An extensive multisensor hyperspectral benchmark datasets of intimate mixtures of mineral powders

AN EXTENSIVE MULTISENSOR HYPERSPECTRAL BENCHMARK DATASETS OF INTIMATE MIXTURES OF MINERAL POWDERS

Bikram Koirala [1], Behnood Rasti [2], Zakaria Bnoulkacem [1], Andréa de Lima Ribeiro [2], Yuleika Madriz [2], Erik Herrmann [2], Arthur Gestels [3], Thomas De Kerf [4], Koen Janssens [3], Gunther Steenackers [4], Richard Gloaguen [2], Paul Scheunders [1]

[1] *Imec-Visionlab, University of Antwerp (CDE), Belgium*
[2] *Helmholtz-Zentrum Dresden-Rossendorf (HZDR), Germany*
[3] *Antwerp X-ray Imaging and Spectroscopy (AXIS), University of Antwerp, Belgium*
[4] *InViLab research group, UAntwerp Industrial Vision Lab, University of Antwerp, Belgium*

ABSTRACT

Since many materials behave as heterogeneous intimate mixtures with which each photon interacts differently, the relationship between spectral reflectance and material composition is very complex. Quantitative validation of spectral unmixing algorithms requires high-quality ground truth fractional abundance data, which are very difficult to obtain.

In this work, we generated a comprehensive hyperspectral dataset of intimate mineral powder mixtures by homogeneously mixing five different clay powders (Kaolin, Roof clay, Red clay, mixed clay, and Calcium hydroxide). In total 325 samples were prepared. Among the 325 samples, 60 mixtures were binary, 150 were ternary, 100 were quaternary, and 15 were quinary. For each mixture (and pure clay powder), reflectance spectra are acquired by 13 different sensors, with a broad wavelength range between the visible and the long-wavelength infrared regions (i.e., between 350 nm and 15385 nm) and with a large variation in sensor types, platforms, and acquisition conditions. We will make this dataset public, to be used by the community for the validation of nonlinear unmixing methodologies (https://github.com/VisionlabUA/Multisensor_datasets)

Index Terms— Hyperspectral, intimate mixtures, multisensor dataset, benchmark, unmixing

1. INTRODUCTION

Because each material interacts differently with incident light, it can be uniquely characterized by its reflectance spectrum. In remote sensing, hyperspectral cameras are employed to detect the reflected sunlight into hundreds of consecutive small wavelength bands in the visible and near-infrared (VNIR, 400-1000 nm) and the shortwave infrared (SWIR, 1000-2500 nm) regions, to produce landcover maps [1]. Due to limitations of the spatial resolution of the sensor, a pixel may contain more than one material, in which case the measured spectral reflectance is generally modeled as a linear mixture of the different materials involved [2]. However, in reality, a material is seldom homogeneous and behaves in itself as an "intimate mixture" with which each photon interacts differently, making the relation between the spectral reflectance and the material composition very complex and highly nonlinear.

With technological progress, small-size, low-cost hyperspectral cameras became available that can be installed on unmanned aerial vehicles, agricultural equipment, conveyor belts, or used in laboratory environments, and even be used as handheld devices [3]. The close-range setting generates higher-quality reflectance data, with the potential to better describe intimate mixtures and to relate the spectral reflectance to the material composition.

Several nonlinear unmixing approaches have been developed [4]. Nonlinear mixing models, describing secondary reflections [5], or higher-order interactions [6] oversimplify the complex interaction of light with intimate mixtures. The study [7] introduced the HapkeCNN, a Convolutional Neural Network (CNN) that incorporates the Hapke model into the learning process. By integrating the physical model into the CNN architecture, the HapkeCNN enhances the understanding and representation of the underlying physical processes involved in the data. Most physics-based mixing models [8, 9] fail when the material grains/particles have a size, smaller or comparable to the wavelength of the light, their shape is non-spherical, and when they behave as anisotropic scatterers. Moreover, these nonlinear mixing models are not invariant to spectral variability, caused by variable acquisition conditions, i.e., variable illumination conditions, distance and orientation from the sensor, or the use of different sensors or white calibration panels.

Some attempts to apply supervised machine learning have been proposed, e.g. by learning a mapping from the spec-

The research presented in this paper is funded by the Research Foundation-Flanders - project G031921N.

tra directly to the fractional abundances, using neural networks and nonlinear regression methods, such as Gaussian Processes [10]. However, since these mappings are unconstrained, they lead to nonphysical results (i.e. negative abundance fractions) [11]. Moreover, most supervised machine-learning methods from the literature fail to deal with spectral variability.

