View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by KITopen

Spectral and spatial unmixing for material recognition in sorting plants

Matthias Michelsburg and Fernando Puente León

Institute of Industrial Information Technology (IIIT),
Karlsruhe Institute of Technology (KIT),
Hertzstr. 16, 76187 Karlsruhe, Germany

Abstract In optical inspection systems like automated bulk sorters, hyperspectral images in the near infrared range are used more and more for identification and classification of materials. However, the possible applications are limited due to the coarse spatial resolution and low frame rate. By adding an additional multispectral image with higher spatial resolution, the missing spatial information can be acquired. In this paper, a method is proposed to fuse the hyperspectral and multispectral images by jointly unmixing the image signals. Therefore, the linear mixing model, which is well-known from remote sensing applications, is extended to describe the spatial mixing of signals originated from different locations. Different spectral unmixing algorithms can be used to solve the problem. The benefit of the additional sensor and the unmixing process is presented and evaluated, as well as the quality of unmixing results obtained with different algorithms. With the proposed extended mixing model, an improved result can be achieved as shown with different examples.

1 Introduction

Recognition and classification of a variety of objects consisting of different materials is a challenging task in automated optical inspection systems as bulk sorters. Such sorting plants are widely used in the fields of mining, food production, and recycling to distinguish between objects according to the material they are made out of, i.e., their chemical composition. Many optical detectors used to discriminate materials work in the near-infrared spectral range (NIR), as the reflected light in this range

gives information on the molecular bindings. This is due to the combination vibrations and its overtones which yield to material-specific spectral signatures. If a precise distinction between different materials is needed, hyperspectral cameras, which usually acquire more than 100 channels simultaneously, are of great importance. These cameras provide a narrowly sampled spectrum for each pixel and thereby allow to recognize different materials at different places.

While the fine spectral resolution provided by the high number of spectral channels is the main advantage of hyperspectral imaging systems, the coarse spatial resolution and low frame rate are its drawbacks. Thereby, light reflected by different objects or several parts of a single object is mixed, which reduces the capabilities of an inspection system. Hence, many industrial constraints concerning speed and resolution cannot be met.

By adding a multispectral (less than 10 channels) or monochrome camera, which meets the requirements for resolution and frame rate, the desired spatial information can be obtained. Therefore, the additional image is fused with the hyperspectral image and a classification in the required spatial resolution is possible. There are several approaches for fusing the different images. When fusing a panchromatic image with a hyperspectral image, the procedure is called pansharpening [1]. A widely used method of this kind is the replacement of a single principal component with the panchromatic image after having transformed the hyperspectral image by principal component analysis. Other methods are based on adding high-frequency components of the monochrome image to the hyperspectral image.

In this paper, a new method for fusing images of different spectral and spatial resolution is proposed. Therefore, the image signals are regarded as mixtures of different material signatures and combined in a common model.

The problem of mixing and its inversion are known as spectral unmixing from the field of remote sensing [2]. Here, mixing coefficients which represent the contribution of each material to the overall signal are assigned to each pixel. In conventional spectral unmixing, the mixing of signals takes place only within each pixel. This approach will be extended by spatial unmixing. This allows images with different spatial resolutions to be merged. The purpose of the proposed method is not the fusion of the different images into one resulting image, but the de-

termination of the mixing coefficients. These coefficients can be used as feature vectors for a subsequent classification process.

This paper is structured as follows. In the following section, the problem is stated in more detail and the procedure of spectral unmixing is introduced. Then, an approach for extending spectral unmixing to several images of different cameras is proposed. The properties of the proposed methods are discussed on the basis of different example signals in the subsequent section.

