Metrical Analysis of Music Signals

Miguel A. ALONSO

{malonso@tsi.enst.fr}

École Nationale Supérieure des Télécommunications (ENST)

Paris, France

What's the song?

- rhythm is essential to music

TELECOM PARIS
école nationale supérieure des télécommunications

What's the song?

- rhythm is essential to music
- *pulse* and *meter* characteristics are very robust to signal transformation
  - transformed
  - original

Objective of this presentation

- Give a general overview about the metrical analysis of music signals as well as some research axis of my PhD work

Presentation content

- Introduction
- Beat-tracking model
- Performance analysis
- Conclusions

Introduction

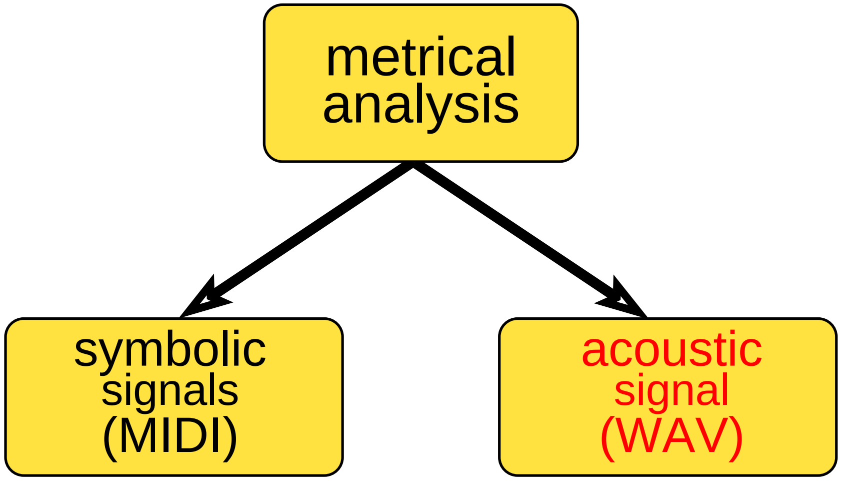
- automatic music analysis is an expanding research area

Introduction

- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
    - rhythm alignement of musical instruments
    - cut and paste operations in audio editing
    - MIR
    - music transcription
    - special effects

Introduction

- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
  - our approach deals with music recordings



Introduction

- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
  - our approach deals with music recordings
  - the most important aspect in rhythm perception is the *tempo*

Introduction

- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
  - our approach deals with music recordings
  - the most important aspect in rhythm perception is the *tempo*
  - automatic estimation is difficult for a broad variety of music

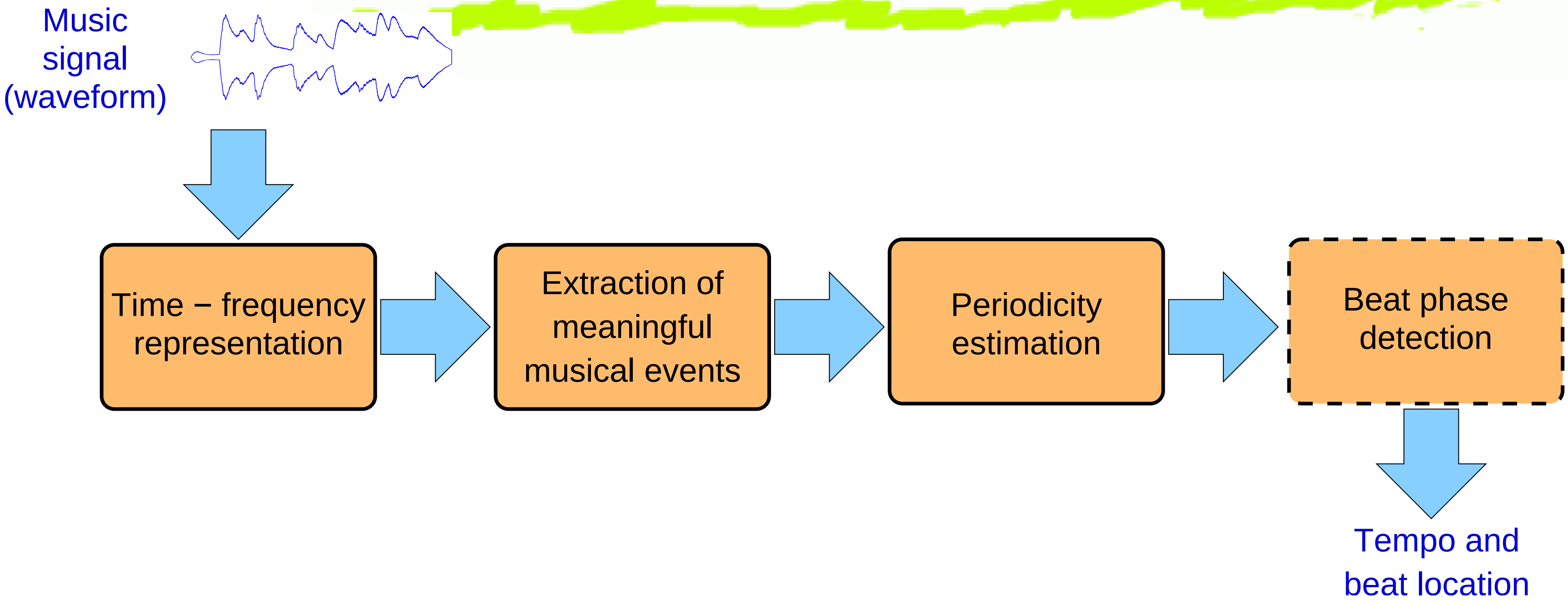
Introduction

- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
  - our approach deals with music recordings
  - the most important aspect in rhythm perception is the *tempo*
  - automatic estimation is difficult for a broad variety of music
- the proposed system aims at various musical genres

