



Object-oriented analysis of pitched musical audio

Wen X. @ C4DM, EE,
Queen Mary, Univ. of London
w. Prof. Mark Sandler
Oct. 2004, Barcelona

Background

- Source separation for polyphonic music?
 - Levels of “separability”
 - Physically separable (most strict)
 - **Computationally separable**
 - Perceptually separable (most practical)
 - Proposal: toward computational separation
 - “Focusing”
-

An example

- Time-frequency focusing as “sinusoid trajectory groups”

Suggested applications

- Music material annotation
 - Instrument tracking, enhancement and suppression
 - Polyphonic transcription
 - A higher SNR
 - Itself as a more meaningful representation
-

Method

- Related work
 - McAulay & Quatieri (1986)
 - Brown & Cooke (1993)
 - Others (Serra, Ellis, Depalle, etc.)
-

Method

- Source models
 - General model: pitched sounds
 - Target model: pre-trained models encapsulating source-specific messages
-

Method

- Front-end
 - Fourier analysis
 - Frequency & amplitude estimation
 - Multi-resolution framework
-

Method

■ Analyzer

- Detect **sinusoid elements** by peak picking
 - Grouping of sinusoid elements into **trajectory groups**, in which sinusoid peaks are connected
 - Vertically under modeled harmonic property
 - Horizontally under modeled time-continuity
-

Method

- Main features
 - Object-oriented
 - Polyphonic
 - Implicit pitch detection
-

Post-processing

- “Focus” is done by locating of object in the spectrogram
 - Filtering-based applications
 - Modeling and separation of target of interest
 - Enhancement / suppression
 - Object-oriented effects
-

Problems in sight

- Time-frequency trade off for peak picking
 - Grouping criteria
 - Masked / missing partials
 - Overlapping partials & overlapping pitches
-



Thank you.

oct. 2004