Borrowing least squares analysis from spectral unmixing to classify plastics in SWIR hyperspectral images

Mehrube Mehrubeoglu*[1], Austin Van Sickle[2], Lifford McLauchlan[3]

[1]Hyperspectral Optical Property Instrumentation (HOPI) Laboratory, Department of Engineering Texas A&M University-Corpus Christi, 6300 Ocean Dr., Corpus Christi, TX 78412, USA
[2]Surface Optics Corporation, 11555 Rancho Bernardo Rd., San Diego, CA 92127, USA
[3] Texas A&M University-Kingsville, 700 University Dr., Kingsville, TX 78363, USA
*corresponding author: ruby.mehrubeoglu@tamucc.edu

ABSTRACT

Plastics have long been receiving attention due to their abundance in daily use, as well as their loss to the environment as debris. Plastic pollution is widely accepted as an environmental crisis, particularly in marine environments as millions of tons of plastics enter the oceans annually. Although some macro plastics can be determined using visible-range or VNIR hyperspectral imaging, microplastics as well as those that are colorless or have similar pigmentation are difficult to differentiate in the visible spectral regions. SWIR or short-wave infrared hyperspectral imaging offers a solution for plastics detection in the near infrared spectrum. This study builds on recent work for detection and identification of plastics using classical feature extraction techniques and spectral indices. Here, we apply least squares analysis borrowed from linear spectral unmixing methods for classification of plastics from SWIR hyperspectral data. In this research, we compare the results of the two approaches. The two methods produce similar results even though the first approach only utilizes a limited number of features and the second approach makes use of the entire spectral bands represented in each scene pixel.

Keywords: SWIR hyperspectral imaging, SWIR imaging spectroscopy, plastic debris, least squares analysis, spectral similarity, hyperspectral classification, semantic segmentation

1. INTRODUCTION

The detection and classification of plastics requires technologies and methods that are capable of discriminating features that appear similar to the naked eye. This is particularly true when the target at hand is of the same color or color free (clear). Hyperspectral imaging in the short-wave infrared (SWIR) region provides such a solution for the differentiation and classification of plastic materials. SWIR hyperspectral imaging allows spatial information at the same time when capturing spectral signatures of the materials in the scene, providing pixel-by-pixel mapping based on the spectral signature at each pixel. SWIR Hyperspectral imaging of thirteen plastic samples in the range 900-1700 nm was described in our recent publication[1]. Classical feature extraction methods and four selected spectral indices in the form of band ratios were used as features. The extracted features were combined with minimum Euclidean distance measure to classify a variety of plastic materials and perform semantic segmentation of the test scene[1]. In this paper, we extend the classification method using least squares analysis borrowed from linear unmixing model to identify plastics based on the identified highest abundance.

1.1 Plastic Pollutants in the Environment

Plastic continues to be an ongoing concern for environmental pollution due to its widespread use in packaging, longevity and durability. Commonly utilized plastics in bags, packaging and other products include: Polypropylene (PP), Low Density Polyethylene (LDPE), High Density Polyethylene (HDPE) polyethylene terephthalate (PET), Polyvinyl Chloride (PVC), and Polystyrene (PS)[2-5]. Plastic pollution is widely accepted as an environmental crisis, particularly in marine environments, as millions of tons of plastics enter the oceans annually. Although some macro plastics can be determined using visible-range or VNIR hyperspectral imaging, microplastics as well as those that are colorless or have similar pigmentation are difficult to differentiate in the visible spectral regions. SWIR or short-wave infrared hyperspectral

imaging offers a solution for plastics detection in the near infrared spectrum. Serranti *et al.*[6] and Hibbitts *et al.*[7] for instance performed studies on identifying plastics in hyperspectral images.