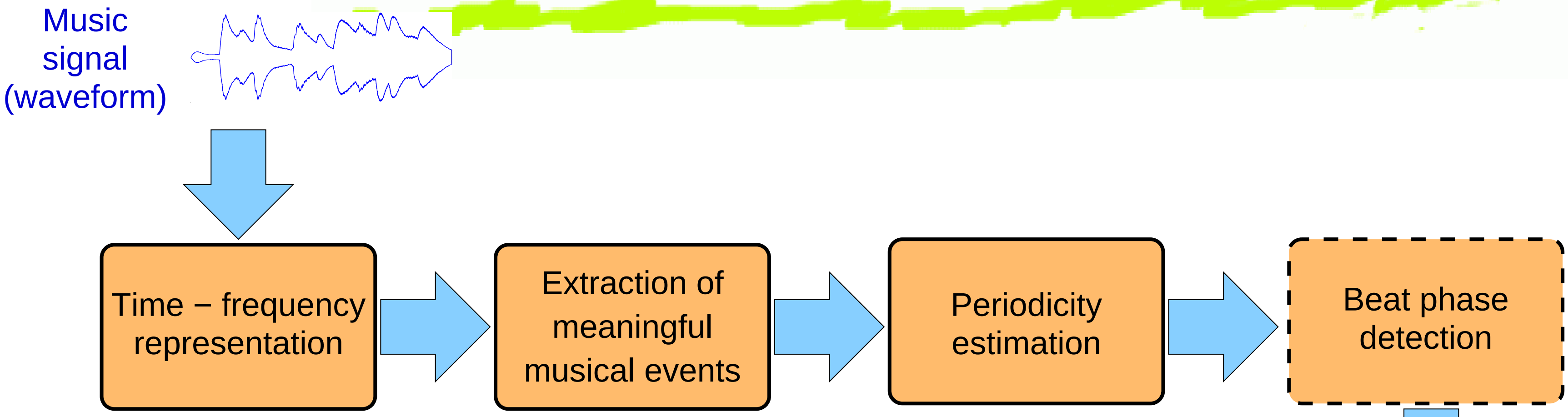
Introduction

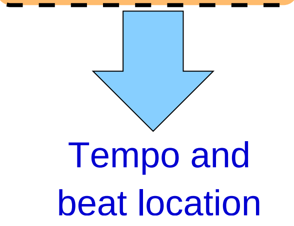
- automatic music analysis is an expanding research area
- metrical analysis is an essential part of this field
  - important for many audio applications
  - our approach deals with music recordings
  - the most important aspect in rhythm perception is the *tempo*
  - automatic estimation is difficult for a broad variety of music
- the proposed system aims at various musical genres
- most algorithms are based on the same general architecture

General architecture



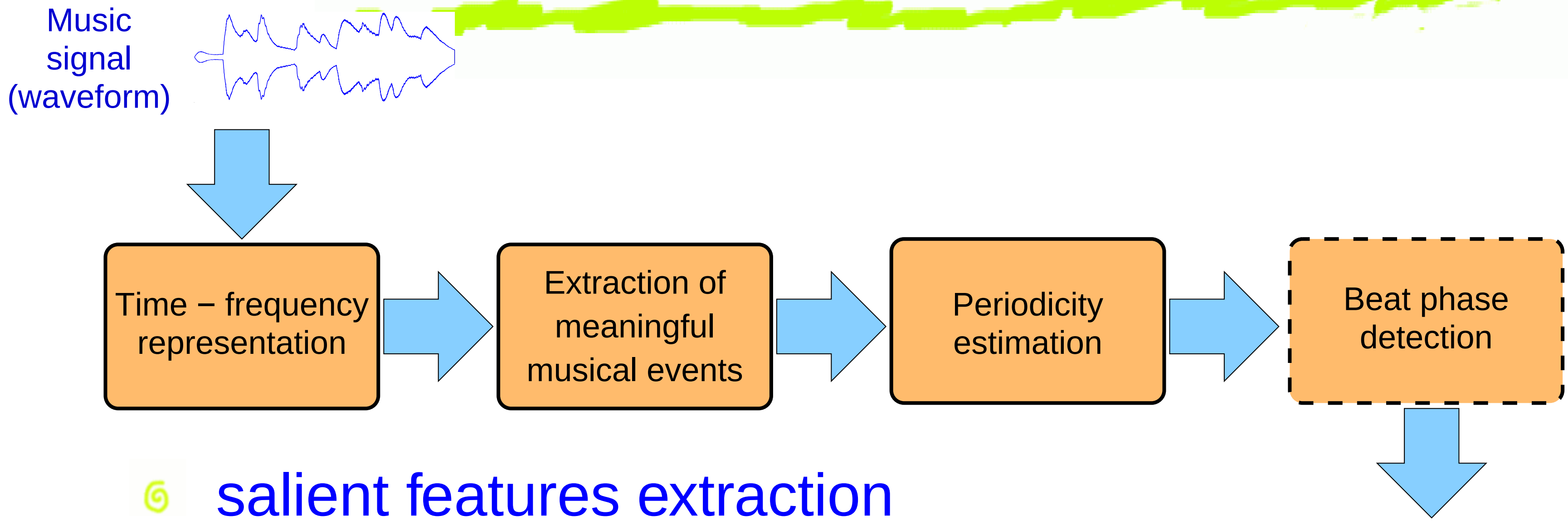
General architecture





- time-frequency representation
  - filter banks
  - STFT
  - wavelets
  - matching pursuit
  - parametric models
  - Wigner-Ville

