Applications of Gabor multipliers to sound characterization and transformation

Adrien Merer

PhD at Laboratoire de Mécanique et d’Acoustique - CNRS Marseille

Supervisors: Sølvi Ystad and Richard Kronland-Martinet

MULAC kick-off meeting - September 2008
Applications of Gabor multipliers to sound characterization and transformation

Adrien Merer

Introduction

Genealogy

Morphing

Conclusion

Plan

1 Introduction

2 Genealogy

3 Morphing

4 Conclusion
Motivation

Characterization of perceptual sound categories

S1 \rightarrow S2 \rightarrow \text{Common points?}

Is it possible to synthesize or transform sounds to give them such a property.

Hypothesis

Brain processing responsible of sounds categorization and comparison are based on signal description aiming at storing the less information as possible and the less ambiguous as possible.

Listeners might use a combination of sound features, so called invariants, to determine specific aspects of sound.
Motivation

Characterization of perceptual sound categories

S1    S2    \(\rightarrow\) Common points?

Is it possible to synthesize or transform sounds to give them such a property.

Hypothesis

Brain processing responsible of sounds categorization and comparison are based on signal description aiming at storing the less information as possible and the less ambiguous as possible.

Listeners might use a combination of sound features, so called invariants, to determine specific aspects of sound.
Motivation

Characterization of perceptual sound categories

S1 → S2 → Common points?
Is it possible to synthesize or transform sounds to give them such a property.

Hypothesis

Brain processing responsible of sounds categorization and comparison are based on signal description aiming at storing the less information as possible and the less ambiguous as possible.

Listeners might use a combination of sound features, so called invariants, to determine specific aspects of sound.
Motivation

Characterization of perceptual sound categories

S1 \rightarrow S2 \rightarrow \text{Common points?}

Is it possible to synthesize or transform sounds to give them such a property.

Hypothesis

Brain processing responsible of sounds categorization and comparison are based on signal description aiming at storing the less information as possible and the less ambiguous as possible.

Listeners might use a combination of sound features, so called invariants, to determine specific aspects of sound.
Motivation

Characterization of perceptual sound categories

S1 \quad S2 \quad \rightarrow \quad \text{Common points?}

Is it possible to synthesize or transform sounds to give them such a property.

Hypothesis

Brain processing responsible of sounds categorization and comparison are based on signal description aiming at storing the less information as possible and the less ambiguous as possible.

Listeners might use a combination of sound features, so called invariants, to determine specific aspects of sound.
Motivation

Example

From recording of same notes played on different saxophones and clarinets:

→ find signal similarities of saxophone sounds and differences with clarinets sounds

→ use those signals characteristics to synthesize saxophone/clarinets sounds.
Example

From recording of same notes played on different saxophones and clarinets:

→ find signal similarities of saxophone sounds and differences with clarinets sounds

→ use those signals characteristics to synthesize saxophone/clarinets sounds.
Classical method
- Computing of signal descriptors from timbre literature (spectral centroid, attack time, spectral variation...)
- Correlation between descriptors and perceptual data (using regression, or machine learning methods)

Problems
- Time depend properties cannot be well described
- Same descriptors are used for all problems whereas each problem is specific and might necessitates new descriptors.

We expect GM to highlight time-frequency patterns responsible of perceptual categorization
Classical method
- Computing of signal descriptors from timbre literature (spectral centroid, attack time, spectral variation...)
- Correlation between descriptors and perceptual data (using regression, or machine learning methods)

Problems
- Time depend properties cannot be well described
- Same descriptors are used for all problems whereas each problem is specific and might necessitates new descriptors.

We expect GM to highlight time-frequency patterns responsible of perceptual categorization
Classical method
- Computing of signal descriptors from timbre literature (spectral centroid, attack time, spectral variation...)
- Correlation between descriptors and perceptual data (using regression, or machine learning methods)

Problems
- Time depend properties cannot be well described
- Same descriptors are used for all problems whereas each problem is specific and might necessitates new descriptors.

We expect GM to highlight time-frequency patterns responsible of perceptual categorization
Applications of Gabor multipliers to sound characterization and transformation

Adrien Merer

Introduction

Classical method
- Computing of signal descriptors from timbre literature (spectral centroid, attack time, spectral variation...)
- Correlation between descriptors and perceptual data (using regression, or machine learning methods)

Problems
- Time depend properties cannot be well described
- Same descriptors are used for all problems whereas each problem is specific and might necessitates new descriptors.

We expect GM to highlight time-frequency patterns responsible of perceptual categorization
Using Gabor Masks

Mask does not provide more information than Gabor transforms.

Estimation of time-frequency shift is necessary but not sufficient.
Mask does not provide more information than Gabor transforms.

Estimation of time-frequency shift is necessary but not sufficient.
From \( \{s_1, \ldots s_n\} \), build a sound that contains common properties.

Characterization is contained both in masks and sound \( S_{1\ldots n} \) (the so called "son").
From \( \{s_1, ... s_n\} \), build a sound that contains common properties.

