Chemometric pre-treatments for correcting the effect of sample morphology in near infrared spectral imaging

C. Esquerre,[a,b] A.A. Gowen,[b] C.P. O'Donnell,[b] J. Burger[c] and G. Downey[a]

[a]*Teagasc, Ashtown Food Research Centre, Dublin 15, Ireland*
[b]*Biosystems Engineering, School of Agriculture, Food Science and Veterinary Medicine, University College Dublin, Dublin 4, Ireland*
[c]*BurgerMetrics SIA, Jelgava, Latvia*

Keywords: spectral imaging, near infrared, chemometric pre-treatment

Introduction

Hyperspectral imaging emerged as a remote sensing technology, but nowadays has become a powerful tool for different fields of science and industry; for example agricultural, pharmaceutical, mining and soil survey. Natural variability in the shape and size of biological materials presents a difficulty in the use of near infrared imaging to assess their quality. As an example, Figure 1a shows a false colour image of an apple from hyperspectral NIR data acquired in the range 950–1650 nm.

It is possible to observe a bright ring and dark areas as result of sample morphology, causing spectral variability according to position. This is more evident in the 1st principal component score image of the same data [Figure 1(b)]. The origin of the variability in the spectra according to the spatial position in non-flat samples could be related to three key factors: (a) shadows and other features produced by uneven illumination (b) the distance and angle between the sample surface and (c) the detector, which cannot be focussed simultaneously for multiple image depths and wavelengths. Recently, different geometrical corrections have been proposed to reduce the morphology effect in hyperspectral images, while preserving the chemical information.[1,2] These methods assume that the sample reflects the radiation as a Lambertian body or that the reflected energy is a function of the height of the sample. The methods successfully correct the morphology effect in the spectra of individual and relatively simple shapes (i.e. spheres and ellipsoids). Another approach is to use chemometric pre-treatment to correct the morphology effect on the spectra. Chemometric pre-treatments work on the spectral dimension of hyperspectral data and can correct the spectra for baseline shift (detrend, derivatives), multiplicative effects (SNV, MSC, EMSC, normalisation) and be used in combination. In this study, a selection of

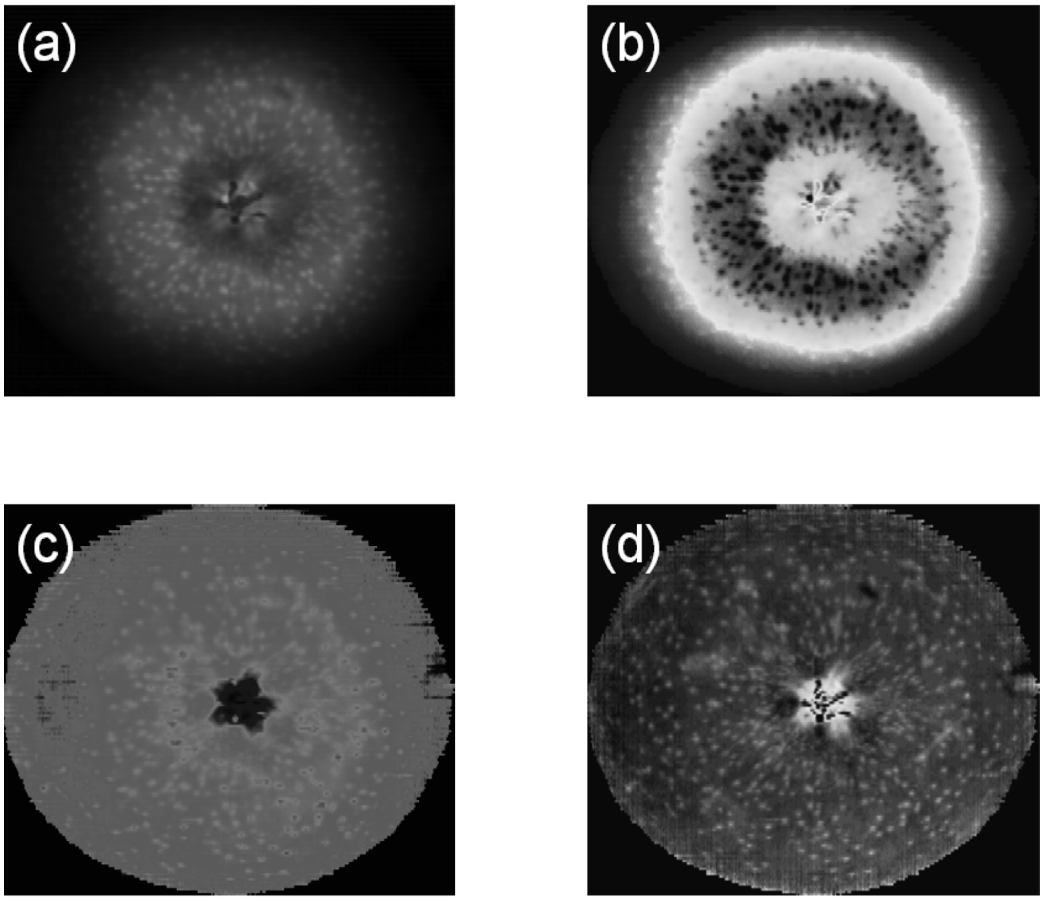

Figure 1. Apple hyperspectral image scanned 950–1650 nm. (a) grey scale image, in which it is possible to observe a shiny central ring and dark edge. (b) PC 1 score image of raw spectra. (c) PC 1 image of median normalised spectra. (d) PC 1 score image of SNV pre-treated spectra.

chemometric pre-treatments were evaluated for morphological correction of the spectra obtained in hyperspectral imaging experiments.

