Vision, Modeling, and Visualization (2013)
Michael Bronstein, Jean Favre, and Kai Hormann (Eds.)

Expressive Spectral Error Visualization for Enhanced Spectral Unmixing

B. Labitzke, F. Urrigshardt and A. Kolb

Computer Graphics Group, Institute for Vision and Graphics (IVG), University of Siegen, Germany

Abstract

A major issue in multispectral data analysis stems from the concept of spectral mixture analysis, i.e. the fact that a pixel does not cover only one material but corresponds to a mixture of materials. Even though many automatic methods for spectral unmixing exist, in many practical applications, domain experts have to verify the result and sometimes have to manually adjust the set of determined materials to achieve proper spectral reconstructions.
In this paper, we propose an approach to enhance the very tedious and time-consuming task of manual verification of the unmixing and optional refinement of the materials. Our visual analysis approach comprises different techniques for an expressive spectral error visualization, efficiently guiding the user towards spectra in the dataset which are potentially missing materials. Here, combined views allow comprehensive, local and global error inspections in parallel. We present results of our proposed approach for two domains.

Categories and Subject Descriptors (according to ACM CCS): I.3.m [Computer Graphics]: Miscellaneous— I.4.m [Image Processing and Computer Vision]: Miscellaneous—

1. Introduction

Multi- and hyperspectral imaging have been applied in the context of various applications on both macroscopic, e.g. remote sensing, and microscopic scales like Raman spectroscopy. Recent technological advances in spectral imaging, like the development of low-cost and compact multispectral imaging cameras [HKW12] or 3D scanning systems that incorporate hyperspectral imaging [KHK*12], show the popularity of this imaging technique. In the following, for the sake of simplicity, multi- and hyperspectral image data are referred to as multispectral data.

In general, multispectral imaging yields three-dimensional datasets, with two spatial dimensions representing the surface position and one spectral dimension that represents the spectral distribution, allowing for a very deep investigation of scene characteristics. Thus, each pixel (x,y) in the multispectral image does not provide only grayscale or color information but in fact has associated a spectrum, i.e. n-dimensional vector $\vec{s}(x,y)$ of values with each value $\vec{s}(x,y,\lambda)$ being the measurement of the reflectance for a specific spectral band λ. Based on this, one can differentiate materials not only on color but also on spectral properties beyond the visible range.

Due to the diversity and growing number of application domains, there is a high demand for efficient generic data analysis and visualization methods for spectral data. Such approaches should enable a user to get access to the relevant information, e.g. constituent spectra, i.e. materials, within the clutter of high-dimensionality. One prominent analysis concept is the technique of linear spectral unmixing (LSU) [BDPD*12, Kes03, PZPM10] that is applied in various applications, e.g. remote sensing [PZPM10]. LSU has been a very active research field in recent years [JMP12]. The general approach is to express all individual spectra in the dataset by convex combination of constituent spectra of the dataset, so-called endmembers. Per pixel coefficients of the endmembers are calculated using inverse operations [Kes03]. Alternatively, the process can be seen as compressing the spectral data, with minimal loss of application specific information.

However, the quality of the final result, i.e. the residual error between the raw and the reconstructed data, strongly depends on the selected endmembers and automatic endmember extraction algorithms can fail to determine a complete set of endmembers [LBK12]. Moreover, one of the most profound source of error in LSU lies in the lack of the ability to account for sufficient temporal and spatial spectral variability, e.g. see Fig. 5 [SATC11]. Thus,

DOI: 10.2312/PE.VMV.VMV13.009-016

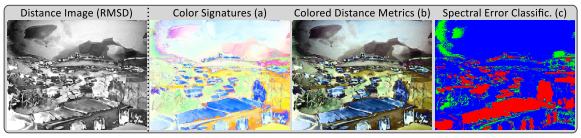

Figure 1: *A typical distance image that expresses the quality of a LSU-result and which can be used as basis for a refinement of an endmember set is shown (left) and is compared to the proposed contributions of our approach (right). Distinct residual errors are easier distinguishable with our* color signatures *(a), which can be combined with common distance images and then are called* colored distance metrics *(b). Additionally, our* spectral error classification *(c) discriminates pixel in terms of missing endmembers (red), variability errors (blue) and neglectable deviations (green).*

manual intervention can be mandatory in many applications to improve the unmixing, leading to the requirement of efficient user guidance and visual exploration methods. For example to identify missing endmember candidates, i.e. missing materials, that have high potential to improve the quality of the unmixing. Evaluating the reconstruction quality always incorporates n-dimensional distance metrics (cf. Tab. 1) [JMP12, PZPM10], leading to a single scalar error value per pixel. This mapping results in severe information loss, since different spectral errors may lead to the same, i.e. ambiguous, scalar error. Even worse, different metrics can lead to different quality impressions. Manual inspection of individual spectral residual errors, on the other hand, is extremely time-consuming and does not provide global information about the distribution of spectral errors.

In this paper, we propose a novel visual analysis approach to enhance LSU-results by *expressive spectral error visualization* to efficiently guide a user to specific spectra for local exploration. In detail, our contributions are as follows:

- We present a qualitative coloring scheme, called *color signatures*, to easier differentiate distinct errors by visually grouping in a global sense. *Colored distance metrics*, the combination of *color signatures* and typical distance images, are used to reduce the ambiguity of scalar error values, see Fig. 1 (a, b).
  The *colored distance metrics* also can be interactively modified by the design of transfer functions to freely highlight or determine error regions of interest.
- A *spectral error classification* is proposed to mainly provide the distinction between errors caused by endmember variability or by missing endmember, see Fig. 1 (c).

