Music Source Separation and its Applications to MIR

Nobutaka Ono and Emmanuel Vincent
The University of Tokyo, Japan
INRIA Rennes - Bretagne Atlantique, France

Tutorial supported by the VERSAMUS project
http://versamus.inria.fr/

Contributions from Shigeki Sagayama, Kenichi Miyamoto, Hirokazu Kameoka,
Jonathan Le Roux, Emiru Tsunoo, Yushi Ueda, Hideyuki Tachibana,
George Tzanetakis, Halfdan Rump, Other members of IPC Lab#1
Outline

- Introduction
- Part I: Brief Introduction of State-of-the-arts
  - Singer/Instrument Identification
  - Audio Tempo Estimation
- Part II: Harmonic/Percussive Sound Separation
  - Motivation and Formulation
  - Open Binary Software
- Part III: Applications of HPSS to MIR Tasks
  - Audio Chord Estimation
  - Melody Extraction
  - Audio Genre Classification
- Conclusions
Introduction

- Focus of the second half of this tutorial is to clarify
  - What source separation has been used for MIR?
  - How does it improve performance of MIR tasks?

- Examples:
  - Multi pitch estimation
    Task itself is tightly coupled with source separation.
  - Audio genre classification
    How source separation is useful? Not straightforward.
Part I: Brief Introduction of State-of-the-arts
Singer Identification

- **Task:** Identify a singer from music audio with accompaniment
- **Typical approach**

![Diagram]

- **Audio** → **Feature Extraction** → **Features** → **Classifier** → **Singer**
Accompaniment Sound Reduction [Fujihara2005]

- Pre-dominant F0 based voice separation

Audio input

1. F0 Estimation by PreFEST [Goto2004]
2. Harmonic Structure Extraction
3. Resynthesis

Feature extraction

Fig.1 [Fujihara2005]
Reliable Frame Selection [Fujihara2005]

- Only reliable frame is used for classification

![Diagram of Reliable Frame Selection](image)

Fig. 1 [Fujihara2005]
**Evaluation by Confusing Matrix**

- **Male/female confusion** is decreased by accompaniment reduction.
- **Combination of reduction and selection** much improves performance.

**Fig. 3** [Fujihara2005]
Vocal Separation Based on Melody Transcriber

- Melody-F0-based Vocal Separation [Mesaros2007]
  - Estimate melody-F0 by melody transcription system [Ryynanen2006].
  - Generate harmonic overtones at multiple of estimated F0.
  - Estimate amplitudes and phases of overtones based on cross correlation between original signal and complex exponentials.

- They evaluate the effect of separation in singer identification performance using by different classifiers.
Evaluation by Identification Rate

Singing to Accompaniment Ratio: -5dB
Singing to Accompaniment Ratio: 15dB

Generated by Table 1 and 2 [Mesaros2007]

Performance is much improved, especially in low singing-to-accompaniment ratio.
Instrument Identification

- Task: Determine instruments present in music piece
- Typical approach
  - Audio to Notes
  - Spectrogram of notes
  - Feature Extraction
  - Features
  - Classifier
  - Instrument
- Important Issue
  - Source separation is not perfect.
  - How to reduce errors?
Feature Weighting [Kitahara2007]

- Feature vectors of each instrument are collected from polyphonic music for training.

- Robustness of each feature is evaluated by ratio of intra-class variance to inter-class variance: Applying Linear discriminant analysis (LDA) for feature weighting.

Modified from Fig. 1 [Kitahara2007]
Effectiveness of Feature Weighting

Feature weighting by LDA improves recognition rate.
Audio Tempo Estimation

- **Task:** Extract tempo from musical audio
- **Typical approach:**

  ![Diagram](diagram.png)

  - Audio
  - STFT or Filterbank
  - Subband signals
  - Onset Detection
  - Detection function
  - Tracking
  - Tempo candidates
  - Periodicity Analysis
  - Tempo
  - Time (t)
Applying Harmonic+Noise Model

- Harmonic+Noise model is applied before calculating detection function [Alonso2007]

![Diagram of source separation based on harmonic + noise model](image)

Source separation based on harmonic + noise model

Detection functions are calculated from both of harmonic component and noise component, and then, they are merged.

Fig. 2 [Alonso2007]
Influence of H+N Model

Separation based on H+N model shows better results.

