Analyzing Musical Rhythm

Karyn Muir (Student)
Department of Mathematics
State University of New York–Geneseo
E-mail: kkm5@geneseo.edu

James S. Walker (Faculty Advisor)
Department of Mathematics
University of Wisconsin–Eau Claire
E-mail: walkerjs@uwec.edu

Abstract

The goal of our project is to find a way to differentiate between different genres of music by analyzing their rhythmic structure. We ultimately would like to have a computer take a sample of music and classify what genre it belongs to by analyzing its rhythm. We report on an initial empirical study of an entropy calculation for measuring differences in the complexity of different rhythm patterns.

Keywords: Musical rhythm, spectrograms, scalograms
AMS 2000 Subject Classification Codes: 42A99, 00A06, 00A69

Introduction

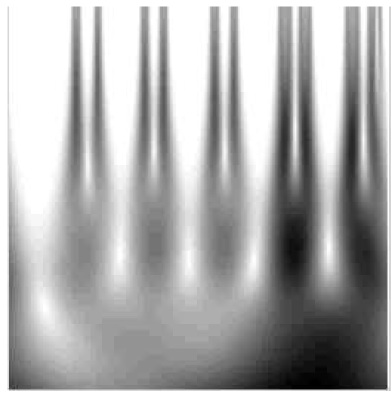
In this paper, we report our work on a project to find a way to differentiate between different genres of music by analyzing their rhythmic structure. We ultimately would like to have a computer take a sample of music and classify what genre it belongs to by analyzing its rhythm. We describe an initial empirical study of an entropy calculation for measuring differences in the complexity of different rhythm patterns.

The paper is divided into two sections. In Section 1, we describe our initial work in analyzing musical passages and deriving a notation for their rhythm. In Section 2, we describe our use of entropy, a measure of complexity, for distinguishing different genres of rhythmic styles. There are two appendices at the end that contain some additional supporting data.

1 Analyzing Musical Rhythm

To analyze various samples of music, we used the computer program FAWAV [9, 10]. FAWAV is a Fourier wavelet analyzer. It takes clips of music in `.wav` format and analyzes them to create spectrograms and percussion scalograms. Spectrograms and percusssion scalograms are described in detail in the book [8, Chaps. 5, 6] and in the article [4]. They provide visual representations of what one hears when listening to the given music samples. A spectrogram is a representation of a music sample in relation to time and frequency [1, 8]. Percussion scalograms are a technique for analyzing musical rhythm, based on a continuous wavelet transform, that was introduced in [8, Chap. 6] and is related to work first described in [6, 7].

The first example we looked at was a drumming passage from a version of the jazz piece, *Sing, Sing, Sing.* Here is the percussion scalogram from this example:



Play video: `http://www.uwec.edu/walkerjs/MandM/SP2008/Sing3.wmv`

The dark vertical bars on the scalogram denote the hits of the drum or other instrument being played in the sample. The shades of grey in the scalogram represent the emphasis on a given hit. The darker shades indicate a harder hit, or more of an emphasis. When looking at these scalograms the hierarchy of the sample can also be seen. When listening to music it is apparent that there are phrases, much like sentence structure in writing. Certain groups of hits or notes are grouped together to create a phrase that makes sense musically and to the listener. Musicians have trouble starting in the middle of these phrases; much like it is difficult for someone to begin reading in the middle of a sentence.

What is interesting about the scalograms is that even though they are computer generated, they show this hierarchy that listeners hear. While the top of the scalogram shows each individual hit, the hierarchies can be seen towards the middle and bottom portion of the scalogram. These scalograms are extremely useful when it comes to looking at the rhythm of a sample. Since they are visual representations of what a sample sounds like, it is possible to notate rhythms by looking at scalograms. By looking across the top of scalograms each hit can be seen, as well as the length of time between each of these hits. When notating rhythms, we used the following symbols to represent different kinds of hits, and different lengths of rests between notes:

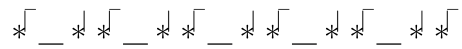
∗♩ represents a more emphasized or accented hit