The main obstacle to real advancements in nonlinear spectral mixture analysis is the lack of proper ground truth data to validate the developed approaches and/or to train supervised approaches. In earlier work [12], we developed a robust nonlinear unmixing method by combining model-based unmixing and machine learning and generated limited high-quality ground-truth data of binary powder mixtures for training and validation purposes.

In this work, we generate an extensive hyperspectral dataset of intimate mixtures of up to 5 mineral powders, acquired by 13 sensors using different acquisition configurations, covering a broad range between the visible and the long-wavelength infrared regions (350 nm - 15000 nm). To our knowledge, this is the first large-scale ground-truth dataset of intimate mixtures. We will make this dataset public, to be used by the community for the validation of nonlinear unmixing methodologies.

2. METHODOLOGY/RESEARCH DESIGN

The study contains five different clay powders: Kaolin, Roof clay, Red clay, mixed clay, and Calcium hydroxide. These clay powders mostly contain Aluminium silicate hydroxide, Aluminium oxide, Silicon dioxide, Calcium hydroxide, and Calcium carbonate. black We generated a total of 325 mixtures by mixing these five pure clay powders. All possible clay combinations of these powders were considered, i.e., 10 binary combinations, 10 ternary combinations, 5 quaternary combinations, and one quinary combination.

The clay powders were homogeneously mixed with different fractional ratios, which are uniformly distributed with a fixed interval of 14,29 %, leading to 325 unique samples: 6 unique fractional ratios for each of the 10 binary clay mixture combinations, 15 unique ratios for each of the 10 ternary mixture combinations, 20 unique ratios for each of the 5 quaternary mixture combinations and 15 unique ratios for the mixture of all 5 clays. In Fig. 1 we display the uniformly sampled fractional abundances for a ternary clay combination. The three clays occupy the corners of the simplex, all binary mixtures lie on the lines connecting two clay's while ternary mixtures lie inside the simplex. Mixtures were produced by weighing and combining the different components. We fixed the weight of each mixture sample to be a total of 10 g, the scale had a precision of 0.001 g. Each 10 g sample was put inside a glass bottle and a homogeneous mixture was produced by rotating the bottle for approximately five minutes. Each sample was then put inside a clear plastic jar with an interior diameter of 3.048 cm and a height of 1.524 cm. Approximately 3 g of mixtures was required to fill the sample holder. The samples were then compacted and smoothened using a stamp compactor.

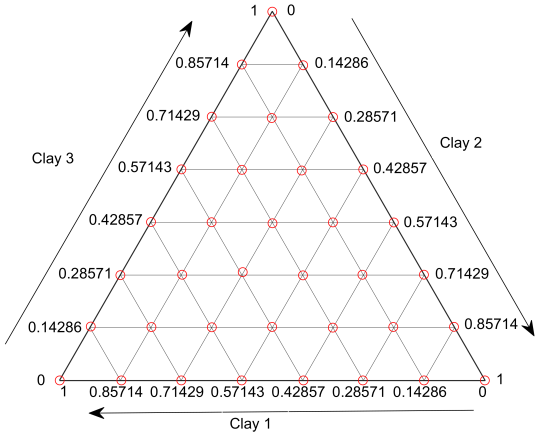


Fig. 1. The ternary diagram of clay mixtures.

Spectral reflectance from all these samples was acquired by 13 different sensors: ASD Spectroradiometer (350-2500 nm), PSR-3500 spectral evolution (350-2500 nm), Specim AsiaFenix (400-2500 nm), Specim sCMOS (400-1000 nm), Senops HSC2 (500-900 nm), Cubert Ultris X20P (350-1000 nm), Specim JAI (440-630 nm), Cubert panchromatic, Specim FX50 (2700-5300 nm), Agilent 4300 Fourier-transform infrared spectroscopy (2500-15385 nm), Telops mid-wave infrared (3000-9000 nm), Specim AsiaOwl (7600-12300 nm), Telops Hypercam long-wave infrared (7400-12500 nm).

For all samples, the ground truth composition is obtained by construction, but to verify that the generated samples are sufficiently homogeneous, X-ray powder diffraction and X-ray fluorescence elemental analysis (Bruker Tornado M4) were performed. To obtain the bulk composition of the samples, the field of view of the X-ray powder diffractometer was fixed to 400 μm, and that of the micro-X-ray fluorescence was set to 170 μm.

