

LOW-RANK BLIND NONNEGATIVE MATRIX DECONVOLUTION

Anh Huy Phan[†], Petr Tichavský^{‡*}, Andrzej Cichocki^{††} and Zbyněk Koldovsky[§]

[†]Brain Science Institute, RIKEN, Wakoshi, Japan

[‡]Institute of Information Theory and Automation, Prague, Czech Republic

[§]Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic

ABSTRACT

A novel blind deconvolution is proposed to seek for basis patterns and their location maps inside a nonnegative data matrix. Basis patterns can have different sizes, and shift in independent directions. Moreover, the location maps can be low-rank or rank-one matrices composed by two relatively small and tall matrices or by two vectors. A general framework to solve this problem together with algorithms are introduced. The experiments on music and texture decomposition will confirm performance of our method, and of the proposed algorithms.

Index Terms— nonnegative matrix deconvolution/factorization, music decomposition, pattern extraction

1. PROBLEM FORMULATION

Convolutional nonnegative matrix factorization (CNMF) or nonnegative matrix deconvolution (NMD) has found a number of applications in music analysis, source detection, image processing [1, 2, 3, 4]. This kind of deconvolution seeks for basis patterns which shift vertically [5] or horizontally [1, 6, 7], or in both directions [8, 9] in the data matrix $\mathbf{Y} = [\mathbf{y}_1 \mathbf{y}_2 \dots \mathbf{y}_J] \in \mathbb{R}_+^{I \times J}$, and can be modelled by

$$\mathbf{Y} \approx \widehat{\mathbf{Y}} = \sum_{p=1}^P \mathbf{A}^{(p)} * \mathbf{X}^{(p)} = \sum_{p=1}^P \widehat{\mathbf{Y}}^{(p)}, \quad (1)$$

where $\mathbf{A}^{(p)} \in \mathbb{R}_+^{R_p \times S_p}$ are nonnegative basis patterns (objects) whose locations and intensities are specified by location maps (matrices) $\mathbf{X}^{(p)} = [\mathbf{x}_1^{(p)} \mathbf{x}_2^{(p)} \dots \mathbf{x}_{L_p}^{(p)}] \in \mathbb{R}_+^{K_p \times L_p}$, $I = R_p + K_p - 1$, $J = S_p + L_p - 1$, $R_p \ll I$, $S_p \ll J$, $p = 1, 2, \dots, P$. The symbol “*” denotes the 2-D convolution defined for $\widehat{\mathbf{Y}}^{(p)} = [\widehat{y}_{i,j}^{(p)}] = \mathbf{A}^{(p)} * \mathbf{X}^{(p)}$ as $\widehat{y}_{i,j}^{(p)} = \sum_{r=-\infty}^{+\infty} \sum_{s=-\infty}^{+\infty} a_{r,s}^{(p)} x_{i-r,j-s}^{(p)}$, where entries $a_{r,s}^{(p)}$ and $x_{k,l}^{(p)}$ are set to zero if they are outside $\mathbf{A}^{(p)}$, $\mathbf{X}^{(p)}$, respectively. The roles of $\mathbf{A}^{(p)}$ and $\mathbf{X}^{(p)}$ are interchangeable within the model.

Normally, pattern sizes ($R_p \times S_p$) are relatively small compared to the data size ($I \times J$). Hence, the sizes of $\mathbf{X}^{(p)}$ ($K_p \times L_p$)

*The work of P. Tichavsky was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project IM0572 and by Grant Agency of the Czech Republic through the project 102/09/1278.

[†]Also affiliated with the EE Dept., Warsaw University of Technology and with Systems Research Institute, Polish Academy of Science, Poland.

are not significantly smaller than that of \mathbf{Y} . This leads to that the NMD model (1) may be expensive due to a large number of parameters to be estimated ($\approx PIJ$).

For spectral analysis of audio or EEG signals, an oscillation usually produces many frequencies or harmonics at the same time and over multi channels. That means the oscillation can be expressed by much smaller dimensional basis spectral objects shifting in both time and frequency. Moreover, due to appearances in the same time intervals and frequency bins, the location maps $\mathbf{X}^{(p)}$ are roughly rank-one matrices, or separable into much smaller matrices with a few components. This fact gives an ability to reduce complexity of the model (1).

