Airborne Hyperspectral Image Analysis for Grain Sorghum Yield Variability Mapping

Qian Du [1], Chenghai Yang [2], James H. Everitt [2]

[1] Department of Electrical and Computer Engineering, Mississippi State University, MS 39762, USA
[2] USDA-Agricultural Research Service, Weslaco, TX 78596, USA

Abstract — **In this paper we will investigate airborne hyperspectral imagery for mapping grain sorghum yield variability as compared with yield monitor data. All the spectral bands are used for yield variability mapping in order to fully use the plenty spectral information in hyperspectral imagery. Yield variability mapping is achieved by estimating the fractional abundance image corresponding to the yield, where a pixel with high gray-scale value represents high yield in the area it covers. To accommodate in-field spectral variation, an unsupervised method is used. Preliminary study demonstrates the feasibility of this technique, although more thorough investigation is needed.**

Keywords: airborne hyperspectral imagery; yield variability mapping; linear mixture model; unsupervised linear unmixing; abundance estimation.

I. INTRODUCTION

Remote sensing has been used to monitor crop growing conditions and estimate crop yields for many years. Quite a few work has been reported, such as in [1-8], where multispectral images were used.

As remote sensing instruments have advanced recently, hyperspectral imaging receives a high degree of interest due to the fact that its high spectral resolution provides potential of more accurate material identification. A hyperspectral sensor collects data at hundreds of very narrow spectral bands, compared to several broad bands used by a multispectral sensor. So it has higher diagnostic power provided that spatial resolution is comparable.

We have investigated the application of Hyperspectral imagery for mapping grain sorghum yield variability [9]. Correlation analysis showed that grain yield was significantly related to the image data for all the bands except for a few in the transitional range from the red to the near-infrared region. Principal component analysis (PCA) indicated that the first few principal components of the hyperspectral image accounted for 99% of variance in the data. Regression analysis based on principal components could quantify the amount of yield variability explained by image data. However, the number of principal components ought to be used is an open problem [10]; some major components may contain more noise than minor components since variance, the ranking criterion of PCA, is not a good criterion for image [11-12]. Stepwise regression analysis also performed directly on the

yield and hyperspectral data identified the optimum bands and band combinations for mapping yield variability. However, optimum bands differed from field to field.

So in this paper, we will investigate a linear mixture model based method for yield variability mapping by using all the spectral bands. It is based on the fact that the rough spatial resolution permits different materials to be present in the area covered by a single pixel. The linear mixture model says that a pixel reflectance in a visible-near infrared multispectral or hyperspectral image is the linear mixture from all independent pure materials (i.e., endmembers) in an image scene [1]. Linear unmixing analysis is a well-known technique in remote sensing image analysis. For an endmember corresponding to the yield, its fractional abundance is linearly related to the yield in the area covered by a pixel.

II. LINEAR MIXTURE MODEL BASED ANALYSIS

Let L be the number of spectral bands and $\mathbf{r}$ an $L \times 1$ column pixel vector in a multispectral or hyperspectral image. Assume that there are P endmembers in an image scene, which construct an $L \times P$ signature matrix $\mathbf{M} = [\mathbf{m}_1 \mathbf{m}_2 \cdots \mathbf{m}_P]$, where $\mathbf{m}_j$ represents the j-th endmember. Assume that $\boldsymbol{\alpha} = (\alpha_1 \alpha_2 \cdots \alpha_P)^T$ is a $P \times 1$ abundance vector associated with $\mathbf{r}$, where α_j denotes the abundance fraction of the $\mathbf{m}_j$ in $\mathbf{r}$. In the linear mixture model, $\mathbf{r}$ is considered as the linear mixture of $\mathbf{m}_1$, $\mathbf{m}_2$, …, $\mathbf{m}_P$ as

$$\mathbf{r} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n} \qquad (1)$$

where $\mathbf{n}$ is included to account for either measurement or model error [1]. If $\mathbf{M}$ is assumed to be known and keeps to be the same for all the pixels, then the problem is to estimate $\boldsymbol{\alpha}$ which is changed with pixel.