3. EXPERIMENTAL RESULTS AND DISCUSSION

In Fig. 2, the spectral reflectance of binary mixtures of Kaolin and Mixed clay, acquired by seven different sensors (in VNIR and SWIR wavelength regions), is shown. It is interesting to note that there is a large spectral variability in the acquired spectra. This variability is likely introduced due to variations in illumination and acquisition angle, and in the distances from the samples to the sensors. In general, band-wise scaling differences can be observed between the acquired spectral reflectances of the different sensors. These effects are caused by variations in illumination and acquisition conditions, the

use of different white calibration panels, and specific differences between the sensors. In general, the spectral features of both Kaolin and Mixed clay are present in these spectra, and gradually change when the fractional abundance of each mineral changes in the mixture. For example, the spectral feature of Kaolin around 1400 nm is clearly visible in the sample Ka-Mi 0.855-0.145 and gradually diminishes with a reduction of the abundance of Ka.

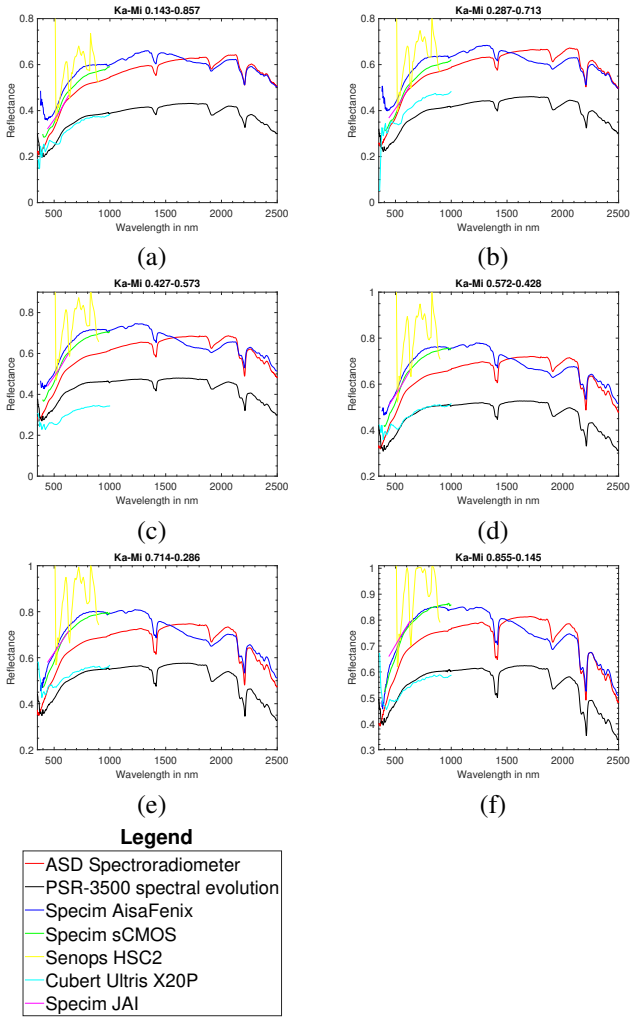


Fig. 2. Spectra of binary mixtures (i.e., a mixture of Kaolin and Mixed clay) acquired by seven different sensors in the VNIR and SWIR; (a) Ka-Mi 14-86; (b) Ka-Mi 28-72; (c) Ka-Mi 43-57; (d) Ka-Mi 57-43; (e) Ka-Mi 72-28; (f) Ka-Mi 86-14.

3.1. Spectral mixture analysis

It can be assumed that the spectral reflectances obtained from the intimate mixtures are nonlinearly related to the ground truth fractional abundances, due to higher-order scattering of the light rays within the powders before reaching the sensor. To demonstrate the impact of these effects on the abundance estimation, the data are linearly unmixed and the deviations of the linearly estimated fractional abundances from the real ground-truth abundances are studied.

In Fig. 3, we display the estimated fractional abundances by FCLSU on the reflectance dataset in the VNIR/SWIR, obtained from the binary and ternary mixtures of the Red Clay, Mixed Clay, and Calcium hydroxide, overlaid on the ternary diagram of the clay mixtures. In the figure, the blue dots denote the estimated abundances, while the red arrows show the real position of the mixtures in the ternary diagram. As can be observed, the error in the estimated fractional abundances of the binary mixtures is significant. Moreover, the linear model projects many of the ternary mixtures onto the faces of the simplex leading to a significant error in the estimated fractional abundances.

