

# ESTIMATION OF TEMPO, MICRO TIME AND TIME SIGNATURE FROM PERCUSSIVE MUSIC

*Christian Uhle*

*Juergen Herre*

Fraunhofer

Fraunhofer

Arbeitsgruppe Elektronische Medientechnologie  
uhle@emt.iis.fhg.de

Institute for Integrated Circuits IIS-A  
hrr@iis.fhg.de

## ABSTRACT

Tempo, micro time and time signature are basic rhythmic features for a wide range of music. Their automated extraction from music is needed for a number of applications, e.g. automatic transcription and music information retrieval systems, authoring tools, entertainment applications, musicological analysis or educational purposes. The presented analysis procedure estimates tempo, micro time and time signature from percussive music. From 117 excerpts of eight seconds length, 84.6% of the tempo values, 83.8% of the micro time and 73.5% of the time signatures were estimated correctly.

## 1. INTRODUCTION

A large amount of research is concerned with the automatic extraction of meta data from musical audio signals. The goal in doing so is twofold: equipping machines with the skills of a human music listener to analyze a piece of music and to derive descriptions of certain musical properties. This opens the door for a number of applications, such as automatic transcription, play-list generation and music information retrieval systems. Additionally, the understanding of the human experience and processing of music is of great interest for researchers in the field of psychology and musicology.

The musical properties regarded in this work belong to the rhythm domain. Tempo, micro time and time signature are basic rhythmic features, and musical pieces can be partly characterized by them. While tempo and time signature are well-established terms among musicians and non-musicians, the term micro time is not. The remainder of this work contains the definitions of the musical terms used herein, an overview of related work, a description of the proposed processing steps, results and conclusions.

## 2. MUSICAL BASICS AND DEFINITIONS

The definition of rhythmic terms is a challenging task, since, as Cooper et.al. point out, “the development of a fruitful approach to the study of rhythm has been hampered by a failure to distinguish clearly among the several aspects of temporal organization itself. The resulting confusion has created a correlative ambiguity of terminology” [1]. According to [2], two components of rhythmic organisation exist, namely grouping and meter. A listener intuitively groups the events contained in a

musical audio signal into distinct units such as motives, themes, and phrases. In parallel, he or she constructs a regular pattern of strong and weak beats to which he or she relates the actual musical sounds. Meter can be defined as the measurement of the number of pulses (whereas a pulse is defined as one of a series of equidistant identical stimuli) between more or less regularly recurring accents. In this sense, it is not limited on the metrical level designated by the time signature and measured by bar lines, because different pulse series occur on different hierarchic levels, called rhythmic strata. Pulses are seldom heard and played in a pure state as a series of undifferentiated events, they are established and supported by phenomenal accents, and the sense of pulse may exist subjectively [1]. The rate of a pulse series describes the frequency of occurrence and the period equals the distance between two subsequent pulses. Figure 1 shows a rhythmic patterns in a simple pattern notation (where distances in the notation between events reflect their temporal distances), which may be played e.g. by a claves player in Brazilian music, and pulse series in different hierarchical levels, ranging from sixteenth to halve notes (for four-four time). Among these strata is the beat level, in this case quarter notes. The beat is the pulse which determines the musical tempo. Another common term for the pulse rate on the beat level is foot-tapping rate. The lowest metrical level is the tatum, a term coined by Bilmes [3]. The tatum may be established by the smallest time interval between two successive notes, but is in general best described by the pulse series that most highly coincidences with all note onsets. Pulses at bar lines constitute a pulse series on a higher level. Although the appearance of more pulse levels is the usual case in a musical piece, these three pulse levels (beat, tatum and bar) are more commonly used than others for the description of rhythmic structures. The rates and periods of the adjacent pulse levels show integer ratios. The term micro time is adopted from Marron [4] for the description of the ratio between beat and tatum period. For a discussion about rhythmic structures, it is advantageous to use distinct terms for the meter on different metrical levels, especially because the time signature is a common used term to describe the meter on the bar level.

