Reconstructing Completely Overlapped Notes from Musical Mixtures

Jinyu Han & Bryan Pardo

Northwestern University, EECS Department, Evanston, IL, USA.

jinyuhan@gmail.com • pardo@northwestern.edu • http://music.cs.northwestern.edu/

Introduction

Systems that perform source separation on recordings of music often encounter situations where a harmonic sound is almost completely overlapped in time and frequency (e.g., when two instruments play one octave apart). Common Amplitude Modulation is an effective method to resolve the problem of overlapped harmonics. It, however, relies on non-overlapped harmonics from the same note being available. We propose a musical sound separation system for monaural recordings that explicitly deals with the “completely overlapped” notes problem, based on Harmonic Temporal Envelope Similarity. We learn a harmonic envelope model for each instrument from the non-overlapped harmonics of the same instrument, wherever they occur in the recording. This model is used to reconstruct the harmonic envelopes for overlapped harmonics.

Review

► Common Amplitude Modulation (CAM)
CAM assumes the harmonic amplitude envelopes from the same note are correlated. CAM holds most of the time for the first few strong harmonics but fails to hold for those with weak energy.

► Harmonic Temporal Envelope Similarity (HTES)
Instead of using CAM to estimate harmonic envelopes, we utilize the property that notes played by the same instrument within a short period of time (e.g., within musical phrase boundaries) have similar harmonic envelopes. We call this Harmonic Temporal Envelope Similarity.

Evaluation

Our proposed method using HTES produces much better envelope estimates for the completely overlapped notes than the CAM does.

► Experimental Setup
• 10 Bach chorale recordings of about 330 seconds of audio.
• Mixtures of two instruments with one instrument (bassoon) playing the bass line and the other playing the alto (clarinet or trumpet) or soprano line (violin) of a Bach chorale.
• The proposed system was compared to a musical sound separation system (denoted LWW) based on CAM

► Experimental Results

<table>
<thead>
<tr>
<th>Mixtures</th>
<th>Proposed</th>
<th>LWW</th>
<th>Proposed</th>
<th>LWW</th>
<th>Proposed</th>
<th>LWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarinet</td>
<td>0.92</td>
<td>0.64</td>
<td>0.90</td>
<td>0.64</td>
<td>0.87</td>
<td>0.64</td>
</tr>
<tr>
<td>Trumpet</td>
<td>0.88</td>
<td>0.49</td>
<td>0.90</td>
<td>0.49</td>
<td>0.87</td>
<td>0.49</td>
</tr>
<tr>
<td>Violin</td>
<td>0.86</td>
<td>0.64</td>
<td>0.88</td>
<td>0.64</td>
<td>0.87</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Table: Performance of the proposed system and the LWW. Numbers in bold indicate the difference between proposed method and LWW are statistically significant.

Conclusion

We proposed a monaural musical sound separation system that explicitly deals with completely overlapped notes in music recordings. Our approach is based on Harmonic Temporal Envelope Similarity, a new assumption on instrumental harmonic envelope we observed from real audio data. Experimental results show the proposed method achieves better separation performance than a state-of-art monaural music separation system that only exploits CAM.

This work was supported by NSF grant numbers 0643752.