



AUN/SEED-Net



8th **AUN/SEED-Net** REGIONAL CONFERENCE ON ELECTRICAL AND ELECTRONICS ENGINEERING

Envision, Enable, and Empower
Smarter and Resilient Societies

co-located with

11th **ERDT Conference** on Semiconductor and Electronics, Information and Communications Technology and Energy

16-17 November 2015
Metro Manila, Philippines



**Proceedings of the 8th AUN/SEED-Net RCEEE 2015 and 11th ERDT Conference
on Semiconductor and Electronics, Information and Communications Technology, and Energy**

Editors:

Dr. Joel Joseph S. Marciano Jr.

Dr. Jhoanna Rhodette I. Pedrasa

Dr. Rhandley D. Cajote

© Copyright 2015 by the Electrical and Electronics Engineering Institute, College of Engineering, University of the Philippines Diliman, Engineering Research and Development for Technology, and ASEAN University Network/Southeast Asia Engineering Education Development Network (AUN/SEED-Net).

All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form without the consent of the editors of the Proceedings of the 8th AUN/SEED-Net RCEEE 2015 and 11th ERDT Conference on Semiconductor and Electronics, Information and Communications Technology, and Energy.

ISBN: 978-616-406-075-3

Published by: ASEAN University Network / Southeast Asia Engineering Education Development Network
(AUN/SEED-Net) JICA Project
Faculty of Engineering, Bldg. 2
Chulalongkorn University, Bangkok
Thailand 10330

Printed in the Philippines by: ERZALAN PRINTING PRESS
45 Cotabato Street, Luzviminda Village, Batasan Hills, Quezon City, Philippines

8th AUN/SEED-Net Regional Conference on Electrical and Electronics Engineering 2015

co-located with

11th ERDT Conference on Semiconductor and Electronics, Information and Communications Technology, and Energy

Envision, Enable and Empower Smarter and Resilient Societies

Published by: ASEAN University Network / Southeast Asia Engineering Education Development Network (AUN/SEED-Net) in partnership with Engineering Research and Development for Technology (ERDT) and University of the Philippines Diliman.

© Copyright 2015

No part of this publication may be reproduced without the consent of the editors of the Proceedings of the 8th AUN/SEED-Net Regional Conference on Electrical and Electronics Engineering 2015 and 11th ERDT Conference on Semiconductor and Electronics, Information and Communications Technology, and Energy.

ISBN: 978-616-406-075-3

STUDY OF HYBRID METHODS FOR MELODY EXTRACTION OF POLYPHONIC SIGNALS FOR PHILIPPINE INDIGENOUS MUSIC

Jason Edward D. Disuanco *, Vanessa H. Tan and Franz A. de Leon

Electrical and Electronics Engineering Institute, University of the Philippines Diliman, PHILIPPINES.

* E-mail: jasondisuanco@gmail.com

ABSTRACT

In this paper, we present a hybrid method for extracting the melody from polyphonic signals. Automatic melody extraction could help ethnomusicologists in music transcriptions and music classification. This is performed by blind source separation, followed by melody extraction.

Blind source separation (BSS) is an algorithm that separates a mixed audio track to its individual sources without any prior information of the track. A BSS algorithm that we considered is Harmonic-Percussive Source Separation (HPSS). In HPSS, the objective is to separate harmonic components from the percussive components of an audio track. Harmonic signals are defined as a sound with pitch. Hence, a song with purely harmonic instrument or voice has contours parallel in the time domain [1]. A problem arises when a percussive instrument is pitched and is mistaken as a harmonic component. This problem can be addressed by separating the harmonic from percussive components through the quality or timbre of the sound produced. An algorithm called Shifted-Non Negative Matrix Factorization (SNMF) clusters the frequencies of each source to extract the pitch of the percussive instrument from the pitches of the harmonic instrument.

We investigate two approaches for melody extraction: salience-based and data-driven. Salience-based approaches start with filtering the audio signal to enhance the frequency content where the melody could be found while attenuating distortions or noise [2]. After preprocessing, a spectral transform is applied to the signal. The two commonly used spectral transforms are the Short-Time Fourier Transform (STFT) and the Multi-Resolution Fast Fourier Transform (MRFFT). Next, a salience function estimates the prominent pitch values over time. The peaks of this function are taken as the possible candidates for the melody. The peaks from the salience function are then grouped as pitch contours or trajectories based on time, pitch and salience continuity [3][4]. Given the pitch contours, the melody is selected by using tracking techniques.