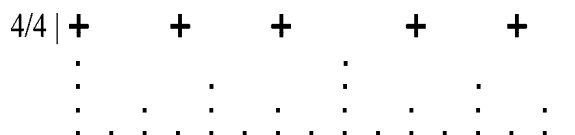


Figure 1: Simple rhythmic pattern (crosses) and underlying pulse levels (dots)

Rhythms can be classified into the categories additive, divisive, meso-periodic and hybrids of the first three [5]. Divisive rhythms are constructed hierarchically such that pulses are divided into parts, usually halves or thirds, to form the lower pulse level. This principle can be found in almost all western music. Additive rhythms are constructed through the concatenation of small rhythmic patterns with a constant pulse. Examples can be found in Turkish, Greek and Eastern European music. The term “meso-periodic” has been adopted from Schloss and describes a phenomenon which can be found in e.g. West-African music. There, a pattern of typically one or two measures length delivers the guideline for the other musicians, comparable to a metronome in western music. This guideline allows the production of polyrhythmic patterns, which is a specific characteristic for West-African music.

### 3. RELATED WORK

Scheirer [6] conducted a psychoacoustic demonstration on beat perception, where a white noise signal was vocoded with sub-band envelopes of a musical signal, i.e. the time-frequency envelopes of the music signal were transferred to the noise signal in sub-bands. Using only a small number of frequency sub-bands, the rhythmic information of the musical signal became evident. He furthermore proposed a beat tracking system (BTS), where the audio signal is split into a small number of frequency bands. The extracted energy envelopes of the band signals are fed into a bank of comb filter resonators. The summed output of the resonators feature peaks at prominent pulse periods.

Goto and Muraoka [7] described a BTS for pop music with drums. It comprises the detection of note onset times and of bass drum and snare drum hits. The detected drum patterns are matched with pre-registered drum patterns in parallel. Their computationally intensive system performs in real-time on a fast parallel computer. In a more recent publication [8], a system is presented that recognizes a hierarchical musical structure of three rhythmic levels in audio signals without drums.

The issue of beat-tracking of performances with expressive timing is addressed by Dixon [9]. The described system detects the most salient note onset times in the audio signal and clusters the inter-onset intervals (IOI), yielding multiple hypotheses for the underlying tempo. For each hypothesis a group of beat tracking agents is created and compared based on three factors: how evenly spaced the chosen beat times are, how often the beat times match the note onset times and their salience.

An off-line algorithm for the estimation of time-varying tempo is described in [10]. An “energy flux” is extracted from the audio signal using a Short Time Fourier Transform (STFT) of small non-overlapping blocks of the audio signal. The magnitudes of the output of the STFT are compressed, a first-order difference function is calculated and summed over all bins in a distinct frequency range and a half-wave rectification is applied. The obtained energy flux is compared to an expected energy flux for all possible tempos and downbeat locations, i.e. a series of equally spaced discrete pulses, yielding a set of candidates for the tempo and downbeat locations and a score obtained from the cross-correlation between calculated and expected energy flux. The final result is identified in a dynamic programming stage, which evaluates the scores of each candidate and a score reflecting the transition cost between the frames.

Gouyon et. al. [11] estimated the tatum period from previous detected note onsets. The most frequent IOI is detected in a histogram of IOIs. The calculation of the IOIs is not limited to successive onsets but rather to all pairs of onsets in a time frame. Tatum candidates are calculated as integer divisors of the most frequent IOI. The candidate is selected which best predicts the harmonic structure of the IOI histogram according to an error function. The estimated tatum period is subsequently refined by calculating the error function between the comb grid derived from the tatum period and the onset times of the signal. Seppänen [12] proposed an IOI-based tatum estimation with a time-varying histogram representation in order to accommodate slight tatum changes (e.g. *accelerando* and *retardando*).

