EE475 Project Report

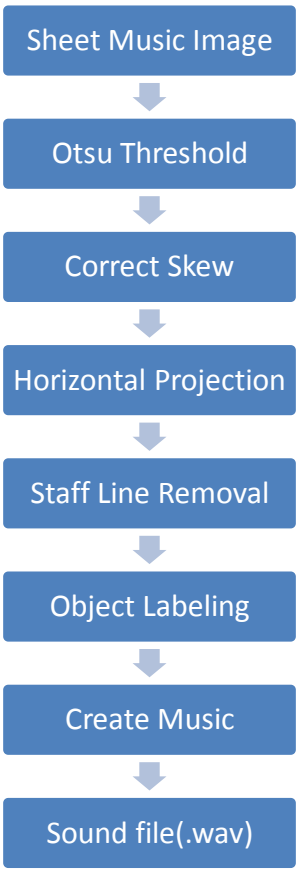
Optical Music Recognition

Ahmet Can Özbek

1)Introduction

In this project, my aim is to design a system that can interpret sheet music or printed scores into editable or playable form.

2)Project Outline



- Otsu Threshold

The binarization is done with Otsu's global thresholding method.

- Correct Skew

Before analyzing the music sheet, we need to fix the skew of the music sheet if there is any.

The algorithm used in skew correction is as follows:
I am assuming the angle of skew is no more than -10 or 10 degrees. In this range, I am traversing along the lines with different slopes . Along this sloped line I am incrementing a «count» value for each black pixel encountered.

I am applying this procedure to the middle one third of the music sheet.
First I am taking the maximum of the "count " values among all the lines which have the same slope. Therefore, there is only one "count" value for each degree value.

Then I take the degree value which has the maximum “count” value. This degree value should be the skew angle of the music sheet because the maximum of the “count” value is reached when the line traverses on one of the staff lines.
Once the skew angle is calculated, “imrotate()” function in MATLAB is used to correct the skew of the music sheet.

- Horizontal Projection

In this step, we are gathering information about the music sheet.
Horizontal projection of a music sheet gives us a lot of information. By investigating the peaks in the horizontal projection we can determine the following information :

-Number of staff(group of 5 staff lines)
-Distance between staffs
-Avarage Staff Line Thickness
-Avarage Staff Spacing (avarage distance between two staff lines in a group of 5)

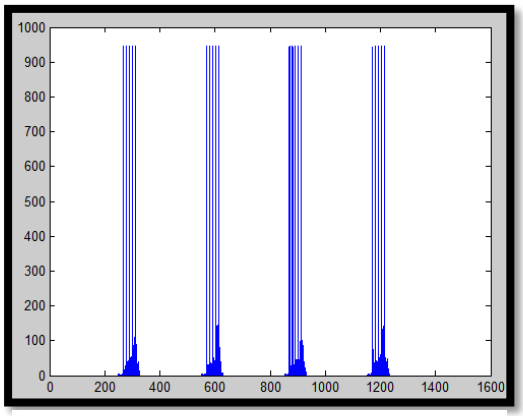


Figure 1 – Horizontal Projections

```
**********Statisctics of The Music Sheet:**********
***Number of Staff Blocks:
     4

***Number of Staff Lines:
    20

***Avarage staff line thickness(n):
    1.4500

***Avarage distance between staff lines(d):
   10.3125

***Avarage distance between staffs:
  258.3333
```

Figure 2 – Output data of “Horizontal Projection” block

These values provides essential information for the next steps.

- Staff Line Removal

Staff Lines connect and intersect almost all other symbols, so their presence disturbs the following recognition of musical objects. It is therefore desirable to remove the staff line segments, which do not overlap any symbols to prevent that intersecting symbols become distorted.

The algorithm for staff line removal used in the project :

This method involves looking at the average thickness of the lines, and go through each single line from left to right, checking if the line is thicker than a certain threshold (typically the average staff line thickness) .The line, at position x, can then be deleted safely if the width of the line is smaller than this threshold. If the line is much thicker it indicates that there is an object going through the line, from above or below, or even an object attached to the line on both sides.

- Object Labeling

In this part, our aim is to identify each object and then identify the pitch of the notes(C4, D4,... etc.).

After staff line removal, I get the locations and the sizes of objects by “bwlabel()” function in MATLAB using 8-connectivity.

To identify the type of objects. I perform template matcing. For Template matching, the Hamming distance is used, since bit-wise comparisons is necessary.

To identify the pitch value of notes, we need the information of the vertical position of note heads relative to the staff line. For this purpose, I look at the horizontal projection of each note.

The peak location in the horizontal projection should give us the center location in vertical direction of the note head.

Later, we can determine the note head’s position relative to the staff line.

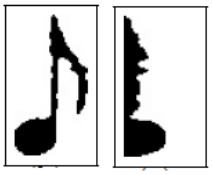


Figure 1 – Horizontal projection of an eight note

However, we cannot apply this method directly to half and whole notes Because inside of half and whole notes are empty. Horizontal projection of them is not going to provide us much information.

Figure 4 -Half note

Figure 2 – Whole note

To overcome this problem, I fill the inside of whole and half notes and then take horizontal projection.

After all objects are labeled and the pitch values of the notes are determined, the objects for music sound are sorted in the way the music is played (starting from left top corner and going in left to right direction until the end of the page). This sorted data is saved and given to “Create Music” block to create and hear the music sound file of the sheet.

- Create Music

In this part, the sorted “object data” is taken as an input from “Object Labeling” block. The types(whole, half, quarter, etc.) and the pitch values(C4,D4, E4, etc.) are saved in this “object data”.

The type of note correspont to its duration, in other words the type of notes determine how long the note will be played.

To create the actual sound of notes, I generate a sinusoidal signal with the frequency of the note. Each note is played according to the order in the object data and the duration of the note will be determined from its type. Hence we create the actual music. The music signal is saved into a “.wav” file. Those files can be found in my Dropbox folder → OutputMusicWavFiles folder.

