

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/266452273>

# Onset detection by means of transient peak classification in harmonic bands

Article · October 2009

---

CITATIONS

4

---

READS

17

1 author:



[Axel Roebel](#)

Institut de Recherche et Coordination Acoustique/Musique

146 PUBLICATIONS 1,113 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



ROUTE (RObot à l'écoUTE) [View project](#)



ChaNTeR [View project](#)

# ONSET DETECTION BY MEANS OF TRANSIENT PEAK CLASSIFICATION

A. Röbel

IRCAM-CNRS –STMS

1, pl Igor-Stravinsky

75004 Paris, France

roebel(at)ircam.fr

## ABSTRACT

The extended abstract describes an onset detection algorithm that is based on a classification of spectral peaks into transient and non-transient peaks and a statistical model of the classification results to prevent detection of random transient peaks due to noise. Compared to the version used for MIREX 2007 and 2009 this algorithm features 2 improvements. The detection of harmonic transients and the use of a slight modification the detection algorithm that avoids fragmentation of onsets as may happen especially when small precursors of the onset are present in the audio.

## 1. INTRODUCTION

In the following article we are going to describe a transient detection algorithm that has been developed for a special application, the detection of transients to prevent transformation artifacts in phase vocoder based (real time) signal transformations [9, 10]. This application requires a number of special features that distinguishes the proposed algorithm from general case onset detection algorithms: The detection delay should be as short as possible, frequency resolution should be high such that it becomes possible to distinguish spectral peaks that are related to transient and non transient signal components, for proper phase reinitialization the onset detector needs to provide a precise estimate of the location of the steepest ascend of the energy of the attack. In contrast to this constraints the application does not require the detection of soft onsets, where a soft onset is characterized by time constants equal to or above the length of the analysis window. This is due to the fact that such onsets are sufficiently well treated by the standard phase vocoder algorithm. False positive detections are not very problematic as long as they appear in noisy time frequency regions. A major distinction is that a single onset may be (and very often is) composed of multiple transient parts, related either to a slight desynchronization of polyphonic onsets or due to sound made during the prepara-

tion of the sound (gliding fingers on a string). While these desynchronized transients are generally not considered as independent onsets they nevertheless constitute transients which should be detected for the intended application.

The evaluation of the transient detection algorithm for onset detection has been evaluated repeatedly in the MIREX evaluation campaigns 2005, [5], 2006 [6], 2007 [7], and 2009 [8] and it has shown very good performance at least in the last 3 evaluations. The analysis of the performance with respect to onset and instrument classes shows that one of the problems of all algorithms is related to the detection of onsets of pitched instruments. Accordingly we have worked on this problem and present here the results of the work. An investigation of the results of the previous algorithms has shown that in some situations individual onsets are fragmented resulting in a reduced attack strength which will in turn have a bad impact on our transient filtering procedure.

## 2. FUNDAMENTAL STRATEGY

There exist many approaches to detect attack transients. For a number of current approaches see [2–4, 13] as well as all algorithms presented in the MIREX campaigns mentioned above. Most of the known algorithms define an onset detection function that is evaluated in different frequency bands. Here we use a similar approach using as detection function a statistical measure related to the time offset (time reassignment) [1] of individual spectral peaks in the standard DFT spectrum. Using a simple threshold for the time reassignment we classify spectral peaks into transient and non transient peaks [9, 10] and use as detection function the change in the transient peak probability in the different spectral bands. The advantage of the implicit peak classification is the fact that for each detected transient we have a precise measure of the time frequency location of the related transient.

The basic idea of the proposed transient detection scheme is straightforward. A peak is detected as potentially transient whenever the center of gravity (COG) of the time domain energy of the signal related to this peak is at the far right side of the center of the signal window. Note, that it can be shown [11] that the COG of the energy of the time signal and the normalized energy slope are two quantities with qualitatively similar evolution and, therefore, the use

of the COG of the energy for transient detection instead of the energy evolution appears to be of minor importance.

### 3. FROM TRANSIENT PEAKS TO ONSETS

Unfortunately not every spectral peak detected as transient indicates the existence of an onset. Further inspection reveals that spectral peaks related to noise signals quite often have a COG far of the center of the window. In contrast to spectral peaks related to signal onsets these false transient peaks in noise are not synchronized in time with respect to each other. This synchronization of a sufficient number of transient peaks is the final means to avoid detection of noise peaks as onsets.

To keep this abstract brief we will not describe the details of the statistical model, and we refer to the description of the first mirex evaluations for further details [11, 12].