The determination of musical meter from audio recordings is addressed in [13] and [14]. Seppänen's method [13] starts with the detection of note onsets from amplitude envelopes of multiple frequency bands. The tatum grid is estimated and a metrical ground, which is an artificial metrical level is derived from the tatum grid. While the tatum may hop from one metrical level to another, the metrical ground does not. The acoustic signal in the neighborhood of each tatum grid point is then characterized by a single normalized value representing the total phenomenal accentuation using a supervised pattern recognition framework. The beat grid estimator is based on probability calculation for finding the most probable periodicity from the phenomenal accent data. Gouyon et. al. [14] investigated various low-level features and their application with an autocorrelation based processing for the determination of meter. Using a set of four descriptors 95% of correct classifications were reached for the classification of the meter into one of the two classes “binary meter” and “ternary meter”.

Klapuri [15] proposed a method for the estimation of tatum, beat and measure length and their phase. Envelope signals are the input to a bank of comb-filter resonators, and a probabilistic model is applied for the interpretation of the detected periodicities. Prior probabilities for tatum, beat and measure were estimated from a set of data. He furthermore proposed a joint estimation of tatum, beat and measure length and compared his results with results obtained with systems by other other authors in their original implementation.

### 4. SYSTEM DESCRIPTION

The general idea of the proposed method is to improve the recognition of the three rhythmic features with a combined estimation, which exploits the mentioned (nearly) integer ratios between them. The processing starts with the decomposition of segments of four seconds length of the audio signal into frequency bands. The envelopes are extracted from each band of the filtered signal. Note onsets are detected from high-pass filtered and half-wave rectified amplitude envelopes. Subsequently, the IOIs are calculated from the note onset times and the tatum period is estimated from an IOI histogram. In the subsequent processing, periodicities in the temporal progression of the amplitude envelopes are detected with an autocorrelation method. The peaks in the autocorrelation function correspond to the time lags where the signal is most similar to itself. Beat period candidates are identified as peaks in the autocorrelation function (ACF), which correspond to (close to) integer multiples of the tatum periods. Similarly, bar length

candidates are detected as peaks corresponding to (close to) integer multiples of the beat period candidates. For each bar length candidate, a score is calculated from the degree of correlation of two subsequent excerpts of the envelopes per band, whereas the length of both excerpt are identical with the bar length candidate. The envelopes of two segments are accumulated beforehand, so that the analysis of a bar length up to four seconds length is possible. This estimation is motivated by the assumptions, that self-similarity exist on the tatum, beat and the bar level. Figure 1 gives an overview of the processing steps.

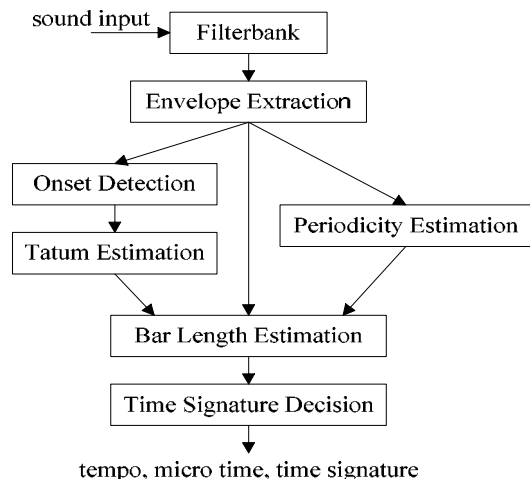


Figure 2: System overview.

#### 4.1. Spectral Decomposition and Envelope Extraction

For the sake of computational efficiency the incoming PCM audio signals are converted to a sampling frequency of 22050 kHz and in the case of stereo signal down-mixed to mono. The input signal is decomposed into a small number of frequency bands using an IIR filter bank. Seven frequency bands were chosen in a range between 62 and 8000 Hz with adjacent borders at 125, 250, 500, 1000, 2000 and 4000 Hz. The current implementation employs sixth-order elliptic filters with 0.5 dB ripple in the pass-band and a minimum stop-band attenuation of 60 dB. Subsequently, envelope signals are generated through a rectify-and-smooth operation for each frequency band. The outputs of the filter-bank are full-wave rectified, then decimated by a factor of 43 and low-pass filtered with a fourth-order elliptic filter with a cut-off frequency at 10 Hz, 3 dB ripple and 60 dB attenuation, yielding the envelope signals  $e_b(t)$ , with  $b$  denoting the band index. Figure 3 shows a normalized envelope signal from the third frequency band over a period of four seconds.

