STAFF-LINE DETECTION ON GREYSCALE IMAGES WITH PIXEL CLASSIFICATION

Jorge Calvo-Zaragoza
PRAI Group
Software and Computing Systems
University of Alicante
Alicante, Spain
jcalvo@dlsi.ua.es

Gabriel Vigliensoni
CIRMMT
Schulich School of Music
McGill University
Montréal, Canada
gabriel@music.mcgill.ca

Ichiro Fujinaga
CIRMMT
Schulich School of Music
McGill University
Montréal, Canada
ich@music.mcgill.ca

ABSTRACT

Staff-line detection is an important processing step in most Optical Music Recognition systems. Traditional methods make use of heuristic strategies based on image processing techniques with binary images. However, binarization is a complex process for which it is difficult to achieve perfect results—especially in ancient musical documents. In this paper we describe a staff-line detection approach that can deal directly with greyscale images. It is devoted to classifying each pixel of the image as *symbol*, *staff*, or *background* by means of machine learning algorithms. In order to perform this classification we use Convolutional Neural Networks. The initial features of each pixel consist of a square patch from the input image centered at the pixel to be classified. Our preliminary results show a promising performance.

1. INTRODUCTION

Optical Music Recognition (OMR) systems have to deal with many aspects of music notation, detecting staff lines being one of the most challenging obstacles [4]. These lines are necessary for human readability, yet they complicate the automatic isolation of music symbols. A few previous OMR approaches took advantage of specific features of printed notation to approach the problem keeping the staff lines [1, 9]; however, most common OMR pipelines include their detection and removal as required steps [10].

A number of methods addressed the staff-line detection and removal stage [3, 6], but most of them were based on hand-engineered transformations applied to binary images. Nevertheless, automatic image binarization is not a trivial matter, and so perfect performance can not be guaranteed. Therefore, the performance of subsequent processing stages, such as staff-line detection and removal, are limited by a poor image binarization step.

A data-driven strategy for staff-line detection and removal was recently proposed in the work of Calvo-Zaragoza et al. [2]. Their approach consisted in training a classifier that learns to discriminate if a pixel belongs to a symbol or to a staff line. Then, the pixels of the image are queried so that those classified as *staff* are removed. In this paper we present an extension of this idea so that staff lines can be directly detected and removed using greyscale images rather than from binary images.

2. STAFF-LINE DETECTION ON GREYSCALE IMAGES WITH PIXEL CLASSIFICATION

Our approach considers staff-line detection as a classification task at pixel level. We use a Convolutional Neural Network [7] trained to distinguish among three categories: *symbols*, *staff lines*, and *background*. We assume that the region surrounding each pixel of interest contains enough information to discriminate the label of each pixel. Hence, the input to the network is a portion of the input image centered at the pixel of interest. The size of this region has to be tuned according to specific features of the musical document, such as the staff line thickness or the space between staff lines. Figure 1 shows a portion of a greyscale musical score with two windows centered at pixels of interest.

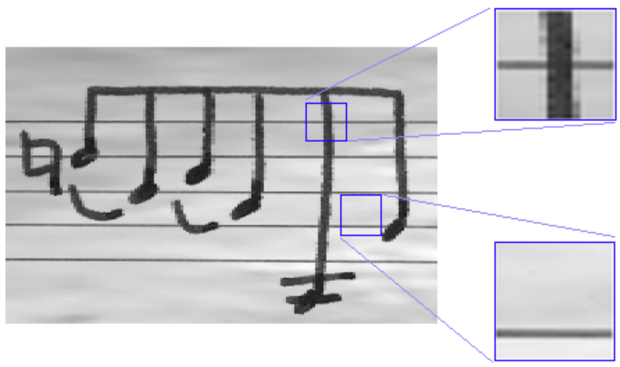


Figure 1. Example of feature extraction with square patches. The pixel to be classified is in the center of the extracted patch of image. Depending on the surrounding information, these pixels will be labeled as *symbol*, *staff line*, or *background*.

