Automatic Chord Detection

After our book was published, we discovered some interesting research on the problem of automatic detection of chords in recorded music. This paper describes our study of the free software that implements the results of this research.

Introduction

It is a challenging problem in audio analysis to identify chords being played from a piece of recorded music. Here we will briefly describe the free software, CHORDINO, which works together with audio programs like AUDACITY to automatically produce transcriptions of the chords within a musical recording. In the final section of this paper, we discuss how to download a free copy of CHORDINO, and how to use it in AUDACITY.

In his Ph.D. thesis (Mauch, 2010), Matthias Mauch describes how to match chords based on the harmonics found within spectrograms of recorded music. There is a concise summary of his approach available at

$$\texttt{http://www.youtube.com/watch?v=gxCbcQk3r68} \tag{1}$$

The basic procedure is to match chord notes using the following model for the amplitude values of the harmonics of each note:

$$A_k = A_0\, c^k, \;\; k = 0, 1, 2, \ldots \tag{2}$$

In this formula, c is a positive parameter called the *spectral shape.* The quantity A_k stands for the amplitude of the $(k+1)^{\text{st}}$ harmonic, with A_0 standing for the amplitude of the note's fundamental. The spectral shape's value lies between 0.5 and 0.9, with a default value of 0.7. On the left of Figure 1, we show an example of (2), using the value 0.7 for the spectral shape parameter. When a fundamental of a second note overlaps with an overtone harmonic for a first note, then this second note can be detected due to the failure of the amplitudes in the spectrogram to follow the single note model. We show an example of this, for the notes C_4 and G_4, on the right of Figure 1.

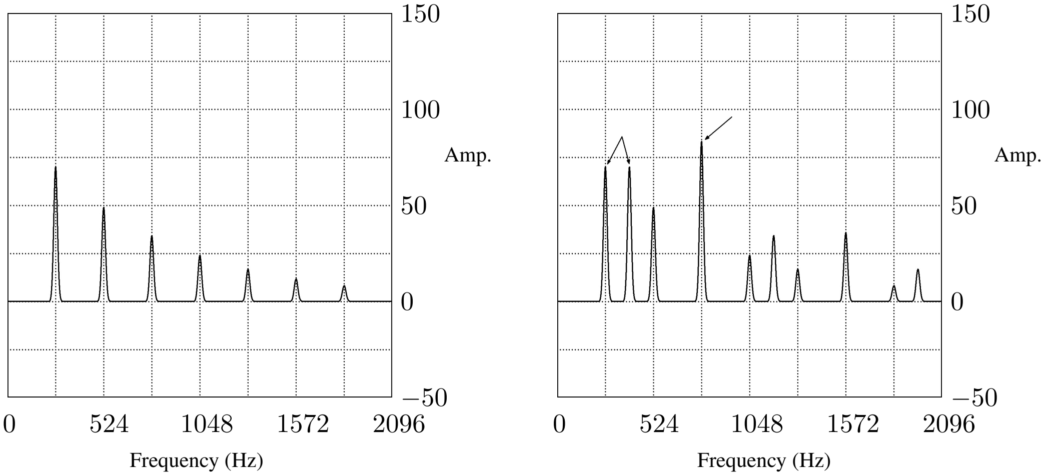

Figure 1. Left: Harmonics for single note C_4. Right: Harmonics for the dyadic chord C^5, having notes C_4 and G_4. The double arrow points to the two fundamentals, while the single arrow points to a larger amplitude harmonic corresponding to a combination of harmonics from the two notes.