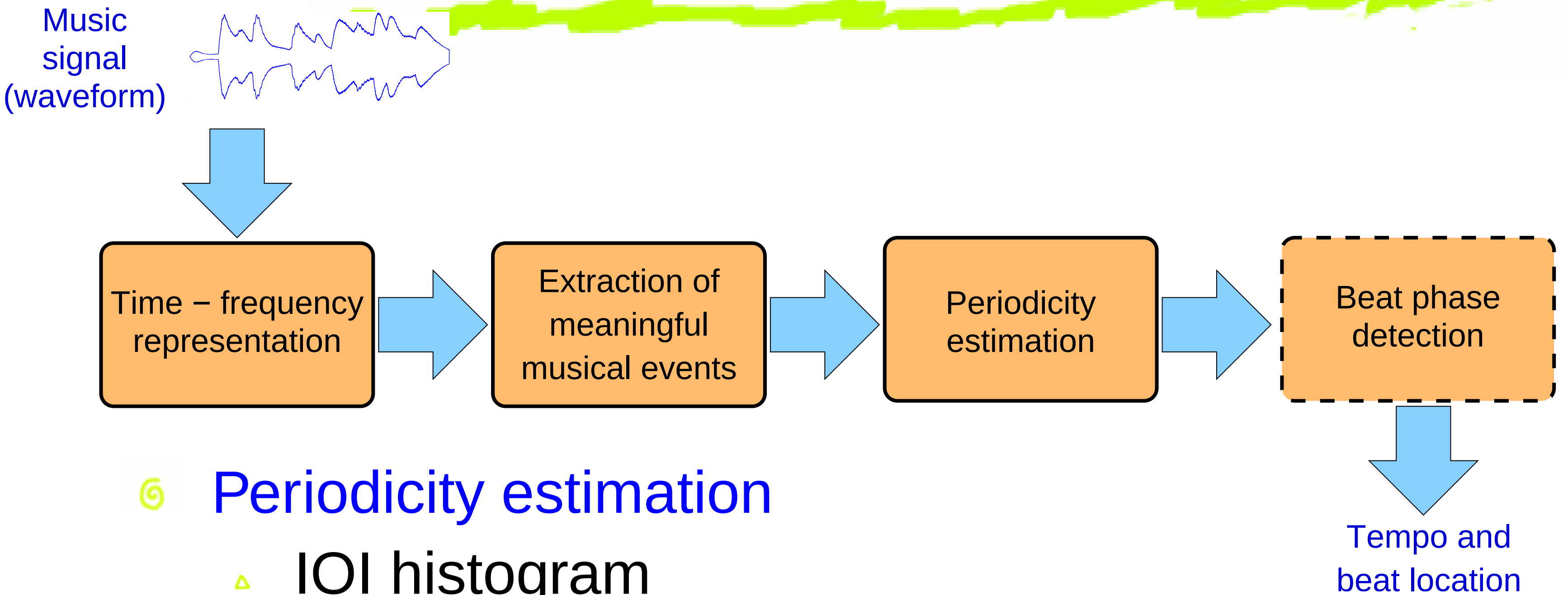
General architecture



- salient features extraction
  - energy content
  - high-frequency content
  - spectral difference
  - phase stability
  - probabilistic models
  - SVM
  - ICA

TELECOM PARIS
école nationale supérieure des télécommunications

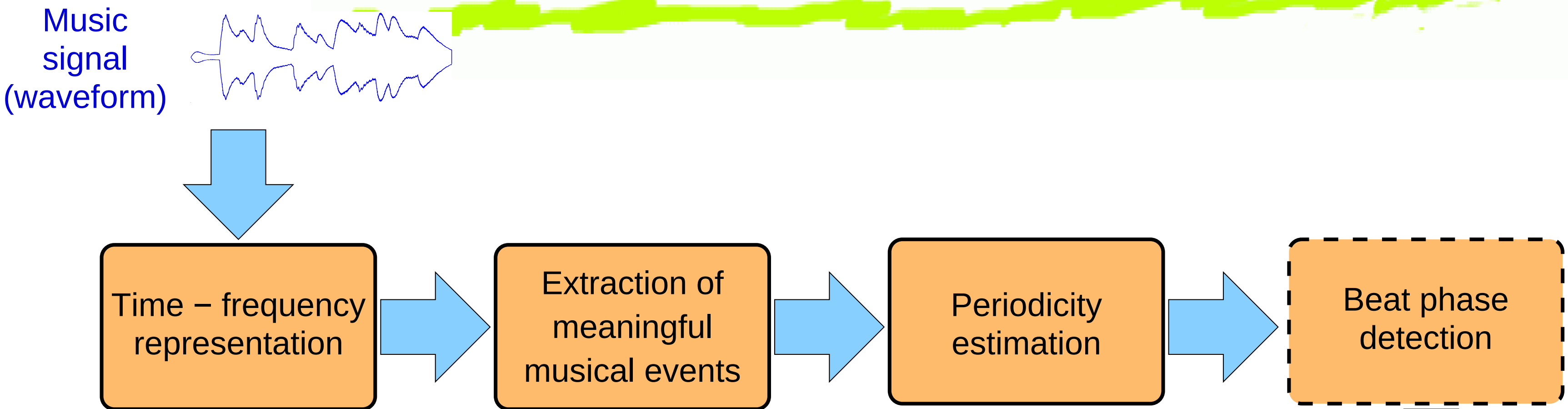
General architecture



Periodicity estimation

- IOI histogram
- pitch estimation methods (ACF, spectral product, YIN, etc.)
- bank of *comb filter resonators*
- probabilistic models (GMM, bayesian neworks)
- resonators based on neural networks
- periodicity transform

