

Subpixel detection of peanut particles in wheat flour using near infrared hyperspectral imaging

Antoine Laborde¹, Benoit Jaillais², Anthony Boulanger¹, Delphine Jouan-Rimbaud Bouveresse³, and Christophe B.Y. Cordella³

¹ GreenTropism,

75116 Paris, France

² StatSC INRA/ONIRIS,

44322 Nantes, France,

³ UMR914 PNCA AgroParisTech,

75231 Paris, France

Abstract Hyperspectral imaging in near-infrared region (NIR) is a powerful tool for characterization and detection in food industry. In particular, the scan of powders is a subject of interest for adulteration evaluation. However, such samples involve intimate mixture hence complex non linear effect in the reflectance signal. In this study, Adaptive Matched Subspace Detector (AMSD) is implemented for detecting peanut flour adulteration in wheat flour. The method consists of a hypothesis test based on the linear mixing model. This is compared with a non supervised technique based on Principal Component Analysis rejection method. Results show that AMSD performs the best by detecting adulterated pixels in samples with global concentration from 20% down to 0.02%. A coefficient of determination of 0.90 is obtained between the number of detected pixels and the global concentration of samples. PCA rejection method shows relevant but insufficient results by detecting much fewer adulterated pixels than AMSD. This study shows that the implementation of AMSD is successful and more efficient than rejection method based on the inner product variability.

Keywords: Subpixel detection, peanut, hyperspectral.

DOI: 10.58895/ksp/1000087509-8 erschienen in:

OCM 2019 - 4th International Conference on Optical Characterization of Materials, March 13th – 14th, 2019, Karlsruhe, Germany

DOI: 10.58895/ksp/1000087509 | <https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000087509/>

1 Introduction

Near-infrared hyperspectral imaging (HSI) is a technique that provides both spatial and spectral information for a sample. The near-infrared range is a particularly interesting since no sample preparation is required and various organic information can be studied. Thanks to this, HSI has been recognized as an emerging tool for controlling food safety for a decade [1]. For example, HSI has been successfully employed to detect melamine adulteration in milk powder [2] [3] [4]. Melamine and milk particles were detected in the case of an intimate mixing and for pixel size bigger than particles. In this context, the resulting spectral signatures of the mixture is nonlinear and complex to explain [5]. However, since melamine and milk powder spectral signatures are very different, they can be directly used for the implementation of a detection method. This is done using Principal Component Analysis (PCA) [3], Partial Least Square Regression (PLSR) [4] or by spectral similarity analysis [2].

As the allergic population seems to grow, allergens contamination is a hot topic for food issue as well. Peanut adulteration in wheat flour has been achieved using HSI and PCA [6]. In this case, peanut particles were bigger than the pixel so that no intimate mixing occurs in the powder. However, the detection of peanut in wheat flour is more difficult since spectral signatures are closer than the melamine and milk situation.

Subpixel detection is a subject of interest in remote sensing application. For Earth observation, pixel field of view are about several meters so that each pixel may contain spectral contributions of several end members. As long as the mixing only occurs in the sensor, linear mixing model theoretically holds. However, for multi-layer samples or intimate mixing structure, the linear assumption is not valid. On the other hand, nonlinear models require a deep knowledge about the samples that are very difficult to obtain [5]. As a consequence, linear mixing model is often employed using additional interaction terms. The Adaptive Matched Subspace Detector (AMSD) is derived from the linear mixing model (LMM) using a hypothesis test. This detector is used at the subpixel context since each pixel is treated individually without spatial information [7].

In this study, we propose to address the detection of peanut particles in wheat flour at the subpixel level. AMSD technique is compared with a PCA model based rejection method to detected adulterated pixels of flour mixing.

2 Theory

2.1 AMSD

The LMM describes each pixel of a hyperspectral image as a linear combination of component spectra or endmembers (cf. equation 8.1).

$$x = \sum_{k=1}^M a_k s_k + w \quad (8.1)$$

where

x is the spectrum of the pixel;

s_k are the spectra of the M endmembers;

a_k are the abundances of the M endmembers;

M is the number of endmembers;

w is a vector that accounts for the lack-of-fit of the LMM.

In the LMM, the subspace S generated by the s_k vectors should be representative of all the chemical species in the mixing. For a detection problem, S is divided in two parts : the background subspace S_b and the target subspace S_t . The matrix S_b contains the spectra of the endmembers considered as background and S_t contains the targeted endmembers.