When there are more notes in a chord, then the pattern of harmonics and their overlaps becomes more complex. For example, suppose the chord is a C-major chord in root position, say C_4-E_4-G_4. Then, the first

6 harmonics match the root's harmonics, and can be detected by their violation of the single note model. We show an example illustrating this on the left of Figure 2. As another example, suppose the chord is a first inversion of a C-major chord, E_4-G_4-C_5. On the right of Figure 2, we show that there is a characteristic pattern for the amplitudes of the harmonics of this chord. These patterns of amplitudes for the harmonics of chords are like "fingerprints," allowing for the identification of these chords within a spectrogram of recorded music. Of course, the identification process is dependent on the amplitudes following the model described by (2). This model is a common generic model for amplitude used in audio engineering. Not all amplitudes for musical tones follow it, however. This makes automatic chord detection a very challenging problem, one that is the subject of ongoing research. Nevertheless, we shall see that CHORDINO did provide good performance on our test examples. These test examples consist of 1) some chords with vertically stacked notes, 2) some arpeggiated chords, and 3) a more rhythmically and harmonically complex set of chords.

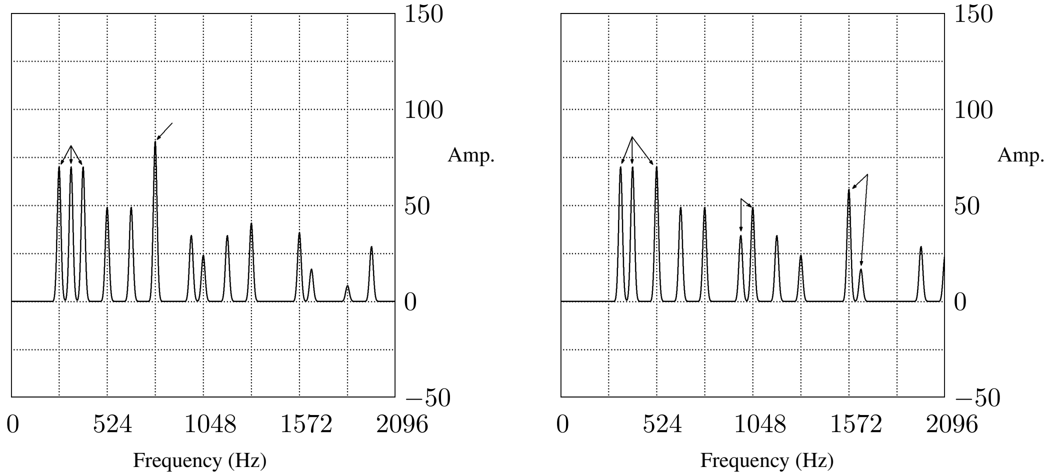

Figure 2. Left: Harmonics for C-major chord in root position, C_4-E_4-G_4. Triple arrow points to the three fundamentals, while the single arrow points to a larger amplitude harmonic corresponding to a combination of harmonics from the C and G notes. Right: Harmonics for first inversion of C-major chord, E_4-G_4-C_5. Triple arrow points to the three fundamentals. Double arrow in the center points to harmonics corresponding to pitches B_5 and C_6. Double arrow on the right points to harmonics corresponding to pitches G_6 and $\mathrm{G}_6^\sharp$.

1.1 Example 1: Vertically Stacked Chords

Our first example uses CHORDINO to analyze four basic chords. These chords are shown at the top of Figure 3.[1] As stated in the caption, the chords are labeled only by their pitch classes. So, for example, the second chord is a first inversion of a G-major chord (G/B in lead sheet notation). Since, its pitch classes are $\mathbf{G}$, $\mathbf{B}$, $\mathbf{D}$, we have simply notated it as G. The spectrogram is shown as well, because it is easy to see the changes in harmonics for the different chords, and because CHORDINO uses information about these harmonics in order to make its chord determinations. As shown at the bottom of Figure 3, CHORDINO was able to identify all the chords correctly (in terms of their pitch classes), using its default value of 0.7 for the spectral shape parameter. (The notation **N** stands for "no chord" at that time.) For our next example, we shall see that CHORDINO is not quite as successful for the much more challenging problem of identifying arpeggiated chords.

[1] The recording can be accessed at

`http://www.uwec.edu/walkerjs/mathematicsandmusic/Nav/ClipsForChordino.htm`

It was created from a MUSESCORE version of the score in Figure 3, saved as a `*.wav` audio file.

