Rice Seed Varietal Purity Inspection using Hyperspectral Imaging

Hai Vu*, Christos Tachtatzis†, Paul Murray†, David Harle †,
Trung Kien Dao *, Thi Lan Le *, Ivan Andonovic † and Stephen Marshall †
*International Research Institute MICA, Hanoi University of Science and Technology
† Dept. of Electronic and Electrical Engineering, University of Strathclyde
Email: {hai.vu}@mica.edu.vn

I. Introduction

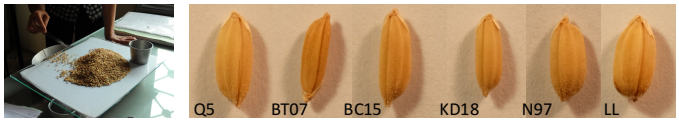
Ensuring rice seed quality is a significant challenge for the large rice export nations such as India, Thailand, USA and Vietnam. The responsibility lies with rice seed producers to ensure high quality seed and a critical procedure is the batch screening and inspection. Conventional methods to inspect seeds, as shown in Figure 1(a), rely on extracting a sample from a batch. The inspection is performed visually to assess the grain properties, such as shape, length, width and size. This task is tedious, laborious, time consuming and requires experienced personnel. Recently, the cost and size of Hyperspectral Imaging (HSI) Systems has reduced significantly. This technology proves to be a useful tool in food sciences and applications. Such systems provide spatial and textural information like other traditional cameras with the added advantage of high resolution spectral signatures for each pixel in the image data acquired. In this paper, we investigate the benefits of analysing the extracted features taken from a HSI system to solve issues of rice seed varietal purity inspection. In this study, the purity of six common rice seed varieties are examined, as shown in Figure 1(b).

Related works: Automatic rice seed inspection systems that employ machine vision and address this challenge have been shown in previous works [1]–[3]. Commonly, shape descriptors of the seed samples are extracted, then statistical classifiers such as Random Forests [3], Neural Networks [1] or Cubic B-Splines shape models [4] are trained. The challenge in comparing and quantifying performance between these approaches, is that each one has been evaluated on different rice seed varieties. It is therefore unclear if the differences in performance come from better feature descriptors or if this is due to varying inter-class/intra-class variations among the examined species. In this study, a HSI system provides both spatial and spectral information about the seed samples. Therefore, inspection techniques that utilize both types of feature and combinations of these are investigated. The use of discriminant analysis techniques and the combinations of both types of features provide significant benefits and potential in HSI offering great advantages for the development of a machine vision system for rice seed quality assessments.

II. Data Acquisition and Methods

A. Rice seed samples preparations and acquisitions

The Near-Infrared (NIR) HSI system used to capture the data was the Inno-Spec™ Redeye 1.7 model (Inno-Spec GmbH, Germany) capturing 256 wavelengths from 950.73 - 1759.4 nm using the well-known pushbroom data acquisition technique. A conveyor platform (the transitional stage) was positioned underneath the camera to allow scanning. Two halogen bulbs were used to illuminate the scene. The HSI system is adjusted so that spatial distortions are avoided, seed samples are fully within FOV of the camera, and light intensity and contrast are suitable. Six rice seed varieties (as shown in Figure 1(b)) were obtained from a seed production company in Vietnam: $BC15$, $BT07$, Khang Dan 18 (KD18), N97, Nep Lang Lieu (LL), and Q5. The selected varieties are the most frequently planted in North Vietnam. The sample population of each variety consisted of 108 seeds with 648 seeds across all varieties. The 108 samples from each species was then divided to 3 batches with 36 samples each. The 36 seeds were positioned on a white sheet of paper constructing a 6×6 matrix. This resulted in 3 hyperspectral datacubes per variety resulting in a total of 36 datacubes. Since the raw reflectance value could vary due to different lighting conditions or manufacturing tolerance of the pixels in the imaging sensor, the data are normalised relative to known min/max reflectance values, which are captured through a calibration procedure [5].



(a) (b)

Fig. 1. (a) A conventional way (human visual) to inspect purity of rice seed samples. (b) Six common rice seed varieties examined in this study.

III. Spatial and Spectral Feature Extractions

A. Physical properties extractions

Rice seed samples are separated from background regions in order to allow the extraction of the physical properties of the grain as well as spectral features. Given an individual rice seed from batch samples, a spatial/morphological feature descriptor f with 6 dimensions is calculated as follows: f_1: is the number of pixels inside a seed sample; f_2, f_3: are the MajorAxisLength and MinorAxisLength respectively; f_4: is the aspect ratio $\frac{f_3}{f_2}$; $f_5 = \frac{Perimeter}{Area}$; and $f_6 = \frac{FociDistance}{MajorAxisLength}$ is the eccentricity, of the ellipse that covers the boundary of the sample seeds.