Data driven approaches use machine learning to train a classifier and estimate the melody directly from the power spectrum [2]. Different features are extracted from the audio track [5][6]. The features are then used to train a Support Vector Machine (SVM) classifier. The last stage for this approach is voicing detection. A global threshold based on the power spectrum is used for the detection [1][5].

We developed two hybrid methods for automatic melody extraction of polyphonic music. The first method includes a salience-based approach and a Harmonic Percussive Source Separation (HPSS) algorithm using STFT and MRFFT for spectral transforms. The second method includes the concept of the data driven approach and the Shifted Non-negative Matrix Factorization (SNMF) algorithm. The algorithms are implemented in MATLAB® and evaluated using samples from the Jose Maceda collection of the University of the Philippines Center for Ethnomusicology. Ten songs were separated through the HPSSSTFT, HPSS-MRFFT and SNMF. The separated tracks were evaluated through listening and determining the dominance of harmonic and percussive components in each separated signal. Results from subjective tests show that SNMF perform better than HPSS for harmonic and percussive source separation. Moreover, objective tests for melody extraction indicate that the salience-based approach has higher accuracy in identifying the melody than the data driven approach.

Keywords: audio signal processing, melody extraction, source separation.

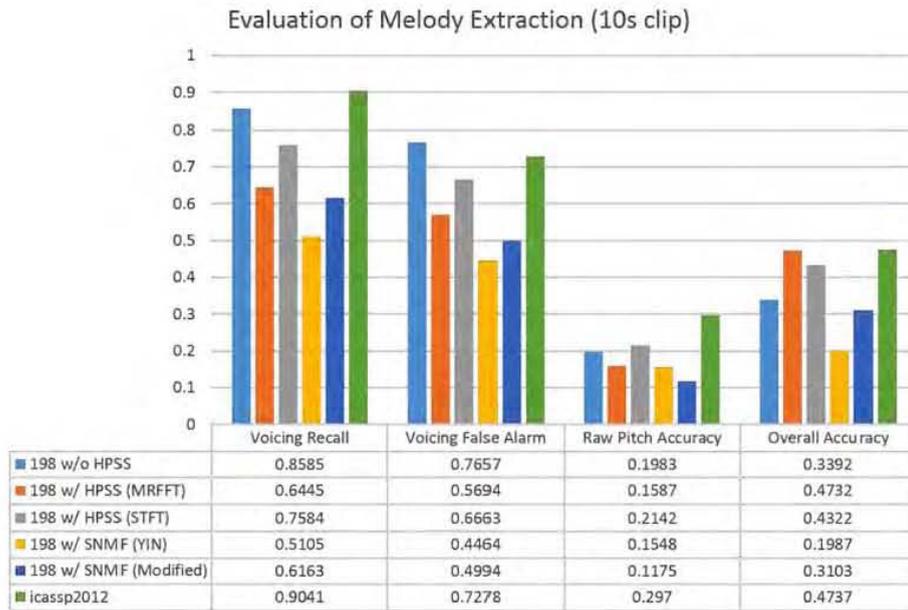


Figure 1. Evaluation of Melody Extraction based on HPSS

References

- [1] N. Ono, K. Miyamoto, J. Roux, H. Kameoka, and S. Sagayama, "Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram," in *Proc. of EUSIPCO*, 2008, pp. 1–4.
- [2] J. Salamon, S. Member, E. Gomez, D. P. W. Ellis, S. Member, G. Richard, and S. Member, "IEEE signal processing magazine 1 melody extraction from polyphonic music signals: Approaches, applications and challenges."
- [3] J. Salamon and E. Gomez, "Melody extraction from polyphonic music signals using pitch contour characteristics," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 6, pp. 1–12, 2012.
- [4] R. Paiva, T. Mendes, and A. Cardoso, "Melody detection in polyphonic musical signals: Exploiting perceptual rules, note salience, and melodic smoothness," *Computer Music Journal*, vol. 30, no. 4, pp. 80–98, Dec. 2006.
- [5] G. Poliner and D. Ellis, "A classification approach to melody transcription," in *Proc. 6th Int. Conf. on Music Inform. Retrieval*, no. 1, 2005, pp. 161–166.
- [6] Y. Uchida and S. Wada, "Melody and bass line estimation method using audio feature database," in *IEEE Int. Conf. Signal Processing, Communications and Computing (ICSPCC)*, 2011.