New vision technology for multidimensional quality monitoring of food processes

Dissing, Bjørn Skovlund; Ersbøll, Bjarne Kjær; Adler-Nissen, Jens

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Dissing, B. S., Ersbøll, B. K., & Adler-Nissen, J. (2011). New vision technology for multidimensional quality monitoring of food processes. Kgs. Lyngby, Denmark: Technical University of Denmark (DTU). (IMM-PHD-2011; No. 256).

New vision technology for multidimensional quality monitoring of food processes

Bjørn Skovlund Dissing

Kongens Lyngby 2011
IMM-PHD-2011-256

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

IMM-PHD: ISSN -

Contents

Summary

Spectroscopy and spectral imaging in combination with multivariate data analysis and machine learning techniques have proven to be an outstanding tool for rapid analysis of different products. This may be utilized in various industries, but especially rapid assessment of food products in food research and industry is of importance in this thesis. The non-invasive spectroscopic imaging techniques are able to measure individual food components simultaneously in situ in the food matrix while pattern recognition techniques effectively are able to extract the quantitative information from the vast data amounts collected. Underlying qualitative features (latent structures) are extracted from multivariate spectral data in order to quantify desired quality parameters properly. Specifically multispectral imaging which has been explored to a lesser extent than ordinary spectroscopy, having the possibility to exploit the inherent heterogeneity that exists in foodstuffs have been investigated here.

An extra feature obtained by combining spectroscopy, imaging and chemometrics is exploratory analysis. This is central in food research, since novel hypotheses about the food systems under observation may be generated using this inductive analytical approach. For the food industry it is an additional advantage that the fast, non-invasive, remote sensing nature of the spectroscopic imaging methods allows on-line measurements. In this way spectroscopic imaging in combination with advanced data analysis meets the high throughput needs for quality control, process control and monitoring. In this Ph.D. project the possibilities provided by spectroscopic imaging and chemometrics have been utilized to improve the analysis and understanding of different food products. The work is presented in seven papers and two additional technical reports which make up the core of the thesis. Furthermore an introduction together with a linking of the contributions is presented in this thesis.

The papers puts an emphasis on the use of multispectral imaging in the baking industry where especially the non-enzymatic browning appearance and features related to this are highlighted. These are features such as colour, water content and internal structure of bread. A paper presenting enzymatic browning in pre stir fried and thawn vegetables is also presented showing that imaging techniques such as the one investigated in this thesis is able to detect even subtle colour changes. The possibility for quantifying early as well as late spoilage in raw pork meat is investigated where use of the heterogenetic structure is utilized to obtain good results on predicting sensory evaluations as well on laboratory analysis.

Colour in other settings such as in the fishery industry is equally important, and a paper describing detection of cartenoid pigment in trouts using spectral images shows promising results.

Finally, two technical papers present possible ways of mapping multispectral images to a visible colour space, as well as how an alternative multispectral imaging system, making use of filters, may be used to design new more broad ranged filters. Fewer filters will increase the speed of such systems. Methods for solving such problems is to the knowledge of the authors rarely covered in the literature.

Resumé

Spektroskopi og spektral billede analyse i kombination med multivariat data analyse teknikker har vist sig at være enestående redskaber til hurtig kvalitets analyse af forskellige typer produkter. Dette bliver allerede udnyttet til en vis grad i industrien men især vurdering af fødevareprodukter i fødevare industrien er af interesse i denne afhandling. Den ikke-invasive spektroskopiske billede optagelses teknik er i stand til at måle individuelle fødevare partikler og komponenter simultant ved et samlebånd, ved at udnytte effektive mønster genendelses metoder til at ekstrahere den egentlige og nyttige information fra de store mængder af data som de spektrale billedeoptagelser udgør. Dette sker ved at finde underliggende latente strukturer i data som beskriver den information, vedrørende bestemte kvalitets parametre, man er ude efter.

Multispektral billedeanalyse af fødevarer er allerede blevet undersøgt meget i litteraturen, dog ikke i et lige så stort omfang som andre hurtige non-invasive metoder såsom f.eks. nær-infrarød og midt-infrarød spektroskopi. En stor fordel ved spektrale billeder i forhold til spektroskopiske målinger er den iboende spatielle information der kan bruges til at identificere præcis de områder på prøven der er vigtige for en given kvalitets parameter.

En anden central egenskab der opnåes ved spektraskopi og spekrale billeder sammen med avancerede dataanalyse teknikker er muligheden for via eksplorativ fremgang at opstille nye hypoteser vedrørende fødevarer. Den ikke-invasive stikprøve tagning er ligeledes i sig selv meget værdifuld for fødevare industrien da det kan nedsætte risikoen for konterminering af fødevaren og tillader meget hurtig vurdering af kvaliteten. På denne måde kan spektroskopiske billeder i kombination med mønstergenkendelse møde de høje produktions krav der sættes i industrien for kvalitets kontrol, process kontrol og overvågning.

I denne afhandling er spektrale billeder og mønster genkendelse blevet brugt til at undersøge og på nogle punkter give forslag til at forbedre analyse og forståelse

af forskellige fødevareprodukter. Det centrale arbejde præsenteres i 7 forskellige artikler (hvoraf 5 beskriver fødevarer og 2 beskriver andre tanker vedrørende spektrale billeder i fødevare industrien) og 2 tekniske rapport som er vedlagt. Ligeledes er en introduktion til emnet samt en sammenbinding af de udarbejdede artiker og konklusion præsenteret i afhandlingen.

Artiklerne sætter primært fokus på brugen af spektrale billeder i bageri industrien og lignende industrier der gør brug af varmebehandlende metoder såsom risting eller ovnbagning. Der sættes fokus på hvoledes spektrale billeder kan bruges til at øge forståelsen af de ikke-enzymatiske brunings processer der foregår under varmebehandlingen og hvorledes disse kan bruges til at forbedre overvågningen af sådanne processor. Overflade farve samt vandindhold og interne strukturer i fødevaren er meget vigtige egenskaber der forklarer brunings processen og derfor er disse undersøgt i forskellige eksperimenter og artikler. En artikel beskrivende en enzymatisk bruning af forarbejdede og frosne grøntsager der har ligget i køleskabet i en periode viser ligeledes hvordan multispektrale billedemetoder kan bruges til at detektere sådanne brunings processer til trods for det er meget svag bruning der foregår.

Andre kvalitets parametre der er blevet undersøgt er muligheden for at bruge spektrale billeder til at kvantificere fordærvelsen i råt hakket svinekød, i tidlig såvel som sen fase. I denne artikel demonstreres bl.a. hvordan den heterogene vurderings egenskab der naturligt ligger i billeder kan bruges til at opnå gode prædiktionsestimater af hvor fordærvet et givent stykke kød er. Til dette eksperiment gøres brug af sensoriske målinger såvel som laboratorie målinger til at relatere de spektrale billeder af kødet til et eventuelt fordærv.

Overflade farve er foruden at være vigtig i bageri industrien også meget vigtig i andre fødevare industrier såsom fiskeindustrien. På samme måde som måinger af overflade farve kan bruges til kvalitetskontrol af bagværk kan det også bruges til kvalitetskontrol af fisk. En artikel beskriver hvorledes pigmentet astaxanthin kan kvantificeres i fiske filleter vha. multispektrale billeder.

Til sidst præsenteres to tekniske artikler der dels beskriver en alternativ metode til at transformere multispektrale billeder til synlige farve rum der kan visualiseres. Dels beskriver de også hvordan et billedeoptagelses system der gør brug af filtre istedet for dioder som det vanlige kamera ellers gør, kan simplificeres fra at bruge mange filtre til at bruge ganske få uden at forringe prædiktions egenskaben af fedt i kobe kød væsentligt.

Preface

This thesis was prepared at Informatics Mathematical Modelling, the Technical University of Denmark in partial fulfillment of the requirements for acquiring the Ph.D. degree in engineering.

The thesis deals with statistical learning and vision technology in relation to various food products. The main focus is on highlighting various ways on estimating a given quality parameter of a food product, based on a set of multispectral images and chemical or sensoric measurements.

The thesis consists of a summary report and a collection of research papers written during the period 2008–2011, and elsewhere published.

Lyngby, May 2011

Bjørn Skovlund Dissing

Papers and technical reports included in the thesis

[A] Bjørn Skovlund Dissing, Mette Stenby, Bjarne Kjær Ersbøll and Jens Adler-Nissen, *Visual Quality Estimation of Oat Cereal*, Journal of Food Control, 2011. Submitted.

[B] Bjørn Skovlund Dissing and Mette Stenby, *Quantifying airbubbles in breadcrumbs*, Technical Report, 2011.

[C] Bjørn Skovlund Dissing and Mette Stenby, *Butter cookie quality estimation*, Technical Report, 2011.

[D] Bjørn Skovlund Dissing, Bjarne Kjær Ersbøll and Jens Adler-Nissen, *Monitoring water content using multispectral imaging and NIR in minced meat*, Proceedings for Near Infrared Spectroscopy International Conference , 2009. Published

[E] Bjørn Skovlund Dissing, Olga Papadopoulou, Chrysoula Tassou, Bjarne Kjær Ersbøll, Jens Michael Carstensen, Stathis Panagou and George-John Nychas, *Using Multispectral imaging for spoilage detection of pork meat* Food Research International, 2011. Submitted

[F] Bjørn Skovlund Dissing, Michael E. Nielsen, Bjarne Kjær Ersbøll and Stina Frosch, *Multispectral imaging for determination of astaxanthin concentration in salmonids* Public Library of Science (PLoS ONE), 2011. Accepted for publication

[G] Bjørn Skovlund Dissing, Line Clemmensen, Hanne Løje, Bjarne Kjær Ersbøll and Jens Adler-Nissen, *Temporal Reflectance change of vegetables*

Proceedings for IEEE Color and Reflectance in Imaging and Computer Vision Workshop, CRICV 2009, October 2009 2009. Published

[H] Bjørn Skovlund Dissing, Jens Michael Carstensen and Rasmus Larsen, *Multispectral Colormapping using Penalized Least Square Regression*, Journal of Imaging Science and Technology 54(3), 2011. Published

[I] Ken-ichi Kobayashi, Bjørn Skovlund Dissing, Masaaki Mori, Ken Nishino, Toshihiro Toyota, and Shigeki Nakauchi, *Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts*, Proceedings of Scandinavian Workshop on Imaging Food Quality, 2011. Accepted for publication

Acknowledgements

I am greatly indebted to the many skilled and visionary people with whom I have had the great privilege to collaborate.

The very competent people in the image and data analysis groups at the department of informatics and mathematical modelling are gratefully thanked for offering numerous interesting challenges, good collaboration, atmosphere and friendship.

I would also like to thank my co-supervisor Jens Adler-Nissen and his group at DTU-FOOD for giving me an insight in the interesting world of Food engineering and for fruitful collaboration .

A thanks also goes to Prof. George Nychas and his group at the agricultural university of Athens for having me during three months.

Also I would like to thank Prof. Jussi Parkkinen and the colour group at university of eastern finland, Joensuu for having me during three months.

Likewise I would also like to thank Michael Engelbrecht and Stina Frosch from DTU Aqua for good discussions, collaborations and guidance.

Finally, I would like to thank my supervisor Bjarne Kjær Ersbœll for his great support and for always having his door open for any kind of discussion.

My friends and family should also have a big thanks for their encouragement and support throughout the period.

Contents

Part I

Summary report

Chapter 1

Introduction

1.1 Motivation and aims

Industrial food production is today a very big industry. Demands on ready made convenience food as well as raw foods increases rapidly and there is a great challenge in keeping up with the demand. Moreover, besides increasing demands on sheer amounts, the consumer is becoming increasingly focused and specific when it comes to quality of foodstuffs. To keep up with these demands it is necessary to aid the food industry with objective quantitative quality inspection methods.

Different foodstuffs consist of discrete particles which during manufacture are heat treated in continuous, open processes. Examples of such foodstuffs are the continuous contact frying of finely chopped meat pieces or vegetables or oven drying of cereals such as corn flakes, nuts or the like. In all cases the heat treatment leads to partly or complete desiccation of the products with a significant colour change as a consequence. During the processing the foodstuff may sustain physical damage by e.g. agglutination of meat due to coagulation or mechanical damages of fragile products such as cereals. Over burning in heat treated products may give rise to occurrence of black particles in the products. A commonality for the above mentioned processes is the visual integrity of the foodstuffs. A correct uniform colour distribution on the surface, proper size distributions and absence of agglutinated or damaged products is crucial for the quality of the food product. The mentioned quality parameters are all visual and are often assessed during the process visually by well trained expert human operators. If the operator notices visible exceptions or anomalies, that person will normally intervene in the process to correct it and thereby recreate the desired visual quality in the concerned foodstuff.

This creates a strong dependence on human operators which has some disadvantages such as late process intervention. Furthermore, manual process control is labour intensive which induces fatigue and affects the performance of the operator. Operators contain a lot of expert knowledge obtained by experience. Such knowledge is very difficult to pass on to others, who will have to make their own experiences.. This creates bottleneck dependencies which also may be a problem. The disadvantages usually lead to quality reduction of the foodstuff which may lead to rejection and in the end loss of profit and thereby a less optimal production line.

The visual judgment of the operator may be supported by operational systems based on experience [1] or different types of online measurement equipment, which in combination with suitable computational algorithms may facilitate an automation. The development of such systems has been slower in the food industry than in other industries, however systems using advanced statistical control algorithms such as fuzzy logic, neural networks or similar types of techniques [2, 3] does exist. Literature study indicates that the introduction of automated process control in many food processes is hindered mainly by the measurement and following data processing and less by the actual control [3].

The introduction of near infra red spectroscopy (NIR) for process surveillance meant a breakthrough in controlling e.g. water and fat content of foodstuffs [4, 5] and speeded up implementation of intelligent systems in the food industry. NIR has a large potential for advanced classification tasks such as sorting fresh, frozen and thawed fish [6].

In spite of NIRs great success in the food industry, it has a drawback of lacking spatial information. This is where camera technology becomes interesting since this has the possibilities of combining spectral and spatial information. A few application examples using camera technology being used for quality inspection of foodstuffs are control of feces on poultry carcasses and detection of defects on apples [7, 8, 9].

The purpose of this project is to investigate the possibility of imitating and possibly improving the manual visual quality assessment of food products in the baking and meat industry, where the process indicates that there might be a demand for online inspection with high reliability of the product. Such processes might benefit from automated process control based on visual inspection.

1.2 Multi/Hyperspectral images

The electromagnetic spectrum (EMS) (Figure 1.1) consists of all possible frequencies of electromagnetic radiation. When electromagnetic radiation interacts with an object some frequencies are absorbed and some are reflected. The reflection depends on the chemistry of the surface of the object it interacts with. The human eye is capable of sensing EM radiation and translate the recorded radiation into colours and images in the brain. However, even though the EMS consists of indefinitely many frequencies, the human eye is only able to absorb radiation in the visible area located from 400 to 700 nm (Figure 1.1). There are four types of photosensitive cells in the human eye. Three of them, called the cones, have absorption properties in different parts of the EMS. The three types of cones are called S, M and L or β, γ and ρ cones, and their spectral sensitivities are seen in Figure 1.2. Basically the cones' absorption properties work as filters, filtering everything but a narrowbanded area of the EM radiation. By combining the information obtained in the three cones, the human brain is able to create a colour sensation. The β-cones are sensible in the blue are, the γ-cones are sensitive in the green area, and the ρ-cones are sensitive in the red area.

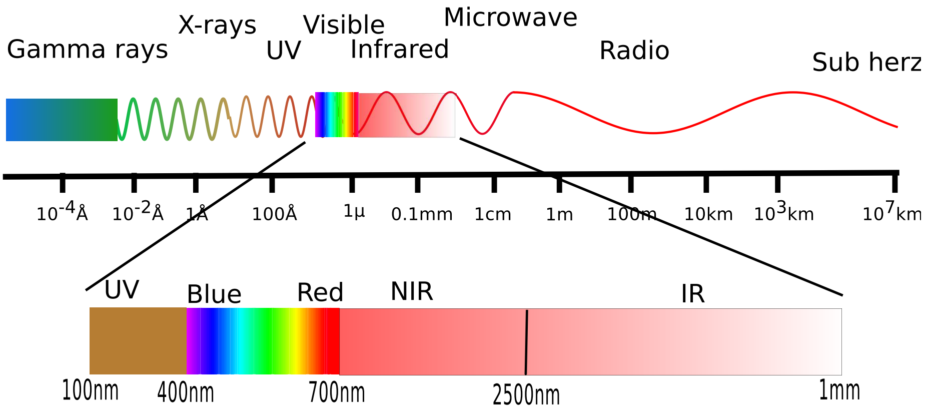


Figure 1.1: The electromagnetic spectrum illustrated from 1 Ångstrom to 10^7km. The UV, visible, NIR and IR area are enhanced. These areas are used for various spectroscopic applications.

By mimicking this filter system and having photosensitive material to absorb the EM radiation a camera is able to record images. The first camera appeared in 1817 and the technology has evolved rapidly ever since. In 1951, the first video tape recorder captured live images from television cameras by converting the information into electrical impulses and saving the information onto mag-

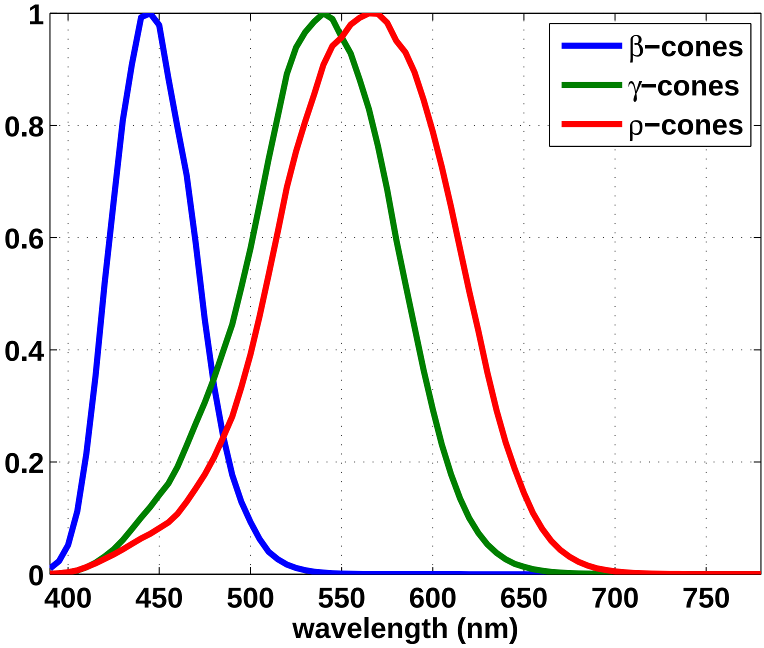


Figure 1.2: Sensitivity for the three types on cones in the human eye.

netic tape - the first commercial digital camera was made available in 1981. The heart of the digital camera is the Charged Coupled Device (CCD) which was invented in 1969 at AT&T Bell Labs and has evolved tremendously since. The CCD is capable of absorbing photons in the visible area of the EMS depicted in Figure 1.1 and convert it to electrical impulses to be interpreted as numerical values. The common CCD basically integrates all the incoming light in the visible area giving rise to monochromatic images. To create color images a set of filters with characteristics similar to those seen in Figure 1.2 a placed in front of the CCD chip. By doing this, three image channels are recorded, one for each filter, which when combined gives a single colour image. Thus, by splitting up the incoming light in three areas using filters and combine the output, colours may be recreated.

Since integration is an irreversible operation, it is impossible to recreate the light spectra when first converted to colour coordinates. Information about the shape of the true spectrum is however exactly what is needed to avoid problems such as metameric failure. Metameric failure can, in short, be described as the case where two objects match colorimetrically under one illumination, but differ under another [10]. For the amateur photographer this might not be a big issue, but when it comes to cameras in industrial settings where reproducibility is one of the most important factors, this might be a problem. By increasing the number of filters and thereby sampling the EM radiation more densely, more information is obtained about the object emitting the light. With increased sampling resolution follows an increased ability to distinguish different types of materials and surface chemistry. This makes dense sampling very usable for a large array of applications. However, its strength is also its weakness, since large amounts of information about the EM radiation also creates large amounts of data which needs to be stored and processed.

Luckily innovation happens concurrently in different fields, so while the detector technology and imaging technology improves and creates ever larger datasets, so does the computing technology. By now, both processing and storing capabilities in computers has made it possible for a larger audience to make use of spectral imaging. Industries and groups which use and develops spectral imaging today include among others; industries using satellite imaging, medical industry, food industry and astronomers.

Spectral cameras today either use 1D or 2D sensor arrays and are usually focused on the visible or near infra red (NIR) area lying between 400 and 2500 nm. However some systems focus on ultra violet (UV) or infrared IR) instead (Figure 1.1). Depending on where in the EM area it is desired to sample frequencies, different types of detectors are available. For 1D array cameras a line scan method is used where a common approach is the push broom method. In a push broom sensor, a line of sensors arranged perpendicular to the scanning direction is used. Either the sensor or the object moves, which as a consequence results in a scanning and thereby recording of a surface.

An alternative approach which only uses a single detector element is the whisk broom approach. Here, a mirror scans across the scanning track, reflecting light into the detector element which collects data one pixel at a time. Whisk broom scanners have the ability to stop the scan and focus the detector on one part of the object, typically capturing greater detail in that area. However the moving parts make this type of sensor expensive and more prone to wearing out. 2D arrays are capable of acquiring a simultaneous snapshot of the incoming photons and is thus a very fast approach.

As described earlier incoming light may be split into a set of measurements using filters. A large amount of different types of filters exists such as acousto-optical tunable filter [11] or liquid crystal tunable filter [12]. More comprehensible filter wheel designs also exists[13, 14]. Obviously for all sensors a light source is necessary to record any type of images. An alternative to filters is by focusing light through an optical dispersing element such as a prism or a grating, splitting the light into several narrow adjacent wavelengths bands. The intensity of each band is then measured by a separate detector.

Independent of the filtering technique used, a light source is necessary. Depend-

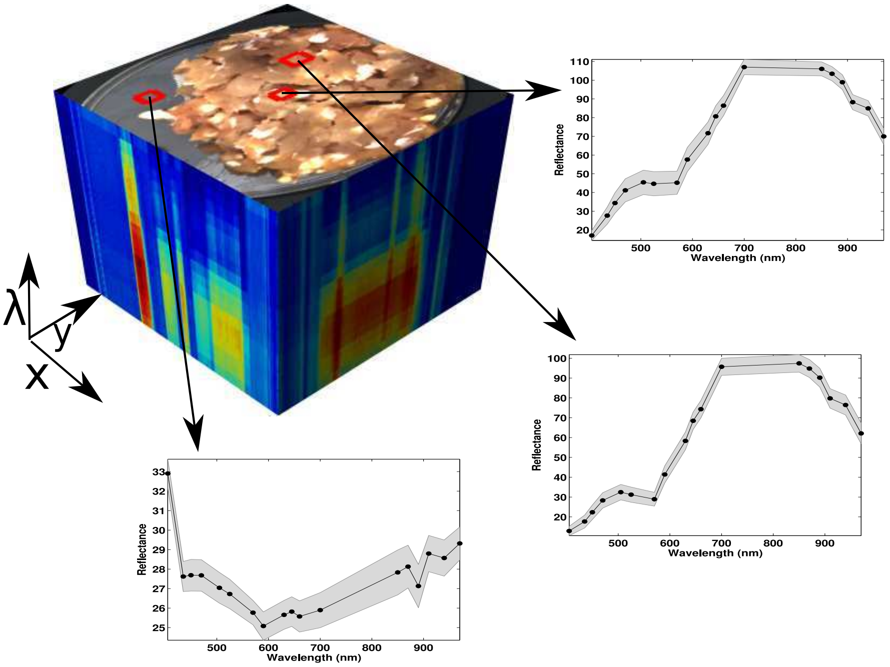


Figure 1.3: A multispectral image of minced pork meat. The image consists of 18 wavelengths scattered over the visual end a small part of the near infra red spectrum (405 - 970 nm). On the sides of the cube spectral values are plotted for the edge pixels. Furthermore the image has been sampled in three different locations where the spectral values are seen to differ. The gray shaded areas represents one standard deviation. The figure illustrates how each pixel in the x,y domain is in fact a spectrum as indicated by λ

ing on the type of system, broad banded light sources such as broad daylight are commonly used. Other light sources are tungsten filament which emit light between 300 and 2500 nm, deuterium arc lamps (190-400 nm) and Xenon arc lamps (160-2000 nm). [15]
An alternative approach is to use light emitting diodes (LED) as light sources which have become increasingly cheaper and better. LEDs are able to emit narrowbanded light, which eliminates the need of using filters or gratings. Such systems are presented in [16, 17]. VideometerLab [17] makes use of this approach together with an integrating sphere coated with a matte paint to obtain diffuse light. This ensures reproducible images with good dynamic range, low scatter- and shadow effect. Typically VideometerLab has around 20 different LEDs mounted which emits light in the visible and the first part of the near infra red (NIR) area. The spatial resolution of VideometerLab is 1280x960 and

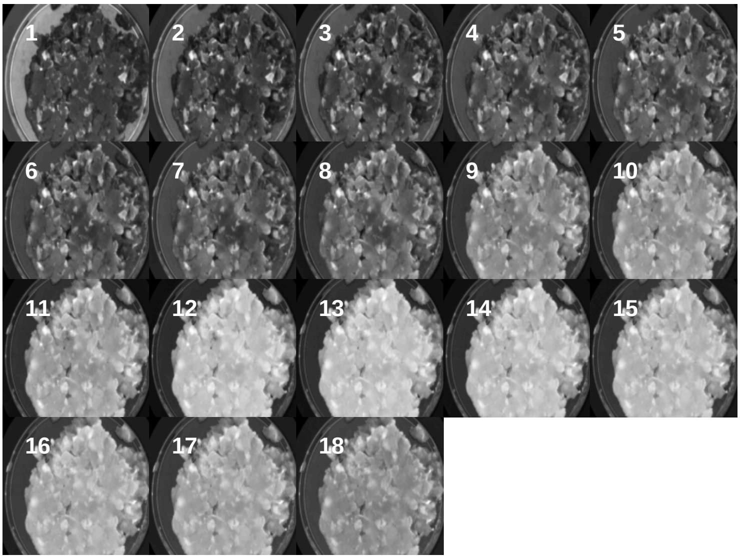


Figure 1.4: Visualization of a multispectral image of a piece of raw pork meat in a petri dish. The image consists of 18 channels from 405 - 970 nm. The wavelength of the channels is printed in the corner of the corresponding tile.

the acquisition time for 20 channels is approximately 8 seconds. The system is geometrically and radiometrically calibrated [18].

Commonly, any contiguous sampling of the EMS in several narrow bands is called multispectral or hyperspectral. The number of bands produced by available sensors range from very few to several hundreds. There is however no universally agreed upon minimum number of bands or bandwidth dimension required for a dataset to be considered either multi or hyperspectral, although hyperspectral images does contain more spectral bands than multi spectral images. A suggestion is that spectral images with more than 20 wavebands is called hyperspectral and multispectral when below 20 [19].
Multispectral images are difficult to visualize in low dimensions, since each pixel in the image represents a large amount of samples. One common approach for visualization is the three dimensional multispectral data cube seen in Figure 1.3. Thinking of the hyperspectral data in this way may have some deficiencies as it might indicate that we assume that the data has similar characteristics in all directions of the data cube, i.e., that correlations are similar in the spatial and spectral directions. This is however not the case and usually the correlation is very high across spectral bands. Figure 1.3 shows a multispectral image acquired by VideometerLab in 18 different channels by minced pork meat. The figure illustrates how each pixel is in fact a spectrum. Samples from different

spatial locations in the image may have very different spectral appearance as illustrated. Even within a small region of interest containing similar tissue, the data is seen to vary as indicated by the gray shaded areas. An alternative visualization of a multispectral image of minced pork meat is seen in Figure 1.4 Page 9. Each Tile in the image represents a single wavelength recording. Naturally such types of visualization for all channels is only possible for multispectral images with relatively low spectral resolution such as VideometerLab images.

1.3 Food technology and Process Analytical Technology

Some of the earliest online univariate systems are found in the petrochemical industry where sensors such as pH meters, oxygen sensors and flow meters has been used. These systems are in fact the cradle of modern Process Analytical Technology (PAT).To completely describe PAT is beyond the scope of this thesis, but this chapter will provide the reader with conceptual knowledge about the ideas behind PAT as well as introducing a possible practical PAT application from the baking industry.

In 2004, the United States Food and Drug Administration (USFDA) motivated a full integration of Process Analytical Chemistry (PAC) [20] in the pharmaceutical industry and termed it Process Analytical technology (PAT)[21].

PAC deals with process monitoring and control of key manifest parameters which are employed to predict and ensure a certain end product quality during the process - mainly chemical or physical analysis of materials in the process. PAC was termed in the beginning of the 20th century and is therefore commonly known in the industry by now. In spite however of the long existence of PAC it has first found its way into common usage in production within the last couple of decades. This is mainly due to the development of very strong computers, fast and very responsive sensors as well as complex multivariate datamining.

PAT encompasses all aspects of a process and is in short: 'Analysis of the process' [22, 20]). Thus, PAT considers a process as a consecutive row of unit operations from raw materials to final products and wastes. The concept of PAT may be boiled down to the definition given by USFDA as of 2004: *"A system for designing, analyzing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality."* In a project such as this thesis, PAT is explored as an investigation of in-process materials. This means certain unit operations in a larger processing chain is investigated and optimized.

Investigating unit operations often requires sampling from a process line. Such sampling when performed in real time requires advanced technology such as vibrational or scattering spectroscopy, chromatography, mass spectrometry, acoustics, chemical imaging or light-induced fluorescence light scattering. Conceptually four different ways of sampling the process line may be set up, as seen in Figure 1.6 Page 13. Off-line sampling is the process of obtaining samples from the process line of interest and transporting them to a central analytical laboratory. Often the results of these analyzes yield spurious information or does not give all of the information about the process that they could have, as the sample properties often changes during transportation and handling - in fact it might even when removed from the actual process line. At-line sampling still requires

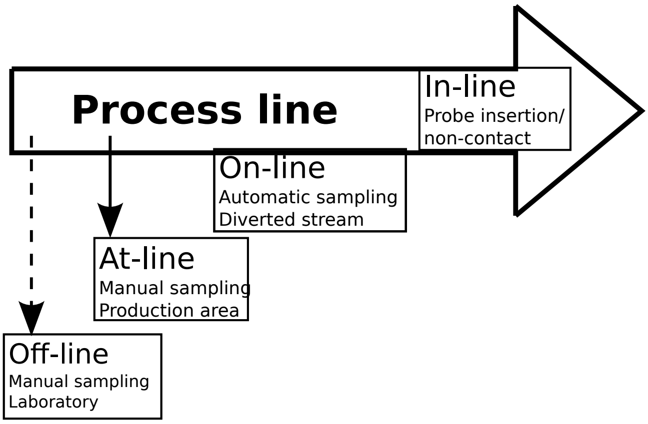


Figure 1.5: Conceptual diagram showing various ways to sample the process line

manual sampling, however instead of transporting samples to a central laboratory, samples may be analyzed in production area. Improving further, methods for automatic sampling exists where a part of the process line may be diverted for sampling and testing. Finally, the superior sampling strategy is performed in-line, which makes it possible to sample major parts of the products on the conveyor belt real time.

Even though offline and online/inline equipment in essence is the same equipment, there are some differences. Equipment in a laboratory is in a temperature-controlled safe environment operated by trained operators. Online/inline equipment may be outdoors or in areas with a lot of uncontrollable factors such as temperature, humidity, pressure etc. affecting the setup. The biggest difference is that in the process-line, the instrument is often running without an operator. There is no personal to check if the proper reagents are present, if the sample is clean, and in the physical state the instrument is expecting etc. In order for a monitoring system to run unsupervised a sufficient amount of system intelligence is necessary. An unsupervised system in production should be able to recalibrate itself, find wrong samples and identify process drifts among other things.

Building such intelligence is often handled using pattern recognition and machine learning methods such as those described in Section 2 Page 17.

1.3.1 PAT in the baking industry

In modern baking industry there is an increased interest in automation of processes, while maximizing the yield at constant quality. [23] described an at-line system for monitoring the leavening process in industrial bread making using NIR spectroscopy. This proof of concept system shows how PAT may be used in the baking industry in a batch based production.

Similarly, a proof of concept system will be described here for the heating processing of butter cookies. Understanding and controlling the non-enzymatic browning processes during heat treatment of baking products is critical to the yield quality of the final baking product. Non-enzymatic browning (described further in Section 3 Page 31) governs the surface colour of heat treated products which is of great importance for consumer acceptance. A way to monitor the browning process is by using digital imaging, either tri-colorimetric or multispectral. A multispectral monitoring system may be set up as shown in Figure 1.6 Page 13.

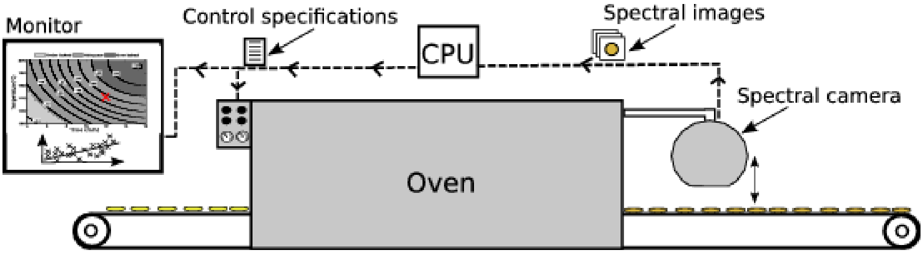


Figure 1.6: Schematic of possible multispectral imaging system setup in a cookie production

A conveyor belt is transporting butter cookies into an oven where they are being heat treated. Upon exit, a multispectral camera such as VideometerLab records images of the cookies and sends the data to a central CPU. The CPU extracts a set of features from the images and stores them together with previous recordings. In this way a time line of features is created which is used to create a set of control parameters for the oven. Furthermore the process state may be displayed on a monitor as shown.

Studies have shown that the Maillard reactions together with caramelization, types of non-enzymatic browning which mainly affect baking products, are very dependent on temperature and water activity [24]. By being able to predict the surface colour, a better control of the product quality is possible. Experimental tests showed that browning kinetics due to Maillard is characterized by a lag phase (detected at $150°C$), followed by an exponential phase and an asymptotic phase (detected at $190°C$) as described in [24]. The same study described a kinetic model for surface browning by modeling the change in surface colour

($\Delta E = \sqrt{L^{*2} + a^{*2} + b^{*2}}$) of bread crust as a function of time at $200°C$ and $250°C$. They concluded that their model was usable at different temperatures up to a ΔE around 30. A more recent study on kinetic browning in biscuits [25] builds a model based on surface colour as well, however based on the RGB colour space. Both kinetic systems estimate parameters from tri chromatic images, which are known to describe the browning process well.
A typical browning progression is shown in Figure 1.7 Page 14 based on tri-chromatic measurements samples from Paper C Page 99. L^* values from 3 surface samples on each of 30 butter cookies baked at $180°C$ are shown as a function of time. The maximum colour change in the experiment is $\Delta E = 41.2$. A clear decrease in lightness is seen as time in the oven increases, which could be quantified directly. The processline in Figure 1.6 Page 13 shows a multi-spectral camera which has some clear advantages over colorimetric imaging as also outlined in Paper C Page 99. Therefore the focus on this application is on multispectral imaging.

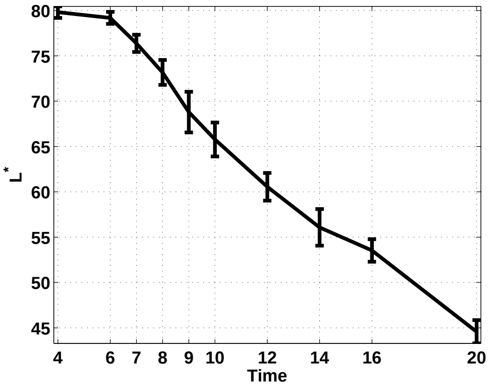
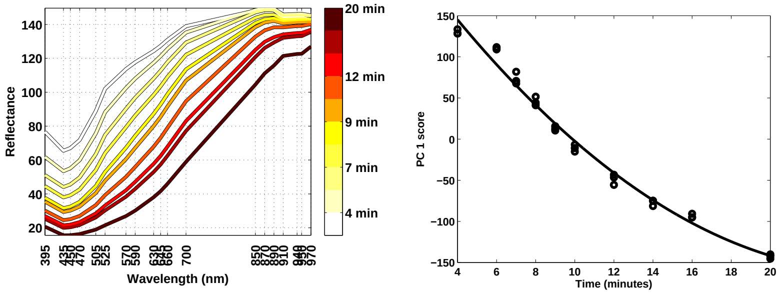


Figure 1.7: Tri-chromatic measurements on butter cookies heat treated at $180°C$. Each cookie was measured at three different locations. The L^* component which is best correlated with the browning is shown here.

Multispectral recordings of the same cookies are shown in Figure C.1 Page 105, coloured according to time spent in the oven. I similar clear trend is here seen as a difference in scaling of the spectra. If these spectra are subjected to a principal component analysis, the browning may be followed accurately in the first principal component as seen in Figure 1.8 Page 15 and correlation between L^* values and PC scores is calculated to be 0.9903. To predict these browning scores a second order model is fitted with an $R^2 = 0.9856$. The papers presented in this thesis elaborate approaches as the one described above which examplifies them as PAT applications.



(a) (b)

Figure 1.8: (a) shows mean spectra for cookies used in Paper C Page 99. (b) shows PC_1 scores from a PCA fitted on the spectra shown in (a). A second order model fits the scores with an $R^2 = 0.986$.

CHAPTER 2

Computational methods

This chapter gives an introduction to the mathematical background for this thesis. The foundation of the type of data present in this thesis and different methods to handle such data are described.

2.1 Symbols

row	Rows in a multispectral image
col	Columns in a multispectral image
q	Number of recorded wavebands in a multispectral image
M	A set of multispectral images
N	Number of samples
p	Number of dimensions in feature space
$\mathbf{X}$	Feature space,covariate matrix,data matrix
$\mathbf{X}_c$	Centered feature space
$\mathbf{y}$	Reference values
β	Coefficients in model

2.2 General experimental data source

In a typical larger study/application of multispectral images of foodstuff where the goal is to be able to model a specific quality parameter of this foodstuff a careful experiment needs to be planned. Briefly, the goal is to vary the parameter of interest so that it covers as large a part of its variation space as possible which in terms have an effect on the foodstuff of interest. At sensible points in this variation space, multispectral images of the foodstuff should be acquired. Designing such experiments properly is a large area [26] and quite central for a topic such as multispectral imaging and PAT (Section 1.3 Page 11) in general. Sampling a quality parameter and corresponding multispectral images of the foodstuff at N points of the reference quality parameter leaves a dataset with N reference values ($\mathbf{y}$) and a set of N multispectral images (M), as seen in Figure 2.1 Page 18. Since each multispectral image consists of ($row \times col$) pixels,

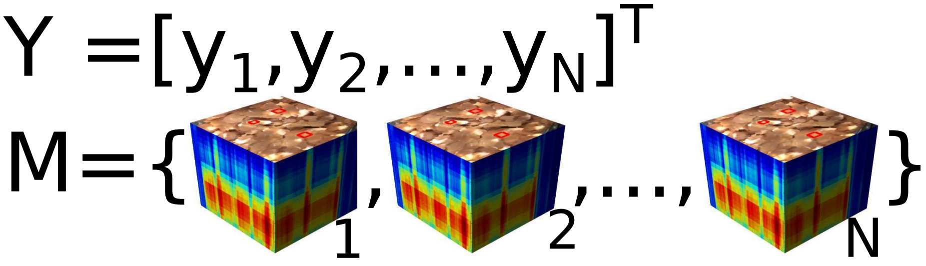


Figure 2.1: A typical calibration set in analysis of foodstuff using spectral imaging

where each pixel is a spectrum with q recorded wavelength bands (Section 1.2 Page 5), there is essentially $row \times col$ observations each of dimensionality q. Thus, essentially for a given multispectral image experiment there is ($row \times col \times N$) image observations but only N reference observations. In order to use conventional machine learning methods to relate the independent variable Y to the set of acquired multispectral images M, a feature space $\mathbf{X}$ of dimensionality p needs to be created from M by some transformation, or more formally stated in Equation 2.1 Page 18.

$$\mathbf{X} = T(M) \tag{2.1}$$

$T(\cdot)$ is a transformation that transforms M to $\mathbf{X}$. $T(\cdot)$ internally consists of a chain of operations on each member of M which are specific for a given problem. Some general approaches to how $T(\cdot)$ may be created is discussed in this section together with ways of relating X to Y.

Having a multispectral image dataset requires a large toolbox of methods to analyze properly. The foundation of modeling such vast data amounts is laid out by machine learning and statistics which seeks to describe methods for analyzing, testing, assessing and understanding all kinds of different types of data. A good introduction and coverage of the field is found in [27].

As described, the initial step of analyzing multispectral images is to determine $T(\cdot)$ (Equation 2.1 Page 18). This transformation may be undertaken in many ways of varying complexity, described in Section 2.3 Page 19. The output of $T(\cdot)$ is a featurespace of dimensionality p with N samples.

Depending on the type of reference values in $\mathbf{y}$, being e.g. grouping or ranking labels from sensory evaluation, measurements from a laboratory or some other parameter such as time, a relevant method needs to be chosen for analysis. The most relevant methods used in this paper to map the featurespace to reference values or similar are described in the following with pointers to alternative methods.

2.3 Region of interest

A part of the operation chain $T(\cdot)$ which is often necessary to perform is creating a region of interest (ROI). A ROI is basically a mask indicating the area or areas in the multispectral image which contain only spectra of interest regarding the reference value.

The amount of ways to create a ROI is tremendous. Some sort of segmentation algorithm is necessary to create a ROI and several different scientific communities exists specializing in different types of segmentation algorithms. Usually a set of operations are required before an image is segmented satisfactory. Naturally the steps involved depends on the type of foodstuff recorded, however a very abstract list of approaches to segment an image is given here.

1. Classification
2. Region-based / Morphological
3. Edge based
4. Partial Differential Equation

Classification based methods is basically anything from the machine learning community, which regards pixels as observations from a larger probability distribution and classifies them solely based on their spectral shapes and not on their spatial location in the image. Examples of such methods are support vector machines [27], neural networks [27], clustering methods [27] and thresholding (adapative) [28].

Edge based methods are those which are usually based on an edge map of the image. Edge maps may be created by calculating finite differences between pixels, or by filtering the images with edge-enhanced filters such laplacian filters [28]. Examples of methods using such edge maps are border tracing [28], edge relaxation [28], Dynamic Programming [29] or the Hough transform [28].

Region based methods aim to distinguish and label different pixels in concordance with their surrounding pixels (their regions). Examples of segmentation methods of this type includes besides the watershed [28] also markov random fields [30] and morphological methods such as tophat and h-domes [31].

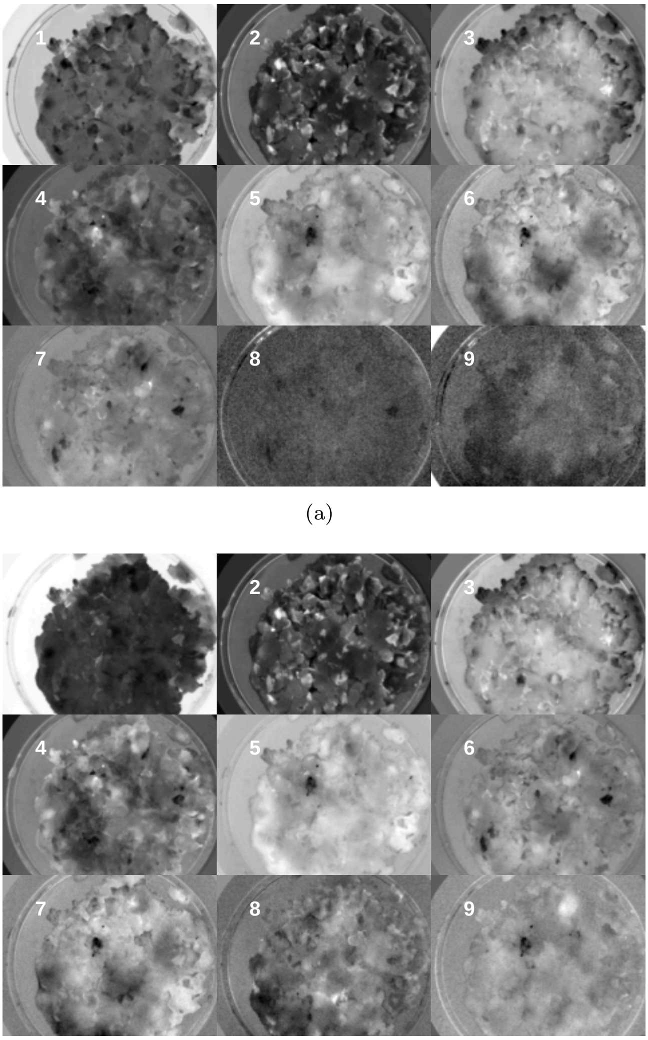
Partial differential equation (PDE) based methods or levelset based methods [32] are advanced methods which creates an initial contour which it evolves based on a set of defined guidance forces which e.g. pulls the contour in directions low image gradients and stops it from going towards directions with high image gradients. PDE methods are implicit and the explicit active contour models (snakes) [32] and the Grenander template model [33].

Basically the distinction of segmentation methods into groups is quite hard since a segmentation is usually a collection of operations where methods from different of the above defined groups are used. An example of this might be the watershedding method [28] which would be grouped under the region growing methods even though it uses edge information in the image.

The majority of the above mentioned techniques are rarely found in literature concerning multispectral images in the food industry. The reason for this is tradition as well as the fact that these methods may be difficult to implement for people without a background in applied mathematics or computer science.

Often encountered ways of creating ROIs in multispectral images is by use of orthogonal methods followed by a labeling step using an adaptive thresholding technique such as otsu's method [34]. Some orthogonal methods includes principal component analysis (PCA Section 2.4 Page 23), canonical discriminant analysis (CDA Section 2.5 Page 24) [27], maximum autocorrelation factors (MAF Section 2.6 Page 24) and minimum noise fraction (MNF) [35]. The theory of PCA, CDA and MAF will be presented briefly later in this section.

An example of creating a ROI for the meat lump seen in Figure 1.4 using MAF and PCA to decompose the multispectral image followed by an adaptive threshold (Otsu's method) follows. Decomposition components found using MAF and PCA are seen in Figure 2.2(b) Page 21. The components for the two transformations are ordered in a descending order according to directions of maximum variance (for PCA) and directions of maximum spatial autocorrelation structure (for MFA). For the present example the MAF is able to project the data in a fashion with creates a large contrast between meet/fat and background, being petri dish etc. The PCA in comparison found a similar component however with lower contrast. By using the first MAF component and subdue it to otsu's algorithm the mask seen in Figure 2.3 Page 22. The presented example needed a transformation of 2 steps in order to segment the background of the multi-



(a)

(b)

Figure 2.2: Examples of PCA (a) and MAF (b) decompositions on the multispectral image seen in Figure 1.4 Page 9. The tiles in each image represents latent components extracted from the multispectral image. Only the first nine components are shown, the remaining show no signal.