General architecture



- Beat phase detection
  - cross-correlation with a pulse-train
  - pick-picking

Presentation content

- Introduction
- Beat-tracking model
- Performance analysis
- Conclusions

Beat-tracking model

- we propose a system based on the same general architecture

Beat-tracking model

- we propose a system based on the same general architecture
  - a novel idea, to perform tempo estimation on the residual part of a harmonic/noise decomposition

Beat-tracking model

- we propose a system based on the same general architecture
  - a novel idea, to perform tempo estimation on the residual part of a harmonic/noise decomposition
  - the residual part was computed using a *noise subspace projection* approach
  - the audio signal is modeled as:

$$x(n) = \sum_{k=1}^{M} \alpha_k e^{i\omega_k n + \phi_k} + e(n)$$

Beat-tracking model

- we propose a system based on the same general architecture
  - a novel idea, to perform tempo estimation on the residual part of a harmonic/noise decomposition
  - the residual part was computed using a *noise subspace projection* approach
  - the audio signal is modeled as:

$$x(n) = \sum_{k=1}^{M} \alpha_k e^{i\omega_k n + \phi_k} + e(n)$$

  - if cannot be modeled as a complex exponential (i.e., onsets, attacks) it is considered as the *residual*

Harmonic plus noise decomposition

- piano example original and residual
- french horn example original and residual
- violin example original and residual

Harmonic plus noise decomposition

- piano example original and residual
- french horn example original and residual
- violin example original and residual
- about subspace-based techniques
  - much more precise than Fourier analysis
  - very robust to high noise levels
  - very short analysis windows can be used
  - not required to subtract the sinusoids

Harmonic plus noise decomposition

- piano example original and residual
- french horn example original and residual
- violin example original and residual
- about subspace-based techniques
  - much more precise than Fourier analysis
  - very robust to high noise levels
  - very short analysis windows can be used
  - not required to subtract the sinusoids
  - very computationally demanding
  - the *model order* must be well estimated

detection of salient features

- what is a *salient feature*?

detection of salient features

- what is a *salient feature*?
  - those timepoints where there is a marked change in any of the perceived psychoacoustical properties of sound, i.e., *loudness, timbre and pitch*

detection of salient features

- what is a *salient feature*?
  - those timepoints where there is a marked change in any of the perceived psychoacoustical properties of sound, i.e., *loudness, timbre and pitch*
- robust detection for polyphonic music is a difficult task

detection of salient features

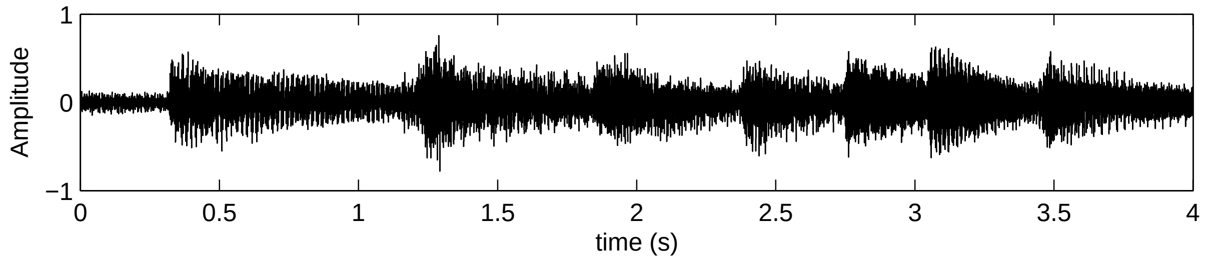
- what is a *salient feature*?
  - those timepoints where there is a marked change in any of the perceived psychoacoustical properties of sound, i.e., *loudness, timbre and pitch*
- robust detection for polyphonic music is a difficult task
- motivated by previous work, we define the *Spectral Energy Flux (SEF)* $E(f, k)$ of an audio signal

Spectral energy flux (1/2)

- Piano example

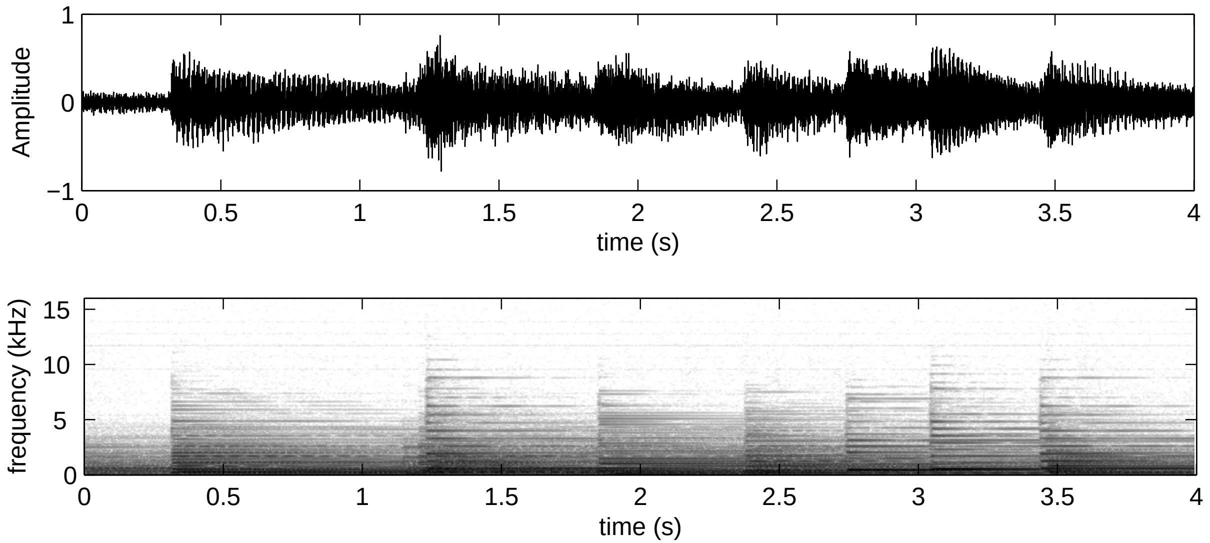
Spectral energy flux (1/2)

