Instrument Equalizer: An Audio Player with Real-time Instrument's Volume Controlling Functions

K. Itoyama(Kyoto Univ.), M. Goto(AIST), H. G. Okuno(Kyoto Univ.)

Background

User's demand changes from Passive Entertainment to Active Entertainment

Passive listening

listen to good sound
- high-quality speakers
- sorround system

easy to enjoy
by technology evolution

Active listening

listen to favorite sound
- composition, arrangement
- perform a instrument

difficult to enjoy
without skill or tools

▶ Technology to change the contents of musical pieces

1. Develop an audio player with instrument's volume controlling functions
2. Separate all musical instrument sounds from polyphonic musical signals

System Implementation

- Playback Slider
- Pan Controller
- Solo/Mute Switch
- Volume Controller

Separated Signals
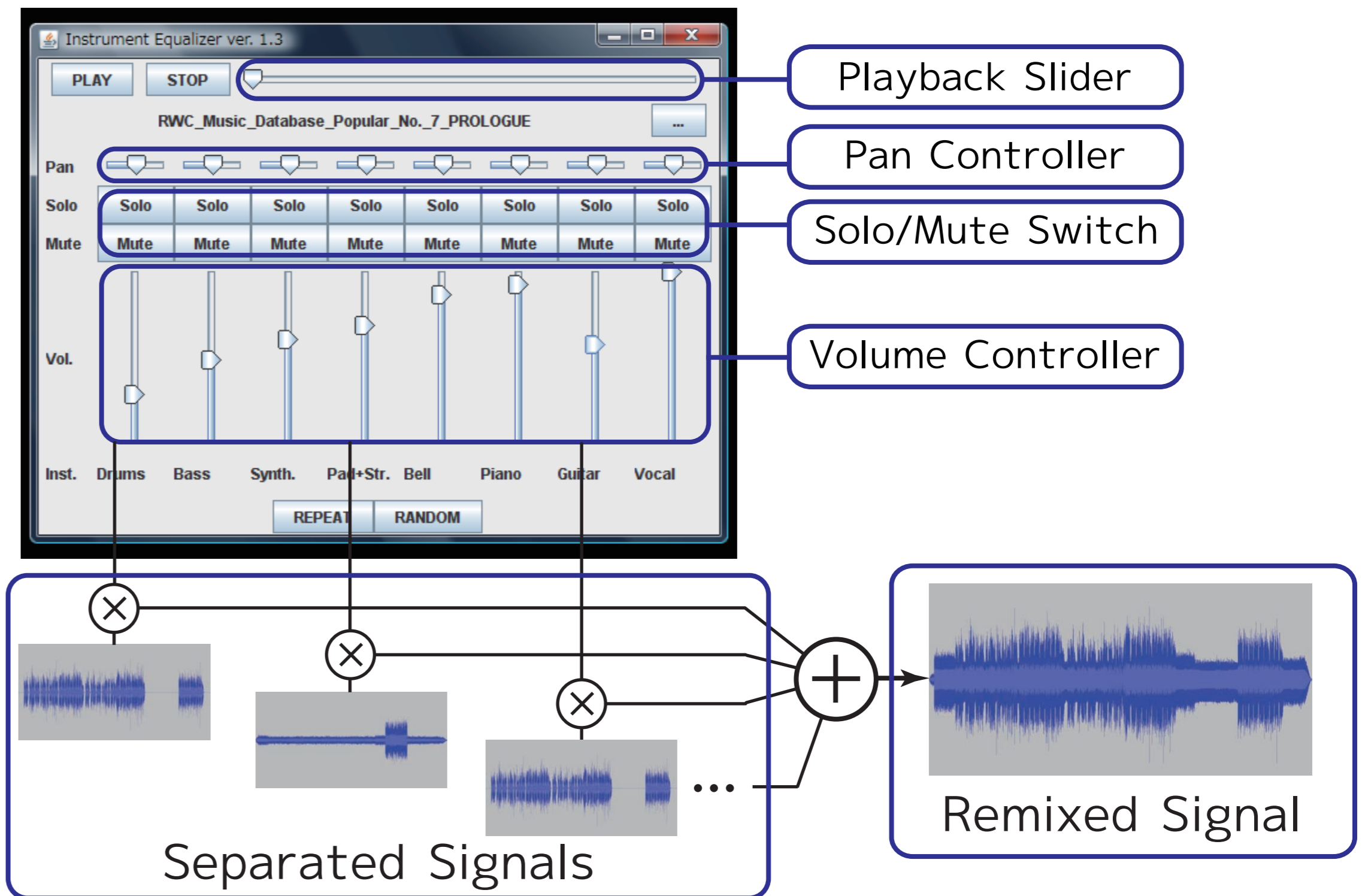
Remixed Signal

Other Applications

- Control volumes of instruments according to one's preference
  Enjoy like arranger
- Cut off a specific instrument
  Create a Minus One Music
- Re-analyze the remixed signal with different mixing balance
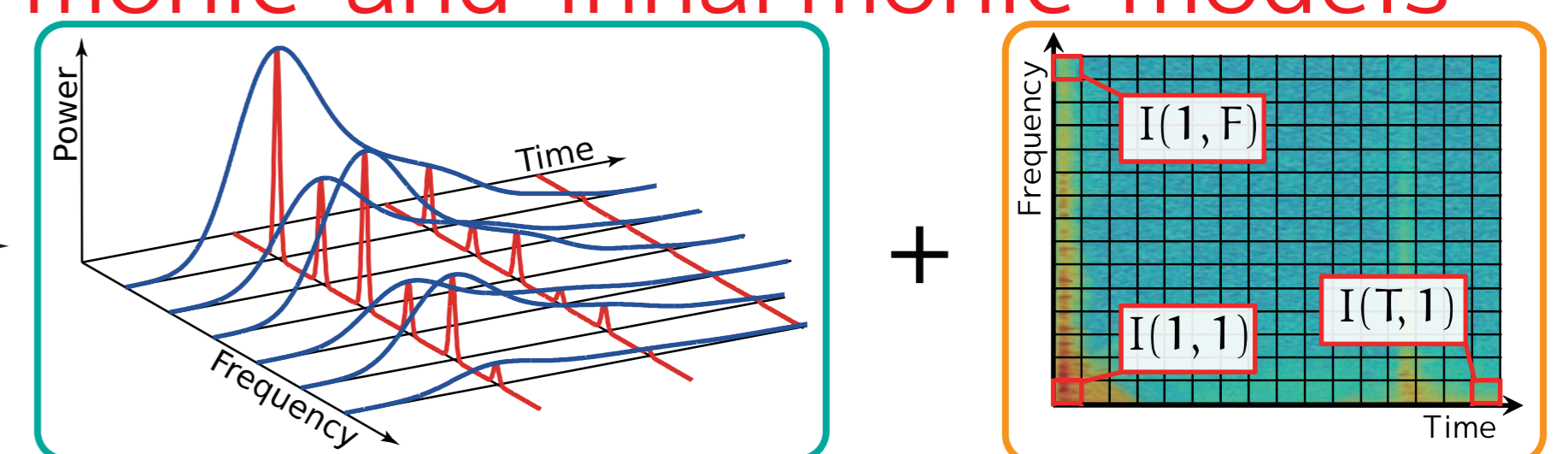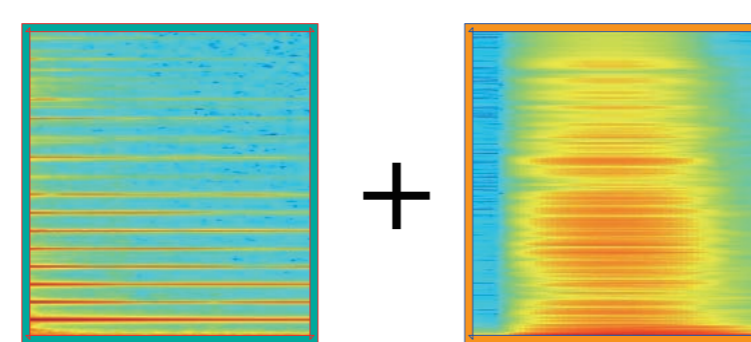 Generate an alternative query for the QBE Retrieval

Technical Problems

Problem 1: Mixture of harmonic and inharmonic sounds

