Employing NIR-SWIR hyperspectral imaging to predict the smokiness of scotch whisky

Julius Tschannerl[1], Jinchang Ren[1,3], Frances Jack[2], Stephen Marshall[1] and Huimin Zhao[3,4]

[1] University of Strathclyde, Hyperspectral Imaging Centre, Dept of EEE 16 Richmond St, Glasgow G1 1XQ, UK
[2] Scotch Whisky Research Institute, The Robertson Trust Building, Research Avenue North, EH14 4AP Edinburgh, UK
[3] Guangzhou Key Laboratory of Digital Content Processing and Security Technologies, Guangzhou, 510665, China
[4] School of Electronic Information, Guangdong Polytechnic Normal University, Guangzhou, China

Abstract Scotch Whisky makes a significant contribution to the UK's food and drinks export. The flavour of this high quality spirit is derived naturally from the whisky making process, with smoky aromas being a key character of certain Scotch whiskies. The level of smokiness is determined by the amount of phenolic compounds in the spirit. Phenols are introduced by exposing the barley malt to peat smoke during the kilning process. The current techniques to determine the levels of phenols, such as High Performance Liquid Chromatography (HPLC), are time consuming as they require distillation of the malt prior to analysis. To speed up this process and enable real-time detection before processing, the possibilities of Near-infrared to Short-wave-infrared (NIR-SWIR) Hyperspectral Imaging (HSI) to detect these phenols directly on malted barley are explored. It can be shown that via regression analysis, various levels of phenol concentration used as working levels for whisky production could be estimated to a satisfying degree. To further optimise industrial application, a hyperspectral band selection algorithm is applied that yields good results and reduces computational cost and may open possibilities to employ multispectral rather than hyperspectral cameras in future applications.

Keywords: hyperspectral imaging, scotch whisky, near-infrared, band selection

1 Introduction

Over the last couple of years, an increasing interest in Hyperspectral Imaging (HSI) for applications other than remote sensing can be observed. Applications include e.g. food inspection [1], medical applications [2] and artwork inspection [3]. The main reason for its popularity is its rapid data acquisition and the non-destructive nature of the same. As opposed to regular spectroscopy techniques, data is acquired at a relatively high frame rate incorporating not only spectral but also spatial information. HSI is able to gain information about the chemical composition of the imaged objects without altering the integrity of the objects. The potential as a non-destructive, real-time chemical analysis technique stirs an increasing industrial interest in HSI as it can be seemlessly integrated into the processing chain.

Scotch whisky is a high quality spirit drink exclusively produced in Scotland in a manner specified by law. According to data collected by the Scotch Whisky Association from January to December 2014, 99 million cases of 12 40% vol. bottles were exported which made up around a quarter of the UK food and drink exports [4]. These figures justify a high interest in maintaining its high quality standards.

One distinct feature of certain Scotch Whiskies are their smoky characteristics. Large parts of Scotland are covered by peat bogs. Dried peat has a long history of being used as a fuel in Scotland. Exposing malted barley that is used for whisky production during the kilning process to the smoke of burning peat introduces the typical smoky aroma. The smoking process can be seen in Figure 1.1. Phenolic compounds in the smoke adhere to the surface of the grain and carry through the process into the spirit. The phenol levels are used as a marker to the degree of smokiness of whisky. The following eight phenolic compounds can be found in scotch whisky. Phenol, Guaiacol, m/p-Cresol, o-Cresol, 4-Methyl-Guaiacol, 4-Ethyl Guaiacol and 4-Ethyl Phenol. The levels are measured in mg/kg or ppm respectively and measured is the total number, i.e. the sum of the concentrations of the individual compounds. Phenol levels are varied depending on desired smokiness. It is important

(a) (b)

Figure 1.1: (a) Burning peat (b) Barley being exposed to peat smoke

from a flavour point of view to be able to control these levels, to avoid having to blend final spirits to obtain the aroma required. Measuring the levels of phenols in malt currently requires a pre-processing step. The malted barley is finely ground, water added and then steam distilled. The levels of phenols in the distillate are measured using High Performance Liquid Chromatography (HPLC) or other colorimetric methods. It would be of significant benefit to the whisky industry to be able to determine the phenol levels directly in the malted barley, before any processing has been applied.