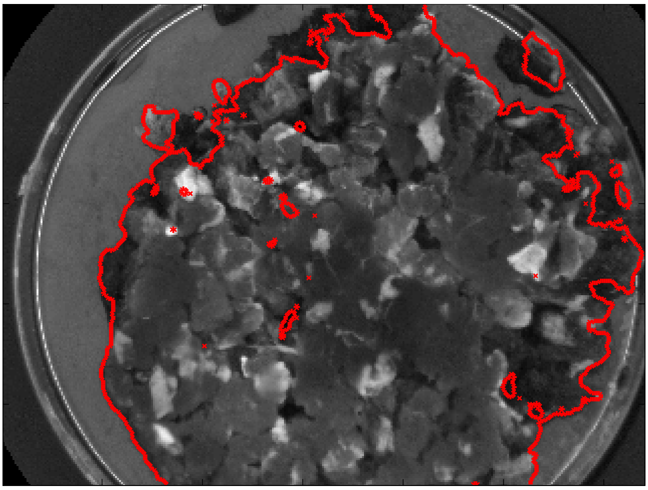


Figure 2.3: The result of using otsu's method on the first MAF component (Figure 2.2(b) Page 21). The graytones show the component intensities and the red contour indicates the found mask.

spectral image containing minced pork meat. More complex examples might need chains or pipelines of several steps to create masks of arbitrary complexity.

2.4 Principal Component Analysis

In high-dimensional data problems such as multispectral images it may be difficult to immediately understand and visualize how the data behaves. Essentially PCA models the original multivariate data using a limited number of latent variables (principal components). First the original data matrix $\mathbf{X}$ is centered by subtracting the mean spectrum $\mathbf{x}$ from each sample spectrum, where after the centered data matrix, $\mathbf{X}_c$, is decomposed to a score matrix $\mathbf{T}$ and a loading matrix $\mathbf{P}$ by consecutive orthogonal subtraction of the largest variation in data until the variation left is unsystematic. The residuals are collected in ϵ (Equation 2.2 Page 23).

$$\mathbf{X}_c = \mathbf{T}\mathbf{P}^T + \epsilon \tag{2.2}$$

To compute the principal components an eigendecomposition of the R-mode scatter matrix ($\mathbf{X}_c^T\mathbf{X}_c$) is performed and sorted according to magnitude of the eigenvalues. The eigenvectors are the orthogonal loadings $\mathbf{P}$ and the score matrix is calculated by projecting the centered data matrix using the loading matrix $\mathbf{T} = X_c\mathbf{P}$, thus creating an space where the variables are uncorrelated. Another way of finding the principal modes of variation is by using a Singular Value Decomposition (SVD) [27]. The SVD is very general in the sense that it can be applied to any $m \times n$ matrix as opposed to the eigenvalue decomposition stated above. The SVD factorizes the covariate-matrix $\mathbf{X}$, as shown in Equation 2.3 Page 23,

$$\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{X}_c \tag{2.3}$$

into $\mathbf{U}$ and $\mathbf{V}$, two orthonormal matrices of size $n \times n$ and $p \times p$ containing the principal components of $\mathbf{X}_c$, and $\mathbf{D}$ which is a diagonal matrix, with singular values on the diagonal $d_1 \geq d_2 \geq \ldots \geq d_p \geq 0$. The principal components contained in $\mathbf{V}$ are equal to those found in the eigenvalue-decomposition of $\widehat{\boldsymbol{\Sigma}}$ as seen in Equation 2.4 Page 23.

$$\begin{aligned} \frac{1}{n}\mathbf{X_c}^T\mathbf{X_c} &= \frac{1}{n}\mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \frac{1}{n}\mathbf{V}\left(\mathbf{D}^T\mathbf{D}\right)\mathbf{V}^T \\ \frac{1}{n}\mathbf{X_c}\mathbf{X_c}^T &= \frac{1}{n}\mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V}\mathbf{D}^T\mathbf{U}^T = \frac{1}{n}\mathbf{U}\left(\mathbf{D}\mathbf{D}^T\right)\mathbf{U}^T \end{aligned} \tag{2.4}$$

where the relationship between the empirical, centered covariance matrix and the factorized matrices are shown. Furthermore the singular values of $\frac{1}{n}\mathbf{D}$ corresponds to the square-roots of the first $\min(n, p)$ eigenvalues of $\widehat{\boldsymbol{\Sigma}}$. Usually when working with spectroscopic data of high spectral resolution it is advantageous to find scores and loadings using SVD.

2.5 Canonical discriminant analysis

The canonical discriminant analysis (CDA)[27], also known as fishers discriminant analysis (FDA) is a orthogonal projection technique used in statistical analysis used for group separation. Given N observations, p variables and k groups, CDA calculates a set of $k-1$ linear functions from the p given variables, known as canonical discriminant function (CDF) or corresponding to loadings. Quadratic discriminant functions may also be used. A projection of the original observations onto the calculated CDF, yields a new set of observations, equivalent to score values. The discriminating functions are calculated by maximizing the separation of the between group scatter ($\mathbf{\Sigma}_B$, Equation 2.5 Page 24), while minimizing the within group scatter ($\mathbf{\Sigma}_w$, Equation 2.5 Page 24).

$$\begin{aligned} \mathbf{\Sigma}_B &= \sum_{j=1}^{k} (\mu_j - \mu)(\mu_j - \mu)^T \\ \mathbf{\Sigma}_W &= \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\mathbf{X}_{ij} - \mu_j)(\mathbf{X}_{ij} - \mu_j)^T \end{aligned} \tag{2.5}$$

where μ_j is the mean of the j'th group. The optimal seperation is found by maximizing Equation 2.6 Page 24

$$R(\mathbf{p}) = \frac{\mathbf{p}^T \mathbf{\Sigma}_B \mathbf{p}}{\mathbf{p}^T \mathbf{\Sigma}_W \mathbf{p}} \tag{2.6}$$

The CDA is the underlying calculation used to check for separation in a Multivariate Analysis of Variance (MANOVA) where Wilk's Λ is used to check the significance of the group separation. CDA was used in Paper A Page 59 and Paper C Page 99 to create a score for the degree of browning in baking products.

2.6 Maximum autocorrelation factors

Proposed originally by [36] as an alternative decomposition technique to PCA, the maximum autocorrelation factors (MAF) creates a set of orthogonal vectors similar to PCA but instead of maximizing the variance, the MAF seeks to minimize the autocorrelation defined by the relationship between neighboring pixels. By incorporating this spatial structure into the transformation, it is optimized to highlight spatially connected areas of with similar spectral shapes. This is often very handy in multispectral imaging. A more general and similar transformation is the maximum noise fractions [35, 37] which searches for latent components by maximizing the signal-to-noise ratio. The MNF is a quite general procedure and the transform is equivalent to a transformation of the data to a coordinate system in which the noise covariance matrix is the identity matrix, followed by a principal components transformation. The criterion being maximized in the MNF is

$$R(\mathbf{p}) = \frac{\mathbf{p}^T \mathbf{\Sigma}_\Delta \mathbf{p}}{\mathbf{p}^T \mathbf{\Sigma} \mathbf{p}} \tag{2.7}$$

where $\mathbf{\Sigma}$ is the covariance structure of the signal also used in PCA and $\mathbf{\Sigma}_\Delta$ is the covariance structure of the estimated noise and $\mathbf{p}$ is the loading vector.
Thus, in order to maximize Equation 2.6 Page 24 the noise covariance needs to be estimated, which in MAF is done by calculating the covariance structure of the differences of neighboring pixels. Practically this is done by shifting the image in two directions and substract the result from the original image: $x_\Delta = x(r) - [x(r+\Delta_h) + x(r+\Delta_v)]/2$. The covariance structure of the autocorrelation may now be calculated from the difference values. Further details are found in [38] where kernel versions of MAF and MNF also are explained for non-linear cases.

2.7 Partial least square regression

The partial least square method (PLS) [27] is a common chemometric tool commonly used in e.g. quantitative spectroscopy which uses principles similar to PCA (Section 2.4 Page 23) to find latent structures in $\mathbf{X}$. PLS has some similarities to the regression version of PCA, principal component regression (PCR) [27]. PLS is like MLR searching for an optimal mapping vector that transforms $\mathbf{X}$ to $\mathbf{Y}$

$$\mathbf{y} = \mathbf{X}\beta_{pls} + \epsilon \tag{2.8}$$

PLS decomposes $\mathbf{X}$ and $\mathbf{v}$ into score and loading matrices similar to PCA (Section 2.4 Page 23).

$$\begin{aligned} \mathbf{X} &= \mathbf{TW} + \epsilon \\ \mathbf{Y} &= \mathbf{UQ} + \epsilon \end{aligned} \tag{2.9}$$

$\mathbf{T}$ and $\mathbf{U}$ contain the score values of $\mathbf{X}$ to $\mathbf{Y}$ while $\mathbf{W}$ and $\mathbf{Q}$ contain the loadings. Using an algorithm such as the SIMPLS or NIPALS algorithm [39] the decomposed variables are used to find β_{pls}.
In PLS the decomposition of $\mathbf{X}$ during regression is guided by the variation in $\mathbf{Y}$. This means the co-variation between $\mathbf{X}$ and $\mathbf{Y}$ is maximized, so that the variation in $\mathbf{X}$ directly correlating with $\mathbf{Y}$ is extracted. In the above $\mathbf{Y}$ is written as a matrix in the general sense, however it often only consist of a vector $\mathbf{y}$.
The PLS model is dependent on how many latent components it should retain

which should be selected by the user according to the guidelines specified in Section 2.9 Page 28.
The reason for PCR and PLS's popularity in chemometrics is their ability to intuitively handle datasets having high collinearity between the independent variables as well as having much more variables than observations. When the reference variable is binary or nominal such as those from sensory panels, the Partial least squares Discriminant Analysis (PLS-DA) is an often preferred tool [40], however this method has not been used in this thesis.

2.8 Least squares and penalization

The multiple least squares regression (MLR) is a classic tool for relating $\mathbf{X}$ to $\mathbf{y}$. The basic idea is to best possible fit a hyperplane in $\mathbf{X}$ such that the distances between the fitted hyperplane and the real values in $\mathbf{y}$ is as small as possible in a euclidean sense.
Formally the MLR is a hyperplane and may be written as in Equation 2.10 Page 26, assuming that $\mathbf{y}$ is centered around 0.

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \tag{2.10}$$

The hyperplane may be found directly by setting up the residual sum of squares (RSS) and setting the gradient to zero. This is possible since RSS is a convex function with a single minimum Equation 2.11 Page 26.

$$RSS(\beta) = \sum (\mathbf{X}\beta - \mathbf{y})^2 \tag{2.11}$$

The result of setting the gradient of RSS to zero and solving is the well known normal equations which yields estimation values (Equation 2.12 Page 26).

$$\widehat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{2.12}$$

The MLR is a simple and powerful approach and often yields good results. However, due to the inversion of $\mathbf{X}^T\mathbf{X}$ MLR has problems with highly correlated and underdetermined systems such as spectral systems which usually have more variables than samples. When this is the case $\mathbf{X}$ will be singular and $\mathbf{X}^T\mathbf{X}$ thereby difficult to invert.
Singularity problems in MLR may be overcome be altering RSS. This is exactly what Ridge Regression (RR) [27] does. By imposing a constraint on RSS to decrease its solution space numerical instability is enforced and a solution may be found even though $\mathbf{X}^T\mathbf{X}$ is singular.

$$RSS(\beta) = \sum (\mathbf{X}\beta - \mathbf{y})^2 \ , \quad s.t. \ ||\beta||_2^2 \leq s \tag{2.13}$$

The quadratic constraint controlled by the complexity parameter s will have the effect that parameters are shrunk towards 0 and each other. Since the problem is a constrained quadratic optimization problem it may be recast into lagrangian form.

$$\beta = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y} \tag{2.14}$$

This problem is solved by fixing λ at different values and selecting the one which performs best (see Section 2.9 Page 28).
An alternative way of regularizing MLR is seen in Equation 2.15 Page 27 which is also known as the lasso [41] which imposes a piecewise linear constraint.

$$RSS(\beta) = \sum (\mathbf{X}\beta - \mathbf{y})^2 \quad , \quad s.t.\ |\beta| \leq s \tag{2.15}$$

The shape of the constraint ensures that "questionable" variables are forced to 0, which thereby gives a sparse solution set or parsimonious model. This means that only the variables significant for the model are left in the model. Due to the non differentiable constraint this regularization has been a large research area for some years. Different solutions for efficient implementations on this optimization problem have been formulated and one popular methods of solving this problem is by using the Least Angle Regression algorithm [42]. Similar to RR Lasso is optimized by calculating the solution for a series of values for s and choose the value which performs best (Section 2.9 Page 28)
A similar regularized method which offers sparse solutions is the elastic net (EN) [43] method Equation 2.16 Page 27. EN is a combination of Lasso and RR which besides offering sparse solutions also groups highly correlated variables.

$$RSS(\beta) = \sum (\mathbf{X}\beta - \mathbf{y})^2 \quad , \quad s.t.\ |\beta| \leq s_1 \quad ||\beta||_2^2 \leq s_2 \tag{2.16}$$

Optimizing EN is slightly more tricky than RR and Lasso since two parameters are involved, however by fixing the ridge constraint s_2 and solving for the entire regularization path of s_1 (That is until the least squares estimate has been reached or a defined maximum of s_1) a value may for s_1 may be selected. This is repeated for a series of s_2 values, thus the parameters needs to be selected on a 2-dimensional surface.
The sparse property of Lasso and EN is a consequence of the piecewise linear constraint. Sparse models is a source for variable selection which may be very strong in some scenarios. In settings such as spectroscopy where there are hundreds of variables it may improve the model if "noise" variables are remove from the set, leaving only the important variables. A common exercise for people doing chemometrics is to assess loadings of their model fit to understand which areas are important for the experiment. By using sparse methods it might become very clear which variables hold the necessary information, as well as ordering the importance of variables. This might be important since not only performance but also model interpretation is important.

Some alternative ways of creating parsimonious models includes forward stepwise selection [27], significance testing by jack-knife estimation of parameter uncertainty [44], interval PLS (iPLS) [45] and genetic algorithms [46]. The elastic net procedure was used in Paper G Page 171.

2.9 Parameter tuning and model assessment

When regression methods such as PLR, PLS, RR, Lasso or EN are used it is necessary to select one or more complexity parameters. Selecting complexity parameters is a challenging task and there are no clear rules for getting the parameter "correct". The problem evolves around the bias/variance trade-off [27]. The bias variance tradeoff describes how the variance of the error will increase when the model complexity increases while for very simple models the variance will by low. On the other hand when the model gets complex the bias will decrease while being larger for more simple models. The concept is depicted in Figure 2.4 Page 28. As an example the RR has a complexity parameter s which

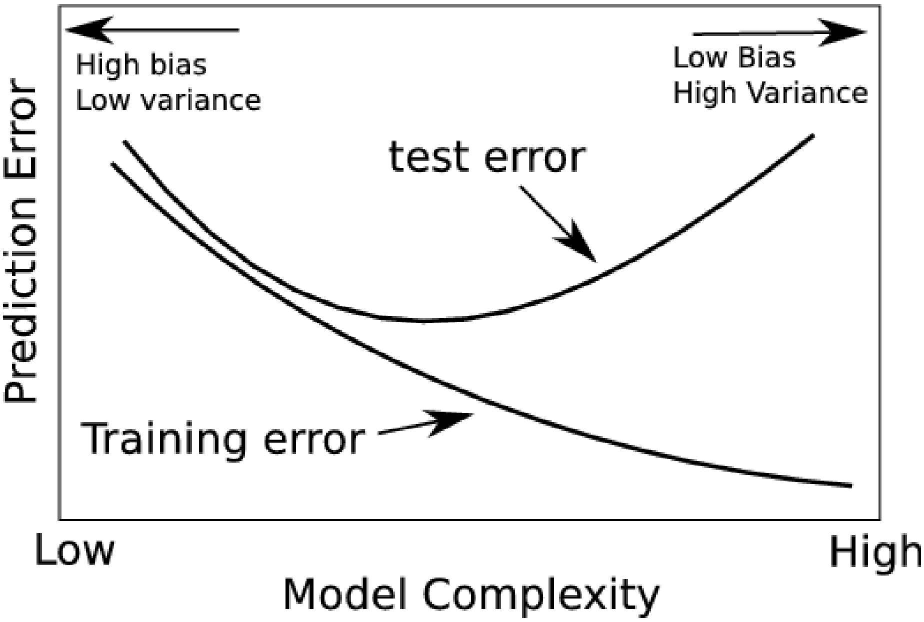


Figure 2.4: Theoretical bias/variance trade-off and effect on training and test error.

increase the solution space for larger values and decreases the solution space for smaller values. This means for larger values of s, the bias decreases while the error variance increases according to Figure 2.4 Page 28. The same is the case for the Lasso.

Figure 2.4 Page 28 furthermore shows how the training error decreases for larger complexity while the test error requires a balanced selection for the optimal model - overfitting the training set. This is a problem when working with small datasets, which often is the case when working with foodstuffs. Creating reference values for models is expensive and difficult. This means data often is scarce and model evaluation is a challenge. To select complexity parameters properly to also maintain a correct generalization ability different methods exists. Two popular methods are the cross validation [27] and the bootstrapping [47] methods. These are both techniques which creates smaller samples from a larger dataset which is used for training and testing. The largest difference is that the cross validation is sampling without resubstitution while bootstrapping is with resubstitution.

The idea behind cross validation is very intuitive. The full set of data is split up into K folds. All folds are looped and each fold is in turn held out while the model is fitted to the remaining folds of data. A test error is then calculated for each left out fold and collected in the end, representing the test or generalization error. The entire procedure is necessary to repeat for each complexity parameter value in order to create a curve similar to the test error curve in Figure 2.4 Page 28. Needless to say the entire procedure grows for larger values of K, where the extreme case is the leave one out cross validation which may be very computationally expensive. The collective error term is calculated as Equation 2.17 Page 29. When doing cross validation there is a trade-off between the number of samples left out and the variance and bias of the estimated performance. Increasing the number of folds decreases the variance in the predicted performance, but at the same time more repetitions of any model fit used in the cross-validation implicitly means higher computational load. Ten-fold cross-validation is usually recommended as a reasonable compromise between both computational cost and the bias/variance trade-off.

$$RMSECV = \sqrt{\frac{||\widehat{Y} - Y||_2^2}{N}} \tag{2.17}$$

Many authors in the literature report $RMSECV$ directly as the result which may be a bit misguiding. Even in starved data settings the cross validation should only be used to select the complexity parameter. Optimally an independent test set should be used to assess the model, but if there really is not enough data to create a test set alternatives such as the two deep cross validation [48] or bootstrapping should be used to assess the model and reported, although it will never be able to substitute having proper amounts of data.

The bootstrapping method suggested by Bradley Efron [49] also poses a very simple and intuitive approach. The dataset is sampled K times with resubstitution to create K new datasets. The model is then trained on each of the K datasets and tested on the original dataset. The bootstrap mimics the cross validation a since when testing on the original dataset, only samples not contained in the sample set should be predicted. Finally, all predictions should be collected. In this thesis bootstrapping have been used in some cases to estimate the accuracy on statistical values.

Chapter 3

Experimental work

Three areas related to food science has been investigated. Browning of foodstuff, mainly non-enzymatic browning of bakery products and meat but also enzymatic browning of vegetables. Cartenoid pigment in salmonids as well as spoilage detection in raw pork meat. All datasets consists of multispectral images recorded with VideometerLab together with one or more independent variables consisting of either sensory labels, chemical measurements or similar. Quality parameters depends on the product, however quality parameters recurs in different contexts/products e.g. it is important to able to predict the water content in both cookies baking in the oven as well as meat being fried in order to infer conclusions on the browning . An overall list of parameters which have been investigated and quantified in this thesis enumerates general surface color, water content, bacterial assessment and fat content. For a better insight in the types of data, a pseudo RGB representation of a random sample from each dataset analyzed in this thesis is seen in Figure 3.1 Page 32.
For each image a region has been extracted and the mean spectrum of this region has been plotted together with one standard deviation of the pixel values in the region, shown as a shaded area. The eight datasets from which the images have been extracted, are also listed in Table 3.1 Page 33 together with information about the reference data and sample sizes.

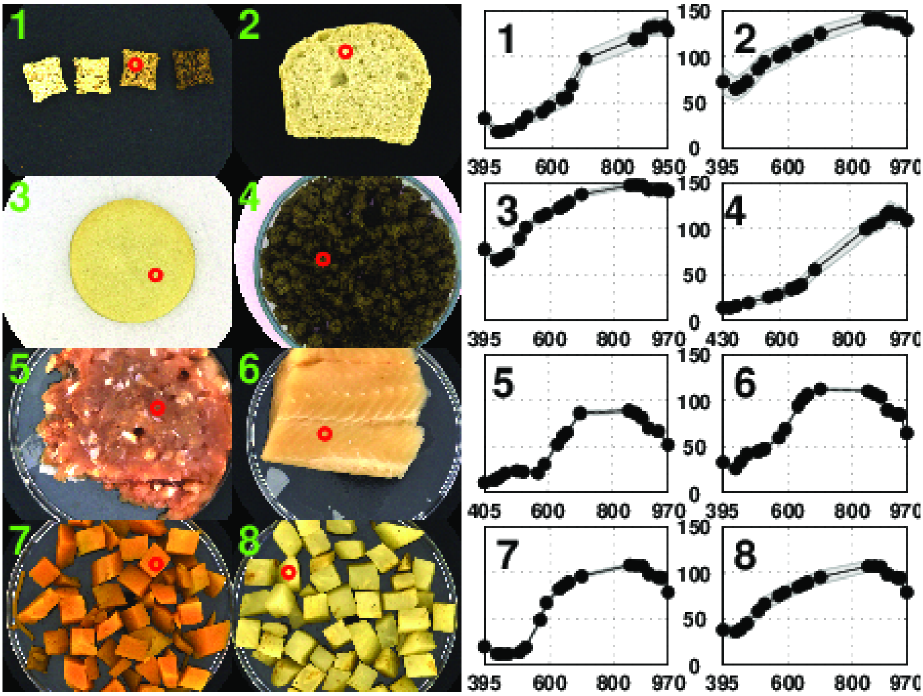


Figure 3.1: Pseudo RGB images of all types of foodstuffs analyzed in this thesis. From upper left; oat flakes, Bread, butter cookie, minced fried beef, minced raw pork, trout, carrots, celariac. In the right side plots shows the mean spectrum of the area indicated with red circles in the rgb images. The shaded area shows one standard deviation for the same area. Note that these spectra cannot be directly compared since different camera setup was used for the different datasets. However, a rough comparison gives an indication of the difference between the spectral shapes.

3.1 (Non-) Enzymatic Browning

3.1.1 Bakery products

The baking industry is facing the challenge of producing products with uniform appearance, physical texture, taste and aroma. These are all quality parameters the consumer evaluates either consciously or unconsciously before or under consumption. Quite naturally the first parameters the consumers evaluate is the appearance and the aroma. The appearance of a food product consists of the colour and the texture product. Consumers automatically associates a lot

#	Foodstuff	# samples	Reference data	App.
1	Oat flakes	90x3 & 11	Water content, time temperature, sensor panel	A Page 59
2	Bread	31x2	Time	B Page 81
3	Butter cookies	35x3	Water content, time temperature, sensor panel	C Page 99
4	Fried minced beef meat	12x3	Water content	D Page 119
5	Raw minced pork meat	167	Microbial count time, temperature storage atmosphere sensor panel	E Page 127
6	Trout Fillet	59	Astaxanthin content Fat content	F Page 151
7	Carrot pieces	6x2+5x2	Time	G Page 171
8	Celariac pieces	6x2+5x2	Time	G Page 171

Table 3.1: Foodstuffs analyzed using VideometerLab in this thesis.

of quality parameters with the appearance such as flavour, nutrition, health and general level of satisfaction. This means the impact of appearance of the baking product on the turnover is significant, and thus a factor that should be optimized.

It turns out that the production parameters related to color change in baking products often are related to the shape and texture as well. Basically the appearance of a baking product may be controlled by “dough” and “baking” parameters. The dough parameters controlling the appearance parameters such as ingredients, amounts of ingredients, mixing time, fermentation time, storage time and environment before baking. External parameters affecting the appearance are such parameters as baking time, baking temperature and oven ventilation.

Visual appearance of baking products is very dependent on the so called browning process. The browning process of baking products describes the transformation from dough to finished product and occurs during heat treatment of the product such as baking in the oven. The browning process undergoing during heat treatment is called non-enzymatic browning and consists of basically three (depending on the “dough” and “baking” parameters) different processes; The Maillard reaction [50], the caramalization [50] and pyrolysis [50]. The reactions are highly dependent on the pH value of the environment, but generally for baking products the pH value is ≤ 7 which is also the scenario that will be assumed for the remainder of the text.

Pyrolysis of a food product is caused by heat alone and is a total scorching which happens at very high temperatures and involves the total loss of water from sugar molecules and the breaking of carbon-carbon bonds [50]. Popularly, pyrolysis is used to remove water from coffee beans without baking them and create unwanted aroma and flavor compounds created by Maillard or caramelization reactions. Pyrolysis leaves the food product as a charred remnant which may lead to a contamination problems in industrial ovens. Pyrolised food remnants may form little particles and contaminate "good" products which lowers the quality of the visual appearance on those products. Paper A Page 59 shows briefly how pyrolized particles attached to the surface of oat flakes may be detected.

The process of caramalization is a complex process which is still not fully understood. The process occurs when reducing sugars such as sucrose, glucose or fructose are heated to and beyond a certain (depending on the sugar type) temperature. The reaction produces flavour and pigment molecules in reaction steps depending on the amount of time the reducing sugar is heated. These molecules are in order isosacchrosan, caramelan, caramelen and caramelin or humin [51]. If carmelization is allowed to proceed to far the taste of the mixture will become less sweet as the original sugar is destroyed. Eventually the flavor will turn bitter. Caramelization often affects the surface colour of baking products to some extent.

The Maillard reaction is like caramelization a complex reaction which only is understood to a certain degree. Since Maillard typically is the browning process which is first triggered, typically when the water activity decreases to 0.4-0.7 and temperature surpasses $105-120°C$ [54], this reaction is responsible for a major part of the browning in baking products and have therefore been given a lot of attention regarding baking products. For the Maillard reaction to trigger basic components such as amino acids, reduced sugars, heat and low water activity is necessary. These are all abundant in baking products which means the Maillard reaction as mentioned before governs a large part of the browning of baking products. The reaction is responsible for both highly appreciated as well as unwanted effects. The reaction is progressing in three steps being the initial, intermediate and final step. In each step very complex reactions are happening which is beyond the scope of this thesis to describe and the reader is referred to works such as [50, 52, 53] for further details. The actual coloration is starting in the second and third stage and is mainly affected by the melanoidin which is responsible for both flavors and color and is the actual end product of the Maillard reaction. Bi products include antioxidants, loss of C-vitamin activity, hydroxymethylfurfural (HMF) and acrylamide where the two latter are toxic compounds which should be kept as low as possible. A quantification of the browning has as described in Paper C Page 99 been studied using light reflection techniques such as cameras and Minolta colorimeters with good results.

Besides understanding the browning of baking products it is equally important to account for subjective consumer demands. Quality is a subjective consumer demand which makes it difficult to quantify. Sensory panels is a common way to quantify subjective measures such as visual appearance and is often used in the food industry for product evaluation. In this project it is proposed how expert knowledge may be incorporated in the quantification of the surface color on cereals and butter cookies as opposed to only quantifying the surface color. Furthermore it is also proposed how the water content of the surface in oat flakes and butter cookies may be quantified. The water content has a clear correlation with the browning process, which makes it interesting to be able to quantify. Finally, as an alternative baking dataset the internal structure of breadcrumbs is investigated as well. The formation of the bread crumb has large impact on the appearance as described in the papers.

Paper A and Paper C revolves around datasets containing oat flakes and butter cookies. In these datasets, the impact of baking time and temperature on the surface colour acceptability based on a sensory panel was investigated. By building a model based on sensory evaluations, an acceptance score is created. The acceptance score is used to predict cookie and oat flake acceptance stages at various times and temperatures. These predictions are then again used to create a quadratic response surfaces to describe the relationship between the involved variables. The acceptance score is essentially a CDA transformation which maps new spectral images into a subspace where three classes chosen by a sensory panel are linearly optimally separated. The three classes represents three states of burn-degree; Under, Adequate and Over burned. The oat flake dataset consisted of 90 samples in three replicates while the buttercookie dataset consisted of 35 samples in three replicates.

Beside time and temperature in the oven, the browning of the product also depends on the water activity [54] The water content for both the oat flakes and butter cookies were measured using the oven method where the weight of a sample was measured before and after insertion in a drying oven at 105°C. For the oat flakes a separate dataset consisting of 11 samples was created while the water content for the butter cookies were measured on the same 35 samples as were used to calculate the response surface for acceptance scores. Both models showed good predictability response, however the dataset from the oat flakes is very small and further investigations should be undertaken to ensure the models validity further.

When the baking product is bread, as in Paper B, the surface color of the crust is also dependent on the water activity. The location of water in the internals of a loaf of bread baking is however very dependent on the structure of the breadcrumb, which means quality assessment of the crumb structure is likewise important. A disadvantage of using a normal 2D camera to assess the crumb structure is that it is necessary to destroy the bread to look at the crumb. A number of publications have described CT scanning of bread to assess the crumb in a non destructive manner. In the work presented in Paper B a set of breads were created to study the development of air bubbles in bread in relation to baking time. A total of 31 breads have been baked with varying baking times ranging between 2 and 26 minutes. The air bubbles were studied as a function of time by extracting spatial features from the images. To observe the air bubbles developed over time, development of extracted features were plotted as a function of time.

Acceptance scores for the oat flakes and butter cookie dataset based on expert knowledge and training samples enabled the creation of acceptance surfaces which indicates areas for acceptable browning. For the oat flake response surface, the temperature showed a larger impact on the acceptance score than the time. This was not the case for the butter cookie response surface. However, temperature and time scaling in the two experiments was set different. The temperature range for the oat flake response surface was from 120°C to 220°C and Butter cookie was from 150°C to 200°C. The time range for the oat flake set was from 0.5 to 9 minutes while the butter cookie was from 4 to 16 minutes. Furthermore the structure of the two products is quite different. While the oat flakes is a hollow oat product, the butter cookies is a solid structure with primarily wheat flour. The loadings used to create the acceptance score showed the most important wavelengths to be 525 and 660 nm corresponding to the green and orange area while for the butter cookie the most significant wavelengths was 395 and 525. The latter corresponds better with the absorption properties of melanoidins formed during the non-enzymatic browning which has absorption in the area around 395 nm [55]. One possible explanation for the significant wavelengths for the oat flakes may be that due to the fact that the oat flakes has already undergone heat treatment in the factory, Maillard reactions have already transformed major parts of the amino acids and reduced sugars on the surface to melanoidins. Another explanation for the difference in wavelength might be the difference in product ingredients. The loadings for water prediction showed larger correspondence as the near infrared region was mainly responsible for predicting water content in both the oat flakes as well as the butter cookies. The dataset for the butter cookies was larger and represented the actual population somewhat better than the oat flake water calibration dataset. Both models were assessed with a bootstrap method and it was found that the SEP for the oat flake water model was approximately 3.8% while the corresponding SEP for the butter cookies was around 13.6%. The butter cookie dataset is a larger dataset and therefore the butter cookie water prediction model will most likely also be more robust. The visualization techniques presented in Paper C presents a good way to assess single cookies on a processline independently from other cookies both for acceptance score but also water content. the two prediction models may be combined to get a better impression of the overall baking state of a given cookie in the processline.

The studied breadcrumbs in Paper B showed that a segmentation of the air bubbles in bread was possible using simple segmentation techniques and that features describing the crumb structure indicated that after 6 minutes baking time at 180° they had reached a steady state. Literature shows that the relationship between the breadcrumb structure and the crust appearance is very significant. The ability for a multispectral imaging system to asses spatial features such as this makes it very powerful tool for creating an overall quality score for baking products based on spectral as well as spatial features.

3.1.2 Minced Meat

The Maillard reaction described in the previous section regarding baking products is also governing the browning of meat products. Meat consists primarily of proteins and thereby amino acids and also has large amounts of glucose and nitrogen-containing compounds. Thus, when exposed to heat at temperatures around 110°C or above the Maillard reactions starts the non-enzymatic browning of the meat while water evaporates. When stir frying meat in the industry e.g. woks such as [56] knowledge about the water content is necessary in order to control the non-enzymatic browning. This is very similar to the problems in the baking industry as described in previous section. A dataset based on work done in [57] were used together with dry matter measurements to create a prediction model for water in meat in Paper D Page 119. A large set of features is created from the multispectral images in order to find band interactions which might describe the water signal better than the pure spectral bands. Instead of using the mean spectrum of the different band interactions, a set of percentiles is calculated based on the histogram of each band interaction. This amounts to a total of 324 variables from which the best variables needs to get selected. A leave one out cross validation was used to select the best possible model which turned out to be 970 nm, exactly at the water absorption peak. The second most significant variable was found to be 850 nm while the final was a ratio between 920 and 910 nm. If a system such as this would be employed at-line, interference from the browning process should in theory not interfere with the water model due to its ability to only focus on specific bands in the near infrared area, where products of the Maillard process do not have any influence on the spectrum [55].

3.1.3 Vegetables

Browning in non heated environments also known as enzymatic browning is a chemical process which occurs in fruits and vegetables by the enzyme polyphenoloxidase, which results in brown pigments [50]. Enzymatic browning may occur in low temperature settings such as in the refrigerator and is a problem for the shelflife of semi-prepared convenience foods such as stir-fried vegetables. In this experiment car- rots and celeriac were investigated over a period of 14 days. Prior to the experiment the vegetables were prefried and frozen at -30°C for four months. During the 14 days the experiment lasted, the vegetables were kept at 5°C in refrigeration. Every second day the carrots and celariac were imaged with a multispectral device. The spectral images were analyzed and revealed a change in the reflection over the 14 day period. This very slight change in reflection over time is most likely due to an enzymatic browning in the plastic bags which were not completely air proof. To determine the reflection change a large featureset was created based on histogram shapes of interactions of wavebands. The shape of the histograms was expressed using percentiles, where the idea is that if certain percentiles in certain waveband interaction changes in accordance with the time, this correlation will be detected by the regression method employed. Due to the very large feature space created, an elastic net method was used to create a sparse solution with the most significant variables. . A watershedding segmentation algorithm was used to separate the particles and treat each of these particles as independent objects. In this way a large set of observations was created. The results were shown as a group plot where each group represented the time spent in the refrigerator. A set of pairwise Bonferroni adjusted t-tests showed a significant reflection change most likely caused by enzymatic browning.

3.2 Microbial Growth

In USA alone the retail market of the beef industry amounted to $76 billion in 2008 [1], yielding efficient storage and distribution of the meat extremely important factors. The heterogeneous nature of the raw meat makes quality assurance difficult during meat production and processing. The chemical composition, technological and sensory attributes are highly influenced by pre-slaughter (e.g., breed, age, environment) intrinsic (e.g., pH, available nutrients) and extrinsic (e.g., storage method, period and temperature of storage) factors [58]. Consequently, in order to keep the quality standards as close as possible to the preference of the consumer, control procedures must be undertaken including sensory, microbiological and physico-chemical analysis.
An EU-project [2] is mapping more than 50 such methods which have been employed for the characterization of microbiologically spoiled or contaminated meat [58, 59]. These methods suffer certain disadvantages as they are time-consuming, destructive, require highly trained personnel, provide retrospective information, and are therefore unsuitable for online monitoring [60, 61]. The same EU-project is investigating methods for rapid, non-invasive probing based on analytical instrumental techniques, such as Fourier transform infrared spec-

---

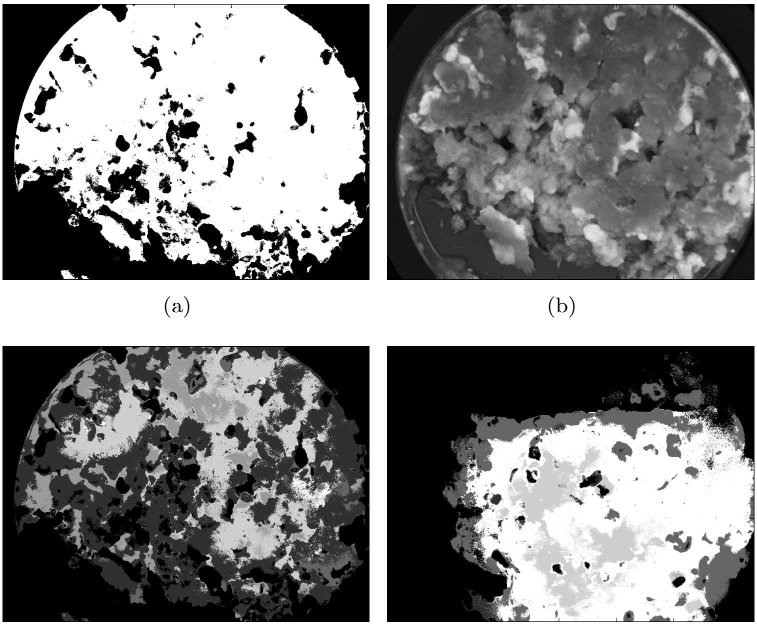
[1] United States Department of Agriculture (USDA), 2008. U.S. Beef and Cattle Industry: Background Statistics and Information. http://www.ers.usda.gov/ news/BSECoverage.htm assessed 16.02.2010.

[2] http://www.symbiosis-eu.net/

troscopy, Raman spectroscopy and near infrared spectroscopy for their potential as reliable meat quality assessment tools. The investigations are largely carried out as academic research papers such as [61]. Besides techniques such as the above mentioned which are based on point measurements, spectral image analysis has a potential due to its ability to utilize information about the heterogeneous nature of the meat. The common underlying concept behind using spectral techniques for assessment of meat spoilage is based on the assumption that the metabolic activity of microorganisms on meat results in biochemical changes with the concurrent formation of metabolic by-products, indicates spoilage. The quantification of these metabolites constitutes a characteristic fingerprint providing information about the type and rate of spoilage [59].

In Paper E Page 127, VideometerLab was used on a set of raw minced pork meat to investigate the ability to track spoilage in different samples of meat. Bacterial counts of a total of 167 meat samples stored for up to 580 hours were quantified using traditional laboratory methods. Meat samples were contained under 2 different storage conditions - aerobic and modified air packages as well as under different temperatures. A large set of bacterial counts including total viable counts, pseudomonads, yeasts and moulds, enterobacteria counts, brochothrix thermosphacta and lactic acid bacteria was performed on each meat sample. Beside the bacterial counts, a sensory panel has judged the spoilage degree of all meat samples into one of three classes on a hedonic scale. The work carried out here primarily focuses on modeling total viable count as well as modeling the sensory evaluation labels.

Previous studies on similar data (beef) used FTIR spectroscopy to predict the spoilage degree [61]. In this work a standard error of prediction of 12% was obtained as the overall regression result. In our work based on pork we were able to obtain a standard error of prediction of 7.5%. To obtain this result the heterogeneity of the meat surface was exploited using preprocessing with a clustering algorithm. Masks indicating areas of meat spoilage for samples were calculated as seen in Figure 3.2 Page 40. Clusters highly correlated with the spoilage degree were used in the further analysis. The cluster center appears to differ in the near infrared area where it has a slightly larger gradient then the remaining clusters. This spatial inference shows the real power of using image analysis instead of averaging methods such as FTIR where sample preparation is necessary. Sensory labels were classified with an overall correctness of 75% which seems quite satisfactory given the fact that laboratory measurement only assessed 80% correct. The largest problem with separation was with the intermediate class. When visually inspecting images of this class it there was large variation in the appearance. The boundaries for this class seems a bit vague and should be given further attention in future studies. An overall conclusion is however that spectral imaging is capable of quantifying spoilage of meat in a setting as the one presented.



(a) (b)

(c) (d)

Figure 3.2: (a) shows a binary mask indicated the area of meat in the image. This piece of meat a total viable count of 9.83 log cfu - indicating heavy spoilage g^{-1}. A grayscale image of the 590 nm waveband for the same piece of meat is seen in (b). (c) shows the spatial distribution of how pixels have been mapped to identified clusters. There are a total of 5 clusters. Dark areas indicate spoiled meat, while lighter areas represent more fresh meat. (d) shows a comparable spatial distribution of clusters in a fresh portion of pork meat. The clusters are colour coded in same order in (b) and (c). The difference in cluster colours in the two images is very clear.

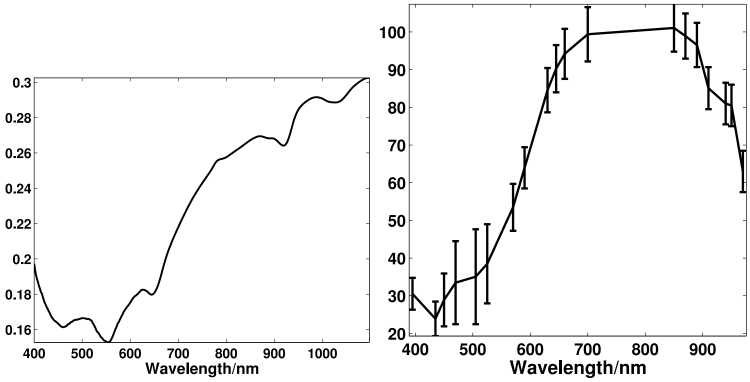
3.3 Cartenoid pigmentation

In the industry of aquaculture homogenization of fish is important for optimal quality. For obvious reasons, homogeneous fish is an impossibility due to natural differences and genes. However by keeping a stable production environment and ensuring as uniform distribution a distribution of fish feed and pigmentation as possible is of great importance for the aquaculture industry.

As is the case for bakery products, color is a high importance when it comes to quality assessment of fish in relation to the commercial production of salmonid fishes. The consumers associate increased intensity of red in salmonid fishes with superior quality, being fresher and having a better flavor [62, 63]. As the change in surface color is the first quality parameter evaluated by the consumer, it is of great economic importance that the color of the salmonid fishes meets consumer preferences. The color of salmonid fishes is caused by deposition of cartenoid pigments in the muscular tissue. Besides being essential for reproduction, proper growth and survival of the fish, carotenoids, primarily astaxanthin and castaxanthin, are also important for the red color in salmonids. Astaxanthin being the single most expensive constituent in the fish feed is sought minimized as much as possible while still maintaining a proper colour.

Rapid assessment of astaxanthin concentration in fish fillets are of great interest for various purposes. One is to ensure the colour of the salmonid at the processing line. Another is to help inspecting the colour of the fish population in the farm.

In Paper F Page 151 a dataset consisting of trout fish has been evaluated for characterization of the concentration of astaxanthin using VideometerLab. 59 fillets of rainbow trout, Oncorhynchus mykiss, were filleted and imaged using VideometerLab for quantitative analysis. The multispectral images were captured prior to determination of the true concentration of astaxanthin. The dataset was split into to parts for training and testing. A predictive model using the partial least squares was developed from the training set and assessed on the independent test set. Previous studies in the literature shows increasing use of multispectral systems for inspection of various parameters of fish [64, 65]. Figure 3.3 Page 42 shows two spectral measurements of astaxanthin. In Figure 3.3(a) Page 42 a NIRSystems 6500 was used to measure the absorbance properties of astaxanthin (See Paper F Page 151 for further details) while Figure 3.3(b) Page 42 shows the reflection of the trout fillet. It should be noted how the largest variation in the trout fillet measurement is situated in the area where astaxanthin has largest absorption properties. This fact is utilized to do the actual quantification of astaxanthin using PLS as described in the paper. Studies using colorimeters to estimate caretenoids showed that the intensity of redness (a*) increases with the carotenoid content in the raw flesh of Atlantic salmon, while lightness (L*) decreases and yellowness (b*) remains unaffected. In our studies we directly estimate the astaxanthin content and find that the



(a) (b)

Figure 3.3: (a) shows reflection properties of pure astaxanthin as recorded by a NIRSystems 6500 absorbance spectrometer and transformed to reflection properties. (b) shows the mean spectrum of all fillets with error bars indicating variation in each wavelength.

most important spectral area is situated from 450 to 570 nm. Measurements from a NIR spectroscopy of pure astaxanthin, also presented in the paper shows that exactly this area represents absorption peaks for astaxanthin. The multispectral images were transformed to color images and a pls model was also trained on these images to compare the performance of the two modalities. This showed that the multispectral images had better prediction power, clear due to the extra information contained in the multispectral images..

3.4 Other considerations regarding spectral imaging

This section contains information not related specifically to food, however indirectly these are techniques that apply to food imaging.

3.4.1 Colormapping

Colour is very important in the food industry as described multiple times in the above. Therefore it is of importance to understand the relationship between spectral images and colour images. Furthermore visualization of true colour

images of spectral recordings is important for scientests, engineers and operators working with such image to get a better understanding of the product.
In order to appreciate the strength of multispectral imaging in the visible area of the electromagnetic spectrum a short introduction to color science follows. As outlined in Section 1.2 Page 5, light is an important aspect of colour. But equally important is the notion of the colour of objects such as green grass, red roses, yellow submarines, etc. The colour of an object is strongly dependent on its spectral reflectance, that is, the amount of the incident light that is reflected from the surface for different wavelengths. A term spectral radiance can be defined as the emitted spectrum of light from a light source or illuminant $l(\lambda)$ for λ being the wavelength. The spectral reflectance in a given surface point of an object is the absorption or reflection properties $r(\lambda)$ at this point. For a given illuminant and object a very simplified definition for the light reflected from this surface point $f(\lambda)$ is given by

$$f(\lambda) = l(\lambda)r(\lambda) \tag{3.1}$$

Thus, reflected colours are products of illumination and surface properties. To quantify this light in the best possible way is to know the shape of the spectrum as close to the truth as possible. A way to sample this spectrum is by imposing filters with certain characteristics $o(\lambda)$ on the spectrum and integrating over the area of the filters as suggested in Equation 3.2 Page 43.

$$P_i = \int_{\lambda} l(\lambda)r(\lambda)o_(\lambda)d\lambda, \quad ,i = 1, 2, ..., p \tag{3.2}$$

Here a system for recording the reflected spectrum in p different intervals are shown. The larger p is, the better is the sampling of the spectrum (given that the filter characteristics are so that the total filter span, spans the entire visible area while each filter is as narrowbanded as possible). The relationship between spectra and color spaces is well defined in the literature [10]. The relationship is basically linear and used in standard colour cameras when the light is filtered in three channels representing red, green and blue (The primary colours in an additive colour space). An example of a mapping from spectral space to a colour space is seen in Equation 3.3.

$$P_i = \int_{\lambda} l_{D65}(\lambda)r(\lambda)o_{XYZ_i}(\lambda)d\lambda, \quad ,i = 1, 2, 3 \tag{3.3}$$

P_i contains the i'th camera response, while the lightsource is set to be a standard light source l_{D65} (standard CIE daylight [10]). r is the reflected light while the filters are set to be a standard observer called $CIEXYZ$ representing a human observer [10].
VideometerLab does not make use of filters. Instead the lightsource is controlled, and the bandwidth of the illumination is relatively narrow. Thus it is assumed that the recording in a spectral band corresponds to the spectral reflection value

at the center point of the filter. Therefore r may be replaced with recordings from VideometerLab and thus a colour sensation may be calculated by integrating from 400 to 700 nm for each observer filter. Naturally there has to be a correspondence between the wavelength recordings of $l_{D65}(\lambda)$, $R(\lambda)$ and O_{XYZ_i}. This may be achieved using numerical interpolation. From P which is the spectral recording mapped to the XYZ colour space, $CIE\ L^*a^*b^*$ colourspace can be directly obtained using the standard nonlinear transformation (Equation 3.4 Page 44)

$$
\begin{aligned}
L &= 116 f_y - 16 \\
a &= 500(f_x - f_y) \\
b &= 200(f_y - f_z) \\
X &= \begin{cases} \sqrt[3]{x_r} & x_r > \epsilon \\ \frac{\kappa x_r + 16}{116} & x_r \leq \epsilon \end{cases} \quad , \quad x_r = \frac{X}{X_r} \\
Y &= \begin{cases} \sqrt[3]{y_r} & y_r > \epsilon \\ \frac{\kappa y_r + 16}{116} & y_r \leq \epsilon \end{cases} \quad , \quad x_r = \frac{Y}{Y_r} \\
Z &= \begin{cases} \sqrt[3]{z_r} & z_r > \epsilon \\ \frac{\kappa z_r + 16}{116} & z_r \leq \epsilon \end{cases} \quad , \quad z_r = \frac{Z}{Z_r}
\end{aligned}
\tag{3.4}
$$

An alternative direct method were created using a special regularized regression method similar to that of (and inspired by) fused lasso [66]. A standard target (Macbeth colorchecker) containing an array of standard colours is used as regression data. By using this method the system can rapidly transform images when the regression has been calibrated since only the only operation which needs to be performed is a simple projection using the obtained weights. A downside with this approach is that a new calibration needs to be carried out if the lightsettings in VideometerLab changes. The weights obtained in the regression shows an intuitive pattern resembling the CIE XYZ color filters compared to a direct linear regression.

