SOUND OBJECT TECHNOLOGIES: UHO ALGORYTHMS
FOR MIREX 2017

Adam Pluta, Rafał Rokicki, Marcin Gawrysz, Katarzyna Rokicka
Sound Object Technologies S.A.
adam.pluta@soundobject.com

ABSTRACT

This extended abstract presents our submission of several
algorithms based on sound object technology. The algo-
rithms operate on unique “Sound Objects” derived from a
specific process of signal filtration. An especially designed
bank of filters aims to extract information from input sig-

nal, process it sample by sample - preserving phase con-

tinuity - and represent it as sinusoidal sound objects, which
are further analyzed, grouped and transformed by the sub-

mitted algorithms.

Our submissions are presented and tested in the follow-
ing categories: Multiple Fundamental Frequency Estima-
tion and Tracking, Audio Key Detection and Audio Onset
Detection. All algorithms are written in C++ and originally
incorporated in the main software of Sound Object Tech-
nologies - Uho. The functionalities described in each task
are perceived as additional plugins of the overall sound de-
composition. For MIREX 2017 they are prepared as sepa-
rate applications serving the purpose of the test. The pro-
cess of signal filtration and examples of the sound objects
are presented at www.soundobject.com.

1. INTRODUCTION

Sound Object approach to sound processing and infor-
mation retrieval is based on an alternative method of sound
decomposition. As a result of a very precise signal filtra-
tion (500 filters in full frequency range which give a reso-

lution of 4 filters per semitone) we obtained the sound rep-

resentation consisting of editable, sinusoidal “sound ob-

jets” that can be grouped, analyzed and modified accord-

ing to their source. With all editable parameters of a sound
object (frequency, amplitude, position and phase) tracked
and preserved sample by sample, we obtain sound vectors
that transmit all the information of the sound occurrences.
The sound depiction achieved in such a process of signal
filtration allows for precise separation of frequencies over-
lapping-in-time.

The algorithms presented for the MIREX evaluation
are extensions of the sound decomposition and aim to
group and separate sound occurrences from their back-
ground. This year submission is oriented towards notes
identification and retrieving all the information that might
be preserved in the vectors. Therefore, all the submissions
will have a common ground from which they evolve.

2. SOFTWARE OPERATIONS

2.1 Preliminary preparation of sound objects

With the sound decomposition performed, the material
analysis begins with sound objects’ preparation. This
phase involves an algorithm which separates the back-
ground noise and unnecessary objects from the content
material needed for further analysis.

The algorithm in the first place retrieves information
about the drift of object frequencies, distinguishes between
homogenous sections and remembers them as units. Sec-
ondly, taking into account the energy of the objects it cor-
rects the drift of frequency resulting in new model objects
prone to further data analysis.

The next step of preparation involves combining ho-

mogenous objects and a correction of the envelope drift,
which leads to clarification of the overall analyzed mate-

rial.

In view of the fact that the current version of the algo-
rithm identifies slides and slurs as several pitches conse-
quent in time, the sources characterized by significant
slides in frequency might be misjudged. Consequently, the
algorithms presented for evaluation are not yet efficient in
transcribing vocals or soft, sliding instruments as well as
might encounter misinterpretations in an analysis of some
percussion instruments.

This document is licensed under the Creative Commons
Attribution-Noncommercial-Share Alike 3.0 License.
http://creativecommons.org/licenses/by-nc-sa/3.0/
© 2010 The Authors
2.2 Data analysis

The next step performed on the material is to gather the information about the objects analyzed – their onset position, energy, amplitude and the slope of their envelope. On the basis of this information another selection and grouping is performed, which leads to notes identification.

2.2.1 Beat Identification

From all the parameters provided, the algorithm begins with onset identification.

While remembering the position in time of each object selected, it also analyses its slope of envelope. This gives the basis for choosing the position of the strongly attacked objects, which consequently provides the “beat road map” to which all weaker objects are segregated. Each beat identified in such a way consists of set of objects belonging to its defined position. The distance between consecutive beats provides the tempo and their energy relations gives the basis for downbeat identification and division into measures.

The next step involving the beat analysis requires object identification with regard to their harmonic relations. In this moment the algorithm chooses the objects that create harmonic relations with other objects and ignores the remaining ones. The final selection is made.

2.2.2 Harmonic grouping

The final step of the software analysis is concentrated upon harmonic relations among the objects belonging to one beat.

On the basis of frequency and amplitude comparison the objects are categorized and bound together into groups – preliminary notes. Then, all the created groups are analyzed from the point of view of their amplitude relationships and undergo an assessment of a decision tree, which determines the frequency of the actual fundamental occurrences – whether the group represents a simple note, a note with a virtual fundamental or numerous harmonic coexisting notes.

Finally, with the objects organized into musical events, the data base of all object parameters is ready for applications. Together with its all statistics it provided the basis for all MIREX 2017 submissions.

As our algorithms are still work in progress, we provide two versions of the software operation as two separate submissions for the Multiple Fundamental Frequency Estimation and Tracking as well as Audio Onset Detection tasks. The difference concerns priorities and directions of the algorithm operations while sorting frequencies as well as it involves a degree of amplitude analysis. The 902 version of the algorithm starts from energetically stronger objects while analyzing harmonic components of each chosen note as well as it searches for coexisting notes within one group. The version 901 does not have this function, which results in fewer notes chosen, as well as it starts its analysis from the lowest frequency groups and moves higher up in registers.

3. MIREX 2017 SUBMISSIONS

3.1 Multiple Fundamental Frequency Estimation and Tracking

The application created for the submission uses the data base of analyzed object parameters as basis for note results. It lists all note occurrences selected by the software in the process of note creation within certain period of time.

3.2 Audio Key Detection

The application for Audio Key Detection submission treats the note statistics provided by the software as a ground for its decision making. With certain mistake tolerance, it employs comparison of the list of note occurrences to a model key data as well as incorporates another decision tree to distinguish between more ambiguous key representations.

3.3 Audio Onset Detection

This submission is the representation of the all the listed onsets of the groups identified as notes.

4. REFERENCES