

Blind multispectral image decomposition by 3D nonnegative tensor factorization

Ivica Kopriva^{1,*} and Andrzej Cichocki²

¹Division of Laser and Atomic Research and Development, Ruđer Bošković Institute, Bijenička cesta 54, HR-10000, Zagreb, Croatia

²Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN 2-1Hirosawa, Wako-shi, Saitama 351-0198, Japan

*Corresponding author: ikopriva@gmail.com

Received March 26, 2009; revised June 19, 2009; accepted June 21, 2009;
posted June 25, 2009 (Doc. ID 109250); published July 14, 2009

α -divergence-based nonnegative tensor factorization (NTF) is applied to blind multispectral image (MSI) decomposition. The matrix of spectral profiles and the matrix of spatial distributions of the materials resident in the image are identified from the factors in Tucker3 and PARAFAC models. NTF preserves local structure in the MSI that is lost as a result of vectorization of the image when nonnegative matrix factorization (NMF)- or independent component analysis (ICA)-based decompositions are used. Moreover, NTF based on the PARAFAC model is unique up to permutation and scale under mild conditions. To achieve this, NMF- and ICA-based factorizations, respectively, require enforcement of sparseness (orthogonality) and statistical independence constraints on the spatial distributions of the materials resident in the MSI, and these conditions do not hold. We demonstrate efficiency of the NTF-based factorization in relation to NMF- and ICA-based factorizations on blind decomposition of the experimental MSI with the known ground truth. © 2009 Optical Society of America

OCIS codes: 100.6890, 100.3190, 150.6910, 100.2960, 170.3880.

Blind or unsupervised multispectral (MSI) and hyperspectral image (HSI) decomposition attract increased attention because of their capability to discriminate materials resident in the MSI/HSI without knowing their spectral profiles [1,2]. However, most blind decomposition schemes rely on two-dimensional (2D) representation of the MSI/HSI, although it is inherently three-dimensional (3D). In this Letter we represent a MSI/HSI as a three-way array or a 3D tensor $\mathbf{X} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times I_3}$ with elements $x_{i_1 i_2 i_3}$ where $i_1 = 1, \dots, I_1$; $i_2 = 1, \dots, I_2$; $i_3 = 1, \dots, I_3$; and \mathbb{R}_{0+} is a real manifold with nonnegative elements. Each index is called a way or mode, and the number of levels on one mode is called the dimension of that mode. An MSI/HSI is a set of I_3 spectral band images with the size of $I_1 \times I_2$ pixels. Two ways of \mathbf{X} are for rows and columns, and one way is for the spectral band. This is standard notation that is adopted for use in multiway analysis [3].

2D representation of MSI has two disadvantages: (i) the 3D tensor \mathbf{X} has to be mapped through three-mode flattening, also called unfolding and matricization, to matrix $\mathbf{X}_{(3)} \in \mathbb{R}_{0+}^{I_3 \times I_1 I_2}$, whereupon the local structure of the image is lost; and (ii) matrix factorization $\mathbf{X}_{(3)} = \mathbf{A}\mathbf{S}$ employed by linear mixing models [1,2] suffers from indeterminacies because $\mathbf{A}\mathbf{T}\mathbf{T}^{-1}\mathbf{S} = \mathbf{X}_{(3)}$ for any invertible \mathbf{T} ; i.e., infinitely many (\mathbf{A}, \mathbf{S}) pairs can give rise to $\mathbf{X}_{(3)}$.

Meaningful solution of the factorization of $\mathbf{X}_{(3)}$ is characterized by $\mathbf{T}\mathbf{T}^{-1} = \mathbf{P}\mathbf{\Lambda}$, where \mathbf{P} is a permutation matrix and $\mathbf{\Lambda}$ is a diagonal matrix. These permutation and scaling indeterminacies are standard for blind decompositions and are obtained by imposing sparseness (orthogonality) constraints on \mathbf{S} by nonnegative matrix factorization (NMF) algorithms [4] and statistical independence constraints by indepen-

dent component analysis (ICA) algorithms [1,5,6]. Orthogonality constraints imply that materials resident in the image do not occupy the same pixel footprint, but that is not a correct assumption, especially in airborne and spaceborne remote sensing. The statistical independence assumption is also not correct for MSI and HSI data, especially when materials are spectrally similar, which occurs in the case of low-dimensional MSI with coarse spectral resolution [7].

