SPECTRAL UNMIXING BASED CONSTRUCTION OF LUNAR MINERAL ABUNDANCE MAPS

V. Bernhardt†, A. Grumpe†, C. Wöhler* *

* Image Analysis Group, TU Dortmund University, D-44227 Dortmund, Germany (vladimir.bernhardt, arne.grumpe, christian.woehler)@tu-dortmund.de

Commission III, WG III/II

KEY WORDS: Spectral unmixing, Mineralogical maps, Moon

ABSTRACT:

In this study we apply a nonlinear spectral unmixing algorithm to a nearly global lunar spectral reflectance mosaic derived from hyperspectral image data acquired by the Moon Mineralogy Mapper (M3) instrument. Corrections for topographic effects and for thermal emission were performed. A set of 19 laboratory-based reflectance spectra of lunar samples published by the Lunar Soil Characterization Consortium (LSCC) were used as a catalog of potential endmember spectra. For a given spectrum, the multi-population population-based incremental learning (MPBIL) algorithm was used to determine the subset of endmembers actually contained in it. However, as the MPBIL algorithm is computationally expensive, it cannot be applied to all pixels of the reflectance mosaic. Hence, the reflectance mosaic was clustered into a set of 64 prototype spectra, and the MPBIL algorithm was applied to each prototype spectrum. Each pixel of the mosaic was assigned to the most similar prototype, and the set of endmembers previously determined for that prototype was used for pixel-wise nonlinear spectral unmixing using the Hapke model, implemented as linear unmixing of the single-scattering albedo spectrum. This procedure yields maps of the fractional abundances of the 19 endmembers. Based on the known modal abundances of a variety of mineral species in the LSCC samples, a conversion from endmember abundances to mineral abundances was performed. We present maps of the fractional abundances of plagioclase, pyroxene and olivine and compare our results with previously published lunar mineral abundance maps.

1. INTRODUCTION

Systematic analyses of the minerals that constitute the lunar surface material based on orbital multispectral and hyperspectral images of the Moon have been performed since the Clementine mission in 1994 (Nozette et al., 1994). Such image-based approaches rely on the framework of reflectance spectroscopy.

Global lunar maps of the abundances of the elements Fe and Ti have been constructed based on Clementine UV/VIS and Kaguya Multiband Imager (MI) multispectral imagery, where the calibration was performed based on laboratory data derived from returned lunar samples (Lucey et al., 2000; LeMouël et al., 2000; Otake et al., 2012). Elemental abundances of the elements Ca, Al, Fe, Mg, Ti and O have been mapped by (Shkuratov et al., 2005) using a regression between Clementine UV/VIS multispectral data and low-resolution Lunar Prospector Gamma Ray Spectrometer (LP GRS) elemental abundance data (Lawrence et al., 1998). These methods rely on the absorption band near 1 $\mu$m wavelength which is due to the presence of mafic minerals such as pyroxene or olivine. An algorithm for estimating the Fe abundance based on the properties of the also pyroxene-related absorption band near 2 $\mu$m has been introduced by (Bhatt et al., 2012) using hyperspectral image data of the Moon Mineralogy Mapper (M3) (Pieters et al., 2009) and SIR-2 point spectrometer (Mall et al., 2009) instruments on the Indian spacecraft Chandrayaan-1. A regression-based approach similar to the one of (Shkuratov et al., 2005) which takes into account a set of spectral parameters describing the 1-$\mu$m and 2-$\mu$m absorption bands in the M3 reflectance spectra has been developed for the elements Fe, Ca, Al and Mg (Wöhler et al., 2014). That method was calibrated with respect to LP GRS elemental abundances. It has been extended to the element Ti (Bhatt et al., 2015).

