IMPLEMENTATION OF HARMONIC-PERCUSSIVE SOUND SEPARATION FOR AUDACITY

Viktor Tamás Erdélyi
National Institute of Informatics
Saarland University[1]
Tokyo, Japan and Saarbrücken, Germany
verdelyi@mpi-sws.org

Nobutaka Ono
National Institute of Informatics
Tokyo, Japan
onono@nii.ac.jp

Shigeki Sagayama
Meiji University
Tokyo, Japan
sagayama@meiji.ac.jp

ABSTRACT

In this demo paper, we present an implementation of Harmonic/percussive sound separation (HPSS) and HPSS-based vocal separation as new effects in Audacity, which is a popular open-source sound processing tool.

1. INTRODUCTION

Harmonic/percussive sound separation (HPSS) [1–3] is a technique for decomposing a music signal into harmonic and percussive components. Applying HPSS twice with different frame sizes in the short-time Fourier transform (STFT) can provide another decomposition for voice (or melody) and other components [4, 5]. These decompositions have been applied to various music information retrieval (MIR)-related tasks such as chord detection, melody extraction, genre classification, and so on [6].

In this work, in order to enable a broader set of people to use the HPSS technique, we implemented the HPSS and HPSS-based vocal separation effects for *Audacity*, which is a popular open-source sound processing tool. In the demonstration, we will show how one can apply HPSS to various music tracks. An example of the HPSS effect is shown in Fig. 1.

2. IMPLEMENTATION

For our implementation, we used the algorithm described in [2]. The update rules of HPSS are applied within a sliding block defined in the STFT domain. In order to strike a good balance between processing speed and code readability, the sliding block analysis is implemented using the queue data type of the C++ STL. We added both HPSS and the HPSS-based vocal separation effect to Audacity as a new effect that appears in the list of effects a user can apply to a given audio signal. In the graphical user interface

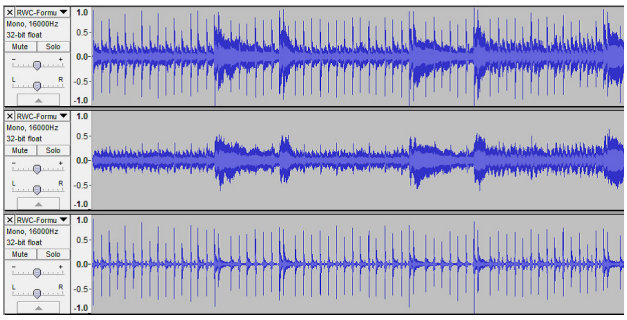

Figure 1. The screenshot of the original (top), harmonic (middle) and percussive (bottom) tracks in Audacity

(GUI) dialog box as shown in Fig. 2 and 3, a user can set the parameters of HPSS such as the frame size, the mask type (binary or Wiener), the output mode (keep the harmonic component only, keep the percussive components only, or, keep both components and create new tracks), and the final amplification factor.

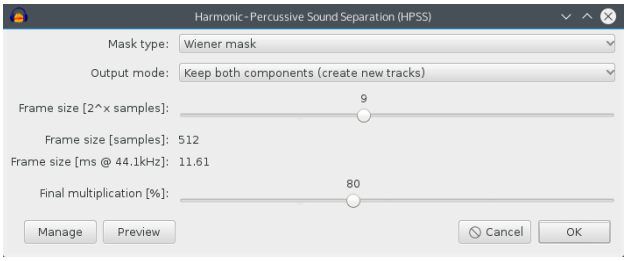

Figure 2. Parameter dialog for HPSS

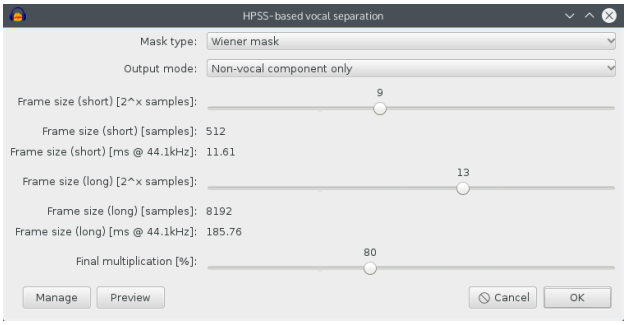

Figure 3. Parameter dialog for HPSS-based vocal separation

[1] work done while at the National Institute of Informatics

3. SUMMARY

Our contributions are as follows:

1. Readable and easy-to-maintain implementation of HPSS in C++
2. Audacity plugin using the new HPSS source code (tested on Windows and Linux)
3. The source code is published on GitHub [7] in the form of a patch against the development branch of the Audacity source code. We sent the patch to the Audacity developer team, and we hope that it will be included in a future public release of Audacity.

4. REFERENCES

[1] N. Ono, K. Miyamoto, J. Le Roux, H. Kameoka, and S. Sagayama, "Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram," in *Proceedings of the European Signal Processing Conference (EUSIPCO)*, IEEE, 2008.

[2] N. Ono, K. Miyamoto, H. Kameoka, and S. Sagayama, "A real-time equalizer of harmonic and percussive components in music signals," in *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, pp. 139–144, 2008.

[3] H. Tachibana, N. Ono, H. Kameoka, and S. Sagayama, "Harmonic/percussive sound separation based on anisotropic smoothness of spectrograms," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 2059–2073, 2014.

[4] H. Tachibana, T. Ono, N. Ono, and S. Sagayama, "Melody line estimation in homophonic music audio signals based on temporal-variability of melodic source," in *Proceedings of the International Conference on Aoustics, Speech and Signal Processing (ICASSP)*, pp. 425–428, IEEE, 2010.

[5] 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 Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 1, pp. 228–237, 2014.

[6] N. Ono, K. Miyamoto, H. Kameoka, J. Le Roux, Y. Uchiyama, E. Tsunoo, T. Nishimoto, and S. Sagayama, "Harmonic and percussive sound separation and its application to MIR-related tasks," in *Advances in music information retrieval* (Z. W. Ras and A. Wieczorkowska, eds.), pp. 213–236, Springer, 2010.

[7] V. Erdélyi, "HPSS source code as published on GitHub." `https://github.com/verdelyi/audacity`.