



5th INTERNATIONAL CONFERENCE ON

Optical Characterization of Materials

MARCH 17th – 18th, 2021
KARLSRUHE | GERMANY

J. BEYERER | T. LÄNGLE (Eds.)



Scientific
Publishing

Jürgen Beyerer | Thomas Längle (Eds.)

OCM 2021

5th International Conference on
Optical Characterization of Materials

March 17th – 18th, 2021
Karlsruhe | Germany

OCM 2021

5th International Conference on
Optical Characterization of Materials

March 17th – 18th, 2021

Karlsruhe | Germany

Edited by

Jürgen Beyerer | Thomas Längle

Veranstalter

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

Dieser Tagungsband ist auch als Onlineversion abrufbar unter
<http://dx.doi.org/10.5445/KSP/1000128686>

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



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



*The cover page is licensed under a Creative Commons
Attribution-No Derivatives 4.0 International License (CC BY-ND 4.0):
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2021 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 2510-7240

ISBN 978-3-7315-1081-9

DOI 10.5445/KSP/1000128686

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 and 2019 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 fifth time in March 2021.

The OCM 2021 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 Automated Visual Inspection of the Fraunhofer Institute of Optics, 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 22 papers as talks, a keynote lecture and several practical demonstrations.

The present book is based on the conference held in Karlsruhe, Germany from March 17–18, 2021. 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 μm), 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 invested a generous amount of their time to suggest possible improve-

ments of the papers. The help of Jürgen Hock and Lukas Dippon 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 2021

Jürgen Beyerer
Thomas Längle

General Chairs

Jürgen Beyerer	Karlsruhe
Thomas Längle	Karlsruhe

Program Chair

Thomas Längle	Karlsruhe
---------------	-----------

Program Committee

Jochen Aderhold	Braunschweig
Oliver Albrecht	Dresden
Johannes Anastasiadis	Karlsruhe
Sebastian Bauer	Madison (Wisconsin)
Mark Bücking	Schmallenberg
Andrea Büttner	Freising
Robin Gruna	Karlsruhe
Michael Heizmann	Karlsruhe
Thomas Hofmann	Würzburg
Olfa Kanoun	Chemnitz
Anna Kicherer	Siebelingen
Andrea Krähmer	Berlin
Félix Salazar	Madrid
Maximilian Schambach	Karlsruhe
Heinar Schmidt	Kulmbach
Henning Schulte	Karlsruhe
Hermann Wotruba	Aachen
Bernhard Zagar	Linz

Contents

Preface	i
---------------	---

Food

Are low-cost, hand-held NIR sensors suitable to detect adulterations of halal meat?	1
---	---

J. Müller-Maatsch, Y. Weese-poel, E. Roetgerink, M. Wijten, and M. Alewijn

Exotic Fruit Ripening Based on Optical Characterisation	11
---	----

A. Scheibelmasser, M. Jeindl, and G. Nakladal

A new approach for evaluation of meat freshness	21
---	----

N. Kasvand, A. Frorip, A. Kuznetsov, T. Püssa, L. Rusalepp, A. Sünter, and D. Anton

Towards the universal assessment of dietary intake using spectral imaging solutions	32
---	----

Y. Weese-poel, F. Daniels, M. Alewijn, M. Baart, J. Müller-Maatsch, G. Simsek-Senel, H. Rijgersberg, J. Top, and E. Feskens

Detection of pyrrolizidine alkaloid containing herbs using hyperspectral imaging in the short-wave infrared	45
---	----

J. Krause, N. Tron, G. Maier, A. Krüner, R. Gruna, T. Längle, and J. Beyerer

Agriculture

Phenoliner 2.0: RGB and near-infrared (NIR) image acquisition for an efficient phenotyping in grapevine research	58
--	----

X. Zheng, J. Krause, B. Fischer, R. Gruna, R. Töpfer, and A. Kicherer

Developing a handheld NIR sensor for the detection of ripening
in grapevine 70
*L. Gebauer, J. Krause, X. Zheng, F. Kronenwett, R. Gruna,
R. Töpfer and A. Kicherer*

Recycling

Fine metal-rich waste stream characterization based on RGB
data: Comparison between feature-based and deep learning
classification methods 84
N. Kroell, K. Johnen, X. Chen, and A. Feil

Other Applications

Improvement of Thermal Fringe Projection for Fast and Accurate
3D Shape Measurement of Transparent Objects 99
M. Landmann, H. Speck, J. T. Schmieder, S. Heist, and G. Notni

Measurement of the coefficient of linear thermal expansion
based on subjective laser speckle patterns 109
A. Spaett and B. Zagar

Algorithms

Vessel extraction from breast MRI 119
M. Gierlinger, D. Brandner, and B. G. Zagar

Generation of artificial training data for spectral unmixing by
modelling spectral variability using Gaussian random variables... 129
J. Anastasiadis and M. Heizmann

Line Spectra Analysis: A Cumulative Approach 141
A. Kehrein and O. Lichtschenko

Quality analysis for the production of pelletised materials using
Fourier Descriptors 153
S. Michlmayr, D. Neumann, and B. Zagar

Sensors

Fiber-Coupled MEMS-based NIR Spectrometers for Material Characterization in Industrial Environments 163
R. Zimmerleiter, P. Gattinger, K. Duswald, T. Reischer, and M. Brandstetter

Improvement of Roughness Measurement in Sub-micron Ranges Using Contrast-based Depolarization Field Components 173
F. Pöller, F. Salazar Bloise, M. Jakobi, J. Dong, and A. W. Koch

Short Paper

Multimodal OCT Imaging 185
B. Heise and I. Zorin and G. Hanneschlaeger, and M. Brandstaetter

A High Quality Image Stitching Process for Industrial Image Processing and Quality Assurance 190
R. Hoffmann

Are low-cost, hand-held NIR sensors suitable to detect adulterations of halal meat?

Judith Müller-Maatsch¹, Yannick Weesepeol¹, Emma Roetgerink¹,
Michiel Wijtten¹, and Martin Alewijn¹

¹ Wageningen Food Safety Research
Wageningen, The Netherlands

Abstract The demand of halal meat products is growing globally. Therefore, it is important to detect adulterations and food fraud attempts in a fast, non-invasive manner for example by using hand-held near-infrared (NIR) devices. In this study, samples of pork, lamb, beef and chicken were measured pure and in mixtures of 2, 5, 10, 25 and 50% pork in the non-pork meat samples, respectively. Five sensors were tested with varying wavelength range: Scio (740-1070 nm), Linksquare (400-1000 nm), Tellspec (900-1700 nm), MicroNIR (900-1650 nm), ASD Labspec 4 High-Res (350-1700 nm). A one-class-classification approach was used for data analysis, applying pork as the target group. For comparison, thresholds of the models were chosen to correctly identify 100% of the pork samples and 75% of all mixtures. Comparing the sensors upon the correct detection of all halal meat samples, i.e., no-pork containing ones, the Scio and the ASD Labspec performed best with an outcome of 34% and 32%, respectively. The Linksquare, MicroNIR and Tellspec were able to correctly identify 27%, 27%, and 10%, respectively, of the halal products. Concluding, the application of these five NIR devices are challenging when it comes to the detection of meat products from different species. Nonetheless, the usage of this application in combination with suitable chemometric approaches may contribute to the detection of food fraud in halal products.

Keywords Near-infrared sensors, pork, lamb, beef, chicken meat, one-class-classification

1 Introduction

The market for halal-certified products increases within Western societies. While halal products have been intended for Muslim consumers, Jewish consumers as well as vegetarians/vegans, and people with various types of allergies or dietary restrictions purchase halal-certified products [1, 2]. When it comes to halal meat, several differences are found to the commercial meat that is available in Western countries. Halal meat may only contain meat from ruminant species like cows or birds like chicken. Horse and pig meat are not considered halal. Besides the species also the feed that is fed to the animals plays an important role. Animals fed with additions of biosolids or animal protein concentrates must undergo a quarantine period with other feed before slaughtering. Moreover, halal meat may only be retrieved from a slaughtering process that renders animals immobile or unconscious, without killing it, prior to the blood drainage [2]. These differences in animal species, feed and slaughtering process have been hard to detect and trace. Hence, several cases of fraud occurred as listed by [3] or illustrated in detail by [4]. Both authors conclude that the main enabling factor of halal meat food fraud is the challenging detection of halal meat authenticity. One possible solution to overcome the issue of halal meat authenticity detection is applying spectroscopy. Spectral data may be collected for example by near-infrared (NIR) sensors. [5], [6], and [7] showed that the discrimination of animal species is possible when using a portable FT-IR or benchtop NIR sensor with a wavelength range reaching between 1000 and 2000 nm. although spectral data acquisition is fast and easy to conduct, the machines trialled in the past were very costly, heavy and bulky [8]. In this paper, we show the application of a one-class-classification (OCC) chemometric approach on data obtained by using several hand-held NIR sensors. OCC describes one specific class as the target class and returns predictions of samples being out or in the respective target class. In the case of halal meat detection, in particular the speciation issue, the target class was set as pork meat, i.e., non-halal meat. That means in that all samples that are “in”, do contain pork and are therefore not halal. On the other hand, all samples that are “out” do not contain pork and may be labelled as halal.

2 Materials and methods

2.1 Materials and sample preparation

Pork, beef, lamb and chicken meat was purchased at local butchers in Wageningen, the Netherlands. 40 samples of each species were purchased in 20 days, being two different meat parts per day per species, i.e., shoulder and leg of lamb, pig and cow or breast and drumstick from chicken. For reference purposes, all purchased, pure samples underwent a real-time PCR assay to validate the species identity. The method used has been described previously by [9]. All samples were purchased as intact meat and minced with a meat mincer Tristar VM-4310 (Smartwares Europe, Tilburg, Nederland). Mixtures of pork with beef, lamb or chicken, respectively, were prepared in the concentrations 2, 5, 10, 25, and 50% pork/ other meat (w/w). A randomised approach was used to make almost every day six mixtures using the two parts of lamb, beef and chicken mixed with pork, leading to 117 sample mixtures. In total 277 samples were measured, being 117 sample mixtures and 160 pure samples. To ensure that all samples have a similar storage period, all freshly prepared samples and mixtures were frozen, stored at -18 °C and thawed for 12h prior to the measurements.

2.2 NIR spectroscopy and data acquisition

Five different hand-held sensors (Fig. 2.1) were studied as follows:

- Scio (Consumer Physics, Herzliya, Israel), wavelength range 740-1070 nm, size 6.8 x 4.0 x 1.9 cm
- Linksquare (Stratio Inc, Paolo Alto, CA, USA), wavelength range 400-1000 nm, size 2.4 x 11.4 x 2.4 cm
- Tellspec (Tellspec Inc., Toronto, Ontario, Canada), wavelength range 900-1700 nm, size 8.2 x 6.6 x 4.5 cm
- MicroNIR (Viavi Solutions Inc, Santa Rosa, CA, USA (former JDSU)), wavelength range 900-1650 nm, size 4.5 x 4.4 x 4.0 cm
- ASD Labspec 4 HighRes (Malvern Panalytica, Almelo, the Netherlands (former ASD)), wavelength range 350-1700 nm, size 12.7 x 36.8 x 29.2 cm

The NIR hardware was calibrated according to the manufacturer instruction with a 99% diffuse reflectance standard. Measurements were done in diffuse reflectance mode by slightly pressing the optical part of the sensor to the sample or by slightly hovering the optical part above the sample in a distance of about 1 cm. For the Scio, Telspec, Linksquare and ASD Labspec sensors the integration time was set automatically by the manufacturer, for MicroNIR the integration time was set at 8 ms with 200 scans. Spectral measurements were conducted at room temperature when the meat had a temperature between 19 and 21 °C. The measurements were repeated four times.



Figure 2.1: Scio, Linksquare, Telspec, MicroNIR, and ASD Labspec (pictures courtesy of the respective hardware manufacturers).

2.3 Data analysis

Outliers were excluded manually using principal component analysis after standard normal variate (SNV) pre-processing in Unscrambler X 10.5 (Camo Analytics AS, Oslo, Norway). Out of the 1108 acquired spectra, 12 of Scio, 10 of Linksquare, 144 of Telspec, 16 of MicroNIR, and 13 of ASD Labspec were excluded, resulting in 1096, 1098, 964, 1092, and 1095 spectra for Scio, Linksquare, Telspec, MicroNIR and ASD Labspec, respectively. It is worth mentioning that the outliers are probably measurement errors, as no sample was found to be an outlier for all sensors. The OCC chemometric approach was conducted in R 3.6.1 (R Core Team, 2018) as described in detail by [10] and [11]. The same R-packages were used as previously reported. Among commonly used pre-processing methods and algorithms the most suited combination was picked according to (area under the receiver operator characteristic) AUROCs of the target class (pork) against the individual other classes, namely beef, lamb and chicken. For each sensor, three models were picked manually. Averaged class distances of four

repetitive measurements (in a row on one occasion) were used for further calculations. Classification results of each model were fused into a final classification using a decision tree, i.e., if more than one out of the three models classified a sample as 'out-of-class' it was classified as 'not containing pork'.

3 Results and discussion

3.1 Raw NIR data

The five sensors used in this study resulted in five very different groups of NIR spectra (Fig. 3.1). Next to the different measurement range of the sensors, also the hardware technology used differs for each sensor, resulting in different responses at each wavelength.

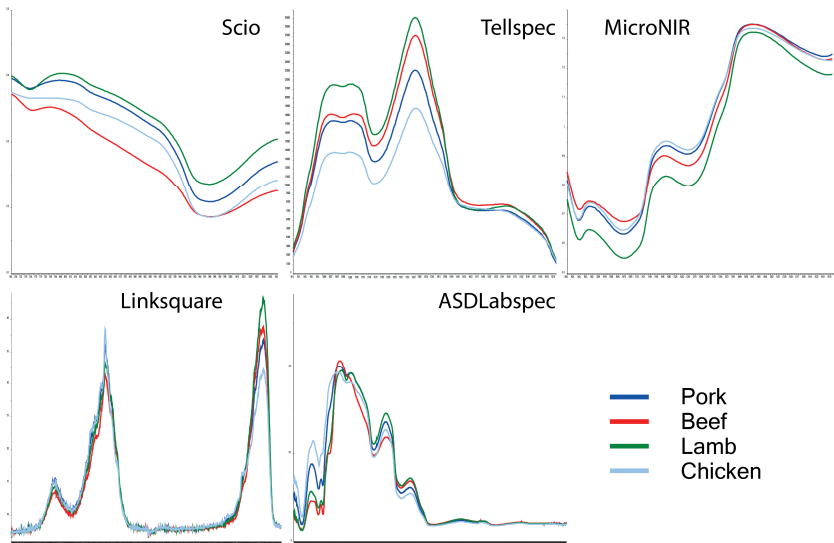


Figure 3.1: Raw data of all pure samples averaged according to meat species (Wavelength ranges were Scio 740-1070 nm, Linksquare 400-1000 nm, TellSpec 900-1700nm, MicroNIR 900-1650 nm and ASD Labspec 4 HighRes 350-1700nm).

3.2 Models picked for the different sensors

- For the Scio sensor all models only used the fourth quarter section from the four sections with equal lengths sorted by increasing wavelength. The first model included SNV as pre-processing and used Soft Independent Modelling of Class Analogies (SIMCA) with three principal components (PCs) as selected by a 5-fold (inner loop) cross-validation. The second model used the 1st derivative (Savitzky-Golay) an 11-point filter length and Principal Components Analysis (PCA) residual, calculating the sample residuals (Q residuals) with three principal components as selected using a 5-fold (inner loop) cross-validation. The third model included the 2nd derivative (Savitzky-Golay, 11-point filter length) and One-Class Support Vector Machine (OCSVM) with radial basis kernel and automatic parameter estimation.
- The Linksquare sensor models included in the first model SNV followed by baseline correction (SNV-DT, Detrend). It used only the fourth section of the spectra and a SIMCA (3PCs). The second model combined the 1st derivative (Savitzky-Golay, 11-point filter length), used only the fourth section of the spectra and calculated the distance to the k-Nearest Neighbor (kNN), with two neighbours as selected using a 5-fold (inner loop) cross-validation. The third model used the 2nd section of the spectra, SNV-DT and kNN as the algorithm.
- All Tellspec models were based on the fourth section of the spectra. The first one included the 1st derivative (Savitzky-Golay, 11-point filter length) and PCA (3PCs), the second the 2nd derivative (Savitzky-Golay, 11-point filter length) and kNN (2 neighbors) and the third SNV and PCA (3 PCs).
- MicroNIR spectra were analysed using only the fourth section of the spectra. The first model included the 2nd derivative (Savitzky-Golay, 11-point filter length), and PCA (3 PCs), the second the same 2nd derivative processing and kNN (2 neighbors) and the third the 1st derivative (Savitzky-Golay, 11-point filter length), and PC (3 PCs).

- The first ASD Labspec model used the fourth section of the spectra, SNV-DT and SIMCA (3PCs). The second SNV, the second section of the spectra and SIMCE (3PCs), and the third the 2nd derivative (Savitzky-Golay, 11-point filter length), the fourth section of the spectra and OCSVM.

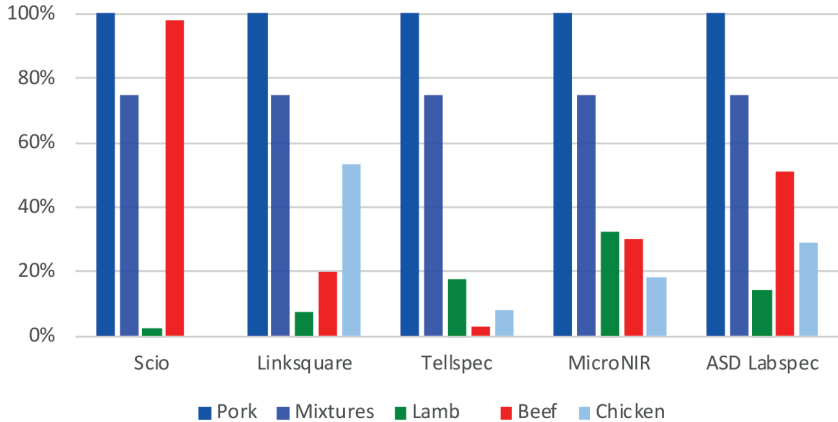


Figure 3.2: Correct classification in % for pork, mixtures (2 to 50% pork in lamb, beef and chicken (w/w)), lamb, beef and chicken samples deriving from the fused classification results of three models per sensor, respectively. Thresholds were set to achieve a correct identification of 100% pork and 75% mixture samples.

3.3 Determination of thresholds after fusion of classification results

The thresholds of the fused classification results were set in the way that 100% of all samples containing 100% pork were correctly identified as “pork” (Fig. 3.2). In that case, with no false negatives, the correct classification of pure beef, chicken, and mutton samples, as well as mixtures, was lower than 100%. The correct classification of mixtures was set by adjusting the thresholds at 75%, whereby the ones containing less than 10% pork were wrongly identified in most cases. In summary, the thresholds were chosen in a way that some halal products, i.e., pure lamb, beef and chicken, were wrongly identified as containing pork but no pork was undetected. In an industrial setting, this approach should

be beneficial as it reduces the number of samples that must be identified by more complex, costly and time-intensive experiments such as polymerase chain reactions (PCR).

3.4 Comparing sensors

Overall the Scio and the ASD Labspec performed best as still 34% or 32%, respectively, of the halal products were classified correctly (Fig. 3.2). Nevertheless, the Scio sensor was only able to discriminate beef samples as 'non-pork' but none of the other meat species such as chicken and lamb. Here, the ASD labspec sensor performed better identifying correctly, 15% of lamb, 51% of beef and 29% of chicken samples. In contrast to the two mentioned sensors, The Linksquare and MicroNIR identified 27% of the halal-products correctly and the Tellspec only 10%. Both the Linksquare and the ASD Labspec performed well on the identification of chicken with 53% and 29% correctly identified. This may be caused by the wider wavelength range that includes part of the visible spectrum 400-1000 nm and 350-1700 nm for Linksquare and ASD Labspec, respectively. Chicken meat has a lighter colour than the other meat species and may be discriminated visually. Also, both sensors included models that used the second section of the spectrum where the visible wavelengths may be found. In literature, the discrimination of meat according to its protein, moisture and fat content was successful using the Scio and another miniaturized NIR device. [12] showed that samples containing fat from 5% to 43%, protein from 12% to 23% and a moisture content of 35% to 69% may be correctly classified. In contrast to the study of [12], the samples used in this study are probably too close to each other concerning their composition. According to the USDA database, raw pork contains 72.6-72.9 g/100 g water, 19.6-20.5 g/100 g protein, 5.4-7.1 g/100 g fat, whereas chicken may contain 73.9-76.8 g /100 g water, 19.4-22.5 g/100 g protein, 2.6-3.7 g/100 g protein fat, lamb 72.5-74.4 g/100 g water, 19.7-21.1 g/100 g protein, 6.5-8.3 g/100 g fat and beef 73.5-73.6 g/100 g water, 20.4-20.5 g/100 g protein, 5.4-6.7 g/100 g fat. These might be the reason why, so few samples of the pure, halal-meat products were correctly identified when the threshold was chosen to correctly detect mixtures in a 75% rate.

4 Conclusion and outlook

The application of fast, cheap and hand-held devices is challenging when it comes to the detection of meat products from different species. Nevertheless, the presented OCC approach enables the application of these devices to contribute to species admixture detection for halal-certified products. Five different devices were tested. The Scio and the ASD Labspec performed best followed by the Linksquare, MicroNIR and Tellspec sensor. Further research is conducted using hyperspectral imaging cameras that cover a wavelength range in the VIS-NIR from 400-1700 nm and include spatial information for classification attempts. Moreover, different slaughtering techniques and feeding will be included.

5 Funding

This work was funded by the Dutch Ministry of Agriculture, Nature and Food Quality (Knowledge base grant).

References

1. J. Pointing, Y. Teinaz, and S. Shafi, "Illegal labelling and sales of halal meat and food products," *The Journal of Criminal Law*, vol. 72, no. 3, pp. 206–213, 2008. [Online]. Available: <https://doi.org/10.1350/jcla.2008.72.3.496>
2. J. M. Regenstein, "Food chemistry and analysis for the purpose of kosher and halal," in *Encyclopedia of Food Chemistry*, L. Melton, F. Shahidi, and P. Varelis, Eds. Oxford: Academic Press, 2019, pp. 679 – 684. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B9780081005965224098>
3. A. Ruslan, N. H. Kamarulzaman, and M. Sanny, "Muslim consumers' awareness and perception of halal food fraud," 2017.
4. G. McElwee, R. Smith, and J. Lever, "Illegal activity in the uk halal (sheep) supply chain: Towards greater understanding," *Food Policy*, vol. 69, pp. 166 – 175, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306919217302750>
5. M. Schmutzler, A. Beganovic, G. Böhler, and C. W. Huck, "Methods for detection of pork adulteration in veal product based on ft-nir spectroscopy for laboratory, industrial and on-site analysis,"

- Food Control*, vol. 57, pp. 258 – 267, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0956713515002364>
6. L. Mamani-Linares, C. Gallo, and D. Alomar, "Identification of cattle, llama and horse meat by near infrared reflectance or transreflectance spectroscopy," *Meat Science*, vol. 90, no. 2, pp. 378 – 385, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0309174011002853>
 7. E. Restaino, A. Fassio, and D. Cozzolino, "Discrimination of meat patés according to the animal species by means of near infrared spectroscopy and chemometrics discriminación de muestras de paté de carne según tipo de especie mediante el uso de la espectroscopia en el infrarrojo cercano y la quimiometria," *CyTA - Journal of Food*, vol. 9, no. 3, pp. 210–213, 2011. [Online]. Available: <https://doi.org/10.1080/19476337.2010.512396>
 8. N. Prieto, O. Pawluczyk, M. E. R. Dugan, and J. L. Aalhus, "A review of the principles and applications of near-infrared spectroscopy to characterize meat, fat, and meat products," *Applied Spectroscopy*, vol. 71, no. 7, pp. 1403–1426, 2017, PMID: 28534672. [Online]. Available: <https://doi.org/10.1177/0003702817709299>
 9. I. M. Scholtens, T. W. Prins, and L. W. van Raamsdonk, "Specificity of a novel taqman pcr method for detection of poultry dna," *Food Control*, vol. 73, pp. 532 – 539, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0956713516304856>
 10. Y. Weesepoel, M. Alewijn, M. Wijtten, and J. Müller-Maatsch, "Detecting Food Fraud in Extra Virgin Olive Oil Using a Prototype Portable Hyphenated Photonics Sensor," *Journal of AOAC INTERNATIONAL*, 07 2020, qsa099. [Online]. Available: <https://doi.org/10.1093/jaoacint/qsaa099>
 11. J. Müller-Maatsch, M. Alewijn, M. Wijtten, and Y. Weesepoel, "Detecting fraudulent additions in skimmed milk powder using a portable, hyphenated, optical multi-sensor approach in combination with one-class classification," *Food Control*, vol. 121, p. 107744, 2021. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0956713520306605>
 12. A. Kartakoullis, J. Comaposada, A. Cruz-Carrión, X. Serra, and P. Gou, "Feasibility study of smartphone-based near infrared spectroscopy (nirs) for salted minced meat composition diagnostics at different temperatures," *Food Chemistry*, vol. 278, pp. 314 – 321, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0308814618319952>

Exotic Fruit Ripening Based on Optical Characterisation

Anton Scheibelmasser¹, Matthias Jeindl¹, and Gunar Nakladal²

¹ Insort GmbH,

Berndorf 166, 8324 Kirchberg an der Raab, Austria,

² Fruturat GmbH,

Fruturstraße 1, 8224 Hartl, Austria

Abstract Re-sellers and customers in super markets expect perfect quality fruit in terms of ripeness, sweetness, taste and a lack of inner and outer defects. Especially exotic fruits provide many challenges due to long transport routes and the logistics of ripening processes. Optical characterisation of the fruits could help ensure the required quality. A combination of visual and infrared light evaluation techniques allows the measurement of quality parameters that support the control system of the reseller's store and the delivery logistics to the super market. For the inspection of chemical fruit characteristics (e.g. dry matter, sugar content), a hyperspectral near-infrared sensor is used. Additionally, an RGB camera is responsible for the visual defect analysis. Based on this measurement principle, a ripening control machine was developed and tested in daily business, successfully.

Keywords Exotic fruits ripening, hyperspectral imaging, Sherlock Food Analyzer

1 Introduction

During the last years the demand of ripened mangos and avocados rose significantly because shops offered more and more „Ready to eat“ products. In former times the consumers found mostly unripened fruit and tried to achieve ripeness by simple methods at home. In most cases the result was unsatisfactory, keeping the consumption demand

on lower levels. In order to improve the situation, some fruit traders started their own ripening processes. The result was an explosion of sale quantities. The company Frutura GmbH [1] has long-time experience in ripening of bananas, leading to the idea of following the ripening trend. After these ripening processes, the result was often an uneven product with several defects and sorting had to be done manually. The market did not yet provide proper sorting equipment for these types of fruit. The Austrian company Insort [2] was selected as project partner, developer and producer of a new kind of sorting machine.

Characterisation of Exotic Fruits In order to deliver perfect exotic fruits, in the appropriate ripening stage, a few requirements have to be met. For mangos, size and weight have to meet a defined range and should be sorted in respective quality classes. The right water content (moisture, dry matter value) and sugar concentration (Brix value) have to be met to ensure a tasty and sweet fruit. The visual appearance of the fruit should have a ripe color and lack damages caused by rot or diseases. The peel should be free of cosmetic defects or bruises. In case of avocados, the oil content given by the dry matter value is the essential quality parameter. These parameters are typically measured on single fruit samples offline in a lab using e.g. destructive sample preparations and high measurement times in the range of a few minutes per fruit. In this project, a nondestructive inline measurement method was required reaching measurement speeds of 5000 mangos per hour on a conveyor belt with a velocity of 1 m/s.

2 Ripeness Sorting Using the Sherlock Food Analyzer

The basic instrumentation idea for the development of a ripeness analyser/sorter was the use of an optical system as an inline measurement device, the Sherlock Food Analyzer (see left image in 2.1). Additionally, in order to bring the fruits to the line of inspection, mechanical infeed machinery is required. Due to the necessity of double sided analysis for proper fruit quality inspection, a turning station was introduced in combination with a second Sherlock Food Analyzer. After analyzing the fruits on the conveyor belt, the Sherlock Food Analyzer acts as a control

device capable of piloting appropriate eject mechanisms for quality class sorting.

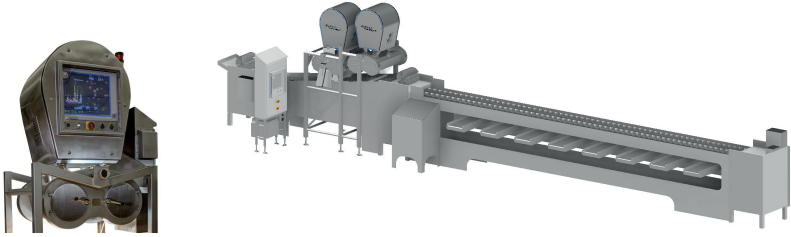


Figure 2.1: Ripening sorter using Sherlock Food Analyzer and eight ejection units.

2.1 Components of the Sherlock Food Analyzer

In order to understand the measurement concept and the idea of the instrumentation, a short introduction to the capabilities and key elements of the Sherlock Food Analyzer will be given in the following paragraphs.

Measurement Principles for Analyzing The device is based on the optical characterisation of food products by means of light spectroscopy over a broad range of wavelengths. On one hand the hyperspectral imaging camera measures light information in the near infrared range, therefore gathering data results on the chemical structure of the product. On the other hand, an RGB camera uses the range of visible light and evaluates visual appearance of the products. In both wavelength ranges line scan cameras are implemented as push broom applications. Regarding the data acquisition, a false color coding method (Insort branding: *CIT*, Chemical Imaging Technology) is used. This means that after a manual modelling step, by means of software, the spectra of the NIR camera are classified and color coded according to class characteristics. In order to have a common software interface, false color coding is also used for the RGB camera images, highlighting visual characteristics of the fruit. In case of the RGB camera, a qualitative

approach is used. Product defects are coded in desired colors based on categories: Good product (e.g. green) or product defects (e.g. red, violet). This coding is then used within the operating software in order to evaluate type, count and area of defects on each fruit. The hyperspectral camera has two modes available for analyzing spectral information. The already mentioned qualitative approach distinguishes between the spectra of the good product in contrast to product defects, which typically differ in chemical composition. Based on this model a false color picture is generated and evaluated by software inline and in real-time. An additional and very important method provided by the hyperspectral camera is the so-called quantitative approach. Based on the reflected light intensity in the selected wave length range (region of interest, ROI), a regression model is established which brings the reflected light intensity in correlation with the laboratory analysis of the chemical components' concentration (e.g. moisture, dry matter). Hence both camera types transfer false color pictures inline and in real-time to the evaluation software.

Optical and Mechanical Sensors The key element in this optical characterisation project is the hyperspectral camera. The near infrared camera *EVK Helios G2* [3] provides 320 pixel spatial resolution and 240 wavelengths per pixel in the range from 900-1700 nm. The line scan camera works in a push broom application and transfers the classified spectra, coded as false color RGB pictures, by means of Ethernet: GigE protocol. The camera needs a model upload upon initialization, together with basic parametrizations (e.g. frame rate). The modelling software *SQALAR* [4] enables the user to generate the model for the fruits and to update the camera.

The used RGB camera [5] provides 2048 pixel and also works in line scan mode. Utilizing a qualitative model, this camera determines fruit size and color defects. The qualitative classification is additionally optimized for black or dark spot area measurement and visible bruise detection. By means of the software *Falco* [6], a false color model is generated and executed on the analysis computer during runtime .

A weight measurement cell [7] is implemented as part of the mechanical infeed system. The load cell is placed after the optical inspection section and before the ejection system. The fruits, placed in cups,

are transported over the load cell and the derived online measurement value is transferred to the analyzing software. The load cell provides an analogue current interface evaluated by a PLC [8]. After calculation and filtering, the fruit weight reading is transferred to the PC by means of an Ethernet connection. The challenge of the weight measurement is the fact, that the fruits in the cup are moving over the weight cell without stopping. In order to get a correct reading (± 2 gram), an adaptive filtering algorithm is necessary to flatten the measurement peak occurring in the moment the cup touches the weight cell with a velocity of 1 m/s.

Another important sensor in the system is the rotary encoder placed on the conveyor chain. Based on this sensor and the appropriate parametrization, the distance between the line of inspection, the weight measurement and the ejectors is determined. The interface from this central position encoder is connected to the PLC. Based on the PLC-PC Ethernet communication the detection of objects (fruits) in the line of inspection and the control of weight measurement and ejector actuating is synchronized.

Illumination Regarding the two types of optical cameras used, two different kinds of illumination are necessary.

On one hand, there is the need of broad band illumination for the hyperspectral camera implemented by means of halogen lamps and a reflector encased into an air cooled steel housing. This illumination is focused onto the line of inspection in order to provide sufficient light for optical evaluation. The fruits reflect this light to the hyperspectral camera enabling spectroscopic evaluations.

On the other hand, additional light in the visible range is needed for the high resolution RGB cameras. A white LED is therefore used to enhance the amount of light in the line of inspection. In order to support the applied machine vision algorithm for the detection of fruits (background segmentation), the conveyor rollers are kept in a unique color (blue). By means of blue light LED illumination, the background of the rollers is filled up homogeneously (see 2.2).

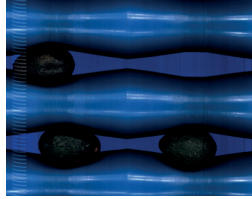


Figure 2.2: Blue conveyor rollers and the active blue LED background.

2.2 Mechanical Infeed for Optical Inspection and Ejector Units

The way of a fruit from the store to one of the eight quality category outputs is controlled by a complex conveyor system. At the entrance the fruit boxes are unloaded and the fruits are transported to an elevating conveyor in order to reach the height of the inspection line. Next, a short conveyor is used to separate the fruits in order to only carry two fruits at once onto the following rollers of the conveyor. The rollers section is used to transport the fruits to the line of inspection of the first Sherlock Food Analyzer. The fruits are detected by the camera system and their path is tracked and calculated in order to synchronize the following steps of the infeed process. The next step is the rotating part of the conveyor where the rollers are used to turn the fruits for the inspection of the rear side at the position of the second Sherlock Food Analyzer. After this inspection step, the fruits are transported to cups. By means of the rotary encoder, the position of the cups is calculated and the weight is measured on the fly in the moment the cup touches the integrated load cell.

Now the analysis PC has to evaluate the sensors and derive the necessary criteria for ejecting the fruit in one of the eight quality outputs based on the implemented rule engine. Again, the position of the electro-mechanical ejectors is derived from the rotary encoder and the rule engine's results. Boxes at all outputs are prepared for receiving the analyzed fruits.



Figure 2.3: Infeed to the Sherlock Food Analyzer with Conveyor System and Ejection Unit.

2.3 Object Analysis and Measurement Quantities

The analysis of fruit is conducted in the context of a so-called sorting program. A sorting program stores all parameters necessary for optical characterisation of the dedicated fruit. Not only different fruits are selectable but, in case of different optical characteristics, selections for different varieties of the fruit are available. In addition to that, the sorting criteria of the implemented rule engine are part of this program.

In terms of optical characterisation there is a distinction between the evaluation on the hyperspectral camera and the RGB image analysis. Nevertheless the results of both will be combined by software and treated as an abstract data structure called: object. The attributes of this object are the physical quantities measured. In the moment an object is generated, the position of the fruit is determined by the position of the rotary encoder. The offset due to calculation time is corrected and the values are then used to determine the position of the fruit on the conveyor. Regarding the measurement of damages and diseases, the amount of pixels and the size of areas in a certain false color coding combined with mechanical parameters (distance, conveyor speed) lead to classification into a certain category.

The false color coding for the RGB pictures is related to the visual optical appearance. Characteristics to be analyzed are: Several diseases, rot, scars, outer bruises, dark spots, color class, size. In a manual parametrization step, classes are generated from the RGB recordings for these categories and coded in different colors using the offline software (Falco). In contrast to this, classification of chemical character-

istics is conducted with the hyperspectral camera images. Quantities to be measured are: Sugar content mango (BRIX, precision $<1^\circ$ Brix), dry matter (% , precision $\pm 1\%$) an inner bruises (mm^2).

In order to reach the demanded quantitative measurement quality, a calibration process followed by a validation step is necessary to generate a linear regression curve. In a first step, a selection of fruits is analyzed in the line of inspection to derive their spectral information. Then, samples are taken from these fruits at distinct surface positions. After packaging and identifications, these samples are sent to the food lab to be analyzed in terms of sugar concentration and moisture. By means of the parametrization software, the concentration values of the fruits are correlated to the magnitude of the reflected spectra in the region of interest.

Hence, a linear regression model is established, ready to be sent to the hyperspectral camera for the following validation step. New fruits are analyzed with the generated model in order to check the quality of the model in terms of accuracy.

For geometrical measurement calibration (e.g. length, area) an object with known dimensions is analyzed at the line of inspection to adjust the size parameters. Typically, the width of the infeed is given by the conveyor dimensions, but the adjustable product speed has an big influence on the image of a line scan camera. Hence, geometric corrections have to be prepared for the particular sorting program.

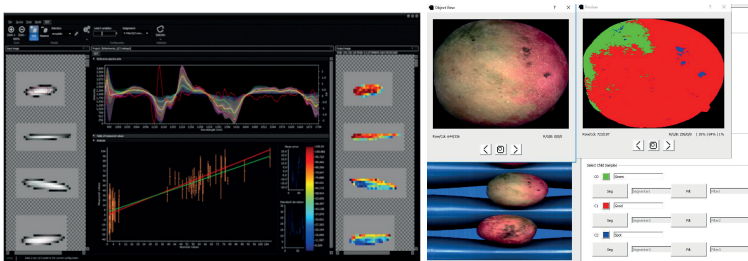


Figure 2.4: Software Tools: Hyperspectral modelling - Sqalar (EVK) and RGB color classification with Falco (KestrelEye).

2.4 Analysis Software

The used software has to fulfill several tasks to enable ripeness sorting. It starts with the data acquisition from two hyperspectral and two RGB cameras located in the Sherlock Food Analyzer. This is done via Ethernet connection and the GigE protocol. The false color pictures are transferred from the network card to the computer's memory. With the availability of the accessed pictures, the next data processing step takes place: background segmentation. Upon successful segmentation, the objects (fruits) are extracted from the conveyor background. As soon as an object is generated, the synchronisation with the conveyor for positioning is started. Now the geometrical and color-based attributes are calculated in order to obtain false color images. This process is repeated on the second Sherlock Food Analyzer after rotating the fruit. Additionally, special algorithms called classifiers are used for the detection of further object attributes (e.g. black spots). Last but not least, the weight of each fruit is derived from the load cell and added to the fruit attributes.

Now all the needed measurement values are available to make the quality decision. The decision making parameters and thresholds are configured by means of the graphical interface of the implemented rule engine. The operator is in charge of selecting the appropriate ejection position based on a defined set of rules. The applied rule can logically combine sugar concentration, moisture, bruises, weight and black spot areas. Based on the selected sorting program, the rule engine is configured to the corresponding parameters and rule settings. The measurements and the rule decisions of every fruit are stored on the machine in a file that is available to be sent to host computers in the plant.

3 Conclusions

The company Insort mastered the challenge and built a unique new sorting machine using their Sherlock Food Analyzer as the key element. Hyperspectral imaging technology in combination with the RGB cameras is able to sort out chemical parameters as well as color differences. The sorting is conducted with a higher precision and velocity than manually before. After the realization of this project, the customer

is able to successfully sort the fruits into different quality categories based on important parameters and supply high quality fresh fruit to the fast growing market. Implementation of the optical characterisation system was done for mango and avocado sorting. Ideas for the sorting of other fruits are being considered.

4 Acknowledgment

The development of the described ripening machine was only possible due to the confidence of the CEOs of Frutura (Mr. Manfred Hohensinner, Mrs. Katrin Hohensinner) and Insort (Mr. Matthias Jeindl) in their team and in the project. Great thanks to all the members of Frutura for the support during commissioning work and to the engineers at Insort who contributed with their knowledge in software, hardware and optical application to the success of this project. Many thanks to Mr. Philipp Staubmann who redacted this paper and transferred the text to LaTeX.

It should be mentioned that the work effort for the generation of this paper was co-funded by the European Union in the framework of Horizon 2020, Grant Agreement: 872770.

References

1. "Frutura, wir sind obst und gemüse." [Online]. Available: www.frutura.com
2. "insort, intelligent selection." [Online]. Available: www.insort.at
3. "Evk helios nir g2 320 class." [Online]. Available: <https://www.evk.biz/de/produkte/hyperspectral-imaging/evk-helios-nir-g2-320-class/>
4. "Evk sqalar, qualitative and quantitative software tool." [Online]. Available: <https://www.evk.biz/de/produkte/analysis-software-tool/evk-sqalar/>
5. "Teledyne dalsa linea." [Online]. Available: <https://www.teledynedalsa.com/en/products/imaging/cameras/linea/>
6. "Kestrel eye, falco." [Online]. Available: <http://kestrel-eye.com/products/>
7. "Hbm an hbk company." [Online]. Available: <https://www.hbm.com/de/>
8. "Siemens sps s7 1200." [Online]. Available: <https://new.siemens.com/global/de/produkte/automatisierung/systeme/industrie/sps/s7-1200.html>

A new approach for evaluation of meat freshness

Nensi Kasvand¹, Aleksandr Frorip¹, Artur Kuznetsov¹, Tõnu Püssa²,
Linda Rusalepp², Alar Sünter¹, and Dea Anton²

¹ Ldiamon AS, Tartu Science Park,
Teaduspargi 7, 51014, Tartu

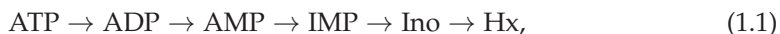
² Estonian University of Life Sciences, Institute of Veterinary Medicine and
Animal Sciences, Department of Food Hygiene
Kreutzwaldi 56/3, 51006 Tartu

Abstract Adenosine triphosphate (ATP) and inosine monophosphate (IMP) dominate in fresh meat at the stage of autolysis, whereas ATP exists only during very first hours after slaughter. Hence, ATP is a highly suitable marker of absolute freshness of rapidly frozen meat/fish. Fast Protein Metabolites Liquid Chromatography (FPMLC) method was used to evaluate freshness of minced pork during storage and compared to results of standard methods - Volatile Fat Acids (VFA) and total count of bacteria. Good agreement between new FPMLC method and standard methods was achieved.

Keywords Meat freshness index, inosine monophosphate, hypoxanthine, fast liquid chromatography

1 Introduction

Many meat and fish freshness/quality assessment indices are based on combination of concentrations of ATP breakdown products:



where ATP, ADP, AMP are adenosine tri-, di- and monophosphate respectively, IMP - inosine monophosphate or inosinic acid, Ino – inosine and Hx – hypoxanthine [1]. ATP and IMP dominate in fresh meat at

the stage of autolysis, whereas ATP exists only during very first hours after slaughter. Hence, ATP is a highly suitable marker of absolute freshness of rapidly frozen meat/fish. IMP is well-known endogenous taste enhancer (E630) and synergic component of umami taste [2]. Hx, on the contrary, lead to the bitter off-taste [3], which makes it a good sign of the onset of meat or fish spoilage. Increased concentration of Hx may also be a matter of health since hypoxanthine is a precursor of hyperuricemia and gout [4] or associated Lesh-Nyhan syndrome [5].

The first historical ATP-linked freshness index K was proposed by Saito et al. [6,7]:

$$K = \frac{[Ino] + [Hx]}{[ATP] + [ADP] + [AMP] + [IMP] + [Ino] + [Hx]} \quad (1.2)$$

Since ATP, ADP and AMP exist in remarkable concentration only during short period of first hours, Karube et al. proposed simplified index K_i for operation in the ordinary time scale of days [8]:

$$K_i = \frac{[Ino] + [Hx]}{[IMP] + [Ino] + [Hx]} \quad (1.3)$$

Both indices, K , and K_i , have been widely used in fish science [1,7], but considerably less for red meat [9], regardless of the fact that these ATP metabolites have been exploited for a long time as freshness markers [10–16]. One of the reasons can be that the indices change only in a limited range of values from 0 to 1 (or 0% - 100%) with no great or dramatic dynamics.

Furthermore, linking data between freshness indices and bacterial contamination of meat is deficient. Mostly researches on ATP metabolites and bacterial contamination have been done separately without any attempt to combining both into one more universal approach. Recently, a two-band freshness index (TBFI) was proposed to monitor bacterial contamination on chicken meat surface using hyperspectral imagery technique [17]. It is crucial to look for correlation of freshness indices and bacterial contamination to expand methodological and technical capabilities in establishment of freshness and early spoilage indications of foods.

