Department of Computer Science
Technical University of Cluj-Napoca

Image Processing

Laboratory activity 2018-2019

Radu Danescu, Diana Borza
OpenCV and C++

Dobner Imola and Jakab Gyongyi Aniko
30233
imola.dobner26@gmail.com and jakabgyongyi8@gmail.com

Musical note recognition from sheets

April, 2019

Contents

1 Abstract

Music has had a important role in human life from time immemorial. MUsic has been shared in two ways aurally and as written documents known as musical notes or musical sheets. Many ancient cultures have used musical symbols to represent melodies and lyrics but none of them is as comprehensive as a written language or document.
Hence to preserve such music to introduce a computerized system to digitalize and to decode the musical symbol images and reconstruct it as a new score which will be in machine readable format.

2 Introduction

Optical music recognition(OMR) has been the subject of research for decades. The main goal of all OMR systems is to automatically decode and interpret the symbols of music notation from scanned images.

Musical symbols are processed in order to convert them into machine readable electronic format suitable to store the semantic information (notes, pitches, dynamics and so on) known as MIDI(Musical Instrument Digital Interface). For this process isolation of musical symbols is important. The unique property of musical sheets is staff lines. The five equally spaced lines over which the musical notes are written. These staff lines itself poses a problem when musical sheet segmentation is considered. Since staff lines touches the musical symbols, their removal will leave musical symbols fragmented with corresponding loss of information.

However the techniques developed to address this problem it still poses many challenges to scientists and researchers today.

The main objective of this project is to develop an application capable to parse music sheets from a given database and to supply playback mechanism for it.

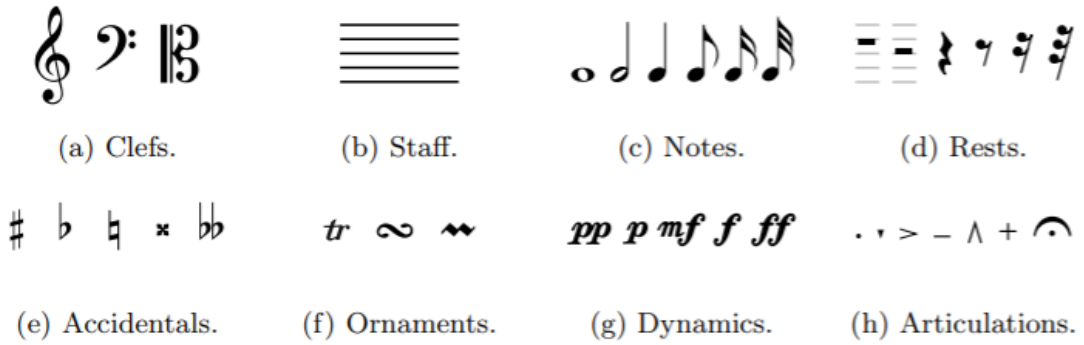
3 Research of Bibliography

As main bibliography source we've chosen is [2], an article about a short introduction to OMR, an overview of what are the main component of a musical sheet and how the different elements are interpreted, and some basis steps of the OMR process itself.

3.1 Music Notation

Music notation has evolved over the period of centuries as the composers and musicians tried to express their musical ideas by written symbols. In the case of CMN (Common Music Notation), there are four types of information involved: a pitch, time, loudness (also dynamics) and timbre

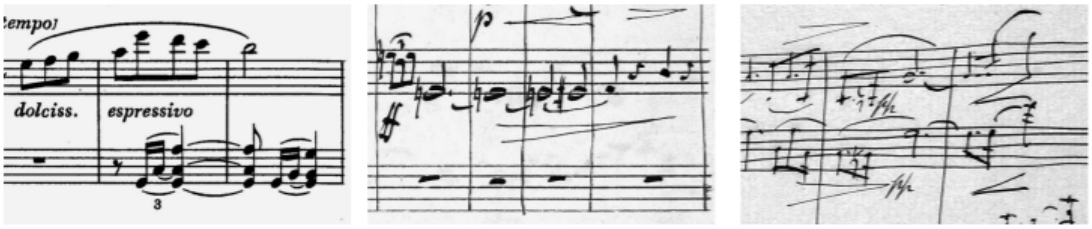
(tone quality).



(a) Clefs. (b) Staff. (c) Notes. (d) Rests.

(e) Accidentals. (f) Ornaments. (g) Dynamics. (h) Articulations.

