Blind Image Separation Using Nonnegative Matrix Factorization with Gibbs Smoothing

Rafal Zdunek* and Andrzej Cichocki**

RIKEN Brain Science Institute, Wako-shi, Saitama, Japan
zdunek@brain.riken.jp

Abstract. Nonnegative Matrix Factorization (NMF) has already found many applications in image processing and data analysis, including classification, clustering, feature extraction, pattern recognition, and blind image separation. In the paper, we extend the selected NMF algorithms by taking into account local smoothness properties of source images. Our modifications are related with incorporation of the Gibbs prior, which is well-known in many tomographic image reconstruction applications, to a underlying blind image separation model. The numerical results demonstrate the improved performance of the proposed methods in comparison to the standard NMF algorithms.

1 Introduction

Nonnegative Matrix Factorization (NMF) [1] attempts to recover hidden nonnegative structures or patterns from usually redundant data. This technique has been successfully applied in many applications, e.g. in data analysis (pattern recognition, segmentation, clustering, dimensionality reduction) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], signal and image processing (blind source separation, spectra recovering) [13, 14], language modeling, text analysis [15, 16], music transcription [4, 17], or neuro-biology (gene separation, EEG signal analysis) [18, 19, 20].

NMF decomposes the data matrix $\boldsymbol{Y} = [y_{ik}] \in \mathbb{R}^{I \times K}$ as a product of two nonnegative matrices $\boldsymbol{A} = [a_{ij}] \in \mathbb{R}^{I \times J}$ and $\boldsymbol{X} = [x_{jk}] \in \mathbb{R}^{J \times K}$, i.e.

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}, \tag{1}$$

where $\forall i, j, k : a_{ij} \geq 0,\ x_{jk} \geq 0$.

Depending on an application, the hidden components may have different interpretation. For example, Lee and Seung in [1] introduced NMF as a method to decompose an image (face) into parts-based representations (parts reminiscent of features such as lips, eyes, nose, etc.). In NMF with application to Blind Source Separation (BSS) [21], the matrix $\boldsymbol{Y}$ represents the observed mixed (superposed)

* Dr. R. Zdunek is also with Institute of Telecommunications, Teleinformatics and Acoustics, Wroclaw University of Technology, Poland.

** Dr. A. Cichocki is also with Systems Research Institute (SRI), Polish Academy of Science (PAN), Warsaw University of Technology, Dept. of EE, Warsaw, Poland.

M. Ishikawa et al. (Eds.): ICONIP 2007, Part II, LNCS 4985, pp. 519–528, 2008.

images, $\boldsymbol{A}$ is a mixing operator, and $\boldsymbol{X}$ is a matrix of true source images. Each row of $\boldsymbol{Y}$ or $\boldsymbol{X}$ is a 1D image representation, where I is a number of observed mixed images and J is a number of hidden (source) components. The 1D representation of a 2D image $\tilde{\boldsymbol{X}} = [\tilde{x}_{mn}] \in \mathbb{R}^{M \times N}$ is obtained as lexicographical ordering of the pixels, i.e. $\tilde{\boldsymbol{x}} = [\tilde{x}_{11}, \tilde{x}_{12}, \ldots, \tilde{x}_{1N}, \tilde{x}_{21}, \ldots, \tilde{x}_{MN}]^T \in \mathbb{R}^{MN}$. The index k denotes the pixel's position in a 1D image representation, and K is a total number of pixels. In BSS, we usually have $K >> I \geq J$, and J is known or can be relatively easily estimated using SVD.

Our objective is to estimate the mixing matrix $\boldsymbol{A}$ and sources $\boldsymbol{X}$ subject to nonnegativity constraints of all the entries, given $\boldsymbol{Y}$ and possibly the prior knowledge on the nature of the true images to be estimated or on a statistical distribution of noisy disturbances.

The basic approach to NMF is the alternating minimization of the specific cost function $D(\boldsymbol{Y}||\boldsymbol{AX})$ that measures the distance between $\boldsymbol{Y}$ and $\boldsymbol{AX}$. Lee and Seung [1] were the first who proposed two types of NMF algorithms. One minimizes the Euclidean distance, which is optimal for a Gaussian distributed additive noise, and the other for minimization of the Kullback-Leibler divergence, which is suitable for a Poisson distributed noise. The NMF algorithms that are optimal for many other distribution of additive noise can be found, e.g. in [22, 21, 23].