3.4.2 Filterdesign

The dataset described in this paper is the only dataset in this thesis which was not recorded by VideometerLab. The camera used here is a narrowbanded NIR camera with 250 filters. As an alternative to selecting single variables using e.g. any cariable selection method described in the introduction chapter, this paper investigates the possibility of creating a subset of filters able to predict the fat content in kobe beef. The filters are designed as ideal bandpass filters with minimum and maximum width constraints which are allowed to overlap. For each filter there are two parameters to be selected, filter width and center point. By testing different amounts of filters and compare the overall predictability by first integrating spectra over filter areas and creating n new response variables, which are used as covariates in an MR regression to predict the fat content. The amount of possible filters in a problem of this is in the order of 10^{4n}, where n is the number of desired filters. Due to the size of the solution space an exhaustive search for the best filter combination is not possible. There is a big dependency problem built into this kind of problem since each filter is only optimal if all other filters are selected optimally. Overall this is an optimization problem with many local minimum. Two search methods are presented to search this very large solution space. One is based on averaging a large set of random guesses while the other is a deterministic greedy search algorithm. There is a huge difference in computational speed between the two methods, while the forward method also seems to obtain lower prediction error. The full spectral sample system investigated in [67] only shows slightly better results indicating that it is indeed possible to come up a set of filters which are able to predict the fat content. The authors have not been able to find similar problems being solved in the literature. An alternative method which would be interesting to test on this problem is genetic algorithms since these algorithms uses a guided stochastic approach.

Chapter 4

Conclusions

The scope of this thesis has been to investigate the potentials of VideometerLab and develop a basis for use in process control of continous baking, roasting and frying processes.
The introduction part of this thesis has described the complexity of quality control in various settings, mainly in the baking industry together with some methods to quantify food quality. Appended papers have clarified even further how different multispectral imaging datasets originating mainly from the baking industry has been used to develop new assment methods. Together with experiments of spoilage detection in raw pork meat and cartenoid pigment detection in trout fish, spectral images seem to be very well suited for quantifying foodstuff.

Foodsuffs are generally heterogeneous at the macroscopic as well as the microscopic level. This variation in the chemical composition cannot be assessed using classical wet chemical analysis. Imaging shows very big potential in utilizing this heterogeneity. Depending on the resolution of the sensing chip in the camera together with the lens, devices such as VideometerLab may be used to detect this heterogeneity and is therefore highly relevant for rapid assessment in such systems. The use of spectroscopic imaging in general is likely to increase in the future as improved imaging techniques emerges and the handling of the large amounts of d$\frac{1}{2}$ata contained in images is no longer limited by lack of computer power. The ability to quantify subjective metrics such as quality with direct quantitative measurements such as the browning of the surface on baking products or the redness of fish fillets furthermore creates use for such devices in the industry as an alternative to or supplement to process operators. Measurement devices such as VideometerLab compared to other multispectral devices is interesting due to the LED technology which is rapidly developing and ever more narrowbanded LED diodes are being are created. The technique is very fast and with upcoming sensor technologies such as black silicon [1] which improves the CCD chips tremendously regarding both light sensitivity as well as detection area, devices such as VideometerLab could be improved to record even more narrobanded reflection areas in a much larger parts of the electromagnetic spectrum which could improve the quantification ability even further.

A look at the literature reveals that the potential of using rapid methods in the baking industry to control the browning process only has been investigated to a certain degree. It has been shown in the literature that in the $CIE\ L^*a^*b^*$ colour system, the overall change in colour (ΔE) and the L^* component are correlated with melanoidins; the product of the Maillard reactions. Comparing this with VideometerLab images has revealed a correlation which together with expert knowledge has been incorporated into an acceptance sccre which may be used to monitor browning and the acceptance stage of baking products. Response surfaces similarly have a potential of being used as charts for automated control of butter cookie or oat flake heating. Visualization techniques has revealed how single elements/particles such as vegetable pieces or butter cookies may be assessed and investigated. The techniques were used to assess humidity and browning development on different types of foodstuff surfaces, which gave an insight in the development of these factors. Furthermore the ability to quantify spoilage detection in raw meat and cartenoid pigment detection in raw fish fillets underlines the versatility of multispectral imaging in the food industry. Further investigations of spoilage detection using VideometerLab could be undertaken to examine the ability to quantify more specific bacterial growth such as e.g. lactic acid or yeasts and mold. The investigations carried out were focused on predicting total viable counts as well as classifying meat according to sensory quality classes. Both experiments turned out successful and demonstrates how the heterogeneity captured by multispectral images may be utilized.

All models were assessed using generalization techniques such as cross validation and bootstrapping to ensure as fair and generalizable as possible. Of course more studies for further assessment is always good, but since reference values are difficult to obtain, techniques such as these are necessary to ensure model validity.

Two studies not involving food products but rather method development were carried out suggesting a novel method to map spectral images into a colorspace as well as a design algorithm for creating more broadbanded filters to predict fat content in kobe beef.

---

[1] http://www.sionyx.com/

To summarize, the two areas of spectroscopic analysis of foods which will presumably expand the most in the near future are on-line analysis and imaging taking both the visual area and thereby colours but also the NIR area and the abundance of information residing here, into account. Focus should be on improving the spectroscopic instrumentation as well as the chemometric tools for these areas of application. In this connection, important chemometric areas will presumably be process analytical chemometrics and multivariate image analysis where various techniques from the machine learning communities will be incorporated to an even greater extent. Already now the field has advanced to incorporating complex methods such as neural networks and support vector machines in various settings, meaning more and more people from the computing communities will be integrated in food technology. This will automatically happen as data amounts increase even further and the need for more precise techniques similarly increases. Computer intensive methods such as bayesian techniques which is of yet relatively uncharted territory in regards to spectral imaging may be seen more in the fields of chemometrics in the future due to its more varied approaches to data analysis. Furthermore as dimensionality in imaging datasets will increase so will the need for using efficient variable selection methods such as those described in the introduction in this paper. In short, multispectral imaging in the food industry has great potential and will certainly be more integrated in the future.

Bibliography

[1] N. Perrot, I. Ioannou, M. Gilles, I. Allais, C. Curt, V. Chevallereau, and G. Trystram. Experiences about operator/technology cooperation using a fuzzy symbolic approach to decision support system design in food processes. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, (1):94–99, 2002.

[2] Timothy A. Haley and Steven J. Mulvaney. Advanced process control techniques for the food industry. Trends in Food Science and Tecnology, 6:103–110, 1995.

[3] S. Linko and P. Linko. Developments in monitoring and control of food processes. Transactions of the Institute of Chemical Engineering, 76, Part C:127–137, 1998.

[4] Hans Büning-Pfaue. Analysis of water in food by near infrared spectroscopy. Food Chemistry, 82(1):107 – 115, 2003. 2nd International Workshop on Water in Foods.

[5] G. Tœgersen, T. Isaksson, B. N. Nilsen, E. A. Bakker, and K. I. Hildrum. On-line nir analysis of fat, water and protein in industrial scale ground meat batches. Meat Science, 51(1):97 – 102, 1999.

[6] Musleh Uddin, Emiko Okazaki, Sandor Turza, Yamashita Yumiko, Munehiko Tanaka, and Yutaka Fukuda. Non-destructive visible/nir spectroscopy for differentiation of fresh and frozen-thawed fish. Journal of Food Science, 70(8):c506–c510, 2005.

[7] B.S. Bennedsen, D.L. Peterson, and A. Tabb. Identifying defects in images of rotating apples. Computers and electronics in agriculture, 48(2):92–102, 2005.

[8] Alan M. Lefcourt, Moon S. Kim, and Yud-Ren Chen. Automated detection of fecal contamination of apples by multispectral laser-induced fluorescence imaging. Appl. Opt., 42(19):3935–3943, Jul 2003.

[9] Bosoon Park, Kurt C. Lawrence, William R. Windham, and Douglas P. Smith. Performance of hyperspectral imaging system for poultry surface fecal contaminant detection. Journal of Food Engineering, 75(3):340 – 348, 2006.

[10] Gunter Wyszecki and W.S. Stiles. Color science: Concepts and Methods, Quantitative Data and Formulae, second edition. Wiley & sons Inc, 2000.

[11] Nahum Gat. Imaging spectroscopy using tunable filters: A review. In Proceedings of SPIE, volume 4056, pages 50–64, 2000.

[12] J. Y. Hardeberg, F. J. Schmitt, and H. Brettel. Multispectral color image capture using a liquid crystal tunable filter. Optical Engineering, 41:2532–2548, October 2002.

[13] Johannes Brauers, Nils Schulte, and Til Aach. Multispectral filter-wheel cameras: Geometric distortion model and compensation algorithms. IEEE Transactions on Image Processing, 17(12):2368–2380, Dec 2008.

[14] Johannes Brauers, Stephan Helling, and Til Aach. Multispectral image acquisition with flash light sources. Journal of Imaging Science and Technology, page in press, 2009.

[15] Stanley R. Crouch, James J. Leary, Douglas A. Skoog, Timothy A. Nieman, and F. James Holler. Principles of instrumental analysis. Brooks Cole, 6 edition, 2006.

[16] Shoji Yamamoto, Norimichi Tsumura, Toshiya Nakaguchi, and Yoichi Miyake. Development of a multi-spectral scanner using led array for digital color proof. Journal of Imaging Science and Technology, 51(1):61–69, 2007.

[17] Jens M. Carstensen and Jørgen Folm-hansen. An apparatus and a method of recording an image of an object, 2000.

[18] J. F. Hansen. On chromatic and geometrical calibration. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 1999.

[19] Paul Geladi, Jim Burger, and Torbjörn Lestander. Hyperspectral imaging: calibration problems and solutions. Chemometrics and Intelligent Laboratory Systems, 72:209–217, 2004.

[20] Jerome Workman, Mel Koch, Barry Lavine, and Ray Chrisman. Process analytical chemistry. Analytical Chemistry, 81(12):4623–4643, 2009. PMID: 19425531.

[21] US Department of Health, Human Services, Center for Drug Evaluation, and Research. Guidance for industry. pat - a framework for innovative pharmaceutical development, manufacturing and quality assurance. Technical report, US Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Rockville, MD 20857, United States, 2004.

[22] Katherine A. Bakeev. Process Analytical Technology. Blackwell Publishing, 2005.

[23] Alessandro Ulrici, Mario Li Vigni, Caterina Durante, Giorgia Foca, Paolo Belloni, Barbara Brettagna, Tommaso De Marco, and Marina Cocchi. At-line monitoring of the leavening process in industrial bread making by near infrared spectroscopy. Journal of Near Infrared Spectroscopy, 16(3):223–231, 2008.

[24] B. Zanoni, C. Peri, and D. Bruno. Modelling of browning kinetics of bread crust during baking. Lebensmittel-Wissenschaft und-Technologie, 28(6):604 – 609, 1995.

[25] Sandra Mundt and Bronislaw L. Wedzicha. A kinetic model for browning in the baking of biscuits: Effects of water activity and temperature. LWT - Food Science and Technology, 40(6):1078 – 1082, 2007.

[26] Douglas C. Montgomery. Design and Analysis of Experiments. John Wiley & Sons, 2009.

[27] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Elements of Statistical Learning: Data mining, Inference and Prediction, 2. edition. Springer-Verlag, 2009.

[28] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. Digital Image Processing. 2 edition, 2002.

[29] Thomas H. Cormen, Charles E. Leierson, Ronald L. Rivest, and Clifford Stein. Introduction to algorithms. The MIT Press, 2003.

[30] Stan Z. Li. Markov Random Field Modeling in Image Analysis. Springer, 2009.

[31] Luc Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Transactions on Image Processing, 2(2):176–201, 1993.

[32] Stanley Osher and Nikos Paragios, editors. Geometric Level Set Methods. Springer, 2006.

[33] Ulf Grenander and Michael I. Miller. Representations of knowledge in complex systems. Journal of the Royal Statistical Society, 56:549–603, 1994.

[34] Nobuyuki Otsu. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9(1):62–66, 1979.

[35] Andrew A. Green, Mark Berman, Paul Switzer, and Maurice D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and remote Sensing, 26:65–74, 1988.

[36] P. Switzer and A. Green. Min/max autocorrelation factors for multivariate spatial imagery. Technical Report 6, Department of statistics, Stanford University, 1984.

[37] J.B. Lee, A.S. Woodyatt, and M. Berman. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. Geoscience and Remote Sensing, IEEE Transactions on, 28(3):295 –304, may 1990.

[38] A.A. Nielsen. Kernel maximum autocorrelation factor and minimum noise fraction transformations. Image Processing, IEEE Transactions on, 20(3):612 –624, march 2011.

[39] Paul Geladi and Bruce R. Kowalski. Partial least-squares regression: a tutorial. Analytica Chimica Acta, 185:1 – 17, 1986.

[40] Matthew Barker and William Rayens. Partial least squares for discrimination. Journal of Chemometrics, 17(3):166–173, 2003.

[41] Robert Tibshirani. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 58(1):267–288, 1996.

[42] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. Annals of Statistics, 32(2):407–499, 2003.

[43] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society, 67(2):301–320, 2005.

[44] Harald Martens and Magni Martens. Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (plsr). Food Quality and Preference, 11(1-2):5 – 16, 2000.

[45] L. Norgaard, A. Saudland, J. Wagner, J. P. Nielsen, L. Munck, and S. B. Engelsen. Interval partial least-squares regression (ipls): A comparative chemometric study with an example from near-infrared spectroscopy. Appl. Spectrosc., 54(3):413–419, Mar 2000.

[46] Riccardo Leardi. Genetic algorithms in chemometrics and chemistry: a review. Journal of Chemometrics, 15(7):559–569, 2001.

[47] Ron Wehrens, Hein Putter, and Lutgarde M. C. Buydens. The bootstrap: a tutorial. Chemometrics and Intelligent Laboratory Systems, 54(1):35 – 52, 2000.

[48] P. Jonathan, W. J. Krzanowski, and W. V. McCarthy. On the use of cross-validation to assess performance in multivariate prediction. Statistics and Computing, 10(3):209–229, 2000.

[49] B. Efron. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics, 7(1):1–26, 1979.

[50] O. R. Fennema. Food Chemistry. Marcel Dekker, New York, 3 edition, 1996.

[51] John Deman. Principles of Food Chemistry. Springer - verlag, 1999.

[52] Harry Nursten. The Maillard reaction: chemistry, biochemistry, and implications. Culinary and Hospitality Industry Publications Services, 2005.

[53] Sara I. F. S. Martins, Wim M. F. Jongen, and Martinus A. J. S. van Boekel. A review of maillard reaction in food and implications to kinetic modelling. Trends in Food Science & Technology, 11(9-10):364 – 373, 2000.

[54] Emmanuel Purlis. Browning development in bakery products - a review. Journal of Food Engineering, 99(3):239 – 249, 2010.

[55] Ji-Sang Kim and Young-Soon Lee. Enolization and racemization reactions of glucose and fructose on heating with amino-acid enantiomers and the formation of melanoidins as a result of the maillard reaction. Amino Acids, 36:465–474, 2009. 10.1007/s00726-008-0104-z.

[56] Jens Adler-Nissen. The continuous wok - a new unit operation in industrial food processes. J. Food Process Engin., 25:435–453, 2002.

[57] Søren Blond Daugaard, Jens Adler-Nissen, and Jens Michael Carstensen. New vision technology for multidimensional quality monitoring of continuous frying of meat. Food Control, 21(5):626 – 632, 2010.

[58] George-John E. Nychas, Panos N. Skandamis, Chrysoula C. Tassou, and Konstantinos P. Koutsoumanis. Meat spoilage during distribution. Meat Science, 78:77–89, 2008.

[59] David I. Ellis and Royston Goodacre. Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. Trends in Food Science & Technology, 12(11):414 – 424, 2001.

[60] D. I. Ellis, D. Broadhurst, D. B. Kell, J. J. Rowland, and R. Goodacre. Rapid and quantitative detection of the microbial spoilage of meat by fourier transform infrared spectroscopy and machine learning. Appl Environ Microbiol, 68(6):2822–2828, June 2002.

[61] A.A. Argyri, E.Z. Panagou, P.A. Tarantilis, M. Polysiou, and G.-J.E. Nychas. Rapid qualitative and quantitative detection of beef fillets spoilage based on fourier transform infrared spectroscopy data and artificial neural networks. Sensors and Actuators, 2009.

[62] T. R. Gormley. A note on consumer preference of smoked salmon colour. Irish Journal of Agricultural and Food Research, 31(2):199–202, 1992.

[63] Stewart Anderson. Salmon color and the consumer. In Microbehavior and Macroresults: Proceedings of the Tenth Biennial Conference of the International Institute of Fisheries Economics and Trade Presentations, July 2000.

[64] Gamal Elmasry and Jens Peter Wold. High speed assesment of fat and water content distributions in fish fillets using online imaging. Journal of Agricultural and Food Chemistry, 56:7672–7677, 2008.

[65] Silje Ottestad, Martin Høy, Astrid Stevik, and Jens Petter Wold. Prediction of ice fraction and fat content in superchilled salmon by non-contact interactance near infrared imaging. Journal of Near Infrared Spectroscopy, 17:77–87, 2009.

[66] Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society Series B, 67(1):91–108, 2005.

[67] Kenichi Kobayashi, Yasunori Matsui, Yosuke Maebuchi, Toshihiro Toyota, and Shigeki Nakauchi. Near infrared spectroscopy and hyperspectral imaging for prediction and visualisation of fat and fatty acid content in intact raw beef cuts. Journal of Near Infrared Spectroscopy, 18(6):301–315, 2010.

Part II

Contributions

Appendix A

Visual Quality Estimation of Oat Cereal

Submitted to Journal of Food Control

Visual Quality Estimation of Oat Cereal

Bjørn Skovlund Dissing[a], Mette Stenby[b], Bjarne Kjær Ersbøll[a], Jens Adler-Nissen[b]

[a]*Informatics and Mathematical Modelling, Technical University of Denmark, Institute for Informatics and Mathematical Modeling, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark*
[b]*National Food Institute, Technical University of Denmark, Søltofts Plads, Building 227, DK-2800 Kgs. Lyngby, Denmark*

Abstract

We consider the possibility of using fast online multispectral imaging in a continuous baking process of oat cereal. Fast multispectral imaging is used as an alternative to pointwise spectroscopy, which is sometimes used in the industry; sacrificing spectral information for increased spatial information. A set of quality parameters which could help automate a manufacturing process is investigated for a specific cereal product. The image analysis is used to quantify the following three quality parameters for oat cereal: color (degree of burn) as a function of time and temperature, detection of coal particles on the surface of the cereals, and water content in the oat flakes. The experiments and data analysis indicate that multispectral imaging may be able to automate the visual inspection previously done by human operators with regard to the parameters under investigation.

Keywords: Multispectral, Quality Control, Chemometrics, Oat, Imaging

1. Introduction

Breakfast cereals are consumed every day, all around the world. The large majority of these breakfast cereals are sold as ready-made batches, which have been prepared in different ways in factories. The preparation very often includes treatment under high temperatures in ovens, which gives the cereals a desired color, flavor, texture and crispness. These factors all play a central role in the consumers' experience of the product, and are therefore vital for the product. However, controlling such factors in large batches of cereals is a difficult task which requires continuous surveillance by experts in the field.

Preprint submitted to Food Control *May 18, 2011*

This is often done using visual inspection and manual sampling. Human experts are an expensive as well as subjective resource, which means a cheaper and less error-prone automation of the surveillance and process control will improve the overall quality and reduce costs.

Automation of surveillance processes in the food industry has been slow, due to inadequate technology. However, in recent years, technology for data acquisition and further reduction of the massive amount of data into meaningful information has evolved rapidly, making research in automation of surveillance processes more feasible. Traditional spectroscopy has been used at-line or in-line and is today the state of the art method for fast non-invasive analysis in the food industry [1]. Its one dimensional nature is a problem however with regard to repeatability and accuracy. This is where imaging becomes interesting, since it adds two-dimensional information to the spectrum and thereby improves the uncertainty issues compared to traditional spectroscopy [2, 3].

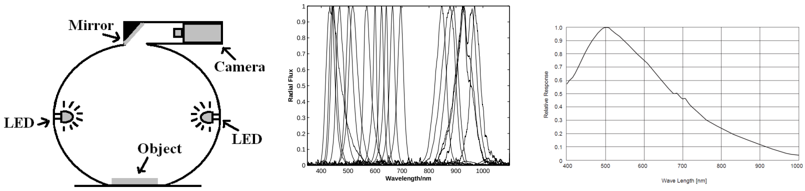
New vision technologies such as VideometerLab [4] are emerging, where a spatial as well as spectral measurement is performed rapidly in form of spectral images. This has the potential of a much closer match to the visual judgment made by the trained process operators, than a 1D measurement.

For full automation of a food preparation facility, it has been shown that it is possible to support the visual inspection of the operator by using learning systems [5] or by replacing different on-line measuring equipment, in combination with appropriate process-control algorithms. Examples are: proportional integral derivative controller(PID) and fuzzy logic or neural networks, and these are beginning to be used more frequently in the food industry [6, 7].

In this work, however, we are only concerned with quality factors of a specific cereal product. To the knowledge of the authors, not much similar work has been performed on cereal products. [8] used texture information from RGB images to classify bulks of different types of cereal. We are presenting ways to quantify surface color, detection of coal particles and moisture in a specific oat cereal product.

1.1. VideometerLab Vision Technology

VideometerLab [4], acquires multi-spectral images at 19 different wavelengths ranging from 385 to 970nm, fully shown in Table 1. The spectral radiation of the 19 bands is seen in a collective plot in Figure 1(b). The acquisition system records surface reflections with a standard camera equipped with a charged coupled device chip with no bandpass filters. The object of interest is placed inside an integrating or so-called Ulbricht sphere, which has its interior coated with a matt-white coating. The coating, together with the curvature of the sphere, ensures a uniform reflection of the cast light, and thereby a uniform light in the entire sphere. At the rim of the sphere, Light Emitting Diodes (LED) are positioned side by side in a pattern which distributes the LEDs belonging to each wavelength uniformly around the entire rim. The LEDs, having the spectral radiant power distributions seen in Figure 1(b), are strobed successively, each resulting in a monochrome image with 32-bit floating point precision. The resulting hyperspectral cube has dimensionality 1280x960x19. The system is first calibrated radiometrically with a diffuse white and dark target, followed by a light setup based on the type of object to be recorded. Furthermore, the system is geometrically calibrated with a geometric target to ensure pixel correspondence across all spectral bands [9]. The homogeneous diffuse light, together with the calibration steps, ensures an optimal dynamic range and minimizes shadows and shading effects as well as specular reflection and gloss-related effects. The system has been developed to guarantee the reproducibility of the images collected. This means it can be used in comparative studies of time series or across a large variety of different samples [10, 11, 12, 13].



(a) (b) (c)

Figure 1: *a) Principle of imaging with integrating (Ulbricht) sphere illumination. The LEDs located in the rim of the sphere ensures narrowband illumination. b) Normalized spectral power distributions of the LEDs located in the VideometerLab. c) Spectral sensitivity of the camera mounted in VideometerLab. It is seen in a) that the camera is placed above the object of interest.*

♯	Wavelength	Color	♯	Wavelength	Color	♯	Wavelength	Color
1	395 nm	UV	7	570 nm	Green	13	850 nm	NIR
2	435 nm	UB	8	590 nm	Green	14	870 nm	NIR
3	450 nm	Blue	9	630 nm	Yellow	15	890 nm	NIR
4	470 nm	Blue	10	645 nm	Red	16	910 nm	NIR
5	505 nm	Blue-Green	11	660 nm	Red	17	940 nm	NIR
6	525 nm	Green	12	700 nm	Red	18	950 nm	NIR

Table 1: *VideometerLab Wavelengths*

2. Materials and Methods

2.1. Havrefras

Havrefras, henceforth referred to as oat flakes, is an oat cereal product for human consumption produced by the company Nakskov Mill Food A/S [1]. The exact recipe is not known, but among other things the oat flakes contain oat flour, wheat bran,sugar, salt and malt. The materials are mixed to a dough, which is extruded [14, 15] into flat squares of about 1 cm^2. The shaped dough is baked in continuous ovens, using hot air. During baking the dough squares inflate into cushions and obtain the desired crispness and color. After baking the oat flakes are cooled, before packing in boxes for storing.

In this work we are interested in a set of quality parameters which are used to identify the state of a given oat flake. The VideometerLab mentioned above is used to make three dimensional measurements, i.e. two spatial and one spectral, and from these measurements a set of quality parameters are quantified and extracted.

The surface color of the oat flakes changes significantly from dough to end product and will give a very clear indication of how burned the cereals are. Coal particles originating from very burned cereals are at a risk of getting stuck on the surface of the newly produced oat flakes, which will lower the quality index of the product. Such coal particles are visually unappealing for consumers and in large amounts potentially harmful. If large amounts are detected in a batch, it should be discarded and the oven should be cleaned. These two quality parameters can be quantified by inspecting the surface

---

[1] http://www.nmfoods.dk

color in the visible wavelengths.

The water content of the oat flakes also indicates the burn-state. The water content must be sufficiently low to obtain the right crisp texture. Therefore, detection of the amount of water in an oat flake is important as another quality parameter.

The oat flakes used for the experiments in this article were bought in a conventional supermarket, thus they have already undergone a preparation treatment in the factory. This treatment, which affects color and waterbinding capabilities, does not matter though, since here we are only interested in basic detection of change in color and water content which are relative measures. Therefore the initial state of the oat flake is irrelevant as long as it is not completely burned.

2.2. Surface burn

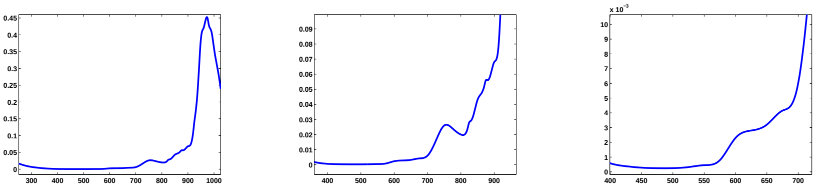
We are interested in identifying the severity of surface burn on a oat flake when baked in the oven. The heat from the oven will trigger Maillard reactions, which have an impact on the surface color. To investigate how time and temperature influence this, a 2-factorial experiment was set up with 6 and 15 levels for temperature and time factors respectively. Single oat flakes were baked in the oven at each of the temperatures 120, 140, 160, 180, 200 and 220°C, for each of the periods 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, 360, 390, 420 and 450 seconds in a randomized order. For each factor combination three repetitions were done, giving a total of 270 samples. To give all cereals equal initial conditions, and thereby not introducing bias, no oat flakes were recycled.

The actual surface color measurement may be done in different ways. Two common approaches are to use either a chroma-meter for very precise color measurements, or a standard digital color camera. The chroma-meter has the disadvantage that it only measures the surface color at a specific point. Both approaches are very dependent on the lighting conditions. Another approach is to use a spectrophotometer, which again is only able to measure the sample at a specific point. A combination of these three approaches is a multispectral camera device which is able to sample the electromagnetic spectrum in narrow intervals, while still obtaining information about the entire surface. Multispectral images of the oat flakes were acquired and saved on the harddisk for further analysis. In order to transform multispectral images to a surface color quantification, a specific transformation needs to be created; the canonical discriminant analysis (CDA)[16], also known as Fisher's

discriminant analysis (FDA) or reduced rank discriminant analysis (RRDA), is a well known and thoroughly tested technique in statistical analysis, used for group separation. Given N observations, p variables and k classes, CDA calculates a set of k-1 linear functions from the p given variables, known as canonical discriminant function(CDF) or loadings. A projection of the original observations onto the calculated CDF yields a new set of observations, often called scores, which optimally maximize the separation of the between-class covariance, while minimizing the within-class label covariance, in the training data. If N is smaller than G-1, where G is the number of groups, the rank of the covariance matrix only supports N loadings. Once the observations are projected, a simple classification technique such as a Baye's classifier or a support vector machine, or even a simple threshold, in the subspace will easily seperate the classes.
Three oat flakes with different surface color were manually picked by experts to represent three classes; A, B and C, i.e. three degrees of surface burn. Multispectral images were acquired of the selected oat flakes. Three regions of interest (ROI) were defined, one for each oat flake.

2.3. Water Content

The water content of oat flakes has a significant influence on crispness, which makes it important to be able to quantify this parameter. As specified in Table 1, VideometerLab contains seven bands in the NIR region, from 850 to 970 nm. Water's absorption properties in the visible and near-infrared region are well known, and shown in Figure 4 for wavelengths between 300 and 1000 nm [17]. A large overtone peak is observed in Figure 2(a) around 970nm. Zooming in around a smaller area of the spectrum, three additional sets of water absorption lines near 820, 730 and 615 nm also appear, as seen in Figures 2(b) and 2(c). In theory these are detectable with a multispectral camera. In order to be able to calibrate a multispectral imaging model for measuring water content in oat flakes, a set of oat flakes with varying water content has been created. To do this we have used five different set-ups with different relative humidity levels: 23, 75, 86, 93 and 97%. The humidity levels were created using different supersaturated salt solutions in closed containers: $KC_2H_3O_2$, $Nacl$, KCl, KNO_3 and K_2SO_4 soluted in water. For each solution a set of six oat flakes was used. Water content has been calculated as the difference in weight between a set of moist oat flakes and the same oat flakes after drying in a drying cabinet. The water content is



(a) Absorption properties of water between 300 and 1000 nm.

(b) Absorption properties of water between 350 and 900 nm.

(c) Absorption properties of water between 400 and 700 nm.

Figure 2: Water absorption properties at wavelengths between 200-1000nm [17]

given in percent, and calculated as in Equation 1.

$$WC(\%) = \left(\frac{W_{wet} - W_{dry}}{W_{dry}}\right) \cdot 100 \tag{1}$$

WC denotes Water Content, W_{wet} is the weight of the moist oat flakes and W_{dry} is the weight of the dry oat flakes . The weight of the oat flakes was measured using a scale (Mettler Toledo AT261 Delta Range Balance (max. 205g/62g d=0.1mg/0.01mg)). Depending on the salt solution, the oat flakes remained in the container a certain amount of hours until no more water was absorbed in the oat flakes. After being taken out from the container, multispectral images of the oat flakes were acquired using the VideometerLab, weighed and subsequently placed in a drying cabinet at 105°C until all the water had evaporated, leaving oat flakes with no residual water. A total of 11 samples were created using the above method, containing from approximately 14 to 37% water, distributed as seen in Figure 3. The multispectral images were segmented to include only spectra from the surface of the oat flakes (Figure 4(a)). A mean spectrum was calculated for all the oat flakes contained in each image, yielding a total of 11 spectra shown in Figure 4(b), which were all centered and standardized before analysis. A logarithmic transformation of the response variable decreased the dependency between the spread and the mean of the water content to obtain a better calibration. In order to relate the multispectral images to the measured water content, a Ridge Regression (RR) [16] has been calibrated. The RR is an alternative to Partial Least Squares (PLS) [16] and Principal Component Regression (PCR)[16] which both make use of latent variables in the data. Due to the

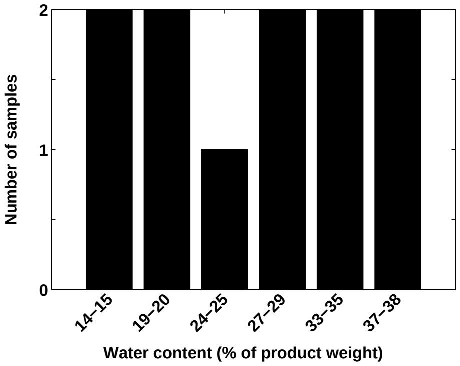
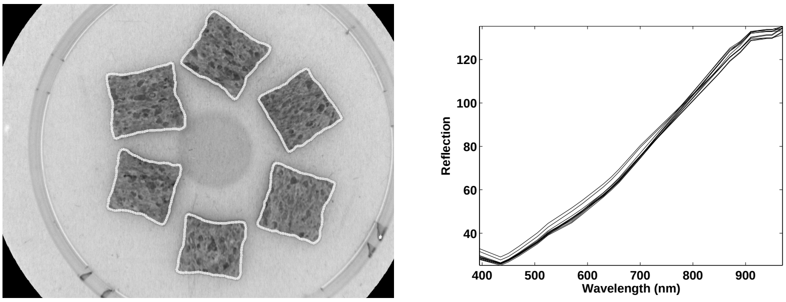


Figure 3: 11 Samples divided into intervals according to water content (percentage of product weight)



(a) (b)

Figure 4: (a) Contours indicating the area of the oat flakes. (b) Mean spectra of the 11 samples. Each spectrum is calculated as a mean value of all spectra contained within 6 oat flakes

ill-posed problem in the dataset containing 11 observations and 19 variables, which has infinitely many solutions, it is necessary to use a modified regression scheme like one of the above. Furthermore, due to the very small dataset of only 11 observations, single observations have a very large influence on the estimated model. RR uses a penalization parameter λ which reduces variance in the model at the cost of introducing bias, the bias-variance tradeoff [16]. Due to the small number of observations, λ was chosen with a cross validation method [18]. The regression model was validated using a 10000-run bootstrapping [19] method to estimate various summation statistics to assess the model. The bootstrap is a Monte Carlo method, which is related to techniques such as the jack-knife and cross-validation.

For model evaluation a set of statistical key values is presented. Ross [20] originally introduced the Bias factor and the Accuracy factor for predictive microbiology. The metrics were originally intended for interpreting models for generation times of bacteria, but it is basically a model assessment tool which is usable for any predictive model. A perfect agreement between the predicted and observed values will lead to a bias factor of 1 while values above 1 indicate that the predictions exceed the observations; a value of 1.1 has a tendency to overpredict by 10% on average. Thus, the further away from 1, the poorer the model. This means a bias factor of 0.5 indicates a poor model.

For the accuracy factor, the larger the value, the less accurate the average estimate. For an accuracy factor of two, the prediction is on average a factor of two different from the observed value - half as large or twice as large.

The mean relative percentage error (MRPE) gives an idea of the overall relative accuracy of the system. The smaller the relative error, the better the predictability. Negative values mean there is a tendency to underpredict while th opposite applies for positive values. A perfect fit gives an MRPE of 0, i.e. there is no bias in the prediction.

The Mean Absolute Percentage Error (MAPE) is similar to MRPE, except the errors do not cancel out. This means the MAPE explains the percentage by which the predicted values deviate from observed values.

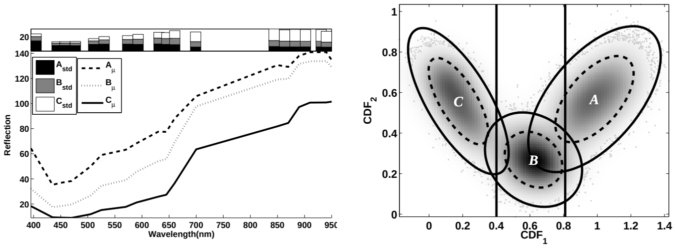
The Root Mean Square Error of Prediction (RMSE) is very similar to MRPE, however here the error is expressed in the same values as the fitted data. Thus, the RMSE shows the uncertainty of the prediction.

The standard error of prediction (SEP) index is a relatively typical deviation of the mean prediction values and expresses the expected average error associated with future predictions. The lower the value of this index is, the better the capability of the model to predict water content in new oat flakes.

3. Results

3.1. Surface burn

The three classes manually picked by experts are seen in Figure 5(a) which shows the mean spectral value of all pixel values across the oat flakes indicated by the ROIs. At the top of the plot the standard deviation of the pixel values is shown as a bar plot. The mean spectra for the three classes are significantly different and are well suited to create a transformation which may convert multispectral images to a surface-burn score. Using all spectra from the defined ROIs, this is verified by a Multivariate Analysis Of Variance (MANOVA), where the value of Wilk's Lambda was found to be 0.04, as shown in Equation 2.



(a) (b)

Figure 5: (a) shows spectral differences between the three classes A, B and C which were manually selected to represent the three degrees of surface color. One oat flake is used for each class, and the mean spectrum is plotted here. In the lower part of the plot, standard deviations are shown as stacked bars for each wavelength. (b) shows a scoreplot for the training data in the two primal directions. The first two CDA loadings have been used to create the scores. The scores are clustered in three seperate groups, where a Gaussian distribution is fitted to each. The contour lines of the estimated Gaussians show the boundaries for 80% (dashed) and 99% (solid) chance of a new observation to fall within this line. The boundaries are calculated by assuming Gaussianity for each class and thereby utilize the relationsship between the χ^2-distribution and Mahalanobis distances. The two vertical lines show the chosen seperation between the classes, based only on information in the first projection vector.

$$\Lambda = \frac{|W|}{|W| + |B|} = 0.04 \qquad (2)$$

W denotes the pooled within-scatter matrix and B denotes the between-scattering matrix. A value of 0.04 for Wilk's Lambda is very close to 0, which indicates a clear separation between the three groups. A corresponding set of CDFs were trained in order to find the transformation which optimally separates the three groups of spectra. The two vectors of this transformation are called loadings and are shown in in Figure 6. The figure shows that the wavelengths 525, 660, 850 and 950 nm. seem to be most important in separating the burn-degree classes. Especially noticeable is a contrast between the two NIR bands at 850 and 950 nm. and a similar contrast between the green band at 525 nm and the red band at 660 nm. The latter, however, has a lower influence, which indicates that, even though we are interested in a color classification, NIR plays an important role when differentiating between the defined classes. Figure 5(b) shows a scoreplot for the $k-1$ vectors calculated for the CDA. The clusters are seen to be fairly well separated at the 80% line, assuming each cluster follows a Gaussian distribution. The intensity in the plot shows the observation density.

If the images of the manually selected oat flakes are projected down on the principal CDF, as seen in Figure 5(b), the vertical lines shown are able to separate the clusters with up to approximately 80% certainty for each class. By only separating in a single dimension, it is possible to create a single score value for an oat flake. The overlap of the clusters comes from the continuous nature of the surface color. The positions of the two boundaries in Figure 5(b) are calculated as the mean of adjacent cluster centers.The boundaries were calculated to 0.40 and 0.81, and the scale goes between 0 and 1, where 0 is an extreme value of class C, a burned oat flake and 1 is a value in class A, a light oat flake. A score for an oat flake is now calculated as the mean of the projection values of all pixels within the oat flake. This

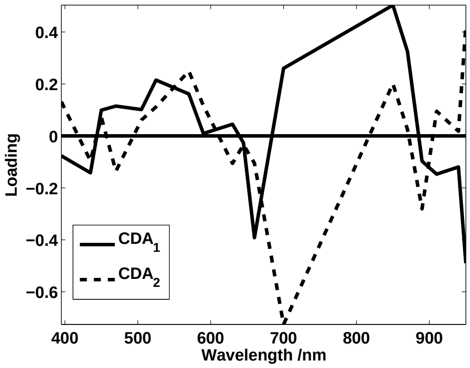


Figure 6: Loadings of canonical discriminant analysis. Especially 525, 660, 850 and 950 nm contains large burn-score information.

is shown in Equation 3, where **X** denotes the collection of pixels in the oat flake. **X** has N rows of pixels and $p = 18$ columns of wavelengths. **c** denotes the weights of the most discriminating CDF which is a single column with p rows.

$$Score = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{p}\mathbf{X_{ij}c_j} \tag{3}$$

All spectral images were converted to burn scores in order to investigate how oven time and temperature affect the surface color. A second order response surface of the form seen in Equation 4, containing first and second order as well as interaction terms, was fitted to the burn scores.

$$y = \beta_0 + \sum_{i=1}^{k}\beta_i x_i + \sum_{i=1}^{k}\beta_{ii}x_i^2 + \sum_{i<j}\sum\beta_{ij}x_i x_j + \epsilon \tag{4}$$

Before fitting the model, the independent parameters were coded such that $x_1 = (temperature - 170)/20$ and $x_2 = (time - 240)/30$ to ease the calculation. All terms in the model were significant, and the final model was found to be as in Equation 5.

$$y = 0.79 - 0.11x_1 - 0.02x_2 - 0.01x_1x_2 - 0.04x_1^2 + 0.01x_2^2 \tag{5}$$

The adjusted R^2 of the model was 0.87, thus accounting for a good amount of variability in the burn-degree. The model found suggests that, at this scale, temperature has more influence on the change in surface color than time. In fact, a change in temperature affects the surface color 5 times as much as an equal change in time. In the experiments conducted, no cereal ever reached state B before the oven was set at $160°C$. The interpretation of the model is complicated by the interaction and second order terms, but it is quite clear from the estimated model that these also have a large influence on the surface color. Figure 7 shows the contours of the fitted quadratic surface. From the gray level coded class areas, which have been superimposed on the figure, it can be seen how class B, corresponding to an intermediate burn-degree, is represented as a wide belt in the upper middle part of the surface. Class A, spanning from 0.81 to 1 is seen as a large area in the bottom, while class C is seen in the upper right corner for very dark surfaces. In a production environment an at-line measurement device should continuously verify that oat flakes are kept within the area corresponding to class B, although these boundaries may be altered for a changed output, if so desired.

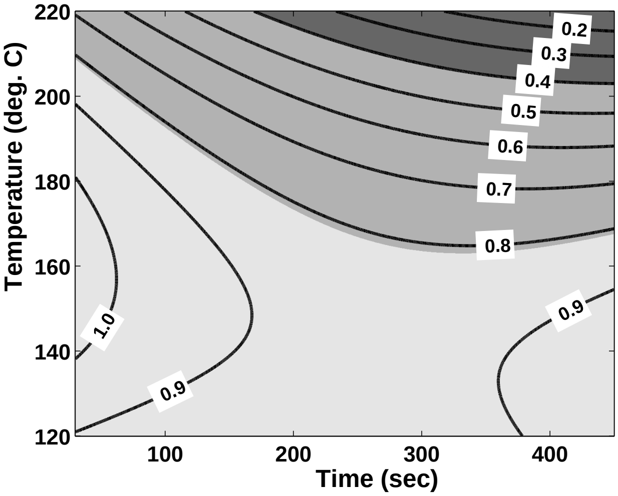


Figure 7: Contour plot of the fitted quadratic function shown in Equation 5. The shaded areas represent the three burn classes A, B and C, where C is seen in the upper right corner as the most dark area. The contour lines represent levels in the continuous burn-state function going from 1 to 0. The boundaries between the classes are at 0.4 and 0.81. It is clearly seen how temperature has a much larger effect on the burn-score, as the gradient is much higher in this direction.

3.2. Localization of Coal Particles

Industrial ovens run continuously over long periods. By chance it happens that some oat flake residue can get stuck in the oven and gets very burned. Particles originating from such oat flakes may contaminate other oat flakes, which is both a health hazard and lowers overall product quality. When inspecting spectral values of such coal particles as seen in Figure 8, a very clear difference is seen compared to the spectrum of over-burned oat flakes represented by class C in the previous subsection. The spectrum generally has much lower reflection properties, which makes it very easy to separate from normal oat flake spectra. To investigate this further, an oat flake with residue particles has been annotated in three areas; coal particle, hole and surface. Holes in the surface generally appear dark, and have spectral appearance closer to coal than normal oat flake surface as seen in Figure 9(a). If Mahalanobis distances are calculated between these three clusters, Table 2,

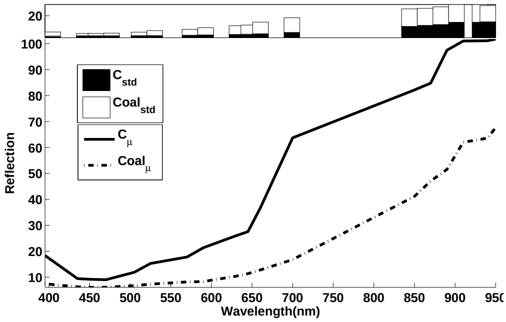


Figure 8: The solid line indicates class C, also shown in Figure 5(a). The dashed line shows the spectral value coal or severely burned particles. The large difference in spectral value indicates that differentiating between coal and even overburned oat flake (class C) is straight forward.

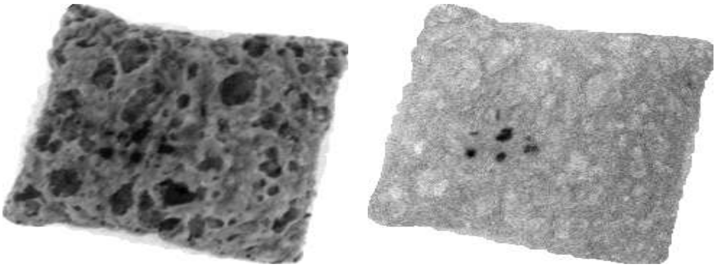
	Surface	Coal	Hole
Surface	0	417.83	58.97
Coal		0	268.55
Hole			0

Table 2: Mahalanobis distances between hole, surface and coal areas in a multispectral image. Coal particles are seen to be located far away from surface and hole spectra.

in spectral space, it is quite clear that coal particles are easily separable. A Multivariate Analysis of Variance (MANOVA) furthermore shows a significant difference between the groups, meaning that it is possible to separate coal from holes with high accuracy. By utilizing the transformation created in the previous subsection to quantify surface color as a score value, it is likewise possible to calculate a score value for coal particles. By transforming the oat flake in Figure 9(a) and visualizing the first component, it is seen how coal particles are clearly highlighted in Figure 9(b).

3.3. Determination of Water Content

The distribution of RMSE of the water prediction model, calculated in the bootstrapping procedure based on the complexity parameter $\lambda = 0.0014$, is seen in Figure 10. An overview of statistics calculated is seen in Table 3. From Table 3 it is seen that the ridge model on average is able to predict the water content with a standard deviation of around 1% in terms of water



(a) (b)

Figure 9: (a) shows a visualization of a spectral band in the blue area in a multispectral image of an oat flake. On the surface, small pieces of coal which look similar to surface holes are seen. The size of the oat flake is approximately 1.5cm. (b) shows the same oat flake projected using the CDF found for the burn-score quantification. A clear highlighting of coal particles appears.

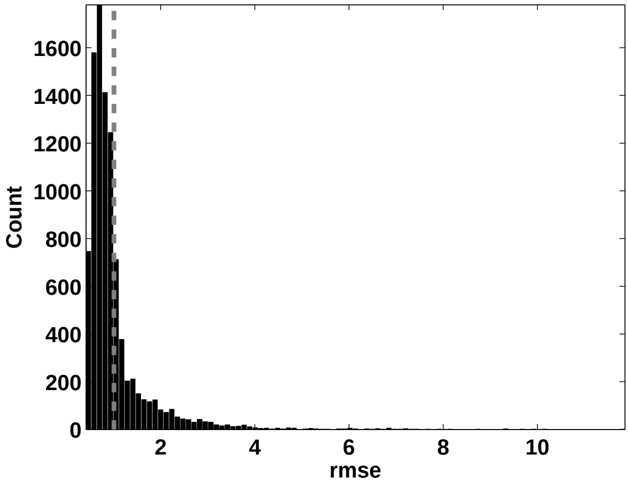


Figure 10: Distribution of RMSE bootstrap residuals. The mean value is indicated with a dashed line

percentage. This however varies as is also seen in Figure 10, where the worst case scenarios with an RMSE value of up to 10 are seen, although very rarely.

