Sheet Music Reader

Sevy Harris, Prateek Verma
Department of Electrical Engineering, Stanford University

Overview

The goal of this project was to design an image processing algorithm that scans in sheet music and plays the music described on the page. Our algorithm takes a digital image of sheet music and segments it by line. For each line, it detects the note types and locations, and computes the frequency and duration. These frequencies and durations are stored in a matrix and fed into the synthesizer, which plays back the music.

Segmentation and Preprocessing

First the image is split into lines by summing the rows and finding the middle of the white space between each staff. The figure to the right shows the row sums after erosion with a long horizontal line. The red lines represent the calculated dividers that split the page into separate lines. An example of a segmented page is shown below.
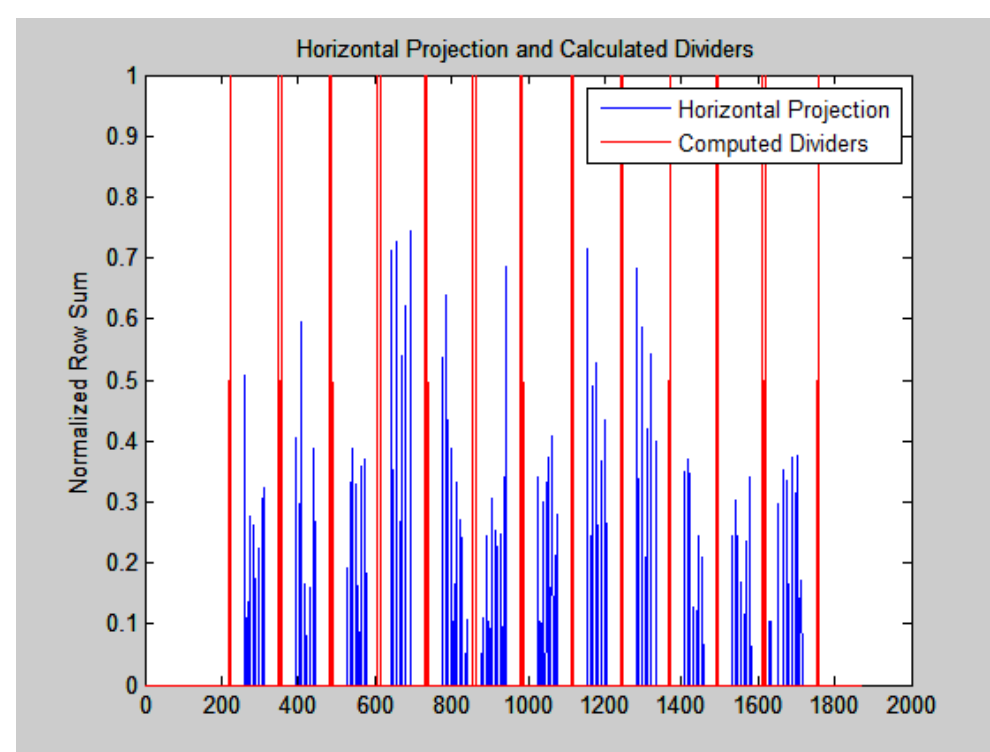
Figure 1: Shows the row sums of a page of sheet music (blue) and the computed dividers between each staff (red).
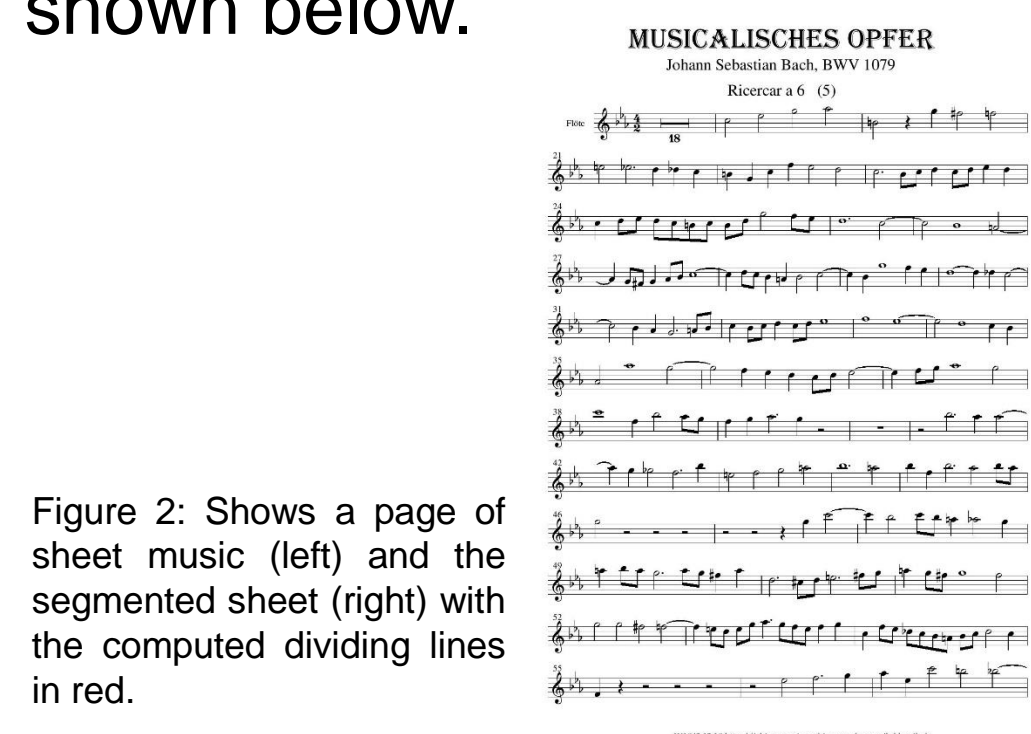
Figure 2: Shows a page of sheet music (left) and the segmented sheet (right) with the computed dividing lines in red.