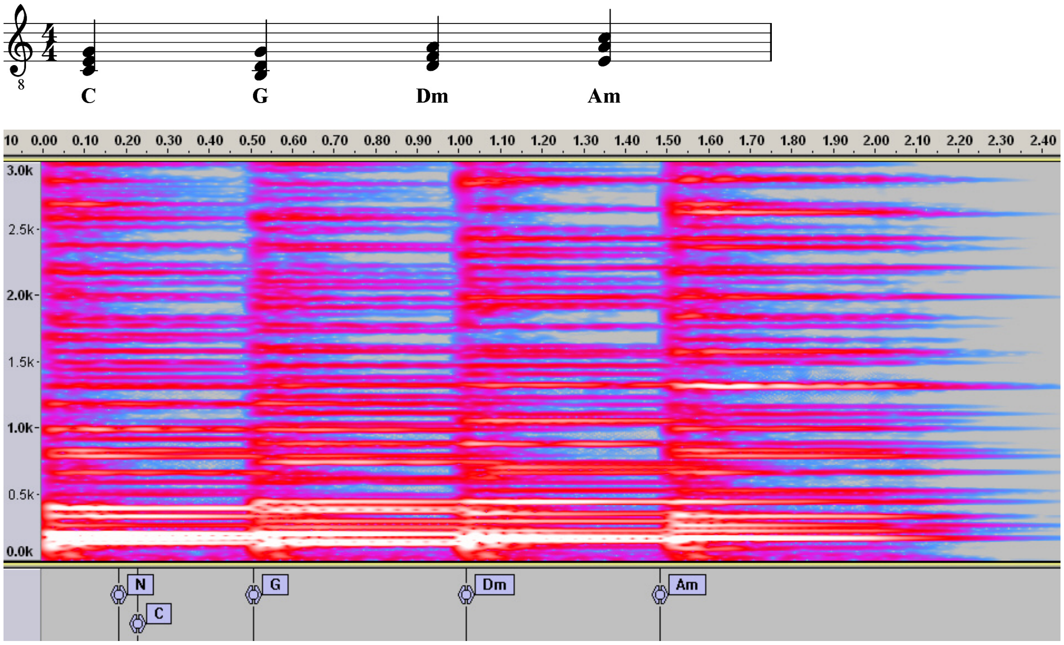

Figure 3. Top: Score of four chords played by a guitar. Chords are labeled only in terms of their pitch classes. Bottom: chord identification using CHORDINO, with spectral shape parameter value 0.7. All four chords are correctly identified.

1.2 Example 2: Arpeggiated Chords

As a second example, we loaded a sound recording of the first three measures of Beethoven's *Moonlight Sonata* into AUDACITY.[2] The score is shown at the top of Figure 4. We have put chord notations directly below the score. These chords are marked in terms of pitch classes only. For example, the first chord is notated as a C$^\sharp$-minor chord, for the following reasons. First, it describes the combination of the arpeggiated notes $\mathrm{G}_3^\sharp$, $\mathrm{C}_4^\sharp$, and E_4, repeated in the four triplets in the first measure of the upper staff. Second, in the lower staff for the first measure, there are the whole notes $\mathrm{C}_3^\sharp$ and $\mathrm{C}_2^\sharp$. Altogether then, the notes in the first measure belong to the pitch classes $\mathbf{C}^\sharp$, $\mathbf{E}$, and $\mathbf{G}^\sharp$, which are the pitch classes for the C$^\sharp$-minor chord. The line extending from the notation $\mathbf{C}^\sharp\textbf{-min}$, underneath the first measure, is meant to indicate that this chord class provides the harmonic structure for that first measure. As another example, the final chord class of $\mathbf{Dmaj}$ with its accompanying line, is meant to indicate that the notes in the second half of the third measure ($\mathrm{F}_1^\sharp$, $\mathrm{F}_2^\sharp$, A_3, D_4, and $\mathrm{F}_4^\sharp$) belong to the pitch classes $\mathbf{D}$, $\mathbf{F}^\sharp$, and $\mathbf{A}$. Those pitch classes describe the chord class $\mathbf{Dmaj}$.