The aim of this work was to study correlations of K_i indices and standardized meat evaluation methods (e.g., Volatile fatty acids VFA and bacterial contamination), validating the FPMLC method in process.

2 Materials and methods

Sampling and materials The study was performed in two sequential experiments. Pork sirloin, one-day-after-slaughter, was minced and samples were packed according to sampling plan in two different conditions: air and vacuum and were stored at 2 – 4 °C for 14 days. Samples were analyzed every other day according to the schedules, 8 times in total.

Volatile fatty acids (VFA), total count of aerobic microorganisms and *Pseudomonas* spp. were determined by the Estonian Veterinary and Food Laboratory (EVFL). The FPMLC device designed by Ldiamon AS has been used for determination of ATP metabolites [18,19]. Determination of ATP metabolites concentration was carried out by LC-DAD (MS) and determination of lipid oxidation level by the TBARS method in Estonian University of life Sciences.

pH determination Measured in homogenized mixtures of 1 g of sample and 9 mL of distilled water with Consort C833 digital pH-meter (Consort, Turnhout, Belgium) at room temperature. Calibration of pH meter was regularly checked.

Determination of ATP metabolites by FPMLC Mixture of 2 g of minced pork with 6 ml TRIS buffer (pH 8.0) in 15 ml test tube was shaken for 11 minutes in a rotator Biosan Multi RS60 (BioSan, Riga, Latvia) and filtered with syringe filter (Whatman, 0.45 µm.). 6 drops of extract was dropped into the PD-10 column of the FPMLC device.

Determination of ATP metabolites concentration by LC-DAD (MS) The liquid chromatographic analysis of meat extracts was carried out by a 1290 Infinity system (Agilent Technologies, Waldbronn, Germany) coupled to an Agilent 6450 Q-TOF mass spectrometer equipped with a Jetstream ESI source. Samples were subjected to a Zorbax 300SB-C18 column 2.1×150 mm; 5 µm, (Agilent Technologies) kept at 40 °C. For the separation of compounds, a gradient of 0.1% of formic acid in water (A) and 5% of water in acetonitrile (B) was used as follows: 0.0 min 1% B, 3.0 min 1% B, 3.01 min 99% B, 11.1 min 99% B, 11.01 min 1% B, regeneration time 8 min. The eluent flow rate was set to 0.3 mL/min

and the injection volume size was 2.5 μL . Mass-spectrometer was working in the negative ionization mode in the mass-to-charge ratio (m/z) range of 100–3200 Da. UV absorbance was measured at 250 nm. Data acquisition and initial data processing were carried out by MassHunter software (Agilent Technologies).

The identification of IMP, Ino and Hx in samples was confirmed by comparing MS/MS and UV spectra with those of analytical standards. IMP, Ino and Hx were quantified by UV absorption at 250 nm using external calibration curve method. Methanolic standard solutions with concentrations of 3.125, 6.25, 12.5, 25, 50 and 100 μM were prepared for all three analytical standards. Calibration curves were characterized by a high correlation coefficient ($R^2 = 1$).

Validation The chain of parallel measurements and comparison of VFA and FPMLC results were considered as validation procedure of FPMLC method. For this purpose, a “clone” samples were tested at the EVFL with the standard method VFA [20]. Official protocols from the EVFL have been compared with results of synchronized measurements carried out by University of Life Sciences and Ldiamon AS.

3 Results and discussion

FPMLC chromatograms of meat consist of two main parts: the sharp protein peak and broad post-protein band (Figure 3.1) that is formed by merging of individual peaks of the main nucleotide actors. During the storage of meat, the ATP is decomposed by the enzymes to the metabolites with smaller molecular weight, as a result, the retention time of the metabolites increases. The major operand of FPMLC analysis - Time, measured in seconds, is the time interval between retention time of the flat band of metabolites and sharp protein peak in the FPMLC chromatograms (Figure 3.1).

The changes in concentrations of ATP degradation products - IMP, Ino and Hx, that occurred in minced pork during air and vacuum storage for 14 days at 2 – 4 $^{\circ}\text{C}$ with correspondently calculated K_i indices are shown in Figure 3.2. In air storage decreasing in concentration of IMP during the meat storage took place more rapidly than in vacuum,

degression 3,14 and 1,42 times, respectively. Changes in K_i indices were steeper in air stored samples for first 7 days of meat storage.

Correlation of new FPMLC method and LC-DAD/MS was satisfying (Figure 3.3). Correlations for both FPMLC indices – Time and K-value with K_i values determined by LC-DAD/MS were linear, with values of R^2 respectively 0,83 and 0,86.

Compatibility of FPMC method and bacterial contamination was good for both storage conditions air and vacuum (Figure 3.4). Correlations of indices Time and K-value with total count of bacteria *Pseudomonas spp.* was linear with values of R^2 0.93-0.94 and 0.79-0.81 respectively for air and vacuum stored samples.

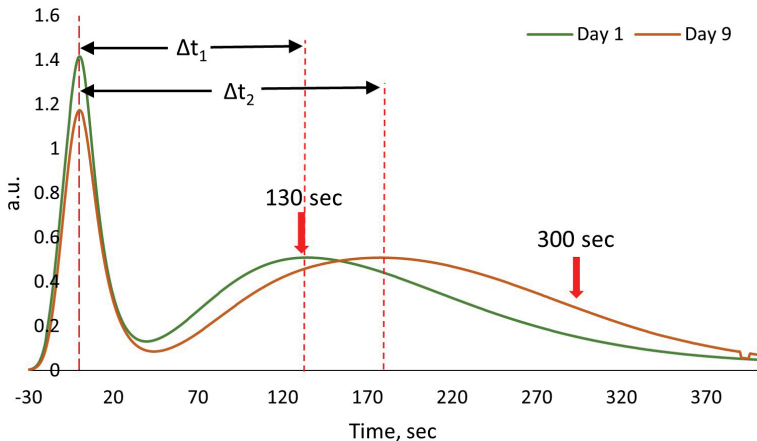


Figure 3.1: Comparison of two FPMLC chromatograms produced at days 2 and 9 during storage of a pork meat.

Validation Validation process of FPMLC method took place in strictly parallel mode of progression of the indices VFA and Time in cooperation with EVFL (Figure 3.5).

The simultaneous steep rises after the 9th (pH=5.6) or 11th (pH=5.3) days mark the end of changes in condition of meat caused by autolysis and onset of changes caused by bacterial activity. For more acidic pork (Figure 3.5b) the level of VFA in the flat interval of predominantly autolytic transformations is somewhat higher than for the meat of pH 5.6.

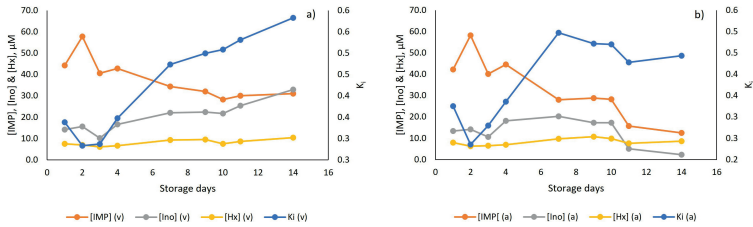


Figure 3.2: Changes in concentrations of ATP metabolites – IMP, Ino and Hx, and corresponding K_i values during the storage of minced pork in vacuum (a) and air (b) at 2 – 4 °C.

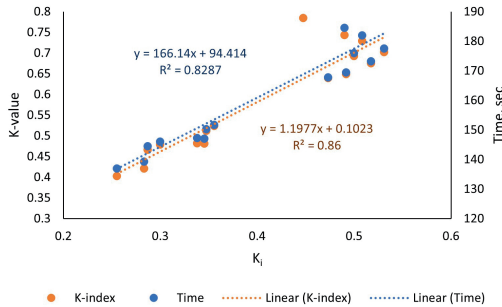


Figure 3.3: Correlations between K_i values determined by LC-DAD (MS) and main indices of FPMLC – Time and K-value for minced pork stored in vacuum and air at 2 – 4 °C.

Therefore, the maximal level of VFA indicated for the pork pH 5.6 on the day 14 is remarkably higher than for acidic one, i.e., 19.2 and 11.78 mg KOH/100g, respectively. In both cases, at the onset of the steep rise the meat had already specific heavy odor and slime.

Correlation coefficients for the paired curves are high: $r = 0.955$ (pH=5.6) and $r = 0.93$ (pH=5.3) sustaining credibility and reliability of the new FPMLC method.

4 Conclusions

Results of FPMLC method were in a good agreement with standardized methods of VFA and total bacterial count. Both indices of FPMLC

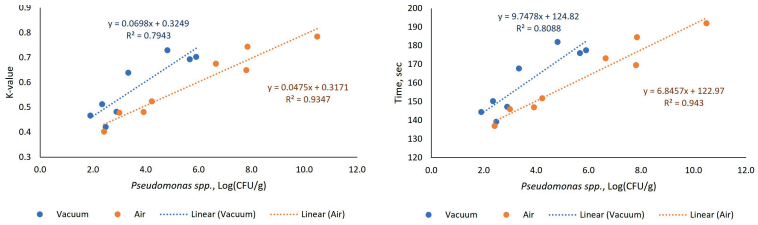


Figure 3.4: Correlations of FPMLC operands – K-value and Time with total count of bacteria *Pseudomonas spp.* (Log (CFU/g)) for minced pork stored in vacuum and air at 2 – 4 °C.

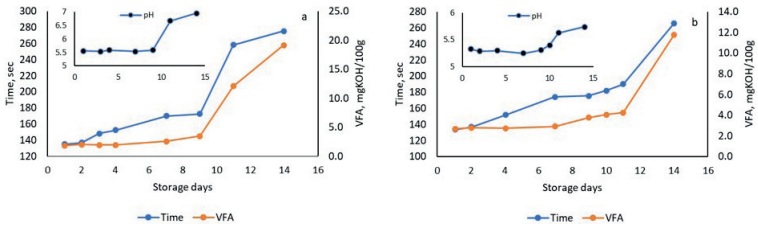


Figure 3.5: Changes in VFA and FPMLC indices Time for air stored minced pork at 2 – 4 °C with different initial pH levels - pH=5.6 (a) and pH=5.3 (b).

– Time and K-value were in a good correlation with bacterial growth of *Pseudomonas spp.*

FPMLC method has promising validity in relation to minced pork freshness control and needs to be studied further with other subjects of meat or fish. The FPMLC method can be applicable for the selection of the raw meat or fish of the highest quality or cold chain management for special control of frozen meat and fish in suspicious cases.

References

1. H. Hong, J. Regenstein, and Y. Luo, "The importance of atp-related compounds for the freshness and flavor of post-mortem fish and shellfish muscle: A review," *Critical Reviews in Food Science and Nutrition*, vol. 57, pp. 1787–1798, 2017.
2. S. Yamaguchi and K. Ninomiya, "What is umami?" *Food Reviews International*, vol. 14, pp. 123–138, 1998.

3. J. Luong, K. Male, C. Masson, and A. Nguyen, "Hypoxanthine ratio determination in fish extract using capillary electrophoresis and immobilized enzymes," *Journal of Food Science*, vol. 57, pp. 77–81, 1992.
4. R. Villegas, Y.-B. Xiang, T. Elasy, W. Xu, H. Cai, Q. Cai, M. Linton, S. Fazio, W. Zheng, and X.-O. Shu, "Purine-rich foods, protein intake, and the prevalence of hyperuricemia: The shanghai men's health study," *Nutrition, Metabolism and Cardiovascular Diseases*, vol. 22, pp. 409–416, 2011.
5. U. N. L. of Medicine. (2013) Lesch-nyhan syndrome. [Online]. Available: <https://ghr.nlm.nih.gov/condition/lesch-nyhan-syndrome>
6. T. Saito, K. Arai, and M. Matsuyoshi, "A new method for estimating the freshness of fish," *Bulletin Of Japanese Society Of Scientific Fisheries*, vol. 24, pp. 749–750, 1959.
7. T. Gill, "Objective analysis of seafood quality," *Food Reviews International*, vol. 6, pp. 681–714, 1990.
8. I. Karube, H. Matsuoka, S. Suzuki, E. Watanabe, and K. Toyama, "Determination of fish freshness with an enzyme sensor system," *Journal of Agricultural and Food Chemistry*, vol. 32, p. 314–319, 1984.
9. N. Batlle, M. Aristoy, and F. Toldrá, "Atp metabolites during aging of exudative and nonexudative pork meats," *Journal of Food Science*, vol. 66, pp. 68–71, 2001.
10. K. Honikel and C. A. Fischer, "Rapid method for the detection of pse and dfd porcine muscles," *Journal of Food Science*, vol. 42, pp. 1633–1636, 1977.
11. M. Flores, E. Armero, M.-C. Aristoy, and F. Toldra, "Sensory characteristics of cooked pork loin as affected by nucleotide content and post-mortem meat quality," *Meat Science*, vol. 51, pp. 53–59, 1999.
12. N. Batlle, M. Aristoy, and F. Toldrá, "Early post-mortem detection of exudative pork meat based on nucleotide content," *Journal of Food Science*, vol. 65, pp. 413–416, 2000.
13. M. Tikk, K. Tikk, M. Tørngren, L. Meinert, M. Aaslyng, A. Karlsson, and H. Andersen, "Development of inosine monophosphate and its degradation products during aging of pork of different qualities in relation to basic taste and retronasal flavor perception of the meat," *Journal of Agricultural and Food Chemistry*, vol. 54, p. 7769–7777, 2006.
14. N. Vijitrothai, S. Noidad, and R. Limsupavanich. (2018) Determination of adenosine triphosphate breakdown products in cooked pork loin by the isocratic hplc-uv method. [Online]. Available: <http://icomst-proceedings.helsinki.fi/papers/2018.10.38.pdf>

15. L. Mora, A. Hernández-Cázares, M.-C. Aristoy, and F. Toldrá, "Hydrophilic interaction chromatographic determination of adenosine triphosphate and its metabolites," *Food Chemistry*, vol. 123, pp. 1282–1288, 2010.
16. M. Dibirasulaev, G. Belozarov, L. Arkhipov, and L. Stoyanova, "Quick and simple spectrophotometric method of identification of the thermal state of meat on the basis of composition and content of free nucleotides," *Food and Nutrition Sciences*, vol. 9, pp. 572–583, 2018.
17. X. Ye, K. Iino, and S. Zhang, "Qmonitoring of bacterial contamination on chicken meat surface using a novel narrowband spectral index derived from hyperspectral imagery data," *Meat Science*, vol. 122, pp. 25–31, 2016.
18. A. Kuznetsov, A. Frorip, R. Korsakov, V. Korsakov, T. Püssa, A. Sünter, and N. Kasvand, "New freshness index, method and device to determine freshness status in fish and meat according to fao recommendations," 31st EFFoST Conference Food Science and Technology Challenges in the 21st Century Poster presentation, 2017.
19. T. Püssa, A. Frorip, A. Kuznetsov, A. Sünter, D. Anton, and P. Raudsepp. (Püssa) New freshness index, method and device to determine freshness status of meat and fish. [Online]. Available: <https://digicomst.ie/wp-content/uploads/2020/05/2018.02.06.pdf>
20. ProAgro. (1980) Meat. methods for chemical and microscopic analysis of freshness(rus). [Online]. Available: <http://www.proagro.com.ua/reference/vedua/uktzed/26.html>

Towards the universal assessment of dietary intake using spectral imaging solutions

Yannick Weesepeel¹, Freek Daniels², Martin Alewijn¹, Mireille Baart³,
Judith Müller-Maatsch¹, Görkem Simsek-Senel², Hajo Rijgersberg²,
Jan Top^{2,4}, and Edith Feskens³

¹ Wageningen Food Safety Research,
Wageningen, The Netherlands

² Wageningen Food and Biobased Research,
Wageningen, The Netherlands

³ Wageningen University and Research, Department of Human Nutrition and
Health,
Wageningen, The Netherlands

⁴ Vrije Universiteit Amsterdam - Faculty of Sciences, Computer Sciences,
Business, Web and Media,
Amsterdam, The Netherlands

Abstract Personal(ised) sensors or “wearables” could, in the future, be applied to provide personalized nutritional advice. A challenge here is to assess dietary intake. To date, this has mainly been done with questionnaires or interviews, for example, food frequency questionnaires or 24-hour recalls. However, these methods are prone to bias due to conscious or unconscious misreports, and more objective measurement methods are desirable. By applying and combining spectral imaging techniques like hyperspectral imaging (HSI) and RGB-depth (RGBD) imaging, information on the macro-composition, identity and quantity of food consumed can be obtained. In this work, we demonstrate that HSI was effective for estimation of the fat content and layer thickness of butter on slices of bread with root mean squared errors of predictions of 4.6 (fat w/w %) and 0.056 mm respectively. Identification and volume estimation of vegetables and preparation methods were successful with RGBD imaging. Using Convolutional Neural Networks, all samples were correctly identified. For volume estimations of vegetables, R-square scores between 0.80 – 0.96 were achieved.

Keywords Non-destructive sensing, measuring and detecting, healthy behavior, consumer science

1 Introduction

Dietary intake data is important in epidemiology, for dietary interventions and for providing dietary advice. Advances in nutritional science have facilitated the shift from general population-based dietary guidelines to more personalized dietary recommendations. To provide personalized advice, an appropriate method for measuring and assessing dietary intake is required. Dietary intake is traditionally assessed with questionnaires or interviews, for example, food frequency questionnaires or 24-hour recalls. However, these methods are prone to bias due to conscious or unconscious misreports. Therefore, more objective measurement methods are desirable. Moreover, the traditional methods are rather time-consuming for both researchers and consumers.

For this reason, we aim to develop an easy-to-use measurement setup that enables consumers to measure their dietary intake themselves, using state-of-the-art vibrational spectroscopy and imaging techniques. The analytical and machine learning challenges for this work comprise several issues:

1. Object-level recognition to identify the type(s) of food products present, to estimate the volume of those objects (before and after eating) and to obtain intake quantities from this information. For this purpose, Red Green and Blue-depth imaging (RGBD) was explored (this work).
2. Molecular-level recognition for the determination of the macro and micro composition and, to a lesser extent, identification of the product. Here a snapshot hyperspectral imaging (HSI) system in the short-wavelength infrared (SWIR) was used (this work).
3. Data fusion methods to connect the data of both imaging sensors to improve decision accuracies and versatility of the set-up.
4. Application of a functional metadata decision layer to deal for example with the within-food heterogeneity, the enormous diversity of products and dishes available [1].

In the current study, a first demonstrator is developed at technology readiness level 4. This means that the demonstrator can be operated in a controlled environment by trained persons. We demonstrate in this study the individual imaging techniques applied to the intake of butter (HSI), and vegetables (RGBD), and their respective data analysis procedures.

2 Materials and Methods

2.1 Materials

Fresh (1 – 2 days after baking) sliced white and wholegrain bread, crispbread, 44 commercially available types of butter and margarine, fresh endive (*Cichorium endivia*), fresh carrots (*Daucus carota* subsp. *sativus*), fresh spinach (*Spinacia oleracea*), frozen spinach and spinach with added cream of a single brand were purchased from a local supermarket in Wageningen, The Netherlands. All chemicals used for macro composition reference analysis were in purity grades following the specifications in the ISO standards used (Section 2.3).

2.2 Sample preparation

To test the ability of dietary intake assessment, the sensing principles were firstly assessed separately in two food cases: (i) (low-fat) margarine and butter applied to a sandwich using HSI, and (ii) raw and prepared vegetables using RGBD imaging. For the first case, standardized amounts (3, 6, 10, 15 and 30 g) of 28 butter types with varying fat concentrations (29.8% - 84.3% w/w) were applied on three different types of sandwiches with standardized sizes (wholegrain and white 10 x 10 cm and crispbread 6.0 x 8.1 cm). A randomized experimental design was made in combining the type of bread, margarine/butter type and layer thickness of the butter/margarine, resulting in 84 samples. As an alternative experiment, butter was applied (16 butter types, fat concentration 28.0% - 82.5% w/w, applied on white and wholegrain bread) using a custom-made triangular applier, resulting in a continuous layer thickness of 10 mm – <0.5 mm (Figure 2.1 A-B).

For the second food case, fresh endive, fresh carrots and fresh spinach were investigated in raw form and after cooking. In addi-

tion to the fresh spinach samples, frozen spinach and frozen spinach with added cream were added to the sample set resulting in 7 different sample types. Fresh samples were cleaned and washed with cold tap water before spectral acquisition. Cooking was performed using a steam oven and specific cooking times for each product. Frozen and frozen creamed spinach were defrosted and heated according to the producers' instruction. All samples were brought to room temperature before spectral acquisition. Sample plates were created by adding a random amount of mass from a product-specific range (carrots 20 - 80 g, endive raw 5 - 15 g, endive cooked 15 - 50 g and spinach 20 - 60 g) to the plate. After every five samples, the plate was emptied and cleaned. Product samples were put on the plate off centre and placed under the RGBD camera such that the centre of mass was roughly at a randomly selected rotation. In total, 585 samples plates were prepared, by incrementally adding vegetables to the same plate (i.e. 5 times for each plate). A single RGB and depth image was captured for each sample (Figure 2.1 C-D).

2.3 Determination of reference values of butter and vegetables

The total moisture content of the butter samples was determined by NEN-EN-ISO 3727-1:2001 / IDF 80-1:2001, the total fat content by NEN-EN-ISO 17189 : 2003 / IDF 194 : 2003. For the vegetables, the sample mass was determined using a standard-issue analytical balance.

2.4 Spectral image setup and acquisition

The HSI and RGBD image equipment was mounted in a customized build setup with an approximate distance to the sample of 30 and 40 cm respectively (Figure 2.1 E-F). A standard-issue white glazed ceramic plate (diameter 20 cm) was used for sample presentation. The HSI camera was mounted dead-centre above the sample, whereas the RGBD camera was mounted on a 45-degree angle. For the HSI system, four standard issue halogen lights were installed at each corner around the camera. Lights were directed dead-centre downwards to avoid direct illumination of the reflecting ceramic plate and surfaces of the samples presented. An Effilux bar light (ELB-400SW, Effilux, Les Ulis, France) with a diffuse window was installed underneath the RGBD camera.

An IMEC SWIR Snapscan hyperspectral imaging system (Interuniversity Microelectronics Centre, Leuven, Belgium) with a spectral range of 1117 – 1670 nm equipped with an Optec 16mm F1.7 SWIR lens (Optec S.p.A., Parabiago, Italy). The system was controlled by the IMEC Snapscan software (version 1.3.0.8, IMEC). Before data acquisition, the camera was calibrated using a 95% reflection calibration white standard (WS) tile sized 200 x 200 mm, using an integration time of 2.5 ms. Raw sample spectra were corrected according to **Equation 2.1** to generate corrected diffuse reflection spectra used in subsequent calculations for the 640 x 512 x 108 ($x \times y \times \lambda$) sized hypercubes.

$$Corr_{(x,y,\lambda)} = (Raw_{(x,y,\lambda)} - Dark_{(x,y,\lambda)}) / (WS_{(x,y,\lambda)} - DarkWS_{(x,y,\lambda)}) \quad (2.1)$$

RGBD images were acquired using an Intel RealSense D415 depth camera (Intel Corporation, Santa Clara, CA, USA) controlled by a custom build GUI. Images were captured in a semi-controlled environment, with light blocking on the left, right and backside of the scene using mat dark plastic plates. Furthermore, a dark A2-sized paper was used as a background on which plates were placed. The exposure time was set to 70 ms to prevent saturation of the image. Intrinsic calibration parameters available on the camera were stored alongside the color and depth image to be able to construct a point cloud of the depth image.

2.5 Data analysis

Multivariate statistics of HSI data

Processing of the HSI images was performed using the hyperSpec package [2] in R 3.6.1 [3]. The corrected images contained hyper-responsive (dead) pixels and non-sample regions, both of which needed to be excluded from further processing (Figure 2.1 B). Dead pixels contained reflection values $> 1e6$ and were easily recognized and removed. To discriminate between the sample (region of interest, ROI) and the background (empty plate, table and showy edges were present), a simple spectral based principal component analysis (PCA) approach was performed, in which scores on Principal Component (PC) 2 and a threshold of 0 determined if any pixel belonged to the

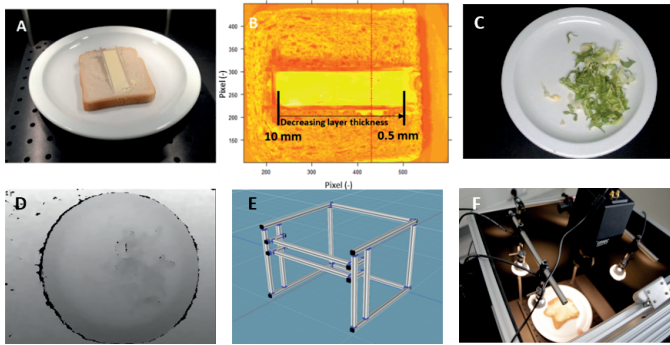


Figure 2.1: (A) Triangular mold for accurate layer thickness estimation of butter; (B) Spectral Image Heat plot of the butter applicator on a wholegrain sandwich; (C) RGB image of raw endive; (D) Depth image of (C) visualized using a gray color map; (E) Frame design of food intake setup; (F) Experimental setup with Hyperspectral SWIR camera (center) and RGBD camera (right bottom).

sample or not. As PCA's signs are non-determined, the portion near the centre of the image – presumed to be sample and not background – determined whether positive or negative PC2 scores contained the ROI. No data preprocessing or further thresholding or area filling was needed nor applied.

To estimate the butter and bread types, and mainly to characterize the butter (fat content, type, fat (un)saturation) and quantity (average layer thickness) of the butter on the bread a straightforward approach was chosen. a more straightforward approach was chosen. First, ROI-spectra were corrected for scattering effects (standard normal variate, SNV). Then the intensity of the signal at 1221 nm was found to correlate well with the amount of butter (fat) [4]. Individual spectra from the ROI pixels were sorted based on the value in this number, and the top and bottom 5% of the spectra were averaged to represent the butter and bread-spectra, respectively. These averaged spectra were input for further (multivariate) classification and regression. Although machine learning approaches per pixel (or pixel cluster) of spectral cube data are a normal approach, this method was not necessary for these relatively simple samples. Bread type (3 types) were classified using the

“bottom 5%” spectra for each of the samples using Soft independent modelling by class analogy (SIMCA), employing cross-validation (25x bootstrapped training sets). Butterfat content (Section 2.3) and butter layer thickness on the bread (5 discrete values, Section 2.2) were fitted using multivariate regression (PLS) using leave-10-out cross-validation. To assess the effect of the interaction of the butter layer thickness on bread on the observed spectra, a custom plexiglass shape was built to help apply butter on bread in known thicknesses. This device was 10 cm long and approximately 2 cm wide, and had a height of 10 mm on one end, linearly decreasing to 0 mm on the other end. When placed on a sandwich, this shape can be filled with butter and smoothed with a knife, after which an HSI will yield spectra with known butter thicknesses on bread -- only marginally affected locally by the foamy structure of the bread, slightly increasing the average thickness of the butter layer.

Analysis of RGBD images

Colour and depth information was used to identify and estimate the mass of vegetables on the plate. Two separate Convolutional Neural Networks (CNNs) were trained on colour images; ResNet50 [5] to classify the vegetable present in the scene and U-Net [6] to semantically segment the scene into pixels belonging to vegetables and the background. The segmentation was used to extract the relevant information from the depth images which were subsequently converted into a point cloud using available intrinsic calibration parameters. From the point cloud the volume was estimated. Finally, linear regression models were constructed for each identified product on the volume to obtain the mass. ResNet was the first network to introduce so called skip connections. These connections allow a layer to learn the identity operator, such that layers can be skipped if they are considered superfluous. U-net is a fully CNN using up sampling layers to obtain the same resolution in the output as layers as the input image. Both networks were trained using PyTorch (version 1.2, www.pytorch.org). The initial learning rate for U-net was upto 0.01, and for ResNet50 to 0.001 as pretrained weights were used for the latter. The learning rate was reduced by factor 10 when no improvement was obtained on the validation set after 200 epoch. Early stopping was applied after 400

epoch of no improvements. Samples were partitioned over a training, validation and test set with each respectively 355, 114 and 116 samples.

3 Results and Discussion

3.1 Bread and butter case using HSI

ROIs were set prior to classification of bread types, fat content estimation and layer thickness determination (Figure 3.1). Bread type classification was performed using SIMCA on the averaged spectra assigned to bread (Table 4.1). While it is clear that the spectra contain sufficient information to reliably distinguish between bread and crispbread underneath the butter layer (regardless of the amount and type of butter applied), the discrimination between white and wholegrain bread was not so clear. Apart from the visual color, their composition (moisture, protein, carbohydrates, fat) is not very different, and the difference in fibers is not strong enough in this part of the near-infrared (NIR) spectrum to make this discrimination reliably in this setup [7].

Table 4.1: Confusion matrix for the classification of bread type.

		Actual		
		White bread	Wholegrain bread	Crispbread
Predicted	White bread	29	17	0
	Wholegrain bread	0	11	0
	Crispbread	0	0	27

The spectra were also used to estimate the fat contents and the (average) thickness of the butter layer applied on the bread. Figure 3.1 gives the predictive capability of the multivariate method applied for both provisions. Note that the fat content of the butter was estimated from all layer thicknesses present in the dataset, and the layer thicknesses were estimated regardless of the butter type and fat content, and these parameters were estimated regardless of the type of bread under the layer. As root mean squared error of predictions (RMSEPs) values of 4.6 (fat w/w %) and 0.056 (mm) were found, which can be considered

quite acceptable for the intended application. Of course, models perform slightly better when separate models are made by determining the bread type and layer thickness first before estimating fat content, or bread type and fat content before estimating layer thickness (data not shown), but these improvements require more data to be calibrated and tested more reliably.

For further development of this model, both in terms of robustness and possibly to reduce the calculation cost (currently approximately 5 seconds per sample on a 2.7 GHz 6-core laptop), it is desirable to understand the interaction that two foodstuffs have on the NIR spectrum as obtained by the hyperspectral camera. Figure 3.1 shows the observed spectra as a function of butter layer thickness as obtained using the butter applicator device. Each line is the average of 20 spectra (± 0.4 mm) around the thickness indicated. Spectra are corrected for scatter-effects (SNV) and transformed using (simplified) Kubelka-Munk (K/S) absorption-over-scattering coefficient-values [8]. It is clear that there are trends in the spectra according to the layer thickness, but there does not seem to be a straightforward solution which explains the trend from 0 to 10 mm of butter on the bread for all wavelengths. The same holds for unprocessed and only scaled data. This implies that the desired information is present in the spectra, but the described multivariate statistics and due calibration are required to extract this information.

3.2 Vegetable case using RGBD imaging

To calculate the volume of vegetables, firstly, PCA was performed on point clouds centered around the origin (Figure 3.2). The vectors found using PCA represent the length, width and height of the three-dimensional objects. Next, points were rotated such that the third principal component was aligned with the vector $(0,0,-1)$. After rotation, the surfaces of the scenes were parallel to the plane $z = 0$. To compensate for noise in the depth image, z -values of the points were subtracted such that the 98th percentile was at $z = 0$. Finally, point clouds were subsampled in a uniform voxel grid of fixed dimensions and estimates of volumes were obtained by summing the z -values.

The CNNs were trained on the training set, early stopping was applied using the validation set and the results are reported on the test

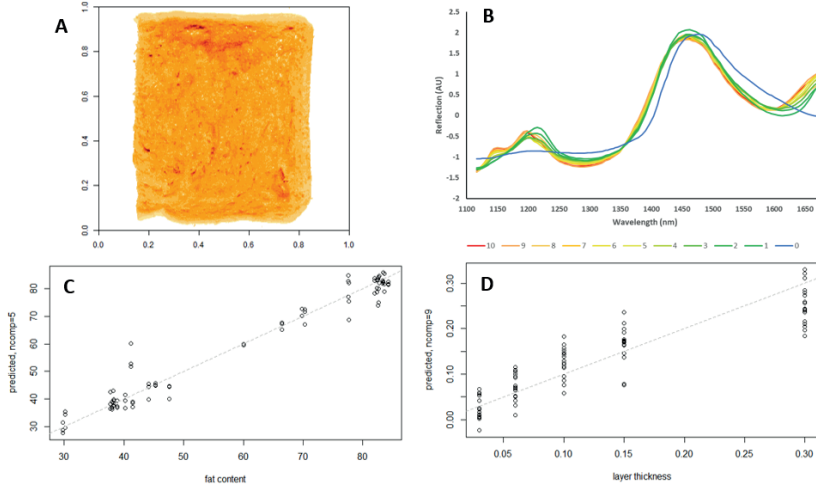


Figure 3.1: (A) A result of ROI detection of the 10 x 10 cm of bread-and-butter detection of HSI data in a false-color image of first PC scores; (B) Average NIR spectrum as a function of the butter layer thickness (0 – 10 mm); (C) determination of the total fat content of butter independent of the bread type and layer thickness; (D) determination of the layer thickness of the butter independent of the bread and butter type.

set. The classification of the selected vegetables and preparations was successful with zero misclassification as shown in the test set. Indeed, there seems to be sufficient variation present in the RGB data in either color or texture for successful identification. When endive is cooked the product darkens whereas the carrots lose the shiny smooth structure after having been prepared. As for the three different types of spinach, these products themselves are visually different. Estimates of the mass were obtained using product-specific regression models on the calculated volumes. Reasonable results were obtained ranging from an R-squared score of 0.80 to 0.96 (Table 4.2). Although the volume estimation method is in principle generic, as it does not assume any shape or form of the product, to get an accurate mass estimation individual models are needed. A single regression model could be constructed based on the estimated volumes when the (average product) densities are known.

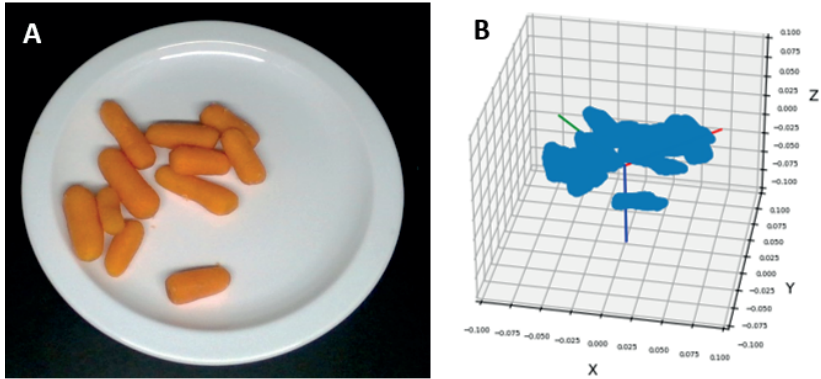


Figure 3.2: (A) Colour image of a scene; (B) Reconstructed point cloud of the segmented vegetables with the first three principal components, corresponding the length (red), width (green), and height or depth (red) axis of the vegetables.

4 Conclusions and outlook

Both sensing systems performed well in the identification and quantification of food products, and the estimation of nutrient composition. To expand the versatility of the setup, a data-fusion experiment will be performed. In this experiment, sandwiches containing multiple elements will be used, such as tomatoes, avocados, nuts/seeds, sprouting vegetables and butter. Also, the application of a functional meta-data layer for improving decision making and reducing the number of reference samples will be worked out. In 2021, a human study will be performed to test and validate the setup. Finally, the ultimate goal is to integrate the used analysis techniques into smaller devices such as wearables or smartphones that can be used by consumers themselves. Data obtained by using these wearables can then be combined with data from other sources such as consumer profiles to assess dietary intake as accurately as possible. Missing data can be supplemented based on user profiles, for example, if a particular food is not optimally oriented concerning the measuring equipment or if the food cannot be adequately determined through detection techniques at all. This will be of great value for providing personalized dietary advice that is op-

timally tailored at the individual consumer, eventually contributing to better health.

Table 4.2: R-squared scores for mass estimation of individual vegetables by RGBD imaging.

Food product	R-squared
Spinach creamed	0.96
Spinach frozen	0.93
Spinach fresh	0.88
Endive raw	0.87
Endive cooked	0.80
Carrots raw	0.91
Carrots cooked	0.83

5 Acknowledgements

This research is part of the Dutch Research Council (NWO) program "Dutch National Research Agenda: Start Impulse Measuring and Detecting of Healthy Behavior" under grant number 400.17.604. www.Metenendeteteren.nl

References

1. J. Marín, A. Biswas, F. Ofli, N. Hynes, A. Salvador, Y. Aytar, I. Weber, and A. Torralba, "Recipe1m+: A dataset for learning cross-modal embeddings for cooking recipes and food images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 187–203, 2021.
2. C. Beleites and V. Sergo, *hyperSpec: a package to handle hyperspectral data sets in R*, 2018, r package version 0.99-20180627. [Online]. Available: <http://hyperspec.r-forge.r-project.org>
3. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019. [Online]. Available: <https://www.R-project.org/>

4. F. R. van de Voort, J. Sedman, and T. Russin, "Lipid analysis by vibrational spectroscopy," *European Journal of Lipid Science and Technology*, vol. 103, no. 12, pp. 815–826, 2001. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/1438-9312%28200112%29103%3A12%3C815%3A%3AAID-EJLT1111815%3E3.0.CO%3B2-P>
5. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv:1512.03385 [cs]*, Dec 2015, arXiv: 1512.03385. [Online]. Available: <http://arxiv.org/abs/1512.03385>
6. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *arXiv:1505.04597 [cs]*, May 2015, arXiv: 1505.04597. [Online]. Available: <http://arxiv.org/abs/1505.04597>
7. T. Ringsted, H. W. Siesler, and S. B. Engelsen, "Monitoring the staling of wheat bread using 2d mir-nir correlation spectroscopy," *Journal of Cereal Science*, vol. 75, pp. 92 – 99, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0733521016303812>
8. F. Boroumand, J. E. Moser, and H. V. D. Bergh, "Quantitative diffuse reflectance and transmittance infrared spectroscopy of nondiluted powders," *Applied Spectroscopy*, vol. 46, no. 12, pp. 1874–1886, 1992. [Online]. Available: <https://doi.org/10.1366/0003702924123502>

Detection of pyrrolizidine alkaloid containing herbs using hyperspectral imaging in the short-wave infrared

Julius Krause¹, Nanina Tron², Georg Maier¹, Andrea Krähmer², Robin Gruna¹, Thomas Längle¹, and Jürgen Beyerer^{1,3}

¹ Fraunhofer IOSB, Karlsruhe, Institute of Optronics, System Technologies and Image Exploitation, Visual Inspection Systems, Fraunhoferstr. 1, 76131 Karlsruhe, Germany

² Julius Kühn-Institut (JKI), Federal Research Centre of Cultivated Plants, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, Königin-Luise-Str. 19, 14195 Berlin, Germany

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

Abstract Plants containing pyrrolizidine alkaloids (PA) are unwanted contaminants in consumer products such as herbal tea due to their toxicity to humans. The detection of these plants or their components using hyperspectral imaging was investigated, with focus on application in sensor-based sorting. For this, 431 hyperspectral images of leaves from three common herbs (peppermint, lemon balm, stinging nettle) and the poisonous common groundsel were acquired. By using a convolutional neural network, a mean F_1 score of 0.89 was obtained for the classification of all four plant products based on the individual spectra. To validate the neural network, significant wavelengths were determined and visualized in an attribution map.

Keywords Hyperspectral Imaging, convolutional neural network, Pyrrolizidine alkaloids, sensor-based sorting

1 Introduction

Pyrrolizidine alkaloids (PA) are secondary plant substances that are toxic to the genome and the liver. They can pose a health hazard, as they are often found as a contaminant in food or medicinal plant products. Just a few plants per hectare e.i. of *Senecio* are sufficient to contaminate the product non marketable. Hence, regular field monitoring and removal of appropriate weeds is necessary. This is a considerable personnel effort that can hardly be afforded economically by the growers. Additionally, for correct identification of toxic herbs and to prevent contamination of the products, the growers need well trained personal. Therefore, methods to identify and remove toxic contaminants after harvest are highly demanded.

Sensor-based sorting is a machine vision application that has found industrial application in various fields. An accept-or-reject task is executed by deflecting single particles from a material stream. The main fields of application of sensor-based sorting are recycling, e.g., removing materials from glass shard streams harmful to the melting process such as stones and ceramic glass [1]. Another field of application is the processing of industrial minerals, mainly to remove unwanted gangue from ore, e.g., copper-gold ore [2]. For ensuring product safety for foodstuff and agricultural products, these methods are used for the detection and removal of fungus-infected wheat kernels [3]. Hence, sensor-based sorting of crops also represents an opportunity to safe the harvest potentially contaminated with PA.

In this paper, three of the economically most important herb cultures in Germany (peppermint, lemon balm, stinging nettle) and the most common contaminant of these cultures (common groundsel) were characterized by near infrared spectroscopy. Optical spectroscopy in near and short-wave infrared is particularly suitable for the detection and differentiation of organic products. Using hyperspectral camera systems, material streams can be monitored and foreign substances can be separated after detection by implementing the techniques in a sensor-based sorting system. Targeting this application, hyperspectral data was acquired for the mentioned herbs and their frequent contamination in an experimental sensor-based sorting system.

2 Material and methods

This section describes the experimental sensor-based sorting system that was used to acquire the data. In addition, insights into the resulting data set are given. Finally, the novel data analysis method for the task at hand is described.

2.1 Experimental sensor-based sorting platform with hyperspectral camera system

In the course of this study, data for different plants are acquired using an experimental sorting platform that is equipped with an hyperspectral camera system, see Fig. 2.1. A thorough description of the experimental platform is provided in [4]. The spectral sensitivity of the InGaAs sensor of the line-scanning hyperspectral camera lies in the range of approximately 1200 to 2200 nm that is sampled into 256 spectral bands and provides 320 pixels locally at a maximal temporal resolution of approximately 300 hz. The implemented illumination consists of two arrays of halogen spotlights, which are locally targeted at the scan-line of the camera. Transportation of the material is realized by means of a conveyor belt that runs at 1.1 ms^{-1} . The material is observed directly after being discharged from the belt, i. e., during a free flight phase. An array of pneumatic nozzles is located behind this scan-line and serves the purpose of material separation. It consists of 16 nozzles that can be triggered individually, each of which covering a width of 10 mm.

2.2 Dataset

The plant material for the measurements was cultivated in beds at the JKI Berlin. The medicinal cultures were harvested twice, each time at the beginning of flowering, as the concentration of essential oils is highest then, using a cutting height of approx. 15 cm. For lemon balm (*Melissa officinalis*) and peppermint (*Mentha x piperita* 'Multimentha') 24 each and for stinging nettle (*Urtica dioica*) 12 and for groundsel (*Senecio vulgaris*) 10 single plants were harvested.

For training, the data set was split by random sample selection. A ratio of 75% was assigned as training and of 25% as validation data.



Figure 2.1: Photo of the experimental sensor-based sorting platform, equipped with the hyperspectral camera system.

Table 5.1: Description of the resulting dataset.

name	trivialis		samples spectra
<i>Senecio vulgaris</i>	groundsel	84	44683
<i>Urtica dioica</i>	stinging nettle	98	34447
<i>Melissa officinalis</i>	lemon balm	141	17747
<i>Mentha piperita</i>	peppermint	108	15970

The entire data set can be described with a matrix $\mathbf{X} := \{\vec{x}_i\}_{i=1..N}$ of N spectra and a matrix $\mathbf{Y} := \{\vec{y}_i\}_{i=1..N}$ of N labels. Each spectrum $\vec{x}_i \in \mathbb{R}^Q$ contains the reflectance of $Q = 256$ spectral bands. The labels $\vec{y}_i \in \mathbb{R}^C$ are coded as one-hot vectors with $C = 4$ classes, which correspond to the plant types. An exemplary visualization of the data is shown in Fig. 2.2.

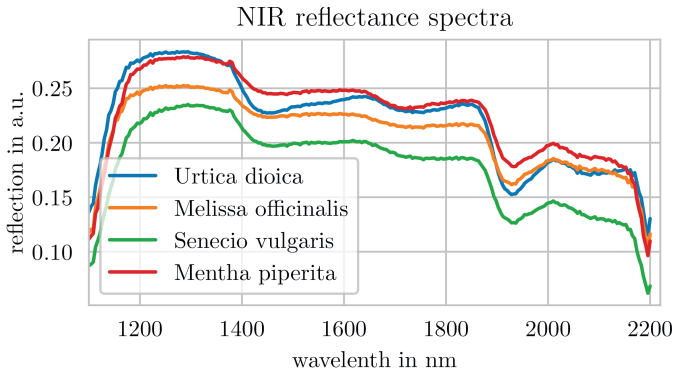


Figure 2.2: Shown are the mean spectra of all four plant species. Through the intensity differences at certain wavelengths, a classification can be done.

