MODAL DECOMPOSITION OF MUSICAL INSTRUMENT SOUND VIA ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Yoshiki Masuyama*, Tsubasa Kusano, Kohei Yatabe, Yasuhiro Oikawa (Waseda University)
*Email: mas-03151102@akane.waseda.jp

waseda university INTERMEDIA ART+SCIENCE The Department of Intermedia Art and Science

Abstract

For musical instrument sounds containing partials, or modes, **the behavior of modes around the attack time is an important factor** for characterizeing it. However, accurately decomposing it around the attack time is not an easy task, especially when the sound has sharp onset at the time. In this paper, **an optimization-based method of modal decomposition** is proposed. The proposed method is formulated as a constrained optimization problem to enforce the perfect reconstruction property for accurate decomposition. In optimization, **the Alternating Direction Method of Multiplier (ADMM)** is utilized, where the update of variables is calculated in closed form.

Motivation

Modal decomposition of musical instrument sounds

In this paper, a sound which consists of attack followed by decaying partials is considered. The parametric modeling of each mode is often considered in sound synthesis. Decaying processes of the modes also important for the timbre of musical instruments, **especially around the attack time.**

Modal decomposition based on filterbank

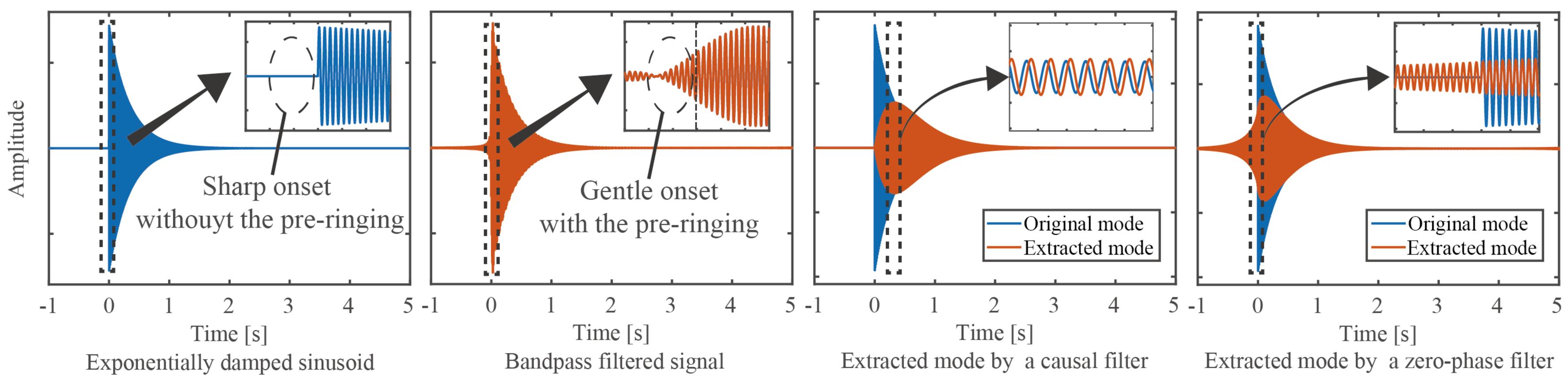
$$\hat{\mathbf{x}}_i = \hat{\mathbf{H}}_i \hat{\mathbf{s}},$$

$\hat{\mathbf{s}}$: The Fourier spectrum of musical instrument sounds
$\hat{\mathbf{H}}_i$: The diagonal matrix whose elements are the frequency response of a filter
$\hat{\mathbf{x}}_i$: The i th mode extracted by a linear filter

Potential issues of bandpass filter

There are three potential issues of modal decomposition by bandpass filter,

- The sharp onset of the extracted modes is corrupted,
- The trade-off between phase delay and pre-ringing cannot avoid,
- Modes are extracted independently (relation of them is not considered).



⇒ An accurate decomposition method must be a non-linear process.

Proposed modal decomposition

In order to overcome these potential issued of linear filtering, we propose the modal decomposition based on constrained optimization.