Instead of solving the deconvolution (1), we consider a low-rank blind deconvolution model to seek for a set of P nonnegative basis objects (patterns) $\mathbf{A}^{(p)} \in \mathbb{R}_+^{R_p \times S_p}$ ($p = 1, 2, \dots, P$), $R_p \leq I$, $S_p \leq J$ and two activating matrices $\mathbf{V}^{(p)} \in \mathbb{R}_+^{K_p \times M_p}$ and $\mathbf{H}^{(p)} \in \mathbb{R}_+^{L_p \times N_p}$ for vertical and horizontal shiftings, $I = R_p + K_p + N_p - 2$ and $J = S_p + L_p + M_p - 2$, respectively

$$\mathbf{Y} \approx \widehat{\mathbf{Y}} = \sum_{p=1}^P \mathbf{V}^{(p)} * \mathbf{A}^{(p)} * \mathbf{H}^{(p)T} = \sum_{p=1}^P \widehat{\mathbf{Y}}^{(p)}, \quad (2)$$

where $\widehat{\mathbf{Y}}^{(p)} = \mathbf{V}^{(p)} * \mathbf{A}^{(p)} * \mathbf{H}^{(p)T}$ are approximations of \mathbf{Y} by $\mathbf{A}^{(p)}$. Pattern sizes should be much smaller than the data size, that is, $R_p \ll I$, $S_p \ll J$, and the number of activating components $1 \leq M_p \leq R_p$ and $1 \leq N_p \leq S_p$. We note that $\mathbf{X}^{(p)} = \mathbf{V}^{(p)} * \mathbf{H}^{(p)T}$ are the location matrices of $\mathbf{A}^{(p)}$ in (1). However, the low-rank NMD in (2) is much cheaper than NMD in (1) due to less fitting parameters.

In the sequel, a general framework is introduced to solve both approximations (1) and (2). Simulations on music and texture will confirm the validity of the proposed model.

2. RELATION WITH OTHER MODELS

When $M_p = N_p = 1, \forall p$, the low-rank NMD becomes the rank-1 NMD. In a particular case when patterns $\mathbf{A}^{(p)} = \mathbf{a}_1^{(p)} \mathbf{a}_2^{(p)T}$ are rank-one matrices, or they consist of only one element $R_p = S_p = 1$, we obtain NMF from rank-1 NMD. NMF can extract only rank-one patterns whereas NMD can retrieve replicate high-rank structures inside a data matrix.

This also indicates that NMD is superior to NMF in the sense of pattern extraction although NMF has a lower complexity than NMD. Based on the distributive law of the convolution, it is straightforward to verify that an NMD in (1) or (2) is in fact an NMF with much higher rank than P . Therefore, high-rank NMF can approximate the data \mathbf{Y} at a fitness comparable with or even higher than NMD. However, NMF cannot incorporate similar components which explain the same structures (such as notes) into the same groups by itself. This characteristic is an important advantage of NMD.

If one of the activating matrices disappears in (2), and the other is a vector, low-rank NMD is simplified into CNMF for horizontal shifting [1, 6, 7], or overlapping NMF (ONMF) for vertical shifting [5]. When $\mathbf{A}^{(p)}$ absorbs into $\mathbf{V}^{(p)}$ or $\mathbf{H}^{(p)}$, low-rank NMD becomes CNMF2D [8]. When objects have the same size, NMD becomes the shift invariant sparse coding model (SCSI) for single channel [8].

We note that all existing nonnegative deconvolution models such as CNMF, ONMF, CNMF2F and SCSI use the truncated convolution, and their basis patterns are of the same size, or one of their dimensions is exactly the data size, that is, $R_p = I$ or $S_p = J, \forall p$. For our models including the generalized NMD (1) and the low-rank NMD (2), patterns can have arbitrary dimensions. This allows us to extract patterns with different models in an approximation. For example, some patterns can be rank-one matrices as in NMF, some patterns can shift along horizontal or vertical direction of the data (CNMF or ONMF), the others can be rank-one or low-rank NMD. That means the low-rank NMD (2) is much more flexible than the other existing models. Finally, NMD can easily impose additional constraints on patterns $\mathbf{A}^{(p)}$ and their activating matrices $\mathbf{V}^{(p)}$ and $\mathbf{H}^{(p)}$.