Results

The sheet music images can be found in the appendix part

Confusion Tables:

Jingle Bells:

	quarter	half	whole	bar	gclef	eight
quarter	40	0	0	0	0	0
half	0	6	0	0	0	0
whole	0	0	3	0	0	0
bar	0	0	0	11	0	1
gclef	0	0	0	0	4	0
eight	0	0	0	0	0	0

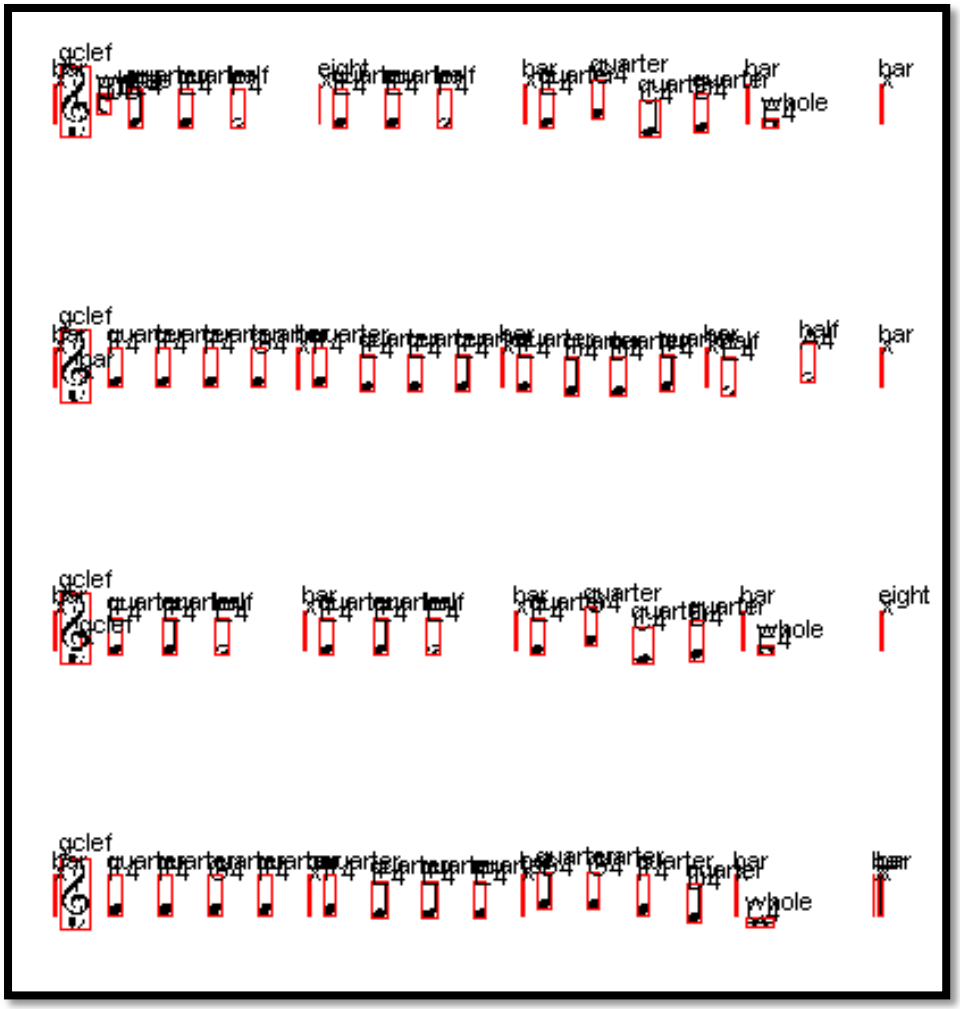


Figure 3 – Jingle Bells labeled

Twinkle Twinkle Little Star:

	quarter	half	gclef	bar	sharp
quarter	36	0	0	0	0
half	0	4	0	0	2
gclef	0	0	3	0	0
bar	0	0	3	10	0
sharp	0	0	0	0	0

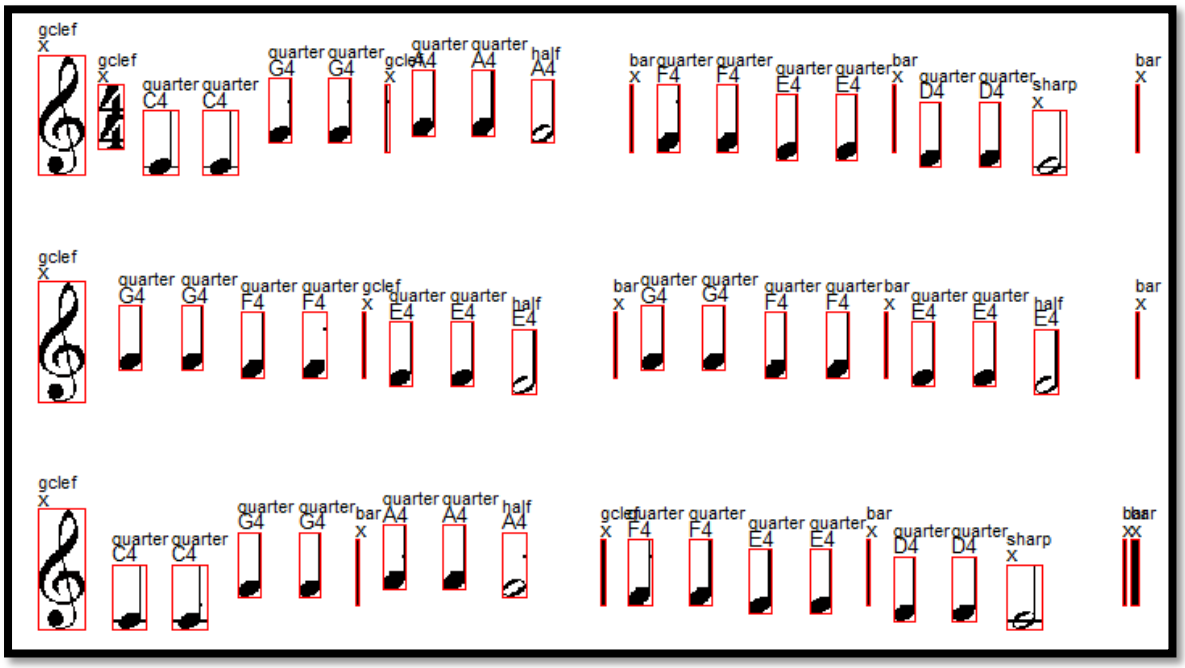


Figure 4 – Twinkle Twinkle Little Star Labeled

Skewed Twinkle Twinkle Little Star:

	quarter	half	gclef	bar	sharp
quarter	36	0	0	0	0
half	0	4	0	0	2
gclef	0	0	3	0	0
bar	0	0	0	13	0
sharp	0	0	0	0	0