∗♪ represents a less emphasized or accented hit

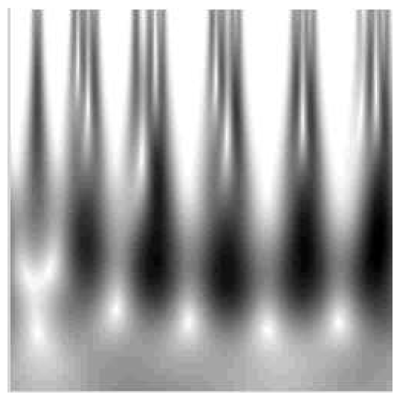
__ represents a rest, with length proportional to the length of the rest.

When looking at the scalograms of the samples, we used these symbols to notate them, listening to the samples for confirmation of their accuracy. The reader may wish to check our work by playing the videos located at the URLs given in the captions for the scalograms shown in this paper.

Look again at the percussion scalogram for *Sing, Sing, Sing.* Across the top we can see each hit of the drum. As we look farther down the scalogram we can see how each set of two hits joins together to show a grouping, or hierarchy. At the bottom of the scalogram we see how the whole sample comes together. By looking and listening to the sample we determined the notation to be:



Another sample we looked at was from the song *Moby Dick* by Led Zeppelin:



Play video: `http://www.uwec.edu/walkerjs/MandM/SP2008/MD.wmv`

Again, we can see the hits of the drum across the top, and the natural grouping of these hits and phrases as we look farther down the scalogram. We determined the notation of this sample to be:

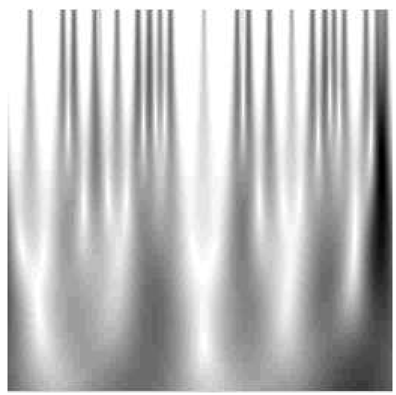
In addition to these samples, many others were looked at as well.

When we looked at the notations for these samples, we wanted to see if there were patterns that could be found between different genres of music. We hoped that these patterns would help us differentiate between genres. The method we used to look for patterns was assigning numbers to parts of the notations. In music, the notes are only part of rhythms. The rests between notes are just as important in determining rhythms, and much of what we hear when listening to music are the rests that create rhythm, patterns, and syncopation.[1] Therefore, we wanted to try to look at the rests of various samples of music to see if we could find patterns between different styles.

[1] As the classical pianist Artur Schnabel is quoted as saying [5]: "*The notes I handle no better than many pianists. But the pauses between the notes — ah, that is where the art resides.*"

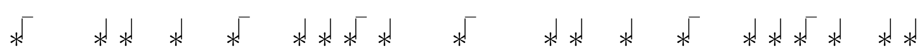
2 Measuring Rhythmic Complexity Using Entropy

Our method of looking for patterns is based on numbering the notations by assigning numbers to the rests and then analyzing those number sequences using the notion of entropy (which we will discuss below). When numbering the notations we looked at every two adjacent hits. If there was no rest (relatively speaking) between two hits, we would assign that part of the sample a 0. If there was a short rest between two hits, we would assign that part of the sample a 1, and if there was a longer rest we would assign that part a 2. We would assign higher numbers for increasing rest lengths. The lengths of rests are looked at proportionally. A 1 in one clip might be the same length of time as a 2 or 3 in another clip. Lengths of rests are measured in relation to the other rests in that specific sample, so the shortest rest will receive a 1 and the other rests will receive numbers in proportion to that shortest rest. To show a numbering example, we will look at a sample from an African drumming song called *Welela:*



Play video: `http://www.uwec.edu/walkerjs/MandM/SP2008/Welela.wmv`

with its notation



From looking at the notation of this sample, we can see that there are two different lengths of rests. Therefore, the shorter rests will receive a 1 and the longer rests will receive a 2. Starting with the first two hits, and then the second and third hits and so on, we see that the numbering of this notation will be:

$$2\,0\,1\,1\,1\,0\,0\,0\,2\,2\,0\,1\,1\,1\,0\,0\,0\,1\,0$$

This representation of the rests is useful in that it shows how the listener will hear the length of time between each drum hit. These various numbers show how the drumming could sound staggered to the listener, instead of being in a steady pattern.