4. CONCLUSIONS

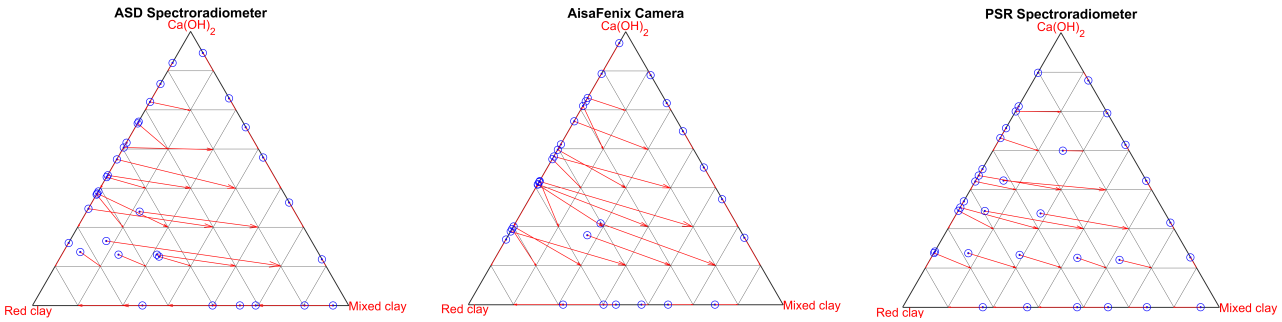
In this work, we generated 325 samples by homogeneously mixing five different clay powders (Kaolin, Roof clay, Red clay, mixed clay, and Calcium hydroxide). Among the 325 samples, 60 mixtures were binary, 150 were ternary, 100 were quaternary, and 15 were quinary. These samples (and pure clay powders) were scanned by 13 different sensors, with a wavelength range between the visible and the long-wavelength infrared regions (i.e., between 350 nm and 15385 nm) to produce a comprehensive hyperspectral dataset of intimate mixtures. The low performance of the linear mixing model to estimate the composition of the mixtures demonstrates that advanced hyperspectral unmixing methods are required that can tackle both spectral variability and nonlinearity.

5. REFERENCES

[1] Pedram Ghamisi, Naoto Yokoya, Jun Li, Wenzhi Liao, Sicong Liu, Javier Plaza, Behnood Rasti, and Antonio Plaza, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, 2017.

[2] José M. Bioucas-Dias, Antonio Plaza, Nicolas Dobigeon, Mario Parente, Qian Du, Paul Gader, and Jocelyn Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[3] Tobias H. Kurz and Simon J. Buckley, "A Review of Hyperspectral Imaging in Close Range Applications," *ISPRS - International Archives of the Photogrammetry,*



(a) (b) (c)

Fig. 3. Linear unmixing results overlaid on the ternary diagram of the mixtures of three clays. In the figure, the blue dots denote the estimated positions of the mixtures while the red arrows show the real positions of the mixtures; (a) ASD Spectroradiometer; (b) AisaFenix Camera; (c) PSR Spectroradiometer.

Remote Sensing and Spatial Information Sciences, vol. 41B5, pp. 865–870, June 2016.

[4] Rob Heylen, Mario Parente, and Paul Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 1844–1868, 2014.

[5] Nicolas Dobigeon, Jean-Yves Tourneret, Cedric Richard, Jose Carlos M. Bermudez, Stephen McLaughlin, and Alfred O. Hero, "Nonlinear unmixing of hyperspectral images," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 82–94, JAN 2014.

[6] Rob Heylen, Vera Andrejchenko, Zohreh Zahiri, Mario Parente, and Paul Scheunders, "Nonlinear hyperspectral unmixing with graphical models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4844–4856, JUL 2019.

[7] Behnood. Rasti, Bikram. Koirala, and Paul. Scheunders, "Hapkecnn: Blind nonlinear unmixing for intimate mixtures using hapke model and convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.

[8] Bruce Hapke, "Bidirectional reflectance spectroscopy: 1. theory," *Journal of Geophysical research*, vol. 86, pp. 3039–3054, 1981.

[9] Bruce Hapke, Robert Nelson, and William Smythe, "The opposition effect of the moon: Coherent backscatter and shadow hiding," *Icarus*, vol. 133, no. 1, pp. 89 – 97, 1998.

[10] Tatsumi Uezato, Richard J. Murphy, Arman Melkumyan, and Anna Chlingaryan, "A novel spectral unmixing method incorporating spectral variability within endmember classes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2812–2831, MAY 2016.

[11] Bikram Koirala, Mahdi Khodadadzadeh, Cecilia Contreras, Zohreh Zahiri, Richard Gloaguen, and Paul Scheunders, "A supervised method for nonlinear hyperspectral unmixing," *Remote Sensing*, vol. 11, no. 20, 2019.

[12] Bikram Koirala, Zohreh Zahiri, Alfredo Lamberti, and Paul Scheunders, "Robust supervised method for nonlinear spectral unmixing accounting for endmember variability," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 9, pp. 7434–7448, 2021.