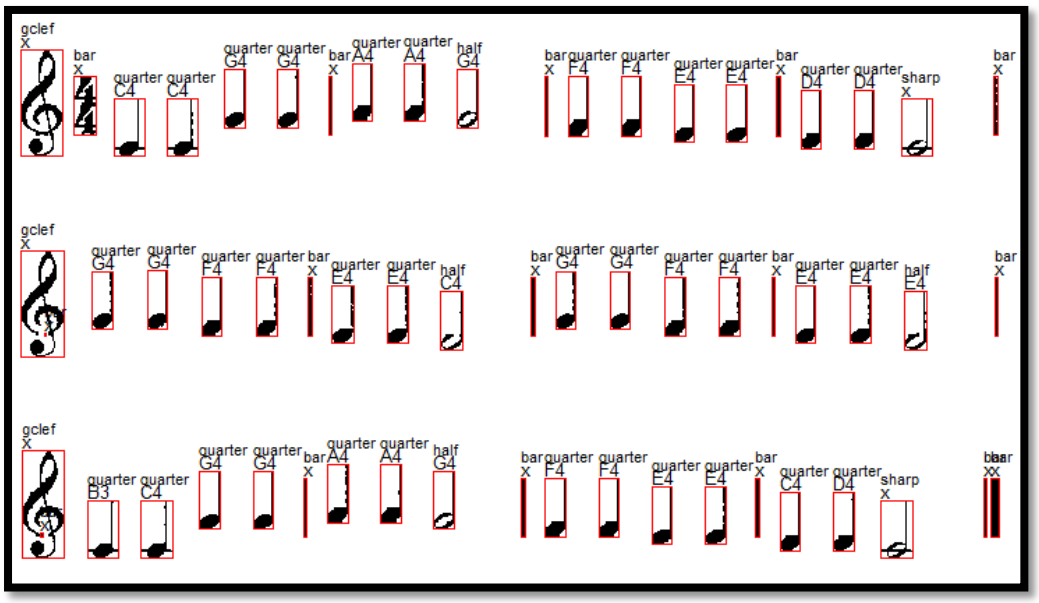


Figure 5 – Skewed Twinkle Twinkle Star Labeled

Old McDonald:

	quarter	half	whole	bar	sharp	gclef
quarter	46	0	0	0	0	0
half	5	0	0	0	0	0
whole	0	0	1	0	0	0
bar	0	0	0	17	0	0
sharp	0	0	4	0	0	0
gclef	0	0	0	0	0	4

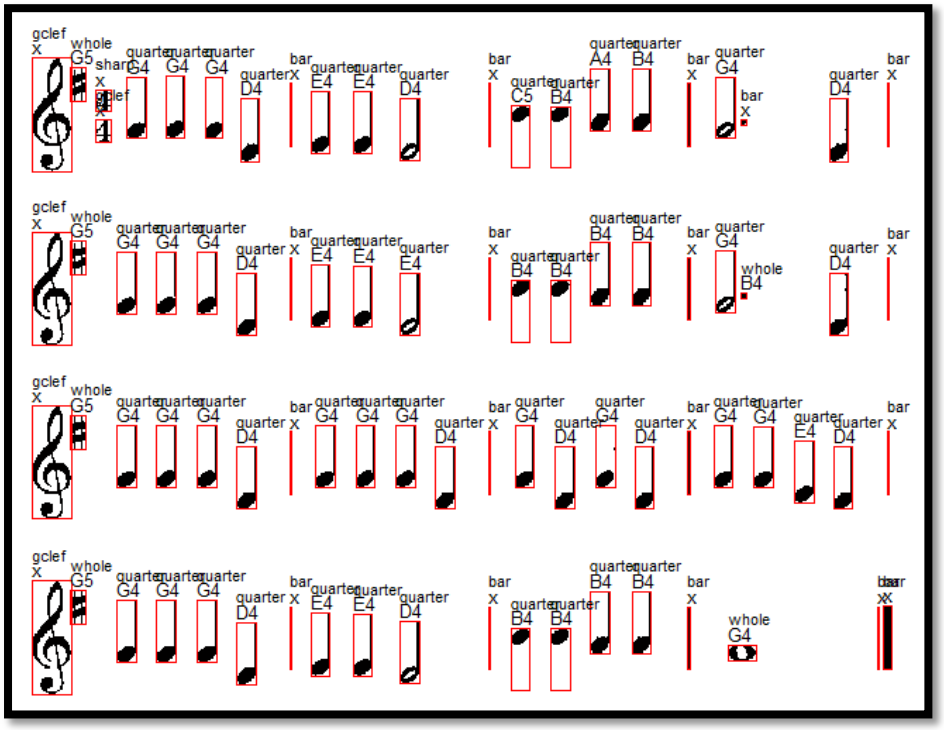


Figure 6 – Old McDonald Labeled

Ode To Joy:

	quarter	half	bar	gclef	sharp
quarter	56	0	0	0	0
half	0	2	0	0	2
bar	0	0	13	0	0
gclef	x	x	x	x	x
sharp	0	0	0	0	0

All the G Clefs in the sheet are fragmented severely due to staff line removal.

