or superior to other methods from the literature, which are analysed in detail in the paper. The method can be reproduced with minimal cost and demonstrates the feasibility of the method for industrial application. The main contribution is that a broader range of wavelengths are considered compared to most previous work, with rigorous analysis using a combination of simple regression and artificial neural networks.

**Keywords** Machine vision, multispectral, lycopene, tomato

## 1 Introduction

Tomatoes have a vital role in food supply, accounting for 16% of global vegetable<sup>3</sup> production during the last decade [1]. Tomatoes are a rich source of nutrients, including vitamins A and C, lycopene, and potassium. Lycopene is one of the most valuable bio-active compounds in

---

<sup>3</sup> Tomatoes are technically fruits but often classified as vegetables in a culinary sense.

tomatoes due to a health stimulating carotenoid with antioxidant properties and helps to prevent cardiovascular diseases, cancers, neurodegenerative maladies, and other conditions [2, 3]. With an estimated global annual production of 180 million tonnes [4] tomatoes are the primary natural source of lycopene in our diets. Lycopene content correlates with the maturity of a tomato [5] and is therefore a critical factor in supply chain logistics for optimising harvesting, transportation and storage.

Humans have a natural ability to assess food quality and safety via a simple analysis of the appearance of the tomato in the visible spectrum. The availability of sensors beyond the visible spectrum and progress in computer vision are extending this basic subjective capability, with 1000s of peer reviewed papers featuring keywords “hyperspectral imaging” and “fruit/vegetable/etc” during the last decade. The latest research is aimed at estimation of properties including ripeness, disease and nutritional value [6].

This paper describes a novel non-destructive method for the estimation of lycopene concentration in tomatoes using multispectral data analysis. The main contribution is that a broad range of wavelengths is considered (15 bands between 365nm and 940nm) and rigorously analysed using a combination of simple regression and artificial neural networks. The outputs offer invaluable information for researchers of automated tomato lycopene estimation (or general ripeness/quality estimation using lycopene as a proxy).

## 2 Related Work

Traditional methods for the precise measurement of lycopene content are high performance liquid chromatography (HPLC), thin layer chromatography (TLC) [7], and spectrophotometric absorbance (SPM) [8]. These chemometric methods have been available for several decades but are time consuming, require hazardous chemicals and destroy the samples.

Non-invasive spectroscopic techniques such as near infrared spectroscopy (NIRS), nuclear magnetic resonance spectroscopy, Raman spectroscopy (RS) and fluorescence spectroscopy are powerful spectroscopic techniques and have been investigated for applications in the

food industry. However, these methods are mostly expensive, are limited to a small number of sample measurement points, and are dedicated for laboratory use only [9,10].

Consequently, computer vision techniques have been explored that deploy reflected or transmitted light to measure lycopene concentration. Some of these methods use the visual spectrum (VIS) in the form of the CIE  $L^*a^*b^*$  colour representation. Other methods use multispectral or hyperspectral techniques, often extended to near-infrared (NIR) and/or ultraviolet (UV) wavelengths.

### **Methods based on the $L^*a^*b^*$ representation of the visual spectrum.**

Aries et al. [5] achieved a promising logarithmic regression correlation of  $R^2=0.96$  between lycopene and the  $a^*$  value from a chroma meter, when averaging 14 spots on the equatorial region of tomatoes. Vazques-Cruz et al. [11], used a similar approach with a point spectrophotometer, to obtain linear regression  $R^2=0.985$  using neural networks (NN) with two hidden layers to map intensities of  $L^*$ ,  $a^*$ ,  $b^*$ ,  $a^*/b^*$  and area of vine leaf to lycopene concentration. Ye et al. [12], claim a lower correlation of  $R^2=0.81$ , but using a handheld camera and ambient lighting, thus showing promise for realistic low-cost applications. The highest result found in the literature was a correlation between  $a^*$  and lycopene of  $R^2=0.985$ , from Barrios et al. [13] using third-grade polynomial regression. In their case, images were taken by a compact camera with white LED illumination and so appears also more practical than some of the earlier methods.

**Spectral methods.** Some works have incorporated non-visible light into computer methods for lycopene estimation, as already stated. The motivation for this is that better-discriminating, and generally richer, data for riper tomatoes may be accessible.

A linear correlation coefficient of  $R^2=0.96$  between predicted and measured lycopene values was published by Polder et al. [14], using a hyperspectral camera with 256 spectral bands. A multispectral approach with 19 wavelengths using LED illumination by Liu et al. [15] gave a lower value of 0.94, but using a set-up more practical for non-laboratory conditions. Tihalun et al. [16] use both VIS/NIR spectrometer and chroma meter for Hunter  $L^*a^*b^*$  representation of VIS. In con-

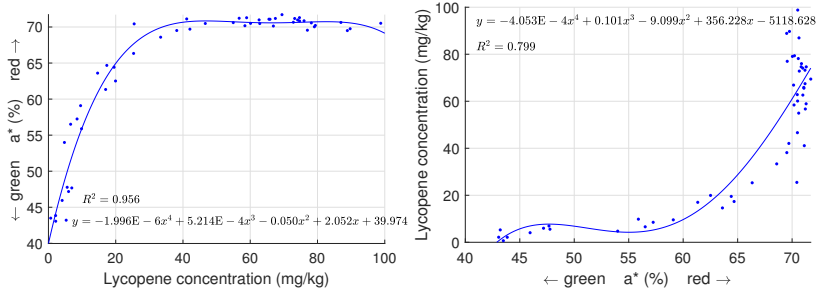
trast to other works, that paper used *transmitted* light passing through the tomato sample rather than reflected light. Results favoured the  $L^*a^*b^*$  method:  $R^2=0.96$  compared to  $R^2=0.85$  with the spectrometer.

**Discussion of the prior work.** The non-destructive lycopene content detection methods considered above are presented in Table 1. The results suggest that non-destructive estimation of the lycopene content by optical sensors is viable. Five methods have  $R^2$  higher than 0.95, of which, four are based on  $L^*a^*b^*$  colour space. Multi/hyper spectral methods have an average correlation of  $R^2=0.916$  compared to  $R^2=0.943$  for the  $L^*a^*b^*$  colour space methods.

**Table 1:** Comparison of previous methods with that proposed in this paper.

| Lycopene detection result/parameter<br>\<br>Authors | Best correlation $R^2$ | Result notes      | Linear correlation |       | Regression method | Colour space  | Light interaction | Spectrum bands /space | Wavelength min | Wavelength max | Ground truth | No. of samples |
|---|------------------------|-------------------|--------------------|-------|-------------------|---------------|-------------------|-----------------------|----------------|----------------|--------------|----------------|
|   |                        |                   | Y                  | N     |                   |               |                   |                       |                |                |              |                |
| Arias et al., (2000)                                | 0.960                  | Chroma meter      | N                  | N     | Logarithmic       | CIE Lab       | Reflect.          | a*                    | VIS            | VIS            | HPLC         | 38             |
| Polder et al., (2004)                               | 0.960                  | Line scan         | Y                  | LOOCV | SIMPLS reg.       | Hyperspectral | Reflect.          | 256                   | 396            | 736            | HPLC         | 37             |
| Vazquez-Cruz et al., (2013)                         | 0.985                  | Spectrophotometer | Y                  | N     | MLP (NN)          | Hunter Lab    | Reflect.          | a*/b*                 | VIS            | VIS            | HPLC         | 48             |
| Liu et al.,(2015)                                   | 0.938                  | Snapshot          | Y                  | LOOCV | BPNN              | Multispectral | Reflect.          | 19                    | 405            | 970            | SPM          | 162            |
| Ye et al., (2018)                                   | 0.810                  | Snapshot          | Y                  | N     | Linear regr.      | CIE Lab       | Reflect.          | (a*/b*) <sup>2</sup>  | VIS            | VIS            | SPM          | 60             |
| Barrios et al., (2018)                              | 0.985                  | Snapshot          | N                  | N     | Polyn.4gr         | CIE Lab       | Reflect.          | a*                    | VIS            | VIS            | SPM          | 60             |
| Tilahun et al., (2018)a                             | 0.960                  | Chroma meter      | Y                  | N     | PLS regr.         | Hunter Lab    | Transmit.         | a*/b*                 | VIS            | VIS            | SPM          | 180            |
| Tilahun et al., (2018)b                             | 0.850                  | Spectrometer      | Y                  | N     | PLS regr.         | Hyperspectral | Transmit.         | 1160                  | 500            | 1100           | SPM          | 180            |
| Mlynarik et al., (2022)a                            | 0.977                  | Snapshot          | Y                  | LOOCV | SNN               | Multispectral | Reflect.          | 15                    | 365            | 940            | SPM          | 50             |
| Mlynarik et al., (2022)b                            | 0.959                  | Snapshot          | N                  | N     | Polyn.4gr         | CIE Lab       | Reflect.          | a*                    | VIS            | VIS            | SPM          | 50             |

The success of the  $L^*a^*b^*$  methods are probably due the  $a^*$  parameter representing a green (chlorophyll) to red (lycopene) transition, reflecting a tomato’s natural colour changes during maturation. Fig. 1, shows the relationship between  $a^*$  and lycopene concentration using data captured for this paper (method described below). That is, an initial rapid transition from green to red as lycopene increases, followed by minimal change in  $a^*$  thereafter. This demonstrates why  $a^*$  alone can be successful, but also that it is not very discriminating for ripe tomatoes. In addition, hardware used for  $a^*$  methods are well established off-the-shelf components with time-proven calibrations, compared to hyperspectral or multispectral systems which are usually bespoke with proprietary calibration methods.



**Figure 1:** Measured relationship between a\* and lycopene concentration. The graphs are identical with polynomial regression, but with axes reversed.

A higher  $R^2$  value for a given regression might be an indicator of a superior fit to the data, but it can also be misleading in terms of achieving a useful model. For example, regression of measured a\* vs ground truth lycopene concentration can be high as  $R^2=0.96$  or as low as  $R^2=0.80$  depending on the somewhat arbitrary axis order (Fig. 1). Further, the regression offers no scientific basis to the underlying relationship.  $R^2$  of a linear regression between estimated and ground truth lycopene is more robust due to its resilience against over-fitting (as is root mean squared error (RMSE)). Unfortunately, not all past methods provide such parameters for comparison.

In addition to accuracy, other important factors for real-world application are practicality, speed and cost. The highest  $R^2$  in L\*a\*b\* methods are detected using multiple points around the sample relying on close proximity of the sensor (e.g. [5, 11]). Such a sampling technique is less practical than a single distant snapshot for high-throughput, high-speed sorting applications. Hyperspectral and multispectral techniques with more bands might increase the complexity of the system further. Therefore, the requirement of our method (and some others) for specialised illumination must be balanced against its benefits of more robust data capture.

### 3 Multispectral method for lycopene estimation

For this research, multispectral light reflections in 15 bands between 365nm and 940nm were used to investigate the precision of the method and its practicality for use in a controlled but non-contact industrial environment. The aim was to attain robustness and high correlation of predicted and measured lycopene content, especially for fully ripe tomatoes, while using commercially available devices that can easily be deployed in industry. The wavelength range was selected based on the assumption that a multispectral system consisting of more than three bands, covering both the full VIS spectrum and beyond, should contain more information than a system just utilising RGB sensor information converted to  $L^*a^*b^*$ . That is, the  $L^*a^*b^*$  data comprise a subset of the broader multispectral data and so should not exceed it in performance.

In this paper, multispectral data capture is optimised in the following ways. (1) Tomatoes were illuminated by dome lighting to avoid shadows and specular reflections. (2) The size and hardware construction were chosen to ensure uniform intensity over the entire fruit 3D surface. (3) The tomato was imaged from four sides to avoid situations where the red pigment is not evenly established during growth. While this arrangement might have limited direct applicability, the aim is to establish a robust baseline on which to build upon in future research.

**Experiment: methods and materials.** Fifty cultivar Saluoso tomatoes were harvested in late-autumn from a hydroponic greenhouse in south-east Slovakia. They were selected randomly, but covered a complete range from fully green to fully red. A multispectral image was captured (see below) for each tomato sample. Each sample was then blended within an hour and dissolved in hexan-etylen-aceton followed by spectrophotometric absorbance measurement at 503nm, in accordance with the method of Anthon and Barrett [17]. This process allowed the acquisition of a ground truth baseline from which comparisons could be made. One sample was later removed due to uncertainty during dissolution.

Multispectral images were captured by a Basler Ace monochromatic and near infrared area-scan camera. For each case, a series of LEDs in the range 365nm to 940nm were used to illuminate the sample in a



bespoke Technomedia dome with 340mm inner diameter. The system was calibrated with a spectralon target plate at seven points to ensure uniformity of image intensity between each wavelength.

Images were then segmented using basic thresholding functions in Halcon software. Next, image processing was split into two paths. (1) Convert the three images corresponding to RGB bands (478nm, 520nm, 635nm) to the  $L^*a^*b^*$  colour space to calculate an average pixel intensity of  $a^*$  for correlation with lycopene concentration. (2) Average segmented image intensities were fed into a shallow neural network (SNN), with five hidden layers, trained using the MATLAB `fitnet` function to map the multispectral data to measured lycopene values.

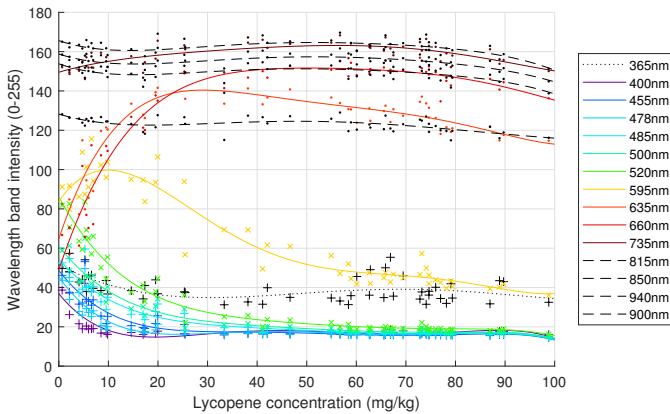
**Tomato surface area involved in computation.** Lycopene is not distributed evenly inside tomatoes, but is almost four times more concentrated in the skin compared to the pulp, and five times higher than the seeds [18]. Further, different parts of a tomato's surface may be more mature than other parts. The multispectral images were therefore taken from four sides: stem, bloom, left and right. Results are shown in Table 2. These  $R^2$  regression results confirm the hypothesis that larger coverage improves correlation.

| Polynomial regression of | Side of tomato for taking image |        |        |        |        |
|--------------------------|---------------------------------|--------|--------|--------|--------|
|                          | All                             | Bloom  | Stem   | Left   | Right  |
| 4 grade                  | 0.9557                          | 0.9531 | 0.9465 | 0.9508 | 0.9336 |
| 3 grade                  | 0.9467                          | 0.9429 | 0.9360 | 0.9431 | 0.9438 |
| 2 grade                  | 0.8847                          | 0.8779 | 0.8735 | 0.8830 | 0.8682 |
| Average above            | 0.9290                          | 0.9246 | 0.9187 | 0.9256 | 0.9152 |
| Log. regression          | 0.8965                          | 0.8926 | 0.8862 | 0.8943 | 0.8760 |

**Table 2:**  $R^2$  correlation between  $a^*$  and ground truth lycopene content for various sides of sample. Logarithmic, 2nd, 3rd and 4th grade polynomial regressions are shown.

**Selected spectra and wavelength bands contribution.** The green colour of unripe tomatoes is due to the prevalence of chlorophyll. During ripening, the synthesis of lycopene results in a red colour. Lycopene has a carotenoid molecular structure of eleven double bonds, allowing it to absorb energy from UV light between 270 and 310nm and blue and green light between 350 and 530nm [19]. In the proposed method therefore, this range is covered with seven spectral bands from 365 to 520nm. This is in addition to three wavelength bands in red spectra to capture the green to red colour shift. In total 15 wavebands were included, including NIR.

In Fig. 2, the measured average intensity of each waveband is plotted as a function of ground-truth lycopene concentration. Polynomial regression lines are also shown for ease of comparison. The figure shows several wavelengths with similar shape, suggesting little benefit of including them all. However, about seven different trends can be recognised. For a well-designed neural network, during training, the weights will become optimised to exploit these trends.

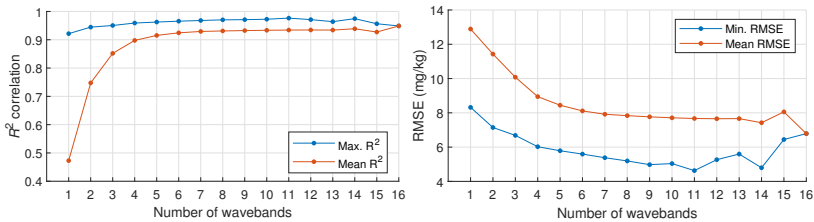


**Figure 2:** Averaged pixel intensity of reflected light as a function of lycopene for all wavelengths. [Colour coding approximately matches wavelength. “+”: ultraviolet/blue, “x”: yellow/green, “.”: red/infrared.]

**Shallow Neural Network (SNN).** The additional information available from multispectral data was incorporated using Levenberg-Marquardt backpropagation SNN with 5 hidden layers. This approach is known to better model the non-linear interaction of sparse data. Modern methods for computer vision typically use convolutional neural networks (CNNs). However, that is deemed unnecessary here since the inputs are single values corresponding to mean intensity measurements for each wavelength (i.e. there is little benefit from setting entire images as inputs, as expected by most CNN architectures). In future work, it might be possible to use CNNs in order to incorporate potentially useful spatial information.

To investigate the influence of the various wavebands on the appearance of lycopene, an SNN was trained for all possible band combinations using identical settings. In addition, one more input to the SNN was added: the physical size of the tomato sample as a 16th possible input. The motivation for this is that, as lycopene is more highly concentrated near the surface, the physical size may affect average concentration levels of the sample. As presented below, the best prediction was, indeed, achieved with that additional input.

For evaluation, the leave one out cross validation (LOOCV) method was used. Given a sample size of 49 therefore, 49 training sessions were performed for each of 65,535 possible combinations of wavebands from 1 to 16 bands. Fig. 3 shows the general effect of the number of bands considered (1,2,...16) in terms of performance.



**Figure 3:** SNN performance expressed in maximum and average  $R^2$  correlation (left) and minimum and average RMSE (right) of prediction against number of input wavelength bands.

Multispectral LOOCV linear regression correlation reached a maximum of  $R^2=0.9765$  for lycopene prediction and measured ground truth concentration. This corresponds to RMS error of prediction of 4.63 mg/kg. A combination of 11 bands gave this result (all those in the legend for Fig. 2 except 485nm, 520nm, 635nm, 850nm).

It was found that the SNN performance does not improve when the number of input bands is above about eight. This might be due to the introduction of noise with additional bands with very similar shape or due to the model over-fitting. Therefore, although the 11 wavebands in the optimal SNN mentioned above had best correlation in experiments, it is likely that almost equally good outputs are possible with fewer (not necessarily identical) inputs.

**Discussion.** Our method allowed us to explore both the multispectral and the  $L^*a^*b^*$  approaches. At best, we found that fitting  $a^*$  against lycopene concentration using 4th grade polynomial regression gave  $R^2=0.9557$ . While this sounds promising, the  $a^*$  value rapidly converges with moderate lycopene concentration, meaning the regression curve has limited use above certain maturity levels. This problem is also apparent in some other research that focuses on  $L^*a^*b^*$  space. Further, high-grade polynomials such as this are widely known to over-fit and should be interpreted with care.

As an alternative to the above approach, where polynomial fitting might be somewhat arbitrary, we have also trained SNNs with varying numbers of hidden layers for all possible combinations of the wavelengths and sample size. Through trial-and-error, it was found that results improved with the number of hidden layers up to about 5, beyond which, little improvement was obtained. For this reason, only results from SNNs with exactly five hidden layers are presented. Results show that the stability and prediction of correlation increase with the number of wavebands, as hypothesised. Additional bands, including those outside the visual spectrum, have proven their contribution to model robustness and preciseness.

The results from both previous works and our own, are shown in Table 1. This indicates that the performance of our method is comparable to others, while maintaining a more reproducible approach and application of cross-validation, which not all others do.

## 4 Conclusion

While previous research has shown promise for lycopene concentration estimation using computer vision, this research offers a more robust grounding with detailed experiments in controlled conditions. This demonstrates what may be possible using intensity analysis at a range of wavelengths in a laboratory setting, which can be reproduced with minimal cost. The limitations of  $L^*a^*b^*$  space are demonstrated and it is shown how our multispectral approach goes some way to overcome these using neural networks. Future work will aim to investigate how the approach can be extended to operate in an agricultural setting.

## References

1. STATISTA.com: Vegetables: Worldwide., <https://www.statista.com/outlook/cmo/food/vegetables/worldwide>, [Accessed: 11 January 2023].
2. "Lycopene: modes of action to promote prostate health," *Archives of Biochemistry and Biophysics*, vol. 430, no. 1, pp. 127–134, 2004.
3. S. Przybylska, "Lycopene – a bioactive carotenoid offering multiple health benefits: a review," *International Journal of Food Science and Technology*, vol. 55, pp. 11–32, 2019.
4. World Food and Agriculture – Statistical Yearbook 2021., <https://doi.org/10.4060/cb4477en>, [Accessed: 11 January 2023].
5. R. Arias, T.-C. Lee, L. Logendra, and H. Janes, "Correlation of lycopene measured by hplc with the L,a,b color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content," *Journal of Agricultural and Food Chemistry*, vol. 48, no. 5, pp. 1697–1702, 2000.
6. J. Wieme, K. Mollazade, I. Malounas, M. Zude-Sasse, M. Zhao, A. Gowen, D. Argyropoulos, S. Fountas, and J. Van Beek, "Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review," *Biosystems Engineering*, vol. 222, pp. 156–176, 2022.
7. W. Wardencki, P. Biernacka, T. Chmiel, and T. Dymerski, "Instrumental techniques used for assessment of food quality," *Proceedings of ECOpole*, vol. 3, 2009.
8. W. W. Fish, P. Perkins-Veazie, and J. K. Collins, "A quantitative assay for lycopene that utilizes reduced volumes of organic solvents," *Journal of Food Composition and Analysis*, vol. 15, no. 3, pp. 309–317, 2002.
9. A. Hussain, H. Pu, and D.-W. Sun, "Measurements of lycopene contents in fruit: A review of recent developments in conventional and novel techniques," *Critical reviews in food science and nutrition*, vol. 59, no. 5, pp. 758–769, 2019.
10. M.-J. Villaseñor-Aguilar, J.-A. Padilla-Medina, J.-E. Botello-Álvarez, M.-G. Bravo-Sánchez, J. Prado-Olivares, A. Espinosa-Calderon, and A.-I. Barranco-Gutiérrez, "Current status of optical systems for measuring lycopene content in fruits," *Applied Sciences*, vol. 11, no. 19, p. 9332, 2021.
11. M. Vazquez-Cruz, S. Jimenez-Garcia, R. Luna-Rubio, L. Contreras-Medina, E. Vazquez-Barrios, E. Mercado-Silva, I. Torres-Pacheco, and R. Guevara-Gonzalez, "Application of neural networks to estimate carotenoid content during ripening in tomato fruits (*solanum lycopersicum*)," *Scientia Horticulturae*, vol. 162, pp. 165–171, 2013.

12. X. Ye, T. Izawa, and S. Zhang, "Rapid determination of lycopene content and fruit grading in tomatoes using a smart device camera," *Cogent Engineering*, vol. 5, no. 1, p. 1504499, 2018.
13. A. Gastélum-Barrios, J. García-Trejoo, G. Soto-Zarazúa, G. Macías-Bobadilla, and M. Toledano-Ayala, "Portable system to estimate ripeness and lycopene content in fresh tomatoes based on image processing," in *Int. Eng. Congress*, 2018, pp. 1–5.
14. G. Polder, G. Van Der Heijden, H. Van der Voet, and I. Young, "Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry," *Postharvest biology and technology*, vol. 34, no. 2, pp. 117–129, 2004.
15. C. Liu, W. Liu, W. Chen, J. Yang, and L. Zheng, "Feasibility in multispectral imaging for predicting the content of bioactive compounds in intact tomato fruit," *Food Chemistry*, vol. 173, pp. 482–488, 2015.
16. S. Tilahun, M. H. Seo, I. G. Hwang, S. H. Kim, H. R. Choi, C. S. Jeong *et al.*, "Prediction of lycopene and  $\beta$ -carotene in tomatoes by portable chromameter and VIS/NIR spectra," *Postharvest biology and technology*, vol. 136, pp. 50–56, 2018.
17. G. Anthon and D. M. Barrett, "Standardization of a rapid spectrophotometric method for lycopene analysis," in *X International Symposium on the Processing Tomato 758*, 2006, pp. 111–128.
18. R. K. Toor and G. P. Savage, "Antioxidant activity in different fractions of tomatoes," *Food Research International*, vol. 38, no. 5, pp. 487–494, 2005.
19. H. Hashimoto, C. Uragami, and R. J. Cogdell, "Carotenoids and photosynthesis," *Carotenoids in nature*, pp. 111–139, 2016.