3. GENERAL FRAMEWORK AND ALGORITHMS FOR NMD

This section presents a general framework to solve NMD. We define Toeplitz matrices $\mathbf{T}_l^{(p)} \in \mathbb{R}^{I \times R_p}, l = 1, \dots, L_p, p = 1, \dots, P$ having $\tilde{\mathbf{x}}_l^{(p)} = [\mathbf{x}_l^{(p)T} \mathbf{0}_{1 \times R_p - 1}]^T$ as the first column and $[\mathbf{x}_l^{(p)}(1) \mathbf{0}_{1 \times (R_p - 1)}]$ as the first row. The shift matrices \mathbf{S}_l are binary matrices of size $J \times J$ with ones only on the l -th superdiagonal for $l > 0$, or on the l -th subdiagonal for $l < 0$, and zeroes elsewhere. Note that $\mathbf{S}_0 = \mathbf{I}_J, \mathbf{S}_l = \mathbf{0}_J$ for $|l| > J$ and $\mathbf{T}_l^{(p)} = [\tilde{\mathbf{x}}_l^{(p)} \tilde{\mathbf{S}}_{-1} \tilde{\mathbf{x}}_l^{(p)} \dots \tilde{\mathbf{S}}_{-R_p+1} \tilde{\mathbf{x}}_l^{(p)}], l = 1, 2, \dots, L_p$, where $\tilde{\mathbf{S}}$ is an $I \times I$ shift matrix. By denoting $\tilde{\mathbf{A}}^{(p)} = [\mathbf{A}^{(p)} \mathbf{0}_{R_p \times (J - S_p)}] \in \mathbb{R}^{R_p \times J}$, the NMD (1) is rewritten as

$$\mathbf{Y} = \sum_{l=1}^{L_p} \mathbf{T}_l^{(p)} \tilde{\mathbf{A}}^{(p)} \mathbf{S}_{l-1} + \left(\sum_{q \neq p} \widehat{\mathbf{Y}}^{(q)} \right) + \mathbf{E}, \quad (3)$$

or

$$\begin{aligned} \text{vec}(\mathbf{Y}) &= \left(\sum_{l=1}^{L_p} (\mathbf{S}_{-l+1} \otimes \mathbf{T}_l^{(p)}) \right) \text{vec}(\tilde{\mathbf{A}}^{(p)}) + \text{vec}(\widehat{\mathbf{Y}}^{(-p)} + \mathbf{E}) \\ &= \mathbf{Z}_A^{(p)} \text{vec}(\tilde{\mathbf{A}}^{(p)}) + \text{vec}(\widehat{\mathbf{Y}}^{(-p)} + \mathbf{E}), \end{aligned} \quad (4)$$

where symbol “ \otimes ” denotes the Kronecker product, \mathbf{E} is the approximation error, and $\widehat{\mathbf{Y}}^{(-p)} = \sum_{q \neq p} \widehat{\mathbf{Y}}^{(q)}$. In order to update $\mathbf{A}^{(p)}$, we fix all other parameters and minimize the following cost function

$$\arg \min_{\mathbf{A}^{(p)} \geq \mathbf{0}} \|\mathbf{Y} - \widehat{\mathbf{Y}}\|_F^2 = \|\text{vec}(\mathbf{Y} - \widehat{\mathbf{Y}}^{(-p)}) - \mathbf{Z}_A^{(p)} \text{vec}(\tilde{\mathbf{A}}^{(p)})\|_2^2. \quad (5)$$

There are plenty of methods to estimate $\mathbf{A}^{(p)}$ from (5) such as the multiplicative update rules, the ALS and recursive ALS update rules, projected gradient methods [10]. By employing one of them we can update $\mathbf{A}^{(p)}, \forall p$. Note that $\mathbf{A}^{(p)} * \mathbf{X}^{(p)} = \mathbf{X}^{(p)} * \mathbf{A}^{(p)}, \mathbf{X}^{(p)}$ is updated using a similar method.