Materials and methods

Samples

A sheet of pink paper and a sheet of blue cardboard were selected because they can be scanned in flat and cylindrical configuration (Fig. 2); appropriate spectral pre-treatments should produce the same spectra in both configurations.

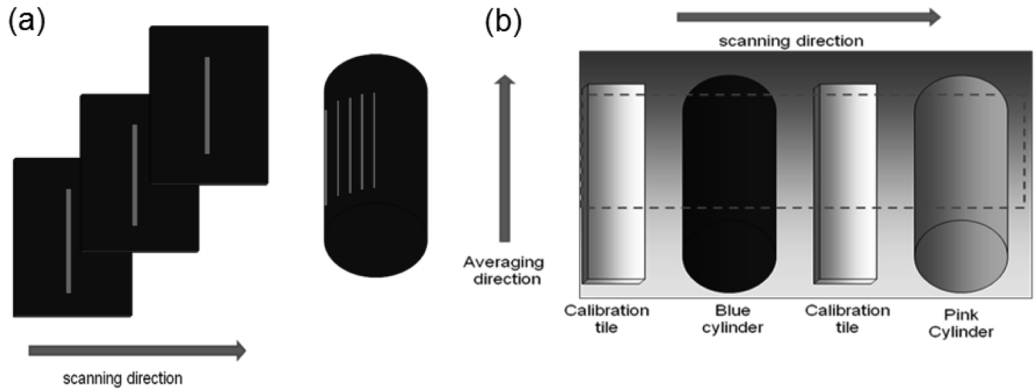

Figure 2. Principal component samples (a) sub-sampling areas from the three flat images and cylinders (b) region where the principal component analysis was performed and averaging direction to obtain the score profile.

The samples were scanned in flat configuration at three height positions: 0, 4 and 6.5 cm from the moving base of the hyperspectral system and were labelled as bottom, middle and top respectively. Afterwards, the samples were scanned in a cylindrical configuration (6.5 cm diameter).

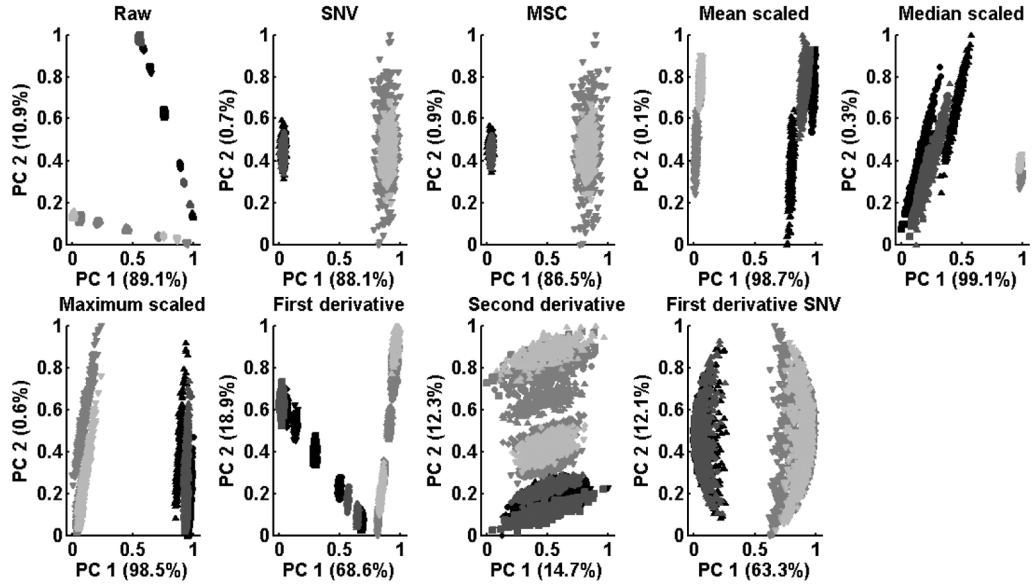

Figure 3. Score plots (PC1 vs PC2) of raw and pre-treated spectra. Colour code: samples in increasing order of brightness: blue cardboard cylinder, blue cardboard flat, pink paper cylinder and pink paper flat. Scores were scaled between 0 and 1.

Hyperspectral system

A line-scanning Vis-NIR hyperspectral imaging system (DV Optics, Italy) was used to acquire spectral data in the range of 400 to 1000 nm at 5 nm intervals, using a cylindrical integrator to produce diffuse illumination over the scanning line.