# An overview of implementing Multispectral Imaging coupled with machine learning for the assessment of microbiological quality and authenticity in foods

Anastasia Lytou<sup>1</sup>, LEMONIA-Christina Fengou<sup>1</sup>, Nette Schultz<sup>2</sup>, Jens Michael Carstensen<sup>2</sup>, Yimin Zhang<sup>3</sup>, Fady Mohareb<sup>4</sup>, and George-John Nychas<sup>1</sup>

<sup>1</sup> Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Human Nutrition, Agricultural University of Athens, Athens, Greece

<sup>2</sup> Videometer

Herlev, Denmark

<sup>3</sup> Lab of Beef Processing and Quality Control, College of Food Science and Engineering, Shandong Agricultural University  
Tai'an, Shandong, 271018, P.R. China

<sup>4</sup> Cranfield University  
Cranfield, UK

**Abstract** Multispectral Imaging is an increasingly applied technique for the estimation of several quality parameters across the food chain. The microbiological quality and safety as well as the detection of food fraud are among the most significant aspects in food quality and safety assessment. MSI analysis was performed using a VideometerLab instrument (Videometer A/S, Videometer, Herlev, Denmark), while more than 9000 food samples were examined in total, for the assessment of microbiological quality and the detection of food fraud. For estimating microbial populations, total aerobic counts (TAC) were determined. Several regression and classification algorithms were employed, including partial least squares regression (PLS-R), support vector machines (SVM), partial least squares discriminant analysis (PLS-DA), tree-based algorithms etc. The slope of the regression line, root mean squared error (RMSE), coefficient of determination (R-squared) and accuracy score were used as metrics

for the evaluation of models' performance. In adulteration case, the prediction of different levels of pork in chicken meat and vice versa yielded high accuracy scores i.e., over 90% , while, using the SVM algorithm, the presence of bovine offal in beef was successfully detected. Additionally, Random Forest algorithm was efficient (accuracy>93% ) in discriminating seabass and seabream fish fillets. Concerning microbiological quality, as indicated by the performance indices, the developed models exhibited satisfactory performance in predicting microbial load in different foods (RMSE<1.00, R-squared>0.80). Indicatively, MSI spectral data combined with PLS-R could satisfactorily predict TAC and *Pseudomonas* spp. counts on the surface of chicken fillets regardless of storage temperature and batch variation based on the performance metrics (R-squared: 0.89, RMSE: 0.88) while, this algorithm presented also satisfactory performance in estimation microbial populations in brown edible seaweed (R-squared: 0.80, RMSE: 0.90). However, in this case, selecting the appropriate analytical approaches and machine learning algorithms is still challenging.

**Keywords** Multispectral Imaging, Food Quality, Machine Learning, Food Fraud

## 1 Introduction

The interest in using optical technologies that are capable of real-time quality, safety and authenticity assessment has been continuously increasing [1]. Food industry, apart from stabilizing the products to avoid food losses and food waste, should also focus to the development of rapid analytical technologies for the estimation of the microbiological quality and freshness. The last few decades there has been a huge effort from stakeholders to investigate alternative methods that are suitable for online, real-time food quality/safety assessment [2]. In recent years, rapid development of non-invasive sensing technologies for food quality contributed to significant transformations in the supply chain [3]. The data acquired from sensors do not indicate anything without processing and conversion into useful information using pattern recognition or prediction models. Towards this direction, machine learning

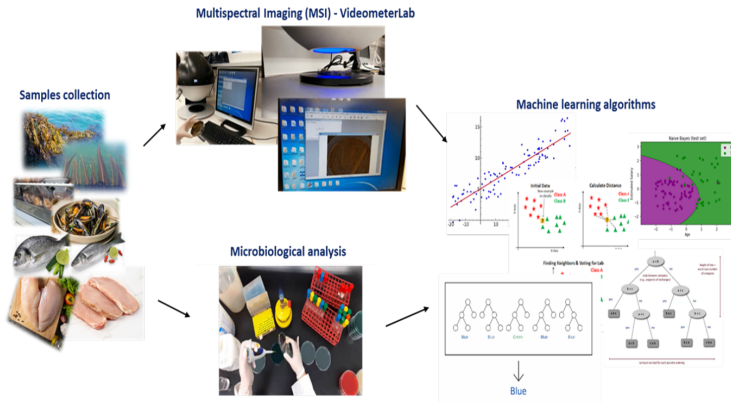


algorithms such as, Partial least squares regression (PLS-R), Linear discriminant analysis (LDA), and Quadratic discriminant analysis (QDA) have been reported as reliable tools for the development of predictive models for quality or adulteration assessment in meat [4], [5]. Moreover, deep learning approaches such as artificial neural networks (ANNs) and support vector machines (SVMs) have been employed, validated, and compared through available online platforms/tools (e.g., sorfML, Metaboanalyst), softwares (e.g., The Unscrambler) or programming languages (R, MatLab, Python), in an attempt to provide accurate predictive models for food spoilage assessment [6], [7]. This work is an overview of studies investigating Multispectral Imaging Analysis, by analyzing various foodstuffs, in an attempt to collect a satisfactory amount of MSI data which in combination with machine learning models can provide significant information about the quality and authenticity of foods.

## 2 Materials and Methods

The whole experimental procedure is briefly shown in Figure 1. The four main steps of the analytical process were 1. Samples' collection, 2. Microbiological analysis, 3. Multispectral Imaging Analysis and 4. Data analysis.

- Samples' collection: Various food samples (9000 samples in total) were collected during the last 5 years. In brief, poultry meat (2300), beef (400), pork (700), fish (1000), pineapple (400), leafy vegetables (500), seaweeds (500), shellfish (500) etc, were subjected to microbiological and MSI analysis, whereas 2000 samples were analysed covering different adulteration scenarios (i.e., chicken vs pork, beef vs offal etc.) In an attempt to increase the diversity of the samples and subsequently the size and the variability of the dataset, apart from the fresh samples, samples that were stored at different temperatures (0, 5, 10, 15°C) for certain time intervals, were also tested. In this way, samples with different microbiological populations and freshness levels were also analysed.
- Microbiological analysis: For the estimation of total aerobic

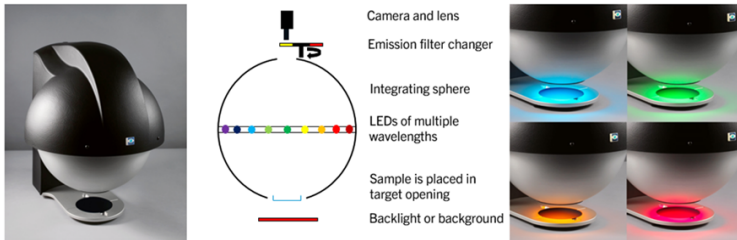


**Figure 1:** Schematic representation of the procedure from samples' collection to data analysis in brief.

counts (TAC), a specific quantity of food sample was transferred aseptically to a stomacher bag, diluted ten times using sterile maximum recovery diluent (MRD) and homogenized in a stomacher (Lab Blender, Seward Medical, London, UK) for 120 s at room temperature. The homogenate was then serially diluted in testing tubes and 0.1 mL of the appropriate dilution was spread in duplicate on the respective culture medium depending on the microbial group. After incubation, colonies were enumerated and their counts were logarithmically transformed ( $\log$  CFU/g).

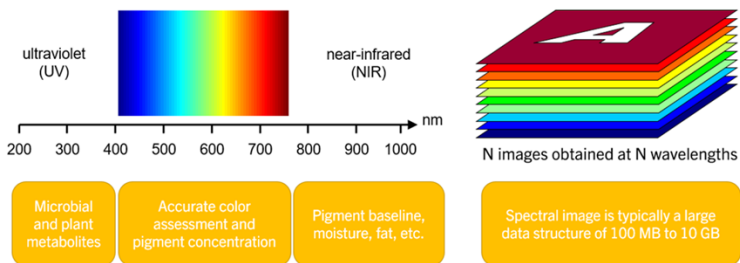
- **Multispectral Imaging Analysis (MSI):** Multi-spectral images (MSI) were captured using a Videometer-Lab instrument (Videometer A/S, Herlev, Denmark) that acquires images in 18 different non-uniformly distributed wavelengths from UV (405 nm) to short wave NIR (970 nm), namely, 405, 435, 450, 470, 505, 525, 570, 590, 630, 645, 660, 700, 850, 870, 890, 910, 940, and 970 nm. LED-based spectral imaging as illustrated in Figure 2 is a fast, non-destructive, and versatile technology for providing high contrast food chemical maps when combined with machine learning methodology. LEDs covering UV, Visual, and NIR wavelengths are sequentially strobed into an integrating sphere with a

superwhite coating. The food sample is placed in the opening of the lower half sphere and receives a very homogenous and diffuse illumination. The built-in calibration and exposure control ensures optimal dynamic range, reproducibility, and traceability.



**Figure 2:** VideometerLab instrument used for spectral imaging of food systems. LED strobes of UV-Vis-NIR wavelengths are used to generate a spectral image. Reflectance and fluorescence modes may be combined in the same imaging sequence.

The spectral image, as illustrated in Figure 3, provides information about a rich set of important food compounds like plant and microbial metabolites, pigments, moisture, and lipids. Further it offers a way to measure or remove effects from physical food properties like scattering, specularity, translucency, and heterogeneity.



**Figure 3:** LED band-sequential imaging for MSI results in a spectral cube data structure that maps many food-related compounds.

- Data analysis: Various algorithms were employed in the analysis of the MSI data, including Partial Least Squares Regression (PLS-R), Support Vector Regression (SVM-R), tree-based algorithms (Random Forests Regression (RF-R) and Extra Trees) k-Nearest Neighbours' Regression (kNN-R), Linear Discrimination (LDA), Quadratic discrimination (QDA) etc. A part of the dataset was used for the training of the model, while an independent, external dataset was used for the validation (testing) of the model. The performance of the developed models was evaluated via the following metrics and indices: root mean squared error (RMSE), correlation coefficient (r), overall accuracy, precision, and recall.

### 3 Results

Some indicative results of the MSI applications using various foods are presented below.

3.1 Estimation of microbial population in chicken fillets regardless of storage temperature and batch variation: A PLS-R model was developed by Spyrelli et al [8] for the estimation of microbial counts in chicken fillets. The model parameters and performance metrics (slope, R-squared, RMSE), for the estimation of the population of TAC and *Pseudomonas* spp. using MSI spectral data, are presented in Table 1.

**Table 1:** Performance metrics of PLS-R models estimating TAC and *Pseudomonas* spp. population of chicken fillets using MSI data.

| TAC                     | Number of samples | slope (a) | R-squared | RMSE |
|-------------------------|-------------------|-----------|-----------|------|
| Calibration             | 330               | 0.74      | 0.86      | 0.73 |
| Cross Validation        | 330               | 0.73      | 0.84      | 0.78 |
| Prediction              | 72                | 0.77      | 0.90      | 0.98 |
| <i>Pseudomonas</i> spp. |                   |           |           |      |
| Calibration             | 330               | 0.73      | 0.85      | 0.83 |
| Cross Validation        | 330               | 0.71      | 0.83      | 0.88 |
| Prediction              | 72                | 0.70      | 0.90      | 1.21 |

For TAC, the RMSE and R-squared values for model calibration and cross validation were 0.73 and 0.78 log CFU/cm<sup>2</sup>, as well as 0.86 and

0.84, respectively, whereas the respective values for the prediction were 0.99 log CFU/cm<sup>2</sup> and 0.90, respectively. The predicted values were mostly observed within the area of  $\pm 1.0$  log CFU/cm<sup>2</sup>, which is considered microbiologically acceptable, while an overestimation for low counts (below 4.0 log CFU/cm<sup>2</sup>) was evident. Concerning the PLS-R model assessing *Pseudomonas* spp. counts, RMSE and R-squared values were 0.83 log CFU/cm<sup>2</sup> and 0.85, respectively, for calibration, while for cross validation they were 0.87 and 0.83 log CFU/cm<sup>2</sup>, respectively. For the prediction of *Pseudomonas* spp. counts, RMSE and R-squared values were estimated at 1.21 and 0.90 log CFU/cm<sup>2</sup> respectively.

3.2 Microbiological quality assessment of seaweed obtained from different geographical areas and harvest years: The prediction model development and validation for the MSI of *A. esculenta* from MI and SAMS samples, from different harvest years are presented below (Table 2), while the findings of this study have been extensively described in [9]. The performance of the model developed in separate for the samples from the different geographical areas was not satisfactory.

**Table 2:** Linear regression fit parameters between actual and predicted TAC values for the different datasets (*A. esculenta* MI, SAMS, MI+SAMS) acquired from MSI analysis.

| MI               | slope (a) | R-squared | RMSE |
|------------------|-----------|-----------|------|
| Cross Validation | 0.67      | 0.67      | 0.96 |
| Prediction       | 0.49      | 0.51      | 0.95 |
| SAMS             |           |           |      |
| Cross Validation | 0.79      | 0.79      | 1.18 |
| Prediction       | 0.56      | 0.40      | 1.83 |
| MI+SAMS          |           |           |      |
| Cross Validation | 0.92      | 0.92      | 0.81 |
| Prediction       | 0.84      | 0.81      | 1.04 |

Extended spectral differences have been observed among the years of harvesting suggesting that maybe the MSI is not suitable for efficient microbial population estimation due to the dependence of this method from the “colour” of the samples that can be misleading for the prediction model. In the case that data from SAMS and MI were

combined, performance statistics values were improved compared to those models developed for each origin in separate (R-squared: 0.80, RMSE: 1.04). Probably by enlarging the size of data, the model was trained/learned better (good performance statistics in cross validation) and the differences in products among the differences in years were more successfully incorporated into the model, while the significance of the visual features (colour related) was degraded.

3.3 Discrimination of fish fillet samples based on different fish species: Several machine learning algorithms were tested for their ability to classify fish fillets to the correct fish species. All the tested models yielded high accuracy scores (>90 % classified to the correct group) for images captured both from the skin and from the flesh side of the fillet (Table 3). Models developed using data from images captured from the skin side, exhibited even better performance (accuracy > 96 %).

**Table 3:** Accuracy scores ( % ) for the discrimination of fish fillets based on species (i.e., seabass, seabream) using different algorithms.

| Accuracy % | SVM   | Extra trees | Random Forest |
|------------|-------|-------------|---------------|
| Skin       | 98.39 | 97.85       | 96.77         |
| Flesh      | 95.65 | 93.48       | 94.57         |

3.4 Detection of meat adulteration: In Table 4 the performance metrics for the external validation and the classification in five classes for the MSI data is presented. The developed models yielded high performances especially for the classes containing higher proportions of chicken (classes 0 and 25% ).

The classification models of SVMs for the detection of the adulteration of beef with bovine offal (bovine hearts) showed higher or equal performance in terms of accuracy scores for the respective cases compared with the pork-chicken adulteration scenario. The overall correct classification (accuracy) for the case of pork in chicken and offal in beef was 90 % and 100.00 % , respectively. These findings are part of results published before [10].

**Table 4:** Linear regression fit parameters between actual and predicted TAC values for the different datasets (*A. esculenta* MI, SAMS, MI+SAMS) acquired from MSI analysis.

|                 | True class |      |      |       |       |
|-----------------|------------|------|------|-------|-------|
|                 | 0 %        | 25 % | 50 % | 75 %  | 100 % |
| Pork in chicken | 0 %        | 25 % | 50 % | 75 %  | 100 % |
| Recall (%)      | 100        | 100  | 100  | 100   | 50    |
| Precision (%)   | 100        | 100  | 100  | 66.67 | 100   |
| Offal in beef   | 0 %        | 25 % | 50 % | 75 %  | 100 % |
| Recall (%)      | 100        | 100  | 100  | 100   | 100   |
| Precision (%)   | 100        | 100  | 100  | 100   | 100   |

## 4 Conclusion

MSI data coupled with machine learning algorithms exhibit potential towards efficient detection of adulteration and microbial counts estimation and could be a rapid and non-invasive tool for the quality assessment in various foodstuffs.

This work has been funded by the project DiTECT (861915).

## References

1. H. Pu, K. M., and D. Sun, "Selection of feature wavelengths for developing multispectral imaging systems for quality, safety and authenticity of muscle foods-a review," in *Trends in Food Science and Technology*, 2015.
2. P. Tsakanikas, A. Karnavas, E. Panagou, and G. Nychas, "A machine learning workflow for raw food spectroscopic classification in a future industry," in *Scientific Reports*, 2020.
3. N. J. Watson, A. Bowler, A. Rady, O. Fisher, A. Simeone, J. Escrig, E. Woolley, and A. Adedeji, "Intelligent sensors for sustainable food and drink manufacturing," in *Frontiers in Sustainable Food Systems*, 2021.
4. A. Ropodi, D. Pavlidis, F. Mohareb, E. Panagou, and G. Nychas, "Multispectral image analysis approach to detect adulteration of beef and pork in raw meats," in *Frontiers in Sustainable Food Systems*, 2017.
5. Y. Kumar and S. Karne, "Spectral analysis: A rapid tool for species detection in meat products," in *Trends in Food Science and Technology*, 2019.

A. Lytou et al.

6. A. M. Jiménez-Carvelo, A. González-Casado, M. G. Bagur-González, and L. Cuadros-Rodríguez, "alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—a review," in *Food research international*, 2019.
7. S. Jaafreh, O. Valler, J. Kreyenschmidt, K. Günther, and P. Kaul, "In vitro discrimination and classification of microbial flora of poultry using two dispersive raman spectrometers (microscope and portable fiber-optic systems) in tandem with chemometric analysis," in *Talanta*, 2019.
8. E. Spyrelli, C. Papachristou, G.-J. Nychas, and E. Panagou, "Microbiological quality assessment of chicken thigh fillets using spectroscopic sensors and multivariate data analysis," in *Foods*, 2021.
9. A. Lytou, P. Tsakanikas, D. Lymperi, and G.-J. Nychas, "Rapid assessment of microbial quality in edible seaweeds using sensor techniques based on spectroscopy, imaging analysis and sensors mimicking human senses," in *Sensors*, 2022.
10. L.-C. Fengou, A. Lianou, P. Tsakanikas, F. Mohareb, and G.-J. Nychas, "Detection of meat adulteration using spectroscopy-based sensors," in *Foods*, 2021.



# Self-supervised Pretraining for Hyperspectral Classification of Fruit Ripeness

Leon Amadeus Varga\*, Hannah Frank\*, and Andreas Zell

University of Tuebingen, Cognitive Systems

Sand 1, 72076 Tuebingen

\* : shared contribution

**Abstract** The ripeness of fruit can be measured in a non-destructive way using hyperspectral imaging (HSI) and deep learning methods. However, the lack of labeled data samples limits hyperspectral image classification. This work explores self-supervised learning (SSL) as pretraining for HSI classification of fruit ripeness. Three state-of-the-art SSL methods, *SimCLR*, *SimSiam*, and *Barlow Twins* are implemented, and augmentation techniques for HSI are developed. A 3D-2D hybrid convolutional network is proposed to support the pretraining procedure. This model is evaluated against a *ResNet-18* and a *HS-CNN*. The pretraining is evaluated on the fruit ripeness prediction task using the proposed second version of the *DeepHS* fruit data set. Besides comparing the classification performance of the pretrained models to only supervised training, the influence of the model architecture and size, pretraining method, and augmentations for SSL is investigated. This work shows that it is possible to transfer the ideas of SSL to HSI. It is possible to extract essential features in an unsupervised manner via this pretraining. Pretraining stabilizes classifier training and improves the classifier performance. Further, it can partially compensate for the need for large labeled data sets in HSI classification.

**Keywords** Self-supervised learning, pretraining, hyperspectral imaging, HSI classification, fruit ripeness

## 1 Introduction

Knowing the ripeness of fruit is of great interest in the food industry. Especially exotic fruit, like avocados, kiwis, or papayas, are harvested when still unripe, kept in storage rooms, and are often shipped for weeks from far away. In addition, those kinds of exotic fruit often have a relatively high price. A reliable estimation of the fruit's ripeness state is required.

For this, usually, chemical and physical indicators like the sugar content and fruit flesh firmness are employed, all of which are obtained by destructive measurement.

It is also possible to predict the ripeness of fruit using hyperspectral imaging (HSI) [1, 2], which is non-destructive and therefore has become increasingly popular in recent years. Current work shows that combining HSI and deep learning can improve those predictions even further [3–5].

However, deep neural networks are usually trained in a supervised manner. Obtaining the actual ripeness state of a fruit still comes with destroying it, making the labeling process tedious and labeled samples scarce. Training networks on small training sets can be challenging, and overfitting becomes likely. Therefore, it is desirable to also use unlabeled fruit recordings that can be obtained without much effort.

Self-supervised learning (SSL) methods have produced astonishing results in computer vision [6–8] and may be applied for pretraining in this particular case of hyperspectral image classification to stabilize the training and potentially improve the network's predictions.

## 2 Experiments

### 2.1 Data Set

This work extended the already publicly available hyperspectral fruit data set, *DeepHS* [5], by additional recordings of avocados, kiwis, mangos, persimmon, and papayas. We used the same measurement setup and proceeding described by Varga et al. [5]. Each fruit was recorded by the *Specim FX 10* with 224 bands (398 nm - 1004 nm) and the *Corning microHSI 410 Vis-NIR Hyperspectral Sensor* with 249 bands (408 nm - 901

nm). Labels (firmness, sugar level, and overall ripeness) were obtained by destructive measurement.

The resulting *DeepHS v2* data set consists of 4671 recordings in total, 1018 labeled. Only the labeled subset was used for classification, while for self-supervised pretraining, also the unlabeled samples were used.

## 2.2 Models

Varga et al. [5] already proposed the *HS-CNN* network, a small convolutional neural network specialized for HSI data and the application for fruit ripeness classification.

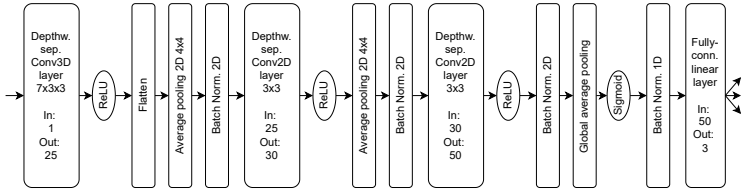


Figure 1: Architecture of the 3D-2D hybrid model.

We suggest a slightly modified variant, a 3D-2D hybrid model, using a 3D convolution instead of a 2D convolution in the first layer – inspired by *HybridSN* [9]. Its architecture is shown in Fig. 1. The backbone consists of a 3D convolutional layer for spectral-spatial feature learning and two 2D convolutional layers for more abstract spatial feature learning. Finally, a fully-connected layer operating on the spectral dimension is used for actual classification. With the hybrid version, we obtained a larger model ( $\approx 20\times$  as many parameters than the baseline).

Additionally, we evaluated our methods using a *ResNet* architecture [10], which is also commonly employed for self-supervised learning (e.g., [6–8]) but has significantly more parameters compared to the other two models.

## 2.3 Self-supervised Pretraining

The model was pretrained using one of the three SSL methods: *SimCLR* [6], *SimSiam* [7], *Barlow Twins* [8].

All employ a siamese network architecture [11] where each branch is built by the encoder, the convolutional part of the classifier model, followed by a projection head. For the latter, we used a MLP with two layers. A *ReLU* non-linearity and batch normalization [12] was applied for each layer. The input dimension was 50 (for the baseline or hybrid model, and 512 for the *ResNet-18*), the hidden dimension was 16, and the embedding dimension was eight. For *SimSiam*, we used an additional prediction MLP, consisting of a single linear layer with input and output dimension of eight. The temperature parameter for *SimSiam* was chosen to be  $\tau = 0.1$ . For *Barlow Twins*, a weighting factor  $\lambda = 0.01$  was used.

A critical component of SSL are the data augmentations. We evaluated 21 augmentation techniques, including four basic image transformations (rotating, flipping, cropping, random noise), two more specific ones (wavelength-dependent noise and pixel-wise intensity scaling), 13 augmentations that modify parts of the hyperspectral cube (i.e., drop or blur specific pixels, channels, or an entire sub-cube [13]), as well as two mixing augmentations (inspired by *MixUp* [14] and *ScaleMix* [15]).

Based on the ablation studies (see Sec. 4), only a subset of the augmentations (random rotations with probability 50%, random cropping with probability 30%, modification of the hyperspectral cube, and mixing with probability 20%) was actually used for pretraining.

The networks were optimized with SGD [16] with a weight decay of  $10^{-4}$ , a momentum of 0.9, and a learning rate of  $10^{-2}$ , decayed with the cosine decay schedule without restart [17]. We trained for 80 epochs with an effective batch size of 32.

## 2.4 Evaluation

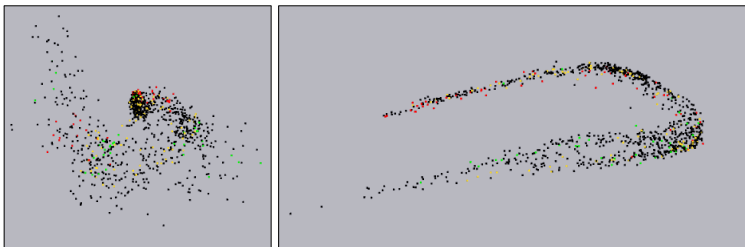
For the evaluation of self-supervised pretraining, the produced embeddings were considered. They were evaluated qualitatively (based on 3D visualizations) and quantitatively (based on the k-Nearest-Neighbor accuracy). For the visualization, the feature values of the embedding were plotted in three-dimensional space, after applying PCA. k-Nearest-Neighbor (k-NN) classification [18] was employed for the embedded labeled samples, using  $k = 5$ , the cosine distance and leave-one-out cross-validation (see, e.g., [7, 19]).

Additionally, we measured the performance for classification without and with pretraining. For the pretrained model, first, the fully-connected part was trained on top of the pretrained backbone, and then all model weights were further fine-tuned on the classification task (e.g., [6–8]). Without pretraining, the randomly initialized model was trained using settings similar to Varga et al. [5].

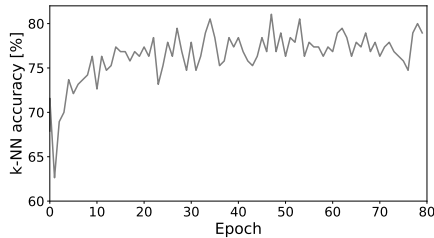
After the supervised training, the model was evaluated on the test set. Test time augmentations [20] were applied with probability 50%.

Using five different seeds each, we conducted experiments for all possible combinations of fruit types, cameras, and categories.

### 3 Results



(a) Embedding, before (left) and after pretraining (right).



(b) k-NN accuracy.

**Figure 2:** (a) 3D visualization of the embedding before and after pretraining via *Barlow Twins* – coloring by ripeness levels: unripe (green), ripe (yellow), overripe (red) and unlabeled (black). (b) k-NN accuracy on the ripeness levels of the labeled samples (train and validation set) during pretraining with *SimCLR*. For the hybrid model and the avocados, recorded by the *Specim* camera.

To evaluate the pretraining per se, we visualized the embeddings in 3D and monitored the k-NN accuracy during pretraining (see Fig. 2).