After looking at a variety of samples, we wanted to analyze these numbered representations in some way to see if we could find some sort of trend. We decided to find the *entropy* of these sequences of numbers. The notion of the entropy of a sequence of symbols, in this case our sequence of numbers for the rest lengths, has been studied for over half a century as a measure of complexity (often called *information*). The more complex the sequence, the higher the entropy, and the more difficult it is to *compress* the data: it contains more information than a less complex sequence [2, 3].

First, we used a simple way to calculate the entropy. The first step to calculating the entropy is to find the probability of each of the numbers appearing in the set. For example, in *Welela* we see that the probability of each of the numbers appearing is:

$$p_0 = \frac{9}{19}\,, \quad p_1 = \frac{7}{19}\,, \quad p_2 = \frac{3}{19}\,.$$

After finding the probability of each of these numbers, the probabilities are put into the equation for entropy:

$$\text{(entropy)} \;=\; p_0 \log_2\left(\frac{1}{p_0}\right) + p_1 \log_2\left(\frac{1}{p_1}\right) + p_2 \log_2\left(\frac{1}{p_2}\right).$$

When we enter the probabilities we found for *Welela* we find that the entropy for this sample is 1.46. We looked at our other samples and calculated the entropy for each of them.

One problem we noticed with this form of entropy is that it does not take patterns of numbers into account. If there was a sample of all 0's and 1's and there were an even number of each, the entropy would always be 1 regardless of how they were lined up in the sample. Order does not matter. If a sample has the same numbers in the same pattern recurring over and over again, it sounds much less complex, but the entropy will still be as high for a sample with the same numbers but in a more random pattern. We decided that it was also important to take these patterns, or the lack of patterns, into account when looking at the entropy of a sample, since they all affect what the listener would hear. Therefore, we decided to use Markov-1 Entropy [3]. This form of entropy looks at the transition from number to number, as well as the frequency of numbers. By taking these patterns into account, it will lower the entropy of a very repetitive pattern. This will give a more accurate representation of the complexity of samples, which is what the idea of entropy is about in the first place.

To calculate this Markov-1 Entropy we begin by taking the sequence of numbers, and calculating the probability of each number occurring, just like before. We then look again at this sequence and consider the transitions from one number to the next. We think about each value as being a different state. For example, looking back at the *Welela* sample we see that there are three different states: 0, 1 and 2. We want to look at the number of times each state transitions to each other state. To do this, we must assume a beginning state in order to transition to the first value in the set. The value we pick for the assumed first state ends up having a very small impact on the final entropy value, so we decided to simply assume a starting state of 0 for all of our samples. Therefore, we now have this set of numbers:

$$(0)\,2\,0\,1\,1\,1\,0\,0\,0\,2\,2\,0\,1\,1\,1\,0\,0\,0\,1\,0$$

We then look at each of the three states, and count the number of times each state moves to each other state in this sample compared to the number of total times it transitions to a new state. Starting with state 0, we see that the relative frequencies of transitions to new states are (using, for example, $\rightarrow 1$ to denote transition to state 1):

State 0 Transition Probabilities

$$\rightarrow 0: \quad \frac{4}{9}$$

$$\rightarrow 1: \quad \frac{3}{9}$$

$$\rightarrow 2: \quad \frac{2}{9}$$

So when beginning in state 0, this sample transitions to state 0 four times, state 1 three times, and state 2 two times. This means that in this sample of *Welela* when there are two adjacent notes without a rest in between them, four out of nine times the next two notes also do not have a rest between them, three out of nine times they have a shorter rest between them, and two out of nine times they have a longer rest between them. Now, for the other two states we find:

State 1 Transition Probabilities

$$\rightarrow 0: \quad \frac{3}{7}$$

$$\rightarrow 1: \quad \frac{4}{7}$$

$$\rightarrow 2: \quad \frac{0}{7}$$

State 2 Transition Probabilities

$$\rightarrow 0: \quad \frac{2}{3}$$

$$\rightarrow 1: \quad \frac{0}{3}$$

$$\rightarrow 2: \quad \frac{1}{3}$$

After looking at these transitions, we take all of these values and use them in this new formula to calculate the Markov-1 Entropy:

$$\text{(Markov-1 Entropy)} = p_0 * (\text{entropy}_0) + p_1 * (\text{entropy}_1) + p_2 * (\text{entropy}_2)$$

where $(\mathrm{entropy}_0)$, $(\mathrm{entropy}_1)$, and $(\mathrm{entropy}_2)$ are calculated by going back to the previous equation for entropy, and calculating each entropy for each state's transition probabilities. We take those values and put them in this new equation, which then multiplies each entropy value by the number of times each number occurs in the set. For *Welala*, the calculation is:

$$
\begin{aligned}
\text{(Markov-1 Entropy)} &= \frac{9}{19} * (\mathrm{entropy}_0) + \frac{7}{19} * (\mathrm{entropy}_1) + \frac{3}{19} * (\mathrm{entropy}_2) \\
&= \frac{9}{19} * \left(\frac{4}{9}\log_2\left(\frac{9}{4}\right) + \frac{3}{9}\log_2\left(\frac{9}{3}\right) + \frac{2}{9}\log_2\left(\frac{9}{2}\right)\right) \\
&\quad + \frac{7}{19} * \left(\frac{3}{7}\log_2\left(\frac{7}{3}\right) + \frac{4}{7}\log_2\left(\frac{7}{4}\right)\right) \\
&\quad + \frac{3}{19} * \left(\frac{2}{3}\log_2\left(\frac{3}{2}\right) + \frac{1}{3}\log_2\left(\frac{3}{1}\right)\right) \\
&= 1.2329.
\end{aligned}
$$

Thus for *Welela* we have found that (Markov-1 Entropy) $= 1.2329$.

This form of entropy is much more useful to us, since it takes into account the patterns that occur in each sample. These patterns are easily distinguishable by the listener, and vary according to the specific genre of music.

We looked at the Markov-1 Entropy in comparison to the previous, simpler form of entropy, and then compared the values to see if any conclusions could be drawn from what we found. Interestingly, we found that when the Markov-1 Entropy was taken of an African style piece, the entropy was above 1.0 (such as *Welela*). The Markov-1 Entropy values for other samples we looked at seemed to be below 1.0 for other genres, such as rock and jazz, as shown in the following table:

Title	**Genre**	**Entropy**	**Markov-1 Entropy**
Wipe Out	Rock	0.567	0.537
Taxman	Rock	1.936	0.543
Sing, Sing, Sing (avg.)*	Jazz	1.317	0.548
Toad	Rock	0.904	0.808
What'd I Say	Rhythm & Blues	1.222	0.968
African Drumming (Clip 1)	African	1.236	1.131
African Drumming Class (avg.)**	African	1.379	1.134
Welela	African	1.460	1.233
African Drumming (Example 4)	African	1.455	1.437

These results show that African drumming generally seems to have a more complex pattern of notes, rests and syncopation than styles such as rock or jazz. Therefore, the Markov-1 Entropy could be useful in determining which genre a certain sample is from. If a computer can take something like the percussion scalograms from the FAWAV program, and use this scalogram to notate the hits (and more importantly, the rests), it might be able to then determine what genre the sample is from, without input from a human. The percussion scalograms and entropy values explicitly represent to a computer what listeners hear when a sample of music is playing. Scalograms show each hit or note, while the entropy values show the patterns and complexity of rests. This could allow a computer to determine musical genres without relying on human ears and interpretation. Although this is an ambitious project for future research, we have found that (at least for our small collection) that Markov-1 Entropy does provide a quantitative, objective measure of differing complexity for various genres of music.

Appendix 1: Some Additional Entropy Values

Here we include some additional entropy values for some of our music samples (referred to by an asterisk * or double-asterisk ** in the table above).

