

VERTEX COMPONENT ANALYSIS GPU-BASED IMPLEMENTATION FOR HYPERSPPECTRAL UNMIXING

*José M. Rodríguez Alves^a, José M. P. Nascimento^{a,b}, Antonio Plaza^c,
Sergio Sanchez^c, José M. Bioucas-Dias^{a,d}, and Vítor Silva^e*

^aInstituto de Telecomunicações, Lisbon, Portugal

^bInstituto Superior de Engenharia de Lisboa, Lisbon, Portugal

^cHyperspectral Computing Laboratory, University of Extremadura, Cáceres, Spain

^dInstituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal

^eInstituto de Telecomunicações, DEEC, University of Coimbra, Coimbra, Portugal

ABSTRACT

Endmember extraction (EE) is a fundamental and crucial task in hyperspectral unmixing. Among other methods vertex component analysis (VCA) has become a very popular and useful tool to unmix hyperspectral data. VCA is a geometrical based method that extracts endmember signatures from large hyperspectral datasets without the use of any a priori knowledge about the constituent spectra.

Many Hyperspectral imagery applications require a response in real time or near-real time. Thus, to met this requirement this paper proposes a parallel implementation of VCA developed for graphics processing units. The impact on the complexity and on the accuracy of the proposed parallel implementation of VCA is examined using both simulated and real hyperspectral datasets.

Index Terms— Hyperspectral Unmixing, Endmember Extraction, Vertex Component Analysis, Graphics Processing Unit, Parallel Methods.

1. INTRODUCTION

Among the remote sensing modalities, hyperspectral imaging has recently emerged as a powerful passive technology [1]. This technology have been widely used in the fields of urban and regional planning, water resource management, environmental monitoring, food safety, counterfeit drugs detection, oil spill and other types of chemical contamination detection, biological hazards prevention, and target detection for military and security purposes [2, 3].

Hyperspectral imaging instruments such as NASA Jet Propulsion Laboratory's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) [4] sample the reflected solar radiation from the Earth surface in the portion of the spectrum extending from the visible region through the near-infrared and mid-infrared (wavelengths between $0.3 \mu\text{m}$ and $2.5 \mu\text{m}$)

in hundreds of narrow (on the order of 10 nm) contiguous bands [5]. This high spectral resolution can be used for object detection and for discriminating between different objects based on their spectral characteristics [6]. However, this huge spectral resolution yields large amounts of data to be processed. For example, AVIRIS collects a 512 (along track) 614 (across track) 224 (bands) 12 (bits) data cube in 5 seconds, corresponding to about 140 MBs. Similar data collection ratios are achieved by other spectrometers [7]. This introduces significant requires from the viewpoint of efficiently processing, transmitting and storing the sheer volumes of data generated in hyperspectral imaging applications [8, 9].

Very often, the resolution cell corresponding to a single pixel in an image, is tens of meters, thus containing several substances (so-called *endmembers*) [5]. In this situation, the scattered energy is a mixing of the endmember spectra. Linear mixture model assume that each pixel is a linear combination of endmember signatures weighted by the correspondent abundance fractions [10]. Over the last decade, several algorithms have been developed to unmix hyperspectral data sets, including vertex component analysis (VCA) [11], automated morphological endmember extraction (AMEE) [12], pixel purity index (PPI) [13], N-FINDR [14], iterative error analysis (IEA) [15], simplex growing algorithm (SGA) [16], sequential maximum angle convex cone (SMACC) [17], alternative volume maximization (AVMAX, SVMAX) [18], among several others [19]. These methods exploit geometric perspective of the linear mixture model and assume that the dataset contains at least one pure pixel of each endmember, *e.g.*, a pixel containing just a single endmember.

VCA has been successfully used to determine endmembers and unmix large hyperspectral data sets without the use of any a priori knowledge of the constituent spectra. Compared with other geometric-based approaches, VCA is an efficient method from the computational point of view. However, most of the above mentioned applications require anal-

This work was supported by FCT project PEst-OE/EEI/LA0008/2011.

ysis methods able to provide a response in (near) real time.

In recent years, high-performance computing systems have become more widespread in remote sensing applications, namely the emergence of programmable graphics processing units (GPUs). Driven by the increasing demands of the videogame industry, GPUs have evolved from expensive application specific units into highly parallel and programmable systems. Specifically several state-of-the-art hyperspectral imaging algorithms can fully benefit from this hardware and take advantage of the extremely high floating-point processing performance, compact size, huge memory bandwidth, and relatively low cost of these units, which make them appealing for onboard data processing [20, 21].

This paper proposes a parallel GPU-based version of VCA that alleviates considerably the computational burden, allowing a significant increase in processing speed. The remainder of the paper is organized as follows. Section 2 presents the fundamentals of the VCA method. Section 3 describes the GPU-based implementation of VCA. Section 4 evaluates the proposed acceleration implementation from the viewpoint of both endmember extraction accuracy and parallel performance. Section 5 concludes with some remarks.