Only very recently tensor factorization methods were employed in MSI/HSI analysis for the purpose of dimensionality reduction, denoising, target detection, and material identification [8–10]. For the purpose of MSI decomposition we adopt two widely used 3D tensor models: the Tucker3 model [11] and the PARAFAC/CANDECOMP model [12,13]. The Tucker3 model is defined as

$$\mathbf{X} \approx \mathbf{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \times_3 \mathbf{A}^{(3)}, \quad (1)$$

where $\mathbf{G} \in \mathbb{R}_{0+}^{J_1 \times J_2 \times J_3}$ is a core tensor, $\{\mathbf{A}^{(n)} \in \mathbb{R}_{0+}^{I_n \times J_n}\}_{n=1}^3$ are factors, and \times_n denotes the n -mode product of a tensor and a matrix $\mathbf{A}^{(n)}$. The result of $\mathbf{G} \times_n \mathbf{A}^{(n)}$ is a tensor of the same order as \mathbf{G} but with the size J_n replaced by I_n . The PARAFAC model is a special case of the Tucker3 model when \mathbf{G} is a super-diagonal tensor with all elements zero except those for which all indices are the same. Compared with PARAFAC, the Tucker3 model is more flexible owing to the core tensor \mathbf{G} , which allows interaction between one factor and any factor in the other modes [14]. In the PARAFAC model factors in different modes can interact only factorwise. However, this restriction enables uniqueness of tensor factorization based on the PARAFAC model within the permutation and scaling indeterminacies of the factors under very mild conditions [15,16] without the need to im-

pose any special constraints on them such as sparseness or statistical independence.

Assuming that $J_1=J_2=J_3=J$ and $J \leq I_3$ the uniqueness condition is reduced to $k_{\mathbf{A}^{(1)}}+k_{\mathbf{A}^{(2)}}+k_{\mathbf{A}^{(3)}} \geq 2J+3$, where $k_{\mathbf{A}^{(n)}}$ is the Kruskal rank of factor $\mathbf{A}^{(n)}$ [15,16]. Because of interaction between the factors there is no such theoretical guarantee on the uniqueness of tensor factorization based on the Tucker3 model. However despite this, the Tucker3 model has been used successfully in HSI analysis for dimensionality reduction, denoising and target detection [8,9]. To identify spatial distributions of the materials resident in the MSI/HSI we refer to the standard linear mixture model used in MSI/HSI data analysis [1,2,10]:

$$\mathbf{X}_{(3)} \approx \mathbf{A}\mathbf{S}, \quad (2)$$

where the columns of $\mathbf{A} \in \mathbb{R}_{0+}^{I_3 \times J}$ represent spectral profiles of the J materials resident in the image, while the rows of $\mathbf{S} \in \mathbb{R}_{0+}^{J \times I_1 I_2}$ represent spatial distributions of the same materials. As already stated, without additional constraints there are infinitely many decompositions satisfying model (2). From the Tucker3 model (1) and the linear mixture model (2), the matrix of spectral profiles and tensor of spatial distributions of the materials are identified as

$$\mathbf{A} \approx \mathbf{A}^{(3)},$$

$$\mathbf{S} \approx \mathbf{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} = \mathbf{X} \times_3 (\mathbf{A}^{(3)})^{-1}, \quad (3)$$

where $\mathbf{S} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times J}$. The second approximation for \mathbf{S} in relation (3) is less sensitive to numerical errors than the first one because only one reconstructed quantity, array factor $\mathbf{A}^{(3)}$, is involved in the reconstruction of \mathbf{S} . We can also express the three-mode flattened version of tensor \mathbf{X} —this is matrix $\mathbf{X}_{(3)}$ —in terms of the three-mode flattened core tensor \mathbf{G} ; this yields matrix $\mathbf{G}_{(3)} \in \mathbb{R}_{0+}^{J \times J \times J}$ and array factors $\{\mathbf{A}^{(n)}\}_{n=1}^3$ as [17]

