Clustering methods for analysis of processed animal proteins adulteration based on FT-NIR microscopic imaging data

Xunpeng Jiang, Chengxu Lü, Zengling Yang, Xian Liu, Lujia Han

College of Engineering, China Agricultural University, Beijing 100083, China
State Key Laboratory of Animal Nutrition, Beijing 100091, China

1. Introduction

Fourier transform near infrared (FT-NIR) microscopic imaging method is presented as a new strategy for detecting feedstuff and ingredients, such as processed animal proteins (PAPs). The challenge is to develop adequate chemometrics methods to extract useful information from FT-NIR imaging data. Clustering, as unsupervised classification methods, is used to identify different components in inhomogeneous data sets. In this work, three kinds of clustering algorithms, namely k-means[1], fuzzy c-means[2], and Gaussian model-based[3], were used and compared to extract different PAPs classification information from FT-NIR microscopic imaging data set.

2. Materials and Methods

2.1. Sample Preparation

One fish meal (FM) and one meat and bone meal(MBM) were collected. FM and MBM particles were pre-arranged on the sample holder as calibration set. Two FM samples contaminated with MBM at concentrations of 30% and 20% by weight, respectively, were prepared as validation set.

2.2. FT-NIR Microscopic Imaging Data Collection

FT-NIR microscopic images were acquired with a PerikinElmer Spotlight400 FT-IR imaging system. The images obtained had a resolution of 25 μm (each pixel covered an area of 25 μm×25 μm). Each spectrum was across the wavenumber range 6000 cm^{-1}-5400 cm^{-1}, with 8 cm^{-1} data resolution.

2.3. Data Processing

Clustering algorithms used the scores given by principal component analysis as input data. All data treatment was performed on MATLAB 2008a (The MathWorks, Inc., USA).

3. Results and Discussion

3.1. Calibration Set

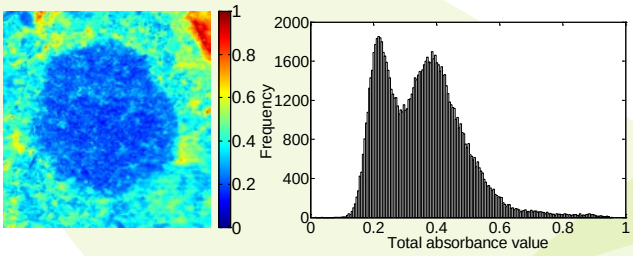


Fig. 1 (Left) Total absorbance image of MBM and FM
(Right) Total absorbance pixel frequency distribution

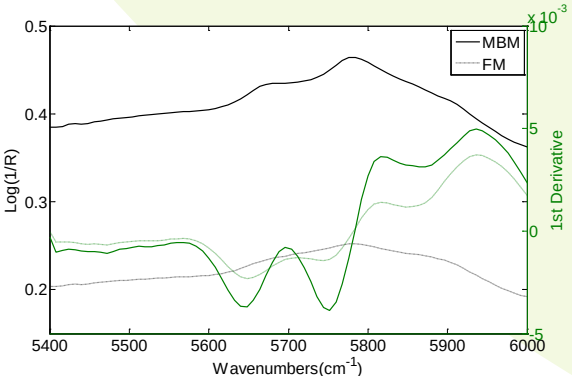


Fig. 2 Raw spectra and first derivative spectra of MBM and FM

As Fig. 1 shows, there was a little difference between FM and MBM at total absorbance image. The two bands in the histogram indicated that FM was distinguishable from MBM. Raw spectra and first derivative spectra of MBM and FM revealed the difference at the spectral range of 6000-5400 cm^{-1} (Fig. 2).

(a) K-means (b) Fuzzy c-means (c) Gaussian model-based

Fig. 3 Binary classification results of FT-NIR microscopic imaging data for pre-arranged sample particles (White pixel – MBM; Dark pixel – FM)

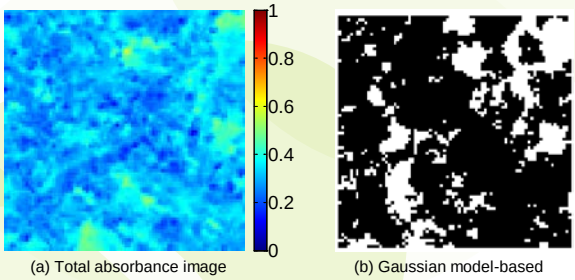
Table 1 Evaluation measures of different clustering algorithms.

	K-means	Fuzzy c-means	Gaussian model-based
sensitivity	0.89	0.88	0.94
specificity	0.98	0.98	0.97
purity	0.92	0.92	0.95

Gaussian model-based method correctly classified 95% of samples in the calibration set, which was better than the other methods (Table 1).

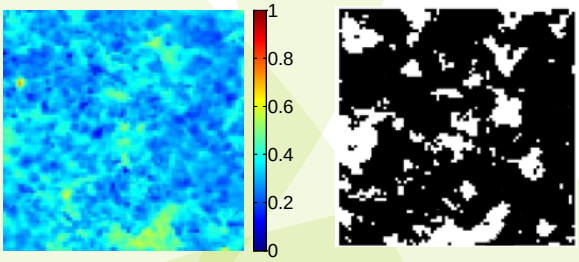
3.2. Validation Set

A total of 26% of pixels were classified as MBM in a sample of FM adulterated with 30% of MBM by using the Gaussian model-based method (Fig.4), and a total of 19% of pixels were classified as MBM in a sample of FM containing 20% of MBM (Fig.5).



(a) Total absorbance image (b) Gaussian model-based

Fig. 4 Binary classification result of FT-NIR microscopic imaging data for FM contaminated with MBM at a concentration of 30% weight (White pixel – MBM; Dark pixel – FM)



(a) Total absorbance image (b) Gaussian model-based

Fig. 5 Binary classification result of FT-NIR microscopic imaging data for FM contaminated with MBM at a concentration of 20% weight (White pixel – MBM; Dark pixel – FM)

4. Conclusions

The results have shown that FM and MBM identification performance using the Gaussian model-based clustering algorithm was more accurate than the results obtained with k-means or fuzzy c-means.

5. References

1. Helmuth Späth. Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples. Translated by J. Goldschmidt. New York: Halsted Press, 1985.
2. James C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press, 1981.
3. Geoffrey McLachlan, and David Peel. Finite Mixture Models. Hoboken, NJ: John Wiley & Sons, Inc., 2000.