The remainder of this paper is organized as follows. Sec. 2 presents a brief overview of the related work and introduces relevant aspects of linear spectral unmixing. Before the details of our contributions are discussed in Sec. 4, the Sec. 3 gives a conceptional overview of the proposed approach. The principle of the interactive exploration is elucidated in Sec. 5. Sec 6 presents results by usage examples for two different domains. Finally we conclude this paper in Sec. 7.

2. Related Work

Multispectral data exploration must deal with the complexity of multispectral data, i.e. high spatial resolution and spectral density. Visual exploration of high-dimensional data in general involves mapping to lower dimensional visual representations, e.g. by scatterplots or parallel coordinates. Here, the major challenge of an analyst is the identification of insightful mappings, which best possible show phenomena contained in the data, like clusters or correlations [TAE*11]. Beside the challenge of high dimensionality, the variety of application domains introduce the need for generic processing tools to gain insights to any multispectral data. Software is mainly available for processing multispectral datasets in the field of remote sensing, like the popular tool MultiSpec [BL02]. Recently Jordan and Angelopoulou [JA10] presented their more universal open-source software Gerbil. Gerbil is highly interactive with combined viewports, e.g. parallel coordinates, and focuses on the analysis of scene reflections and exploration of the relationship between spectral as well as topological information. Also, the work of [CRHW09] and [KZD*10] utilizes user-interaction as a key feature to analyze application related multispectral data. The approach proposed here also is a user driven exploration approach. But, in contrast to all mentioned approaches, we focus on linear spectral unmixing to analyze and express the compositions of all pixels.

Linear spectral unmixing (LSU) is a popular approach in multispectral data analysis to explore subpixel details [Kes03], which usually consists of two steps (please also see Fig. 2):

1. *Endmember Extraction:* This step identifies the set $V = \{\vec{v}_1, \ldots, \vec{v}_q\}$ of constituent spectra (endmembers), where q is the number of endmembers.
2. *Computation of Abundances:* Each spectrum $\vec{s}(x,y)$ of the dataset is expressed as a linear combination with respect to V, assuming a linear superposition of the endmembers in each pixel, i.e.

$$\vec{s}(x,y) = \sum_{j=1}^{q} \alpha_j(x,y) \cdot \vec{v}_j + \vec{n}(x,y) \qquad (1)$$

$$\text{with } \sum_{j=1}^{q} \alpha_j(x,y) = 1,\ 0 \leq \alpha_j(x,y) \leq 1,$$

   where α_j are the abundances and $\vec{n}(x,y)$ is a noise vector of the current pixel (x,y).

In Eq. (1), typically the abundance non-negativity constraint (ANC) and the abundance sum-to-one constraint (ASC) are enforced to ensure physical plausibility. In this case the LSU is called fully constrained LSU (FCLSU). When only ANC is enforced, then we call it non-negative constrained LSU (NCLSU). Based on Eq. (1), inverse operations, e.g. [SMPC10], are applied to compute the coefficients α_j.

Plaza et al. [PZPM10] state that good reconstruction in FCLSU relies on the correct identification of the endmember set. But the determination of a proper set of endmembers is the hardest and most crucial part of the unmixing problem [Kes03]. Here, the proposed method of Chang and Du [CD04] maybe is a helpful preprocess to initially estimate the number of endmembers that are present in the data, but the results strongly depend on the signal to noise ratio of the data. Endmember extraction algorithms usually search for the extreme spectra in a dataset. This process can also be seen as identifying spectra which generate the convex hull of all spectra in the whole dataset. Common extraction algorithms are for instance the Pixel Purity Index (PPI) [BKG95], Automated Morphological Endmember Extraction (AMEE) [PMPP02] or the Orthogonal Subspace Projection (OSP) [HC94]. Many more algorithms can be found in the literature, please refer to [BDPD*12, PZPM10] for an overview.

Since these algorithms typically only extract a single standard endmember spectrum for each endmember class, they usually do not incorporate the spectral variability within an endmember class (see Fig. 5), which has been identified as one of the most profound sources of error in the estimation of abundances [SZPA12]. The recent literature presents solutions to improve the abundance estimation, e.g. the multiple endmember spectral mixture analysis (MESMA) [RGC*98] algorithm, please see the review of Somers et al. [SATC11] for further approaches and details. However, the premise of these solutions is the availability of a spectral library that allows the modeling of the endmembers variability, thus introducing limitations when measurements are not available or incomplete [SZPA12]. Furthermore, experimental results comparing different endmember variability reduction techniques are very scarce and this makes it hard to identify the most robust and most effective technique to build an understanding of how to match application and endmember reduction strategies [SATC11].