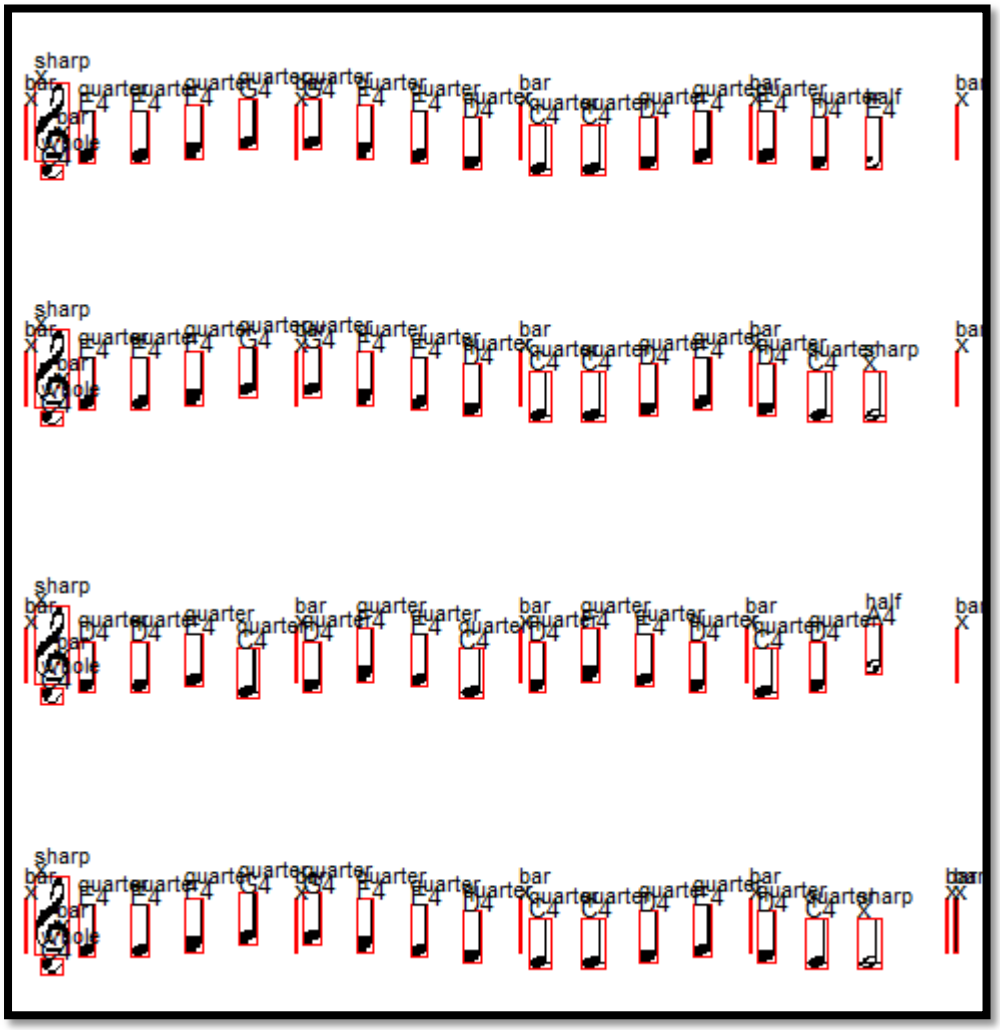


Figure 7 – Ode Yo Joy Labeled

Handwritten Sheet Music :

	quarter	half	bar	gclef	eight
quarter	36	0	0	0	0
half	0	6	0	0	0
bar	0	0	5	1	3
gclef	x	x	x	x	x
eight	0	0	0	0	0

All the G Clefs in the sheet are fragmented severely due to staff line removal.