B. Spectral feature extractions

A hyperspectral datacube contains spectral information for every pixel of the seed regions. For each wavelength, the mean normalised intensity across all pixels in the seed region can be computed resulting in 256 spectral features per seed. The 256

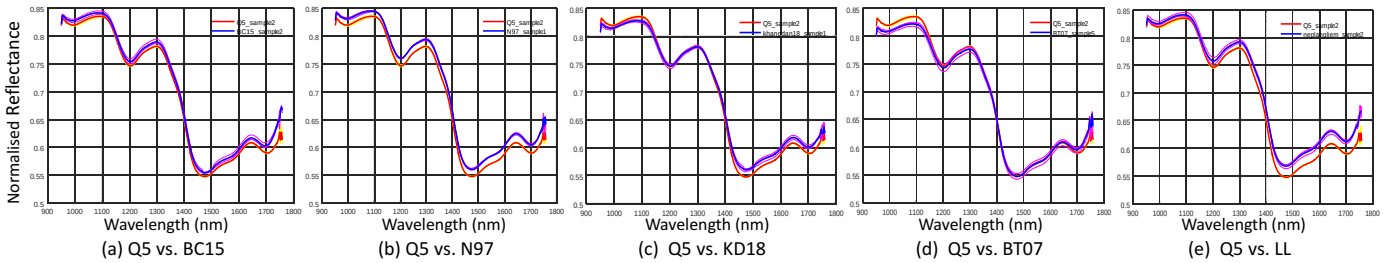


Fig. 2. Comparison wavelength profiles of a species (Q5) with others. The analysis utilized the hyperspectral datacubes of six examined rice seed species

spectral features of the mean normalised intensity of pixels in the seed leads to the "Curse of Dimensionality" [6] and dimensionally reduction techniques are commonly applied in spectral data analysis to avoid overfitting, while reducing redundancy and co-linearity of spectral data. This also facilitates the construction of simple, stable and practical classification models. In particular, we use Principal Component Analysis (PCA) [6] to transform the original data into a small number of uncorrelated variables.

C. Discriminant analysis and classification

Inspection techniques that utilize both types of feature are investigated. The wavelength profile of each species is averaged based on a hyperspectral datacube collected from 108 seed samples. Pair comparisons of the spectral profiles between one species with others are shown in Figures 2(a)-(e). We formulate the purity inspection problem as six one-versus-rest binary classifiers, which are built using SVM and RF techniques [6]. Both approaches are compared.

IV. Result and Discussion

In addition to each set of features (spatial, spectral individually) on the collected dataset, we also evaluate the performance of classifying two different combination schemes: (1) combining the spatial and all the spectral features together, where a feature vector consists of $256+6$ dimensions; (2) using the 6 spatial features along with the 10 principal components. To validate the proposed method, leave-p-out-cross-validation was utilized. For each classifier, 50 seed samples were collected randomly as positive samples, the negative samples were collected in a balanced fashion from all other species so that total negative samples are equal 50 (in other words, 10 from each other species). The experimental results show that the RF classifiers are slightly better than the SVM. Average performances of the RF classifiers are shown in Fig. 3. The performance increases from 77-78% precision for spatial only and spectral only features to 81% when combined as scheme (1), and achieved the highest performance with precision at 84% according to combination scheme (2).

Although the precision obtained from spatial features is lower compared to that reported in [3] it must be noted that the imaging acquisition in [3] is a high resolution camera to extract rice seed shape properties. Going forward, we propose to combine data from registered high resolution images from a high resolution CCD camera with spectral images from the HSI system. Moreover, we believe that utilizing the spectral data at each pixel rather than analyzing and classifying only the mean spectrum on all of the pixels of the seed regions

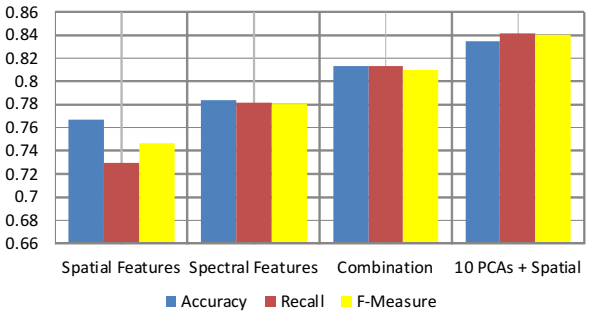


Fig. 3. Average results of the Random Forest classifiers.

can be useful to investigate chemical features of a seed and therefore, discrimination of species would improve.

V. Conclusion

This paper describes a HSI system supporting rice seed varietal purity inspection. The proposed system combines a hardware camera setup and a tool for extracting features from the collected hyperspectral datacubes. We have confirmed that by taking advantage of a HSI system on both spatial and spectral features, we achieve very promising results on eliminating varietal impurity of species from large seed samples.

Acknowledgements

The authors would like to thank the Newton Fund for providing funding (NRCP1516/1/65) to support this internationally collaborative program of work. Thanks also to Adam Polack for his support and guidance in all aspects of optimizing the acquisition of hyperspectral data.

References

[1] Z. Liu, F. Cheng, Y.Ying, and X.Rao, "Identification of rice seed varieties using neural network," *Journal of Zhejiang University SCIENCE*, vol. 11, pp. 1095–1100, 2005.

[2] A. G. OuYang, R. j. Gao, Y. d. Liu, X. d. Sun, Y. y. Pan, and X. ling Dong, "An automatic method for identifying different variety of rice seeds using machine vision technology," in *2010 Sixth International Conference on Natural Computation*, vol. 1, Aug 2010, pp. 84–88.

[3] P. T. T. Hong, T. T. T. Hai, L. T. Lan, V. T. Hoang, V. Hai, and T. T. Nguyen, "Comparative study on vision based rice seed varieties identification," in *Knowledge and Systems Engineering (KSE), 2015 Seventh International Conference on*, Oct 2015, pp. 377–382.

[4] C. Peralta, J. Pabico, and V. Mariano, "Modeling shapes using uniform cubic b-splines for rice seed image analysis," in *2016 Sixth International Conference on Communications and Electronics*, July 2016.

[5] P.Geladi, J.Burger, and T.Lestander, "Hyperspectral imaging: calibration problems and solutions," *Chemometrics and Intelligent Laboratory Systems*, vol. 72, pp. 209–217, 2004.

[6] R.O.Duda, P.E.Hart, and D.G.Stork, *Pattern Classification*. Wiley, 2001.