PARAMETRIC APPROXIMATION OF PIANO SOUND BASED ON KAUTZ MODEL WITH SPARSE LINEAR PREDICTION

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

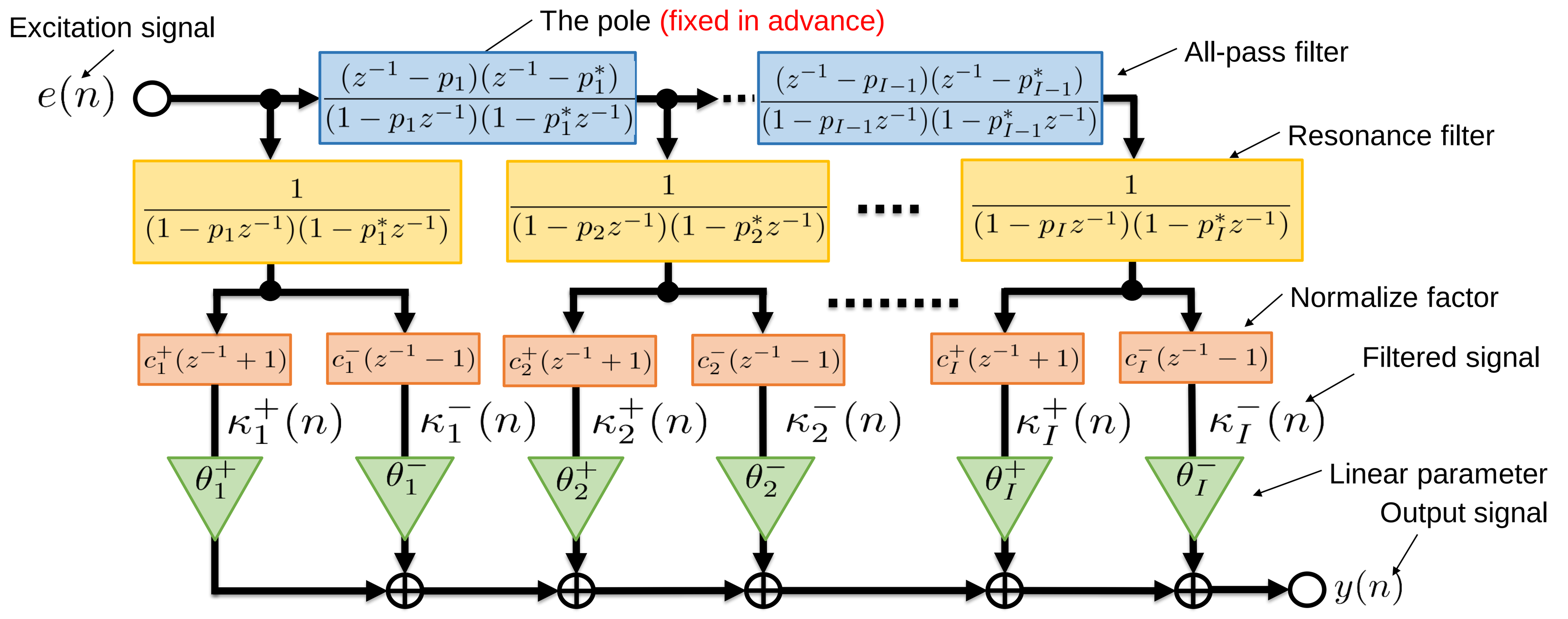
Kenji Kobayashi, Daiki Takeuchi, Mio Iwamoto,
Kohei Yatabe and Yasuhiro Oikawa (Waseda University, Japan)

Introduction

- The modal based methods using IIR (infinite impulse response) filters are popular method to approximate the piano sound and preferable for real-time synthesis.
- Estimating parameters of IIR filters is not easy because of non-linearity of parameters.
- The Kautz model can be optimized quite easily because of linearity in parameters.
- However, the Kautz model is not flexible because the poles and excitation signal have to be fixed in advance.
- The proposed method reduces such unwanted properties of the Kautz model for approximating piano sound and real-time synthesis.