At the bottom of Figure 4, we show how CHORDINO transcribed the chords from the recorded sound file, using the default value of 0.7 for the spectral shape parameter. The chord notations used by CHORDINO are standard ones in lead-sheet notation. The spectrogram is shown as well, because it is easier to see the boundaries for the different measures (compared to the initial waveform display), and because CHORDINO uses information about harmonics in order to make its chord determinations. For identifying arpeggiated chords, CHORDINO uses a rather sophisticated probabilistic inference model, known as a *Hidden Markov Model.* The basic idea is that a particular note (or interval) establishes a *context* for subsequent notes. This context consists of a set of probabilities for subsequent notes to form various chords. As subsequent notes

[2] The recording can be accessed at

`http://www.uwec.edu/walkerjs/mathematicsandmusic/Nav/ClipsForChordino.htm`

It was created using MUSESCORE to save the first three measures as a `*.wav` audio file. Its dynamics and tempo are not the same as a good human performer would produce, but we thought it would still make an interesting test case for CHORDINO.

are played, Chordino uses the Viterbi decoding algorithm for the *Hidden Markov Model* to estimate the most likely chord being arpeggiated (or forming the harmonic context). The mathematical details are given in Mauch's thesis (Mauch, 2010).

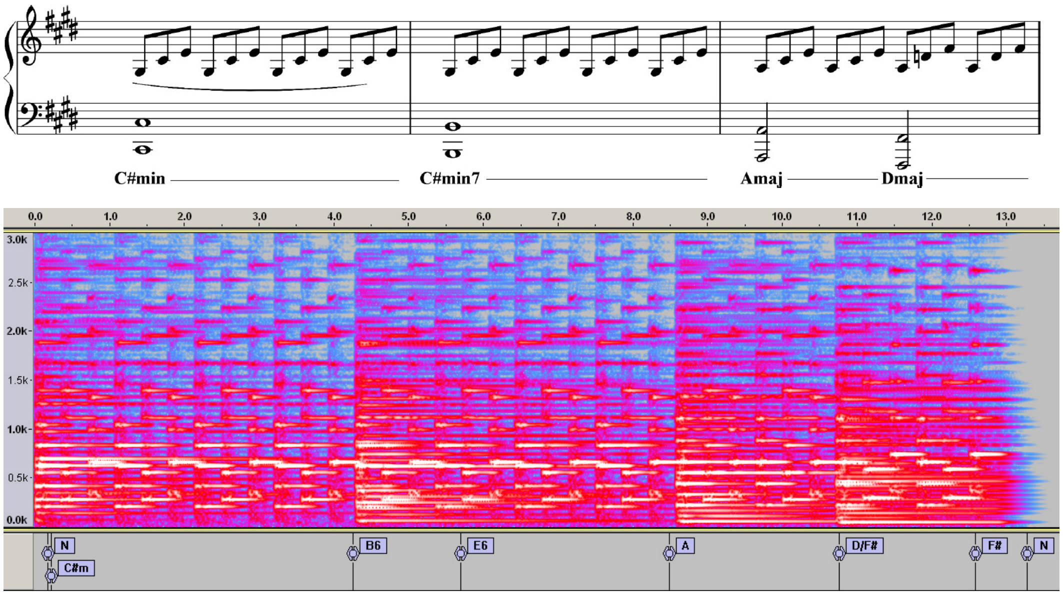

Figure 4. Top: Score of first 3 measures of Beethoven's *Moonlight Sonata.* Chords are shown in pitch-class form only, and they are marked as indicators of the harmonic structure of the passage. Bottom: chord identification using Chordino, with spectral shape parameter value 0.7.

In Table 1, we summarize how well Chordino performed in identifying the underlying chords in the music. In this table, we use a lower-case letter to denote a minor chord, and an upper-case letter to denote a major chord. So, for example, c$^{\sharp 7}$ denotes the c$^{\sharp}$-minor seventh chord, and A denotes the A-major chord.

