The Trumpet Shall Sound: De-anonymizing jazz recordings

Janet Lazar
Rutgers University
New Brunswick, NJ, USA
janetlazar@icloud.com

Michael Lesk
Rutgers University
New Brunswick, NJ, USA
lesk@acm.org

We are experimenting with automated techniques to identify performers on jazz recordings by using stylistic measures of acoustic signals. Many early jazz recordings do not identify individual musicians, leaving them under-appreciated. We look at individual notes and phrasing for recognition of jazz trumpeters as an example.

Jazz, performer identification, music analysis.

1. INTRODUCTION
For much of the 20th century jazz recordings did not contain full listings of the performers; attributions would only name a group such as “Count Basie and his All American Rhythm Section” or “Duke Ellington and his Orchestra”. Who were the actual performers? Our goal is to recognize them automatically, using jazz trumpeters as an example. The pictures below are all from Wikipedia.

Identification of flamenco singers and classical pianists has been studied before [Kroher 2014, Saunders 2008]; the jazz problem is more complex because there is no written score to be aligned with the notes played. However, experienced human listeners can recognize the performers, so the problem is feasible. Some researchers have invested in manual creation of the score [Abesser 2015] followed by a complex separation of the playing of each performer. We’ve been looking at solo passages, identified by ear, but hoping to recognize them mechanically in the future.

Why not try this as a very general machine learning problem? One could feed all the data into WEKA and sit back and watch. However, there isn’t enough data: we have at most hundreds, not millions, of samples. Worse yet, there are many “accidental” properties of the acoustic signals. For example, different recording studios used microphones with different frequency limits. Until the 1950s many microphones recorded only up to 10kHz [Ford]. We would not wish to train a system on whether a recording was made at RCA in Camden, NJ or at Columbia in New York.

What features would be characteristic of musical style? The diagram below is from [Ramirez 2007] and shows the intensity contour of a single note:
What features might be exploited for machine classification? Single-note features include:

- **Vibrato** – are notes steady or wavering?
- **Tone complexity** – are the notes simple tones or many additional frequencies?
- **Onset speed**: do the notes rise quickly or slowly in intensity?
- **Decay speed**: do the notes stop quickly or does the performer “tail off” each note?

Multi-note features, derived from phrasing, include:

- **Staccato/legato** – notes separated or continuous?
- **Beat timing**: regularly spaced notes or “ragged” time.

2. **ST. LOUIS BLUES**

For demonstration purposes, and to test software, we are using recordings of W. C. Handy’s *St. Louis Blues*, written in 1914 and recorded more than 100 times. Here are sound spectrograms for snippets of sound by Louis Armstrong, Harry James and Wynton Marsalis. The software used in this paper includes BeatRoot [Dixon] and the MIR Toolbox [Lartillot]; we thank the creators and maintainers of these programs.

Armstrong has the most complex sound (least dominated by the main note frequency) while Marsalis played fewer tones in each note. Marsalis’ playing is the most staccato; Armstrong and James played more continuously. Looking at frequency stability, Marsalis plays with the most stable notes, i.e., the least vibrato, while James is a bit more variable and Armstrong still more.

For another comparison, Figure 2 shows sound spectrograms for about 0.2 seconds (a single note, roughly) taken from three different places for each performer. All are again *St. Louis Blues*. Look here at the extent to which the pure note and its overtones dominate the signal. Marsalis is playing with the least sound beyond the specific note; James has a more complex note, with extra overtones; and Armstrong has much more in the way of low frequency components in the notes.
The Trumpet Shall Sound: De-anonymizing jazz recordings
Janet Lazar & Michael Lesk

Figure 2: Single-note sound spectrograms.

Figure 3: Single-note, Benny Goodman (top), Harry James (bottom). What would we see if we compared two different clarinetists? The next pair of spectra, in Figure 4, show Benny Goodman above and Artie Shaw.

Figure 4: Benny Goodman (top), Artie Shaw (bottom). Compared to the trumpet both are weighted to lower frequency and simpler in structure. Comparing these two, Benny Goodman’s notes are “purer” and contain fewer frequencies.
3. CLARINET AND HARP

What happens if we look at other instruments? Figure 3 shows a comparison of Benny Goodman (above, clarinet) with Harry James (below, trumpet). Note the generally lower frequency spectrum of the clarinet and the complexity of the trumpet notes in terms of frequencies.

As another example, we took sound spectra of four different harpists. In Figure 5, the top left spectrum is Lucia Bova, top right is Csilla Gulyas, bottom left is Maria Graf and bottom right is Judy Loman. They are all playing C. P. E. Bach’s *Harp Sonata in G major*, Wq 139.

![Figure 5: Four harpists. Left column: Lucia Bova, Maria Graf. Right column: Csilla Gulyas, Judy Loman.](image)

We then calculated the basic tempo for each and the attack time, measuring off the sound spectra, using two samples for each player. Below is a plot showing that the performers differ but each tends to repeat her characteristic choices.

![Figure 6: Distribution of tempi and attack time.](image)

4. CONCLUSION

The longer-run purpose of this work is to help with cataloging old recordings. Since music had no requirement for compulsory deposit in the United States until the 1970s, the Library of Congress has an unusually incomplete collection. Rutgers University, at its Institute of Jazz Studies in Newark, NJ, holds more than 100,000 sound recordings, and this is the largest jazz repository. Unfortunately, practical difficulties, such as fragility of records, and legal difficulties, such as copyright ownership of recordings made by companies that may be long out of business, impede the study of these recordings. We hope that by automating the creation of metadata we can help the scholars and bring recognition to artists whose contributions are fading from memory and insufficiently documented.

5. REFERENCES