- Piano example



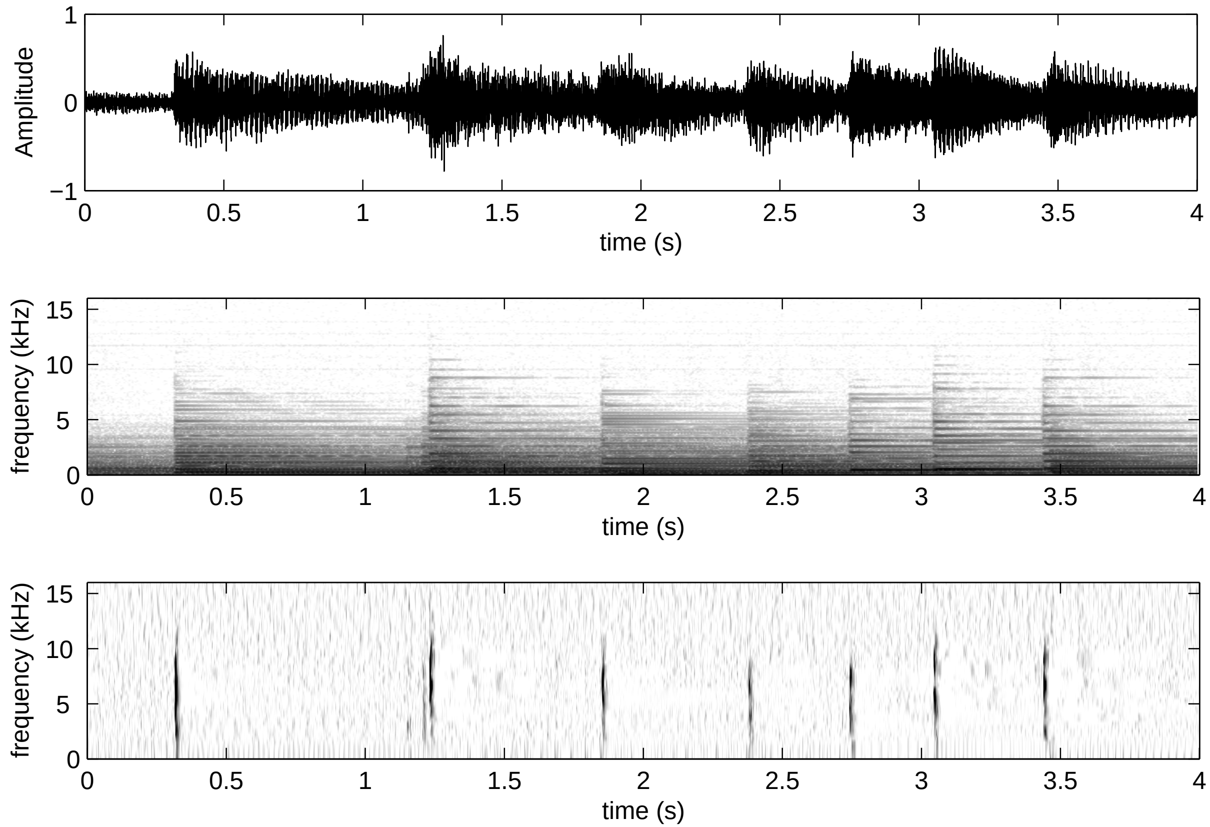
Spectral energy flux (1/2)