Table 1 Matches of Chordino chord transcriptions, using 0.7 for spectral shape

Score:	c$^\sharp$	c$^{\sharp 7}$	c$^{\sharp 7}$	A	D	D
Chordino:	c$^\sharp$	B^6	E^6	A	D/F$^\sharp$	F$^\sharp$
Match?	Yes	No	No	Yes	Yes	No

Table 1 shows that Chordino was only partially successful in identifying the chords. Three chords were mistakenly transcribed. These mistaken chords seemed to have resulted from an overemphasis on the bass notes. For example, the B^6 chord (which denotes an interval of **B**-**G**$^\sharp$), corresponds to the B$_1$ and B$_2$ notes, together with the G$^\sharp_3$ note, at the start of the second measure. The immediately following C$^\sharp_4$ and E$_4$ notes, in the triplet in the upper staff, are not included by Chordino in the chord. The mistakenly transcribed chord E^6 is more difficult to explain. This chord denotes an interval of **E**-**C**$^\sharp$. So it appears, based on its position relative to the spectrogram, that Chordino has grouped the first E$_4$ note and the second C$^\sharp_4$ note in measure 2 together, while ignoring the intervening G$^\sharp_3$ note. Finally, Chordino has mistakenly identified the final cluster of three F$^\sharp$ notes (two half notes F$^\sharp_1$ and F$^\sharp_2$, and an F$^\sharp_4$ note at the end of the last triplet) as an F$^\sharp$-major chord. Although this is a mistake, it is an interesting one. As discussed in Chapter 1 of our book, a single note will contain harmonics for a major chord within its first 6 harmonics. In this case, these F$^\sharp$ notes will contain as their third, fifth, and sixth harmonics, the harmonics corresponding to the notes C$^\sharp$ (for the third and sixth harmonics) and A$^\sharp$ (for the fifth harmonic). Those notes together, F$^\sharp$-A$^\sharp$-C$^\sharp$, form an F$^\sharp$-major chord. So, although there is not an explicit chord being played here, Chordino detects one. Our ears sometimes do the same, as pointed out in our discussion of a passage of Wagner on p. 219 of our book, where a single bass note of $\text{A}_1^\flat$ induces the sound of an $\text{A}_1^\flat$-major chord.

These errors are mostly eliminated by choosing a different value for the spectral shape parameter. In Figure 5, we show the results we found using 0.9 for this parameter. Table 2 shows that CHORDINO was much more successful in this case. Only one chord was misidentified. CHORDINO mistakenly listed an E/B chord. This is essentially the same error that it made above, when it mistakenly listed a B^6 chord. The chord that CHORDINO lists as $\text{C}^\sharp$-minor at about 6.6 seconds can be viewed as valid, if the listener no longer hears the B notes in the bass (which the spectrogram shows have substantially faded away at that time). This is a judgement call as to whether the B notes are still contributing to any musical tension here, and so we have marked that $\text{C}^\sharp$-minor chord as correctly identified. To summarize, when the value of 0.9 is used for the spectrum shape parameter, CHORDINO does a reasonably good job at reading the musical context and identifying the arpeggiated chords.