Although elemental abundance maps are very helpful for examining the lunar mineralogy, the detected elements cannot be unambiguously assigned to specific minerals. For example, clinopyroxene in mare regions and plagioclase in highland regions both contain the element Ca, and orthopyroxene and olivine may both contain Mg (Papike et al., 1991). Complementary knowledge about the surface minerals in addition to elemental composition can be obtained by the technique of spectral unmixing (Keshava and Mustard, 2002). Spectral unmixing aims at reconstructing the observed spectrum based on a set of spectra of basis materials, the so-called spectral endmembers, using a linear or nonlinear mixing model, where a nonlinear mixing model is required for intimate mixtures of different constituents such as the lunar regolith (Keshava and Mustard, 2002). A commonly applied approach to nonlinear spectral unmixing is based on the Hapke model (Hapke, 1984; Hapke, 2002), which allows for a transformation of the nonlinear unmixing of the spectral reflectance into a linear unmixing of the single-scattering albedo spectrum (Mustard and Pieters, 1989). A difficult problem in spectral unmixing is the selection of an appropriate set of endmembers for the spectrum under study from a large catalogue of potential endmembers. For a limited set of six potential endmembers, (Dobigeon and Tourneret, 2009) used a reversible jump Markov chain Monte Carlo algorithm. Another approach is to impose a sparseness constraint on the unmixing result, such that solutions in which only a few endmembers have non-zero abundances are preferred (Iordache et al., 2010, 2014). Alternatively, a sparse set of endmembers can be obtained by subsequently removing individual endmembers from an initially large set (Iordache et al., 2011).
The genetic algorithm based endmember selection scheme has been proposed by (Rommel et al., 2017), which yields the combination of endmembers that minimizes a fitness function depending on the root-mean-square difference between the observed and the reconstructed single-scattering albedo spectrum and the deviations between several spectral parameters derived from the observed and reconstructed reflectance spectrum, respectively.

Maps of the distribution of lunar key minerals have been constructed by (Lucey, 2004) based on Clementine UV/VIS spectral reflectance data, relying on a database that contains about 85000 spectra of mixtures of varying mineral and elemental composition, regolith grain size and degree of spaceweathering, which were computed using the Hapke model. Observed Clementine spectra were compared by (Lucey, 2004) with the precomputed model spectra in order to infer the mineral composition, where only immature soils that have not undergone strong spaceweathering processes were considered. A similar algorithm has been applied to a global Kaguya MI multispectral mosaic by (Lemelin et al., 2016).

In this paper we utilize a version of the framework of (Rommel et al., 2017) in which the genetic algorithm is replaced by the multi-population population-based incremental learning algorithm (Baluja, 1994). Folly and Venayagamoorthy, 2013. Based on the obtained spectral unmixing results, nearly global maps of the fractional abundances of key lunar minerals are constructed.

### 2. ENDMEMBER SELECTION FOR SPECTRAL UNMIXING

#### 2.1 Population-based incremental learning (PBIL)

The method of population-based incremental learning (PBIL) (Baluja, 1994) is conceptionally closely related to the genetic algorithm (GA) (Marsland, 2009). In this study we follow the approach of (Baluja, 1994) which is outlined in this paragraph. In contrast to the GA, the binary population is replaced in the PBIL algorithm by an array of probabilities having a length that is equal to the length of the catalogue of potential endmembers. Each position is assigned the probability that the corresponding endmember is contained in the observed spectrum. At the early stages in the optimization procedure of a GA, the population is very diverse and the individuals cover a large area in the search space. Correspondingly, the PBIL probability vector shows the highest diversity if all array positions have a value of 0.5. Towards the end of the evolutionary process, the individuals of the GA resemble each other closely. This is due to the fact that the fittest individuals have procreated and shaped the populations of the next generations. Likewise, the PBIL probability vector has to represent a prototype of a fit individual, i.e., the probabilities of the array entries have converged towards 1 or 0. In contrast to the GA, where operations like mating and mutation are imposed on the population, PBIL updates the probability vector directly.

