

Music Notation Handwriting Recognition

An Experimental System for Online Recognition of Handwritten Music Notation

User-Drawn Input

The user draws a collection of conventional Western music notation (CWMN) symbols. Ink can be entered in any order or combination of discrete symbols. The symbols are recognized as they are drawn.



A Blank Slate

The user is presented with a blank five-line staff upon which to draw.

Motivation

CWMN is fundamentally graphical and non-linear, making input a difficult problem. Presently, a user typically enters music notation with a combination of a mouse, alphanumeric keyboard, and/or a piano style keyboard. The system presented here is an attempt to develop a transparent user interface not unlike pencil and paper. It is best suited for gestural interfaces using either a stylus or a finger for input, but can be used with a mouse or trackpad.

Testing Framework

In development, a testing framework can record and replay serialized ink, as well as take bitmap captures for iterative test comparisons and identification of regressions. This framework was used to generate some of the images on this very poster.

Recognition in Three Steps

- 1. Segmentation:** which combination of ink segments form a symbol.
- 2. Classification:** what is the most likely symbol for a collection of ink.
- 3. Language model:** how the discrete symbols fit together in a meaningful way.



1. Segmentation

The user-drawn ink is segmented to determine which combination of strokes are most likely to combine to form a symbol. A stroke is defined as a vector of points from pen down to pen up. The resulting collection of strokes will then be compared to a collection of hand-trained templates.

These four strokes of ink:
1) 2) 3) 4)



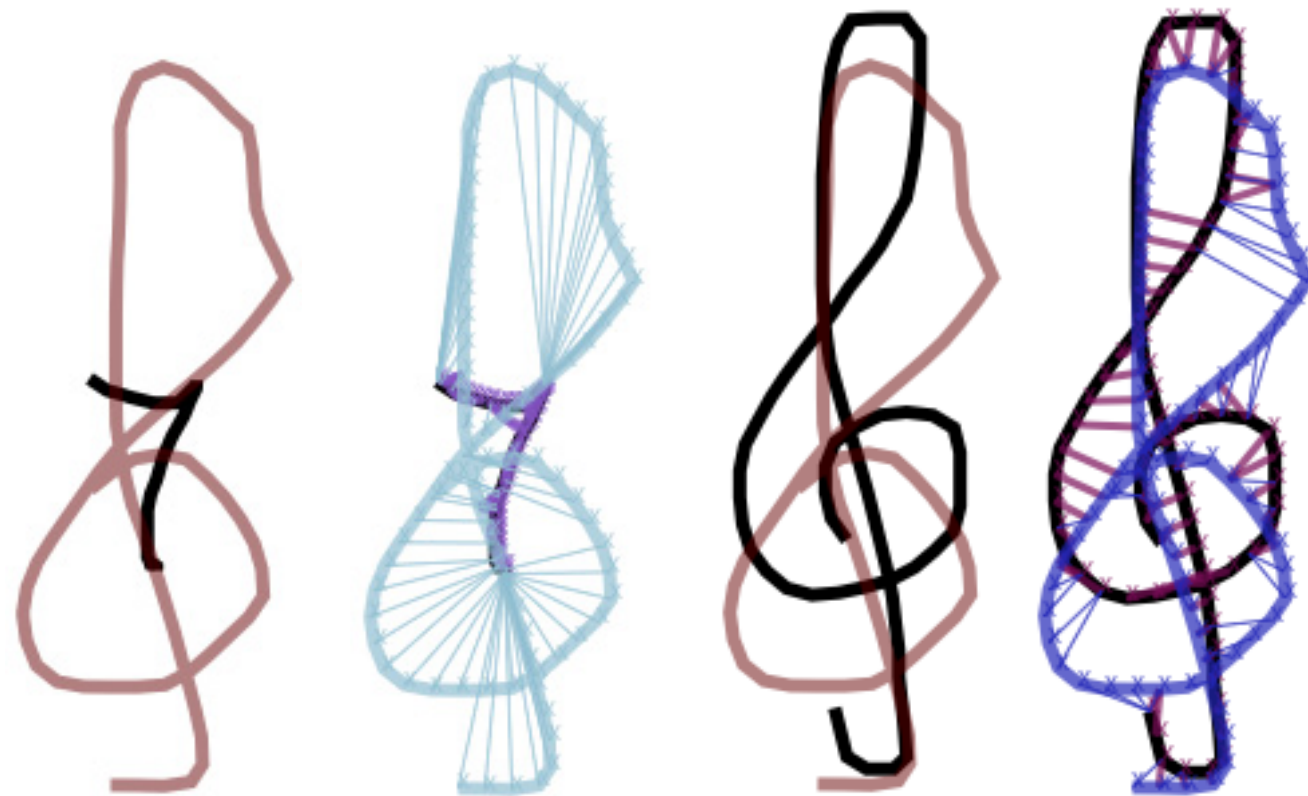
Combine to form a sharp:



2. Classification

After normalizing each collection of strokes by scale and to an equidistant fixed number of points, template matching is performed by calculating the mean of the distances from each drawn point to the nearest template point and vice-versa. In other words, this comparison is bidirectional between drawn ink and the template.

Drawn ink of a treble clef matched against:
1) an eighth note template & 2) a treble clef template



1) loses with the larger mean distance & 2) wins with the smaller.

3. Language Model

A hand-coded collection of bigrams define how likely one symbol is to follow another.



Typeset Result

The resulting music notation typeset without the user-drawn ink visible.

The system is trained on these symbols:



Credits

Geoff Chirgwin is a Music Informatics master's student working on this project under the guidance of Professors Christopher Raphael and Larry Yaeger.

iQuil.com

Demonstration Video: demo.iquil.com

Since all code for the system is either markup or interpreted JavaScript, the above web address effectively constitutes a reference to a full source code listing of the present implementation.

User Correction of Symbols