2.3 Data analysis using the convolutional neural network *AnniNet*

A new neural network architecture, *AnniNet*, has been developed at Fraunhofer IOSB. This architecture is particularly suitable for the analysis of near-infrared spectra. *AnniNet* consists of three components, namely an encoder network for feature extraction, a decoder network to improve feature extraction and a classification network that estimates the class membership of a spectrum. All three networks are trained in parallel through multi-task learning. The feature extraction structure does not require any type of spectral data pre-processing. The individual parts of *AnniNet* and multi-task learning are presented in the following.

Encoder network

The encoder layer is used for feature extraction. For this purpose, a layer with different one-dimensional convolution kernels is trained and applied to the spectrum. This layer can be compared with wavelet-based methods that are already successfully used in chemometrics [5]. It has been shown that wavelets are successful in reducing noise [6] and suppressing background effects [7]. This enabled improved results [8], even the transfer of chemometric models was improved [9]. The features determined by means of the convolutional layer are sent to

a pooling layer. This layer has a very practical effect besides reducing the parameters for the following layers. The feature extraction becomes more robust against shifts in the wavelength. This is especially a disturbing effect known as smiley keystone distortion in hyperspectral cameras.

Decoder network

End-to-end learning in artificial neural networks with high dimensional inputs such as hyperspectral measurements requires a large amount of labeled training samples due to a high count of weights which are iteratively adjusted. Autoencoders are used to learn the representation of a dataset in an unsupervised manner. Adding a decoder network creates such an autoencoder within *AnniNet* and improves training results without the need for additional labels.

Classification network

To evaluate the overlapping information from the individual absorption bands, fully connected (dense) layers were used. In correspondence with the exponential behavior of absorption processes (Beer-Lambert law), the Scaled Exponential Linear Units (SELU) activation function

$$f(x) = \begin{cases} \lambda\alpha(e^x - 1), & \text{if } x < 0 \\ \lambda x, & \text{otherwise} \end{cases} \quad (2.1)$$

was chosen for the first dense layer. Finally, a fully connected layer with the softmax activation function is applied. The output of the classification network is a normalized probability density function of the estimated class memberships.

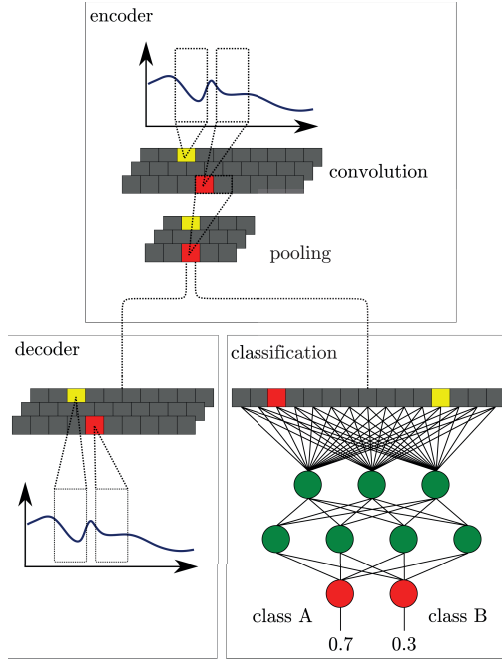


Figure 2.3: *AnniNet* consists of three components. First, an encoder network extracts spectral features using a convolutional layer. During training, a decoder network is used as an autoencoder to improve feature extraction. The classification network analyses the non-linear and overlapping spectral features.

Multi-task-learning

Using multi-task-learning, all four classes and the reconstruction of the autoencoder were trained in parallel. The categorical cross entropy

$$\text{CCE} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{i,j} \log \hat{y}_{i,j} \quad (2.2)$$

was used as a loss function to optimize the classification results for all classes C . The reconstruction of the spectral data using the autoencoder

was optimized by minimizing the Root Mean Square Error

$$\text{RMSE} = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{\sum_{j=1}^Q (\hat{x}_{i,j} - x_{i,j})^2}{n}} \quad (2.3)$$

for the autoencoder, which improves the feature extraction.

3 Results and discussion

The classification was carried out using *AnniNet* without any spectral pre-processing. To validate the results in the following, only spectra from samples that were not included in the training were used. To check the plausibility of the classification results, a sensitivity analysis was performed and compared with the spectral data.

3.1 Classification results

The results of the classification are shown in the confusion matrix in Fig. 3.1. For all classes, the classification achieves a rather high accuracy. The application of a majority vote for the respective samples has led to a completely correct mapping in the validation data set. The quality of the classification results can be quantified by the so-called F_1 score

$$F_1 = \frac{\text{tp}}{\text{tp} + 0.5(\text{fp} + \text{tp})}, \quad (3.1)$$

which is the harmonic mean of precision and recall. The evaluation calculation is based on the false (fp) and true (tp) positive classification results. For the present four-class classification, the F_1 score was determined for the individual classes and averaged. The score determined in this way is $F_1 = 0.89$.

3.2 Spectral attribution map

Neural networks cannot be interpreted by humans due to their high number of parameters and high-dimensional transformations. However, when evaluating hyperspectral data, it is interesting to see which

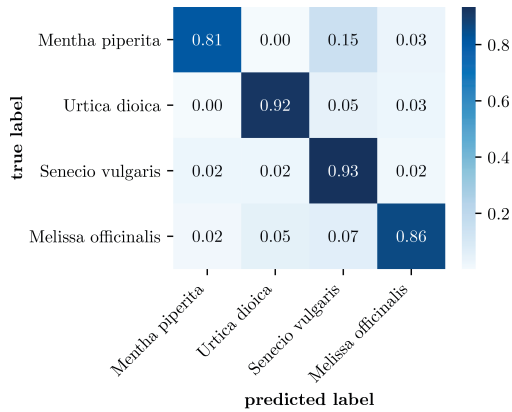


Figure 3.1: The normalised confusions matrix shows the probabilities of the estimated class memberships in the individual fields. The estimation of the class membership by single spectra is mostly correct. It is important here that by majority vote of all spectra of a sample, all test samples could be correctly classified.

absorption bands have an influence on the classification result. One possibility for investigating neural networks is the creation of a so-called attribution map. For this purpose, individual areas of the spectrum are successively masked and then the classification quality is evaluated. For the classification task at hand, such attribution maps were determined and are visualized in Fig. 4.1. As can be seen, the classification result is mainly depend on data in the range between 1400 nm and 1900 nm. This range is between the absorption's by water and can therefore be considered robust. In addition, areas are selected or weighted differently for the different class memberships, which is another indicator of robust classification.

4 Summary

Leaves from four different plants were recorded with a hyperspectral camera built in a sensor-based sorting system. The data were used without spectral pre-processing to train a neural network designed for this classification task. The trained network was able to predict individual spectra of the validation data of all four classes at a high accuracy. Us-

ing a majority vote per sample, all validation samples were correctly classified. An attribution map was used to investigate which spectral ranges dominate the classification decision. The result shows different spectral ranges for the individual classes, apart from absorption by water.

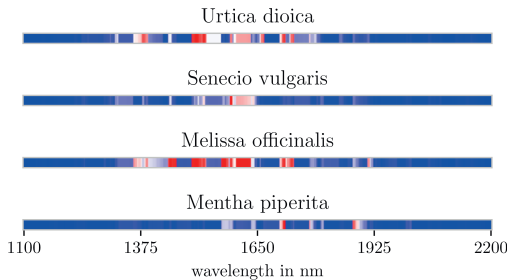


Figure 4.1: The attribution maps show the classification error caused by masking out individual spectral channels, highlighting the importance of information in the range of 1400 nm and 1900 nm.

Acknowledgement

The authors of this work were supported by the joint project "Detektion und Entfernung von Pyrrolizidinalkaloid-haltigen Unkräutern aus Kulturpflanzen nach der Ernte – PA-NIRSort" which is funded by the German Federal Ministry of Food and Agriculture and by the Fachagentur für Nachwachsende Rohstoffe e. V. (FNR) (FKZ 220132165) on the basis of a resolution of the German Bundestag and the Fraunhofer Center for Machine Learning within the Fraunhofer Cluster of Excellence Cognitive Internet Technologies CCIT.

References

1. G. Bonifazi and S. Serranti, "Imaging spectroscopy based strategies for ceramic glass contaminants removal in glass recycling," *Waste Management*, vol. 26, no. 6, pp. 627–639, Jan. 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0956053X05001649>

2. T. Phiri, H. J. Glass, and P. Mwamba, "Development of a strategy and interpretation of the NIR spectra for application in automated sorting," *Minerals Engineering*, vol. 127, pp. 224–231, Oct. 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0892687518303522>
3. S. R. Delwiche, T. C. Pearson, and D. L. Brabec, "High-Speed Optical Sorting of Soft Wheat for Reduction of Deoxynivalenol," *Plant Disease*, vol. 89, no. 11, pp. 1214–1219, Nov. 2005. [Online]. Available: <http://apsjournals.apsnet.org/doi/abs/10.1094/PD-89-1214>
4. G. Maier, F. Pfaff, C. Pieper, R. Gruna, B. Noack, H. Kruggel-Emden, T. Längle, U. D. Hanebeck, S. Wirtz, V. Scherer *et al.*, "Experimental evaluation of a novel sensor-based sorting approach featuring predictive real-time multiobject tracking," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 2, pp. 1548–1559, 2020.
5. C. B. Singh, R. Choudhary, D. S. Jayas, and J. Paliwal, "Wavelet analysis of signals in agriculture and food quality inspection," *Food and Bioprocess Technology*, vol. 3, no. 1, pp. 2–12, 2010.
6. C. R. Mittermayr, S. G. Nikolov, H. Hutter, and M. Grasserbauer, "Wavelet denoising of Gaussian peaks: A comparative study," *Chemometrics and Intelligent Laboratory Systems*, vol. 34, no. 2, pp. 187–202, 1996.
7. H. Tan, S. T. Sum, and S. D. Brown, "Improvement of a standard-free method for near-infrared calibration transfer," *Applied Spectroscopy*, vol. 56, no. 8, pp. 1098–1106, 8 2002.
8. X. Fu, G. Yan, B. Chen, and H. Li, "Application of wavelet transforms to improve prediction precision of near infrared spectra," *Journal of Food Engineering*, vol. 69, no. 4, pp. 461–466, 2005.
9. J. Trygg and S. Wold, "PLS regression on wavelet compressed NIR spectra," *Chemometrics and Intelligent Laboratory Systems*, vol. 42, no. 1-2, pp. 209–220, 1998.

Phenoliner 2.0: RGB and near-infrared (NIR) image acquisition for an efficient phenotyping in grapevine research

Xiaorong Zheng¹, Julius Krause², Benedikt Fischer², Robin Gruna², Reinhard Töpfer¹, and Anna Kicherer¹

¹ Julius Kühn-Institut, Federal Research Centre of Cultivated Plants, Institute for Grapevine Breeding Geilweilerhof, 76833 Siebeldingen, Germany

² Fraunhofer IOSB Institute of Optronics, System Technologies and Image Exploitation, Visual Inspection Systems Fraunhoferstr. 1, 76131 Karlsruhe

Abstract In grapevine research, phenotyping needs to be done for different traits such as abiotic and biotic stress. This phenotypic data acquisition is very time-consuming and subjective due to the limitation of manual visual estimation. Sensor-based approaches showed an improvement in objectivity and throughput in the past. For example, the ‘Phenoliner’ a phenotyping platform, based on a modified grape harvester, is equipped with two different sensor systems to acquire images in the field. It has so far been used in grapevine research for different research questions to test and apply different sensor systems. However, the driving speed for data acquisition has been limited to 0.5 - 1 km/h due to capacity of image acquisition frequency and storage. Therefore, a faster automatic data acquisition with high objectivity and precision is desirable to increase the phenotyping efficiency. To this aim, in the present study a prism-based simultaneous multispectral camera system was installed in the tunnel of the ‘Phenoliner’ with an artificial broadband light source for image acquisition. It consists of a visible color channel from 400 to 670 nm, a near infrared (NIR) channel from 700 to 800 nm, and a second NIR channel from 820 to 1,000 nm. Compared to the existing camera setup, image recording could be improved to at least 10 images per second and a driving speed of up to 6 km/h. Each image is geo-referenced using a real-time-kinematic (RTK)-GPS system. The setup of the sensor system was tested on seven varieties (Riesling, Pinot Noir, Chardonnay, Dornfelder, Dapako,

Pinot Gris, and Phoenix) with and without symptoms of biotic stress in the vineyards of Geilweilerhof, Germany. Image analysis aims to segment images into four categories: trunk, cane, leaf, and fruit cluster to further detect the biotic stress status in these categories. Therefore, images have been annotated accordingly and first results will be shown.

Keywords Multispectral camera, image acquisition, geo-information, *Vitis vinifera*, field phenotyping

1 Introduction

Vitis vinifera (Grapevine) is considered to be one of the most economically important fruit crops worldwide with 7.4 million ha and an annual production of 78 million tonnes in 2018. In France and Germany, 99% of the grapes are grown for wine production [1]. The yield and vine health, including wine quality are regarded as the most important economic indicators for viticulturists [2]. Grapevine is highly susceptible to several diseases, such as powdery mildew and downy mildew, of which the infection can result in a significant reduction of total soluble solids in berries and further cause a negative effect on total glucose and fructose content, total red pigments, and thus alcohol content in wine [3,4]. Therefore, the grapevine breeding activities mainly focus on selecting new varieties with high abiotic and biotic stress resistance and high quality characteristics [5]. As a perennial woody plant, observation of phenology, analysis of growth habits (grapevine architecture), and evaluation of yield are very time-consuming and can only be conducted in the field. Most of these characteristics need to be evaluated within a narrow time window resulting in limitation of phenotyping workload and somewhat limited efficiency of grapevine breeding. Therefore, faster sensor assisted phenotyping methods with high objectivity and precision have been intensively investigated in recent years to screen breeding material, such as number and size of berries or clusters [6,7], berry skin characteristics [8], leaf area [9], cane mass [10], diseases recognition [11–13] etc.

Due to the improvement of pattern recognition algorithms, computation capability, and image quality in recent years, computer vision is

being applied in agriculture [14] for disease detection in rice, maize, coffee, grapevine etc. [15–18]. To develop a more precise and faster tool for identification of plant disease by computer vision, the convolutional neural network (CNNs) techniques are numerously used. So far, most of the published methods applied for symptom detection are based on digital images acquired using wavelength of visible and near-infrared range of the spectrum [19]. With increasing optimization of algorithms, sensor-based approaches showed an improvement in objectivity and throughput in disease sensing in the past. In grapevine research, several sensor techniques have already been tested for various applications, including pre-symptomatic and symptomatic disease sensing [13,20,21]. Both ground-based hyper- and multispectral imaging are proven to be suitable for detection of foliar symptoms of esca trunk disease and powdery mildew infection in vineyard achieving overall accuracies of 82-83% and 87%, respectively [13,22]. By combining the biophysical- (e.g. content of chlorophyll, carotenoid and dry matter) and texture parameters during classification, the accuracy to predict esca disease was increased from 81% to 99% [23]. However a hyperspectral sensor, which covers a broader wavelength range to acquire more information, is usually expensive and bulky. Therefore, using multispectral or Red Green Blue (RGB)-cameras for image capture would be more feasible in agricultural applications.

Meanwhile, several platforms, such as *PHENObot* [24], Phenoliner [25], grape-picking robot [26] or unmanned aerial vehicle [27], have been developed to carry the sensor systems for sensing in vineyard or machinery picking of fruit. However, the operating speeds of these platforms is either too slow to adapt to the usual operating speed of field working machines or provides images with low resolution. Therefore, to improve the efficiency of machinery phenotyping, an update of the former screening platform 'Phenoliner' was conducted in the present study to optimize the automated image acquisition in the field, data management, and data analysis.

2 Material and methods

As described by [25], the ERO-Grapeliner SF200 (ERO Gerätebau, Simmern, Germany), of which all units for harvesting, including the hy-

draulic system were removed, served as a carrier for the sensor system. A generator driven by the vehicle was conducted to provide the energy to operate the sensors, light units, and computers.

2.1 Plant Materials

Field trials were conducted in September 2020 in the experimental vineyards at the Julius Kühn-Institut Geilweilerhof, Siebeldingen, Germany (49°13'06.0"N, 8°02'48.2"E). Seven varieties including red and white grapevine cultivars, i.e., Riesling, Pinot Noir, Chardonnay, Dornfelder, Dapako, Pinot Gris, Calardis Musqué and Rieslaner with and without symptoms of biotic stress were used for testing.

2.2 SmartVision Camera System

In the present study, the embedded image processing system *SmartVision* from Fraunhofer IOSB was used. The *SmartVision* system includes the user-interface, camera control, image acquisition and real-time image processing unit with artificial intelligence. Furthermore, *SmartVision* provides a machine-to-machine interface based on OPC-UA (Open Platform Communications Unified Architecture) using the open source implementation *open62541*. This allows the system to be controlled remotely and additional sensors can be added.

The system is equipped with a prism-based simultaneous multispectral camera system (Fusion Series FS-3200T-10GE-NNC, JAI A/S, Germany) with a 8 mm lens (VS-0818H/3CMOS, Pyramid Imaging, USA) and was installed in the right part of the tunnel of the 'Phenoliner' with an artificial light source for image acquisition as in figure 2.1. This camera system consists of a visible color channel from 400 to 670 nm, a near infrared (NIR) channel from 700 to 800 nm, and a second NIR channel from 820 to 1,000 nm. The camera system has 2048 x 1536 active pixels and single/multi-readout mode for each channel, which provides a high resolution and speed with lower processing loads. To minimize the direct sunlight interference, curtains were used on the opening of the tunnel, and therefore, five LED light lamps (Lumimax LB500-44-W, IIM AG, Germany) were installed to achieve efficient light conditions during image acquisition. In order to achieve a better image of NIR and further increase the driving speed, two broadband NIR LED bars

(EFFI-Flex, EFFILUX GmbH) were installed in addition at the side of the camera. The distance of the camera to the grapevine plants and the height of the captured region were illustrated in figure 6.1(e). The camera system was ported to a separated mini computer (MAGNUS EN072070S, Zotac, Hongkong, China) extended with WLAN access point, which serves as local host and was placed on the top platform of the harvester to process the image on-board. A 1 T of fast solid-state disc drives (EMTEC X200 Portable SSD, 450MB/s transfer speed) was used for storage of data.

To reference the geographic information to each grapevine, a real-time-kinematic GPS system (SPS852, Trimble, Sunnyvale, USA) was installed. The GPS antenna is located on top of the vehicle (Fig. 6.1(d)) and the receiver provides standardized National Marine Electronics Association (NMEA) strings for the camera system as described by Kicherer *et al.* (2017).

3 Image acquisition

The SmartVision system provides a WLAN access point and the user interface is accessible using an internet browser. Both signal intensity and visible images from both RGB- and NIR-channels are shown (Fig. 3.1), which gave the operators more information to evaluate the quality of image acquisition in the field. The developed program also provides an automatic evaluation of exposure of the images. To gain a better quality of the image several camera parameters, such as cycle time, integration time, and exposure balance were adjusted accordingly. In the end, a cycle time of 100 ms and an integration time of 9 ms were set to achieve a good image acquisition with 10 images per second at a driving speed of up to 6 km/h. This operating speed is much faster compared to the former Phenoliner setup, which is only able to move at a speed of 0.5 - 1 km/h when capturing the images with a frequency of 5 images per second [25]. Meanwhile, GPS information was taken and stored.

During the image acquisition, the program displayed the status of the application, i.e., number of image files, remaining disk space, camera frame rate etc. as well. The acquired images were saved in a hierarchical data format (hdf5) and further archived in folders with metadata

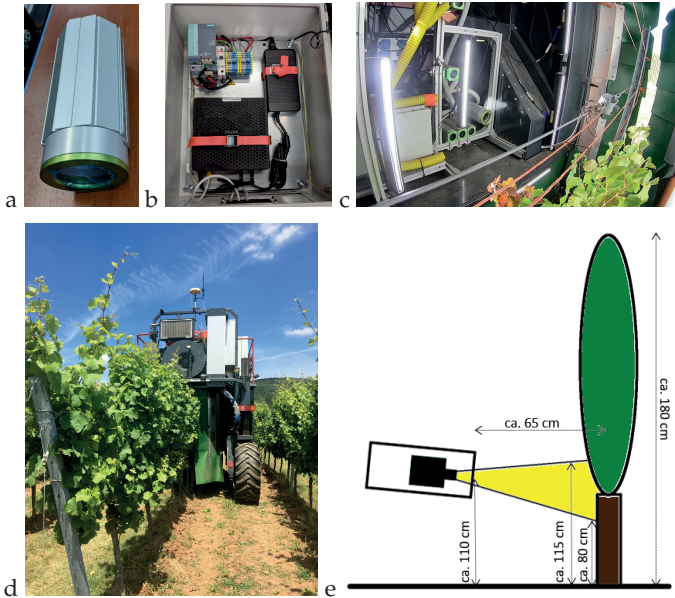


Figure 2.1: Construction of sensor system on Phenoliner. a) container for the camera, b) computer including WLAN access point for image processing, c) sensor and light units in the tunnel of Phenoliner, d) image acquisition by Phenoliner in the vineyard, e) scheme of distance.

correspondingly. The acquired images will be further processed to be segmented into four classes (trunk, cane, leaf, and fruit cluster) manually using software 'Labeltool' (Fraunhofer IOSB). Under classes of cane, leaf, and fruit cluster, at least two sub-classes, healthy and diseased (including different disease severity levels) will be given.

4 Prospective phenotyping pipeline of Phenoliner 2.0 in viticultural and grapevine breeding

As the acquisition of images with high quality and at normal working driving speed has been successfully tested in vineyards in the present study, an opportunity to apply this sensor system in two different areas

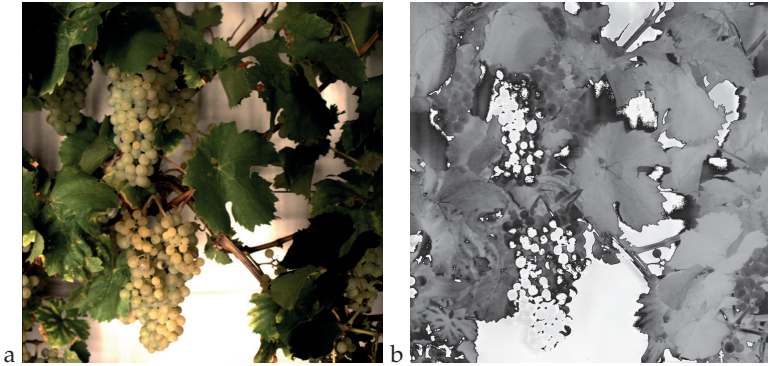


Figure 3.1: RGB Image (a) and normalized difference vegetation index (NDVI) image using RGB and NIR data (b).

is expected. On the one hand the use in grapevine breeding to describe and screen breeding material and on the other hand to improve mechanical harvest in viticulture.

Both applications have a slightly different workflow and in some steps different parameters of interest. The pipeline for both applications is shown in Figure 4.1. For mechanical harvesting a trained model based on automatic detection is planned to help the grape harvester make a yes or no decision for harvesting. This could be achieved by controlling the on and off function of the shaking units, and therefore the diseased fruit clusters could be excluded to increase the harvest quality. The application of this sensor system in grapevine breeding aims at screening selection criteria like yield and plant health fast and objective in the field. Besides the estimation of yield, the distinction of different diseases, as well as the disease incidence and severity, are relevant for the description and selection of breeding material. For digital documentation of phenotyping data, the processed results referenced to geographic information will be recorded and saved in a database. Based on the recorded data, a map with information of yield, the health status of the individual grapevine can be generated. In case of the breeding application it opens the opportunity to evaluate different breeding material over several years comparably, also retrospective if new phenotyping tools may be available.

In further experiments, this constructed sensor system could be installed in front of a commercial grape harvester and used under natural sunlight that does not provide a stable light condition compared to the Phenoliner, to investigate whether it enables the sensing of the grapevine status in real-time during harvesting. This will open up the opportunity to use such a sensor system in viticulture in the future to help the wine grower to improve his plot performance, adjust vineyard management and increase his wine quality.

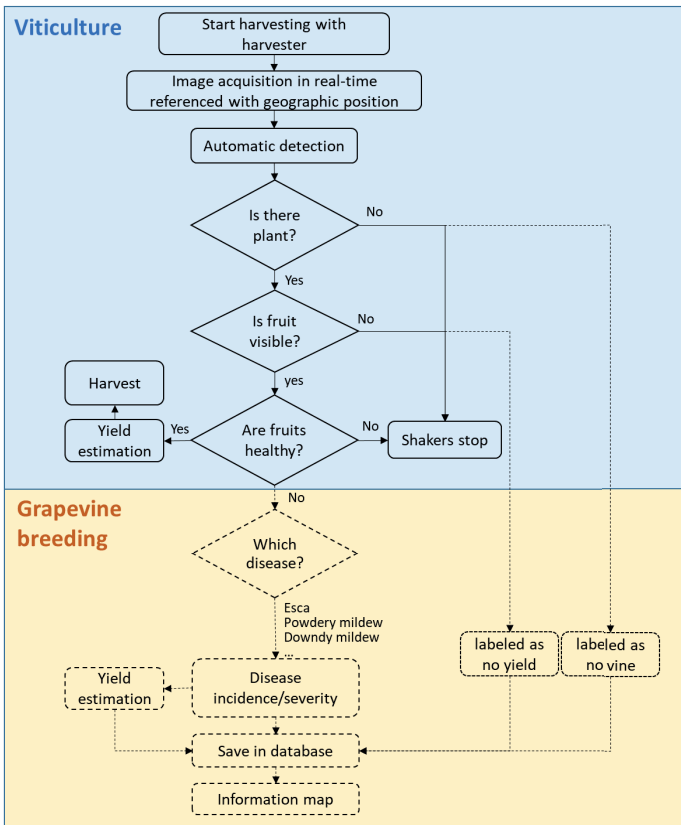


Figure 4.1: Phenotyping pipeline for grape growers and breeders during harvesting. Dotted frames and lines are considered by grape breeders only.

Acknowledgement

The study is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE), within the frame work DigiVine (FKZ: 28DE113A18).

References

1. OIV, "2019 statistical report on world vitiviniculture," 2019. [Online]. Available: <http://www.oiv.int/public/medias/6782/oiv-2019-statistical-report-on-world-vitiviniculture.pdf>
2. I. Santiago-Brown, A. Metcalfe, C. Jerram, and C. Collins, "Sustainability assessment in wine-grape growing in the new world: economic, environmental, and social indicators for agricultural businesses," *Sustainability*, vol. 7, no. 7, pp. 8178–8204, 2015.
3. D. M. Gadoury, R. C. Seem, R. C. Pearson, W. F. Wilcox, and R. M. Dunst, "Effects of powdery mildew on vine growth, yield, and quality of concord grapes," *Plant disease*, vol. 85, no. 2, pp. 137–140, 2001.
4. B. E. Stummer, I. L. Francis, A. J. Markides, and E. S. Scott, "The effect of powdery mildew infection of grape berries on juice and wine composition and on sensory properties of chardonnay wines," *Australian Journal of Grape and Wine Research*, vol. 9, no. 1, pp. 28–39, 2003.
5. R. Töpfer, L. Hausmann, M. Harst, E. Maul, E. Zyprian, and R. Eibach, "New horizons for grapevine breeding," *Fruit, Vegetable and Cereal Science and Biotechnology*, vol. 5, pp. 79–100, 2011.
6. A. Kicherer, R. Roscher, K. Herzog, S. Šimon, W. Förstner, and R. Töpfer, "Bat (berry analysis tool): A high-throughput image interpretation tool to acquire the number, diameter, and volume of grapevine berries," *VITIS - Journal of Grapevine Research*, vol. 52, no. 3, pp. 129–135, 2013.
7. A. Kicherer, R. Roscher, K. Herzog, W. Förstner, and R. Töpfer, "Image based evaluation for the detection of cluster parameters in grapevine," *Acta Horticulturae*, no. 1082, pp. 335–340, 2014.
8. K. Herzog, R. Wind, and R. Töpfer, "Impedance of the grape berry cuticle as a novel phenotypic trait to estimate resistance to botrytis cinerea," *Sensors (Basel, Switzerland)*, vol. 15, no. 6, pp. 12498–12512, 2015.

9. M. Klodt, K. Herzog, R. Töpfer, and D. Cremers, "Field phenotyping of grapevine growth using dense stereo reconstruction," *BMC bioinformatics*, vol. 16, p. 143, 2015.
10. A. Kicherer, M. Klodt, S. Sharifzadeh, D. Cremers, R. Töpfer, and K. Herzog, "Automatic image-based determination of pruning mass as a determinant for yield potential in grapevine management and breeding," *Australian Journal of Grape and Wine Research*, vol. 23, no. 1, pp. 120–124, 2017.
11. A. Adeel, M. A. Khan, T. Akram, A. Sharif, M. Yasmin, T. Saba, and K. Javed, "Entropy-controlled deep features selection framework for grape leaf diseases recognition," *Expert Systems*, 2020.
12. N. Bendel, A. Kicherer, A. Backhaus, J. Köckerling, M. Maixner, E. Bleser, H.-C. Klück, U. Seiffert, R. T. Voegele, and R. Töpfer, "Detection of grapevine leafroll-associated virus 1 and 3 in white and red grapevine cultivars using hyperspectral imaging," *Remote Sensing*, vol. 12, no. 10, p. 1693, 2020.
13. N. Bendel, A. Kicherer, A. Backhaus, H.-C. Klück, U. Seiffert, M. Fischer, R. T. Voegele, and R. Töpfer, "Evaluating the suitability of hyper- and multispectral imaging to detect foliar symptoms of the grapevine trunk disease esca in vineyards," *Plant methods*, vol. 16, no. 1, p. 142, 2020.
14. E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine vision systems in precision agriculture for crop farming," *Journal of Imaging*, vol. 5, no. 12, p. 89, 2019.
15. T. V. Price, R. Gross, W. J. Ho, and C. F. Osborne, "A comparison of visual and digital image-processing methods in quantifying the severity of coffee leaf rust (*hemileia vastatrix*)," *Australian Journal of Experimental Agriculture*, vol. 33, no. 1, p. 97, 1993.
16. D. Sena Jr, F. Pinto, D. Queiroz, and P. Viana, "Fall armyworm damaged maize plant identification using digital images," *Biosystems Engineering*, vol. 85, no. 4, pp. 449–454, 2003.
17. Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017.
18. R. Marani, A. Milella, A. Petitti, and G. Reina, "Deep neural networks for grape bunch segmentation in natural images from a consumer-grade camera," *Precision Agriculture*, 2020.
19. J. G. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," *SpringerPlus*, vol. 2, p. 660, 2013.

20. S. Bellow, G. Latouche, S. C. Brown, A. Poutaraud, and Z. G. Cerovic, "Optical detection of downy mildew in grapevine leaves: daily kinetics of autofluorescence upon infection," *Journal of experimental botany*, vol. 64, no. 1, pp. 333–341, 2013.
21. E.-C. Oerke, K. Herzog, and R. Toepfer, "Hyperspectral phenotyping of the reaction of grapevine genotypes to plasmopara viticola," *Journal of experimental botany*, vol. 67, no. 18, pp. 5529–5543, 2016.
22. U. Knauer, A. Matros, T. Petrovic, T. Zanker, E. S. Scott, and U. Seiffert, "Improved classification accuracy of powdery mildew infection levels of wine grapes by spatial-spectral analysis of hyperspectral images," *Plant methods*, vol. 13, p. 47, 2017.
23. H. Al-Saddik, A. Laybros, B. Billiot, and F. Cointault, "Using image texture and spectral reflectance analysis to detect yellowness and esca in grapevines at leaf-level," *Remote Sensing*, vol. 10, no. 4, p. 618, 2018.
24. A. Kicherer, K. Herzog, M. Pflanz, M. Wieland, P. Ruger, S. Kecke, H. Kuhlmann, and R. Topfer, "An automated field phenotyping pipeline for application in grapevine research," *Sensors (Basel, Switzerland)*, vol. 15, no. 3, pp. 4823–4836, 2015.
25. A. Kicherer, K. Herzog, N. Bendel, H.-C. Kluck, A. Backhaus, M. Wieland, J. C. Rose, L. Klingbeil, T. Labe, C. Hohl, W. Petry, H. Kuhlmann, U. Seiffert, and R. Topfer, "Phenoliner: A new field phenotyping platform for grapevine research," *Sensors (Basel, Switzerland)*, vol. 17, no. 7, 2017.
26. J. Xiong, Z. Liu, R. Lin, R. Bu, Z. He, Z. Yang, and C. Liang, "Green grape detection and picking-point calculation in a night-time natural environment using a charge-coupled device (ccd) vision sensor with artificial illumination," *Sensors (Basel, Switzerland)*, vol. 18, no. 4, 2018.
27. M. Kerkech, A. Hafiane, and R. Canals, "Vine disease detection in uav multispectral images using optimized image registration and deep learning segmentation approach," *Computers and Electronics in Agriculture*, vol. 174, p. 105446, 2020.

Developing a handheld NIR sensor for the detection of ripening in grapevine

Lucie Gebauer¹, Julius Krause², Xiaorong Zheng¹, Felix Kronenwett², Robin Gruna², Reinhard Töpfer¹, and Anna Kicherer¹

¹ Julius Kühn-Institut, Federal Research Centre of Cultivated Plants, Institute for Grapevine Breeding, Geilweilerhof, 76833 Siebeldingen

² Fraunhofer IOSB Institute of Optronics, System Technologies and Image Exploitation, Visual Inspection Systems, Fraunhoferstr. 1, 76131 Karlsruhe

Abstract It has already been proven that near infrared (NIR) reflectance spectroscopy can be used to measure the maturity of grapes by the determination of the sugar and acid content. Until now, winegrowers frequently collect a random one hundred berries sample per plot, to measure these parameters destructively for the estimation of the ideal harvest time of the gained product. Meanwhile, inexpensive sensors are available, to build convenient instruments for the non-destructive, low-priced and fast control of ripening parameters in the vineyard. For this, a small handheld device including a NIR sensor (1350 nm – 2600 nm) was built from a Raspberry Pi 3 single-board computer and a NIR sensor. Spectra of individual berries, sampled from Riesling (*Vitis vinifera*, L.) were collected. Corresponding reference data were determined with high performance liquid chromatography (HPLC). Samples were taken from different fruit as well as cluster zones and from the beginning of véraison until after harvest, to ensure a broad range of ingredients and the ripening properties of different berries from the vine. This study is the first that systematically investigates the ripening parameters of a whole vineyard with a handheld sensor. The sensor can be used in viticulture practice to detect the ripening progress and determining the ideal harvest time effective, simple, and non-destructively.

Keywords *Vitis vinifera*, Near Infrared Spectroscopy (NIRS), reflectance, ripening parameters, soluble solids, sugar, acid, handheld sensor, individual berry measurement

1 Introduction

Viticulture is of great economic importance worldwide with an estimated average production of 258 million of hectoliters (Mio hl) of wine in 2020. In relation to global production, 60% were produced in Europe (156 Mio hl), while the USA (24.3 Mio hl), China (8.3 Mio hl) and Russia (4.6 Mio hl) are the leading countries outside of the European Union. Within Europe, Italy (47.5 Mio hl) is the largest wine producer followed by France (42.1 Mio hl), Spain (33.5 Mio hl), and Germany (9.0 Mio hl). Riesling is one of the widely grown white grapevine varieties worldwide, and in Germany the most important cultivar of about 23% (2019) of the wine grapes area. [1,2]

One important factor to hold a large market share is the quality of the produced wine. While the proportion of table wine in Germany is relatively stable, the production of quality wine increases rapidly [1]. In order to obtain high quality grapes, the harvest time is of great importance, since after reaching the peak of ripening, over-ripening results in a decrease in quality. The growth of grape berry takes place in three phases consisting of two growth cycles and one lag-phase [3,4]. After the lag-phase, from véraison onwards, acids (mainly malic acid) slowly start to decrease while sugars increase rapidly. The presence and amount of mainly sugars and acids determines the quality and later characteristics of the wine, whereby the sugars are mainly represented by glucose and fructose and 70% to 90%, respectively, and the acid being mainly tartaric acid and malic acid [5–8]. Sugars determine the alcohol content of the later wine and acidity contributes to its fruity character. The balance of sugar and acidity are determining factors of typicality and wine style which is increasingly affected by climate change. [9] Due to uneven ripening on the cluster and on the grapevine, a winegrower in general collects one hundred berries, to estimate the average maturation of his vineyard. This process is destructive, laborious and to some extent error prone.

NIR sensors are already widespread in a wide variety of disciplines. They are used in medicine (blood oxygen, diabetes, intracranial bleeding), pharmacy or agriculture (particle measurements), and to ensure food quality. Till now, several vibrational spectroscopic methods were discovered to analyse different materials, based on the interaction of radiation with matter. The most prominent method is the FTIR, which is

based on the lowering of the radiation intensity due to infrared-active (IR-active) bonds. In contrast to the passing of light through the material, near infrared spectroscopy relies on the principle of reflection and absorption of radiation from the visible and near infrared spectrum of light (400 nm – 2500 nm). The IR-active bonds lower the reflection, resulting in an increase of the extinction coefficient (the degree of radiation attenuation), depicted as a peak in the reflection spectrum.

In order to investigate a feasible technique to detect quality attributes in crops, several sensors using IR radiation have been built for several fruits and vegetables. [10–12] Whilst some sensors measure the transmission spectrum of destroyed berries/bunches [13], other measure the diffuse reflection but with complex and expensive built ups with lamps and sensors. [14,15] Goisser *et al.* 2019 [16] developed a handheld NIR-sensor which measures diffuse reflection, by simply placing the fruit on the sensor. This technique could be used to classify the firmness and to determine sugar content of tomatoes.

Here, we built and investigated a handheld sensor to determine the most important quality attributes of an entire vineyard without having to repeatedly destroy grapes. With a few further developments, such as an app that can process the spectra immediately, we can give winemakers a simple and quick-to-use tool.

2 Material and methods

2.1 SmartSpectrometer system

The embedded spectrometer system *SmartSpectrometer* from Fraunhofer IOSB includes the spectrometer control and a real-time spectral processing unit with artificial intelligence. Furthermore, *SmartSpectrometer* provides a machine-to-machine interface based on OPC-UA (Open Platform Communications Unified Architecture) using the open source implementation open62541. This allows the system to be controlled remotely and additional sensors can be added. The built system consists of a FT-NIR sensor (NeoSpectra-Micro SWS 62231, Neospectra, SI-Ware, La Canada, California) mounted on a Raspberry Pi 3 (Raspberry Pi Foundation, Cambridge, United Kingdom) single-board computer. The sensors measures near infrared radiation in the range of 1350 nm - 2600 nm with a resolution of 16 nm.

The spectra of the berries were measured and recorded in the laboratory. A white and a black calibration was done with a Spectralon[®] reflection standard (Sphereoptics, Labsphere, Inc., North Sutton, NH), to minimize influence of the ambient light and background noise. Individual berries were dried with a paper towel, placed on the sensor, and five spectra from different sides of each berry were taken using a graphical user interface.

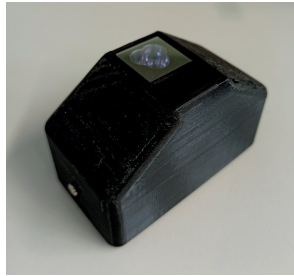


Figure 2.1: Handheld *SmartSpectrometer* device based on the NeoSpectra Micro development Kit. To determine the degree of maturity, the reflection spectrum is recorded non-destructively and evaluated directly in the embedded device.

2.2 Dataset

Grapevines of the variety Riesling (*Vitis vinifera*, L.) from four different locations in the individual vineyard site “Mütterle” in Wollmesheim (WH, Landau, Rhineland Palatinate, Germany) were used (see Fig. 2.2, Table 7.1). All vines were healthy, leaves and berries were free of damages. In the plots, vine and row spacing were comparable, the greening in the tracks was well-tended and rows were oriented north-south. All vines were trained in semi-arched canes and there were either one or two fruit zones, depending on height of the canopy, with a height of 30 cm to 50 cm per zone. Per vine and fruit zone two or four berries respectively were collected.

In total 512 individual berries were harvested. The sampling took place each week, beginning with véraison (08-17-2020) until harvest (09-28-2020) from defined vines, and one week after harvest (10-05-2020) from two to four vines in the field. Berries were taken from

defined but random selected vines, according to their exposure to the sun (sun exposed, shaded), position in the cluster (near or far from the peduncle) and position at the vine (Table 7.1). Additionally, from each plot one hundred berries were randomly picked once a week from 08-10-2020 till harvest 09-28-2020, in order to determine the average maturity of the vineyard. Berries were transported to laboratory with a cooling box equipped with ice packs.

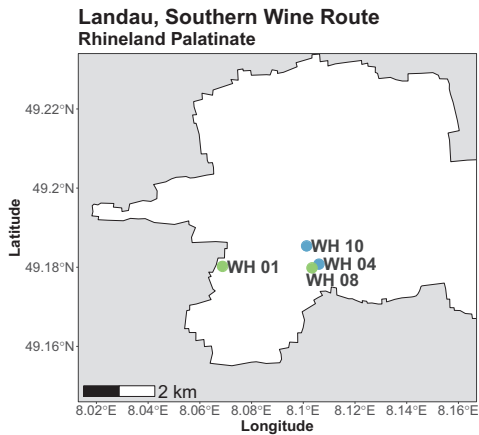


Figure 2.2: Wine growing areas in Wollmesheim (WH) near Landau, Rhineland Palatinate, Germany. Cultivar: *Vitis vinifera* (L.); variety: Riesling, individual vineyard site: Mütterle; farming practice: organic (green), conventional (blue).

Table 7.1: Rating data from the used Riesling areas in 2020 (Figure 2.2), created by the cooperative Deutsches Weintor; farming practice: organic (Org), conventional (Con); n.d.: not documented.

Abbre- viation	Farming practice	Area (ha)	Planting year	Root- stock	Canopy height	Defolia- tion
WH01	Org	13.32	2005	5C	1.10 m - 1.30 m	n.d.
WH08	Org	48.60	1991	Binova	1.10 m - 1.30 m	none
WH04	Con	48.67	1991	5C	> 1.30 m	one-sided
WH10	Con	40.48	2006	5C	> 1.30 m	one-sided

2.3 Reference data

The juice for the measurement of the one hundred berries sample was obtained by mixing them with a commercial available mixer (BL 6280, Grundig, Germany). Individual berries were destroyed in a Falcon tube by shaking them with four metal bullets in a paint shaker (SK450 Fast and Fluid Management, Sassenheim, Netherlands). All Falcons were centrifuged at $25,419 \cdot g$ in a cooled centrifuge (Sigma 6–16ks, Sigma, Kawasaki, Japan) to discard the cell debris. The juice was transferred, centrifuged with $12,100 \cdot g$ (Minispin Eppendorf, Hamburg, Germany) again and 1:3 diluted with double distilled and filtered (pore size 0.2 nm) water. Amount of sugars and acids (glucose, fructose, malic and tartaric acid) was determined using high performance liquid chromatography (Agilent 12900 Infinity 2, Agilent Technologies, Inc., Santa Clara, California) with ion-exchange ligand-exchange HiPlex H column (Agilent Technologies, Inc., Santa Clara, California), a refractive index detector for detection of carbohydrates and a diode array detector for acids. A standard series ranged from 1.5 to 90 g/L and 0.25 to 15 g/L for sugars and acids, respectively, was used. Recorded data were processed with Agilent OpenLab CDS Chemstation Software (Agilent Technologies, Inc., Santa Clara, California).

2.4 Least Squares Support Vector Regression

Due to the Beer-Lambert law, there is a non-linear relationship between the optical signal and a concentration of compounds, therefore Least Squares Support Vector Regression (LSSVR) with an RBF kernel was chosen as a non-linear regression method. The so-called kernel trick enables the estimation of non-linear correlations using LSSVR.

For the basic LSSVR the linear relation $y = \mathbf{w}\mathbf{x} + b$ between the regressors x and the dependent variable y is calculated by solving the following optimization problem

$$Q_{\text{LSSVM}} = J(\mathbf{w}, \mathbf{e}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \gamma \sum_{i=1}^n e_i^2 \quad (2.1)$$

subject to the equality constraints

$$y_i - \mathbf{w}^T \mathbf{x}_i - b = e_i, \quad i = 1, 2, \dots, l \quad (2.2)$$

where γ is the regularization parameter and e_i is the regression error. The advantage of the LSSVR over the regular SVM is provided via converting a quadratic programming problem into a set of linear equations. By constructing the Lagrange function with the Lagrange multipliers α , a set of conditions for optimality can be derived. This leads to a set of linear equations that needs to be solved to obtain the parameters α and b . The resulting LSSVM model used in function estimation is

$$y_i = \sum_{i=1}^n \alpha_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (2.3)$$

with the RBF kernel function K . For the model calculation the regularization parameter γ and the bandwidth of the RBF kernel σ needs to be chosen [17].