### 4. PITCHED TRANSIENTS

The onset detection algorithm that is presented here is based on the detection of multiple synchronous events in the detection bands. The bands that have been used in the previous versions of the algorithm were always covering continuous frequency regions. In a polyphonic setting this band organization is a drawback for soft pitched onsets, because these onsets will be characterized by synchronous partials distributed over a large and non continuous frequency band. This systematic problem can be countered easily by means of allowing non continuous observation bands. In the present case we consider observation bands that are formed by a collection of bands with harmonically related center frequencies and a common bandwidth additionally to the continuous bands that have been used before. The level of confidence of the change in transient peak probability that is required for the detection of a transient event in the non continuous bands can be selected independently of the confidence that is required for the continuous bands. This should allow us to configure the algorithm for different types of sound signals.

### 5. AVOIDING ONSET FRAGMENTATION

The algorithm compares the distribution of the COG of spectral peaks in the current time segment to the distribution in a history. Whenever the the distribution of the COG shift to higher values an onset is detected. In the previous version the time distribution of the peaks having large COG in the current time segment has not been taken into account. In the new version (marked with keepsv) the centroid of the bins presenting transient activity will be calculated and a reset of the transient is deferred to after the centroid. Like this a transient event will less likely be fragmented into multiple independent transients with relatively weak attack strength.

### 6. DIFFERENCES IN THE 4 SUBMITTED ONSET DETECTION ALGORITHMS

The submissions use 4 different configurations that can be produced with the 2 algorithmic extensions: harmonically related transient evaluation bands, and deferring transient resets.

- *roebel\_onset\_noharm* does not make use of any extensions,
- *roebel\_onset\_harm* uses extended set of evaluation bands including harmonically related bands,
- *roebel\_onset\_noharm\_keepsv* uses the deferred transient reset mechanism, and
- *roebel\_onset\_harm\_keepsv* uses both the deferred reset algorithm as well as extended harmonically related evaluation bands.

### 7. REFERENCES

- [1] F. Auger and P. Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method. *IEEE Trans. on Signal Processing*, 43(5):1068–1089, 1995.
- [2] J. Bonada. Automatic technique in frequency domain for near-lossless time-scale modification of audio. In *Proceedings of the International Computer Music Conference (ICMC)*, pages 396–399, 2000.
- [3] C. Duxbury, M. Davies, and M. Sandler. Improved time-scaling of musical audio using phase locking at transients. In *112th AES Convention*, 2002. Convention Paper 5530.
- [4] P. Masri and A. Bateman. Improved modelling of attack transients in music analysis-resynthesis. In *Proceedings of the International Computer Music Conference (ICMC)*, pages 100–103, 1996.
- [5] Mirex audio onset detection evaluation results. <http://www.music-ir.org/evaluation/mirex-results/audio-onset/index.html>, September 2005. ISMIR 2005, London, Great Britain.
- [6] Mirex audio onset detection evaluation results. [http://www.music-ir.org/mirex2006/index.php/Audio\\_Onset\\_Detection\\_Results](http://www.music-ir.org/mirex2006/index.php/Audio_Onset_Detection_Results), October 2006. ISMIR 2006, Victoria, Canada.
- [7] Mirex audio onset detection evaluation results. [http://www.music-ir.org/mirex/2007/index.php/Audio\\_Onset\\_Detection\\_Results](http://www.music-ir.org/mirex/2007/index.php/Audio_Onset_Detection_Results), September 2007. ISMIR 2007, Vienna, Austria.
- [8] Mirex audio onset detection evaluation results. [http://www.music-ir.org/mirex/wiki/2009:Audio\\_Onset\\_Detection\\_Results](http://www.music-ir.org/mirex/wiki/2009:Audio_Onset_Detection_Results), September 2009. ISMIR 2009.

- [9] A. Röbel. A new approach to transient processing in the phase vocoder. In *Proc. of the 6th Int. Conf. on Digital Audio Effects (DAFx03)*, pages 344–349, 2003.
- [10] A. Röbel. Transient detection and preservation in the phase vocoder. In *Proc. Int. Computer Music Conference (ICMC)*, pages 247–250, 2003.
- [11] A. Röbel. Onset detection in polyphonic signals by means of transient peak classification. <http://www.music-ir.org/evaluation/mirex-results/articles/onset/roebel.pdf>, September 2005. ISMIR 2005, London, Great Britain.
- [12] A. Röbel. Onset detection in polyphonic signals by means of transient peak classification. [http://www.music-ir.org/evaluation/MIREX/2006\\_abstracts/OD-roebel.pdf](http://www.music-ir.org/evaluation/MIREX/2006_abstracts/OD-roebel.pdf), October 2006. ISMIR 2005, London, Great Britain.
- [13] X. Rodet and F. Jaillet. Detection and modeling of fast attack transients. In *Proc. Int. Computer Music Conference (ICMC)*, pages 30–33, 2001.