For low-rank or rank-1 NMD, patterns $\mathbf{A}^{(p)}$ and activating matrices $\mathbf{V}^{(p)}$ and $\mathbf{H}^{(p)}$ can be updated by minimizing a similar cost function as (5). Based on the commutative and associative laws of the convolution, the low-rank NMD (2) is rewritten in 3 equivalent NMD forms

$$\mathbf{Y} = \mathbf{V}^{(p)} * (\mathbf{A}^{(p)} * \mathbf{H}^{(p)T}) + \widehat{\mathbf{Y}}^{(-p)} + \mathbf{E} = \mathbf{V}^{(p)} * \mathbf{X}_{\mathbf{V}^{(p)}} + \mathbf{E}_p \quad (6)$$

$$= \mathbf{A}^{(p)} * (\mathbf{V}^{(p)} * \mathbf{H}^{(p)T}) + \mathbf{E}_p = \mathbf{A}^{(p)} * \mathbf{X}_{\mathbf{A}^{(p)}} + \mathbf{E}_p \quad (7)$$

$$= \mathbf{H}^{(p)T} * (\mathbf{V}^{(p)} * \mathbf{A}^{(p)}) + \mathbf{E}_p = \mathbf{H}^{(p)T} * \mathbf{X}_{\mathbf{H}^{(p)}} + \mathbf{E}_p. \quad (8)$$

By solving similar problems as in (5) we can alternatively update $\mathbf{A}^{(p)}, \mathbf{V}^{(p)}$ and $\mathbf{H}^{(p)}$.

The following derivation is for the multiplicative algorithm for low-rank NMD. By applying the multiplicative Least-Square update rule to (5), $\tilde{\mathbf{A}}^{(p)}$ ($p = 1, 2, \dots, P$) are updated as follows

$$\text{vec}(\tilde{\mathbf{A}}^{(p)}) \leftarrow \text{vec}(\tilde{\mathbf{A}}^{(p)}) \otimes (\mathbf{Z}_A^{(p)T} \text{vec}(\mathbf{Y})) \oslash (\mathbf{Z}_A^{(p)T} \text{vec}(\widehat{\mathbf{Y}})),$$

where symbols “ \otimes ”, “ \oslash ” denote the Hadamard product and division. That leads to the update rules for $\mathbf{A}^{(p)}$

$$\mathbf{A}^{(p)} \leftarrow \mathbf{A}^{(p)} \otimes \left(\sum_{l=1}^{L_p} \mathbf{T}_l^{(p)T} \mathbf{Y}_{I_l} \right) \oslash \left(\sum_{l=1}^{L_p} \mathbf{T}_l^{(p)T} \widehat{\mathbf{Y}}_{I_l} \right), \quad (9)$$

where $I_l = [l, l+1, \dots, l+S_p-1]$, and $\mathbf{Y}_{I_l} = [\mathbf{y}_l \mathbf{y}_{l+1} \dots \mathbf{y}_{l+S_p-1}]$ is a subset of S_p columns of \mathbf{Y} with $\mathbf{y}_l \equiv \mathbf{0}$ for $l \geq J$. From (6), (7) and (8), we construct Toeplitz matrices from columns of $\mathbf{X}_{\mathbf{V}^{(p)}} = \mathbf{A}^{(p)} * \mathbf{H}^{(p)}, \mathbf{X}_{\mathbf{A}^{(p)}} = \mathbf{V}^{(p)} * \mathbf{H}^{(p)T}$ and $\mathbf{X}_{\mathbf{H}^{(p)}} = \mathbf{V}^{(p)} * \mathbf{A}^{(p)}$, and alternatively update $\mathbf{A}^{(p)}, \mathbf{V}^{(p)}$ and $\mathbf{H}^{(p)}, \forall p$.

In the same manner, algorithms for NMD and low rank NMD can be straightforwardly extended from ones for NMF. For example, we can derive the recursive (Q)ALS algorithm for NMD whose truncated version for CNMF is introduced in [7]. Note that any algorithm for NMD can be applied to NMF, CNMF, ONMF and CNMF2D. That means we have new multiplicative algorithms for CNMF, ONMF and CNMF2D deduced from (9). Although there are some existing multiplicative algorithms for CNMF and ONMF [1, 3, 6], for CNMF2D and SCSI [8, 9], the proposed algorithm with update rules (9) is different from them in the sense of mechanism and flexibility.