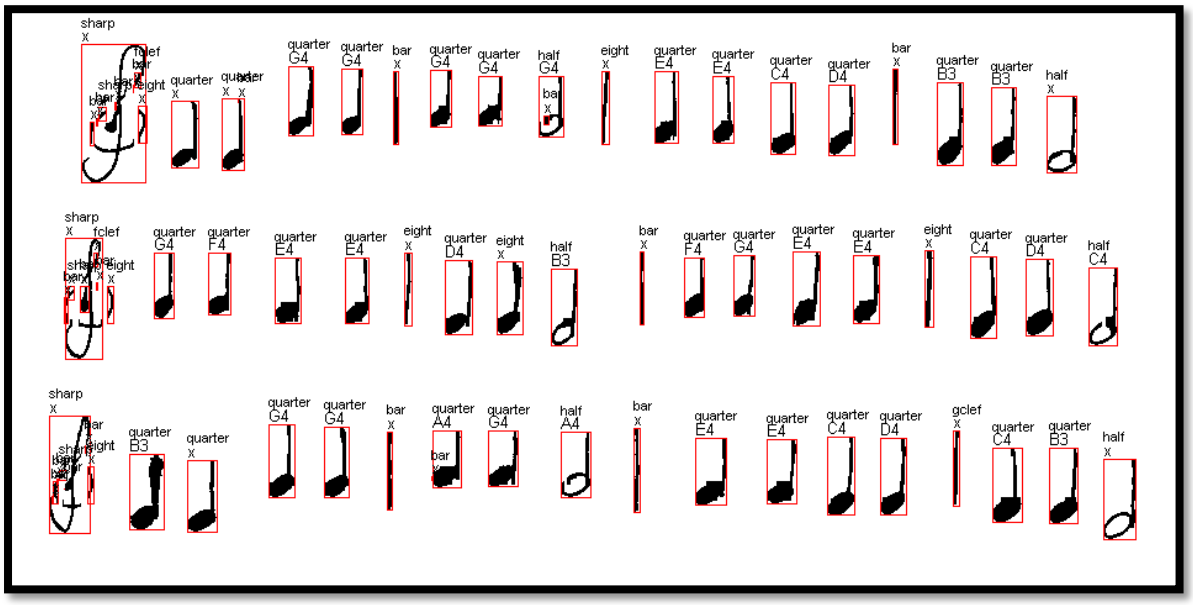


Figure 8 – Handwritten sheet music image labeled

An Additional Result

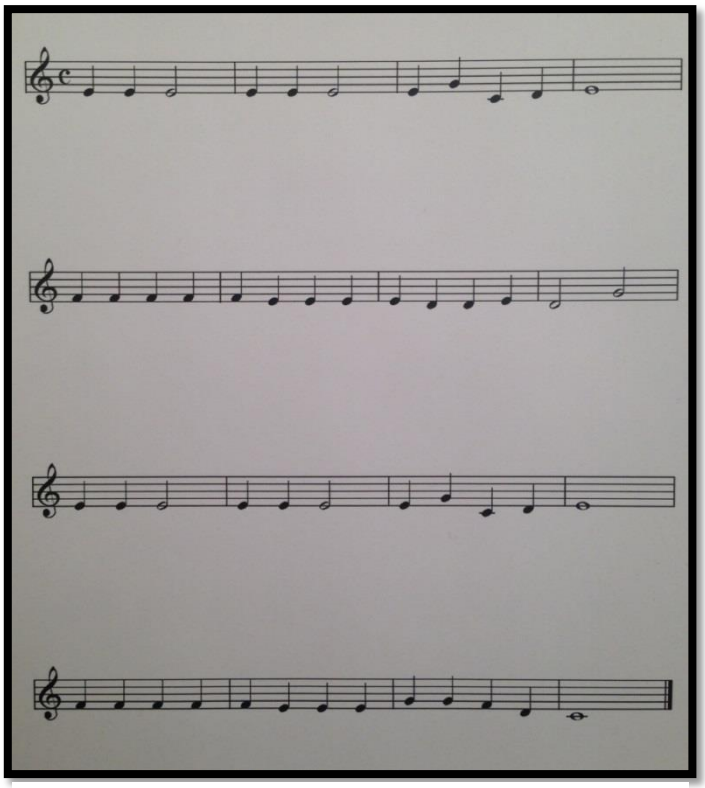


Figure 12 – Jingle Bells sheet music paper photo taken with a cell phone camera

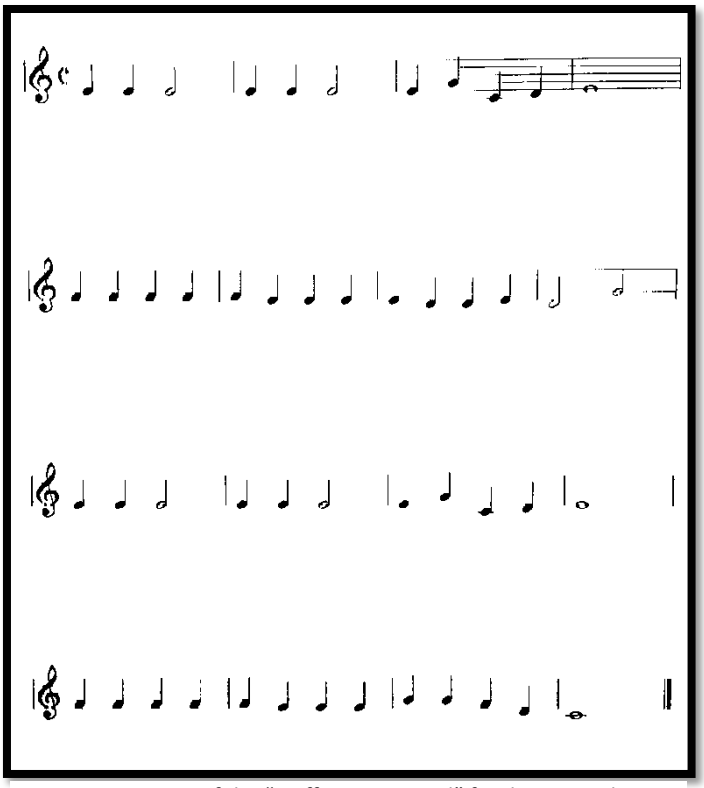


Figure 13 – Output of the "Staff Line Removal" for the music sheet which is taken with a cell phone camera

I also tried my project with a sheet music image which is taken with a cell phone camera. Left image is the sheet music image taken with a cell phone camera and the Right image is the output of my project. However, as we can see my system failed at "Staff Line Removal" stage. There are some staff lines are remaining at the right top of the page.

The reason why "Staff Line Removal" stage failed may be that right top corner of the staff lines are a little bit slanted in the image. Therefore, I conclude that the algorithm used in "Staff Line Removal" is susceptible to very slight curvatures.