#### 4.2. Note Onset Detection

Note onsets are detected for a subsequent estimation of the tatum periods. Several algorithms for note onset detection examine the slope of the envelope [3][5][16]. The slope  $s_b(t)$  is derived via high-pass filtering with a customized IIR Filter of first order and half-wave rectification of the envelope signals  $e_b(t)$ . Figure 3 illustrates the signal, envelope and slope signals for the third

frequency band. Note onset candidates for each band are calculated from the  $s_b(t)$  using a peak picking method. Subsequently, candidates which are closer than 50ms to a candidate with a larger value of  $s_b(t)$  are eliminated.

Some of the remaining candidates may occur because of non-monotonically increasing or decreasing envelope signals. In audio recordings with transient signals and a strong degree of reverb, first reflections may cause spurious peaks in  $s_b(t)$ . Two additional conditions have to be satisfied by an onset candidate to be accepted. The first condition is that  $s_b(t)$  exceeds a dynamic threshold generated by a decay function from a value which is set on the detection of a preceding note onset to half of this value after approximately 200ms. This is motivated by the assumption that in a musical situation a change from fortissimo to pianissimo does not happen instantly. The second condition is that  $e_b(t)$  drops between two subsequent note onsets below a value of 75% of the minimum value of the two peaks of  $e_b(t)$  corresponding to the onset times.

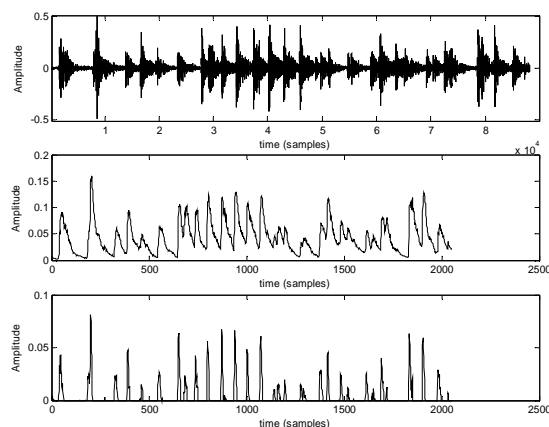


Figure 3: Signal (upper plot), envelope signal and slope signal (lower plot) for the third band

#### 4.3. Tatum Period Estimation from Note Onset Times

The note onset times are passed to the tatum estimation module, which estimates the smallest rhythmic pulse from a histogram of inter-onset intervals. Its implementation closely follows the approach proposed in [11] with slight modifications. Inter-onset intervals are calculated between all pairs of notes if less than a predefined number of onsets occur between these notes and if the temporal distance between the two notes is below a predefined threshold. The IOI histogram is smoothed by a convolution with a Gaussian function with a standard deviation of 9 ms. Peaks are detected in the smoothed IOI histogram and the most frequent IOI (MFIOI) is obtained as the IOI corresponding to the maximum value of the histogram. Candidates for the tatum period are calculated as the time spans which are integer divisors of MFIOI and exceed a predefined minimum value for the tatum which is set to 75 ms. The time span which corresponds to 66% of the MFIOI is added to the candidates. For each tatum candidate a harmonic comb grid is generated and a two-way mismatch error function (TWM) is computed between the grid elements and the histogram peaks. The first error measures the

deviations between the comb grid and the nearest histogram peak, whereas the second error represents the deviations between each histogram peak and the nearest comb grid element. These two functions are linearly combined to form the TWM, the best tatum period candidate corresponds to its minimum value. A final refinement and a phase estimate (which is not used here) are achieved by applying the TWM in the time domain to a comb grid corresponding to different tick candidates and the detected note onsets.