Spectral error evaluation can be done by comparing the results to accurate ground truth references, when available. But typically the accuracy of LSU is mainly quantified based on the fit between the reconstructed data S_{LSU}, obtained by using Eq. (1), and the raw data S_{RAW} [SATC11]. Here, several metrics (cf. Table 1) can be applied to evaluate the quality by pairwise computation of the distances between $\vec{s}_{RAW}(x,y)$ and $\vec{s}_{LSU}(x,y)$ for each pixel, resulting in a distance image, which may suffer from the ambiguity of distance values. Only a few approaches exist in the literature that are aiming the evaluation of spectral unmixing quality in a comprehensive way. Recently, Jimenez et al. [JMP12] developed a comprehensive tool, called HyperMix, that allows performing all steps of the spectral unmixing chain and doing quantitative comparison of algorithms, e.g. AMEE or OSP. Result analysis is done by using several distance metrics. Labitzke et al. [LBK12] have shown that visual analysis concepts can be applied meaningfully to approve a determined endmember set and to interactively refine/correct the set, if necessary. However, for the error evaluation they also focus on the common scalar-valued distance metrics.

Metric	Formula
Spectral Angle Distance (SAD)	$\mathrm{SAD}(\vec{x},\vec{y}) = \arccos\left(\frac{\langle\vec{x},\vec{y}\rangle}{\|\vec{x}\|_2\|\vec{y}\|_2}\right)$
Spectral Gradient Angle (SGA)	$\mathrm{SGA}(\vec{x},\vec{y}) = \mathrm{SAD}\left(\mathrm{SG_x},\mathrm{SG_y}\right)$ with $\mathrm{SG_x} = (x_2-x_1, x_3-x_2, \ldots, x_n-x_{n-1})$, $\mathrm{SG_y}$ analog
Normalized Euclidean Distance (NED)	$\mathrm{NED}(\vec{x},\vec{y}) = \left\|\frac{\vec{x}}{\|\vec{x}\|_2} - \frac{\vec{y}}{\|\vec{y}\|_2}\right\|_2$
Root Mean Square Deviation (RMSD)	$\mathrm{RMSD}(\vec{x},\vec{y}) = \frac{1}{n}\sum_{i=1}^{n}\left[(\vec{x}_i-\vec{y}_i)^2\right]^{\frac{1}{2}}$

Table 1: *Distance metrics applied to two n-dimensional spectra (vectors) $\vec{x}$ and $\vec{y}$.*

Both approaches map the spectral residual errors to single scalar error values, incorporating a significant loss of information, and do not allow detailed local (spectral) explorations. In contrast, we propose an enhanced spectral error visualization using color information, *colored distance metrics*, to express spectral errors and an *spectral error classification* scheme. Thus, giving a global (complete data) impression of the spectral residual error, which allows the guided identification of pixels with similar error behavior.

3. Overview

Compared to typical spectral visualization and analysis approaches, the aim of our approach is to verify the quality of the LSU-result, i.e. the completeness of the endmember set, and to optimize the set of endmembers in an interactive way, if necessary. Here, it is vital to examine and understand the influence of each endmember for both, the complete dataset (global) and also single pixels (local), e.g. to compare $\vec{s}_{RAW}(x,y)$ and $\vec{s}_{LSU}(x,y)$. Beside the quality-verification, our approach also allows the exploration of the LSU-result to facilitate the knowledge gathering process of previously unknown data, e.g. by exploring the distribution or the influence of endmembers in mixed spectra.

A conceptual overview of our approach is depicted in Fig. 2. Starting from a LSU-result, that has been determined by the typical LSU, the exploration process is driven by local and global visualizations. While the applied local visualizations allow the detailed examination of the reconstruction quality and the composition of a spectrum, the global views provide visual guidance by *expressive spectral error visualizations* to sophisticatedly guide an analyst to pixels of interest. All visualizations are linked to allow effective knowledge extraction. If necessary, the set of endmembers can be refined in a refinement step.

As a simple example, consider a typical LSU-result: Usually the unmixing quality is proved by applying similarity measures, which can lead to different quality impressions and the results can be ambiguous since different spectral errors can lead to same error values (see Sec. 2). Moreover, errors based on endmember variability can lead to conspicuous distance values. In a typical distance image these distance values cannot be distinguished from distance values that are

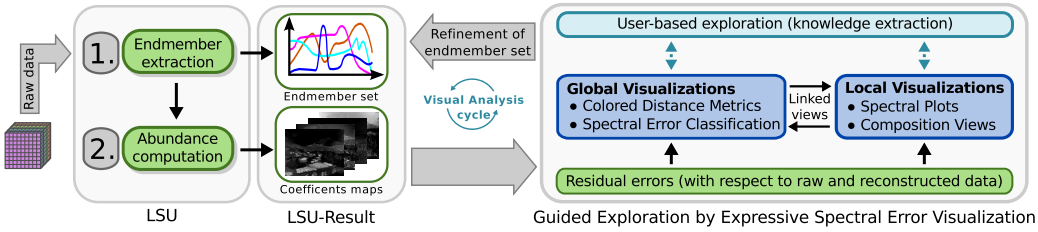

Figure 2: *Conceptual overview of the proposed visual analysis approach. The analysis starts with the typical process of LSU. In the guided exploration process the global* expressive spectral error visualizations *sophisticatedly guide an analyst to pixels of interest for local verification and exploration of the reconstruction quality of a spectrum. All visualizations are linked to allow effective knowledge extraction. If necessary, the set of endmembers can be instantly refined.*

originate from missing endmembers. Therefore, endmember variability can compromise the detection of endmembers which are missing completely. In an interactive approach this at the end would mean, in a worst-case, that an analyst has to explore all pixels to validate the unmixing result, which is undesired. In order to encounter these challenges and to reduce the interaction effort mainly two global *expressive spectral error visualizations* are presented.