According to (Baluja, 1994), the general approach is to generate an initial population by using a uniformly distributed probability vector at the beginning of the optimization procedure. Subsequently, a loop is entered which terminates if a specified number of iterations has been reached. In each iteration cycle a population of individuals is created, and the best and worst individuals are determined. After evaluating the generated population, the values of the probability vector are shifted towards representing the fittest individual in that generation. Hence, the probability of the $i$-th entry is updated based on the best individual according to

$$p_i \leftarrow (1 - \mu)p_i + \mu w_i$$

with $\mu$ as the learning rate and $w_i$ as the probability value of the fittest individual. In order to exploit also knowledge about the bad individuals, we have slightly adapted the original PBIL algorithm by using Equation 1 with a negative learning rate in order to update the probability vector away from the worst individual if the worst individual has a different value at position $i$ than the fittest individual.

Since the generated population is discarded after every probability vector update, elite individuals can in principle be introduced by copying the best individuals of the current generation into the population of the next generation, but this approach was not applied due to its tendency to drive the algorithm towards a local optimum.

#### 2.2 Multi-population PBIL (MPBIL)

In order to improve the performance of the standard PBIL, the influence of using multiple populations in the learning procedure has been explored in (Folly and Venayagamoorthy, 2013). They link the tendency of PBIL to converge towards local optima to a loss of diversity. In this paragraph the main aspects of the work of (Folly and Venayagamoorthy, 2013) are summarized.

The standard PBIL procedure relies on a single probability vector for which an optimal solution has to be found in a search problem. To increase diversity within the solution space, a second probability vector is introduced. Both probability vectors are initialized with uniform values of 0.5, and two populations are created using the respective probability vector. To make this approach comparable to the standard PBIL procedure, the population sizes $n_1$ and $n_2$ of the newly created populations is set to the identical values $n_1 = n_2 = n_{pop}/2$ with $n_{pop}$ as the total number of individuals. All following steps are performed for both probability vectors separately. As the learning process proceeds, one of the two probability vectors will yield better results than the other. To reward the better performing probability vector, the size of its population is increased, while the population size of the other probability vector is decreased. The best individuals of both populations compete for the number of individuals the corresponding probability vector is allowed to generate in the next iteration. The minimum and maximum number of individuals a population is allowed to have is restricted to 0.4 $n_{pop}$ and 0.6 $n_{pop}$, respectively. Eventually, the probability vectors are mutated, using a “forgetting factor” to shift them towards the neutral state with $p_i = 0.5$.

The optimization is terminated once a predefined number of iteration cycles has been completed.

#### 2.3 MPBIL-based spectral unmixing

For each set of endmembers proposed by the MPBIL, linear unmixing of the spectrum of Hapke’s single-scattering albedo (Hapke, 2002) is performed as proposed by (Mustard and Pieters, 1989), corresponding to nonlinear unmixing of the reflectance spectrum. In the MPBIL-based optimization procedure, the fitness function introduced by (Rommel et al., 2017) is employed. It corresponds to a weighted sum of the mean squared difference between the observed and the reconstructed single-scattering albedo spectrum and the squared differences between spectral parameters describing the absorption bands around 1 $\mu$m and 2 $\mu$m of the observed and the reconstructed reflectance spectrum.
3. CONSTRUCTION OF NEARLY GLOBAL LUNAR MINERAL ABUNDANCE MAPS

