High Resolution Hyperspectral Imagery for Estimating Sweetness Content in Soybean Crops

1. Introduction
In this paper, we investigate the performance of artificial neural network (ANN) based models to predict sweetness content in soybean crops using high resolution hyperspectral data. The sweetness of green vegetable soybeans (edamame) can be estimated by its sucrose and glucose contents.

2. Hyperspectral Imagery Data
Hyperspectral data was acquired locally producing high spatial resolution images without interfering in the crop growth process. We utilized a hyperspectral line sensor, ImSpector, coupled with a CCD camera and computer controller mounted on the tip of a crane. Advantages of the crane based system for agricultural data analysis of localized crop fields are noteworthy, e.g., higher spatial resolution, increased data accuracy, and reduced atmospheric effects. The hyperspectral data acquired comprises the visible to the near-infrared range of the spectrum, wavelengths from 400nm to 1000nm, with approximately 5nm between bands.

The raw radiance data captured by the hyperspectral camera needs to be converted to reflectance. A three dimensional averaging filter is then applied to reduce the noise of the sample dataset. After this preprocessing, image regions containing vegetation are identified using the normalized difference vegetation index (NDVI) [1].

3. Neural Network Prediction Model
We implemented multilayer perceptron (MLP) networks in order to obtain a nonlinear model. The hidden-layer was designed to have 10 neurons. Along with the hyperspectral images, ground truth data was also collected from different varieties of soybeans in the crop field. For each labeled region, the concentration values of sucrose and glucose were measured using liquid chromatography of freeze-dried soybean samples. The actual values were used as target data for the supervised training of the ANN prediction models. The training method for the MLP was the Levenberg-Marquardt algorithm using early stopping [2].

4. Results
A total of 13 different varieties of green vegetable soybeans were analyzed, thus providing 13 levels of concentration for training of the ANN prediction model. Each MLP network was tested over 10 independent runs. The best performing networks were retained, i.e., networks presenting lower MSE value and higher R-value.

5. Conclusion
Reasonable prediction accuracy was obtained by the MLP model. In the range of wavelengths investigated, which covers from the visible to the near infrared, the glucose concentration was predicted more accurately than sucrose. The proposed method permits the non-destructive forecast of sweetness in leguminous crops from high-resolution hyperspectral imagery data.

6. References