Phenolic compounds in grapes have previously successfully been detected by HSI [5]. This work aims to explore possibilities of determining phenol levels in malted barley by means of Near-infrared to Short-wave-infrared (NIR-SWIR) HSI in a potentially industry based application. In addition to a detection before processing of the malt, HSI offers a potential real-time detection of phenol levels, whereas HPLC needs at least a couple of days in a lab to be performed. To show this, the following things are explored in this study: By means of HSI and multivariate data analysis, it is attempted to determine different levels of phenol concentration in barley malt. To further optimise the process for industrial application, a band selection algorithm is evaluated for suitability on this data that enables faster processing and multispectral data acquisition in

future applications.

2 Sample preparation and data acquisition

As mentioned above, the peatiness of the barley is measured in ppm. This study looked initially at difference between extreme samples; unpeated malt with 0ppm phenols versus heavily peated malt containing 120-130ppm total phenols. The desired working levels for the actual whisky production ranges from 0 to about 50ppm where these levels are usually divided into three classes as follows:

1. **Lightly peated:** 1 - 5ppm
2. **Medium peated:** 6 - 15ppm
3. **Heavily peated:** 16 - 50ppm

These levels are created by mixing unpeated barley malt with the very heavily peated to equivalent parts. The concentration of the very heavily peated malt is determined by HPLC. Where HPLC measures the average concentration of phenols in a batch of barley grains, HSI can only measure spectra on various spatial points. As seen in Figure 1.1, the barley is spread out in a large pool and smoked from below and each individual grain is exopsed to the smoke to a different degree. The amount of phenols that adhere to the surface also vary within one grain depending on its orientation and batches with a low amount of phenols only include very little grains with phenols on the surface. This needs to be considered when measuring the concentration with an imaging device as only the surface of the grains can be measured.
For initial tests of detectability, a batch of grains with no phenols and barley with a concentration of about 120ppm have been produced and imaged. Three different batches of barley with 3, 12 and 30ppm have then been created to cover the three working levels of phenols. The barley has been placed in a tray as seen in Figure 1.2 and imaged with a pushbroom NIR-SWIR HSI camera. The camera has a spectral range of 900 - 1700 nm and scans 320 pixels with 256 bands per line. The samples have been illuminated with two halogen lightbulbs on each side which have been diffused to minimise specular reflections and shadows. However, due to the curved surface of the grains, differences in intensity and different

spectral reflections cannot be completely avoided. Six different samples of each level have been imaged. Since the measured phenols will most likely vary strongly between the pixels, the mean of subsets of all pixels per sample was taken. The subsets consist of randomly selected pixels within each samples to avoid localised effects of phenol accumulation or absence. 50 subsets per sample have been generated, resulting in 300 observations per phenol level.

3 Spectral pre-processing

To compensate for variations in the lighting between different samples and spatial variations along the scanned line, the raw measured signal **S** can be calibrated and converted to percent reflectance spectra **R** with the following formula, shown in [6]:

$$\mathbf{R} = \frac{\mathbf{S} - \mathbf{D}}{\mathbf{W} - \mathbf{D}} \times 100\% \tag{1.1}$$

D is a dark current image acquired by minimising the camera's sensor exposure to any radiation. This is to estimate the sensor's shot noise. **W** is a white reference imaged acquired by imaging e.g. a Spectralon plaque which exposes almost lambertian scattering over the desired spectral range. This helps to approximate the maximum reflectance measured at each wavelength.