2. VERTEX COMPONENT ANALYSIS

Linear mixing model considers that a mixed pixel is a linear combination of endmember signatures weighted by the correspondent abundance fractions. Thus, each observed spectral vectors (pixel) $\mathbf{y} \in \mathbb{R}^L$ of an hyperspectral image (L is the number of bands) is given by

$$\mathbf{y} = \mathbf{M}\mathbf{s} + \mathbf{n}, \quad (1)$$

where $\mathbf{M} \equiv [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p]$ is a full-rank $L \times p$ mixing matrix (\mathbf{m}_j denotes the j th endmember signature), p is the number of endmembers present in the covered area (with $p < L$), $\mathbf{s} = [s_1, s_2, \dots, s_p]^T$ is the abundance vector containing the fractions of each endmember, and \mathbf{n} is additive noise vector (notation $(\cdot)^T$ stands for vector transposed).

To be physically meaningful [6], abundance fractions are subject to nonnegativity and full additivity constraints, *i.e.*, abundance fractions are in the $p-1$ probability simplex: $\{\mathbf{s} \in \mathbb{R}^p : s_j \geq 0, \sum_{j=1}^p s_j = 1\}$. Thus, under the linear mixing model the observed spectral vectors in a given scene are in a simplex whose vertices correspond to the endmembers. Assuming that each endmember has at least one pure pixel on the dataset, endmember extraction amounts to find those pure pixels.

VCA is an unsupervised method to unmix linear mixtures of hyperspectral sources and is based on the geometry of convex sets. It exploits two facts: 1) the endmembers are the vertices of a simplex and 2) the affine transformation of a simplex is also a simplex. VCA is a fully automatic algorithm and it works with and without dimensionality reduction pre-processing step. The algorithm iteratively projects data onto

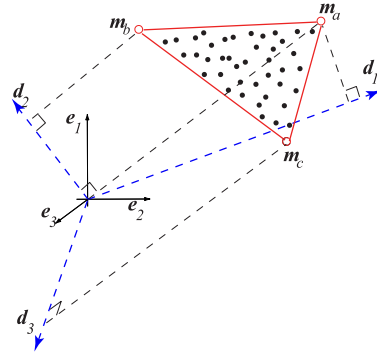


Fig. 1. Hyperspectral mixture of three endmembers illustrating the VCA algorithm.

a direction orthogonal to the subspace spanned by the endmembers already determined. The new endmember signature corresponds to the extreme of the projection. The algorithm iterates until all endmembers are exhausted. Fig. 1 illustrates the VCA method working on a simplex defined by a mixture of three endmembers. In the first iteration, data is projected onto the first direction \mathbf{d}_1 . The extreme of the projection corresponds to endmember \mathbf{m}_a . In the next iteration, endmember \mathbf{m}_b is found by projecting data onto direction \mathbf{d}_2 , which is orthogonal to \mathbf{m}_a . Finally, a new direction \mathbf{d}_3 , orthogonal to the subspace spanned by \mathbf{m}_a and \mathbf{m}_b is generated and the endmember \mathbf{m}_c is found by seeking the extreme of the projection of the dataset onto \mathbf{d}_3 . Full VCA algorithm can be found in [11].

3. GPU-BASED VCA

As mentioned in the previous section VCA iteratively project all pixels to a direction orthogonal to the subspace spanned by the endmembers found so far. At each iteration the pixels projection are independent from each other, thus it is highly desirable that this procedure be parallel implemented in order to increase in processing speed. Figure 2 present the flow chart of VCA. The dataset is read to the memory, where is shared by the GPU cores, the direction determination, which is composed by small matrix operations, is made on the CPU, whereas the pixel projections are made in the GPU in parallel fashion.

4. EXPERIMENTAL RESULTS

The impact on the complexity and the accuracy of the proposed parallel implementation of VCA is evaluated using simulated data based on laboratory spectra available from the United States Geological Survey (USGS)¹ and real hyperspectral data collected by the AVIRIS sensor over the Cuprite

¹<http://speclab.cr.usgs.gov/spectral-lib.html>

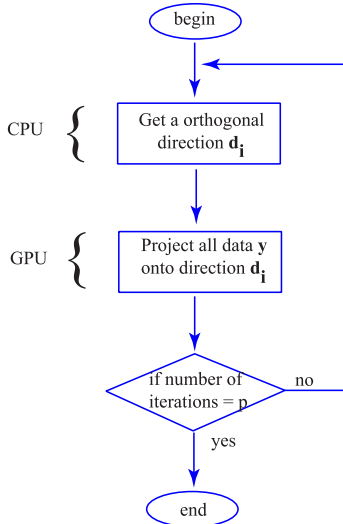


Fig. 2. Parallel VCA flowchart.