The spatial arrangement in the 3D space correlates with the ripeness level; samples of the same ripeness level are brought closer together. This fits the development of the k-NN accuracy, which increases as pretraining advances and finally converges towards 80%. This shows that pretraining can extract meaningful features and find useful representations for the data, without using label information.

**Table 1:** Classification accuracies (median, IQR) for regular classifier training versus *SimCLR* pretraining plus fine-tuning, for the *HS-CNN* (baseline) and hybrid model. One example for the five different fruit: Avocado (ripeness, *Specim*), kiwi (sugar, *Specim*), mango (firmness, *Specim*), kaki (sugar, *Specim*), papaya (ripeness, *Corn-ing*), and over all fruit, categories and camera types. Highest accuracies in **bold**.

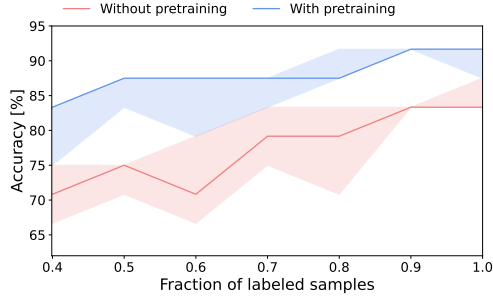
|          |                     | Avocado                         | Kiwi                            | Mango                     | Kaki                            | Papaya                           | Overall                          |
|----------|---------------------|---------------------------------|---------------------------------|---------------------------|---------------------------------|----------------------------------|----------------------------------|
| Baseline | Without pretraining | 83.3%<br>( $\pm 4.2\%$ )        | 65.2%<br>( $\pm 4.3\%$ )        | 50.0%<br>( $\pm 33.3\%$ ) | 50.0%<br>( $\pm 4.3\%$ )        | 77.8%<br>( $\pm 11.1\%$ )        | 55.6%<br>( $\pm 32.2\%$ )        |
|          | With pretraining    | <b>87.5%</b><br>( $\pm 0.0\%$ ) | <b>73.9%</b><br>( $\pm 8.7\%$ ) | 50.0%<br>( $\pm 8.3\%$ )  | <b>66.7%</b><br>( $\pm 8.7\%$ ) | <b>88.9%</b><br>( $\pm 0.0\%$ )  | <b>58.3%</b><br>( $\pm 32.2\%$ ) |
| Hybrid   | Without pretraining | 75.0%<br>( $\pm 4.2\%$ )        | 73.9%<br>( $\pm 13.0\%$ )       | 50.0%<br>( $\pm 33.0\%$ ) | 58.3%<br>( $\pm 13.0\%$ )       | <b>88.9%</b><br>( $\pm 11.1\%$ ) | 54.2%<br>( $\pm 33.3\%$ )        |
|          | With pretraining    | <b>91.7%</b><br>( $\pm 4.2\%$ ) | <b>78.3%</b><br>( $\pm 4.3\%$ ) | 50.0%<br>( $\pm 16.7\%$ ) | 58.3%<br>( $\pm 4.3\%$ )        | <b>88.9%</b><br>( $\pm 11.1\%$ ) | <b>58.3%</b><br>( $\pm 36.1\%$ ) |

Further, the pretrained model was evaluated on the downstream classification task. Especially, classification performance with pretraining and additional fine-tuning was compared to classification without pretraining.

We present the classification accuracy per fruit in Tab. 1.

The pretraining led, for all examples, to a performance improvement. We achieved an overall classification accuracy of 58.3%. Comparing the baseline model initially designed for pure classification to our newly proposed hybrid model with pretraining, overall, we could observe an improvement of approx. 3% in classification accuracy. For some fruit, it could be increased by more than 10%. Where this was not the case, the IQR was reduced, indicating that pretraining increased stability.

Further, experiments, visible in Fig. 3, show that pretraining even could compensate for the need for large amounts of labeled samples.

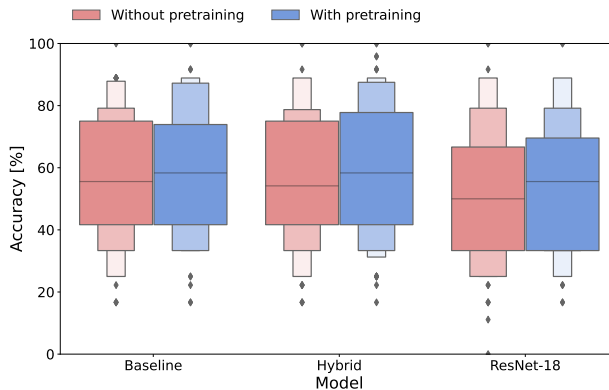


**Figure 3:** Classification accuracy (median and IQR) versus fraction of labeled samples used for classifier training for the baseline model with default classifier training (red) and hybrid model with pretraining (via *SimCLR*) plus fine-tuning (blue). Example: Avocado, *Specim* camera, ripeness classification.

## 4 Ablation Study

### 4.1 Classifier Model

For each of the three models, the classification accuracy with and without pretraining is visualized in Fig. 4.



**Figure 4:** Classification accuracies for the *HS-CNN* baseline, hybrid and *ResNet-18* model, without pretraining (red) and with pretraining via *SimCLR* (blue).

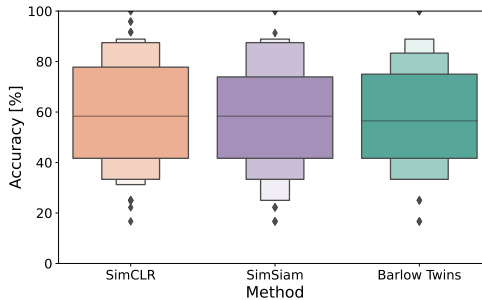
For classification without pretraining, the *HS-CNN* performs best among all three models (55.6% accuracy). With pretraining, the performance can be improved only by a small amount, probably due to the affected backbone extracting only spatial and no spectral features.

The hybrid model, with 54.2% accuracy, performs slightly worse for classification without pretraining than the baseline, possibly due to overfitting. However, more importantly, with pretraining, the accuracy improved by a larger amount – reaching equal accuracy (58.3%) and indicating that a more powerful backbone makes pretraining more effective for the hybrid variant.

The *ResNet-18* performs worse than the other two models without and with pretraining. Again, this is probably due to overfitting and spatial feature extraction. However, it has the most significant improvement (more than 5%) by pretraining.

Overall, pretraining improved the classification accuracy relative to classification without pretraining. This improvement is more significant for larger models. We claim that pretraining can prevent overfitting and enables the training of larger models.

## 4.2 Self-supervised Pretraining Method



**Figure 5:** Classification accuracies for pretraining via *SimCLR*, *SimSiam*, *Barlow Twins* using the hybrid model. Over all fruit, categories and both cameras.

Secondly, we compare the three pretraining methods employed [6–8].

Although their approaches are very different, the classification performance is rather similar (visualized in Fig. 5). Overall, *SimCLR* per-



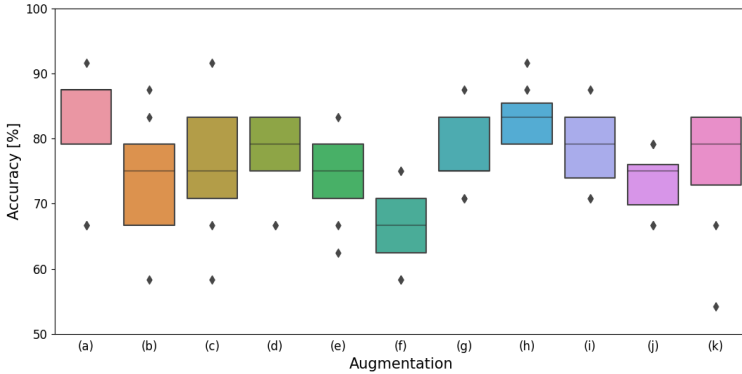
formed best, slightly better than *SimSiam*, which both have a median classification accuracy of 58.3%. *Barlow Twins* obtains only 56%.

### 4.3 Augmentations

Further, we evaluated the influence of the 21 proposed data augmentation techniques, by grouping them and using only one group for pretraining, respectively. Fig. 6 shows the resulting classification accuracies for the avocado fruit as a representative example.

The basic augmentations (rotating, flipping, cropping, and cutting) showed the highest accuracy ( $> 80\%$ ) and therefore seemed to be most important. The pixel augmentations, like the modification of edge pixels and dropping random or consecutive pixels, were also helpful for pretraining. On the other hand, dropping multiple consecutive channels led to the worst classification accuracy ( $< 70\%$ ). Also, dropping or blurring visible color channels decreased performance.

In general, distorting the spectrum resulted in low classification ac-



**Figure 6:** Classification accuracies for self-supervised pretraining (via *SimCLR*) using only the group of (a) basic augmentations, (b) noise augmentations, (c) augmentations that blur or drop random pixels, (d) drop consecutive pixels, (e) blur or drop random channels, (f) drop consecutive channels, (g) drop a sub-cube, (h) blur or drop edge pixels, (i) blur or drop edge channels, (j) blur or drop visible color information channels, and (k) mixing augmentations. Over all three SSL methods. Example: Avocado, *Specim*, ripeness classification.

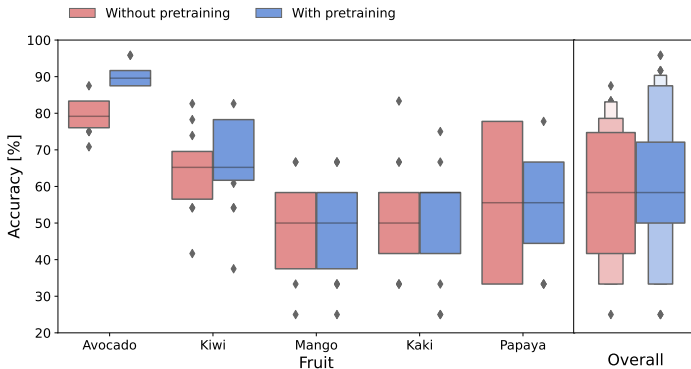
curacy. We found that, for hyperspectral image data, introducing noise systematically instead of entirely random is more valuable.

## 5 Conclusion

In this work, the hyperspectral data set of ripening fruit was extended by two new measurement series and three new fruit types.

Further, we show that it is possible to transfer the ideas of SSL to hyperspectral data. SSL pretraining extracts essential features in an unsupervised manner and allows using larger models. It can stabilize classifier training and improves the classification accuracy in some situations. Therefore, pretraining can partially compensate for the need for large labeled data sets in HSI classification.

Fig. 7 shows the improvements achieved using SSL pretraining for the ripeness classification for the five different fruit. The classification accuracy could be boosted by more than 10% for the avocados and also for the kiwis. For mangos, kakis, and papayas, the classification itself is not stable, but for the papayas as well as overall, pretraining could reduce the variability. Summarizing, the pretraining allows a more reliable ripeness classification for specific exotic fruit.



**Figure 7:** Classification accuracies for the baseline model without pretraining (red) versus the hybrid model with *SimCLR* pretraining (blue). For the *Specim* camera and the five different fruit (avocado, kiwi, mango, kaki, papaya), classified by all three categories (ripeness, firmness and sugar content).

## References

1. O. O. Olarewaju, I. Bertling, and L. S. Magwaza, "Non-destructive evaluation of avocado fruit maturity using near infrared spectroscopy and PLS regression models," *Scientia Horticulturae*, vol. 199, pp. 229–236, 2016.
2. J. Pinto, H. Rueda-Chacon, and H. Arguello, "Classification of hass avocado (*persea americana* mill) in terms of its ripening via hyperspectral images," *TecnoLogicas*, vol. 22, pp. 111 – 130, 08 2019.
3. Z. Gao, Y. Shao, G. Xuan, Y. Wang, Y. Liu, and X. Han, "Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning," *Artificial Intelligence in Agriculture*, vol. 4, pp. 31–38, 2020.
4. C. Manliguez and J. Y. Chiang, "Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation," *Sensors*, vol. 21, p. 1288, 02 2021.
5. L. A. Varga, J. Makowski, and A. Zell, "Measuring the ripeness of fruit with hyperspectral imaging and deep learning," in *International Joint Conference on Neural Networks, IJCNN 2021, Shenzhen, China, July 18-22, 2021*. IEEE, 2021, pp. 1–8.
6. T. Chen, S. Kornblith, M. Norouzi, and G. E. Hinton, "A simple framework for contrastive learning of visual representations," in *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, ser. Proceedings of Machine Learning Research, vol. 119. PMLR, 2020, pp. 1597–1607.
7. X. Chen and K. He, "Exploring simple siamese representation learning," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*. Computer Vision Foundation / IEEE, 2021, pp. 15750–15758.
8. J. Zbontar, L. Jing, I. Misra, Y. LeCun, and S. Deny, "Barlow twins: Self-supervised learning via redundancy reduction," in *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event*, ser. Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., vol. 139. PMLR, 2021, pp. 12310–12320.
9. S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-d-2-d CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote. Sens. Lett.*, vol. 17, no. 2, pp. 277–281, 2020.
10. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*. IEEE Computer Society, 2016, pp. 770–778.

11. J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Säcker, and R. Shah, "Signature verification using a "siamese" time delay neural network," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 7, no. 4, pp. 669–688, 1993.
12. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, ser. JMLR Workshop and Conference Proceedings, F. R. Bach and D. M. Blei, Eds., vol. 37. JMLR.org, 2015, pp. 448–456.
13. J. M. Haut, M. E. Paoletti, J. Plaza, A. Plaza, and J. Li, "Hyperspectral image classification using random occlusion data augmentation," *IEEE Geosci. Remote. Sens. Lett.*, vol. 16, no. 11, pp. 1751–1755, 2019.
14. H. Zhang, M. Cissé, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," in *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.
15. X. Wang, H. Fan, Y. Tian, D. Kihara, and X. Chen, "On the importance of asymmetry for siamese representation learning," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*. IEEE, 2022, pp. 16 549–16 558.
16. J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 462–466, 1952.
17. I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," *CoRR*, vol. abs/1608.03983, 2016.
18. E. Fix and J. L. Hodges, "Discriminatory analysis. Nonparametric discrimination; consistency properties," USAF School of Aviation Medicine, Randolph Field, Texas, Tech. Rep., 1951.
19. Z. Wu, Y. Xiong, S. X. Yu, and D. Lin, "Unsupervised feature learning via non-parametric instance-level discrimination," *CoRR*, vol. abs/1805.01978, 2018.
20. A. G. Howard, "Some improvements on deep convolutional neural network based image classification," in *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2014.

# Thermographic Techniques to Explore Small-Scale Processes at Water Surfaces

Bernd Jähne<sup>1,2</sup>, Lucas Warmuth<sup>1</sup>, Roman Stewing<sup>1</sup>, and Kerstin E. Krall<sup>1</sup>

<sup>1</sup> Heidelberg University, Institute for Environmental Physics  
Im Neuenheimer Feld 229, 69120 Heidelberg

<sup>2</sup> Heidelberg University, Interdisciplinary Center for Scientific Computing  
Berliner Straße 43, 69120 Heidelberg

**Abstract** Techniques based on thermography are well-established for destruction-free material inspection. A similar technique was invented independently in environmental sciences to explore exchange processes at air-water interfaces. The analysis was, however, limited to one-dimensional vertical transport assuming a horizontally homogeneous and stationary exchange process on average. In this contribution, first steps pursuing a true spatio-temporal approach are presented. This allows much faster measurements, identification of the transport mechanisms and has the prospect to even measure the shear stress right at the water surface, which drives exchange processes at a windy water surface.

**Keywords** Thermography, Lock-In Technique, Heat Transport, Interface

## 1 Introduction

Lock-in thermography and heat flux thermography are well-established techniques for destruction-free material inspection [1, 2]. A periodically varying or flashed heat flux is applied at the surface of an object and the temperature response of the surface is captured with a thermographic camera. The applied heat at the surface diffuses into the material of the object. Above cracks, holes or other material inhomogeneities with lower heat conduction, the material surface remains

warmer. In this way, it is possible to look below the surface of opaque materials.

It is less known that similar techniques were invented independently in environmental sciences [3,4] to explore exchange processes on ocean, lake, and river surfaces or in laboratory simulation facilities such as wind-wave tunnels. Water would be a perfectly homogeneous material without any flow, because the applied heat at the water surface just diffuses into the bulk of the water body. In reality, turbulent transport processes cause inhomogeneous heat flow at the surface.

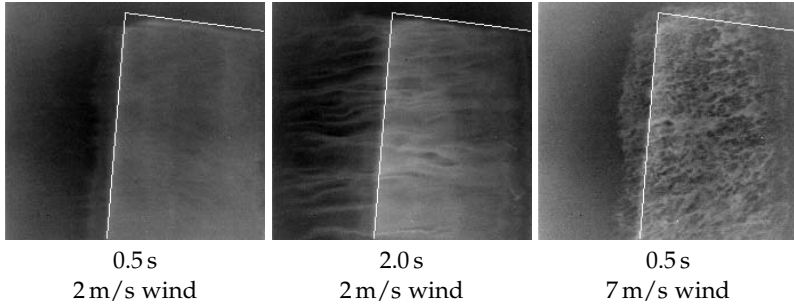
Section 2 briefly explains the basics of thermography to explore turbulent transport processes across the air-water interface and the established technique with periodic heating. Then, two new approaches are discussed: a direct analysis of the intermittent transport process under spatially constant irradiation (Section 3) and a line-shaped irradiation to measure the water surface velocity and the gradient of the shear flow (Section 4).

## 2 Basics

The basic characteristic of transport processes across interfaces is that turbulent transport becomes less efficient closer to the interface because turbulent fluctuations (“eddies”) become smaller in size. Below a certain scale, turbulent fluctuations are even damped by viscosity. This leads to the formation of a viscous boundary layer. Therefore, the final transport to the interface can only take place by molecular diffusion.

This basic characteristic of the transport process can be seen in thermographic images, taken after a constant heat flux density was applied to the interface for a certain time. This can be done, for instance, by irradiating the water surface using a CO<sub>2</sub> laser beam expanded to an area of up to a square meter. The radiation penetrates only 14 μm into the water. That means that the controllable heat flux density is placed directly at the surface. An MWIR thermal camera images the water surface temperature over a slightly deeper layer [5]. The 10.6 μm laser radiation is not directly detected in the surface temperature images, because the camera is sensitive only in the 3–5 μm wavelength region.

After 0.5 s, at a low turbulence level with a wind speed of 2 m/s, the heat has penetrated only such a short distance into the water, that it is



**Figure 1:** Temperature increase at the water surface in the Heidelberg Aeolotron wind-wave tank. The area heated by a  $\text{CO}_2$  laser (about  $25\text{ cm} \times 25\text{ cm}$ ) is marked by white outline. The time after switching on the laser and the wind speed applied to the water surface is given below the images.

still inside the viscous boundary layer. Because heat conduction into the water is driven only by molecular diffusion, the surface temperature in the heated area is uniform (Figure 1, left image). After a four times longer time span (2 s, Figure 1, middle image), the heat has penetrated about twice the distance into the water. Now the influence of the turbulent heat transport in deeper layers starts to become visible. At a higher turbulence level with a wind speed of 7 m/s, the turbulent structures can already be seen 0.5 s after switching on the heat flux and exhibit a much finer scale and different patterns (Figure 1, right image). With a higher wind speed, the induced velocity gradient at the water surface is steeper and turbulence comes closer to the interface.

Previous research of the controlled flux technique has not looked into the evolution of these structures, but rather used it for fast measurements of the speed of heat exchange, expressed by the *transfer velocity*  $k$  (units m/s), in wind-wave facilities [6] and at sea [7]. This is because heat can be used as a proxy tracer for environment- and climate-relevant trace gases exchanging across the atmosphere-ocean interface with this technique. All other field measuring techniques integrate and average over much larger spatial and temporal scales [5]. By considering the different diffusion coefficients of heat and gases dissolved in water, the transfer velocity of gases can be computed from those for heat [8].

The periodic variation of the heat flux by a CO<sub>2</sub> laser — or lock-in technique — has the advantage that all the information about the response of the system is contained in the switching frequencies and its higher harmonics. Constant or randomly fluctuating heat flux densities by sensible heat transfer, latent heat transfer (evaporation) or radiative cooling into the sky, are the more suppressed, the longer the amplitude variation is measured.

At low switching frequencies, the heat response at the surface can follow the applied heat flux density  $j$  and reaches a constant temperature increase of

$$\Delta T = \frac{j}{\rho c_p k} \quad \rightsquigarrow \quad k = \frac{j}{\rho c_p \Delta T}, \quad (1)$$

so that  $k$  can be determined if the heat flux density is known;  $\rho$  is the density and  $c_p$  the specific heat capacity of water. If the switching frequencies are increased beyond a critical frequency  $\nu_c$ , the amplitude of the temperature response starts to decrease. Finally, the penetration depth becomes so shallow that the response is no longer determined by turbulence but only by molecular diffusion. Then the temperature amplitude response  $\Delta T$  is given by [4]

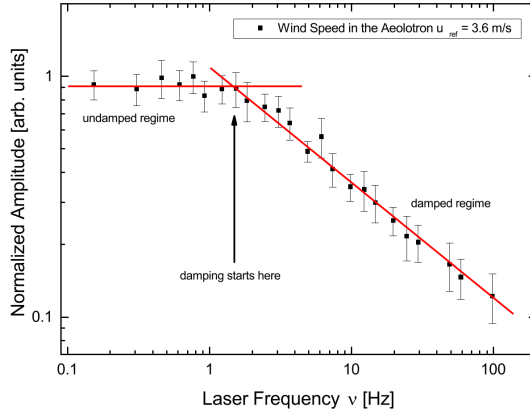
$$\Delta T = \frac{j}{\rho c_p (2\pi\nu D_h)^{1/2}}. \quad (2)$$

$D_h$  is the molecular diffusion coefficient for heat in water (thermal diffusivity). The frequency response is therefore similar to a low-pass filter. However, the amplitude response for higher frequencies does not decrease with  $\nu^{-1}$  but slower, only with  $\nu^{-1/2}$ . The asymptotic constant and the damped parts of the amplitude response curve meet at the critical frequency  $\nu_c$  (Figure 2). Eqs. (1) and (2) yield

$$\nu_c = \frac{k^2}{2\pi D_h} \quad \rightsquigarrow \quad k = \sqrt{2\pi\nu_c D_h}. \quad (3)$$

This means that the transfer velocity  $k$  can also be computed from the measurement of the amplitude response without any knowledge about the heat flux density  $j$ . Figure 2 also shows that transport across the thin heat boundary layer at the water surface is quite fast. Up to frequencies of 1 Hz the amplitude response shows no damping.





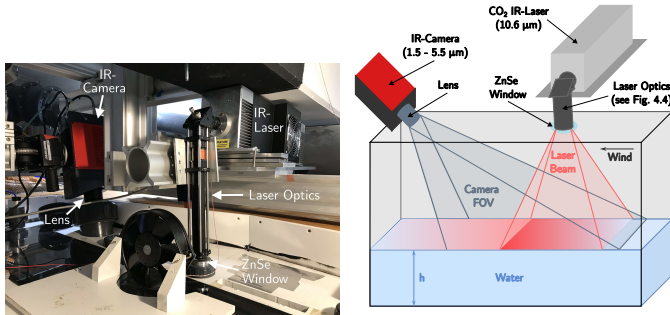
**Figure 2:** Frequency response of the heat boundary layer at the water surface for frequencies between 0.1 to 100 Hz; from [9].

### 3 Analysis of intermittency

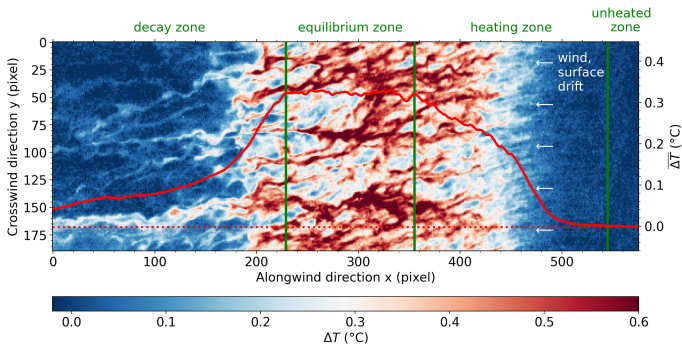
The approach discussed so far has still two deficits. Firstly, the measurements are still quite slow, because averaging over several periods of the periodic heating and a frequency sweep are required. Secondly, horizontal averaging over the heated footprint is performed. The averaging over both temporal and spatial scales misses all the important information contained in the patterns.

In this paper, two first steps into a true spatio-temporal analysis are described. The setup used for these measurements is shown in Figure 3. At the water surface, the camera images an area larger than the area heated by the CO<sub>2</sub> laser. Because of the drift of the water induced by the wind, a characteristic temperature profile averaged perpendicular to the wind direction and time establishes (red line in Figure 4). There is a heating zone characterized by an increase in temperature followed by an equilibrium zone with more or less constant temperature. After the water leaves the heated zone, the mean temperature decays again.