1.2 Hyperspectral Imaging for Characterization of Plastics

Identification of plastic in images will aid in the collection and recycling of plastics found in the environment. The observed spectrum for a pixel in a hyperspectral image will be a mixture or combination of various spectral signatures observed at the pixel location in the image. The spectral unmixing problem has garnered much interest in remote sensing as well as other imaging applications. Spectral unmixing aids in the determination of what is represented in the image at a particular pixel location by its unmixed spectral signatues[8-15]. Marinoni, Plaza and Gamba[8] developed synthetic data sets to test the efficacy of a fully constrained least squares (FCLS) approach to unmix spectral content in a pixel. The results on the synthetic data sets demonstrate the accuracy of the approach. Chen *et al.* investigated a more general spectral unmixing framework utilizing an objective function[9]. The proposed framework was tested on a simulated data set as well as Hyperion data, with the objective function based on one of the following four measures: Euclidean Distance based on the FCLS method, Spectral Angle Mapping (SAM), Spectral Correlation Measure (SCM) and Spectral Information Divergence (SID). In their research, the SCM, SID and SAM measure-based methods performed better than the Euclidean distance based one for the tested data sets. Kwan *et al.* studied the use of Nonnegatively Constrained Least Squares (NCLS) as well as the Orthogonal Subspace Projection (OSP) method for determining the composition of a mixture of biological and chemical compounds[10]. The NCLS performed better than the OSP on the tested data as it produced more realistic abundance values. Ibarrola-Ulzurrun *et al.*[14] studied the FCLS, scaled constrained least squares unmixing (SCLSU), Extended linear mixing model (ELMM) and Robust ELMM (RELMM) methods for spectral unmixing for images from Teide National Park.

SWIR hyperspectral images of thirteen 3" x 4" plastic samples captured individually for training are used to classify plastics and perform semantic segmentation in an independent scene with multiple plastic materials as part of testing. The images were acquired outdoors under natural light using SOC710-SWIR hyperspectral imaging system[16]. The purpose of this study is to extend four-feature minimum Euclidean distance-based classification to full spectral least squares method-based classification, compare the results, and extend the discussion to similarity among spectral signatures of multiple plastic material signatures. Section 2, Methodology, provides the list of plastic samples used in the study. SWIR hyperspectral imaging equipment and data analysis methods, including the least squares method (LMS) for spectral unmixing, are also summarized in this section. Image of the tested scene with multiple plastic samples as well as hyperspectral data are reproduced in Section 3. Plastic sample detection and classification results and discussion are presented in Section 4. Conclusions are summarized in Section 5.

2. METHODOLOGY

Table 1 lists the plastic materials included in the study[1]. These materials were chosen for their abundance in everyday use as well as their contribution to plastic waste.

2.1 Equipment and Collection of Hyperspectral Data Cubes

The equipment and data acquisition were described in detail in Mehrubeoglu *et al.*[1] Here we provide a short overview for completeness: All hyperspectral data cubes were acquired using the SWIR hyperspectral imaging system (SOC710-SWIR, Surface Optics, San Diego, CA)[16]. The system had a pixel dynamic range of 12 bits, and span a spectral range from 900 to 1700 nm. 696 samples and 512 lines formed the 2D scene, with 288 image frames, each frame representing a spectral band. A high-quality gray panel was used as a reference panel to compute modified reflectance, R_M, as follows:

$$R_M(x,y,\lambda) = I_i(x,y,\lambda) / I_g(\lambda), \quad (1)$$

where (x,y) is the index to the 2D spatial location in each image frame, λ represents spectral band, or frame number associated with the spectral band, and $I_i(x,y,\lambda)$ refers to the observed or measured value at the hyperspectral image voxel, $I_g(\lambda)$ is computed as the spatial average of observed values for the gray panel at spectral band λ. R_M was used for feature extraction with selected wave bands (method 1) and as full-length reflectance spectra with all bands $R_M(x,y)$ (method 2) in classification and semantic segmentation.

Table 1. Plastic materials[2,3] investigated using SOC710-SWIR HSI due to their abundant use and contribution to plastic waste[4-6] in alphabetical order (recreated from Mehrubeoglu *et al.*[1])

SAMPLE	MATERIAL	ACRONYM	VISIBLE COLOR
1	High-Density Polyethylene	HDPE	Black
2	High-Density Polyethylene	HDPE	White
3	High-Temperature Chlorinated Polyvinyl Chloride	High-Temperature CPVC	Gray
4	Kydex Acrylic Polyvinyl Chloride	Kydex Acrylic PVC	Gray
5	Low-Density Polyethylene	LDPE	White
6	Polycarbonate	PC	Clear
7	Polyetheretherketone	PEEK	Tan
8	Polyethylene Terephthalate Modified with CHDM	PETG	Clear
9	Polyester [unspecified]	Polyester	White
10	Polypropylene	PP	White
11	Polystyrene	PS	Amber
12	Polystyrene	PS	White
13	Polyvinyl Chloride	PVC	Black

2.2 Material Classification, and Semantic Segmentation

In the previously-described[1] method 1, four spectral indices from eight spectral bands were selected from modified reflectance spectra as effective features to achieve semantic segmentation of the plastic test samples. Individual plastic samples' average spectral signatures were used as reference spectra to extract the four spectral indices as features. These features were then used as a multi-dimensional vector representing cluster centers associated with each sample. Minimum Euclidean distance was computed from the unlabeled tested pixel features to classify then map each pixel in the test scene. Pixels whose minimum distance from a cluster center exceeded a maximum threshold were not labeled as one of the plastic samples, but identified otherwise.