For the AMSD, the detection problem is a hypothesis test. In each case, a LMM is assumed according to the presence of the target. For the null hypothesis, the spectrum of the pixel is assumed to lie in the background subspace. The alternative hypothesis assumes x lies in the union of both subspaces S_t and S_b (cf. 8.2).

$$\begin{aligned} H_0 : x &= S_b a_b + w \\ H_1 : x &= S_t a_t + S_b a_b + w \end{aligned} \quad (8.2)$$

where we assume $w \hookrightarrow \mathcal{N}(0, \sigma_w^2 I)$.

Designing a detector for the hypothesis test can be done using the likelihood ratio (LR). Since the abundance vectors a_t and a_b as well as the noise variance σ_w^2 are unknown, the generalized likelihood ratio is used [7]. As a result, the ratio between the error sum of squares (SSE) under each hypothesis is used as a detector (cf. 8.3).

$$d_{AMSD}(x) = \frac{SSE_{H_0}}{SSE_{H_1}} - 1 \quad (8.3)$$

where

$$\begin{aligned} SSE_{H_0} &= (x - S_b a_b)^T (x - S_b a_b), \\ SSE_{H_1} &= (x - S_t a_t + S_b a_b)^T (x - S_t a_t + S_b a_b). \end{aligned}$$

2.2 PCA rejection method

In this study, PCA provides orthogonal components for modeling wheat pure sample: V_b . Only the k first components are kept to retain the main spectral variance: V_{b_k} . As a result, an orthogonal projector can be constructed (cf. 8.4).

$$P_{b_k}^\perp = 1 - V_{b_k}^T V_{b_k} \quad (8.4)$$

Using the projector $P_{b_k}^\perp$, the residual of each pixel with respect to the PCA model based on the pure wheat sample can be computed. Finally, the sum of squares (SS) of these residuals are used as a metric for the pixel rejection method (cf. 8.5).

$$d_{rejection}(x) = SS(xP_{b_k}^\perp) \quad (8.5)$$

3 Material and methods

3.1 Samples

Defatted peanut flour and white wheat flour were bought and mixed together in different proportions. A total of 8 weight concentrations of peanut were obtained: 20%, 10%, 5%, 2%, 1%, 0.5% 0.2% and 0.02%.

For each concentration, 3 replicates of approximately 13 g of flour were measured as well as one pure peanut flour sample and one pure wheat flour sample. Each sample was prepared using a precision balance (accuracy of 0.01g) and by manually shaking the resulting mix. Samples were put into a polylactic (PLA) sample holder of 7 mm depth for the hyperspectral measurement.

3.2 Hyperspectral acquisition

A SWIR camera from SPECIM (SPECIM, Spectral Imaging Ltd., Oulu, Finland) was used for hyperspectral measurement. The samples were lightened by four halogen lamps at 58cm from the camera. Black measurement and white reference were acquired in order to obtain the normalized reflectance.

3.3 Detector design

For designing the AMSD, non negative matrix factorization (NMF) was applied on the pure wheat flour and peanut flour sample. The k first components were used to design the matrix S_b and S_t . When the detector is applied on a test image, a non negative least square resolution is applied on each pixel to calculate the corresponding abundance vectors a_t and a_b . Then, the estimation of x under each hypothesis is used for the calculation of SSE_{H_0} and SSE_{H_1} . The resulting value of $d_{AMSD}(x)$ must be finally compared to a threshold value η which is obtained according to equation 8.6. This boils down to minimizing the probability of false alarm on the test images.

$$\eta = \max_{x \in X_0} d_{AMSD}(x) \quad (8.6)$$

where X_0 is the matrix containing the spectra from the reference pure wheat flour image.

Similarly, the PCA rejection method gives a final detection by thresholding $d_{rejection}(x)$. The threshold value is obtained from 8.6 by replacing $d_{AMSD}(x)$ with $d_{rejection}(x)$.

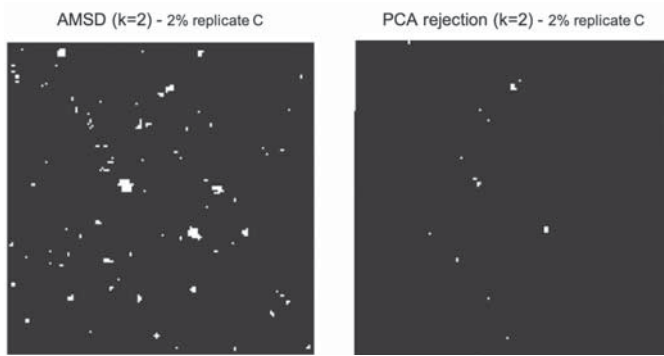


Figure 8.1: Detection map for wheat flour mixed with 2% of peanut flour (replicate C) using the AMSD (left) and the PCA rejection method (right). White pixels: detection, black pixels: no detection.

