A CONNECTED PATH APPROACH FOR STAFF DETECTION ON A MUSIC SCORE

Jaime S. Cardoso[*], *Artur Capela*[†], *Ana Rebelo*[‡], *Carlos Guedes*[§]

ABSTRACT

The preservation of many music works produced in the past entails their digitalization and consequent accessibility in an easy-to-manage digital format. Carrying this task manually is very time consuming and error prone. While optical music recognition systems usually perform well on printed scores, the processing of handwritten musical scores by computers remain far from ideal. One of the fundamental stages to carry out this task is the staff line detection. In this paper a new method for the automatic detection of music staff lines based on a connected path approach is presented. Lines affected by curvature, discontinuities, and inclination are robustly detected. Experimental results show that the proposed technique consistently outperforms well-established algorithms.

***Index Terms*—** Music, optical character recognition, document image processing, image analysis

1. INTRODUCTION

The Universal Declaration on Cultural Diversity adopted by the General Conference of UNESCO on 2001 asserts that cultural diversity is as necessary for humankind as biodiversity is for nature, and that policies to promote and protect cultural diversity thus are an integral part of sustainable development. Being music a pivotal part of our cultural heritage, its preservation, in all of its forms, must be pursued. Frequently, the preservation of many music works entails their digitalization and consequent accessibility in a format that encourages browsing, analysis and retrieval. In fact, many music works produced during the last centuries still exist only as original manuscripts or as photocopies. The digitalization of these works is therefore a highly desirable goal. Unfortunately, the ambitious goal of providing generalized access to handwritten scores that were never published has been severely hampered by the actual state-of-the-art of handwritten music recognition. The manual process required to recognize handwritten musical symbols in scores and to put them in relationship with the spine structure of the score is very time consuming. This justifies the research around the definition of reliable optical music recognition (OMR) algorithms.

Staff line detection is one of the fundamental stages of the OMR process, with subsequent processes relying heavily on its performance. The reasons for detecting and removing the staff lines lie on the need to isolate the musical symbols for a more efficient and correct detection of each symbol presented on the score.

The detection of staves is complicated due to a variety of reasons. The handwritten staff lines are rarely straight and horizontal, and are not parallel to each other. For example, some staves may be tilted one way or another on the same page or they may be curved. These scores tend to be rather irregular and determined by a person's own writing style. Moreover, if we consider that most of these works are old, the quality of the paper in which it is written might have degraded throughout the years, making it a lot harder to correctly identify its contents.

In this paper a method for the automatic detection of staff lines based on a connected path approach is presented. The proposed paradigm uses the image as a graph, where the staff lines result as the connected path between the two margins of the image. Our previous work [1] was a first effort to explore this concept. We now present a complete, principled solution, with a strong experimental validation of the proposed approach. This introduction is concluded with a brief review of the work done in this area. In section 2 the proposed algorithm is described. In section 3, the proposed algorithm is evaluated experimentally using a well-known dataset of music scores. Finally, conclusions are drawn and future work is outlined in section 4.

1.1. Related Works

The problem of staff line detection is often considered simultaneously with the goal of their removal, although exceptions exist [2, 3]. The simplest approach consists on finding local maxima in the horizontal projection of the black pixels of the image [4]. These local maxima represent line positions, assuming straight and horizontal lines. Several horizontal projections can be made with different image rotation angles, keeping the image in which the local maxima are bigger. This eliminates the assumption that the lines are always horizontal. An alternative strategy for identifying staff lines is to use vertical scan lines [5]. More recent works present a more or less sophisticated use of a combination of projection techniques to improve on the basic approach [6].

Fujinaga [7] incorporates a set of image processing techniques in the algorithm, including run-length coding (RLC), connected-component analysis, and projections. After applying the RLC to find the thickness of staff lines and the space between the staff lines, any vertical black runs that are more than twice the staff line height are removed from the original. Then, the connected components are scanned in order to eliminate any component whose width is less than the staff space height. After a global deskewing, taller components, such as slurs and dynamic wedges are removed.

Other techniques for finding staff lines include the application of mathematical morphology algorithms [8], rule-based classification of thin horizontal line segments [9], and line tracing [10, 11]. The methods proposed in [2, 3] operate on a set of 'staff segments', with methods for linking two segments horizontally and vertically and merging two segments with overlapping position into one.