Colored Distance Metrics enhance the typical distance images to achieve a visual grouping of pixels with a similar error behavior. Thus, reduces the ambiguity of distance values by coloring.

Spectral Error Classification is done to classify pixels in three terms: neglectable deviations and spectral errors due to endmember variability or missing endmembers.

Based on these global impressions, the analysts sophisticatedly can identify pixels of interest for comprehensive local validation.

4. Expressive Spectral Error Visualization

This section describes the details of both proposed global *expressive spectral error visualization* methods.

4.1. Colored Distance Metrics

Focusing on typical distance images, the differentiation of error values is improved by a proposed qualitative coloring of residual spectral errors, so-called *color signatures*, to visually group comparable residues. Thus, reducing ambiguity scalar values by color. In Fig. 3 (right) we can see a typical challenge of grayscaled distance images. Because of ambiguous distance values, distinct residual errors are indistinguishable. In contrast, the proposed *color signatures* allow a qualitative distinction of the residual errors by their colors. In order to best possibly express the distinct residual errors by a color, the variance of the residuals is used to achieve optimal color distribution and saturation. An enhanced distance image, where residual errors with same error behavior are visually grouped, is achieved by weighting a *color signature* with a common distance image.

In the following the calculation of the *color signatures* is discussed. Based on both, $\vec{s}_{RAW}(x,y)$ and $\vec{s}_{LSU}(x,y)$, the residual error $\vec{r}(x,y) = \vec{s}_{RAW}(x,y) - \vec{s}_{LSU}(x,y)$ and the absolute residuum $\vec{r}_{abs}(x,y) = |\vec{r}(x,y)|$ are calculated for each pixel. Then the $\vec{r}_{abs}(x,y)$ is considered as a spectrum and transformed into a *RGB*-color. Here, one of our goals is an intuitive interaction concept. Therefore, we propose a way for the spectrum transformation, which is intuitive and allows a user to assume a residual characteristic from its color. Thus, we compute the *RGB*-values

$$B = \frac{1}{u_B} \sum_{\lambda=k_B}^{l_B} \vec{r}_{abs}(x,y,\lambda),\ k_B = 1,\ l_B = u_B \tag{2}$$

$$G = \frac{1}{u_G} \sum_{\lambda=k_G}^{l_G} \vec{r}_{abs}(x,y,\lambda),\ k_G = l_B + 1,\ l_G = k_G + u_G \tag{3}$$

$$R = \frac{1}{u_R} \sum_{\lambda=k_R}^{l_R} \vec{r}_{abs}(x,y,\lambda),\ k_R = l_G + 1,\ l_R = k_R + u_R \tag{4}$$

by splitting the whole residual spectrum $\vec{r}_{abs}(x,y)$ in three intervals to correspond with the *RGB* color model. The deviations of each interval are summed up and averaged to achieve the *RGB*-values. In order to achieve the maximum brightness, the color values R, G and B are normalized by dividing through the maximal component value $\max(R,G,B)$. To achieve the best possible color distribution and saturation the size of each interval (u_R, u_G and u_B) is selected in the

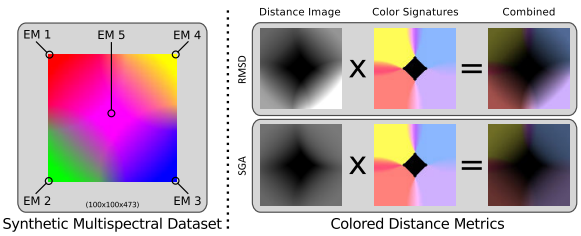

Figure 3: *Based on a synthetic dataset (left), consisting of five endmembers, the benefit of the* colored distance metrics *is exemplary depicted for two metrics (right). In this example, only the centered endmember (EM 5) was selected. Thus, the four missing endmembers lead to distinct residual errors. While the residuals are indistinguishable in the typical grayscale images, they are easily distinguishable in the* color signatures *enhanced distance images (combined).*

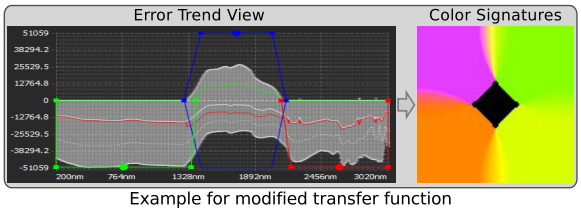

Figure 4: *The* error trend view*, where the white line is the average residuum, the green and red line represents the positive and negative averages of the residues. The dashed white line represents the local residuum of the current pixel. Here, the TFs of the* color signatures *of Fig. 3 are modified.*

way that the standard deviations σ_λ of the spectral bands

$$\sigma_\lambda = \sqrt{\frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} (\vec{r}_{abs}(x,y,\lambda) - \mu_\lambda)^2}, \quad (5)$$

$$\text{where } \mu_\lambda = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} \vec{r}_{abs}(x,y,\lambda) \quad (6)$$

are consistent in all three intervals

$$\frac{1}{3} \sum_{\lambda=1}^{n} \sigma_\lambda \approx \sum_{\lambda=1}^{u_B} \sigma_\lambda \approx \sum_{\lambda=u_B+1}^{u_B+u_G+1} \sigma_\lambda \approx \sum_{\lambda=u_B+u_G+1}^{u_B+u_G+u_R+1=n} \sigma_\lambda. \quad (7)$$

Here, $\vec{r}_{abs}(x,y,\lambda)$ is the residual value of band λ at pixel (x,y) and X, Y are the spatial resolution.