Chemometric pre-treatments

The chemometric pre-treatments compared were: Normalisation by mean, median and maximum of each spectrum, standard normal variate (SNV), multiplicative scatter correction (MSC), first derivative (1st derv), second derivative (2nd derv), and, as an example of a combination, first derivative (1st derv) on SNV pre-treated data. The data analysis was performed with Matlab (The MathWorks, Inc. USA). Principal component analysis was the selected method to evaluate the performance of the pre-treatments; standard deviation of pre-treated data was not practical because pre-treated data were on different scales. The coefficient of variability was not useful either, because some bands had means close to zero.

Sixteen sub-samples of 3×100 pixels were selected, one for each flat hypercube and five from each cylinder as is shown in Figure 2(a). For a further comparison of pre-treatments performance, the average score of 100 pixels from the hyperspectral data of both cylinders was calculated as a score profile [Figure 2(b)].

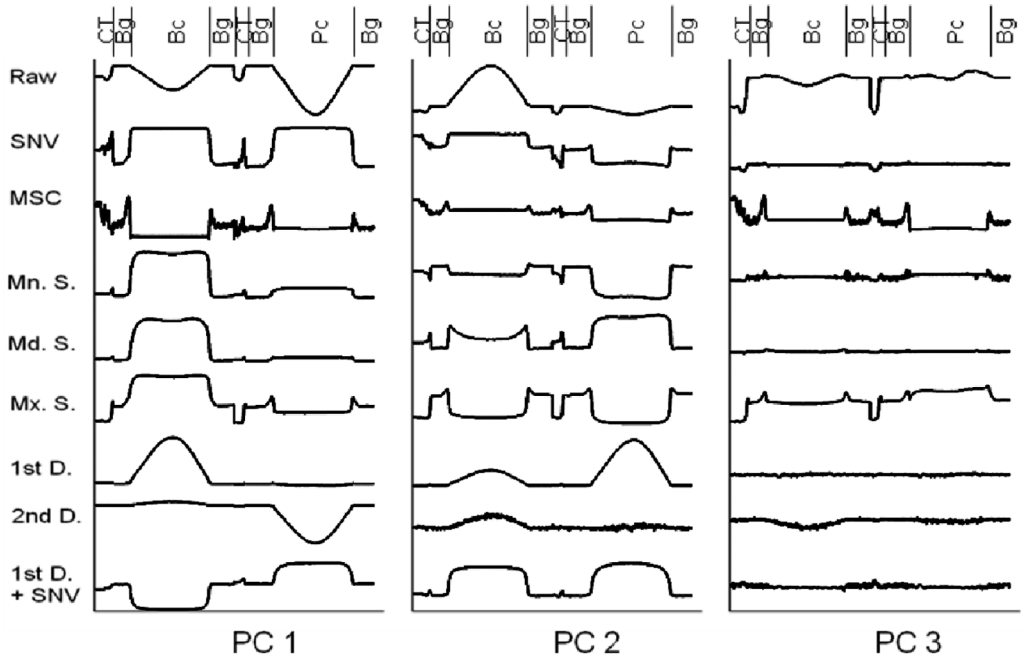

Figure 4. Score profiles of average raw and pre-treated spectra. CT = calibration tile; Bg = background; Bc = blue cylinder; Pc = pink cylinder. Scores were scaled between 0 and 1.

Results and discussion

The normalised (between one and zero) principal component one and two score plots for the sixteen subsamples are shown in Figure 3 for raw and pre-treated data.

The light blue clusters correspond to pink paper in flat configuration at the top, middle and bottom position. The dark blue clusters correspond to the five sub-samples of the hypercube of the pink paper cylinder. The score plot of raw spectra shows that the main source of variability described by PC1 is more related to the distance between the target and the detector than to the difference between materials, representing 89.1 % of the variability in the raw data set. All pre-treatments studied, with the exception of second derivative, produced corrected spectra in which the main source of variability in the data (PC 1) may be related to the difference between the materials studied.

Score profiles obtained from raw and pre-treated data are shown in Figure 4, in which it is possible to observe that SNV and MSC pre-treated spectra of the cylinders have sharper edges compared with the score profiles of normalised spectra; thus, these pre-treatments (SNV and MSC) could better correct the spectra of the borders of the samples.

Conclusion

Chemometric pre-treatments successfully corrected the morphology effect in the spectra from NIR spectral images, preserving chemical information without having previous information of shape or size.

Acknowledgement

Financial support from the Irish Department of Agriculture, Fisheries and Food under the FIRM Programme is gratefully acknowledged.

References

1. W. Wang and J. Paliwal, in *Correction of Curvature-Induced Spectral Variability in Hyperspectral Images of Wheat Kernels*, ASAE Annual Meeting, (2005).
2. J. Gómez-Sanchis, E. Moltó, G. Camps-Valls, L. Gómez-Chova, N. Aleixos and J. Blasco, *J. Food Eng.* **85,** 191 (2008).