$$\mathbf{X}_{(3)} \approx \mathbf{A}^{(3)} \mathbf{G}_{(3)} [\mathbf{A}^{(2)} \otimes \mathbf{A}^{(1)}]^T, \quad (4)$$

where \otimes denotes Kronecker's product. In direct comparison between relations (2) and (4) we arrive at

$$\mathbf{A} \approx \mathbf{A}^{(3)},$$

$$\mathbf{S} \approx \mathbf{G}_{(3)} [\mathbf{A}^{(2)} \otimes \mathbf{A}^{(1)}]^T = (\mathbf{A}^{(3)})^{-1} \mathbf{X}_{(3)}. \quad (5)$$

Again, a numerically more accurate approximation of \mathbf{S} is obtained from the second part of relation (5).

Various cost functions can be used as discrepancy measures between a tensor and its model. In this Letter we employ α -divergence because it is adaptable to noise statistics [17] and because it has been demonstrated in [18] that it outperforms NTF based on least-square error function [19]. We refer to Appendix A and to [17] for α -divergence-based update NTF algorithms [Eqs. (13)–(15)].

We have compared the performance of α -NTF, second-order (SO) NMF [20], and dependent component analysis (DCA) algorithms on a real world example: blind decomposition of the experimental

weak-intensity fluorescent red–green–blue (RGB) image of the skin tumor shown in Fig. 1(b). For the purpose of tumor demarkation it is of interest to extract a spatial map of the tumor as accurately as possible. The DCA algorithm combines the joint approximate diagonalization of eigenmatrices ICA algorithm [6] and innovation transform-based preprocessing [21] to enhance statistical independence among materials present in the MSI and improve the accuracy of the ICA. The MSI contains three spectral bands and three materials: tumor, surrounding tissue, and the ruler added to the scene to give perspective about the size of the tumor. A high-intensity image of the same tumor is shown in Fig. 1(a). It was used for estimation of the binary spatial maps of tumor and surrounding tissue necessary for the estimation of the receiver-operating-characteristic (ROC) curves. Spatial maps of the tumor extracted from Fig. 1(b) are shown in Figs. 1(c)–1(e). Implementation of the Tucker3 α -NTF algorithm was based on MATLAB Tensor Toolbox provided in [22]. The tensor of spatial distributions of materials was identified by means of the second part of relation (3). SO NMF and DCA algorithms were based on the 2D the MSI representation of relation (2).

According to the ROC curves shown in Fig. 2 the α -NTF algorithm exhibited best performance, i.e., yielded largest area under the ROC curve. A