Based on the *color signatures*, a user can intuitively estimate from the colors in which of the three intervals the error occurs, e.g. red color means last third or magenta means error in the first and third range. Errors in nearly the complete range will lead to white, while a black color will mean, that the reconstruction error is low in all ranges. The *color signatures* are computed for all pixels, leading to an color signature image that element-wise can be weighted by arbitrary distance images, resulting in a *colored distance metric*, see Fig. 3 for an example. Since different metrics can lead to different quality impressions of results, our approach allows an analyst to simultaneously overview the distance images of all implemented metrics as well as the *color signature* image at the same time allowing an intuitive combination.

Modification of Color Signatures by Transfer Functions So far the result of the *color signatures* depends on the variance-based sizes of the three intervals R, G and B. Sometimes it is desirable to freely highlight or determine residual error characteristics of interest. Moreover, since light emission often is not uniform and sensors provide smaller intensity values in the infrared range, an adjusted weighting is useful to manually compensate such issues.

Based on residual statistics, that illustrate the error trend, a global view is proposed that allows the interactive design of transfer functions (TF) to fine-tune the initial intervals of R, G and B, see Fig. 4. Moreover, color signatures also can be completely changed, since the number of TFs is not limited to three. The *error trend view* shows the current residuum $\vec{r}(x,y)$ as well as several global average residual errors. In detail, the average residual error $\vec{r}_{avg}$ and the average positive $\vec{r}^{+}_{avg}$ as well as average negative $\vec{r}^{-}_{avg}$ residual error

$$\vec{r}_{avg} = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} \vec{r}(x,y) \text{ and} \quad (8)$$

$$\vec{r}^{+}_{avg} = \frac{1}{P^+} \sum_{x=1}^{X} \sum_{y=1}^{Y} \vec{r}^{+}(x,y),\ \vec{r}^{-}_{avg} \text{ analog} \quad (9)$$

are included with respect to all pixels, where

$$\vec{r}^{+}(x,y,\lambda) = \max\{\vec{r}(x,y,\lambda), 0\},\ \vec{r}^{-}(x,y,\lambda) \text{ analog.} \quad (10)$$

Here, P^+ and P^- are the number of positive and negative error values, respectively. Furthermore, for each band also the min- and maximum error value is depicted to show the maximum variance at the same time. The TFs can be defined via trapezoids and are related to the bands. Here, the width and the height of a trapezoid defines the wavelength interval and the weighting factor, respectively.

4.2. Spectral Error Classification

The main aim of the *spectral error classification* is to distinguish errors introduced by endmember variability from errors due to missing endmembers, to more purposefully guide an analyst to missing endmembers. Beside missing endmembers, as mentioned in Sec. 2, the endmember variability also is one of the most profound sources of error in the estimation of abundances [SZPA12]. The reason for this is, that typically FCLSU is applied to achieve a physically reasonable result. But, for instance when in a pixel low coefficient values $\alpha_j(x,y) < 1$ are necessary because of the variability and the sum of all coefficients is not the unity, then the enforced ASC introduces errors in the abundance estimation. An example of this issue is illustrated in Fig. 5 (right), where an optimal $\sum \alpha_j(x,y)$ is 0.72. But, because of the ASC in FCLSU the individual $\alpha_j(x,y)$ are scaled to sum up to one, which leads to an error that is introduced by the variability. As can be seen, when only ANC is enforced the reconstructed spectrum $\vec{s}_{LSU}(x,y)$ is comparable to $\vec{s}_{RAW}(x,y)$. This circumstance serves as idea to identify errors based on endmember variability. In detail the *spectral error classification* is based on two results of LSU, i.e. NCLSU and FCLSU.

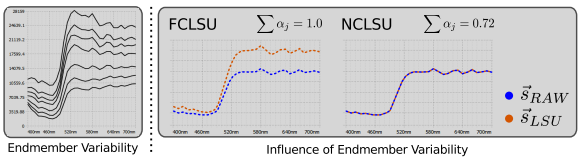

Figure 5: *Example for the variability in an endmember class (left). Illustration of the variability based abundance estimation error in FCSLU, in comparison to NCLSU (right).*

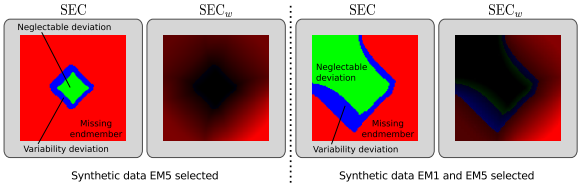

Figure 6: *Two examples of the proposed spectral error classification. First, based on the LSU-result already presented in Fig. 3 (left). Because of low variability errors in the first example, also a second scenario is depicted for more obvious illustration of the weighting in SEC$_w$ (right).*