The staff lines for each line of music are found by summing the rows and finding the 5 equally spaced rows that maximize the sum. These rows are then subtracted from the image. Results are shown in the images below.

Figure 3: Staff lines highlighted in red.

Figure 4: A line of music with the staff lines removed.

Algorithm Flow

Segmentation and Preprocessing

Scan Image   Split Lines   Find and Erase Barlines

Morphological Image Processing for Detection

Key and Clef   Solid Notes   Notes with Holes   Rests

Music Synthesis

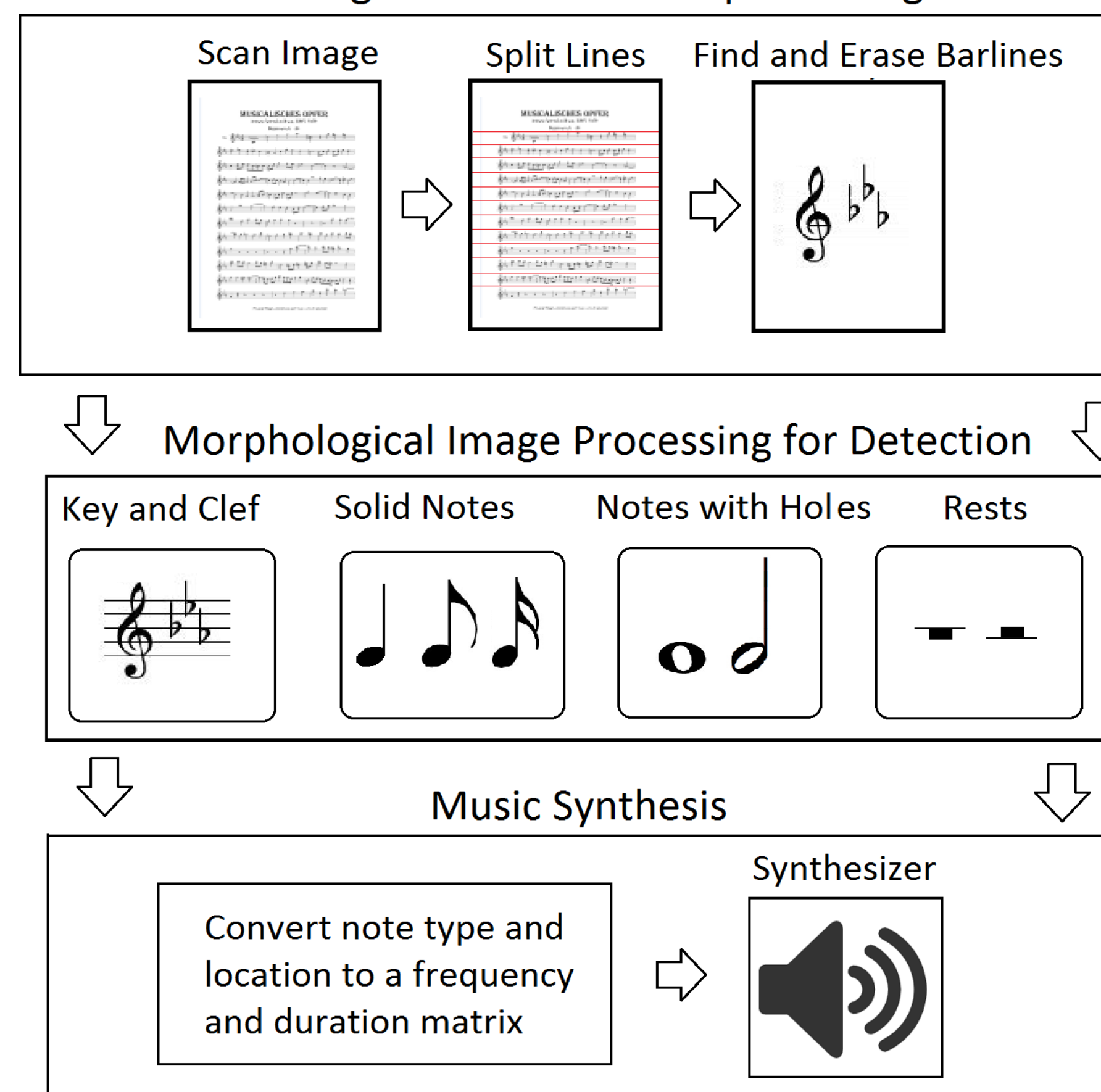Convert note type and location to a frequency and duration matrix

Synthesizer