A number of spectral pre-processing techniques have been analysed in [7] for NIR spectra including various techniques for de-noising [8]. Most popular ones include conversion to Standard Normal Variate (SNV) and 1st and 2nd derivatives using Savitzky-Golay filters. SNV transforms the data to zero mean and unit variance. This removes an additive baseline and can in theory be used to compensate for slight intensity variations due to shadow effects on uneven surfaces and compensate for scattering effects [7]. In practice however, no general optimal pre-processing procedure can be established and most researchers try various combinations and use the ones that yield the best results for the specific application. In this case, pre-treatment with SNV yielded the best regression results, as is later shown in Section 5. The SNV of a reflectance spectrum **R** is defined as follows:

$$\mathbf{R}_{SNV} = \frac{\mathbf{R} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \tag{1.2}$$

where $\boldsymbol{\mu}$ represents the mean and $\boldsymbol{\sigma}$ the standard deviation of all measured spectra.

4 Band selection

One major disadvantage of hyperspectral data is its high dimensionality. While it is often desired to maintain a relatively high spatial resolution for optimal identification of the measured objects, the spectral information can in most cases be reduced. Hyperspectral cameras cover a wide, continuous range of spectral bands but in many applications, only certain wavelengths are of particular interest. Additonally, adjacent bands usually carry redundant information and the well known curse of dimensionality may even decrease the discriminability if too many bands are included. Different feature extraction techniques [9–11], predominantly Principal Component Analysis (PCA) and variations [12] have widely been successfully applied to hyperspectral data to reduce the dimensionality. The disadvantage of feature extraction however is that new features are generated by e.g linearly combining spectral bands and the resulting features are no longer physically interpretable and linked to specific chemical properties of the imaged objects.

Various feature selection techniques have been developed over the last decades [13] that try to select an optimal subset of features for a specific application. In [14] a new unsupervised band selection algorithm for hyperspectral images is introduced. Based on information theory, the following criterion is defined to evaluate the fitness of a selected subset of features:

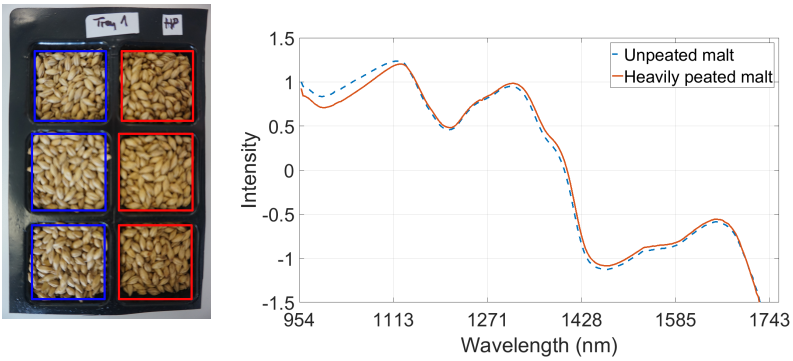
$$max \left(\sum_{m=1}^{s} H(X_{i_m}) - \frac{2}{s-1} \sum_{1 \leq m_1 < m_2 \leq s} I(X_{i_{m_1}}; X_{i_{m_2}}) \right) \tag{1.3}$$

where s is the number of desired features, $H(X_{i_m})$ the entropy of the mth feature in the subset X_i and $I(X_{i_{m_1}}; X_{i_{m_2}})$ the mutual information between the two features. The criterion maximises the information carried by each feature in the subset individually while minimising the redundant information carried by the whole subset. The criterion is therefore called Maximum-Information-Minimum-Redundancy (MIMR). To optimise the maximisation function, an adapted version of the Clonal Selection Algorithm (CSA) is employed as a heuristic. As shown in [14], the

MIMR-CSA algorithm outperforms state of the art algorithms in classification applications for hyperspectral remote sensing data. Its suitability for regression is to be tested here.

5 Data analysis and results

To visually inspect if there are any spectral differences in the NIR detectable, three batches of unpeated malt and heavily peated barley with a phenol concentration of about 120ppm have been imaged and the mean spectra pre-processed with SNV have been plotted. This can be seen in Figure 1.2. The spectra show a very similar shape but differ in intensity in some spectral regions, especially in bands between 950 and 1120nm. These differences are likely to be the result of slight shift into "red" of the peated barley. The smoked grains tend to be a bit more brown than the unsmoked grains.