Interpretation of linear filtering as least squares method

This interpretation will illustrate the relation between the proposed method and the ordinary linear filtering.

$$\hat{\mathbf{x}}_i = \hat{\mathbf{H}}_i \hat{\mathbf{s}}, \quad \Rightarrow \quad \hat{\mathbf{H}}_i^{-1} \hat{\mathbf{x}}_i = \hat{\mathbf{s}}, \quad \Rightarrow \quad \min_{\mathbf{x}_i} \frac{1}{2} \|\hat{\mathbf{H}}_i^{-1} \hat{\mathbf{x}}_i - \hat{\mathbf{s}}\|_2^2.$$

Proposed formulation

The proposed modal decomposition is formulated as minimization of weighted energy of modes and the attack component:

$$\min_{\hat{\mathbf{x}}} \; \frac{1}{2} \|\mathbf{W}\hat{\mathbf{x}}\|_2^2 \quad \text{s.t.} \quad \hat{\mathbf{s}} = \sum_{i=1}^{N+1} \hat{\mathbf{x}}_i, \quad [\mathrm{IFFT}(\hat{\mathbf{x}}_i)]_n = 0 \;\; (n < \tau_A),$$

where n is the time index and τ is the time index corresponding to the attack time. In the proposed formulation,

- **the perfect reconstruction property eliminates the phase delay,**
- **the causality constraint eliminates the pre-ringing.**

Main points of proposed modal decomposition

- The frequency responses are replaced by an arbitrary diagonal matrixes.
- Considering residual which is expected to be the attack component.
- Data fidelity is considered in the perfect reconstruction constraint.

⇒ The proposed method solves potential issues of the linear filtering.

Proposed algorithm utilizing ADMM

Reformulation of proposed modal decomposition

In this paper, ADMM is adopted for proposed modal decomposition. For applying ADMM, the proposed modal decomposition is reformulated as:

$$\min_{\hat{\mathbf{x}}, \hat{\mathbf{z}}} \; \frac{1}{2} \|\mathbf{W}\hat{\mathbf{x}}\|_2^2 + \iota_{C_1}(\hat{\mathbf{x}}) + \iota_{C_2}(\hat{\mathbf{z}}) \quad \text{s.t.} \quad \hat{\mathbf{x}} = \hat{\mathbf{z}},$$

where

$$C_1 = \left\{ \hat{\mathbf{x}} \in \mathbb{C}^{(N+1)L} \;\middle|\; \hat{\mathbf{s}} = \textstyle\sum_{i=1}^{N+1} \hat{\mathbf{x}}_i \right\}, \quad \text{No phase delay}$$

$$C_2 = \left\{ \hat{\mathbf{z}} \in \mathbb{C}^{(N+1)L} \mid [\mathrm{IFFT}(\hat{\mathbf{z}}_i)]_n = 0 \;\; (n < \tau_A) \right\}. \quad \text{No pre-ringing}$$

The ADMM procedure

The proposed modal decomposition is implemented through the following procedure:

$$\hat{\mathbf{x}}^{k+1} = \arg\min_{\hat{\mathbf{x}}} \; \frac{1}{2} \|\mathbf{W}\hat{\mathbf{x}}\|_2^2 + \iota_{C_1}(\hat{\mathbf{x}}) + \frac{1}{2\rho} \|\hat{\mathbf{x}} - \hat{\mathbf{z}}^k + \hat{\mathbf{u}}^k\|_2^2, \quad \text{Eliminate phase delay}$$

$$\hat{\mathbf{z}}^{k+1} = \arg\min_{\hat{\mathbf{z}}} \; \iota_{C_2}(\hat{\mathbf{z}}) + \frac{1}{2\rho} \|\hat{\mathbf{x}}^{k+1} - \hat{\mathbf{z}} + \hat{\mathbf{u}}^k\|_2^2, \quad \text{Eliminate pre-ringing}$$

$$\hat{\mathbf{u}}^{k+1} = \mathbf{u}^k + \mathbf{x}^{k+1} - \mathbf{z}^{k+1}$$

where k is the iteration index, and the $\boldsymbol{x}$-update corresponds to the extension of linear filtering, and the $\boldsymbol{z}$-update considers causality. In ADMM procedure, **all updates are calculated in closed form.**

The $\boldsymbol{x}$-update calculation

The $\boldsymbol{x}$-update can be interpreted as a **linear filtering without phase delay which satisfies the perfect reconstruction property.**

$$\hat{x}_{i\xi} = \frac{\prod_{j \neq i} |w_{j\xi}|^2}{\sum_{l=1}^{N+1} \prod_{j \neq l} |w_{j\xi}|^2} \hat{s}_{\xi}, \quad \Leftrightarrow \quad \hat{\mathbf{x}}_i = \mathbf{G}_i \hat{\mathbf{s}}.$$