Figure 5: Shows the logical flow of the image processing algorithm.
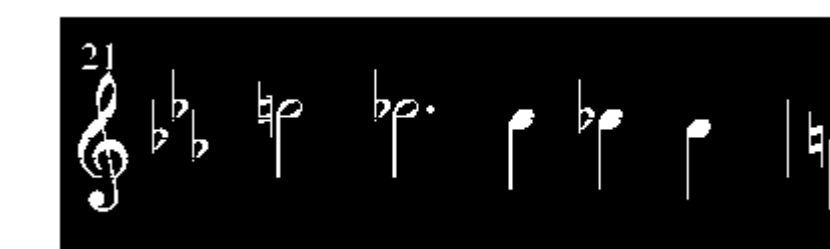
Note Detection

Filled notes and unfilled notes are detected using the method shown in the figure below. The note stems, flags, and dots are detected in later stages to determine the type.

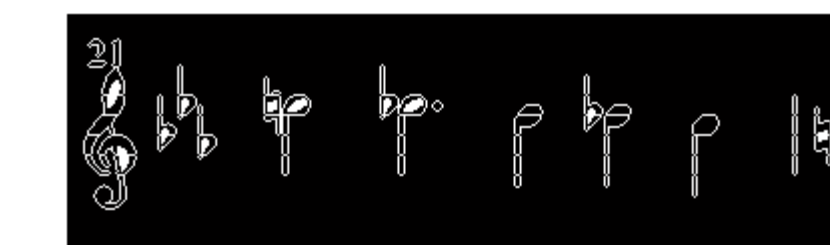Quarter and Eighth Note Detection (Filled)

Whole and Half Note Detection (Unfilled)

After Image Erosion

After filling 'holes' (dark regions not connected to the main background)

After Eccentricity and Size Thresholding

After Erosion, and Eccentricity and Size Thresholding

Figure 6: Shows the intermediate morphological image processing stages involved in note detection.

