Entropy and the Estimation of Musical Ability

Andrew Russell
Carnegie Mellon University
andrewru@andrew.cmu.edu

Roger Dannenberg
Carnegie Mellon University
rbd@cs.cmu.edu

ABSTRACT

Musical ability determines how well a musician, or a band, can perform. From a listener's perspective, a musician's ability is often judged by the subjective overall impression of the music. However, others, such as teachers or talent seekers, use more specific criteria to determine a musician's ability. Regardless of who is evaluating, judgements of musical ability are still subjective and require the judges to listen to recordings or live performances of the musicians in question. Automatic estimation techniques would greatly decrease the time required to determine a musician's ability. Automatically estimating a musician's ability would also be very useful for online communities of musicians to help users find other users with a similar musical talent.

In this work, the automatic estimation of musical ability is explored. More specifically, two features, both based on the concept of entropy, are proposed. The first feature looks at the rhythmic consistency of a recording, while the second looks at the tonal consistency. The performance and relative importance of each feature is studied by correlating the results of the feature with data that was manually labelled by 12 musicians. Using the rhythmic feature, a Pearson's correlation coefficient of -0.55 with a p-value of 0.00052 was found, whereas the pitch feature had a coefficient of -0.32 and a p-value of 0.056.

1. INTRODUCTION

Musical ability is defined as how well a musician can play their instrument. Is a musician able to play rhythmically on tempo with a band? Can that same musician sing a note in key? Musical ability is also commonly used to describe how well a musician can emotionally express themselves during a performance. However, we do not have a clear definition of ability. For example, if a listener greatly enjoys a musical performance for whatever reason, they might rate musical ability as high.

How musical ability is defined can also change drastically between genres. A musician can perform with a very high ability for one genre, but have a beginner's level performance for a second genre. Therefore, evaluating the musical ability of a musician across genres is a difficult task. We focus on just western rock and pop music to simplify the process. We also assume that we are analyzing polyphonic recordings.

The original motivation for this work is finding musical partners. Good collaboration often takes place when the collaborators believe that they all have a similar skill level [2-3]. Measuring musical ability would be useful for online communities to better suggest musicians that would perform well together. Additionally, measuring the musical ability would be useful for the music industry to help discover new talent.

We propose two new features that help measure musical ability, one based on the rhythm of the performance, and one based on the pitch of the performance.

2. ENTROPY

Since the performances that we intend to use have no known reference score, we cannot measure performance accuracy directly. Instead, we estimate accuracy and consistency in the rhythm and in the pitch. To do so, we measure entropy [1] over a spectrogram as an indicator of pitch accuracy, and over a beat histogram at an indicator of rhythmic accuracy. The equation for entropy is defined in (1).

$$Entropy = -\sum_{x} x * \log x \qquad (1)$$

3. ESTIMATION METHODS

3.1 Rhythm

To extract the rhythm feature, we use the following steps:

1) *Detect Onsets:* We find the onsets of the entire audio recording.

2) *Calculate Onset Diff Histogram:* We subtract the timestamp of each onset with the timestamp of its neighbor and store the results in a histogram.

3) *Parzen Smoothing:* We smooth the histogram using Parzen Smoothing.

4) *Calculate Entropy:* We calculate the entropy of the smoothed histogram.

3.2 Pitch

To extract the pitch feature, we use the following steps:

1) *Calculate Frequency Spectrum:* We first split the input audio into chunks of 64,000 samples, using an overlap of 50%. We then apply a Hamming window and calculate the Fourier transform. The magnitude of each bin is summed across every chunk and averaged so we end up with a single 64,000 bin frequency spectrum.

2) *Loglog spectrum:* We transform the spectrum using decibels for the magnitude and log10 for the frequencies.

3) *Calculate Entropy:* We calculate the entropy of the loglog spectrum.

4. EVALUATION

To evaluate the new features, we had raters rate 42 30-second clips of songs containing both amateur and professional performances. We then adjusted for raters bias and calculated an overall rating for each song. We then correlated the ratings with the calculated entropy values for each feature using Pearson's correlation coefficient. The results of the correlation can be found in Table 1.

	r-value	p-value
Rhythm	-0.55	5.2e-4
Pitch	-0.32	0.056

Table 1. The correlation results of the two features.