Appendix

Jingle Bells

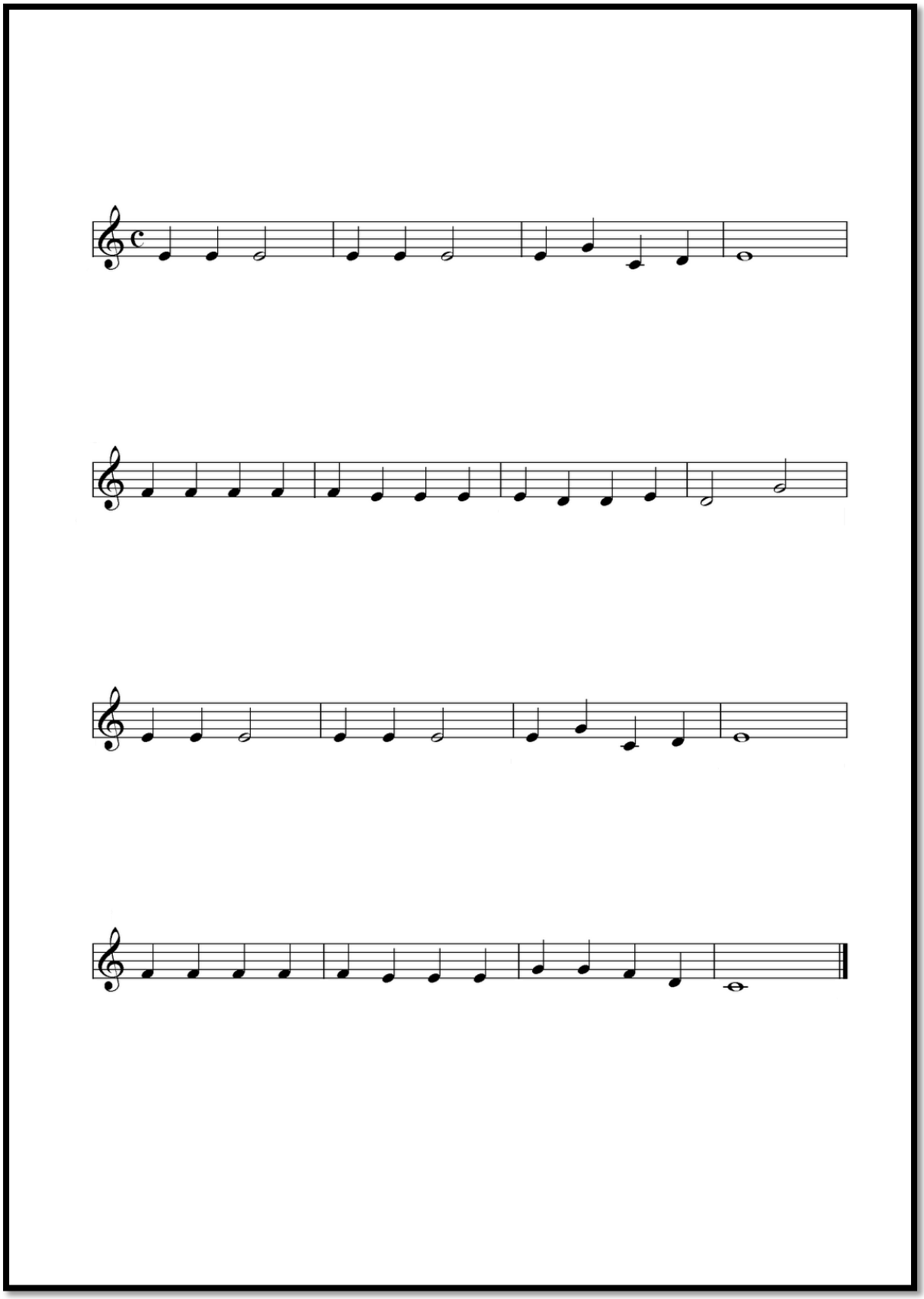


Figure 9 – Jingle Bells Sheet music image

Twinkle Twinkle Little Star

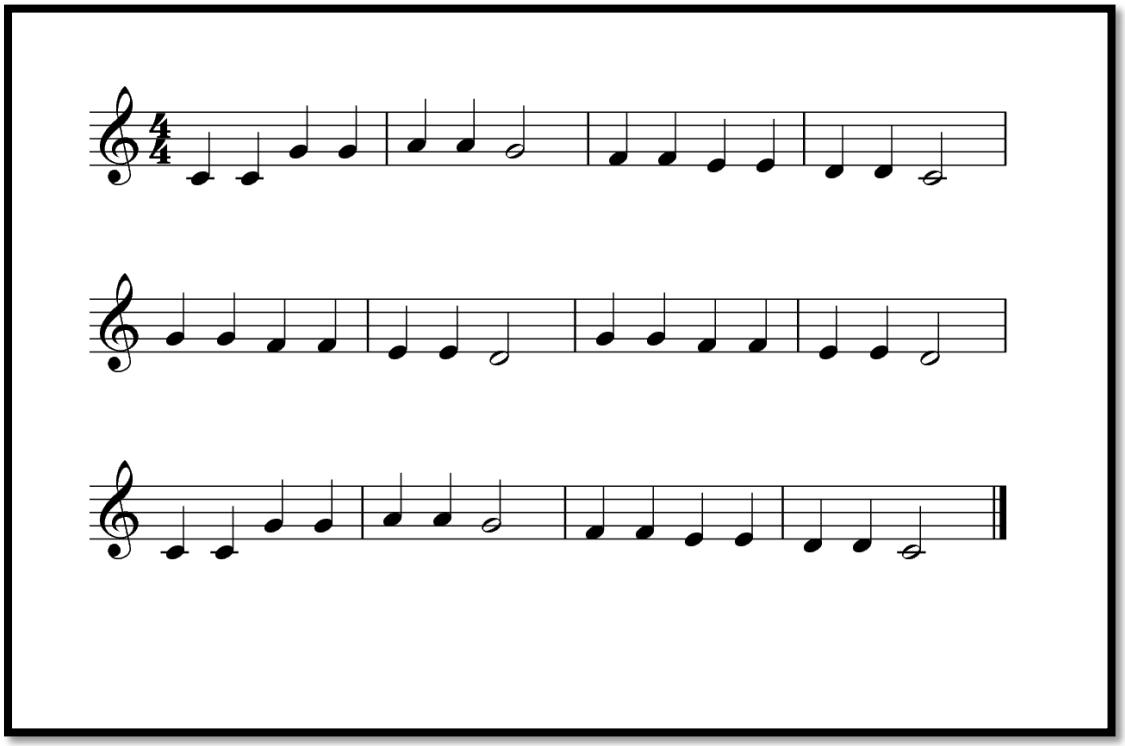


Figure 10 – Twinkle Twinkle Little Star sheet music image

Skewed Twinkle Twinkle Little Star(Scanned with 300 dpi)

