instrument transcription system based on probabilistic latent component analysis, and performed a conservative transcription pre-processing step in order to detect notes with a high confidence. Based on the initial transcription, the spectra of the detected notes are collected, processed, and are used in order to create a new dictionary that closely matches the spectral characteristics of the input instrument source(s). Both single-instrument and multi-instrument variants of the proposed method are presented and evaluated, in terms of multi-pitch detection and instrument assignment. Experimental results using the MAPS and Bach10 datasets show that there is a clear and consistent performance improvement when using the proposed template adaptation method, especially when there is a large discrepancy between the original dictionary and the spectral characteristics of the test instrument sources.

In the future, we will evaluate the proposed system using multiple-instrument recordings with more than two instruments. Parametric models (such as source-filter models) will also be investigated for updating the note templates, along with adaptive methods for deriving the conservative transcription threshold. We also plan to evaluate the proposed system in the next MIREX evaluations [1]. Finally, the proposed template adaptation steps will also be evaluated in the context of score-informed source separation using spectrogram factorization models [9].

6. REFERENCES

[1] Music Information Retrieval Evaluation eXchange (MIREX). `http://music-ir.org/mirexwiki/`.

[2] E. Benetos, S. Cherla, and T. Weyde. An efficient shift-invariant model for polyphonic music transcription. In *6th Int. Workshop on Machine Learning and Music*, Prague, Czech Republic, September 2013.

[3] E. Benetos, S. Dixon, D. Giannoulis, H. Kirchhoff, and A. Klapuri. Automatic music transcription: challenges and future directions. *J. Intelligent Information Systems*, 41(3):407–434, December 2013.

[4] J. J. Carabias-Orti, T. Virtanen, P. Vera-Candeas, N. Ruiz-Reyes, and F. J. Cañadas-Quesada. Musical instrument sound multi-excitation model for non-negative spectrogram factorization. *IEEE J. Selected Topics in Signal Processing*, 5(6):1144–1158, 2011.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Society*, 39(1):1–38, 1977.

[6] A. Dessein, A. Cont, and G. Lemaitre. Real-time polyphonic music transcription with non-negative matrix factorization and beta-divergence. In *11th Int. Society for Music Information Retrieval Conf.*, pages 489–494, 2010.

[7] Z. Duan, B. Pardo, and C. Zhang. Multiple fundamental frequency estimation by modeling spectral peaks and non-peak regions. *IEEE Trans. Audio, Speech, and Language Processing*, 18(8):2121–2133, 2010.

[8] V. Emiya, R. Badeau, and B. David. Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle. *IEEE Trans. Audio, Speech, and Language Processing*, 18(6):1643–1654, 2010.

[9] S. Ewert, B.Pardo, M. Müller, and M. D. Plumbley. Score-informed source separation for musical audio recordings. *IEEE Signal Processing Magazine*, 31(3):116–124, May 2014.

[10] B. Fuentes, R. Badeau, and G. Richard. Harmonic adaptive latent component analysis of audio and application to music transcription. *IEEE Trans. Audio, Speech, and Language Processing*, 21(9):1854–1866, 2013.

[11] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. RWC music database: music genre database and musical instrument sound database. In *International Conference on Music Information Retrieval*, October 2003.

[12] H. Kirchhoff, S. Dixon, and Anssi Klapuri. Missing template estimation for user-assisted music transcription. In *IEEE Int. Conf. Audio, Speech and Signal Processing*, pages 26–30, 2013.

[13] C. Schörkhuber and A. Klapuri. Constant-Q transform toolbox for music processing. In *7th Sound and Music Computing Conf.*, Barcelona, Spain, July 2010.

[14] M. Shashanka, B. Raj, and P. Smaragdis. Probabilistic latent variable models as nonnegative factorizations. *Computational Intelligence and Neuroscience*, 2008. Article ID 947438.

[15] P. Smaragdis and G. Mysore. Separation by "humming": user-guided sound extraction from monophonic mixtures. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 69–72, New Paltz, USA, October 2009.

[16] D. Tidhar, M. Mauch, and S. Dixon. High precision frequency estimation for harpsichord tuning classification. In *IEEE Int. Conf. Audio, Speech and Signal Processing*, pages 61–64, Dallas, USA, March 2010.

[17] F. Weninger, C. Kirst, B. Schuller, and H.-J. Bungartz. A discriminative approach to polyphonic piano note transcription using supervised non-negative matrix factorization. In *IEEE Int. Conf. Audio, Speech and Signal Processing*, pages 6–10, May 2013.

[18] T. B. Yakar, P. Sprechmann, R. Litman, A. M. Bronstein, and G. Sapiro. Bilevel sparse models for polyphonic music transcription. In *14th Int. Society for Music Information Retrieval Conf.*, pages 65–70, 2013.

[19] K. Yoshii and M. Goto. Infinite composite autoregressive models for music signal analysis. In *13th Int. Society for Music Information Retrieval Conf.*, pages 79–84, October 2012.