

OPTICAL MEASURE RECOGNITION IN COMMON MUSIC NOTATION

Gabriel Vigliensoni
gabriel@music.mcgill.ca

Gregory Burlet
gregory.burlet@mail.mcgill.ca

Ichiro Fujinaga
ich@music.mcgill.ca

Abstract

We present an optical measure recognition algorithm capable of recognizing the physical location of barlines in a wide range of scores of common Western music notation (CWMN), deriving the bounding boxes for the measures, and storing these elements in a symbolic music file format (MEI).

When testing the performance of the algorithm on a ground-truth dataset of 100 images of music scores with manually annotated measures, our algorithm obtained an *f*-score of 91 percent with the optimal set of parameters.

Measure recognition algorithm Approach

Our technique for locating the bounding boxes of measures within a music score relies on applying several image processing functions to the key stages of a decomposed OMR situation: image preprocessing and normalization, staffline identification and removal, musical object location, and musical reasoning.

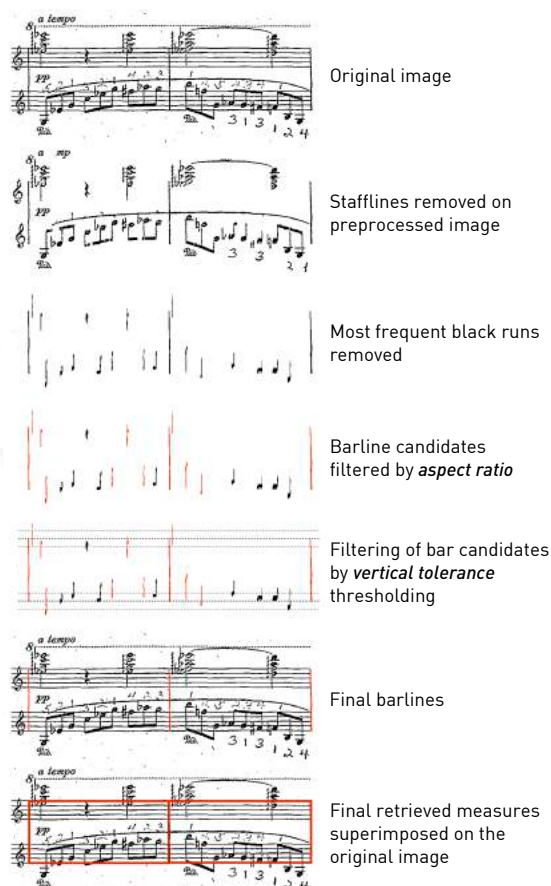


Figure 1. Process of extracting barline candidates.

Implementation and critical parameters

Our implementation uses *Gamera* as the software framework to perform all pre-processing and filtering in the images, the *MusicStaves* Gamera toolkit to detect and remove staff lines, and MEI as the output symbolic file format. Barline candidates are discarded according to two parameters:

- **Aspect ratio:** relation between the width and the height of a glyph
- **Vertical tolerance:** allowed upper bound of the ratio between the height of a bar candidate and the system height where it belongs

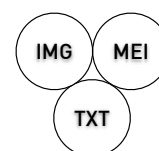
Staff-group hint

As we want to properly encode the output of our measure recognition algorithm into a symbolic music file, our approach requires prerequisite information supplied by humans, that encodes the structure of the staves on the music score being processed. We called this information the *staff-group hint*.

Algorithm evaluation Ground-truth dataset

As there are no OMR datasets with annotated measure bounding boxes we created our own to evaluate the performance of our system. It consists of:

- 100 pages of common Western music notation randomly extracted from the International Music Score Library Project (IMSLP).
- MEI files with the absolute position of each measure occurring in the entire dataset, annotated by two musically-trained students with a custom-made application.
- TXT files with the staff-group hint for each image.



Metrics and evaluation

The performance of our optical measure recognition algorithm was evaluated by computing precision, recall, and *f*-score statistics for the automatically recognized measures on each page of music in the ground-truth dataset.

Since we are not only concerned with the number of retrieved measures but also their size and physical position on the page, a measure is considered correctly recognized if its bounding box coordinates are within a quarter of an inch of its corresponding bounding box in the ground-truth measure annotations (measurements in inches are converted to pixels using the *pixels per inch* parameter from the metadata of each image in the dataset).

Parameter tuning

Experiments were conducted to investigate the impact of different values of the *aspect ratio* and the *vertical tolerance* parameters. We iterated over a set of 100 combinations of these parameters.

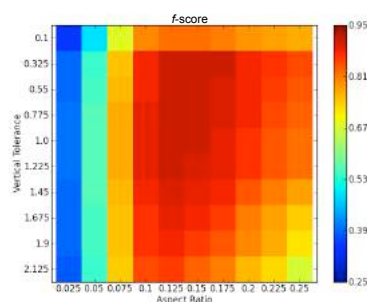


Figure 2. *f*-score results of the optical measure recognition algorithm. The x-axis displays different values of the *aspect ratio threshold* parameter. The y-axis displays different values of the *vertical tolerance threshold* parameter.

Results

- *Aspect ratio* parameter highly influences the performance of the algorithm
- *Vertical tolerance* parameter does not have a strong impact, especially in the optimal value of the *aspect ratio*.
- With the best settings for the two parameters, our implementation yielded an average of *f*-score of 0.91 for the entire dataset.

Acknowledgements

The authors would like to thank our great development team for their hard work: Nick Esterer, Wei Gao, and Xia Song. Special thanks also to Alastair Porter for his invaluable insights at the beginning of the project, and Andrew Hankinson for his thorough proofreading.