Signal Approximation by Kautz model

- The structure of the Kautz model for set of complex poles $\{p_1, \cdots, p_I\}$



- The output signal was defined by linear combination of filtered wave:

$$y(n) = \sum_{i=1}^{I} \left[\theta_i^+ \kappa_i^+(n) + \theta_i^- \kappa_i^-(n) \right].$$

- The parameters of the Kautz model are easily estimated by the least squares method:

$$\underset{\{\theta_i^+,\theta_i^-\}_{i=1}^{I}}{\text{Minimize}} \sum_{n=1}^{N} \left| s(n) - \sum_{i=1}^{I} \left[\theta_i^+ \kappa_i^+(n) + \theta_i^- \kappa_i^-(n) \right] \right|^2. \qquad s(n): \text{target signal}$$

Pro:
- It is easy to estimate linear parameter $\theta_i^{\pm}$ if poles are fixed in advance.

Cons:
- The suitable poles must be fixed in advance.
- The excitation signal has to be fixed beforehand.

The proposed method solved these two problems of the Kautz model.

Proposed method

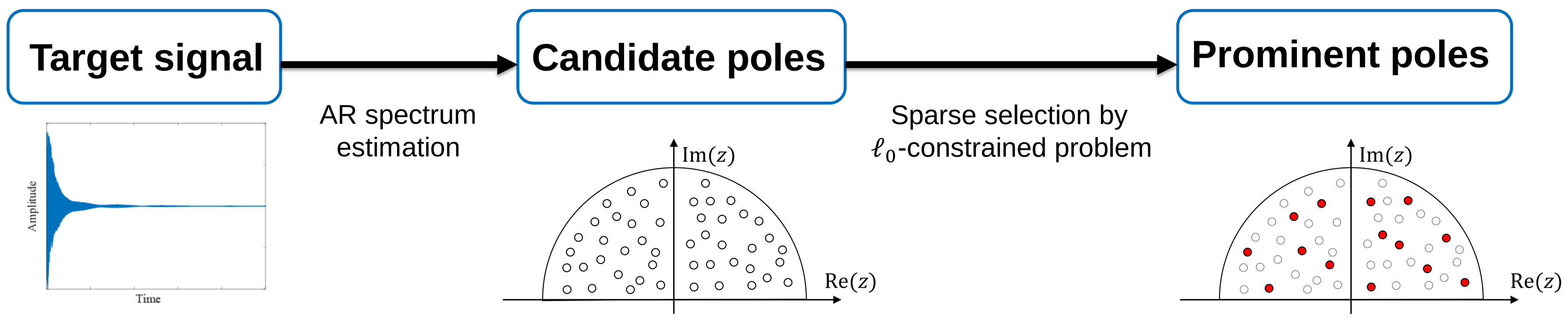
- The easiness of the Kautz model on its parameter estimation comes with a price of restriction on the poles and excitation signal.
- Generally, the computational cost increases according to the number of poles.
- In addition, the excitation signal should have small non-zero elements for saving the computational resources.

We propose method for estimating poles and excitation signal from the target signal.

STEP1 Get audio signal of the target piano sound $s(n)$.
STEP2 Apply sparse linear prediction to the target sound to obtain a prediction residual .
STEP3 Generate a large number of candidates of poles for the Kautz modeling from piano sound by an AR spectral analysis technique.
STEP4 Construct the Kautz filter $\{\Psi_i^{\pm}\}$ from obtained poles in the previous step.
STEP5 Input the excitation signal $e(n)$ obtained in **STEP2** to the filters for calculating $\boldsymbol{\kappa}(n)$.
STEP6 Solve ℓ_0-constrained least square problem to obtain P poles and the corresponding parameters $\{\theta_i^{\pm}\}$.

Sparse selection of the prominent poles

- The Kautz model itself cannot optimize poles within the framework of the model.
- We propose a method of generating several candidates poles and a method of sparsely selecting the prominent poles from them.



