Near infrared spectroscopy for classification of apples using K-mean neural network algorism

MASAHIRO MURAMATSU*, YOSHIYASU TAKEFUJI* and SUMIO KAWANO**

*Graduate School of Media and Governance, Keio University, 5322 Endo, Fujisawa, 252-0816, Japan

**National Food Research Institute, Ministry of Agriculture, Forestry and Fisheries, 2-1-2 Kannondai, Tsukuba, 305-8642, Japan

To develop a nondestructive quality evaluation technique of fruits, a K-mean algorism is applied to near infrared (NIR) spectroscopy of apples. The K-mean algorism is one of neural network partition methods and the goal is to partition the set of objects O into K disjoint clusters, where K is assumed to be known a priori. The algorism introduced by Macqueen draws an initial partition of the objects at random. It then computes the cluster centroids, assigns objects to the closest of them and iterates until a local minimum is obtained.

The advantage of using neural network is that the spectra at the wavelengths having absorptions against chemical bonds including C-H and O-H types can be selected directly as input data. In conventional multiple regression approaches, the first wavelength is selected manually around the absorbance wavelengths as showing a high correlation coefficient between the NIR 2nd derivative spectrum and Brix value with a single regression. After that, the second and following wavelengths are selected statistically as the calibration equation shows a high correlation. Therefore, the second and following wavelengths are selected not in a NIR spectroscopic way but in a statistical way. In this research, the spectra at the six wavelengths including 900, 904, 914, 990, 1000 and 1016nm are selected as input data for K-mean analysis. 904nm is selected because the wavelength shows the highest correlation coefficients and is regarded as the absorbance wavelength. The others are selected because they show relatively high correlation coefficients and are revealed as the absorbance wavelengths against the chemical structures by B. G. Osborne.

The experiment was performed with two phases. In first phase, a reflectance was acquired using fiber optics. The reflectance was calculated by comparing near infrared energy reflected from a Teflon sphere as a standard reference, and the 2nd derivative spectra were used for K-mean analysis. Samples are intact 67 apples which are called Fuji and cultivated in Aomori prefecture in Japan. In second phase, the Brix values were measured with a commercially available refractometer in order to estimate the result of K-mean approach. The result shows a partition of the spectral data sets of 67 samples into eight clusters, and the apples are classified into samples having high Brix value and low Brix value. Consequently, the K-mean analysis realized the classification of apples on the basis of the Brix values.