Metric	**Expression**	μ_{BS}	σ_{BS}
Bias Factor (B_f)	$10^{\sum \frac{\log(\widehat{y}/y)}{n}}$	1.0016	0.0210
Accuracy Factor (A_f)	$10^{\sum \frac{\lvert\log(\widehat{y}/y)\rvert}{n}}$	1.0261	0.0202
Mean Relative Percentage Error	$\frac{1}{n}\sum \frac{100\cdot(y-\widehat{y})}{y}$	-0.2901	2.3194
Mean Absolute Percentage Error	$\frac{1}{n}\sum \frac{100\cdot\lvert y-\widehat{y}\rvert}{y}$	2.6144	2.1681
Root Mean Square Error of Prediction	$\sqrt{\frac{1}{n}\sum(y-\widehat{y})^2}$	1.0152	0.7121
Standard Error of Prediction	$\frac{100}{\bar{y}}\sqrt{\frac{1}{n}\sum(y-\widehat{y})^2}$	3.8278	2.6849

Table 3: Summation statistics for the water content prediction. All values were estimated using bootstrapping. μ_{BS} is the mean value of the bootstrap residuals and σ_{BS} is the standard deviation value of the bootstrap residuals.

The standard deviation of the RMSE is 0.71. These cases are due to the very small data amount and “unlucky” sampling in the bootstrap routine. In these cases the training set has very low rank because of many sample replications.

As seen in Table 3, B_f is very close to 1, and has a standard deviation of 0.021, showing an overall low bias in the fitted models. The B_f shows a slight overprediction tendency, which is also seen in the MRPE showing an overprediction of 0.29%. In addition, the accuracy factor indicates that the predicted water content is 2.61% different from the measured values, with a deviation of approximately 2%. The values in the MAPE of 2.61 verify the accuracy factor. The value of SEP was less than 4% and in total the bootstrapped model statistics indicate a good general performance. Figure 11 shows the loadings of the ridge regression, which clearly shows the 970 nm wavelength to be of most importance in the discrimination between oat flakes with different water content. Since this is an absorption peak and the camera device has captured reflectance images, this means more wet oat flakes will have lower numerical values than more dry oat flakes at the 970 nm. band. This also causes the 970 nm loading to be negative.

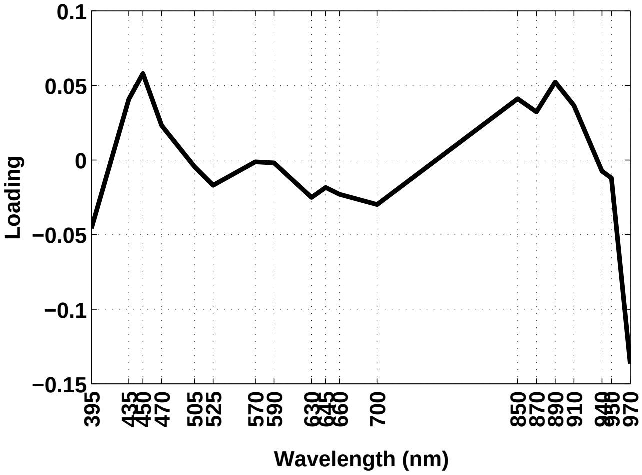


Figure 11: Loadings of the ridge regression with penalization parameter λ selected as 0.0014 by cross validation. The most discriminating wavelengths are found to be in 970 nm. and in the areas around 890 and 450 nm.

4. Conclusion

In this paper, we have shown that multispectral imaging is a non-invasive and rapid alternative for quantification of certain quality parameters for oat flake. Quality parameters highlighted were the degree of burn, ability to detect coal particles on the surface, and water content of the oat flakes. To quantify the degree of burn when the oat flakes are baked in the oven, a burn score was created as a numerical index. A CDA was used to transform multispectral images to surface color quantification, using three oat flakes with different surface color, manually picked by experts. A MANOVA confirmed a significant difference in the surface color of these three oat flakes with a Wilks Lambda of 0.04. Loadings of the CDA showed that the wavelengths 525, 660, 850 and 950 nm seemed to be responsible for the best seperation. To understand the dependency between baking time, baking temperature and surface color, a quadratic surface was fitted to experimental data with an adjusted R^2 of 0.87. This model suggested that temperature has more influence on the change in surface color than time. A response surface demonstrated the area of adequate baking conditions and an at-line measurement device continuously ensured that new samples were processed within this area.

Detection of severely burned residual oat flake particles was shown to be an extension of the burn-degree transformation by considering these particles as extreme observations. A MANOVA furthermore showed a significant difference between holes in the surface and black particles.

The ability for the multispectral camera to differentiate between oat flakes with different water content was tested with a dataset of 11 samples of varying water content, ranging from approximately 14 to 37% water. A Ridge regression was tuned using cross validation and based on a logarithmic transformation of the response variable and mean spectra of the multispectral images. The performance of the model was validated using a bootstrapping method. Summation statistics showed a good fit, with B_f very close to 1 and RMSE of 1.0152. Loadings of the regression model indicated that 970 nm had a large influence on the model corresponding well with known properties of water absorption.

In total, satisfactory results could be achieved for all measurements for identifying surface burn degree, detecting coal particles as well as water content, thus pointing to the fact that multispectral imaging is feasible for objective quantification of certain parameters in an oat flake production environment. Since the techniques used are in essence general techniques, they are also likely to be applicable to other food matrices.

References

[1] H. Huang, H. Yu, H. Xu, Y. Ying, Journal of Food Engineering 87 (2008) 303 – 313.

[2] T. Brosnan, D.-W. Sun, Journal of Food Engineering 61 (2004) 3–16.

[3] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, J. Frias, Trends in Food Science & Technology 18 (2007) 590 – 598.

[4] J. M. Carstensen, M. E. Hansen, N. K. Lassen, P. W. Hansen, in: B. Ersbøll, T. M. Jørgensen (Eds.), 9th Medical Image Computing and Computer Assisted Intervention (MICCAI) - Workshop on Biophotonics Imaging for Diagnostics and Treatment, Institute of Mathematical Modelling - Technical Report-2006-17, 2006.

[5] N. Perrot, I. Ioannou, M. Gilles, I. Allais, C. Curt, V. Chevallereau, G. Trystram, Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (2002) 94–99.

[6] T. A. Haley, S. J. Mulvaney, Trends in Food Science and Tecnology 6 (1995) 103–110.

[7] S. Linko, P. Linko, Transactions of the Institute of Chemical Engineering 76, Part C (1998) 127–137.

[8] S. Majumdar, D. S. Jayas, Journal of Agricultural Engineering Research 73 (1999) 35 – 47.

[9] J. F. Hansen, On chromatic and geometrical calibration, Ph.D. thesis, Department of Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 1999.

[10] B. S. Dissing, L. H. Clemmesen, H. Løje, B. K. Ersbøll, J. Adler-Nissen, In Proceedings of Institute of Electrical and Electronics Engineers (IEEE) Color and Reflectance in Imaging and Computer Vision Workshop, Kyoto, Japan. (2009).

[11] L. H. Clemmensen, M. E. Hansen, B. K. Ersbll, Machine Vision and Applications (2009).

[12] D. D. Gomez, L. H. Clemmensen, B. K. Ersbøll, J. M. Carstensen, Computer Vision and Image Understanding 106 (2007) 183–193.

[13] M. E. Hansen, B. K. Ersbøll, J. M. Carstensen, A. A. Nielsen, in: Lecture Notes in Computer Science, LNCS3540, Lecture Notes in Computer Science, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005, pp. 1228–1237.

[14] N. D. Frame (Ed.), The technology of extrusion cooking, Aspen Publishers, Inc., 1999.

[15] S. Singh, S. Gamlath, L. Wakeling, International Journal of Food Science & Technology 42 (2007) 916–929.

[16] T. Hastie, R. Tibshirani, J. Friedman, Elements of Statistical Learning: Data mining, Inference and Prediction, 2. edition, Springer-Verlag, 2009.

[17] G. M. Hale, M. R. Querry, Appl. Opt. 12 (1973) 555–563.

[18] P. Jonathan, W. J. Krzanowski, W. V. McCarthy, Statistics and Computing 10 (2000) 209–229.

[19] R. Wehrens, H. Putter, L. M. C. Buydens, Chemometrics and Intelligent Laboratory Systems 54 (2000) 35 – 52.

[20] T. Ross, Journal of Applied Microbiology 81 (1996) 501–508.

Appendix B

Quantifying airbubbles in breadcrumbs

Technical report IMM-Technical Report-2011-06

Abstract

Spatial inference of multispectral images are studied in this technical report with a focus extracting features from images of bread crumbs. The features extracted describe the distribution of air bubbles in the crumb as well as information about the solid material in the crumb. An experiment where 31 loaves of bread were prepared has been carried out, with a total of 63 multispectral images. From each image air bubbles were segmented and four different features were extracted based on the segmentation. The features are considered in collective plots as a function of baking time in the oven. It is found that after six minutes in the oven the features agree to reach a steady state. However, there are some fluctuations seen in the timeline. Furthermore, a small review of previous work is presented.

Introduction

A very important quality parameter for solid foods such as bread is the crispness and the physical texture. These factors are typical consumer purchasing decision factors [1]. Bread crust crispness is lost when water migrates from crumb (Figure B.1) to crust during storage or when the bread is kept in humid environments. Thus the crispness of bread is known to be very dependent on the bread crumb and its water transportation capabilities. Furthermore the crumb structure contributes to the texture mechanical strength and perceived product freshness of the bread as well as to its visual appearance. When scoring the bread crumb, the visual texture accounts for approximately 20% of the weighting used in judging bread quality [2].
Beside the crumbs relationship to the mechanical and visual properties, the bread crumb is furthermore important due to its inherent relation to a set of parameters. These are the volume of the bread, the symmetry of the bread and the texture and taste during eating [3].
By knowing which parameters defines the crumb properties and combining this knowledge with the knowledge of how raw materials affect the breadcrumb it is possible to predict many of the quality attributes of the above mentioned. which is why study of the breadcrumb is very important.

The formation of the breadcrumb is dependent on the air bubbles in the crumb and the steps in which these air bubbles or cells are formed is seen in Figure B.2 (inspired by a similar figure in [4]). The bread making process is divided into three main stages: mixing, fermentation/proving and baking. It has been showed that the air entrained during the early stage of mixing is the source of the cells. There are discussions about how much cells are expanding in this

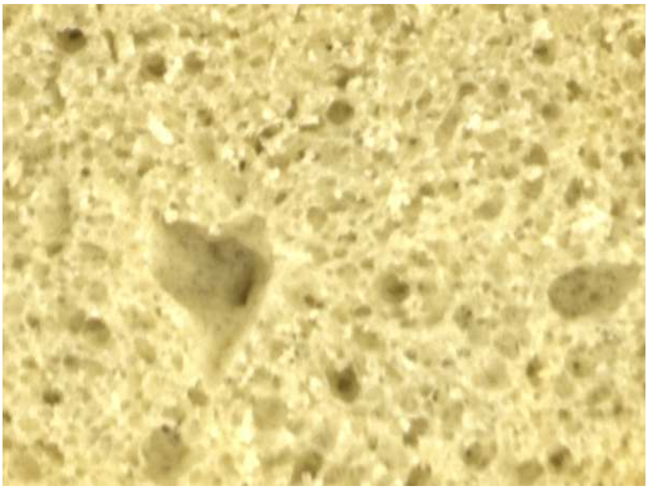


Figure B.1: Pseudo RGB image of bread crumb from a piece of bread

phase, but since there is a continuous production of CO_2 from the yeast, with a net transport of gas from the surrounding dough into the cells, the hypothesis is that there is cell expansion from very early stages of the bread making. Fermentation and baking does not form new cells, this only expands existing cells [5]. From left side in Figure B.2, when the dough is mixed, kneaded and is proving, small air bubbles are trapped inside the dough. When put in the oven (step 2) H_2O vaporizes and together with yeast which is producing CO_2 under heat, makes the existing cells expand. This creates a fragile foam. Cells keep expanding until the foam reaches a temperature of around 60°C. At this point, the gluten-starch contained within the thin walls between the expanded cells will gelatinize. This means the cells will loose their elasticity and maintain their size and structure. This prevents further cell growth. The CO_2 contained in the bubbles needs to escape and will cause the cells to rupture since they can no longer expand. The ruptured cells has thus transformed to large networks as seen in step 3, with a sponge like appearance. The remaining crumb cell walls consist of partly gelatinized starch.

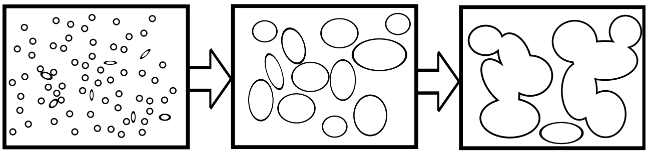


Figure B.2: Formation of air bubbles/cells in the bread crumb.

By controlling the numbers, sizes and uniformity of distribution of the crumb cells it is possible to largely control the above mentioned bread quality parameters. This may be done as early as in the mixing stage [6]. In general, the gas bubbles have a relatively small size; for example, at the end of mixing their mean diameter is about 75 μm, but they expand to reach a mean diameter as large as a few millimeters [7]. A smaller pore size is expected to slow down moisture migration [8], which will increase the shelflife of the bread. Furthermore, for fine-celled white bread, it is generally accepted by consumers that small holes that are uniformly distributed throughout the crumb are required and that large holes or irregular cell distributions are undesirable [6]. For a good product, the crumb cell wall, or matrix, is required to be as thin as possible [2]. However it must be resilient enough to recover from modest deformation, such as squeezing and pressing which is the two most common ways by which consumers assess product freshness [9]. Thus, the examined parameters are crumb fineness (open versus closed cells), uniformity, cell shape, and cell wall thickness (CWT).
The traditional method for texture scoring or inspection is qualitative and subjective in nature since it relies on human vision. This is known to be inconsistent among different experts [10] and can vary over a period of time even for the same expert. This problem necessitates the use of more objective quantification methods that are fast, precise, consistent and reliable such as digital image analysis or similar. Litterature shows various attempts to study the cellular structure of bread crumb using different methods such as [11, 12] who used Haar transform [28] and textual features to describe the surface of the breadcrumb. [14] compared k-means clustering and a set of adaptive thresholding methods to assess their differences. [15] used 2D (flatbed scanner) and 3D (X-ray tomography) images of breadcrumbs together with granulometric methods of six bread crumbs from three different recipes and three baking procedures. The conclusion in this paper was that 2D images could be used to quantify the cellular structure for the calculation of mechanical properties. The use X-ray tomography to assess 3D models of breadcrumbs have also been investigated in [16, 17, 18] which all shows that this is also a promising technique for investigation of the breadcrumb, however more for scientific purposes while the use of digital images also may more easily be implemented in a plant.

Materials and Methods

Data collection

31 loaves of bread were prepared using the same recipe. The recipe consisted of the following ingredients

1. 1000 grams of wheat flour
2. 10 grams of sugar
3. 15 grams of salt
4. 30 grams of dry yeast
5. 570 grams of water

The ingredients were mixed together in a mixer for three minutes. After mixing the dough is allowed to prove and ferment for 1 hour. After fermentation, bread of 100 g. are formed and are again allowed to prove for further 35 minutes at approximately 35°C at high humidity. As the only variable parameter the loaves are then kept in the oven for different amounts of time. The oven temperature is maintained at 175°C. Before slicing and image acquisition the the bread is cooled for 2 hours.

The baking times of the loaves ranges between 2 and 26 minutes. Other parameters regarding dough composition, weight, pan dimensions, rising time, oven temperature, ventilation degree and cooling time before slicing, have been kept equal for all loaves. The baking was done in three batches, named A (10 loaves), B (10 loaves) and C (11 loaves). Slicing and image acquisition procedure were as follows. After cooling, the bread was divided at half length. One half was used and sliced again 1.5 cm from the "center" end of the half bread to form a 1.5 cm wide slice. After slicing multispectral images were acquired of both sides of all slices using the VideometerLab.

Computational techniques

Bread as a solid is "soft" and like many other foodstuffs is comprised at a macroscopic level, of two phases- a fluid (air) and a solid (cell wall material).

Quantification of such two phase systems generalises to a generic area in digital image analysis. This area focuses on quantification and assessment of particles, and depending on the type of particles there are various ways of dealing with such problems. Well studied examples of such systems are those of assessing cells in various biological systems [19, 20, 31].

Mainly there are three ways of quanitifying such phase systems in digital images. These are; morphological/granulometric methods [28], textural methods [28] or by using segmentation [28]. Regarding image analysis on the breadcrumb thorough investigation has been carried out since the early 1990's. Different image analysis techniques and statistical methods has been tested for crumb quantification [3]. granulometric methods for crumb analysis was carried out by [15], while textural methods has been in vestigated in [11, 12]. Actual spatial segmentation of cells, meaning splitting the breadcrumb into each of the two phases; air and cell-wall was demonstrated in [18, 22]. The advantage of utilizing spatial segmentation for feature extraction is that, in addition to it being able to provide an estimation of crumb fineness, it can also accurately measure various structural parameters of the bread crumb such as cell size, number of cells per unit area, CWT, cell shape etc. and provide distributions of these for each crumb. Basically, using segmentation techniques gives the possiblity to assess each cell, while the granulometric and textural approaches only measures a macroscopic score. The downside however is that by using segmentation techniques there is a possibility of over or under segmenting either of the two bread phases.

Spatial segmentation may become better if guided by multiple variabels which provide the segmentation algorithm with more information e.g. several variables of wavelengths. However, after thorough investigation it has been concluded here that combination of multiple spectral measurements with VideometerLab does not increase conrast and highlight neither the air nor the cell wall phase more than single channels. This was tested by visual inspection and by use of Principal Component Analysis [27], Minimum Noise Fraction [35], Canonical Discriminant Analysis [27] and non-linear versions of these. It was found that the best segmentation results were based on a single channel, blue (395 nm.), from the multispectral image (Figure B.4(a)).

A single channel may be considered as a two-dimensional discrete functions $f(x, y)$. Mathematical Operations on two-dimensional discrete functions representing images are in abundance [28]. Often relatively simple procedures are able to provide good results in a robust manner. To enhance the contrast of the blue channel an adaptive histogram matching technique was used [25]. Histogram matching is a technique which enhances contrast in an image by remapping the brightness values such that the histogram of the gray level values in the image are mapped to a model histogram of choice. If such an equalization is performed on subsets of the image i.e. sub tiles of the image, the transformation works locally which in many cases enhances the contrast further than a global histogram matching. However there is a tradeoff, since to small image subtiles will yield no contrast enhancement. Thus a proper tile size needs to be selected. In this study a tilesize of 8 was used.

Finding the outline of the crust is simply done using a threshold, while the crust is removed using erotion with disk shaped structuring element, removing

30 pixels from the edge.
For the actual cell marking an H-dome [31] technique was used. H-domes is a morphological operation which makes heavy use of a technique called grayscale reconstruction [31] to find regional maxima. Regional maxima in its strict sense, not to be confused with local maxima is defined as a set of connected pixels M (either 4 or 8 connection [31]) which has a certain value k and no pixels in its neighborhood with higher gray level values than k. However, in practical applications (digital images) where it is desirable to find patches of similar gray level values, a more loose definition is required. Thus, the H-dome technique allows a regional maximum with up to k+H gray levels to be a regional maximum instead. The principle of the H-domes algorithm allowing H values difference in the regional maxima is shown in Figure B.3.

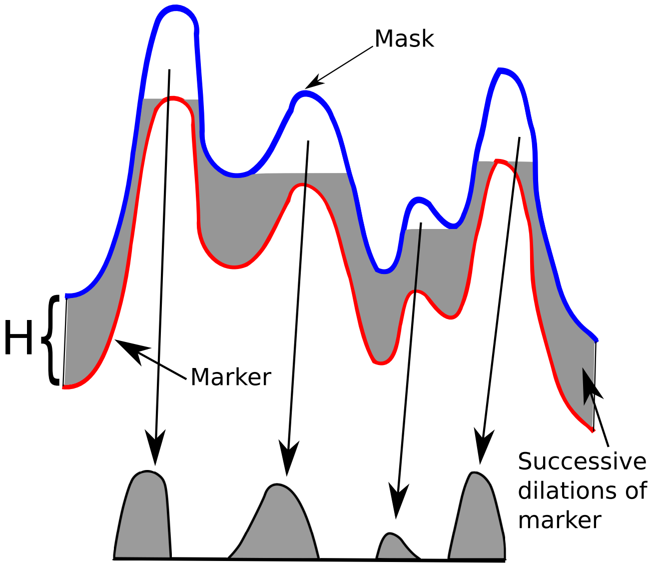


Figure B.3: 1 dimensional conceptual drawing of the H-domes technique. The mask signal (blue), which represents the image function $f(x, y)$ is used to create a marker signal (red) by substracting a value H. Using morphological reconstruction, which is a series of successive morphological dilations, on the mask and marker image, regional maxima are found.

The idea is that two images are used, a mask and a marker image. The mask image, which is the original gray level image is used to create the marker image by substracting H. A series of gray level dilations are performed on the marker image until new dilations does not change the marker level. The result is an image where similar connected pixels now has the exact same value i.e. regional maxima. By substracting this reconstructed image from the original image, the

H-domes are kept. If the entire procedure is flipped to find regional minima, we are able to find H-basins instead. An example of H-basins found on the contrast enhanced blue channel image using the inverted H-domes technique is seen in Figure B.4(d).

Using simple boolean arithmetics, the binary H-dome mask can be inverted to a mask representing the cell-walls. To quantify CWT, [22] propsed to locate cell centroids and draw vectors between neighboring centroids. The width of the corresponding wall was then calculated as the euclidean distance between the two intersection points of the vector and the wall. In this work an alternative method, making use of an euclidean distance map (EDM) [28] and a medial axis transformation (MAT) [28]. The MAT is used to calculate a skeleton representation of the cell-wall structure which locates the median of all cell-walls in the image (Figure B.4(e)). The actual CWT is looked up in the EDM (Figure B.4(f)) using the skeleton indexes, which means that the CWT is actually calculated as the distance from the middle of the wall to the nearest cell and multiplied with 2.

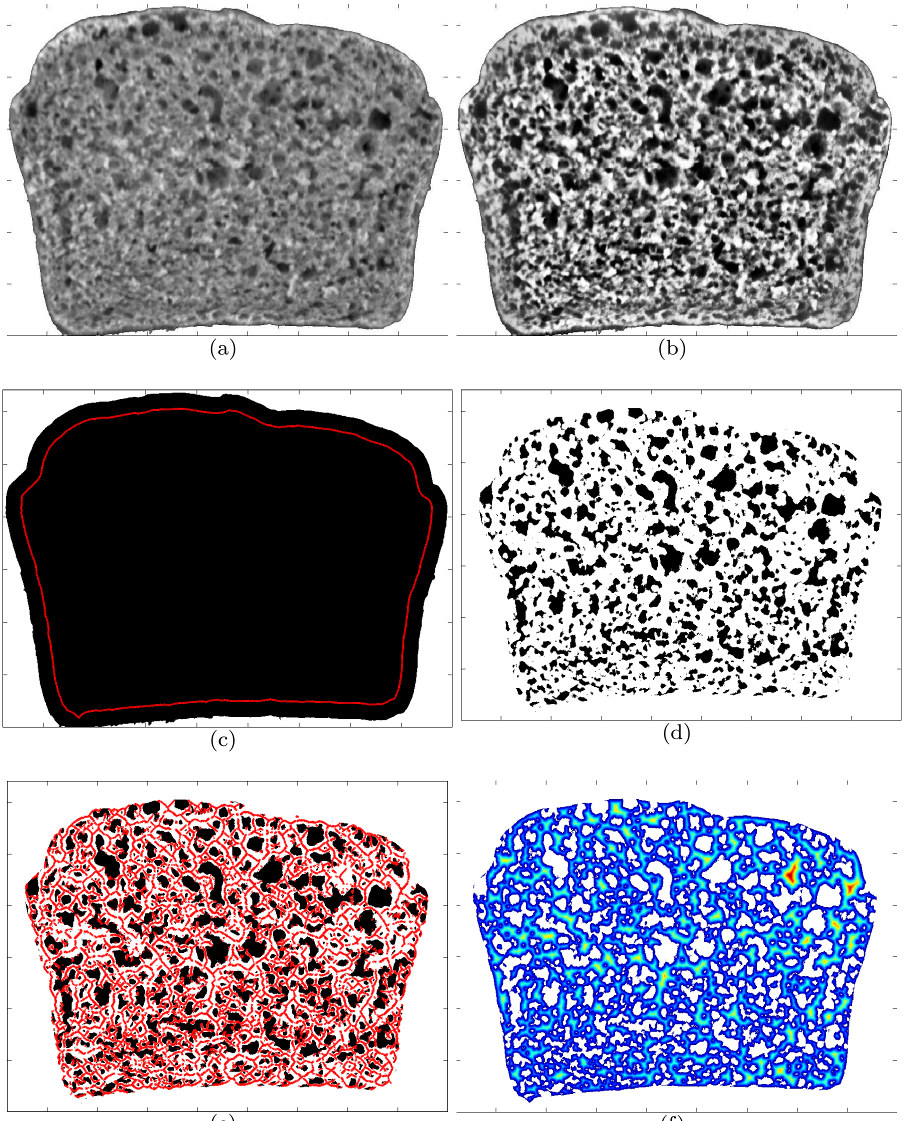
From the extracted masks of cell-walls and cells a set of features are extracted:

1. Total number of cells per area
2. Mean gray level value of cells (indicating the depth)
3. Mean Area
4. Mean Wall Thickness

Results and Discussion

Cell segmentation of a piece of bread which has been in the oven for 21 minutes (Figure B.5(a)) is shown in Figure B.5(b). The actual decision of whether this segmentation is good is very subjective. However, it seems that when assessed visually by comparing with Figure B.5(a) that the segmentation process has segmented the cells very well. According to the segmented results the piece of bread consist of a large number of very small cells; 88% of the detected cells have an area smaller than 1.5 mm^2. However this still makes up a total of 16.8 cm^2 out of an entire piece of bread of approximately 26 cm^2. The largest cell is found to be about 20 mm^2 while the mean cell size is found to be 0.65 mm^2. For CWT the values appear to be more gaussian like distributed with a mean value of 0.84 mm. The distribution of cell area and CWT are shown in Figure B.5(c) and Figure B.5(d), and the cell area distribution is as expected skewed.

Features has been extracted from all images and are shown in Figure B.6. The samples were as explained in the materials section prepared in three batches and

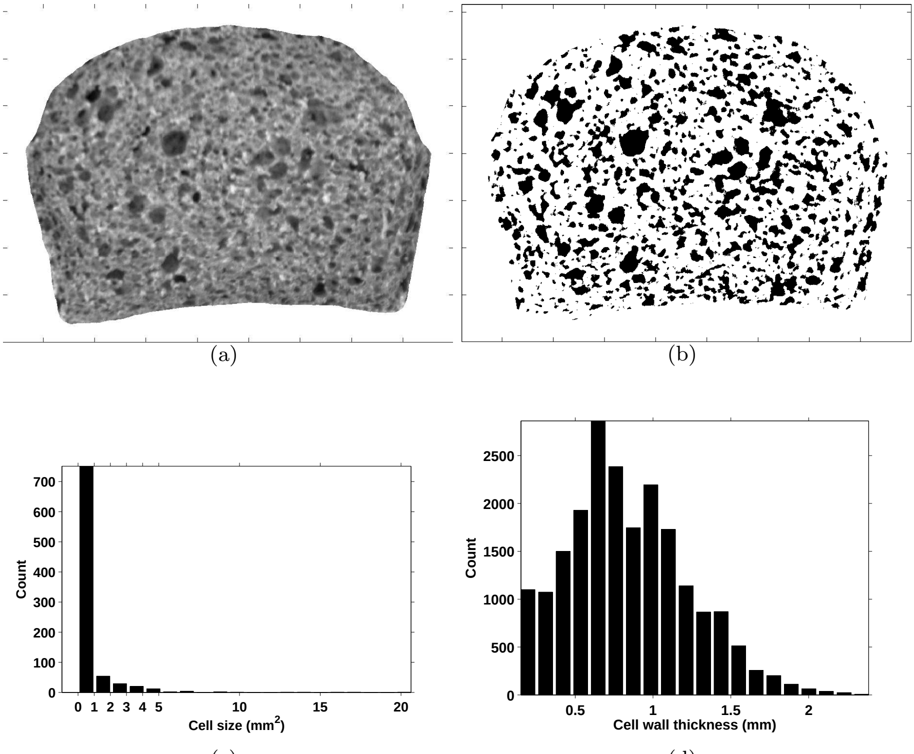


(a) (b)

(c) (d)

(e) (f)

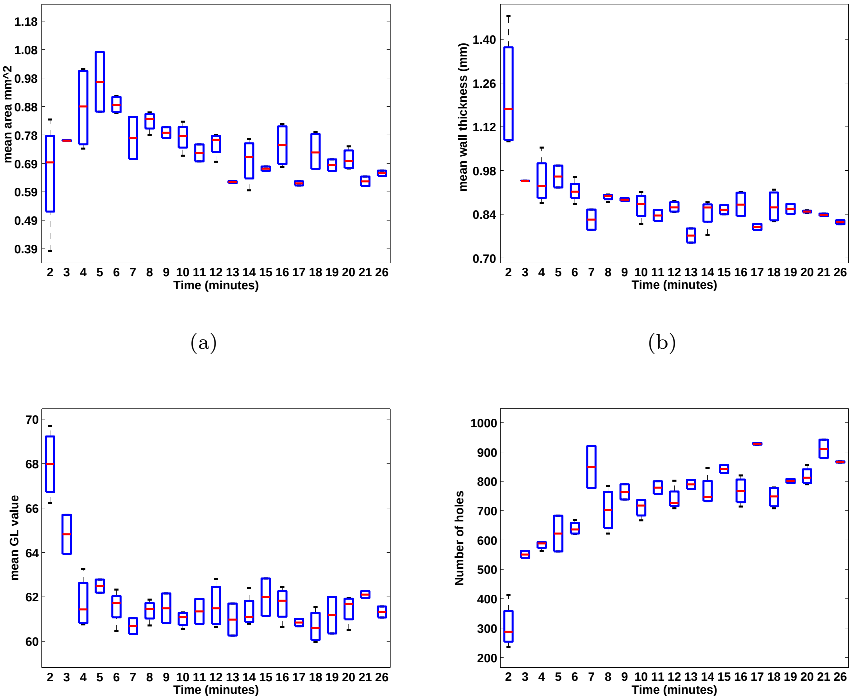
Figure B.4: The steps of how the cells in the breadcrumb are localized and segmented. All 6 images are from the same piece of bread (time in oven: 11 minutes). (a) shows the blue channel from the multispectral image (395 nm.), here the cells are fairly visible. (b) shows the the result of (b) after processing with an Adaptive Histogram Equalization algorithm. Transitions from cell to wall is now better highlighted. (c) A simple threshold is enough to locate the outline of the crust. An erosion ensures the crust to be removed. (d) shows the result of treatment of the contrast enhanced crumb image. To segment the cells an H-dome algorithm was applied together with an adaptive threshold. (e) outlines the "skeleton" of the walls. This is used together with the eucldian distance map shown in (f) to calculate the mean CWT. The skeleton was created using the Medial Axis Transformation.



(a) (b)

(c) (d)

Figure B.5: Assessment of a single piece of bread is shown here for a bread which has been in the oven for 21 minutes. The total size of the bread is 58x45 mm. (a) shows a raw image at 395 nm. of the bread. (b) shows the segmented cells where holes are represented with black. (c) and (d) shows the distribution of the area of all cells and the distribution of CWT.



(a) (b)

(c) (d)

Figure B.6: Results of extracted features across all images. Observations are plotted as a boxplot where for odd numbered minutes there are 2 observations and for even numbered minutes there are 4 observation (See batch A, B and C in the materials section). All plots are plotted as feature as a function of time. (a) shows the mean area size for a breadcrumb. (b) shows the mean CWT. (c) shows the mean grayvalue of the blue channel i.e. an approximation to mean depth of a cell in the crumb. (d) shows the total cellcount in a crumb.

the bread were sampled every second minute. Two batches were sampled at even numbers while one batch was sampled at odd numbers. This means that on the plots shown in Figure B.6 there are two samples (on for each side of the bread slice) for odd numbers and four for even numbers. The extracted features are shown as boxplots which shows the mean value of the repetitions as a red bar. As expected the tendency is that the mean CWT decreases as the number of holes increases. There is a tendency for the area of the holes to increase during the first 5 or 6 minutes while it then starts to decrease. The GL values decreases rapidly within the first minutes and then goes to a steady state. Intuitively one would expect the mean GL value feature to have the same behaviour as the mean area feature since these both describe area although in different dimensions (GL value represent the depth of the cells). The mean area feature should however be regarded as more correct since there are more certainty of this area estimate than the GL value area estimate. A common trend for all calculated features is that they seem to reach some steady state after approximately 6 minutes in the oven. This is of course dependent on the ingredients used in the dough etc. however for the given samples the main change in air bubbles appear to happen in the beginning of the baking time where the dough is still spongy and the cell walls still have not gelatinized.

Conclusion

The ability of a multispectral imaging system such as VideometerLab to quantify the cell distribution in bread baked over different time spans was investigated. Similar studies have been performed before, however to the knowledge of the author no published work presents the use of digital images for assessing the development of cells in the breadcrumb over time. Most studies focuses on looking at bread prepared with different mixing times, recipes etc. but baking time and digital image analysis has not been combined in published work. Computed parameters describing the crumb development include total number of cells per area,cells depth, mean Area and mean wall thickness. Morphological methods together with edge enhancing methods were used to extract the cells, which is able to scale very well, meaning it will be able to assess the cell population in different breads. Basically the methodology should in theory apply equally well to virtually any type of baked good product where internal appearance characteristics such as “holes” help to define product quality.
A system such as this could be used in a laboratory to assess the effects of new or existing ingredients or processes on the crumb. Another is direct quality control at the process line (obviously destructive sampling). The studies carried out in this work has mainly been proof of concept and explorative work. However it shows that after approximately 6 minutes in the oven, the structures

in the breadcrumb starts to stabilize. The ability for the multispectral device to assess spatial structures such as cells in the breadcrumb furthermore shows the ability for spatial inference while other experiments shows the ability for spectral inference.

Acknowledgements

The author would like to express his gratitude to Mette Stenby for preparing the bread and insightfull discussions and to Nicolas Otso who in a small period helped baking and doing initial dataanalysis as part of a course [02507] January 2011.

References

[1] hannemieke luyten, johan j. plijter, and ton van vliet. crispy/crunchy crusts of cellular solid foods: a literature review with discussion. Journal of Texture Studies, 35(5):445–492, 2004.

[2] E.J. Pyler. Baking Science & Technology Volume 2. Sosland publishing co., 4 edition, 2009.

[3] M. G. Scanlon and M. C. Zghal. Bread properties and crumb structure. Food Research International, 34(10):841 – 864, 2001.

[4] E.N.C. Mills, P.J. Wilde, L.J. Salt, and P. Skeggs. Bubble formation and stabilization in bread dough. Food and Bioproducts Processing, 81(3):189 – 193, 2003. Cereal Processing.

[5] J. C. Baker and M. D. Mize. The origin of the gas cell in bread dough. Cereal Chemistry, 18:19–33, 1941.

[6] Stanley P. Cauvain and Linda S. Young. Technology of breadmaking. Springer Science+Business Media, 2007.

[7] y. shimiya and k. nakamura. changes in size of gas cells in dough and bread during breadmaking and calculation of critical size of gas cells that expand. Journal of Texture Studies, 28(3):273–288, 1997.

[8] T. P. Labuza and C. R. Hyman. Moisture migration and control in multi-domain foods. Trends in Food Science & Technology, 9(2):47 – 55, 1998.

[9] Ronald A. Hebeda and Henry F. Zobel, editors. Baked Goods Freshness. Marcel Dekker, 1996.

[10] Jian Wang and Graeme D. Coles. Objective measurement of bread crumb texture. In George E. Meyer and James A. DeShazer, editors, SPIE Proceedings, Optics in Agriculture, Forestry and Biological Processing, volume 2345, pages 85–94. SPIE, 1995.

[11] D. Bertrand, C. Le Guerneve, D. Marion, M. F. Devaux, and P. Robert. Description of the textural appearance of bread crumb by video image analysis. Cereal Chemistry, 69:257–261, 1992.

[12] Inna Y. Zayas, James L. Steele, G. Weaver, and D. E. Walker. Breadmaking factors assessed by digital imaging. volume 2064, pages 135–151. SPIE, 1993.

[13] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. Digital Image Processing. 2 edition, 2002.

[14] Ursula Gonzales-Barron and Francis Butler. A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumb features by digital image analysis. Journal of Food Engineering, 74(2):268 – 278, 2006.

[15] Nejla Lassoued, Perrine Babin, Guy Della Valle, Marie-Françoise Devaux, and Anne-Laure Réguerre. Granulometry of bread crumb grain: Contributions of 2d and 3d image analysis at different scale. Food Research International, 40(8):1087 – 1097, 2007.

[16] P. M. Falcone, A. Baiano, F. Zanini, L. Mancini, G. Tromba, F. Montanari, and M. A. Del Nobile. A novel approach to the study of bread porous structure: Phase-contrast x-ray microtomography. Journal of Food Science, 69(1):FEP38–FEP43, 2004.

[17] C. Primo-Martín, G. van Dalen, M.B.J. Meinders, A. Don, R.H. Hamer, and T. van Vliet. Bread crispness and morphology can be controlled by proving conditions. Food Research International, 43(1):207 – 217, 2010.

[18] Gerard van Dalen, Peter Nootenboom, Lucas J Van Vliet, Lennard Voortman, and Erik Esveld. 3-d imaging, analysis and modelling of porous cereal products using x-ray microtomography. Image Analysis and Stereology, 26(3):169–177, 2007.

[19] Meng Wang, Xiaobo Zhou, Fuhai Li, Jeremy Huckins, Randall W. King, and Stephen T.C. Wong. Novel cell segmentation and online svm for cell cycle phase identification in automated microscopy. Bioinformatics, 24(1):94–101, 2008.

[20] Vinh-Thong Ta, Olivier Lézoray, Abderrahim Elmoataz, and Sophie Schüpp. Graph-based tools for microscopic cellular image segmentation. Pattern Recogn., 42:1113–1125, June 2009.

[21] Luc Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Transactions on Image Processing, 2(2):176–201, 1993.

[22] H. D. Sapirstein, R. Roller, and W. Bushuk. Instrumental measurement of bread crumb grain by digital image analysis. cereal chemistry, 71:383–391, 1994.

[23] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Elements of Statistical Learning: Data mining, Inference and Prediction, 2. edition. Springer-Verlag, 2009.

[24] Andrew A. Green, Mark Berman, Paul Switzer, and Maurice D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and remote Sensing, 26:65–74, 1988.

[25] Karel Zuiderveld. Contrast limited adaptive histogram equalization, pages 474–485. Academic Press Professional, Inc., San Diego, CA, USA, 1994.

Appendix C

Butter cookie quality estimation

Technical report IMM-Technical Report-2011-07

Abstract

Methods for studying and visualizing the browning process of butter cookies using mutispectral images is presented. A sensory panel consisting of six people visually evaluates a set of butter cookies and divides them into three groups. An acceptance score is developed based on the panel evaluations as well as multispectral images, which is used to visualize how the browning develops over time and temperature in a quadratic response surface while also indicating in which areas the surface colour is acceptable. Furthermore the acceptance score is used to visualize spatially of the browning progresses within a cookie. A water prediction model is similarly created to track changes in water activity. The water prediction model is found to correctly estimate the actual water content with an error margin of 13.6%. The propagation was visualized similar to the acceptance score with intuitive results.

Introduction

In an industrial setting, when preparing baking products such as cookies in the oven, surface colour is a critical index for the baking state. Together with texture and aroma, the surface colour is the first impression a consumer gets of a baking product [2]. This puts a considerable importance on the surface colour of cookies as a quality parameter. The surface colour of cookies depends both on the physicochemical characteristics of the raw dough such as water content, pH, reducing sugars and amino acid content and the operating conditions applied during baking; temperature, time, air speed and relative humidity. Thus there are many parameters which may be tuned to obtain a desired surface colour.
In the baking production line it is of high importance to keep a constant control of the surface colour to ensure a homogeneous appearance of the baking product within an acceptable region. Naturally this is done by continuously inspecting the product for the surface colour and texture and adjusting the above mentioned parameters to in order for the suface colour not to get outside this accept region. Traditionally such control and quality checking has been done by human expert operators. In spite of having experts inspecting the production line, which has many qualities, there are also heavy disadvanteges. Besides difference in the vision of the experts, they are in nature subjective and inconsistent and can vary over periods of time even for the same expert [3].

Maillard reactions [53] is one type of non-enzymatic browning corresponding to a set of reactions occurring between amines and carbonyl compounds, especially reducing sugars such as glucose. The Maillard reactions are known to occur in heated, dried, or stored foods and thus play an important role in the browning of cookies. Even though the browning of the surface colour is mainly controlled by Maillard reactions, the extent of Maillard reactions are controlled by the physicochemical characteristics and operating conditions mentioned earlier. However, the chemistry underlying the Maillard reaction is very complex and is notoriously difficult to control [53]. Beside adding pleasant aroma, good taste and good looking brown appearance, a product of Maillard reactions is the formation of toxic mutagenic compounds also known as acrylamide. The ability to keep a stable cookie production with homogeneus acceptable surface colour and texture as well as a minimum of acrylamide requires precise and objective control of the baking production line. Such objective measurement can be achieved through rapid assessment tools such as spectroscopy or digital images together with intelligent algorithms which are able to cope with some of the disadvanteges of using human experts.

In the work [5], the authors use spectral imaging to quantify enzymatic browning in apple slices.

In this work focus is put on the possibility of using a multispectral imaging system to quantify the browning of cookies by building an acceptance score based on expert selected cookies in different browning stages. The ability to combine browning progress together was expert panel knowledge in a single score is to the knowledge of the authors not investigated in previous studies. Furthermore, the ability to quantify water content from the same images as the acceptance score is estimated is investigated. The intention is to show that it is possible to extract different quality parameters from the same multispectral image. Furthermore the images are used to investigate how time and temperature affect the browning and moisture content. Control of initial moisture content and moisture migration is critical to the quality and safety of foods. Ideally, food manufacturers develop products with defined moisture contents to produce a safe product with optimum shelf-life.

Materials and Methods

A total of two datasets were created. The first set (set 1) was created as a training set to be used partly in a sensory evaluation of cookie appearance and partly to be imaged using a multispectral device. These images were then used to create an acceptance score which indicates the stage of browning a specific cookie is in.

The second set was used to build a response surface describing how time and temperature affects the acceptance score. All cookies were made in independent triplicates.

The color of all cookies were also measured in three points around the center of the cookie using a Minolta Colorimeter.

Baking procedure

Cookies with a thickness of approximately 6 mm were created in the laboratory. The cookies were created according to a recipe from Haas Meincke[1]. A table top “Teddy” mixer [2] was used to mix the following ingredients into a smooth and even dough.

Icing sugar

Margarine or butter

Vanilla extract

Whole egg

Skimmed milk powder

Salt

Sodium bicarbonate

Water

Wheat starch

Cake flour - wheat

Set 1 was baked in the oven at 180°C for different of time periods which ensured a good sampling of the entire browning process. Specifically the times in the oven was 4, 6, 7, 8, 9, 10, 12, 14, 16 and 20 minutes.
Set 2 were baked in the oven at the temperatures: 150, 160, 170, 180 and $200°C$, for each of the periods: 4, 6, 8, 10, 12, 14 and 16 minutes.
For both other parameters regarding dough composition, weight, pan dimensions, rising time, oven temperature, ventilation degree and cooling time before slicing, have been kept constant.

---

[1] http://dfemeincke.com/

[2] http://www.varimixer.com

Sensory cookie evaluation

Sensory evaluations were performed by the staff at the laboratory. A total of six persons participated. The cookies from set 1 were presented to the sensory panel. Each individual was asked to divide the cookies into three classes independently from the other members in the panel, based only on appearance. The classes are described as underbaked, adequately baked and overbaked. The sensory individuals were blinded to the bakingtime, so that the only factor playing a role in the decision was the appearance. The evaluation was done in a daylight setting.

Naturally a sensory panel varies significantly depending on the people it consists of. The variance depends on peoples preference to the degree of burning/baking a product has which is very individual. However the preferences depends largely on the culture of the people in the panel as well as their geographical origin. All persons in the present panel were people in mid twenties to thirties of danish origin.

Color conversion

In order to transform VideometerLab images to CIE $L^*a^*b^*$ images a numerical integration was done to calculate a response for each of the three types of color sensitive cones in the human eye [10]. The general relation between spectral reflectance and color is shown in Equation C.1.

$$P_i = \int_\lambda l_{D65}(\lambda)R(\lambda)O_{XYZ_i}(\lambda)d\lambda, \quad ,i = 1,2,3 \tag{C.1}$$

P_i contain the ith cone response, l_{D65} is the lightsource, R is the surface reflection as recorded by VideometerLab and O_{XYZ_i} is a human observer. It is necessary to transform the spectra to the XYZ colorspace [10] first since there is a linear relation between this colorspace and reflection spectra. A series of standard illumination sources exist and here the D65 corresponding to outdoor daylight illumination was used [10]. Since VideometerLab records relatively sparsely over the visual area, the spectra were interpolated using cubic splines. From the XYZ colorspace a non-linear standard transformation to the CIE $L^*a^*b^*$ colorspace was carried out [10].

Calculation of acceptance score

The acceptance score based on multispectral images and sensory evaluations is created using fishers discriminant analysis [27]. The images in data set 1 are classified into groups by the sensory panel and in turn transformed to multivariable data matrices containing only pixel values from the surface of the cookies. The within group and between group scattering matrices used in the fishers analysis were estimated based on all pixel values across all cookie surfaces in set one. Using the two scattering matrices an optimal projection vector is found such that each group when projected onto the vector will contract will the three groups will be scattered as much as possible. This means in the acceptance score space, which we call the space in which the cookies are projected into, the group separation is maximized in order to put a weight on the group membership.