- Candidate poles can be estimated by an AR spectral estimation technique.
- The proposed method imposes ℓ_0-constraint into least square problem for selecting prominent poles from the candidate poles.

$$\underset{\boldsymbol{\theta}}{\text{Minimize}} \sum_{n=1}^{N} \left| s(n) - \boldsymbol{\kappa}(n)^T \boldsymbol{\theta} \right|^2 \ \underbrace{\text{subject to } \|\boldsymbol{\theta}\|_0 \le P.}_{\text{Induces sparsity of } \boldsymbol{\theta}}$$

$$\boldsymbol{\theta} = [\theta_1^+, \theta_1^-, \theta_2^+, \theta_2^-, \ldots, \theta_I^+, \theta_I^-]^T$$
$$\boldsymbol{\kappa}(n) = [\kappa_1^+(n), \kappa_1^-(n), \ldots, \kappa_I^+(n), \kappa_I^-(n)]^T$$
$\|\cdot\|_0$: number of non-zero elements

P prominent poles are selected from candidate poles.

Generating excitation signal based on sparse LPC

- linear prediction approximates signal into the prediction filter and the residual:

$$x(n) = \sum_{l=1}^{L} a_l x(n-l) + e(n). \qquad a_l: \text{coefficient of the linear prediction model}$$

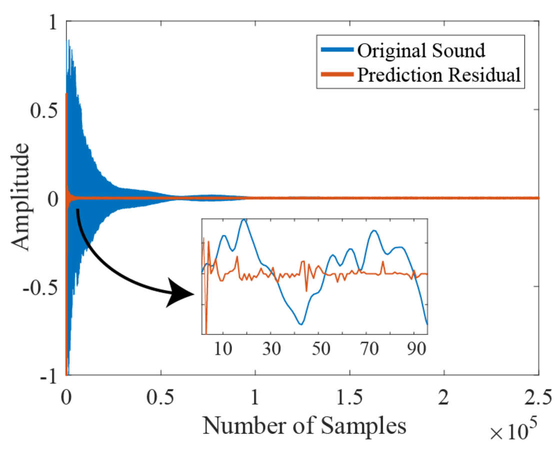
- The residual can be regarded as the component of signal that is difficult to be approximated by a filter.

It may be suitable for the excitation signal.

- We propose to use the sparse linear prediction technique based on ℓ_1-norm:

$$\underset{\{a_l\}_{l=1}^{L}}{\text{Minimize}} \sum_{n=1}^{N} \left| s(n) - \sum_{l=1}^{L} a_l s(n-l) \right|.$$

- This figure shows residual signal using sparse LPC to piano signal of A4 (441 Hz).



The residual has quite small values at most of the duration, and the remaining energy was concentrated around the first 10 samples.

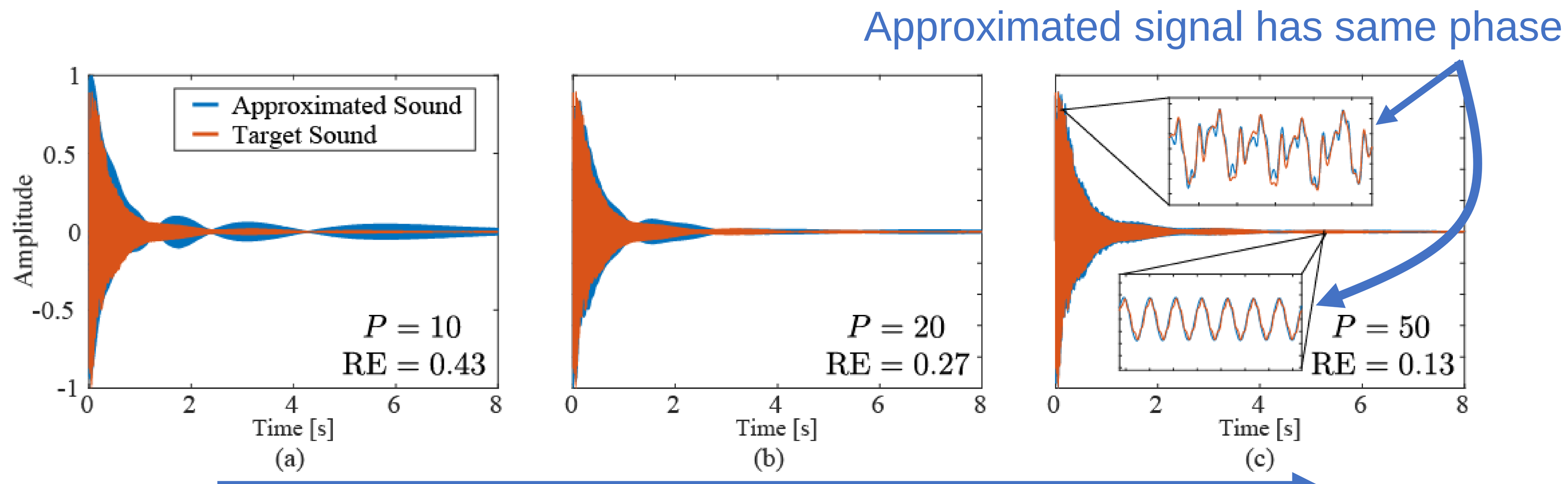
The residual signal can be utilized for excitation signal of the Kautz model.