The analysis here is limited to the equilibrium zone averaged only over 25 images taken with a frame rate of 600 Hz. This arrangement



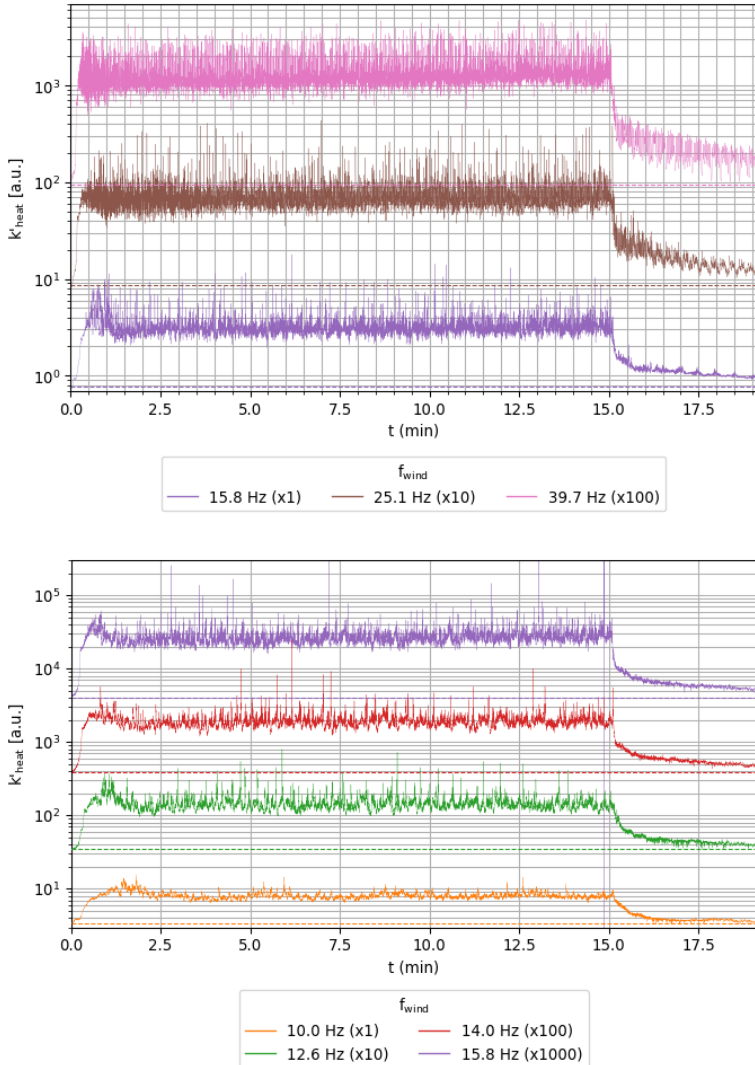
**Figure 3:** Setup of thermography at the ceiling of the Heidelberg Aeolotron; from [10]



**Figure 4:** Temperature response at water surface by heating an area of about  $60\text{ cm} \times 60\text{ cm}$ . Wind direction is right to left; from [10].

made it possible to measure the transfer velocity instantaneously according to Eq. (1) with a temporal resolution of  $0.042\text{ s}$ . This is faster than the time constant of the transfer process.

A few seconds after the measurements were started, the wind was switched on and within several seconds the transfer velocity jumped up (Figure 5). At the lowest wind speed, the transfer velocity remains quite constant, whereas with increasing wind speed more and more spikes with up to 10 times higher transfer velocity show up. They could be related to extensive turbulent mixing at the surface caused by micro-scale wave breaking events (wave breaking without bubble en-



**Figure 5:** Instantaneous transfer velocities  $k$  measured at different wind speeds (indicated by the drive frequency of the wind fans) in the Heidelberg Aeolotron with a water depth of 32 cm. The wind was switched on a few seconds after the start of the measurement and was kept on for 15 min; from [10].

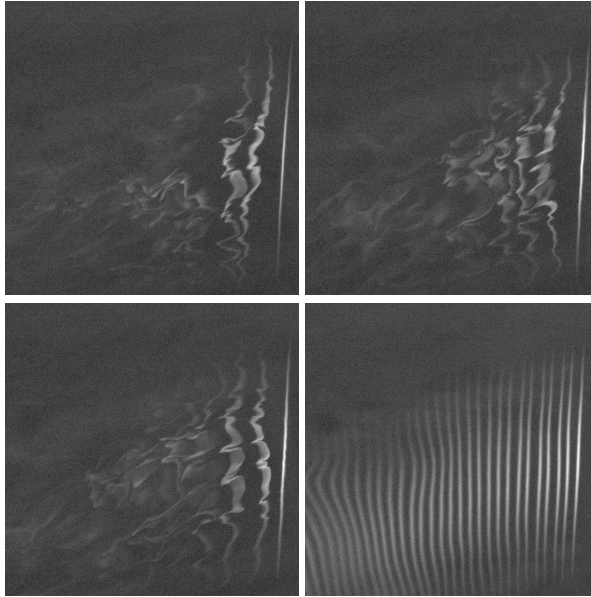
traintment). After the start of the wind, the wind-wave field gradually evolves from small ripples to larger and larger gravity waves. Except for the initial waves at medium wind speeds, where a clear overshoot of the transfer velocity is observed, the transfer velocity is remarkably insensitive to status of the wind wave field. When the wind is stopped after 15 min, the transfer velocity immediately decreases.

#### **4 Analysis of the shear current at the interface**

The measurements shown above, clearly demonstrate that the wind is the main driver of the transport process. The wind induces a shear flow at the water interface within the aqueous viscous boundary layer. This shear layer can also be investigated using thermography. The key idea is to heat up only a line perpendicular to the wind direction at the water surface with a penetration depth for the radiation of about one millimeter matching the thickness of the viscous mass boundary and to apply a short pulse of a few milliseconds, which yields a very thin heated line. If only the surface was heated up by a CO<sub>2</sub> laser, the line would quickly disappear because of vertical diffusion into the water. With the deeper penetration depth used here, vertical diffusion is not dominant so that the horizontal transport in the shear layer can be studied. An Erbium fiber laser with a wavelength in the near infrared (1568 nm) is used, which has a penetration depth of 1.0 mm.

Stewing [11] showed that the widening of the lines at the water surface is proportional with the diffusion of heat in horizontal direction, as long as there is no shear current at the water surface, but only the water body as a whole moves in the water channel of a wind-wave facility (Figure 6, lower left image). This is already the case a few seconds after the wind is turned off. Because of inertia, the water body continues to move and decreases its velocity only slowly [12].

With a wind-induced shear current at the water surface, the situation is completely different (Figure 6, first three images). Because of the velocity gradient at the water surface, different parts of the heated line move with different velocities. Although only the heated line at the water surface is seen, the slower moving parts now diffuse also vertically towards the surface. The result is that the line widens much faster in flow direction and its temperature drops much faster. The complexity



**Figure 6:** Evolution of heated lines produced by a 100 W 1540 nm fiber laser with 10 ms duration every 200 ms at a low wind speed in the Heidelberg Aeolotron; lower left thermal image seconds after the wind has been turned off; image sector about 20 cm  $\times$  20 cm.

of the velocity field at the water surface influenced by a wind-induced shear current together with wind-induced waves can be seen and studied in these images. The flow field at the surface is turbulent and there are thin streaks in wind direction with much higher velocity.

## 5 Conclusions and outlook

The active thermography techniques described here show how powerful this optical inspection methods are. They allow a detailed analysis of complex flow fields and transport processes at free interfaces and can look below the surface. This progress in experimental techniques for environmental research may also stimulate new approaches in engineering sciences and material inspection.

## References

1. O. Breitenstein, W. Warta, and M. Langenkamp, *Lock-in Thermography*, 2nd ed. Springer Berlin Heidelberg, 2010.
2. M. Sackewitz, Ed., *Handbuch zur industriellen Bildverarbeitung*, 3rd ed. Fraunhofer Verlag, 2017.
3. P. Libner, B. Jähne, and E. Plate, "Ein neue Methode zur lokalen und momentanen Bestimmung der Wiederbelüftungsraten von Gewässern," *Wasserwirtschaft*, vol. 77, no. 5, pp. 230–235, 1987.
4. B. Jähne, P. Libner, R. Fischer, T. Billen, and E. J. Plate, "Investigating the transfer process across the free aqueous boundary layer by the controlled flux method," *Tellus*, vol. 41B, no. 2, pp. 177–195, 1989.
5. B. Jähne, L. Warmuth, and K. E. Krall, "Exploring the mechanisms of air-sea gas transfer using thermographic techniques," in *Air-Water Gas Exchange: Cross-linking field measurements, laboratory measurements and modeling*, B. Jähne, K. E. Krall, and C. Marandino, Eds. Heidelberg University Publishing, 2023, in press.
6. J. Kunz and B. Jähne, "Investigating small scale air-sea exchange processes via thermography," *Front. Mech. Eng.*, vol. 4, p. 4, 2018.
7. L. Nagel, K. E. Krall, and B. Jähne, "Measurement of air-sea gas transfer velocities in the Baltic Sea," *Ocean Sci.*, vol. 15, pp. 235–247, 2019.
8. L. Nagel, K. E. Krall, and B. Jähne, "Comparative heat and gas exchange measurements in the Heidelberg Aeolotron, a large annular wind-wave tank," *os*, vol. 11, pp. 111–120, 2015.
9. J. Kunz, "Active Thermography as a Tool for the Estimation of Air-Water Transfer Velocities," Ph.D. dissertation, Institut für Umweltphysik, Fakultät für Physik und Astronomie, Univ. Heidelberg, 2017.
10. L. Warmuth, "Spatio-temporal analysis of air-sea gas transfer with active thermography," Master's thesis, Institut für Umweltphysik, Universität Heidelberg, Germany, 2022.
11. R. Stewing, "Optimierung thermografischer Messmethoden zur Geschwindigkeits- und Geschwindigkeitsgradientenmessung," Bachelor's thesis, Institut für Umweltphysik, Universität Heidelberg, Germany, 2023.
12. M. Bopp, "Luft- und wasserseitige Strömungsverhältnisse im ringförmigen Heidelberger Wind-Wellen-Kanal (Aeolotron)," Master's thesis, Institut für Umweltphysik, Universität Heidelberg, Germany, 2014.

# Sensitivity enhanced glucose sensing by return-path Mueller matrix ellipsometry

Chia-Wei Chen<sup>1,2</sup>, Matthias Hartrumpf<sup>2</sup>, Thomas Längle<sup>2</sup>, and Jürgen Beyerer<sup>1,2</sup>

<sup>1</sup> Karlsruhe Institute of Technology (KIT), Vision and Fusion Laboratory (IES),  
Haid-und-Neu-Str. 7, 76131 Karlsruhe, Germany

<sup>2</sup> Fraunhofer Institute of Optronics, System Technologies and Image  
Exploitation IOSB,  
Fraunhoferstraße 1, 76131 Karlsruhe, Germany

**Abstract** Diabetes is a worldwide public health problem. According to the survey of the Robert Koch Institute, in Germany, at least 7.2 percent population (aged between 18 to 79 years) have diabetes. Therefore, the demand for glucose monitoring is increasing, especially for non-invasive glucose monitoring technology. In this work, we proposed a novel method to enhance the sensitivity of glucose monitoring by return-path ellipsometry with a quarter-wave plate and mirror. The coaxial design improves the sensitivity and reduces the complexity of optical system alignment by means of a fixed quarter-wave plate. The proposed system showed higher sensitivity compared to the transmission configuration.

**Keywords** Glucose measurement, Mueller matrix, return-path ellipsometry, optical polarimetry

## 1 Introduction

Diabetes is a worldwide public health problem. According to the survey of the Robert Koch Institute, in Germany, at least 7.2 percent population (aged between 18 to 79 years) have diabetes [1]. Diabetes patients cannot regulate their blood glucose levels when their blood sugar goes up. High blood sugar levels staying too long in the bloodstream cause

serious health problems, such as nerve damage, vision loss, and kidney disease. Therefore, regular self-monitoring of blood glucose (SMBG) is essential in managing diabetes.

SMBG can be categorized into two types: invasive and non-invasive methods. The former methods include blood glucose monitoring and skin-attachable glucose sensors. However, these methods might cause discomfort and skin irritation which increase the risk of skin or tissue damage. Hence, the development of non-invasive glucose monitoring has been increasing in recent years. In the literature, the non-invasive methods of SMBG found are optical polarimetry [2], optical coherence tomography [3], Raman spectroscopy [4] and surface plasmon resonance [5]. Compared to these methods, the advantages of optical polarimetry are wide detection range, simple setup and capability of high scattering effects and weak signals. Nevertheless, the limitation of optical polarimetry is the resolution of glucose concentration. According to the guideline from Food and Drug Administration (FDA) in the United States, a minimum accuracy of 12 mg/dl is required for blood glucose monitoring test systems [6]. Phan and Lo used the Stokes-Mueller matrix polarimetry system to measure glucose concentration and claimed the limitation was 20 mg/dl [7]. Mukherjee et al. achieved a sensitivity of 20 mg/dl by a Mueller matrix polarimeter with dual photoelastic modulators [8]. Al-Hafidh et al. developed multireflection polarimetry which used micromirrors to enlarge the optical path length. They can achieve a 30-fold enhancement with 11 reflections [9]. However, their system required 11 mirrors which increase the complexity of assembly, alignment and calibration. In this work, we proposed a simple method to enhance the sensitivity of glucose monitoring by means of a quarter-wave plate and mirror. The design is based on a coaxial design which can be easily applied to current optical polarimetry.

## 2 Measurement principle

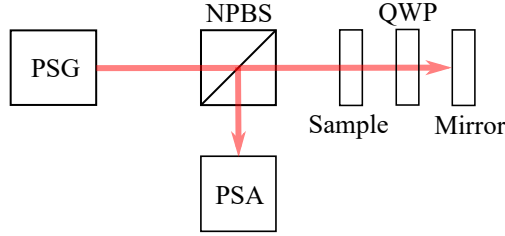
The principle of optical polarimetry is based on the property of optical activity of glucose solution, i.e., the change of optical rotation is related to the concentration of the glucose concentration. The phenomenon can be described as [9]

$$\alpha = CL[\alpha]_{\lambda}^T, \quad (1)$$



where  $\alpha$  is the measured optical rotation,  $C$  is the concentration of the solution,  $L$  is the optical path length and  $[\alpha]_{\lambda}^T$  is the rotation power of the chiral material (e.g., sugar and glucose) which is related to temperature  $T$  and wavelength  $\lambda$  of the light source. Therefore, for low concentrations of glucose, high accuracy and sensitivity measurements for optical rotation are required.

Inspired by the concept of Chen et al. [10], we improve the measurement sensitivity of the optical rotation for glucose solution by return-path ellipsometry (RPE) [11]. In the configuration of RPE, the light beam transmits through the sample and returns by reflecting optical elements. Compared to conventional ellipsometry, the main feature is that RPE has a higher sensitivity to the optical properties of samples because of the double reflection from the sample.



**Figure 1:** The schematic of the proposed return-path ellipsometry.

Figure 1 shows the schematic of the proposed return-path ellipsometer, which consists of a polarization state generator (PSG), non-polarizing beamsplitter (NPBS), quarter-wave plate (QWP), mirror and polarization state analyzer (PSA). The polarization effect of optical elements or interaction at boundaries can be described by Stokes vectors and Mueller matrices [12]. Stoke vectors  $\mathbf{S}$  describe the polarization state of light beams.  $s_0$  represents the total intensity.  $s_1$ ,  $s_2$  and  $s_3$  denote the relative difference (linear or circular). Mueller matrices  $\mathbf{M}$  represent the characteristics of the altering of Stokes vectors when light interacts with matter.

$$\mathbf{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix}, \mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}. \quad (2)$$

The PSG can generate light with different polarization states  $\mathbf{S}_{\text{PSG}}$  and the PSA can measure the state of polarization of light  $\mathbf{S}_{\text{PSA}}$ . Then, the measured Mueller matrix can be obtained by

$$\mathbf{S}_{\text{PSA}} = \mathbf{M}_{\text{meas}} \cdot \mathbf{S}_{\text{PSG}}. \quad (3)$$

The measured Mueller matrix  $\mathbf{M}_{\text{meas}}$  in the return-path ellipsometry can be described as

$$\mathbf{M}_{\text{meas}} = \mathbf{M}_{\text{BS}}^t \cdot \mathbf{M}_{\text{S}}(\alpha) \cdot \mathbf{M}_{\text{QWP}}(-\theta) \cdot \mathbf{M}_{\text{M}} \cdot \mathbf{M}_{\text{QWP}}(\theta) \cdot \mathbf{M}_{\text{S}}(\alpha) \cdot \mathbf{M}_{\text{BS}}^i, \quad (4)$$

where  $\mathbf{M}_{\text{BS}}$ ,  $\mathbf{M}_{\text{QWP}}(\theta)$  and  $\mathbf{M}_{\text{M}}$  are the Muller matrices of the NPBS, QWP and mirror, and  $r$ ,  $t$  and  $\theta$  denote the reflection and transmission of the NPBS and fast-axis orientation angle of the QWP. It should be noted that the Mueller matrix of optically active medium is the same for propagation and propagation back to the medium [13]. If every optical element is ideal,  $\mathbf{M}_{\text{BS}}^i$  and  $\mathbf{M}_{\text{M}}$  are diagonal matrices, where the diagonal elements are 1, 1,  $-1$ , and  $-1$ .  $\mathbf{M}_{\text{BS}}^t$  is a diagonal matrix with diagonal elements 1, 1, 1, and 1. For simplicity, the Mueller matrix of an optically active medium can be treated as a circular retarder [8]

$$\mathbf{M}_{\text{S}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

The QWP whose retardance is  $90^\circ$  can be expressed as

$$\mathbf{M}_{\text{QWP}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta & \cos 2\theta \sin 2\theta & \sin 2\theta \\ 0 & \cos 2\theta \sin 2\theta & \sin^2 2\theta & -\cos 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \end{bmatrix}. \quad (6)$$

If the fast axis  $\theta$  is 0, the measurement result can be simplified as

$$\mathbf{M}_{\text{meas}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha & 0 \\ 0 & \sin 2\alpha & -\cos 2\alpha & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (7)$$

From Eqs. 5 and 7, it is clear that the measured rotation angles by the return-path ellipsometry are twofold compared to the measured rotation angles by the transmission configuration because the optical path length increases twice. Therefore, with the return-path configuration, we can enhance the sensitivity two times. In Eq. 7, the rotation angle of the glucose concentration can be calculated by

$$\arctan \frac{m_{32}}{m_{22}} = \arctan \frac{-m_{23}}{m_{33}} \quad (8)$$

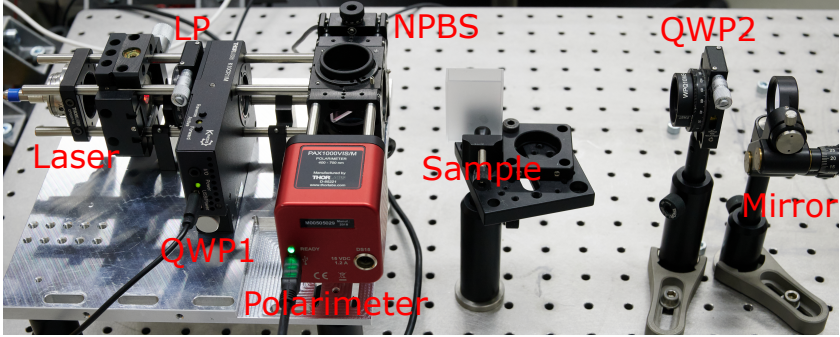
It is worth noting that if the QWP in the configuration is removed, the measured Mueller matrix becomes a  $4 \times 4$  identity matrix, i.e., the sensor cannot measure the rotation angle induced by the optically active medium.

### 3 Experiment setup

Figure 2 shows a prototype of a return-path ellipsometer. The principle is based on dual rotating-compensator [14] and return-path Mueller matrix ellipsometry. Therefore, the ellipsometer can measure full Mueller matrices [15] and the optical rotation can be solved by the measured matrices. The setup consists of a laser with a wavelength of 638 nm from Integrated Optics, a linear polarizer (LPVISE100-A, Thorlabs, Inc.), an NPBS, two QWPs (WPQ10ME-633, Thorlabs, Inc.), a silver mirror (PF10-03-P01, Thorlabs, Inc.) and a Stokes polarimeter (PAX1000VIS, Thorlabs, Inc.). QWP1 is mounted on a stepper motor rotation mount (K10CR1, Thorlabs, Inc.). The sample is a cuvette with an optical path length of 30 mm.

### 4 Experimental results

Before the measurements of glucose concentration, the NPBS and QWP2 need to be calibrated first. The NPBS has strong polarization distortions which induce polarization changes in the measurements and cause calculation errors. The calibration procedure of the NPBS can be found in Ref [16]. The measured Mueller matrix of the NPBS is



**Figure 2:** Photograph of the return-path ellipsometer, where LP, QWP and NPBS are linear polarizer, quarter-wave plate and non-polarized beamsplitter, respectively.

shown as

$$\mathbf{M}_{\text{NPBS}} = \begin{bmatrix} 1 & -0.167 & -0.004 & 0.002 \\ -0.175 & 1.010 & 0.006 & 0.002 \\ 0.002 & -0.005 & -0.981 & -0.251 \\ -0.003 & 0.005 & 0.261 & -0.945 \end{bmatrix}. \quad (9)$$

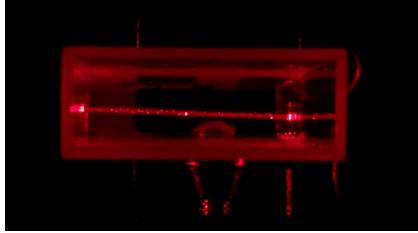
As can be seen, the NPBS is not a perfect element. Therefore, careful calibration of each optical element in the system is necessary and important. In Section 2, the fast axis of the QWP should be adjusted to  $0^\circ$ . Then the product of  $\mathbf{M}_{\text{QWP}}$ ,  $\mathbf{M}_{\text{M}}$  and  $\mathbf{M}_{\text{QWP}}$  is a  $4 \times 4$  identity matrix. After the fast axis adjustment of the QWP, We obtained the Mueller matrix as

$$\mathbf{M}_{\text{QWP}} \cdot \mathbf{M}_{\text{M}} \cdot \mathbf{M}_{\text{QWP}} = \begin{bmatrix} 1 & -0.003 & 0.014 & -0.009 \\ 0.010 & 0.992 & -0.003 & -0.004 \\ 0.009 & 0.004 & 0.996 & 0.034 \\ -0.009 & 0.004 & -0.041 & 0.993 \end{bmatrix}. \quad (10)$$

The result is very close to the ideal condition ( $4 \times 4$  identity matrix). The error sources might be the alignment and wavelength mismatch between the laser and the QWP.

In glucose concentration measurements, the glucose solution of 5% from B. Braun SE was first placed in a quartz cuvette with an optical

path length of 30 mm and a wall thickness of 10 mm. Deionized water was used to dissolve the glucose concentration to 50 mg/ dl, 117 mg/dl and 150 mg/dl. An additional sample with deionized water was prepared for reference. An ultrasonic bath was used to speed up the dissolving process. Figure 3 shows the measurement of the glucose concentration. For the transmission measurements, the laser beam only passes the cuvette once. For the return-path measurements, the laser beam passes the cuvette forward and backward.



**Figure 3:** Photograph of the glucose measurements.

Figure 4 shows the measurement results for optical rotation angles with different glucose concentrations by the transmission and return-path ellipsometers. Table 1 demonstrates the fitting result (linear fitting) of the measurements. It can be seen that the slope of the return-path configuration (0.0047) is higher than the slope of the transmission configuration (0.0014), which proves the concept of sensitivity enhancement for glucose sensing. The coefficients of determination ( $R^2$ ) in both methods are close to 1, i.e., the polarization model derived in Section 2 can well explain the optical rotation for different glucose concentrations.

**Table 1:** Fitting results for optical rotation angles with different glucose concentrations by the transmission and return-path ellipsometers.

| Configuration | Linear fitting         | $R^2$ |
|---------------|------------------------|-------|
| Transmission  | $y = 0.0014x + 0.0116$ | 0.98  |
| Return-path   | $y = 0.0047x - 0.0501$ | 0.93  |

However, the accuracy of the return-path configuration is lower than the accuracy of the transmission configuration. The reasons might be

the alignment of the cuvette and the temperature of the glucose concentration. Because of the return-path configuration, the laser beam will pass the cuvette twice with four boundaries. If there is a small alignment error, the cuvette might induce polarization errors. As shown in the literature [13], the glucose concentration is sensitive to the temperature which was not controlled in the experiments. In addition, a pipette is used to transport a measured volume of the deionized water and glucose solution to the cuvette. The maximum permissible systematic error and random error of the pipette are  $\pm 0.5\%$  and  $\pm 0.15\%$  which might lead deviations of the concentration.

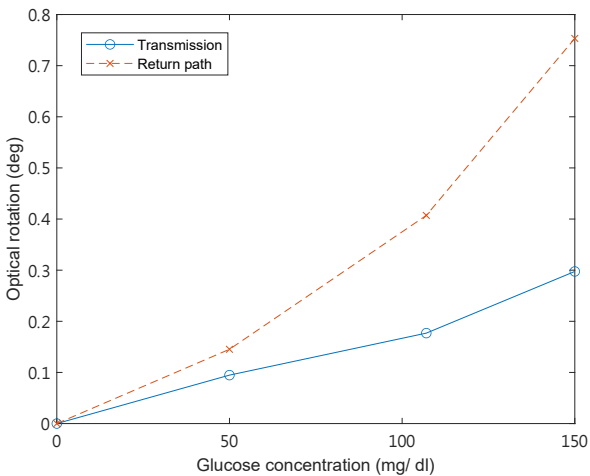


Figure 4: Measurement results of optical rotation for different glucose concentrations.

## 5 Conclusion

In this work, we proposed a novel glucose sensor by return-path Mueller matrix ellipsometry. Compared to the work from Phan and Lo and Mukherjee et al. (transmission Mueller matrix ellipsometry), the sensitivity of the measured rotations angle increases two times because the light passes the sample forward and backward. In principle,

if the return-path configuration is applied to their systems, the sensitivity of their systems can be enhanced to 10 mg/dl which fulfills the FDA regulation. The proposed sensor uses a coaxial design, decreasing the optical system alignment's complexity. The measurement sensitivity is enhanced by using a fixed QWP (fast axis 0) and a mirror, i.e., the optical path length is twofold. For high-speed measurements, a liquid crystal or a division-of-amplitude photopolarimeter can be used to achieve several  $\mu\text{s}$  per Stokes vector. Currently, we only use the glucose concentration which has no scattering and depolarization effect. For real applications, both effects should be taken into account. Therefore, we will add intralipid with different glucose concentrations for the next step. In the future, we plan to evaluate the sensitivity, accuracy and uncertainty of the glucose sensor and study the calibration and stability of the system.