Algorithms of periodicity detection

Fig. 14 [Alonso2007]
Applying PLCA

- PLCA (Probabilistic Latent Component Analysis), NMF-like method is applied.
- It increases much candidates of tempo.
- They report its effectiveness.

Fig. 1 [Chordia2009]
Part II: Harmonic/Percussive Sound Separation
Motivation and Goal of HPSS

- **Motivation**: Music consists of two different components

![Spectrogram example of a popular music (RWC-MDB-P034)](image)

- **Goal**: Separation of a monaural audio signal into harmonic and percussive components

- **Target**: MIR-related tasks
  - multi-pitch analysis, chord recognition… H-related
  - beat tracking, rhythm recognition… P-related
Related Works to H/P Separation

- Source separation into multiple components followed by classification
  - ICA and classification [Uhle2003]
  - NMF and classification [Helen2005]

- Steady + Transient model
  - Adaptive phase vocoder
  - Subspace projection
  - Matching pursuit
  - ...etc

  Good review is provided in [Daudet2005]

- Bayesian NMF [Dikmen2009]
Point: Anisotropy of Spectrogram

- **horizontally smooth**
- **vertically smooth**

- **harmonic component**
- **percussive component**
H/P Separation Problem

Problem:
Find \( H_{t,\omega} \) and \( P_{t,\omega} \) from \( W_{t,\omega} \) on power spectrogram

Requirements:
1) \( H_{t,\omega} \) : horizontally smooth
2) \( P_{t,\omega} \) : vertically smooth
3) \( H_{t,\omega} \) and \( P_{t,\omega} \) : non-negative
4) \( H_{t,\omega} + P_{t,\omega} \) : should be close to \( W_{t,\omega} \)
Formulation of H/P Separation (1/2)

Formulation as an Optimization Problem:

Objective function to minimize

\[ J(H, P) = D(W, H + P) + C_H(H) + C_P(P) \]

Under constraints:

- \( H_t, \omega \geq 0 \)
- \( P_t, \omega \geq 0 \)

In MAP estimation context, they are corresponding likelihood term and prior term, respectively.
Formulation of H/P Separation (2/2)

- **Closeness cost function:** I-divergence

\[
D(W, H + P) = \sum_{\omega, \tau} \left\{ W_{\omega, \tau} \log \frac{W_{\omega, \tau}}{H_{\omega, \tau} + P_{\omega, \tau}} - W_{\omega, \tau} + H_{\omega, \tau} + P_{\omega, \tau} \right\}
\]

- **Smoothness cost function:** Square of difference

\[
C_H(H) = \sum_{\omega, \tau} \frac{1}{2\sigma_H^2} (H_{\omega-1, \tau}^{\gamma} - H_{\omega, \tau}^{\gamma})^2 \quad \gamma = 0.5 \\

C_P(H) = \sum_{\omega, \tau} \frac{1}{2\sigma_P^2} (P_{\omega-1, \tau}^{\gamma} - P_{\omega, \tau}^{\gamma})^2
\]

Weights to control two smoothness

- A variance modeling based separation using
  - Poisson observation distribution
  - Gaussian continuity priors

[Miyamoto2008, Ono2008, etc]
Update Rules

- Update alternatively two kinds of variables:
  - H and P:
    \[
    H_{\omega,\tau} \leftarrow \left( \frac{b_{H,\omega,\tau} + \sqrt{b_{H,\omega,\tau}^2 + 4a_{H,\omega,\tau}c_{H,\omega,\tau}}}{2a_{H,\omega,\tau}} \right)^2
    \]
    \[
    P_{\omega,\tau} \leftarrow \left( \frac{b_{P,\omega,\tau} + \sqrt{b_{P,\omega,\tau}^2 + 4a_{P,\omega,\tau}c_{P,\omega,\tau}}}{2a_{P,\omega,\tau}} \right)^2
    \]
  - Auxiliary variables:
    \[
    m_{P,\omega,\tau} = \frac{P_{\omega,\tau}}{H_{\omega,\tau} + P_{\omega,\tau}}
    \]
    \[
    a_{P,\omega,\tau} = \frac{2}{\sigma_P^2} + 2
    \]
    \[
    b_{P,\omega,\tau} = \frac{(\sqrt{P_{\omega-1,\tau}} + \sqrt{P_{\omega+1,\tau}})}{\sigma_P^2}
    \]
    \[
    c_{P,\omega,\tau} = 2m_{P,\omega,\tau}W_{\omega,\tau}
    \]
    \[
    m_{H,\omega,\tau} = \frac{H_{\omega,\tau}}{H_{\omega,\tau} + P_{\omega,\tau}}
    \]
    \[
    a_{H,\omega,\tau} = \frac{2}{\sigma_H^2} + 2
    \]
    \[
    b_{H,\omega,\tau} = \frac{(\sqrt{H_{\omega,\tau-1}} + \sqrt{H_{\omega,\tau+1}})}{\sigma_H^2}
    \]
    \[
    c_{H,\omega,\tau} = 2m_{H,\omega,\tau}W_{\omega,\tau}
    \]
## Separation Examples

