Experimental Data and Results*

Yi-Chieh Wu

Abstract

Training and testing of our instrument recognition system.

1 Experimental Data

1.1 Training

For our training set, we used purely monophonic recordings to ease manual classification. One full chromatic scale was recorded for each of our three instruments. We note that our training is weak, as we only provided one recording for each instrument. By covering the full range of the instrument, we give roughly equal weights to every note, whereas most instruments have a standard playing range and rarely play in the lower or upper limits or their range. Since the spectrum is more apt to skewing effects in the extreme ranges, our average spectral envelope and training features are also negatively affected.

Finally, if we wanted our training set to perform better with polyphonic recordings, we would in practice also provide a few polyphonic recordings as part of our training set. This would allow features unique to polyphonic environments to be modeled as well. For example, a clarinet and trumpet usually cover the melody and are therefore more predominant than a tenor saxophone.

1.2 Testing

One short monophonic tune per instrument was recorded, as well as two short polyphonic tunes with each instrument combination (clarinet + saxophone, clarinet + trumpet, saxophone + trumpet, all), generating a total of 9 recordings.

2 Results

2.1 Self-Validation

We first tested our GMM with the training set to determine how accurate it would be at classifying the data that trained it. The confusion matrix is shown below. (Our confusion matrix shows the actual classification at the left, and the predicted classification at the top.)

*Version 1.1: Dec 14, 2005 7:00 pm -0600

	Clarinet	Saxophone	Trumpet
Clarinet	90.0%	7.5%	2.5%
Saxophone	2.9%	92.3%	4.9%
Trumpet	0.9%	11.5%	87.5%

Table 1: **Table 1:** Confusion matrix for instrument recognition with training data.

2.2 Monophonic Recordings

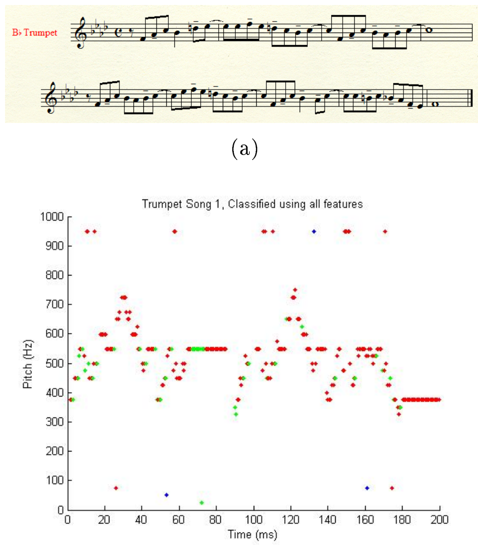
Satisfied that our GMM could classify the training data accurately, we then tested it on a new set of monophonic recordings.

	Clarinet	Saxophone	Trumpet
Clarinet	67.0%	15.1%	17.9%
Saxophone	19.7%	73.0%	7.3%
Trumpet	1.0%	14.9%	84.1%

Table 2: **Table 2:** Confusion matrix for instrument of single notes from monophonic recordings.

Average instrument identification using our GMM was 75%, whereas pure guessing would land us at 33.3%. We also see that in our test data, the clarinet and saxophone are confused the most often and can therefore be considered the most similar. This makes sense as both belong to the same instrument family (woodwinds), whereas a trumpet is a brass instrument. In contrast, the clarinet and the trumpet were confused the least often, which is also as expected since their spectrum represent the two extremes within our tested instruments. We are unsure of why the clarinet is often mistaken as a trumpet, but a trumpet is not mistaken as a clarinet, but we believe part of the problem may lie again in our training data, as the self-validation tests showed that the clarinet and trumpet were almost exclusive of one another, and our GMM may have started to memorize the training data.

The following figures show the performed piece of music and the results of our detection and classification algorithm. We note that some discrepancies are due to player error (key fumbles, incorrect rhythmic counting, etc). We follow our coloring scheme of blue representing clarinet, green saxophone, and red trumpet.



(a)

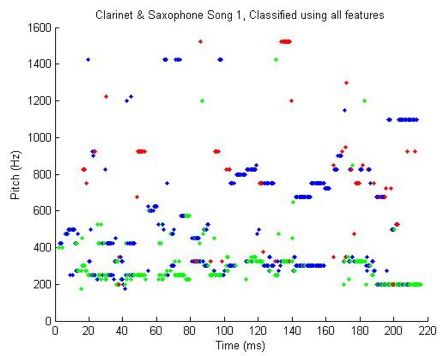
(b)

Figure 1: Original score versus output from our algorithm for a monophonic trumpet tune.

2.3 Polyphonic Recordings

Finally, we input some polyphonic recordings and compared the experimental outputs to the input music. Quantitative validation is not provided, as it would require us to manually feed into the validation program which instruments at what time. Visually however, we can clearly see that our algorithm correctly separates the melody line, as played by the clarinet, from the lower harmony line, as played by the tenor saxophone.

(a)



(b)

Figure 2: Original score versus output from our algorithm for a polyphonic piece.