In spite of the variety of methods available, they all suffer from some limitations. In particular, lines with some curvature or discontinuities are inadequately resolved. The dash detector [12] is one of few works that try to handle discontinuities. The dash detector is an

[*]INESC Porto, Faculdade de Engenharia, Universidade do Porto, Portugal, email: jaime.cardoso@inescporto.pt

[†]INESC Porto, Faculdade de Engenharia, Universidade do Porto, Portugal, email: gcapela@inescporto.pt

[‡]INESC Porto, Faculdade de Ciências, Universidade do Porto, Portugal, email: arebelo@inescporto.pt

[§]INESC Porto, Escola Superior de Música e Artes do Espectáculo, Portugal, email: carlosguedes@mac.com

algorithm that searches the image, pixel by pixel, finding black pixel regions that it classifies as stains or dashes. Then, it tries to unite the dashes to construct lines.

2. A CONNECTED PATH APPROACH FOR STAFF LINE DETECTION

A staff line can be considered as a connected path from the left side of the music score to the right side. As staff lines are almost the only extensive black objects on the music score, the path we are looking for is the shortest path between the two margins **if** paths (almost) entirely through black pixels are favoured. More formally, let s and t be two pixels of the image and $\mathcal{P}_{s,t}$ a path over the image connecting them. We are interested in finding the path $\mathcal{P}$ that optimizes some predefined distance $d(s,t)$. This criterion should embed the need to favour black pixels.

In the work to be detailed, the image grid is considered as a graph with pixels as nodes and edges connecting neighbouring pixels. The weight of each arc, $w(p,q)$, is a function of pixels values and pixels relative positions. A path from vertex (pixel) v_1 to vertex (pixel) v_n is a list of unique vertices $v_1, v_2, \ldots, v_n$, with v_{i-1} and v_i corresponding to neighbour pixels. The *path cost* is the sum of each arc weight in the path $\sum_{i=2}^{n} w(v_{i-1}, v_i)$.

As mentioned before, a staff line corresponds to a path from (almost) the left margin of the image to (almost) the right side of the image, (almost) always through black pixels. If the weight assigned to an edge captures the intensity of the path of the adjacent pixels, finding the best path between a point s on the left margin and a point t on the right margin translates into computing the minimum accumulated weight along all possible connected curves connecting s and t:

$$d(s,t) = \min_{\mathcal{P}_{s,t}} \sum w(p,q). \tag{1}$$

Staff lines are best modelled as paths between two regions Ω_1 and Ω_2, the left and right margins of the score. The shortest path between two regions Ω_1 and Ω_2 is defined as a path $\mathcal{P}_{s,t}$, with $s \in \Omega_1$ and $t \in \Omega_2$ and cost equal to

$$d(\Omega_1, \Omega_2) = \min_{s\in\Omega_1, t\in\Omega_2} d(s,t). \tag{2}$$

One may assume that staff lines do not zigzag back and forth, left and right. Therefore, one may restrict the search among connected paths containing one, and only one, pixel in each column of the image[1]. Formally, let I be a $n \times m$ image and define an admissible staff to be

$$\mathbf{s} = \{(j, y(j))\}_{j=1}^{n} \text{ , s.t. } \forall j \ |y(j) - y(j-1)| \leq 1,$$

where y is a mapping $y : [1, \cdots, n] \rightarrow [1, \cdots, m]$. That is, a staff line is an 8-connected path of pixels in the image from left to right, containing one, and only one, pixel in each column of the image.

Given the weight function $w(p,q)$, one can define the cost of a staff as $C(\mathbf{s}) = \sum_{i=2}^{n} w(v_{i-1}, v_i)$. The optimal staff line that minimizes this cost can be found using dynamic programming. The first step is to traverse the image from the second column to the last column and compute the cumulative minimum cost C for all possible connected staff lines for each entry (i,j):

$$C(i,j) = \min \begin{cases} C(i-1, j-1) & + \quad w(p_{i-1,j-1}; p_{i,j}) \\ C(i-1, j) & + \quad w(p_{i-1,j}; p_{i,j}) \\ C(i-1, j+1) & + \quad w(p_{i-1,j+1}; p_{i,j}) \end{cases}$$

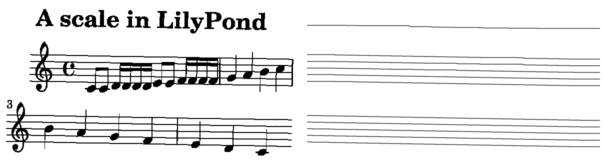
[1]This assumption imposes a maximum detectable 45 rotation degree.