2.3 Linear Mixing Model and Least Squares Method

In this paper, we describe method 2 for mapping pixels to labeled plastic materials: Utilizing the Linear Mixing Model (LMM) to represent the observed spectrum (mixed spectrum with noise), $\boldsymbol{S_T}$, for a 2D spatial scene pixel in the hyperspectral data cube, one can define the following:[9-11, 13-14]

$$\boldsymbol{S_T} = \sum_{j=1}^{c} a_j \boldsymbol{S_j} + \boldsymbol{E}, \tag{2}$$

where c is the total number of classes ($c = 13$ in this research since there are 13 different plastics investigated), $\boldsymbol{S_j}$ is the spectrum for the j^{th} plastic sample (pure spectrum) and a_j the abundances for each $\boldsymbol{S_j}$. Noise or error, $\boldsymbol{E}$, associated with this approximation for the observed or mixed spectrum is assumed in this model. In this research, a constraint on the abundance values dictates that the sum of the abundance values, a_j, should be 1. The abundance vector is defined as $\boldsymbol{a} = [a_1\ a_2 \ldots\ a_j]$. In this study, since the abundance values, a_j, represent real values, the percentage of the pure spectrum, $\boldsymbol{S_j}$, in the observed pixel spectrum, $\boldsymbol{S_T}$, the abundance values are positive and should sum to 1 for the c values of a_j in the abundance vector $\boldsymbol{a}$ as shown in Eq. (3).[9, 11, 13]

$$\sum_{\boldsymbol{j=1}}^{\boldsymbol{c}} a_j = 1 \tag{3}$$

Thirteen plastics, $c = 13$ and their corresponding pure spectra are utilized in this research. Based on the Linear Mixing Model, the abundance values can be adapted to minimize the cost function, $f(\boldsymbol{a})$, by using the Least Squares Method as in Eq. (4).[9-11, 14] With the abundance vector constraint, this is known as the fully constrained least squares (FCLS).

$$f(\boldsymbol{a}) = \min_{\boldsymbol{a}} \frac{1}{2} \left\| \boldsymbol{S_T} - \sum_{j=1}^{c} a_j \boldsymbol{S_j} \right\|^2, \quad (4)$$

subject to $\boldsymbol{a_j} \geq 0$, and the constraint of Eq. (3).

Minimizing the cost function will minimize the error in the linear mixing model approximation, for the observed spectrum, $\boldsymbol{S_T}$. After cost function minimization, a higher abundance value, a_j, for a given plastic spectrum, $\boldsymbol{S_j}$, implies that the observed pixel spectrum has more spectral contribution from plastic j's spectrum, $\boldsymbol{S_j}$. In this work, the highest abundance plastic is used to label the observed spectrum, that is, the unknown pixel spectrum.

3. DATA

Figure 1 shows the RGB color image of the test scene with plastic labels and sample numbers from Table 1. For testing and mapping the same scene with multiple plastics was used when acquiring the SWIR hyperspectral images.