3 Results and discussion

Reference data were evaluated and depicted using R (Version 3.6.1) [18] and R-Studio [19], as well as the packages ggplot2 [20] and Rmisc [21]. Spectral data analysis and modelling was performed using the spectraltoolbox framework from Fraunhofer IOSB based on Python 3.8.

3.1 Reference data for spectral data processing

The sugar (glucose, fructose) and acid (tartaric -, malic acid) contents of berries were measured with high performance liquid chromatography (HPLC) and results of hundred berries samples are depicted in Figure 3.1.

In individual berries, sugar contents ranged from 6.59 g/l to 115.58 g/l and from 11.52 g/l to 108.83 g/l for glucose and fructose, respectively, over the entire ripening process. For acids, ranges were smaller and high values were gained at the begin of the ripening with 15.45 g/l for tartaric and 21.79 g/l for malic acid, respectively. While the malic acid was partially almost completely metabolized, minimum tartaric acid content was 3.93 g/l.

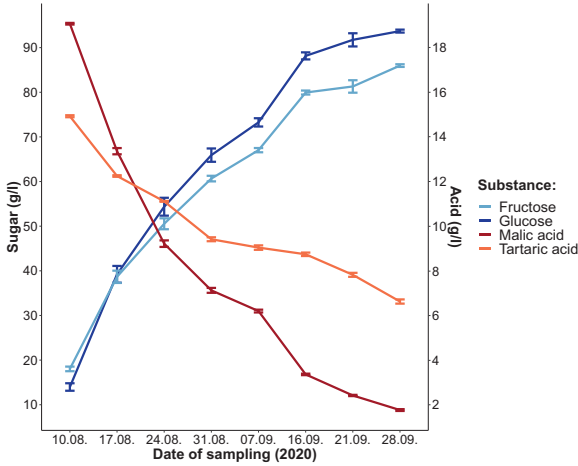


Figure 3.1: Sugar and acid contents in the hundred berries samples, measured with HPLC, error bars correspond to standard errors, each data point represents the mean of the four parcels.

3.2 NIR data analysis

The data set contains spectral data from four plots at different locations. The validation is performed like the later use-case of the ripeness estimation of a plot at a specified date. Therefore, the plot used for validation was not included in the training and the median of all predictions is used.

The evaluation of the spectral data was done in several steps. First, spectra with a low signal strength (average intensity below 0.04) were discarded. Subsequently, the intensities of the spectral data were normalised using Standard Normal Variate (SNV) [22]:

$$x_{\text{snv}} = \frac{x - \bar{x}}{\sigma_x} \quad (3.1)$$

As a metric to evaluate the regression model, the RMSE and R^2 score is used. The RMSE score

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3.2)$$

estimates the standard deviation of the prediction of a regression model. A distinction can be made between the RMSE of prediction and the RMSE of calibration, depending on y_i used during calculation. In addition, the R2 score

$$R^2 = 1 - \frac{\sum_{i=1}^n (e_i^2)}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.3)$$

indicates how well the independent variables are suited to explain the variance of the dependent variables. In both formulas n is the number of observations.

Due to the lack of sample preparation, the measurement in reflection with low signal strength and the use of miniaturised low-cost sensors, there is an enhanced measurement uncertainty. Because the prediction of the entire plot at a point in time are relevant for the winemaker, stochastic errors and outliers can be reduced very well by determining the median of the single berry predictions. The quality of the prediction of individual grape berries in the training and the median of the prediction of the validation plot at each time point is shown in Table 7.2.

It can be clearly seen that the median of the single berry prediction leads to very good results. More precisely, Figure 3.2 shows the median of acid and sugar of the measured individual berries at different time points. With this model we were able to determine glucose and fructose content with 87% accuracy (± 7.59 g/l and ± 6.57 g/l, respectively) and tartaric, as well as malic acid with 89% and 78% accuracy (± 0.52 g/l and ± 1.89 g/l, respectively). Compared to previous studies that measured the total soluble solids, we were able to predict two individual sugars with a high degree of accuracy [23,24]. The two selected acids could also be predicted separately, with a better forecast for tartaric acid, which is the prevalent acid at harvest [23,25]. Higher precision could eventually be achieved with other models, for which more data are to be collected.

Table 7.2: Results of the spectral evaluations using LSSVR. Shown are root mean square errors of calibration ($RMSE_C$) from the training set, the root mean square error of the median prediction ($RMSE_P$) from validation set, the degree of determination (R^2).

Validation		training		validation	
Plot	Compound	individual grape berries	median prediction		
		$RMSE_C$	R^2	$RMSE_P$	R^2
WH 01	Glucose	5.29	0.94	9.90	0.87
	Fructose	4.48	0.94	8.75	0.86
	Tartaric acid	0.51	0.93	1.80	0.54
	Malic acid	1.02	0.93	2.11	0.84
WH 04	Glucose	5.21	0.94	7.59	0.87
	Fructose	4.48	0.94	6.57	0.87
	Tartaric acid	0.68	0.90	0.52	0.89
	Malic acid	1.06	0.93	1.89	0.78
WH 08	Glucose	5.34	0.94	8.70	0.78
	Fructose	4.63	0.94	7.98	0.77
	Tartaric acid	0.61	0.91	1.17	0.62
	Malic acid	0.98	0.94	1.40	0.86
WH 10	Glucose	5.67	0.93	10.14	0.80
	Fructose	4.83	0.93	8.47	0.81
	Tartaric acid	0.63	0.91	1.25	0.59
	Malic acid	1.09	0.93	1.37	0.86

4 Summary

Sugar (alcohol) and acid content of grapes have a huge impact on the sensory perception of wine. Measuring of these ingredients is a proxy for ripening of a vineyard.

Proof of concept was provided on applying a miniature sensor to measure the maturity of Riesling in vineyards with high accuracy. In view of the application in the vineyard, the prototype offers an option for monitoring the ripening in a time efficient way. It is expected that the sensor will be manageable with one hand. If an App for mobile phones is developed, the results of the measurement will become immediately apparent. The alternative to invasive measurements being laborious e.g. a handheld sensor to quantify sugars and acids non-invasively on hundreds of berries is a step forward towards digitalization and precision viticulture.

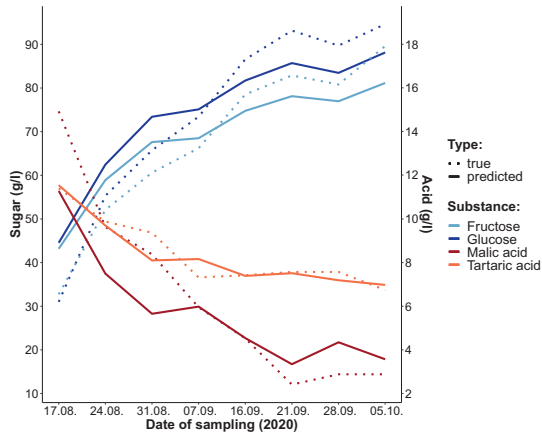


Figure 3.2: Median of determined (true, solid line) and estimated (predicted, dashed line) sugar and acid contents in single berries of the plot WH04.

Acknowledgement

The study is supported by funds of the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE), within the frame work DigiVine (FKZ: 28DE113A18). We also thank Ann-Kathrin Ickert and Maximilian Mathes for technical support and Michael Straube from the cooperative Deutsches Weintor e.G. for arranging experimental areas from winegrowers and corresponding rating data.

References

1. Deutsches Weininstitut GmbH, "Deutscher Wein Statistik 2020/2021," 2020. [Online]. Available: www.deutscheweine.de, www.germanwines.de
2. OIV, "Weinerzeugung 2020, statusbericht der oiv - erste schätzungen." [Online]. Available: <http://www.oiv.int/public/medias/7548/de-weinerzeugung-2020-oiv-erste-sch-tzungen.pdf>
3. B. G. Coombe, "Relationship of growth and development to changes in

- sugars, auxins, and gibberilins in fruit of seeded and seedless varieties of *vitis vinifera*," *Plant Physiology*, vol. 35, no. 2, pp. 241–250, 1960.
4. B. G. Coombe and C. R. Hale, "The hormone content of ripening grape berries and the effects of growth substance treatments," *Plant Physiology*, no. 51, pp. 629–634, 1973.
 5. C. Conde, P. Silva, N. Fontes, A. C. P. Dias, R. M. Tavares, M. J. Sosua, A. Agasse, S. Delrot, and H. Gerós, "Biochemical changes throughout grape berry development and fruit and wine quality," *Global Science Books*, vol. 1, no. 1, pp. 1–22, 2007.
 6. H.-F. Liu, B.-H. Wu, P.-G. Fan, S.-H. Li, and L.-S. Li, "Sugar and acid concentrations in 98 grape cultivars analyzed by principal component analysis," *Journal of the Science of Food and Agriculture*, vol. 86, no. 10, pp. 1526–1536, 2006.
 7. R. B. Beelman and J. F. Gallander, "Wine deacidification," ser. *Advances in Food Research*. Elsevier, 1979, vol. 25, pp. 1–53.
 8. H. P. Ruffner, "Metabolism of tartaric and malic acids in *vitis*. a review - part a," *Vitis*, vol. 21, pp. 247–259, 1982.
 9. R. Mira de Orduña, "Climate change associated effects on grape and wine quality and production," *Food Research International*, vol. 43, no. 7, pp. 1844–1855, 2010.
 10. D. Cozzolino, R. Damberg, L. Janik, and W.U. Cynkar and M. Gishen, "Review: Analysis of grapes and wine by near infrared spectroscopy," *Journal of Near Infrared Spectroscopy*, no. 14, pp. 279–289, 2006.
 11. D. Cozzolino, W. Cynkar, N. Shah, and P. Smith, "Technical solutions for analysis of grape juice, must, and wine: the role of infrared spectroscopy and chemometrics," *Analytical and bioanalytical chemistry*, vol. 401, no. 5, pp. 1475–1484, 2011.
 12. C. A. T. dos Santos, M. Lopo, R. N. M. J. Páscoa, and J. A. Lopes, "A review on the applications of portable near-infrared spectrometers in the agro-food industry," *Applied spectroscopy*, vol. 67, no. 11, pp. 1215–1233, 2013.
 13. J. Fernández-Novales, M.-I. López, M.-T. Sánchez, J. Morales, and V. González-Caballero, "Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines," *Food Research International*, vol. 42, no. 2, pp. 285–291, 2009.
 14. A. J. Daniels, C. Poblete-Echeverría, U. L. Opara, and H. H. Nieuwoudt, "Measuring internal maturity parameters contactless on intact table grape

- bunches using nir spectroscopy," *Frontiers in plant science*, vol. 10, p. 1517, 2019.
15. H. Abdul Rahim, K. S. Chia, and R. Abdul Rahim, "Visible and near infrared (vir-nir) spectroscopy: Measurement and prediction of soluble solid content of apple," *Sensors & Transducers Journal*, vol. 121, no. 10, pp. 42–498, 2010.
 16. S. Goisser, J. Krause, M. Fernandes, and H. Mempel, "Determination of tomato quality attributes using portable nir-sensors: Conference paper," 2019.
 17. J. A. K. Suykens, L. Lukas, and J. Vandewalle, "Sparse approximation using least squares support vector machines," in *2000 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 2, 2000, pp. 757–760.
 18. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019. [Online]. Available: <https://www.R-project.org>
 19. RStudio Team, *RStudio: Integrated Development Environment for R*, RStudio, Inc., Boston, MA, 2019. [Online]. Available: <http://www.rstudio.com/>
 20. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. [Online]. Available: <https://ggplot2.tidyverse.org>
 21. R. M. Hope, *Rmisc: Ryan Miscellaneous*, 2013, r package version 1.5. [Online]. Available: <https://CRAN.R-project.org/package=Rmisc>
 22. A. 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.
 23. F. Antonucci, F. Pallottino, G. Paglia, A. Palma, S. D'Aquino, and P. Mene-satti, "Non-destructive estimation of mandarin maturity status through portable vis-nir spectrophotometer," *Food and Bioprocess Technology*, vol. 4, no. 5, pp. 809–813, 2011.
 24. S. Lee, S. Sarkar, Y. Park, J. Yang, and G. Kweon, "Feasibility study for an optical sensing system for hardy kiwi (*actinidia arguta*) sugar content estimation," *Journal of Agriculture & Life Science*, vol. 53, no. 3, pp. 147–157, 2019.
 25. R. Beghi, A. Mena, V. Giovenzana, R. Civelli, S. Best, L. F. Leòn Gutiérrez, and R. Guidetti, "Quick quality evaluation of chilean grape by a portable vis/nir device," *Acta Horticulturae*, no. 978, pp. 93–100, 2013.

Fine metal-rich waste stream characterization based on RGB data: Comparison between feature-based and deep learning classification methods

Nils Kroell, Kay Johnen, Xiaozheng Chen, and Alexander Feil

Department of Anthropogenic Material Cycles, RWTH Aachen University,
Wüllnerstraße 2, 52062 Aachen

Abstract *Background:* Material compositions in the recycling industry are currently determined by manual sorting, which is time intensive and shows subjective influences. For an automated, sensor-based material flow characterization a particle-based material classification is necessary. *Aim:* The classification of metal-containing fine-fractions based on RGB images with different machine learning (ML) techniques is investigated on two created datasets A (12,480 images) and B (19,498 images). *Method:* Two approaches are compared: In approach I, images are firstly pixel- and then object-based classified with six different ML models on three color spaces. In approach II, images are classified by six different convolutional neural networks (CNNs). *Results:* The classification of dataset A was possible with high accuracy ($> 99.8\%$) for both approaches and chosen ML algorithms were of minor importance. For dataset B, approach I achieved an accuracy of $78.2\% \pm 2.0\%$, and chosen ML algorithms were of higher importance for object-based classification. In approach II, the best-performing CNN achieved an accuracy of $80.4\% \pm 4.2\%$ and a top-3 score of $94.2\% \pm 2.6\%$. *Conclusion:* Results from existing studies for coarser particle sizes can be transferred to fine fractions. Further research is needed to improve the classification of dataset B, e. g. by adding instances to less frequent classes and applying deeper CNNs.

Keywords Sensor-based material flow characterization, fine fractions, recycling, metals, RGB, classification, machine learning, deep learning, convolutional neural networks.

1 Introduction

In the European Union, around 15.7 million Mg of non-ferrous (NF) metal-containing waste was generated in 2016 [1]. Recycling of valuable materials inside substitutes primary resources and therefore, achieves environmental benefits [2]. For example, recycling of aluminium saves 95 % of the used energy and reduces greenhouse gas emissions by 92 % compared to primary production [3].

Currently, about one-third of all NF metals produced originate from secondary raw material sources [4]. With rising climate change [5], decreasing valuable material contents of primary resources [6, p. 47] and an increasing demand for raw materials from emerging and developing countries [7], a transition towards a circular economy and thus an increased exploitation of secondary raw materials is necessary [8].

During the processing of metal scrap, metal-containing fine fractions (≤ 20 mm) are generated as a by-product [9, p. 263f], which can contain up to 30 wt% metals [10] but are insufficiently recovered to date [10]. For an optimal recovery of the contained metals as well as a quality assurance of generated product streams, the material composition needs to be known. Currently, samples are analysed by manual sorting [11, p. 51], which is time and cost intensive and often shows subjective influences from the manual sorter [12]. As a result, material flows are often characterized on an irregular basis and sample amounts are statistically limited.

A sensor-based material characterization (SBMC) offers the possibility to overcome these limitations, as it minimizes subjective influences, and material flows can be monitored in nearly real-time. The first step of such an SBMC is the classification of individual particles to given material classes. In a second step, the predicted material classes can be combined with individual particle masses in order to derive mass-based material compositions for e. g. process control or automated quality assurance.

NF metals can be classified by different sensors, e. g. X-ray transmission or fluorescence, laser induced breakdown spectroscopy, induction (IND) or color sensors (RGB as well as hyperspectral imaging (HSI)) [13]. In comparison to other available sensors, RGB sensors are available in high resolution, technically matured and often by one order of magnitude or more cost-effective and thus focus of this study.

2 Related Work

While sensor-based sorting of metals is state of the art [13], the SBMC of metal-containing waste streams is an active area of research. For coarse particle sizes (> 20 mm), the classification based on RGB images has been researched in several studies:

- Wang et al. (2019) [14] demonstrated the classification of aluminium and copper from shredded end-of-life vehicles based on particle related features. With a support vector machine (SVM), an object-based classification score of 96.64 % was achieved.
- Dang et al. (2019) [15] compared the classification of metal objects with convolutional neural networks (CNNs). With an AlexNet-architecture [16], the material classes copper, brass and other metals could be classified with an object-based accuracy of 97.15 %.
- Karbasi et al. (2018) [17] studied the classification of particles from shredded waste of electrical and electronic equipment (WEEE) using a ResNet-architecture [18]. For the three material classes printed circuit boards (PCB), plastics and cables, an object-based accuracy of 98 % was achieved.

Studies on the classification of metal-containing fine fractions (≤ 20 mm) until now have focused on the application of hyperspectral imaging or a combination of multiple sensors, see Table 8.1. The classification of these fine fractions only based on RGB images has not been investigated yet, which is the research gap addressed here.

Table 8.1: Existing studies on optical classification of metal-containing fine fractions; 3DLT: 3D laser triangulation, * sorting rate, ** product purity.

Study	Date	Material		Classification accuracy:		
		origin	Sensor	pixel-based	object-based	
Candiani et al.	2017	PCB	HSI	87.7 %	97.2 %	[19]
Barnabé et al.	2015	WEEE	HSI	83.9 %	96.0 %	[20]
Huang et al.	2010	NF metals	RGB + 3DLT	-	98.0 %*	[21]
Picon et al.	2009	WEEE	HSI	71.7 %	98.4 %	[22]
Kutilla et al.	2006	NF metals	RGB + IND	-	80.0 %**	[23]

3 Material and Methods

Two machine learning (ML) approaches (I: "feature-based classification" and II: "deep learning") are compared. For training and evaluation of the models, two datasets A and B were created.

3.1 Data Acquisition

Particle images were taken with a custom made test rig on a white paper background by a high-resolution RGB-camera (IDS U3-3800CP-C-HQ Rev.2) with a color depth of 12-bit and a resulting spatial resolution of 0.11 mm/Pixel. The recording area of 100 mm × 100 mm was illuminated from top by white LED-stripes (Samsung 2835 120 LED 12V IP20 4000K). For each particle, a front and back image was recorded.

3.2 Samples and Datasets

Dataset A contains standard components made of copper (Cu), grey metals (GM), brass (Msg) and black plastics (KS). Characteristic particle shapes and sizes are generated by comminuting the particles in a hammer mill and sieving them into the investigated particle size range of 3.15 mm to 10 mm ($\Delta t_{\text{sieving}} = 90$ s). In total, there are $n = 12,480$ images in the dataset ($n_{\text{Cu}} = 4,016$; $n_{\text{GM}} = 2,120$; $n_{\text{Msg}} = 3,056$; $n_{\text{KS}} = 3,288$), see Figure 3.1. After acquisition, the dataset was split in a train and test dataset with a train-test-split-ratio of 80 % : 20 %.

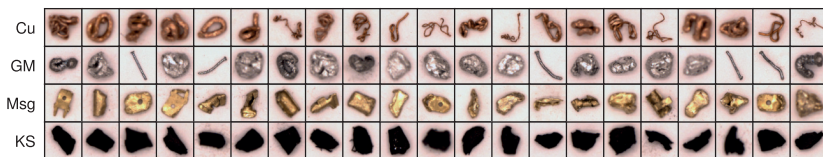


Figure 3.1: Exemplary images (cropped) for the four material classes of dataset A.

For dataset B, samples in the particle range 3 mm to 9 mm were collected from a metal-recovery plant in Germany according to LAGA PN 98 [24]. The samples were manually sorted in 22 material classes with a two stage process (see [12] for details). The final dataset contains

19,498 images, see Table 8.2 and Figure 3.2, and was split in a train and test dataset with a train-test-split-ratio of 75 % : 25 % to guarantee enough instances for less frequent classes in the test set.

Table 8.2: Number of images n in dataset B. +: Composite of the mentioned materials.

Material class	Abbreviation	n	Material class	Abbreviation	n
Copper	Cu	3,123	GM wires	GMKaAd_oS	38
Grey metals	GM	2,169	GM wires + plug	GMKaAd_mS	2
Brass	Msg	2,080	PCB	PCBA	864
Plastics	KS	2,453	GM + NM	GM_NM	195
Glass	Glas	48	Cu + NM	Cu_NM	893
Minerals	KSP	50	Msg + NM	Msg_NM	63
Residual	Rest	390	GM + Cu	GM_Cu	1,702
Cu wires	CuAd_oS	898	GM + Msg	GM_Msg	134
Cu wires + plug	CuAd_mS	174	Cu + Msg	Cu_Msg	291
Cu cables	CuKa_oS	30	GM + Cu + Msg	GM_Cu_Msg	62
Cu cables + plug	CuKa_mS	5	Complex	Komplex	54
				Σ	19,498

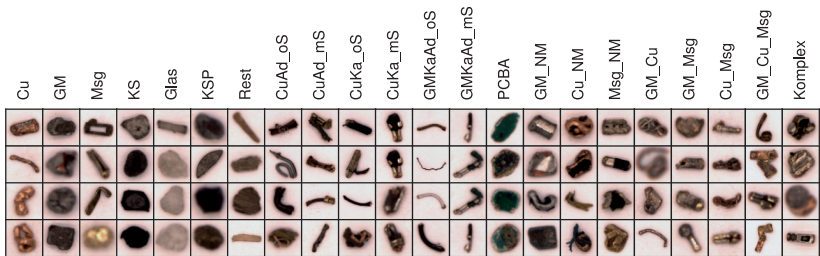


Figure 3.2: Exemplary images (cropped) for the 22 material classes of dataset B.

3.3 Preprocessing

The captured images were then preprocessed to improve the color representation. Preprocessing algorithms were implemented in *Python 3.7* in a *scikit-image* v0.17.1 [25] framework. Firstly, a Gaussian blur filter

with a standard deviation of $\sigma = 0.35$ was applied to reduce small artefacts from demosaicing during image acquisition. Secondly, a white balance based on a white reference image was applied to guarantee a neutral color representation. Thirdly, an spatial illumination correction based on the reference image ensured an uniform illumination.

To extract individual pixels for color classification, each captured image was segmented into background, shadows and particle trough thresholding the gray scale image and improving the segmentation result by a combination of morphological operations. For the color classification, 1,000 pixels per particle as well as 1,000 background pixels (850 pixels of white background and 150 shadows pixels) were randomly selected from each segmented image.

For the image classification with CNNs, quadratic image sections of the particle were extracted (cf. Figure 3.1 and 3.2). A minimum padding of 25 % of the respective particle bounding box size was ensured between particle and the border of the image section. All extracted image sections were rescaled to a size of 64 Pixel \times 64 Pixel.

3.4 Approach I: Feature-Based Classification

The color classification was tested for three color spaces (RGB, HSV and L*a*b*). For dataset A, each color class represents a material class. For dataset B, the color classes BG, Cu, GM, Msg, KS, Glas, KS and Rest ('pure fractions') were defined, as the color of composite material classes (e. g. Cu + Msg) overlapped strongly with the pure fractions. Besides, characteristic colors of PCBA and KaAD (cables and wires) were extracted with a tomek link technique [26, p. 57]. All color classifiers were trained on 4,000 pixels per material class and 8,000 background pixels to avoid overfitting.

For object-based classification, shares of each classified color class as well as 50 dimensionless shape factors were considered, which were generated by normalizing shape measurements by the particle area A or its square root (\sqrt{A}). Shape measurements were extracted from the false-color image by a self-developed Python module named *eidoss*¹.

For both pixel- and object-based classification, six ML classification algorithms, implemented in *scikit-learn* [27], were compared: k nearest

¹<https://github.com/nilskroell/eidos>

neighbor (kNN), decision trees (DT), random forests (RF), AdaBoost (AB), SVM, neural networks (NN). For each model and classification task, a systematic grid search with a k -fold-cross-validation ($k = 5$) on the training dataset was conducted to find the optimal hyperparameter settings. All hyperoptimized models were evaluated on the test dataset.

3.5 Approach II: Deep Learning

For image classification, six different CNN architectures (Figure 3.3) were compared, implemented in *Keras* [28]. All hidden layers were equipped with a ReLU activation function and the final network output was normalized by a softmax function. During training, an Adam [29] optimizer with a learning rate of $\eta = 10^{-3}$ was used. Dataset A and B were trained for 15 and 20 epochs respectively with a mini batch size of 256 instances per iteration. L2-norm regularization ($\lambda = 10^{-3}$) as well as dropout layers with an dropout rate of 0.5 were applied to avoid overfitting, and the classification performance was evaluated by a categorical cross entropy cost function. To reduce the imbalance, downsampling was applied to material classes with more than 2,000 instances. All trained CNNs were evaluated on the test dataset.

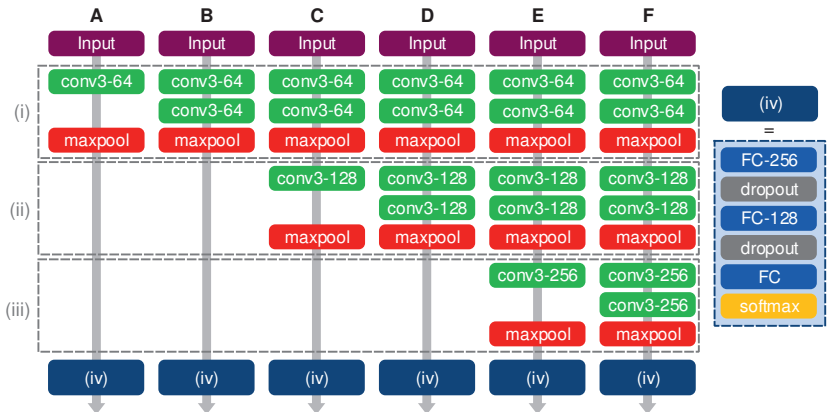


Figure 3.3: Investigated CNN architectures A-F consisting of input (64×64), fully connected (FC- n), 3×3 convolutional (conv3- n), 2×2 max pooling, dropout (0.5) and softmax activation layers; n : number of neurons/feature maps.

4 Results and Discussion

4.1 Approach I: Feature-Based Classification

The color classification of dataset A was possible with similar accuracies of $\geq 79.8\%$ for all models and colorspace, see Table 8.3. On average, the L*a*b* color space with 86.1 % and a SVM model with 87.0 % achieved the best classification results. Because of the superior color classification, an object-based classification with 100 % accuracy could be achieved, e. g. by a plain kNN model, for dataset A.

Table 8.3: Accuracy for color classification of dataset A after hyperparameter optimization; 95 % confidence interval for all accuracies: $\pm 0.08\%$ to $\pm 0.18\%$.

	kNN	DT	RF	AB	SVM	NN	mean
RGB	85.4 %	79.8 %	85.2 %	84.1 %	87.3 %	87.3 %	84.9 %
HSV	82.7 %	82.9 %	86.7 %	85.5 %	87.2 %	84.4 %	84.9 %
L*a*b*	85.3 %	85.2 %	86.9 %	85.6 %	86.7 %	87.1 %	86.1 %
mean	84.5 %	82.6 %	86.3 %	85.1 %	87.0 %	86.3 %	85.3 %

Based on the results above, the L*a*b* color space was chosen for the color classification of dataset B (Table 8.4). Here, the classification was more challenging with achieved test scores between 47.2 % (DT) and 51.1 % (kNN, SVM). As for dataset A, the chosen classification algorithm was of minor importance. Significant differences were found in the prediction time: The fastest model (DT) was with 13 ms/MPixel 714 times faster than the slowest one (AB; 9,278 ms/MPixel). Overall, a NN model showed the best combination of prediction time and accuracy and was thus the basis for subsequent object-based classification.

Test scores of the object-based classification of dataset B, based on extracted false-color shares and ten shape factors with highest feature importance (determined by χ^2 -test on training data) are shown in Table 8.5. Here, the chosen algorithm was of higher importance with achieved accuracies ranging from 66.5 % (kNN) and 78.2 % (RF).

A closer look on the best-performing model (RF) in Figure 4.1a shows that the predictions differ significantly between different material classes. As expected, material classes with a higher number of instances can be significantly better classified than less frequent classes.

Table 8.4: Test scores and prediction time (Δt_{pred}) for color classification of dataset B after hyperparameter optimization in $L^*a^*b^*$ color space; CI: confidence interval.

	kNN	DT	RF	AB	SVM	NN
Accuracy	51.1 %	47.2 %	50.8 %	48.3 %	51.1 %	51.0 %
95 %-CI	$\pm 2.9 \%$	$\pm 0.6 \%$	$\pm 1.8 \%$	$\pm 3.1 \%$	$\pm \mathbf{0.6 \%$	$\pm 0.7 \%$
Δt_{pred} [ms/MPixel]	1,234	13	3,250	9,278	5,589	196
95 %-CI [ms/MPixel]	± 29	$\pm \mathbf{4}$	± 290	± 83	± 24	± 5

Table 8.5: Test scores for object-based classification of dataset B; CI: confidence interval.

	kNN	DT	RF	AB	SVM	NN
Accuracy	66.5 %	67.4 %	78.2 %	77.2 %	72.9 %	72.6 %
95 %-CI	$\pm 3.6 \%$	$\pm 2.0 \%$	$\pm \mathbf{2.0 \%$	$\pm 1.9 \%$	$\pm 1.8 \%$	$\pm 3.4 \%$

On average, pure fractions with an accuracy of 62.5 % could be far better classified than composite fractions (16.7 %). Providing different numbers of shape factors to the model (Figure 4.1b) revealed that the shape factors contributed little to the classification performance, which is primarily achieved by the extracted false color components.

4.2 Approach II: Deep Learning

As for the feature-based approach, the classification of dataset A with CNNs was possible with high accuracy ($> 98.8 \%$), see Table 8.6. Even simple CNN architectures showed a high accuracy.

Table 8.6: Test scores of the CNN architectures A-F for dataset A; CI: confidence interval.

CNN	A	B	C	D	E	F
Accuracy	99.59 %	99.82 %	99.65 %	99.76 %	99.65 %	98.88 %
95 %-CI	$\pm 1.19 \%$	$\pm \mathbf{0.84 \%$	$\pm 1.21 \%$	$\pm 0.87 \%$	$\pm 0.87 \%$	$\pm 1.61 \%$

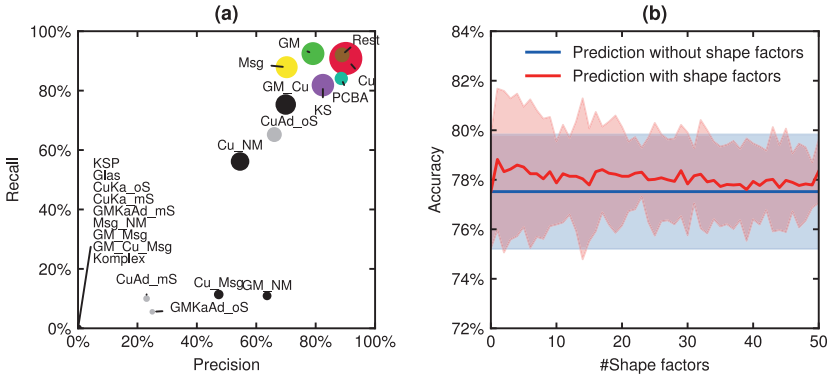


Figure 4.1: Object-based classification of the trained RF model on dataset B. **(a)** Accuracy by material class (circle area \sim #instances), **(b)** influence of shape factors.

In contrast, training CNNs on dataset B turned out to be more challenging. For all six CNN architectures, the error rate did not converge after the initial planned 20 training epochs (see Figure 4.2), resulting in accuracies of $\leq 33.5\%$ when training all 22 material classes.

To optimize the classification result the architecture and training of CNN F was further adapted. Firstly, the number of neurons in the fully connected layer was increased to 1,024, which boosted the resulting accuracy to $48.9\% \pm 3.6\%$. Secondly, the number of training epochs was increased from 20 to 400. As shown in Figure 4.2a, the optimized CNN F* converged after about 200 training epochs and achieved an overall classification accuracy of $80.4\% \pm 4.2\%$. Furthermore, a strong correlation between the number of images in the training set and the classification performance (F1-score) per material class could be confirmed, see Figure 4.2b.

The activations in the six convolutional layers of CNN F* (Figure 4.4) show how CNN F* is able to abstract the particle characteristics with each layer. When considering the given probabilities for each prediction, CNN F* achieved a top-3 score of $94.2\% \pm 2.6\%$, i. e. for 94.2% of the predictions the right material class was in the three material classes with the highest probability.

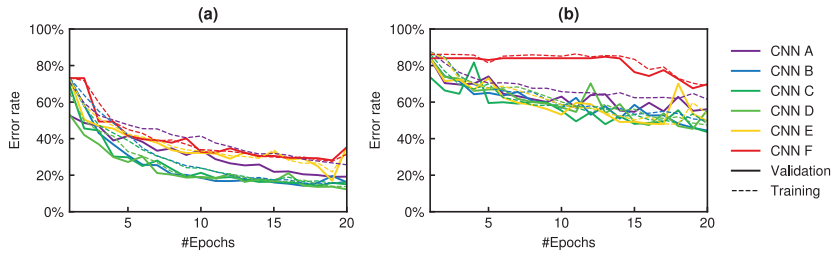


Figure 4.2: Training of CNN A-F for the classification of dataset B (a) with four selected material classes (Cu, GM, Msg, KS) and (b) all 22 material classes.

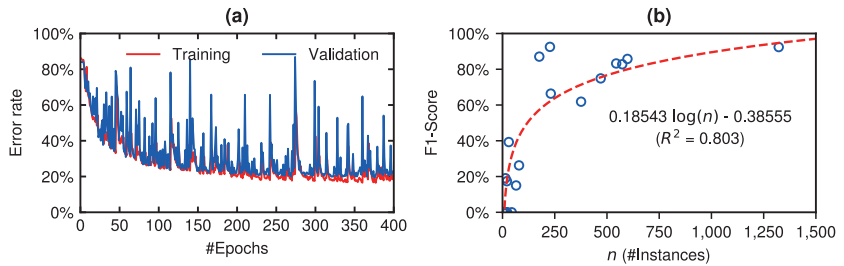


Figure 4.3: (a) Training of CNN F* with dataset B over 400 epochs, (b) correlation between number of instances per material class and F1-score for dataset B.

5 Conclusions and Outlook

In this research, the classification of two datasets A (12,480 images in four material classes) and B (19,498 images in 22 material classes) with a feature-based and deep learning approach was investigated. Our results show that a nearly perfect ($> 99.8\%$) classification for all four material classes of dataset A could be achieved with both approaches, as particles of the different material classes can be clearly differentiated by their color in the RGB images. This indicates that classification results from coarse fraction can in general be transferred to fine fractions. The achieved classification results are better than existing studies on coarse fractions, most likely due to missing influences of different metal alloys or post-consumer waste characteristics, as comminuted standard components were used to create dataset A.

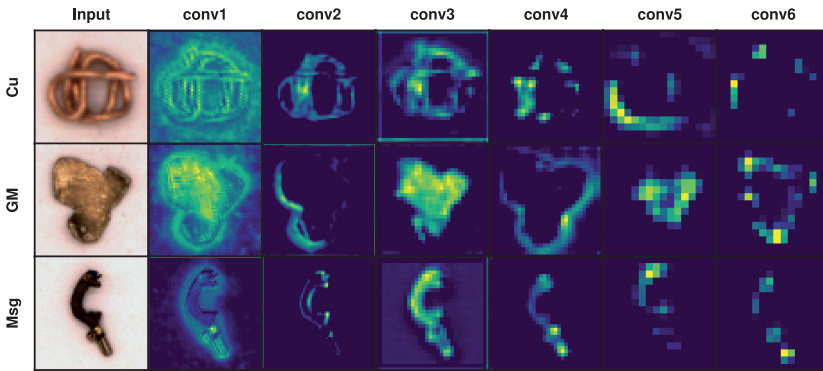


Figure 4.4: Selected convolutional feature maps of CNN F* when classifying dataset B.

In contrast, the classification of dataset B turned out to be quite challenging. Our best feature-based approach consisting of a pixel-based color classification with a neural network in the $L^*a^*b^*$ color space and a subsequent object-based classification with a random forest, achieved a classification accuracy of $78.2\% \pm 2.0\%$. A slightly better classification result was achieved by the deep learning approach with a 14 layer CNN with an accuracy of $80.4\% \pm 4.2\%$ and an top-3 score of $94.2\% \pm 2.6\%$. To improve the classification of dataset B in the future we see potential for future research in the following points:

1. **Extension of dataset B:** For both classification approaches a strong correlation between the number of instances and the classification accuracy of material classes was found. Thus, extending dataset B with more images of less frequent material classes may improve the overall classification performance.
2. **Training of deeper CNNs:** Deeper CNN architectures may be advantageous in adapting to the complex particle characteristics of dataset B. Furthermore, existing CNN architectures or transfer learning may be applied to enhance the classification.

3. **Application of multi sensors:** While this study has focused on RGB images, new distinctive features for classification might be obtained by utilizing multiple sensors, e. g. RGB and induction or X-ray fluorescence/transmission.

References

1. Eurostat, "Abfallaufkommen nach Abfallkategorie, Gefährlichkeit und NACE Rev. 2 Tätigkeit." [Online]. Available: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_wasgen&lang=de
2. B. K. Reck and T. E. Graedel, "Challenges in metal recycling," *Science (New York, N.Y.)*, vol. 337, no. 6095, pp. 690–695, 2012.
3. S. Grimes, J. Donaldson, and G. C. Gomez, "Report on the Environmental Benefits of Recycling." [Online]. Available: https://www.mgg-recycling.com/wp-content/uploads/2013/06/BIR_CO2_report.pdf
4. UNEP – International Resource Panel, "Report 2: Recycling Rates of Metals – A Status Report."
5. V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, and P. R. Shukla, "Global Warming of 1.5°C: An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty."
6. M. Kranert, *Einführung in die Kreislaufwirtschaft*. Wiesbaden: Springer Fachmedien Wiesbaden, 2017.
7. A. Bloodworth and G. Gunn, "The future of the global minerals and metals sector : issues and challenges out to 2050," *Geosciences: BRGM's journal for a sustainable Earth*, vol. 15, pp. 90–97, 2012.
8. European Commission, "A new Circular Economy Action Plan for a cleaner and more competitive Europe." [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0098>
9. H. Martens, *Recyclingtechnik: Fachbuch für Lehre und Praxis*. Heidelberg: Spektrum Akademischer Verlag, 2011.
10. K. Johnen and A. Feil, "Untersuchung von Feinfraktionen aus metallhaltigen Abfallströmen," in *14. Recy & DepoTech-Konferenz : Montanuniversität Leoben, Österreich*.

11. "Technische Anleitung zur Verwertung, Behandlung und sonstigen Entsorgung von Siedlungsabfällen: TA Siedlungsabfall," 14.05.1993.
12. K. Johnen, N. Kroell, and A. Feil, "Entwicklung einer Methodik zur Wertstoffgehaltsbestimmung von feinkörnigen Abfällen," in *15. Recy & DepoTech-Konferenz : Montanuniversität Leoben, Österreich*.
13. A. Feil, T. Pretz, J. Julius, N. Go, M. Bosling, and K. Johnen, "Metal waste," in *Waste*, T. M. Letcher and D. A. Vallerio, Eds. Elsevier, 2019, pp. 211–223.
14. C. Wang, Z. Hu, Q. Pang, and L. Hua, "Research on the classification algorithm and operation parameters optimization of the system for separating non-ferrous metals from end-of-life vehicles based on machine vision," *Waste management (New York, N.Y.)*, vol. 100, pp. 10–17, 2019.
15. T. L. Dang, T. Cao, and Y. Hoshino, "Classification of metal objects using deep neural networks in waste processing line," *International Journal of Innovative Computing, Information and Control*, no. 15, pp. 1901–1912, 2019.
16. 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.
17. H. Karbasi, A. Sanderson, A. Sharifi, and C. Wilson, "Robotic Sorting of Shredded E-waste: Utilizing Deep Learning," in *ICAI '18*, H. R. Arabnia, D. d. La Fuente, E. B. Kozerenko, J. A. Olivas, and F. G. Tinetti, Eds. United States of America: CSREA Press, 2018, pp. 119–123.
18. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." [Online]. Available: <https://arxiv.org/pdf/1512.03385>
19. G. Candiani, N. Picone, L. Pompilio, M. Pepe, and M. Colledani, "Characterization of Fine Metal Particles Derived from Shredded WEEE Using a Hyperspectral Image System: Preliminary Results," *Sensors (Basel, Switzerland)*, vol. 17, no. 5, 2017.
20. P. Barnabé, G. Dislaire, S. Leroy, and E. Pirard, "Design and calibration of a two-camera (visible to near-infrared and short-wave infrared) hyperspectral acquisition system for the characterization of metallic alloys from the recycling industry," *Journal of Electronic Imaging*, vol. 24, no. 6, p. 061115, 2015.
21. J. Huang, T. Pretz, and Z. Bian, "Intelligent solid waste processing using optical sensor based sorting technology," in *2010 3rd International Congress on Image and Signal Processing*. IEEE, 2010, pp. 1657–1661.
22. A. Picon, O. Ghita, P. F. Whelan, and P. M. Iriondo, "Fuzzy Spectral and Spatial Feature Integration for Classification of Nonferrous Materials in

- Hyperspectral Data," *IEEE Transactions on Industrial Informatics*, vol. 5, no. 4, pp. 483–494, 2009.
23. M. Kutila, J. Viitanen, and A. Vattulainen, *Scrap Metal Sorting with Colour Vision and Inductive Sensor Array: Proceedings 28-30 November 2005, Vienna, Austria*. Los Alamitos Calif.: IEEE Computer Society, 2006.
 24. Länderarbeitsgemeinschaft Abfall, "LAGA PN 98: Richtlinie für das Vorgehen bei physikalischen, chemischen und biologischen Untersuchungen im Zusammenhang mit der Verwertung/Beseitigung von Abfällen," 2001.
 25. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 2014.
 26. M. Kubat, *An Introduction to Machine Learning*, 2nd ed. Cham: Springer International Publishing, 2017.
 27. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
 28. F. Chollet, K. Allison, M. Wicke, S. Bileschi, P. Bailey, A. Gibson, J. J. Allaire, F. Rahman, F. Branchaud-Charron, T. Lee, and G. de Mermiesse, "Keras," 2015. [Online]. Available: <https://keras.io>
 29. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." [Online]. Available: <https://arxiv.org/pdf/1412.6980>

Improvement of Thermal Fringe Projection for Fast and Accurate 3D Shape Measurement of Transparent Objects

Martin Landmann^{1,2}, Henri Speck¹, Jan Till Schmieder^{1,2},
Stefan Heist¹, and Gunther Notni^{1,3}

- ¹ Fraunhofer Institute for Applied Optics and Precision Engineering IOF,
Albert-Einstein-Str. 7, 07745 Jena, German
- ² Friedrich Schiller University, Institute of Applied Physics, Abbe Center of
Photonics, Albert-Einstein-Str. 6, 07745 Jena, Germany
- ³ Ilmenau University of Technology, Department of Mechanical Engineering,
Gustav-Kirchhoff-Platz 2, 98693 Ilmenau, Germany

Abstract Structured light sensors for three-dimensional (3D) shape measurements working in the visible up to the short-wave infrared range have been intensively investigated in the last decades. Reliable measurements require diffuse reflection of the projected patterns. However, this condition is strongly limited or not fulfilled for transparent, translucent, shiny, or absorbing materials. Based on the scanning from heating approach, we have developed a two-step optical method for the measurement of such uncooperative objects. In this contribution, we present a simplified and robust projection technique of a focused single thermal fringe. Due to the irradiation of only locally strongly restricted areas, we obtain considerably higher intensities which enables us to reduce the total measurement time to or even below the second range while increasing the measurement accuracy. We show measurement results of non-opaque surfaces of objects made of a single material as well as of metal objects. Our customizable setup is of interest for quality assurance, bin picking, digitization, and many other areas of applications.

Keywords 3D shape measurement, thermal fringe projection, surface, infrared, transparent

1 Introduction

Structured light sensors for 3D shape measurements working in the visible up to the short-wave infrared range have become more and more important in recent years. Diffuse reflection of the projected patterns is required to obtain reliable results. However, this condition is strongly limited or not fulfilled for transparent, translucent, shiny, or absorbing materials. There have been several approaches [1–3] to measure such uncooperative objects. Based on the scanning from heating approach proposed by Pelletier et al. [4], Eren et al. [5], and Meriaudeau et al. [6], we have developed a two-step method [7–11].