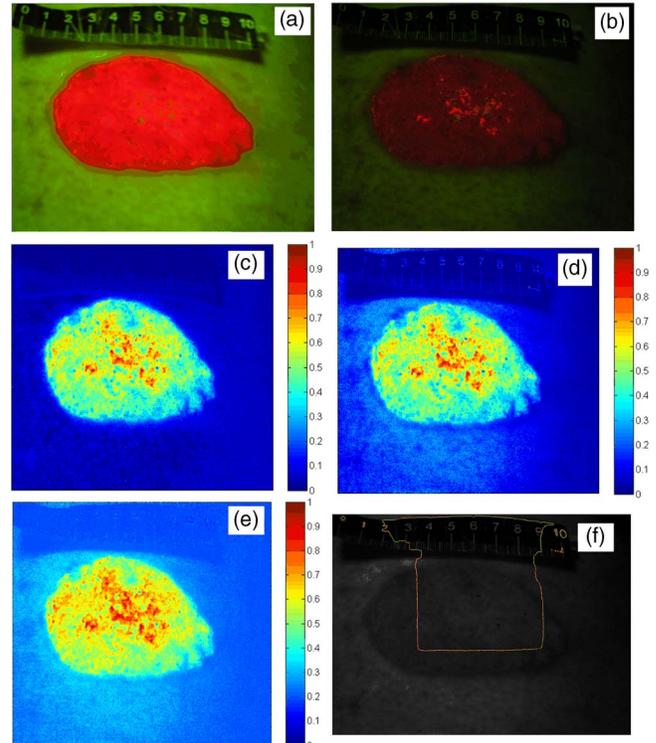


Fig. 1. (Color online) Experimental fluorescent MSI (RGB) image of skin tumor: (a) high-intensity version, (b) low-intensity version. Spatial maps of the tumor extracted from Fig. 1(b) by means of (c) α -NTF algorithm [18] with $\alpha=0.1$, (d) SO NMF algorithm [5], (e) DCA algorithm [7,21]. (f) Evolution curve calculated by level set method on Fig. 1(b) after 1000 iterations. Dark red color indicates that tumor is present with probability 1, while dark blue color indicates that tumor is present with probability 0.

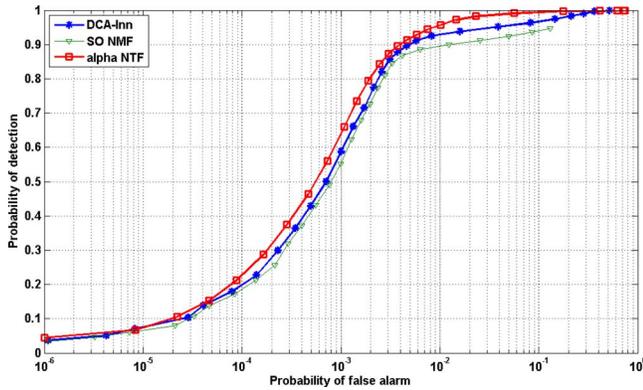


Fig. 2. (Color online) ROC curves calculated for spatial maps of the tumor shown in Figs. 1(c)–1(e): open squares, α -NTF algorithm based on Tucker3 model with $\alpha=0.1$; asterisks, DCA algorithm [6,21]; open triangles, SO NMF algorithm.

state-of-the-art intensity-based image segmentation algorithm—that is, the level set method [23]—has also been applied to Fig. 1(b). The result in Fig. 1(f) shows the evolution curve after 1000 iterations. Because of weak boundaries the method failed to converge.

α -NTF, SO NMF, and DCA algorithms were implemented in MATLAB on a 2.4 GHz Intel Core 2 Quad Processor Q6600 based desktop computer with 4 GB RAM. Computation times are given, respectively, as 4783 s, 30 s, and 3.6 s. In the implementation of the innovation-based DCA algorithm a tenth-order linear prediction filter was used.

Appendix: Elements of Tucker3 α -NTF Algorithm

Multiplicative update rules for core tensor \mathbf{G} and factors $\{\mathbf{A}^{(n)}\}_{n=1}^3$ in relation (1) are obtained by minimizing α -divergence as

$$\mathbf{G} \leftarrow \mathbf{G} \odot \left\{ \frac{(\hat{\mathbf{X}}/\hat{\mathbf{X}})^{\cdot\alpha} \times_1 \mathbf{A}^{(1)\text{T}} \times_2 \mathbf{A}^{(2)\text{T}} \times_3 \mathbf{A}^{(3)\text{T}}}{\mathbf{E} \times_1 \mathbf{A}^{(1)\text{T}} \times_2 \mathbf{A}^{(2)\text{T}} \times_3 \mathbf{A}^{(3)\text{T}}} \right\}^{.1/\alpha}, \quad (\text{A1})$$

$$\mathbf{A}^{(n)} \leftarrow \mathbf{A}^{(n)} \odot \left\{ \frac{[(\hat{\mathbf{X}}/\hat{\mathbf{X}})^{\cdot\alpha}]_{(n)} \mathbf{G}_{\mathbf{A}}^{(n)\text{T}}}{\mathbf{1}^{\text{T}} \mathbf{G}_{\mathbf{A}}^{(n)\text{T}}} \right\}^{.1/\alpha}, \quad (\text{A2})$$

where \odot denotes element-wise multiplication and $/$ denotes element-wise division. In relation (A1) \mathbf{E} is a tensor with all elements equal to 1. In relation (A2) $\mathbf{1}$ denotes a vector whose every element is 1. The numerator in (A2) is calculated as