- Piano example



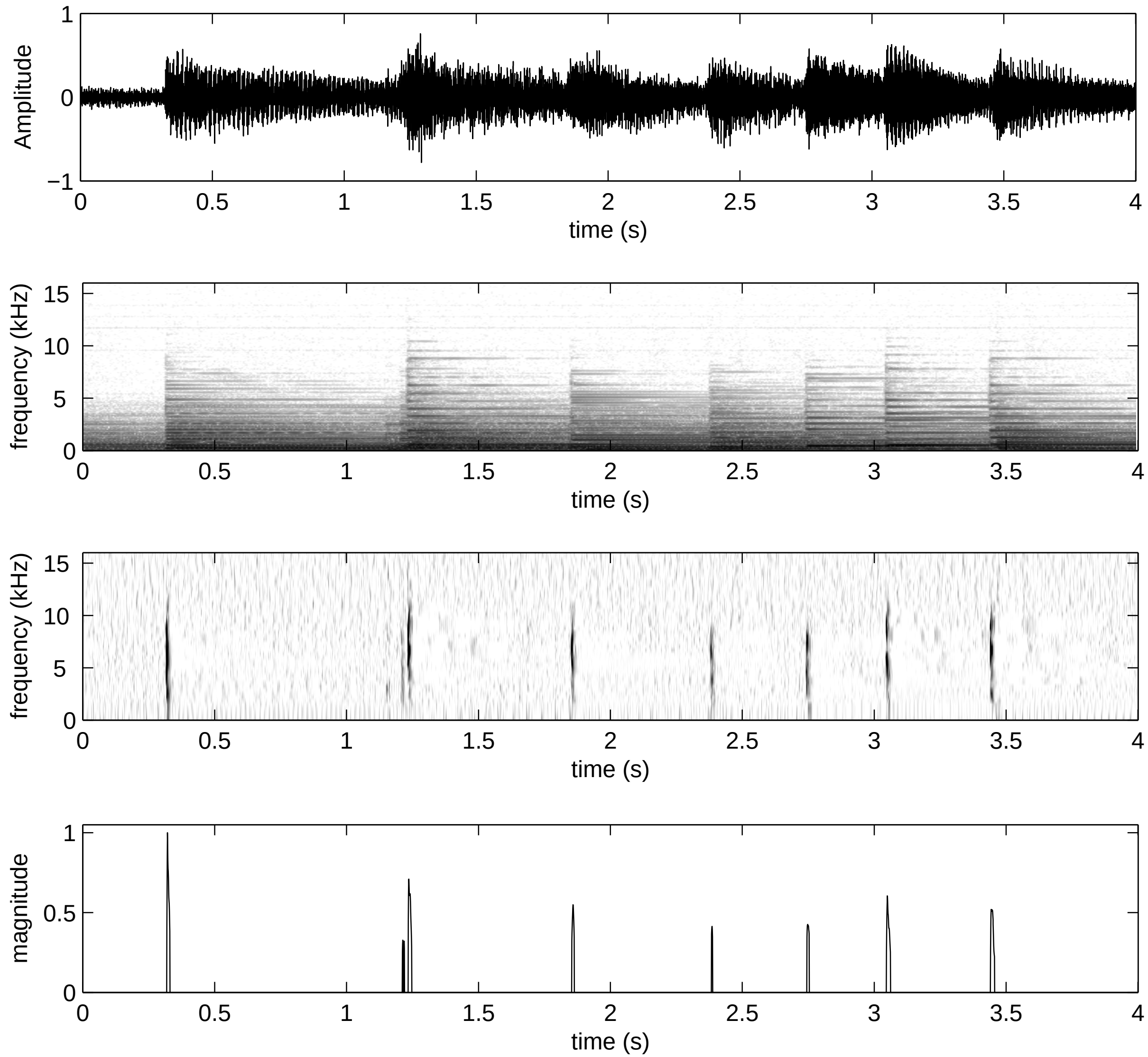
Spectral energy flux (1/2)

- Piano example



Spectral energy flux (1/2)

- Piano example

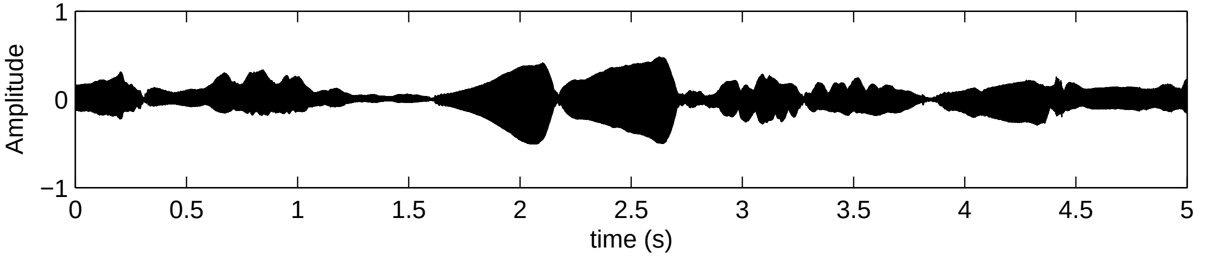


Spectral energy flux (2/2)

- Violin example

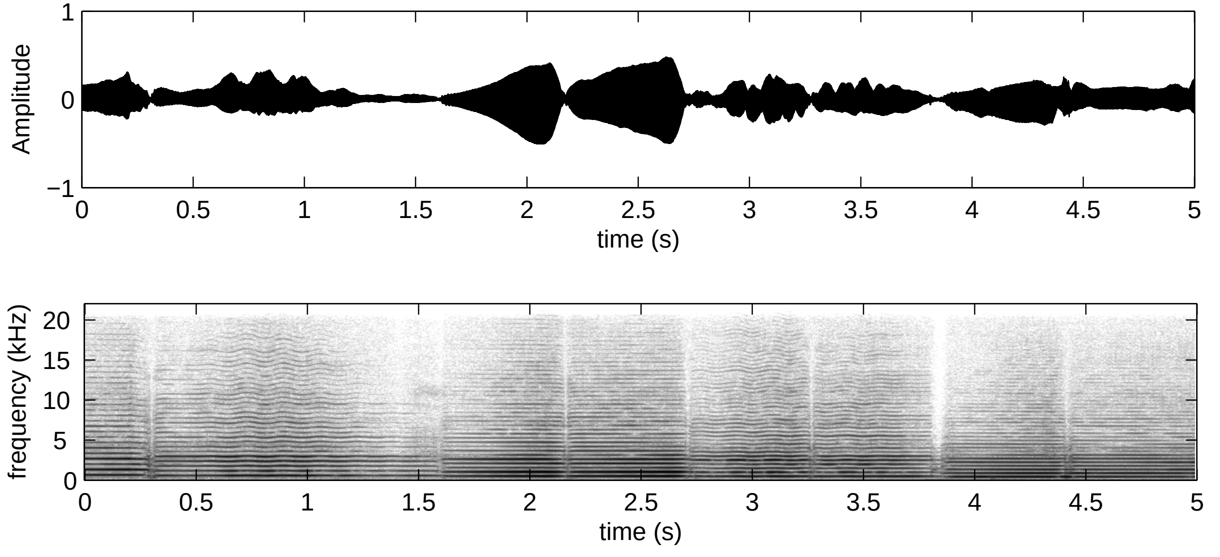
Spectral energy flux (2/2)