(a) (b)

Figure 1.2: (a) Barley samples. Unpeated grains are on the left column and very heavily peated on the right (b) Spectra acquired by taking the mean over all three samples of each class and applying SNV as pre-treatment

Support Vector machines have been applied successfully to HSI data for classification in the past [15] and new techniques including multi-kernel learning [16] and deep-learning based approaches [17] show that there is more potential for improved algorithms. To estimate the phenol

concentration, Support Vector Regression (SVR) has been applied. SVR has been chosen as it has shown good results for regression analysis on hyperspectral data in the past [18–20], is especially capable of handling non-linearities within the data [21] and only needs a relatively small training data set to deliver robust results [22]. A Radial Basis Function kernel (RBF) has been used whose parameters have been tuned via grid search. The goal here is not to determine the most effective regression method but to prove the feasibility of phenol level estimation through regression analysis. Although the division of phenol concentrations in three levels would imply a classification approach, future applications may include a more detailed estimation of the concentration rather than grouping it in high, medium and low levels. Thus, it is to determine the effectiveness of regression for this application.

To evaluate the regression results, the Coefficient of Determination (r^2) has been calculated as follows:

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)} \tag{1.4}$$

where y_i is the actual value, $\hat{y}_i$ the predicted value, $\bar{y}_i$ the mean of all original values and n the number of samples. Additionally, the Root Mean Squared Error (RMSE) has been calculated. A 20/80% training/testing set has been used and split randomly. To compensate for statistic variations between the selected training sets, this process has been repeated 10 times. MIMR-CSA has been applied with an increasing number of features starting from two to 256. MIMR-CSA has been applied three times on each of those selected sets resulting in a total number of 30 repetitions per number of selected features. The results of the regression can be seen in Figure 1.3.

6 Conclusion and future work

While it could be shown that some spectral differences between peated and unpeated barley can be detected with NIR-SWIR HSI, it is not yet clear if these differences derive from chemical absorption of the phenol, from physical scattering or from colour differences. Scattering effects have been tried to be minimised via spectral pre-processing. It could be shown that the measured spectra of three different phenol levels can

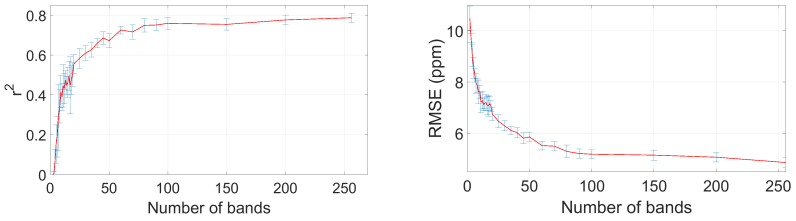


Figure 1.3: Regression results with increasing amount of selected features for phenol level prediction with SVR. (a) r^2 values stabilise around a value of 0.77 (b) The RMSE stabilises around a value of around 5ppm

to a certain extent be detected via regression analysis. It is likely that the differences detected are caused by colour information rather than spectral absorption of the phenols, but this cannot be verified at this stage. These three levels represent general working levels for whisky production. A minimum RMSE of about 5ppm is still a relatively high number considering the working range of the three classes. Additional samples, especially in more levels of concentration in between the current ones are likely to decrease this value. Applying a state-of-the-art band selection algorithm for hyperspectral data on this dataset showed that most information needed for regression lies in only about 75 of the 256 bands. This number is still too high for determining selected bands for a multispectral camera, but the computational cost can be drastically reduced. The presented results encourage us to put more work into the phenol detection via HSI.

7 Acknowledgements

This work is partially funded by the Institute of Brewing and Distilling. The images in Figure 1.1 are courtesy of the Scotch Whisky Research Institute.