2 Spectral unmixing

Spectral unmixing is based on the assumption that the reflected spectrum of a pixel is composed of a mixture of different signals originated from different endmembers. These endmembers are usually pure materials. By spectral unmixing of hyperspectral images, the ratios of the endmembers are determined. This can be done in a supervised or unsupervised way. For supervised unmixing, the spectral signatures of the endmembers must be known in advance. Whereas, the spectral signatures are extracted from the hyperspectral image when using unsupervised techniques. There are also methods which do not need endmember spectra at all. Different measures like the pixel purity index or the volume of the simplex spanned by the endmember spectra can be used to determine the endmember signatures. A comparison of different endmember extraction methods can be found in [3]. Instead of extracting endmembers from the image, they also can be taken from several databases. The number of endmembers must be known a priori or can be specified with several methods such as the concept of virtual dimensionality [4].

Spectral unmixing is often used in remote sensing to investigate the earth surface and its geological composition, its development, and vegetation. The areas viewed by a single pixel of a hyperspectral imaging system can be several meters due to the long distance between the image sensor based in an airplane or satellite and the observed object. Hence, different objects and materials are usually found in the field of view of one pixel. Similar effects can be observed when using hyperspectral images in inspection systems like automated bulk sorters. While the distance between sensor and object is small, signal

components are also mixed due to the high speed of the objects and the low frame rate of the camera. By spectral unmixing, the mixture is attempted to be inverted and, as a result, the relative contribution of each material is determined.

There are different mixing models which are based on different constraints. The linear mixing model is the simplest and most widely used model. Here, additively mixed signal compounds are assumed. The linear mixing model is presented in the next section. Other mixing models, as the bilinear mixing model, allow for more complex mixtures caused by scattering and other nonlinear effects taking place in the material [5].

2.1 Linear mixing model

In the following, all signals are regarded as discrete variables for mathematical descriptions. Thereby, they can be written in matrix notation.

The linear mixing model assumes a signal $\mathbf{y}$ to conform

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{a} + \mathbf{n}\,, \tag{17.1}$$

where $\mathbf{X}$ is a $N \times M$ matrix whose columns $\mathbf{x}_i$ represent the spectra of the M endmembers. N stands for the number of channels of the sensor, $\mathbf{n}$ is a noise term which combines model errors and sensor noise. Vector $\mathbf{a}$ consists of the mixing coefficients. There are two restrictions for these coefficients. The coefficients need to be non-negative

$$a_i \geq 0 \qquad \text{for} \qquad i = 1, \ldots, M\,, \tag{17.2}$$

and all coefficients of one pixel need to sum up to one:

$$\sum_{i=1}^{M} a_i = 1\,. \tag{17.3}$$

The different endmembers contribute only positively to the overall signal, which is ensured by the first constraint. The second restriction accounts for the signal is being fully described by the endmembers. The assumptions yield

$$a_i \leq 1 \qquad \text{for} \qquad i = 1, \ldots, M. \tag{17.4}$$

All possible combinations of mixing coefficients are found within an M dimensional simplex with edges of length one.

2.2 Unmixing algorithms

The inversions of the mixing problem, i.e., the estimation of the mixing coefficients $\hat{\mathbf{a}}$, can be done by different approaches. Methods that minimize the reconstruction error

$$e(\hat{\mathbf{a}}) = \|\mathbf{y} - \mathbf{X}\,\hat{\mathbf{a}}\|^2 \tag{17.5}$$

are widely used. The least-squares method, which can be extended to fulfill the requirements of (17.2) and (17.3), belongs to this kind of algorithms. The nonnegativity constraint of the mixing coefficients is ensured by the nonnegativity constrained least-squares algorithm (NNLS) [6], the normalization is ensured by the sum-to-one constrained least-squares algorithm (SCLS) [7]. Both methods can be combined into the fully constrained least-squares method (FCLS), which meets both requirements.

Beside the least-squares approaches, there are methods based on stochastic models. Here, the unmixing problem is resolved by a maximum-likelihood estimator or by hierarchical Bayesian models [8]. The nonnegative matrix factorization (NMF) determines the endmember spectra and the mixing coefficients simultaneously and, hence, does not need any endmember spectra at all [9].