Unfortunately, the alternating minimization does not provide a unique solution, and often some additional constraints must be imposed to select a solution that is close to the true one. For example, finding such $\boldsymbol{P} > 0$ for which $\boldsymbol{P}^{-1} > 0$, we have: $\boldsymbol{AX} = (\boldsymbol{AP}^{-1})(\boldsymbol{PX}) = \tilde{\boldsymbol{A}}\tilde{\boldsymbol{X}} = \boldsymbol{Y}$, where $\boldsymbol{A} \neq \tilde{\boldsymbol{A}}$ and $\boldsymbol{X} \neq \tilde{\boldsymbol{X}}$. Obviously, $\boldsymbol{P}$ could be any permutation matrix. Also, the alternating minimization is not convex with respect to both sets of the arguments $\{\boldsymbol{A}, \boldsymbol{X}\}$, even though the cost function is expressed by a quadratic function. To relax the ambiguity and non-convexity effects, the common approach is to incorporate some penalty terms to the cost function, which adequately regularizes the solution or restricts a set of all admissible solutions. Such regularization has been widely discussed in the literature with respect to various criteria for selection of the desired solution. The penalty terms can enforce sparsity, smoothness, continuity, closure, unimodality, orthogonality, or local rank-selectivity. A widely-used approach in many NMF applications is to apply sparsity constraints [24, 22, 25, 26, 27].

In the paper, we apply the penalty term that enforces local smoothness in the estimated 2D images. This case may take place in many BSS applications with locally smooth features. This paper is motivated by the preliminary results obtained in [28], where we have proposed the NMF algorithm for blind separation of locally smooth nonnegative signals.

The penalty term, which we use in the paper, is motivated by the Markov Random Field (MRF) models that are widely applied in image reconstruction. Such models, which are often expressed by the Gibbs prior, determine local roughness (smoothness) in the analyzed image with consideration of pair-wise interactions among adjacent pixels in a given neighborhood of a singe pixel. Thus, a total smoothness in an image can be expressed by a joint Gibbs distribution with

a nonlinear energy function. In our approach, we use the Green's function for measuring strength of the pair-wise pixel interactions. Using a Bayesian framework, we get the Gibbs regularized Euclidean cost function that is minimized with a gradient descent alternating minimization technique subject to nonnegativity constrains that can be imposed in many ways. One of them is achieved with standard multiplicative updates that were used, e.g. by Lee and Seung [1]. Another approach is to apply the projected Alternating Least Squares (ALS) algorithms [27], which are generally more efficient to NMF problems than standard multiplicative algorithms.

2 Gibbs Regularized Algorithms

Since in practice a Gaussian noise occurs the most often in BSS applications, we restrict our considerations only to the following joint multivariate normal likelihood model:

$$p(\boldsymbol{Y}|\boldsymbol{X}) \propto \exp\left\{-\frac{1}{2}\operatorname{tr}\{(\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X})\}\right\}, \tag{2}$$

where each sample $\boldsymbol{n}_k$ from the residual (noise) matrix $\boldsymbol{N} = \boldsymbol{Y} - \boldsymbol{A}\boldsymbol{X} = [\boldsymbol{n}_1, \ldots, \boldsymbol{n}_K]$ is assumed to follow the same statistics with the covariance matrix $\boldsymbol{\Sigma}$.

Let us assume the prior information on total smoothness of the estimated images is given by the following Gibbs distribution

$$p(\boldsymbol{X}) = \frac{1}{Z}\exp\left\{-\alpha U(\boldsymbol{X})\right\}, \tag{3}$$

where Z is a partition function, α is a regularization parameter, and $U(\boldsymbol{X})$ is a total energy function that measures the total roughness in the object of interest. The function $U(\boldsymbol{X})$ is often formulated with respect to the Markov Random Field (MRF) model that is commonly used in image reconstruction to enforce local smoothing.

The prior can be incorporated into the likelihood function with the Bayesian framework:

$$p(\boldsymbol{X}|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\boldsymbol{X})p(\boldsymbol{X})}{p(\boldsymbol{Y})}, \tag{4}$$

where $p(\boldsymbol{Y})$ is a marginal likelihood function. Thus the Gibbs regularized Euclidean cost function can be expressed in the form:

$$\Psi = -2\ln p(\boldsymbol{X}|\boldsymbol{Y}) = ||\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}||_F^2 + 2\alpha U(\boldsymbol{X}) + c, \tag{5}$$

where c is a constant.