- Violin example



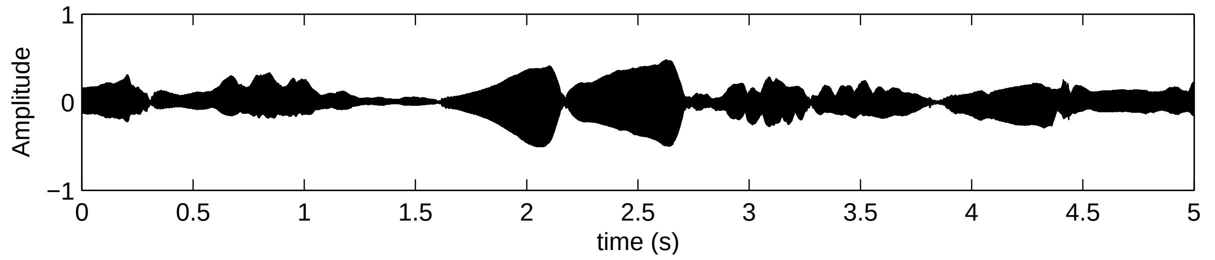
Spectral energy flux (2/2)

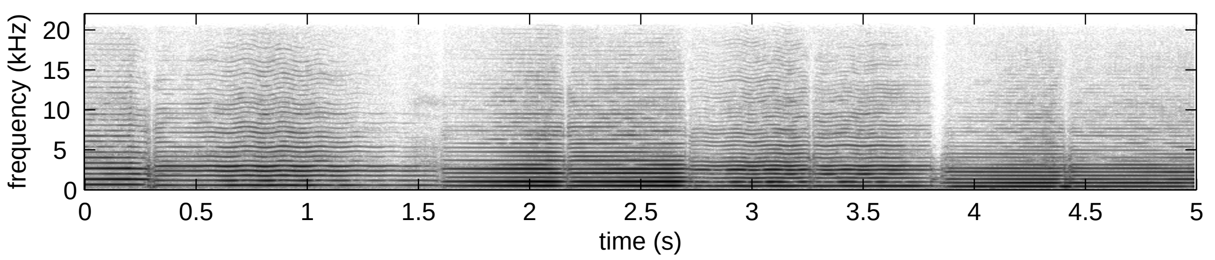
- Violin example

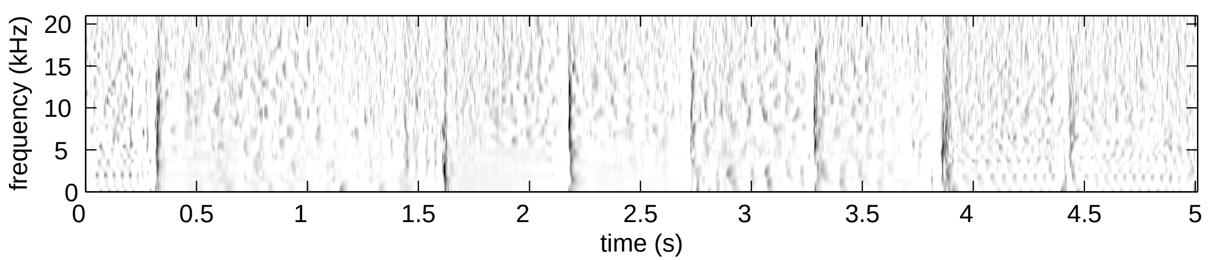


Spectral energy flux (2/2)