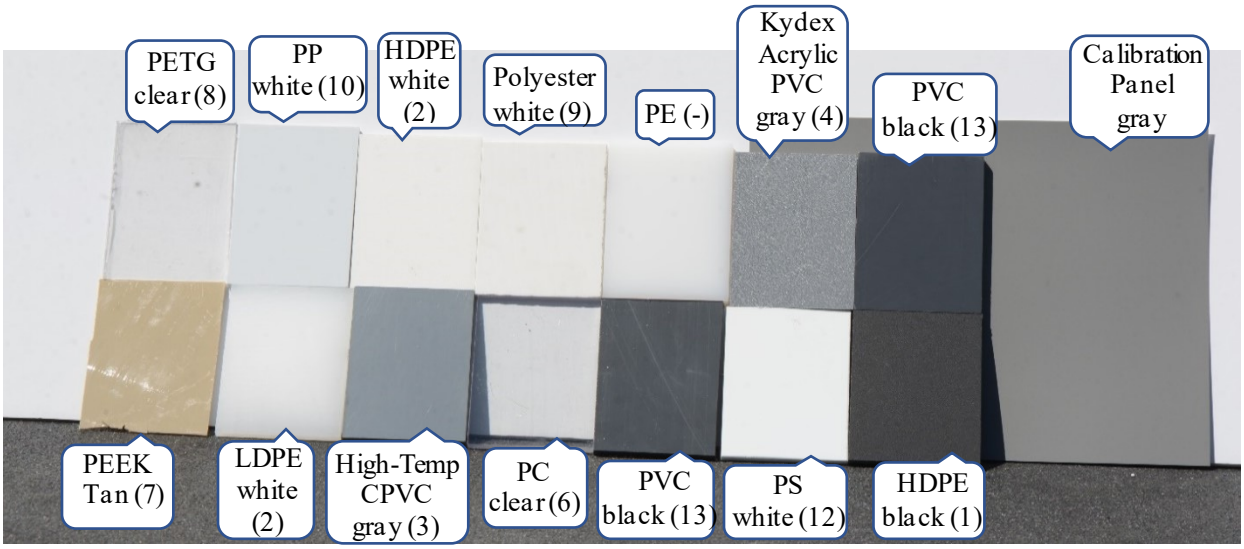


Figure 1. RGB color image of tested plastic samples. This image also represents the test scene which is also acquired using SWIR HSI system. (PE: Polyethylene - not available during training)

Average modified reflectance spectra for plastic samples are reproduced in Figure 2 for reference[1]. Each spectrum represents the average of 500 plastic material spectral signatures from thirteen samples. We note that some modified reflectance spectral values are greater than one, since the reference panel was gray and not white. Based on the average spectra, in method 1, four band ratios identified as features included: (1) ratio_1 = 985:1041, (2) ratio_2 = 1139:1223, (3) ratio_3 = 1419:1461, and (4) ratio_4 = 1545:1670.

