

Music Signal Separation by Supervised Nonnegative



Matrix Factorization with Basis Deformation

Kazunobu Kondo, Yu Takahashi (Yamaha Corporate Research & Development Center, Shizuoka, Japan)



5. Evaluation experiment

To confirm the effectiveness of the proposed algorithm, we compared the conventional method (SNMF) and our SNMF with basis deformation. Experimental condition

Experimental condition					
Target instruments	Flute, Clarinet, Piano, Trombone				
Observed signal	Mixing two sources selected from four sources with the input SNR of 0dB				
Supervision sound (MIDI)	Artificial MIDI sounds of the target instruments that consists two octave notes, which cover all notes of the target signal				
Number of bases	Supervision bases: 100, Other bases: 50				
Number of iterations of NMF	Training process: 500, Separation process: 400				
Parameters	Experimentally determined				
Evaluation scores [9]	Signal to distortion ratio (SDR: quality of extracted signal), Source to interference ratio (SIR: degree of separation), Sources to artifact ratio (SAR: absence of distortion)				

The target signals were recorded with actual musical instruments.

The supervision sounds were generated by MIDI synthesizer.

Flute	67		1		2		٢	\$	
Clarinet	\$1	r .				_	<u>f</u>	ż	
Piano	<u>\$</u> ‡	ينسب			+		P,	*	ŧ
Trombone	ツ肴	5.0	ľ	٢		٢	r.	۲	

Target signals

Experimental results

Target	Other	Conventional method			Proposed method		
sound	sound	SDR	SIR	SAR	SDR	SIR	SAR
Piano	Clarinet	1.0	6.1	3.5	7.4	13.0	9.1
Piano	Trombone	1.6	13.7	2.1	12.3	23.5	12.7
Clarinet	Flute	0.1	1.8	7.2	0.6	2.4	7.3
Clarinet	Trombone	-0.5	12.8	0.0	9.4	21.5	9.7
Flute	Piano	4.2	14.2	4.8	6.4	16.6	6.9
Trombone	Clarinet	0.7	12.5	1.2	5.4	17.0	5.8

Average scores



Example of spectrograms



References

D. D. Lee, H. S. Seung, "Algorithms for non-negative matrix factorization," Neural Info. Process. Syst., vol.13, pp.556–562, 2001. T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," IEEE Trans. Audio, Speech and Language Processing, vol.15, pp.1066-1074, 2007. S. A. Raczynski, N. Ono, S. Sagayama, "Multipitch analysis with harmonic nonnegative matrix approximation," Proc. ISMIR, pp.381-386

[4] A. Ozerov, C. Fevotte, M. Charbit, "Factorial scaled hidden Markov model for polyphonic audio representation and source separation," Proc. WASPAA, 2009.

[] W. Wang, A. Cichocki, J. A. Chambers, "A multiplicative algorithm for convolutive non-negative matrix factorization based on squared Euclidean distance," IEEE Trans. on Signal Processing, vol.57, no.7, pp.2858-2864, 2009. H. Kameoka, M. Nakano, K. Ochiai, Y. Imoto, K. Kashino, S. Sagayama, "Constrained and regularized variants of non-negative matrix factorization incorporating music-specific constraints," Proc. ICASSP, pp.5365–5368, 2012. P. Smaragdis, B. Raj, M. Shashanka, "Supervised and semi-supervised separation of sounds from single-channel mixtures," Proc.

LVA/ICA 2010, LNCS 6365, pp.140-148, 2010. K. Yagi, Y. Takahashi, H. Saruwatari, K. Shikano, K. Kondo, "Music signal separation by orthogonality and maximum-distance constrained

nonnegative matrix factorization with target signal information," Proc. Audio Engineering Society 45th International Conference, 2012. E. Vincent, R. Gribonval, C. Fevotte. "Performance measurement in blind audio source separation," IEEE Trans. Audio, Speech and Language Processing, vol.14, no.4, pp.1462–1469, 2006.