











- [8] P.-S. Huang, S. D. Chen, P. Smaragdis, and M. Hasegawa-Johnson, "Singing-voice separation from monaural recordings using robust principal component analysis," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, pp. 57-60, 2012.
- [9] C.-L. Hsu and J.-S. R. Jang, "On the improvement of singing voice separation for monaural recordings using the MIR-1K dataset," *IEEE Trans. on Audio, Speech, and Language Process.*, Vol. 18, No. 2, pp. 310-319, 2010.
- [10] C.-C. Hsu, T.-E. Lin, J.-H. Chen, and T.-S. Chi, "Voice activity detection based on frequency modulation of harmonics," *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, pp. 6679-6683, 2013.
- [11] D. Klein, and S. A. Shamma, "The case of the missing pitch templates: how harmonic templates emerge in the early auditory system," *J. Acoust. Soc. Am.*, Vol. 107, No. 5, pp. 2631-2644, 2000.
- [12] H. Lei, B. T. Meyer, and N. Mirghafori, "Spectro-temporal Gabor features for speaker recognition," *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, pp. 4241-4244, 2012.
- [13] Y. Li and D. Wang, "Separation of singing voice from music accompaniment for monaural recordings," *IEEE Trans. on Audio, Speech, and Language Process.*, Vol. 15, No. 4, pp. 1475-1487, 2007.
- [14] B. C. J. Moore: *An Introduction to the Psychology of Hearing 5<sup>th</sup> Ed.*, Academic Press, 2003.
- [15] A. Ozerov, P. Philippe, F. Bimbot, and R. Gribonval, "Adaptation of Bayesian models for single channel source separation and its application to voice / music separation in popular songs," *IEEE Trans. on Audio, Speech, and Language Process.*, special issue on Blind Signal Proc. for Speech and Audio Applications, Vol. 15, No. 5, pp. 1564-1578, 2007.
- [16] Z. Rafii and B. Pardo, "Repeating Pattern Extraction Technique (REPET): A Simple Method for Music/Voice Separation," *IEEE Trans. on Audio, Speech, and Language Process.*, Vol. 21, No. 1, pp. 73-84, 2013.
- [17] P. Sprechmann, A. Bronstein, and G. Sapiro, "Real-time online singing voice separation from monaural recordings using robust low-rank modeling," *Proc. of the Int. Soc. for Music Inform. Retrieval Conf.*, pp. 67-72, 2012.
- [18] R. M. Stern and N. Norgan, "Hearing is believing: biologically inspired methods for robust automatic speech recognition," *IEEE Signal Process. Mag.*, Vol. 29, No. 6, pp. 34-43, 2012.
- [19] H. Tachibana, N. Ono, and S. Sagayama, "Singing Voice Enhancement in Monaural Music Signals Based on Two-stage Harmonic/Percussive Sound Separation on Multiple Resolution Spectrograms," *IEEE/ACM Trans. on Audio, Speech, and Language Process.*, Vol. 22, No. 1, pp. 228-237, 2014.
- [20] S. Vembu and S. Baumann, "Separation of vocals from polyphonic audio recordings," *Proc. of the Int. Soc. for Music Inform. Retrieval Conf.*, pp. 337-344, 2005.
- [21] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. on Audio, Speech, and Language Process.*, Vol. 14, No. 4, pp. 1462-1469, 2006.
- [22] Y. Yang, "Low-rank representation of both singing voice and music accompaniment via learned dictionaries," *Proc. of the Int. Soc. for Music Inform. Retrieval Conf.*, pp. 427-432, 2013.
- [23] [http://bass-db.gforge.inria.fr/bss\\_eval/](http://bass-db.gforge.inria.fr/bss_eval/)
- [24] <http://www.isr.umd.edu/Labs/NSL/nsl.html>