To construct a global M^3 spectral reflectance mosaic, we started from the M^3 level-1B spectral radiance data set (download from the Planetary Data System (PDS) at [https://pds-imaging.jpl.nasa.gov/volumes/m3.html](https://pds-imaging.jpl.nasa.gov/volumes/m3.html)), which we resampled to a resolution of 20 pixels per degree, corresponding to about 1.5 km per pixel at the lunar equator. The M^3 hyperspectral image data cover the wavelength range of 0.42–3.0 µm by 85 spectral bands, where the image resolution in global mode corresponds to 140 m per pixel (Pieters et al., 2009). The method of (Wöhler et al., 2014) was used to estimate the temperature of the lunar surface and subtract the thermal emission component from the spectral radiance. The solar illumination direction and the viewing direction are available on the PDS together with the spectral radiance data. The nearly global topographic map GLD100 (Scholten et al., 2012) was used to determine the surface normal vector for each pixel of the mosaic, which allows for computing the pixel-wise solar incidence angle and the emission angle. The smallest details reliably recovered in the GLD100 have extensions of about 1.5 km (Scholten et al., 2012), which fits with the resolution of the resampled M^3 mosaic. All spectral radiance measurements that are available for a specific pixel were used to determine the pixel-specific single-scattering albedo according to the Hapke model (Hapke, 1984; Hapke, 2002), where all other parameters of the Hapke model were chosen according to the first global lunar solution of (Ware, 2004). The spectral reflectances were then normalised to the standard illumination and viewing geometry of 30° incidence angle and 0° emission angle (Pieters, 1999), based on the Hapke model. The result of these operations is a nearly global M^3 reflectance mosaic of 20 pixels per degree resolution that has been corrected for the effects of thermal emission, topography and changing illumination and viewing geometry during image acquisition. This mosaic is the basis for the spectral unmixing analyses in this paper.

Although the MPBIL algorithm is computationally less expensive than the GA, it is not possible to apply spectral unmixing to all pixels of the nearly global M^3 mosaic. Hence, the mosaic was clustered into 64 spectral classes in an unsupervised manner using a self-organizing map (SOM) (Kohonen, 2001). As features used for SOM clustering the pixel-specific abundances of the elements Fe, Ca, Mg and Ti were used, which were computed for the complete mosaic based on the regression-based algorithm of (Wöhler et al., 2014; Bhattacharyya et al., 2015). To each of the 64 cluster prototypes the MPBIL-based spectral unmixing algorithm with a population size of n_{pop} = 20 was applied, yielding 64 cluster-specific sets of endmembers. As a catalogue of potential endmembers for spectral unmixing, the set of 19 laboratory spectra of lunar mare and highland samples provided by the Lunar Soil Characterization Consortium (LSCC) (Taylor et al., 2001; Taylor et al., 2010) was used. Only the 20–45 µm grain size fraction was taken into account. This data set is well suited for spectral unmixing on large scales because it consists of soil sample spectra exhibiting similar spaceweathering effects as remotely sensed spectra of the lunar soil (Taylor et al., 2001).

Each pixel of the mosaic was assigned to the prototype having the smallest Euclidean distance in the space spanned by the elemental abundances of Fe, Ca, Mg and Ti used for clustering. With the set of endmembers previously determined for the assigned prototype, pixel-wise linear unmixing of the single-scattering albedo spectrum (Mustard and Pieters, 1989) was performed. The obtained endmember fractions were normalized to a sum of 1. The LSCC compositional laboratory analyses state the modal abundances of a variety of mineral species for all 19 samples used as spectral endmembers (Taylor et al., 2001; Taylor et al., 2010). These data allow for a conversion from endmember fractions into mineral fractions. Here the agglutinitic glass component was excluded from the mineral fractions under the assumption that the agglutinitic glass has the same mineral composition as the non-glassy component of the sample. It is noteworthy that the obtained mineral fractions cannot be directly interpreted as mass fractions because the fractional endmember abundances obtained by spectral unmixing are not equal to the endmember mass fractions.

4. RESULTS

Nearly global maps of the minerals plagioclase, pyroxene and olivine are shown in Figures 1–3. They were obtained using the MPBIL algorithm with a population size of n_{pop} = 20.

The plagioclase abundance map (Figure 1) reveals rather uniform values slightly below 90% throughout the lunar highlands and around 35%–40% in the maria. The large South Pole Aitken (SPA) basin situated on the southern lunar hemisphere around 180° longitude exhibits intermediate plagioclase fractions of about 60% outside the basaltic mare patches of SPA. Our map is similar to the plagioclase abundance map of (Lucey, 2004), with the main difference that in that work similarly low plagioclase abundances for maria and SPA are derived. The plagioclase map of (Lemelin et al., 2016) shows plagioclase abundance variations of nearly 20% between farside and nearside highlands, which neither appear in our map nor in the map of (Lucey, 2004). Unfortunately, the plagioclase abundance map of (Lemelin et al., 2016) is scaled in a way that no exact abundance values can be extracted for the maria.