First, multi-fringe thermal patterns are projected onto the measurement object. The object absorbs the energy resulting in a temperature on the surface. In a second step, the stereo recording of the thermal radiation, which is re-emitted by the object surface, takes place. To generate a thermal pattern onto the (object) surface, a metal mask with vertical slits and strips working is used. The mask works like a slide. Basically, this setup can be used to measure the shape of, e.g., transparent glasses. However, both the measurement time, ranging from a few tens of seconds to minutes, and the measurement accuracy are improvable. A big disadvantage is that at least 50 % of the incoming intensity is blocked. The envelope irradiance is Gaussian distributed with the consequence that at the horizontal edges of the measurement field, the irradiance drops down to only 28 % of the value in the center.

In this contribution, we present a simplified and robust projection approach of a focused single thermal fringe which is quickly scanned over the object surface. The temperature contrast between adjacent dark (cold) and bright (warm) infrared image areas is the key quantity for the thermal 3D measurement quality. The contrast increase during the irradiation period is counteracted by the thermal diffusion within the object material. Due to the irradiation of only locally strongly restricted areas, we obtain considerably higher irradiances than with multi-fringe projection. This enables us to reduce the irradiation period to only a few milliseconds. Within such short irradiation periods, the effect of thermal diffusion is drastically reduced. Simultaneously to the projection, we record the stereo thermal image. In order to measure the entire surface of the object, a galvanometer scanner sequentially shifts the laser sheet to specified positions at which the irradiation and

recording process is repeated. Although we have a higher number of sequential projections, the total measurement time is substantially reduced to or even below the seconds range while increasing the measurement accuracy. We show measurement results of a transparent glass as well as a metal objects. Our customizable design is attractive for quality inspection, bin picking, digitization and many other applications.

2 Setup of MWIR 3D system based on sequential thermal fringe projection

Instead of projecting several two-dimensional patterns as in the multi-fringe projection technique ([10,11]), we now scan the object surface with a single thermal fringe. The irradiance of this fringe is by factors higher than in our previous setup. To obtain a sufficiently high contrast, the irradiation period can be drastically reduced. Within such short irradiation periods, the effect of thermal diffusion is significantly reduced. A schematic setup of our MWIR 3D system based on the projection of sequential thermal fringes is shown in Fig. 2.1.

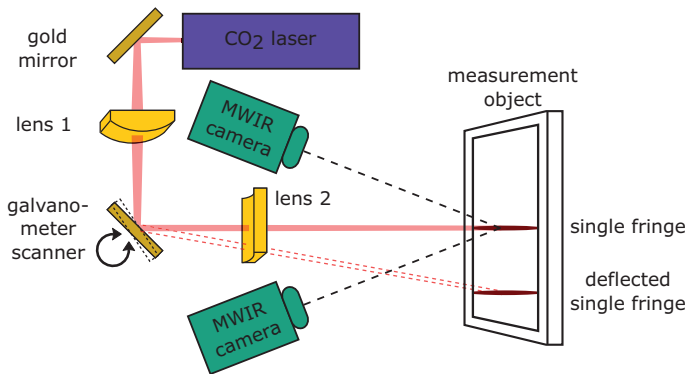


Figure 2.1: Schematic setup of MWIR 3D system based on projection of sequential thermal fringes consisting of projection unit (CO₂ laser, gold mirror, galvano-meter scanner, and horizontal ZnSe cylinder lens 1 and vertical ZnSe cylinder lens 2), measurement object, and two MWIR cameras. The projection unit irradiates the measurement object with a single fringe which is sequentially deflected by the galvano-meter scanner.

The setup shows two cylindrical ZnSe lenses to transform the CO₂ laser beam into a thin laser line sheet. The convex lens 1 is focusing the beam in horizontal direction while the concave lens 2 is used to irradiate the object in vertical direction. We can adjust the fringe width on the object surface by shifting lens 1 between the gold mirror and the galvanometer scanner or by replacing it with another lens of another focal length. Equivalently, we can move or replace lens 2 to adjust the height of the illuminated measurement field. By changing the beam width or height, the irradiance increases or decreases by the same factor.

The galvanometer scanner enables the sequential horizontal scan of the entire measurement field with the single fringe. The scanning range is flexible and can be adjusted to match the horizontal width of the object. The projected laser line is absorbed by the object and re-emits thermal radiation according the temperature distribution which we record with two MWIR cameras.

The measurement procedure shown schematically in Fig. 2.2 is similar to the one for the multi-fringe projection setup [12]. The main difference is the significant reduction of the irradiation period. The measurement starts with the first irradiation of a single fringe which lasts for the irradiation period t_{irr} in the milliseconds range. At the same time, both cameras synchronously start to capture the diffused and re-emitted thermal pattern from the object surface. The recording period t_{rec} is the sum of the camera integration time t_{int} and the camera read out time t_{ro} . Immediately after the end of the irradiation period, the galvanometer scanner moves the fringe to the next position (t_{trans}). We set the irradiation period t_{irr} to the maximum (camera recording period t_{rec} minus translation time t_{trans}). For the total measurement time $t_{\text{meas}} = Nt_{\text{rec}}$, the camera frame rate is the limiting factor.

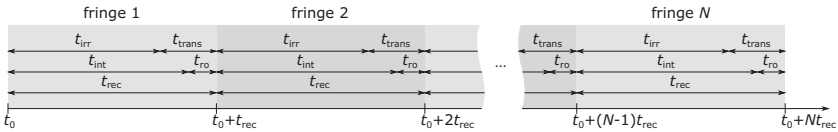


Figure 2.2: Representation of the measurement procedure with sequence length N , irradiation period t_{irr} , translation time t_{trans} , and recording period t_{rec} as the sum of camera integration time t_{int} and the camera read out time t_{ro} .

3 Experimental investigations

After extensive simulation tests with promising results, we designed a laboratory setup according to Fig. 2.1 for the new approach of sequential fringe projection. The distance between the sensor system and the measurement field of $160 \times 128 \text{ mm}^2$ is about 450 mm. The desired vertical width is $w_{\text{ver}} = 120 \text{ mm}$.

With our MWIR 3D system based on sequential fringe projection, we analyzed the 3D result quality dependent on the sequence length N and fringe width w_{hor} . The following results were measured by a recording time of $t_{\text{rec}} = 8 \text{ ms}$. The translation per step took $t_{\text{trans}} = 0.5 \text{ ms}$ and the irradiation period was $t_{\text{irr}} = 7.5 \text{ ms}$. The measurements were conducted on a parallel-sided borosilicate glass plate with a thickness of 3 mm. The determined 3D standard deviation (standard deviation of the 3D points to a fitted plane) refers to an area of $40 \times 40 \text{ mm}^2$.

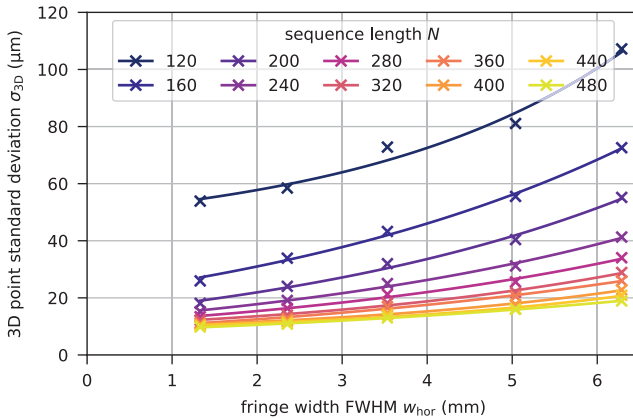


Figure 3.1: 3D point standard deviation σ_{3D} for sequential fringe projection as a function of the horizontal fringe width w_{hor} (1.3 mm...6.2 m) for different sequence lengths N (120...480) and an integration time of $t_{\text{int}} = 7.8 \text{ ms}$.

Figure 3.1 illustrates the 3D standard deviation as a function of the horizontal fringe width w_{hor} and the sequence length N . The measurements were taken for five different fringe widths (from 1.3 mm to 6.2 mm) for ten different sequence lengths (120...480). These differ-

ent horizontal fringe widths were created by changing the position of the horizontal focus lens or replacing them by another one. With a recording time $t_{\text{rec}} = 8 \text{ ms}$, 120 sequentially projected sequences correspond to a measuring time of 0.96 s. The lowest standard deviation was achieved with the smallest fringe width of 1.3 mm. By increasing the number of fringes per measurement, the measured object is scanned more precisely and more images are used for the 3D result, which results in a smaller standard deviation but in an increased measurement time.

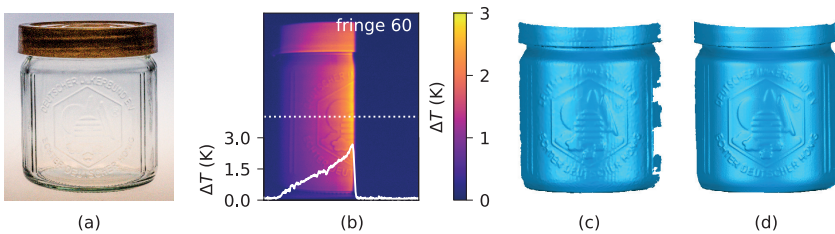


Figure 3.2: 3D measurement example: (a) photograph of a honey jar, (b) left MWIR camera image and temperature profile curve, (c) 3D result for $N = 125$ and $t_{\text{meas}} = 1 \text{ s}$, (d) 3D result for $N = 500$ and $t_{\text{meas}} = 4 \text{ s}$.

4 3D measurement example

We measured a honey jar with a plastic cap with $t_{\text{rec}} = 8 \text{ ms}$ ($t_{\text{irr}} = 7.5 \text{ ms}$, $t_{\text{int}} = 7.8 \text{ ms}$), $w_{\text{hor}} = 1.3 \text{ mm}$, and $N = 10 \dots 500$. Figure 3.2 shows the measurement example with a photograph, a typical MWIR camera image as well as an exemplary 3D results for $N = 125$ and $N = 500$. It is already possible to obtain reasonable 3D results with measurement times of just one second (Fig. 3.2(c)). By increasing the sequence length and the measuring time from 1 s to 4 s, the 3D result can be improved considerably (Fig. 3.2(d)). Depending on the application, the setup can optionally measure very accurately (measurement accuracy about $10 \mu\text{m}$) or very quickly (3D acquisition in less than 1 s). In the case of bin picking, short measurement times are required, the contour of the objects must be determined, and details are mostly unimportant. In contrast, in quality assurance, a high mea-

surement accuracy is required. Moreover, objects with metal surfaces (e.g., metal cup, see Fig. 4.1) could also be measured. The previous measurement system based on multi-fringe projection was not able to provide reliable measurement results (Fig.4.1(c)) due to three disadvantages that occur with metallic materials: (1) high reflectivity of projected irradiance, (2) high thermal conductivity, and (3) low emissivity in MWIR.

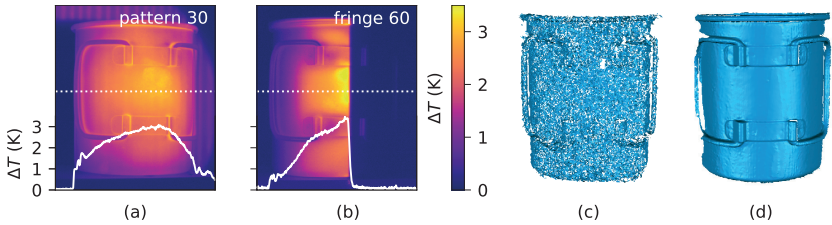


Figure 4.1: 3D measurement example metal cup: (a) left MWIR camera image and temperature profile curve for multi-fringe projection, (b) left camera MWIR image and temperature profile curve for sequential fringe projection, (c) 3D result for $t_{\text{meas}} = 42\text{ s}$ for multi-fringe projection, (d) 3D result for $t_{\text{meas}} = 4\text{ s}$ for sequential fringe projection.

5 Improvement of thermal fringe projection

In additions to the improvements due to higher irradiances, the sequential fringe projection with a galvanometer scanner has several advantages over the multi-fringe technique:

1. The first improvement is the strong increase in uniform radiation. In contrast to the multi-fringe technique (at the edges of the measurement field, the irradiance drops down to only 28 % of the value in the center), with the sequential fringe technique we can project an almost homogenous fringe over the entire measurement field (horizontal edges to 98.8 % of the value in the center).
2. By using only a few lenses and mirrors, we (are able to) remove the blocking mask, which absorb at least 50 % of the incoming intensity, and increase the total intensity on the measurement field.

3. Due to the simple flexible structure of the projection system, individual lenses can be moved or exchanged in order to adapt quickly to different conditions like measuring time, accuracy, or different object sizes.
4. The sequential fringes are not imaged onto the measurement object. This makes it easy to vary the distance between the target and the measuring system. By simply shifting the cylindrical lenses, the fringe width and the vertical extension can be modified.
5. With the knowledge of the measurement object geometry, the range of the galvanometer scanner can be changed to decrease measurement time while maintaining the accuracy. If higher measurement accuracy is required in certain areas, the fringe density can be increased by utilizing special algorithms.

6 Conclusion

In this contribution, we presented the new sequential fringe projection system based on the scanning from heating approach. Instead of projecting several two-dimensional patterns, we propose to scan with a single infrared line across the object. After the general system design, we described the measurement procedure. For this new projection system, we only need few components, such as a radiation source for the generation of the heat fringe, some gold mirrors and lenses for shaping and redirecting the laser beam, and a fast turning mirror for scanning over the object. With the resulting thin fringes, we can drastically reduce the irradiation period to a few milliseconds. The previous system based on multi-fringe technique has measuring times ranging from a few tens of seconds to minutes. We presented different parameters and their influences to the measurement quality. An increase in measurement quality can be achieved both by increasing the number of projected fringes and by using very thin fringes.

We also demonstrated the advantages of the sequential fringe technique by objects with high thermal diffusion. The previous measurement system based on multiple fringes was not able to provide reliable measurement results of a metal cup. With the sequential fringe

technique, we are able to measure such objects in significantly shorter time. The experimental studies showed promising results, as not only the measurement accuracy (about $10\ \mu\text{m}$) is increased, but also the measurement time (less than 1 s) is significantly reduced. With these improvements to our system, this measurement technology based on scanning from heating can be applied, e.g., in industrial applications such as quality assurance or bin picking.

For faster measurements, the camera recording time is the limiting factor. In further investigations, the use of high-speed infrared cameras can exceed this limit. Integrating a rotary table into the setup will enable us to perform 360° measurements.

Funding

This research project (2018 FGR 0098, project acronym “3DWÄRME”) is supported by the Free State of Thuringia, the European Social Fund (ESF) of the European Union, and the Thüringer Aufbaubank (TAB).

References

1. I. Ihrke, K. N. Kutulakos, H. P. A. Lensch, M. Magnor, and W. Heidrich, “Transparent and Specular Object Reconstruction,” *Computer Graphics Forum*, vol. 29, no. 8, pp. 2400–2426, 2010. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-8659.2010.01753.x>
2. F. Mériaudeau, R. Rantson, D. Fofi, and C. Stolz, “Review and comparison of non-conventional imaging systems for three-dimensional digitization of transparent objects,” *J. Electron. Imaging*, vol. 21, no. 2, p. 021105, 2012. [Online]. Available: <https://doi.org/10.1117/1.JEI.21.2.021105>
3. S. Lee and H. Shim, “Skewed stereo time-of-flight camera for translucent object imaging,” *Image Vis Comput.*, vol. 43, pp. 27–38, 2015,]. [Online]. Available: <https://doi.org/10.1016/j.imavis.2015.08.001>
4. J.-F. Pelletier and X. Maldague, “Shape from heating: a two-dimensional approach for shape extraction in infrared images,” *Opt. Eng.*, vol. 36, no. 2, pp. 370–375, 1997. [Online]. Available: <https://doi.org/10.1117/1.601210>
5. G. Eren, O. Aubreton, F. Meriaudeau, L. S. Secades, D. Fofi, A. T. Naskali, F. Truchetet, and A. Ercil, “Scanning from heating: 3D shape estimation of transparent objects from local surface heating,” *Opt.*

- Express*, vol. 17, no. 14, pp. 11457–11468, 2009. [Online]. Available: <https://doi.org/10.1364/OE.17.011457>
6. F. Meriaudeau, L. A. S. Secades, G. Eren, A. Erçil, F. Truchetet, O. Aubreton, and D. Fofi, “3-D scanning of nonopaque objects by means of imaging emitted structured infrared patterns,” *IEEE Trans. Instrum. Meas.*, vol. 59, no. 11, pp. 2898–2906, 2010. [Online]. Available: <https://doi.org/10.1109/TIM.2010.2046694>
 7. A. Brahm, C. Rößler, P. Dietrich, S. Heist, P. Kühmstedt, and G. Notni, “Non-destructive 3D shape measurement of transparent and black objects with thermal fringes,” *Proc. SPIE*, vol. 9868, p. 98680C, 2016. [Online]. Available: <http://dx.doi.org/10.1117/12.2227093>
 8. A. Brahm, E. Reetz, S. Schindwolf, M. Correns, P. Kühmstedt, and G. Notni, “3D shape measurement with thermal pattern projection,” *Adv. Opt. Technol.*, vol. 5, no. 5-6, pp. 405–413, 2017. [Online]. Available: <https://www.degruyter.com/view/j/aot.2016.5.issue-5-6/aot-2016-0052/aot-2016-0052.xml>
 9. A. Brahm, S. Schindwolf, M. Landmann, S. Heist, P. Kühmstedt, and G. Notni, “3D shape measurement of glass and transparent plastics with a thermal 3D system in the mid-wave infrared,” *Proc. SPIE*, vol. 10667, p. 106670D, 2018. [Online]. Available: <https://doi.org/10.1117/12.2304777>
 10. M. Landmann, S. Heist, A. Brahm, S. Schindwolf, P. Kühmstedt, and G. Notni, “3D shape measurement by thermal fringe projection: optimization of infrared (IR) projection parameters,” *Proc. SPIE*, vol. 10667, p. 1066704, 2018. [Online]. Available: <https://doi.org/10.1117/12.2304761>
 11. M. Landmann, S. Heist, P. Kühmstedt, and G. Notni, “3D shape from thermal patterns: investigation of projection parameters in simulation and experiment,” *Proc. SPIE*, vol. 11056, p. 1105615, 2019.
 12. M. Landmann, S. Heist, P. Dietrich, H. Speck, P. Kühmstedt, A. Tünnermann, and G. Notni, “3D shape measurement of objects with uncooperative surface by projection of aperiodic thermal patterns in simulation and experiment,” *Opt. Eng.*, vol. 59, no. 9, p. 094107, 2020.

Measurement of the coefficient of linear thermal expansion based on subjective laser speckle patterns

Alexander Spaett¹ and Bernhard G. Zagar¹

Johannes Kepler University Linz, Institute for Measurement Technology,
Altenberger Straße 69, A-4040 Linz

Abstract We present an improved processing of laser speckle intensity signals aimed at attaining spatial displacement resolutions well in the sub-pixel domain, which is necessary for measuring the minute displacements caused by thermal expansion of metallic specimens. The signal processing revolves around the cross power spectral density and the coherence function of two discrete signals, which allows for a continuously varying displacement result as opposed to a discretized, if determined in the spatial domain via the application of the cross-correlation function. The presented technique is then applied to estimate the coefficient of linear thermal expansion (CLTE) of a brass sample.

Keywords Coefficient of linear thermal expansion, subjective laser speckles, strain measurement, displacement estimator, unbiased minimum variance estimator

1 Introduction

Laser speckle patterns are caused by the interference of coherent light reflected off an optically rough surface. When it comes to speckle imaging, one differentiates between subjective and objective laser speckle patterns. In the former an optical setup is used and the speckle pattern in the image plane is evaluated. The latter describes the speckle pattern in the object plane that is captured directly by a camera [1]. Successively captured one-dimensional speckle pattern can be seen in Fig. 4.1.

Although laser speckles are often an unwanted source of noise they can also be utilized in different applications, since they represent something similar to a fingerprint of the illuminated surface. For instance, the measurement of the shift of the imaged laser speckle pattern may be used to determine the surface strain of a sample. Knowledge about the surface strain, in turn, allows to deduce the coefficient of linear thermal expansion (CLTE).

The presented contactless method provides the possibility to measure mechanical properties of samples for which common approaches, like strain gauges, are not applicable. As an example one could name the measurement of strain of materials which are small in at least one dimension, e.g. natural fibres.

2 Speckle shift in the image and its relation to the CLTE

In order to give a better understanding of the digital signal processing applied, a theoretical model linking the speckle shift in the image with the CLTE of the sample will be reviewed. However, the following considerations will be limited to telecentric imaging systems. A broader treatment can be found in [2].

When a telecentric imaging system is used to observe the speckles, the speckle shift captured by a camera is given by the following equation first derived by Yamaguchi in 1981 [3]

$$A_y = \underbrace{\frac{a_x}{M} \frac{\Delta L}{L_S} l_{Sx} l_{Sy}}_{A_I} - \underbrace{\frac{a_y}{M} \left[1 + \frac{\Delta L}{L_S} (1 - l_{Sy}^2) \right]}_{A_{II}} + \underbrace{\frac{a_z}{M} \frac{\Delta L}{L_S} l_{Sx} l_{Sz}}_{A_{III}} + \underbrace{\frac{\Delta L}{M} \cdot (\varepsilon_{xy} l_{Sx} + \varepsilon_{yy} l_{Sy} - g(\Omega))}_{A_{IV}} + E. \quad (2.1)$$

A_y describes the resulting shift of the laser speckle pattern in the image plane in y_I -direction and ε is the desired strain tensor of the sample. As can be seen, both the components ε_{xy} and ε_{yy} of the strain tensor can - in theory - be observed directly. Instead, in order to avoid multiple sources contributing to A_y , the shift a_y of the diffracting surface will be estimated. The strain in y -direction ε_{yy} can then be deduced

from this quantity. The general setup leading to Eqn. 2.1 is depicted in Fig. 2.1.

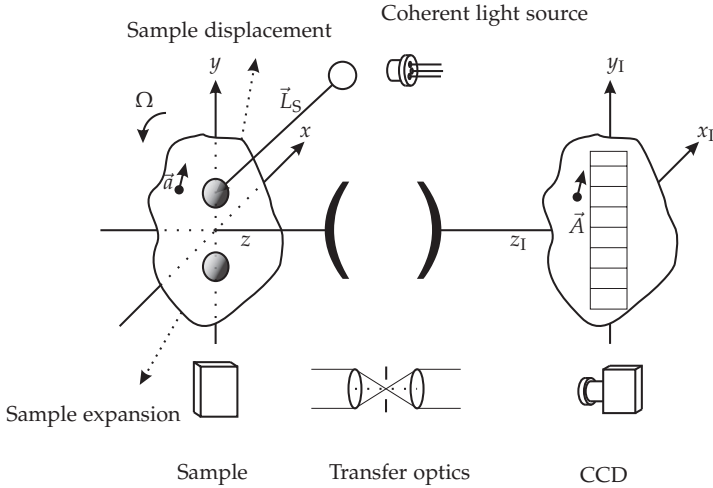


Figure 2.1: General geometrical setup for modelling the pattern shift in the image plane.

The sample lying in the x - y -plane on the left is illuminated by a coherent light source. The distance L_S then describes the distance between the ideal point source and the illuminated spot on the surface and therefore the radius of the curvature of the laser beam. The vector pointing from the source to the spot on the surface can be described by $\vec{L}_S = L_S \cdot \vec{l}_S$ with $\vec{l}_S = [l_{Sx} \ l_{Sy} \ l_{Sz}]^T$ denoting the unit vector, representing the direction of the laser beam. The demagnification factor M , given by the optics used, as well as the distance between the focal plane of the telecentric optics and the sample ΔL influence the speckle shift in the image plane.

Tilting the sample also translates to a shift in the image. This is stated by the function g of the rotation vector Ω . The additional error term E summarizes non ideal or not sufficiently well known dimensional parameters, like temperature dependant expansion of the sensor and overall build-up, causing speckle pattern shifts not distinguishable from a true strain contribution [4].

When the strain estimation is solely based on a_y , it becomes evident that any term in Eqn. 2.1 but A_{II} is negatively influencing the accuracy of the estimate and hence should be minimized. By placing the sample near or ideally in the focus, both ΔL and $\frac{\Delta L}{L_s}$ become small, resulting in the contributions of A_I , A_{III} as well as A_{IV} to be negligible. The impact of A_I and partially A_{IV} can be reduced further by choosing the position of the coherent light source, which in our case is a laser diode (LD) with collimating optics, carefully. When the LD is placed at the same height, meaning having the same y -component, as the spot on the surface, then for the parameter l_{Sy} it follows $l_{Sy} = 0$.

Having taken all measures mentioned above, Eqn. 2.1 reduces to

$$A_y = -\frac{a_y}{M} + E. \quad (2.2)$$

For the calculation of the strain and the CLTE two vertically aligned laser spots are utilized. When $A_{y,Upper}$ is the shift that the upper spot experiences and $A_{y,Lower}$ the one of the lower spot, then the engineering strain can be described as

$$\varepsilon = \frac{-M \cdot (A_{y,Upper} - A_{y,Lower})}{d_{ul}}, \quad (2.3)$$

where d_{ul} is the initial distance between the two spots. Here, the error term E has been neglected.

The instantaneous CLTE is defined as

$$\alpha_L = \frac{1}{l_{sample}} \frac{dl_{sample}}{dT}, \quad (2.4)$$

where T is the temperature in K and l_{sample} denoting the initial length of the sample in m [5]. Since dealing with homogenous materials which are unconstrained and have zero stress components, the thermal strain is the only strain to be observed. Therefore, we can simply rewrite the equation above in order to get the following relationship between the estimated strain and the CLTE

$$\alpha_L = \frac{d\varepsilon}{dT}. \quad (2.5)$$

Applying the results from equation 2.3 it follows that the CLTE is given by

$$\alpha_L = \frac{d\left(\frac{-M \cdot (A_{y,Upper} - A_{y,Lower})}{d_{ul}}\right)}{dT}. \quad (2.6)$$

3 Measurement setup

The measurement setup, depicted in Fig. 3.1, consists of four components which correspond to the ones presented in the general setup in Fig. 2.1. Two LDs with maximum power outputs of 5mW serve as coherent light sources. Each diode illuminates a small circular region on the sample. One of these regions is located in the upper and one in the lower half of the sample. The upper LD has a center wavelength of 657 nm whereas the center wavelength of the lower one was determined to be at 655 nm. The measurement for the center wavelength has been carried out with the help of a Hamamatsu TM-CCD series mini-spectrometer which has a wavelength resolution of 1 nm.

The two illuminated spots are imaged through a telecentric optical system with demagnification $M = 1.000$. Therefore, according to equation 2.3, every shift of the illuminated surface directly translates to the same shift in the image plane although in opposing direction. The telecentric imaging system also ensures that the minimum speckle size in the image plane is large enough to not cause spatial aliasing effects. The minimum speckle size is defined by the smallest optical aperture and the laser wavelength.

The camera is a 3000 pixel line scan CCD camera with a pixel pitch, the distance between the center of two of adjacent pixel, of $7 \mu\text{m}$. The CCD is driven directly by a Linux system running the `preempt_rt` patch, enabling real-time functionalities, for Ubuntu 20.04 in order to minimize variations of the exposure time which can also lead to falsely detected shifts of the patterns. Variations of $\Delta t_{\text{exposure}} < 1 \mu\text{s}$ have been achieved. This limit is only exceeded in very few occasions.

The sample, depicted on the right side of Fig. 3.1, is mounted on a 2-dimensional axis so that positioning of the surface in the focus of the optics can be achieved. In order to measure the samples temperature a Pt-1000, placed in a bore in the sample, is utilized.

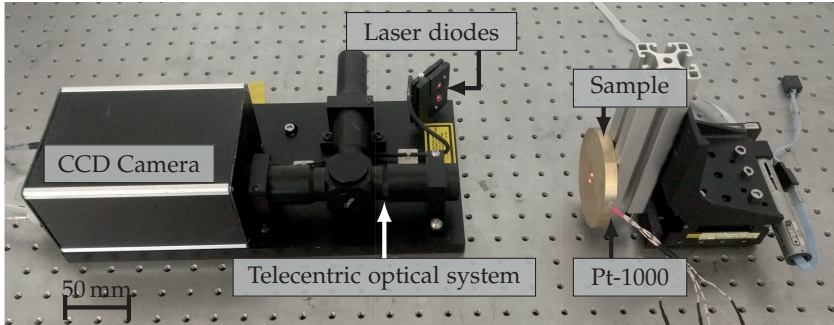


Figure 3.1: Measurement setup using telecentric imaging.

4 Signal processing and results

In this section the signal processing applied in order to estimate the CLTE will be discussed.

In Section 2 it becomes evident, that for the CLTE estimation high resolution, in the range of 100 nm, of the speckle pattern shift is crucial. Hence, the main goal of the applied signal processing techniques is to determine the pattern shift as accurately as possible. The routine described below is executed for both spots independently. Therefore, a region of interest (ROI) for each spot is chosen. The ROI for the upper spot, which is a subset of one line captured by the CCD, is depicted in Fig. 4.1.

In order to accurately estimate the pattern shift, the cross power spectral density (CPSD) $S_{xy}(f)$ as well as the coherence function $\rho(f)$ are utilized. According to the Fourier shift theorem, the argument of the complex valued CPSD encodes the space-lag of the image. The CPSD can be calculated by the following equation

$$S_{xy}(f) = E \{ \mathcal{F} \{ x(m) \} \mathcal{F} \{ x(m+i) \}^* \}, \quad (4.1)$$

with the expected value $E \{ \dots \}$ and $\mathcal{F} \{ \dots \}$ denoting the (fast) Fourier transform (FFT) of two different lines $x(m)$ and $x(m+i)$, captured by the CCD camera at two different points in time m and $m+i$ [6]. In order to minimize the variance of the estimate of the CPSD, Welch's

method is applied. This method includes averaging over multiple periodograms of potentially overlapping sub-signals. In Fig. 4.1 this concept of splitting the signals into i sub-signals of length N_{FFT} and N_{OL} overlapping pixels is shown. However, there is a trade-off to be made. The more sub-signals i are used, the lower the variance of the resulting coefficients is. In turn, since the signal only consists of a finite amount of sample points, the number of frequencies that are calculated are reduced, too [7].

It follows, that the spatial shift A_y of either of the two spots based on the phase $\theta = \arg(S_{xy})$ of the CPSD is given by

$$A_y = d_{\text{pixel}} \cdot \frac{\theta N_{\text{FFT}}}{2\pi f k}. \quad (4.2)$$

This can also be interpreted as a spatial group-delay, which can be assumed to be constant for the application to shifts in speckle patterns. Therefore, higher spatial frequencies also experience a larger phase-delay. For the evaluation of the space-lag based on the k -th spectral line, the pixel pitch d_{pixel} and the number of points used for the calculation of the FFT N_{FFT} are needed.

In order to improve the estimation, weighting of the spectral lines of the CPSD is introduced. Each spectral line of the CPSD is weighted with its corresponding squared coherence value $0 \leq |\rho(f_k)|^2 \leq 1$. The coherence function between two signals can be calculated by the relationship

$$\rho(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f) \cdot S_{yy}(f)}}, \quad (4.3)$$

where $S_{xx}(f)$ and $S_{yy}(f)$ are the power spectral densities for line $x(m)$ and $x(m+i)$, respectively. The coherence function is a measure of the degree to which both lines in question are related by a linear time-invariant transform [6]. Especially in the case at hand, one could alternatively view a spectral line with low coherence as having less cause and effect but more noise to determine the actual phase value. Hence, additionally to weighting, a spectral line with a coherence value undercutting a predefined minimum, here $\rho_{\text{min}} = 0.99$, is not being considered further for the calculation of the shift.

Further improvements can be made, if for the calculation of the shift $A_y(m, m + 1)$ between line m and line $m + 1$ not only these two lines but all lines $m - j, m - j + 1, \dots, m, \dots, m + j - 1, m + j$ are considered. In Fig. 4.1 lines $m - 2, \dots, m + 3$ are depicted. Since the time between capturing each of these lines is comparably small it is assumed that the associated shift between the lines is (i) linear as a function of time and (ii) sufficiently small. The pattern shifts $A_y(m, m - j) \dots A_y(m, m + j)$

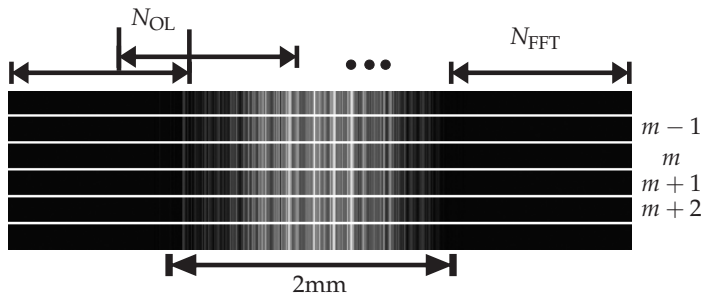


Figure 4.1: Speckle pattern in the ROI of the upper spot, taken at different points in time $m + j$.

can then be used to estimate a polynomial of first order describing the shift over time. This in turn, evaluated at t_{m+1} results in the final estimate of the shift between line m and $m + 1$. This concept is visualized in Fig. 4.2. The empty circles (left) represent the pattern shifts $A_y(m, m - j) \dots A_y(m, m + j)$ whereas the dashed line is the polynomial of first order. The resulting shift then transfers to the total shift over time depicted on the right side.

Applying the previously described signal processing routine, the results in Fig. 4.3 have been obtained. A brass sample was heated up to about 110°C using a heat-gun. Then, as can be seen in the left graph, the sample slowly cooled down to about room temperature. As expected, the specimen shrinks at approximately the theoretical rate, depicted in grey. However, as can be seen, the estimated shift is always lower than the theoretical shift. The reason for this lies within an observable bias of the estimator, which still needs further research and potentially the introduction of a more precise model of the influence of the CCD camera. Despite the bias, the estimated averaged CLTE $\alpha_{L,25-110}$

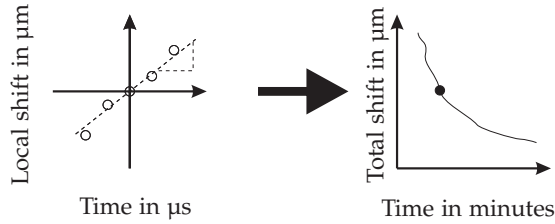


Figure 4.2: Local pattern shifts $A_y(m, m + j)$ translate to a point in the total pattern shift when evaluation of the linear polynomial at $t_{m, m+1}$ is carried out.

over the whole temperature range, based on the linearised total shift, seen on the right of Fig. 4.3, is in good agreement with the CLTE found in literature. For a smaller temperature range of 25°C – 50°C a result closer to the theoretical one has been obtained. The values are as follows: $\alpha_{L, \text{Theo}} = 18.7 \text{ ppm}/^\circ\text{C}$ [8], $\alpha_{L, 25-110} = 17.8 \text{ ppm}/^\circ\text{C}$, $\alpha_{L, 25-50} = 18.2 \text{ ppm}/^\circ\text{C}$.

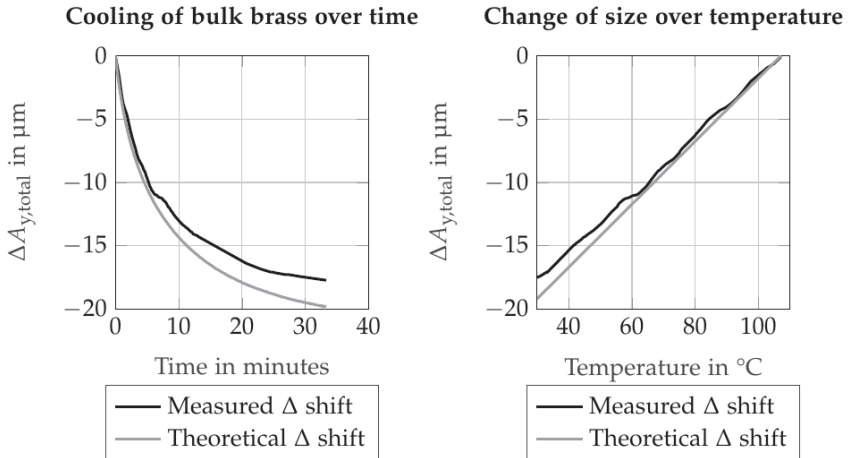


Figure 4.3: Comparison of the estimated difference $\Delta A_y = -A_{y, \text{Upper}} + A_{y, \text{Lower}}$ based on measurements vs. theoretical values for a bulk brass specimen. ΔA_y is essential for the calculation of the CLTE.

5 Conclusions

We have shown, that it's possible to get good estimates for the CLTE of metals evaluating the CPSD weighted by the cause-and-effect measuring coherency function. The described signal processing routine leads to acceptable results presented for the case of bulk brass and can be extended to thin specimen as well. It has been observed that with the presented estimator resulting shifts are always smaller than the ones obtained by literature based calculations. Further research is needed in order to describe this phenomenon accurately and also improve the estimate.

References

1. B. G. Zagar and C. Kargel, "A laser-based strain sensor with optical preprocessing," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 1, pp. 97–101, 1999.
2. R. S. Sirohi, *Speckle metrology*. New York: Dekker, 1993.
3. I. Yamaguchi, "A laser-speckle strain gauge," *Journal of Physics E: Scientific Instruments*, vol. 14, no. 11, pp. 1270–1273, 1981.
4. B. Zagar and J. Lettner, "A discussion of sources of error in laser-speckle based systems," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 8306, pp. 5–, 09 2011.
5. Tipler, *Physik*. Springer Berlin Heidelberg, 2015.
6. W. A. Gardner, *Introduction to random processes: With applications to signals and systems*, 2nd ed. New York: McGraw-Hill, 1990.
7. P. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
8. R. Corruccini and J. Gniewek, *Thermal Expansion of Technical Solids at Low Temperatures: A Compilation from the Literature*, ser. Monograph 29 Series. U.S. Department of Commerce, National Bureau of Standards, 1961.

Vessel extraction from breast MRI

Marco Gierlinger, Dinah Brandner, and Bernhard G. Zagar

Johannes Kepler University Linz, Institute for Measurement Technology,
Altenberger Straße 69, AT-4040 Linz

Abstract We present an extension of the previous work, where a multi-seed region growing (MSRG) algorithm was shown, that extracts segments from breast MRI. The algorithm of our extended work filters elongated segments from the segments derived by the MSRG algorithm to obtain vessel-like structures. This filter is a skeletonization-like algorithm that collects useful information about the segments' thickness, length, etc. A model is shown that scans through the solution space of the MSRG algorithm by adjusting its parameters and by providing shape information for the filter. We further elaborate on the usefulness of the algorithm to assist medical experts in their diagnosis of diseases relevant to angiography.

Keywords Feature extraction, breast MRI, region-based, image segmentation

1 Introduction

Magnet resonance (MR) angiography is a diagnostic tool to depict vessels, e.g. blood vessels. Also, potential malignant masses like cancer can be analyzed, since the spread of cancer is supported by excreting angiogenesis factor to prompt vessel growth into the mass to provide nutrients and oxygen [1]. Information about location, size and morphology of vessels may provide clinicians with useful information before surgery during e.g. a neoadjuvant therapy. Blood vessels in breast MRI are determined subjectively by radiologists, which is considered as the gold standard as research shows [2,3]. Rarely accessible data-sets and their expensive annotation by experts hinder the progress of vessel extraction in medical fields with deep-learning [4]. In recent years,

many researchers were motivated to provide clinicians with useful algorithms for vessel segmentation for medical diagnostics [5].

In previous work, we developed a fast multi-seed region growing (MSRG) algorithm [6], that is capable of extracting homogeneous regions even in images exhibiting noise artefacts. While the algorithm is still in active development, it has shown promising results on which we rely on in this current contribution. The algorithm generates a finite two-dimensional solution space of extracted homogeneous regions from a given input image. We extend the previous work with a skeletonization-like algorithm that is capable of filtering elongated segments from this solution space.

In Section 2, the algorithm of the previous work is explained in detail. The subsequent sections will describe the skeletonization-like algorithm and show the results of allegedly extracted vessels.

2 Previous Work

The multi-seed region growing (MSRG) algorithm is described for n -bit RGB images in $\mathbb{N}_0^{w \times h \times 3}$ as follows: Multiple seeds are employed and aligned as grid with a user-defined width u and height v with $u, v \in \mathbb{N}$, $u \leq w$ and $v \leq h$ as schematically depicted in red ($u = v = 3$) on the blue pane of Fig. 2.1. The algorithm walks through pixel space in an 8-adjacency flood-fill manner independently for each of the k seeds as follows: Let $\vec{p} \in \mathbb{N}_0^5 = (p_0, p_1, p_R, p_G, p_B)$ be a pixel, where (p_0, p_1) is its position in pixel space and let $I_R(\vec{p}) = p_R$, $I_G(\vec{p}) = p_G$, and $I_B(\vec{p}) = p_B$ be its intensity values. A bucket queue is used to hold data objects (\vec{p}, q) , where $q \in \mathbb{N}_0$ with $q < 2^n$ is the bucket's index. Initially, the queue is filled with the seed only, which is a single pixel only. An iteration i is initiated by polling a bucket B_i . $\forall \vec{a} \in B_i \forall \vec{b} \in N'(\vec{a})$, a cost function $\delta_f : (\vec{a}, \vec{b}) \mapsto \{r \in \mathbb{N}_0 \mid r < 2^n\}$ is applied, where $N'(\vec{a})$ denotes all non-visited neighbors of \vec{a} . Cost function δ_f is defined as:

$$\delta_f(\vec{a}, \vec{b}) = \left\lfloor \sqrt{\frac{(I_R(\vec{a}) - I_R(\vec{b}))^2 + (I_G(\vec{a}) - I_G(\vec{b}))^2 + (I_B(\vec{a}) - I_B(\vec{b}))^2}{3}} \right\rfloor. \quad (2.1)$$

At the end of an iteration, each result is added to the queue as $(\vec{b}, \delta_f(\vec{a}, \vec{b}))$.

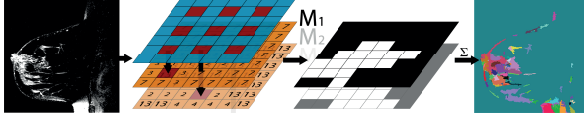


Figure 2.1: MSR applied to a single slice of breast MRI dataset [7].

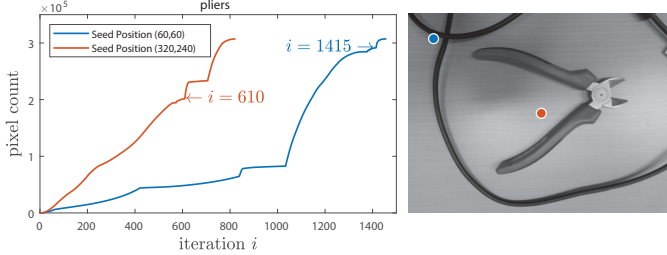


Figure 2.2: Left: Cumulative frequency of number of assigned pixels m_i on 'pliers' image with $(w, h) = (640, 480)$ and largest step m_{\max} is found at highlighted iteration i . Right: Image with arbitrary seed positions indicated.

2.1 Heuristic

Additionally, the number of newly assigned pixels m_i is tracked for each iteration i . A map $M_s \in \mathbb{N}_0^{w \times h}$, drawn as an orange pane in Fig. 2.1, is used for the s^{th} of k seeds and every $r_{a_0, a_1} \in M_s$, where (a_0, a_1) is the position of \vec{a} , is set to $m_{\max}(i) = \max(m_0, m_1, \dots, m_i)$ at iteration i . Finally, when the queue is empty, i.e. every pixel was visited, M_s is converted into a binary map by

$$r \in M_s = \begin{cases} 0 & r < m_{\max}(i) \\ 1 & \text{otherwise} \end{cases} \quad (2.2)$$

Then, all k binary maps are added up elementwise to $\sum_{s=1}^k M_s$, and normalized to the range from zero to $2^n - 1$ to suppress regions that

occur less often than others. An example is highlighted in Fig. 2.1 (right) with random colors assigned for regions containing the same numerical value. Growth is depicted in Fig. 2.2 and Fig. 2.3 for two arbitrary seed positions.

Furthermore, an image quantization factor $g \in \mathbb{R}$ with $0 < g \leq 1$ is used to decrease the image's bit depth as a preprocessing step. A user can tune the grid size (u, v) , i.e. the density of dispersed seeds, and image quantization factor g until a suitable solution is found as shown in Fig. 2.4.