<table>
<thead>
<tr>
<th>Music piece</th>
<th>original</th>
<th>H</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>RWC-MDB-P-7 “PROLOGUE”</td>
<td>🎧</td>
<td>🎧</td>
<td>🎧</td>
</tr>
<tr>
<td>RWC-MDB-P-12 “KAGE-ROU”</td>
<td>🎧</td>
<td>🎧</td>
<td>🎧</td>
</tr>
<tr>
<td>RWC-MDB-P-18 “True Heart”</td>
<td>🎧</td>
<td>🎧</td>
<td>🎧</td>
</tr>
<tr>
<td>RWC-MDB-P-25 “tell me”</td>
<td>🎧</td>
<td>🎧</td>
<td>🎧</td>
</tr>
<tr>
<td>RWC-MDB-J-16 “Jive”</td>
<td>🎧</td>
<td>🎧</td>
<td>🎧</td>
</tr>
</tbody>
</table>
Real-Time Implementation

- Sliding Block Analysis

Iterations are applied only within sliding block
Open Software: Real-time H/P equalizer

Available at http://www.hil.t.u-tokyo.ac.jp/software/HPSS/

- Control H/P balance of audio signal in real time
- Simple instructions:
  1) Click “Load WAV” button and choose a WAV-formatted audio file.
  2) Click “Start” button, and then, audio starts.
  3) Slide H/P balance bar as you like and listen how the sound changes.

Available at http://www.hil.t.u-tokyo.ac.jp/software/HPSS/
Part III: Applications of HPSS to MIR Tasks

III-1: Audio Chord Detection
Audio Chord Detection

- **Task:** Estimate chord sequence and its segmentation from music audio
Typical Approach: Chroma Feature + HMM

- **Feature:** chroma [Fujishima1999]
  - Chroma observation probability \( p(x_t | c_t) \)
- **Transition:** chord progression
  - Bigram probability \( p(c_t | c_{t-1}) \)
- **Maximum a Posteriori Chord Estimation** [Sheh2003]
  - Viterbi algorithm for …

\[
\text{argmax}_{c} p(x_0|c_0)p(c_0) \prod_{t=1}^{T} p(x_t|c_t)p(c_t|c_{t-1})
\]

Initial prob. emission transition

\[
p(x_t | c_t)
\]

\[
p(c_t | c_{t-1})
\]
Feature-refined System [Ueda2009]

Feature Extraction

Audio → STFT, HPSS, iSTFT

Percussion-suppressed Audio → Harmonic-suppressed Audio

Constant Q Transform Chroma Transform

Chroma Vector Candidates

Tuning Compensation

Tuned Chroma Vectors

Regression Analysis → Delta Chroma Vectors

DFT → Fourier-Transformed Chroma Vector

HMM-based chord recognition

24 dim. features

HMM training

Viterbi decoding

acoustic model language model

Recognized chord sequence

training → recognition
 Suppressing Percussive Sounds

- Percussive sounds are harmful in chord detection

Emphasize harmonic components by HPSS
Fourier-transformed Chroma

- Covariance matrix of chroma
  - Highly correlated components: diagonal-only approximation infeasible
    - Caused by harmonic overtones or some pitches performed at the same time
    - Results in large number of parameters

- Covariance matrix is near circulant
  - Assuming …
    - Harmonic overtones of all pitches have the same structure
    - The amount of occurrence of the same intervals is the same
  - Circulant matrix diagonalized by DFT