Results

- Half Note
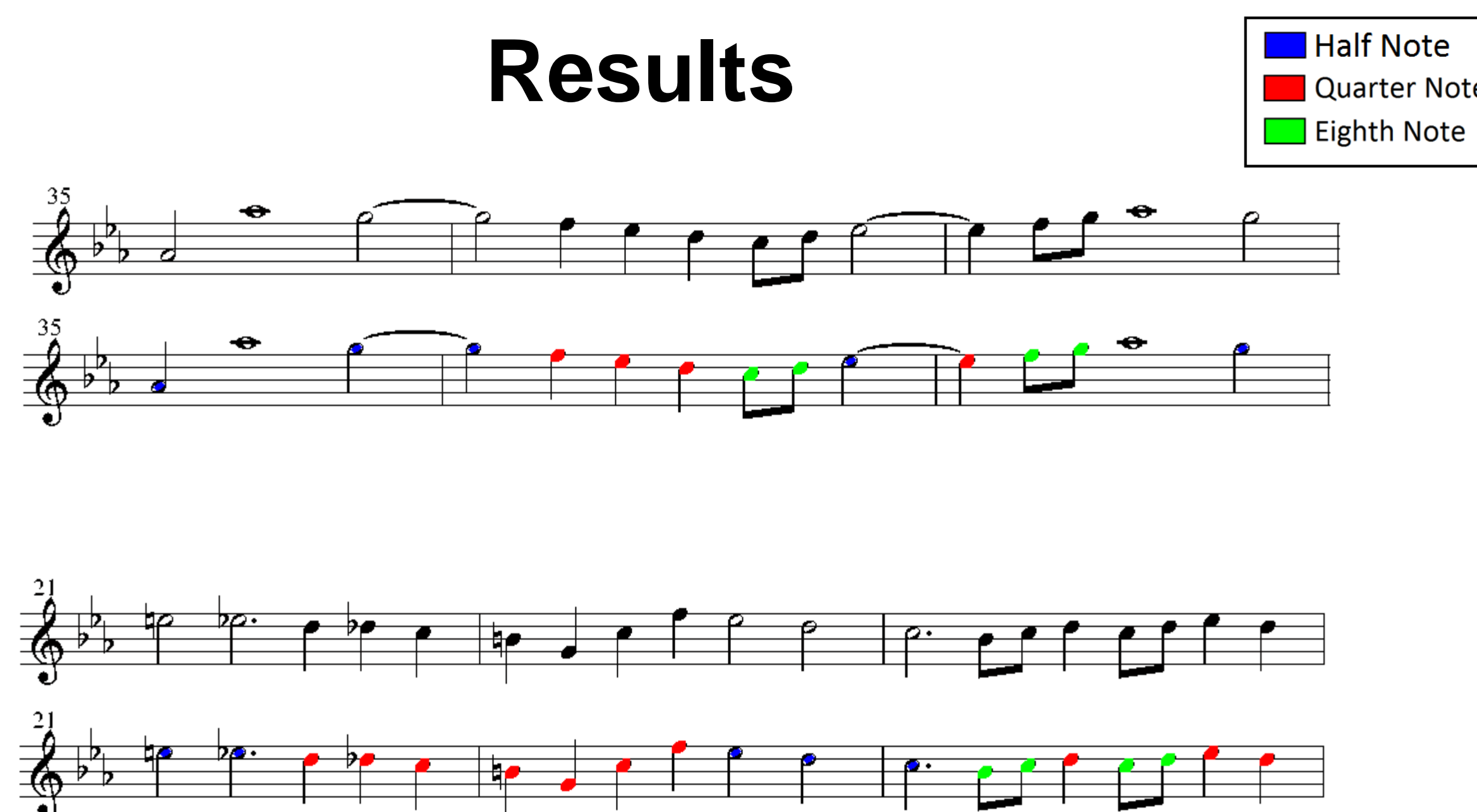- Quarter Note
- Eighth Note

Figure 7: Shows two lines of music and the types of notes detected.

Future Work

- Handling Accidentals Outside of the Key Signature
- Handling of Shorter Rests
- Finding Slurs
- Improving Robustness of Dotted Notes
- Triplets and Other Groupings
- Dynamics
- Accents