Music Score Recognition
Literature Review

Piotr Wilczyński

November 16, 2017

1 Introduction

The purpose of this literature review is to illustrate what previous research has been done on the subject and to identify the possible approaches to the problem. It is also to demonstrate the gaps in the referenced literature and how the state-of-the-art methodology can be improved.

2 State-of-the-art approach

As described by Homenda (2005), sheet music has a logical and a geometrical interpretation: it is a time line of events of different lengths but it has to be broken up into pieces (systems) to be printed on paper. It is also comprised of three dimensions: time, frequency and part/instrument. Essentially, it constitutes a set of notes with certain properties. Every part has a stave (five horizontal lines) and a clef that determines the location of pitches on it. A stave is split by horizontal lines into bars, each having equally many beats. Bar lines are also used to indicate repetition. Bars are stretched horizontally to fit the page width rather than separated between systems. Empty staves (for voices/instruments that are currently not involved) are usually dropped. There is also textual information on the sheet - the title, composer, copyright etc.

First, if the image to be processed has been obtained by scanning, it has to be binarized so that there are only black and white pixels left (Tardón et al., 2009). Since stave lines occupy a relatively large fraction of the page width, they are easy to identify by horizontal projection (counting the number of pixels in every row and considering the highest numbers). As described by Rebelo et al. (2009), this projection can also be used to straighten the scanned image (by maximising the histogram maxima). Then, the symbols are categorized into several classes: ones with fixed positions (clef, time signature), ones with easily identifiable features (vertical lines - could be detected with vertical projection), ones whose location is determined by the location of other symbols (dots, beams connecting quavers) and randomly placed ones (clef changes, ornamentation).

Symbols are assigned different labels so that it is possible to distinguish them at later stages. According to He et al. (2008), there are four main types of labelling algorithms: multi-scan, two-scan, hybrid and tracing-type. Labelling is performed by traversing the image with a mask and looking at 8-connectivity. In the end, symbols are provided with bounding boxes so that their dimensions are known. Some of them, such as clefs, will have sizes relative to the stave. During the recognition, symbols can be considered in their original form or with the stave lines removed. Algorithms used for detection include neural networks, statistical classifiers and classification trees.

Rebelo et al. (2009) describe how neural networks can be used for recognition. The specific architecture involved in the process is called multi-layer perceptron. Moreover, a data set has to be collected that is then split randomly into a training set and a test set.

Music Recognition is essentially about constructing a logical model representing the notation on the sheet (Rebelo et al., 2009). Once this has been done, the model can be automatically converted into an audio file (Tardón, 2009).

Byrd et al. (2010) demonstrate that due to the logical structure of sheet music, its recognition is much more difficult than alphanumeric character detection. There are symbols related to several other symbols (signatures, accidentals, accents) and some symbols might take different shapes (joined quavers). Sheet music is also characterized by overlapping symbols and zones of high symbol density (Rebelo et al., 2009). There are multiple ways of testing a recognition algorithm, e.g. by considering the number of wrong pitches, wrong lengths, added notes, missing notes etc.

3 Rationale for further research

According to Homenda (2005), a lot of improvement can still be achieved in this area. For instance, recognizing complex musical symbols and elaborate sheet music is a difficult task and the current approaches do not deal with it as effectively as with simple music notation (Doermann and Tombre, 2014). More research can also be done into handwritten music recognition (Rebelo et al., 2009) and old music recognition to ensure its preservation (Tardón, 2009).

Homenda (2005) claims that stave line removal is a difficult and time-consuming process, however it has been already implemented as part of this project and proves to be working quite well so far. Therefore, it is useful to experiment with different approaches to music score recognition to find out what works most accurately and efficiently.

References

Byrd, D., Guerin, W., Schindele, M., Knopke, I. (2010) *OMR Evaluation and Prospects for Improved OMR via Multiple Recognizers.*

Doermann, D. and Tombre, K. (2014) *Handbook of Document Image Processing and Recognition*. London: Srpinger-Verlag.

He, L., Chao. Y., Suzuki, K. and Wu, K. (2008) Fast connected-component labeling. In: E. Hancock, *Pattern Recognition*. ScienceDirect.

Homenda, W. (2005) *Optical Music Recognition: The Case Study of Pattern Recognition*. Warsaw: Faculty of Mathematics and Information Sci, Warsaw University of Technology.

Rebelo A., Capela G. and Cardoso, J.S. (2009) *Optical recognition of music symbols*. Springer-Verlag.

Tardón, L.J., Sammartino S., Barbancho I., Gómez, V. and Oliver A. (2009) *Optical Music Recognition for Scores Written in White Mensural Notation*. Malaga, Departamento de Ingeniería de Comunicaciones, E.T.S. Ingeniería de Telecomunicación, Universidad de Málaga.