- Diagonal approximation of FT-Chroma covariance
  - Reduces the number of model parameters (statistically robust)
Tuning Compensation

- Tuning difference among songs
  - Neglecting this may blur chroma features
- Choose best tuning from multiple candidates
  - Find maximum chroma energy (sum of all bins of chroma)
  - Assume: tuning does not change within a song
Delta Chroma Features

- Improve chord boundary accuracy
  - by features representing chord boundaries

- Chord tones largely changes at chord boundary
  - Delta chroma: derivative of chroma features
  - Cf. Delta cepstrum (MFCC): Effective features of speech recognition

- Calculated by regression analysis of $\delta$ sample points
  [Sagayama&Itakura1979]

- Robust to noise

\[
\Delta C(i,n) = \frac{\sum_{k=-\delta}^{\delta} k w_k C(i, t+k)}{\sum_{i=-\delta}^{\delta} k^2 w_k}
\]

\[
i = 1, \ldots, 12
\]
Multiple States per Chord

- Chroma changes from “onset” to “release”
  - capture the change by having multiple states per chord
  - tradeoff between data size and the number of states

![Diagram of multiple states per chord](image)
Experimental Evaluation

Test Data
- 180 songs (12 albums) of The Beatles (chord reference annotation provided by C. Harte)
- 11.025 kHz sampling, 16bit, 1ch, WAV file
- Frequency range: 55.0Hz-1661.2Hz (5 octaves)

Labels
- $12 \times \text{major/minor} = 24 \text{ chords} + N \text{ (no chord)}$

Evaluation
- Album filtered 3-fold cross validation
  - 8 albums for training, 4 albums for testing
- Frame Recognition Rate
  - $\frac{\#\text{correct frames}}{\#\text{total frames}}$
- Sampled every 100ms
Chord Detection Results

HE: harmonic sound emphasized
TC: tuning compensation
FT: FT chroma (diagonal covariance)
DC: Delta chroma

Err Reduc Rate 28.1%
Err Reduc Rate 11.0%
MIREX2008 best score [Uchiyama2008]

HPSS improves chord detection performance

Chroma HE HE+TC HE+TC+FT HE+TC+DC
1 state
2 states
3 states
Part III: Applications of HPSS to MIR Tasks

III-2: Melody Extraction
Melody Extraction

- Task: Identify a melody pitch contour from polyphonic musical audio
- Typical approach:

![Diagram showing the process of melody extraction](attachment:attachment)

- Singing voice enhancement will be useful pre-processing.
Singing Voice in Spectrogram

RWC-MDB-P-25 “tell me”

A. Vertical component: Percussion
B. Horizontal component: Harmonic instrument (piano, guitar, etc.)
C. Fluctuated component: Singing voice
Is Voice Harmonic or Percussive?

 Depends on spectrogram resolution (frame-length)

- On short-frame STFT domain, voice appears as “H” (time direction clustered).
- On long-frame STFT domain, voice appears as “P” (frequency direction clustered).
**HPSS Results with Different Frame Length**

Example

Frame length: 16ms

Frame length: 512ms
Two-stage HPSS [Tachibana2010]

Original

HPSS with short frame

Sinusoidal Sound

Percussive Sound

HPSS with long frame

Stationary-sinusoidal Sound

Fluctuating-sinusoidal Sound (≒singing voice)
Spectrogram Example

Original signal (from LabROSA dataset)
Spectrogram Example

Voice-enhanced signal (by two-stage HPSS)
# Separation Examples

<table>
<thead>
<tr>
<th>title</th>
<th>original</th>
<th>Extracted Vocal</th>
<th>Vocal Cancelled*</th>
<th>Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>“tell me”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>F, R&amp;B</td>
</tr>
<tr>
<td>“Weekend”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>F, Euro beat</td>
</tr>
<tr>
<td>“Dance Together”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>M, Jazz</td>
</tr>
<tr>
<td>“1999”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>M, Metal rock</td>
</tr>
<tr>
<td>“Seven little crows”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>F, Nursery rhyme</td>
</tr>
<tr>
<td>“La donna è mobile” from Verdi’s opera “Rigoletto”</td>
<td></td>
<td>○</td>
<td>○</td>
<td>M, Classical</td>
</tr>
</tbody>
</table>

Aug. 9, 2010 ISMIR2010 Tutorial 1 48
Melody Tracking by DP [Tachibana2010]

- Estimating hidden states by dynamic programming

Observation (Voice-enhanced-Spectrum)

State (Pitch series)
Example of Melody Tracking

- train06.wav, distributed by LabROSA database
Results in MIREX 2009

- Data: 379 songs, mixed in +5 dB, 0dB, and -5 dB.