#### 4.4. Periodicity Estimation from Envelope Signals

The periodicity estimation kernel is derived from the concept of the Meddis/Hewitt model for pitch perception by applying it to the task of beat tracking [17]. The ACF of the averaged slope signal  $s_b(t)$  is computed in the range of interest between 0.1 and 4 seconds (corresponding to 15 – 600 bpm). This is done for consecutive analysis windows of four seconds length. An efficient way of implementation employs an FFT and its inverse (IFFT) of the square of the amplitudes of the complex spectral coefficients.

The ACF shows peaks at lags corresponding to period lengths of pulse series or to periods between note values played in an accentuated way, which may appear as single events or parts of larger patterns. Because the ACF is biased, a trend correction is performed. A line that fits the ACF in a least square sense is calculated using linear regression and subtracted from the ACF. All peaks of the trend-corrected ACF (TCACF) are detected with a three-point-running window method. Multiple peaks may occur between zero-crossings of the TCACF if no distinct rhythm is apparent in the signal, rapidly varying tempo or not precisely played notes. For each area between two zero-crossings, the largest peak is recognized as pulse period candidate  $PC$ . All  $PC$ s and their corresponding intensities  $PI$  are stored for further processing.

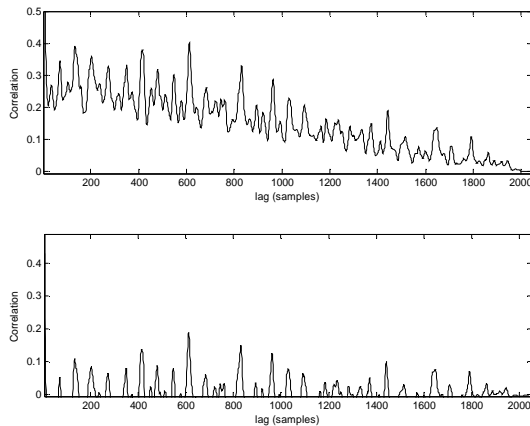


Figure 4: Outputs of the ACF before (upper plot) and after trend correction

A measure for the beat strength is computed and used for the estimation of the bar length period, there for the joint analysis of two or more segments one tatum period is chosen. The derived beat strength serves as a criterion for the selection of the tatum

period among the tatum estimates of each segments (see Section 4.5). The measure for the beat strength  $m_k$  is derived from the number of peaks  $n_k$  and of zero-crossings  $z_k$  in the TCACF and its maximum value for each segment  $k$ .

$$m_k = \frac{z_k}{p_k} \cdot \max(\text{TCACF}_k) \quad (1)$$

The beat strength measure differs from the two proposed beat strength measures in [18], where the beat strength of a longer musical piece is estimated from a “beat histogram”, which is derived from the accumulation of ACFs of subsequent analysis windows. The evaluation of the ratio  $z/p$  serves to indicate segments which are played exactly in the sense of a (nearly) constant tempo on all pulse levels, which is a prerequisite for a reliable tatum estimation.

#### 4.5. Bar Length Estimation

The estimation of the bar length proceeds after a periodicity estimation of at least two subsequent analysis windows is finished. The  $PC$ s and  $PI$ s for both excerpts are gathered in a joint histogram, so that  $PI$ s of identical  $PC$ s are summed up. In each of the analysis segments, a tatum period was estimated. For the following processing, the tatum is chosen, for which the former calculated beat strength measure  $m_k$  is maximized.

In the first step of the bar length estimation, all (nearly) integer multiples of the tatum period among the  $PC$ s are identified as hypotheses for the beat period. For each beat period candidate, a hypothesis about the bar length is obtained, if a (nearly) integer multiple of the beat period is among the  $PC$ .