Since the FCLSU-coefficients can be directly obtained from the NCLSU-coefficients by scaling the coefficients to fulfill the ASC, the computational effort is comparable to the typical FCLSU calculation. Based on both results, two corresponding RMSD-values are computed for each pixel

$$d^{NC}(x,y) = RMSD(\vec{s}_{RAW}(x,y), \vec{s}_{NCLSU}(x,y)) \quad (11)$$

$$d^{FC}(x,y) = RMSD(\vec{s}_{RAW}(x,y), \vec{s}_{FCLSU}(x,y)). \quad (12)$$

The two calculated deviation values are used to classify the type of deviation, neglectable deviation (ND), variability deviation (VD) and missing endmember (ME):

$$sec(x,y) = \begin{cases} \text{ND} & \text{if } 0 \leq d^{NC}(x,y),\ d^{FC}(x,y) < v \\ \text{VD} & \text{else if } 0 \leq d^{NC}(x,y) < v \\ \text{ME} & \text{else,} \end{cases} \quad (13)$$

where v is a user defined quality threshold. When both deviation values are smaller than v a pixel is classified as neglectable deviation, while in all other cases $d^{NC}(x,y)$ is used to distinguish the remaining types of deviations. Applying $sec(x,y)$ to all pixels generates a classification image SEC, see Fig. 6 for an example. Moreover, SEC can be weighted resulting in SEC$_w$

$$sec_w(x,y) = \begin{cases} sec(x,y)\dfrac{d^{NC}(x,y)}{\max(d^{NC}(\text{all pixels}))} & \text{if ND} \\ sec(x,y)\dfrac{f(x,y)}{\max(f(\text{all pixels}))} & \text{else} \end{cases} \quad (14)$$

$$\text{with } f(x,y) = \text{abs}\left(1 - \sum_{j=1}^{q} \alpha_j^{NC}(x,y)\right) \quad (15)$$

to easier discover pixels that have major deviations. Here, q is again the number of endmembers and for a pixel that is classified as *missing endmember* the normalized $d^{NC}(x,y)$ is used as weighting factor for $sec(x,y)$. In all other cases, the endmember variability is expressed in $sec_w(x,y)$ by the difference between the sum of the coefficients of the endmembers of the NCLSU $\alpha_j^{NC}(x,y)$ and the FCLSU, which is one. Consider Fig. 6 for an example of this weighting scheme.

5. Interactive Exploration

The graphical user interface of our visual analysis approach is shown in Fig. 7 and consists of several linked views. Beside several 2D data visualizations, e.g. sRGB color representation, the *global visualizations* mainly consists of the two proposed *expressive spectral error visualizations* for global guidance to identify pixels for further local investigations. The *local visualizations* allow the analysts to explore the currently selected pixel to validate the reconstruction, e.g. by exploring the composition of a mixed spectrum. In the reconstruction and composition views the white line shows the raw spectrum $\vec{s}_{RAW}(x,y)$ and the blue line/area represents the reconstruction $\vec{s}_{LSU}(x,y)$. The visualization of the residual error $\vec{r}(x,y)$ helps to overview in which wavelength ranges errors occur and which quantity they have, if present. Moreover, the visualization of the endmember set provides the opportunity to examine the spectral distribution of all endmembers V at the same time. By this, the analyst can see how different the distributions are, e.g. to see if they partly converge or diverge. Here, also the currently selected spectrum is visualized to see easily which endmember resembles the current spectrum best. Or the other way around, to see how distinct the current spectrum is compared to V. Thus, to see how good it can serve as a new endmember. The view of the *coefficients maps* helps to overview in which region which endmember has influence and to what extent. Each map is colored in the respective color of the corresponding endmember, which can be modified.

All visualizations allow zooming into details and are linked. Thus, zooming or clicking on a pixel in a global view leads to an direct update in all views. In each global view the current pixel position is highlighted. The user can visually explore the LSU-result step-by-step to gain insight. Our approach also allows the *interactive refinement* of endmember sets, in the manner of Labitzke et al. [LBK12], which has turned out to be a good improvement. Thus, a user can directly use the perceived knowledge in case of wrong or incomplete sets to refine the LSU-result, by removal or insertion of an endmember.

6. Results

Before we discuss limitations of our approach, we first show usage examples for two domains, confocal Raman microscopy and multispectral scene data.

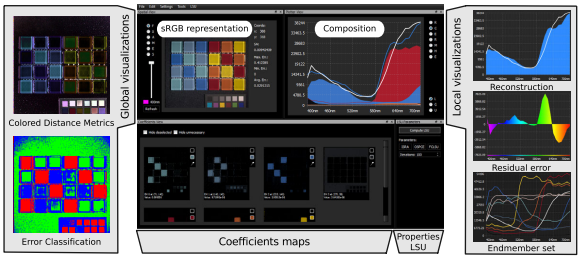

Figure 7: *The graphical user interface, that focuses on several linked views: coefficients maps, global spectral error views and per-pixel views for detailed local investigations.*

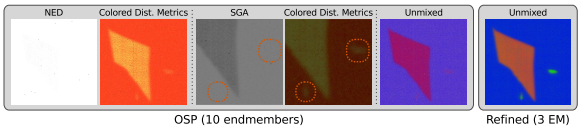

Figure 8: *The OSP-result for the graphene dataset. Compared to the usual distance images, error values are easier to distinguish in the proposed* colored distance metrics *(SGA images are zoomed in). The quality improvement of interactive refinement is shown by the unmixed images.*

6.1. Usage Example

Since, the focus is the validation of LSU-results and not the endmember detection algorithms itself, we will not concentrate on details of the detection process. In the following we mainly utilized two common methods, OSP and PPI. The inverse operation is done by using the implementation of the image space reconstruction algorithm of Sánchez et al. [SMPC10].