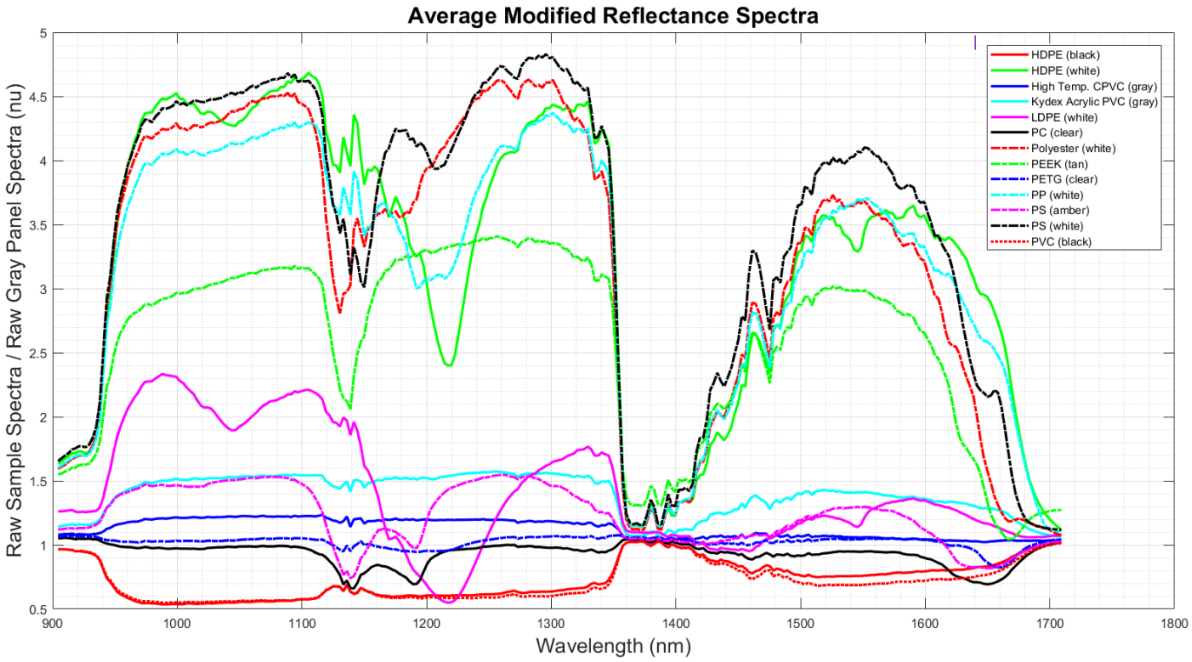


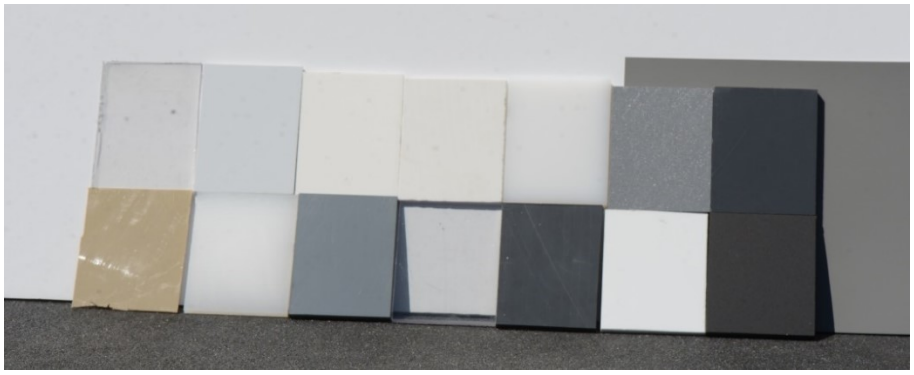
Figure 2. Average modified reflectance spectra[1] used as labeled spectra for LSM as reference.

4. RESULTS AND DISCUSSION

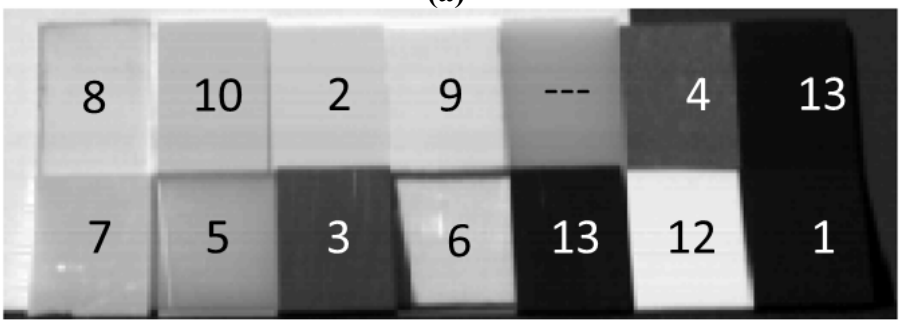
Figure 3(a) shows the RGB color image of the scene for comparison. In Figure 3(b) a single hyperspectral image frame representing the test scene and obtained from the hyperspectral data cube is shown. This image shows fourteen plastic samples overlaid with true sample numbers from Table 1. It is note that Polyethylene (PE) included in the test scene was not available as one of the ground truth samples, therefore it did not have a true reference among the tested samples. In addition, there were two independent samples of PVC black. Finally, PS amber was not in the test scene although it was included among the pure spectra. Figure 3(c) displays the classification and mapping results based on method 1: minimum Euclidean distance. Figure 3(d) depicts the same using method 2: least squares method.

As can be seen in Figure 3 and summarized in Table 2, both classification methods produced comparable results, fully or partially mapping 10 (minimum Euclidean distance) and 11 (LSM) out of the 14 samples in the test scene. Both methods identified HDPE black the same as PVC black. Upon reviewing the modified reflectance spectra for these two materials in Figure 2, it is clear that these two materials display very similar spectral signatures, which could explain the difficulty in distinguishing these materials. Eight of the plastic samples (2, 3, 4, 5, 7, 8, 13(a), 13(b)). Sample 9 was correctly identified by minimum Euclidean distance method. Polyethylene, a plastic material that was not in the list of pure spectra of Figure 2, was identified uniquely, though not matched to one of the ground truth spectra using the same classifier. Sample 10 was incorrectly classified the same as sample 5, and sample 12 was incorrectly classified the same as sample 9. Sample 6 was incorrectly identified as sample 6. (Fig 3(b) and 3(c)). On the other hand, using LSM, samples 6, 10 and 12 were correctly classified, whereas sample 9 was incorrectly identified as sample 10. Polyethylene sample was identified as sample 5.