References

1. D.-W. Sun, *Hyperspectral imaging for food quality analysis and control*. Academic, 2010.
2. G. Lu and B. Fei, "Medical hyperspectral imaging: a review." *Journal of biomedical optics*, vol. 19, no. 1, p. 10901, 2014.
3. A. Polak, T. Kelman, P. Murray, S. Marshall, D. J. Stothard, N. Eastaugh, and F. Eastaugh, "Use of infrared hyperspectral imaging as an aid for paint identification," *Journal of Spectral Imaging*, vol. 5, no. 1, pp. 1–10, 2016.
4. Scotch Whisky Association. (2015) Facts & Figures. Accessed: 09/12/2016. [Online]. Available: http://www.scotch-whisky.org.uk/what-we-do/facts-figures/
5. M. J. Jara-Palacios, F. J. Rodríguez-Pulido, D. Hernanz, M. L. Escudero-Gilete, and F. J. Heredia, "Determination of phenolic substances of seeds, skins and stems from white grape marc by near-infrared hyperspectral imaging," *Australian Journal of Grape and Wine Research*, vol. 22, no. 1, pp. 11–15, 2016.
6. H. Yao and D. Lewis, "Spectral Preprocessing and Calibration Techniques," in *Hyperspectral Imaging for Food Quality Analysis and Control*, 2010, pp. 45–78.
7. Å. Rinnan, F. van den Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," pp. 1201–1222, 2009.
8. T. Qiao, J. Ren, Z. Wang, J. Zabalza, M. Sun, H. Zhao, S. Li, J. A. Benediktsson, Q. Dai, and S. Marshall, "Effective Denoising and Classification of Hyperspectral Images Using Curvelet Transform and Singular Spectrum Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 119–133, 2016.
9. J. Ren, J. Zabalza, S. Marshall, and J. Zheng, "Effective feature extraction and data reduction in remote sensing using hyperspectral imaging [Applications Corner]," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 149–154, 2014.
10. J. Zabalza, J. Ren, J. Zheng, H. Zhao, C. Qing, Z. Yang, P. Du, and S. Marshall, "Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging," *Neurocomputing*, vol. 185, pp. 1–10, 2016.
11. J. Zabalza, J. Ren, J. Zheng, J. Han, H. Zhao, S. Li, and S. Marshall, "Novel Two-Dimensional Singular Spectrum Analysis for Effective Feature Extraction and Data Classification in Hyperspectral Imaging," *IEEE*

Transactions on Geoscience and Remote Sensing, vol. 53, no. 8, pp. 4418–4433, 2015.

12. J. Zabalza, J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han, "Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 112–122, 2014.
13. V. Kumar and S. Minz, "Feature Selection: A literature Review," *Smart Computing Review*, vol. 4, no. 3, pp. 211–229, 2014.
14. J. Feng, L. Jiao, F. Liu, T. Sun, and X. Zhang, "Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images," *Pattern Recognition*, vol. 51, pp. 295–309, 2016.
15. F. Melgani and L. Bruzzone, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, 2004.
16. L. Fang, S. Li, W. Duan, J. Ren, and J. A. Benediktsson, "Classification of Hyperspectral Images by Exploiting Spectral & Spatial Information of Superpixel via Multiple Kernels," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6663–6674, 2015.
17. J. Han, D. Zhang, X. Hu, L. Guo, J. Ren, and F. Wu, "Background Prior-Based Salient Object Detection via Deep Reconstruction Residual," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 8, pp. 1309–1321, 2015.
18. T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin, and S. Marshall, "Quantitative prediction of beef quality using visible and nir spectroscopy with large data samples under industry conditions," *Journal of Applied Spectroscopy*, vol. 82, no. 1, 3 2015.
19. ——, "Singular spectrum analysis for improving hyperspectral imaging based beef eating quality evaluation," *Computers and Electronics in Agriculture*, vol. 115, pp. 21–25, 2015.
20. M. Sun, D. Zhang, Z. Wang, J. Ren, B. Chai, and J. Sun, "What's Wrong with the Murals at the Mogao Grottoes: A Near-Infrared Hyperspectral Imaging Method." *Scientific reports*, vol. 5, p. 14371, 2015.
21. I. Steinwart and A. Christmann, *Support vector machines*. Springer, 2008.
22. C. J. C. Burges, *Geometry and invariance in kernel based methods*. Cambridge, MA: MIT Press, 1999.