mining district in Nevada, a widely used test site for evaluating the accuracy of endmember extraction and spectral unmixing algorithms. The sequential VCA is implemented in C programming language running on a computer platform equipped with a Intel i7-2600 (3.4GHz), with 16 Gbyte memory and the parallel VCA is a CUDA-based implementation (version 4.0) on a GPU card equipped with a GTX-590 from NVidia (with 1.5GB of memory). To evaluate the performance of the algorithm the spectral angle distance (SAD) [6] is used. This widely known metric measures the shape similarity between the i th endmember signature \mathbf{m}_i and its estimate $\hat{\mathbf{m}}_i$. Based on this metric, we define a spectral mean angle error (SMAE), given by:

$$\text{SMAE} \equiv \sqrt{\frac{1}{p} \sum_{i=1}^p \left[\arccos \left(\frac{\mathbf{m}_i^T \hat{\mathbf{m}}_i}{\|\mathbf{m}_i\| \|\hat{\mathbf{m}}_i\|} \right) \right]^2}. \quad (2)$$

4.1. Evaluation with Simulated Data

In this section, we apply the sequential and the parallel VCA to simulated scenes. To evaluate speedup of the parallel versus sequential versions of VCA several datasets are created with different number of pixels. Each pixel is generated according to expression (1), where 5 spectral signatures are selected from the USGS digital spectral library [22]. The reflectances contain 224 spectral bands covering wavelengths from 0.38 to 2.5 μm with a spectral resolution of 10 nm . The abundance fractions are generated according to a Dirichlet distribution which enforces positivity and full additivity constraints (see [23] for details).

Figure 3 illustrates the processing time of sequential and parallel implementations of VCA as a function of the number of pixels of the image. As expected the parallel version is more fast achieving a speedup higher than 15.

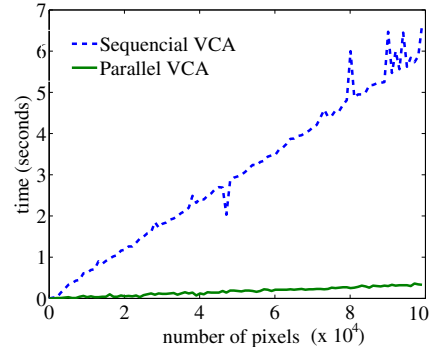


Fig. 3. Processing time of parallel and sequential versions of VCA as a function of the number of pixels.

Table 1. SMAE results for VCA as a function of SNR and of p . ($N = 10^4$)

SNR	SMAE	p	SMAE
$p = 5$		SNR= 30 dB	
20	0.97	3	0.17
30	0.63	5	0.22
50	0.01	10	0.40
∞	0.00	15	0.63

To evaluate the accuracy of both versions of VCA, several scenarios were created, where the number of endmembers is set to 3, 5, 10, and 15, and the SNR defined by $\text{SNR} \equiv 10 \log_{10} \left(\frac{\mathbb{E}[(\mathbf{M}\mathbf{s})^T \mathbf{M}\mathbf{s}]}{\mathbb{E}[\mathbf{n}^T \mathbf{n}]} \right)$, is set to 20, 30, and 50 dB. Table 1 shows the performance measure, SMAE, results as a function of the number of endmembers (right columns) and as a function of the SNR (left columns). As expected the performance of both methods is the same since the orthogonal directions are chosen in the same manner, giving the same projections. Note that the accuracy improves with higher SNR and with smaller number of endmembers.

4.2. Evaluation with Real Data

In this section, the proposed method is applied to real hyperspectral data collected by the AVIRIS sensor. A subset of the Cuprite dataset² containing 350×350 pixels with 187 spectral bands (noisy and water absorption bands were removed) is considered. This site has been extensively used for remote sensing experiments over the past years and its geology was previously mapped in detail [24]. Fig. 4 shows band 30 (wavelength $\lambda = 647.7\text{nm}$) of the subimage of AVIRIS Cuprite Nevada dataset. Table 2 shows the evaluation metric SMAE, for five minerals of interest where the signatures estimated by the different methods are compared with the nearest laboratory spectra. This results show how accurate is VCA method. It is noteworthy to mention that the sequential method spent 6.03 seconds whereas the parallel method spent 0.15 seconds, *i.e.*, a speedup higher than 40.