3 Extended signal model

The linear mixture model is extended to represent also spatial mixtures. Therefore, the linear mixture model is defined for the whole image instead of only for a single pixel. The linear mixture model in (17.1) is assumed to be valid for each pixel of the image. Hence, all signals $\mathbf{y}$ and the corresponding mixing coefficients $\mathbf{a}$ can be combined into matrices. This yields

$$\mathbf{Y} = \mathbf{X}\,\mathbf{A}\,, \tag{17.6}$$

where the columns of $\mathbf{Y}$ and $\mathbf{A}$ represent the signals and mixing coefficients of the single pixels. Here and in the following, the noise term is not mentioned for simplicity.

The individual images are acquired with different spectral and spatial resolutions and are combined in one common model. The effects

caused by the different resolutions are called spectral and spatial mixing, respectively.

Spectral mixing The spectral resolution describes the spectral sensitivity of the single sensor channels. This resolution is referred to as a spectral base resolution. For convenience, the spectral channels of the hyperspectral image sensor are used as the base resolution. Matrix $\mathbf{X}$ consists of the spectral signatures of the endmembers at base resolution. The specific spectral sensitivity of each channel of a camera is modeled as a linear combination of the channels in $\mathbf{X}$. This yields an adapted matrix of endmember signatures for each sensor

$$\mathbf{X}_i = \mathbf{C}_i \cdot \mathbf{X} \,. \tag{17.7}$$

Here, matrix $\mathbf{C}_i$ consists of the relative spectral sensitivities of camera i.

Spatial mixing The spatial resolution of a sensor is affected by the region, out of which a signal of a pixel is composed. The signal of a pixel is written as linear combination of signals of a high spatial base resolution. For the linear mixing model, this can be regarded as a linear combination of the mixing coefficients $\mathbf{A}$. For the coefficients at lower resolution, this results in

$$\mathbf{A}_i = \mathbf{A} \cdot \mathbf{B}_i \,. \tag{17.8}$$

Matrix $\mathbf{B}_i$ describes the mixing of signals originating from different locations and can be derived from the point-spread function of the sensor.

3.1 Combination of spatial and spectral mixing

For each sensor with the spectral and spatial mixing effects described above, one has

$$\mathbf{Y}_i = \mathbf{X}_i \mathbf{A}_i = \mathbf{C}_i \mathbf{X} \cdot \mathbf{A} \mathbf{B}_i \,, \tag{17.9}$$

rearranging by using the Kronecker product yields

$$\mathrm{vec}\{\mathbf{Y}_i\} = \left(\mathbf{B}_i^{\mathrm{T}} \otimes \mathbf{C}_i \mathbf{X} \right) \cdot \mathrm{vec}\{\mathbf{A}\}. \tag{17.10}$$

Here, the operator $\mathrm{vec}\{\mathbf{Z}\}$ denotes the column representation of matrix $\mathbf{Z}$ by stacking the columns of $\mathbf{Z}$ into a single column vector, and $\otimes$ stands for the Kronecker product.

In this representation, the signals of multiple sensors can be combined by stacking the vectors and matrices column-wise:

$$\underbrace{\begin{bmatrix}\mathrm{vec}\{\mathbf{Y}_1\}\\ \mathrm{vec}\{\mathbf{Y}_2\}\\ \vdots\\ \mathrm{vec}\{\mathbf{Y}_k\}\end{bmatrix}}_{\overline{\mathbf{y}}} = \underbrace{\begin{bmatrix}\mathbf{B}_1^{\mathrm{T}}\otimes\mathbf{C}_1\mathbf{X}\\ \mathbf{B}_2^{\mathrm{T}}\otimes\mathbf{C}_2\mathbf{X}\\ \vdots\\ \mathbf{B}_k^{\mathrm{T}}\otimes\mathbf{C}_k\mathbf{X}\end{bmatrix}}_{\overline{\mathbf{X}}} \cdot \underbrace{\mathrm{vec}\{\mathbf{A}\}}_{\overline{\mathbf{a}}}. \tag{17.11}$$

This yields a common linear model

$$\overline{\mathbf{y}} = \overline{\mathbf{X}}\,\overline{\mathbf{a}} \tag{17.12}$$

that needs to be solved. As this problem is similar to the linear mixing model, the approaches for spectral unmixing described above can be used to solve it.