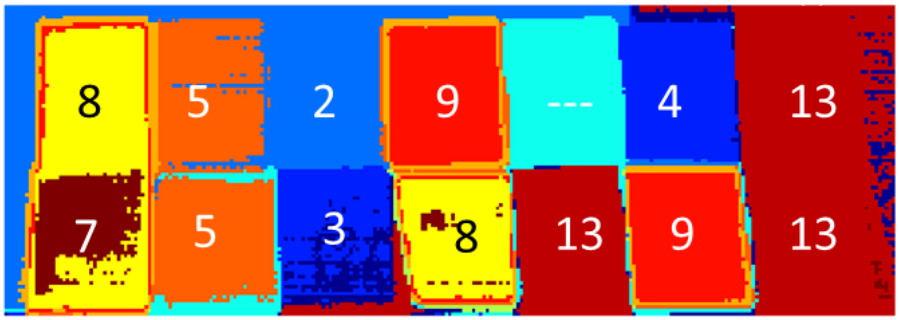
In an effort to explore similarities among spectra based on determined abundances through the LSM, Figure 4 was generated. In this figure, the matrix shows the abundance of pure spectra representing plastic materials in the tested spectrum from each spatial pixel location. It is interesting to note the abundance values along the diagonal, which are used to label the test pixel's spectrum, and the values off diagonal suggest similarities with other pure spectra.



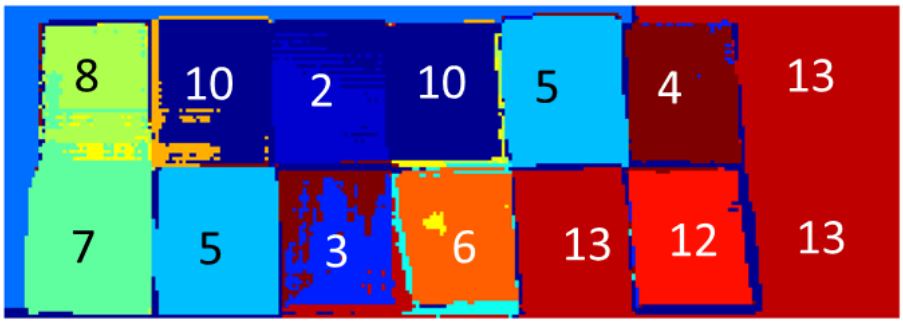
(a)



(b)



(c)



(d)

Figure 3. Semantic segmentation results. **(a)** RGB color image of the test scene; **(b)** single frame (band) image of the scene from the SWIR hyperspectral data cube; **(c)** classification results from method 1: minimum Euclidean distance; **(d)** classification results from method 2: LSM

Table 2. Analysis of plastics classification results from Figure 3 showing plastic materials and classification success using minimum Euclidean distance measure and LSM.

SAMPLE	PLASTIC MATERIAL	Successful Classification Yes: √√ Partial: √ No: X	
		Minimum Euclidean Distance	Least Squares Method
1	HDPE black	X (identified as PVC black)	X (identified as PVC black)
2	HDPE white	√√	√√
3	High-Temperature CPVC gray	√ (partially identified)	√ (partially Identified)
4	Kydex Acrylic PVC gray	√√	√√
5	LDPE white	√√	√√
6	PC clear	X (identified as PETG clear)	√√
7	PEEK tan	√√	√√
8	PETG clear	√√	√√
9	Polyester white	√√	X (identified as PP white)
10	PP white	X (identified as LDPE white)	√√
11	PS amber	N/A (not in the test scene)	N/A (not in the test scene)
12	PS white	X (identified as Polyester white)	√√
13 (a)	PVC black (a)	√√	√√
13 (b)	PVC black (b)	√√	√√
Untrained	Polyethylene (PE) white	√ (uniquely identified but not classified)	X (identified as LDPE white)