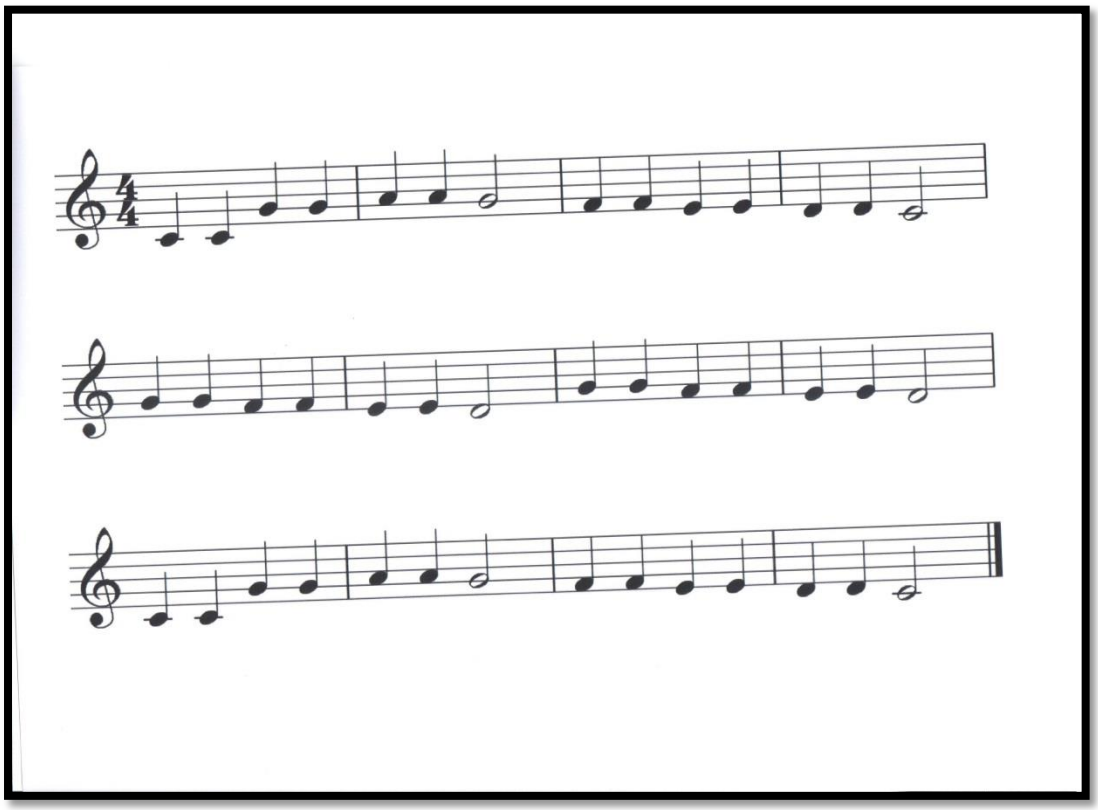


Figure 11 – Skewed Twinkle Twinkle Little Star sheet music image

Old McDonald

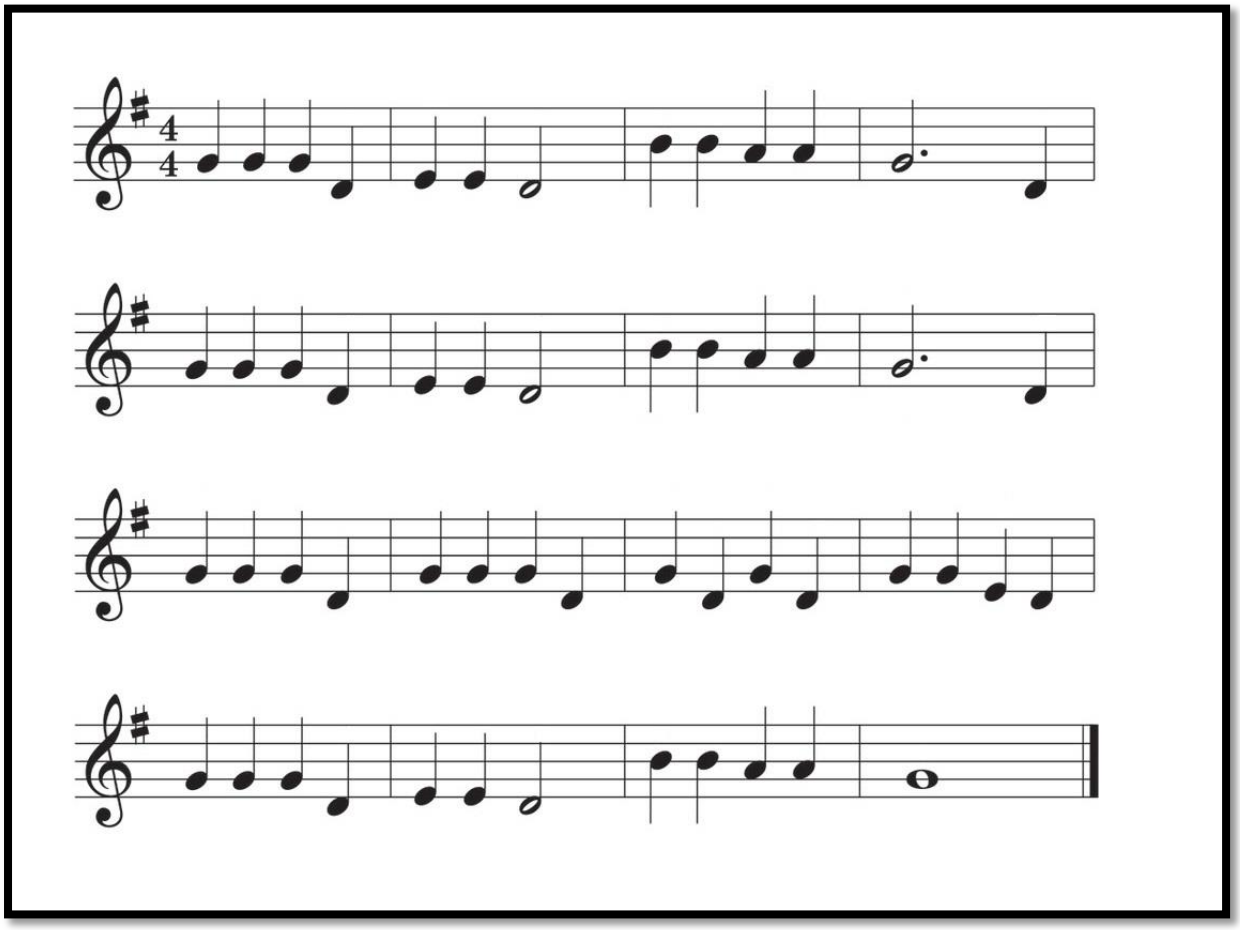


Figure 12 – Old McDonald sheet music image

Handwritten Sheet Music(Scanned with 300 dpi)