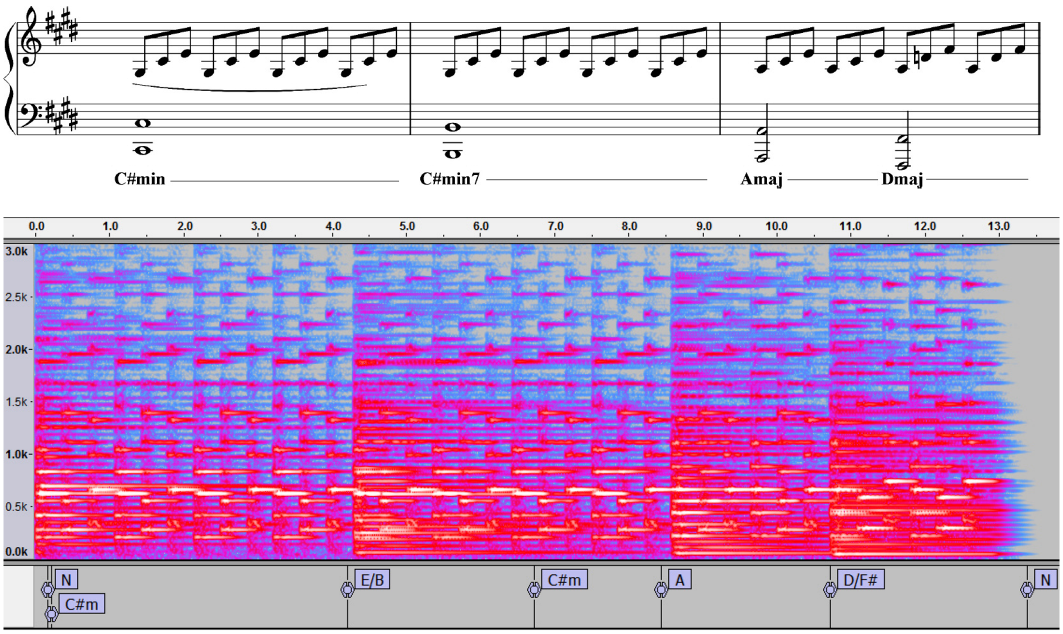

Figure 5. Top: Score of first 3 measures of Beethoven's *Moonlight Sonata.* Bottom: chord identification using CHORDINO, with spectral shape parameter value 0.9.

Table 2 MATCHES OF CHORDINO CHORD TRANSCRIPTIONS, USING 0.9 FOR SPECTRAL SHAPE

Score:	$\text{c}^\sharp$	$\text{c}^{\sharp 7}$	$(\text{c}^\sharp)$	A	D
CHORDINO:	$\text{c}^\sharp$	E/B	$\text{c}^\sharp$	A	$\text{D}/\text{F}^\sharp$
Match?	Yes	No	Yes	Yes	Yes

1.3 Example 3: A More Rhythmically and Harmonically Complex Set of Chords

For our final example, we look at how CHORDINO analyzed the first four measures of Debussy's *Sarabande.* We used a recording of the piece by Claudio Arrau (Arrau, 2006). The score for the first four measures is shown at the top of Figure 6. The spectrogram for the piece, along with CHORDINO's analysis of its chords is shown at the bottom of Figure 6.

The score shows that this piece consists mostly of stacked chords, although there are two rapidly arpeggiated $\text{G}^\sharp$-minor chords (notated as $\text{g}^\sharp$, using the lower-case convention for minor chords). The rhythm is more complex than in the Beethoven example, and there are more notes composing the chords. Our harmonic analy-