The stationary points of Ψ can be derived from the gradients of Ψ with respect to $\boldsymbol{X}$ and $\boldsymbol{A}$. Thus:

$$\nabla_X\Psi = 2\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}-\boldsymbol{Y}) + 2\alpha\nabla_X U(\boldsymbol{X}) \equiv 0, \tag{6}$$

$$\nabla_A\Psi = (\boldsymbol{A}\boldsymbol{X}-\boldsymbol{Y})\boldsymbol{X}^T \equiv 0. \tag{7}$$

2.1 NMF Algorithms

From (6)–(7), we have:

$$\frac{[\boldsymbol{A}^T\boldsymbol{Y} - \alpha\nabla_X U(\boldsymbol{X})]_{jk}}{[\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{X}]_{jk}} = 1, \quad \frac{[\boldsymbol{Y}\boldsymbol{X}^T]_{ij}}{[\boldsymbol{A}\boldsymbol{X}\boldsymbol{X}^T]_{ij}} = 1. \tag{8}$$

Using multiplicative updates, we get the Gibbs regularized multiplicative NMF algorithm:

$$x_{jk} \leftarrow x_{jk}\, \frac{\left[[\boldsymbol{A}^T\boldsymbol{Y}]_{jk} - \alpha[\nabla_X U(\boldsymbol{X})]_{jk}\right]_\varepsilon}{[\boldsymbol{A}^T\boldsymbol{A}\,\boldsymbol{X}]_{jk}}, \tag{9}$$

$$a_{ij} \leftarrow a_{ij}\, \frac{[\boldsymbol{Y}\,\boldsymbol{X}^T]_{ij}}{[\boldsymbol{A}\,\boldsymbol{X}\,\boldsymbol{X}^T]_{ij}}, \quad a_{ij} \leftarrow \frac{a_{ij}}{\sum_{j=1}^{J} a_{ij}}, \tag{10}$$

where $[x]_\varepsilon = \max\{\varepsilon, x\}$ is a nonlinear operator for projection onto a positive orthant (subspace $\mathbb{R}_+$) with small ε (*eps*). Typically, $\varepsilon = 10^{-16}$. The normalization in (10) additionally constrains the basis vectors to a unit l_1-norm, which relaxes the intrinsic scaling ambiguity in NMF.

It is easy to notice that for $\alpha = 0$ in (9), the updating rules (9)–(10) simplify to the standard Lee-Seung algorithm that minimizes the Euclidean distance (Frobenius norm).

The algorithm (9)–(10) can also be improved by replacing the step (10) with a more exact updating rule. It is well-known that multiplicative algorithms are slowly-convergent, and the system of linear equations to be solved in the step (10) is highly over-determined. Hence, the update (10) can be successfully replaced with the projected Moore-Penrose pseudo-inverse [27] or the quasi-Newton approach [26]. For simplicity, we consider only the former approach, thus from (7) we have

$$\boldsymbol{A} \leftarrow \left[\boldsymbol{Y}\boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{X}^T)\right]_\varepsilon. \tag{11}$$

2.2 Markov Random Field Model

MRF models have been widely applied in many image reconstruction applications, especially in tomographic imaging. In our application, MRF models motives the definition of the total energy function in the Gibbs prior (3). Thus

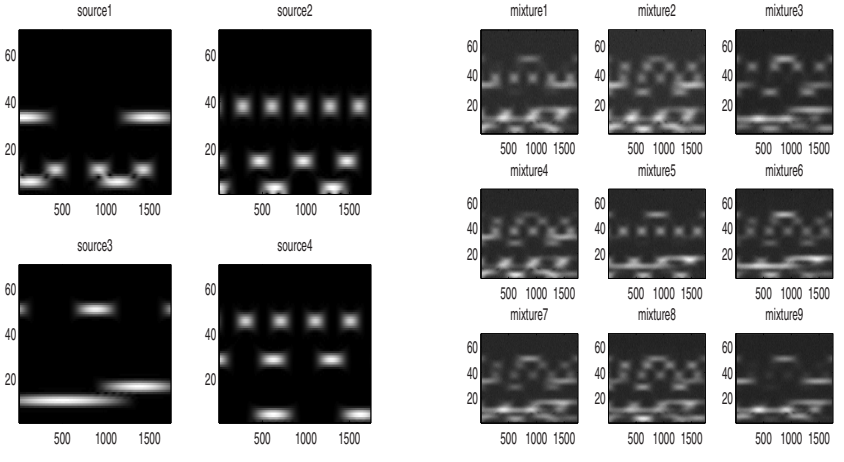
$$U(\boldsymbol{X}) = \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l\in S_k} w_{kl}\psi\left(x_{jk} - x_{jl}, \delta\right), \tag{12}$$

where S_k is a set of pixels in the neighborhood of the k-th pixel, w_{kl} is a weighting factor, δ is a scaling factor, and $\psi\left(\xi, \delta\right)$ is some potential function of ξ, which can take various forms. Exemplary potential functions are listed in Table 1.