The pyroxene abundance map (Figure 2) shows fractions around 27% in the nearside and farside maria, with extended local maxima reaching 35%–40% in northwestern Oceanus Procellarum, northeastern Mare Imbrium and northern and northwestern Mare Serenitatis. The pyroxene content in SPA outside the basaltic patches corresponds to 15%–18%. Our pyroxene abundance map is similar to a sum of the clinopyroxene and orthopyroxene maps of (Lucey, 2004), with the main difference that similarly high total pyroxene abundances are inferred for the nearside maria and SPA by (Lucey, 2004). Qualitatively, our pyroxene map is also similar to the sum of low-Ca pyroxene and clinopyroxene derived by (Lemelin et al., 2016), but the summed pyroxene fractions of (Lemelin et al., 2016) in the nearside maria correspond to 60% and more. In contrast, the laboratory analyses of (Taylor et al., 2001) of lunar mare samples have shown that the total pyroxene content of these samples is always less than 34% and in many cases below 20%.

The olivine abundance map (Figure 3) shows relatively high fractions around 6.5% in western Oceanus Procellarum, western Mare Imbrium and parts of Mare Serenitatis, Mare Tranquillitatis, Mare Crisium, Mare Australe and small parts of SPA. The olivine abundances in the other maria are slightly lower (5%–6%) and show variations between 4% and 4.5% in the highlands. Several of the areas in Oceanus Procellarum indicated as olivine-rich in our map have also been found to have a high olivine content in (Varatharajan et al., 2014; Staid et al., 2016). The high-olivine basaits of western Oceanus Procellarum and western Mare Imbrium have also been mapped by (Lucey, 2004), however, with
olivine abundances of 40%–50%. The olivine content of the other maria hardly differs from that of the highlands in the map of (Lucey, 2004). The olivine abundance map of (Lemelin et al., 2016) shows the highest olivine abundances of 15%–20% in western Oceanus Procellarum, western Mare Imbrium, Mare Serenitatis and Mare Crisium, while all other maria (except a few small patches) and the highlands have olivine fractions below 3%. The olivine content of SPA is near zero in the map of (Lemelin et al., 2016).

5. SUMMARY AND CONCLUSION

In this paper we have described an approach to endmember selection for spectral unmixing that relies on multi-population population-based incremental learning (MPBIL). This framework has been applied to a nearly global M³ mosaic of 20 pixels per degree resolution. The mosaic has been divided into 64 clusters, and the MPBIL-based spectral unmixing framework has been applied to each of the cluster prototypes. For each individual pixel of the mosaic, spectral unmixing is performed using the set of endmembers determined for the cluster prototype which is most similar to the pixel under consideration. As a catalog of potential spectral endmembers, laboratory spectra of the 19 LSCC mare and highland samples (Taylor et al., 2001) have been used. Based on the known fractional abundances of key minerals in these samples, the endmember abundances obtained by spectral unmixing were converted to mineral abundances. We have presented abundance maps of the minerals plagioclase, pyroxene and olivine. Our maps are qualitatively consistent with previously published mineral maps that were inferred from Clementine and Kaguya multispectral image data (Lucey, 2004) [Lemelin et al., 2016], but there are differences regarding regional structures as well as the absolute mineral abundance values. Future work will involve a detailed analysis of the similarities and discrepancies between our mineral maps and previously published mineral abundance data sets.

REFERENCES


Figure 1. Map of the fractional abundance of plagioclase (in percent). Black pixels denote missing data.
Figure 2. Map of the fractional abundance of pyroxene (in percent). Black pixels denote missing data.
Figure 3. Map of the fractional abundance of olivine (in percent). Black pixels denote missing data.