▶ Construct an integrated tone model consisting of harmonic and inharmonic models

Integrated Tone Model:

$$J(t,f) = H(t,f) + I(t,f)$$

Harmonic Tone Model

$$H(t,f) = \sum_{m=0}^{M-1} \sum_{n=1}^{N} E(m,t) F(m,n,t,f)$$

Two-dimensional Gaussian Mixture Model

Temporal Power Envelope
$$E(m,t) = \frac{u(m)}{\sqrt{2\pi}\varphi} \exp\left(-\frac{(t-\tau-m\varphi)^2}{2\varphi^2}\right)$$

Frequency Component
$$F(m,n,t,f) = \frac{v(m,n)}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(f-n\omega(t))^2}{2\sigma^2}\right)$$

Inharmonic Tone Model

$$I(t,f) = \begin{bmatrix} I(1,1) & \cdots & I(T,1) \\ \vdots & & \vdots \\ I(1,F) & \cdots & I(T,f) \end{bmatrix}$$

Nonparametric Model

Problem 2: Undesirable local minumal values in the iterative algorithm

▶ Design an objective function with multiple constraints for the iteration algorithm

Observed Spectrogram: $\Delta(k,t,f)$
Integrated Model: $J_k(t,f)$
Distribution Function: $X(t,f)$
(separate $\Delta(k,t,f)$ to each note)

Minimize the Kullback-Leibler Divergence between $J_k(t,f)$ and $X(t,f)$
and costs corresponding to constraints
- Smoothness of the inharmonic model
- Intra-instrument timbre consistency
- Continuity of F0 trajectory

Input
Sound Mixture
Power Spectrogram
Musical Score

Onset Time, Pitch, Duration, Timbre

Initialize

Middle Representation
Integrated Models

Separation

Adaptation

Output
Separated Sounds