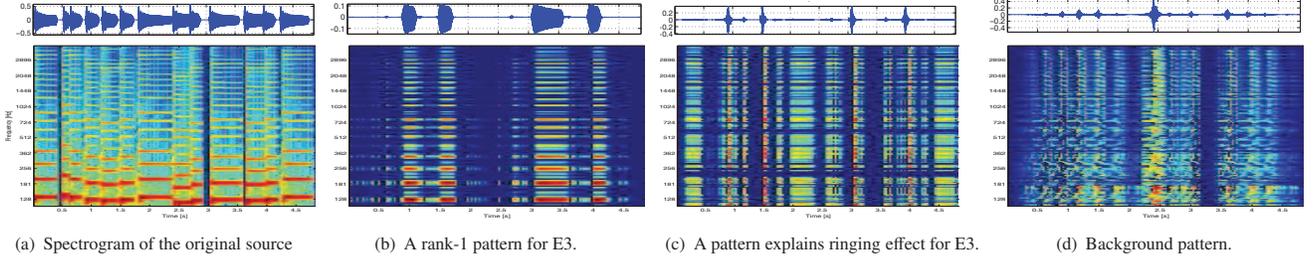


Fig. 1. Waveforms and log-frequency spectrograms of the observed sequence and of 3 spectral patterns which explain stripe patterns, ringing effect of the note E3, and background of \mathbf{Y} .

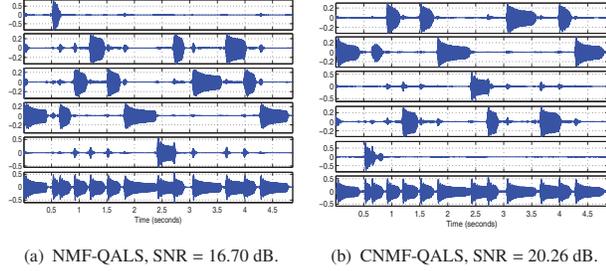


Fig. 2. Waveforms of basis sequences constructed from patterns and their location matrices estimated by NMF, CNMF. The reconstructed sequences (in the bottom) are summation of basis sequences.

4. EXAMPLE

4.1. Analysis of Patterns in Music

In this example, we perform decompositions for multi harmonic frequencies played on a guitar. The sampled song “London Bridge” is composed of five notes A3, G3, F3, E3 and D3 (sampled at frequency rate of 8 kHz in 5 seconds) (see Chapter 3 [10]). The log-frequency spectrogram \mathbf{Y} (364×151) illustrated in Fig. 1(a) was converted from the linear-frequency spectrogram with a quality factor $Q = 100$ and in the frequency range from $f_0 = 109.4$ Hz (bin 8) to $f_l = f_s/2 = 4000$ Hz (bin 257). We note that the lowest approximation error was 27.56 dB when there was no decomposition. Algorithms were applied to extract the notes, and to explain the observed sequence through basis audio sequences. The approximation signals were reconstructed from basis patterns, and normalized to have the same energy as the original signal. For CNMF and ONMF, the proposed multiplicative algorithm (oMLS) and the oQALS algorithm which sequentially update objects (patterns) were compared with the average multiplicative LS algorithm (aMLS) [1], the simplified multiplicative LS algorithm (MLS) [9].

As seen on Fig. 1(a), the major parts of the notes are thin stripes each of which is composed by one spectral component and one temporal component. However, spectrograms of the notes not only consist of stripes, but also comprise background patterns which are strong at low frequencies, and weaker at higher frequencies. Moreover, boundaries of the stripes are smeared at beginnings of notes or at transitions between them, especially at low frequencies, for example, at beginning of E3 around 128 Hz. NMF can extract the rank-one parts (stripes) but cannot capture fully the notes by only

Table 1. Performance comparison for various NMDs for Examples 4.1 and 4.2.

Ex.	Model	Algorithm	Pattern		No. Params.	SNR (dB)	
			$P - \mathbf{V}^{(p)}$	$\mathbf{A}^{(p)} * \mathbf{H}^{(p)T}$			
4.1	NMF	QALS	5 - (364×1) * (1×151)		2575	16.70	
			10 - (364×1) * (1×151)		5150	23.91	
	CNMF	oMLS	5 - (364×10) * (1×151)		18955	11.81	
			5 - (364×10) * (1×151)		18955	19.42	
			5 - (364×10) * (1×142)		18910	19.13	
			5 - (364×10) * (1×142)		18910	20.26	
	ONMF	oQALS	5 - (364×1) * (10×151)		9370	17.84	
			5 - (364×1) * (10×151)		9370	19.24	
	4.2	NMF	QALS	5 - (355×1) * (10×151)		9325	19.90
				10 - (355×1) * (10×151)		18650	24.27
				10 - (355×1) * (10×151)		18650	25.64
		NMD	oMLS	5 - (362×2) * (2×2) * (2×149)		5130	18.46
5 - (360×3) * (3×3) * (3×147)					7650	18.74	
10 - (362×2) * (2×2) * (2×149)					10260	24.59	
5 - (10×10) * (355×142)					252550	27.27	
5 - (364×1) * (1×151)							
5 - (362×2) * (2×2) * (2×149)							
1 - (354×2) * (10×10) * (2×141)					20089	26.54	
4.2	NMF	QALS	5 - (383×1) * (1×385)		3840	25.08	
			10 - (383×1) * (1×185)		7690	28.03	
	NMD	oMLS	5 - (374×1) * (10×10) * (1×376)		3400	28.09	