$$[(\hat{\mathbf{X}}/\hat{\mathbf{X}})^{\cdot\alpha}]_{(n)} \mathbf{G}_{\mathbf{A}}^{(n)\text{T}} = [(\hat{\mathbf{X}}/\hat{\mathbf{X}})^{\cdot\alpha} \times_{m \neq n} \mathbf{A}^{(m)\text{T}}]_n \mathbf{G}_{(n)}^{\text{T}},$$

where $\mathbf{G}_{(n)}$ represents the n -mode flattened version of the core tensor \mathbf{G} . The denominator in (A2) is computed as

$$\mathbf{1}^{\text{T}} \mathbf{G}_{\mathbf{A}}^{(n)\text{T}} = [\mathbf{G} \times_{m \neq n} \mathbf{1}^{\text{T}} \mathbf{A}^{(m)}]_{(n)}^{\text{T}},$$

where $\mathbf{G} \times_{m \neq n} \mathbf{1}^{\text{T}} \mathbf{A}^{(m)}$ denotes m -mode products between core tensor \mathbf{G} and matrices $\mathbf{1}^{\text{T}} \mathbf{A}^{(m)}$ for all $m=1, \dots, N$ and $m \neq n$.

This work was partially supported through grant 098-0982903-2558 funded by the Ministry of Science, Education and Sports, Republic of Croatia.

References

1. C.-I. Chang, S.-S. Chiang, J. A. Smith, and I. W. Ginsberg, *IEEE Trans. Geosci. Remote Sens.* **40**, 375 (2002).
2. T. Tu, *Opt. Eng. (Bellingham)* **39**, 897 (2000).
3. H. A. L. Kiers, *J. Chemom.* **14**, 105 (2000).
4. A. Cichocki, R. Zdunek, and S. Amari, *IEEE Signal Process. Mag.* **25**, 142 (2008).
5. A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing* (Wiley, 2002).
6. J. F. Cardoso and A. Soulomniac, *Proc. IEE F* **140**, 362 (1993).
7. J. M. P. Nascimento and J. M. Dias Bioucas, *IEEE Trans. Geosci. Remote Sens.* **43**, 175 (2005).
8. N. Renard, S. Bourennane, and J. Blanc-Talon, *IEEE Geosci. Remote Sens. Lett.* **5**, 138 (2008).
9. N. Renard and S. Bourennane, *IEEE Trans. Geosci. Remote Sens.* **46**, 2407 (2008).
10. Q. Zhang, H. Wang, R. J. Plemons, and V. P. Pauca, *J. Opt. Soc. Am. A* **25**, 3001 (2008).
11. L. R. Tucker, *Psychometrika* **31**, 279 (1966).
12. J. D. Carroll and J. J. Chang, *Psychometrika* **35**, 283 (1970).
13. R. A. Harshman, *UCLA Working Papers in Phonetics* **16**, 1 (1970).
14. E. Acar and B. Yener, *IEEE Trans. Knowl. Data Eng.* **21**, 6 (2009).
15. J. B. Kruskal, *Linear Algebr. Appl.* **18**, 95 (1977).
16. N. D. Sidiropoulos and R. Bro, *J. Chemom.* **14**, 229 (2000).
17. Y. D. Kim, A. Cichocki, and S. Choi, presented at the *2008 IEEE ICASSP International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, Nevada, March 30–April 4, 2008, pp. 1829–1832.
18. A. H. Phan and A. Cichocki, *Lect. Notes Comput. Sci.* **5264**, 772 (2008).
19. C. A. Anderson and R. Bro, *Chemom. Intell. Lab. Syst.* **52**, 1 (2000).
20. R. Zdunek and A. Cichocki, *Signal Process.* **87**, 1904 (2007).
21. A. Hyvärinen, in *Proceedings of the International Conference on Artificial Neural Networks (ICANN'98)* (IEEE, 1998), pp. 541–546.
22. B. W. Bader and T. G. Kolda, *MATLAB Tensor Toolbox version 2.2.*, <http://csmr.ca.sandia.gov/~tkolda/TensorToolbox>.
23. Ch. Li, Ch. Xu, Ch. Gui, and M. D. Fox, in *IEEE International Conference on Computer Vision and Pattern Recognition* (IEEE, 2005), pp. 430–436.