Results and Discussion

Multispectral images from set 1 were treated to first remove the background, leaving only pixels with cookie surface. This was done using simple adaptive thresholding techniques. Calculating mean spectra for all cookies in set 1, revealed a clear correlation between time spend in the oven and spectral shape as seen in Figure C.1. Each spectrum is plotted as spectral reflectance value as a function of wavelength and colured according to the time spent in the oven; light colours are short time in the oven and dark colours are long time spent in the oven. Mainly there is a scaling difference between the spectra meaning the lightness. The brightness difference is largest between 400 and 700 nm (the visible wavelengths) while in the near infra-red area the scaling difference is smaller. Furthermore it is seen how there is a non-linear difference in the short wave end of the spectrum in the area between 400 and around 500 nm, which is the purple/blue area of the visible spectrum. It is well known that melanoids has absorption properties in this area, specifically at 420 nm [55]. As melanoids are formed during baking of the cookies it is natural that a spectral difference will occur in this area. The cookies were measured using a Minolata colorimeter in order to establish a correlation between the Minolta measurements and the VideometerLab measurements. Previous work has shown a correlation between Minolta measurements and the browning of cookies [7] and a correlation will serve as an establishing factor for using VideometerLab as an instrument for measuring the browning of cookies. Figure C.2 shows plots of Minolta measurements as well as those of VideometerLab. The CIE $L^*a^*b^*$ values show a clear decreasing trend as the baking time increases. Especially the L^* component, representing the lightness shows this trend clearly. In spectral terms the L^* component may be interpreted as the scaling mentioned earlier, which also showed a significant decreasing trend in Figure C.1. The a^* component representing position between red/magenta and green (negative values indicate green while positive values indicate magenta), shows an ascend from negative values to positive values showing the change from green values in the spectrum

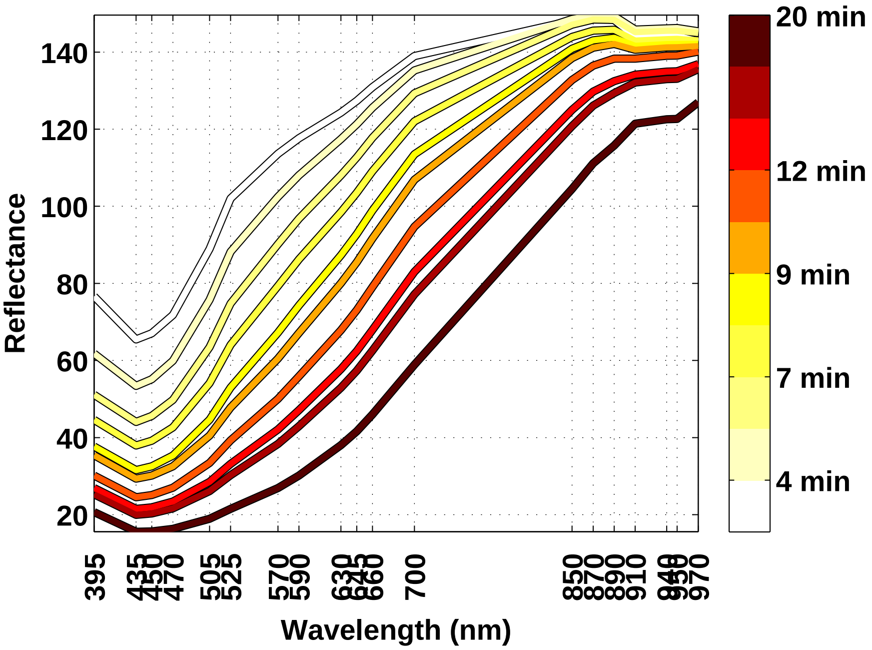
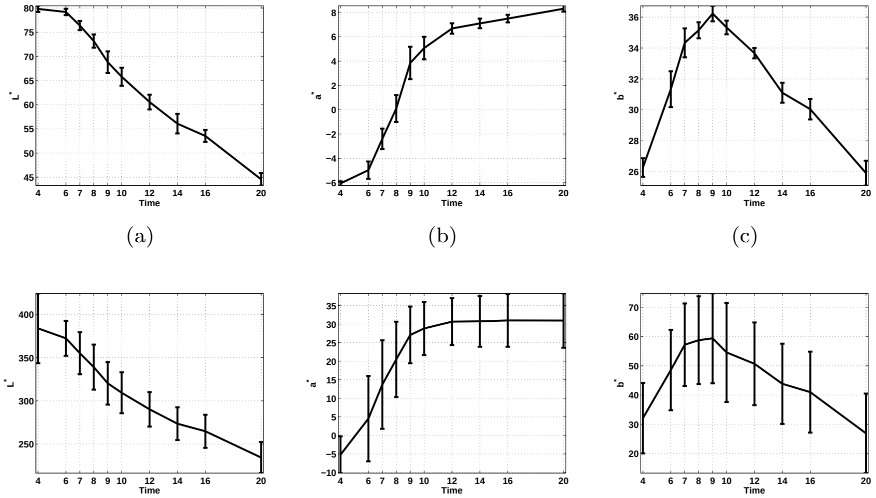


Figure C.1: Spectra of training data. Each spectrum represents the mean mean spectral shape of all cookie pixels in a multispectral image. The shapes show a clear correlation with baking time.

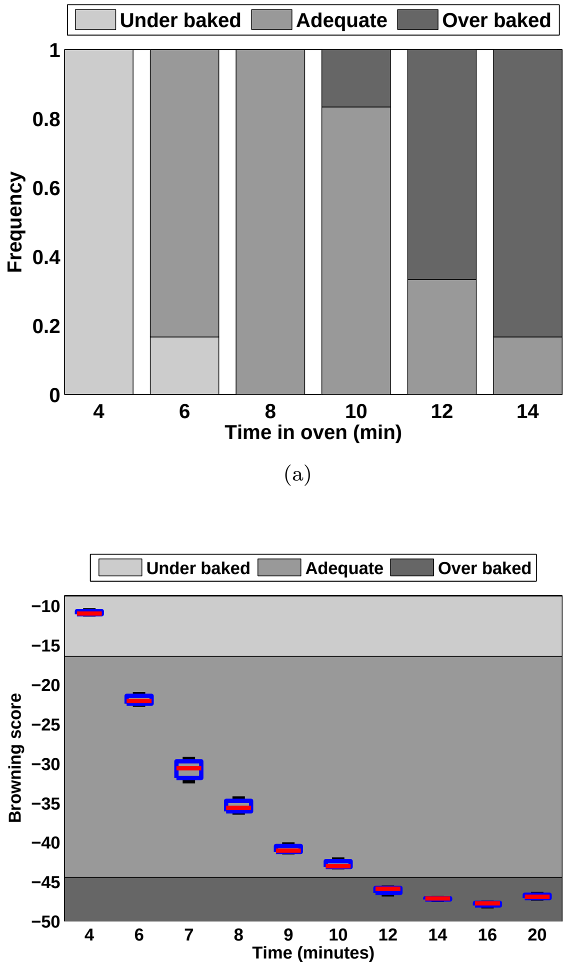
to more magenta values. This shift corresponds to a change around 510 nm as well as in the red and violet areas, corresponding to 400 and 700 nm. The a^* component seems to settle after approximately 10 minutes balking time. Finally the b^* component representing the position between yellow and blue (b*, negative values indicate blue and positive values indicate yellow) shows an increase in yellow until 10 minutes, where after the cookies get to brown and the yellow decreases again. Clearly the standard deviation of the transformed VideometerLab CIE $L^*a^*b^*$ values, which is averaged over a large quanitity of pixels has large standard deviation than the measured CIE $L^*a^*b^*$ values. This is expected as the multispectral images capture all reflection information on the surface and thus a large amount of variance while the CIE $L^*a^*b^*$ measurement only captures three distinct points. Sensory evaluations of set 1 is shown in Figure C.3(a). Each baking time was evaluated by each individual in the panel which shows that the adequate baking state is reached after six minutes. There is slight disagreement in the panel, however 80% agrees that 6 minutes baking time is enough for adequate baking. There is more disagreement on the transi-



(a) (b) (c)

(d) (e) (f)

Figure C.2: (a), (b), (c) shows the CIE $L^*a^*b^*$ from a Minolata colorimeter while (d), (e), (f) shows calculated CIE $L^*a^*b^*$ values for the corresponding VideometerLab images. All values are plotted against baking time. The errorbars indicate 1 standard deviation calculated from the repetitions. Minolta measurements has 3 samples measured on 3 distinct places on the cookie surface. The VideometerLab plots somewhat large deviations, however each timesamples is represented with approximately 400000 measurements corresponding to the amount of pixels on the surface of each cookie. Beside the color variation in each cookie there is of course also some noise. The mean value is plotted as a line, and there is good correspondence between VideometerLab measurements and Minolta measurements. The scale on the Y-axis differ due to the scaling in the calculation of CIE $L^*a^*b^*$ values from VideometerLab. Videometer images shows larger standard deviations since there for each cookie measurement is around $4x10^5$ observations and much larger probability for extreme values.



(a)

(b)

Figure C.3: (a) shows the evaluation of a sensory panel of six persons. Each person classified a set of cookies into each of the stages under baked, adequately baked and over baked. The histogram shows a normalized frequency for each group as a function of time in the oven. (b) shows a boxplot of mean values of acceptance score. Each box contains three samples. The background is shaded according to the classes defined by the sensory panel. The limits between the groups has been calculated as the mean euclidean distance between border boxes.

tion from adequate to the overbaked stage. 80% agrees that 10 minutes baking times produces adequate cookies while 40% agrees that also 12 minutes produce adequate cookies. There is no incident of all three classes occurring at the same baking time, which indicates very clear separation between the extreme classes. This clear separation is obvious in Figure C.3(b) where the browning stage has been quantified.

Here the large gap between the extreme classes is seen as euclidean distances based on an acceptance score. The browning score is calculated based on data from one of the three repetitions of set 1. A CDA was trained and a projection vector ensuring maximum separation between the three class was found. Figure C.4 shows the two most significant projection vectors or loadings from the CDA training. The most significant wavelengths in the first loading are 395 and 525, which both are seen to attribute largely to the variance in Figure C.1. The peaks in at these wavelengths might be influence by the acrylamide formed

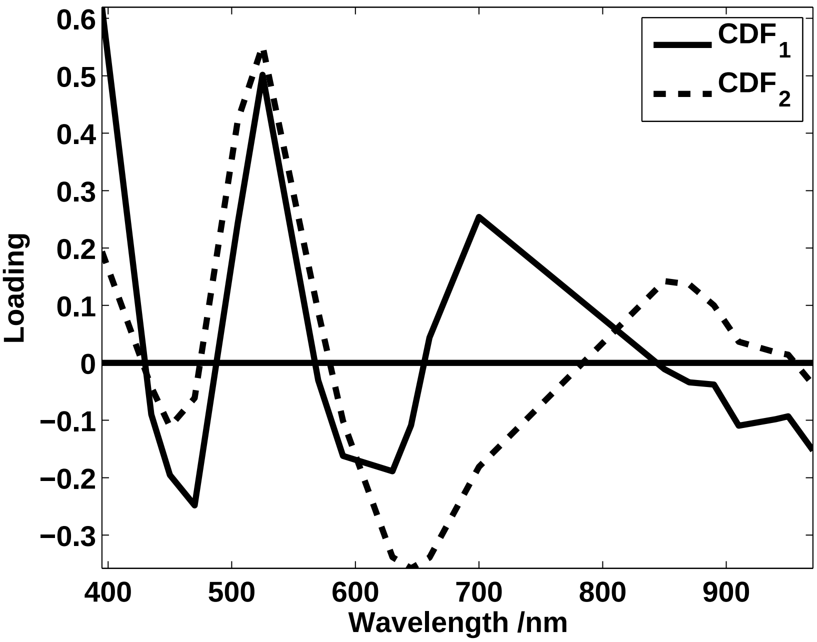


Figure C.4: The loadings from the canonical discriminant analysis used to create the acceptance score. It is clearly seen how the area from 400 to 550 plays a significant role in the transformation vectors.

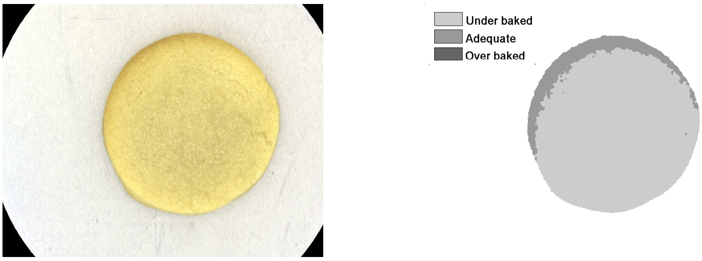
during baking which influences the area around 420 and 550 nm.

Each box in Figure C.3(b) contains mean score values of the cookies used to train the CDA model together with mean score values of the remaining two replicates. The spread of the values is relatively low, while the development over time shows a significant decrease. The limit between the classes were calculated as the mean euclidean distance between 4 and 6 minute acceptance score and the 10 and 12 minute acceptance score. The acceptance score can be applied to multispectral recordings of new cookies to predict which stage of browning they are in. This was done in Figure C.5. By applying the CDA transformation to the cookie images, it is possible to assess pixelwise acceptance scores. This enables a spatial investigation of each cookie to see if the browning is uniformly distributed across the cookie. Such assessment showed in Figure C.5 reveals how the browning propagates from the exterior toward the center. Underbaked areas were only detected in the cookie with 4 minutes baking time which corresponds well to the sensory evaluations. An overall acceptance score for each cookie is obtained by calculating the mean acceptance score of all pixels.

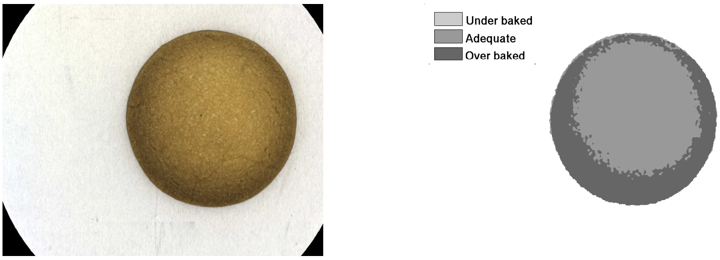
Mean acceptance score was calculated for all cookies in set 2 to assess the browning over a set of different times and temperatures. A quadratic surface (Equation C.2) was fitted to the acceptance scores in Figure C.2.

$$y = 0.32 - 0.12x_1 - 0.13x_2 + 0.008x_1x_2 + 0.02x_1^2 + 0.02x_2^2 \tag{C.2}$$

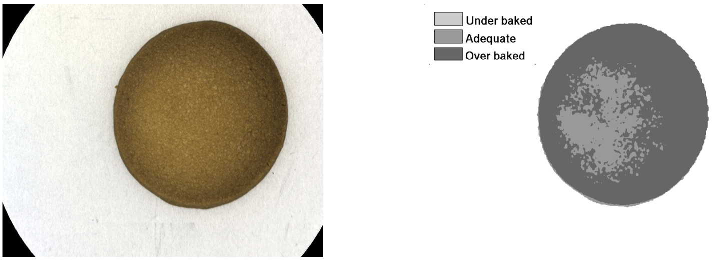
When looking at the response surface the gradient is more or less equal in both axes. This is further verified by looking at the model parameters for x_1 and x_2 which are almost equal. This means that with a cookie made from the ingredients listed in the materials section, the surface color changes equally fast in both directions. It is however still seen how the over baked state naturally dominates the upper right corner. Furthermore it is seen that the transition from under baked to adequately baked is just below 180°C which corresponds quite well with the sensory data.



(a) 4 min - RGB (b) 4 min - Labels



(c) 10 min - RGB (d) 10 min - Labels



(e) 12 min - RGB (f) 12 min - Labels

Figure C.5: pseudo RGB representations of the multispectral images of cookies taken at different baking times. For each RGB image a classified image is shown where Pixels has been classified to one of the three browning stages. The predicted images were smoothed with a gaussian filter before thresholding in order to remove noise. When observing images from top and down it is seen how the browning of the cookies starts from the edge and works towards the center. Only the first cookie with a baking time of 4 minutes contain areas of underbaked surface.

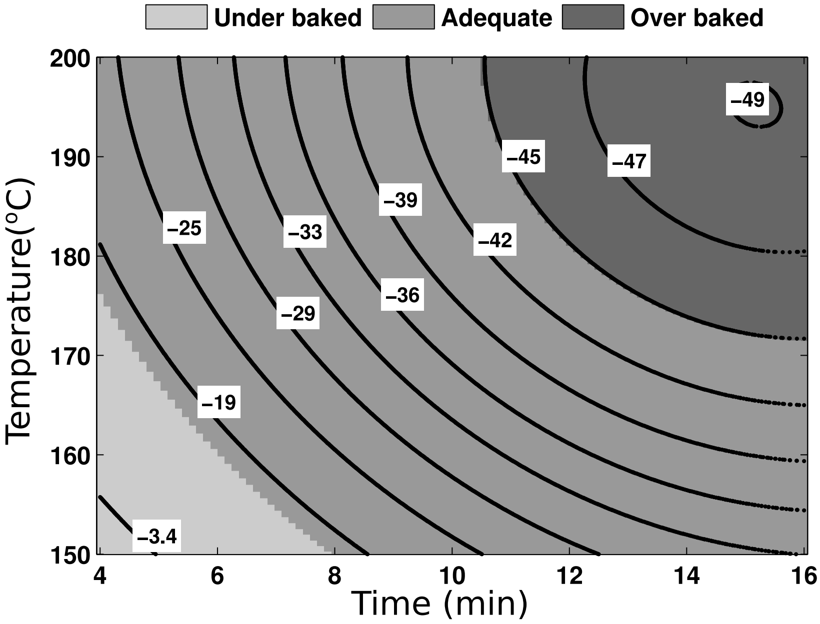


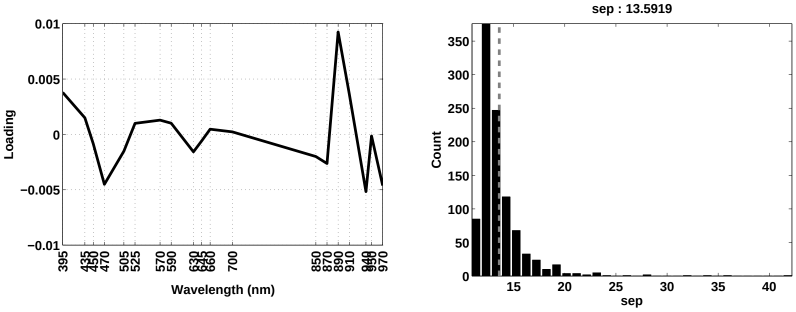
Figure C.6: The surface colour score as a function of time of temperature for cookies. Zones indicates areas created by the sensory evaluations, corresponding to under, adequate and over baked surface colour.

Determining water content

Literature shows that browning mainly happens as a consequence of temperature and water activity [53]. The non-enzymatic browning which is responsible for creation of among other things the brown nitrogenous polymers melanoidins is dependent on low water activity in the surface of the baking product. Due to an initial dehydration step in the Maillard reaction a certain temperature is needed to get the proper amount of activation energy to start the process. As the surface temperature increases and the water evaporates the rate of the Maillard reactions increase. Therefore a knowledge of the water content will give additional information on the browning process in the product. Generally when water activity decreases to 0.4-0.7 and temperature surpasses $105 - 120°\mathrm{C}$ [54].

NIR Multispectral imaging is capable of detecting water shifts in the NIR area of the electromagnetic spectrum, disregarding the surface color. The brown

colour on the surface products is known to have high absorbance properties in the UV and VIS area [55], quite far away from the water peaks at 970, 1440 and 1930 nm. The author was not able to find information in the litterature about NIR absorption on melanoids and other Maillard products, however when observing the spectra in the first part of the near infra red region it seems there is only slight variance here(700-970 nm.). This variance might be attributed to different molecules, however since water has an absorption peak in this area (970 nm) the surface water most certainly will contribute to this variation.
To investigate the relation between water content and the recorded spectra, a prediction model was created based on meanspectra of cookies and measured water $(kg_W kg_{DW}^{-1})$.
A ridge regression method was used to relate the spectral information directly to the measured water in the cookies. The spectral images used for this were the same as those used to create the browning surface in precious section. To select the complexity parameter of the ridge regression a ten fold cross validation was used, while the assessment of the model was based on a bootstrap method with 1000 resamplings. Figure C.7(a) shows the loadings of the best regression model. Basically the model seems to put importance to the near infra red area as well

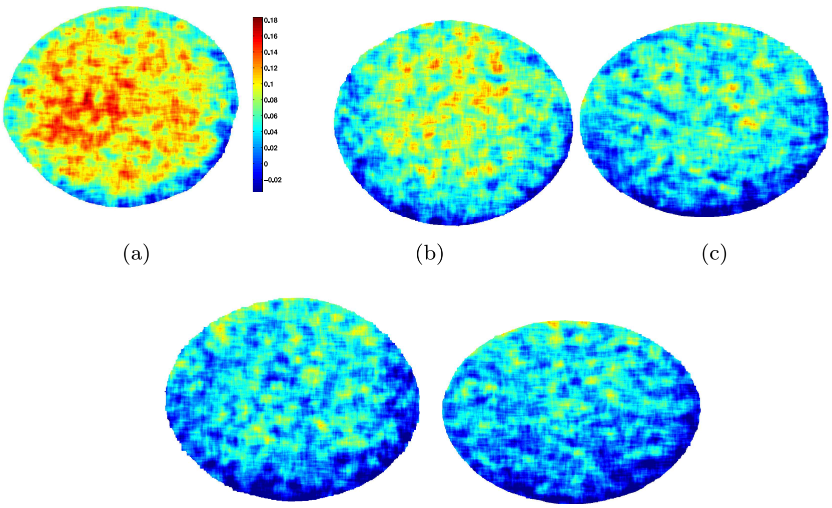


(a) (b)

Figure C.7: (a) shows the loadings of the water prediction model. The near infra red area shows significant peaks together with the area around 470 nm. (b) shows values of standard error of prediction calculated in a bootstrap procedure. The average of the SEP values is calculated to 13.6, meaning on average the prediction method has an errorrate of 13.6%.

as the area around 470 nm. The large loading around 470 nm might be because of a correlation between an increasing browning (shift in the area around 470 nm) and a decreasing water content in the surface (shift in the near infra red area). The dimensionless parameter, standard error of prediction (SEP) which is basically a rescaling of the well-known Root Mean Square Error (RMSE) is used as to evaluate the model. The distribution of SEP based on the bootstrap

is seen in Figure C.7(b), and the mean value is calculated to be 13.6%. Thus on average the predictions may vary with up to 13.6% from the real value.
The spatial distribution of water in each cookie may similarly be assessed by projecting each recorded spectrum in the cookie surface (one per pixel) with the calculated loading (Figure C.7(a)). Figure C.8 shows spatial predictions on cookies baked in the oven at 150°C in different time periods. Table C.1 shows the corresponding predictions for the same cookies. The visual impression is very appealing. Due to noise the predicted images were averaged with a mean filter with a size of 30 pixels. A clear negative gradient is seen going from image C.8(a) to image C.8(e), corresponding to 4 and 12 minutes baking time. This is also confirmed by the actual water content as measured with the oven method in Table C.1. Furthermore it is seen how the water is relatively uniformly distributed with a slight tendency to cluster in the center. By averaging all predicted pixels an overall estimate of the water content in the cookie can be calculated, which shows errors with limits within the calculated SEP. The largest errors is the most moist cookie at 0.0107 $kg_W kg_{DW}^{-1}$ corresponding to approximately 10% deviation within the range of the dataset.



(a) (b) (c)

(d) (e)

Figure C.8: Water content for five cookies have been predicted for each pixel on the cookie surface and smoothed to remove noise. The average water prediction for the five cookies has been calculated and presented in Table C.1. The colormap used in the visualization of the five surface is the same where blue values correspond to low water content and red values correspond to high water content.

Table C.1: Real and predicted water content for the cookie surfaces shown in Figure C.8. The real values have been measured using a drying cabinet while the predicted values were calculated as average values of all pixels on the surface.

Time	Temperature	Real ($\frac{kg_W}{kg_{DW}}$)	Predicted ($\frac{kg_W}{kg_{DW}}$)	Error	Figure
4	150	0.1037	0.093	0.0107	C.8(a)
6	150	0.0684	0.0596	0.0088	C.8(b)
8	150	0.0475	0.0399	0.0076	C.8(c)
10	150	0.0370	0.0316	0.0054	C.8(d)
12	150	0.0244	0.0255	-0.0011	C.8(e)

Conclusion

Studying browning in cookies using advanced equipment is of importance for the baking industry to improve the quality inspection at the processline. The development of non-enzymatic browning, Maillard reactions and caramelization has been a topic for academic investigation for many years. However only recently people have started using first colorimeters and now digital cameras. In this work assessment of the browning process using multispectral imaging was proposed. The multispectral system proposed records multispectral images in the visible and first part of the near infra red area of the electromagnetic spectrum. An acceptance score based on the evaluation of a set of cookies by a sensory panel were developed using canonical discriminant analysis. The sensory panel consisting of 6 persons decided to vote cookies with a baking time of 4 minutes as under backed while cookies with a backing time from (and including) six minutes to (and including) 10 minutes as adequately baked, while all times above as over baked. The evaluations were purely based on the visual appearance of the cookies, thus by colour and visual texture. The discriminant analysis revealed that in order to map a spectrum to an acceptance score was 395 and 525 nm, corresponding very well with the absorption properties of melanoids. From the acceptance score a response surface was created to build a model which relates the influence time and temperature has on the browning. It was found that on the scale of the eternal parameters, time and temperature, that the change in surface color happened at the same rate per unit step. Furthermore a broad acceptance band was seen, also corresponding very well with the sensory evaluations. A demonstration of the browning distribution on the surface of the cookies was presented which revealed how the browning started from the edges of the cookie and worked its way to the interior. Such information may be used to control the uniformity of the browning of each cookie.

Besides assessing the browning on the surface of the cookies, the surface water content was similarly evaluated. An approach similar to the acceptance score approach was used, however to estimate the water content a ridge regression

technique was used. The ridge regression needs to select a complexity parameter to control the regularization. This complexity parameter was chosen using ten fold cross validation. Similar to the CDA, the ridge regression also projects new observations using a loading vector. The loading vector revealed that in order to map spectra to water content the NIR infrared area as well as 470 nm was of significance. Having the NIR area in a water prediction model was expected due to the water peak around 970 nm, however the importance of the 470 nm may be due to the correlation of the browning and the evaporating water which are negatively correlated.
A bootstrapping method was used to assess the standard error of prediction - a dimensionless error parameter expressed in percent, of the water prediction model. The Standard error of prediction showed an error of approximately 13.6% on average. The water prediction model was demonstrated to predict the distribution of the water activity on 5 cookies. An intuitive visualization result was obtained where the model showed how water evaporated from the surface over a period of up to 12 minutes baking time. Similar to the browning of the cookies, the evaporation propagated from the edges towards the interior as the time elapsed. Since the Maillard reactions will be stronger with less water up til a certain point, the visualization of the acceptance score and the water content both propagating from the edge towards the interior seems to be a good result which also is in accordance with the literature.

Acknowledgements

The author would like to thank Mette Stenby for producing the dataset used in this technical report.

References

[1] Emmanuel Purlis. Browning development in bakery products - a review. Journal of Food Engineering, 99(3):239 – 249, 2010.

[2] Fergus M. Clydesdale. Color perception and food quality. Journal of Food Quality, 14(1):61–74, 1991.

[3] A. C. Little. American Society for Testing and Materials, chapter Color evaluation of foods, correlation of objective facts with subjective impressions, pages 109–127. 1973.

[4] Sara I. F. S. Martins, Wim M. F. Jongen, and Martinus A. J. S. van Boekel. A review of maillard reaction in food and implications to kinetic modelling. Trends in Food Science & Technology, 11(9-10):364 – 373, 2000.

[5] Loredana Lunadei, Pamela Galleguillos, Belén Diezma, Lourdes Lleó, and Luis Ruiz-Garcia. A multispectral vision system to evaluate enzymatic browning in fresh-cut apple slices. Postharvest Biology and Technology, 60(3):225 – 234, 2011.

[6] Gunter Wyszecki and W.S. Stiles. Color science: Concepts and Methods, Quantitative Data and Formulae, second edition. Wiley & sons Inc, 2000.

[7] Ella Pagliarini, Monica Vernille, and Claudio Peri. Kinetic study on color changes in milk due to heat. Journal of Food Science, 55(6):1766–1767, 1990.

[8] Ji-Sang Kim and Young-Soon Lee. Enolization and racemization reactions of glucose and fructose on heating with amino-acid enantiomers and the formation of melanoidins as a result of the maillard reaction. Amino Acids, 36:465–474, 2009. 10.1007/s00726-008-0104-z.

APPENDIX D

Monitoring water content using multispectral imaging and NIR in minced meat

---

Published in Proceedings for Near Infrared Spectroscopy International Conference, NIR 2009, November 2009

Monitoring water content using multispectral imaging and NIR in a minced meat preparation process

Bjørn S. Dissing,[a] Bjarne K. Ersbøll[a] and Jens Adler-Nissen[b]
[a]*Department of Informatics and Mathematical Modeling*
[b]*National Food Institute, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark. bdi@imm.dtu.dk*

Introduction

Online quality inspection of food process control is today often done by human expert operators who have many years of experience. However, the trend seems to point towards fast non-invasive inspection methods, such as near infrared (NIR) technology for quality inspection in different food process control tasks, as either replacement for, or supplement to the human operators.

We are investigating the potential of using multispectral imaging in the visible as well as the NIR area of the electromagnetic spectrum instead of human operators and as an alternative to standard NIR measurement methods. A drawback of spectroscopic methods is its one dimensional nature. A spectroscope measures everything within its field of view as an average measurement over the area registered by the measuring device. By employing imaging instead of point measurements it is possible to record much larger spatial areas, and thereby gain spatial as well as spectral information. This makes it possible to assess chemical as well as spatial quality features at-line, such as water content, surface color, fat content, particle sizes, texture etc.

In this study, we are specifically investigating the ability of a multispectral camera to predict the water content in minced meat after it has been processed in a continuous frying process at different times and temperatures. Other similar investigations have been done.[1] The camera used is called a VideometerLab,[1] and records multispectral images in a set of predefined wavelengths.

Some absorption bands of water lies in the vis and NIR area around 640 nm, 752–756 nm, 960 nm and 1152–1160 nm. VideometerLab overlaps a large part of this region, which is what we want to utilise to quantify the amount of water in the surface of fried minced meat, by correlation to dry-matter measurements of the same sample.

---

[1] http://www.videometer.com

Experimental setup and data

The study subject of this paper is minced beef with 15–18% fat. The meat was purchased through a wholesale supplier, Inco Danmark a.m.b.a, Copenhagen in a frozen state. The meat blocks of approximately 2 kg. were stored at –30°C. For the experiment the meat blocks were crushed with a hammer in coarse pieces below 200 g. A portion of about 1 kg was chopped in an industrial meat chopper (Kilia 0.57 m diameter) at the lowest speed step, to prevent heating. The chopping was continued (about 2–3 min.) until the frozen meat was disintegrated with no large lumps left.

Of the disintegrated, still frozen meat, 800 g was fed in consecutive portions of 100 g each to a continuous frying machine at pre-selected temperatures and frying times. Samples were prepared at the temperatures 200°C, 225°C and 250°C; the frying time varied from 120 s to 240 s in 40 s intervals.

The actual water content was determined using a standard dry-matter method (oven drying at 105°C for 24 h) where the mass of evaporated water of the samples (about 2 g each) was measured by weighing. All measurements were done in triplicate to get more stable measurements, where the mean of the three replications is seen in Table 1.

Standard deviations of the replicates were estimated between 0.18 and 0.42.

The multispectral images were acquired using a VideometerLab, see Figure 1, which records 18 different reflectance spectra corresponding to the wavelengths; (430 nm, 450 nm, 470 nm, 505 nm, 565 nm, 590 nm, 630–645 nm, 660 nm, 700 nm, 850 nm, 870 nm, 890 nm, 910 nm, 920 nm, 940 nm, 950 nm and 970 nm).

The VideometerLab uses an LED technology, which means that no filtering of the incoming light is needed. Furthermore, the camera is equipped with an integrated sphere coated with a matte material, which ensures uniform lighting, avoids highlights and makes it easy to optimise the dynamic range in low contrast areas. VideometerLab technology is a low cost way of acquiring multispectral images since it uses standard silicium chip technology.

For the entire experiment, only the longest 8 wavelengths, the NIR channels, were used for the analyses, in order to avoid e.g. confounded variables in the data modeling.

Table 1. Water Content determined by the dry-matter method.

	200 °	225 °	250 °
120 s	54,32 %	53,17 %	51,01 %
160 s	52,66 %	53,96 %	46,29 %
200 s	51,49 %	52,55 %	49,70 %
240 s	51,16 %	51,27 %	48,28 %

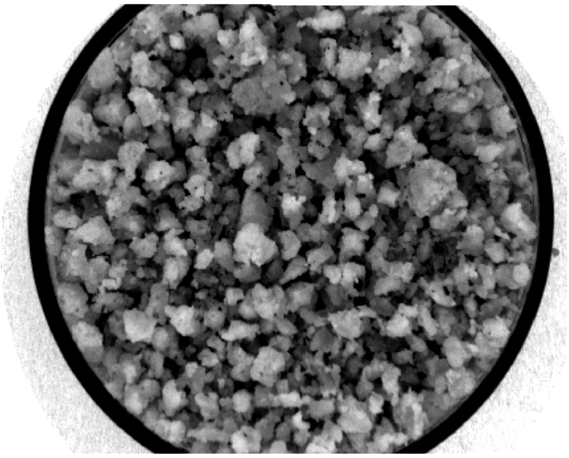


Figure 1. 850 nanometers: unprocessed image.

Data analysis

Having a set of 8 NIR images measured between 850 and 970 nanometers (both included), with corresponding pixels, leads to application of multivariate and chemometric methods. Before applying a multivariate method and calibrating the water system, an expansion of the multispectral images was done. All possible ratios of all eight NIR wavelengths were derived in order to find better features to describe the water content of the sample. The new multispectral image including the ratio set had a total of p^2 channels, where p is the number of original channels. For each of

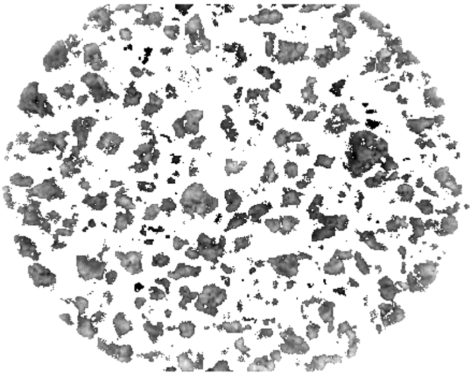


Figure 2. 850 nanometers: processed image: only local maxima remain.

the channels in this new image, a mask was created to include only the local maxima of the meat granules using a mathematical morphological operation, called the h-domes technique.[2] The very coarse surface of the minced meat gives rise to shadows with less information, which is why only the fully illuminated areas of the images are considered in the analysis, as seen in Figure 2.

The preprocessed p^2 channels were then converted to intensity densities, where the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 99th percentiles were calculated. The percentiles give information about the shape of each of the p^2 density estimations, and are used as covariates in a now very ill posed linear regression problem of the form

$$y=Ax+b+\varepsilon$$

The actual water content from is used as the dependent variable y, and the expanded feature space of the multispectral images is used as the independent variable space *x*.

A large amount of the calculated variables are nearly linearly dependent, and bring no actual information. This means a standard PLS calibration would give us a very large and complex model. In order to get a more parsimonious and interpretive model, a sparse method is utilised to select relevant features. Many sparse methods exist to solve such problems, where a very well known and intuitive method is the stepwise selection method, chosen to solve this problem.

Results and discussion

The very sparse set of NIR measurements made this an interesting study. Even though the water absorption band at 850 and 970 nanometers are relatively weak compared to bands at lower frequencies we managed to get good correlation results with independent chemical measure-

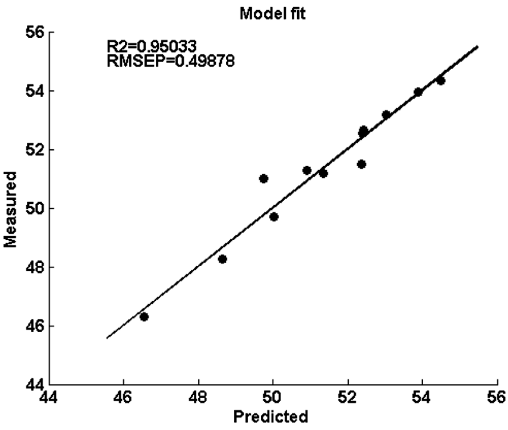


Figure 3. The resulting model shows good correlation between measured and predicted values of water.

ments. A value of R^2 of 0.95 was obtained between the predictions and the measured response in is seen in Figure 3.

The final model selected by the stepwise regression was

$$y = \alpha_1 q(\lambda_{970}, 50\%) + \alpha_2 q(\lambda_{890}, 99\%) + \alpha_1 q\left(\frac{\lambda_{920}}{\lambda_{910}}, 1\%\right) + \varepsilon$$

α_i denotes the coefficients, q denotes the quantile function with two parameters, the first being the variable from the expanded basis, and the second being the quantile-number. In this study, the absorption bands at 970 and 890 nanometers play the most important role in the prediction model, which had a p-value of 0.03. The model performance was calculated using a Leave One Out Cross Validation (LOOCV) scheme, due to the small amount of observations.

Due to water bands high sensitivity to temperature, it is of course important to emphasise that the measurements in this experiment were performed at 20°C. This means that for on-line use, a calibration needs to be performed in order to compensate for the temperature.

References

1. S.B. Daugaard, J. Adler-Nissen, J.M. Carstensen *Food Control* **21(5),** 626 (2009).
2. Luc Vincent. In *IEEE Transactions on image processing*, **2(2),** 1993.

Appendix E

Using Multispectral imaging for spoilage detection of pork meat

Submitted to Food Research International

Using Multispectral imaging for spoilage detection of pork meat

Bjørn Skovlund Dissing[a,], Olga Papadopoulou[b,c], Chrysoula Tassou[c], Bjarne Kjaer Ersbøll[a], Jens Michael Carstensen[a], Stathis Panagou[b], George-John Nychas[b]

[a]*Informatics and Mathematical Modelling, Technical University of Denmark, Institute for Informatics and Mathematical Modeling, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark*
[b]*Laboratory of Microbiology and Biotechnology of Foods, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, 11855 Athens, Greece*
[c]*Institute of Technology of Agricultural Products, National Agricultural Research Foundation, Sofokli Venizelou 1, Lycovrissi, Greece, GR-141 23*

Abstract

A large experiment regarding quality in stored minced pork meat has been performed to test the ability of a rapid multispectral imaging device to quantify spoilage degree. Bacterial counts of a total of 167 meat samples stored for up to 580 hours have been quantified using traditional laboratory methods. Meat samples were contained under 2 different storage conditions - aerobic and modified air packages as well as under different temperatures. Besides bacterial counts, a sensory panel has judged the spoilage degree of all meat samples into one of three classes. We find that the multispectral imaging device is able to classify 76.13% of the meat samples correctly according to the defined sensory scale. Furthermore we find that using the multispectral camera device, we are able to predict total viable count with a standard error of prediction of 7.47%. We conclude that there is a good possibility for a setup like the one investigated will work for detection of spoilage degree in minced pork meat.

Keywords: Multispectral Imaging, Meat spoilage, Baranyi modeling, Chemometrics, Aerobic storage, MAP storage

---

Email address: `bdi@imm.dtu.dk` (Bjørn Skovlund Dissing)

Preprint submitted to Meat Science *May 8, 2011*

1. Introduction

Colour is among the most important factors playing a significant role in the evaluation of meat quality. Muscle colour at the point of purchase is an indicator of freshness and anticipated palatability for the consumer [1, 2]. Microbial growth in fresh meat is the main factor associated with deterioration of meat quality, spoilage and economic loss, and is closely associated with the colour and appearance of the meat[3]. During conventional slaughter procedures and further processing necessary to prepare meat for consumption, microorganisms are introduced into and onto carcasses.The type of spoilage is affected by intrinsic and/or extrinsic parameters [4]. These could be the characteristics of the meat or the storage or transportation environment as well as processing parameters such as how the meat is minced or grinded [3, 5]. These parameters influence the establishment of the particular microbial association and determine the rate of attainment of a climax population so called 'Ephemeral Spoilage micro-Organisms' (ESO) [3, 6]. It has been shown that pseudomonads are the ESO under aerobic storage conditions while Lactic acid bacteria or Brochothrix thermosphacta have been named as such for storage under Modified Atmosphere Packaging (MAP) conditions. So far more than 50 chemical, physical and microbiological methods have been proposed for the detection and measurement of bacterial safety or spoilage in meat. However most of these methods are time-consuming and provide retrospective information and thus they can not be used on- or at-line. Additionally the changes and development in technologies for food processing and preservation [e.g., vacuum packaging (VP), modified atmosphere packaging (MAP), active packaging, etc] make it evident that the important and urgent task of identifying safety and spoilage indicators is a complicated proposition. On the other hand, the meat industry needs rapid analytical methods or tools for quantification of these indicators in order to determine and select suitable processing procedures for their raw material and to predict the remaining shelf life of their products. Furthermore, inspection authorities need reliable methods for control purposes, while the wholesale and retail sectors need these valid methods to ensure the freshness and safety of their products and to resolve potential disputes between buyers and sellers. The use of microbial metabolites as a consequence of microbial growth in meat has been continuously recognized as a potential means for assessing meat quality [3, 7].

Multispectral imaging techniques are a natural extension to normal colour cameras. Whereas normal colour cameras integrates electromagnetic radiation over three broad bands in the visual range, multispectral cameras are able to record electromagnetic information in more narrow bands. This means multispectral cameras are able to record spectral reflection properties in narrow bands, which thereby makes it possible to assess the composition of surface chemistry of the object of interest. Such recordings may thus be used to extract intrinsic chemical and molecular information such as water, fat, protein or other hydrogen-bonded constituents. Sometimes multispectral images are also referred to as surface chemistry maps [8] or hypercubes. In short, multispectral images can provide not only spatial information, as regular imaging systems, but also spectral information for each pixel in an image. Thus, using hyper-spectral images, it is possible to assess physical and geometric characteristics such as colour, size, shape, and texture. Several publications have been written on the subject of using multispectral imaging for food control [9, 10, 11].

Multispectral images are a natural source of massive high dimensional datasets, which may be analysed using specific techniques. A discipline which is gaining ground in the area of chemical and biological analysis is the discipline of machine learning. When used in conjunction with massive modern datasets of chemical or biological character it is sometime referred to as chemometrics, which focuses more on specific machine learning methods than others. Machine learning covers advanced statistical and numerical machine learning methods such as support vector machines, cluster analysis, neural networks, partial least squares and logistic regression. Both supervised as well as unsupervised statistical methods are considered, and may be used in conjunction with multispectral image analysis to relate multispectral images to chemical reference measurements values or sensory labels. Clustering algorithms based on data from multispectral images have been thoroughly investigated to assess food quality [12].

In this study, the potential of multispectral imaging techniques was exploited for assessment of spoilage degree in pork meat. The specific objective was to evaluate pork meat quality based on spectral as well as spatial information using various chemometric techniques; regression and clustering. The aim was to make a method that would be able to evaluate meat spoilage at different storage temperatures (0, 5, 10, 15, and $20°C$) and package type (aerobic and modified atmosphere). The evaluation was based on predicting microbial growth as well as by classifying between different quality classes as evaluated by a sensory panel.

2. Materials and Methods

2.1. Sample preparation

Fresh minced pork was purchased from the central meat market in Athens and transported under refrigeration to the laboratory within 30 min,. Minced pork was divided in portions of 75 g and packed individually either aerobically or under modified atmosphere (MAP). For aerobic storage, meat samples were placed on foam trays which were subsequently wrapped with air-permeable polyethylene plastic film. Moreover, samples packaged in MAP were enclosed inside plastic pouches with oxygen permeability of $6cm^3m^{-2}24h^{-1}$ at $20°C$ and 50% RH, flushed with a gas mixture of 60% CO_2 / 20% O_2 / 20% N_2 and heat sealed with a HencoVac machine (Howden Food Equipment BV, The Netherlands). Meat samples were stored under controlled isothermal conditions at 0, 5, 10, 15, and $20°C$ in high precision ($\pm 0.5°C$) incubators (MIR-153, Sanyo Electric Co., Osaka, Japan) for an overall period of 580 h, depending on storage temperature, until spoilage was pronounced.

2.2. Microbiological analyses

Minced meat samples (25 g) were weighed aseptically, added to sterile quarter strength Ringers solution and homogenized in a stomacher (Lab Blender 400, Seward Medical, London UK) for 60 s at room temperature. Further decimal dilutions were prepared with the same medium and duplicate 0.1 or 1 ml samples of the appropriate dilutions were spread or mixed on the following agar media: Plate Count Agar (PCA, Biolife 4021452, Milano, Italy) for total viable counts (TVC), incubated at $30°C$ for 48h; Pseudomonas Agar Base CN selective supplement (PAB, Biolife 401961, Milano, Italy) for Pseudomonas spp, incubated at $25°C$ for 48-72h; Streptomycin Thallous Acetate-Actidione Agar (STAA, Biolife 402079, , Milano, Italy) for Brochothrix thermosphacta, incubated at $25°C$ for 72h; Rose Bengal Chloramphenicol agar (RBC, LabM36, supplement X009, LabM, London, England) for yeasts and moulds, incubated at $25°C$ for 72h; Violet Red Bile Glucose Agar (VRBGA, Biolife, 402185, Milano, Italy) for Enterobacteriaceae counts, overlaid with the same medium and incubated at $37°C$ for 18-24h; de Man-Rogosa-Sharp medium with pH adjusted to 5.7 (MRS, Biolife, 4017282, Milano, Italy) for lactic acid bacteria, overlaid with the same medium and incubated at $30°C$ for 48-72h. Duplicate packages from each storage temperature and packaging condition were analysed at appropriate time intervals to allow for efficient kinetic analysis of different microbial groups. Growth data from plate counts were log transformed and fitted to the primary model of Baranyi and Roberts [13] using the DMFit /program (available at www.combase.cc) to determine the kinetic parameters of microbial growth (maximum specific growth rate and lag phase duration). In parallel with microbiological analyses, the pH value of minced pork meat was recorded with a digital pH-meter (Metrohm pH Lab, Switzerland), the glass electrode of which was immersed in the homogenised meat sample after the end of microbiological analysis.

2.3. Sensory analysis

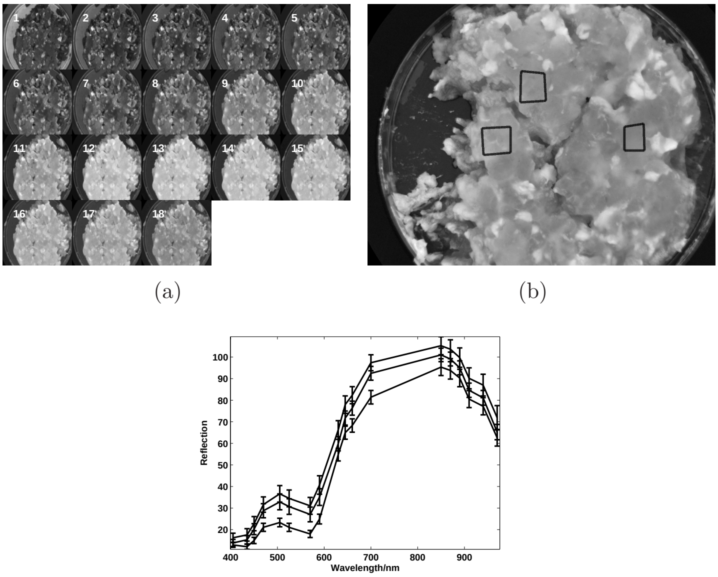
Sensory evaluation of pork samples was performed during storage by a sensory panel composed of three members (in-house trained staff from the laboratory) at the same time intervals as for microbiological analyses as described elsewhere [14]. The same trained persons were used in each evaluation, and all were blinded to the meat sample tested. The sensory evaluation was carried out in artificial light and the temperature of the packed product was close to ambient. The descriptors selected were based on the perception of colour, smell, and taste. The first two descriptors were assessed before and after cooking for 20 min at $180°C$ in a preheated oven, while the last descriptor was evaluated only after cooking. Each sensory attribute was scored on a three point hedonic scale corresponding to: Fresh, Semi-fresh and Spoiled. The first vague indications of meat spoilage occurred at the point labeled Semi-fresh. Odour characteristics of minced pork, as determined by special samples kept frozen and thawed prior to each sensory evaluation, were considered as fresh. putrid, sweet, sour, or cheesy odours were regarded as indicative of microbial spoilage and classified the samples as spoiled. Bright colours typical of fresh oxygenated meat were considered fresh, whereas a persistent dull or unusual colour rendered the sample spoiled [15, 16, 14]. Overall, 167 minced pork meat samples were scored by the taste panel and discriminated into the pre-defined groups as fresh (22), semi-fresh (65), and spoiled (80).