5 components. That’s why the signal reconstructed by NMF has a signal-to-noise ratio $\text{SNR} = 16.70$ dB. Waveforms of basis and approximate sequences are shown in Fig. 2(a).

CNMF and ONMF were set to extract 5 patterns with 10 components. The results given in Table 1 indicate that with the same number of patterns, the convolutive models yield better approximation than NMF because they explain the notes by more parameters. Increasing the number of components in NMF can improve the approximation, but NMF cannot incorporate similar components which explain the same notes. Waveforms of basis sequences and the reconstructed signal by oQALS for CNMF are shown in Fig. 2(b).

Signal reconstructed from 5 NMD patterns of size 2×2 with two activating matrices of size 362×2 and 149×2 achieved an $\text{SNR} = 18.46$ dB. That means low-rank NMD explains the notes better than NMF. Extracting 10 NMD patterns of the same size (2×2) needs to estimate only 10300 parameters and significantly improves the approximation signal up-to an $\text{SNR} = 24.59$ dB. Note that in order to retrieve 5 patterns of 10 columns, CNMF algorithms need to estimate 18910 parameters, while ONMF requires 9370 parameters to fit the spectrogram. The number of parameters for ONMF will be

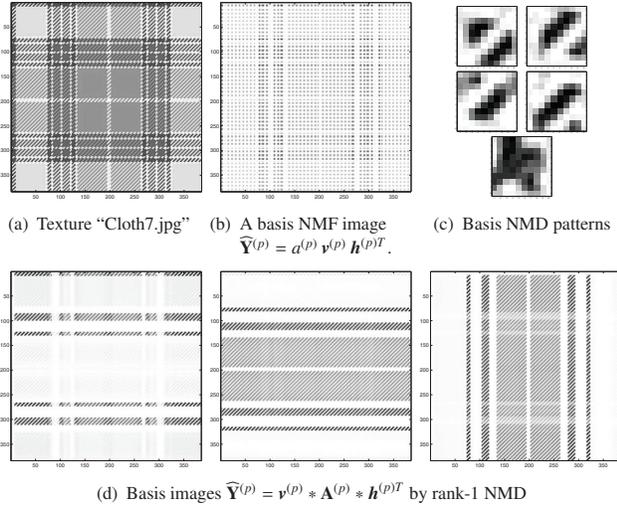


Fig. 3. Pattern extraction from “Cloth7” for Example 4.2.

dramatically increased for analysis of longer sequences.

An approximation by full NMD with 5 patterns $\mathbf{A}^{(p)} \in \mathbb{R}_+^{10 \times 10}$ and 5 $\mathbf{X}^{(p)} \in \mathbb{R}_+^{355 \times 142}$ achieved a perfect reconstruction with SNR = 27.27 dB. However, this deconvolution is extremely expensive due to estimation of 252.550 parameters.

An extra simulation illustrates the flexibility of NMD. The major parts of the notes are rank-1 matrices (stripes). Ringing effects near the note boundaries can be modelled by relatively small patterns which shift along both directions (time and frequency) of the spectrogram. The background can be modelled by larger patterns such as tall patterns covering the whole frequency band and shifting along the time axis, or long patterns shifting along the frequency axis. As a result, we decomposed \mathbf{Y} into 5 rank-one patterns (NMF), 5 low-rank patterns of size 2×2 , 1 pattern of size 10×10 , 2 tall patterns of size 364×10 and 2 long patterns of size 151×10 . This produced a nearly optimal reconstruction with SNR = 26.54 dB from 20089 arguments. The rank-1 spectrogram of the note E3 shown in Fig. 1(b) is composed by a spectral component and a temporal component, where the ringing parts appear at beginning of E3 illustrated in Fig. 1(c) is constructed from a 2×2 pattern and two rank-2 activating matrices. Fig. 1(d) illustrates background of \mathbf{Y} composed by a tall pattern of size 364×10 .