²Available online at <http://aviris.jpl.nasa.gov/html/aviris.freedata.html>

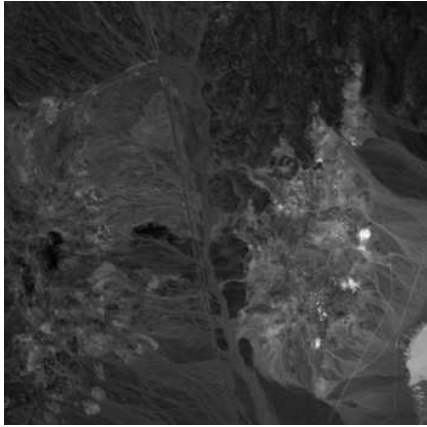


Fig. 4. Band 30 (wavelength $\lambda = 647.7nm$) of the subimage of AVIRIS Cuprite Nevada dataset.

Table 2. SMAE results of VCA for the Cuprite dataset.

Mineral name	VCA (SMAE)
Alunite	5.15
Buddingtonite	5.71
Calcite	7.36
Kaolinite	2.24
Muscovite	4.90

5. CONCLUSIONS

In this paper a parallel implementation of VCA developed for graphics processing units is proposed. The significant speedup reported in experiments will bridge the gap towards real-time spectral unmixing of hyperspectral datasets, which is a highly desirable goal in the remote sensing community. Further research will also include experiments with multi-GPU systems and clusters of GPUs.

6. REFERENCES

- [1] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J.A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Rem. Sens. of Environ.*, vol. 113, pp. 110–122, 2009.
- [2] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," *Proc. of the 2nd IEEE WHISPERS*, pp. 1–4, 2010.
- [3] J. M. Bioucas-Dias and A. Plaza, "Hyperspectral unmixing: geometrical, statistical, and sparse regression-based approaches," *Proc. of SPIE*, vol. 7830, pp. 1–15, 2010.
- [4] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Rem. Sens. of Environ.*, vol. 65, no. 3, pp. 227–248, 1998.
- [5] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Sig. Proc. Mag.*, vol. 19, pp. 12–16, 2002.
- [6] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Sig. Proc. Mag.*, vol. 19, no. 1, pp. 44–57, 2002.
- [7] S.G. Ungar, J.S. Pearlman, J.A. Mendenhall, and D. Reuter, "Overview of the Earth Observing One (EO-1) mission," vol. 41, no. 6-1, pp. 1149–1159, 2003.
- [8] A. Plaza and C.-I Chang, *High Performance Computing in Remote Sensing*, Taylor & Francis: Boca Raton, FL, 2007.
- [9] A. Plaza, J. Plaza, A. Paz, and S. Sanchez, "Parallel hyperspectral image and signal processing," *IEEE Sig. Proc. Mag.*, vol. 28, no. 3, pp. 119–126, 2011.
- [10] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 42, no. 3, pp. 650–663, 2004.
- [11] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 43, no. 4, pp. 898–910, 2005.
- [12] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 40, no. 9, pp. 2025–2041, 2002.
- [13] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," *Proc. JPL Airborne Earth Science Workshop*, pp. 23–26, 1995.
- [14] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Proc. of SPIE*, vol. 3753, pp. 266–277, 1999.
- [15] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," *Proc. 21st Canadian Symp. Rem. Sens.*, pp. 21–24, 1999.
- [16] Chein-I Chang, Chao-Cheng Wu, Wei min Liu, and Yen chieh Ouyang, "A New Growing Method for Simplex-Based Endmember Extraction Algorithm," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 44, no. 10, pp. 2804–2819, 2006.
- [17] John H. Gruninger, Anthony J. Ratkowski, and Michael L. Hoke, "The sequential maximum angle convex cone (smacc) endmember model," *Proc. of SPIE*, vol. 5425, pp. 1–14, 2004.
- [18] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 49, no. 11, pp. 4177–4193, 2011.
- [19] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Applied Optics*, vol. 47, pp. 77–84, 2008.
- [20] C. A. Lee, S. D. Gasster, A. Plaza, C.-I Chang, and B. Huang, "Recent developments in high performance computing for remote sensing: A review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 3, pp. 508–527, 2011.
- [21] A. Plaza, Q. Du, Y.-L.-Chang, and R. L. King, "High performance computing for hyperspectral remote sensing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 3, pp. 528–544, 2011.
- [22] R. N. Clark, G. A. Swayze, A. Gallagher, T. V. King, and W. M. Calvin, "The U.S. Geological Survey Digital Spectral Library: Version 1: 0.2 to 3.0 μm ," Open file report 93-592, U.S. Geological Survey, 1993.
- [23] J. M. P. Nascimento and J. M. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of dirichlet components," *IEEE Trans. on Geosc. and Rem. Sens.*, vol. 50, no. 3, pp. 863–878, 2012.
- [24] G.A. Swayze, R.N. Clark, S. Sutley, and A. Gallagher, "Ground-Truthing AVIRIS Mineral Mapping at Cuprite, Nevada," in *Summaries of the Third Annual JPL Airborne Geosciences Workshop*, 1992, pp. 47–49.