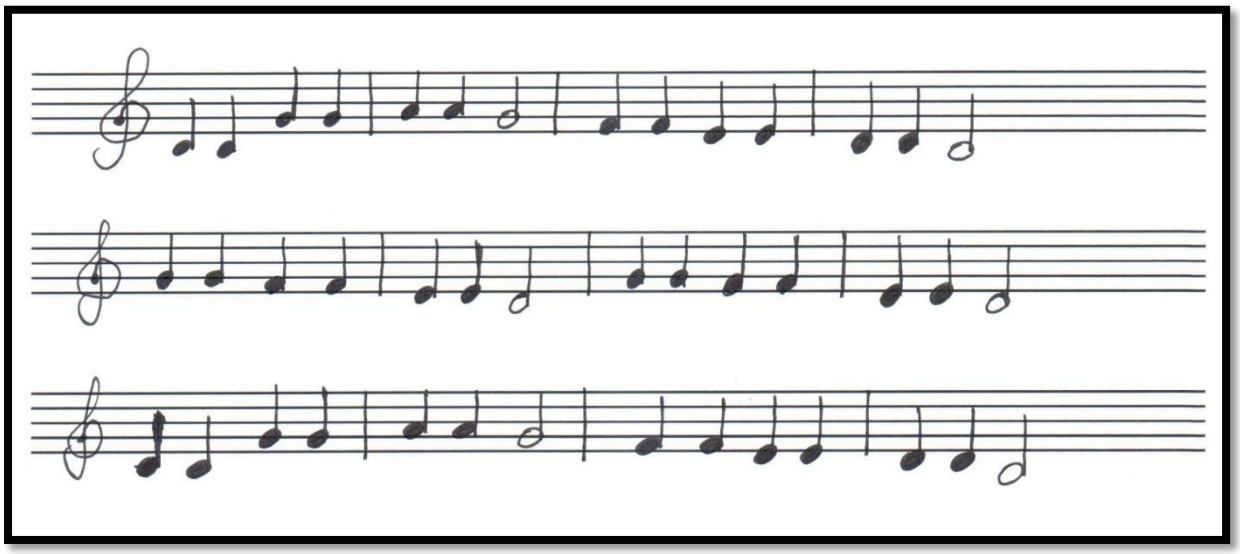


Figure 13 – Handwritten sheet music image

Ode To Joy

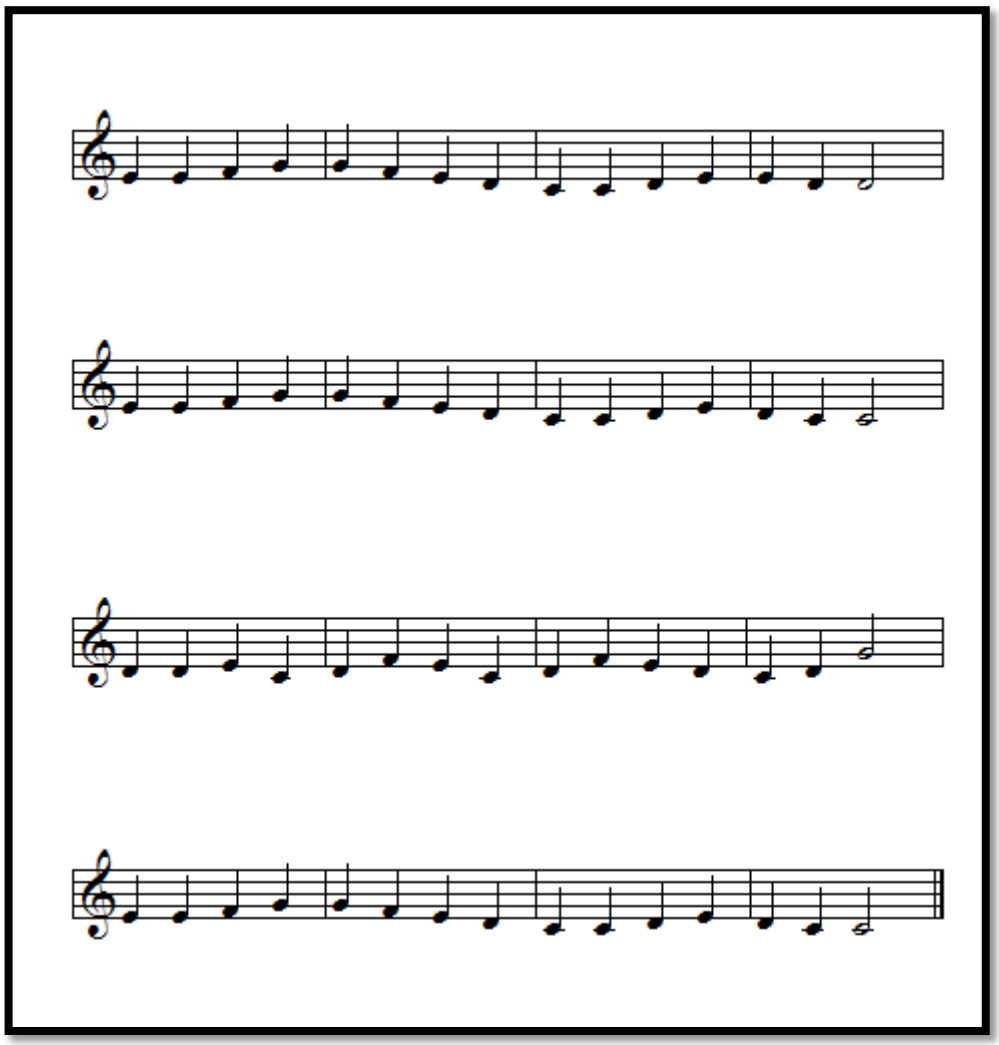


Figure 14 – Ode To Joy sheet music image