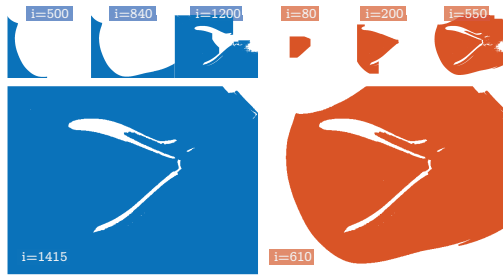


Figure 2.3: Growth of blue seed at $(60, 60)$ and orange seed at $(320, 240)$ on 'pliers'.

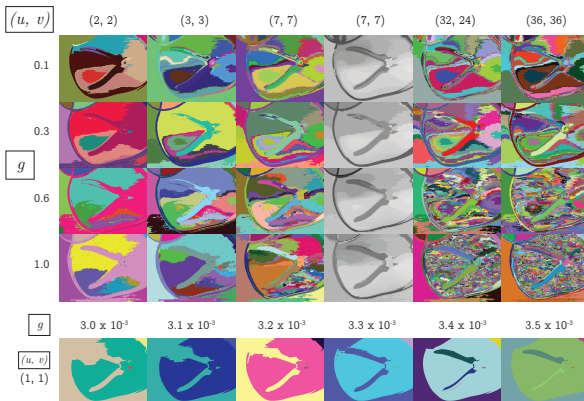


Figure 2.4: Solution space of MSRG [6].

3 Proposed Method

In this section we initially detail an algorithm that filters elongated segments from the result set of the multi-seed region growing (MSRG) algorithm. Furthermore, in the second subsection, a model is depicted to show how these two algorithms interact to obtain vessel-like structures from an input breast MRI stack. The stack is defined as follows: Let S_i be the i^{th} image on the breast MRI stack, that is an 8-bit grayscale image, we define C_i as an RGB image, that is composed of S_i as red channel, S_{i+1} as green channel and S_{i+2} as blue channel. For example, the top of the second stack from left (C_{24}) in Fig. 3.1 shows such an 8-bit 256×256 RGB image as one of 34 possible slices.

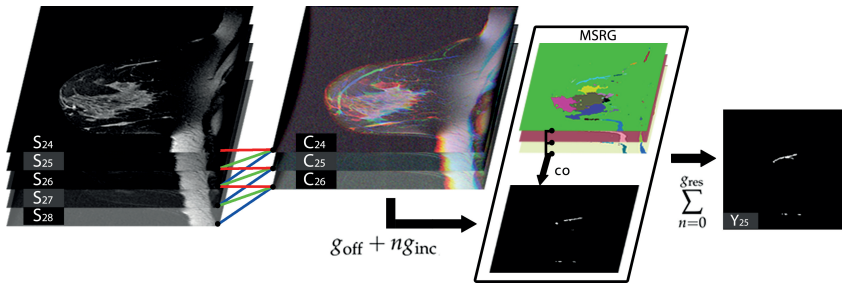


Figure 3.1: From left to right: Five of a set of 36 Breast MRI slices of a single session \rightarrow RGB image composed by three grayscale images from the left stack \rightarrow Filtered MSRG algorithm $\text{fi}(C, g_{\text{off}} + n \cdot g_{\text{inc}}) \rightarrow$ Result stack showing a single result slice.

3.1 Algorithm

Elongated vessel-like segments are filtered out from the set of segments derived by the MSRG. We apply a skeletonization-like algorithm $\text{sk}(\text{msrg}(C, g, u, v)) = M_{\text{sk}}$, where C is an arbitrary RGB image and g, u, v the other MSRG input parameters as described in the previous section. Result M_{sk} is a binary map containing the filtered segments. The algorithm $\text{sk}(\text{msrg}(C, g, u, v))$ is described for a single segment of the MSRG result set as follows:

Initially, we select a random pixel \vec{p} from given segment A and 8-adjacently flood-fill until every pixel has been 'filled'. Let

$N(\vec{p})$ be the set of adjacent filled pixels of \vec{p} in pixel space of A . Each iteration, we prioritize from the set of potential new pixels $N(A) = \{r \in A \mid \neg \text{isfilled}(r) \wedge N(r) \neq \emptyset\}$ the subset $N_{\max}(A) \subseteq N(A)$, defined as the largest group of adjacent pixels as highlighted in Fig. 3.2 left.

This rule leads to a fast method such that the segment is flood-filled nearly homogeneously along an elongated segment as shown in Fig. 3.2 right. Each iteration and until every pixel has been examined, we derive an updated set of $N_{\max}(A)$. We use this set for statistical analysis: E.g., $|N_{\max}(A)|$ gives a rough estimation of the thickness of an elongated shape. This statistical information is useful to filter out segments that are too short, too thick, or have a too large variance in thickness or intensity along the potential vessel or along its cross section. With this filter, the number of parameters increases but the parameters could be adjusted by medical experts to fit the realistic requirements of the vessels in question.

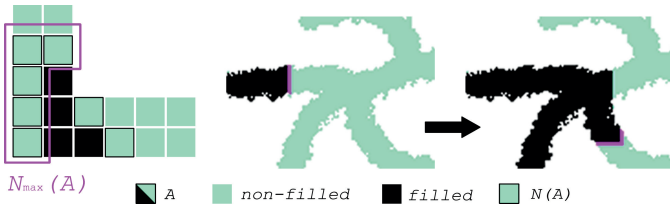


Figure 3.2: A Flood-fill algorithm that promotes nearly homogeneous flooding by prioritizing the largest adjacent group $N_{\max}(A) \subseteq N(A)$.

3.2 Model

We generate with the MSRG a one-dimensional solution space: The seed grid size (u, v) is set to $(4, 4)$, i.e. we select a column as shown in Fig. 2.4. These values are selected based on user experience with the MSRG. We define within this subsection the filtered MSRG algorithm $\text{fi}(C, g) = \text{sk}(\text{msrg}(C, g, 4, 4))$, where $\text{sk}(\text{msrg}(C, g, 4, 4))$ denotes the filter algorithm from the previous subsection. The algorithm $\text{fi}(C, g)$ is applied to RGB (3 channels) image $C \in \mathbb{N}_0^{w \times h \times 3}$ with image width w and image height h and results in a binary map $M_{\text{fi}} \in \mathbb{N}_0^{w \times h}$. The

image quantization factor $g \in \mathbb{R}$ with $0 < g \leq 1$ is now the only parameter that influences the result.

To keep the computation time low, we do not iterate over all possible values of the image quantization factor g . We limit the range from g_{off} to $g_{\text{off}} + g_{\text{inc}} \cdot g_{\text{res}}$, where g_{off} denotes the lower bound, g_{res} the resolution, and $g_{\text{off}} + g_{\text{inc}} \cdot g_{\text{res}}$ the upper bound of the dimension. Again, the values are selected based on user experience of the MSRG and should be selected such that the desired features are found within these bounds.

The model is schematically depicted in Fig. 3.1. The results are binary images with elongated structures that potentially reveal vessels. Formally, we generate a result stack (see Fig. 3.1 right) for all $j \in \mathbb{N}$ with $1 < j \leq (\text{MRI stack size} - 3)$, defined as:

$$Y_j = \sum_{n=0}^{g_{\text{res}}} \text{co}(C_{j-1}, C_j, C_{j+1}, g_{\text{off}} + n \cdot g_{\text{inc}}) \quad (3.1)$$

where $\text{co}(R_1, R_2, R_3, g)$ returns a single binary map that is composed of the union of the two adjacent slice pairs. Formally,

$$\text{co}(R_1, R_2, R_3, g) = \text{fi}(R_1, g) \wedge \text{fi}(R_2, g) \vee \text{fi}(R_2, g) \wedge \text{fi}(R_3, g) \quad (3.2)$$

where \wedge and \vee denote element-wise binary operators. Finally, as a post processing step, we set

$$r \in Y_j = \begin{cases} 0 & r < t \\ 1 & \text{otherwise} \end{cases} \quad (3.3)$$

where t is a user-defined threshold.

With this technique, it is possible to differ the contours of the skin from elongated segments that are within the region of interest. We assume that the color gradient more likely varies along the cross section as shown in Fig. 3.3. This can be detected by the algorithm of the previous subsection by statistical analysis of $N_{\text{max}}(A)$.

4 Results and Evaluation

We apply the proposed method of the previous section to two different breast MRI stacks that were acquired from the same patient (and same

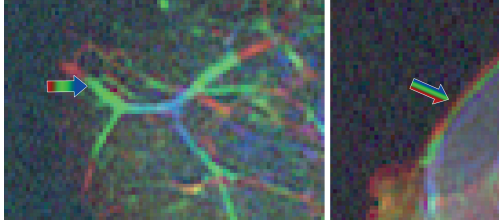


Figure 3.3: Left: A potential vessel with color gradient varying lengthwise. Right: Breast contour with color gradient varying along the cross section.

breast) during two sessions. Several weeks lay between these two sessions. Algorithm parameters are set for both stacks equally. The result is shown in Fig. 4.1 for four adjacent slices of each result stack. Slices of these two stacks are shifted such that they roughly match each other from column to column in Fig. 4.1. We derive similar results for both stacks, although the region of interest is shaped differently as seen in the two RGB images of Fig. 4.1. However, these binary results do not cover all elongated structures as visible in the RGB images. A better strategy is required to efficiently look for elongated segments in the solution space of the MSRG algorithm to derive more results. Also, a slice represents 3 mm thickness and obviously, a much higher resolution would improve the results for thinner vessel-like structures.

Due to the breast varying in shapes in different postures, it is very challenging to detect common features, however, this method might be used for automatic image registration.

5 Conclusions

With our proposed filter and combined with the MSRG algorithm from the previous work, it seems possible to assist clinicians in detecting vessel-like segments. The employed algorithm filtered elongated structures from the solution space of the MSRG algorithm. The results seem promising, however, they could not be evaluated for the correct detection of vessels without medical expertise. Instead, the proposed method was applied to two input MRI stacks from the same patient (and same breast), that were acquired during two sessions. The evalua-

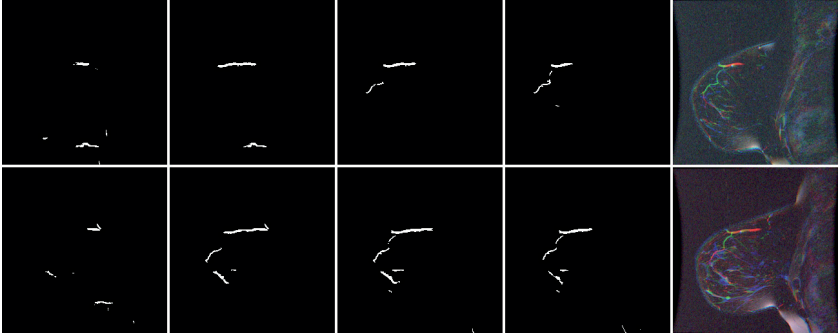


Figure 4.1: Each row contains four adjacent slices of the result stack as shown in Fig. 3.1 right. Each row's input stack (Fig. 3.1 on the left) was acquired from the same patient but several weeks lay between these sessions.

tion showed how barely the results deviate from each other. Although vessels will be positioned slightly differently between two sessions due to shaping the breast dependent on the actual posture in the MRI scanner, this method seemingly extracted similar elongated segments for both input MRI stacks. This suggests the assumption that the extracted segments are not artefacts.

Based on user experience, the solution space of the MSRG algorithm was searched through. Due to the complexity and computation time, it is impractical to search through the whole solution space. It is required to adjust the MSRG parameters accordingly, however, the MSRG algorithm is not guaranteed to be complete. Further investigations are required for more efficient search strategies.

Acknowledgement

This work has been supported by the LCM K2 Center within the framework of the Austrian COMET-K2 Programme.

References

1. Nishida N., Yano H., Nishida T., Kamura T., Kojiro M., "Angiogenesis in cancer." *Vasc. Health Risk Manag.* 2, 213–219, 2006.
2. Kahala G., Sklair M., Spitzer H., "Multi-scale blood vessel detection and segmentation in breast mris." *Journal of Medical and Biological Engineering volume 39*, 424–430, 2019.
3. Lin M., Chen J.H., Nie K., Chang D., Nalcioglu O., Su M.Y., "Algorithm-based method for detection of blood vessels in breast mri for development of computer-aided diagnosis." *J Magn Reson Imaging 2009*; 30: 817-824, 2009.
4. Lundervold A.S., Lundervold A., "An overview of deep learning in medical imaging focusing on mri." *Zeitschrift für Medizinische Physik*, 29 (2) (2019), pp. 102-127., 2019.
5. Moccia S., De Momi E., El Hadji S., Mattos L.S., "Blood vessel segmentation algorithms review of methods, datasets and evaluation metrics." *Comput. Methods Programs Biomed.*, 158 (2018), pp. 71-91, 2018.
6. Gierlinger M., Brandner D., Zagar B., "Multi-seed region growing algorithm for medical image segmentation." *M. Heizmann and T. Längle: Proceedings of FORUM BILDVERARBEITUNG 2020, KIT Scientific Publishing, Seite(n) 267 - 278.*, 2020.
7. D. Newitt, N. Hylton, on behalf of the I-SPY 1 Network and ACRIN 6657 Trial Team, "Multi-center breast dce-mri data and segmentations from patients in the i-spy 1/acrin 6657 trials." *The Cancer Imaging Archive.* <http://doi.org/10.7937/K9/TCIA.2016.HdHpgJLK>, 2016.

Generation of artificial training data for spectral unmixing by modelling spectral variability using Gaussian random variables

Johannes Anastasiadis and Michael Heizmann

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

Abstract A stochastic method how artificial training data for spectral unmixing can be generated from real pure spectra is presented. Since the pure spectra are modelled as Gaussian random vectors, spectral variability is also considered. These training data can in turn be used to train an artificial neural network for spectral unmixing. Non-negativity and sum-to-one constraints are enforced by the network architecture. The approach is evaluated using real mixed spectra and achieves promising results with the used datasets.

Keywords Spectral unmixing, spectral variability, Gaussian random variables, convolutional neural network

1 Introduction

Optical measurement techniques are often used for process monitoring in industry because they are non-contact and non-destructive. An important task is to determine the relative proportions of the components in substance mixtures. There are applications in many fields, such as food industry, medical technology, as well as in the processing of bulks. This task cannot be solved sufficiently with conventional colour images, because these only contain three colour channels (red, green, blue) per pixel, whereas hyperspectral images have a finely gained spectrum in each pixel that characterizes the materials much better [1]. If information of different materials is contained in one measured pixel,

spectral unmixing (SU) is needed to get the relative proportions, the abundances, of the pure materials in the pixel [2]. This is often done model-based [3]. However, these models usually assume a single spectrum for each pure substance involved. Actually, the spectra of the pure substances and also of the mixtures vary quite a lot [4,5].

Artificial neural networks (ANNs) achieve excellent results in many domains and have the benefit of not requiring model knowledge. Particularly for SU, the use of ANNs has further advantages [6]. First, the non-negativity and the sum-to-one constraints can be enforced by a normalising output layer. Second, spectral variability can be considered if it is contained in the training data. However, ANNs need a lot of significant training data to perform well, which are often not available in the domain of hyperspectral imaging.

In this approach we model the mixing characteristics including spectral variability of substances and use these models to generate training data for an ANN used for SU. Several spectra of each pure substance are needed for this approach. Even without the availability of real mixed spectra, most of the advantages of SU using an ANN can be exploited. Furthermore, performance can be improved if additional mixed spectra are available. This approach is therefore suitable for use in an industrial environment where the pure substances involved are known and hyperspectral images can be acquired in advance. We have already trained ANNs with artificially generated mixed spectra in a preceding work [7]. There, pure spectra randomly drawn from a set were mixed using models to obtain mixed spectra. In contrast, the approach in this paper models the spectra of the pure substances as Gaussian distributed random vectors. There is another approach, where SU is accomplished by directly applying Gaussian process regression [8], but not for training data generation. In this contribution, to generate the training data, the random vectors are combined for many different sets of abundances using the linear mixing model (LMM), the Fan Model (FM), the generalized bilinear model (GBM), and the linear quadratic model (LQM) [2,9–11].

The rest of the paper is structured as follows: In Section 2 the basics of SU are summarized. Section 3 then provides the probabilistic equivalents of the mixing models used. In Section 4 the experimental results are shown. Finally the paper is summarized and conclusions are drawn in Section 5.

2 Spectral unmixing

The estimation of the abundances in a mixture of substances based on its spectrum is called SU [2]. If the spectra of the pure substances involved are known, the term supervised SU is used. This is the case here. The estimation $\hat{\mathbf{a}} = [\hat{a}_1, \dots, \hat{a}_P]^T \in \mathbb{R}^P$ of the true abundances $\mathbf{a} = [a_1, \dots, a_P]^T \in \mathbb{R}^P$ of the up to P pure substances involved has to fulfil constraints in order to ensure physical validity. These are the non-negativity constraint and the sum-to-one constraint:

$$\hat{a}_p \geq 0 \quad \forall p, \quad \sum_{p=1}^P \hat{a}_p = 1. \quad (2.1)$$

Many of the methods used for SU are model-based. The models used here approximate the mixed spectra using a parametric function. The most commonly used mixing model is the LMM, which works well for many applications [2,5,12,13]:

$$\mathbf{y} = \sum_{p=1}^P \mathbf{m}_p a_p + \boldsymbol{\omega} = \mathbf{M} \mathbf{a} + \boldsymbol{\omega}. \quad (2.2)$$

Here $\mathbf{y} \in \mathbb{R}^\Lambda$ is a discrete spectrum sampled at Λ wavelength channels and $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_P] \in \mathbb{R}^{\Lambda \times P}$ are the spectra of the up to P involved pure substances. Differences between \mathbf{y} and the weighted sum of the pure material spectra are considered by $\boldsymbol{\omega} \in \mathbb{R}^\Lambda$. The LMM is based on the assumptions that mixing takes place on a macroscopic scale and that photons interact with only one material before they hit the sensor [14]. There are also non-linear mixing models that are summarized in [3]. In this paper the GBM [10]

$$\mathbf{y} = \sum_{p=1}^P \mathbf{m}_p a_p + \sum_{p=1}^{P-1} \sum_{q=p+1}^P \gamma_{pq} a_p a_q \mathbf{m}_p \odot \mathbf{m}_q + \boldsymbol{\omega} \quad (2.3)$$

and the LQM [11]

$$\mathbf{y} = \sum_{p=1}^P \mathbf{m}_p a_p + \sum_{p=1}^P \sum_{q=1}^P b_{pq} \mathbf{m}_p \odot \mathbf{m}_q + \boldsymbol{\omega} \quad (2.4)$$

are used, where γ_{pq} and b_{pq} are the non-linearity coefficients and \odot is the element-wise product. For $\gamma_{pq} = 1 \forall p, q$ the GBM is equivalent to the FM [9]. These mixing models also consider light that has interacted with up to two substances before hitting the sensor.

A commonly used approach considering (2.1) is the Fully Constrained Least Squares (FCLS) algorithm [15]. Here, the Lagrangian $L : \mathbb{R}^{P+1} \rightarrow \mathbb{R}$ with Lagrange multiplier $l \in \mathbb{R}$ is optimized:

$$L(\mathbf{a}, l) = \|\mathbf{y} - \mathbf{M} \mathbf{a}\|_2^2 - l \left(\sum_{p=1}^P a_p - 1 \right). \quad (2.5)$$

It is an iterative process that removes negative \hat{a}_p and the corresponding pure spectra. For estimations based on non-linear mixing models, gradient based line search methods can be used [10].

The presented mixing models assume that each pure substance is represented by a single spectrum. In reality, however, pure substances may have varying spectra [4]. These variations are referred to as spectral variability, which is primarily caused by surface structure. There are also unmixing algorithms considering the spectral variability such as the extended linear mixing model (ELMM) [16]. The ELMM extends the LMM by a set of parameters that allow scaling of the pure material spectra.

The presented approach considers spectral variability by including it in the generated training data. The process of generating these data is described in the following section.

3 Stochastic mixing models

For the approach used in this paper the spectra \mathbf{m}_p of the pure substances are modelled as Gaussian distributed random variables $M_p \in \mathbb{R}^\Lambda$. These can be entirely described by their mean vector $\boldsymbol{\mu}_{M_p} = \boldsymbol{\mu}_p \in \mathbb{R}^\Lambda$ and their covariance matrix $\boldsymbol{\Sigma}_{M_p, M_p} = \boldsymbol{\Sigma}_{p,p} \in \mathbb{R}^{\Lambda \times \Lambda}$. Mean vectors and covariance matrices of pure spectra are estimated using real data, which is why a set of measured spectra is required for each pure substance. Since the pure spectra of different pure materials do not depend on each other, they are assumed to be stochastically independent and therefore the cross-covariance matrix is $\boldsymbol{\Sigma}_{p,q} = \mathbf{0}$

$p \neq q$. A mixed spectrum $\mathbf{Y} \in \mathbb{R}^\Lambda$ is also modelled as a Gaussian distributed random vector with mean vector $\boldsymbol{\mu}_Y \in \mathbb{R}^\Lambda$ and covariance matrix $\boldsymbol{\Sigma}_{Y,Y} \in \mathbb{R}^{\Lambda \times \Lambda}$. The following moments result for the LMM:

$$\boldsymbol{\mu}_Y = \sum_{p=1}^P a_p \boldsymbol{\mu}_p, \quad \boldsymbol{\Sigma}_{Y,Y} = \sum_{p=1}^P a_p^2 \boldsymbol{\Sigma}_{p,p}. \quad (3.1)$$

Appropriate non-linearity coefficients γ_{pq} and b_{pq} must be determined for the GBM and the LQM. In this paper a constant value is used for this purpose, which means for all p and q : $\gamma_{pq} = \gamma$ and $b_{pq} = b$. This is a limitation, but in practice it allows better results compared to the LMM and the FM. To determine suitable values, a small validation dataset with some real mixed spectra is required. The mean vector $\boldsymbol{\mu}_Y$ for the GBM can be obtained by replacing \mathbf{m}_p by $\boldsymbol{\mu}_p$ in (2.3). The following covariance matrix results for the GBM:

$$\begin{aligned} \boldsymbol{\Sigma}_{Y,Y} &= \sum_{p=1}^P a_p^2 \boldsymbol{\Sigma}_{p,p} + \sum_{p=1}^{P-1} \sum_{q=p+1}^P \gamma^2 a_p^2 a_q^2 \boldsymbol{\Sigma}_{p \odot q, p \odot q} \\ &+ \sum_{p=1}^P \sum_{\substack{q=1 \\ q \neq p}}^P \gamma a_p^2 a_q (\boldsymbol{\Sigma}_{p, p \odot q} + \boldsymbol{\Sigma}_{p \odot q, p}) \\ &+ \sum_{p=1}^P \sum_{\substack{q=1 \\ q \neq p}}^{P-1} \sum_{\substack{l=q+1 \\ l \neq p}}^P \gamma^2 a_p^2 a_q a_l (\boldsymbol{\Sigma}_{p \odot q, p \odot l} + \boldsymbol{\Sigma}_{p \odot l, p \odot q}). \end{aligned} \quad (3.2)$$

The (cross-)covariance matrices are calculated by:

$$\begin{aligned} \boldsymbol{\Sigma}_{p \odot q, p \odot q} &= \mathcal{D}(\boldsymbol{\mu}_p) \boldsymbol{\Sigma}_{q,q} \mathcal{D}(\boldsymbol{\mu}_p) + \mathcal{D}(\boldsymbol{\mu}_q) \boldsymbol{\Sigma}_{p,p} \mathcal{D}(\boldsymbol{\mu}_q) + \boldsymbol{\Sigma}_{p,p} \odot \boldsymbol{\Sigma}_{q,q}, \\ \boldsymbol{\Sigma}_{p, p \odot q} + \boldsymbol{\Sigma}_{p \odot q, p} &= \boldsymbol{\Sigma}_{p,p} \mathcal{D}(\boldsymbol{\mu}_q) + \mathcal{D}(\boldsymbol{\mu}_q) \boldsymbol{\Sigma}_{p,p}, \\ \boldsymbol{\Sigma}_{p \odot q, p \odot l} + \boldsymbol{\Sigma}_{p \odot l, p \odot q} &= \mathcal{D}(\boldsymbol{\mu}_q) \boldsymbol{\Sigma}_{p,p} \mathcal{D}(\boldsymbol{\mu}_l) + \mathcal{D}(\boldsymbol{\mu}_l) \boldsymbol{\Sigma}_{p,p} \mathcal{D}(\boldsymbol{\mu}_q). \end{aligned} \quad (3.3)$$

Here, the output of operator $\mathcal{D}(\cdot)$ is a diagonal matrix with the elements of the input vector as its diagonal elements. If the input is a matrix, the output is a vector whose elements are the diagonal elements of the input matrix. The moments for the FM can be obtained by setting $\gamma = 1$.

The LQM also takes element-wise squared spectra into account. The following mean vector formula results for the LQM after transforming:

$$\boldsymbol{\mu}_Y = \sum_{p=1}^P a_p \boldsymbol{\mu}_p + b \boldsymbol{\mu}_{p \odot p} + \sum_{p=1}^{P-1} \sum_{q=p+1}^P 2b \boldsymbol{\mu}_p \odot \boldsymbol{\mu}_q, \quad (3.4)$$

where $\boldsymbol{\mu}_{p \odot p} = \boldsymbol{\mu}_p \odot \boldsymbol{\mu}_p + \mathcal{D}(\boldsymbol{\Sigma}_{p,p})$. The covariance matrix of a mixed spectrum is not shown for the LQM because it is similar to the one of the GBM. As for the other mixing models, all covariance and cross-covariance matrices are taken into account.

Using the Gaussian random variables calculated with the mixing models, training data for an ANN can now be generated. If training is performed over several epochs, new samples of mixed spectra can be drawn in every epoch. After data sampling, the spectra generated with non-linear mixing models are normalized. For this purpose they are multiplied by d_{GBM} or d_{LQM} , respectively:

$$d_{\text{GBM}} = \left(1 + \sum_{p=1}^{P-1} \sum_{q=p+1}^P \gamma a_p a_q \right)^{-1}, \quad d_{\text{LQM}} = \left(1 + P^2 b \right)^{-1}. \quad (3.5)$$

This is necessary because, unlike the LMM, the prefactors do not add up to 1. However, since the light that contributes to the linear component cannot contribute additionally to the quadratic component, this normalization is useful [7]. Thus, values for γ and b greater than 1 are also reasonable. In the next section the described approach is evaluated with real hyperspectral datasets.

4 Experimental results

To evaluate our approach we use real hyperspectral data acquired in our image processing lab. All datasets have 91 wavelength channels with an average width of 4 nm from 450 nm to 810 nm and 400 spectra per mixture. Two of the datasets consist of mixtures of coloured quartz sand. The first of them (quartz-3) contains 45 mixtures of maximum 3 components including pure substances, which vary in abundance steps of 0.125. The second one (quartz-4) contains 56 mixtures of maximum

4 components including pure substances, which vary in abundance steps of 0.2. The third dataset consists of 56 mixtures of colour powders (colour-4), which also have a maximum of 4 components each. Here, the components are varied in steps of 0.2, too. Only pure spectra are used for training. For all datasets, 12 mixtures are used for validation and the rest (30 respectively 40) to test. We train our ANN using the artificial mixtures generated with the approach in Section 3. The validation datasets are used to determine good non-linear coefficients for the GBM and the LQM. After training we test the ANN with the real mixtures in our test datasets. The root-mean-square error

$$\Delta_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{n=1}^N \frac{1}{P} \sum_{p=1}^P (\hat{a}_{pn} - a_{pn})^2}, \quad (4.1)$$

based on all N spectra including all mixtures of a test dataset, is used as a measure of performance. The results are compared to the FCLS algorithm and the ELMM based SU.

The ANN we used for the evaluation is the convolutional neural network (CNN) from [17], which is the one dimensional version of [7]. It has three convolutional layers and two fully connected layers. The length of the convolution kernels is 3 and the numbers of feature maps from input layer to output layer are 1, 16, 32, 64, 64, and 1. The number of epochs depends on the dataset. Therefore, the CNN is trained with quartz-3 dataset for 251 epochs, quartz-4 dataset for 61, and colour-4 dataset for 31 epochs.

Different training datasets are generated using all presented mixing models and different step sizes $s \in \mathbb{R}$ for the abundances \mathbf{a} (see Fig. 4.1). In every epoch there are 400 spectra per mixture drawn from the previously determined Gaussian random vectors of the mixtures. Figure 4.1 shows the results for the different methods. The prefix CNN indicates that a CNN was trained for SU in order to get this result. The training data were generated with the corresponding stochastic mixing model.

It is evident that in almost all cases the presented approach leads to an improvement of the results compared to the FCLS or the ELMM based approach. This is due to spectral variability being modelled based on data and taken into account by the CNN. Which mixing

model yields the best results depends on the unmixing task. In the datasets used here the non-linear mixing models perform better. The FM delivers quite good results, although like LMM, it does not need any additional parameters. The colour-4 dataset shows, however, that the choice of the right mixing model can achieve a significant improvement. Smaller step sizes and thus more large training datasets lead in most cases to an improvement.

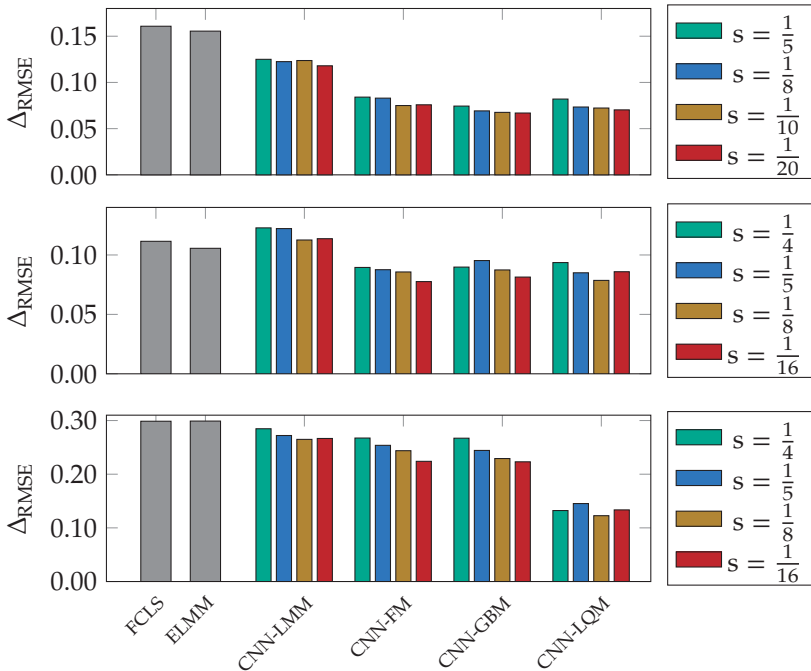


Figure 4.1: Root-mean-square error of the abundances for the compared methods applied on the quartz-3 (top), the quartz-4 (middle), and the colour-4 (bottom) dataset, respectively.

5 Summary

In this paper an approach is presented where artificial mixed spectra are generated using stochastic mixing models. These mixed spectra can then be used to train a CNN for SU. In this way, spectral variability is considered. Compared to methods from literature, which in part also include spectral variability, better results can be achieved with regard to the error of the estimated abundances. The choice of the mixing model in dependence of the problem significantly influences the quality of the estimation. Finer step sizes in the specified abundances for the training data can lead to an additional improvement.

In the future, the approach could be extended in such a way that the mean vectors and covariance matrices of the random vectors of mixed spectra are determined based on data instead of models. However, larger training datasets would be needed to train these networks.

References

1. A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590–598, 2007.
2. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 44–57, 2002.
3. N. Dobigeon, Y. Altmann, N. Brun, and S. Moussaoui, "Linear and non-linear unmixing in hyperspectral imaging," in *Data Handling in Science and Technology*, C. Ruckebusch, Ed. Elsevier, 2016, vol. 30, pp. 185–224.
4. R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, and C. Jutten, "Spectral variability in hyperspectral data unmixing: A comprehensive review," *arXiv preprint arXiv:2001.07307*, 2020.
5. W. Krippner, S. Bauer, and F. Puente León, "Considering spectral variability for optical material abundance estimation," *tm – Technisches Messen*, vol. 85, no. 3, pp. 149–158, 2018.
6. J. Anastasiadis and F. Puente León, "Spatially resolved spectral unmixing using convolutional neural networks (German paper)," *tm – Technisches Messen*, vol. 86, no. s1, pp. 122–126, 2019.

7. J. Anastasiadis, P. Benzing, and F. Puente León, "Generation of artificial data sets to train convolutional neural networks for spectral unmixing (German paper)," *tm – Technisches Messen*, vol. 87, no. 9, pp. 542–552, 2020.
8. Y. Altmann, N. Dobigeon, S. McLaughlin, and J. Tourneret, "Nonlinear spectral unmixing of hyperspectral images using Gaussian processes," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2442–2453, 2013.
9. W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.
10. A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, 2011.
11. I. Meganem, P. Déliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear-quadratic mixing model for reflectances in urban environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 544–558, 2013.
12. S. Bauer, J. Stefan, and F. Puente León, "Hyperspectral image unmixing involving spatial information by extending the alternating least-squares algorithm," *tm – Technisches Messen*, vol. 82, no. 4, pp. 174–186, 2015.
13. W. Krippner and F. Puente León, "Band selection and estimation of material abundances using spectral filters (German paper)," *tm – Technisches Messen*, vol. 85, no. 6, pp. 454–467, 2018.
14. R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *Journal of Geophysical Research: Solid Earth*, vol. 89, no. B7, pp. 6329–6340, 1984.
15. D. Heinz, C.-I. Chang, and M. L. Althouse, "Fully constrained least-squares based linear unmixing," in *IEEE 1999 International Geoscience and Remote Sensing Symposium.*, vol. 2. IEEE, 1999, pp. 1401–1403.
16. M. A. Vezon, L. Drumetz, G. Tochon, M. Dalla Mura, A. Plaza, J. Bioucas-Dias, and J. Chanussot, "A new extended linear mixing model to address spectral variability," in *2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2014, pp. 1–4.
17. J. Anastasiadis and M. Heizmann, "CNN-based augmentation strategy for spectral unmixing datasets considering spectral variability," in *Image and Signal Processing for Remote Sensing XXVI*, L. Bruzzone, F. Bovolo, and

E. Santi, Eds., vol. 11533, International Society for Optics and Photonics. SPIE, 2020, pp. 188–199.

Line Spectra Analysis: A Cumulative Approach

Achim Kehrein¹ and Oliver Lischtschenko²

¹ Rhine-Waal University of Applied Sciences,
Faculty of Technology and Bionics,
Marie-Curie Str. 1, 47533 Kleve, Germany

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

Abstract An optical spectrometer uses detector pixels that measure the integrated intensity over a certain interval of wavelengths. These integrated pixel values are divided by the interval width and then interpreted as estimates of function values of the wanted spectral irradiance. Hence each pixel measurement constitutes an averaging process. But, averaging biases at maxima: pixel data feature lower maxima. This paper proposes the conceptual use of a cumulated spectrum to estimate spectral data. The integrated quantities are placed in their natural habitat. The motivation originates from the fact that pixel data as integrated quantities are exact values of the cumulated spectrum. Averaging becomes obsolete. There is no information loss. We start with a single spectral line. This “true” spectrum is blurred mimicking the instrument function of the spectrometer optics. For simplicity we consider the instrument function to have Gaussian shape. We integrate the blurred spectrum over subintervals to simulate the pixel measurements. We introduce a cumulated spectrum approach. We compare the cumulated approach with the approach that interpolates the averaged function value estimates of the non-cumulated spectrum. The cumulated approach requires only basic mathematical concepts and allows fast computations.

Keywords Spectrum, spectral fitting, cumulative spectrum, data processing, plasma, line emission, spectroscopy

1 Introduction

The analysis of emitted line spectra provides insight into the qualitative and quantitative composition of materials. Virtually all analysis is carried out by using non-linear fitting of (multi) Gaussians to the recorded data. This paper replaces the all-purpose non-linear fitting by direct computation of the peak parameters. This reduces dramatically the computational effort and eliminates the dependence on intelligent initial parameter guesses.

Probability theory or statistics knows two perspectives on Gaussians, i.e. on normal distributions. On the one hand there is famous bell-shaped curve, the probability density function (PDF). This is how we usually visualize the normal distribution. On the other hand there is the tabulated cumulative distribution function (CDF) with which we perform almost all calculations. The PDF possesses non-negative values and finite area. The CDF is a non-decreasing function.

A spectrum is also a function with non-negative spectral irradiance values and finite area which corresponds to the finite irradiance. This is the usual visualization of a spectrum. In this paper we add the equivalent perspective of a cumulated spectrum to perform our calculations. After all the irradiance of the spectrum is nothing but the cumulated spectral irradiance (integrated with respect to wavelength). This paper introduces how the cumulated approach can be used to fit Gaussian peaks to discrete spectral data in a simple fashion.

In 2020 the global market volume of applications relying on spectral line emission analyses is estimated at a trillion US-Dollar. Examples of these applications are semi-conductor production, quality control of physical coating processes [1], minimally invasive cancer surgery with live tissue diagnostics [2], natural resource exploration [3], and magnetic plasma fusion research [4].

2 Mathematical Model

We use a simple model. An object sampled by a spectrometer possesses a “true spectrum” which is a function that assigns to each wavelength a certain non-negative value, the spectral irradiance. The wavelength can have any positive value, so mathematically, the true spectrum is a

function S (signal) defined on the interval $(0, \infty)$ with non-negative values. We restrict our attention to a true spectrum with a single spectral line which we approximate by a Dirac delta function,

$$S(\lambda) = S_\mu \cdot \delta(\lambda - \mu). \quad (2.1)$$

2.1 Blurring - Spectral Lines Become Gaussians

A spectrometer blurs the true spectrum S by convoluting it with the so-called instrument function I . We call the result of this convolution $I * S$ the blurred spectrum,

$$B(\lambda) = \int_0^\infty I(\lambda - \ell) \cdot S(\ell) d\ell. \quad (2.2)$$

This blurring depends on the used spectrometer. In this paper we assume that the instrument function is sufficiently well approximated by the Gaussian shape

$$I(\lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\lambda^2}{2\sigma^2}} \quad (2.3)$$

with the width parameter³ $\sigma > 0$. Selecting a Gaussian is purely illustrative. Any shape can be cumulated and treated analogously.

The blurring convolutes the line spectrum, equation (2.1), into the Gaussian peak

$$B(\lambda) = S_\mu \cdot I(\lambda - \mu). \quad (2.4)$$

2.2 Pixel Detectors - Spectral Irradiances Are Locally Integrated

A spectrometer measures the blurred spectrum on a finite interval of wavelengths $[\lambda_0, \lambda_R]$. Its detector consists of finitely many pixels that detect adjacent parts of the blurred spectrum. Each pixel measures the cumulated spectral irradiance on a subinterval of wavelengths. Let the r th pixel cover the subinterval $[\lambda_{r-1}, \lambda_r]$, $1 \leq r \leq R$. Then the spectral irradiance of the blurred spectrum measured by the r th pixel is

$$B_r = \int_{\lambda_{r-1}}^{\lambda_r} B(\lambda) d\lambda. \quad (2.5)$$

³We do not want to call σ standard deviation, since this is a non-statistical use.

We divide these cumulated values by their subinterval widths to convert B_r into the same unit as the blurred spectrum and obtain the averaged values

$$B_r^* = \frac{B_r}{\lambda_r - \lambda_{r-1}} = \frac{1}{\lambda_r - \lambda_{r-1}} \cdot \int_{\lambda_{r-1}}^{\lambda_r} B(\lambda) d\lambda . \quad (2.6)$$

A spectrometer with R pixels provides the finitely many values B_1^* , B_2^* , ..., B_R^* that approximate the blurred spectrum over the interval $[\lambda_0, \lambda_1] \cup \dots \cup [\lambda_{R-1}, \lambda_R] = [\lambda_0, \lambda_R]$.

Each pixel value B_r^* – though obtained by averaging – can be interpreted as a function value estimate at the pixel's mean wavelength

$$\lambda_r^* = \frac{\lambda_{r-1} + \lambda_r}{2} . \quad (2.7)$$

Then the measured spectrum is a function whose graph consists of the finitely many points (λ_1^*, B_1^*) , (λ_2^*, B_2^*) , ..., (λ_R^*, B_R^*) .

3 Proof of Concept - Matching a Gaussian Peak

As an example we consider a single blurred spectral line shown in Figure 3.1. This is a Gaussian peak centered at wavelength $\mu = 500$ nm with width parameter $\sigma = 1$ nm and irradiance $100 \mu\text{W}/\text{cm}^2$ (integrated spectral irradiance). We assume that the wavelength interval from $\lambda_0 = 497$ nm till $\lambda_{11} = 503$ nm is detected by a row of eleven equidistant pixels. Their averaged values B_1^* , ..., B_{11}^* are shown as function value estimates. Figure 3.1 illustrates in particular the effect of averaging the concave function segment close to the maximum: averaging underestimates a maximum value.

3.1 Non-linear All-purpose Fitting of the Function Value Estimates

As control case for our example we use the all-purpose Matlab command `fit` to fit a Gaussian peak to the pixel values (λ_1^*, B_1^*) , ..., $(\lambda_{11}^*, B_{11}^*)$. The model function to be fitted is

$$\hat{B}(\lambda) = a \cdot e^{-\frac{(\lambda-b)^2}{c}} \quad (3.1)$$

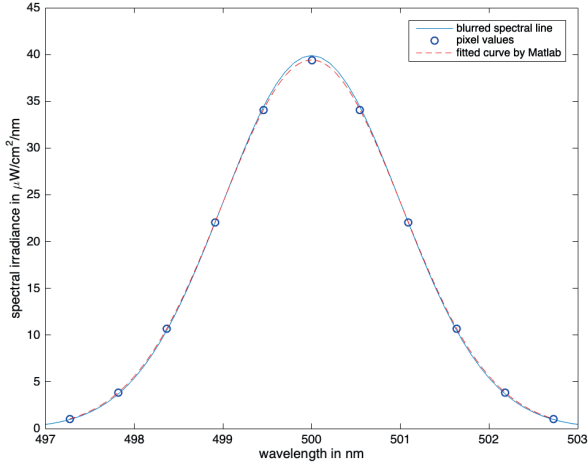


Figure 3.1: A Gaussian peak models a blurred spectral line. The wavelength interval [497 nm, 503 nm] is divided into 11 equidistant subintervals (pixels). The pixel values are placed at the midpoints of the subintervals. The dashed curve is the fitted result by Matlab. The maximum of the fit is lower than the maximum of the blurred spectrum.

with model parameters a , b , and c . The underlying algorithm for non-linear least squares requires intelligent initial parameters to produce convergence. We provided the initial parameters $a_0 = 39.4$ (maximal pixel value), $b_0 = 500$, and $c_0 = 2$.

Matlab finds $\hat{a} = 39.41$, $\hat{b} = 500$, and $\hat{c} = 1.432$ with 95% confidence intervals (39.40, 39.41), (500, 500), and (1.432, 1.432) respectively. The estimated model parameters translate into: the line is located at wavelength 500 nm with width parameter $\sigma = \hat{c} / \sqrt{2} = 1.0126$ nm and irradiance $\hat{S}_\mu = \hat{a} \cdot \sigma \sqrt{2\pi} = 100.0287 \mu\text{W}/\text{cm}^2$. Figure 3.1 also shows the fitted curve. It fits the blurred spectrum well, but features the biased smaller maximum. The estimated irradiance corresponds to the area under the curve. The fit overestimates the irradiance.

3.2 Cumulative Spectrum - Natural Use of Pixel Values

This section presents the alternative way to match a Gaussian peak to the pixel data. Instead of interpreting the pixel data as function value estimates, we use them exactly as the integrated quantities that they are. To do so, we consider the blurred spectrum cumulated over the measured interval

$$C(\lambda) = \int_{\lambda_0}^{\lambda} B(\ell) d\ell \quad , \lambda_0 \leq \lambda \leq \lambda_R \quad , \quad (3.2)$$

and the cumulated pixel values

$$C_r = \sum_{i=1}^r B_i \quad , \quad r = 0, 1 \dots, R. \quad (3.3)$$

The cumulated pixel values are exact samples of the cumulated blurred spectrum at the subinterval endpoints, $C_r = C(\lambda_r)$, $r = 0, 1 \dots, R$. For our example, these quantities are shown in Figure 3.2 and in Table 13.1.⁴

Table 13.1: Cumulated pixel values of example.

k	λ_k	C_k	k	λ_k	C_k	k	λ_k	C_k	k	λ_k	C_k
0	497.00	0.00	3	498.64	8.50	6	500.27	60.61	9	501.91	97.05
1	497.55	0.57	4	499.18	20.53	7	500.82	79.20	10	502.45	99.16
2	498.09	2.68	5	499.73	39.12	8	501.36	91.23	11	503.00	99.73

3.3 Estimating Parameters from Cumulative Values

The cumulative normal distribution function

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(\zeta-\mu)^2}{2\sigma^2}} d\zeta \quad (3.4)$$

⁴The blurred spectrum B (spectral irradiance) is the derivative of the cumulative spectrum C (irradiance). Consistently, the earlier introduced averaged values $B_r^* = (C_r - C_{r-1})/(\lambda_r - \lambda_{r-1})$ located at the midpoint λ_r^* turn out to be the centered differences known from numerical differentiation.