The criteria for the decision about the bar length is the cross correlation of two subsequent excerpts whose lengths correspond to the bar length candidate. For each frequency band, the cross correlation  $c_i(\tau)$  of the two excerpts is calculated according to

$$c_b(\tau) = \sum_{n=0}^{N-1} s_{1,b}(n) s_{2,b}(n+\tau) \quad (2)$$

where  $s_{1,b}(n)$  and  $s_{2,b}(n)$  denoting the two envelope excerpts for one frequency band  $b$ . The maximum value within a small area surrounding the middle of  $c_b(\tau)$  represents the best correlation of the two vectors,  $s_{1,b}(n)$  and  $s_{2,b}(n)$ . The small area is chosen with respect to tempo variations in the audio signal and inaccuracies in the estimation of the bar length candidates. The decision criterion is computed as the sum of the weighted best correlation values. The best band-wise correlations are weighted, so that lower frequencies influence the decision criteria more than upper ones. This is motivated through the observation, that low sounding instruments are often played with less variations, and improves the recognition performance in the final experiments. The weights  $w_b$  for each frequency band were manually adjusted.

Subsequently, the estimated bar length was corrected in the following manner. In cases, where the ration of bar length and tatum period is larger than a predefined number, a smaller bar length candidate with a corresponding next best criterion were chosen, and likewise, short bar length are replaced with the next best longer candidate. This is motivated through the assumption, that common values for micro time and time signature are in the

range of 2...8, and that in these cases pulse period on a higher respectively lower level was more salient.

In case of additive rhythms, strong correlation is obtained for the bar length as well as for the lengths of the smaller "atoms" of the rhythm. To improve the robustness of the estimation of the bar length, an algorithm for the identification of additive rhythms is implemented. The period lengths of the salient bar candidates are evaluated, and if one period length among them is equivalent to the sum of two smaller ones of salient candidates, the item is identified as additive and the larger beat candidate is chosen.

#### 4.6. Decision about the Time Signature

After the estimation of the bar length, still multiple hypotheses about tempo, micro time and time signature may exist, i.e. various combinations of micro time and tempo are possible for a specific bar length. In these ambiguous cases, a criterion for the decision among the remaining beat candidates is calculated from the corresponding values from the ACF, which are weighted with probability values for the beat period. Prior probabilities for beat period length are reported in e.g. [13]. For the special case of percussive music, the probabilities used here were estimated from test data.

The decision about the beat period length is problematic especially in cases where the bar length is identified to correspond to 12 or 24 tatum periods. Here, a six-eight time with binary micro time, or a four-four time with ternary micro time are identified.

### 5. RESULTS

Performance of the algorithm was tested on a database with 117 excerpts of percussive music of 8 seconds length. The test data included music of various origins and cultures, i.e. West-African, Brazilian and Japanese folkloristic music and solo drum-set performance. The reference values for tempo, time signature and micro time were gathered by one person with musical background and in ambiguous cases average values of the judgments of up to four amateur musicians were taken. It turned out, that, especially for West-African music, an assignment of the rhythmic features is often ambiguous. This is not surprising, because at least two rhythms are developed in parallel in some of the test items [20].

For the evaluation of the tempo estimates, deviations up to 4% were tolerated. 84.6% of the tempo values, 83.8% of the micro time and 73.5% of the bar length were estimated correctly. For 5.1% of the test items the double, and for 2.6% the half of the reference tempo was estimated. A ternary four-four time or a binary six-eight time were the reference value for 11.9% of all test items. For 14.2% of these items (which equals 1.7% of all items), the rhythmic features were estimated wrongly because of the confusion of binary six-eight time and ternary four-four time. The results for these items are acceptable, in the sense that the determination of the reference values was ambiguous too.

After closer examination of the results and the test items, different sources of error became evident. The appearance of solo voices decreases the recognition performance. If a drastic change in the musical situation occurs, the estimation of the bar length fails. A drastic change may be the entry of an instrument, the change of a prominent pattern or the transition between two

different parts. Additionally, the estimation of the tatum period may fail in cases where the tempo changes are distinct or polyrhythm occur.

### 6. CONCLUSIONS

The combined estimation of tempo, tatum and bar length is motivated by the simple relation between them and improves the robustness of the system. Still, automatic estimation of rhythmic features does not reach the performance of a human listener. The recognition results strongly rely on the robustness of the tatum estimation. A prior segmentation for the identification of regions with drastic changes in the musical situation seems to be necessary. These segment boundaries are well suited for the definition of the beginning of an analysis window.