4 Studies

The proposed method for fusing different images is illustrated and evaluated with an example. Therefor, images of different spectral and spatial resolution need to be created.

4.1 Example data

Five different materials with spectra shown in Fig. 17.1 are evaluated. The spectra were extracted from a single hyperspectral image of minerals. For further evaluation, a simulated image is created, which consists of 100×100 pixels at the highest spatial resolution. The spatial distribution of the five mixing coefficients is shown in Fig. 17.2 and were chosen based on the examples in [3]. The brighter a pixel the higher the contribution of the endmember to the overall signal of the pixel. Every material is represented by a pure pixel. The other pixels are mixtures of multiple materials. The mixing coefficients fulfill the constraints in (17.2) and (17.3).

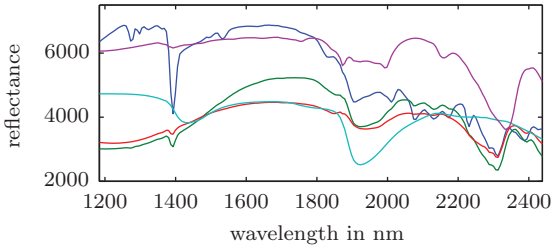

Figure 17.1: Spectra of endmembers.

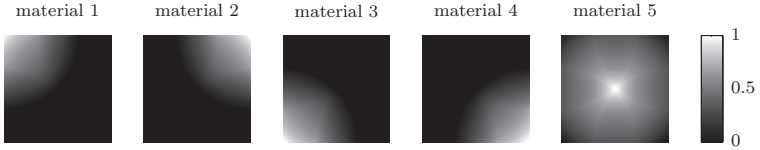

Figure 17.2: Spatial distribution of mixing coefficients for the different materials.

Different signals can be simulated with the spectra shown in Fig. 17.1 and the mixing coefficients shown in Fig. 17.2. Gaussian noise is added to the signal to account for sensor noise. The signal-to-noise ratio (SNR) is the ratio of half the mean signal value to the standard deviation of the noise (compare [3]). Unless otherwise stated, the SNR is 50 : 1.

4.2 Evaluation

Different measures can be used to evaluate the results of spectral unmixing algorithms. As simulated data is used, measures can be determined on the error of the mixing coefficients and on the reconstructed image signal. Therefor, the root-mean-square error of the estimated mixing coefficients can be used. The ERGAS index is another widely used error measure for the fusion of images of different spatial and spectral resolutions [10]:

$$\mathrm{ERGAS} = 100 \cdot \frac{h}{l} \cdot \sqrt{\frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{RMSE}(Y_k))^2}{\overline{Y}_k^2}} \,. \tag{17.13}$$

Here, h is the spatial resolution of the high-resolution image and l the resolution of the low-resolution image. RMSE(Y_k) stands for the RMSE of the k-th channel of the reconstructed image, $\overline{Y}_k$ denotes the mean value of a channel. Unlike the RMSE of the mixing coefficients, the ERGAS index is related to the reconstitute image and not to the mixing coefficients themselves. The lower the values of the ERGAS index the better the fusion of the two images.

4.3 Impact of the resolution

A scenario with one hyperspectral and one multispectral camera is evaluated. The spectral resolution of the hyperspectral camera is the same as the one of the base resolution, i.e., 200 channels. The spatial resolution is smaller than the base resolution. A single pixel of the hyperspectral image is composed of 6×6 pixels of the base resolution. The number of channels of the multispectral camera and its spatial resolution are modified. To minimize the spatial correlation of the mixing coefficients, the image pixels are randomized spatially. Figure 17.3 shows the RMSE of the mixing coefficients for different resolutions of the multispectral sensor with a varying number of spectral channels.