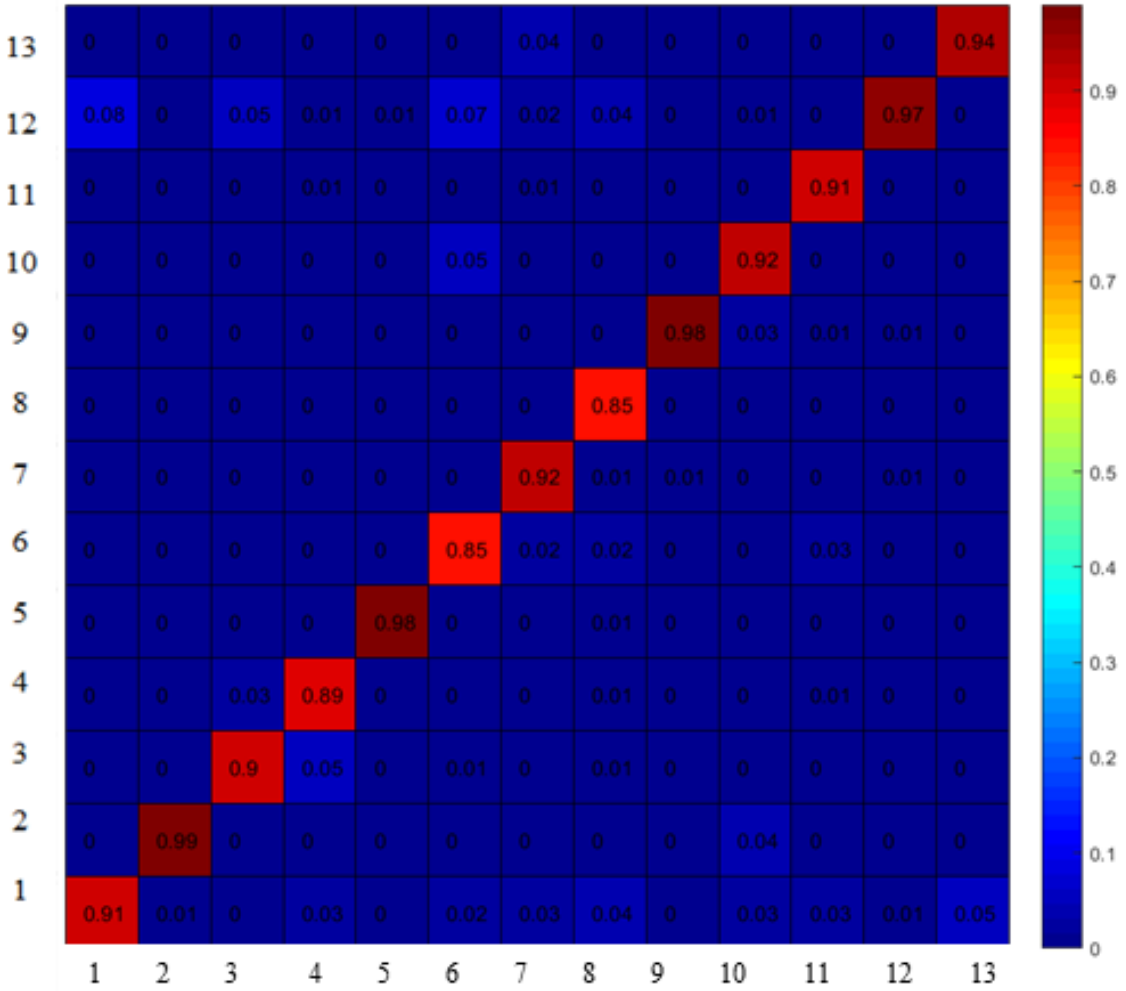


Figure 4. LSM results of abundances. Each column represents the average abundance of 500 pure spectra for samples 1 through 13, as indicated in the labeling below the matrix. The abundances in the diagonal represents the maximum abundances that were used to label a pixel. Off-diagonal values show the abundance of other materials along the same column, suggesting similarities to labeled test spectra identified as the maximum abundance in the diagonal.

5. CONCLUSION

Fourteen plastic samples representing thirteen different materials have been analyzed using SOC170-SWIR hyperspectral imaging. SWIR hyperspectral imaging has been combined with minimum Euclidean distance measure applied to four spectral indices selected as effective features, and least squares method (LSM) applied to the entire spectrum to classify the material represented in each spatial pixel. Both methods produced comparable results. Using the minimum Euclidean distance metric, ten out of fourteen samples in the test scene were identified correctly, whereas using LSM, 11 out of 14 samples were successfully classified. HDPE (black) and PVC (black) samples were identified as the same material through both classifiers. An average abundance matrix was generated based on the abundance computations from 500 pure spectra representing thirteen plastic samples each of which was compared to the average labeled spectra shown in Figure 2 using the linear unmixing model through LSM. Some pure material spectra were identified to have abundances of other pure spectra, suggesting similarities in the tested spectra in the least squares sense, and require further investigation.