Robustness to large singer-to-accompaniment ratio is greatly improved.

<table>
<thead>
<tr>
<th>Accompaniments</th>
<th>+5dB</th>
<th>0dB</th>
<th>-5dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>original</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>processed</td>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>

HPSS-based method

Noise Robust 😊

Sensitive

RawPitchAccuracy
Part III: Applications of HPSS to MIR Tasks

III-3: Audio Genre Classification
Audio Genre Classification

- **Task:** estimate genre from music audio
  - Blues, classical, jazz, rock, ...

- **Typical approach**

- **Example of features** [Tzanetakis2001]
  - Timbral information (MFCC, etc.)
  - Melodic information
  - Statistics about periodicities: Beat histogram
New Features I: Percussive Patterns

Musical Audio Signal → Harmonics / Percussions Separation

Percussion-Emphasized Spectrogram → Rhythmic Structure

Rhythm Map → Updating Rhythmic Structure

Rhythmic Structure → Unit Patterns

Unit Patterns → Updating Unit Patterns

Updating Unit Patterns → Bar-long Patterns

Feature Extraction [Tsunoo2009]
Motivation for Bar-long Percussive Patterns

- Bar-long percussive patterns (temporal information) are frequently characteristic of a particular genre

- Difficulties
  1) Mixture of harmonic and percussive components
  2) Unknown bar-lines
  3) Tempo fluctuation
  4) Unknown multiple patterns

A A A A B A A A C C C C
Rhythmic Structure Analysis by One-pass DP algorithm

- Assume that correct bar-line unit patterns are given.
- Problem: tempo fluctuation and unknown segmentation
  - Analogous to continuous speech recognition problem
  - One-pass dynamic programming algorithm can be used to segment
Dynamic Pattern Clustering [Tsunoo2009]

- Actually, unit patterns also should be estimated.
  - Chicken-and-egg problem
  - Analogous to unsupervised learning problem
- Iterative algorithm based on k-means clustering
  - Segment spectrogram using one-pass DP algorithm
  - Update unit patterns by averaging segments
- Convergence is guaranteed mathematically
Example of “Rhythm Map”

Rhythm 1 (Fundamental)

Rhythm 2 (Fill-in)

Rhythm 3 (Interlude)

Rhythm 4 (Climax)

One-pass DP alignment

Fundamental melody

Climax

Interlude

FullSong
Necessity of HPSS in Rhythm Map

With HPSS

Without HPSS

Rhythm patterns and structures are not extracted without HPSS!
Extracting Common Patterns to a Particular Genre

- Apply to a collection of music pieces
- Alignment calculation by one-pass DP algorithm
  - Use same set of templates
- Updating templates by $k$-means clustering
  - Use whole music collection of a particular genre

![Diagram showing templates and alignment between two songs](image)
Features and Classifiers

- Feature Vectors:
  Genre-pattern Occurrence Histogram (normalized)
- Classifier: Support Vector Machine (SVM)
Experimental Evaluation

- **Dataset**
  - (standard)
    - GTZAN dataset
    - 22050Hz sampling, 1ch
    - 30 seconds clips
    - 10 genres
      - {blues, classical, country, disco, hiphop, jazz, metal, pop, reggae, rock}
    - 100 songs per genre: total 1000 songs
  - (rhythm-intensive)
    - Ballroom dataset
    - 22050Hz sampling, 1ch
    - 30 seconds clips
    - 8 styles
      - {chacha, foxtrot, quickstep, rumba, samba, tango, viennesewaltz, waltz}
    - 100 songs per style: total 800 songs

- **Evaluation**
  - 10-fold cross-validation
  - Classifier: linear SVM (toolkit “Weka” used)
Extracted Percussive Patterns

- **Pattern set**
  - Divided the datasets into 2 parts and obtained 2 sets of 10 templates for each genre