4 Results and discussion

Even though the two methods are very different, they both provide consistent results. Figure 8.1 illustrates the fact that the detection maps obtained on a same sample show similar spatial results. Indeed, even if the PCA rejection method detects much fewer pixels than AMSD does, the position of the detected pixels correspond to detected clusters on the AMSD detection map.

Table 8.1 exhibits a high correlation between the number of detected pixels and the expected concentration of peanut for the AMSD method. Results also show that AMSD is robust to the choice of k . On the other side, the PCA rejection method exhibits lower R^2 score as well as a high dependency to the value of k .

These results are explained by the fact that the PCA rejection method generally detects much fewer pixels than AMSD. This can be seen as an example on Figure 8.1. Indeed, in most cases, PCA rejection works well for location where clusters of peanut particle are present. In this case, there is more chance that a pixel has an almost full contribution of peanut. More generally, PCA rejection method has the advantage to not be dedicated to one adulterant spectral signature. In return, the detection method is less sensitive to the peanut adulteration. Adding

more than 2 components to the rejection model (choosing $k > 2$) makes R^2 worst. This is because the accepted variability increases so that peanut particles are comprised in it and thus, not detected.

For the subpixel detection problem, the contribution of the target may be hidden by the background. If their spectral signatures are similar, the detection may be difficult. In this case, supervised target detection is necessary. Indeed, the change in the spectral signal may be of the same order than the background variability which makes unsupervised method fail.

Table 8.1: Coefficient of determination (R^2) between the number of detected pixels in each image and the corresponding concentration of peanut flour.

k	1	2	3
R^2_{AMSD}	0.90	0.90	0.89
R^2_{PCA}	0.55	0.61	0.19

5 Conclusion

The study shows the AMSD approach is relevant for subpixel detection problem. Even if the complexity of the mix may involve non linear effects on the signal mixture, the hypothesis test based linear mixing models is relevant. In contrast, the PCA rejection method gives interesting results but appears to be insufficient to detect the majority of the adulterated pixels. Indeed, in this model, outliers are assumed to have higher spectral differences than the inner product variability. In the case of peanut and wheat flour, this assumption appears to be questionable. Finally, AMSD enables to find adulterated pixels down to a global concentration of 0.02% but the detector sensitivity at the pixel level is still to be investigated. The high R^2 score obtained between the number of detected pixels and the true sample concentration shows that AMSD provide a successful and relevant detection of peanut in wheat at the subpixel level.

References

1. A. Gowen, C. O'Donnell, P. Cullen, D. G., and J. Frias, "Hyperspectral imaging an emerging process analytical tool for food quality and safety control," *Trends in Food Science and Technology*, vol. 18, pp. 590–598, 2007.
2. X. Fu, M. S. Kim, K. Chao, J. Qin, J. Lim, L. H., A. Garrido-Varo, P.-M. D., and Y. Ying, "Detection of melamine in milk powders based on nir hyperspectral imaging and spectral similarity analyses," *Journal of Food Engineering*, vol. 124, pp. 97–104, 2014.
3. J. A. Fernandez Pierna, D. Vincke, P. Dardenne, Z. Yang, L. Han, and B. V., "Line scan hyperspectral imaging spectroscopy for the early detection of melamine and cyanuric acid in feed," *Journal of Near Infrared Spectroscopy*, vol. 22, pp. 103–112, June 2014.
4. J. Lim, G. Kim, C. Mo, M. S. Kim, K. Chao, J. Qin, X. Fu, I. Baek, and B. Cho, "Detection of melamine in milk powders using near-infrared hyperspectral imaging combined with regression coefficient of partial least square regression model," *Talanta*, vol. 151, pp. 183–191, 2016.
5. J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
6. P. Mishra, P. Herrero-Langreo, A. and Barreiro, J. Roger, B. Diezma, N. Gorretta, and L. Lleo, "Detection and quantification of peanut traces in wheat flour by near infrared hyperspectral imaging spectroscopy using principal-component analysis," *Journal of Near Infrared Spectroscopy*, vol. 23, pp. 15–22, February 2015.
7. D. Manolakis, C. Siracusa, and S. G., "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1392–1409, July 2001.