

An Onset Detection Algorithm by ODF Fusion

Chunta Chen

Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan
chun-ta.chen@mirlab.org

ABSTRACT

We present a framework for improving onset detection. It combines multiple onset detection functions (ODF) of different frame sizes, which are fused by alignment peaks of ODF in temporal order.

1. INTRODUCTION

Onset detection is commonly used as a crucial pre-processing in high-level music processing techniques. An onset is usually defined as the starting time of a sound event. The general procedure [1, 2] of onset detection is to compute an onset detection function (ODF) from the input audio and then pick local maxima from ODF as onsets. ODF is supposed to have larger change of value in the vicinity of a music onset.

2. ALGORITHM

We proposed a framework of ODF fusion by peak alignment. We regard the peaks of different ODFs within a scanning window as the same group. Then, for each group, we extract features around each peak and cascade these features to be a feature vector. We train a classifier to distinguish if it is an onset or not.

2.1 Peak picking and peak alignment

Peak picking is first applied to two ODFs yielding two lists of onset candidates. Onsets from the two lists occurring within a fixed temporal tolerance window will be accepted.

2.2 Feature extraction

Because the peaks of ODFs are aligned, we can extract features around these peaks and cascade the features of the same group. These features are classified into two labels, onset or not, and can be used to train a classifier by machine learning technique.

2.3 SVM training

We choose SVM to train a classifier for onset detection. The SVM derives its advantage from the optimization process, which maximizes the distance between decision boundaries. The parameters were optimized by using a grid search on the training set.

3. MIREX RESULTS

4. REFERENCES

- [1] J.P. Bello, L. Daudet, S. Abdallan, C. Duxbury, and M. Davies: "A tutorial on onset detection in music signals," *IEEE Transactions on Audio, Speech, and Language Processing*, volume 13, 2005.
- [2] S. Dixon: "Onset detection revisited," *Proc. of the 9th Int. Conference on Digital Audio Effects (DAFx'06)*, 2006.
- [3] A. Holzapfel, Y. Stylianou, A.C. Gedik, and B. Bozkurt: "Three dimensions of pitched instrument onset detection," *IEEE Transactions on Audio, Speech, and Language Processing*, 2010.
- [4] S. Böck and G. Widmer: "Maximum filter Vibrato Suppression for Onset Detection," *Proc. of the 16th Int. Conf. on Digital Audio Effects (DAFx)*, 2013
- [5] S. Böck and G. Widmer: "Local group delay based vibrato and tremolo suppression for onset detection," *Proc. of the 14th Int. Soc. for Music Information Retrieval Conference (ISMIR)*, 2013.