CS 766 - Music Symbol Segmentation in OMR

Stephen Sheen (Email: sheen2@wisc.edu, UW ID: 9067861477),
Akshaya Kalyanaraman (Email: kalyanarama3@wisc.edu, UW ID: 9080068340),
Abbinaya Kalyanaraman (Email: akalyanaram2@wisc.edu, UW ID: 9080066682)

1 Problem Statement

Optical music recognition (OMR) is a process of transforming music scores into digital formats.

The pipeline for OMR is usually divided into four steps: Image pre-processing, such as binarization, noise removal, and etc.

1. Music symbol segmentation and recognition
2. Music symbol semantic recognition
3. Reconstruction of music sheet into a digitized format

The second step, music symbol segmentation and recognition, is further divided up into three sub-steps: staff detection and removal, symbol segmentation, and symbol recognition [Rebelo et al., 2010].
In this project, we will focus mainly on the sub-step of music symbol segmentation. More specifically, we will like to look at the specific problem of handwritten music symbol segmentation, since this is an area of research that is still far from perfect.

2 OMR: Importance and an emerging area of interest

OMR has been an area of active research in the past decades. [[Blostein and Baird, 1992],[Rossant and Bloch, 2007],[Toyama et al, 2006]]. One major reason for this active research is because of the preservation of cultural heritage. Before the age of technology, music scores had always been handwritten on paper. Due to them being paper-based, the problem of natural paper degradation cannot be avoided. In order to prevent these music scores from being lost, they will need to be stored differently, and digitalization is one of the best ways to do that.

Apart from the preservation of cultural heritage, OMR is also important to the people today. With the existence of the Internet, file sharing has become an easily achievable task. As a result, the access of music, in the form of digitized music sheets, has been made a lot easier. For modern composers and other music-related industries, being able to edit music scores that were once in paper-based form and reprint them can be very beneficial. In order to achieve all of the tasks described above, however, a robust optical music recognition pipeline needs to be developed. [Novotny and Pokorny, 2015]

3 State-of-the-art in Music Symbol Segmentation

The current state-of-the-art in music symbol segmentation uses a hierarchical decomposition approach [Novotny and Pokorny, 2015]. Hierarchical decomposition is a top-down technique commonly used to analyze a complex system, by iteratively breaking down the system into smaller and smaller subcomponents that are easier to understand. In terms of OMR, this means first locating where the staves are and removing them. After the staves are removed, symbols are located and segmented into vertical slices. Finally, from the vertical slices, symbols are extracted and stored for the

next symbol recognition step [Mehta and Bhatt, 2015].

Apart from hierarchical decomposition, other methods, such as Mahoneys candidate sets and Carters linear adjacency graphs, also exist [Rebelo et al., 2012]. Although these methods work well for typeset music scores, they often do poorly on handwritten scores. The reason for this is because handwritten scores often contain a number of variabilities which typeset scores do not contain. For example, there are countless of different handwriting styles, which can cause music symbols to differ in sizes and shapes. Variabilities in the appearances of compound music notes also create problems for these methods [Bar et al, 2016]. Therefore, current research are focused on improving the performance of handwritten music symbol segmentation (and OMR in general).

4 Project Approach

For this project, we would like to propose a new approach to the problem of music symbol segmentation. However, we would like to first implement the algorithm proposed in [Bar et al, 2016], which is a learning-free algorithm to segment and recognize handwritten music notes. The algorithm first detects music notes as graphic primitives (i.e. stems, bars, and note-heads), then uses a hierarchical clustering algorithm to group these graphics primitives together, and finally recognizes notes with pre-written rules. After testing the implementation using publicly open dataset of handwritten scores, we would like to come up with our own approach that uses machine learning.

5 Justification of proposed approach

As seen in [Bar et al, 2016], their algorithm only works moderately well for handwritten music scores. More specifically, the precisions and recalls of their algorithms are around 0.4 to 0.8. We believe that with a learning-based approach, we should be able to achieve precisions and recalls that are within a higher range.

6 Evaluation

For the experiments, we will be selecting a dataset containing handwritten music notes which are to be segmented. The experimental results can be formulated into a table, where the first column indicates the music pages that have been used, the second column indicates whether the score is polyphonic or monophonic, the third and fourth columns show the detection of note-heads, whereas the last two columns show the detection of music notes (e.g. half, quarter, 8th note, etc.).

The metrics used are the precision (number of correctly detected elements divided by the number of detected elements), and recall (number of correctly detected elements divided by the number of elements in the dataset). As seen in [Bar et al, 2016], the mean precision and recall is around 52 %, and we propose to achieve better results in our approach.

7 Timeline

1. February 20 - Project webpage set-up
2. March 8 - Algorithm as proposed in [Bar et al, 2016] implemented
3. March 22 - New learning-based algorithm proposed
4. April 12 - New learning-based algorithm implemented
5. April 19 - New learning-based algorithm tested

8 References

[Novotny and Pokorny, 2015] Novotny, J. and Pokorny, J. (2015). Introduction to optical music recognition: Overview and practical challenges.

[Mehta and Bhatt, 2015] Mehta, A. A. and Bhatt, M. S. (2015). Optical music notes recognition for printed piano music score sheet.

[Rebelo et al., 2012] Rebelo, A., Fujinaga, I., Paszkiewicz, F., Marcal, A. R. S., Guedes, C., and Cardoso, J. S. (2012). Optical music recognition: state-of-the-art and open issues.

[Rebelo et al., 2010] Rebelo, A., Capela, G., and Cardoso, J. (2010). Optical recognition of music symbols - A comparative study.

[Bar et al, 2016] Bar, A., Riba, P., and Forns, A. (2016). Towards the recognition of compound music notes in handwritten music scores.

[Blostein and Baird, 1992] Blostein, D. and Baird, H. S. (1992). A critical survey of music image analysis.

[Rossant and Bloch, 2007] Rossant, F. and Bloch, I. (2007). Robust and adaptive omr system including fuzzy modeling, fusion of musical rules, and possible error detection.

[Toyama et al, 2006] Toyama, F., Shoji, K., and Miyamichi, J. (2006). Symbol recognition of printed piano scores with touching symbols.