## References

1. C. Heidemann and C. Scheidt-Nave, "Prevalence, incidence and mortality of diabetes mellitus in adults in germany – a review in the framework of the diabetes surveillanc," in *Journal of Health Monitoring*. Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung, 2017, vol. 2.
2. T.-T.-H. Pham and Y.-L. Lo, "Extraction of effective parameters of turbid media utilizing the mueller matrix approach: study of glucose sensing," *Journal of Biomedical Optics*, vol. 17, no. 9, pp. 97 002–1, 2012.
3. R. O. Esenaliev, K. V. Larin, I. V. Larina, and M. Motamedi, "Noninvasive monitoring of glucose concentration with optical coherence tomography," *Optics letters*, vol. 26, no. 13, pp. 992–994, 2001.
4. A. M. Enejder, T. G. Scecina, J. Oh, M. Hunter, W. Shih, S. Sasic, G. L. Horowitz, and M. S. Feld, "Raman spectroscopy for noninvasive glucose measurements," *Journal of biomedical optics*, vol. 10, no. 3, p. 031114, 2005.
5. W. Lam, L. Chu, C. Wong, and Y. Zhang, "A surface plasmon resonance system for the measurement of glucose in aqueous solution," *Sensors and Actuators B: Chemical*, vol. 105, no. 2, pp. 138–143, 2005.
6. "Blood glucose monitoring test sys. for prescription point-of-care use." [Online]. Available: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/blood-glucose-monitoring-test-systems-prescription-point-care-use>

7. Q.-H. Phan and Y.-L. Lo, "Stokes-mueller matrix polarimetry system for glucose sensing," *Optics and Lasers in Engineering*, vol. 92, pp. 120–128, 2017.
8. P. Mukherjee, N. Hagen, and Y. Otani, "Glucose sensing in the presence of scattering by analyzing a partial mueller matrix," *Optik - International Journal for Light and Electron Optics*, vol. 180, pp. 775–781, 2019.
9. M. H. Al-Hafidh, A. Glidle, R. Wilson, A. E. Kelly, J. Reboud, and J. M. Cooper, "Multireflection polarimetry in microfluidics," *IEEE Sensors Letters*, vol. 3, no. 10, pp. 1–4, 2019.
10. C.-W. Chen, M. Hartrumpf, T. Längle, and J. Beyerer, "Sensitivity enhanced roll-angle sensor by means of a quarter-waveplate," *tm-Technisches Messen*, vol. 88, no. s1, pp. s48–s52, 2021.
11. R. Azzam, "Return-path ellipsometry and a novel normal-incidence null ellipsometer (nine)," *Optica Acta: International Journal of Optics*, vol. 24, no. 10, pp. 1039–1049, 1977.
12. H. Fujiwara, *Spectroscopic ellipsometry: Principles and applications*. Chichester, England and Hoboken, NJ: John Wiley & Sons, 2007.
13. D. H. Goldstein, *Polarized light*, third edition ed. Boca Raton, Fla. and London and New York: CRC Press, 2011.
14. P. S. Hauge, "Recent developments in instrumentation in ellipsometry," *Surface Science*, vol. 96, no. 1-3, pp. 108–140, 1980.
15. C.-W. Chen, M. Hartrumpf, T. Längle, and J. Beyerer, "Retroreflex ellipsometry for isotropic substrates with nonplanar surfaces," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 38, no. 1, p. 014005, 2020.
16. X. Chen, J. Liao, H. Gu, Y. Shi, H. Jiang, and S. Liu, "Proof of principle of an optical stokes absolute roll-angle sensor with ultra-large measuring range," *Sensors and Actuators A: Physical*, vol. 291, pp. 144–149, 2019.



# Fluorescence Imaging of Concentration Fields of Dissolved Gases at Water Interfaces

Dennis Hofmann<sup>1</sup> and Bernd Jähne<sup>1,2</sup>

<sup>1</sup> Heidelberg University, Institute of Environmental Physics  
Im Neuenheimer Feld 229, 69120 Heidelberg

<sup>2</sup> Heidelberg University, Interdisciplinary Center for Scientific Computing  
Berliner Straße 43, 69120 Heidelberg

**Abstract** Characterization of materials at interfaces includes also transfer processes taking place there. This is an ubiquitous phenomenon in the environment, technical processes, and living species. Because at least one of the two phases at the interface is mobile, these processes are characterized by a complex interplay between molecular diffusion and turbulent transport. In this paper, a new technique is introduced for fluorescence imaging of the mass transfer across the air-water interface.

**Keywords** Fluorescence Imaging, Interface, Mass Transfer

## 1 Introduction

The characterization of materials by optical techniques, as it was presented at all previous OCM conferences [1] contains a wide range of material properties, wavelengths from x-rays to thermal infrared, and a remarkable wealth of different optical effects, e. g., the refractive index, reflectance, emission, absorption, fluorescence, elastic and inelastic scattering. With this wealth of techniques quite different material properties can be investigated. This includes the concentration of various chemical species, classification of different materials for sorting, 3-D surface shape and surface contamination, to name just a few.

Dynamic material properties are so far missing. An important class of dynamic processes is the exchange of mass across the interface from one medium into another. Here the question is how fast is this process

and which factors are controlling its mechanisms. At first glance this property might appear quite exotic, but it is actually an ubiquitous process:

**Environment** In environmental sciences, the exchange of mass between the compartments of the planet earth, i. e., land, oceans, lakes, rivers, and atmosphere is an important process [2]. What controls evaporation from water or land surfaces? How much of the climate relevant gas species emitted by human activities into the atmosphere are transferred into oceans, biosphere and finally into sediments? The most prominent example of this kind is the global carbon cycle.

**Engineering** In chemical engineering transfer processes are relevant for any gas-liquid, gas-solid, liquid-liquid (immiscible liquids) and liquid-solid reactions. Here the essential question is how to design the corresponding systems in order to maximize the transfer rate and to increase the efficiency [3].

**Biology** Any living species requires food in order to win energy and to take in the required chemical species to live and grow. Plants transport water and minerals from the soil via roots and xylem to leaves, where they take up carbon dioxide and convert it by photosynthesis into organic material. Animals take up oxygen via lungs or gills. Oxygen carried by blood cells is transported with the blood flow to reach finally each cell, where oxygen is taken as a energy source for metabolism. Waste of the cell metabolism, including carbon dioxide, has to be transported away, and is often chemically converted and finally segregated.

Common to all these processes is that they are of complex nature because of two basic facts. Firstly, the transport is often accompanied by chemical reactions. Secondly, at least one of the two phases at the interface is not solid. Therefore mass is not only transported by molecular diffusion but also by flow. Except for microfluidic systems, the flow is turbulent. This gives rise to viscous boundary layers at the interface, in which molecular diffusion is dominant. Outside of the boundary layer, the transport is controlled by turbulent velocity fluctuations.

In the past, most measuring techniques for mass transfer were non-imaging and non-optical. But already almost forty years ago, it be-

came evident that only contactless imaging techniques can resolve the mechanisms controlling them [4, 5]. From 2005–2015 the joint DFG research unit GRK 1114 “Optical Techniques for Measurement of Interfacial Transport Phenomena” of the Technical University Darmstadt and Heidelberg<sup>3</sup> helped to advanced imaging techniques.

In this paper we focus on one of the most complex problem, the gas transfer across air-water interface, which is undulated by wind waves. Under these conditions, the aqueous mass transfer boundary layer, which is the bottleneck for the transfer, is just 10–350  $\mu\text{m}$  thick [6]. Therefore it is obvious that absorption techniques will not work, but fluorescence imaging may work. Nevertheless, serious experimental challenges have to overcome even in laboratory facilities:

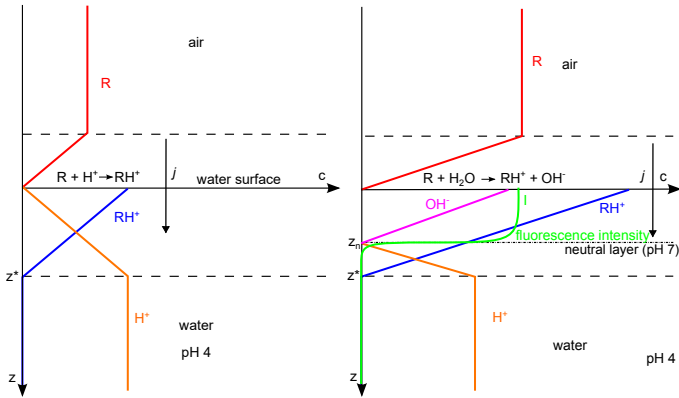
1. The concentration of a gas dissolved in water has to be made visible by a suitable fluorescence technique.
2. The fluorescence intensity of the thin layer will be weak and because of the fast movements of the water surface by waves and the shear flow only sub-ms exposure times are possible. Therefore very bright light sources and sensitive cameras are required.
3. Since sub-mm structures have to be resolved, it is impossible to focus over the range of height variations caused by wind waves. Therefore either a wave-following imaging system or permanent refocusing is required.

The paper is organized as follows. After a brief historic description of fluorescence imaging for mass transfer in Section 2, the basic principles of a newly designed and optimized fluorescence technique are explained (Section 3) and first test results from a small linear wind-wave facility are shown (Section 4). The paper closes in Section 5 with an outlook on the planned setup at the large Heidelberg Air-Sea Interaction Facility, the Aeolotron<sup>4</sup> and 4-D (3 spatial and one time coordinate) imaging of the imaged concentration fields.

---

<sup>3</sup> <https://gepris.dfg.de/gepris/projekt/462057?language=en>

<sup>4</sup> <https://www.youtube.com/watch?v=UNOWLx90w9Q&t=25s>



**Figure 1:** Sketch of the boundary layer thickness imaging technique proposed by Hiby, when an alkaline gas is absorbed by an acid liquid: low flux density with neutral layer at the surface (left) and higher flux with the neutral layer within the mass boundary layer (right).

## 2 Historical development

To the best knowledge of the authors, the chemical engineer Julius W. Hiby (RWTH Aachen) was the first to use fluorescence imaging for mass transfer studies. He studied absorption of acid or alkaline gases in falling films and reported already in 1966 the usage of fluorescent dyes which are either only fluorescent in the alkaline or acid region [7]. His work was widely overlooked because he published only a few German language papers and just a single late English publication in 1983 [8].

Figure 1 illustrates the fluorescent technique proposed by Hiby. In order to explain the basic idea, it is sufficient to assume that a) the mass boundary layers on both sides of the interface are layers of constant thickness with only molecular diffusion taking place there and b) the process is stationary with a constant flux density  $j$  from air to water. Outside of the boundary layer the turbulent mixing should be so strong that the concentrations are constant. This simplification is known as the film model.

The water is slightly acid (pH 4) and a low concentration of an alkaline gas  $R$  is put into the air space. At the acid interface it reacts

with the  $H^+$ -ions. Therefore the concentration of the gas  $[R]$ , is zero at the water surface forcing a constant flux density  $j$ , which is given according to Fick's first law for stationary diffusion as

$$j = \frac{D_a}{z_a} \Delta c = k_a [R]. \quad (1)$$

The quantity  $k$  has the dimension of a velocity and is known as the *transfer velocity*,  $D_a$  the diffusion coefficient of R in air, and  $z_a$  the thickness of the mass boundary layer in air. The  $H^+$ -ions are converted at the interface into  $RH^+$ -ions. Therefore a coupled counter diffusion takes place in the aqueous mass boundary layer:  $H^+$ -ions diffuse upwards and  $RH^+$ -ions downwards. The left figure shows the limiting condition, when the  $H^+$ -ions become zero at the interface. Because the flux density  $j$  remains constant

$$j = \frac{D_w}{z_w} \Delta c = k_w [RH^+] = k_w [H^+]. \quad (2)$$

If the concentration of R is further increased, no more  $H^+$ -ions are available at the water surface and the alkaline gas now reacts with water to produce  $OH^-$ -ions. These ions diffuse downwards and at a neutral layer within the boundary layer react with the  $H^+$ -ions to water again (left part of Figure 1). Assuming that the coupled diffusion coefficients remain the same, half of the boundary layer thickness becomes alkaline, if the concentration of R is the double of the limiting case shown in the left figure. With a pH indicator which fluoresces only in the alkaline region, the total fluorescence intensity is then proportional to the alkaline fraction of the boundary layer thickness.

In this way, the thickness of the mass boundary layer can be measured by the fluorescence intensity. Fluorescence starts, when the  $H^+$ -ion concentration becomes zero at the interface. By comparing Eqn. (1) and (2), the concentration in the air space must reach the following value

$$[R] = \frac{k_w}{k_a} [H^+]. \quad (3)$$

At a pH value of 4 the  $H^+$ -ion concentration is  $10^{-4}$  Mol/L. Because of the much slower diffusion in liquids,  $k_w$  is typical three orders of

magnitude lower than  $k_a$ . Therefore fluorescence starts already at air concentrations of R higher than about  $10^{-7}$  Mol/L, which corresponds to a partial pressure of only 2.5 ppm (parts per million). Therefore this technique is remarkably sensitive.

However, it has also two significant disadvantages:

1. The transfer process is governed by an interplay between turbulent and molecular transport. Therefore it is not stationary. This means that in regions where  $k_w$  is significantly higher than the average, there will be no fluorescence at all and it can not be determined how large the transfer velocity is in this regions.
2. The concentration of the dye must be at least an order of magnitude lower than the  $H^+$ -ion concentration. Otherwise the dye would no longer be an indicator but influence the chemical reactions. The low indicator concentrations hindered so far, to perform measurements at higher wind speeds where the mass boundary layer is correspondingly thin and therefore the fluorescence intensity is too low.

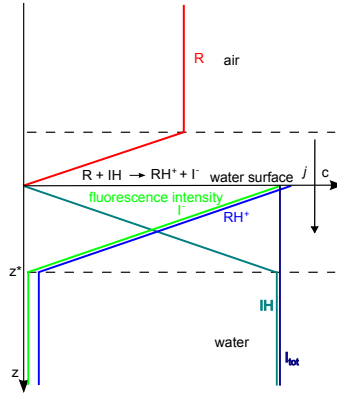
### 3 Basic principle of the new pH indicator method

In order to overcome the weaknesses of previous techniques, a new pH indicator method has been developed [9,10]. Its principle is based on a direct chemical reaction with the indicator itself. When an alkaline trace gas R enters the water, it immediately undergoes an acid-base reaction with the pH indicator IH at the water surface



In this way an invisible gas R is replaced at the air-water interface by the alkaline form of the fluorescent dye  $I^-$ , which diffuses together with  $RH^+$  across the boundary layer (Figure 2). Two basic prerequisites must be met for the technique to work:

1. The concentration of the fluorescent dye has to be much higher than those of the  $H^+$  and  $OH^-$ -ions. This ensures for the alkaline trace gas to predominantly deprotonate the pH indicator according to reaction (4).



**Figure 2:** Sketch of the new fluorescence imaging with a sufficient high pH indicator concentration to replace a trace gas via an acid-base reaction at the surface.

2. The pK value of the trace gas should be significantly above the pH value in the water-sided boundary layer to guarantee all gas molecules are protonated when dissolving in the water and the equilibrium of reaction (4) is strongly on the left side.

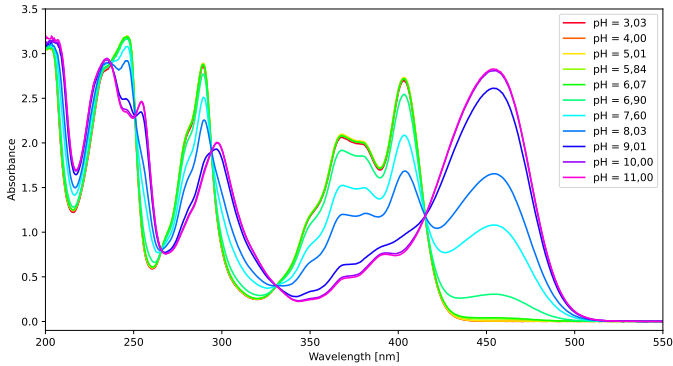
Both conditions jointly result in a linear relationship between the concentrations of the trace gas dissolving in water and the pH indicator's alkaline form

$$[I^-] \propto [R]_w. \quad (5)$$

For experimental realization of the requirements on the new chemical system, we work with an indicator concentration  $[I_{\text{tot}}]$  of about  $10^{-4}$  Mol/L. Then in a pH range from 5 to 9

$$[I_{\text{tot}}] \gg [H^+], [OH^-]. \quad (6)$$

The fluorescent dye pyranine (Trisodium 8-hydroxypyrene-1,3,6-trisulfonate) has proven to be ideal, with a pK value close to the neutral range. We determined the pK value to  $7.89 \pm 0.01$  from absorption measurements of pyranine (Figure 3). Compared to the formerly used ammonia [9] with pK = 9.24, ethylamine and other amines are planned to be used instead, as they have the advantage of a significantly higher alkalinity with a pK value larger than 10.6.



**Figure 3:** Absorption spectra of a  $10^{-4}$  molar pyranine solution at pH values as indicated. Only the alkaline form of pyranine absorbs in the range of 440–500 nm.

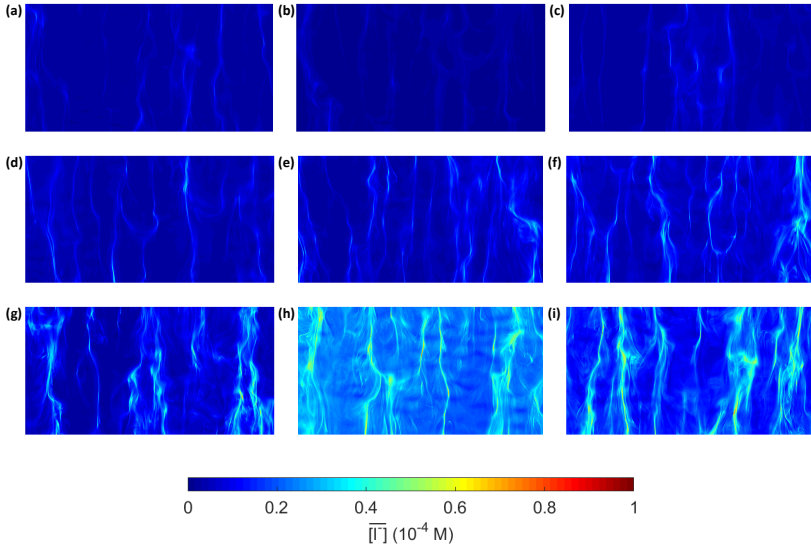
## 4 First results

In a measurement the pH value of the water is initially adjusted to 5, causing a large proportion of the pyranine to be in its acidic form IH. Subsequently, an alkaline gas is added to the gas space, which increases the alkaline form of pyranine  $I^-$  as it invades into water.

Both forms of pyranine are fluorescent, but only the alkaline form absorbs light at wavelengths larger than 440 nm (Figure 3). Therefore the fluorescence will be according to Eq. (5) proportional to the concentration of the dissolved gas, when fluorescence is excited at 450 nm. At the starting pH value of 5 about a permille of pyranine is already existent in its alkaline form, so the water bulk generates a non-negligible background fluorescence. To suppress this, the dye tartrazine is additionally added, which absorbs the excitation light and prevents it from penetrating into deeper water layers. Consequently, the detected fluorescence pattern displays only the concentration fields of the gas in the uppermost centimeter of the water-side boundary layer.

The new method has already been tested in a small linear wind-wave facility and proven to work as expected. With increasing flux density  $j$  of the alkaline gas, the patterns just get brighter, but there is no threshold effect as with the Hiby method (Figure 4).





**Figure 4:** Example images taken in a small linear wind-wave facility with the new pH indicator method [9]. The applied flux of the alkaline gas ammonia increased from images (a) to (h) and started to decrease again at image (i).

## 5 Outlook

The technique is ready to be used in the Heidelberg Aeolotron, an annular wind-wave-facility 10 m in diameter [11]. The fluorescence is stimulated by four light sources radiating from above through a glass window onto the channel's water surface with a total optical peak power of 250 W irradiating about  $0.25 \text{ m}^2$  at the water surface. Seven Lucid Vision Atlas 10GigE ATX051S cameras image the fluorescence patterns at the water surface from underneath through a bottom glass window at 500 fps and a resolution of  $1224 \times 1024$  pixel.

This arrangement makes 3-D imaging possible to reconstruct the shape of the water surface as well and to distinguish the thin boundary layer at the water surface from structures swept down into the bulk water by surface renewal events. A light field imaging approach, similar to the technique of Wanner and Goldlücke [12] to separate reflective and transparent surfaces, will be used.

## References

1. J. Beyerer and T. Längle, Eds., *OCM 2021 - Optical Characterization of Materials: Conference Proceedings*. KIT Scientific Publishing, 2021.
2. P. S. Liss and M. T. Johnson, Eds., *Ocean-Atmosphere Interactions of Gases and Particles*. Springer, 2014.
3. H. D. Baehr and K. Stephan, *Wärme- und Stoffübertragung*, 10th ed. Springer, 2019.
4. B. Jähne, "Transfer processes across the free water interface," Habilitation thesis, Institut für Umweltphysik, Fakultät für Physik und Astronomie, Heidelberg University, 1985.
5. B. Jähne, "Impact of quantitative visualization and image processing on the study of small-scale air-sea interaction," in *Air-Water Gas Transfer, Selected Papers, 3rd Intern. Symp. on Air-Water Gas Transfer*, B. Jähne and E. Monahan, Eds. Hanau: AEON, 1995, pp. 3–12.
6. B. Jähne, "Air-sea gas exchange," in *Encyclopedia of Ocean Sciences*, 3rd ed., J. K. Cochran, H. J. Bokuniewicz, and P. L. Yager, Eds. Academic Press, 2019, vol. 6, pp. 1–13.
7. J. W. Hiby, "Eine Fluoreszenzmethode zur Untersuchung des Transportmechanismus bei der Gasabsorption im Rieselfilm," *Wärme- und Stoffübertragung*, vol. 1, pp. 105–116, 1968.
8. J. W. Hiby, "The chemical indicator: a tool for the investigation of concentration fields in liquid," *Ann. N. Y. Acad. Sci.*, vol. 404, pp. 348–349, 1983.
9. M. Papst, "Development of a method for quantitative imaging of air-water gas exchange," Master's thesis, Institut für Umweltphysik, Universität Heidelberg, Germany, 2019.
10. D. Hofmann and B. Jähne, "Imaging of concentration fields within the mass boundary layer," in *Air-Water Gas Exchange: Cross-linking field measurements, laboratory measurements and modeling*, B. Jähne, K. E. Krall, and C. Marandino, Eds. Heidelberg University Publishing, 2023, in press.
11. K. E. Krall and B. Jähne, "Annular wind-wave tanks," in *Air-Water Gas Exchange: Cross-linking field measurements, laboratory measurements and modeling*, B. Jähne, K. E. Krall, and C. Marandino, Eds. Heidelberg University Publishing, 2023, in press.
12. S. Wanner and B. Goldluecke, "Reconstructing reflective and transparent surfaces from epipolar plane images," in *Pattern Recognition*. Springer, 2013, pp. 1–10.

# Material Characterization using a Compact Computed Tomography Imaging Spectrometer with Super-resolution Capability

Simon Amann<sup>1\*</sup>, Mazen Mel<sup>2\*</sup>, Pietro Zanuttigh<sup>2</sup>, Tobias Haist<sup>1</sup>,  
Markus Kamm<sup>3</sup>, and Alexander Gatto<sup>3</sup>

<sup>1</sup> University of Stuttgart, Institut für Technische Optik,  
70569 Stuttgart, Germany

<sup>2</sup> University of Padova, Department of Information Engineering,  
35131 Padova, Italy

<sup>3</sup> Sony Europe B.V., Stuttgart Technology Center,  
70327 Stuttgart, Germany

**Abstract** Computed Tomography Imaging Spectrometer (CTIS) systems are snapshot hyperspectral imaging devices capable of capturing dense spectra of static as well as dynamic scenes. A three-dimensional hyperspectral cube is smeared across the spatial dimension via Diffractive Optical Element (DOE) and projected across multiple angles forming a two-dimensional compressed sensor image. In this paper we demonstrate material characterization and classification capability of a compact CTIS system leveraging spectral signatures. Then we propose an approach to simultaneously reconstruct and segment into regions corresponding to different materials hyperspectral images with enhanced spatial resolution from CTIS sensor measurements.

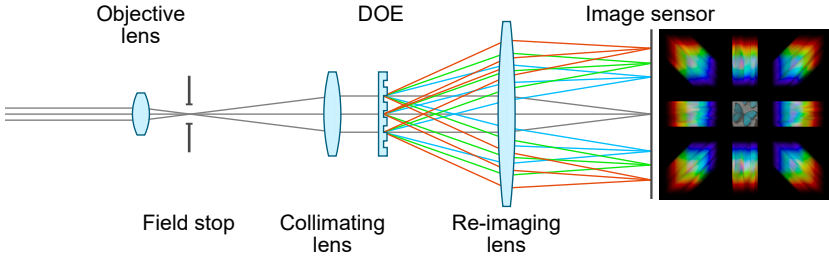
**Keywords** CTIS, spectral reconstruction, super resolution, optical characterization

## 1 Introduction

Hyperspectral Imaging (HSI) plays an important role in the field of optical characterization of materials [1]. It allows, for example, to distin-

---

\* Authors contributed equally.



**Figure 1:** Optical layout of a commonly used CTIS system. Image based on [4].

guish or identify materials that look almost identical in a monochrome or color image. HSI-devices acquire a complete spectrum for each imaged object point. The resulting hyperspectral cube has three dimensions: the two spatial ones and the spectral dimension.