* The *Sing, Sing, Sing* entropy values were found from averaging several clips:

Clip identifier	Entropy	Markov-1 Entropy
ssshs9	1	0
ssshs13	1.459	0.66667
ssshs14	1.299	0.951205
ssshs16	1.3921	0.360568
ssshs18	1.4355	0.760964

** The *African Drumming Class*'s entropy values were also found from averaging clips:

Clip identifier	Entropy	Markov-1 Entropy
adc1	1.3788	0.902534
adc2	1.4153	1.36468

Two other versions of *Sing, Sing, Sing* were also looked at:

Sing, Sing, Sing (Pro. Version 1)	Entropy	Markov-1 Entropy
sss1	0.8905	0.885367
sss2	0.873981	0.860541
sss3	0.940286	0.838043
Average	**0.90159**	**0.861317**

Sing, Sing, Sing (Pro. Version 2)	Entropy	Markov-1 Entropy
	1.520	1.468

The drummer in this second version of *Sing, Sing, Sing* was using much more of an African style than in the other two versions, which supports our hypothesis of African style drumming having a higher complexity, and therefore a Markov-1 Entropy value higher than 1.0.

Appendix 2: Notes on Producing Percussion Scalograms

As pointed out in the derivation of percussion scalograms given in [8, Chap. 6] and [4], it is often necessary to use only frequencies in the spectrogram that are above some threshold. When creating the percussion scalograms for the clips described in Appendix 1 and in the table in Section 2, ssshs13 was edited by using only frequencies above 700 Hz in the spectrogram, ssshs14 was edited by using only frequencies above 1000 Hz, ssshs16 was edited by using only frequencies above 800 Hz, and ssshs18 was edited by using only frequencies above 1000 Hz. The *Wipe Out* clip was edited by using only frequencies above 1000 Hz. The *Moby Dick* clip was edited by using only frequencies above 1000 Hz. The *African Drumming Class* clips were both edited by using only frequencies above 2300 Hz.

All of the clips we analyzed are available at:

`http://www.uwec.edu/walkerjs/MandM/SP2008/MusicFiles.htm`

Acknowledgement

This research was supported by the National Science Foundation Grant 144-2974/92974 (REU site SUREPAM program).

References

[1] J.F. Alm and J.S. Walker. Time-frequency analysis of musical instruments. *SIAM Review,* **44** (2002), 457–476. Available at

`http://www.uwec.edu/walkerjs/media/38228[1].pdf`

[2] R.B. Ash. *Information Theory.* Dover, NY, (1990).

[3] T.C. Bell, J.G. Cleary, and I.H. Witten. *Text Compression.* Prentice-Hall, Englewood Cliffs, NJ, (1990).

[4] X. Cheng, J.V. Hardt, J.S. Walker. Time-Frequency Analysis of Musical Rhythm. Submitted. Available at

`http://www.uwec.edu/walkerjs/media/TFAMR.pdf`

[5] A. Schnabel, quoted in the *Chicago Daily News,* June 11, 1958.

[6] L.M. Smith. Modelling rhythm perception by continuous time-frequency analysis. *Proceedings of the 1996 International Computer Music Conference*, Hong Kong, 392–395.

[7] L.M. Smith. A multiresolution time-frequency analysis and interpretation of musical rhythm. Thesis, University of Western Australia, (2000).

[8] J.S. Walker. *A Primer on Wavelets and their Scientific Applications, Second Edition.* Chapman & Hall/CRC Press, Boca Raton, FL, (2008). Material referenced from Chapters 5 and 6 is available at

`http://www.uwec.edu/walkerjs/primer/Ch5extract.pdf`

and

`http://www.uwec.edu/walkerjs/primer/Ch6extract.pdf`

[9] J.S. Walker. Free software FAWAV. Available at

`http://www.uwec.edu/walkerjs/FAWAVE/`

[10] J.S. Walker. Document on using AUDACITY and FAWAV. Available at

`http://www.uwec.edu/walkerjs/tfamr/DocAudFW.pdf`