In practice, certain symbols have almost unlimited variations in representation. The most used CMN symbols and their graphical aspects are listed e.g. in the Essential Dictionary of Music Notation.



(a) Entirely printed. (b) Preprinted staff lines. (c) Entirely handwritten.

3.2 Image Processing

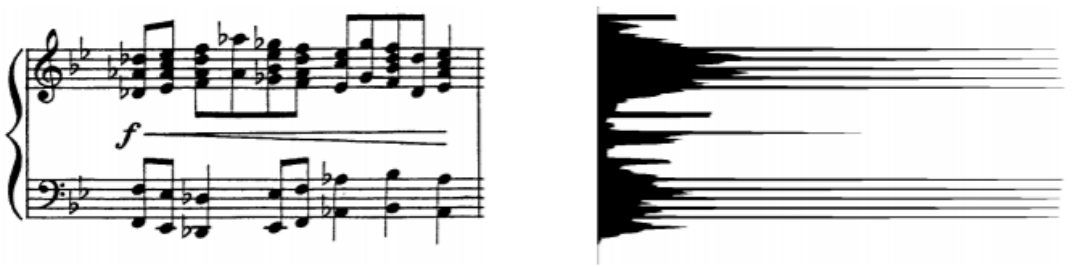
The main goal of the preprocessing phase is to adjust the scanned image to make the recognition process more robust and efficient. Different methods are typically used: enhancement, blurring and morphological operations and noise removal, deskewing and binarization, being the most crucial step in the OMR pipeline.

Binarization algorithms convert the input image into a binary one, where objects of interest (music symbols, staves, etc.) are separated from the background. This is motivated by the fact, that music scores have inherently binary nature. It facilitates the subsequent tasks by reducing the volume of information that is needed to be processed.

3.3 Segmentation

3.3.1 Staff Lines

Staff line detection is fundamental in OMR, because the staff creates a two dimensional coordinate system essential to understand the CMN. Unfortunately, staff lines are not guaranteed to be perfectly horizontal, straight or of uniform thickness in scanned images. Precise staff detection is a tricky problem that still represents a challenge.



3.3.2 Symbol segmentation

Once the staff lines have been detected, the music primitives must be located and isolated. This can be performed in two manners: either remove the staff lines or ignore them. Although the majority of researchers remove the staff lines in order to isolate the musical symbols as connected components, there are some authors who suggest the opposite.

3.4 Object Recognition

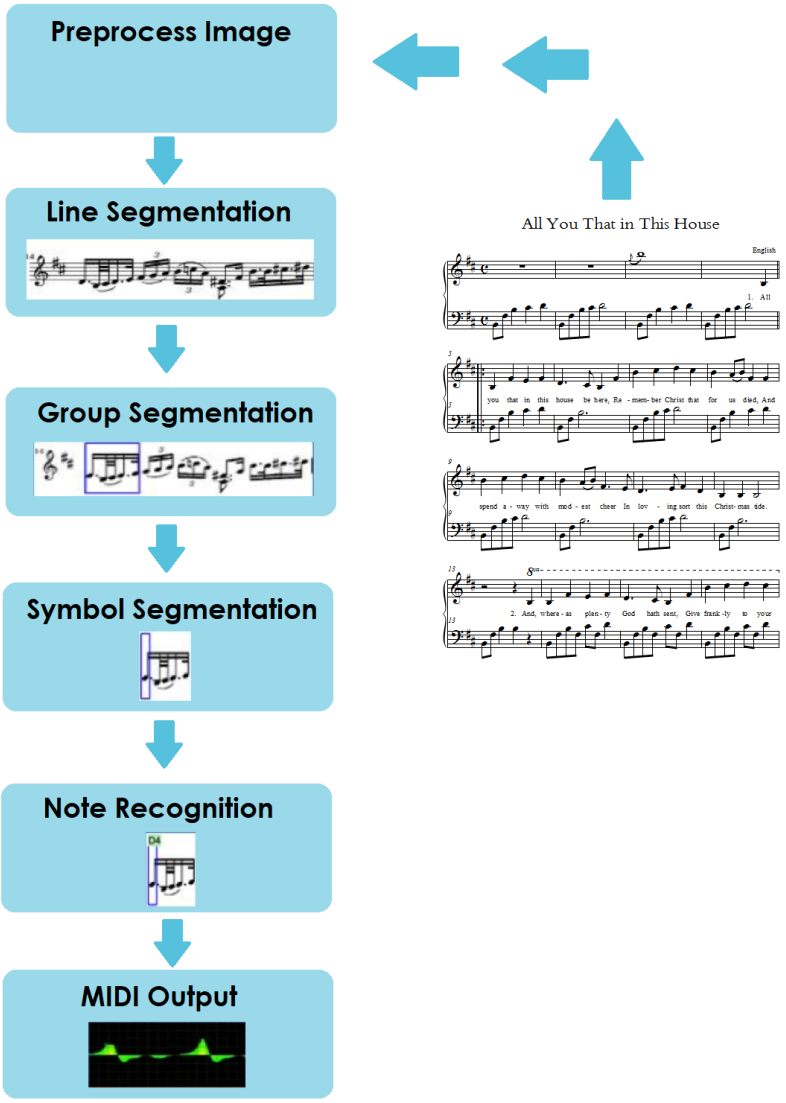
Segmented symbols are further processed and given to the classifier that tries to recognize them. Unfortunately, music shapes are inherently complex — they are often formed by several touching and overlapping graphical components. In addition, the staff line removal can break some objects. Hence, the object recognition phase is very delicate and it is usually combined with the segmentation step.