Overall, the results of the two simple classification methods for semantic segmentation of plastics show that the SWIR hyperspectral imaging is capable of differentiating different plastic materials. More advanced classifiers are warranted for improved accuracy in classification results and is the topic of ongoing research.

ACKNOWLEDGMENT

This project is an extension of the work performed as part of a grant from Texas A&M University-Corpus Christi - TCRF: *Integrated Characterization and Simulation System for Microplastics in Coastal Watersheds*. The authors thank Dr. Emmett Ientilucci and his graduate students at the Center for Imaging Science Lab at RIT for the least squares spectral unmixing algorithm in MATLAB.

REFERENCES

[1] M. Mehrubeoglu, A. Van Sickle, J. Turner, "Detection and identification of plastics using SWIR hyperspectral imaging," SPIE Proceedings Volume 11504, Imaging Spectrometry XXIV: Applications, Sensors, and Processing; 115040G (2020) https://doi.org/10.1117/12.2570040

[2] Cope Plastics, Inc., Engineering materials, Products and applications guide. https://pdf4pro.com/cdn/products-amp-applications-guide-cope-plastics-220784.pdf (last accessed: 09-27-2020).

[3] Professional Plastics, Inc., Acronyms for Plastics – (abbreviations).

[4] The 5 Most Common Plastics & Their Everyday Use. https://www.cutplasticsheeting.co.uk/blog/uncategorized/the-5-most-common-plastics-their-everyday-uses/ (last accessed: 09-27-2020).

[5] Serranti, S., "Plastic waste monitoring and recycling by hyperspectral imaging technology," Proc. SPIE 11197, SPIE Future Sensing Technologies, 1119706 (November 2019). https://doi.org/10.1117/12.2549670

[6] Serranti, S., Fiore, L., Bonifazi, G., Takeshima, A., Takeuchi, H., and Kashiwada, S., "Microplastics characterization by hyperspectral imaging in the SWIR range," Proc. SPIE. 11197, SPIE Future Sensing Technologies (Nov. 2019).

[7] Hibbitts, C. A., Bekker, D., Hanson, T., Knuth, A., Goldberg, A., Ryan, K., Cantillo, D., Daubon, D., and Morgan, F., "Dual-band discrimination and imaging of plastic objects", Proc. SPIE 11012, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIV, 1101211 (22 May 2019). https://doi.org/10.1117/12.2519014

[8] A. Marinoni, A. Plaza and P. Gamba, "On the detection of linear mixtures in hyperspectral images," *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, 2016, pp. 6990-6993.

[9] J. Chen, X. Jia, W. Yang and B. Matsushita, "Generalization of Subpixel Analysis for Hyperspectral Data With Flexibility in Spectral Similarity Measures," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2165-2171, July 2009.

[10] C. Kwan, B. Ayhan, G. Chen, Jing Wang, Baohong Ji and Chein-I Chang, "A novel approach for spectral unmixing, classification, and concentration estimation of chemical and biological agents," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 2, pp. 409-419, Feb. 2006.

[11] Xingmei Li, Liang Chen and Min Yang, "A priori fully constrained least squares spectral unmixing based on sparsity," *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, 2016, pp. 2139-2143.

[12] A. Karami, R. Heylen and P. Scheunders, "Denoising of hyperspectral images using shearlet transform and fully constrained least squares unmixing," *2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Los Angeles, CA, 2016, pp. 1-5.

[13] A. Marinoni, A. Plaza and P. Gamba, "A Novel Pre-unmixing Framework for Efficient Detection of Linear Mixtures in Hyperspectral Images," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4325-4333, Aug. 2017.

[14] E. Ibarrola-Ulzurrun, L. Drumetz, J. Marcello, C. Gonzalo-Martín and J. Chanussot, "Hyperspectral Classification Through Unmixing Abundance Maps Addressing Spectral Variability," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4775-4788, July 2019.

[15] K. T. Shahid and I. D. Schizas, "Unsupervised Kernelized Correlation-Based Hyperspectral Unmixing With Missing Pixels," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4509-4520, July 2019.

[16] Surface Optics Corporation, SOC 710 Series Hyperspectral Imaging Systems. https://surfaceoptics.com/products/hyperspectral-imaging/soc710-portable-hyperspectral-camera/ (last accessed: 07/31/2020).