Graphene: The graphene dataset was acquired by using a confocal Raman microscope which provides a very high spectral density, 510 bands, from 322 nm to 870 nm. This dataset has a spatial resolution of 256×256 pixels and contains some imperfections, e.g. peaks because of cosmic rays, that are quite challenging for automatic algorithms. As shown in Fig. 8, all distance images present high error values, especially for the normalized euclidean distance (NED). While two error regions are already distinguishable in the distance image of the spectral gradient angle (SGA), the NED-image reflects a worst-case, where no errors are distinguishable. The interpretation of both distance images are improved by the combination with our proposed *color signatures*. As a result of this, different error regions are getting visible in all *colored distance metrics*. Also two other small error parts, see marked areas of SGA, are now noticeable. By means of our approach, the automatically estimated set of endmembers (applying OSP) were refined from ten to only three spectra. The quality improvement is shown in the comparison of the unmixed images.

Peppers: The peppers dataset provides 31 spectral bands (400-700 nm), has a spatial resolution of 512×512 pixels and is available as free download from the CAVE multispectral image database (`http://www.cs.columbia.edu/CAVE/databases/multispectral`). This dataset mainly consists of two red, two green and two yellow peppers. For each color, one artificial pepper was used. Fig. 9 shows the initial reconstruction quality for both algorithms, PPI and OSP. Please note, in both cases the algorithms have not found the dark background. Thus, we manually added a background-spectrum in each case to facilitate the upcoming explanation of the verification process. In both depicted results the discriminability of the error values is improved by the proposed *colored distance metrics*. Moreover, the spectral error classification view supports the analyst in the determination of missing constituent spectra. Comparing the marked areas of the OSP-result, it is getting obvious that not all high error values are missing endmembers. On the one hand, the magenta colored area shows error values that are quite high, but the reason for this is the endmember variability and not a missing constituent. On the other hand, the errors in the yellow area are due to missing endmembers, which are distinguishable in the *colored distance metrics*. The quality of the reconstructions, based on the OSP-, PPI- and the refinement, are illustrated in Fig. 10 by showing the sRGB transformed reconstructed multispectral datasets. In order to judge the quality, also the sRGB-image of the raw data is included.

6.2. Limitations

As shown before, the common grayscaled distance images can be improved with our proposed *color signatures* to achieve *colored distance metrics*. Here, residues are roughly subdivided into three intervals. It may happen that different errors can produce comparable colors, when e.g. one error has deviations in the first part of an interval and another error has deviations in the second part of the same interval, while the remaining parts of both errors are comparable. The analyst can still notice the wavelength range of interest, but can not distinguish the different errors so easily. Here, the manually modification of the *colored distance metrics* by TF design in the error trend view can be used to enhance the analysis and highlight the affected bands separately. However, a complete change of TFs may result in a less intuitive color representation, especially when mixed colors are applied as representative TF-colors.

7. Conclusion

In this paper, we have presented a visual analysis approach for interactive exploration, verification and optimization of

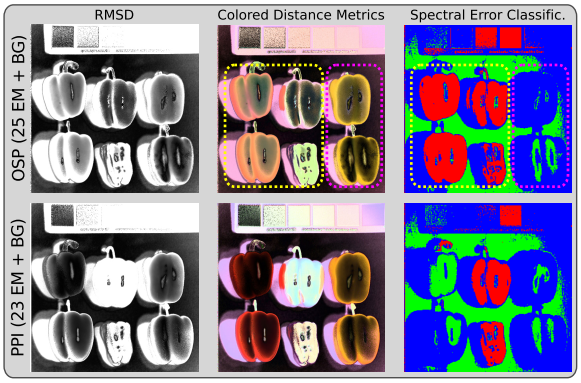

Figure 9: *Results of OSP and PPI for the peppers data, both with added background. The interpretation of the results is facilitated by the two proposed global visualizations,* colored distance metrics *and* spectral error classification*, in terms of both, easier discriminability of different residual errors and easier identification of missing endmembers.*

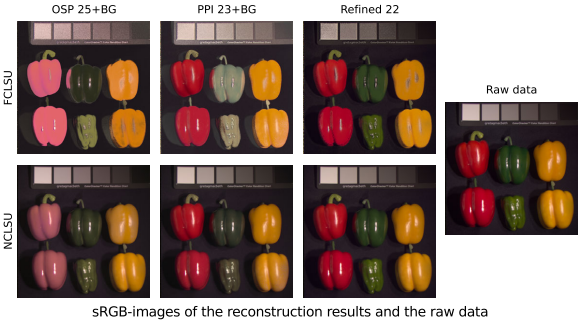

Figure 10: *The reconstruction quality is shown by sRGB transformed datasets of the respective results and the raw data. As can be seen, mainly the green peppers are not well represented in the initial results of OSP and PPI. In case of OSP, also the red peppers are not satisfying reproduced.*

spectral unmixing results. Here, two global *expressive spectral error visualizations* are introduced to efficiently guide a user to specific spectra for local exploration. *Color signatures* are used to enhance common distance metrics, so-called *colored distance metrics*, to reduce the ambiguity of distance values by coloring. Furthermore, the *color signatures* are freely adjustable by the design of transfer functions, based on residual statistics. In addition, the global error impressions are facilitated by a *spectral error classification* view. All these means help to make correlations visible, thus guide a user to interesting pixels for detailed local investigation. Future work will be dedicated to use the proposed approach to gather new insights about LSU, e.g. to develop more advanced analytical data processing methods.