Since the Green's function [34] satisfies all the properties mentioned in [35], i.e. it is nonnegative, even, 0 at $\xi = 0$, strictly increasing for $\xi > 0$, unbounded,

Table 1. Potential functions

Author(s) (Name)	Reference	Functions: $V(\xi, \delta)$
(Gaussian)		$\left(\frac{\xi}{\delta}\right)^2$
Besag (Laplacian)	[29]	$\left\|\frac{\xi}{\delta}\right\|$
Hebert and Leahy	[30]	$\delta \log\left[1 + (\frac{\xi}{\delta})^2\right]$
Geman and McClure	[31]	$\frac{16}{3\sqrt{3}} \frac{(\xi/\delta)^2}{(1 + (\xi/\delta)^2)}$
Geman and Reynolds	[32]	$\frac{\lvert\xi/\delta\rvert}{1 + \lvert\xi/\delta\rvert}$
Stevenson and Delp (Hubert)	[33]	$\min\{\lvert\frac{\xi}{\delta}\rvert^2, 2\lvert\frac{\xi}{\delta}\rvert - 1\}$
Green	[34]	$\delta \log[\cosh(\xi/\delta)]$

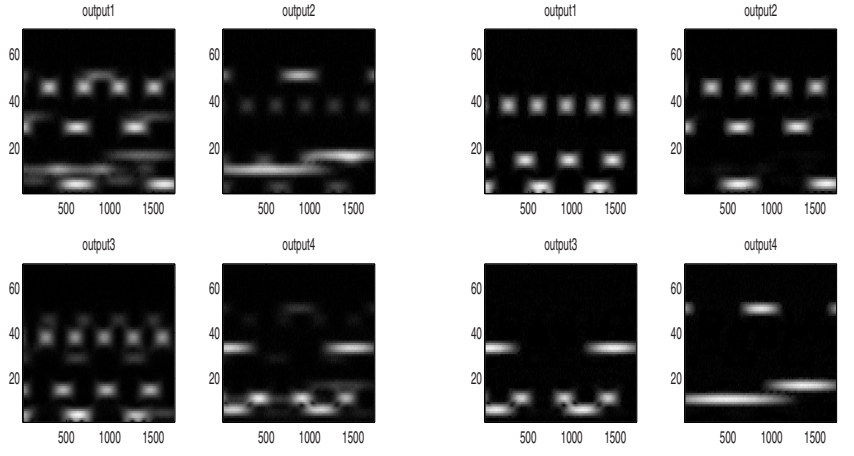


(a) (b)

Fig. 1. (a) Original 4 smooth source images; (b) Observed 9 very noisy mixed images with $SNR = 10$[dB])

convex, and has bounded first-derivative, we decided to select this function to our tests. Thus

$$\psi(\xi, \delta) = \delta \log[\cosh(\xi/\delta)], \tag{13}$$



(a) (b)

Fig. 2. Estimated sources with: (a) standard multiplicative (Lee-Seung) NMF algorithm (9)–(10) at $\alpha = 0$ ($SIR_X = 7.1, 11.7, 12.6, 13.1$[dB], $SIR_A = 12.3, 7.6, 14.8, 13.3$[dB] respectively); (b) Gibbs regularized algorithm given by (9)–(10) with parameters $\alpha = 0.2$ and $\delta = 10^{-3}$ ($SIR_X = 18.5, 18.3, 17.9, 18$[dB], $SIR_A = 29.6, 39.7, 27.4, 31.2$[dB], respectively)

which leads to

$$[\nabla_X U(\boldsymbol{X})]_{jk} = \sum_{l \in S_k} w_{kl} \tanh\left(\frac{x_{jk} - x_{jl}}{\delta}\right). \tag{14}$$

The set S_k and the associated weighting factors w_{kl} are usually defined by the MRF model. Taking into account the nearest neighborhood, $w_{kl} = 1$ for pixels adjacent along a horizontal or vertical line, and $w_{kl} = \frac{1}{\sqrt{2}}$ for pixels adjacent along a diagonal line.