2.4. Multispectral Imaging System

The data acquisition was done using VideometerLab [17], which acquires multi-spectral images in 19 different wavelengths ranging from 405 to 970nm. The spectral radiation of the 19 bands is not uniformly distributed over the sampling area but rather at wavelengths 405, 435, 450, 470, 505, 525, 570, 590, 630, 645, 660, 700, 850, 870, 890, 910, 940 and 970 nm. The acquisition system records surface reflections with a standard monochrome charged coupled device chip, nested in a Point Grey Scorpion camera. The object of interest is placed inside an integrating or Ulbricht sphere in which the camera is top mounted. The sphere has its interior coated with a matte coating. The coating together with the curvature of the sphere ensures a uniform reflection of the cast light, and thereby a uniform light in the entire sphere. At the rim of the sphere, Light Emitting Diodes (LED) with narrow-band spectral radiation distribution are positioned side by side. The LEDs are placed in a pattern which distributes them uniformly around the entire rim. When an image is acquired the LEDs are turned on successively and the reflection from that specific wavelength is recorded by the top mounted camera. The result is a monochrome image with 32 bit floating point precision for each LED type, giving in the end, a hyperspectral cube of dimensionality 1280x960x19. The system is first calibrated radiometrically and geometrically using well-defined standard targets, followed by a light setup based on the type of object to be recorded [18]. The homogeneous diffuse light together with the calibration steps ensures an optimal dynamic range and minimizes shadows and shading effects as well as specular reflection and gloss-related effects. The system has been developed to guarantee the reproducibility of collected images which means it can be used in comparative studies of time series or across a large variety of different samples [19, 20, 21, 22].

2.5. Image processing and Data Analysis

An example of a recorded multispectral image is seen in Figure 1(a), where the channels are listed according to wavelengths mentioned in previous section. Moreover, Figure 1(b) and 1(c) illustrate the mean reflectance spectrum with error-bars indicating 1 standard deviation for 3 different locations in the same piece of meat. The mean spectra are calculated as mean values of all the pixels within each of the three squares indicated in Figure 1(b). The diversity of spectra contained within each image may be appreciated when looking at the difference between these mean-spectra. The largest difference between the spectra is a scaling, which basically indicates how much light is reflected in general in a point, or in other words the luminosity. Scaling differences has many causes, but two essential causes are shadow effects as well as light scattering effects due to the topology of the surface. In order to compensate for such influences, a simple preprocessing step commonly known as autoscaling or simple standardization of data [23] which centers all spectra and ensures unit variance has been used on the images. Such a preprocessing step will help enhance the true differences in the spectra, and thereby improve the later signal processing. Figure 2 shows an overview of the entire



(a) (b)

(c)

Figure 1: (a) All channels, ranging from 410 nanometer to 970 nanometer. (b) Channel recorded at 525 nm with squared annotation areas. (c) Mean Spectra of square areas in (b)

data processing pipeline which has been performed in order to quantify the spoilage degree. The images are initially segmented or divided into regions of interest. This means dividing the images into foreground and background, meaning only the meat area except fat-areas is to be considered in the further statistical analysis. It is common to use orthogonal transformations such as Principal Component Analysis to find similar tissues in the multispectral image space. However, PCA creates orthogonal projection vectors based on a variance maximizing criterion and does not take spatial information into account. The Maximum Noise Fraction (MNF) [24] is a related method also belonging to the orthogonal transformation function family which seeks to maximize the Signal to Noise ratio (SNR) instead of the variation. This is done by estimating the covariance matrix of the spectra, Σ as well as the covariance matrix of the noise Σ_Δ [25] and finding vectors, b, that maximizes the ratio of these (the Rayleigh quotient)

$$R(b) = \frac{b^T \Sigma_\Delta b}{b^T \Sigma b} \tag{1}$$

When the noise covariance structure is estimated as the covariance of the difference of neighboring pixels, the MNF reduces to the Maximum Autocorrelation Factor (MAF). A maximization of the signal to noise ratio is then achieved by minimizing the autocorrelation between neighboring pixels. This optimally finds projection directions of similar neighboring reflection properties. Two components are used to cut away background as well as fat tissue, where an adaptive thresholding technique, [26], is used to transform components to masks which indicate pure meat. Before clustering the spectra, a standardization of the spectra as previously described has been performed. All pixels are finally mapped to identified clusters to indicate areas of spoilage in order to improve predictability of the entire sample image.

3. Results and Discussion

3.1. Development of microbial association

The changes in the population of Total Viable Counts (TVC) during storage of minced pork meat storage at different temperatures and packaging conditions is presented in Figure 4, whereas the estimated kinetic parameters for the Total Viable Counts after fitting the primary model of Baranyi and Roberts are shown in Table 1 for aerobic and MAP packaging. The model fitted the experimental data well as can be inferred by the low values of the standard error of fit and the high values of R^2. However, in some cases of data derived from MAP packaging (0 and $5°C$), no typical sigmoidal growth curves could be obtained (Figure 4). A lag phase was observed at 0 and $5°C$ in aerobically stored meat samples, the duration of which was greatly reduced or not observed at all at higher temperatures. However, in MAP packaging no lag phase was observed in most of the cases (Table 1). A progressive increase of maximum specific growth rate (μ_{max}) was observed for all members of the microbial association with increasing storage temperature. The influence of packaging type was also evident on the calculated values of μ_{max}. In fact, at the five different storage temperatures assayed, the microbial

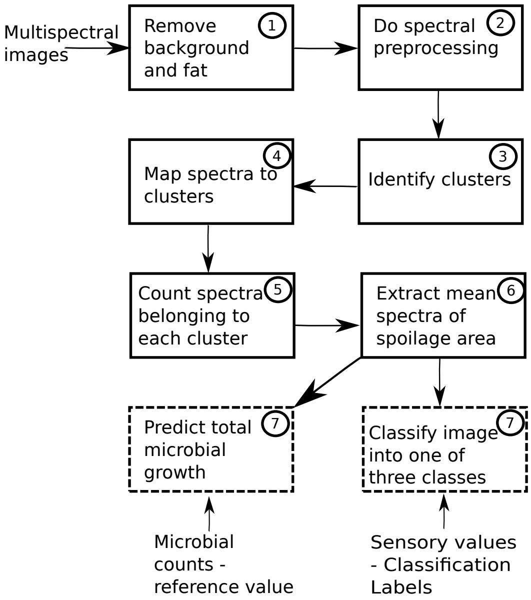


Figure 2: Schematic structure of processing pipeline

association presented lower values of growth rates for MAP compared to aerobically packaged samples. In general, aerobic storage of minced pork at all temperatures allowed the members of microbial association to reach higher population levels as can be concluded by the final population values (y_{end}) determined by the Baranyi and Roberts model, Table (1). With air packaging, Pseudomonas spp. were the dominant microorganism followed by Br. thermosphacta, whereas lactic acid bacteria and Enterobacteriaceae remained at lower levels. On the other hand, with MAP storage, lactic acid bacteria became the dominant bacteria throughout storage together with Br. thermosphacta. This is on line with the already existing data from previous studies dealing with the meat spoilage and the contribution of the ESO [3, 27] The estimated kinetic parameters by the Baranayi model are presented in Table 1. From Figure 4 it is clearly seen how the lag phase decreases significantly as temperature increases. Initial total viable counts ranged between 5.38 and 5.95 log cfu g^{-1} while the end measurements ranged between 9.65 and 9.84 for aerobic packages and 7.19 to 8.68 log cfu g^{-1} for

MAP packages indicating a clear difference in microbial growth in the two environments. Based on preprocessed spectra originating from the pure meat area, a large variation in the pixels is still found, however more subtle. By empirically looking at spectra as well as colour transformed images, it was found that 5 different types of meat was the optimal subdivision of meat types existing across the meat samples. These types of meat were basically very bright areas with high fat content, more pure meat which had been oxidized, areas which appeared very spoiled with abnormal meat colours and two intermediate types. A K-means algorithm [23] was used to identify the cluster centers shown in Figure 3 as normalized spectra. A certain cluster center is shown as a dashed line, which indicates areas of meat with a spoiled appearance. Carefully inspecting the characteristics of this spectrum reveals a higher response in the area of shorter wavelengths corresponding to blue, and a lower response from 600 to 800 nm resembling the reddish colours. Furthermore it then seems to shift in the high end of near infrared area, indicating a higher response here compared to the remaining four cluster centers. These characteristics resemble quite well an intuitive understanding of the appearance of spoiled meat as being less red and more green/blue.

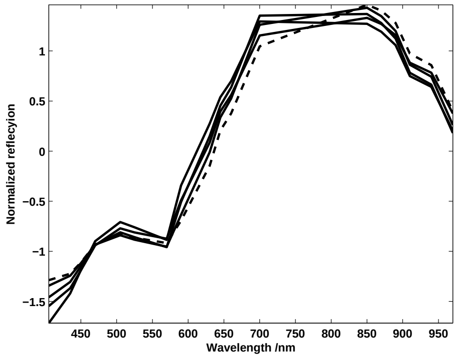


Figure 3: Spectral characteristics of the 5 meat types. The dotted line indicates the spectrum for spoiled meat types

Having normalized images as well as identified cluster centers, a mapping of each recorded spectrum in each picture may be done by calculating either true or approximated euclidean distances and assigning each pixel to its

Temp.	Pack.	lag phase(h)	y_0 (pred.)	y_{end} (pred.)	μ_{max}	se	r^2
0	AIR	74.138	5.951	9.680	0.045	0.332	0.947
5	AIR	21.667	5.627	9.640	0.093	0.255	0.974
10	AIR	2.306	5.398	9.347	0.146	0.369	0.933
15	AIR	1.767	5.372	9.513	0.205	0.404	0.930
20	AIR	3.813	5.881	9.429	0.325	0.327	0.938
0	MAP	6.067	5.579	7.207	0.008	0.258	0.830
5	MAP	0.000	5.652	7.780	0.027	0.246	0.902
10	MAP	7.893	5.464	7.949	0.082	0.251	0.929
15	MAP	1.789	5.503	8.264	0.126	0.266	0.928
20	MAP	0.000	5.362	8.326	0.218	0.366	0.877

Table 1: Estimated kinetic parameters of total viable counts (TVC) by the Baranyi model

closest cluster. This process creates a further segmented image, where it is possible to estimate the distribution of different types of meat surface in each image. This distribution may then directly or indirectly be used for e.g. classifying the image as being fresh, semi-fresh or spoiled, or for quantifying the amount of bacterial growth on the meat by predicting the total viable counts, which both are shown as the two output boxes in Figure 2. The actual classification of images and prediction of total viable count was done using a logistic regression model [23] and a partial least squares model [23], respectively.

Figure 5(a) shows a binary mask of the meat sample in Figure 5(b). The mask indicates the areas which represent only meat, i.e. no background or fat. This mask was created using a maximum autocorrelation transformation, which decomposes a multispectral image into similar tissue types, the first step in the spectral process line (Figure 2). This specific piece of meat is a very spoiled one having a total viable count of 9.83 log cfu g^{-1}. A graphical representation of the image after each pixel has been assigned to a cluster center is shown in Figure 5(c). The image contains 5 levels of graytones, each corresponding to a cluster center. The very dark area represents spoiled meat, while the very bright area represents less spoiled meat. For comparison a fresh piece of meat is seen in Figure 5(d) with TVC measured to 5.63 log cfu g^{-1}. A clear difference is seen as the meat area generally appears much brighter in the fresh sample; Figure 5(d), compared to the spoiled image in Figure 5(d). Thus the majority of pixels in the 2 images have been assigned to different clusters. The dark edges of the fresh piece of meat might be due

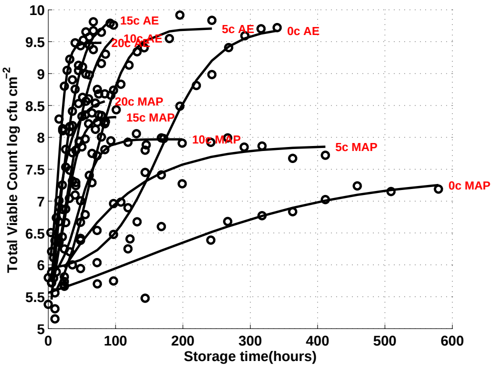
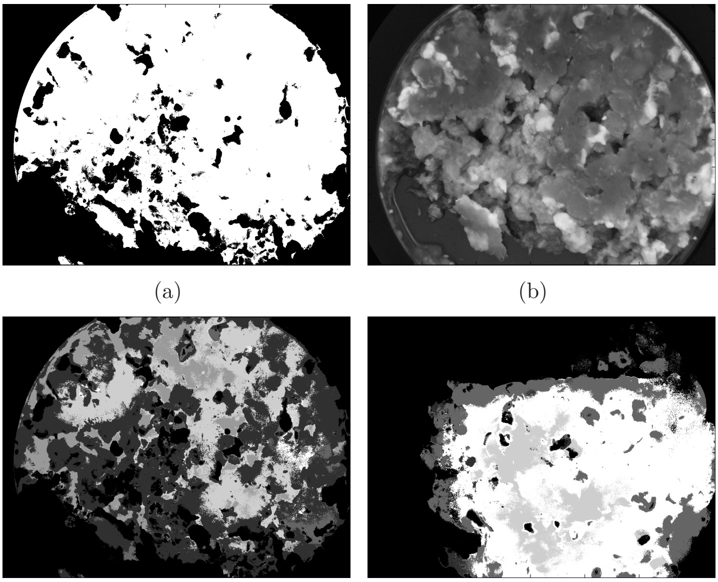


Figure 4: Changes in the population of total viable counts in pork stored in aerobic and modified atmosphere at 0, 5, 10, 15 and 20°C. The curve represents models fitted with the Baranyi bacterial growth model.

to an initial spoilage, but may also be due to scattering effects as well as shadow effects caused by the change in topology in these regions.

Being able to spatially determine areas of spoilage enables spatial inference on the images. This means e.g. it is possible to count how many pixels occurrences of each meat type there exist in the image, and estimate the total area percentage covered by a specific meat type. Doing this for all types of area on all images, and plotting it as a function of total viable count is seen in Figure 6. In Figure 6 the meat area for cluster 1 shows an increase in area as TVC increases while the area for cluster 5 shows a linear decrease as TVC increases. The area size for the intermediate clusters shows little or no development as TVC increases. The trends of cluster 1 and 5 are interpreted as a spoiled meat cluster and a fresh meat cluster. In order to extract the best features for a prediction model of TVC as well as for a classification model of sensory labels, the mean spectrum of the area of spoiled meat in all images is used as a feature space. A logistic regression model using the extracted features as covariates was used to classify the pork samples into three sensorial categories (fresh, semi-fresh, spoiled). For comparison a logistic regression was similarly used to classify the same samples into the same categories using the measured four microbial measurements as independent variables. Due to the small amount of samples, both classification models were assessed using leave one out cross validation in order to obtain realistic and generalizable



(a) (b)

(c) (d)

Figure 5: (a) shows a binary mask for a piece of spoiled piece of meat having TVC of 9.83 log cfu g^{-1}. A grayscale image of this same piece of meat is seen in (b), where the recording at 590nm. is shown. The binary mask indicates where in the image meat is located, which is used in further analysis. (c) shows the spatial distribution of how pixels have been mapped to identified clusters. There are a total of 5 clusters. Dark areas indicate spoiled meat, while lighter areas represent more fresh meat. (d) shows spatial distribution of clusters in a fresh portion of pork meat. The clusters are colour coded in same order as in (c), ie. light clusters being fresh meat. The difference in cluster colours in the two images is very clear.

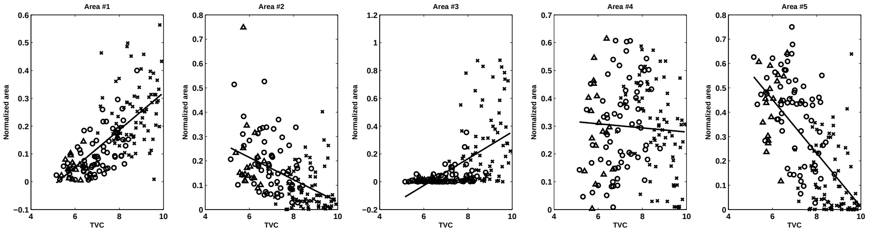


Figure 6: Meat area size as a function of total viable count with correlations 0.68, -0.49, 0.55, -0.058 and -0.67. Each plot represents a meat area assigned to a similar euclidean cluster as presented in Figure 8. Δ symbolizes fresh samples, o symbolizes semi-fresh and x symbolizes spoiled samples.

	Fresh	Semi-fresh	Spoiled	Row total	Sensitivity(%)
Fresh	10	7	1	18	55.5
Semi-fresh	7	46	11	64	71.8
Spoiled	1	10	62	73	84.9
Column total	18	63	74	155	
Specificity(%)	55.56	73.02	83.8		
Total classification rate: 76.13%, Cohens Kappa value: 0.598					

Table 2: Logistic regression on sensory labels using meat area estimations

results. The classification accuracy of the classification models is seen as confusion matrices together with their sensitivity, specificity and total correctly classified percentage. The performance of the predictability of TVC for each meat sample analyzed was determined by the bias and accuracy factors [28], the mean relative percentage residual, the mean absolute percentage residual [29], and finally by the root mean squared error of prediction and the standard error of prediction.

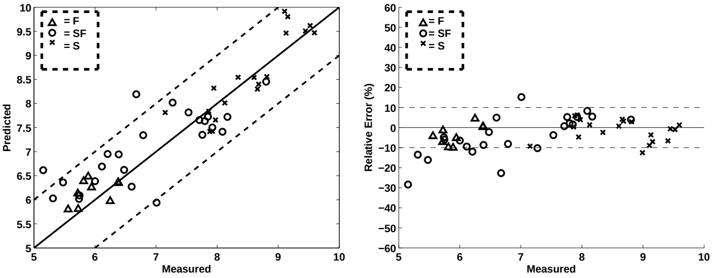
A classification rate of 73.6% for overall correct classification as seen in Table 2 was obtained. The sensitivity of the model was found to be 55.5%, 71.8% and 84.9% for fresh, semi-fresh and spoiled meat samples, respectively. This indicates fairly high uncertainty for the fresh samples, which are frequently identified as semi-fresh and vice versa. The sensitivity increases for semi-fresh and spoiled, which has several reasons. The sensory labels were based on other organoleptic senses than visual appearance, such as taste and odor which are very subjective factors. The amount of fresh sam-

	Fresh	Semi-fresh	Spoiled	Row total	Sensitivity(%)
Fresh	12	6	0	18	66.6
Semi-fresh	6	46	12	64	71.8
Spoiled	0	7	66	73	90.4
Column total	19	61	75	155	
Specificity(%)	63.2	75.4	88		
Total classification rate: 80%, Cohens Kappa value: 0.662					

Table 3: Logistic regression on sensory labels using microbial reference data

ples compared to the amount of semi-fresh and spoiled is very small, which gives a high probability that even a small subset of semi-fresh samples which overlaps fresh samples in feature space will affect the fresh class sensitivity significantly. For comparison, the actual microbial counts were likewise used for classifying the sensory labels, which is seen in Table 3. Slightly better results are seen for this set of predictors with sensitivity of 66.6%, 71.8% and 90.4%. Thus, the same problem of differentiating between fresh and semi-fresh classes exists for these predictors. Cohens Kappa value [30] was calculated to be 0.59 and 0.66, meaning that both values lie in the region of substantial agreement between predicted and observed classes. The prediction of total viable counts as seen in Figure 7(a) shows very small errors around the line of equality, $y = ax + b$, $a = 1$, $b = 0$, all within ±1 log unit area shown with stippled lines. However, three samples did fall outside the area of 1 log unit. A Partial Least Squares model was used for TVC prediction, which requires a separate training and test set for selecting model parameters and to evaluate the model. The selection of parameters was done using leave one out cross validation on 2/3 of randomly selected data from the entire set of 155 samples. The remaining 52 samples are used for validating the model. It is very crucial that the data which was used for selecting model parameters is not used for validating the model, as this clearly will give a biased and over-fitted model.

The predictive performance of the TVC model is furthermore presented in Figure 7(b) where the relative error in percent is plotted as a function of the observed microbial population. The errors are seen to be distributed equally around 0, with 83% of the predicted microbial counts included within the ±10% zone of relative errors. Further statistical values for the TVC prediction model are presented in Table 4. Ross [28] introduced Bias factor and Accuracy factor as interpretable indices for average deviation or the spread



(a) (b)

Figure 7: (a) Regression results for Total Viable Counts using a Partial Least Square Model trained using cross validation on a training set and validated on an unknown test set. The plotted values are predicted versus total viable counts. The lines show the line of equality i.e. a perfect fit, while the dashed lines represent ±1 log cfu unit. (b), Relative errors in % between observed and predicted total viable counts (TVC) during storage of pork meat at different temperatures, atmospheres and time-spans (F: fresh; SF: semi-fresh; S: spoiled meat samples).

of the results about the prediction. Perfect agreement between the predicted and observed values will lead to bias factors of 1, while bias factors above 1 indicates 'fail-dangerous' models and below 1, 'fail-safe' models. E.g. a value of 1.1 indicates that the predictions exceed the observations, on average, by 10%. Thus, the further away from 1, the more poor the model, i.e. a bias factor of 0.5 indicates a poor model. For the accuracy factor, the larger the value, the less accurate is the average estimate. For an accuracy factor of two, the prediction is on average a factor two different from the observed value - half as large or twice as large. The bias factor shows values slightly below 1, indicating a very small tendency of under estimation for all types of meat, including the overall bias factor. The accuracy factor furthermore shows that on average the predictions were ≈ 11.3%, ≈ 20.8% and ≈ 10.3% above the observed values for fresh, semi-fresh, and spoiled meat samples, respectively, while in total ≈ 15%. This is also confirmed by the mean absolute percentage error, representing the average deviation between observed and predicted counts. The mean relative percentage residual index confirmed the under-prediction for all classes since they are all below 0. The standard error of prediction index is a relative typical deviation of the mean prediction values and expresses the expected average error associated with future predictions. The model shows good predictive performance, i.e. below

	Expression	Fresh	Semi-fresh	Spoiled	All
Bias Factor	$10^{\sum \frac{\log(\widehat{y}/y)}{n}}$	0.923	0.920	0.973	0.940
Accuracy Factor	$10^{\sum \frac{\lvert\log(\widehat{y}/y)\rvert}{n}}$	1.114	1.209	1.103	1.152
MRPE	$\frac{1}{n}\sum \frac{100\cdot(y-\widehat{y})}{y}$	-3.662	-4.153	-1.331	-3.005
MAPE	$\frac{1}{n}\sum \frac{100\cdot\lvert y-\widehat{y}\rvert}{y}$	4.821	8.563	4.348	6.337
RMSEP	$\sqrt{\frac{1}{n}\sum(y-\widehat{y})^2}$	0.348	0.666	0.466	0.551
SEP	$\frac{100}{\bar{y}}\sqrt{\frac{1}{n}\sum(y-\widehat{y})^2}$	5.842	9.716	5.410	7.478

Table 4: Key values for the predictive performance of the TVC model for all classes, including overall predictive performance. The abbreviations are; MRPE: Mean relative Percentage Error, MAPE: Mean Absolute Percentage Error, RMSEP: Root Mean Square Error of Prediction, SEP: Standard Error of Prediction

10% standard error of prediction for all three classes, although especially for spoiled and fresh samples with only a percentage standard error of 5.4% and 5.8%.

4. Conclusion

In this article a large experiment regarding quality of stored pork meat has been performed in order to assess the ability of a rapid multipectral imaging device to quantify sensory labels as well as microbial count. The multispectral imaging device recorded spectra in the visible and beginning of the near infrared area. Recordings were used to investigate microbiology on pork meat stored in two different types of atmospheres under five different temperatures (0, 5, 10, 15 and $20°C$) in varying timespans up to almost 600 hours. Various machine learning and vision techniques were used to analyse the multispectral images. Features were extracted to evaluate the spoilage of the meat by predicting Total Viable Counts as well as classifying meat pieces into one of three classes; fresh, semi-fresh or spoiled, with ground truth being set by a sensory panel. For the multispectral images, an overall classification performance of 76.13% was achieved. For the microbial counts an overall classification performance of 80% was achieved. Thus, considering the fact that the electromagnetic area was sampled in only 18 distinct areas in mainly the visible region, a classification error of 76.13% is a relatively good performance A very good discrimination between spoiled and fresh pieces of meat was achieved while semi-fresh meat caused bigger problems. It should be noted that due to the spoilage nature of pork meat, there was an unequal distribution of samples across the three spoilage classes, with a significantly larger amount of samples in the spoiled class than in the fresh class. A set of seven images were manually removed as outliers due to their very large residual variation. The problem of correctly classifying semi-fresh pork meat may have an explanation in the fact that actual definition of the boundaries between classes are rather fuzzy for the sensory panel. This might indicate that an automated method may perform better if based on quantification of quantitative units (TVC) rather then qualitative units (Sensory labels). Spoilage classes might then be set, based on knowledge about total viable counts, as also done in the past [31, 32]. The difference here is that a much cheaper and more rapid equipment, requiring no sample preparation has been used. Furthermore, as opposed to the previous works, additional parameters have been included in this work such as time, temperature and atmosphere. As indicated in Figures 7(a)-7(b) and Table 4 the prediction performance on an unknown test set yielded good results especially for fresh and spoiled samples (standard error of prediction under 6%). This could help laboratory workers speed up their work significantly by providing non-invasive, and low cost microbiological analyses [32]. The values of the bias factor were close to unity indicating good agreement between predictions and observations. The calculated values of the bias factor are within the range of 0.9 - 1 which are considered adequate [28]. Thus a setup asthe one presented in this paper could in the future be used to satisfactorily predict microbial counts as well as sensory labels of pork meat.

References

[1] M. Livingston, M. S. Brewer, J. Killifer, B. Bidner, F. McKeith, Shelf life characteristics of enhanced modified atmosphere packaged pork, Meat Science 68 (2004) 115 – 122.

[2] P. Singh, A. A. Wani, S. Saengerlaub, H.-C. Langowski, Understanding critical factors for the quality and shelf-life of map fresh meat: A review, Critical Reviews in Food Science and Nutrition 51 (February 2011) 146–177(32).

[3] G.-J. E. Nychas, P. N. Skandamis, C. C. Tassou, K. P. Koutsoumanis, Meat spoilage during distribution, Meat Science 78 (2008) 77–89.

[4] G.-J. Nychas, D. Marshall, J. Sofos, Food Microbiology Fundamentals and Frontiers 3rd Edition, Chapter 6; Meat Poultry and Seafood, Eds Doyle, Beuchat and Montville ASM Press, 2007.

[5] J. H. J. H. in't Veld, Microbial and biochemical spoilage of foods: an overview, International Journal of Food Microbiology 33 (1996) 1 – 18. Specific Spoilage Organisms.

[6] D. Ercolini, F. Russo, E. Torrieri, P. Masi, F. Villani, Changes in the spoilage-related microbiota of beef during refrigerated storage under different packaging conditions, Appl. Environ. Microbiol. 72 (2006) 4663–4671.

[7] R. H. Dainty, Chemical/biochemical detection of spoilage, International Journal of Food Microbiology 33 (1996) 19 – 33. Specific Spoilage Organisms.

[8] J. M. Carstensen, M. E. Hansen, N. K. Lassen, P. W. Hansen, Creating surface chemistry maps using multispectral vision technology, in:

B. Ersbøll, T. M. Jørgensen (Eds.), 9th Medical Image Computing and Computer Assisted Intervention (MICCAI) - Workshop on Biophotonics Imaging for Diagnostics and Treatment, Institute of Mathematical Modelling - Technical Report-2006-17, 2006.

[9] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, J. Frias, Hyperspectral imaging - an emerging process analytical tool for food quality and safety control, Trends in Food Science & Technology 18 (2007) 590 – 598.

[10] S. B. Daugaard, J. Adler-Nissen, J. M. Carstensen, New vision technology for multidimensional quality monitoring of continuous frying of meat, Food Control 21 (2010) 626 – 632.

[11] M. Taghizadeh, A. Gowen, P. Ward, C. P. O'Donnell, Use of hyperspectral imaging for evaluation of the shelf-life of fresh white button mushrooms (agaricus bisporus) stored in different packaging films, Innovative Food Science & Emerging Technologies 11 (2010) 423 – 431.

[12] T. N. Tran, R. Wehrens, L. M. Buydens, Clustering multispectral images: a tutorial, Chemometrics and Intelligent Laboratory Systems 77 (2005) 3–17.

[13] J. Baranyi, T. A. Roberts, A dynamic approach to predicting bacterial growth in food, International Journal of Food Microbiology 23 (1994) 277 – 294.

[14] A. Argyri, E. Panagou, P. Tarantilis, M. Polysiou, G.-J. Nychas, Rapid qualitative and quantitative detection of beef fillets spoilage based on fourier transform infrared spectroscopy data and artificial neural networks, Sensors and Actuators (2009).

[15] P. N. Skandamis, G.-J. E. Nychas, Preservation of fresh meat with active and modified atmosphere packaging conditions, International Journal of Food Microbiology 79 (2002) 35–45.

[16] M. S. Ammora, A. Argyri, G.-J. E. Nychas, Rapid monitoring of the spoilage of minced beef stored under conventionally and active packaging conditions using fourier transform infrared spectroscopy in tandem with chemometrics, Meat Science 81 (2009) 507–514.

[17] J. M. Carstensen, J. Folm-hansen, An apparatus and a method of recording an image of an object, 2000.

[18] J. F. Hansen, On chromatic and geometrical calibration, Ph.D. thesis, Department of Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 1999.

[19] B. S. Dissing, L. H. Clemmesen, H. Løje, B. K. Ersbøll, J. Adler-Nissen, Temporal reflectance changes in vegetables, In Proceedings of Institute of Electrical and Electronics Engineers (IEEE) Color and Reflectance in Imaging and Computer Vision Workshop, Kyoto, Japan. (2009).

[20] L. H. Clemmensen, M. E. Hansen, B. K. Ersbll, A comparison of dimension reduction methods with application to multi-spectral images of sand used in concrete, Machine Vision and Applications (2009).

[21] D. D. Gomez, L. H. Clemmensen, B. K. Ersbøll, J. M. Carstensen, Precise acquisition and unsupervised segmentation of multi-spectral images, Computer Vision and Image Understanding 106 (2007) 183–193.

[22] M. E. Hansen, B. K. Ersbøll, J. M. Carstensen, A. A. Nielsen, Estimation of critical parameters in concrete production using multispectral vision technology, in: Lecture Notes in Computer Science, LNCS3540, Lecture Notes in Computer Science, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005, pp. 1228–1237.

[23] T. Hastie, R. Tibshirani, J. Friedman, Elements of Statistical Learning: Data mining, Inference and Prediction, 2. edition, Springer-Verlag, 2009.

[24] A. A. Green, M. Berman, P. Switzer, M. D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal, IEEE Transactions on Geoscience and remote Sensing 26 (1988) 65–74.

[25] A. A. Nielsen, An extension to a filter implementation of a local quadratic surface for image noise estimation, in: 10th International Conference on Image Analysis and Processing, ICIAP'99., pp. 119–124.

[26] N. Otsu, A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man, and Cybernetics 9 (1979) 62–66.

[27] A. I. Doulgeraki, S. Paramithiotis, D. M. Kagkli, G.-J. E. Nychas, Lactic acid bacteria population dynamics during minced beef storage under aerobic or modified atmosphere packaging conditions, Food Microbiology 27 (2010) 1028 – 1034.

[28] T. Ross, Indices for performance evaluation of predictive models in food microbiology, Journal of Applied Microbiology 81 (1996) 501–508.

[29] A. Palanichamy, D. Jayas, R. Holley, Predicting survival of escherichia coli o157:h7 in dry fermented sausage using artificial neural networks., Journal of Food Protection 71 (2008) 6–12.

[30] J. Cohen, A Coefficient of Agreement for Nominal Scales, Educational and Psychological Measurement 20 (1960) 37.

[31] D. I. Ellis, R. Goodacre, Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends, Trends in Food Science & Technology 12 (2001) 414 – 424.

[32] D. I. Ellis, D. Broadhurst, D. B. Kell, J. J. Rowland, R. Goodacre, Rapid and quantitative detection of the microbial spoilage of meat by fourier transform infrared spectroscopy and machine learning., Appl Environ Microbiol 68 (2002) 2822–2828.

Appendix F

Multispectral imaging for determination of astaxanthin concentration in salmonids

Accepted in Journal of Public Library of Science (PLoS ONE)

Multispectral imaging for determination of astaxanthin concentration in salmonids

Bjørn S. Dissing[1*], Michael E. Nielsen[2], Bjarne K. Ersbøll[1] and Stina Frosch[2*]
1 Department of Informatics and Mathematical Modeling, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark
2 National Food Institute, Division of industrial food technology, Technical University of Denmark, DK-2800, Kgs. Lyngby, Denmark
∗ E-mail: bdi@imm.dtu.dk, sfr@aqua.dtu.dk

Abstract

Multispectral imaging has been evaluated for characterization of the concentration of a specific cartenoid pigment; astaxanthin. 59 fillets of rainbow trout, *Oncorhynchus mykiss*, were filleted and imaged using a rapid multispectral imaging device for quantitative analysis. The multispectral imaging device captures reflection properties in 19 distinct wavelength bands, prior to determination of the true concentration of astaxanthin. The samples ranged from 0.20 to 4.34 μg per g fish. A PLSR model was calibrated to predict astaxanthin concentration from novel images, and showed good results with a RMSEP of 0.27. For comparison a similar model were built for normal color images, which yielded a RMSEP of 0.45. The acquisition speed of the multispectral imaging system and the accuracy of the PLSR model obtained suggest this method as a promising technique for rapid in-line estimation of astaxanthin concentration in rainbow trout fillets.

Introduction

Color is a highly important quality parameter in relation to the commercial production of salmonid fishes. The consumers associate increased intensity of red in salmonid fishes with superior quality, being fresher and having a better flavor [1, 2]. As the change in surface color is the first quality parameter evaluated by the consumer, it is of great economic importance that the color of the salmonid fishes meets consumer preferences. The color of salmonid fishes is caused by deposition of cartenoid pigments in the muscular tissue. Besides being essential for reproduction, proper growth and survival of the fish, carotenoids, primarily astaxanthin and castaxanthin, are also important for the red color in salmonids. As fish cannot synthesize carotenoids de novo their intake rely on the content of cartenoids in the feed. Wild salmonids obtain the cartonids from intake of e.g. crusteceans, krill and other sources rich in carotenids whereas carotenoids primarily astaxanthin is added to the feed of farmed salmonids. The primary use of astaxanthin within aquaculture is as a feed additive to ensure that farmed salmon and trout achieve a coloration that comply with the consumers preferences.

Astaxanthin is the single most expensive constituent in salmonid fish feed. Even though astaxanthin constitutes less than 20% of the total fish feed costs, control and optimization of the concentration of astaxanthin from feed to fish is of paramount importance for a cost effective salmonid fish production. Traditionally, astaxanthin content in fish is determined by spectrophotometric analysis or high-performance liquid chromatography (HPLC) analysis. In both methods astaxanthin is extracted from the minced sample into a suitable solvent such as acetone or hexane before further analysis. U.S. Food and Drug Administration (21 CFR 73.185) and Canadian Food Inspection Agency (Registration no. 990535) have accepted the method based on HPLC analysis for determining astaxanthin content of a product. Both methods have several drawbacks. First, the method based on spectrophotometric analysis overestimate the astaxanthin as other compounds such as lutein, canthaxanthin and astacene are falsely included. This means they absorb light at the same wavelength as astaxanthin and thereby increase the signal. Second, both methods are time consuming, labor demanding and sample destructive.

For quality assessment of salmonid color there are two widely accepted color standards in the salmonid industry, which are used by quality inspectors in their visual assessment of fillets, the SalmoFan$^{\text{TM}}$card and the SalmoCard$^{\text{TM}}$(Hoffmann-La Roche Basel, Switzerland). Both methods enable an inspector to score the color of a salmonid fillet into one of 15 red color categories ranging from 20 (pale red) to 34 (dark red). This method has the advantage of being a very straight forward, intuitive and cheap. It is easily applicable and does require intensive expert training. In spite of these advantages there may be reasons to inspect the color quality of fish fillets using other methods. A human operator is required in order to use the SalmoFan / SalmoCard, which means such a color-evaluation will be subject to operator bias and fatigue while also being time-consuming, costly and relatively labor-expensive.

Other instruments previously used for color evaluation are tricolorimeters, spectrophotometers and standard trichromatic charged coupled device cameras. These devices probe the visual spectrum in order to in some sense imitate human visual perception and objectively quantify colors. A Colorimeter (e.g. Minolta Chroma Meter II-CR200, Hunterlab Miniscan) makes use of a stable light source such as Xenon to illuminate a small surface patch of roughly $1cm^2$, and measures the reflection of the surface in this area. The reflection is then integrated according to the CIE-XYZ [3] tristimulus curves and transformed to the uniform L*, a*, b* color space [3]. The L*, a*, b* color space is a three dimensional color space, where L* represents the lightness of the color (100 being diffuse white), a* the mix of red and green and b* the mix of yellow and blue. Examples of studies where a colorimeter was used in conjunction with studies of fish color include [4–6] where the latter established that the intensity of redness (a*) increases with the carotenoid content in the raw flesh of Atlantic salmon, while lightness (L*) decreases and yellowness (b*) remains unaffected. While colorimeters acquire very accurate colors, they do not contain any spatial information, and therefore no information on surface texture and structure / shape. On the other hand chromatic images measure a larger spatial area of reflected photons and thereby provide color as well as spatial information. A review of vision technology and color cameras in the food industry may be found in [7]. The actual color evaluation ability of a trichromatic camera in regard to fish quality inspection was investigated in [8], where comparisons between trichromatic camera images and SalmoFan evaluations were performed on fillets of the Atlantic salmon (Salmo salar). The comparison was based on measurements from five different locations, more or less uniformly spatially distributed across the fillet surface. Here the authors found that there was no significant statistical difference between SalmoFan and camera-based evaluations. Similar experiments were performed in [9], where the authors found a correlation of 0.95 between sensory panel SalmoFan evaluations and computer vision based color evaluations.

Current state-of-the art vision systems for quality and process control in the fish processing industries are typically based on traditional trichromatic (Red Green Blue) imaging. In this study we are interested in going one step further, by quantifying the astaxanthin content and thereby indirectly also the color of the fillet. The relative presence of some wavelengths and absence of others is a specific characteristic of many material properties. Consequently the adaption of multispectral imaging technology can reveal relevant information and measurement of more biological quality parameters such as fat, astaxanthin and cartilage content, simultaneously. A multispectral image may also be referred to as a surface chemistry map [10] where a set of neighboring spectra are recorded, revealing information about the surface chemistry to a larger degree than in a trichromatic image. Thus, multispectral imaging is well suited for applications where it is crucial to detect small differences in texture, color and surface chemistry [11–15]. It is expected that vision systems based on multispectral imaging will be employed to a much larger extent in the near future [16–20]. Aquaculture and the fish processing industries are areas where the added information in a multispectral image can be exploited to improve the general quality and/or reduce the production cost. In this study we investigate the use of multispectral images for estimating natural astaxanthin concentration in rainbow trout fillets. In order to justify the use of multispectral images we calibrate and compare multivariate models for multispectral as well as traditional color images of trout fillets to predict astaxanthin content. Furthermore we illustrate shortly how the calibrated model can be used to predict all spectra in an image in a pixel-wise manner, in order to visualize the predicted spatial astaxanthin distribution within the fish fillet.

Materials and Methods

Sample preparation

The Rainbow trout (*Oncorhynchus mykiss*) were from the organic farm at Bisserup Havbrug and harvested in November 2009. The fish were fed with commercial organic trout feed of approximately 1.5% of body weight per day throughout the entire rearing in accordance with commercial practise. According to legislation, the fish feed were coated with natural astaxanthin [21]. The fish were slaughtered at 2 years of age, with an average weight of 1.1 kg. The fish were filleted and trimmed by hand at Bisserup Havbrug the day after slaughtering and transported to the Technical University of Denmark on ice at the same day. The fillets were stored overnight on ice in a $2°C$ chill room. The fillets were then subsequently cut into three pieces (Figure 1). The middle piece was used as the experimental sample in the further analysis. The samples were placed in plastic petri dishes (90 mm diameter) and stored on ice in styrofoam boxes. Multispectral images of the samples were captured 30 minutes after placement in the styrofoam box. Right after image capture each sample was minced and subsequently frozen at $-40°C$. After 14 days of storage at $-40°C$ the astaxanthin concentrations were determined using chemical extraction.

Chemical determination of astaxanthin content

Astaxanthin content of the minced fillets was determined in duplicate from the lipid extracts of the fish meat using an Agilent 1100 series HPLC (Agilent Technologies, Palo Alto, CA), equipped with a UV diode array detector. The fillet sample were minced, and 10 g in duplicates was used for extraction using chloroform and methanol according to the modified protocol of Bligh and Dyer [22]. A fraction of the lipid extract was evaporated under nitrogen and redissolved in 2mL of n-heptane before injection.Astaxanthin content was determined after injection of an aliquot (50 μL) of the n-heptane fraction onto a LiChrosorb Si60-5 column (100 mm x 3 mm, 5 μm) equipped with a Cromsep Silica (S2) guard column (10 mm x 2 mm; Chrompack, Middelburg, The Netherlands) and eluted with a flow of 1.2 mL min-1 using n-heptane/acetone (86:14, v/v) and detection at 470 nm. Concentrations of astaxanthin were calculated using authentic standards from Dr. Ehrenstprfer GmbH (Augsburg, Germany).

Reflection characteristics of astaxanthin

The reflection properties of natural astaxanthin [22] in a solution of fishoil was recorded by a NIRSystems 6500 absorption spectrometer and transformed to reflection values using the standard relation $A = -log(R)$, where A is absorption values and R is the reflection values.

Multispectral Imaging System

Data acquisition was done using a VideometerLab [10], which obtains multi-spectral images at 19 different wavelengths ranging from 385 to 970 nm, fully shown in Table 1. The acquisition system records surface reflections with a standard monochrome charged coupled device chip, nested in a Point Grey Scorpion camera. Figure 2 shows the principal setup of the system where the object of interest is placed inside an integrating or so called Ulbricht sphere, with a matte white coating. The coating, together with the curvature of the sphere, ensures a uniform reflection of the cast light and thereby a uniform light in the entire sphere. At the rim of the sphere Light Emitting Diodes (LED) are positioned side by side in a pattern which distributes the LEDs belonging to each wavelength uniformly around the entire rim. The system is first calibrated radiometrically using both a diffuse white and dark target followed by a light setup based on the type of object to be recorded. The system is geometrically calibrated with a geometric target to ensure pixel correspondence for all spectral bands [23]. The homogeneous diffuse light, together with the calibration steps, ensures an optimal dynamic range and minimizes shadows and shading effects as well as specular reflection and gloss-related effects. The system has been developed to guarantee the reproducibility of collected images which means it can be used in comparative studies of time series or across a large variety of different samples [24–27].

Color Images

The advantage of going from color to multispectral images is illustrated by comparing models calibrated using either of the two types of images. To be able to compare results from the two models we have transformed the multispectral images to RGB images. In this paper we used a spectral reconstruction technique [28] in order to estimate the reflectance spectrum in each pixel with 5 nm spacing. Each spectrum was then integrated over the entire recorded spectral range in 3 different intervals, according to a CIE 1931 2° Standard Observer [3]. The resulting color images were then transformed to standard RGB images using a transformation formula described by Wyszecki, G. and Stiles [3].

Image segmentation and data extraction

Segmenting images into distinct regions is a very important preprocessing step in image analysis before further analyzing the images. Having a specific region representing only the area of the image which should be analyzed is called a region of interest (ROI). Segmentation of images may be done in a large variety of ways, where we in this work made use of statistical orthogonal methods or so called decomposition techniques to highlight desired features for easy extraction. Specifically we have used a Maximum Noise Fraction (MNF) [29] transformation to remove the image background material (petri dish and cardboard under the petri dish). Canonical Discriminant Analysis (CDA) [30] was then used to remove areas assumed to be fat and collagen. The decomposed result with desired features highlighted was then segmented easily using an adaptive thresholding technique known as Otsus adaptive thresholding method [31]. Having segmented the image into a ROI, the image could be transformed to a spectrum based on a mean calculation. Thus each image contributed with a single spectrum for the model calibration.

Data Analysis

Partial least square regression (PLSR) [32] was used to estimate calibration models between the extracted spectra and reference values (chemically determined) using LOOCV; a total of 59 samples were included in the analysis; 20 for training and 39 for validation. Models were calibrated using leave one out cross validation (LOOCV) [33] using a training set and validated using a testset. The quality of the models was determined based on the coefficient of determination (R^2), the prediction error expressed as the root mean square error of prediction (RMSEP) and the standard error of the fit. Spectra were centered by substracting the mean calculated from each wavelength followed by a scaling with the standard deviation calculated from each wavelength, commonly known as autoscaling or standardization [34]. More formally this is calculated as

$$\begin{aligned}
\mu_{X_j} &= \frac{1}{N}\sum_{i=1}^{N} X_i, \quad j = 1,2,..,p \\
\sigma_{X_j} &= \frac{1}{N-1}\sum_{i=1}^{N} (X_i - \mu_{X_j})^2 \\
X_{std_{ij}} &= \frac{X_{ij} - \mu_{X_j}}{\sigma_{X_j}}, \quad i = 1,2,..,N \wedge j = 1,2,..,p
\end{aligned} \tag{1}$$

X is here a matrix of size Nxp containing all the spectra in the calibration, where N is the number of samples and p is the number of sampled wavelengths. A pixel-wise astaxanthin prediction of the images was done. The loadings of the astaxanthin PLS model (as described above) were used to project acquired preprocessed images into a subspace highlighting the distribution of the astaxanthin contentbased on electromagnetic reflection properties in each fillet. Principal Component Analysis (PCA) was used for visualizing trends in the multivariate dataset, and identifying outliers. All extraction, image analysis routines, color transformations, pixel based predictions and calibration analyses were programmed in Matlab 7.8 (The Mathworks Inc., Natick, MA, USA).

Results and Discussion

Reflection characteristics of astaxanthin

The reflection spectrum recorded by the NIRSystems 6500 instrument, seen in Figure 3, shows large reflection properties starting from around 600 nm as well as large absorption properties from around 400 nm to 600 nm. This corresponds to having high absorption in the cyan, green and yellow area while the red and blue area is highly reflected, giving astaxanthin its characteristic dark red/purple color. The present measurements are well in accordance with previous absorption measurements of astaxanthin [35].

Reference data

The results from the chemical measurements of the astaxanthin content are presented in Figure 4. The astaxanthin concentration in the samples ranges from 0.2 to 4.34 μg per g fish with a mean of 1.69 μg per g fish and a standard deviation of 0.95 μg per g fish.

Acquired images and segmentation results

An example of a recorded multispectral image is presented in Figure 5 with the channels listed according to their wavelength number in Table 1. It is clearly shown that the general brightness of the image increases as the wavelength increases and that some features are more pronounced at certain wavelengths that others e.g. fat and collagen in meat structure are primarily pronounced at low wavelengths (395 nm to 570 nm). An example of the final segmentation results is shown in Figure 6, where the mask indicates the segmentation of fillet from background the using the MAF transformation and fat the CDA transformation. The MAF transformation is a contrast between the extreme bands, ultra blue (385 nm) and NIR (970 nm), and the middle color bands in the blue / green area of the visible spectrum (630-700 nm). The segmentation of meat based on CDA transformation relies primarily on the blue part of the spectrum (430-470 nm).