3.5 Semantic Reconstruction

The inevitable task of all OMR systems is to reconstruct the musical semantics from previously recognized graphical primitives and store the information in a suitable data structure. This

necessarily requires an interpretation of spacial relationships between objects found in the image. Relations in CMN are essentially two dimensional and the positional information is very critical.

4 Proposed Method



4.1 Loading Image & Pre-processing

The images used for this system are stored in the folder in the path of the project, there are 20 images used for system testing of different resolutions, number of staff lines and musical notes. In case of loading a specific image a segment of code is needed to be modified, specifically the name of images that is needed to be loaded. The preprocessing part of this system consist of a grayscale convertion from the original colored image and an adaptive binarization step, where the

threshhold value used for the global automated binarization based on a specific error value. This algorithm is based on processing of the histogram of an image by searching the best middle value between the two maximums of the image histogram (one maximum representing the background and another one representing the objects).

4.2 Line segmentation

The first segmentation type of the system is applied on the entire image. By applying horizontal projection on the image we can identify where are or are not white spaces. White spaces separate the staff lines and musical notes on them, by allimminating the white spaces we segmentate the image into multiple smaller parts. At this step of the algorithm for each separated line there is created a bounding box of color yellow, so that it can be represented on the final image.

4.3 Identifying staff lines & Group segmentaion

For every and each staff line image we use a combination of erosion and dilation for destroying the staff lines, but also extract the coordinate values of each staff line. The staff line information will be stored in an array on 5 elements. After the staff lines are destroyed by using erosion (a structure that has the form of a rectangle), the musical notes and signs are reconstructed using dilation to almost the original form.