Usually, the potential functions in (12) are parameter-dependent. At least, one parameter (in our case, the parameter δ) must be set up in advance, or simultaneously with the estimation. Generally, this can be regarded as a hyperparameter, and consequently estimated with maximization of the marginal likelihood function $p(\boldsymbol{Y})$ in (4). However, a direct estimation of the parameter from the data usually involves a high computational complexity, and it is not absolutely needed if we operate on one class of data for which preliminary simulations can be performed. We notice that for our class of data, the parameter has a very slight impact on the estimation in quite a wide range of its values. Thus, we set $\delta = 10^{-3}$ in all the tests in the paper.

3 Numerical Tests

The proposed algorithms have been extensively tested for various sets of the parameters (α and δ), and the algorithms are compared with the standard NMF

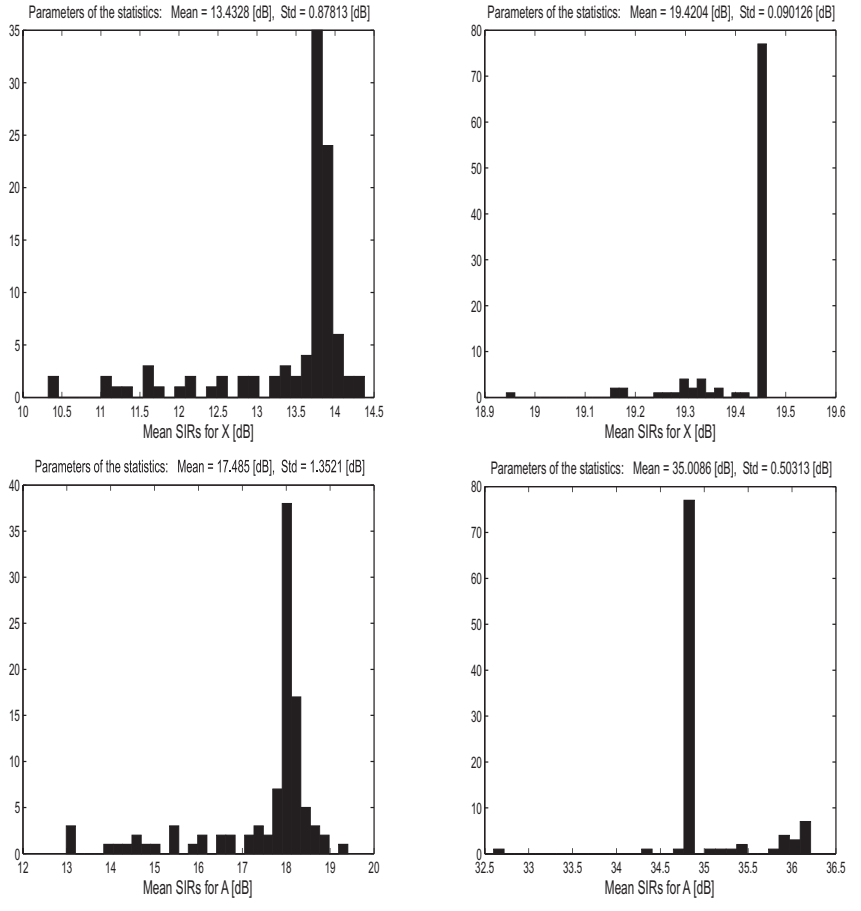


Fig. 3. Histograms from 100 mean-SIR samples generated with the following algorithms: (left) standard multiplicative (Lee-Seung) NMF algorithm; (right) Gibbs regularized algorithm; (top) estimation of $\boldsymbol{X}$ (sources); (bottom) estimation of columns in mixing matrix $\boldsymbol{A}$

algorithm. For the numerical tests we have used the benchmark of 4 smooth original images (Fig. 1(a)) which are mixed with the dense random mixing matrix $\boldsymbol{A} \in \mathbb{R}^{9\times 4}$ uniformly distributed ($\mathrm{cond}(\boldsymbol{A}) = 4.11$). The mixtures are then corrupted with the Gaussian noise of $SNR = 10$[dB]. Fig. 1(b) presents the noisy mixed images. The estimated images with the standard Lee-Seung algorithm (the updates (9)–(10) at $\alpha = 0$) are shown in Fig. 2(a). The results obtained with the improved Gibbs regularized NMF algorithm given by (9)–(10) are illustrated in Fig. 2(b) for $\alpha = 0.2$. The updating process for each algorithm has been terminated after 1000 alternating steps. The estimations are also quantitatively assessed with the standard Signal-to-Interference Ratio (SNR).