Extracted spectra

Figure 7 shows the mean spectrum of each of the 59 recorded trout fillets. A general scaling difference is seen in the spectra which have been removed using autoscaling preprocessing, thereby highlighting the nonlinear differences. The scaling of a spectrum in the visible range is in general an expression for the brightness of the sample, which means some fish among the samples set appear brighter than other. A clear difference in the intensity of the spectra, which becomes very apparent after treating with autoscaling in Figure 8, is seen in the area around 450 to 525 nm. This area corresponds quite well to a known absorption area of astaxanthin, which also is seen in Figure 3. Further a deviation in the spectra is seen in the NIR area and below 435 nm.

Calibration models

Figure 9 shows a score plot (PC1 versus PC2) of a PCA of the entire dataset after autoscaling. The plot shows a clear trend in the first component describing 80% of the variation with few outliers. Based on outlier diagnostics seven samples were categorized as outliers and removed from the data set prior to further analysis. All outliers were characterized by bad filleting. Table 2 shows the results of the final PLSR model for astaxanthin prediction based on a multispectral image. The reported RMSEP of 0.26 from a 7 component PLSR model is based on an independent test set while the model itself was cross validated on a trainingset. The cross validation showed a minimum generalization error when using 7 PLS components (Figure 10), which together with a total variance description on the response variable of 91% led to the choice of 7 components in the model. The variance decription percentages for the 7 components in the response variable were 48, 69, 75, 83, 86, 89 and 91% which shows that the performance of the model drastically increases with the first 4 components. The loadings for these components are shown in Figure 11. The first 2 components clearly show high response in the area around the absorption peak of astaxanthin in the blue/green area of the visual spectrum. The components naturally reflect the areas of largest variation in the preprocessed spectra shown in Figure 8. Table 2 also contain a PLSR model fitted on the same data after transformation to sRGB images. This means only a total of 3 variables exists for the regression problem which could basically be handled using a full multiple linear regression. This was tested together with a PLS model, which was found to yield equal results. The RGB model in Table 2 is seen to have a higher RMSEP value, indicating reduced prediction abilities than the multispectral mode. Among the two models the multispectral models has best variance description with an R^2 value of 0.86 versus and R^2 value of 0.66 for the RGB model. Furthermore the variance in the residuals is seen to be smaller for the multispectral model with a standard error of 0.02 against 0.05 for the RGB model. Previously similar techniques has been investigated using VIS spectroscopy and digital photography in [36], for calibration against chemically measured astaxathin. VIS spectroscopy was here found to have a correlation coefficient of 0.92 with a RMSEP of 0.42. For digital photography a correlation of 0.92 was found together with an RMSEP of 0.41. They reported a cross validation error based on all samples, meaning this error was used to chose the correct number of components. Compared to our results, we managed to achieve an RMSEP of 0.27 on an independent testset, to get a truly unbiased model, for the multispectral images.

The calibrated PLS model makes it possible to predict the astaxanthin concentration in each pixel of the image - a so called pixel-wise prediction. The pixel-wise prediction can therefore be used to estimate the astaxanthin distribution in the fillet. An illustration of this is seen in Figure 12, where a multispectral image of a trout piece is projected pixel-wise, in order to get an impression of the spatial distribution of the astaxanthin concentration. The pixel-wise prediction is color-coded according to the amount of astaxanthin predicted in each pixel, so that pixels with high values of astaxanthin appears red, while low value astaxanthin pixels appear blue. The projected image clearly shows that the upper part of the fillet contain the largest concentration of astaxanthin. This technique is well suited for visualization purposes. However, since the PLS model in this study was based on the preprocessed mean spectra from the entire salmonid pieces, the accuracy of this pixel-wise predictions remains to be validated properly in a further study.

Conclusion

In this paper an experiment was conducted to examine the possibilities of using multispectral imaging to assess the concentration of cartenoids, with focus on astaxanthin, in rainbow trout fillets. The recorded images ranged spectrally throughout the visible area and up into the first part of the near infra red area. The astaxanthin concentration of the investigated fillets ranged from 0.2 to 4.34 μg per g fish with a mean of 1.69 μg per g fish. A total of 7 images were classified as outliers using PCA scoreplots for identification. A PLSR model was calibrated based on mean values of spectral value in a region of interest in the image. A training set was used for model training in a leave one out cross validation scheme, while a separate test set was used to evaluate the model in terms of RMSEP and and R^2. The result was compared to a similar model based on color images extracted from the multispectral images, in order to motivate the use of multispectral images in a study like this. As a consequence of offering more spectral information about the sample, it is possible to gain more knowledge about which area of the spectrum yields the information we are interested in. The RMSEP obtained from the test set was 0.27 for the multispectral images and 0.45 for the color images, showing a somewhat higher prediction certainty for the multispectral images. Furthermore, the goodness of fit (R^2) was similarly somewhat better for the multispectral model, being 0.86. The most significant components of the PLSR model revealed that the area between 470 and 525 nm. carried the largest amount of variation, which corresponds very well with absorption peaks of pure astaxanthin being in the vicinity of 450 to 600 nm. In conclusion, the current study has shown that multispectral imaging is a promising method for rapid analysis of the astaxanthin concentration of rainbow trout, and thereby a qualified candidate for replacement of ordinary laborious and destructive sampling of astaxanthin for concentration prediction.

Acknowledgments

The expert technical assistance of Ms Heidi Olander Petersen is gratefully acknowledged.

References

1. Gormley TR (1992) A note on consumer preference of smoked salmon colour. Irish Journal of Agricultural and Food Research 31: 199–202.
2. Anderson S (2000) Salmon color and the consumer. In: Proceedings of the Tenth Biennial Conference of the International Institute of Fisheries Economics and Trade Presentations. Corvallis, Oregon USA.
3. Wyszecki G, Stiles W (2000) Color science: Concepts and Methods, Quantitative Data and Formulae, second edition. Wiley & sons Inc.
4. Skrede G, Risvik E, Huber M, Enersen G, Blumlein L (2006) Developing a color card for raw flesh of astaxanthin-fed salmon. Journal of Food Science 55: 361–636.
5. Skjervold PO, Rørå AMB, Fjæra SO, Vegusdal A, Vorre A, et al. (2001) Effects of pre-, in-, or post-rigor filleting of live chilled atlantic salmon. Aquaculture 194: 315 - 326.
6. Skrede G, Storebakken T (1986) Instrumental colour analysis of farmed and wild atlantic salmon when raw, baked and smoked. Aquaculture Vol.53 Issue.3-4: 279-286.
7. Brosnan T, Sun DW (2004) Improving quality inspection of food products by computer vision - a review. Journal of Food Engineering 61: 3-16.
8. Misimi E, Mathiassen J, Erikson U (2007) Computer vision-based sorting of atlantic salmon (salmo salar) fillets according to their color level. Journal of Food Science 72.
9. Quevedo RA, Aguilera JM, Pedreschi F (2008) Color of salmon fillets by computer vision and sensory panel. Food and Bioprocess Technology .
10. Carstensen JM, Hansen ME, Lassen NK, Hansen PW (2006) Creating surface chemistry maps using multispectral vision technology. In: Ersbøll B, Jørgensen TM, editors, 9th Medical Image Computing and Computer Assisted Intervention (MICCAI) - Workshop on Biophotonics Imaging for Diagnostics and Treatment. Institute of Mathematical Modelling - Technical Report-2006-17.

11. Hansen ME (2005) Indexing and Analysis of Fungal Phenotypes Using Morphology and Spectrometry. Ph.D. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby.
12. Maletti GM (2003) Novelty detection in dermatological images. Ph.D. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby.
13. Gomez DD (2005) Development of an image based system to objectively score the severity of psoriasis. Ph.D. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby.
14. Clemmensen LH, Ersbøll BK (2006) Multispectral recordings and analysis of psoriasis lesions. In: MICCAI '06 - Workshop on Biophotonics Imaging for Diagnostics and Treatment. 9th MICCAI Conference, pp. 15-18.
15. Hansen ME, Lund F, Carstensen JM (2003) Visual clone identification of penicillium commune isolates. Journal of Microbiological Methods 52: 221-229.
16. Sivertsen AH, Heia K, Nilsen H (2008) Detection of parasites in fish developing an industrial solution. Infofish International 3.
17. Sivertsen AH, Chu CK, Wang LC, Godtliebsen F, Heia K, et al. (2009) Ridge detection with application to automatic fish fillet inspection. Journal of Food Engineering 90: 317-324.
18. Heia K, Sivertsen AH (2008) Automatic cod fillet inspection by imaging spectroscopy. In: 38th West European Fish technologists Association Annual Meeting, Florence, Italy.
19. Heia K, Sivertsen AH (2008) Fish freshness assessment as part of an automatic fillet inspection. In: 38th West European Fish technologists Association Annual Meeting, Florence, Italy.
20. Heia K, Sivertsen AH, Stormo SK, Elvevoll E, Wold JP, et al. (2007) Detection of nematodes in cod (gadus morhua) fillets by imaging spectroscopy. Journal of Food Science 72: E011 - E015.
21. Nielsen ME, Mikkelsen H, Nielsen LB, Joensen O (2005) By-product based production of natural astaxanthin (nax). In: 7th Joint meeting: 50th Annual Atlantic Fisheries Technology Conference and 29th Annual Seafood Science and Technology Society of the Americas.
22. Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Canadian journal of biochemistry and physiology 37: 911-917.
23. Folm-Hansen J (1999) On chromatic and geometrical calibration. Ph.D. thesis, Department of Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby. URL `http://www2.imm.dtu.dk/pubdb/p.php?1210`.
24. Dissing BS, Clemmesen LH, Løje H, Ersbøll BK, Adler-Nissen J (2009) Temporal reflectance changes in vegetables. In Proceedings of Institute of Electrical and Electronics Engineers (IEEE) Color and Reflectance in Imaging and Computer Vision Workshop, Kyoto, Japan .
25. Clemmensen LH, Hansen ME, Ersbøll BK (2009) A comparison of dimension reduction methods with application to multi-spectral images of sand used in concrete. Machine Vision and Applications .
26. Gomez DD, Clemmensen LH, Ersbøll BK, Carstensen JM (2007) Precise acquisition and unsupervised segmentation of multi-spectral images. Computer Vision and Image Understanding 106: 183–193.

27. Hansen ME, Ersbøll BK, Carstensen JM, Nielsen AA (2005) Estimation of critical parameters in concrete production using multispectral vision technology. In: Lecture Notes in Computer Science, LNCS3540. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby: Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Lecture Notes in Computer Science, pp. 1228-1237. URL `http://www2.imm.dtu.dk/pubdb/p.php?3653`.

28. Zhao Y, L A Taplin MN, Berns RS (2005) Using the matrix r method for spectral image archives. In: 10th Congress of the International Colour Association.

29. Green AA, Berman M, Switzer P, Craig MD (1988) A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and remote Sensing 26: 65-74.

30. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Annals of Eugenics 7: 179–188.

31. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9: 62-66.

32. Brown SD (1995) Chemical systems under indirect observation: Latent properties and chemometrics. Appl Spectrosc 49: 14A–31A.

33. Stone M (1977) Asymptotics for and against cross-validation. Biometrika 64: 29-35.

34. Hastie T, Tibshirani R, Friedman J (2009) Elements of Statistical Learning: Data mining, Inference and Prediction, 2. edition. Springer-Verlag.

35. Buchwald M, Jencks WP (1968) Optical properties of astaxanthin solutions and aggregates. Biochemistry 7: 834-843.

36. Folkestad A, Wold JP, Rørvik K, Tschudi J, Haugholt K, et al. (2008) Rapid and non-invasive measurements of fat and pigment concentrations in live and slaughtered atlantic salmon (salmo salar l.). Aquaculture 280: 129-135.

Figure Legends

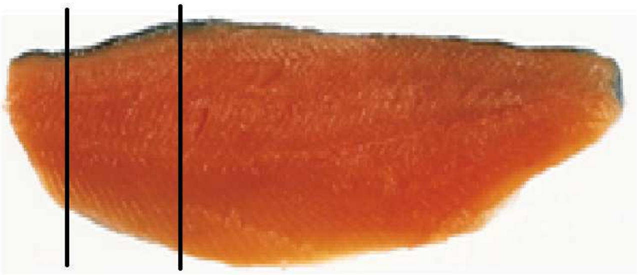


Figure 1. Trout fillet. This image shows how the fillets were cut in order to fit under the camera. The middle piece is used in further analysis.

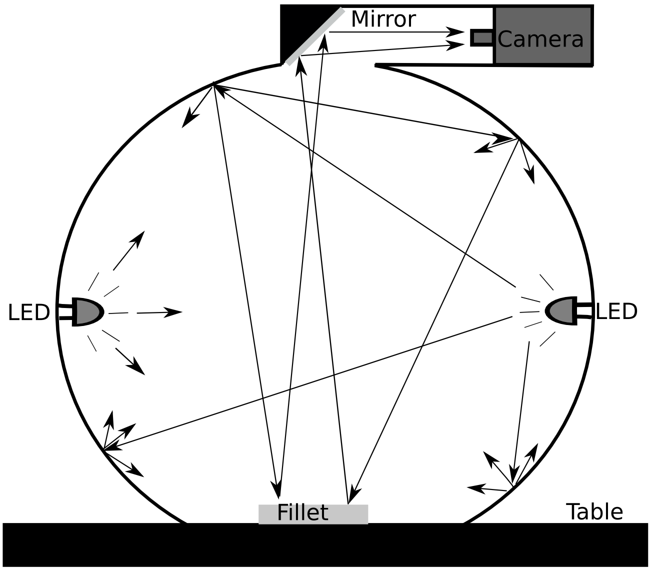


Figure 2. Principal setup of the multispectral system. An integrating sphere coated with a matte white coating ensures optimal lighting conditions. In the rim of the sphere a set of narrow band light emitting diodes ranging from 395 to 970 nm. are mounted. The image acquisition is performed by a monochrome grayscale CCD camera mounted in the top of the sphere. The arrows illustrate how the light is distributed inside the sphere to uniformly illuminate the fillet.

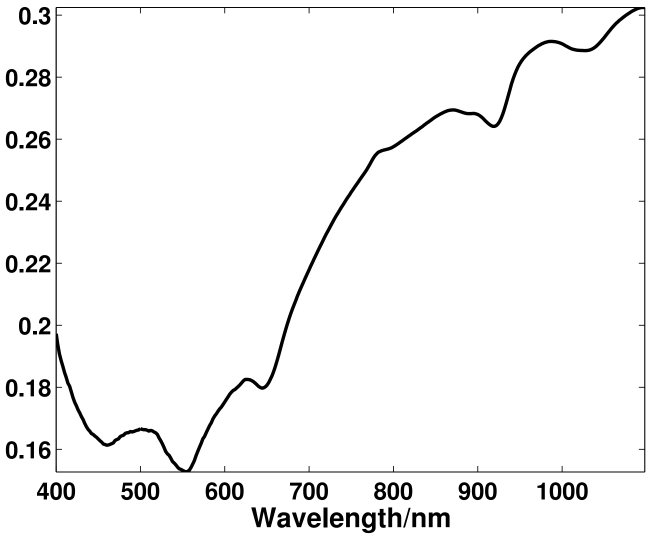


Figure 3. Reflection properties of astaxanthin. These reflection properties has been recorded using an absorbance spectrometer and transformed to reflection properties. The axes shows amount of light reflected as a function of wavelength.

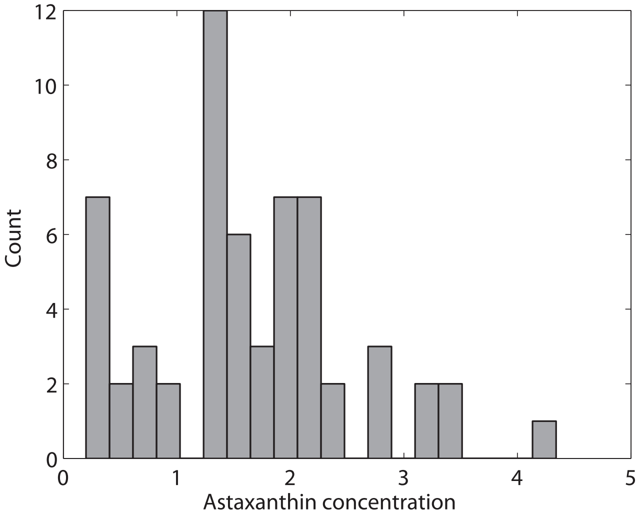


Figure 4. Distribution of measured astaxanthin. A histogram of the reference data shows the sample count as a function of μg astaxanthin per fish, revealing a high number of observations around 2μg.

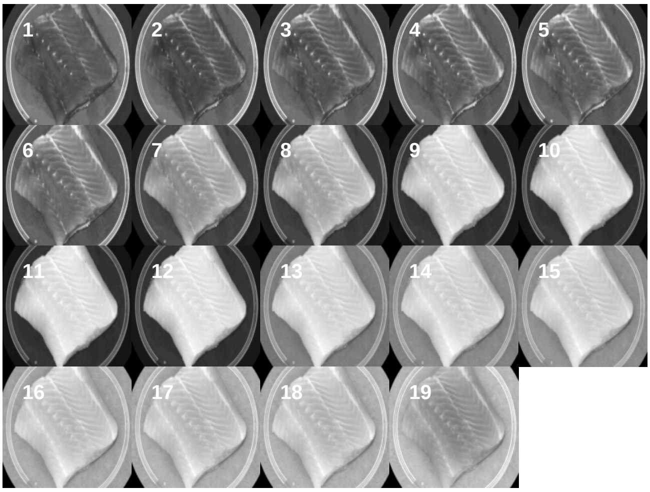


Figure 5. Channels ranging from 395 nanometer to 970 nanometer. A multispectral image of a trout fillet is here shown where the reflected light is seen for each narrowband LED, which gives a 19 dimensional spectrum for each pixel in the image.

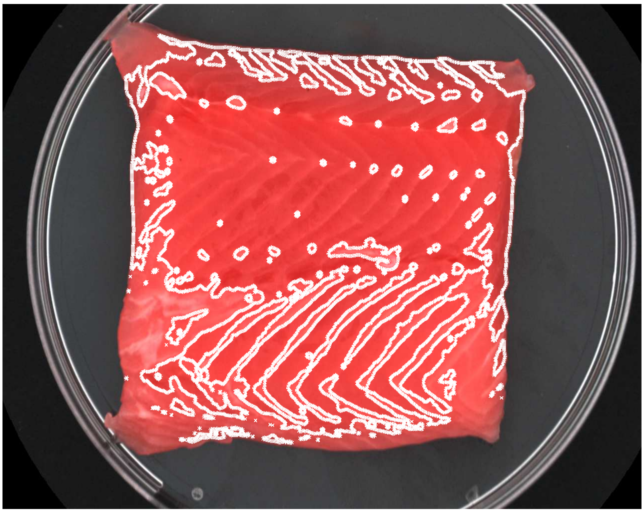


Figure 6. Example of fillet used in analysis with region of interest indicated as contours. The colors of the image are reconstructed from the multispectal image, while the mask is created using the maximum autocorrelation decomposition and otsu's threshold method.

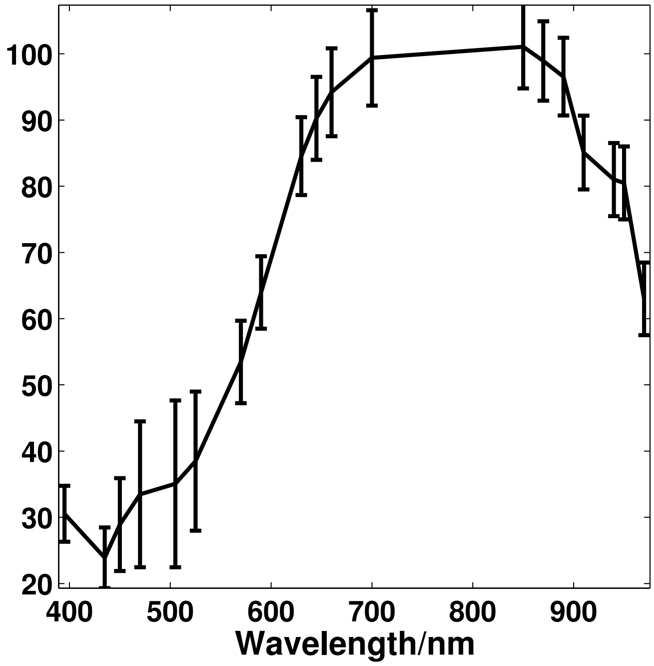


Figure 7. Autoscaling of mean spectra. The mean of all mean spectra is seen together with errorbars indicating one standard deviation of all mean spectra. Each mean spectrum is calculated as the mean of all pixels within the region of interest in a multispectral image.

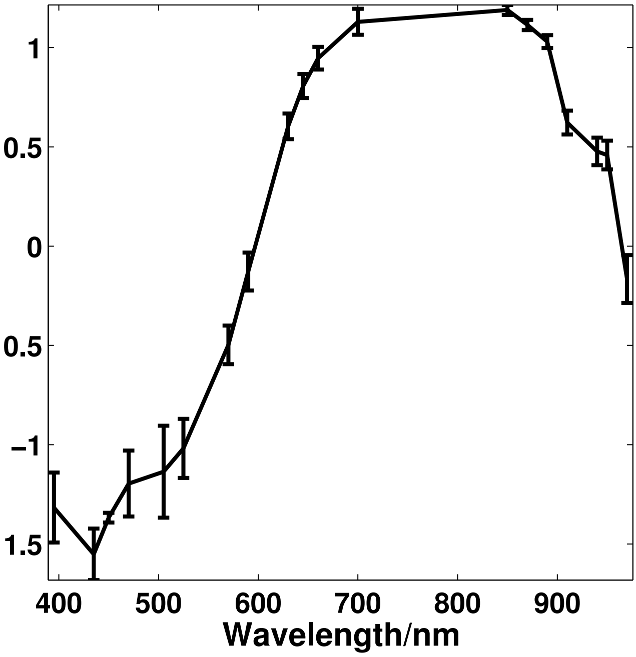


Figure 8. Autoscaling of mean spectra. Autoscaled spectra show significant lower variation in general except for the area between 400 and 500 nm which is now highlighted

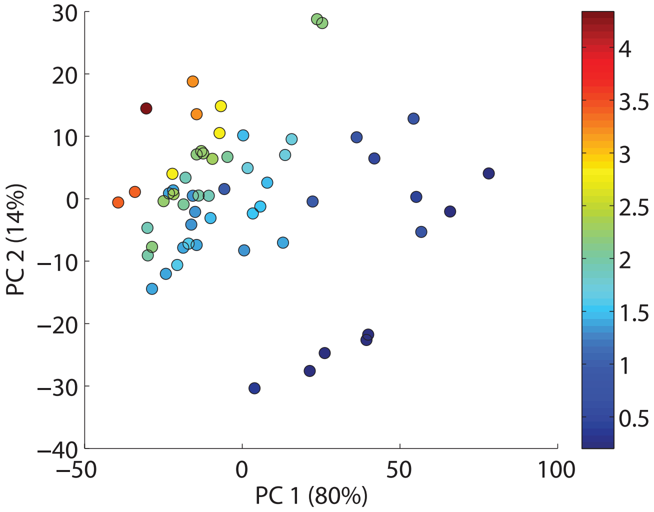


Figure 9. Directions of maximum variation in the autoscaled meanspectra. A scoreplot for a principal component analysis of the 19 dimensional mean spectra shows a definite trend in the the data along the first principal component which accounts for 80% of the variation in the dataset. The second principal component on the y axis accounts for a total of 14%. Each label in the plot represents a multispectral image of a trout fillet, and the number is the concentration in the corresponding fillet in μg.

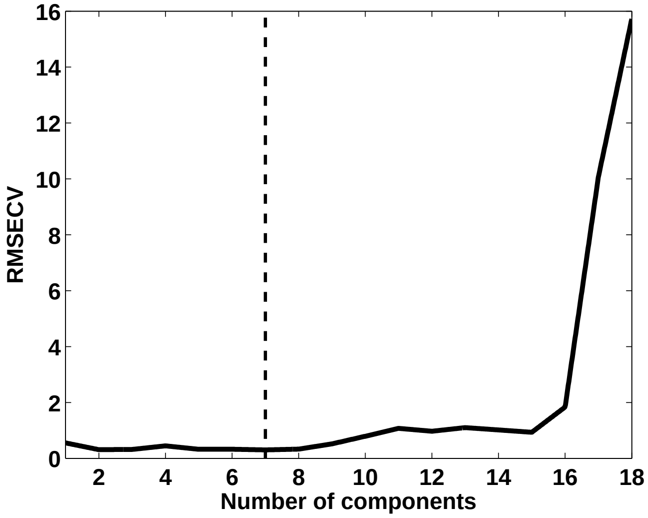


Figure 10. Generalization error calculated as RMSECV. To select a proper model, a leave one out cross validation scheme has been used, where the sum of squared errors (RMSECV) are shown here as a function of components included in the model. The lowest error is indicated with a vertical line, corresponding to a total of 7 components.

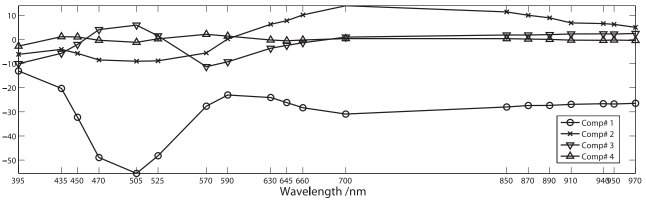


Figure 11. Loadings from PLS model of the multispectral images. The PLS model used to make astaxanthin predictions is based on seven loadings where the first four are shown here. Each loading is shown with a unique symbol, and indicates that especially the area in the beginning of the visible spectrum is important for astaxanthin prediction.

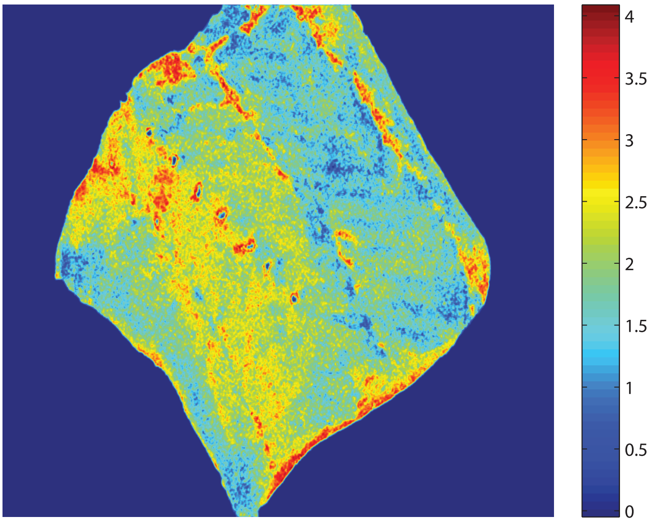


Figure 12. Projected PLS image. Chemically measured average astaxanthin content for the entire piece: 2.26μg The multispectral image is unfolded and projected using the loadings from calibrated prediction model. The result is reshaped to an image which then gives a spatial overview of the astaxanthin within the fillet, based on the light properties of the surface.

Tables

Table 1. Spectral bands of VidemeterLab

01) 395 nm.	09) 630 nm.	17) 940 nm.
02) 435 nm.	10) 645 nm.	18) 950 nm.
03) 450 nm.	11) 660 nm.	19) 970 nm.
04) 470 nm.	12) 700 nm.	
05) 505 nm.	13) 850 nm.	
06) 525 nm.	14) 870 nm.	
07) 570 nm.	15) 890 nm.	
08) 590 nm.	16) 910 nm.	

Narrowbanded lightsources of VideometerLab. The wavelength values shown are the peak values of all light emitting diodes mounted in the sphere. The diodes cover the visible and the first part of the near infrared spectrum.

Table 2. Prediction results obtained using multispectral and sRGB images

	Multispectral	**sRGb**
R^2 (testset)	0.86.	0.66
RMSEP (testset)	0.27	0.45
Std. Error (testset)	0.02	0.05

Appendix G

Temporal Reflectance change of vegetables

Published in Proceedings for IEEE Color and Reflectance in Imaging and Computer Vision Workshop, CRICV 2009, October 2009

Temporal reflectance changes in vegetables

Bjørn S. Dissing[1] , Line H. Clemmesen[1], Hanne Løje[2], Bjarne K. Ersbøll[1] and Jens Adler-Nissen[2]
[1]Department of Informatics and Mathematical Modelling
Technical University of Denmark, DK-2800 Lyngby, Denmark `{bdi,lhc}@imm.dtu.dk`
[2]National Food Institute of Denmark
Technical University of Denmark, DK-2800 Lyngby, Denmark
`{halo,jadn}@food.dtu.dk`

Abstract

Quality control in the food industry is often performed by measuring various chemical compounds of the food involved. We propose an imaging concept for acquiring high quality multispectral images to evaluate changes of carrots and celeriac over a period of 14 days. Properties originating in the surface chemistry of vegetables may be captured in an integrating sphere illumination which enables the creation of detailed surface chemistry maps with a good combination of spectral and spatial resolutions. Prior to multispectral image recording, the vegetables were pre-fried and frozen at $-30°C$ *for four months. During the 14 days of image recording, the vegetables were kept at +5° C in refrigeration. In this period, surface changes and thereby reflectance properties were very subtle. To describe this small variation we employed advanced statistical techniques to search a large featurespace of variables extracted from the chemistry maps. The resulting components showed a change in both the carrot and celeriac samples. We were able to deduct from the resulting components that oxidation caused the changes over time.*

1. Introduction

Quality assessment of food products is a non trivial task which has been approached in different ways over time. Depending on the food product, different parameters are considered important for the overall quality estimation of the food product. Parameters such as surface color, texture and appearance are very general, and should be assessed in most quality estimation scenarios.

Online quality inspection for food process control is today often done by human expert operators who have many years of experience. However, the trend seems to point towards fast non-invasive inspection methods such as Near Infra Red (NIR) technology for quality inspection in different food process control tasks instead. We propose the use of multispectral imaging in the visible as well as the NIR area of the electromagnetic spectrum to quantify chemical properties of food, and thereby stating its level of quality, instead of human operators and as an alternative to standard NIR measurement methods. By employing imaging instead of point measurements it is possible to gain more spatial information about the process, which makes it possible to assess non-chemical as well as chemical quality features. Non-chemical quality features are evaluations of e.g. piece-size, shape and texture.

In this study, we are specifically investigating the quality loss of meal elements for professionally prepared meals with regards to change in surface color after super-chilling and during thawing at $+5°$C over a period of 14 days. Meal elements are robust semi-prepared convenience components based typically on meat, fish or vegetables and meant for professional use. The authors have recently shown that pre-fried vegetable meal elements have promising properties with respect to high culinary quality and robustness towards freezing and thawing, thereby potentially solving a major hindrance for the use of heat treated vegetables as meal elements [3]. Super chilling involves a partial freezing of the products, which slows down quality deterioration [4]. In [16] an experiment of celeriac stored in an refrigerated environment was carried out and various parameters were measured using traditional methods. Celeriac and carrots were the subjects of this study, where we measured the reflectance properties using a multispectral imaging device called VideometerLab which will be described in the next section. The pre-fried vegetables were produced by a new process for continuous stir-frying in industrial scale, which has been introduced for producing convenience high-quality vegetables [1]. The pre-fried vegetables have a low fat content (typically 1%-2% of the product weight), a texture and flavor similar to what can be achieved in the kitchen, and vitamins are preserved almost 100% [5]. In subsequent studies it was observed that the products may be frozen and re-

2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops

heated on a frying pan or in a convection oven without any exudation of excess water, which is a major advantage over existing quick-frozen vegetables [2].

2. Materials and methods

In the following, the experimental design, the acquisition of digital images of the vegetables and further post-processing of these are described.

2.1. Experimental setup

In the present work the quality of pre-fried vegetables (celeriac and carrots shaped as cubes of size approximately $0.5\ cm^3$) were evaluated (e.g. by means of change in color surface) after freezing and thawing. In a pilot plant, the raw products were pre-fried using a special frying machine ”the continuous wok” [1]. After frying, the products were packed in 500g portions in plastic bags and frozen at $-30°C$. After four months of freezing, the bags with the pre-fried vegetables were removed from the freezer and thawed up to 14 days at $+5°C$ in refrigeration. On each day of analysis (day 2, 4, 8, 10, 12 and 14) one plastic bag was taken out from the refrigerator and the contained vegetables were digitized. For both types of vegetables, the samples were digitized using two petri dishes to create a test and training set. The multispectral images were segmented in two steps, first isolating all vegetable-piece in the image, and then separating the pieces from each other. After segmentation, ratios were calculated for all combinations of wavelengths to remove shadow effects and possibly get better baseline separation in different spectral bands. For each ratio in each vegetable-piece, the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 99th percentiles were calculated. This yielded a total of 3249 variables, in a test and training set having 193 and 192 observations respectively for the carrot data. For the celeriac data, similar datasets were created yielding a total of 3249 variables with 207 and 206 observations in the test and training set respectively. Obviously we need a way to figure out which ratios best describe changes over time. For this task a penalized LS algorithm called LARS-EN which is described later was employed to find a set of optimal components. Subsequently statistical tests were performed to evaluate if the identified changes were significant.

2.2. VideometerLab

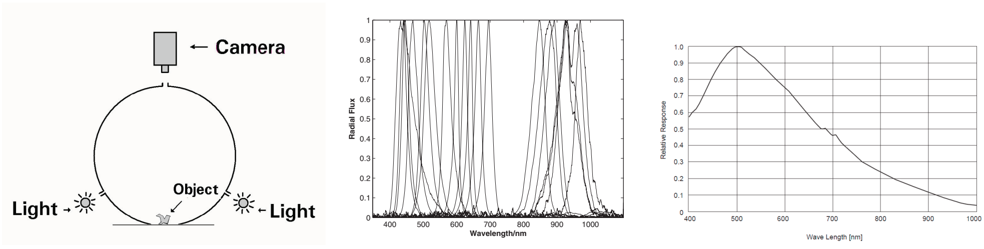
The acquisition of data was done using VideometerLab(http://www.videometer.com) which acquires multi-spectral images in up to 20 different wavelengths ranging from 430 to 970 nm. The camera setup is seen in Figure 1a. The object, in this paper, vegetables, is placed inside an integrating or Ulbricht sphere which has its interior coated to obtain high diffuse reflectivity for optimal light conditions. In the top of the sphere a camera is located with the sensitivity spectrum seen in Figure 1c. The sensitivity decays towards the near-infrared area, which means that the illuminating diodes in this area needs more power to achieve the same level of intensity as the visible bands. The LEDs, having the spectral radiant power distributions seen in Figure 1b, are strobing successively, resulting in an image for each LED of dimensionality 1280x960. These are calibrated radiometrically as well as geometrically to obtain the optimal dynamic range for each LED as well as to minimize distortions in the lens and thereby pixel-correspondence across the spectral bands. The well defined and diffuse illumination of the optically closed scene aims to avoid shadows and specular reflections. Furthermore, the system has been developed to guarantee the reproducibility of the collected images. This allows for comparative studies of time series of images [8].

2.3. Segmentation of the images

In the experiment we considered one vegetable-piece as an observation. In order to extract each vegetable-piece of the multispectral image seen in Figure 2, a relative difficult segmentation problem is at hand. This is caused by the fact that the individual pieces were not placed in a systematic manner where they were isolated, but instead lie in lumps, touching each other. This means that they cast shadows on each other as seen in Figure 2. Furthermore the pieces have very similar spectral fingerprints which means they cannot be discriminated using purely spectral values.

As an initial step, the background, meaning everything but the vegetables was isolated. This was done using Otsu's method [13] on the multispectral image, projected onto a hyperplane. The projection function used to carry out this projection, optimally separates pixels coming from one of two populations. These populations are described by two types of labels which were manually annotated. One set of labels contained spectra of petri dish and general background, while the other of vegetable surface, either carrot or celeriac. The labeled data was used to calculate the projection function by means of a Canonical Discriminant Analysis (CDA) [11].

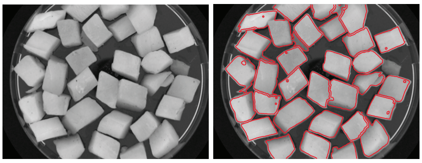
In order to isolate each piece, spatial information is needed, especially gradient information. This information acts as a good guide for the segmentation, and together with morphological transformations employed under a marker-controlled Watershed segmentation as described in [9] we were able to do an automatic segmentation of the vegetables. The result of the segmentation is seen in Figure 2. This seems to be a relatively good result, although flaws are present. In the upper left corner of the petri dish, two pieces are merged together as a result of bad gradient information. In the middle right side it is seen how a dim piece has been



(a) (b) (c)

Figure 1. a) Principle of imaging with integrating (Ulbricht) sphere illumination. The LEDs located in the rim of the sphere ensures narrowband illumination. b) Normalized spectral power distributions of the LEDs located in the VideometerLab. c) Spectral sensitivity of the camera mounted in VideometerLab. It is seen in a) that the camera is placed above the object of interest.

totally ignored by Otsu's method due to its dark appearance. These flaws might be avoided by using an alternative segmentation technique but they were not crucial for the task at hand.



(a) (b)

Figure 2. Both figures show band 10, corresponding to 645 nm of the carrot sample, on day 2. a) is the pure image and b) has the segmentation result superimposed on the image.

2.4. Feature selection method

Having a feature space with n observations and p variables, there are different ways of using this space to describe a variable depending on it. A common technique used to relate the dependent variable and the feature space in a well-posed problem is by using Ordinary Least Square (OLS). Here we have chosen the dependent variable to be the number of the day the observation belongs to, while the independent variables as mentioned earlier are the ratios of the recorded spectrum.

If a problem is well-posed it means among other things that the solution of the problem is unique. Some problems are however not well-posed, which is why many have looked into solving so called ill-posed problems [10] where the covariate matrix does not have full rank. This will always be the case when there are more variables than observations ($p > n$). If such a problem is to be solved properly using Least Squares (LS), some sort of regularization is necessary. Typically this involves including additional assumptions, such as smoothness of the solution. Tikhonov regularization [15] also known as ridge regression [12] is one of the most common ways of regularizing a linear ill-posed problem or an overdetermined system. The ridge regression minimizes the residual sum of squares like an OLS, but in addition it penalizes the L2-norm of the model coefficients. This means all variables are kept in the model but in a smoothed manner. However, in some situations where $p >> n$, ridge regression is not well suited because it creates very complex and thus very little interpretable models. This also means that if some variables contain none or little information regarding the dependent variable, they will still contribute to the final model and thus induce noise. Another approach to solve $p >> n$ problems is by using subset selection or stepwise selection. These methods choose variables having largest partial correlation with the dependent variable, and discards the rest. This type of model is also sometimes known as a parsimonious model and is often much more interpretable, although unfortunately often yields lesser prediction ability.

The Least Absolute Shrinkage and Selection Operator (LASSO), proposed by Tibshirani in [14] was created to solve this problem. Here an L1-norm penalization of the coefficients is used instead of the L2-norm. This means that a sparse solution, as is the case with stepwise selection, is obtained while still continuously smoothing the coefficients to some degree for good prediction. This approach proved to be an improvement of the ridge regression in many cases, while boosting regression and forward stagewise regression both were invented as alternative methods approximately thereafter. These are all described in [11].

A method able to obtain the solution of all these methods in a computationally fast manner is the Least Angle Regression (LARS)[7], proposed by Efron. This regression method gives rise to at most the same amount of calculations as an ordinary LS. An alternative regularization and variable selection method is the elastic net (EN) by Zou [17], which often outperforms forward stagewise regression as well as lasso regression. The elastic net can be incorporated into the LARS regression, commonly known as

LARS-EN, and penalizes the L1 as well as the L2 norm of the coefficients; see Equation (1).

$$L(x,\theta) = \sum_{i=1}^{n}\left(\sum_{j=1}^{p}(\theta_j x_{ij}) - y_i\right)^2$$
$$s.t. \sum_{i=1}^{p}\theta^2 \leq s_1 \text{ and } \sum_{i=1}^{p}|\theta| \leq s_2 \qquad (1)$$

L denotes the loss function, which is the residual sum of squares. θ are the model coefficients and y is the dependent variable, in this case the experimental days. s_1 and s_2 are the constraint bounds on the LASSO and ridge constraints respectively, which together gives the Elastic Net constraint.

The Contours as well as constraints of a simple 2 dimensional problem, simulated as an example of Equation 1 is also seen graphically in Figure 3. This regression scheme is especially suited to solve $p >> n$ problems due to the stability of the L2 norm, and the sparsity property of the L1 norm, which is also shown in [6]. Here the LARS-EN efficiently manages to select a set of suiting variables to detect water in different types of sand, which also is the reason why we have chosen to use it to solve the problem in this paper.

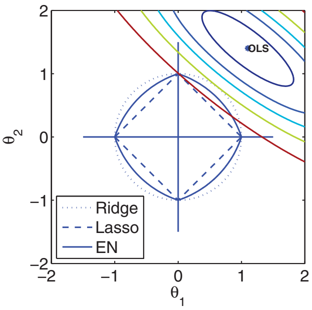


Figure 3. Contours of the residual sum of squares function with the Ordinary LS solution defined as the minimum. The ridge, LASSO and Elastic Net constraints are similarly illustrated. Where the respective constraints and contours intersect, the ridge, LASSO and Elastic Net solutions will be defined.

3. Results and discussion

In order to generalize the model as much as possible a leave one out cross validation (LOO-CV)[11] was used on a training set (A) to estimate the model, and a separate test set (B) was used to evaluate the performance of the model. To further check the repeatability of the model, the training and test set were switched and a new model estimated. Predictions of the estimated models are seen in Figure 4. The boxes in the figures are standard type boxplots and have lines at the 25th, 50th and 75th percentiles. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers. By visually inspecting the boxplots in Figure 4 there seems to be a tendency that the celeriac models having MSE: $(5.88, 4.77)$ respectively, generally have a better prediction ability than the carrot models having MSE: $(17.31, 13.88)$. It also seems that there is a difference between some of the groups in each of the four models, which generally increases slightly in the beginning and then flattens out towards the end. Specifically for the carrots it seems that after day 4, the predictions starts to oscillate, as if an equilibrium has been reached. The same is the case for the celeriac after day 8.

Figure 5 shows the result of all pairwise two-sided t-tests between all days in each model. The t-tests test the H_0-hypothesis, that two groups can be assumed to come from the same population at the 5% level of significance. The statistical tests show that for carrots we are able to verify a significant change in the mean from day 2 to day 4. After day 4 we are not able to verify a significant change in mean for the carrots, which could indicate a steady state has been reached. However, for the celeriac we are able to significantly track a change from day to day until day 12, with the exception of day 8 where some uncertainty appears.

	D2	D4	D8	D10	D12	D14
D2	0	2	2	2	2	2
D4	2	0	0	0	2	1
D8	2	0	0	0	2	0
D10	2	0	0	0	2	1
D12	2	2	2	2	0	2
D14	2	1	0	1	2	0

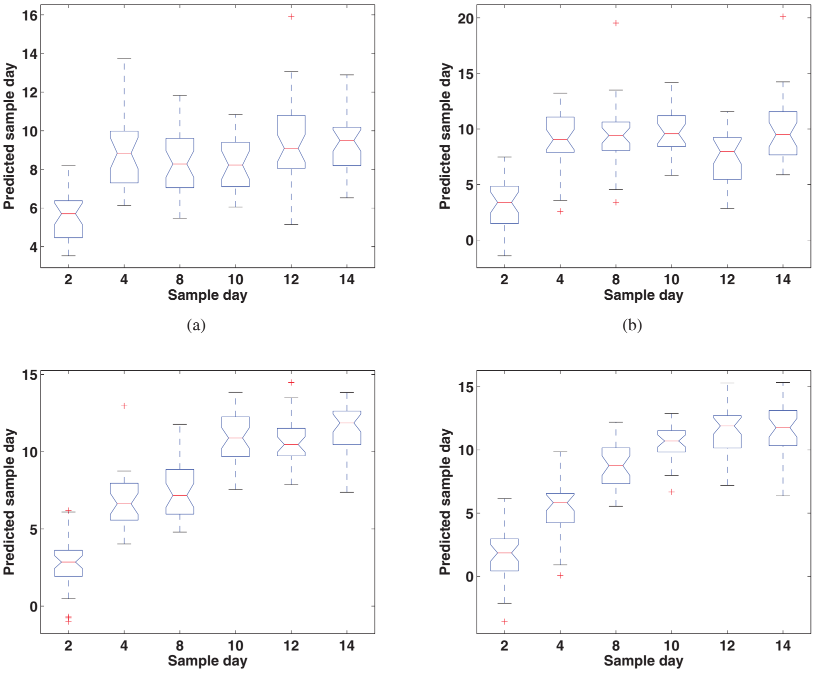
(a)

	D2	D4	D8	D10	D12	D14
D2	0	2	2	2	2	2
D4	2	0	1	2	2	2
D8	2	1	0	2	2	2
D10	2	2	2	0	1	1
D12	2	2	2	1	0	0
D14	2	2	2	1	0	0

(b)

Figure 5. Both tables show all pairwise 2-sided t-tests between all groups of the carrot model(a) and the celeriac model(b). 0 indicates neither model can reject the H_0-hypothesis (i.e. we accept that the means are equal). 1 indicates one of the two models rejects the H_0-hypothesis, and 2 indicates both models reject the H_0-hypothesis(i.e. the groups are significantly different at a 5% level). Both tables are symmetrical respectively. The diagonal contain only zeros.

As mentioned in the section about the experimental setup, the vegetables were kept in the refrigerator in plastic bags.



(a)

(b)

(c)

(d)

Figure 4. All figures show predictions grouped by true sample day. The top two plots show predictions for the carrot samples, while the bottom two plots show the predictions of the celeriac. The two leftmost figures show the model trained with LOO-CV on dataset A and tested on dataset B, while the rightmost figures show the model trained with LOO-CV on dataset B and tested on dataset A.

Plastic bags are not able to isolate oxygen molecules, which is why we believe the change in the spectra is caused by oxidation of the vegetables. An oxidation causes browning/graying of celeriac and carrots to become more pale. An increasing brown/gray color is a change in a wide range of the spectrum, and is essentially a change of brightness. The most significant components describing the celeriac consists of wavelengths from the entire visible spectrum, which coincides with a general shift in brightness. For the carrots the components seem to have a tendency to lie in the red/NIR area, which also coincides with a general more pale appearance, or removal of redness/orangeness, which essentially is an oxidation of the beta-carotene. This is exactly the color change to expect in an oxidation process of these vegetables.

4. Conclusions

An objective measure of the quality change of carrots and celeriac was proposed which uses multispectral image analysis. Six images were recorded over 14 days, for two different data sets. Each carrot or celeriac piece was isolated using a combination of a Canonical Discriminant Analysis and watershed algorithm, for a total of around 200 pieces per training and test set, for both carrots and celeriac pieces respectively. A set of 3249 features were extracted for each vegetable piece, giving rise to a very ill posed $p >> n$ problem. A special regression technique, Least Angle Regression-Elastic Net, was performed on both the carrot and celeriac data sets. Test and training were interchanged, resulting in two models per vegetable type. These were estimated to check the repeatability and statistical tests were performed to check if it in fact was possible to discriminate between the different days predicted on behalf of the estimated models.

The results showed that the celeriac predictions were somewhat better than the carrots, although a trend was seen in both. We see that there is a large change from day 2 to day 4 in the reflectance spectrum for both carrots and celeriac, and for the celeriac we see the change continuing until day 12. The pairwise two sided t-tests showed exactly that these changes were statistically significant at a 5% level of significance. The corresponding sensory tests showed no difference over the 14 days, which makes it the more important that we are able to detect minor changes using multispectral imaging.