- Violin example

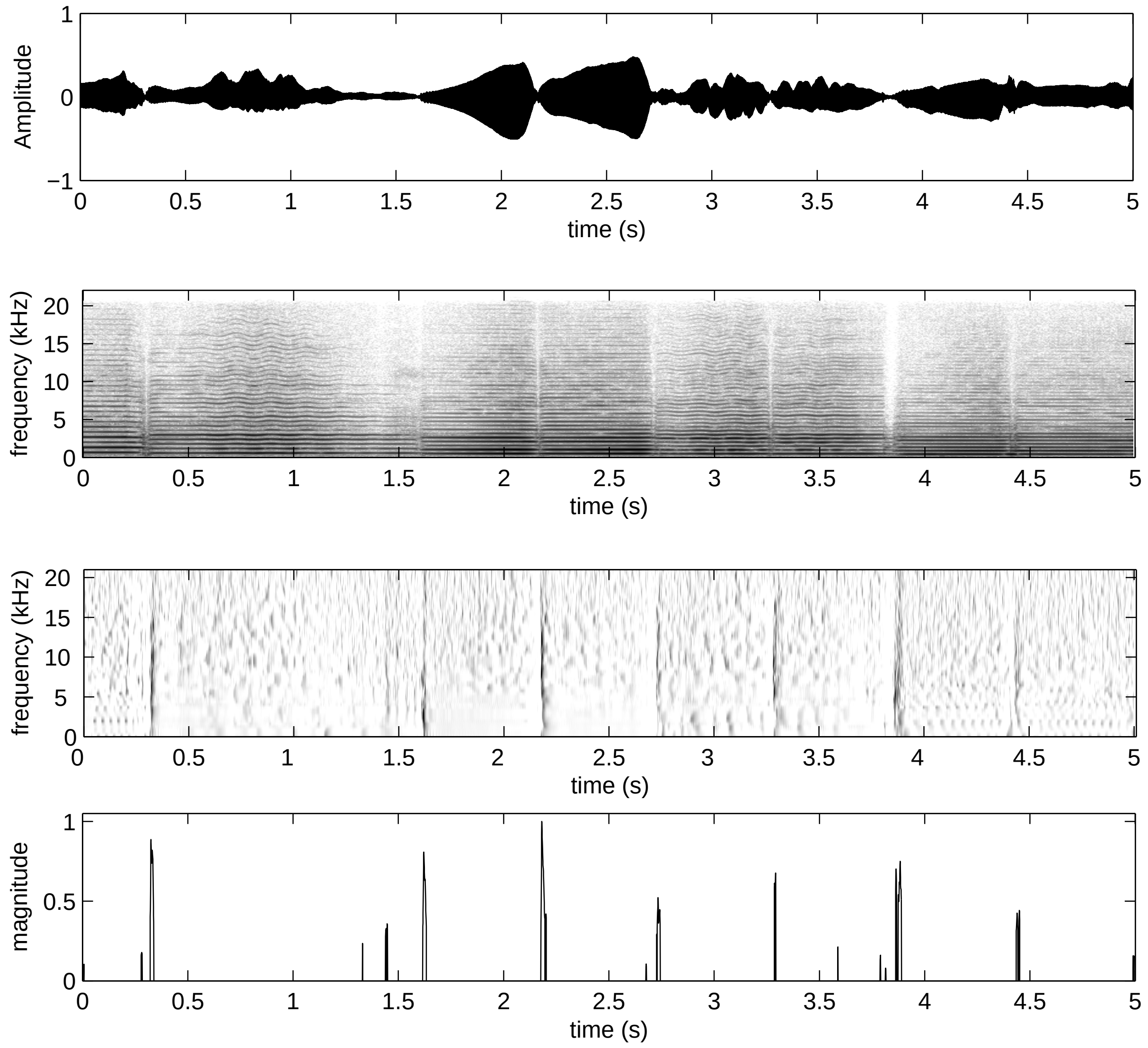






Spectral energy flux (2/2)

- Violin example



Periodicity estimation and beat location

- the output of the salient feature stage is a quasiperiodic pulse-train like signal

Periodicity estimation and beat location

- the output of the salient feature stage is a quasiperiodic pulse-train like signal
- detection function periodicity is found using two different methods
  - *spectral product*
  - *autocorrelation function*

Periodicity estimation and beat location

- the output of the salient feature stage is a quasiperiodic pulse-train like signal
- detection function periodicity is found using two different methods
  - *spectral product*
  - *autocorrelation function*
- beat location is found via a cross-correlation with an artificial pulse-train

Presentation content

- Introduction
- Beat-tracking model
- Performance analysis
- Conclusions

Performance analysis

- evaluation using a corpus of 489 musical excerpts

Performance analysis

- evaluation using a corpus of 489 musical excerpts
- wide diversity of musical genres

Genre	Pieces	Percentage
classical	137	28.0 %
jazz	79	16.2 %
latin	37	7.6 %
pop	40	8.2 %
rock	44	9.0 %
reggae	30	6.1 %
soul	24	4.9 %
rap, hip-hop	20	4.1 %
techno	23	4.7 %
other	55	11.2 %
total	**489**	**100 %**

Performance analysis

- evaluation using a corpus of 489 musical excerpts
- wide diversity of musical genres
- wide variety of instruments, dynamic range, etc.

Performance analysis

- evaluation using a corpus of 489 musical excerpts
- wide diversity of musical genres
- wide variety of instruments, dynamic range, etc.
- tempi in the 50 to 200 BPM range

Performance analysis

- evaluation using a corpus of 489 musical excerpts
- wide diversity of musical genres
- wide variety of instruments, dynamic range, etc.
- tempi in the 50 to 200 BPM range
- the tempo of each musical piece was manually annotated and cross-validated by at least two musicians

Results

- the algorithm was compared to our previous work
- it was also compared to our own implementation of the methods proposed by Paulus[1] and Scheirer[2]
- overall recognition rate for the evaluated systems

Method	Recognition rate
Paulus	56.3 %
Scheirer	67.4 %
SP .	63.2 %
AC .	73.6 %
SP using SEF.	84.0 %
AC using SEF	89.7 %

[1] Paulus J. and Klapuri A.,"*Measuring the similarity of rhythmic patterns*", Proc. ISMIR 2002.

[2] Scheirer, E.D., "*Tempo and beat analysis of acoustic music signals*", JASA, January 1998.

Sound examples

- example rock
- example country music
- example soul
- example salsa
- example guitarre
- example jazz 1
- example jazz 2
- example musique classique 1
- example musique classique 2

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings
- the concept of Spectral Energy Flux was used to derive an onset detector

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings
- the concept of Spectral Energy Flux was used to derive an onset detector
  - effective for a large range of audio signals
  - straightforward to implement
  - relatively low computational cost

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings
- the concept of Spectral Energy Flux was used to derive an onset detector
  - effective for a large range of audio signals
  - straightforward to implement
  - relatively low computational cost
- the performance was evaluated on a large database containing 489 musical pieces

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings
- the concept of Spectral Energy Flux was used to derive an onset detector
  - effective for a large range of audio signals
  - straightforward to implement
  - relatively low computational cost
- the performance was evaluated on a large database containing 489 musical pieces
- global success rate of 89.7%

Conclusions

- rhythm detection is performed on the residual signal
- efficient beat tracking algorithm for audio recordings
- the concept of Spectral Energy Flux was used to derive an onset detector
  - effective for a large range of audio signals
  - straightforward to implement
  - relatively low computational cost
- the performance was evaluated on a large database containing 489 musical pieces
- global success rate of 89.7%
- the system works off-line