Acknowledgments

Activities leading to this work have been funded by the German Research Foundation (DFG) in the context of the Research Training Group 1564 Imaging New Modalities.

References

[BDPD*12] BIOUCAS-DIAS J. M., PLAZA A., DOBIGEON N., PARENTE M., DU Q., GADER P., CHANUSSOT J.: Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 5*, 2 (2012), 354–379. 1, 3

[BKG95] BOARDMAN J. W., KRUSE F. A., GREEN R. O.: Mapping target signatures via partial unmixing of AVIRIS datas. *Summaries of the Fifth Annual JPL Airborne Geoscience Workshop 1* (1995). 3

[BL02] BIEHL L., LANDGREBE D.: Multispec: a tool for multispectral–hyperspectral image data analysis. *Computers and Geosciences 28* (2002), 1153–1159. 2

[CD04] CHANG C.-I., DU Q.: Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Trans. Geoscience and Remote Sensing 42*, 3 (2004), 608–619. 3

[CRHW09] CUI M., RAZDAN A., HU J., WONKA P.: Interactive hyperspectral image visualization using convex optimization. *IEEE Trans. Geoscience and Remote Sensing 47*, 6 (2009), 1673–1684. 2

[HC94] HARSANYI J. C., CHANG C.-I.: Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. *IEEE Trans. Geoscience and Remote Sensing 32*, 4 (1994), 779–785. 3

[HKW12] HABEL R., KUDENOV M., WIMMER M.: Practical spectral photography. *Computer Graphics Forum 31*, 2pt2 (2012), 449–458. 1

[JA10] JORDAN J., ANGELOPOULOU E.: Gerbil - A Novel Software Framework for Visualization and Analysis in the Multispectral Domain. In *Vision, Modeling & Visualization (VMV)* (2010), vol. 1, pp. 259–266. 2

[JMP12] JIMENEZ I., MARTIN G., PLAZA A.: A new tool for evaluating spectral unmixing applications for remotely sensed hyperspectral image analysis. *Int. Conf. Geographic Object-Based Image Analysis (GEOBIA), Rio de Janeiro, Brazil* (2012), 644–648. 1, 2, 3

[Kes03] KESHAVA N.: A survey of spectral unmixing algorithms. *Lincoln Laboratory Journal 14*, 1 (2003), 55–78. 1, 2, 3

[KHK*12] KIM M. H., HARVEY T. A., KITTLE D. S., RUSHMEIER H., DORSEY J., PRUM R. O., BRADY D. J.: 3D imaging spectroscopy for measuring hyperspectral patterns on solid objects. *ACM Trans. Graphics (Proc. SIGGRAPH 2012) 31*, 4 (2012), 38:1–11. 1

[KZD*10] KIM S. J., ZHUO S., DENG F., FU C.-W., BROWN M.: Interactive visualization of hyperspectral images of historical documents. *IEEE Trans. Visualization and Computer Graphics 16*, 6 (2010), 1441–1448. 2

[LBK12] LABITZKE B., BAYRAKTAR S., KOLB A.: Generic visual analysis for multi- and hyperspectral data. *Data Mining and Knowledge Discovery, Special Issue: Intelligent Data Visualization* (2012), 117–145. 1, 3, 6

[PMPP02] PLAZA A., MARTÍNEZ P., PÉREZ R., PLAZA J.: Spatial/spectral endmember extraction by multidimensional morphological operations. *IEEE Trans. Geoscience and Remote Sensing 40*, 9 (2002), 2025–2041. 3

[PZPM10] PLAZA A., ZORTEA M., PLAZA J., MARTÍN G.: Recent developments in spectral unmixing and endmember extraction. In *Optical Remote Sensing - Advances in Signal Processing and Exploitation Techniques* (2010), vol. 3 of *Augmented Vision and Reality*, Springer, p. 300 p. 1, 2, 3

[RGC*98] ROBERTS D., GARDNER M., CHURCH R., USTIN S., SCHEER G., GREEN R.: Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment 65*, 3 (1998), 267–279. 3

[SATC11] SOMERS B., ASNER G. P., TITS L., COPPIN P.: Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment 115*, 7 (2011), 1603–1616. 1, 3

[SMPC10] SÁNCHEZ S., MARTÍN G., PLAZA A., CHANG C.-I.: GPU implementation of fully constrained linear spectral unmixing for remotely sensed hyperspectral data exploitation. vol. 7810, SPIE, (2010), p. 78100G. 2, 7

[SZPA12] SOMERS B., ZORTEA M., PLAZA A., ASNER G. P.: Automated extraction of image-based endmember bundles for improved spectral unmixing. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 5*, 2 (2012), 396–408. 3, 5

[TAE*11] TATU A., ALBUQUERQUE G., EISEMANN M., BAK P., THEISEL H., MAGNOR M., KEIM D.: Automated analytical methods to support visual exploration of high-dimensional data. *IEEE Trans. Visualization and Computer Graphics 17*, 5 (2011), 584–597. 2