Characterization is contained both in masks and sound \( S_1 ... n \) (the so called ”son”).
GM is considered as the transformation that makes the sound ”thinner”

We want to keep only common points

1: x0: initialization;
2: repeat
3: Compute m1;
   Compute m2;
   \[ x_0 = \frac{m_1 \cdot x_1 + m_2 \cdot x_2}{2} \]
   update error
4: until error < threshold

\[ error = \| x_0^i - x_0^{i+1} \|^2 \]
Algorithm

GM is considered as the transformation that makes the sound "thinner"

We want to keep only common points

1. $x_0$: initialization;
2. repeat
3. Compute $m_1$;
   Compute $m_2$;
   $x_0 = \frac{m_1 x_1 + m_2 x_2}{2}$
   update error
4. until $\text{error} < \text{threshold}$

$\text{error} = \|x_0^i - x_0^{i+1}\|^2$
Questions

- Highly dependent on initialization
  → Many different informations can be introduced with initialization sound (we can use harmonic comb or impulse train)
- We can also initialize using two masks!
  → Problem must be better defined.
  - Necessitate to know exactly what is expected as the ”son”
• Highly dependent on initialization
  → Many different informations can be introduced with initialization sound (we can use harmonic comb or impulse train)

• We can also initialize using two masks!
  → Problem must be better defined.
  - Necessitate to know exactly what is expected as the ”son”

...
- Classical methods are based on analysis/synthesis process and simple interpolation between synthesis parameters.

Examples
- Additive synth: Impacts Glass → Metal
- Using Gabor Multipliers Impacts Glass → Metal 2
Morphing

- Classical methods are based on analysis/synthesis process and simple interpolation between synthesis parameters.

Examples

- Additive synth: Impacts Glass → Metal
- Using Gabor Multipliers Impacts Glass → Metal 2
Morphing using GM

N-root of mask

\[ x_a \rightarrow m \rightarrow x_b \]

- \( x_{i\frac{ab}{N}} \neq x_{i\frac{ba}{N}} \) even for \( i = N/2 \) (and even values of N)
- For noisy signals, we can hear numerical artifacts.

*Initial / morphing*
Morphing using GM

N-root of mask

\[ x_a \rightarrow \sqrt{m} \rightarrow x_{1/2}ab \rightarrow \sqrt{m} \rightarrow x_b \]

- \[ x_{i \frac{N}{ab}} \neq x_{i \frac{N}{ba}} \] even for \( i = N/2 \) (and even values of N)
- For noisy signals, we can hear numerical artifacts.

*Initial / morphing*
Morphing using GM

N-root of mask

- $x_{i \frac{N}{2}}^{ab} \neq x_{i \frac{N}{2}}^{ba}$ even for $i = N/2$ (and even values of $N$)
- For noisy signals, we can hear numerical artifacts.

Initial / morphing
Applications of Gabor multipliers to sound characterization and transformation

Morphing using GM

N-root of mask

$x_a \rightarrow \sqrt{m} \rightarrow x^{1/2}_{1/2} \rightarrow x_b$

• $x_{i \frac{ab}{N}} \neq x_{i \frac{ba}{N}}$ even for $i = N/2$ (and even values of N)

• For noisy signals, we can hear numerical artifacts.

Initial / morphing
Applications of Gabor multipliers to sound characterization and transformation

Adrien Merer

Introduction

Genealogy

Morphing

Conclusion

Morphing using GM

N-root of mask

\[ x_a \rightarrow \sqrt{m} \rightarrow x_{1/2}ab \rightarrow x_b \]

- \[ x_{i \frac{ab}{N}} \neq x_{i \frac{ba}{N}} \text{ even for } i = N/2 \text{ (and even values of } N) \]
- For noisy signals, we can hear numerical artifacts.

Initial / morphing
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity → Depends on windows size

Normal morphing Without phase S/D separation
\( x_1(\text{Sax}) \quad x_2(\text{Clarinet}) \)

To do: separation for only parts of each modulation
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity → Depends on windows size

Normal morphing Without phase S/D separation
\[ x_1(Sax) \quad x_2(\text{Clarinet}) \]
To do: separation for only parts of each modulation
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity \[ \rightarrow \text{Depends on windows size} \]

Normal morphing Without phase S/D separation
\( x_1(\text{Sax}) \quad x_2 (\text{Clarinet}) \)

To do: separation for only parts of each modulation
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity → Depends on windows size

Normal morphing Without phase S/D separation

$x_1(Sax)$  $x_2$ (Clarinet)

To do: separation for only parts of each modulation
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity $\rightarrow$ Depends on windows size

Normal morphing Without phase $\rightarrow$ S/D separation

$x_1(Sax)$ $\quad x_2$ (Clarinet)

To do: separation for only parts of each modulation
Deterministic and Stochastic parts

Idea is not to use phase for reconstruction of stochastic part of the signal, hence to define and separate it.

How to quantify randomness in a frequency band?

Randomness

- Linear regression of phase
- Standard deviation of phase
- Define a threshold of linearity $\rightarrow$ Depends on windows size

Normal morphing Without phase S/D separation

$x_1(Sax), x_2(\text{Clarinet})$

To do: separation for only parts of each modulation
Combining morphing and genealogy, an idea is to build a prototype of a given sound category as a weighted morphing between all the sound.
Combining morphing and genealogy, an idea is to build a prototype of a given sound category as a weighted morphing between all the sound.
Perceptual data

The concept of prototype is often used in study about categorization and it is possible to ask listeners to judge a level of "prototypicity" for each sound inside a category.

Such rating can be used as a weight/distance for each sound.

Problem

How to morph sounds together?
Initialize $P$ and all morphings between each sound and $P$?
Perceptual data

The concept of prototype is often used in study about categorization and it is possible to ask listeners to judge a level of "prototypicity" for each sound inside a category. Such rating can be used as a weight/distance for each sound.

Problem

How to morph sounds together?

Initialize $P$ and all morphings between each sound and $P$?
Perceptual data

The concept of prototype is often used in study about categorization and it is possible to ask listeners to judge a level of "prototypicity" for each sound inside a category. Such rating can be used as a weight/distance for each sound.

Problem

How to morph sounds together?
Initialize $P$ and all morphings between each sound and $P$?