Isolation Of Individual Tracks From Polyphonic Music (Research Paper)

Nick Starr

10/31/2008
Computer Systems Lab 2008-2009
Period 1

Source separation is an example of a classic situation in computer science: even the average human can perform it extraordinarily well, but it's extremely challenging to automate the process for a computer to solve. Our purpose is to develop a source separation program to apply to musical input. First, we convert the signal to the frequency domain; then, we split up the frequency into observed components via Singular Value Decomposition (SVD) and apply Independent Component Analysis (ICA) to that to retrieve individual frequency components. Third, those components are classified somehow, and finally, they are recombined and sent back to the time domain, but now separated by the classification in step 3. We hope to be able to classify music based on what instrument is playing it, which would enable us to extract any instrumental part from a piece of music.

The key technique behind our program is the Independent Component Analysis. solves the problem of being given a vector **x** observation signals, another vector **s** of source signals, and a square matrix **A** mixing constants, such that **s** = **Ax**. The goal of ICA is to determine **A**'s inverse, so that **A**^-1**s** = **A**^-1**Ax** = **x**. However, ICA requires that there are as many sources as a observed signals, but in reality, the best we will usually get for observed signals is 2 (stereo music), when we may be looking for many different sources. So, we use the Fourier Transform combined with Singular Value Decomposition. The Fourier Transform takes our signal and transforms it to the frequency domain. However, the values in the frequency domain are complex numbers, with a magnitude and a phase. Only the magnitude is necessary for the ICA computation, so the phases are simply saved for the recombination step. So, given the list of magnitudes, we apply the technique of Singular Value Decomposition. For some matrix $\mathbf{X}$, SVD gives us $\mathbf{X}^T = \mathbf{U}*\mathbf{D}*\mathbf{V}^T$. $\mathbf{D}$is a diagonal matrix of singular values in decreasing order, $\mathbf{U}$= ($\mathbf{u}$, $\mathbf{u}$, $\mathbf{u}$ ..... $\mathbf{u}$) where the $\mathbf{u}$are the eigenvectors of $\mathbf{X}\mathbf{X}^T$, and $\mathbf{V}$ = ($\mathbf{v}$, $\mathbf{v}$, $\mathbf{v}$, ... $\mathbf{v}$) where the $\mathbf{v}$are the eigenvectors of $\mathbf{X}^T\mathbf{X}$. Now that we have useful individual frequency components, we can perform ICA on those components. After we retrieve the sources, we classify them somehow into different categories. The paper that we are basing this off of looked at which frequencies in music are typically contributed by percussion instruments, and splitting the components based on that. Once the components have been separated, each group of components is recombined (using the phase data) with an Inverse Fourier Transform, resulting in the original musical track being split up into the desired components.