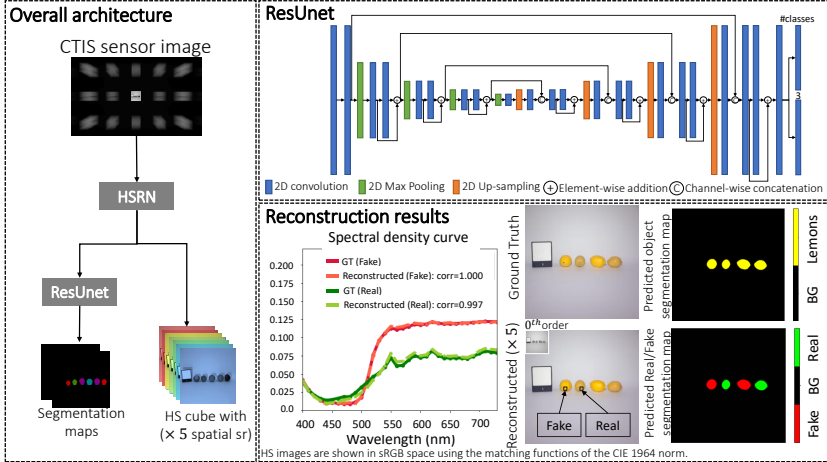
A Computed Tomography Imaging Spectrometer (CTIS) is based on a non-scanning (snapshot) technique [2]. Other methods in this area are the multi-aperture filtered camera and the pixel-level filter array camera [3]. They are both based on spectral filters. CTIS, on the other hand, uses a diffractive optical element (DOE) in combination with computational imaging algorithms. Figure 1 shows an optical layout of a commonly used CTIS system. The objective lens images the scene on the left to an intermediate image plane. There, it is cropped by a field stop, which defines the system’s field of view. The collimating lens collimates the light, which is then spectrally dispersed by a diffractive optical element. A re-imaging lens creates the final sensor image. An example is shown on the right. It contains several higher diffraction orders arranged around the undiffracted zeroth order image of the scene. The higher diffraction orders are spectrally smeared. Blue light hits the sensor closer to the center than its red counterpart.

A reconstruction algorithm is needed to get the hyperspectral image from this spatio-spectral smeared sensor image. It solves a similar inverse problem as the reconstruction algorithms needed for computed tomography scanners. The different diffraction orders can be conceived of as two-dimensional projections of the three-dimensional hyperspectral-cube onto the image sensor. The Expectation-Maximization (EM) algorithm has been predominantly

used in CTIS image reconstruction [5]. The EM iteratively solve for the latent hyperspectral cube starting from an initial estimate. EM cannot handle priors and it is sensitive to the presumed noise and system model leading sometimes to poor reconstruction quality. Deep learning-based approaches have been devised to tackle the shortcomings of the EM solver: In [6] the authors used a sequential approach with a CNN followed by an EM solver wherein the CNN provides the initial estimate for the EM stage. Zimmermann *et al.* [7] proposed an end-to-end learning approach performing customized reshaping operations at the beginning to get an input shape suitable for 3D processing of high dimensional input data that is followed by a U-Net like architecture used to refine the estimated hyperspectral cube. We have recently proposed HSRN [8] tackling for the first time spectral reconstruction and spatial super-resolution from CTIS measurements. It allows to achieve a higher spatial resolution than that of the zeroth diffraction order while reconstructing accurate spectral information.

## 2 Method

We propose a two-stage approach for object classification using hyperspectral data captured by a CTIS system (see Figure 2). In the first stage we train our HSRN [8] architecture for hyperspectral reconstruction and spatial super resolution with up to  $\times 5$  the resolution of the zeroth diffraction order for synthetic data. In the second stage, the reconstructed hyperspectral cubes are used to train a ResUnet [9] to perform semantic segmentation. The network produces two segmentation maps, one corresponding to object classes and the other underlining whether those objects are real or fake. Note that the two networks are trained separately. In more details, we use slightly modified architectures of both networks for better reconstruction quality and to avoid over-fitting. For HSRN [8] we increase the number of filters within the refinement network from 64 to 128 for all convolution layers and set the super-resolution factor to 5 for synthetic data and 2 for real data while keeping the rest of the architecture unchanged. For ResUnet [9] we use the modified architecture shown in Figure 2, the network has two output layers, one for each segmentation task. We train both networks for 500 epochs and use the training settings of HSRN suggested



**Figure 2:** Left: Proposed two-stage architecture for hyperspectral image reconstruction and semantic segmentation, the two networks are trained separately. Upper right: The slightly modified ResUNet architecture used to learn object class and real/fake segmentation maps. Lower right: A reconstruction example with  $\times 5$  spatial super-resolution and the corresponding segmentation maps, we also show spectral density curves of two selected image regions (real and fake lemons) along with the Pearson correlation coefficient to assess the accuracy of the reconstructed spectra.

in [8]. The cross-entropy loss is used to train the ResUNet.

### 3 Datasets

**Synthetic data** We use Fourier optics to simulate CTIS sensor images using hyperspectral cubes from FVgNet dataset [10] containing 252 labeled scenes of real and fake fruits and vegetables. A DOE that generates a structure with  $5 \times 3$  diffraction orders is used in the simulation (see Figure 2). The simulated zeroth order has a spatial resolution of  $102 \times 102$  pixels while the ground truth hyperspectral cubes have  $510 \times 510$  pixels which corresponds to a  $\times 5$  spatial super-resolution of the reconstructed cube. As in [10], we use a spectral range of  $[400nm, 730nm]$  with 34 spectral bands. We chose randomly 80% of the scenes as training data and the rest for testing, random vertical and

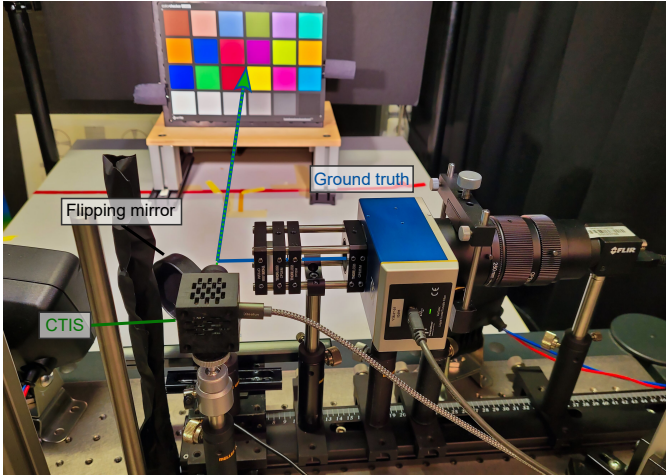
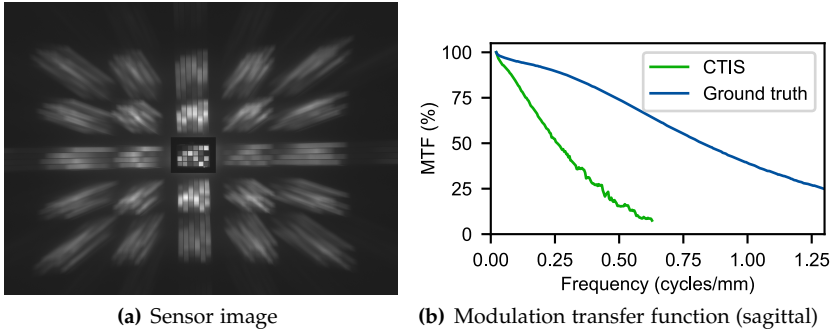


Figure 3: Photo of the miniaturized prototype together with the ground truth setup.

horizontal flipping is used as data augmentation.

**Real data** We have implemented a setup to validate that our reconstruction method also works on real world CTIS data. A photo of the system is shown in Figure 3. For the dataset needed to train our model, we always acquire a CTIS measurement together with a ground truth measurement. Our CTIS system is built with off-the-shelf lenses, a computer-generated hologram, a commercial smartphone lens and a 13 MP monochrome smartphone image sensor. The dimensions of the prototype are only  $36.0 \text{ mm} \times 40.5 \text{ mm} \times 52.8 \text{ mm}$ . This small size is achieved by using a Galilean instead of the commonly used Keplerian beam expander. Its diagonal field of view is  $29^\circ$ . The DOE creates a  $5 \times 5$  arrangement of the diffraction orders. The zeroth order image size is  $420 \times 312$  pixels, which corresponds to around 10% of the horizontal and vertical sensor size. Filters are used to limit the captured spectral range from 470 nm to 700 nm. Each CTIS measurement is made of two images captured with different exposure times (7.8 ms and 250 ms). This is needed to get an image with a well exposed zeroth order and one with well exposed higher diffraction orders. Our

prototype is therefore not a single-shot camera. Figure 4(a) shows a sample acquisition of a ColorChecker. The zeroth order part of the image taken with the longer exposure time is exchanged with that of the shorter exposure time. More information about a similar system can be found in [11]. Amann et al. [11] use the same prototype, just with a different shortpass filter.



**Figure 4:** Sensor image of the CTIS prototype and MTF measurement results comparing the CTIS prototype with the ground truth setup.

To capture the ground truth data, we built a hyperspectral camera based on a VariSpec tunable color filter. The hyperspectral image is captured time-sequentially. We use a flip mirror to bypass light into this reference system. This way, it sees the object from the same point of view as the CTIS system. The VariSpec filter has a bandwidth of 7 nm. We therefore capture our scenes in 7 nm steps and also reconstruct the CTIS images with this channel width. The camera captures the scene with a spatial resolution that is around  $\times 4$  higher (in each dimension) than that of the zeroth order image of the CTIS prototype. Figure 4(b) shows a modulation transfer function (MTF) of the CTIS system compared to the ground truth system. This has been determined using a measurement of a Siemens star. It shows that we have a three times better imaging quality with the ground truth system than with the CTIS system (zeroth order). It thus can be used to train our network accounting for super-resolution.



## 4 Experimental Results

**Synthetic Data** Spectral reconstruction, as well as semantic segmentation results, are presented in this section. To highlight the contribution of spectral information for object classification, we compare results obtained by training the ResUnet using the reconstructed hyperspectral cubes from CTIS measurements with the ones obtained using RGB images extracted from the reconstructed hyperspectral cubes. Quantitative results are shown in Tables 1 and 2, while the qualitative are in Figures 5 and 6. From Table 1 and Figure 5 it can be seen that the

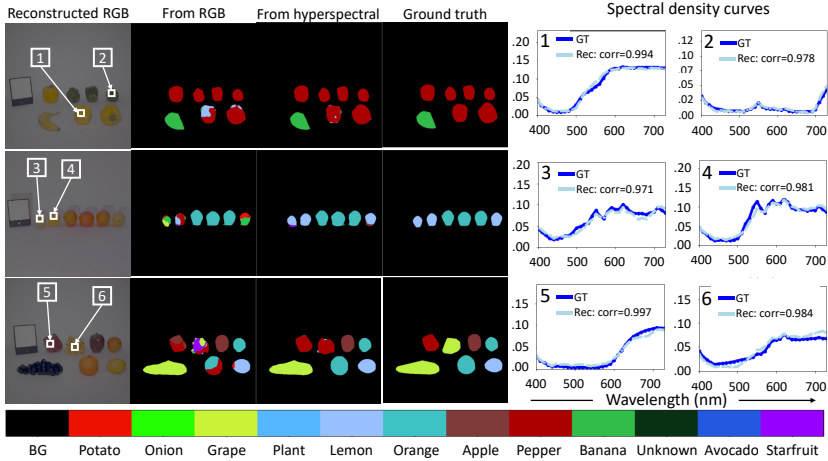
**Table 1:** Quantitative metrics for spectral reconstruction and image super-resolution on FVgNet [10].

| Split | PSNR (dB) | SSIM  | MAE ( $1e^{-3}$ ) |
|-------|-----------|-------|-------------------|
| Train | 51.943    | 0.995 | 1.5               |
| Test  | 51.781    | 0.995 | 1.6               |

**Table 2:** Quantitative metrics for semantic segmentation on the test set of FVgNet [10]: *Obj* refer to the semantic segmentation task on object classes meanwhile *R/F* refer to the task of classifying real and fake objects (better in bold).

| Input         | mIoU (%)     |              | F1           |              | Precision    |              | Recall       |              |
|---------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
|               | Obj          | R/F          | Obj          | R/F          | Obj          | R/F          | Obj          | R/F          |
| RGB           | 78.61        | 91.54        | 0.878        | 0.953        | 0.862        | <b>0.966</b> | 0.907        | 0.941        |
| Hyperspectral | <b>86.63</b> | <b>91.95</b> | <b>0.926</b> | <b>0.956</b> | <b>0.902</b> | 0.958        | <b>0.957</b> | <b>0.954</b> |

model produces acceptable reconstructions both spatial and spectral-wise with  $\times 5$  super-resolution factor. Figure 5 shows how semantic segmentation using only RGB data fails sometimes to learn correct pixel labels due to the limited information carried out by the three color components, instead the network might rely heavily on semantic cues. In the case of semantic segmentation from spectral data, results are much better for both classification tasks, in particular achieving a gain of more than 8% on the objects' semantic segmentation task. Although segmentation metrics for Real/Fake classification task using spectral data is only slightly better than the one using RGB as shown in Table 2 and Figure 6, such behavior can be due to the network capability to better leverage semantic cues in the latter case.



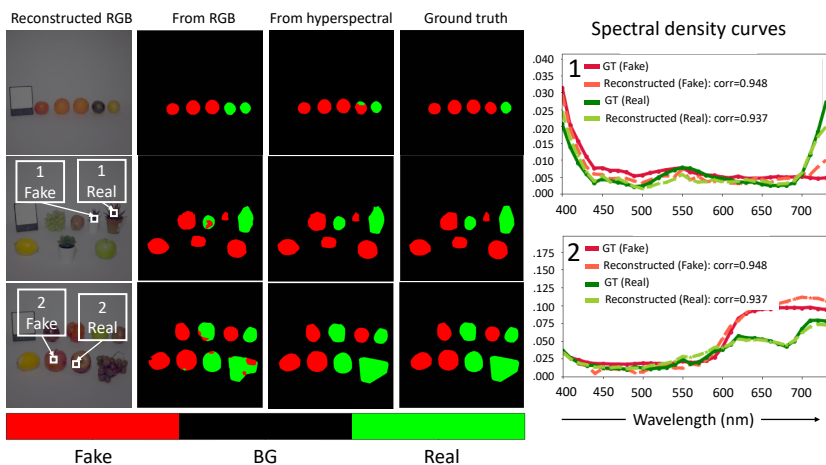
**Figure 5:** Qualitative results on hyperspectral reconstruction and semantic segmentation of various objects. We show also spectral density curves of some chosen image regions.

**Real Data** In this section we present reconstruction results on real data captured by our compact CTIS system. Figure 7 shows a few reconstructed images in sRGB space and some selected individual spectral bands along with spectral density curves of some image regions to highlight the discrepancies between the spectrum of real and fake red peppers.

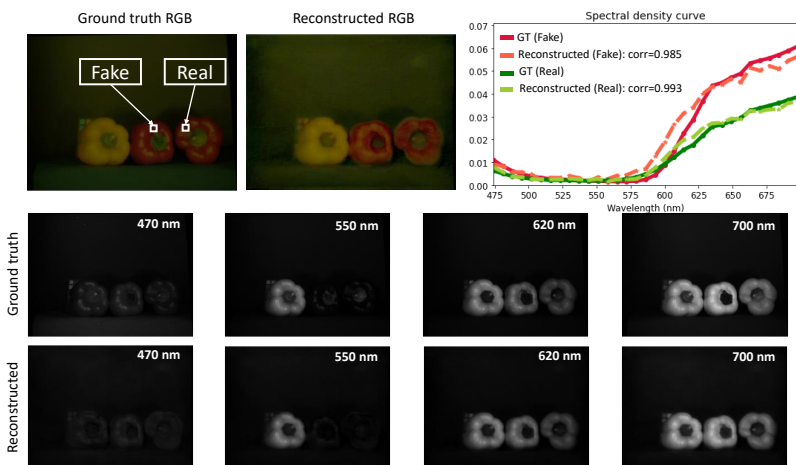
## 5 Conclusion

We presented a compact CTIS prototype using a Galilean design and a ground truth acquisition apparatus that allows to capture high quality hyperspectral images. We showcased spectral reconstruction and material classification capability from CTIS measurements using a deep learning based approach to reconstruct spatially super-resolved hyperspectral cubes and perform semantic segmentation of fake and real fruits and vegetables leveraging their spectral signature.

## Material characterization using a compact CTIS with SR capability



**Figure 6:** Qualitative results on Real/Fake semantic segmentation. We also show spectral density curves of some chosen image regions.



**Figure 7:** Qualitative reconstruction of a real CTIS scene containing real and fake red peppers. The reconstruction image has  $\times 2$  the resolution of the zeroth diffraction order.

## References

1. M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14 118–14 129, 2018.
2. T. Okamoto and I. Yamaguchi, "Simultaneous acquisition of spectral image information," *Optics Letters*, vol. 16, no. 16, p. 1277, aug 1991.
3. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering*, vol. 52, no. 9, p. 090901, sep 2013.
4. Y. Monno, S. Kikuchi, M. Tanaka, and M. Okutomi, "A practical one-shot multispectral imaging system using a single image sensor," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3048–3059, oct 2015.
5. C. E. Volin, "Portable snapshot infrared imaging spectrometer," Ph.D. dissertation, The University of Arizona, 2000.
6. M. J. Ahlebaek, M. S. Peters, W.-C. Huang, M. T. Frandsen, R. L. Eriksen, and B. Jørgensen, "The hybrid approach—convolutional neural networks and expectation maximization algorithm—for tomographic reconstruction of hyperspectral images," *arXiv preprint arXiv:2205.15772*, 2022.
7. M. Zimmermann, S. Amann, M. Mel, T. Haist, and A. Gatto, "Deep learning-based hyperspectral image reconstruction from emulated and real computed tomography imaging spectrometer data," *Optical Engineering*, vol. 61, no. 5, p. 053103, 2022.
8. M. Mel, A. Gatto, and P. Zanuttigh, "Joint reconstruction and super resolution of hyper-spectral ctis images," in *33rd British Machine Vision Conference 2022*. BMVA Press, 2022.
9. F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data," *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020.
10. M. Makarenko, A. Burguete-Lopez, Q. Wang, F. Getman, S. Giancola, B. Ghanem, and A. Fratolocchi, "Real-time hyperspectral imaging in hardware via trained metasurface encoders," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 12 692–12 702.
11. S. Amann, T. Haist, A. Gatto, M. Kamm, and A. Herkommer, "Design and realization of a miniaturized high resolution computed tomography imaging spectrometer," *EPJ Web of Conferences*, vol. 266, p. 02001, 2022.

# Sulfur Dioxide Fluorescence Imaging

Bernd Jähne<sup>1,2</sup>, Rada Beronova<sup>1</sup>, and Kerstin E. Krall<sup>1</sup>

<sup>1</sup> Heidelberg University, Institute for Environmental Physics  
Im Neuenheimer Feld 229, 69120 Heidelberg

<sup>2</sup> Heidelberg University, Interdisciplinary Center for Scientific Computing  
Berliner Straße 43, 69120 Heidelberg

**Abstract** Sulfur dioxide is an ideal tracer to study the partitioning of the resistance of gas transfer across the water interface between air and water because the pH value in water controls the effective solubility of sulfur dioxide. Friman and Jähne [1] already demonstrated that it is possible to measure sulfur dioxide concentration profiles with laser induced fluorescence (LIF), but the best excitation wavelength under standard atmospheric conditions was not known. Here, we report the result of our investigation to select the best excitation wavelength for sulfur dioxide fluorescence to reach maximum intensity with the lowest possible absorption.

**Keywords** Sulfur Dioxide, Fluorescence Imaging

## 1 Introduction

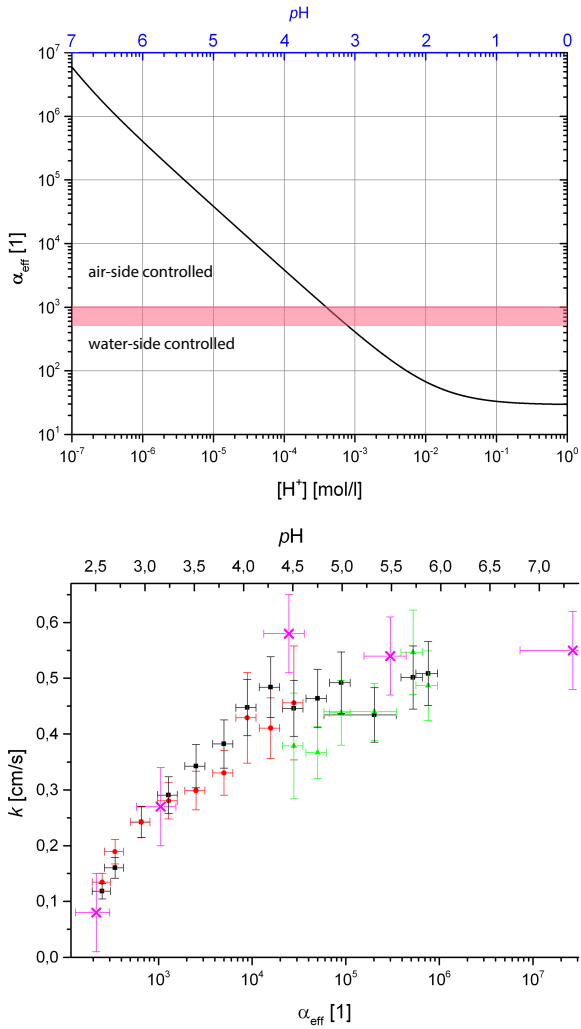
Fluorescence imaging has two specific advantages. Firstly, it allows to measure concentration fields. The simplest setup is to stimulate the fluorescence by a light sheet to obtain a planar cross-section of a 3-D concentration field. Secondly, by using the right combination of the stimulation wavelength and spectral range, it is very specific and can be tuned to measure the fluorescence of a single chemical component. Therefore fluorescence imaging has become very useful in life sciences, fluid dynamics and combustion research. In this paper we describe fluorescence imaging of sulfur dioxide. It nicely demonstrates that all details must carefully be considered to set up an optimal measuring system.

Our interest in sulfur dioxide is caused by the fact that sulfur dioxide is an ideal tracer to study the partitioning of the resistance of gas transfer across the air-water interface. The dimensionless solubility expresses how much of a dissolved species is contained per volume unit in water as compared to air. The solubility of a volatile species or gas in water decides whether it can be transported more easily in water or in air. A species with a low solubility experiences a high concentration difference in water compared to the concentration difference in air, because not much of the dissolved species can be transported by a volume element in water. The transport experiences then a high resistance, i. e., concentration difference in water. In this case the transport processes in water control the speed of transfer and not those in the air space. For a high solubility in water, it is the other way round. At a wind-driven water surface the transition between water-side to air-side control occurs at a solubility between 500 and 1000 [2, 3].

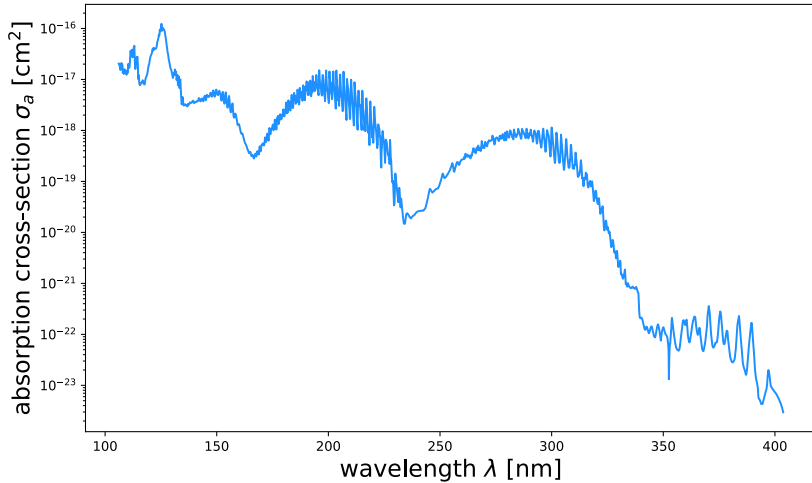
The physical solubility of sulfur dioxide is about 29 at room temperature [4]. At pH-values larger than 1, sulfur dioxide reacts with water to form hydrogen sulfite. Therefore, the effective solubility increases tenfold per pH-value (Figure 1, top). At pH values higher than 4.5, the solubility reaches such high values that sulfur dioxide is transported better in water than in air. At a pH-value of about 3.3, the air-side and water-side resistances are expected to be equal, which means that the transfer is about half as fast as at high pH-values with pure air-side and negligible water-side resistance. Niegel [5] verified this in a small linear wind-wave facility (Figure 1, bottom).

The transfer resistance can therefore be shifted from water-side to air-side control when the pH-value changes from 2.5 and 4.5 and any ratio of the transfer resistance between air and water can be set by the pH-value. This allows a detailed investigation of the partitioning of the transfer resistance between air and water, which has not yet been performed at all. Of special interest is the direct measurement of the concentration sulfur dioxide reaches in air right at the water surface. This value directly yields the partitioning ratio of the resistance between air and water. It has never been observed yet to which extend this ratio fluctuates and which parameters control these fluctuations.

Such a measurement, however, requires to measure vertical sulfur dioxide profiles in the air down to the wavy water surface using a fluorescence technique. Friman and Jähne [1, 6] demonstrated that it is



**Figure 1:** Top: Effective solubility of sulfur dioxide depends on the pH-value of water; Bottom: Measured transfer velocities of sulfur dioxide at different pH-values in a small wind-wave facility [5].



**Figure 2:** UV absorption spectra of sulfur dioxide (absorption cross-section) at wavelengths from 100 to 400 nm [7].

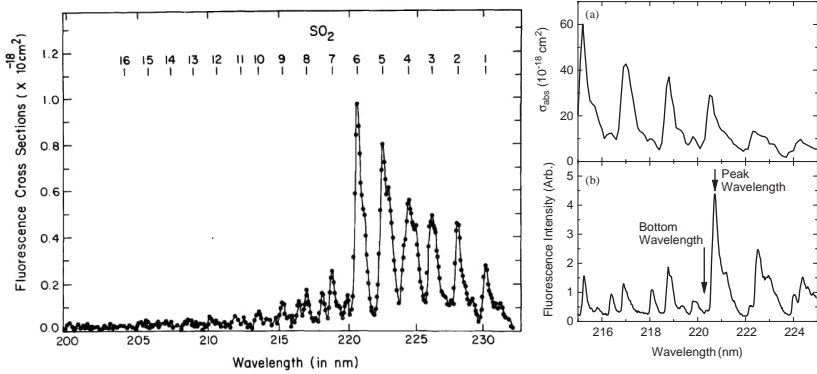
possible to measure sulfur dioxide concentration profiles with laser induced fluorescence (LIF), although only a suboptimal fixed excitation wavelength of 223.7 nm was available. Because sulfur dioxide has a complex absorption spectrum, the best excitation wavelength was unknown. Competing processes such as fluorescence quenching or dissociation of the sulfur dioxide molecule lower the fluorescence quantum yield and must be considered.