One can see that the effect of the additional image is higher the higher

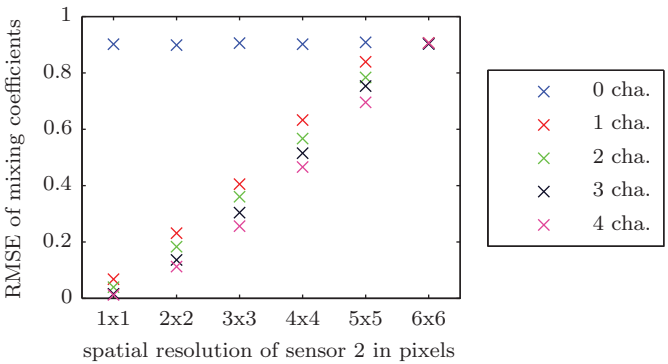

Figure 17.3: RMSE as a function of the spatial resolution of the multispectral sensor and its number of spectral channels in combination with a hyperspectral sensor with a spatial resolution of 6×6 pixels.

the spatial resolution. For all spatial resolutions, the RMSE is lower than when using only the hyperspectral sensor. The RMSE decreases with an increasing number of channels of the multispectral sensor.

4.4 Impact of the algorithms

To evaluate the different algorithms, a different scenario is used. Here, a hyperspectral image with 200 channels is combined with a multispectral image with 3 channels. The spatial resolution of the hyperspectral image is 3×3 pixels and 1×1 for the multispectral image, respectively. Four different algorithms are compared. Firstly, the unconstrained least-squares algorithm (UCLS), NNLS and FCLS, all described above. The fourth algorithm is a successive algorithm (sFCLS), which at first solves the hyperspectral image with low resolution and subsequently solves the high resolution image constrained by the first result. The results in terms of the RMSE and ERGAS index are listed in Tab. 17.1.

	UCLS	NNLS	FCLS	sFCLS
ERGAS	1.13	1.12	0.81	0.81
RMSE	0.16	0.16	0.11	0.11

Table 17.1: ERGAS index and RMSE for unmixing results with different algorithms. The results were obtained with the combination of a hyperspectral image with resolution 3×3 and a multispectral image with 3 channels and a resolution of 1×1.

UCLS and NNLS yield similar results. The same holds for FCLS and sFCLS. The unmixing result of the FCLS algorithms is better than the one of the unconstrained algorithm. There is no big difference of the combined and the successive variant of the FCLS algorithm. However, the advantage of the combined method is the possibility to apply it to more than two images.

5 Summary

An extended mixing model based on the linear mixing model and spectral unmixing was proposed. It can be used to jointly unmix image sig-

nals of different spatial and spectral resolutions. Thereby, an improvement of the unmixing result can be achieved in applications where there is sufficient spectral information, but too low spatial information. The proposed method is fully based on the mixing model which has been established in remote sensing applications for object classification. Hence, many remote sensing methods can also be used for analysis of the material composition in visual inspection.

The proposed sensor model can be easily extended to multiple sensors. An extension by image registration and other spatial transformations is also possible. Different sensor noise models can be considered when solving the unmixing problem. The possibility of applying the methods to non-linear mixing models and the benefit for the final classification will be further investigated.

References

1. I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *EURASIP Journal on Advances in Signal Processing*, vol. 1, no. 79, pp. 1–22, 2011.
2. N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55–78, 2003.
3. A. Plaza, P. Martínez, R. Pérez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 650–663, Mar. 2004.
4. C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
5. A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
6. R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," *Journal of Chemometrics*, vol. 11, no. 5, pp. 393–401, Sep. 1997.
7. E. Ashton and A. Schaum, "Algorithms for the detection of sub-pixel targets in multispectral imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 64, no. 7, pp. 723–731, 1998.

8. N. Dobigeon, J.-Y. Tourneret, and C.-I. Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 2684–2695, 2008.
9. V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra and its Applications*, vol. 416, no. 1, pp. 29–47, Jul. 2006.
10. L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?" in *Proceedings of the third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, T. Ranchin and L. Wald, Eds., 2000, pp. 99–103.