A typical method to estimate $\boldsymbol{\alpha}$ is the least squares approach. The estimate from the least squares solution is the one that minimizes the estimation residual

$$\min_{\boldsymbol{\alpha}} (\mathbf{r} - \mathbf{M}\boldsymbol{\alpha})^T (\mathbf{r} - \mathbf{M}\boldsymbol{\alpha}). \qquad (2)$$

In order for the estimated abundance vector $\boldsymbol{\alpha}$ to faithfully represent an image pixel vector $\mathbf{r}$, two constraints are generally imposed on $\boldsymbol{\alpha}$ in Eq. (1): (a) abundance sum-to-one constraint,

Report Documentation Page

1. REPORT DATE **25 JUL 2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Airborne Hyperspectral Image Analysis for Grain Sorghum Yield Variability Mapping	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical and Computer Engineering, Mississippi State University, MS 39762, USA**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001850, 2005 IEEE International Geoscience and Remote Sensing Symposium Proceedings (25th) (IGARSS 2005) Held in Seoul, Korea on 25-29 July 2005. , The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**3**	

referred to as ASC, $\sum_{p=1}^{P} \alpha_p = 1$; and (b) abundance non-negativity constraint, referred to as ANC, $\alpha_p \geq 0$ for all $1 \leq p \leq P$. There is no closed-form solution to such a constrained linear unmixing problem. So an iterative method generally is used.

Because pixel spectral signatures are changed during the sharpening process, an unsupervised method is used to estimate the endmember signatures in **M** as well as their abundances. The unsupervised fully constrained least squares linear mixture (UFCLSLU) algorithm is used for this purpose [8]. Initially, any arbitrary pixel vector can be selected as an initial desired object denoted by $\mathbf{m}_0$. However, a good choice may be a pixel vector with the maximum length, i.e., the brightest pixel in the image scene. It is assumed that all other pixel vectors in the image scene are pure pixels made up of $\mathbf{m}_0$ with 100 % abundance. Of course, this is generally not true. So a pixel vector that has the largest least square error (LSE) between itself and $\mathbf{m}_0$ is found and selected as a first object pixel vector denoted by $\mathbf{m}_1$. Because the LSE between $\mathbf{m}_0$ and $\mathbf{m}_1$ is the largest, it can be expected that $\mathbf{m}_1$ is most distinct from $\mathbf{m}_0$. Then a signature matrix $\mathbf{M} = [\mathbf{m}_0 \mathbf{m}_1]$ is formed. The FCLSLU algorithm is used to estimate the abundance fractions for $\mathbf{m}_0$ and $\mathbf{m}_1$, denoted by $\hat{\alpha}_0^{(1)}(\mathbf{r})$ and $\hat{\alpha}_1^{(1)}(\mathbf{r})$ for each pixel **r** respectively as the estimates from the first iteration. Now an optimal constrained linear mixture of $\mathbf{m}_0$ and $\mathbf{m}_1$, $\hat{\alpha}_0^{(1)}(\mathbf{r})\mathbf{m}_0 + \hat{\alpha}_1^{(1)}(\mathbf{r})\mathbf{m}_1$, is computed to approximate the **r**. Then the LSE between **r** and this estimate linear mixture is calculated for all image pixel vectors **r**. Once again a pixel vector that yields the largest LSE will be selected to be a second object pixel vector $\mathbf{m}_2$. As expected, such a selected object pixel is the most dissimilar to $\mathbf{m}_0$ and $\mathbf{m}_1$. The same procedure with $\mathbf{M} = [\mathbf{m}_0 \mathbf{m}_1 \mathbf{m}_2]$ is repeated until the resulting LSE is below a prescribed error threshold η or enough endmembers are extracted. If there is partial knowledge of **M** available, then it can be used as initial $\mathbf{m}_0$ or $\mathbf{M}_0$.

The ground truth data about the spectral signature of mature grain sorghum will be compared with signatures within matrix **M** using spectral angle mapper (SAM) [14]. The one closest to the ground truth is considered as the endmember related to grain sorghum yield, and its fractional abundance image will be used for yield mapping. Yield mapping will be achieved by finding the linear relationship between the fractional abundances to the collected yield data.

III. EXPERIMENT

An airborne hyperspectral image was acquired using a CCD camera-based hyperspectral imaging system from a grain sorghum field during the 2000 growing season, and yield data were also collected from the fields using a yield monitor. The hyperspectral images contained 128 bands covering a spectral range from 457.2 to 921.7 nm with a band width of 3.63 nm. The images had a swath width of 640 pixels and each pixel had a gray level between 0 and 4095. The raw images were corrected for the geometric distortion caused by the motion of the aircraft. The corrected images were rectified to the UTM coordinate system with 1m resolution. The calibrated image data were then aggregated into images with a cell size of 9m. Figure 1 shows blue, green, red, and near-infrared bands in a subimage scene of size 450×450.