- **Example of learned templates**

<table>
<thead>
<tr>
<th>classical</th>
<th>country</th>
<th>disco</th>
<th>hiphop</th>
<th>jazz</th>
<th>metal</th>
<th>pop</th>
<th>reggae</th>
<th>rock</th>
</tr>
</thead>
</table>

10 templates learned from “blues”
Genre Classification Accuracy

- Percussive pattern feature only

<table>
<thead>
<tr>
<th>Features [number of dim.]</th>
<th>GTZAN dataset</th>
<th>Ballroom dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (Random)</td>
<td>10.0%</td>
<td>12.5%</td>
</tr>
<tr>
<td>Rhythmic (from template set #1) [10/8]</td>
<td>43.6%</td>
<td>54.0%</td>
</tr>
<tr>
<td>Rhythmic (from template set #2) [10/8]</td>
<td>42.3%</td>
<td>55.125%</td>
</tr>
</tbody>
</table>

- Merged with timbral features
  - Statistic features such as MFCC, etc. (68 dim.) [Tzanetakis 2008]
  - Performed well on audio classification tasks in MIREX 2008

<table>
<thead>
<tr>
<th>Features [number of dim.]</th>
<th>GTZAN dataset</th>
<th>Ballroom dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing (Timbre) [68]</td>
<td>72.4%</td>
<td>57.625%</td>
</tr>
<tr>
<td>Merged (from template set #1) [78/76]</td>
<td>76.1%</td>
<td>70.125%</td>
</tr>
<tr>
<td>Merged (from template set #2) [78/76]</td>
<td>76.2%</td>
<td>69.125%</td>
</tr>
</tbody>
</table>

Classification accuracy is improved by combining percussive pattern features.
New Features II: Bass-line Patterns

[Tsunoo2009]
Examples of Extracted Bass-line Patterns

Jazz

Blues

Rock

Hiphop
Genre Classification Accuracy

<table>
<thead>
<tr>
<th>Features</th>
<th>GTZAN dataset</th>
<th>Ballroom dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (random classifier)</td>
<td>10.0%</td>
<td>10.0%</td>
</tr>
<tr>
<td>Only bass-line (400 dim.)</td>
<td>42.0%</td>
<td>44.8%</td>
</tr>
<tr>
<td>Existing (timbre, 68 dim.)</td>
<td>72.4%</td>
<td>72.4%</td>
</tr>
<tr>
<td>Merged (468 dim.)</td>
<td>74.4%</td>
<td>76.0%</td>
</tr>
</tbody>
</table>
Another Application of HPSS [Rump2010]

- Autoregressive MFCC Model applied to Genre Classification
- HPSS increases the number of channels mono -> three (original, harmonic, percussive) and improves performance
Conclusions

Source separation techniques used to MIR
- F0-based harmonic separation
- Non-negative matrix factorization or PLCA
- Sinusoid + Noise model
- Harmonic/percussive sound separation

Source separation is useful
- To enhance specific components
- To increase the number of channels and the dimension of feature vectors
- To generate new features
Future Works

- Application of source separation to other MIR tasks
  - Cover song identification, audio music similarity,...
- Improvement of separation performance itself by exploiting musicological knowledge
- Using spatial (especially stereo) information
  - Current works are limited to monaural separation
- Feature weighting technique for overcoming errors due to imperfect source separation
Reference Book Chapter


Available Separation Softwares

- Harmonic Percussive Sound Separation (HPSS)
  - http://www.hil.t.u-tokyo.ac.jp/software/HPSS/

- ICA Central: Early software restricted to mixtures of two sources

- SiSEC Reference Software: Linear modeling-based software for panned or recorded mixtures

- QUAERO Source Separation Toolkit: Modular variance-modeling based software implementing a range of structures: GMM, NMF, source-filter model, harmonicity, diffuse mixing, etc
  - To be released Fall 2010: watch the music-ir list for an announcement!
Advertisement: LVA/ICA 2010

- LVA/ICA 2010 will be held in St. Malo, France on September 27-30, 2010.
- More than 20 papers on music and audio source separation will be presented.
References

Singer/Instrument Identification


References

Audio Tempo Estimation


Related Works to H/P Separation

References

Harmonic/Percussive Sound Separation

References

- Applications of HPSS to MIR Tasks
Applications of HPSS in MIR Tasks