The same algorithms are also tested with the Monte Carlo (MC) analysis where for each run the initial conditions are randomly set. Fig. 3 presents the histograms obtained from 100 mean-SIR samples generated with the MC analysis for the above-mentioned NMF algorithms: unregularized version (left) and Gibbs regularized version (right).

4 Conclusions

In the paper, we derived the new algorithm for NMF, which may be useful for estimation of locally smooth images in BSS applications. The algorithm exploits the information on pair-wise interactions between adjacent pixels, which is motivated by MRF models in tomographic image reconstruction. Incorporating such a prior information to the NMF updating rules (especially for $\boldsymbol{X}$) is also very profitable for relaxing NMF ambiguity and non-convexity effects. The numerical results demonstrate the robustness of the proposed algorithm, especially for highly noisy data. The algorithm is much less sensitive to initialization in comparison to the standard NMF algorithms. This is confirmed with the MC simulations shown in Fig. 3. The proposed approach can be further extended with additional constraints or different updating rules. Also, another extension may concern the application of data-driven hyperparameter estimation techniques, especially for the regularization parameter.

The proposed algorithm has been implemented in Matlab Toolbox for Nonnegative Matrix Factorization: NMFLAB for Signal and Image Processing [36].

References

1. Lee, D.D., Seung, H.S.: Learning the parts of objects by nonnegative matrix factorization. Nature 401, 788–791 (1999)
2. Guillamet, D., Vitrià, J., Schiele, B.: Introducing a weighted nonnegative matrix factorization for image classification. Pattern Recognition Letters 24, 2447–2454 (2003)
3. Ahn, J.H., Kim, S., Oh, J.H., Choi, S.: Multiple nonnegative-matrix factorization of dynamic PET images. In: ACCV, p. 5 (2004)
4. Lee, J.S., Lee, D.D., Choi, S., Lee, D.S.: Application of nonnegative matrix factorization to dynamic positron emission tomography. In: 3rd International Conference on Independent Component Analysis and Blind Signal Separation, San Diego, CA, pp. 556–562 (2001)
5. Li, H., Adali, T., Wang, W., D.E.: Non-negative matrix factorization with orthogonality constraints for chemical agent detection in Raman spectra. In: IEEE Workshop on Machine Learning for Signal Processing, Mystic, USA (2005)
6. Carmona-Saez, P., Pascual-Marqui, R.D., Tirado, F., Carazo, J.M., Pascual-Montano, A.: Biclustering of gene expression data by non-smooth non-negative matrix factorization. BMC Bioinformatics 7 (2006)
7. Pascual-Montano, A., Carazo, J.M., Kochi, K., Lehmean, D., Pacual-Marqui, R.: Nonsmooth nonnegative matrix factorization (nsNMF). IEEE Trans. Pattern Analysis and Machine Intelligence 28, 403–415 (2006)