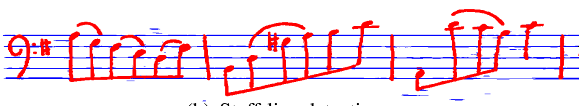
No further feature extraction is performed on this portion of the image because this task is expected to be assumed by the CNN itself [8]. Once pixels have been labeled appropriately, it is easy to perform the line removal by just removing every pixel that has not been labeled as *symbol*.

3. PRELIMINARY RESULTS

We carried out a preliminary experiment with images of the CVC-Muscima dataset [5]. This corpus contains pairs of scores with and without staff lines. Also, the scores are distributed in binary as well as greyscale formats. Therefore, this dataset provides us with readily-available data for training the network, as well as testing data for evaluating our approach.

Figure 2 shows an example of the classification process for both staff-line detection and removal. The resulting image shows each category highlighted in a different color.

(a) Greyscale score



(b) Staff-line detection

(c) Staff-lines removal

Figure 2. Staff-line detection and removal on greyscale images. Pixels are colored according to their predicted category: *symbol* as red, *staff* as blue, and *background* as white.

Although some pixels were mislabeled in the classification step (especially background pixels predicted as *staff*), the result seems quite accurate for the staff-line removal process. Note that, if that is the final objective, it is irrelevant if the classifier confuses background and staff pixels.

These preliminary results are quite encouraging, however the biggest value of this approach is that it is very easy to adapt since the only parameters to change are the training data as well as the window size for the region of interest. Therefore, it can be effortlessly extended to other types of musical documents such as printed scores or handwritten manuscripts, in black and white or colored images, as long as appropriate training examples are available.

4. CONCLUSIONS

This work studies the staff-line detection and removal of music scores following a classification approach at pixel level. Our idea is to extend this strategy to deal with greyscale images, in order to avoid errors produced by an imperfect binarization. Our preliminary results show that this approach is an interesting option to solve the task from greyscale images directly, with additional advantages related to the machine learning paradigm.

5. ACKNOWLEDGMENTS

This work has been supported by the Social Sciences and Humanities Research Council of Canada and the Spanish Ministerio de Educación, Cultura y Deporte through a FPU Fellowship (Ref. AP2012–0939).

6. REFERENCES

[1] Jorge Calvo-Zaragoza, Isabel Barbancho, Lorenzo J. Tardón, and Ana M. Barbancho. Avoiding staff removal stage in optical music recognition: Application to scores written in white mensural notation. *Pattern Anal. Appl.*, 18(4):933–43, 2015.

[2] Jorge Calvo-Zaragoza, Luisa Micó, and Jose Oncina. Music staff removal with supervised pixel classification. *International Journal on Document Analysis and Recognition (IJDAR)*, pages 1–9, 2016.

[3] Jaime S. Cardoso, Artur Capela, Ana Rebelo, Carlos Guedes, and Joaquim Pinto da Costa. Staff detection with stable paths. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(6):1134–9, 2009.

[4] Christoph Dalitz, Michael Droettboom, Bastian Pranzas, and Ichiro Fujinaga. A comparative study of staff removal algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(5):753–66, 2008.

[5] Alicia Fornés, Anjan Dutta, Albert Gordo, and Josep Lladós. CVC-MUSCIMA: A ground truth of handwritten music score images for writer identification and staff removal. *International Journal on Document Analysis and Recognition (IJDAR)*, 15(3):243–51, 2012.

[6] Thierry Géraud. A morphological method for music score staff removal. In *Proceedings of the 21st International Conference on Image Processing (ICIP)*, pages 2599–603, Paris, France, 2014.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–105, 2012.

[8] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–44, 2015.

[9] Laurent Pugin. Optical music recognition of early typographic prints using hidden markov models. In *ISMIR 2006, 7th International Conference on Music Information Retrieval*, pages 53–6, 2006.

[10] Ana Rebelo, Ichiro Fujinaga, Filipe Paszkiewicz, André R. S. Marçal, Carlos Guedes, and Jaime S. Cardoso. Optical music recognition: State-of-the-art and open issues. *IJMIR*, 1(3):173–90, 2012.