Band 5 (B)   Band 25 (G)
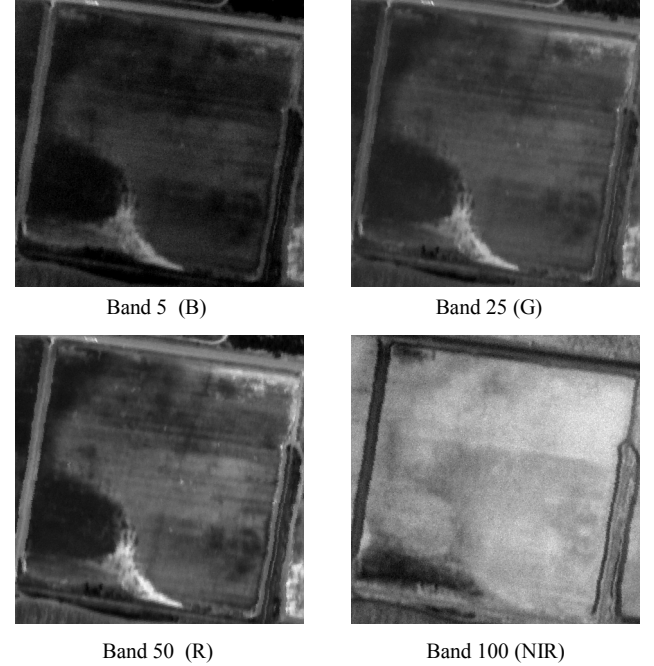
Band 50 (R)   Band 100 (NIR)

Figure 1. Original subimage of size 450×450.

Assume there are ten distinctive endmembers present in this scene. Figure 2 lists ten endmember signatures extracted from the image scene using the unsupervised methods described in Section 2.
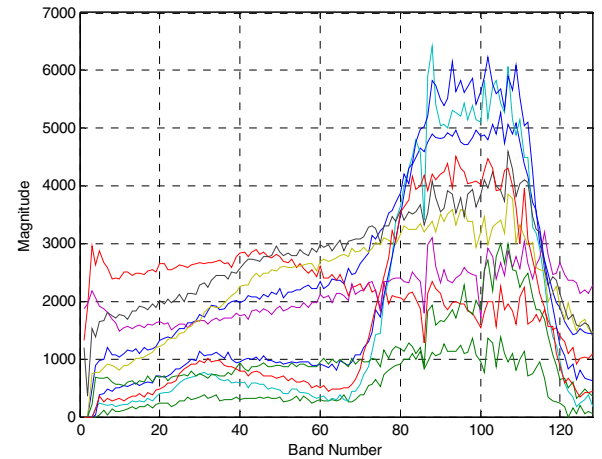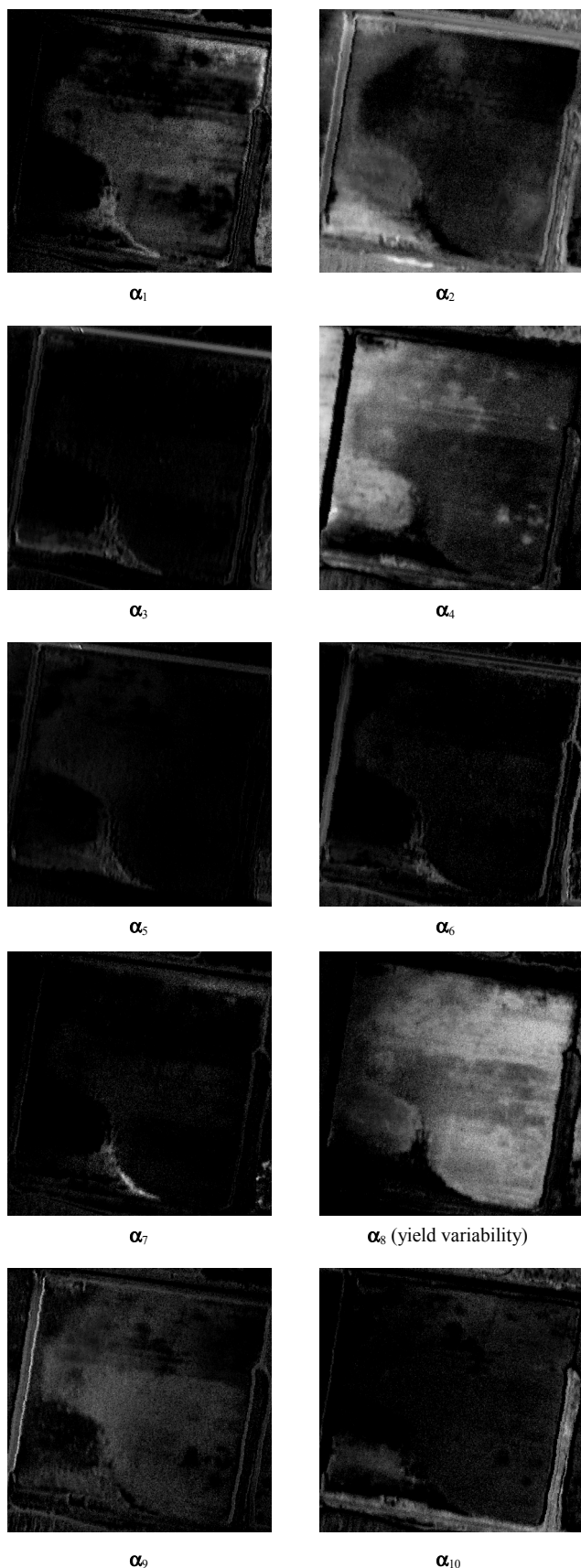