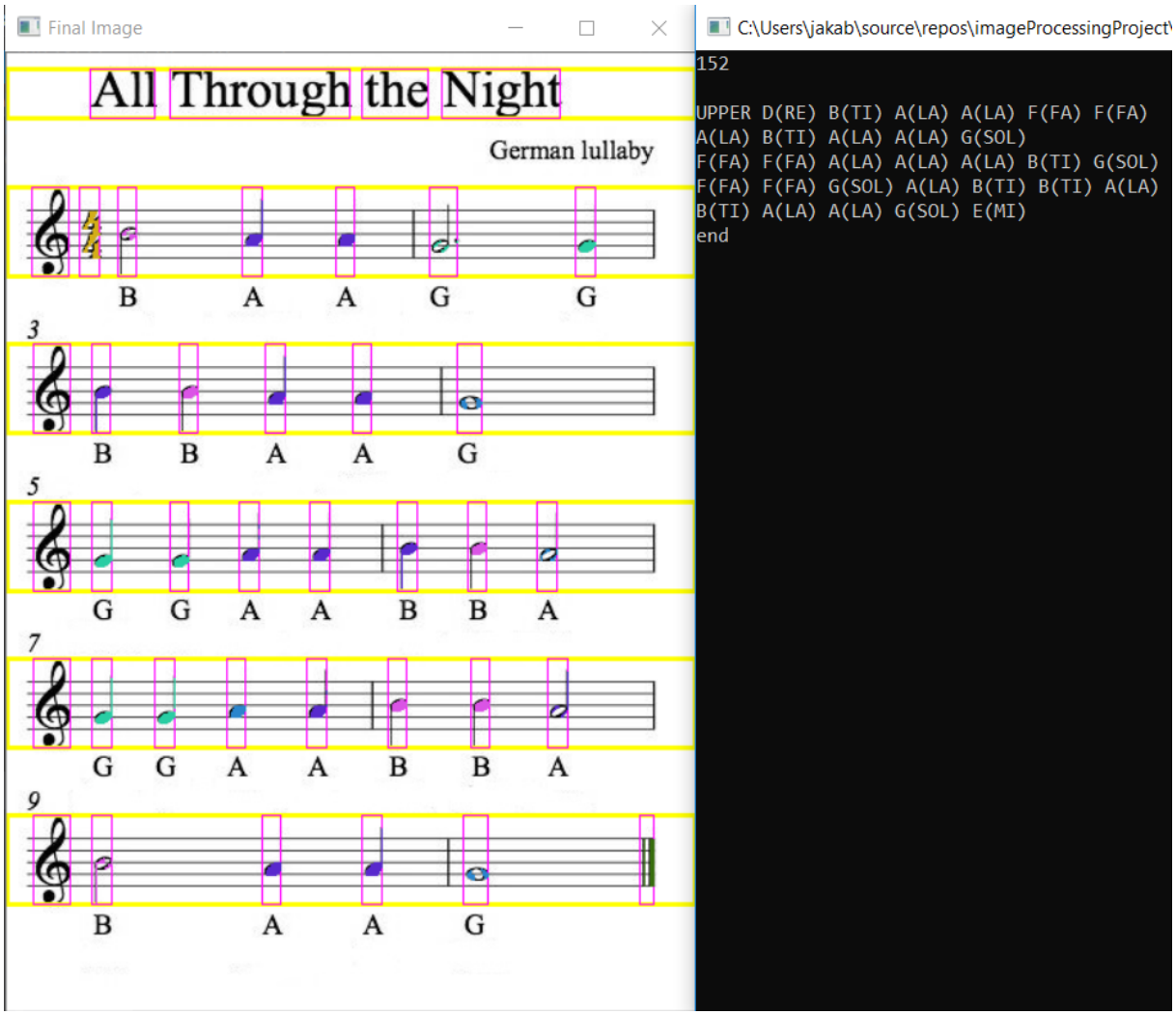
4.4 Note segmentation

Form each image representing a single line there is applied vertical projection to detect white spaces. The image in which the white spaces are detected has no staff line, just musical notes and signs. After the regions with no black pixels are detected each range will represent a group of musical note. This smaller image may or may not contain multiple musical notes/ signs. After the whole group segmentation of the original image is finished, an analysis of these elements are made for determining the median width of a group. Further more each group image is checked if its width is smaller or bigger than the average one, if smaller it is considered a single note, otherwise a group of notes that need further segmentation. Segmentation of a group is made using vertical projection as the line to group segmentation, however not the white spaces, but the average black pixel segments are detected.

4.5 Note recognition

The specific musical note recognition analysis the image and the staff line coordinates send as function pharameters. The system stores the can detect whether the musical note is in the upper, normal or lower category and the musical notes name, for the moment there is no specific method

how to detect the type of the note - whole, half or other.



4.6 Music reproduction

At the beginning of the project we talked about representing the musical notes in form of sounds after musical note detection. Because we made to steps for detecting the type of musical note, we don't now the duration of the musical note when the music is played.

5 Experimental results

To get better result for musical note detection we've realized that experiments are needed to be made. The threshhold value needed to be automatized in this process, because in the most of the images the musical notes had a gray surrounding that in case of binarization whould made the

note unrecognisable. Other error values that may need addaption to a specific image are the error values in the line and group segmentation.

6 Conclusions and Improvements

At the current state the musical note processing system is not able to replay the music from a musical sheet, but can identify musical notes in 70% of the cases. In the future steps will be made to make the system mode adaptive and precise.

7 References

[1] Musical Notes Reader
Anna Shmushkin & Lior Abramov
http://www1.idc.ac.il/toky/ImageProcAndroid-15/Musical%20Notes/report.pdf

[2] Introduction to Optical Music Recognition: Overview and Practical Challenges
Jirı Novotný & Jaroslav Pokorný
http://ceur-ws.org/Vol-1343/paper6.pdf

[3] Musical Sheet Segmentation
Vinaya V.& Dhanya M. Dhanalakshmy
https://pdfs.semanticscholar.org/bdc7/8eccc132590984d270ced14bbb952623c516.pdf