Experiments

Signal estimation by Kautz model

- The proposed method was applied to a real piano sound of A4 (441 Hz).
- First 10 samples of the residual of sparse LPC as the excitation signal.
- P important poles were selected from estimated poles.

Approximated signal has same phase



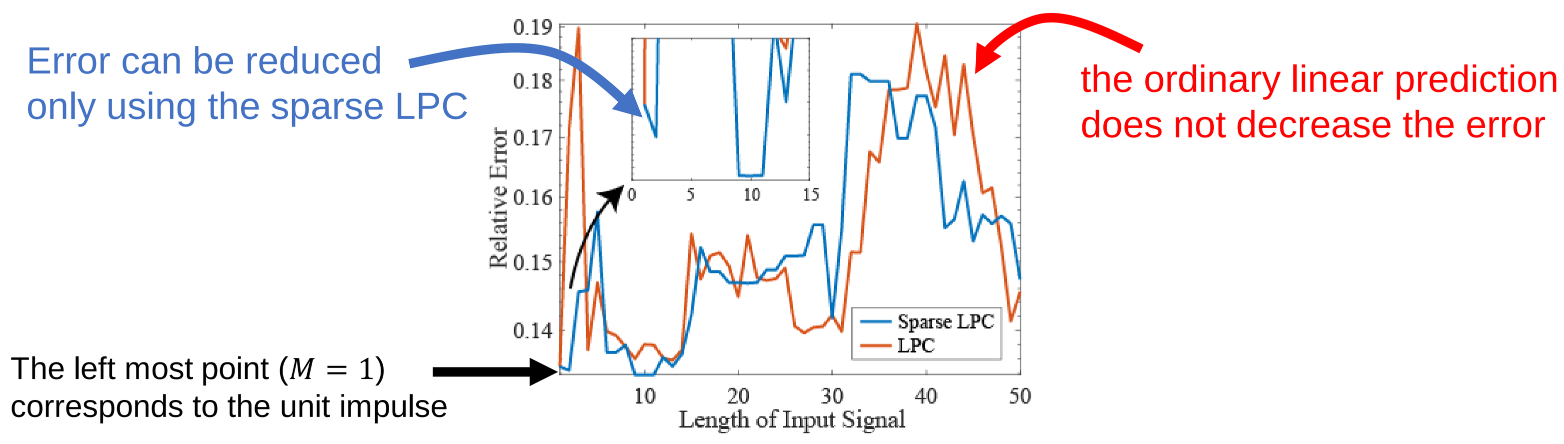
Approximation accuracy increases

Comparison with sparse LPC and ordinary LPC

- The relative error was compared with the ordinary linear prediction using least square method:

$$\underset{\{a_l\}_{l=1}^{L}}{\text{Minimize}} \sum_{n=1}^{N} \left| s(n) - \sum_{l=1}^{L} a_l s(n-l) \right|^2.$$

- First M samples of residual were used as the input excitation of the Kautz model.



Conclusion

- In this paper, a new modal based method for approximating piano sound by the Kautz model was proposed.
- The proposed method aims to resolve the two issues of the Kautz model by two sparsity-aware optimization: Sparse selection of poles using ℓ_0-constrained least square problem and Generating excitation signal based on sparse LPC.
- To see the effect of the proposed method, a real piano sound was approximated, and the sparse LPC was compared with the ordinary LPC.
- By applying the proposed method to a real piano sound, it was confirmed that the two kinds of sparsity are important for approximating piano sound.