Figure 2. Ten endmember signatures in matrix **M**.

Figure 3. Fractional abundance images.

Figure 3 shows the ten fractional abundance images, which corresponded to different materials. Obviusly, the eighth image α_8 is for yield variability. On this image, a pixel with high gray-scale value means the area it covers has high yield. We also compared it with yield data, and found out that gray-scale values in this image is positively proportional to the collected yield data as expected.

IV. DISCUSSION

The preliminary study demonstrates the feasibility of unsupervised linear unmixing in grain sorghum yield variability mapping. But more work is needed, in particular, on the influence of the number of endmembers used in the algorithm.

REFERENCE

[1] S. A. Morain and D. L. Williams, "Wheat production estimates using satellite images," *Agronomy Journal*, vol. 67, no. 3, pp. 361-364, 1975.

[2] C. J. Tucker, B. N. Holben, and J. H. Elgin, Jr., "Relationship of spectral data to grain yield variation," *Photogrammetric Engineering and Remote Sensing*, vol. 46, no. 5, pp. 657-666, 1980.

[3] C. L. Wiegand and A. J. Richardson, "Leaf area, light interception, and yield estimates from spectral components analysis," *agronomy Journal*, vol. 76, no. 4, pp. 543-548, 1984.

[4] A. J. Richardson, M. D. Heilman, and D. E. Escobar, "Estimating grain sorgham yield from video and reflectance based PVI measurements at peak canopy development," *Journal of Imaging Technology*, vol. 16, no. 3, pp. 104-109, 1990.

[5] S. J. Birrell, S. C. Borgelt, and K. A. Sudduth, "Crop yield mapping: comparison of yield monitors and mapping techniques," *Proceedings of Site-specific Mangement for Agricultural Systems*, pp. 15-31, 1995.

[6] C. Yang and G. L. Anderson, "Mapping grain sorghum yield variability using airborne digital videography," Precision Agriculture, vol. 2, pp. 7-23, 2000.

[7] C. Yang, J. M. Bradford, and C. L. Wiegand, "Airborne multispectral imagery for mapping varibale growing conditions and yields of cotton, grain sorghum, and corn," Transactions of the ASAE, vol. 44, no. 6, pp. 1983-1994, 2001.

[8] C. Yang and J. H. Everitt, "Relationships between yield monitor data and airborne multidate multispectral digital imagery for grain sorgham," Precision Agriculture, vol. 3, pp. 373-388, 2002.

[9] C. Yang, J. H. Everitt, and J. M. Bradford, "Airborne hyperspectral imaging and yield monitoring of grain sorghum yield variability," Proceedings of ASAE Annual Meeting, 2002.

[10] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geoscience Remote Sensing*, vol. 42, pp. 608-619, 2004.

[11] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geoscience Remote Sensing*, vol. 28, no.3, 295-304, 1990.

[12] J. B. Adams and M. O. Smith, "Spectral mixture modeling: a new analysis of rock and soil types at the Viking lander 1 suite," *J. Geophysical Res.*, vol. 91, pp. 8098-8112, 1986.

[13] D. Heinz and C.-I Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529-545, 2001.

[14] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon and A. F. H. Goetz, "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, no.2-3, pp. 145-163, 1993.