A New N-FINDR Algorithm and the unmixR *Package for Spectral Unmixing*

C. Beleites[1], C. McManus[2], S. Fuller[3], B. Hanson[4], C. Krafft[1], J. Popp[1,5]

[1]Leibniz Institute of Photonic Technology Jena, Jena, Germany
[2]National University of Ireland Maynooth, Maynooth, Ireland
[3]Computer Science Dept., National University of Ireland Maynooth, Maynooth, Ireland
[4]Institute of Chemistry & Biochemistry, DePauw University, Greencastle IN, USA
[5]Institute of Physical Chemistry and Abbe Center of Photonics, University Jena, Germany

In spectral unmixing, a data set $\boldsymbol{X}$ composed of n observed spectra of chemical mixtures with p wavelengths or spectral bands is decomposed to identify pure component spectra. Each spectrum is assumed to be a linear mixture of a limited number m of pure component spectra, in this context also referred to as endmembers. m is also known as chemical rank of the spectra matrix $\boldsymbol{X}$. In matrix notation, this is the bilinear model

$$\boldsymbol{X}^{(n \times p)} = \boldsymbol{A}^{(n \times m)} \boldsymbol{E}^{(m \times p)} + \varepsilon.$$

The concentrations or abundances $\boldsymbol{A}$ can be depicted in the mixture diagram of the m components, forming an (m - 1) simplex. Also the spectra $\boldsymbol{X}$ lie in an (m – 1) simplex. If noise is low and pure component spectra are present in the data, this decomposition can be obtained by finding the corners of the simplex. Two well-known algorithms to achieve this are N-FINDR [1] and Vertex Component Analysis [2]. N-FINDR is a rather slow iterative algorithm. While VCA is much faster, we find that it does not always yield as good decompositions. A number of improved algorithms are known [3, 4]. In addition, we present a new N-FINDR algorithm that uses a fast projection and is particularly suitable for implementation in high-level languages such as R or Matlab where optimized and parallelized BLAS routines can be used but loops are computationally expensive.

unmixR (http://github.com/Chathurga/unmixR) provides different N-FINDR and VCA algorithms as an R package. We demonstrate the correctness of the algorithms by reproducing the results of [5].

Acknowledgements: CM implemented algorithms [1-4] during Google Summer of Code 2013, supervised by CB, SF and BH. CB now maintains the package. The new algorithm was developed by CB, who is funded via the BMBF project RamanCTC (13N12685).

References

[1] M.E. Winter et al.. N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. *ProcSPIE* 3753, 266-275 (1999).
[2] J.M.P. Nascimento and J.M.B. Dias. Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data. *IEEE Trans. Geosci. Rem. Sens* 43, 898-910 (2005).
[3] Dowler et al.. Reducing the complexity of the N-FINDR algorithm for hyperspectral image analysis. *IEEE Trans Image Process* 22, 2835-2848 (2013).
[4] Lopez et al.. A Low-Computational-Complexity Algorithm for Hyperspectral Endmember Extraction: Modified Vertex Component Analysis. *IEEE Geosci. Rem. Sens. Letters* 9, 502-506 (2012).
[5] Hedegaard et al.. Spectral unmixing and clustering algorithms for assessment of single cells by Raman microscopic imaging. *Theor Chem Acc, Springer* 130, 1249-1260 (2011).