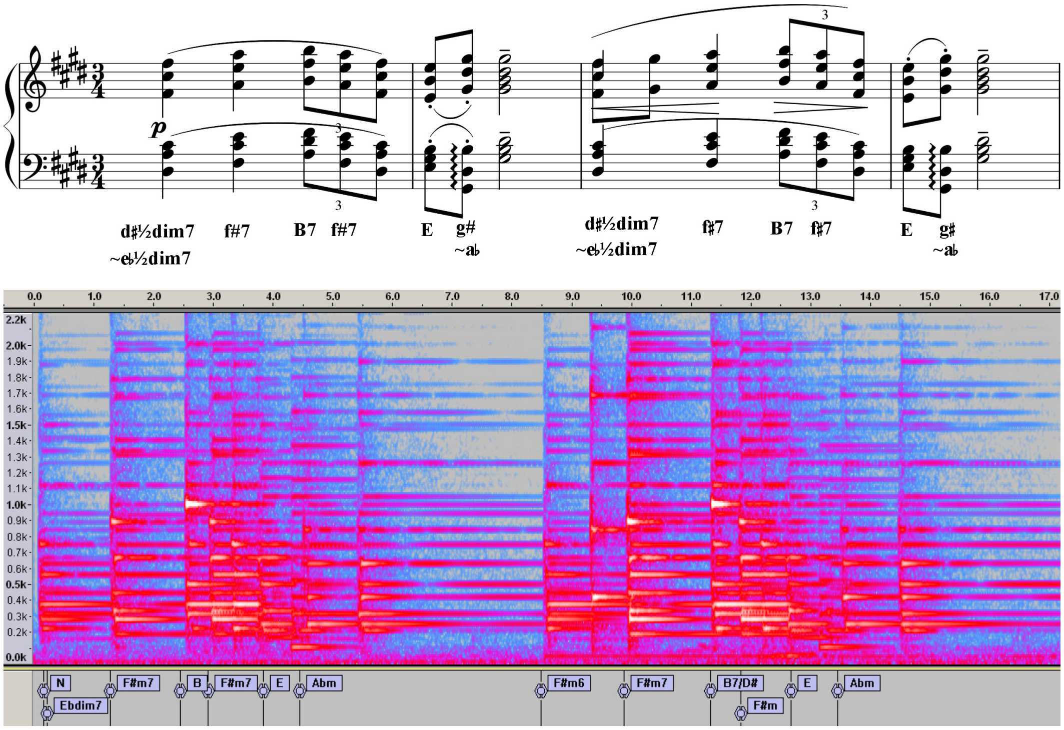

Figure 6. Top: Score of first 4 measures of Debussy's *Sarabande.* Bottom: chord identification using CHORDINO, from a recording by Claudio Arrau. The spectral shape parameter value is 0.5.

sis, written at the bottom of the score, was aided by the analysis done by CHORDINO. Comparing our harmonic analysis with the one provided by CHORDINO, we can see that CHORDINO performed a nearly perfect analysis (if we allow for its enharmonic naming of the g$^\sharp$ chord as an $a^\flat$ chord). However, the first chord is misnamed as a diminished chord, when it is half-diminished. (Its root is also listed as E$^\flat$, but that is enharmonic with D$^\sharp$.) It did take some experimenting with the spectral shape parameter in order to find the value of 0.5 that worked best.

Conclusion

For our three examples, we found that CHORDINO did perform reasonably well. It's two biggest weaknesses are 1) the need to set the spectral shape parameter by hand rather than automatically, 2) difficulty with identifying arpeggiated chords (i.e. with identifying chords that make up a harmonic context rather than explicit stacked chords). Mauch does not claim that CHORDINO is the final answer to the problem of automatic chord identification. Further research is needed. Nevertheless, in its present form, it seems to provide a valuable aid to our ears for identifying chords in recorded music.

Downloading and installing CHORDINO

CHORDINO can be downloaded from the following web page:

`http://isophonics.net/nnls-chromahttp://isophonics.net/nnls-chroma` (3)

The download will be an archived, `*.zip`, file. After extracting the files from this archive, you need to put those files into a folder where AUDACITY can use them. Here are the ways you would do that, depending on whether you are using a WINDOWS computer or a MAC:

- WINDOWS: Put the extracted files in a subfolder of the `Program Files (x86)` folder. This subfolder is named `Vamp Plugins`.

- MAC: Put the extracted files in a subfolder of the `Audio` section of the `Library` folder. This subfolder is named `Vamp`. (Note: if the `Vamp` subfolder of `Library\Audio` does not exist, then you should create it.)

Once you have placed those extracted files in their proper location, then AUDACITY can use CHORDINO. After loading an audio file, you select `Analyze`, and then `Chordino: Chord Estimate`.

References

C. Arrau. (2006). *The Final Sessions. Debussy: Pour le Piano - 2. Sarabande.* D7/Tr5. Decca Music Group.

M. Mauch. (2010). Automatic Chord Transcription from Audio Using Computational Models of Musical Context. Ph.D. Thesis, School of Electronic Engineering and Computer Science, Queen Mary College, University of London.

M. Mauch and S. Dixon. (2010). Approximate Note Transcription for the Improved Identification of Difficult Chords. Centre for Digital Music, Queen Mary College, University of London.

G.W. Don and J.S. Walker. (2013). *Mathematics and Music: Composition, Perception, and Performance,* CRC Press.

James S. Walker
Department of Mathematics
University of Wisconsin–Eau Claire