4.2. Texture Decomposition

In this example, NMD was applied to seek for basis complex (high-rank) patterns inside textures. The observed RGB texture “Cloth7” of size 383×385 was taken from the library <http://textures.forrest.cz/index.php?spgmGal=fabric&spgmPic=9>. We considered only the luminance component shown in Fig.3(a). The texture comprises small diagonal structural patterns, and flat regions in the center. NMF can explain the texture by basis images each of which is composed by a vertical component and a horizontal component. Fig. 3(b) illustrates one of basis rank-1 images obtained by NMF with 5 components. It is clear that NMF cannot retrieve

diagonal patterns because they are not rank-1 matrices in spite of being simple structures. By applying rank-1 NMD with $P = 5$ patterns $\mathbf{A}^{(p)} \in \mathbb{R}_+^{10 \times 10}$ and $\mathbf{v}^{(p)} \in \mathbb{R}_+^{374}$ and $\mathbf{h}^{(p)} \in \mathbb{R}_+^{376}$, we can obtain the diagonal patterns as illustrated in Fig. 3(c). Some basis images $\widehat{\mathbf{Y}}^{(p)}$ shown in Fig. 3(d) intuitively reveal replication of patterns. Comparison between NMF and rank-1 NMD for this texture is given in Table 1. Increasing the number of NMF components can improve the approximation but we cannot obtain complex patterns by NMF.

5. CONCLUSIONS

New blind deconvolution model has been proposed to seek for patterns inside nonnegative data matrices. Patterns can have arbitrary sizes, and can shift independently in the data. Moreover, NMD and low-rank NMD allow multiple models in an approximation. That means our models are much more flexible than the existing convolutive models. The paper also presented a general approach to estimate basis patterns and the location matrices in NMD. As a result, we can convey an arbitrary NMF update rules into the new model. In addition, a simple multiplicative algorithm has been introduced to verify the proposed approach. The experimental results have confirmed the validity of our model and of our algorithms for both music and texture deconvolution. Especially, the proposed algorithms outperformed other multiplicative algorithms for the same model.

6. REFERENCES

- [1] P. Smaragdis and J.C. Brown, “Nonnegative matrix factorization for polyphonic music transcription,” in *IEEE WASPAA*, New Paltz, NY, October 2003, pp. 177–180.
- [2] P. Smaragdis, “Convolutive speech bases and their application to supervised speech separation,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 1, pp. 1–12, 2007.
- [3] A. Ozerov and C. Fevotte, “Multichannel nonnegative matrix factorization in convolutive mixtures, with application to blind audio source separation,” in *IEEE ICASSP 2009*, '09, pp. 3137–3140.
- [4] A. T. Cemgil, U. Simsekli, and Y. C. Subakan, “Probabilistic latent tensor factorization framework for audio modeling,” in *WASPAA*, 2011.
- [5] J. Eggert, H. Wersing, and E. Koerner, “Transformation-invariant representation and NMF,” USA, 2004, Proceedings of IJCNN.
- [6] P. D. O’Grady and B. A. Pearlmutter, “Discovering speech phones using convolutive non-negative matrix factorisation with a sparseness constraint,” *Neurocomput.*, vol. 72, pp. 88–101, 2008.
- [7] A.-H. Phan, A. Cichocki, P. Tichavský, and Z. Koldovský, “On connection between the convolutive and ordinary nonnegative matrix factorizations,” in *LVA/ICA, Springer LNCS-7191*, pp. 288–296, 2012.
- [8] M. Mørup, M.N. Schmidt, and L.K. Hansen, “Shift invariant sparse coding of image and music data,” Tech. Rep., 2008.
- [9] M.N. Schmidt and M.Mørup, “Nonnegative matrix factor 2-D deconvolution for blind single channel source separation,” *Springer LNCS*, vol. 3889, pp. 700–707, 2006.
- [10] A. Cichocki, R. Zdunek, A.-H. Phan, and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Chapter 3, Wiley, Chichester, 2009.