8. Shahnaz, F., Berry, M., Pauca, P., Plemmons, R.: Document clustering using nonnegative matrix factorization. Journal on Information Processing and Management 42, 373–386 (2006)
9. Okun, O., Priisalu, H.: Fast nonnegative matrix factorization and its application for protein fold recognition. EURASIP Journal on Applied Signal Processing Article ID 71817, 8 (2006)
10. Wang, Y., Jia, Y., Hu, C., Turk, M.: Non-negative matrix factorization framework for face recognition. International Journal of Pattern Recognition and Artificial Intelligence 19, 495–511 (2005)
11. Liu, W., Zheng, N.: Non-negative matrix factorization based methods for object recognition. Pattern Recognition Letters 25, 893–897 (2004)
12. Spratling, M.W.: Learning image components for object recognition. Journal of Machine Learning Research 7, 793–815 (2006)
13. Sajda, P., Du, S., Brown, T.R., Shungu, R.S.D.C., Mao, X., Parra, L.C.: Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. IEEE Trans. Medical Imaging 23, 1453–1465 (2004)
14. Cichocki, A., Zdunek, R., Amari, S.: New algorithms for non-negative matrix factorization in applications to blind source separation. In: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP2006, Toulouse, France, pp. 621–624 (2006)
15. Dhillon, I.S., Modha, D.M.: Concept decompositions for large sparse text data using clustering. Machine Learning J. 42, 143–175 (2001)
16. Berry, M., Browne, M., Langville, A., Pauca, P., Plemmons, R.: Algorithms and applications for approximate nonnegative matrix factorization. Computational Statistics and Data Analysis 52, 55–173 (2007)
17. Cho, Y.C., Choi, S.: Nonnegative features of spectro-temporal sounds for classification. Pattern Recognition Letters 26, 1327–1336 (2005)
18. Brunet, J.P., Tamayo, P., Golub, T.R., Mesirov, J.P.: Metagenes and molecular pattern discovery using matrix factorization. In: PNAS, vol. 101, pp. 4164–4169 (2000)
19. Rao, N., Shepherd, S.J., Yao, D.: Extracting characteristic patterns from genome – wide expression data by non-negative matrix factorization. In: Proc. of the 2004 IEEE Computational Systems Bioinformatics Conference (CSB 2004), Stanford, CA (2004)
20. Rutkowski, T.M., Zdunek, R., Cichocki, A.: Multichannel EEG brain activity pattern analysis in time-frequency domain with nonnegative matrix factorization support. International Congress Series 8611, 266–269 (2007)
21. Cichocki, A., Zdunek, R., Amari, S.: Csiszar's divergences for non-negative matrix factorization: Family of new algorithms. In: Rosca, J.P., Erdogmus, D., Príncipe, J.C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 32–39. Springer, Heidelberg (2006)
22. Dhillon, I., Sra, S.: Generalized nonnegative matrix approximations with Bregman divergences. In: Neural Information Proc. Systems, Vancouver, Canada, pp. 283–290 (2005)
23. Kompass, R.: A generalized divergence measure for nonnegative matrix factorization. Neural Computation 19, 780–791 (2006)
24. Hoyer, P.O.: Non-negative matrix factorization with sparseness constraints. Journal of Machine Learning Research 5, 1457–1469 (2004)

25. Kreutz-Delgado, K., Murray, J.F., Rao, B.D., Engan, K., Lee, T.W., Sejnowski, T.J.: Dictionary learning algorithms for sparse representation. Neural Computation 15, 349–396 (2003)
26. Zdunek, R., Cichocki, A.: Nonnegative matrix factorization with constrained second-order optimization. Signal Processing 87, 1904–1916 (2007)
27. Cichocki, A., Zdunek, R.: Regularized alternating least squares algorithms for nonnegative matrix/tensor factorizations. In: Liu, D., Fei, S., Hou, Z., Zhang, H., Sun, C. (eds.) ISNN 2007. LNCS, vol. 4493, pp. 793–802. Springer, Heidelberg (2007)
28. Zdunek, R., Cichocki, A.: Gibbs regularized nonnegative matrix factorization for blind separation of locally smooth signals. In: 15th IEEE International Workshop on Nonlinear Dynamics of Electronic Systems (NDES 2007), Tokushima, Japan, pp. 317–320 (2007)
29. Besag, J.: Toward Bayesian image analysis. J. Appl. Stat. 16, 395–407 (1989)
30. Hebert, T., Leahy, R.: A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. IEEE Transactions on Medical Imaging 8, 194–202 (1989)
31. Geman, S., McClure, D.: Statistical methods for tomographic image reconstruction. Bull. Int. Stat. Inst. LII-4, 5–21 (1987)
32. Geman, S., Reynolds, G.: Constrained parameters and the recovery of discontinuities. IEEE Trans. Pattern Anal. Machine Intell. 14, 367–383 (1992)
33. Stevenson, R., Delp, E.: Fitting curves with discontinuities. In: Proc. 1-st Int. Workshop on Robust Computer Vision, Seattle, Wash., USA (1990)
34. Green, P.J.: Bayesian reconstruction from emission tomography data using a modified EM algorithm. IEEE Trans. Medical Imaging 9, 84–93 (1990)
35. Lange, K., Carson, R.: EM reconstruction algorithms for emission and transmission tomography. J. Comp. Assisted Tomo. 8, 306–316 (1984)
36. Cichocki, A., Zdunek, R.: NMFLAB for Signal and Image Processing. Technical report, Laboratory for Advanced Brain Signal Processing, BSI, RIKEN, Saitama, Japan (2006)