The paper is organized as follows. Section 2 reviews the knowledge about the absorption spectra and fluorescence of sulfur dioxide. Then the setup to measure sulfur dioxide fluorescence is explained (Section 3) and the results are discussed in Section 4.

## 2 Sulfur Dioxide Absorption Spectra and Fluorescence

Sulfur dioxide has a complex absorption spectrum in the UV (Figure 2), which is caused by electronic transitions together with changes of the vibration and rotation state. Measurements of sulfur dioxide by ab-





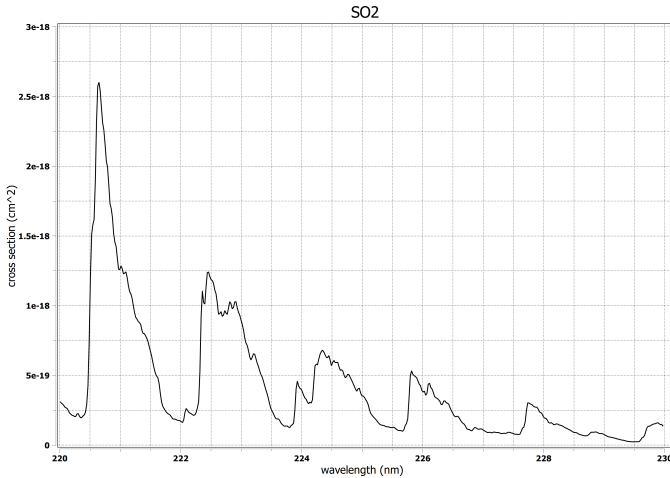
**Figure 3:** Left: Fluorescence absorption cross-section measured at 5–13  $\mu\text{bar}$  pure sulfur dioxide [8]; right: Absorbance and fluorescent intensity of 10 ppbv sulfur dioxide in air at 13 mbar [10].

sorption spectroscopy are possible in the band between 260 and 310 nm or with a tenfold increased sensitivity in the deep UV around 200 nm. It is known from literature [8] that the quantum yield of sulfur dioxide fluorescence excited in the weaker second absorption band between 260 and 310 nm is low even in pure sulfur dioxide gas at low pressures. The quantum efficiency for fluorescence is only high in this band at high temperatures. Sick [9] used it for fluorescence imaging of sulfur dioxide in flames.

In the deep UV, radiation can dissociate the sulfur dioxide molecule. Hui and Rice [11] observed that the high quantum efficiency for fluorescence at 0.13 mbar decreases from about one at 225.8 nm down to zero at 215.24 nm by this effect. This is in agreement with the findings of Ahmed and Kumar [8], who observed that the fluorescence absorption cross-section (absorption cross-section times fluorescence quantum yield) shows a strong decrease (Figure 3, left), even though the absorption cross-section still increases.

Matsumi et al. [10] used fluorescence to measure atmospheric sulfur dioxide concentrations. They found a maximum fluorescence intensity with an excitation wavelength of 220.8 nm (Figure 3, right). The pressure in the measuring chamber was reduced to 13 mbar.

No data about sulfur dioxide fluorescence could be found in air at



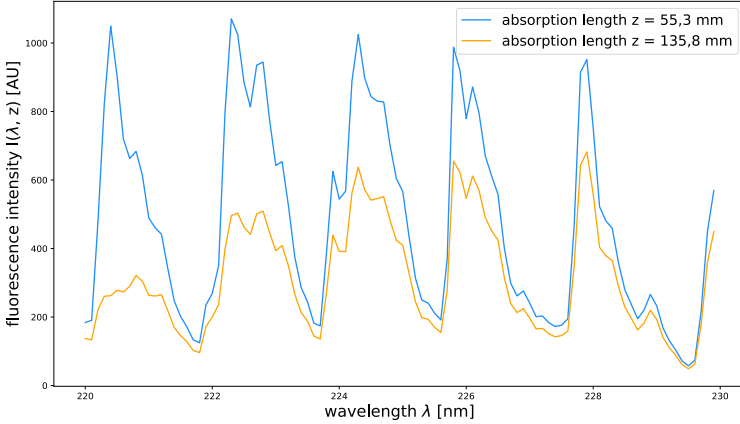
**Figure 4:** Absorption spectrum of sulfur dioxide from [12] smoothed to the line width of the InnoLas SpitLight Compact OPO-355, data collection by [13].

atmospheric pressures. Therefore the optimum excitation wavelength under this condition was unclear and a new investigation was required.

### 3 Experimental Setup

Fluorescence was excited by an InnoLas SpitLight Compact OPO-355 with UV extension to tune the excitation wavelength between 220 and 230 nm with a pulse energy of about 4 mJ at 20 Hz. Within this wavelength range, the pulse energy of the OPO remained constant. As an absorption reference we used the measurements from Rufus et al. [12] as made available by the MPI Mainz UV-VIS spectral atlas [13]. The high-resolution data were smoothed to the line width of the OPO (Figure 4). The absorption cross-section at 220.7 nm is about ten times larger than at 227.8 nm.

A flow of 20 NL/min of dry air set by a mass flow controller was mixed with a flow of 28.8 NmL/min of sulfur dioxide set by a second mass flow controller to obtain a sulfur dioxide concentration of 1440 ppm in air at atmospheric pressure. The mixed flow was directed



**Figure 5:** Fluorescence spectrum of sulfur dioxide in air measured by Beronova [14] at absorption lengths of 55.3 mm and 135.8 mm.

through a Duran glass tube with a diameter of about 6 cm. The OPO laser beam entered the tube at one end through a quartz glass window and the fluorescent light was imaged with a PCO edge 4.2 UV back illuminated UV sensitive camera using a Linos inspec.x 2.8/50 UV-VIS APO prototype lens. The imaging system covered an absorption distance between 55.3 mm and 135.8 mm, i. e., a laser beam length of 80.5 mm. Further details about the experimental setup can be found in Beronova [14].

## 4 Results and Discussion

In contrast to Matsumi et al. [10] (Figure 3), we found that the fluorescence intensity is about the same at all absorption peaks after the laser beam intensity has already been attenuated slightly by an absorption distance of 55.3 mm at 1440 ppm sulfur dioxide concentration (Figure 5). After a further distance of 80.5 mm, the fluorescence is even about two times higher at 227.8 nm than at 220.8 nm, because the absorption there is significantly lower (Figure 4).

For experiments in wind-wave facilities the laser beam has to travel

a distance of about 1 m in air before it reaches the water surface. In this experiments it is planned to use sulfur dioxide concentrations of only 100 ppm. Therefore laser beam experiences about the same attenuation and the absorption peak at 227.8 nm is the then the best choice for maximum fluorescence intensity close to the water surface.

It could be demonstrated that sulfur dioxide fluorescence measurements are possible in air at atmospheric pressure and an optimum excitation wavelength of 227.8 nm was be found. The higher fluorescence intensity at higher wavelengths in contrast to the results of Matsumi et al. [10] (Figure 3) is obviously caused by additional fluorescence quenching because of more frequent collisions of sulfur dioxide with other molecules in air. The quenching appears to be higher at lower excitation wavelengths.

## 5 Acknowledgments

Funding of this research by the German Science Foundation (DFG) Koselleck Project Grant JA 395/19-1 "Quantifying the Mechanisms of Air-Sea Gas Exchange and Bridging Laboratory and Field by Imaging Measurements" is gratefully acknowledged.

## References

1. S. I. Friman and B. Jähne, "Investigating SO<sub>2</sub> transfer across the air–water interface via lif," *Exp. Fluids*, vol. 60, p. 65, 2019.
2. P. S. Liss and P. G. Slater, "Flux of gases across the air-sea interface," *Nature*, vol. 247, pp. 181–184, 1974.
3. C. Kräuter, "Aufteilung des Transferwiderstands zwischen Luft und Wasser beim Austausch flüchtiger Substanzen mittlerer Löslichkeit zwischen Ozean und Atmosphäre," Diplomarbeit, Institut für Umweltp Physik, Fakultät für Physik und Astronomie, Univ. Heidelberg, 2011.
4. R. Sander, "Compilation of Henry's law constants (version 4.0) for water as solvent," *Atmos. Chem. Phys.*, vol. 15, pp. 4399–4981, 2015.
5. D. Niegel, "Messung des Transferwiderstands von Schwefeldioxid an der Luft-Wasser-Grenzfläche," Master's thesis, Institut für Umweltp Physik, Universität Heidelberg, Germany, 2013.

6. S. I. Friman, "Laboratory investigations of concentration and wind profiles close to the wind-driven wavy water surface," Dissertation, Institut für Umweltphysik, Fakultät für Physik und Astronomie, Univ. Heidelberg, Heidelberg, 2020.
7. S. L. Manatt and A. L. Lane, "A compilation of the absorption cross-sections of SO<sub>2</sub> from 106 to 403 nm," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 50, no. 3, pp. 267 – 276, 1993.
8. S. Ahmed and V. Kumar, "Quantitative photoabsorption and fluorescence spectroscopy of so<sub>2</sub> at 188-231 and 278.7-320 nm," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 47, pp. 359–373, 1992.
9. V. Sick, "Exhaust-gas imaging via planar laser-induced fluorescence of sulfur dioxide," *Applied Physics B: Lasers and Optics*, vol. 74, pp. 461–463, 2002.
10. Y. Matsumi, H. Shigemori, and K. Takahashi, "Laser-induced fluorescence instrument for measuring atmospheric SO<sub>2</sub>," *Atmospheric Environment*, vol. 39, no. 17, pp. 3177–3185, 2005.
11. M. Hui and S. A. Rice, "Decay of fluorescence from single vibronic states of SO<sub>2</sub>," *Chem. Phys. Let.*, vol. 17, no. 4, pp. 474–478, 1972.
12. J. Rufus, G. Stark, P. L. Smith, J. C. Pickering, and A. P. Thorne, "High-resolution photoabsorption cross section measurements of so<sub>2</sub>, 2: 220 to 325 nm at 295 k," *J. Geophys. Res.*, vol. 108, 2003.
13. H. Keller-Rudek, G. K. Moortgat, R. Sander, and R. Sörensen, "The MPI-Mainz UV/VIS spectral atlas of gaseous molecules of atmospheric interest," *Earth Syst. Sci. Data*, vol. 5, pp. 365–373, 2013. [Online]. Available: [www.uv-vis-spectral-atlas-mainz.org](http://www.uv-vis-spectral-atlas-mainz.org)
14. R. Beronova, "Optimierung der Aufnahme hochauflösender SO<sub>2</sub>-Konzentrationsprofile für Gasaustauschmessungen basierend auf LIF," Bachelor's thesis, Institut für Umweltphysik, Universität Heidelberg, Germany, 2022.



# Imaging radar systems for non-destructive material testing

## An overview of the state of the art, the limitations and the opportunities of radar technology.

Dirk Nüßler, Sven Leuchs, and Christian Krebs

Fraunhofer Institute for High Frequency Physics and Radar Techniques

FHR, Fraunhoferstraße 20, 53343 Wachtberg, Germany

**Abstract** Radar systems have been used for over 100 years to measure distances and angular positions accurately. Radar systems benefit from relatively long wavelengths, which means that most absorption and scattering mechanisms do not have a relevant influence on the propagation conditions of the emitted electromagnetic waves. As a result, radar systems were and are used primarily for measurements under poor environmental conditions. Today, we usually find applications that work with waves in the meter to millimeter wave range. Especially in the millimeter wave range, the influence of the atmosphere can no longer be neglected. Communication systems, in particular, with their need for large bandwidths, are driving the development of components in the millimeter wave range, thus opening up further fields of application. In this context, imaging radar systems are increasingly important in various application areas. This paper will look at the possible applications in industrial process monitoring [1] [2] [3] [4] [5]. The monitoring of production processes benefits from the phenomenon's importance that many non-conductive materials are partially transparent to an electromagnetic wave. Radar systems thus allow a view below the surface and can therefore measure the material thickness of, e.g. plastics in extruders. This paper will investigate the advantages and disadvantages of radar technologies and procedures and their suitability for use in production lines.

**Keywords** Non-destructive-testing, industrial, application, in-

line, radar, imaging, synthetic-aperture-radar, MIMO, coherent, portal-scanner, high-frequency, conveyor belt

## 1 Distance measurement

Before we look at imaging systems, however, let us first consider how a radar system measures the distance to an object in the first place. Usually, explanations use the concept of pulsed radar systems. In the transceiver path, pulses are generated and emitted. The pulse propagates until it reflects off an object, and the signal is beamed back to the radar. The time between the transmission of the pulse and the reception of the reflected pulse is twice the distance between the radar and the target. If there are several targets in the direction of propagation, the radar system measures the different echoes, provided the pulse is short enough. This approach, still used in many air surveillance systems, is unsuitable for industrial applications. System concepts which can create extremely short pulses to generate a sufficiently high-range resolution are expensive. While resolutions in the centimeter or meter range are sufficient for long distances, industrial applications usually require resolutions in the centimeter to the millimeter-wave range, sometimes even down to the micrometer range. However, the generation of extremely short pulses with simultaneously high energy and the necessary back-end structures with high sampling rates are uneconomical for industrial applications.

For this reason, the basis of almost all low-cost systems are approaches based on frequency-modulation. Here, a frequency ramp is emitted. As with the pulsed concept, the transmitted signal is reflected at the target and radiated back to the radar. The received signal is mixed with the currently transmitted signal at the receiver. Since the frequency modulation is continuous, the signal's transit time to the target and back means that the currently transmitted frequency no longer corresponds to the received frequency (Fig. 1). A constant ramp slope results in constant frequency  $\omega_a$  of the output signal  $s_a$ :

$$s_a \approx A \cdot \cos(\underbrace{\dot{\omega} \tau t}_{\omega_a}), \text{ with } \dot{\omega} = 2\pi \frac{B}{T}$$



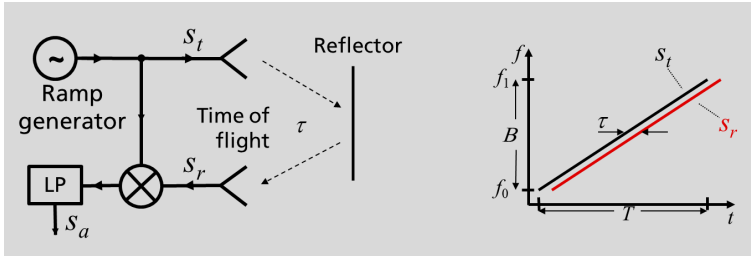
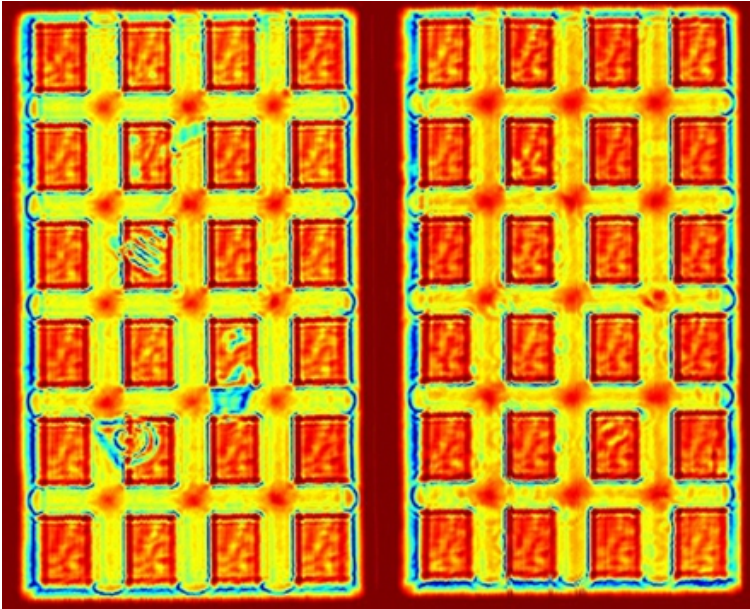


Figure 1: FMCW Principle.

The IF frequency  $\omega_a$  is directly proportional to distance. In contrast to a pulsed system, the system does not measure time but the frequency shift. This concept allows more precise measurements than a comparable pulsed system. Another advantage is that the transmitter emits continuously, so the total transmission power is not bundled into one short pulse. As a result, a much lower maximum transmission power is required to achieve the same system dynamics than a single pulse.

## 2 Mechanical scanners

Close-range applications usually use focusing optics with the object to be viewed in the focal point. If the object is moved in the focal point, it can be imaged two-dimensionally. The wavelength of the measuring frequency used determines the achievable lateral resolution. For a system at 300 GHz, focussing to below 500  $\mu\text{m}$  can theoretically be achieved with a short focal length. Since radar systems allow phase and time-of-flight measurement, objects can be reconstructed two- and three-dimensional. Here, a distinction must be made between resolution and measurement accuracy. The resolution determines the ability of a radar to separate two neighbouring objects from each other. The bandwidth of the radar system determines the minimum distance between two objects to be divided. It is usually a maximum of 10% to 30% of the centre frequency of the radar system. For the sake of simplicity, a distance resolution of 2 mm is assumed. If there is only one scattering center in this range cell, e.g. a flat surface, the range to this surface can be determined much more precisely via the phase information in



**Figure 2:** Transmission image of a bar of chocolate with (left) and without (right) impurities.

a coherent radar. Usually, the longitudinal measurement accuracy is higher by a factor of 100 than the lateral resolution of a corresponding system. Theoretically, packaged products can be inspected in this way (Fig. 2), but the measuring time could be faster for use in a conveyor line, so the technology is more suitable for single-piece inspection. This is especially true for moulded plastic parts where the composition and structure of internal layers need to be imaged. A fast imaging system with a single channel requires a quick mechanical scanning process and a high measuring speed of the sensor. High-frequency systems typically do not use detector concepts that allow continuous wave measurements with update rates between several thousand and hundred thousand measurements per second. Most scanning methods are based on a linear motorised XY scanner. The most significant disadvantage of 2D scanner systems is the low scanning speed, so a scan of an area of a DIN A4 sheet can take up to one hour. Faster motor

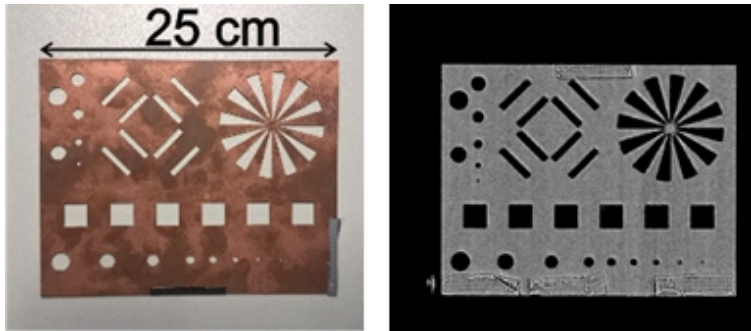


**Figure 3:** Comparison of the scan paths for a classic XY scanner (left) and a rotating scanner approach (right).

concepts with a lower positioning accuracy can realise such a measurement in one to two minutes. But even with this speed improvement, the mechanical 2D scanner concepts are far from the measurement time needed for inline quality control systems in production lines. The time loss is mainly caused by braking and acceleration of the linear motor stages. The change in direction causes a time gap that slows down the entire measuring system. A promising approach to speed up the measurement is to change from a linear motor concept to a rotating scanner concept (Fig. 3, right). A transmission measurement is carried out with these systems, such as the T-Sense. The device under test (DUT) passes between the two rotating probes. In the current generation of devices, 30,000 measuring points are scanned per second with this fast scanning method. This concept makes it possible, for example, to check a DIN A4 envelope within a few seconds.

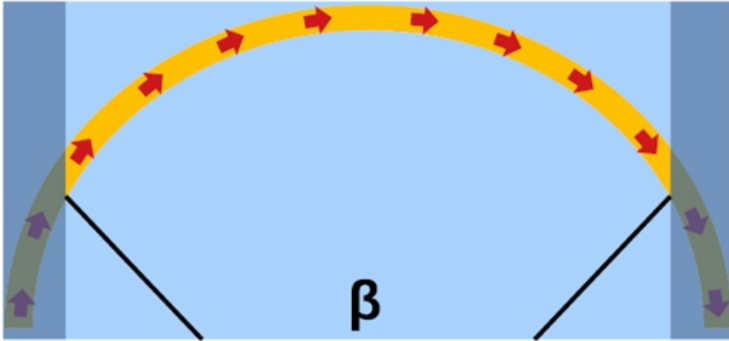
### 3 Illustration with SAR method

However, these measurement methods are unsuitable for larger structures such as window frames or wind turbine blades. For more complex 3D structures, synthetic aperture techniques (SAR) are often used. With these, the object to be examined is scanned at a greater distance with a coherent radar, and a synthetic aperture is created. In this case, no strongly bundling antenna concepts are used, as in the case of close-range scanning, but rather antennas with a particularly wide antenna lobe. A SAR radar processor stores all amplitudes and the corresponding phase position of the echo signals of all pulse repetition periods over a time  $T$  from all positions where the section to be observed is located in the antenna's footprint. During scanning, the individual reflection points of the object to be measured are detected at different

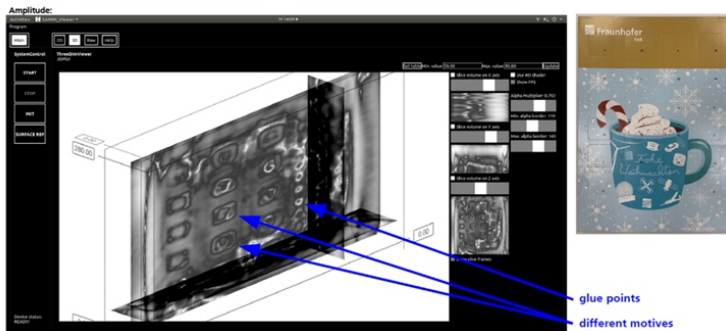


**Figure 4:** Test sample and the corresponding SAR image at 120 GHz.

angles, and a focussed image is generated by mathematical methods such as the “back projection” algorithm. When using a synthetic aperture in an endless motion, the numerical aperture of the image is determined only by the aperture angle of the antenna. As the distance from the target increases, the size of the synthetic aperture also increases so that the spatial resolution is independent of distance. For this reason, satellite-based radar systems often use SAR methods for Earth observation. However, they are also excellently suited for close-range applications and are used today, particularly for security scanners (Fig. 4). If you want to use a 3D SAR approach in an inline measurement configuration, you can use a TX/RX line and the conveyor belt’s movement to span the virtual aperture. A fully populated array is technologically complex due to the high number of channels required. In this context, MIMO lines with reduced TX/RX channels have recently been investigated. However, hybrid approaches can also be used that combine mechanical scanner concepts with the assembly line configuration. There is also the possibility of moving a single-channel system for slow belt speeds. Here again, a rotating scanning approach is a reasonable alternative [6]. In the implementation presented, the antenna rotates at a frequency of 10 Hz, so the duration per cycle ( $360^\circ$ ) is 100 ms. For a SAR configuration, the belt movement should ideally be orthogonal to the direction of movement of the antenna. Unfortunately, this is no longer guaranteed in the side ranges, as the direction of movement of the antenna corresponds to the direction of movement of the conveyor



**Figure 5:** The path of movement of the antenna (yellow), the measuring range of the semicircle ( $\beta$ ) and the side edges of the semicircle where no measurements are recorded (dark blue area).

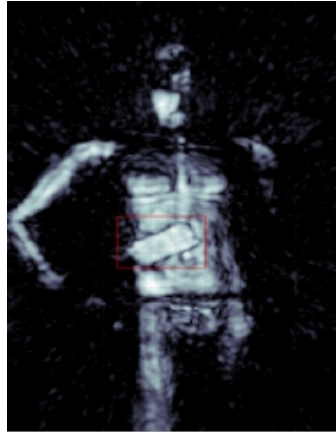


**Figure 6:** Visualisation of the 3D point cloud using the example of an advent calendar.

belt. Therefore, the measuring range is limited to the middle (Fig. 5, measuring range marked in light blue). Any sectional planes can now be placed in the resulting 3D point cloud to precise search for product defects (Fig. 6).

## 4 Imaging through MIMO radar systems

However, SAR methods require the movement of either the sensor or the object to be examined. Therefore, research is currently focusing on the development of radar-based camera systems. Since fully occupied antenna arrays are still too costly, MIMO systems are used. MIMO stands for Multiple-Input Multiple-Output. It is a system consisting of several transmitting and receiving antennas. MIMO systems can be developed for different operating modes, the most common being the design in which each transmitting antenna transmits a time-delayed transmission signal independently of the other transmitting antennas. The basic idea of this concept is to use an array of transmitters (TX array) to illuminate the object under test and an array of receivers (RX array) to detect the backscattered radiation coherently. This concept creates a virtual far-field antenna between the transmitter and each receiver antenna. The thinning of the array is achieved by design. By folding the TX and RX arrays, a fully occupied antenna array can thus be simulated. To simulate a fully occupied array with 100 elements, one needs ten transmitters and ten receivers in the best case and ten times the measurement time since all transmitters must be switched through one after the other. The virtual antenna elements' arrangement is usually made so that the resulting virtual array corresponds to the geometry of a fully occupied antenna array. The best-known application for this technology is the body scanner, which is now installed at numerous airports worldwide [7] [8]. The illustration (Fig. 7) shows a typical MIMO image of a person as created with comparable security scanners. When set up in one location, the MIMO radar system resembles a phased array antenna with a thinned-out antenna array. Each radiator has its transmit-receive module and A/D converter. But in a phased array antenna, each radiator transmits only one (possibly time-delayed) copy of a transmit signal generated in a central waveform generator. In a MIMO system with sequential control of the system, the measurement time increases according to the number of transmission channels. For this reason, MIMO systems are often used in which each radiator has its own waveform generator with which an individual signal form can be emitted. This unique waveform forms the basis for assigning the echo signals to their source. For more effective radar signal processing, each individual transmit signal can then be specif-



**Figure 7:** Radar image of a person taken with a MIMO system at 15 GHz.

ically modified ("adaptive waveform") to improve the signal-to-noise ratio (SNR) for each target in the subsequent sampling. Furthermore, suppose the generation of the respective waveform in the transmitters is synchronous with each other, i.e. based on a synchronising clock from a central "mother generator". In that case, this is referred to as coherent MIMO. By increasing the frequency of such systems and combining them with low-cost silicon technology, highly integrated radar cameras can be developed. The first compact prototypes already exist [9], but this development is still in its infancy and requires further steps, especially with regard to integration and the evolution towards higher frequencies. In the long term, however, 300 GHz radar cameras could be used in a wide range of industrial areas.