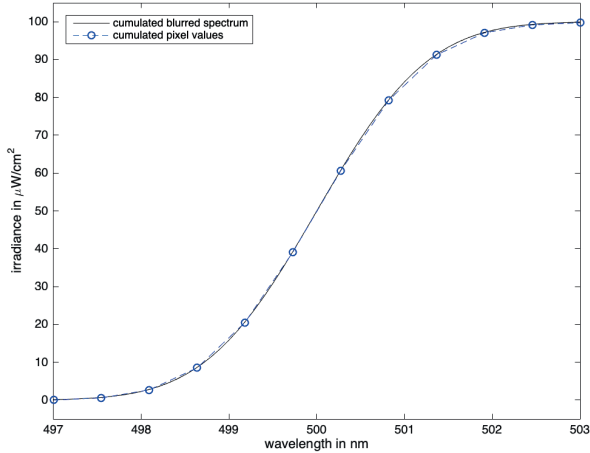


Figure 3.2: The cumulative Gaussian peak function (solid curve) and the cumulated pixel values at the respective endpoints of the subintervals. The dashed curve is the piecewise linear interpolation. The cumulated pixel values are exact values of the cumulative Gaussian peak function, not just estimates.

satisfies $F(\mu) = 0.5$ and can be normalized, $z = (x - \mu)/\sigma$, into the standard normal distribution $\Phi(z)$. In particular, we obtain the tabulated value

$$F(\mu - \sigma) = \Phi(-1) = 0.1587. \quad (3.5)$$

We use these properties to estimate the parameters of the blurred spectrum.

The irradiance of the spectral line is estimated as the area under the measured blurred spectrum, which equals a difference in the cumulated spectrum,

$$\hat{S}_\mu = C_R - C_0 = \int_{\lambda_0}^{\lambda_R} B(\lambda) d\lambda \approx \int_0^\infty B(\lambda) d\lambda \approx S_\mu. \quad (3.6)$$

In Figure 3.2 this is $C_{11} - C_0 = 99.73 - 0 = 99.73$. After we will have obtained a first estimate for all parameters S_μ , μ , and σ , we will check for consistency and update them.

Now we estimate the wavelength μ of the spectral line from the cumulated pixel values. First, we find the subinterval $[\lambda_{j-1}, \lambda_j]$, in which the spectral line is located,

$$C_0 < \dots < C_{j-1} < 0.5 \cdot \hat{S}_\mu \leq C_j < \dots < C_R . \quad (3.7)$$

In Figure 3.2 this is $[\lambda_5, \lambda_6]$. We estimate $\hat{\mu}$ by linear interpolation,

$$\hat{\mu} = \lambda_{j-1} + \frac{0.5 \cdot \hat{S}_\mu - C_{j-1}}{C_j - C_{j-1}} (\lambda_j - \lambda_{j-1}) . \quad (3.8)$$

Since the cumulative normal distribution function is almost linear at the half-height value, linear interpolation is sufficient. However, the interpolation accuracy depends on the pixel width $[\lambda_{j-1}, \lambda_j]$. A narrower pixel width produces a better approximation. Figure 3.2 shows a good match between the cumulative blurred spectrum and the linear interpolation on $[\lambda_5, \lambda_6]$. We compute $\hat{\mu} = 500$ nm.

Now we estimate σ from the cumulated pixel values. We find the subinterval $[\lambda_{k-1}, \lambda_k]$, in which $\mu - \sigma$ is located (see equation 3.5),

$$C_0 < \dots < C_{k-1} < 0.1587 \cdot \hat{S}_\mu \leq C_k < \dots < C_R . \quad (3.9)$$

For the following interpolation we need $k \geq 2$ which is a reasonable assumption on the number and widths of the pixels covering the spectral line. In Figure 3.2 this is the interval $[\lambda_3, \lambda_4]$. The curvature over this interval is not negligible. Linear interpolation would systematically overestimate σ due to the convexity. Thus we compute $\hat{\mu}$ by quadratic interpolation through (λ_{k-2}, C_{k-2}) , (λ_{k-1}, C_{k-1}) , and (λ_k, C_k) .⁵ We use Newton's form of the interpolating parabola. The required divided differences are

$$[C_k, C_{k-1}] = \frac{C_k - C_{k-1}}{\lambda_k - \lambda_{k-1}} \quad , \quad [C_{k-1}, C_{k-2}] = \frac{C_{k-1} - C_{k-2}}{\lambda_{k-1} - \lambda_{k-2}} \quad , \quad (3.10)$$

and

$$[C_k, C_{k-1}, C_{k-2}] = \frac{[C_k, C_{k-1}] - [C_{k-1}, C_{k-2}]}{\lambda_k - \lambda_{k-2}} \quad . \quad (3.11)$$

⁵We use the indices $k-2, k-1, k$ instead of $k-1, k, k+1$ for numerical reasons.

The interpolating parabola is

$$P(\lambda) = C_k + [C_k, C_{k-1}](\lambda - \lambda_k) + [C_k, C_{k-1}, C_{k-2}](\lambda - \lambda_k)(\lambda - \lambda_{k-1}) \quad (3.12)$$

and we compute $\hat{\mu} - \hat{\sigma}$ from $P(\hat{\mu} - \hat{\sigma}) = 0.1587 \cdot \hat{S}_\mu$. We plug in

$$C_k + [C_k, C_{k-1}](\hat{\mu} - \hat{\sigma} - \lambda_k) + [C_k, C_{k-1}, C_{k-2}](\hat{\mu} - \hat{\sigma} - \lambda_k)(\hat{\mu} - \hat{\sigma} - \lambda_{k-1}) = 0.1587 \cdot \hat{S}_\mu \quad (3.13)$$

and sort for the powers of the unknown $\hat{\sigma}$,

$$\begin{aligned} 0 = & C_k + [C_k, C_{k-1}](\hat{\mu} - \lambda_k) - 0.1587 \cdot \hat{S}_\mu \\ & + [C_k, C_{k-1}, C_{k-2}](\hat{\mu} - \lambda_k)(\hat{\mu} - \lambda_{k-1}) \\ & + ([C_k, C_{k-1}] + [C_k, C_{k-1}, C_{k-2}](\hat{\mu} - \lambda_k + \hat{\mu} - \lambda_{k-1})) \cdot \hat{\sigma} \\ & + [C_k, C_{k-1}, C_{k-2}] \cdot \hat{\sigma}^2 \end{aligned} \quad (3.14)$$

In the template $0 = a\hat{\sigma}^2 + b\hat{\sigma} + c$ we obtain the following coefficients for our example.

$$\begin{aligned} a &= [C_k, C_{k-1}, C_{k-2}] = 10.730 \\ b &= 22.278 + 10.730(0.82 + 1.36) = 45.669 \\ c &= 20.53 + 22.278 \cdot 0.82 - 15.827 + 10.730 \cdot 0.82 \cdot 1.36 = 34.937 \end{aligned} \quad (3.15)$$

From the geometry, the solution of the quadratic formula with the difference is the relevant one. The result rounded to four significant digits is

$$\hat{\sigma} = \frac{b - \sqrt{b^2 - 4ac}}{2a} = 1.000. \quad (3.16)$$

3.4 Updating the Estimates

Now that we have a complete set of estimates, \hat{S}_μ , $\hat{\mu}$, and $\hat{\sigma}$ we use them to update the estimate of the irradiance S_μ . By design our first estimate of S_μ is an underestimate. We do not know how much of its area we actually considered. Now $\hat{\mu}$ and $\hat{\sigma}$ allow us to assess this underestimate. We can compute the area covered by the interval $[\lambda_0, \lambda_R]$ using

the tabulated normal distribution. This becomes particularly important when we have covered a wavelength range significantly smaller than $[\mu - 3\sigma, \mu + 3\sigma]$ (which corresponds to 99.7% of the area). This is also useful when we have an asymmetric interval about the location of the spectral line.

Such an update becomes important when spectral lines are located close to one another and their areas partially overlap. A similar problem arises when a spectral line is close to an endpoint λ_0 or λ_R so that a significant part of the blurred spectrum lies outside our observed interval.

For our simple example the numbers will not change much, but let us describe the updating process. Our interval was $[\lambda_0, \lambda_{11}] = [497 \text{ nm}, 503 \text{ nm}]$. With respect to our estimates $\hat{\mu} = 500 \text{ nm}$ and $\hat{\sigma} = 1 \text{ nm}$ the interval endpoints correspond to the standardized z-values -3 and 3 respectively (computed from $\lambda_i = \hat{\mu} + z\hat{\sigma}$). From the tabulated normal distribution we read off $\Phi(-3) = 0.0013$ and $\Phi(3) = 0.9987$. Therefore, the area covers $0.9987 - 0.0013 = 0.9974 = 99.74\%$ of the irradiance. Hence we update the irradiance estimate

$$\hat{\sigma}_\mu = \frac{\text{original estimate}}{\text{covered area}} = \frac{99.73}{0.9974} = 99.99 . \quad (3.17)$$

With this updated estimate we recalculate $\hat{\mu}$ and $\hat{\sigma}$. If the updated values deviate significantly from the original ones, we iterate this updating process until the estimates become stationary.

4 Outlook

We have only considered the simple example of an isolated spectral line in a sufficiently large wavelength interval. The results of the cumulated view point are promising. Only linear and quadratic equations were required in combination with only two tabulated values of the normal distribution. However, many aspects are still in need of further investigation:

- The whole process starts with the determination of an interval $[\lambda_0, \lambda_R]$ which covers most of the irradiance of a blurred spectral line. We have to automate the partitioning of a spectrum with several spectral lines into such intervals.

- We need to study examples of spectral lines whose blurred spectra overlap. Which separation of lines is required for the method to work properly? How will the method react when the lines are too close? Similarly, how do we handle spectral lines that are close to the endpoints of the measured interval?
- We want to transfer this method to other templates like a Lorentzian or a Voigt profile instead of a Gaussian peak.

The presented approach significantly reduces the calculation complexity of spectral matching by avoiding all-purpose non-linear fit routines. Due to the simplicity of the computation steps, the presented cumulative approach can be implemented in FPGAs. This allows immediate on-camera real-time data processing.

This affects many applications. Control circuits for online plasma analysis will benefit from shorter latencies. Processes like laser induced breakdown spectroscopy (LIBS) that so far require a complex post mortem analysis have the potential to be treated live. Adapting the approach to complex fusion spectra will significantly reduce the amount of currently needed a-priori knowledge. The whole analysis becomes more deterministic and less dependent on intelligent initial estimates. The only information to be determined in advance is the cumulative distribution of the instrument function.

References

1. A. Mackus, S. Heil, E. Langereis, H. Knoops, M. Van de Sanden, and W. Kessels, "Optical emission spectroscopy as a tool for studying, optimizing, and monitoring plasma-assisted atomic layer deposition processes," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 28, no. 1, pp. 77–87, 2010.
2. R. Kanawade, F. Mehari, C. Knipfer, M. Rohde, K. Tangermann-Gerk, M. Schmidt, and F. Stelzle, "Pilot study of laser induced breakdown spectroscopy for tissue differentiation by monitoring the plume created during laser surgery?an approach on a feedback laser control mechanism," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 87, pp. 175–181, 2013.
3. A. J. R. Bauer and S. G. Buckley, "Novel applications of laser-induced breakdown spectroscopy," *Applied spectroscopy*, vol. 71, no. 4, pp. 553–566, 2017.

4. M. G. von Hellermann, G. Bertschinger, W. Biel, C. Giroud, R. Jaspers, C. Jupén, O. Marchuk, M. O'Mullane, H. Summers, A. Whiteford *et al.*, "Complex spectra in fusion plasmas," *Physica scripta*, vol. 2005, no. T120, p. 19, 2005.

Quality analysis for the production of pelletised materials using Fourier Descriptors

Sebastian Michlmayr¹, Dominik Neumann², and Bernhard Zagar¹

¹ Johannes Kepler University Linz
Altenberger Straße 69 A-4040 Linz

² ECON GmbH,
Biergasse 9, A-4616 Bergern

Abstract We evaluate the effectiveness of Fourier Descriptors for an in-line characterization of the quality of pelletised materials. For the quality analysis we evaluate the significance of information conveyed by a limited number of low order Fourier Descriptors. Further, we investigate the influence of image resolution and shape on the outcome.

Keywords Fourier analysis, fourier descriptors, surface analysis, digital image processing

1 Introduction

Plastic pellets may be produced by pushing molten prime material through die plates into a water stream and pelletizing by cutting the extrusion with a rotating cutter to the preselected length [1]. Examples of resulting pellets are shown in Figure 1.1. The water stream simultaneously cools the pellets and transports them to the next stage of production – the dryer.

The quality of these pellets is influenced, among other things, by the temperature of the molten material, the sharpness of the cutter, and the cleanliness and temperature of the water. Some polymers, for instance, may crystallize with higher temperatures and, subsequently, change their optical appearance and become opaque [2]. Furthermore, the shape of the pellets is highly dependent on the sharpness of the



Figure 1.1: Images of two different sets of pellets. On the left-hand side the pellets are comparatively smooth. The observable difference in Gray-scale values is due to temperature differences during production which causes some pellets to crystallize. On the right-hand side the shape of the pellets is less homogeneous which might indicate a blunt cutting tool.

cutter's blade. Given the plasticity of the molten filaments the separation process of the pellets may lean more on the side of the preferred cutting instead of tearing them off in the cutting tool, which indicates wear on the tool or some other quality diminishing conditions. A blunt blade may create pellets as shown on the right-hand side of Figure 1.1 where the main body of the pellet has thin appendices or strands.

To devise an in-line system able to observe the extrusion process and characterize the resultant quality with low reaction time is not a trivial task. Since the quality is dependent on multiple parameters and only partly controlled by the process itself but always dependent on the processed materials, to attain an optimized result requires quick reaction times of the quality assertion process. In our proposed scheme pellets were sampled at random right after the dryer and shape information, as one of the early indicators of wear on the cutter, was gathered with a camera to draw conclusions regarding some of the process parameters.

2 Measurement Setup

A B&R Vision System [3] camera was used as it allows an easy integration into the PLC system of the production line. The gray-scale-camera is complemented by LEDs with different colours which by successive illumination of an object allows to gain also colour information about the analysed specimen. That information allows to detect a beginning

cystallisation of the specimens, which can be accompanied by a change in opacity [2]. In Figure 1.1 the crystallised pellets are recognisable due to their different transparency.

The overall image acquisition and processing system is complemented by an on-board image processing system. The on-board system is capable to evaluate – amongst other things – the area in pixels, the mean gray value, the rectangularity, the circularity, and the anisometry of a detected object. In combination with a B&R PLC, additional image processing algorithms can be implemented.

The measurement setup used for acquiring the images is shown in Figure 2.1. A funnel is used to concentrate the falling pellets into the focal plane of the camera.

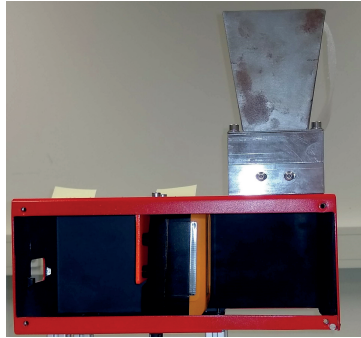


Figure 2.1: Picture of setup for testing.

3 Fourier Descriptors augmenting the feature space

The camera's on-board image parameters showed strong limitations when trying to analyse the produced pellets for production quality. Figure 3.1 shows two pellets which objectively have a very different shape, yet a distinction based on the parameters rectangularity, circularity and anisometry with values of 79/69/13 for the first and 80/69/13 for the second blob is almost impossible.

To overcome those limits the usefulness of Fourier Descriptors to assess the quality of pellets was investigated.



Figure 3.1: Images taken of two different pellets. The red line represents the detected contour of the blobs.

Image Processing In a first step the gray-scale image acquired by the camera is transformed into a binary image by thresholding the image. With this the regions, also called blobs (binary large objects), that represent pellets can easily be separated from the dark background. After labelling the regions, the boundary of each region is identified. [4] proposes a simple tracing algorithm, that returns the region's boundary as a list of points with their x - and y -components.

Following [5] the contour is transformed into a complex valued signal

$$s[n] = x[n] + iy[n], \quad (3.1)$$

with $n = 0, 1, 2, \dots, M - 1$ where M is the number of contour pixels. The discrete Fourier transform of $s[n]$ returns the complex valued spectrum

$$S[n] = a[n] + ib[n], \quad (3.2)$$

the Fourier Descriptors, with $n = 0, 1, 2, \dots, M - 1$ where M is the number of contour pixels.

The component $S[0]$ of the nonsymmetric spectrum represents the centre of the contour. As it is irrelevant for the evaluation of the contour's shape we won't consider it for the following calculations.

The remaining spectral components can be interpreted as rotating pointers with different rotational speeds, amplitudes and phases. The components $S[k]$ and $S[M - k]$, with $k = 1, 2, 3 \dots (M - 1)/2$ for an uneven M and $k = 1, 2, 3 \dots M/2 - 1$ for an even M , have the same rotational speed but opposite rotational directions and are in the following identified as k -th order Fourier Descriptors. For an even M there is only a single component of order $M/2$.

For the following analysis we consider only the Fourier Descriptors' absolute value

$$C[n] = |S[n]|, \quad (3.3)$$

as it describes the magnitude of each spectral component.

Exemplification Figures 3.2 and 3.3 show two different pellets. The red line represents the original contour and the green one the result of inversely transforming only the lowest 4, 20, and 32 (left to right) Fourier Descriptors.

Figure 3.2 shows a pellet with a smooth surface and no appendices. The simple contour implies that higher order Fourier Descriptors have small magnitudes. This can also be observed in the Figure, as the inverse transformation of only the first four descriptors already gives a very close approximation of the original contour. As the number of Fourier Descriptors for the inverse transformation is increased the green contour changes only slightly.

Figure 3.3 shows a pellet with a complex shape due to appendices resulting from a low quality production process.

The complexity implies larger magnitudes in higher order Fourier Descriptors, therefore the inverse transform using higher order Fourier Descriptors provide a better approximation of the contour. This is clearly visible in the figure as an inverse transform of the first four Fourier Descriptors results in an inadequate approximation of the contour. The use of 20 and 32 Fourier Descriptors results in an increasingly better approximation.

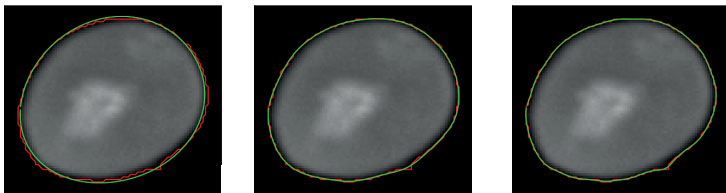


Figure 3.2: Images of a smooth pellet with its original contour in red and in green the result of inversely transforming the lowest 4, 20, and 32 Fourier Descriptors.

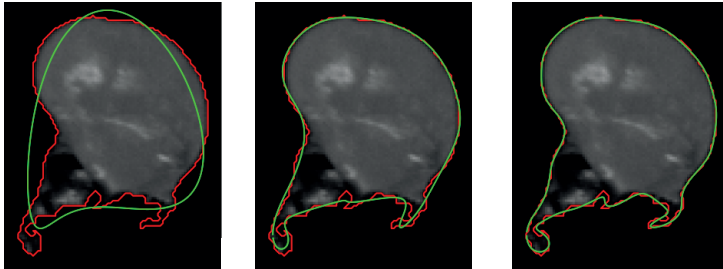


Figure 3.3: Images of complex blob with its original contour in red and in green the result of inversely transforming the lowest 4, 20, and 32 Fourier Descriptors.

Application as Shape-Descriptor In [6] the absolute values of the single descriptors are used in order to obtain information on the size of the object to be measured. Our goal, however, is not to estimate the size of the object, but rather the smoothness of the surface.

As mentioned above the smoother the surface the lower are the values in the higher order Fourier Descriptors in comparison to lower order ones. Hence, the ratio of lower order Fourier Descriptors to all descriptors would be comparatively large and, for example, for a perfect circle would be 1. Therefore we propose

$$v = \frac{\sum_{n=0}^N (C_{low}[n])}{\sum_{n=1}^{M-1} (C[n])} \quad (3.4)$$

as a measure of the smoothness with $C_{low} \in C$. C_{low} is a selection of N lowest order Fourier Descriptors.

Influence of Resolution The above considerations are made under the assumption of shapes with an infinitely high resolution. In a real world application the resolution is always limited, due to the quantisation carried out by a camera. This quantisation leads to a stepped contour of objects in images, which lowers their smoothness and leads to larger magnitudes in higher order Fourier Descriptors.

For increasingly lower resolutions – measured in pixels per blob – the influence of the quantisation on the smoothness increases as the steps in the contour become increasingly bigger relative to the contour.

This dependency is depicted in Figure 3.4 which shows the calculated smoothness values for a quantised ideal circle (left) and a quantised ideal square (right) for increasing pixels per blob. The edges of the square were rotated by 45° to the edges of the pixels to observe the influence of the quantisation. The four different lines in each plot represent each a different combination of the lowest three order Fourier Descriptors, which will be discussed in the next part.

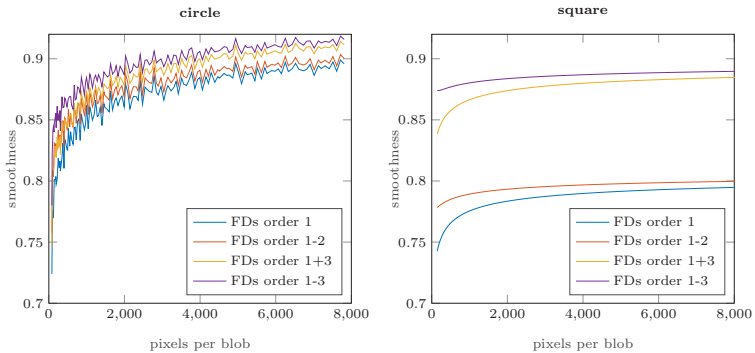


Figure 3.4: Smoothness values for circles (left) and squares (right) with different areas. The different lines represent different combinations of the three lowest order Fourier Descriptors for calculating the smoothness.

It can be observed that in the graph representing the circle, the smoothness value is comparatively low for small resolutions and increases for higher ones. The smoothness of the square on the other hand stays comparatively constant for the range of resolution, when the second order Fourier Descriptors are included. When the first and the first and third order Fourier Descriptors are used the graph shows similar properties to the circle's.

The use of different cameras, the positioning of the camera to the falling pellets and the size of the pellets have an influence on the resolution of the detected pellets images. Consequently, the influence of the resolution on the smoothness have to be taken into account when making assumptions based on the result of Equation 3.4.

Required order of Fourier Descriptors to calculate smoothness Depending on the aimed form and the needed accuracy of the detection of faulty pellets different orders of Fourier Descriptors for the nominator of Equation 3.4 will be necessary.

Considering again the graphs in Figure 3.4, it is observable for both objects, that the second order Fourier descriptors have a relatively limited influence on the smoothness value. In both cases the addition of the third order Fourier Descriptors yields a higher change of the smoothness value, as they contain information on the rectangularity of the object. For the circle this is probably due to the quantisation with square pixels.

The above implies that generally the use of first and third order Fourier Descriptors for the calculation of the smoothness will return a high value for smooth pellets, as they will most likely result in a rectangular, circular or oval projection in the images.

If high precision for detection deviations from spherical and ellipsoidal shapes is the goal, the use of only first order Fourier Descriptors is suggested by Figure 3.4, as they contain only a low quadratic share in comparison to e.g. cylindric ones. Any information in the third order Fourier Descriptors will be considered erroneous and will lead to lower smoothness values.

4 Results

In order to evaluate the application of the smoothness value on real images of pellets, we took 30 images of 11 different samples of pellets. The images were taken with the setup in Figure 2.1. Of every detected pellet the smoothness value was calculated. Figure 4.1 shows the result of every sample for different combination of orders of Fourier Descriptors as boxplots. Each boxplot shows the median as red bar, the 16%- and 84% percentile as the edges of the blue box and the minimum and maximum value as black bars. Thus, the blue box contains 68% of the detected pellets.

Samples one to nine are roughly the same size with about 4000 pixels per blob. Samples ten and eleven have a size of around 900 pixels per blob. Samples one and two are also depicted in Figure 1.1. Samples two, six and nine are considered bad quality, as they have appendices

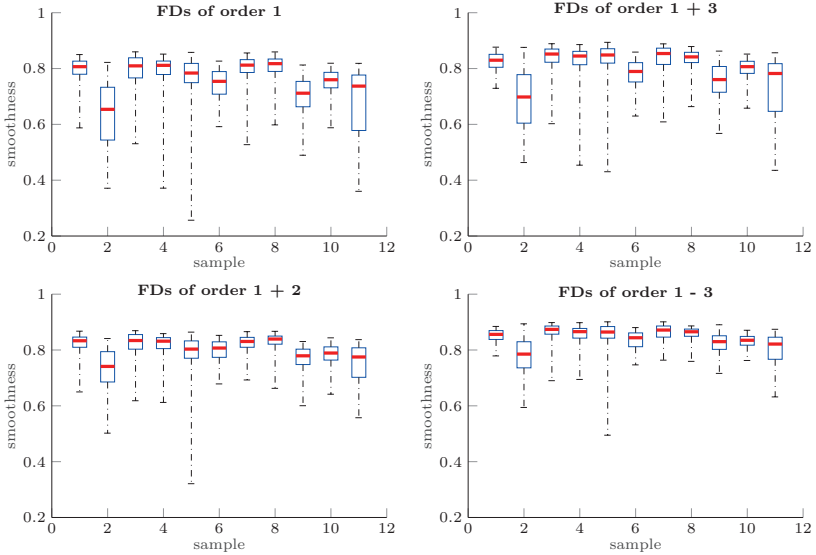


Figure 4.1: Boxplots of the smoothness values of different samples with a different combination of lower order Fourier Descriptors. The red bars show the median, the upper and lower edges of the blue boxes the 16%- and 84% percentiles and the black bars the maximum and minimum value of the smoothness values for every sample.

or show other kinds of deformations. Sample 11 contains pellets of high and low quality.

As implied by the results of the idealized forms above, the smoothness value shows the best results for the evaluation of production quality with the use of first and first and third order Fourier Descriptors. This is observable in Figure 3.3, as the 84% percentile of the low quality pellets is lower than or close to the 16% percentile of the high quality pellets in the upper two graphs. In the lower two graphs a clear distinction of samples 6 and 9 from the others is not possible.

For samples two and five the best results are obtained using first and third order Fourier Descriptors. This is due to the fact that those samples contain pellets with cylindrical shapes.

The above confirms that the best results for the evaluation of a pellets

quality is obtained using first or first and third order Fourier Descriptors for the calculation of the smoothness value.

5 Conclusion

Fourier Descriptors are a very useful tool to analyse the complexity of the contour of a blob. By analysing the information in the lower order Descriptors a single value can be calculated for a rough classification of the production quality. Flaws from the production process, as appendices, can easily be detected, with the proposed value for smoothness.

Acknowledgement

This work has been supported by the Pro²Future-K1 Center within the framework of the Austrian COMET-K1 Programme.

References

1. ECON, "Underwater pelletizing eup, process water and drying system ewt." [Online]. Available: <https://www.econ.eu/eup.html>
2. Y. Lin, E. Bilotti, C. W. Bastiaansen, and T. Peijs, "Transparent semi-crystalline polymeric materials and their nanocomposites: A review," *Polymer Engineering & Science*, vol. 60, no. 10, pp. 2351–2376, aug 2020.
3. B&R, "Vison systems." [Online]. Available: <https://www.br-automation.com/en-gb/products/vision-systems/>
4. R. E. W. Rafael C. Gonzalez, *Digital Image Processing, Global Edition*. Pearson, 2018. [Online]. Available: https://www.ebook.de/de/product/30712961/rafael_c.gonzalez_richard_e.woods_digital_image_processing_global_edition.html
5. A. Jain, *Fundamentals of digital image processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
6. K. K. L. Kandlbauer, R. Sarc, "Charakterisierung von partikeln gemischten gewerbeabfalls über partikeldescriptoren zur sensorischen messung der korngroesse," *Recy & DepoTech 2020*, 2020.

Fiber-Coupled MEMS-based NIR Spectrometers for Material Characterization in Industrial Environments

Robert Zimmerleiter¹, Paul Gattinger¹, Kristina Duswald¹,
Thomas Reischer², and Markus Brandstetter²

¹ RECENDT – Research Center for Non-Destructive Testing GmbH,
Altenbergerstraße 69, 4040 Linz

² Metadynea Austria GmbH,
Hafenstraße 77, 3500 Krems an der Donau

Abstract Recently emerged near-infrared (NIR) spectrometers based on micro-electromechanical systems (MEMS) are a highly compact, rugged and cost-efficient alternative to infrared spectrometers conventionally used in industrial environments. The majority of the devices currently available are designed for measurements in diffuse reflection geometry in close contact with the sample, with built-in low-power halogen light sources – usually intended for consumer applications. However, for most material characterization applications in the industrial environment such a measurement configuration is neither feasible, nor practical. Using light transmitting optical fibers in combination with a measurement probe and a high power fiber-coupled light source is often preferable. In this contribution we compare various fiber-coupled NIR-spectrometers based on different MEMS technologies and demonstrate the applicability of MEMS spectrometer technology in industrial environments.

Keywords NIR Spectroscopy, MEMS Technology, inline monitoring, fiber-coupled, partial least squares regression

1 Introduction

Near infrared (NIR) spectroscopy is one of the most frequently used tools in industry to gain real-time information about the chemical or

physical state during production processes. It allows for accurate monitoring and tight process control to achieve maximum process efficiency and product quality. Besides its inline-capability, non-destructive nature and low running costs, it allows to measure multiple process parameters simultaneously through the use of multivariate data analysis techniques [1]. One major obstacle for the implementation of NIR spectroscopy remains the relatively high hardware costs of conventionally used process spectrometers. Modern MEMS-based NIR spectrometers offer a rugged and compact alternative to conventional process spectrometers at a fraction of the price (about 1/10th) without any moving parts, which is an additional advantage in industrial environments. This relatively new technology has already proven capable of replacing conventional spectrometers in certain demanding applications [2–4]. However, most of the MEMS-based NIR spectrometers available today are designed for measurements in reflection geometry using built-in low power halogen light sources. Performance comparisons for the different types of these spectrometers can be found in the literature [5]. While this measurement geometry is suitable for solid samples that can be directly accessed e.g. in handheld measurements, it is often not suitable or practical for industrial applications. Often high power light sources in combination with inline measurement probes connected via light transmitting fibers are necessary or much more convenient.

In this contribution different fiber-coupled MEMS-based spectrometers are compared in terms of covered wavelength range, metrological properties and overall performance. Furthermore, the applicability of MEMS-based spectrometers in an industrial setting is demonstrated. This should shed some light on the upsides and downsides of the different available technologies and measurement principles when used in a fiber-coupled configuration.

2 Experimental

Seven individual ultra-compact fiber-based NIR-spectrometers that use three different MEMS-based technologies for spectral light acquisition were tested. Two spectrometers rely on digital light processing (DLP), one on Fourier transform infrared (FTIR) spectroscopy and four on tunable Fabry-Pérot (FP) filter technology. All spectrometers were simply

connected via USB to a PC and require no additional power source. For the comparison of these different MEMS-spectrometers, a simple setup was built consisting of a fiber-coupled halogen light source, a sample holder and two low-OH SMA-coupled fibers for connecting the light source and MEMS-spectrometer, respectively. A schematic drawing of the experimental setup is shown in Fig.2.1.

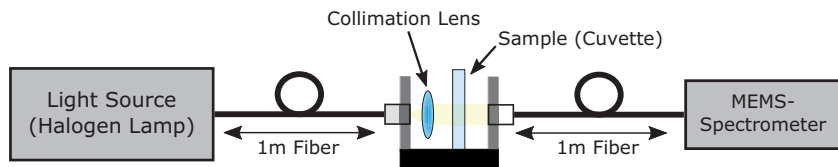


Figure 2.1: Schematic drawing of the measurement setup used to compare the different MEMS-based spectrometers. For details see text.

The fiber on the left was kept the same for all measurements (low-OH, 600 μ m core diameter). The same fiber was used on the right for the MEMS-FTIR and the FP spectrometers, but was exchanged for a 7-to-1 round to linear fiber for the DLP spectrometers (low-OH 200 μ m core diameter for each fiber). This was done to ensure maximum light coupling into the grating-based DLP-spectrometers which have an entrance slit that be illuminated much more effectively using a linear arrangement of fibers on the spectrometer end. For the measurement of the 100%-lines (Fig.3.1), no sample was inserted and the light of the fiber-coupled light source was measured directly, while for the sample spectrum shown in Fig.3.2, ethanol was put into a Infrasil quartz cuvette with a light path inside the sample of 1mm. Measurement settings for each spectrometer were individually set to get an acquisition time of 2 seconds for each spectrum.

3 Results and Discussion

The setup shown in Fig. 2.1 without any sample cuvette was used to acquire 101 individual spectra for each spectrometer. These spectra were used to calculate 100%-lines $A^{100\%}$, where the first measured spectrum

S_0 is used as background which results in 100 lines $A_i^{100\%}$ given by:

$$A_i^{100\%} = \log_{10}(S_0) - \log_{10}(S_i) \tag{3.1}$$

The resulting 100%-lines for all tested spectrometers are plotted in Fig.3.1. Therein also the root mean square (RMS) of the 100%-lines for each of the spectrometers is indicated to allow for a better performance comparison. The RMS is obtained by averaging all the squared entries for all 100 obtained lines $A^{100\%}$:

$$\text{RMS} = \frac{1}{100} \sum_{i=1}^{100} \left[\frac{1}{N_\lambda} \sum_{j=1}^{N_\lambda} [A_i^{100\%}(\lambda_j)]^2 \right]^{1/2} \tag{3.2}$$

where N_λ is the total number of spectral points acquired with the respective spectrometer.

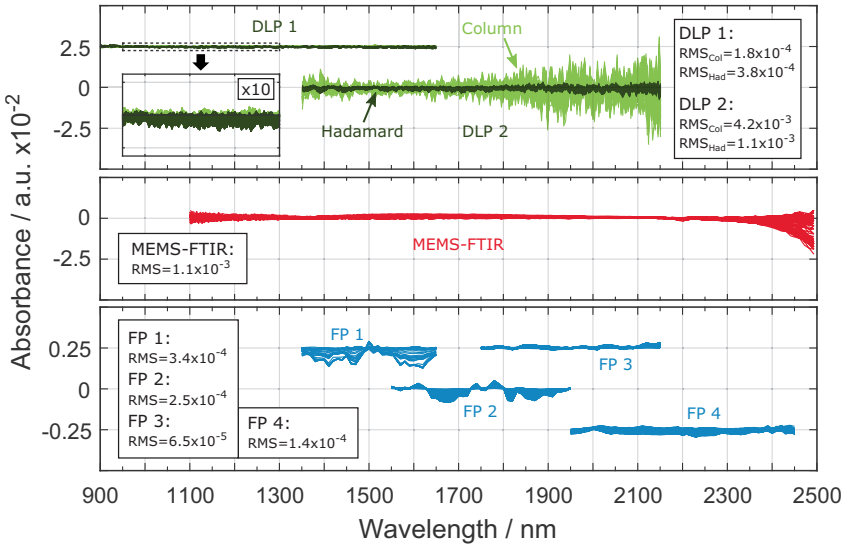


Figure 3.1: 100%-lines recorded with the different MEMS NIR spectrometers. The data for some of the spectrometers have been vertically shifted (DLP 1, FP1, FP 3, FP 4) for better visibility. RMS values for the spectrometers are given to allow for easier performance comparison.

The uppermost graph in Fig.3.1 shows the 100%-lines for the two investigated DLP-spectrometers. The first one (DLP 1) covers the wavelength region from 900nm-1650nm which is a very common region for NIR-spectrometers since it is the spectral response range of typical InGaAs detectors. The second DLP device (DLP 2) covers the spectral region from 1350nm-2150nm which requires an extended InGaAs detector. Also the sizes of both devices differ, with DLP 2 having a size of only $59.5 \times 47.5 \times 24.5 \text{ mm}^3$ while DLP 1 has the dimensions $75 \times 58 \times 26.5 \text{ mm}^3$ ($\approx 66\%$ bigger). Due to the measurement principle used in both of these devices, different measurement modes can be used, namely the *Column* mode and *Hadamard* mode. The latter has a multiplex advantage resulting from collecting light from 50% of all available wavelengths at a time using Hadamard patterns on the digital mirror device (DMD) instead of single columns that project only a small wavelength band onto the photodetector. Both measurement modes were tested using the same measurement parameters (resolution, exposure time etc.) and are shown in Fig.3.1 in bright green (Column) and dark green (Hadamard), respectively. It can be seen that the shift to longer wavelengths and the more compact form factor does not come without trade-offs when it comes to noise performance. The RMS of the 100%-line is one or two orders of magnitude smaller for DLP 1 for Hadamard and Column mode, respectively. Interestingly, the multiplex advantage does not affect the performance of DLP 1, in fact the noise performance when measuring in Column mode is even better than in Hadamard mode, suggesting other noise sources that are not influenced by the multiplex advantage e.g. influence of temperature drifts. A clear demonstration of the benefits of multiplexing can be seen in the data for DLP 2, where the detector noise is the main noise source. Here the RMS is about four times better when using Hadamard mode instead of Column mode.

The middle graph visualizes the data obtained with the MEMS-FTIR device. This device covers the broadest spectral range of all tested spectrometers, namely 1100nm-2500nm using an extended InGaAs detector and has the largest size of $76 \times 57 \times 49 \text{ mm}^3$. Also the FTIR measurement principle used in this device has multiplex advantage, which allows for good noise performance similar to the performance using the multiplexing Hadamard mode with the DLP 2 spectrometer. The larger noise level at wavelengths above 2300nm can be partially explained by

the low light intensity in this wavelength range. Not only does the halogen light source (2850K color temperature) emit most of its intensity at shorter wavelengths, but also the used low-OH fibers show significant light absorption for wavelengths longer than 2300nm.

The third kind of tested MEMS-spectrometers relies on tunable FP-filters. This measurement principle results in the smallest wavelength coverage, which is why four individual spectrometer modules are required to cover the wavelength range from 1350nm-2450nm. An advantage of this technology is that it can be realized in an extremely compact housing (25x25x25mm³) and also has the lowest hardware costs. As visible in the bottom graph in Fig.3.1, another upside of this technology is the superior noise performance over the whole wavelength region. All tested FP-spectrometers have RMS values in on the order of 10⁻⁴ or lower similar to the performance of DLP 1, but also at longer wavelengths where commonly lower light intensities are encountered.

To get a better idea about how the measurement technology shapes acquired absorption spectra ethanol was used as a sample since it shows multiple absorption bands across the whole NIR spectral range. As background the measured spectrum without any sample was used. As a reference, the absorption spectra of ethanol measured with a Bruker Vertex 70 benchtop FTIR spectrometer using the same cuvette (without any fibers and using its internal light source) are shown in gray. The acquired data can be seen in Fig.3.2.

Spectra acquired with the individual MEMS-spectrometers overlap very nicely and show the same features at the same spectral positions. Furthermore good agreement with the reference spectrum measured using the benchtop FTIR was achieved. The spectra recorded with the FP-filter spectrometers (blue) tend to show flattened spectral features due to the lower spectral resolution of approximately 15nm-30nm, compared to the MEMS-FTIR (5nm-13nm) and the DLP spectrometers (10nm-15nm).

The biggest differences between the recorded spectra can be seen in the wavelength region above 2200nm. Here the FP-based spectrometer strongly underestimates the strength of the absorption of ethanol. However, the main spectral features can still be seen albeit with significantly lower resolution, such as the shoulder at 2350nm resulting from the sharp absorption band at this wavelength visible in both the MEMS-FTIR and benchtop FTIR spectrum. The lack of multiplex ad-

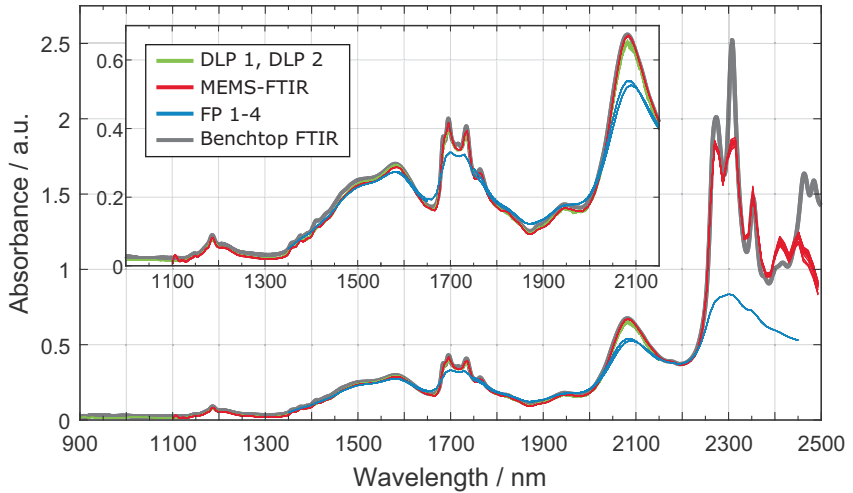


Figure 3.2: Absorption spectrum of ethanol recorded with the different MEMS NIR spectrometers (green, red, blue) and a benchtop FTIR (gray). The inset shows a zoom of the wavelength region 1000nm-2150nm.

vantage that comes with the FP-based measurement technique can also cause problems in this wavelength region. Since the light intensity is very low due to low emission from the light source, high absorption from the used fibers and additionally strong absorption from the ethanol sample, the measured signal is rather weak and non-linearities of the used detector or electronics can cause deviations from the real absorption spectrum.

Slight differences can also be seen between the MEMS-FTIR and benchtop FTIR caused by the lower resolution of the MEMS-FTIR which leads to the flattening of the sharp absorption bands around 2300nm. Furthermore, the much lower light intensity at long wavelengths due to the absorption in the used fibers (not present for the benchtop device) does not allow for a good measurement of the ethanol absorption bands above 2400nm. This problem could be solved e.g. by using more expensive zirconium fluoride (ZrF_4) fibers to reduce light attenuation.

4 Inline Process Monitoring using Fiber-Based MEMS-Spectrometer Technology

In this section measurement data from an industrial application of a FP filter based MEMS-spectrometer will be presented to demonstrate the potential of MEMS technology to replace commonly used FTIR-process spectrometers in the industrial environment. As a use-case, NIR-monitoring of a batch-wise melamine formaldehyde resin production plant was chosen. At Metadynea Austria GmbH, the resin production is routinely monitored and controlled in real time using NIR spectroscopy in combination with partial least squares (PLS) regression [6]. The combination of NIR spectroscopy and PLS regression is e.g. used to determine the batch end point. This leads to a significant reduction of necessary offline measurements and helps to avoid out of spec batches.

To demonstrate the usability of FP-filter based MEMS spectrometer technology for the determination of the batch endpoint, the optical fiber was simply switched from the routinely used FTIR spectrometer to a suitable MEMS device and a PLS regression model was calibrated and validated using the data acquired with the FP spectrometer. The data is summarized in Fig.4.1. As can be seen from the data, the measurement quality is very similar for the two different spectrometers. The root mean squared error of prediction (RMSEP), which is an important parameter to judge the measurement performance, is higher for the MEMS spectrometer (0.044) when compared to the FTIR (0.024), but still clearly below the threshold value (0.1) set by the company. This means the low-cost MEMS-based spectrometer can be used to replace the conventional, much larger FTIR spectrometer while still delivering satisfying results.