A separation of the single instrument tracks is conceivable and would increase the robustness of the estimation regarding solo playing and polyrhythmic. Another improvement of the recognition results is expected for the analysis of longer excerpts of a musical piece and a subsequent evaluation of the results of the single segments.

The evaluation of the proposed method on a larger database of test items is required for the acquisition of more reliable recognition results.

### 7. REFERENCES

- [1] Grosvenor Cooper and Leonard B. Meyer, *The Rhythmic Structure of Music*, University Of Chicago Press, 1960.
- [2] Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music*, MIT Press, 1983.
- [3] Jeffrey A. Bilmes, *Timing Is Of Essence*, MSc Thesis, Massachusetts Institute Of Technology, 1993.
- [4] Eddy Marron, *Die Rhythmiklehre*, (in German), AMA Verlag, 1990.
- [5] Walter A. Schloss, *On the Automatic Transcription of Percussive Music*, PhD Thesis, Stanford University, 1985.
- [6] Eric D. Scheirer, "Tempo and Beat Analysis of Acoustic Musical Signals", *Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 588-601, 1998.
- [7] Masataka Goto and Yoichi Muraoka, "A Beat Tracking System for Acoustic Signals of Music", in *Proc. of the Second ACM Intl. Conf. on Multimedia*, pp. 365-372, 1994.
- [8] Masataka Goto and Yoichi Muraoka, "Real-Time Rhythm Tracking for Drumless Audio Signals-Chord Change Detection for Musical Decisions", in *Proc. of the Workshop on Computational Auditory Scene Analysis- Int. Joint Conf. on Artificial Intelligence*, pp. 135-144, 1997.
- [9] Simon Dixon, "Automatic Extraction of Tempo and Beat from Expressive Performances", in *J. New Music Research* vol. 30, no. 1, pp. 39-58, 2001.
- [10] Jean Laroche, "Efficient Tempo and Beat Tracking in Audio Recordings", in *J. Audio Eng. Soc.*, Vol. 51, No. 4, pp. 226-233, 2003.
- [11] Fabien Gouyon, Perfecto Herrera and Pedro Cano, "Pulse-Dependent Analysis of Percussive Music", in *Proc. Workshop on Digital Audio Effects (DAFx-98)*, Barcelona, Spain, pp. 188-191, 1998.

- [12] Jarno Seppänen, “Tatum Grid Analysis of Musical Signals”, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New York, USA, pp. 101-104, 2001.
- [13] Jarno Seppänen, *Computational Problems of Musical Meter Recognition*, MSc Thesis, Tampere University of Technology, 2001.
- [14] Fabien Gouyon and Perfecto Herrera, “Determination of the Meter of Musical Audio Signals: Seeking Recurrences in Beat Segment Descriptors”, in *Proc. AES 114<sup>th</sup> Convention*, Amsterdam, Netherlands, 2003.
- [15] Anssi P. Klapuri, “Musical Meter Estimation and Music Transcription”, presented at *Cambridge Music Colloquium 2003*, Cambridge University, UK, 2003.
- [16] Anssi P. Klapuri, “Sound Onset Detection by Applying Psychoacoustic Knowledge”, in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, vol. 6, pp. 3089-3092, 1999.
- [17] Eric D. Scheirer, “Pulse Tracking with a Pitch Tracker”, in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk, 1997.
- [18] George Tzanetakis, Georg Essl and Perry Cook, “Human Perception and Computer Extraction of Musical Beat Strength”, in *Proc. 5<sup>th</sup> Int. Conf. on Dig. Audio Effects (DAFX02)*, Hamburg, Germany, pp. 257-261, 2002.
- [19] Leon van Noorden and Dirk Moelants, “Resonance in the Perception of Musical Pulse”, in *J. New Music Research*, vol. 28, no. 1, pp. 43-66, 1999.
- [20] John Miller Chernoff, *African Rhythm and African Sensibility*, University of Chicago Press, 1979.