Experiments

Simulation

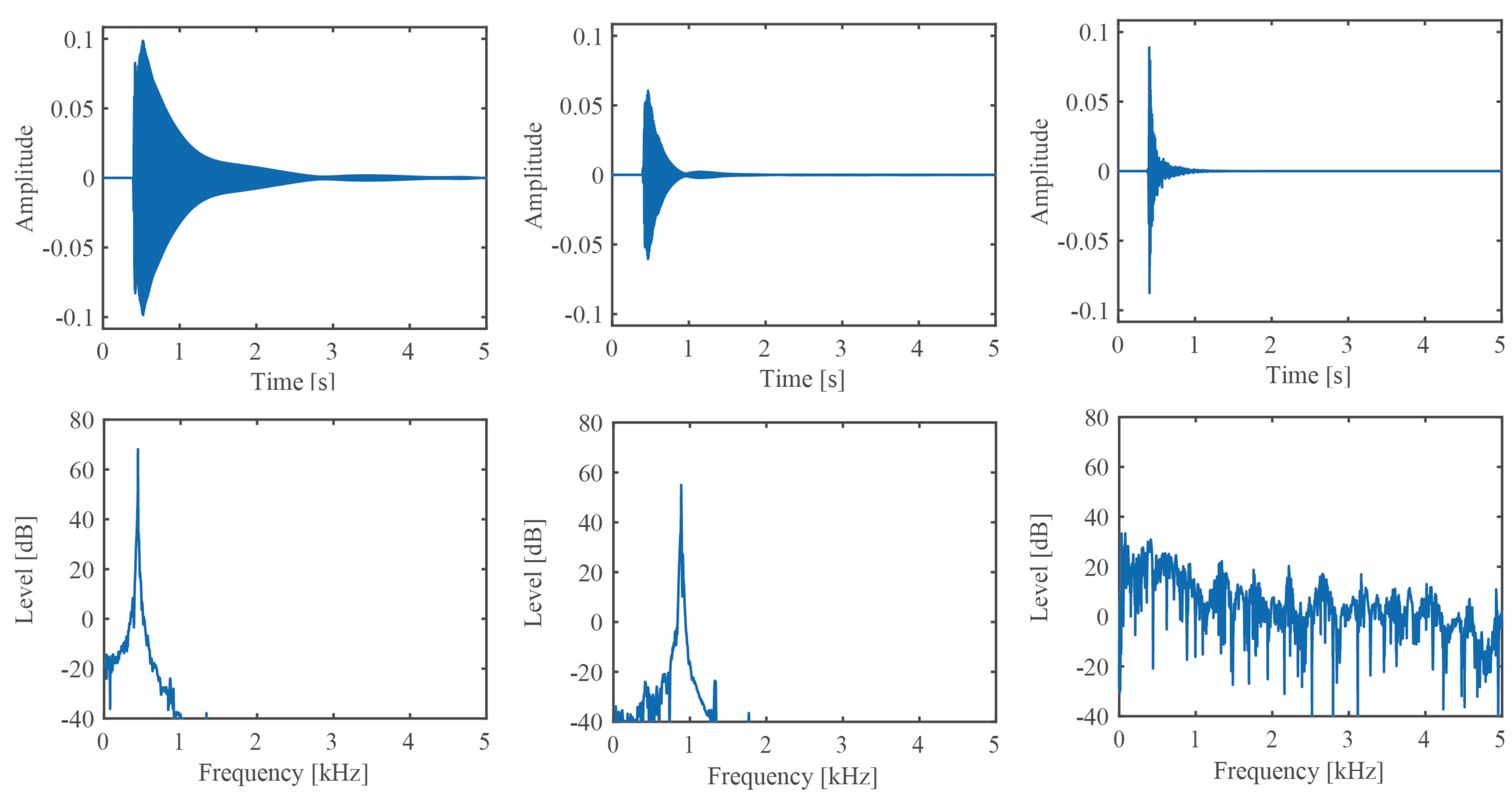
SDR of the decomposed modes of the simulated signal.

Modes	SDR [dB] 1st	2nd	3rd	4th
Causal filter	50.6	2.8	2.0	2.7
Zero-phase filter	33.9	2.7	2.0	2.7
STFT(a)	35.4	26.4	24.3	24.0
STFT(b)	43.4	28.7	27.2	26.3
Proposed method	**111.1**	**97.4**	**96.2**	**93.5**

- The simulated signal given by sum of four impulse responses of resonance filters was decomposed.
- SDR of resonance filters are corrupted by mode-mixing, especially in 3rd mode, but the proposed method kept high SDR in 3rd mode.

Application to piano sound decomposition

The piano sound of A4 was decomposed into 16 modes and the attack component. The piano sound was also approximated with 1000 order AR model where the sampling rate was 96000 Hz, and λ = 0.01. **The proposed method achieves modal decomposition with sharp onset.**



(a) the 1st mode (b) the 2nd mode (c) residual

Conclusion

The modal decomposition method of musical instrument sounds is proposed. By interpreting a filtering process as the least squares method, the proposed method is formulated as a constrained optimization problem which enables to incorporate two constraints. The proposed optimization problem is solved by the ADMM algorithm. In ADMM procedure, all updates are calculated in closed form, which allows an easy and fast update of the variables.