5 Conclusion

In this contribution, seven individual fiber-coupled MEMS-based NIR spectrometers were compared in terms of their covered wavelength range, resolution and noise performance. The tested MEMS-spectrometers based on different measurement technologies, namely DLP, FTIR and tunable FP-filters. Each technology has its advantages

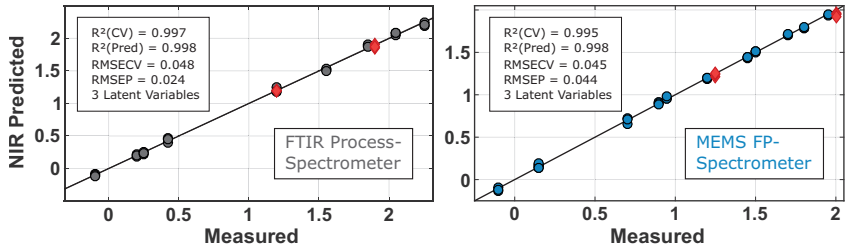


Figure 4.1: Data comparison between a conventional FTIR process spectrometer (left) and a low-cost FP spectrometer (right). Normalized values calculated from the PLS regression models are plotted versus offline measurements. Calibration data is shown in gray (FTIR) and blue (FP), respectively. Validation data is shown in red.

and disadvantages. While MEMS-spectrometers based on tunable FP-filters tend to have the best noise performance, they also have lower spectral resolution and spectral coverage (300nm-500nm) and lack the multiplex advantage that can be exploited with DLP and FTIR technology. The broadest spectral coverage can be achieved using the FTIR measurement technique, while still maintaining very reasonable noise performance, but at higher hardware costs (about 3x). MEMS-spectrometers based on DLP-technology offer a good alternative by offering good spectral coverage of about 800nm while having comparable noise performance to the MEMS-FTIR when the multiplex advantage is exploited by measuring in Hadamard mode. When the right instrument is chosen for the specific application, fiber-based NIR MEMS-spectrometers show huge potential to allow for much more cost effective spectroscopic measurement solutions in the industrial environment without significant sacrifices in measurement performance. For this the presented application for inline batch monitoring at a resin production facility is a good example. There it was successfully shown that the narrow band FP devices were a suitable substitute for the conventional broadband FTIR spectrometer.

Acknowledgements

Financial support was provided by the Austrian research funding association (FFG) under the scope of the COMET programme within the

research project “PhotonicSensing for Smarter Processes (PSSP)” (contract no. 871974). This programme is promoted by BMK, BMDW, the federal state of Upper Austria and the federal state of Styria, represented by SFG.

References

1. E. Stark, K. Luchter, and M. Margoshes, “Near-infrared analysis (NIRA): A technology for quantitative and qualitative analysis,” *Applied Spectroscopy Reviews*, vol. 22, no. 4, pp. 335–399, 1986.
2. R. Zimmerleiter, J. Kager, R. Nikzad-Langerodi, V. Berezinskiy, F. Westad, C. Herwig, and M. Brandstetter, “Probeless non-invasive near-infrared spectroscopic bioprocess monitoring using microspectrometer technology,” *Analytical and Bioanalytical Chemistry*, vol. 412, no. 9, p. 2103–2109, 2020.
3. C. R. Avila, J. Ferré, R. R. de Oliveira, A. de Juan, W. E. Sinclair, F. M. Mahdi, A. Hassanpour, T. N. Hunter, R. A. Bourne, and F. L. Muller, “Process monitoring of moisture content and mass transfer rate in a fluidised bed with a low cost inline mems nir sensor,” *Pharmaceutical Research*, vol. 37, no. 84, 2020.
4. R. Zimmerleiter, E. Leiss-Holzinger, E. M. Wagner, K. Kober-Rychli, M. Wagner, and M. Brandstetter, “Inline biofilm monitoring based on near-infrared spectroscopy with ultracompact spectrometer technology,” *Journal of Pharmaceutical and Biomedical Analysis*, vol. 31, no. 7-8, pp. 9–13, 2020.
5. H. Yan and H. W. Siesler, “Quantitative analysis of a pharmaceutical formulation: Performance comparison of different handheld near-infrared spectrometers,” *Journal of Pharmaceutical and Biomedical Analysis*, vol. 160, pp. 179–186, 2018.
6. P. Geladi and B. R. Kowalski, “Partial least-squares regression: a tutorial,” *Analytica Chimica Acta*, vol. 185, pp. 1 – 17, 1986.

Improvement of Roughness Measurement in Sub-micron Ranges Using Contrast-based Depolarization Field Components

Franziska Pöller¹, Félix Salazar Bloise², Martin Jakobi¹, Jie Dong¹, and Alexander W. Koch¹

¹ Technical University of Munich, Institute for Measurement Systems and Sensor Technology,

Theresienstr. 90, 80333 Munich, Germany

² Universidad Politécnica de Madrid, ETSI Minas y Energía,
Calle de Ríos Rosas 21, 28003 Madrid, Spain

Abstract The characterization of a surface, especially the roughness, is of great importance for industrial applications. There are already various optical methods for roughness measurement, which usually do not consider the depolarization effects of the scattered beam caused by the sample. This work aims out improving the roughness measurement method described in [1] by taking into account depolarization effects, due to the rough sample. For validation, other instruments are employed to compare the results in the measurement interval for which this technique is applied.

Keywords Roughness measurement, optical method, speckle, depolarization, contrast

1 Introduction

Surface roughness characterization is essential in many industrial applications: for monitoring manufacturing processes in the optical industry [2], for determining the roughness of pharmaceutical pills to ensure their effectiveness [3], and for analyzing fatigue in materials [4]. Due to the production of novel structured materials and the need to ensure their functionality without damaging them, the demand for

purely optical characterization of roughness in industry is also increasing. Several methods have been provided for determining the surface roughness. Often profile-based methods, such as the stylus method, are used because they provide reliable results [5]. The diamond tip of the stylus instrument, however, may not be useful for soft surfaces like rubbers and some plastics because it destroys the surface under test. Moreover, these techniques are very time-consuming.

There are also optical techniques, such as the white light interferometer [6] or the confocal microscope [7]. One advantage of these optical methods is their accuracy, but they are associated with high acquisition and maintenance costs. This motivates research of alternative optical measuring techniques for roughness determination, which are fast, accurate, and cost-effective. At the same time, non-destructive properties must not be neglected. Most of the different optical procedures for roughness measurements do not take into account depolarization effects in the light-matter interaction process [1], which limits the applicability range of these methods. Other techniques based on light scattered by the rough surface, however, take into consideration changes in polarization, extending their possibilities in application [8,9]. Another way to determine the roughness is to consider the depolarization processes in interferometric fringe speckle patterns by analyzing two different contrasts in two interferograms corresponding to the respective polarization field components. For this purpose, the interferogram is spatially evaluated in regions. This has already been the subject of previous investigations [10]. The method promises a higher accuracy compared with [1] of the estimated surface roughness and is valid for small roughness values ($15 \text{ nm} < Rq < 40 \text{ nm}$), where different industrial applications may be found. To verify the presented method, measurements are carried out for different rough samples and compared with measurement results from a stylus instrument, a white light interferometer, and a laser confocal microscope to verify the roughness values measured with the interferometer. The results are also evaluated concerning the objectives used in the confocal microscope and the white light interferometer. To investigate the current applicability of the method in industry, the suitability of the commercial measuring instruments for the presented technique is also evaluated. The roughness of the samples is measured in each case, and the accuracy and applicability of those methods are analyzed.

2 Methodology

The theory of the method is explained below, before further detailing the operation principle of the roughness measurement by contrast-based depolarization field components.

2.1 Theory

The presented method is based on two-beam interference of two waves. This superposition results in the typical speckle pattern with the interference fringes. The two averaged intensities in the interferogram, $\langle I_A \rangle$ and $\langle I_B \rangle$, visible on the camera can be expressed as follows [10]:

$$\begin{aligned} \langle I_A \rangle &= I_{ys} + I_{ym} + I_{xs} \\ &+ 2\sqrt{I_{ys}I_{ym}} \cdot \exp\left(-\frac{(2ku_{\text{geom}}\sigma)^2}{2}\right) \cdot \cos(u_{\text{geom}}kf(x, y)), \end{aligned} \quad (2.1)$$

$$\begin{aligned} \langle I_B \rangle &= I_{xs} + I_{xm} + I_{ys} \\ &+ 2\sqrt{I_{xs}I_{xm}} \cdot \exp\left(-\frac{(2ku_{\text{geom}}\sigma)^2}{2}\right) \cdot \cos(u_{\text{geom}}kf(x, y)), \end{aligned} \quad (2.2)$$

where $\langle I_A \rangle$ is the intensity on camera part A and $\langle I_B \rangle$ is assigned to camera part B, where the quarter-wave plate is inserted in front of the reference mirror. I_{ys} and I_{xs} are the y component and the x component of the total intensity I_S of the rough surface, where the same is assumed for the total intensity on the reference mirror I_M . u_{geom} is the geometrical factor that describes the scattering at the surface, defined here as $u_{\text{geom}} = 2$, since the rough surface and the reference mirror are reflective. $k = \frac{2\pi}{\lambda}$ corresponds to the wave number with the laser wavelength λ , σ is the standard deviation of the heights on the surface of the target and $f(x, y)$ defines the shape of the reference, which here corresponds to an inclination around the x - or y -axis, resulting in parallel interference fringes. The Michelson contrast C_M , which describes the fringe visibility in the interferogram [1], is composed of the maximum

intensity I_{\max} and the minimum intensity I_{\min} as follows:

$$C_M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (2.3)$$

The two different contrasts C_A and C_B in the interferogram of the two camera parts can be determined under the assumption for the total intensity on the reference mirror formula $I_M = I_{xm} \approx I_{ym}$ [10]:

$$C_A \approx \frac{2\sqrt{I_{ys}I_{ym}} \exp\left(-\frac{(2u_{\text{geom}}k\sigma)^2}{2}\right)}{I_{xs} + I_{ys} + I_M}, \quad (2.4)$$

$$C_B \approx \frac{2\sqrt{I_{xs}I_{xm}} \exp\left(-\frac{(2u_{\text{geom}}k\sigma)^2}{2}\right)}{I_{xs} + I_{ys} + I_M}, \quad (2.5)$$

which result from only one experiment due to the quarter-wave plate being with one half inserted into the reference beam and provide the separated information on the camera. This divided information and the independent knowledge of the intensity components I_{xs} and I_{ys} , where $I_S = I_{xs} + I_{ys}$ applies, is required for the roughness determination of the sample, since the sample depolarizes the light with $I_{xs} \neq 0$. By rearranging the previous formulae the following expression for the standard deviation of the height of the object surface σ results in [10]:

$$\sigma = \frac{\lambda}{4\pi} \sqrt{\ln\left(\frac{4I_M I_S}{(I_{xs} + I_{ys} + I_M)^2 (C_A^2 + C_B^2)}\right)}, \quad (2.6)$$

where σ can be equated with the root mean square (RMS) height $Sq = \sigma$ [11], which allows a comparison of the measurement results with the interferometer setup and the results of the measurements with the commercial measuring instruments. For the validity of the presented method, the standard deviation σ of the height must be Gaussian-distributed on the object surface, the surface itself must be a weak scatterer, i.e., $Rq < \lambda/4$ [12], so that we still find the necessary interference fringes and the speckles are not fully developed, and the rough object surface must be flat.

2.2 Roughness Measurement by Contrast-based Depolarization Field Components

For the measurements, we use a modified Michelson interferometer with a quarter-wave plate (QWP) set with one half in the reference beam (figure 16.1(a)). For the Ar⁺-laser, we choose the wavelength $\lambda = 488 \text{ nm}$ to verify the roughness measurement by contrast-based depolarization field components because it is in the middle of the spectrum of the laser and has a high intensity. To guarantee the initial polarization conditions, i.e., linear polarization of the laser light, a polarizer (PO) is placed in front of the Ar⁺ laser. A 50:50 beam splitter (BS) divides the beam into an object and reference path. The laser beam, linearly polarized in the y -direction, is directed to the tilted plane mirror (M) in the reference path, where a QWP is placed in one of the two halves, called here part B, whose optical axis is set to 45° , which enables the required two different polarization states of the reference beam, providing the different information about part A and part B. The rough surface (RS) scatters the laser light, creating different interferometric fringe patterns on both camera parts A and B. The camera (Photonfocus) has a resolution of 1312×1082 pixels (12 bit) with a pixel size of $8 \mu\text{m} \times 8 \mu\text{m}$. The achromatic lens (AL) and the adjustable aperture (AA) produce a focused image of the object (RS) and the reference (M) on the camera. The magnification of the measuring system is 1.5. To determine the RMS height S_q of the rough surface, the roughness measurement by contrast-based depolarization field components consists of three measurements. In the first measurement, we determine the total intensity of the rough surface I_S , for which only the object path in figure 16.1(a) is considered. The light backscattered from the surface (RS) is recorded by the camera (see figure 16.1(b)) and we can calculate the intensity I_S reflected from the rough surface. In a second measurement, we measure the total intensity of the reference mirror (M) I_M , assuming approximately equal intensity on both parts of the plane mirror $I_M = I_{xm} \approx I_{ym}$. To ensure that the intensity of the reference mirror is the same, we place a glass in front of the mirror (in part A) for additional radiation absorption, since the QWP absorbs about 8 % of the radiation, in all directions of propagation. The glass must have the same absorption properties as the QWP. For the measurement, only the reference path in the Michelson interferometer is

considered in figure 16.1(a) and the intensity reflected from the mirror (M) is recorded and shown in figure 16.1(b). In the third measurement, both paths of the Michelson setup in figure 16.1(a) are now considered, producing two interference patterns with different contrasts C_A and C_B generated on the camera image (see figure 16.1(b)). These two contrasts are calculated according to equation 2.3 by the Michelson contrast for part A and part B. If now the measured values for I_S , I_M , C_A , and C_B are inserted into equation 2.6, the RMS height Sq of the rough surface results.

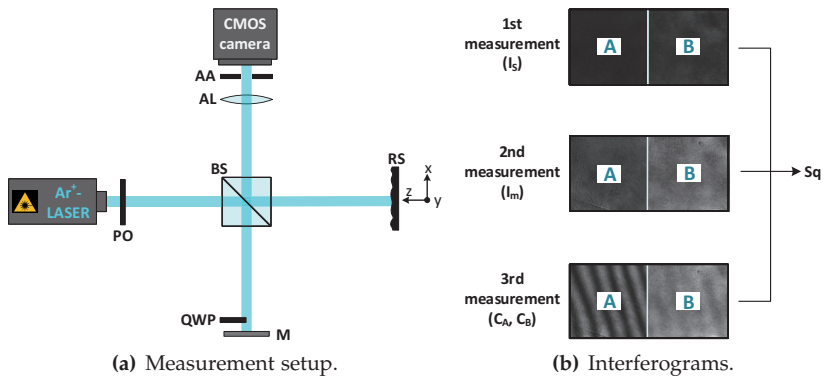


Figure 2.1: Interferometer setup and measuring procedure of the roughness measurement by contrast-based depolarization field components for the surface under test with RMS roughness $Rq_{stylus} = 28 \text{ nm}$ at a wavelength of $\lambda = 488 \text{ nm}$.

3 Experimental Results

To verify the suitability of the measurement instruments, the roughness of five different samples is evaluated using the theory explained above. The measurements with the interferometer setup in figure 16.1(a) were performed with the wavelength $\lambda = 488.0 \text{ nm}$. Since the Michelson interferometer results (according to equation 2.6) may vary slightly from one part of the surface to another, the $Sq_{\text{michelson}}$ values from nine different areas for each sample were calculated and averaged. Besides, in order to compare this new technique with the basic methodology

shown in [1], a calculation of the roughness without considering the depolarization of the scattered light was performed (Sq_{MWODSL}). To validate the proposed procedure, measurements using other instruments were carried out. In addition, these measurements also serve to analyze the suitability of each technique for industrial applications. To this objective, a stylus instrument, a white light interferometer, and a laser confocal microscope were used.

We measured the samples twenty times in different directions using a stylus instrument (SURFCOM FLEX 50 A with a measuring force of 0.75 mN and a stylus tip radius of 2 μm) and then averaged them. The measurements with the white light interferometer (Wyko NT3300) have been performed in the VSI mode [13], i.e., by vertical scanning interferometry. To generate the magnification of a 40X objective, a 20X objective and an additional adjusted filter with a field of view (FOV) of 2 were used, and for the 10X objective, a 20X objective and an additional adjusted filter with a FOV of 0.5 were placed. With the confocal microscope (SENSOFAR PL μ 2300), two different microscope objectives were employed (50X and 100X).

By analyzing the experimental results shown in table 16.1 we find that the new method gives different values ($Sq_{michelson}$) than those obtained by the technique without considering depolarization effects Sq_{MWODSL} [1]. These series of values better confirm with some of the other methods employed. The stylus instrument and the confocal microscope (100X) seem to be adequate to compare the results. The Michelson interferometer working without depolarized scattered light (MWODSL) [1] provides values that do almost not change in the interval of the roughness values of samples 1 to 5. On the contrary, the $Sq_{michelson}$ varies approximately for S1 to S3 according to the stylus and the confocal microscope (100X). It shows that the method proposed may be a good estimate for small roughnesses up to 40 nm, which is an indicator for considering the depolarization effects of rough surfaces. It is interesting to note the discrepancies of the results among the techniques used. The results vary with the method and also with respect to the magnification of the objectives, then it is necessary to be careful when comparing experimental values. It is clear from table 16.1 that the confocal microscope with a 100X objective, and the stylus (radius of 2 μm), are suited to compare results within the investigated interval. The white light interferometer, even though has a very good resolution,

it is not useful, in the context to validate the new technique. Exemplary captured images of sample 2 of the confocal microscope with a 50X objective and of the white light interferometer with a 40X objective are shown in figure 3.1.

Table 16.1: Experimental results for five samples (in nm) measured by the stylus profiler, the white light interferometer, the confocal microscope, the Michelson interferometer without considering depolarization of the scattered light (MWODSL), and the interferometer setup considering depolarization effects ($Sq_{\text{michelson}}$). The uncertainties of the results and the measuring instruments are also given.

	S 1	S 2	S 3	S 4	S 5
Rq_{stylus}	28 ± 2	31 ± 2	37 ± 2	45 ± 2	116 ± 4
$Sq_{\text{white light (10X)}}$	9 ± 1	6 ± 1	52 ± 1	91 ± 1	260 ± 1
$Sq_{\text{white light (40X)}}$	6 ± 1	4 ± 1	21 ± 1	43 ± 1	198 ± 1
$Sq_{\text{confocal (50X)}}$	15 ± 1	18 ± 1	29 ± 1	47 ± 1	106 ± 1
$Sq_{\text{confocal (100X)}}$	26 ± 1	28 ± 1	52 ± 1	49 ± 1	116 ± 1
Sq_{MWODSL}	40 ± 2	41 ± 2	44 ± 2	44 ± 2	49 ± 1
$Sq_{\text{michelson}}$	27 ± 1	31 ± 1	32 ± 2	28 ± 3	31 ± 4

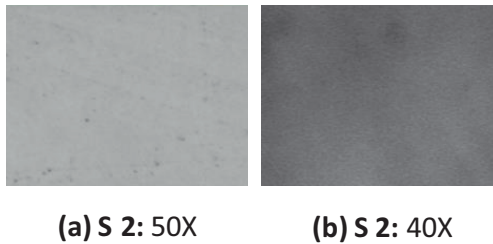


Figure 3.1: Captured images for sample 2 of the confocal microscope with the objective 50X ((a)) and of the white light interferometer with the objective 40X ((b)).

To determine the uncertainties of the experimental results, the expanded uncertainty is calculated [14]:

$$U(Sq_{\text{michelson}}) = k_{\text{cov}} u_c(Sq_{\text{michelson}}) \tag{3.1}$$

with the coverage factor $k_{\text{cov}} = 2$ [10] and the combined standard un-

certainty $u_c^2(Sq_{\text{michelson}})$ of the method [14]:

$$u_c^2(Sq_{\text{michelson}}) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i), \quad (3.2)$$

where f corresponds to the RMS height $Sq_{\text{michelson}}$ of the surface and $u(x_i)$ is the standard uncertainty for each input λ , I_M , I_S , C_A , and C_B (see equation 2.6). The uncertainties of the commercial measuring instruments, which are included in table 16.1, are given by the instrument and are determined according to the manufacturer's information.

From the investigation shown, we can conclude that the new technique, which includes depolarization effects, improves the results obtained with the usual Michelson interferometer. The confocal microscope may also be suitable for industrial applications due to its resolution in the same interval ($15 \text{ nm} < Rq < 40 \text{ nm}$) but is associated with high acquisition costs of tens of thousands of euros. In the same way, the white light interferometer has great capabilities in a wide interval of roughnesses, but at significant costs. Thus, if only low roughness values of $< 40 \text{ nm}$ are to be examined on materials, the modified Michelson interferometer may be preferable after a cost estimate.

4 Summary and Outlook

By the roughness measurement considering depolarization processes in interferometric fringe speckle patterns by analyzing two distinct contrasts in two patterns corresponding to the respective polarization field components, higher accuracy of the estimated surface roughness is possible for small roughnesses in the range of $Rq = 15 \text{ nm}$ to $Rq = 40 \text{ nm}$. Commercially available measuring instruments, such as the confocal microscope, can be used for this method, and thus the technique can be currently applied in industry. Due to the short measurement time of a few seconds and the increased accuracy by considering the depolarization of the sample, the presented method is suitable for all reflective materials as a cost-effective method for optical characterization of surfaces in industry and as a supplement to (optical) measurement systems for roughness measurement.

References

1. P. J. Chandley, "Determination of the standard deviation of height on a rough surface using interference microscopy," *Optical and Quantum Electronics*, vol. 11, no. 5, pp. 407–412, 1979.
2. G. Kloppenburg, M. Knöchelmann, and A. Wolf, "Additive Fertigung transparenter Optiken," in *Additive Manufacturing Quantifiziert*, R. Lachmayer and R. B. Lippert, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 163–174.
3. F. Salazar Bloise, L. Aulbach, M. Jakobi, and A. W. Koch, "Rauheitsmessung an pharmazeutischen Tabletten mittels Angularer Speckle-Korrelation," *tm - Technisches Messen*, vol. 81, no. 6, pp. 289–295, 2014.
4. P. Zuluaga-Ramírez, Á. Arconada, M. Frövel, T. Belenguer, and F. Salazar, "Optical sensing of the fatigue damage state of CFRP under realistic aeronautical load sequences," *Sensors*, vol. 15, no. 3, pp. 5710–5721, 2015.
5. A. Wennerberg, R. Ohlsson, B.-G. Rosén, and B. Andersson, "Characterizing three-dimensional topography of engineering and biomaterial surfaces by confocal laser scanning and stylus techniques," *Medical Engineering & Physics*, vol. 18, no. 7, pp. 548–556, 1996.
6. M. J. Giraldez, C. García-Resúa, M. Lira, M. E. C. D. Real Oliveira, and E. Yebra-Pimentel, "White light interferometry to characterize the hydrogel contact lens surface," *Ophthalmic & Physiological Optics*, vol. 30, no. 3, pp. 289–297, 2010.
7. X. Teng, F. Li, and C. Lu, "Visualization of materials using the confocal laser scanning microscopy technique," *Chemical Society Reviews*, vol. 49, no. 8, pp. 2408–2425, 2020.
8. L. Liu, X. Li, and K. Nonaka, "Light depolarization in off-specular reflection on submicro rough metal surfaces with imperfectly random roughness," *The Review of Scientific Instruments*, vol. 86, no. 2, pp. 1–9, 2015.
9. Z. Liu, Q. Zhong, X. Yu, Q. J. Wang, and Y. Zhang, "Surface roughness measurement by depolarization method," *Applied Optics*, vol. 54, no. 18, pp. 5686–5690, 2015.
10. F. Pöller, F. Salazar Bloise, M. Jakobi, S. Wang, J. Dong, and A. W. Koch, "Non-Contact Roughness Measurement in Sub-Micron Range by Considering Depolarization Effects," *Sensors*, vol. 19, no. 10, pp. 1–12, 2019.
11. I. Yamaguchi and K. Kobayashi, "Measurement of surface roughness by speckle correlation," *Opt. Engineering*, vol. 43, no. 11, pp. 2753–2761, 2004.

12. R. Jones and C. Wykes, *Holographic and speckle interferometry: A discussion of the theory, practice, and application of the techniques*, 2nd ed. Cambridge: Cambridge University Press, 1989, vol. 6.
13. J. Czajkowski, T. Prykäri, E. Alarousu, J. Palosaari, and R. Myllylä, "Optical coherence tomography as a method of quality inspection for printed electronics products," *Optical Review*, vol. 17, no. 3, pp. 257–262, 2010.
14. "JCGM 100:2008: Evaluation of measurement data - Guide to the expression of uncertainty in measurement," 2008.

Multimodal OCT Imaging

Bettina Heise¹ and Ivan Zorin¹ and Guenther Hanneschlaeger, and
Markus Brandstaetter¹

RECENDT- Research Center for Non-Destructive Testing GmbH
Altenberger Str 69, 4020 Linz, Austria

Abstract We demonstrate OCT as an inspection tool for monitoring and validating 3d printed AM specimens. We will discuss the potential and challenges in this applications. Furthermore, we illustrate extensions like spectroscopy and speckle imaging.

Keywords Optical Coherence Tomography, imaging, multimodal, speckles, quality inspection, reliability, additive manufacturing

1 Introduction

Multimodal imaging has emerged in recent years. Its advantage is the agglomeration of information through image and data fusion, which allows deeper insights into structural, functional or chemical information of samples. In addition, different spatial or temporal scales can be combined. Here, we mainly limit ourselves to imaging by optical coherence tomography (OCT) to extract structural information from samples and to enable optical characterization in a non-destructive and not hazardous way. The combination of OCT with spectroscopic information is a valuable tool, as it provides insights into the chemical and structural composition. Furthermore, with the advent of novel light sources, such as supercontinuum sources; covering spectral ranges from the visible to the near and mid infrared, this range can be facilitated in a very comprehensive way [1]. Additionally, also combinations of spectral probing range of OCT in near infrared and THz imaging are exploitable for advanced material characterization. OCT imaging itself also allows the use of different contrast mechanisms partly based on statistical optics approaches [2]. Therefore, the almost ubiquitous speckles in OCT

imaging are not only noise but can also serve as a source of information. Their variance, higher order moments or local autocorrelation can provide information about the dynamics inherent in the sample under investigation [3]. We will present our latest results with a multimodal OCT system combining structural and spectroscopic material characterization. In addition, statistical optics/speckles based characterization for scattering samples with dynamics on different time scales will be presented.

2 Methods

OCT imaging is currently mainly realized in Fourier Domain OCT (FD-OCT) configurations due to its increased sensitivity. Both versions, Spectral Domain OCT (SD-OCT) and Swept Source OCT (SS-OCT) are the most predominant versions nowadays. However, to extend these FD-OCT setting to other spectral ranges is challenging. However such extension offer promising combinations and multi-modal imaging approaches, especially such as spectroscopy or speckle (variance) imaging. We will demonstrate comparatively particular features of FD-OCT in NIR and MIR spectral range. Thereby we can cover a spectral range covering from about 2 to 5 μm currently, beyond the conventional NIR range of 0.8 to 1.5 μm .

3 Results

In particular the inspecting and monitoring additive manufacturing (AM) products and processes already during manufacturing or in intermediate states by OCT seems very promising currently to introduce it with ab industrial background. The 3d printing of layers in micron scales represents a feasible approach for OCT imaging. The 3d printed materials hereby comprehends on one hand polymeric substances, on the other hand also the wide range of ceramics 3D printed specimens is of interest. OCT can enable to monitor in future such printing, or allows to inspect the green parts. That give valuable insights on micro-defects already before e.g. sintering processes start. We will demonstrate different OCT settings and results for visualization and inspection of different 3d printed materials exhibiting micro-defects. How-

ever, one has also to know the specific features of OCT imaging, where pigments and fillers can modify appearance of cross-sectional images. This can result that defects are mitigated or suppressed in their visibility by scattering sites. It is essential therefore to consider the scattering features of specimens under test to define and select suitable wavelength and probing ranges for OCT imaging and sensing.

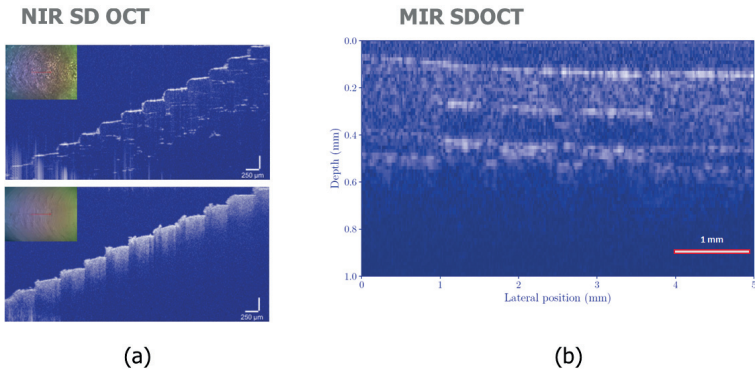


Figure 3.1: OCT cross-sectional images (B-scans), captured from 3d printed specimens. They represent examples both for (a) 3d filament printed polymer samples and (b) lithographically AM ceramics.

4 Acknowledgement

We acknowledge the funding of this work by the Austrian FFG program within FFG-project DIQACAM. Additionally, this work was supported by the strategic economic- and research program Innovative Upper Austria 2020 of the province of Upper Austria and the European Union's Horizon 2020 research and innovation programme under grant agreement No 820661 FLOIM. Furthermore, we would like to thank Lithoz GmbH in Vienna for providing the ceramics test samples.

References

1. I. Zorin, R. Su, B. Heise, B. Lendl, and M. Brandstetter, "Correlative infrared optical coherence tomography and hyperspectral chemical imaging," *J.*

- Opt. Soc. Am. A*, vol. 37, no. 9, pp. B19–B26, Sep 2020. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-37-9-B19>
2. M. Münter, M. vom Endt, M. Pieper, M. Casper, M. Ahrens, T. Kohlfaerber, R. Rahmanzadeh, P. König, G. Hüttmann, and H. Schulz-Hildebrandt, “Dynamic contrast in scanning microscopic oct,” *Opt. Lett.*, vol. 45, no. 17, pp. 4766–4769, Sep 2020. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-45-17-4766>
 3. B. Heise, S. E. Schausberger, B. Buchroithner, M. Aigner, I. Milosavljevic, P. Hierzenberger, S. Bernstein, S. Häuser, A. Jesacher, S. Bernet, M. Ritsch-Marte, and D. Stifter, “Full field optical coherence microscopy for material testing: contrast enhancement and dynamic process monitoring,” in *Optical Methods for Inspection, Characterization, and Imaging of Biomaterials*, P. Ferraro, M. Ritsch-Marte, S. Grilli, and D. Stifter, Eds., vol. 8792, International Society for Optics and Photonics. SPIE, 2013, pp. 194 – 199. [Online]. Available: <https://doi.org/10.1117/12.2020283>

A High Quality Image Stitching Process for Industrial Image Processing and Quality Assurance

Rolf Hoffmann

Technische Universität Ilmenau, Group for Quality Assurance and Industrial Image Processing,
Gustav Kirchhoff Platz 2, 98693 Ilmenau, GERMANY

Abstract The size of the recording area of a camera is limited. The resolution of a camera image is also limited. To capture larger areas, a wide angle lens can be used, for example. However, the image resolution per unit area decreases. The decreased image resolution can be compensated by image sensors with a higher number of pixels. However, the use of a high pixel number of image sensors is limited to the current state of the art and availability of real image sensors. Furthermore the use of a wide angle lens has the disadvantage of a stronger distortion of the image scene. Also the viewing direction from a central location is usually problematic in the outer areas of a wide angle lens. Instead of using a wide angle lens, there is still the possibility to capture the large image scene with several images. This can be done either by moving the camera or by using several cameras that are positioned accordingly. In case of multiple image captures, the single use of the required image is a simple way to evaluate a limited area of a large image scene with image processing. For example, it can be determined whether a feature limited by the size is present in the image scene. The use of this simple variant of a moving camera system or the use of single images makes it difficult or even impossible to use some image processing options. For example, determining the positions and dimensions of features that exceed a single image is difficult. With moving camera systems, the required mechanics add to the effort, which is subject to wear and tear and introduces a time factor. Image stitching techniques can reduce many of these problems in large image scenes. Here, single images

are captured (by one or more cameras) and stitched together to fit. The original smaller single images are merged into a larger coherent image scene. Difficulties that arise here and are problematic for the use in industrial image processing are, among others: the exact positioning of the single images to each other and the actual joining of the images, if possible without creating disturbing artifacts. This publication is intended to make a contribution to this.

Keywords Image processing, lens, image stitching, large image scene

1 Introduction

With a single camera system, capturing a scene is limited in the size of the scene and the resolution of the image. Furthermore, the image may contain distortion and perspective errors that become larger toward the edge of the shot and are especially prevalent with wide angle lenses. Multi camera array systems (Fig. 1.1) can reduce such problems. The difficulty, however, is to combine the individual images from the multi camera system (Fig. 1.2) into a complete image in high quality and without errors. Typical stitching algorithms have problems if the individual images show major distortions, depict different perspectives or do not have exactly the same brightness [1]. A typical stitching result can be seen in Fig 1.3. The idea presented here is to enhance the individual images before the actual stitching process. This creates better preconditions to enable a high quality stitching process.

2 Homogeneity Correction

A problem that can occur with the single images is that within an image a non-uniform illumination can occur (Fig. 2.1). This happens, for example, due to uneven illumination or vignetting.

Proposed solution: Determine the inhomogeneity of the brightness distribution and correct it in LAB color space. Since the brightness distribution is different in different brightness scenarios, it must be

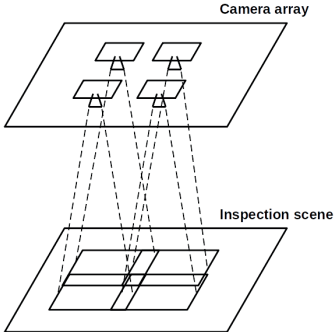


Figure 1.1: Multi camera array system.

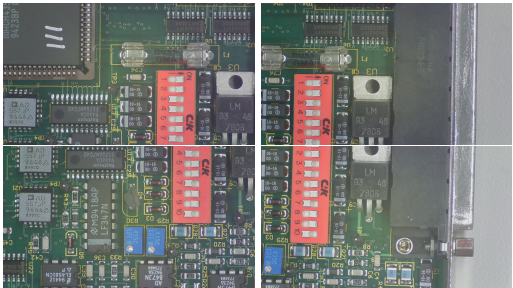


Figure 1.2: Single images from multi camera array system.

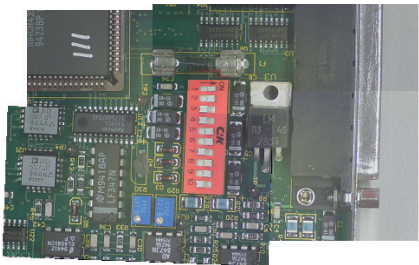


Figure 1.3: Not optimal stitching result.

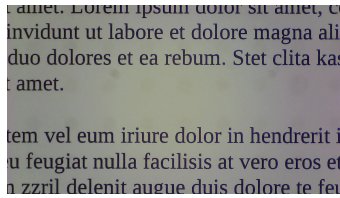


Figure 2.1: Uneven brightness distribution due to uneven illumination and vignetting.

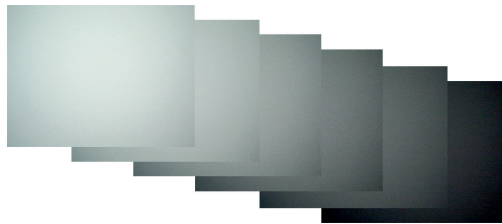


Figure 2.2: Determined brightness distributions of a neutral white body in bright, medium and dark scenarios.

determined for bright, medium and dark scenarios (Fig. 2.2). This is done by means of a neutral white body.

To correct a single image (Fig. 2.3), the brightness scenario is first determined by determining the average brightness. This determines the corresponding correction data set to be used from the neutral white body (Fig. 2.4). By addition/subtraction, the uneven brightness distribution of the image can be improved in the brightness channel L of the LAB color space (Fig. 2.5). In addition, the overall brightness of the image can be brought to a uniform brightness level for all individual images of the multi camera system at the same time in this step.

3 Distortion and Perspective Correction

Another major problem with stitching is the distortion of the individual images (Fig. 3.1). Likewise, different perspectives can already occur within an image, especially at the edges compared to the center of the image. In industrial image processing, systematic distortion correction

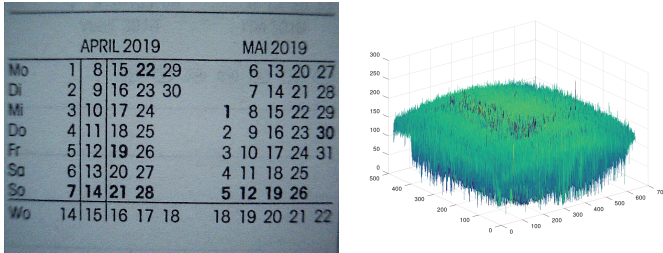


Figure 2.3: Image acquisition with uneven brightness distribution.

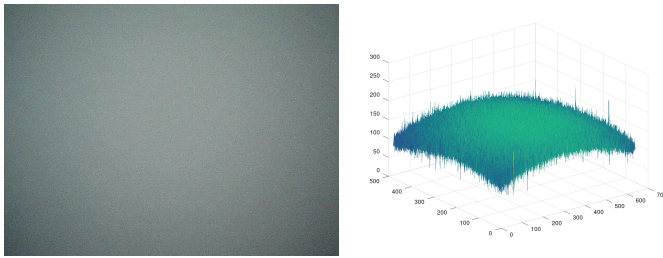


Figure 2.4: Selected correction data set from the neutral white body according to the determined brightness scenario.

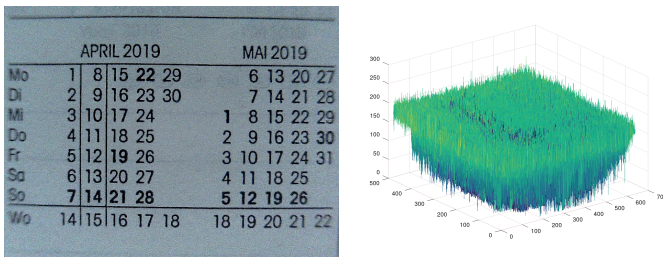


Figure 2.5: Brightness distribution corrected with the correction data set.

can be determined and compensated. This also performs a perspective correction at the same time.

Proposed solution [2]: An areal correction is calculated by means of polynomial interpolation. Deviations from the ideal position can

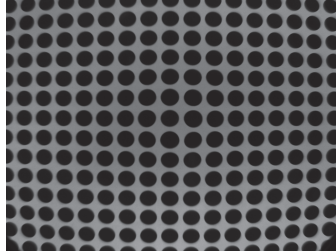


Figure 3.1: Distorted image that may also contain perspective errors.

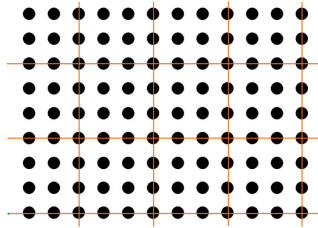


Figure 3.2: Decomposition of the total area into smaller areas to be corrected.

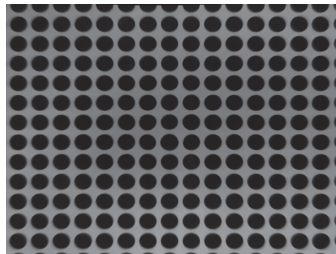


Figure 3.3: Distortion and perspective errors minimized by correction.

be determined and subsequently corrected by grid shaped support points. To enable a highly accurate correction, many grid points are necessary. In order to handle such a system better, it is advisable to divide the total area into individual smaller areas (Fig. 3.2).

Support points are arranged in the grid to form n columns and m

rows. Distorted actual values $(x_k, y_k) k \in (0, l)$, for $l + 1$ interpolation points are determined by image processing. Associated ideal coordinates $(x_{Soll,k}; y_{Soll,k})$ are known. Place polynomials of the $(n - 1)$ th or $(m - 1)$ th degree over interpolation points.

$$x_{Ideal,k}(x_k, y_k) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} a_{ij} * x_k^i * y_k^j \quad k \in (0; l) \quad (3.1)$$

$$y_{Ideal,k}(x_k, y_k) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} b_{ij} * x_k^i * y_k^j \quad k \in (0; l) \quad (3.2)$$

System of equations in matrix notation:

$$\begin{pmatrix} x_{Ideal,0} \\ x_{Ideal,1} \\ \vdots \\ x_{Ideal,l} \end{pmatrix} = \begin{pmatrix} 1 & y_0 \dots y_0^{m-1} & x_0 \dots x_0^{n-1} & x_0 y_0 & x_0^2 y_0^2 \dots x_0^{n-1} y_0^{m-1} \\ 1 & y_1 \dots y_1^{m-1} & x_1 \dots x_1^{n-1} & x_1 y_1 & x_1^2 y_1^2 \dots x_1^{n-1} y_1^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & y_l \dots y_l^{m-1} & x_l \dots x_l^{n-1} & x_l y_l & x_l^2 y_l^2 \dots x_l^{n-1} y_l^{m-1} \end{pmatrix} * \begin{pmatrix} a_{00} \\ a_{01} \\ \vdots \\ a_{n-1m-1} \end{pmatrix} \quad (3.3)$$

X_{Ideal}

$$\begin{pmatrix} y_{Ideal,0} \\ y_{Ideal,1} \\ \vdots \\ y_{Ideal,l} \end{pmatrix} = \begin{pmatrix} 1 & y_0 \dots y_0^{m-1} & x_0 \dots x_0^{n-1} & x_0 y_0 & x_0^2 y_0^2 \dots x_0^{n-1} y_0^{m-1} \\ 1 & y_1 \dots y_1^{m-1} & x_1 \dots x_1^{n-1} & x_1 y_1 & x_1^2 y_1^2 \dots x_1^{n-1} y_1^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & y_l \dots y_l^{m-1} & x_l \dots x_l^{n-1} & x_l y_l & x_l^2 y_l^2 \dots x_l^{n-1} y_l^{m-1} \end{pmatrix} * \begin{pmatrix} b_{00} \\ b_{01} \\ \vdots \\ b_{n-1m-1} \end{pmatrix} \quad (3.4)$$

Y_{Ideal}

4 Image Stitching

The stitching process requires multiple source images that have overlapping areas. Due to the previous corrections of the individual images (Fig. 4.1), the conditions are very good for obtaining a very good stitching result [3]. Fig 4.2 shows the high quality result of the stitching process with OpenCV.

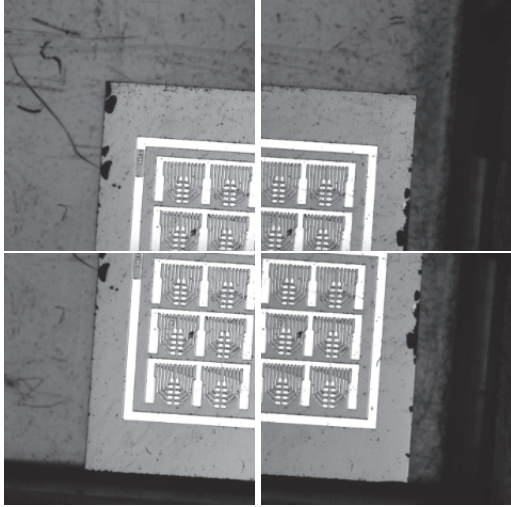


Figure 4.1: Single images optimized by previous corrections serve as the basis for the subsequent stitching process.

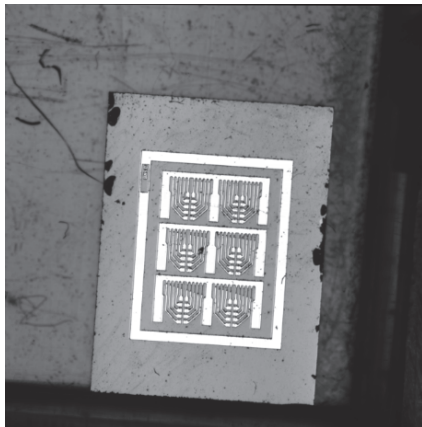


Figure 4.2: High quality result of the stitching process.

5 Results and Conclusions

Camera arrays can be used to inspect large areas. The individual images from all cameras must be combined to form one large overall image. This stitching process can be significantly improved if each individual image is enhanced before stitching. In this paper, homogeneity correction and distortion and perspective correction are demonstrated before the stitching process.

6 Acknowledgement

The research work presented in this paper is part of the research support program "European Regional Development Fund" (ERDF), Thuringia, Germany

References

1. Y. e. a. Li, "Automatic image stitching using sift," *International Conference on Audio, Language and Image Processing. IEEE*, 2008.
2. L. Haar, "Review and improvement of the correction process for distortions of lenses in the macro range," *Master thesis, Technische Universität Ilmenau*, 2014.
3. M. Chen, "Design and implementation of a method for the creation of a digital overview map in industrial image processing," *Master thesis, Technische Universität Ilmenau*, 2020.

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-2021) 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 Automated Visual Inspection of the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (Fraunhofer IOSB).

ISBN 978-3-7315-1081-9



9 783731 510819 >

ISSN 2510-7240
ISBN 978-3-7315-1081-9

Gedruckt auf FSC-zertifiziertem Papier