Figure 1 and Figure 2 plot each song's rating versus its entropy value. As expected, the professional performances tend to have higher user ratings and lower entropy values while the amateur performances tend to have lower user ratings and higher entropy values. However, it is interesting to note that not all professional performances scored well, both by the raters and by the entropy features. Additionally, some amateur performances scored well by both the raters and the entropy features.

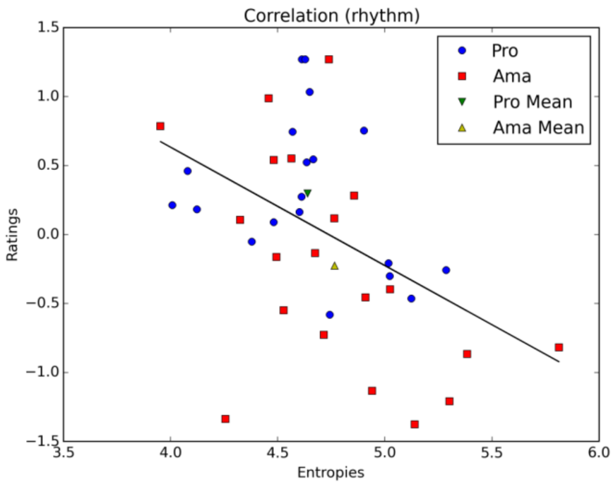

Figure 1. Amateur and professional rhythm entropies vs. ratings.

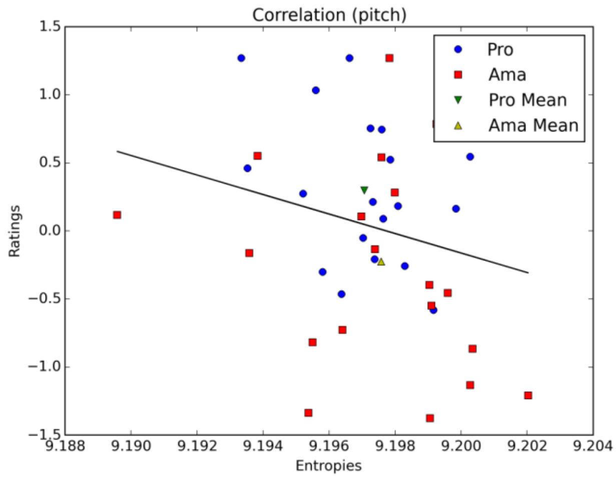

Figure 2. Amateur and professional pitch entropies vs. ratings.

5. CONCLUSION

Of the two introduced features, rhythm has a strong correlation against the subjective ratings (r = -0.55, p = 5.2e-4). Pitch appears to have a strong correlation (r = -0.32) but the correlation was not significant in this study (p = 0.056). However, there is much future work that could be done to enhance both features. The rhythm feature currently penalizes intentional rhythm and tempo changes, as well as intentionally playing ahead of or behind the beat. The pitch feature currently penalizes recordings that change key, as well as instruments played with vibrato.

In spite of these shortcomings, we find it both encouraging and interesting that some degree of musical ability can be determined by relatively simple means. We believe this the first study [4] of features for estimating musical ability without reference to a score or target performance.

6. REFERENCES

[1] K. Ekštein and T. Pavelka: "Entropy and entropy-based features in signal processing," *Proceedings or PhD workshop systems & control*, 2004.

[2] P. J. Guinan, J. G. Cooprider, and S. Faraj: "Enabling software development team performance during requirements definition: A behavioral versus technical approach," *Information Systems Research*, Vol. 9, No. 2, pp. 101-125, 1998.

[3] N. Katira, L. Williams, E. Wiebe, C. Miller, S. Balik, and E. Gehringer: "On understanding compatibility of student pair programmers," *ACM SIGCSE Bulletin*, Vol. 36, No. 1, pp. 7-11, 2004.

[4] A. Russell: *Entropy and the Estimation of Musical Ability*, M.S. thesis, Carnegie Mellon University, 2016.