## 5 Conclusion

In recent years, radar systems have developed into indispensable sensor systems in the industrial environment. Their application area focuses on measurement environments with very harsh environmental conditions. At the moment, however, other advantages of radar systems are coming to the fore. In addition to the high measurement

speed, research focuses on imaging processes with high spatial resolution. 3D-SAR concepts are a promising approach. These future works apply in particular to real-time capability with simultaneous high assembly line speeds.

## References

1. C. Baer, T. Jaeschke, P. Mertmann, N. Pohl, and T. Musch, "A mmwave measuring procedure for mass flow monitoring of pneumatic conveyed bulk materials," in *IEEE Sensors*, 2014.
2. L. Piotrowsky, T. Jaeschke, S. Kueppers, J. Siska, and N. Pohl, "Enabling high accuracy distance measurements with fmcw radar sensors," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 12, pp. 5360-5371, doi: 10.1109/TMTT.2019.2930504, December 2019.
3. M. Vogt, "Radar sensors (24 and 80 ghz range) for level measurement in industrial processes," in *IEEE MTT-S International Conference on Microwaves for Intelligent Mobility*, 2018.
4. A. Bhutani, S. Marahrens, M. Gehringer, B. Göttel, M. Pauli, and T. Zwick, "The role of millimeter-waves in the distance measurement accuracy of an fmcw radar sensor," in *Sensors*, doi: 10.3390/s19183938, September 2019.
5. S. Ayhan, S. Scherr, P. Pahl, T. Kayser, M. Pauli, and Z. T., "High-accuracy range detection radar sensor for hydraulic cylinders," in *IEEE Sens.*, 14:734–746, doi: 10.1109/JSEN.2013.2287638, 2014.
6. C. Schwaebig, S. Wang, and S. Guetgemann, "A real-time sar image processing system for a millimetre wave radar ndt scanner," in *tm - Technisches Messen*, Band 88, Heft 7-8, June 2021.
7. R. Appleby and R. N. Anderton, "Millimeter-wave and submillimeter-wave imaging for security and surveillance," in *Proc. IEEE*, vol. 95, no. 8, pp. 1683-1690, August 2007.
8. D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," in *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 9, pp. 15811592, September 2001.
9. B. Baccouche and et al, "Three-dimensional terahertz imaging with sparse multistatic line arrays," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 23, no. 4, pp. 1-11, Art no. 8501411, doi: 10.1109/JSTQE.2017.2673552, August 2017.



# Quick-and-Dirty Computation of Voigt Profiles, Classification of Their Shapes, and Effective Determination of the Shape Parameter

Achim Kehrein<sup>1</sup> and Oliver Lischtschenko<sup>2</sup>

<sup>1</sup> Rhine-Waal University of Applied Sciences,  
Marie-Curie Str. 1, 47533 Kleve, Germany

<sup>2</sup> Ocean Insight - A Brand of Ocean Optics B.V.,  
Maybachstr. 11, 73760 Ostfildern, Germany

**Abstract** A spectral line is modeled by a Voigt profile, which is a convolution of a Gaussian and a Lorentzian. The width of the Gaussian is described by the standard deviation  $\sigma$ ; the width of the Lorentzian, by its lower quartile  $\gamma$ . One common method of computing a Voigt profile uses the real part of the complex-valued Faddeeva function, which is conceptually demanding and whose evaluation is computationally expensive. Other computational methods approximate Voigt profiles by simpler functions. We show that the shape of a Voigt profile only depends on the ratio  $\rho = \gamma/\sigma$  and, consequently, introduce a one-parameter family of standardized Voigt profiles. Then we present a conceptually simple and efficient numerical method for computing these standardized Voigt profiles – we only require basic numerical integration. Next we compute the second derivative by a finite-difference formula and determine empirically the relationship between the shape parameter  $\rho$  and the location of the inflection points described by their quantiles. This empirical relationship suffices to determine the parameters of a Voigt profile directly from data points and thus avoids the use of computationally costly, time-consuming, and sometimes failing general iterative fitting methods. In particular, this new and faster approach allows more real-time analyses of spectral data.

**Keywords** Voigt profile, classification, standardization, computation, line spectra analysis, spectroscopy

## 1 Introduction

The centered Voigt profile is defined as the convolution

$$V(x; \sigma, \gamma) = \int_{-\infty}^{+\infty} G(x - z; \sigma) L(z; \gamma) dz \quad (1)$$

of a centered Gaussian and a centered Lorentzian,

$$G(x; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad \text{and} \quad L(x; \gamma) = \frac{\gamma}{\pi(x^2 + \gamma^2)} \quad (2)$$

with width parameters  $\sigma > 0$  and  $\gamma > 0$ . For any pair of parameters, the total area of the Voigt profile is one,

$$\int_{-\infty}^{\infty} V(x; \sigma, \gamma) dx = 1 \quad . \quad (3)$$

Thompson reviews some computational algorithms [1]. Based on work by Johnson, Wuttke provides a library in which the Voigt profile is computed via the complex Faddeeva function [2].

Section 2 briefly reviews the geometries of the Gaussian and the Lorentzian. The section particularly stresses that up to scaling and shifting there is only one shape of a Gaussian - the standardized Gaussian is the shape prototype. Moreover, the inflection point of the Gaussian reveals the width parameter. Section 3 shows that the shape of a Voigt profile depends only on the ratio of the parameters  $\rho = \gamma/\sigma$ . Therefore Voigt profiles form a one-parameter family of the standardized form  $V(x; 1; \rho)$  with *shape parameter*  $\rho > 0$ . Then, Section 4 presents an elementary numerical method to compute these standardized Voigt profiles. Finally, Section 5 applies numerical differentiation to the computed standardized Voigt profiles and establishes an empirical relationship between the location of the point of inflection and the ratio parameter  $\rho$ . This empirical relationship shows how  $\rho$  and eventually the parameters  $\gamma$  and  $\sigma$  can be read of a graph of a Voigt profile.

The relationship between the inflection point and the shape parameter allows to match Voigt profiles to line spectra directly without having to use general iterative fitting algorithms. Section 6 sketches a procedure to do so.

## 2 Geometries of the Gaussian and Lorentzian

Of course, the Gaussian does not need an introduction. We review only briefly the aspects relevant to our treatment of the Voigt profile.

Any Gaussian can be transformed into any other Gaussian by a linear transformation. So, the tabulated standard Gaussian is the shape prototype of all Gaussians. See Figure 1.

The transformation rule

$$G(x; \sigma/\alpha) = \frac{\alpha}{\sigma\sqrt{2\pi}} e^{-\alpha^2 x^2 / (2\sigma^2)} = \alpha \cdot G(\alpha \cdot x; \sigma) \quad (4)$$

with scaling parameter  $\alpha > 0$  is of particular interest. For example, for  $\alpha > 1$  the expression on the right-hand side describes that the graph is compressed horizontally and stretched vertically by the factor  $\alpha$ . The area stays the same. This has the same effect as, on the left-hand side, dividing the standard deviation  $\sigma$  by  $\alpha$ , i.e. the effect of consistently compressing the width parameter. Consequently, for all Gaussians, the inflection points are invariantly one standard deviation away from the maximum. Also, the inflection points are invariantly located at the quantiles 0.1587 and 0.8413.

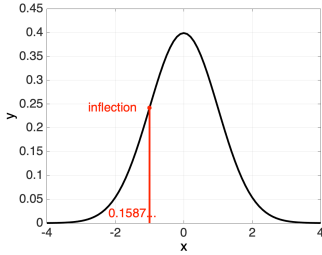
A Lorentzian also looks bell-shaped. See Figure 2. However, a Lorentzian approaches the horizontal asymptote  $y = 0$  so slowly that the improper integrals for the expected value and the standard deviation diverge. Regardless of the symmetry about zero, the expected value and the standard deviation are undefined. We need another quantity to describe the width of a Lorentzian.

The values  $\pm\gamma$  are the upper and lower quartiles. They are the locations that cut off the top and bottom 25% of the area under the Lorentzian.

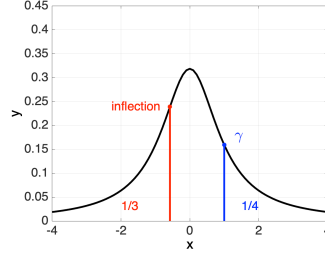
As for the Gaussian we have the transformation rule

$$\begin{aligned} L(x; \gamma/\alpha) &= \frac{\gamma/\alpha}{\pi(x^2 + \gamma^2/\alpha^2)} = \frac{\gamma/\alpha}{\pi/\alpha^2((\alpha \cdot x)^2 + \gamma^2)} \\ &= \frac{\alpha\gamma}{\pi((\alpha \cdot x)^2 + \gamma^2)} = \alpha \cdot L(\alpha \cdot x; \gamma) \end{aligned}$$

for  $\alpha > 0$ . For example, halving the parameter  $\gamma$  (left-hand side), compresses the Lorentzian horizontally by the factor two and doubles it



**Figure 1:** The Gaussian with  $\sigma = 1$ . The inflection points deviate at  $\pm\sigma$ . The left inflection point has the quantile rank  $\approx 0.1587$ .



**Figure 2:** The Lorentzian with  $\gamma = 1$ . The upper and lower quartiles are at  $\pm\gamma$ . The left inflection point has the quantile rank  $1/3$ .

vertically (right-hand side). The area stays the same. The parameter  $\gamma$  is a sensible width parameter and has an invariant geometric meaning.

### 3 Standardization and Classification of Voigt Profiles

Let  $\alpha > 0$ . For a Voigt profile we obtain the transformation rule

$$\begin{aligned} V(x; \sigma/\alpha, \gamma/\alpha) &= \int_{-\infty}^{\infty} G(x-z; \sigma/\alpha) L(z; \gamma/\alpha) dz \\ &= \int_{-\infty}^{\infty} \alpha \cdot G(\alpha \cdot (x-z); \sigma) \alpha \cdot L(\alpha \cdot z; \gamma) dz \\ &= \int_{-\infty}^{\infty} \alpha^2 \cdot G(\alpha x - \alpha z; \sigma) L(\alpha \cdot z; \gamma) dz \end{aligned}$$

Substitute  $u = \alpha \cdot z$ , hence  $du = \alpha dz$ ,

$$\begin{aligned} &= \alpha \int_{-\infty}^{\infty} G(\alpha x - u; \sigma) L(u; \gamma) du \\ &= \alpha \cdot V(\alpha \cdot x; \sigma, \gamma) \end{aligned}$$

In particular, we get for  $\alpha = \sigma$  a *standardized* expression with Gaussian width parameter 1,

$$V(x; 1, \gamma/\sigma) = \sigma \cdot V(\sigma \cdot x; \sigma, \gamma) \quad . \quad (5)$$

Equivalently, every Voigt profile is a suitably scaled standardized Voigt profile,

$$V(u; \sigma, \gamma) = \frac{1}{\sigma} \cdot V\left(\frac{u}{\sigma}; 1, \frac{\gamma}{\sigma}\right) \quad . \quad (6)$$

The shape of a Voigt profile only depends on the ratio  $\rho = \gamma/\sigma$ . The Voigt profiles can be classified into different shapes with respect to the single parameter  $\rho > 0$ .

Now we show that  $V(x; \sigma, \gamma)$  and  $V(x; \sigma/\alpha, \gamma/\alpha) = \alpha \cdot V(\alpha \cdot x; \sigma, \gamma)$  with homogeneously scaled parameters have the inflection points at the same quantiles. Let  $p$  denote the  $x$ -coordinate of an inflection point of  $f(x) = V(x; \sigma, \gamma)$ , so  $p$  is a zero of the second derivative  $f''$ . The second derivative of the scaled function satisfies

$$\frac{d^2}{dx^2}(\alpha \cdot V(\alpha \cdot x; \sigma, \gamma)) = \frac{d^2}{dx^2}(\alpha \cdot f(\alpha \cdot x)) = \alpha^3 \cdot f''(\alpha \cdot x) \quad , \quad (7)$$

which possesses the correspondingly scaled zero  $p/\alpha$ . The quantile rank at this position is given by

$$\int_{-\infty}^{p/\alpha} \alpha \cdot f(\alpha \cdot x) dx = \int_{-\infty}^{\alpha \cdot p/\alpha} f(u) du \quad , \quad (8)$$

where we substituted  $u = \alpha \cdot x$  and  $du = \alpha \cdot dx$ . The right-hand side describes the quantile rank of the unscaled function at the inflection point  $p$ . The quantile rank of the inflection point is a scaling invariant.

Section 5 establishes empirically an increasing relationship between the shape parameter  $\rho$  and the quantile rank of the smaller inflection point. There is a one-to-one correspondence between the Voigt profile shapes and the parameter  $\rho = \gamma/\sigma$ .

## 4 Quick-and-Dirty Computation of Voigt Profiles

We compute a standardized Voigt profile  $V(x; 1, \rho)$  approximately by suitably truncating the improper convolution integral and by numerically integrating the remaining definite integral.

Due to the symmetry of the Gaussian,  $G(x - z; \sigma) = G(z - x; \sigma)$ , the Voigt profile value at  $x$  equals the integral with respect to  $z$  over the

product of the Gaussian with mean  $x$  and the centered Lorentzian.

$$\int_{-\infty}^{\infty} G(x - z; 1) L(z; \rho) dz = \int_{-\infty}^{\infty} G(z - x; 1) L(z; \rho) dz \quad (9)$$

We know that the values of the Gaussian are very close to zero outside  $[\mu - 4\sigma, \mu + 4\sigma]$ , so a sensible truncation is

$$\int_{-\infty}^{\infty} G(z - x; 1) L(z; \rho) dz \approx \int_{x-4}^{x+4} G(z - x; 1) L(z; \rho) dz \quad (10)$$

Since both functions, the Gaussian and the Lorentzian, can be approximated quite accurately by polynomials on reasonably small intervals, a piecewise low-degree numerical integration formula is sufficient for practical accuracy. We use the iterated trapezoid rule and iterated midpoint rule so that the proximity of the two estimates indicates how accurate they are. Moreover, the arithmetic mean of these values produces the result of the iterated trapezoid rule with twice as many subintervals. Finally, a weighted average of the two iterated trapezoid values coincides with Simpson's rule. These steps are the beginning of Romberg's scheme and can be extended, if more accuracy is needed.

To set up the iterated integration rules we divide  $[x - 4, x + 4]$  into  $n$  equidistant subintervals of length  $\Delta z = 8/n$ . The trapezoid rule uses the nodes  $z_k = x - 4 + k \cdot \Delta z$  with  $0 \leq k \leq n$ .

$$\begin{aligned} V(x; 1, \rho) &\approx T_n(x; \rho) = \left( \frac{G(z_0 - x; 1) L(z_0; \rho)}{2} + \right. \\ &\quad \left. \sum_{k=1}^{n-1} G(z_k - x; 1) L(z_k; \rho) + \frac{G(z_n - x; 1) L(z_n; \rho)}{2} \right) \cdot \Delta z \\ &= \left( \frac{G(-4; 1) L(x - 4; \rho)}{2} + \sum_{k=1}^{n-1} \frac{1}{\sqrt{2\pi}} e^{-(z_k - x)^2/2} \frac{\rho}{\pi(z_k^2 + \rho^2)} + \right. \\ &\quad \left. \frac{G(4; 1) L(x + 4; \rho)}{2} \right) \cdot \frac{8}{n} \\ &= \frac{8\rho}{n\pi\sqrt{2\pi}} \cdot \left( \frac{e^{-8}}{2((x - 4)^2 + \rho^2)} + \right. \\ &\quad \left. \sum_{k=1}^{n-1} \frac{e^{-(-4+8k/n)^2/2}}{(x - 4 + 8k/n)^2 + \rho^2} + \frac{e^{-8}}{2((x + 4)^2 + \rho^2)} \right) \end{aligned}$$

On the other hand, let  $m_k = x - 4 + (k - 1/2)\Delta z$  with  $0 \leq 1 \leq n$  denote the midpoints of the subintervals. The iterated midpoint rule is

$$\begin{aligned} V(x; 1, \rho) &\approx M_n(x; \rho) = \sum_{k=1}^n G(m_k - x; 1) L(m_k; \rho) \cdot \Delta z \\ &= \sum_{k=1}^n \frac{1}{\sqrt{2\pi}} e^{-(m_k - x)^2/2} \frac{\rho}{\pi(m_k^2 + \rho^2)} \cdot \frac{8}{n} \\ &= \frac{8\rho}{n\pi\sqrt{2\pi}} \sum_{k=1}^n \frac{e^{-(-4+8(k-1/2)/n)^2/2}}{(x - 4 + 8(k - 1/2)/n)^2 + \rho^2} \end{aligned}$$

The trapezoid value with twice as many subintervals is the arithmetic mean

$$T_{2n}(x; \rho) = (T_n(x; \rho) + M_n(x; \rho))/2 \tag{11}$$

and Simpson's rule is the weighted average

$$S_n(x; \rho) = \frac{4 \cdot T_{2n}(x; \rho) - T_n(x; \rho)}{3} \approx V(x; 1, \rho) \quad . \tag{12}$$

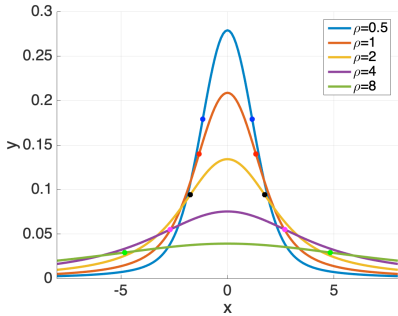
Figure 3 shows some computed Voigt profiles for various ratio parameters  $\rho$  that have been computed using the above formulas with  $n = 32$  subintervals at the equidistant arguments  $x \in \{-16.0, -15.9, -15.8, \dots, 16.0\}$ . We use equidistant arguments to prepare for the consistent use of a finite-difference formula to determine numerically the second derivative of the Voigt profile.

## 5 Empirical Relationship between the Shape Parameter and the Points of Inflection

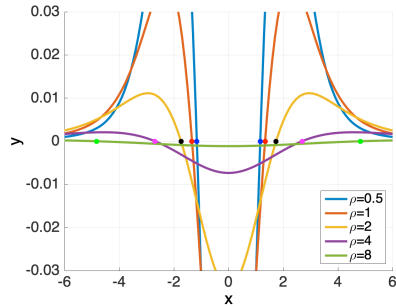
To approximate the second derivative of a Voigt profile based on the equidistant samples we use the finite difference formula

$$\frac{d^2}{dx^2} V(x; 1, \rho) \approx \frac{V(x - h; 1, \rho) - 2V(x; 1, \rho) + V(x + h; 1, \rho)}{h^2} \quad . \tag{13}$$

Figure 4 shows the second derivatives of Voigt profiles for various parameters  $\rho$  that are computed by the finite difference formula. We



**Figure 3:** The standardized Voigt profiles  $V(x; 1, \rho)$  for various  $\rho = \gamma/\sigma$  ratios and their numerically determined inflection points.



**Figure 4:** The numerically determined second derivatives of  $V(x; 1, \rho)$  for various  $\rho$ . The zeros (inflection points of  $V$ ) depend monotonically on the shape-parameter.

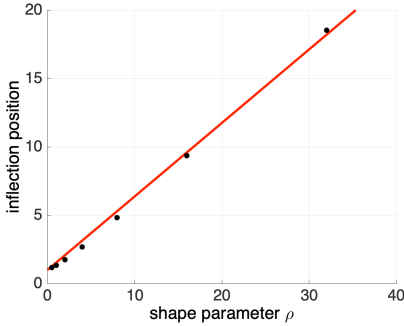
**Table 1:** The positions of the inflection points for various shape parameters  $\rho$ . The limiting quantile rank for  $\rho \rightarrow \infty$  seems to be  $1/3$ , see Figure 2.

| Index $\nu$ | Shape Param. $\rho_\nu$ | Inflection Points    | Quantile Rank $Q_\nu$ |
|-------------|-------------------------|----------------------|-----------------------|
| Gauss       | 0                       | $(\pm 1, 0.242)$     | 0.1587                |
| 1           | 0.5                     | $(\pm 1.16, 0.179)$  | 0.2190                |
| 2           | 1                       | $(\pm 1.34, 0.140)$  | 0.2550                |
| 3           | 2                       | $(\pm 1.74, 0.094)$  | 0.2922                |
| 4           | 4                       | $(\pm 2.69, 0.055)$  | 0.3178                |
| 5           | 8                       | $(\pm 4.83, 0.029)$  | 0.3288                |
| 6           | 16                      | $(\pm 9.35, 0.015)$  | 0.3321                |
| 7           | 32                      | $(\pm 18.53, 0.007)$ | 0.3330                |
| Lorentz     | $\infty$                | n.a.                 | $1/3$                 |

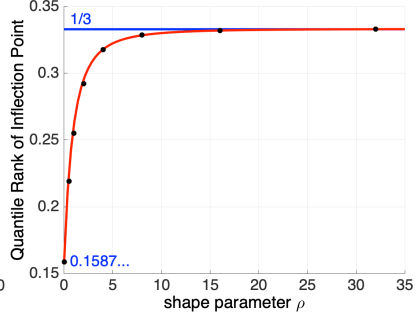
see qualitatively that the deviation of the inflection points from the mean increases with the parameter  $\rho$ . We compute estimates of these positions by finding the pair of neighboring second-derivative values with a sign change to which we apply linear interpolation. The function value estimates of the inflection points are also computed as linear interpolations of the neighboring already computed function values. The results are documented in Table 1.

According to the scatter plot in Figure 5 we start with the linear





**Figure 5:** The location of the points of inflection depends monotonically on the shape-parameter  $\rho$ . A linear fit with theoretically prescribed intercept 1 provides a reasonable fit.



**Figure 6:** Relationship between quantiles of the smaller point of inflection and the shape parameter. The shown function is given by  $QR = 1/3 - 1/(\rho + C)^{2.3}$  with  $C = (1/3 - 0.1587)^{-1/2.3}$ .

model

$$x = 1 + m \cdot \rho \quad , \quad (14)$$

in which we choose the intercept 1 from the limiting case as the position of the inflection point of the Gaussian. Based on a least squares approximation for the data points  $(\rho_\nu, x_\nu)$ ,  $1 \leq \nu \leq n$ , we compute the slope estimate

$$\hat{m} = \frac{\sum_{\nu=1}^n \rho_\nu (x_\nu - 1)}{\sum_{\nu=1}^n \rho_\nu^2} \approx 0.536 \quad . \quad (15)$$

There is another useful relationship. We pair the shape parameter  $\rho$  with the quantile rank of the left inflection point. We have already computed estimates of the symmetrically located points of inflection. Now we numerically integrate the Voigt profile between the inflection points, subtract this estimated area from one, and divide it by half to obtain the quantile. The numerical integration uses the iterated Simpson rule on the equidistant nodes between the inflection points and, separately, computes the trapezoids from the inflection points to the neighboring node inside. The widths of these trapezoid are smaller than the

equidistant stepsize since we estimated the position (and the value) of the inflection point by linear interpolation. The resulting quantiles are listed in Tab 1 and the relationship is shown in Figure 6. By inspired guessing we have found

$$QR \approx \frac{1}{3} - \frac{1}{(\rho + C)^k} \quad \text{with } C = (1/3 - 0.1587)^{-1/k} \text{ and } k \approx 2.3. \quad (16)$$

## 6 Application to Line Spectra

To analyze a line spectrum of Voigt profiles we propose the following procedure. First, numerically compute the first and second derivatives of the spectral data. A spectral line consists of a subinterval  $[\ell, m]$  with positive first derivative and a subinterval  $[m, r]$  with negative first derivative. Integrate the original data over  $[\ell, r]$ , keeping track of the integral values from  $\ell$  to (1) the first sign change of the second derivative, (2) the sign change of the first derivative at  $m$ , (3) the second sign change of the second derivative, and (4) the right endpoint  $r$ . Use asymmetries such as “the value at (4) is not twice the value at (2)” or “the values at (1) and (3) are not symmetric about (2)” to determine overlapping spectral lines and suitably adjust the values. The adjusted ratio (1)/(4) determines the shape parameter, the adjusted horizontal difference between the location of the maximum and the inflection points determines the parameter  $\sigma$ , and, finally, the adjusted value (4) determines the required vertical scaling of the Voigt profile.

The details of this procedure, especially the necessary adjustments for significantly overlapping spectral lines are the subject of current research.

## References

1. W. J. Thompson, “Numerous neat algorithms for the voigt profile function,” *Computers in Physics*, vol. 7, no. 6, pp. 627–631, 1993.
2. J. Wuttke and S. G. Johnson, “libcerf, numeric library for complex error functions,” <https://apps.jcns.fz-juelich.de/man/voigt.html>, [Accessed: 2022-12-17].



# International Conference on Optical Characterization of Materials

Each material has its own specific spectral signature independent if it is food, plastics, or minerals. New trends and developments in material characterization have been discussed as well as latest highlights to identify spectral footprints and their realizations in industry.

## Conference topics

- **Agriculture**
- **Food Inspection**
- **Recycling and Environment**
- **Spectral Applications**
- **Spectral Data Processing**
- **Spectral Sensors**

The International Conference on Optical Characterization of Materials (OCM-2023) was organized by the Karlsruhe Center for Spectral Signatures of Materials (KCM) in cooperation with the German Chapter of the Instrumentation & Measurement Society of IEEE.

KCM is an association of institutes of the Karlsruhe Institute of Technology (KIT) and the business unit Inspection and Optronic Systems of the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (Fraunhofer IOSB).

ISBN 978-3-7315-1274-5



9 783731 512745

ISSN 2510-7240  
ISBN 978-3-7315-1274-5

Printed on FSC-certified paper