5. Acknowledgements

The authors would like to thank Rene Thrane and Peter Reimer Stubbe for carrying out the manual experiments in the laboratory at the National Food Institute of Denmark. This study was financially supported by The Danish Food Industry Agency

References

[1] Jens Adler-Nissen. The continuous wok - a new unit operation in industrial food processes. *J. Food Process Engin.*, 25:435–453, 2002.

[2] Jens Adler-Nissen. Industrial stir frying. *Asia Pacific Food Industry*, 17(5):32–34, 2005.

[3] Jens Adler-Nissen. Continuous wok-frying of vegetables: Process parameters influencing scale up and product quality. *Journal of Food Engineering*, 83:54–60, 2007.

[4] Huynh N. Duy Bao, Sigurjon Arason, and Kristin A. Porarinsdottir. Effects of dry ice and superchilling on quality and shelf life of arctic charr (salvelinus alpinus) fillets. *International Journal of Food Engineering*, 3(3):art.7, 2007.

[5] Maria G. Burgaard, A. Matzen, and Jens Adler-Nissen. Kontinuerlig wok til industriel brug. *Plus Proces*, (6):24–26, 2004.

[6] Line H. Clemmesen, Michael E. Hansen, and Bjarne K. Ersbøll. A comparison of dimension reduction methods with application to multi-spectral images of sand used in concrete. *Machine Vision and Applications*, 2009.

[7] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Annals of Statistics*, 32(2):407–499, 2004.

[8] David D. Gomez, Line H. Clemmensen, Bjarne K. Ersbøll, and Jens M. Carstensen. Precise acquisition and unsupervised segmentation of multi-spectral images. *Computer Vision and Image Understanding*, 106(2-3):183–193, 2007.

[9] Rafael C. Gonzalez, Richard E. Woods, and Steven L. Eddins. *Digital Image Processing*. 2 edition, 2002.

[10] Per C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, 1998.

[11] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *Elements of Statistical Learning: data mining, inference and prediction*. Springer-Verlag, 2001.

[12] A. E. Hoerl and R. W. Kennard. Ridge regression. biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[13] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[14] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[15] Andrey N. Tikhonov. Solutions of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, 4:1035–1038, 1963. English translation of Dokl. Akad. Nauk. SSSR, 151:501-504, 1963.

[16] Sonia Z. Vina and Alicia R. Chaves. Texture changes in fresh cut celery during refrigerated storage. *Journal of the Science of Food and Agriculture*, 83:1308–1314, 2003.

[17] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, 67(2):301–320, 2005.

APPENDIX H

Multispectral Colormapping using Penalized Least Square Regression

---

Published in Journal of Imaging Science and Technology 54(3)

Multispectral Colormapping using Penalized Least Squares Regression

Bjørn Skovlund Dissing,* Jens Michael Carstensen, and Rasmus Larsen
Department of Informatics and Mathematical Modelling, Technical University of Denmark.

(Dated: March 12, 2010)

We propose a novel method to map a multispectral image into the device independent colorspace CIE-XYZ. This method provides a way to visualize multispectral images by predicting colorvalues from spectral values while maintaining interpretability and is tested on a Light Emitting Diode based multispectral system with a total of 11 channels in the visible area. To obtain interpretable models, the method estimates the projection coefficients with regard to their neighbors as well as the target. This results in relatively smooth coefficient curves which are correlated to the CIE-XYZ color matching functions. The target of the regression is a well known color chart, and the models are validated using leave one out cross validation in order to maintain best possible generalization ability. We compare the method with a direct linear regression, and see that the interpretability improves significantly, but comes at the cost of slightly worse predictability.

I. INTRODUCTION

The majority of today's color images are acquired with a CCD or CMOS chip equipped with a Bayer filter i.e. a mosaic filter which splits the incoming photons into three broad primary channels representing the colors or variables; Red, Green and Blue(RGB). Exactly three variables were according to Grassman [1, 2] enough to describe a color sensation.

The usual camera model, assuming lambertian surfaces, is modelled as a linear transformation, Equation 1,

$$P_i = \int_\lambda Q_i(\lambda) R_i(\lambda) E_i(\lambda) d\lambda + \epsilon \tag{1}$$

integrating the lightsource spectrum E, the surface reflection R and the sensor spectral sensibility Q over the visible region of the electromagnetic spectrum for the i'the camera channel, $i = 1, 2, 3$ for standard color images. ϵ is the molling error, assumed gaussian.

This way of capturing color has definitely proven itself usable, but unfortunately it also has some drawbacks. The rough splitting of the photons has the consequence that the intensity recorded in each channel is an integration over a large range of wavelengths. This means that the spectral radiant power distribution of the scene remains hidden for the camera and can lead to metameric failure. Metameric failure can shortly be explained as when two objects match colorimetrically under one illumination, but differ under another. This is because the spectral radiant power distribution of the two objects are different, but the rough splitting of photons fail to observe this. Another drawback of the traditional RGB image acquisition technique is that the colors recorded are device dependent. This means that all cameras records the same scene slightly different, in their own color space.

One way to overcome the problems with metameric failure as well as device dependent colors is by using multispectral imaging systems. In a multispectral image system, the electromagnetic spectrum is sampled more often and in more narrow banded intervals than the three broad intervals used in standard RGB imaging. This means that an approximation of the true distribution of incoming photons is known, for each pixel in the image. There are different ways of creating multi-spectral images. One approach is to use a set of narrow-band filters which basically makes a more delicate grouping of the reflected light from the scene. A setup used often, e.g. [3] and more recently in [4], is the filter wheel approach, where a turnable wheel with different filters is mounted between the lens and the CCD chip. Instead of using a filter wheel, another solution is the crystal liquid tunable filter [5, 6]. Videometer[7] has commercialized a multispectral imaging system based on a Light Emitting Diode (LED) technology. Here, a set of LEDs are strobing successively, and an image is recorded for each LED. Further description of this camera can be found in next section.

In the well established ICC color profile system, color is transferred between different devices by use of a profile connection space (PCS) which is an independent color space, either CIE-XYZ or CIE-LAB. The mapping from a device dependent color space to PCS is well investigated and describeded in e.g. [3, 8–11], where the most common methods used are direct linear regression or regression using polynomials of various degree. Similarly, when using a multispectral device, such as a multispectral imaging system, there is a need to be able to map the multispectral images into PCS. Such a mapping routine is not a trivial task and is the motivation for this paper. In [12] the authors introduce the Spectral Image Processing System (SIPS), where a visualization is based on Spectral Angle Mapping (SAM) and [13] creates an extension to the SAM based visualization algorithm. Furthermore [13] gives a thorough investigation and description of the requirements for successful mapping from a multispectral space to a three dimensional color space in a set of nine visualization goals. As an alternative to multispectral images, which may have problems with acquisition speed and proper calibration, spectral reconstruction methods offer a way to estimate the reflection spectrum of the

---

*bdi@imm.dtu.dk

sample object. Many different methods has been proposed to create this reconstruction, some of them being [14–19]. As a well established method in this area is the Wiener method [19] which makes use of apriori knowledge to build a reconstruction matrix which is multiplied onto the camera response to reconstruct the reflection spectre. Normally spectral reconstruction techniques are used together with normal RGB cameras, but certainly the precision increases with the amount of bands recorded.

In general there are two approaches to transform native camera response to device independent colorspace. These approaches are using an analytical method as described above, or by first estimating the spectral functions as also described above, and then convert it to an independent colorspace. This paper uses an analytical method by transforming a multispectral image to the independent colorspace CIE-XYZ using a regression model, as done in previous studies including [3, 20–22]. Here, direct regression or regression of a polynomial basis expansion of the camera response was used for the mapping. As analternative to these previous studies we here make use of a regression technique which penalizes the curvature of the regression coefficients instead of direct regression. The idea of this penalization, inspired by [23, 24], is to be able to get more smooth, less noisy and more interpretable models.

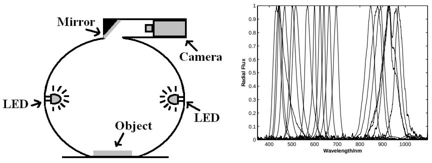
A set of training data is needed in order to calibrate the model properly. Different standards could be used for this, such as the NCS colorsystem which covers a vast amount of colors.We have made use of a well known color rendition chart, *X-rite color checker standard*[25][26], containing 24 squares of different spectral simulations of various common colors as e.g. light and dark human skin.

II. IMAGE ACQUISITION AND DATA

The acquisition of data is done using VideometerLab which acquires multi-spectral images in up to 20 different wavelengths ranging from 385 to 970 nm. VideometerLab is a multipurpose camera often used for scientific purpose and proof of concept applications, which is why diodes emitting light outside as well as inside the visible area of the electro magnetic spectrum is mounted. Previously this device has been used in many different vision applications. A few examples are quality estimation of mink fur, analysis of psoriasis lesions, classification of fungi and temporal change detection in reflectance of vegetables [27–29].

The camera setup is seen in Figure 1 a). The object is placed inside an integrating Ulbricht sphere which has its interior coated with a matte coating to obtain high diffuse reflectivity for optimal light conditions. By optimal light conditions is meant conditions which avoids shadows to a certain degree as well as highlights/reflections. In the top of the sphere a Point Gray Scorpion camera is mounted. The LEDs, having the spectral radiant power distributions seen in Figure 1 b) are strobing successively, and for each LED an image of dimensionality 1280x960 is acquired, which in the end yields a multi-spectral image. In general, not many multispectral imaging systems using LEDs exists, however [30] introduces a camera system which employs a LED array coupled with a photodiode array to measure reflectance spectra. Their system is evaluated by its ability to estimate reflection spectra using 928 colors in the ISO12642 IT8/3 chart, using a clustering and polynomial regression method.

An image of the X-rite ColorChecker standard(Macbeth) has been recorded in this manner using the VideometerLab system. Since only 11 of the bands created by the camera resides in the visible spectrum, these are the only ones containing color information, which is why the rest are discarded, weighted down, and not used further in this study. Finally the acquired multispectral image is seen in Figure 2 with 11 distinct wavelengths at; 430, 450, 470, 505, 525, 565, 590, 630, 645, 660 and 700 nanometer.



(a) (b)

FIG. 1. *a) Principle of imaging with integrating (Ulbricht) sphere illumination. The LEDs located in the rim of the sphere ensures narrowband illumination. b) Normalized spectral power distributions of the LEDs located in the VideometerLab. c) Spectral sensitivity of the camera mounted in VideometerLab. It is seen in a) that the camera is placed above the object of interest.*

As seen, the x-rite standard contains a matrix of colors, four rows and six columns, which sums to a total of 24 squares. These squares are extracted from the image manually, and a median value is calculated for each square thus yielding a total of 24 spectral samples. The real values of the colors in sRGB and CIE-Lab space are known, and can thus be used for calibration, which will be discussed further in the theory section.

III. THEORY

The relationship between spectral space and the three dimensional CIE-XYZ space is known to be linear, Equation 1. Therefore we are able to map the spectra of the multispectral image seen in Figure 2 into the CIE-XYZ using a linear model. For the duration of this section, the size of the spectral space will be denoted p, and the amount of samples is denoted N. A standard linear model is written on the form 2.

$$y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \epsilon \tag{2}$$

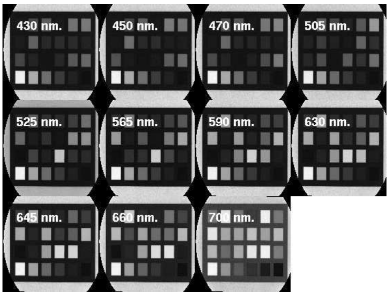


FIG. 2. *11 Channels of a multispectral image containing a total of 20 channels. The wavelength at which the channel was recorded is shown on each image. The image was recorded using the VideometerLab with an LED technology.*

Here y denotes the response, which we want to be able to predict, in this case a known CIE-XYZ value. x denotes the independent covariates, here a sparsely recorded electromagnetic spectrum. Since we do not know the exact relation between x and y, it is desirable to compute a set of adjustable parameters β in such a way that ϵ becomes as small as possible using a set of measured/observed spectra and a known target. If there exists more spectra samples than recorded wavelengths for each sample, the system is overdetermined, and several solutions exists, which is why the best approximate solution have to be determined.

Equation 2 resembles a plane in a hyper dimensional space, which we want to fit by minimizing the euclidian distances between a given set of observations and the plane. This fit is often written as the residual sum of squares, and is normally known as the Ordinary Least Square (OLS) method, Equation 3.

$$\widehat{\beta} = \arg\min_{\beta} \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \right\} \tag{3}$$

The fastest way to minimize Equation 3 is by setting the derive to 0 and solve for beta since this is a strictly convex problem. Written in matrix form, normally referred to as *the normal equations*, this may be written as

$$\widehat{\beta} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{X}^T\mathbf{y} \tag{4}$$

$\widehat{\beta}$ is a $px3$ projection vector which is capable of projecting new x observations in p dimensional space into the three dimensional color space; $\widehat{y} = \mathbf{X}\beta$, where $\mathbf{X}$ is the multispectral data and $\mathbf{y}$ is the matrix containing the true color values.

According to the Gauss-Markov theorem, the estimate of β in an OLS is the best linear unbiased estimator (BLUE), which means that for all unbiased solutions, the OLS solution is the one with smallest variance. However, the OLS problem may be modified to get biased estimators, and thereby get an even better solution. A well known method used to modify the OLS in order to be able to solve ill-posed problems of overdetermined systems where the Gramian matrix$\left(\mathbf{X}^T\mathbf{X}\right)$ is singular is the Tikhonov regularization [23, 31, 32] method - also known as ridge regression. This regularization basically penalizes the L2-norm or euclidian length of the parameter vector, as seen below

$$\begin{aligned} \widehat{\beta}^{\text{ ridge}} &= \arg\min_{\beta} \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \right\}, \\ &s.t. \sum_{j=1}^{p} \beta_j^2 \leq s \end{aligned} \tag{5}$$

s is thus a parameter that basically controls the bias-variance tradeoff as best as possible, meaning that this regularization makes it possible to find a biased solution with smaller variance than the BLUE estimate. The parameter s is chosen so that the generalization error is minimized, which in other terms means the best fit which is not an overfit. How this is done is described in the end of this section.

Since the Tikhonov regularization is a constraint version of the convex OLS problem, this leads to a quadratic constrained optimization problem. Such a problem may be solved in different ways e.g. using quadratic programming with constraints or by solving the Lagrangian problem, by introducing a lagrange multiplier λ in Equation 5. After λ has been introduced, the formulation may again be recast to matrix form.

$$\widehat{\beta} = \left(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}\right)^{-1} \mathbf{X}^T\mathbf{y} \tag{6}$$

$\mathbf{I}$ is the $p \times p$ identity matrix. The penalty term of the Ridge Regression shrinks the coefficients towards zero, and thereby weighs down badly influencing variables more than others.

Instead of shrinking the length of the coefficient vector, we are interested in penalizing large variance between neighboring coefficients in order to obtain a more smooth coefficient/weight curve. We present a term based on the common finite difference operators, which are used to approximate numerical derivatives of arbitrary order. A penalization of the gradient of the coefficients β, is thus equivalent to a penalization of the backward finite difference of the coefficients

$$\nabla_h \left[f\right](x) = f(x) - f(x-h) \tag{7}$$

where h is set to step a single unit. Rewritten into a constraint, this is given as

$$\sum_{j=1}^{p-1} (\beta_j - \beta_{j+1})^2 \leq s \tag{8}$$

Similarly a smoothing term could be incorporated as

$$\sum_{j=1}^{p-2} (\beta_j - 2\beta_{j+1} + \beta_{j+2})^2 \leq s \tag{9}$$

Both terms should have the effect of preventing large fluctuations in the coefficients, and make it easier to understand how the model predicts. In [24], the authors proposed a similar penalization, however with the L1-norm instead, which gives rise to piecewise linear sparse models which not is interesting in the type of problem we are dealing with in this paper. Before this regularization may be used we have to assume an ordering of the coefficients β, which comes naturally for spectral data, since the coefficients represents neighboring wavelengths. Building the regularization into the OLS of a curvature penalization , yields

$$\widehat{\beta} = \arg\min_{\beta} \left\{ \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 \right\}, \qquad s.t. \sum_{j=1}^{p-2} (\beta_j - 2\beta_{j+1} + \beta_{j+2})^2 \leq s \tag{10}$$

Again we can rewrite this into lagrangian form

$$\widehat{\beta} = \left(\mathbf{X}^T\mathbf{X} + \lambda \mathbf{A}^T\mathbf{A}\right)^{-1} \mathbf{X}^T\mathbf{y} \tag{11}$$

where A is a $p \times p$ tridiagonal matrix.

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} \tag{12}$$

Due to the sparse amount of observations, Leave One Out Cross Validation(LOOCV) is used to evaluate the model, and calculate the generalization error [33]. This means the model is trained and evaluated n times, one time for each sample, with all observations except the left out sample which is used to generate a test-error residual. Finally a total testerror can be calculated. A finite difference or gradient descend scheme was used to select proper parameter values for the regularization parameters.

IV. EXPERIMENTAL DATA AND RESULTS

Having a multispectral image of a colorchecker with 24 different color patches recorded at 11 different wavelengths, means we have $n = 24$ observations and $p = 11$ variables, and thus an overdetermined system. 24 patches are very few observations in a regression method, and care should be taken not to overfit the data, which we have chosen to avoid by minimizing the testerror with LOOCV and a finite difference scheme, as described in the previous section. Before the calibration, all real color patch values were transformed from CIE-Lab to CIE-XYZ using

$$\begin{aligned}
Y &= \begin{cases} Y_n f_y^3 & \text{for } f_y > \delta \\ f_y - \frac{16}{116} 3\delta^2 Y_n & \text{otherwise} \end{cases} , \quad f_y \equiv \frac{L* + 16}{116} \\
X &= \begin{cases} X_n f_x^3 & \text{for } f_x > \delta \\ f_x - \frac{16}{116} 3\delta^2 X_n & \text{otherwise} \end{cases} , \quad f_x \equiv f_y + \frac{a*}{500} \\
Z &= \begin{cases} Z_n f_z^3 & \text{for } f_z > \delta \\ f_z - \frac{16}{116} 3\delta^2 Z_n & \text{otherwise} \end{cases} , \quad f_z \equiv f_y - \frac{b*}{200} \\
&, \ \delta = \frac{6}{29}
\end{aligned} \tag{13}$$

with the D50, 2 degree observer white spot reference. Four different regression schemes were compared, i.e. Least Squares, Ridge Regression, Gradient Ridge Regression (GRR) and Curvature Ridge Regression (CRR), all presented in the theory section. In order to compare the results, we report the min, max and mean ΔE of the color reconstructions.

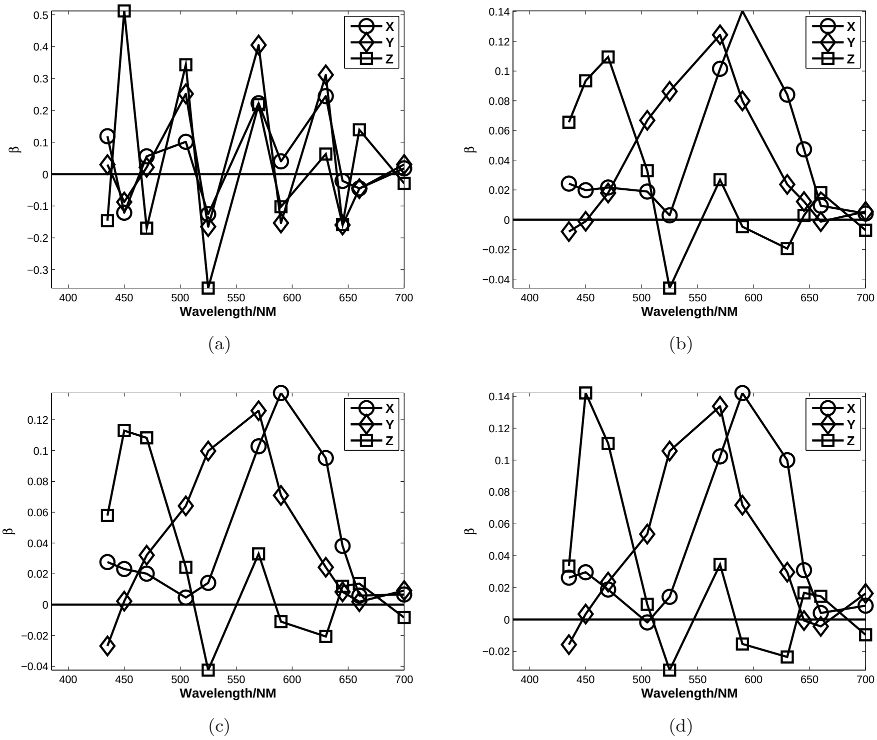
$$\Delta E = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2} \tag{14}$$

as well as mean euclidian distance in XYZ space, denoted $\overline{\Delta XYZ}$ and finally the test Root Mean Square Error (RMSE).
For the penalized regression schemes, a λ was chosen for each channel, resulting in a total of three λ per regression scheme. Table I shows a compilation of the results, the best being underlined.

	ΔE_{MIN}	ΔE_{MAX}	$\overline{\Delta E}$	$\overline{\text{RMS}}$	$\overline{\Delta XYZ}$
OLS	$\underline{0.24431}$	$\underline{3.9309}$	$\underline{1.2912}$	0.75088	$\underline{0.58368}$
RR	0.26119	4.6468	1.3993	0.66173	0.64521
GRR	0.25967	4.5934	1.3191	$\underline{0.64494}$	0.60110
CRR	0.26384	4.6233	1.3249	0.64973	0.60863

TABLE I. *Resulting colorimetric errors using 4 different regression methods on 11 different wavelengths. All errors are generalization errors calculated using Leave One Out Cross Validation.*

From Table I it is seen that the OLS method attains the lowest $\overline{\Delta E}$ error of 1.29. This is however not the case with the $\overline{\text{RMS}}$. This is worth noticing, and is caused by the fact that all the applied procedures minimizes RMS and not $\overline{\Delta E}$, which has a nonlinear relationship. Aside from OLS, the other procedures attains more or less equal error magnitude in CIE-Lab space. It is furthermore worth noticing that all methods attain mean LAB differences below 3, which according to [3] as a rule of thumb is *Hardly perceptible* by the human eye. Figure 3 shows plots of the calculated coefficients for each of the four tested regression methods. OLS clearly stand out, with coefficient the does not show any particular pattern. The three regularized regression schemes however



(a) (b) (c) (d)

FIG. 3. a) Ordinary Least squares fit b) Ridge Regression fit b) Curvature Penalization fit of first derivative c) Curvature Penalization fit of second derivative

shows a clear correlation with the CIE-XYZ tristimuli curves, which would make sense to obtain in a regression of a problem as the one posed in this paper. The blue tristimuli model seems however a bit oscillating, while the red and green models seems to be shifted a bit in the spectrum compared to the real tristimuli functions. The coefficients of the three regularized methods seem quite similar, especially the penlization of the gradient and curvature methods, which naturally both seem to be a bit more smooth than ridge regression. In case we had densely sampled hyperspectral images we would be able to directly use the theoretical XYZ tristimulus values to map the spectra into the XYZ colorspace. However, having sparsely sampled spectra this is not an option, and as an alternative to reconstructing spectra using any existing method, the smooth regularization shown in Figure 3 shows that we are able to create sensible substitution weighting curves to the theoretical XYZ tristimulus curves. Thus, these curves are able to map sparsely sampled multispectral images into XYZ space in the same way the theoretical curves maps densly sampled spectra using Equation 1.

V. CONCLUSION

The focus of this paper has been to develop an alternative method to map a multispectral image into the independent colorimetric colorspace, CIE XYZ . The method is built to support a multispectral image of arbitrary spectral dimensionality, above 3 and specifically tested on a multispectral system with 11 wave bands in the visible area. To create a projection vector, a regularized regression technique was utilized and evaluated on the x-rite ColorChecker Chart containing 24 different color patches using leave one out cross validation. The cross validation enabled us to assess the generalization ability in spite of the few color samples available. The colorimetric prediction ability was reported in $\overline{\Delta E}$,$\overline{\text{RMS}}$ and $\overline{\Delta XYZ}$ and yielded good and interpretable color reproduction results. A drawback of this method is that in order to calculate proper weights of the projection vector, β, it is necessary to do the regression on reflection spectre from a training image with the same light setup as the target image. By light setup for the used multispectral system, is meant the amount of power going through the light emitting diodes, as well as the strobing time which is optimized for best dynamic range in a local area of a given surface chemistry. However, for any other multispectral system, the training and target image would have to be recorded under the same lighting conditions, which usually is possible.

For correct light setups, the reported results, all with ΔE values below 3, indicates a hardly perceptible error between the real colorchecker values and the estimation provided by the mapping method. Plots of the estimated regularized models motivated a regularization approach to achieve higher interpretability, while colorimetric measures showed that Ordinary Least Square better predictions performed slightly better.

VI. REFERENCES

---

[1] Gunter Wyszecki and W.S. Stiles. Color science: Concepts and Methods, Quantitative Data and Formulae, second edition. Wiley & sons Inc, 2000.

[2] H. Grassman. Zur theorie der farbenmischung. Annalen der Physik, 165(5):69–84, 1853.

[3] J.Y Hardeberg. Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches. PhD thesis, Ecole Nationale Superieure des Telecommunications, 2001.

[4] Johannes Brauers, Nils Schulte, and Til Aach. Multispectral filter-wheel cameras: Geometric distortion model and compensation algorithms. IEEE Transactions on Image Processing, 17(12):2368–2380, Dec 2008.

[5] J. Y. Hardeberg, F. J. Schmitt, and H. Brettel. Multispectral color image capture using a liquid crystal tunable filter. Optical Engineering, 41:2532–2548, October 2002.

[6] Nahum Gat. Imaging spectroscopy using tunable filters: A review. In Proceedings of SPIE, volume 4056, pages 50–64, 2000.

[7] URL: http://www.videometer.com.

[8] Christine Connolly and Thomas Fliess. A study of efficiency and accuracy in the transformation from rgb to cielab color space. IEEE Transactions on Image Processing, 6(7):1046–1048, 1997.

[9] P.C. Hung. Colorimetric calibration for scanners and media. In W.-C. Chang and J. R. Milch, editors, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 1448 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pages 164–174, June 1991.

[10] Guowei Hong, M. Ronnier Luo, and Peter A. Rhodes. A study of digital camera colorimetric characterization based on polynomial modeling. Color Research & Application, 26(1):76–84, 2001.

[11] T. L. V. Cheung, S. Westland, D.R. Connah, and C. Ripamonti. A comparative study of the characterization of colour cameras by means of neural networks and polynomial transforms. Journal of Coloration Technology, 120(1):19–25, 2004.

[12] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz. The spectral image processing system (sips) - interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment, 44:145–163, 1993.

[13] Nathaniel P. Jacobson and Maya R. Gupta. Design goals and solutions for display of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing, 43:2684–2692, 2005.

[14] Alamin Mansouri, Tadeusz Sliwa, Jon Y. Hardeberg, and Yvon Voisin. Representation and estimation of spectral reflectances using projection on pca and wavelet bases. COLOR research and application, 33(6):485–493, March 2008.

[15] Jeffrey M. DiCarlo and Brian A. Wandell. Spectral estimation theory: beyond linear but before bayesian. Journal Of Optical Society of America, 20(7):1261–1270, July 2003.

[16] Norimichi Tsumura, Hideki Sato, Takayuki Hasegawa, Hideaki Haneishi, and Yoichi Miyake. Limitation of color samples for spectral estimation from sensor responses in fine art painting. Optical Review, 6(1):57–61, 1999.

[17] Y. Zhao, M. Nezamabadi L. A. Taplin, and R. S. Berns. Using the matrix r method for spectral image archives. In 10th Congress of the International Colour Association, 2005.

[18] D. R. Connah and J. Y. Hardeberg. Spectral recovery using polynomial models. In R. Eschbach and G. G. Marcu, editors, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 5667 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pages 65–75, December 2004.

[19] P. Stigella, K. Miyatab, and M. Hauta-Kasaric. Wiener estimation method in estimating of spectral reflectance from rgb images. Pattern Recognition and Image Analysis, 17(2):233–242, June 2007.

[20] Daniel Nystrom. Colorimetric and multispectral image acquisition using model-based and empirical device characterization. In Bjarne K. Ersboll and Kim Steenstrup Pedersen, editors, SCIA, volume 4522 of Lecture Notes in Computer Science, pages 798–807. Springer, 2007.

[21] J. Conde, H. Haneishi, M. Yamaguchi, N. Ohyama, and J. Baez. Cie-xyz fitting by multispectral images and mean square error minimization with a linear interpolation function. Revista Mexicana de Fisica, 50:601–607, December 2004.

[22] J. Y. Hardeberg, F. J. Schmitt, and H. Brettel. Multispectral image capture using a tunable filter. In R. Eschbach and G. G. Marcu, editors, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 3963 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pages 77–88, December 1999.

[23] A. E. Hoerl and R. W. Kennard. Ridge regression. biased estimation for nonorthogonal problems. Technometrics, 12(1):55–67, 1970.

[24] Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society Series B, 67(1):91–108, 2005.
[25] URL: http://www.xrite.com.
[26] C. S. McCamy, H. Marcus, and J. G. Davidson. A color-rendition chart. Journal of Applied Photographic Engineering, 2(3):95–99, 1976.
[27] Bjørn S. Dissing, Line H. Clemmesen, Hanne Løje, Bjarne K. Ersbøll, and Jens Adler-Nissen. Temporal reflectance changes in vegetables. In Proceedings of IEEE Color and Reflectance in Imaging and Computer Vision Workshop, Kyoto, Japan., 2009.
[28] Line H. Clemmensen and B. K. Ersbøll. Multispectral recordings and analysis of psoriasis lesions. In MICCAI '06 - Workshop on Biophotonics Imaging for Diagnostics and Treatment, pages 15–18. 9th MICCAI Conference, October 2006.
[29] M. E. Hansen, B. K. Ersbøll, J. M. Carstensen, and A. A. Nielsen. Estimation of critical parameters in concrete production using multispectral vision technology. In Lecture Notes in Computer Science, LNCS3540, Lecture Notes in Computer Science, pages 1228–1237, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005. Informatics and Mathematical Modelling, Technical University of Denmark, DTU.
[30] Shoji Yamamoto, Norimichi Tsumura, Toshiya Nakaguchi, and Yoichi Miyake. Development of a multispectral scanner using led array for digital color proof. Journal of Imaging Science and Technology, 51(1):61–69, 2007.
[31] Andrey N. Tikhonov. Solutions of incorrectly formulated problems and the regularization method. Soviet Mathematics Doklady, 4:1035–1038, 1963. English translation of Dokl. Akad. Nauk. SSSR, 151:501-504, 1963.
[32] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. Elements of Statistical Learning: data mining, inference and prediction. Springer-Verlag, 2001.
[33] Harald A. Martens and Pierre Dardenne. Validation and verification of regression in small data sets. Chemometrics and Intelligent Laboratory Systems, 44(1-2):99 – 121, 1998.

APPENDIX I

Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts

Accepted in Proceedings of Scandinavian Workshop on Imaging Food Quality

Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts

Ken-ichi Kobayashi[1], Ken Nishino[1], Bjørn Skovlund Dissing[2], Masaaki Mori[3], Toshihiro Toyota[1], and Shigeki Nakauchi[1]

[1] Department of Computer Sciences, Toyohashi University of Technology,
1-1 Hibarigaoka, Tenpaku-cho, Toyohashi 441-8580, Japan
{kobayashi09,nishino06,toyota}@vpac.cs.tut.ac.jp,nakauchi@tut.jp
http://www.vpac.cs.tut.ac.jp/
[2] Department of Informatics and Mathematical Modelling,
Technical University of Denmark, DK-2800 Lyngby, Denmark
bdi@imm.dtu.dk
[3] Mie Prefecture Livestock Research Institute,
1444-1 Ureshino-cho, Matsusaka 515-2324, Japan
morim05@pref.mie.jp

Abstract. Quantification of specific compounds in a food-matrix is a very important factor for an overall quantification of the quality. Near infrared (NIR) hyperspectral imaging is a powerful technique to quantify specific constituents as well as its spatial distribution of the food-matrix. Hyperspectral imaging is however very expensive. We propose a way to design a simple measurement system consisting of a NIR sensitive monochrome camera together with a small set of optical filters to estimate and visualize a specific food compound without requiring a full hyperspectral device. Based on a set of hyperspectral measurements of beef and physical and chemical analysis of the fat within the beef, we propose a method to design a set of ideal Band Pass Filters (BPF), as small as possible while still maintaining predictability of fat content. The results show that 2 filters is a suitable amount of filters for prediction.

Keywords: NIR hyperspectral imaging, Optical filter, Beef, Content

1 Introduction

Traditionally quality evaluation of food has been done using visual inspection, chemical measurements or sensory testing. These methods are destructive, time-consuming and/or subjective, which calls for other quantification methods. Recently non-destructive methods for evaluation of food quality as well as visualization of the spatial distribution of constituents, by using (NIR) hyperspectral information have emerged [1]. Although hyperspectral image data is very versatile and contains much information, the measurement system is extremely expensive to install in a food factory. As another approach based on hyperspectral data, a method designing the optical transmission for the optical filter to modulate a

RGB camera's spectral sensitivity and to highlight an object's spectral features is proposed [2][3].

We propose a simple measurement system consisting of a NIR monochrome camera together with a small set of optical filters to estimate and visualize a specific food compound without use of a hyperspectral device. We use the fat content in raw beef as the target. Currently in Japan, the quality evaluation of beef carcasses is performed manually by a grader. In this grading, only visual inspection is used. Marbling, which is the amount and distribution of fat in the meat is the most important factor. Based on a set of hyperspectral measurements and physical and chemical analysis of fat within the beef, we propose a method to design a set of optical Filters, which accurately is able to predict the amount and distribution of this fat.

2 Materials and methods

2.1 Samples and measurements

A total of 126 meat samples consisting of various parts from three 25-month-old Japanese black cattle were collected. After about 60 days of ageing at $0-5°$C, the beef carcasses were kept at $-25°$C to maintain the fat properties during storage and transportation.

The fat content used for reference values was analyzed by physical and chemical method. Automated Soxhlet extraction equipment (Soxtherm416, Gerhardt, Germany) was used to obtain the fat percentage.

The hyperspectral measurements were performed by a NIR hyperspectral imaging system consisting of a linear image sensor (Spectral Camera SWIR; SPECIM Spectral Imaging Ltd, Finland), a linear slide table and halogen light sources (MH-M15, 150 W; Hataya Ltd, Japan). The hyperspectral camera works in the wavelength range of 970-2500 nm with a bandwidth of 6.3 nm at a resolution of 320 pixels (X-axis). We acquired samples at a resolution of $380\mu m/pixel$ over a rectangular region of $120\times130mm$ by moving the slide table. The exposure time was 3.0 ms.

2.2 Calculation of filter transmission intensity

The MATLAB 7.5 (R2007b; The MathWorks Inc., Natick, MA, USA) software package was used to analyze the hyperspectral image data. Optical filters were designed as ideal (rectangle-shaped) BPF and an assumption was made that a measurement using an optical filter consists of three images; a dark current image (I_{Dark}), a white standard image (I_{White}) and a sample image (I_{Sample}). To remove the effects of dark current, spectral features produced by the light source, and flat field inhomogeneities, we use I_R as a model parameter calculated from measured images or hyperspectral images by

$$I_R = \frac{I_{Sample} - I_{Dark}}{I_{White} - I_{Dark}} = \int_{\lambda_{short}}^{\lambda_{long}} \frac{I_{Sample}(\lambda) - I_{Dark}}{I_{White}(\lambda) - I_{Dark}} d\lambda$$

Where $\{I_{Dark}, I_{White}(\lambda), I_{Sample}(\lambda)\}$ is hyperspectral data, λ_{short} and λ_{long} are the wavelength edges of the BPF. When calculating I_R , the spectra $\{ I_{Sample}(\lambda), I_{White}(\lambda) \}$ were interpolated by cubic spline to 1,000 wavelength points between λ_{short} and λ_{long} .

2.3 Design of optical filter property

The filter properties were modeled by using the center wavelength (λ_c) and the half-bandwidth (w_h). We limited the minimum bandwidth of BPF to 50 nm, because too narrow BPFs cannot obtain enough luminance, which will cause reduction of the signal-to-noise ratio. The maximum bandwidth was limited to 1,000 nm, because very wide BPFs are hard to implement as a real optical filter. The wavelength range was also limited from 1,000 nm to 2,300 nm, because shorter/longer wavelength ranges of hyperspectral data could not provide sufficient intensity. With a spectral resolution of about 6.3 nm, meaning a total number of wavelength-points of 206. Even if the edges of BPFs are limited to these wavelength-points, every possible combination of n BPFs is $\simeq 10^{4n}$. Therefore "brute-force search" is not suitable for more than 2 or 3 filters in terms of searching time.

Multiple Linear Regression (MLR) was used to estimate parameters for linear models using filter transmission intensities as variables. To create and evaluate the estimation models, samples were divided into calibration and validation sets. Calibration samples were selected randomly ($N_c = 84$) and remainder were used as validation samples ($N_v = 42$). These sample sets were fixed to compare the results of different feature selection method.

Filter feature selections were done using leave-one-out cross validation, to minimize the root mean square error of cross-validation ($RMSECV$) given by

$$RMSECV = \sqrt{\frac{\sum (y_c - \widetilde{y}_c)^2}{N_c}}$$

where y_c is the reference value, and $\widetilde{y_c}$ is the predicted value of the calibration-set in cross validation. Furthermore the standard error of calibration (SEC) , the root mean square error of calibration ($RMSE_c$) and the standard error of prediction (SEP) were calculated as

$$SEC = \sqrt{\frac{\sum (y_c - \hat{y}_c)^2}{N_c - n - 1}}, \quad RMSE_c = \sqrt{\frac{\sum (y_c - \hat{y}_c)^2}{N_c}}, \quad SEP = \sqrt{\frac{\sum (y_v - \hat{y}_v)^2}{N_v}}$$

where $\hat{y}_c$ is the predicted value of the calibration-set using the model, n is the number of filters, y_v is the reference value of the validation-set, and $\hat{y_v}$ is the predicted value of the validation-set using the model.

We compared the following three feature selection methods.

Stepwise random selection In this method one needs to define the number of filters. A scoremap and a countmap is maintained for each filter which is used for deciding the final properties for the corresponding filter:

1. Generate the $\{m, (m+1), (m+2), ..., n\}$-th filters randomly.
2. Calculate n filter outputs of each calibration sample.
3. Make a MLR model by using the calculated filter outputs and the corresponding reference values of the calibration-set.
4. Calculate the RMSECV for the calibration-set.
5. Add the RMSECV value to the n points in the scoremap. Also add 1 to the n points in the countmap.

6. Repeat step 1 to 5 sufficiently many times.

The scoremaps and the countmaps are made for each m-th filter individually, and have coordinate points corresponding to every filter feature (λ_c, w_h).

7. Remove 0-count points both in the m-th countmap and the m-th scoremaps. Then make a mean scoremap by dividing the m-th scoremap by the m-th countmap.
8. Choose a point that minimizes the mean scoremap. This is then fixed as the m-th filter (λ_{c_m}, w_{h_m})
9. Repeat steps 1 to 8 n times with the $\{1, 2, ..., (m-1)\}$-th fixed filters.

Finally, local optimization (constrained nonlinear optimization) is performed using the result as a starting guess, minimizes

$$min \left(\sqrt{\frac{\sum_i^{N_c} (y_i - \sum_j^n f(\lambda_{short}(j), \lambda_{long}(j), x_i))^2}{N_c}} \right), f(\lambda_{short}, \lambda_{long}, x) = \int_{\lambda_{short}}^{\lambda_{long}} x d\lambda$$
$$s.t. \quad (1000 \leq \lambda_{short}, \lambda_{long} \leq 2300), \quad (50 \leq \lambda_{long} - \lambda_{short} \leq 1000)$$

where x is reflectance (hyperspectral data) of a sample. In this method, a filter feature which has small average error in many trials of various feature combinations, is assumed to contain useful information for estimation.

Forward selection method Inspired by the classic variable selection technique also known as forward selection [4], the forward selection method calculates $RMSECV$ for all possible filters and selects the filter (λ_{C_1}, w_{h_1}) which minimizes this metric. This filter (λ_{C_1}, w_{h_1}) is fixed at the found feature and the procedure is then repeated for next filter (λ_{C_2}, w_{h_2}) . The m-th filter (λ_{C_m}, w_{h_m}) is chosen with $(m-1)$ fixed filters. The wavelength of the BPF edges are discrete values.

At the m-th filter selection step, this method consider the combination of $(m-1)$ filters already fixed and m-th filter. Contrary to this the stepwise random selection method considers the combination of n filters at every step.

Forward selection with local optimization After each forward selection step, local optimization (same to stepwise random selection method) is performed using the result of the forward selection method as a starting guess.

3 Results and discussion

Table 1 shows the statistics of results for the calibration and validation samples found by the physical and chemical analyses. The samples have a rather wide distribution.

Table 1. Statistics for reference values in each data set

	Number	Mean[%]	SD[%]	Min[%]	Max[%]
Entire set	126	30.92	14.66	5.25	71.23
Calibration set	84	30.36	13.99	5.25	63.50
Validation set	42	32.04	16.03	8.98	71.23

Figure 1 shows an example of the result of stepwise random selection for n=6 with the area-averaged reflectance spectra from all 126 samples. The number of repetitions for each selection step is 10^5. Figure 3 shows the mean scoremaps. The minimum point (the selected condition) is indicated by a magenta mark in each of the maps. Figure 2 shows an example of the plot of the evaluation values for $n=$ 1 to 10. In this result, local optimization decreases the $RMSECV$, and increases the SEP .

Figure 4 shows the result of the forward selection method, while figure 5 shows a plot of the evaluation values. Although the $RMSE_c$ continues to decrease at least until 10 filters, the SEP begin to increase after 6 filters, and the $RMSECV$ is not improved after 3 filters. Also the 4-7th filter is exactly similar, it's impractical to implement a real optical filter individually.

Figure 6 shows the modifications achieved by local optimization after each forward selection step. It appears that there are few or no modifications. Figure 7 shows a plot of the evaluation values. Local optimization decreases the $RMSECV$ slightly, however the SEP is increased. This might indicate that the local optimization causes overfitting.

In summary, local optimization decreases the $RMSECV$, however, it does not necessarily mean the model's accuracy improves. In this case, 2 filters might be enough for prediction.

In previous work [1], which uses hyperspectral data and PLS1 regression, the SEP is 4.81. In that study, spectral correction (multiplicative scatter correction) was used. Because the purpose is to implement a real optical filter, we use raw reflectance spectra without any spectral correction. We achieve a minimum SEP around 5.0. This could indicate that our results demonstrate sufficient accuracy.

In this study, similar $RMSECV$ curves but different SEP curves are obtained. Accuracy has not improved even we increase the number of filters to more than 2. This might mean that the prediction of the fat content is rather easy, because it was high in fat.

In future work, we will apply these methods to the estimation of fatty acid content, which is more difficult to predict than fat content because the content is much lower. Also, we will apply imaging to visualize the food composition. To improve the accuracy, we will add a combination of filter output values to the estimation model, and perform selection of them.

References

1. Kobayashi, K., Matsui, Y., Maebuchi, Y., Toyota, T., Nakauchi, S.: Near infrared spectroscopy and hyperspectral imaging for prediction and visualisation of fat and fatty acid content in intact raw beef cuts. J. Near Infrared Spectrosc.,18,301–315(2010)
2. Nishino, K., Nakamura, M., Matsumoto, M., Tanno, O., Nakauchi, S.: Optical Filter for Highlighting Spectral Features Part I: Design and Development of the Filter for Discrimination of Human Skin With and Without an Application of Cosmetic Foundation. Optics Express, 19, 6020–6030(2011)
3. Nishino, K., Nakamura, M., Matsumoto, M., Tanno, O., Nakauchi, S.: Optical Filter Highlighting Spectral Features Part II: Quantitative Measurements of Cosmetic Foundation and Assessment of their Spatial Distributions under Realistic Facial Conditions. Optics Express, 19, 6031–6041(2011)
4. Wilkinson, L., Dallal, G.E.: Tests of significance in forward selection regression with an F-to enter stopping rule. Technometrics.,23,377–380(1981)

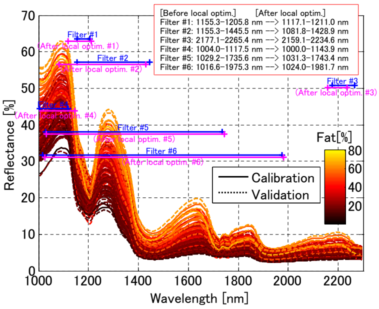


Fig. 1. An example of the result of stepwise random selection (n = 6)

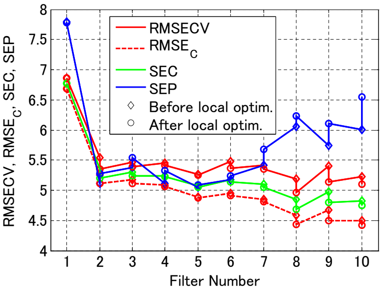


Fig. 2. An example of individual error measures (for n=1-10) for stepwise random selection

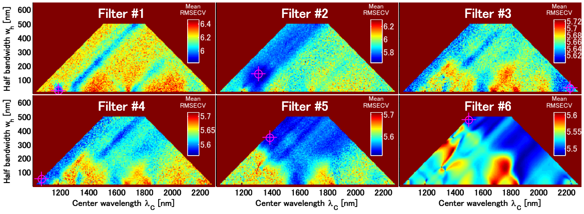


Fig. 3. An example of an evaluation map of stepwise random selection (n = 6)

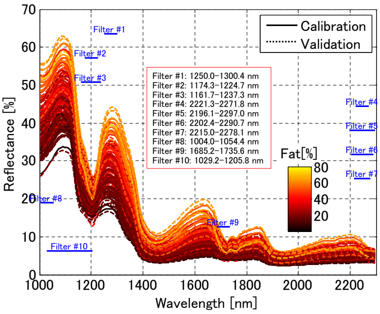


Fig. 4. The result of forward selection (n=10)

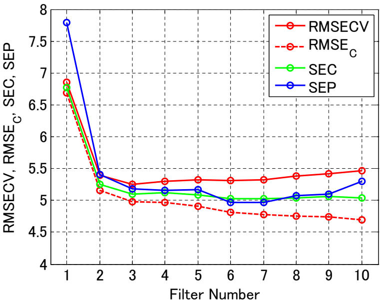


Fig. 5. Error measures for forward selection

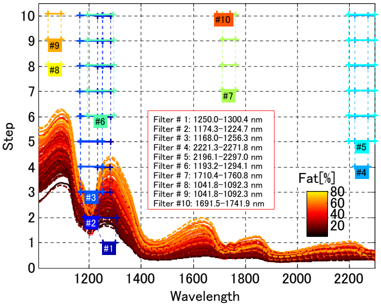


Fig. 6. Modifications of the BPFs found by local optimization after each forward selection step (n=1-10)

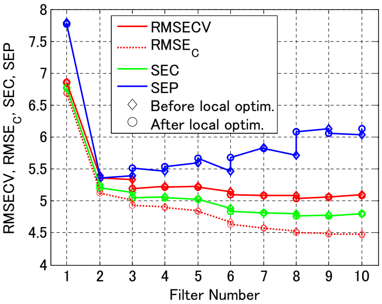


Fig. 7. Error measures for forward selection with local optimization