At the end of this process, the minimum value of the last column in C will indicate the end of the minimal connected staff. Hence, in the second step one backtrack from this minimum entry on C to find the path of the optimal staff.

2.1. Algorithm outline

Assume one wants to find all staff lines present in a score. This can be approached by successively finding and removing the shortest path from the left to the right margin of the score. The removal operation is required to ensure that a staff is not detected twice[2].

Consider the music score presented in Figure 1(a). In Figure 1(b) the first 11 shortest paths are traced. This example shows that music symbols placed on top of staff lines do not interfere with the detection of the staff lines. Moreover, the example also makes clear that slight skewed scores do not pose any problem to the proposed approach.



(a) Skewed staff lines with music symbols. (b) The first shortest paths between left and right margins.

Fig. 1. An exemplificative example.

Our first naive effort to apply the shortest path foundation to staff line detection in [1] did it by computing the shortest path between *two pixels at the same height* on the left and right margin. That approach was not robust enough to tilted scores, leading the detected paths to jump between consecutive staff lines.

Nonetheless, two main issues are still visible with the current methodology and need to be conveniently addressed. A criterion is needed to stop the iterative detection of the shortest paths (staff lines) and the initial and the final parts of a path should be trimmed.

2.2. Proposed Algorithm

To detect the staff lines, the proposed overall algorithm starts by estimating the staff space height, `staffspaceheight`, and staff line height, `stafflineheight`. These lengths are used as reference lengths on subsequent operations. Robust estimators are already in common use: the technique starts by computing the vertical run-lengths representation of the image. If a bit-mapped page of music is converted to vertical run-lengths coding, the most common black-runs represents the staff line height and the most common white-runs represents the staff space height [7].

After estimating the reference lengths, the proposed approach applies the main step of the framework, by successively finding the shortest path between the left and right margin, adding the path found to the list of staff lines and removing it from the image. The weight $w(p,q)$ was experimentally set to $w(p,q) = 2 + (I_p + I_q)/255$, with $I_p, I_q \in \{0, 255\}$, for pixels in a 4-neighbourhood or $\sqrt{2}$ times that value for 8-neighbours. The removal operation sets to white the pixels on a vertical strip of height= $2\times$`stafflineheight`, centred on the detected staff

[2]We implemented the removal operation by setting to white the pixels on the detected staff; image resizing could be a valid alternative [13].

line. To stop the iterative staff line search, a sequence of (arguably) sensible rules is used to validate the last found path; if it does not pass the checking, the iterative search is broken. Two validation rules were applied, both assessing features with respect to the first detected staff line (assumed to be the most perfect one). If the last path does not have a percentage of black pixels above a threshold, the search is broken (a threshold of `blackperc` = 0.8 of the percentage of black pixels in the first staff line was used in the experiments). Likewise, if the shape of the last detected path differs too much from the shape of the first detected path (measured as the average y-distance between both paths, after removing the means), the iteration is broken. A threshold $\texttt{shapediff} = 4 \times \texttt{staffspaceheight}$ was experimentally selected.

After the main search step, detected staff lines are post-processed. Although true staff lines never intersect, the above algorithm may occasionally create intersecting lines. That may be due to a local low quality of a line, leading the shortest path to jump between consecutive lines; the next iteration will then follow the remaining segments, intersecting with the previous detected line. To preclude such final, undesired state, lines are post-processed to remove intersections. That is easily and efficiently accomplished by, for each image column, sorting on y the pixels of the detected lines and assigning the i-pixel to the i-line. After this simple process, lines may touch but they do not intersect. Each retained line is then trimmed at the beginning and at the ending. As visible in the previous example (refer to Figure 1(b)), before meeting with a staff line, a path travels through a sequence of white pixels. Likewise, after the end of the staff line, the path goes again through a sequence of white pixels until it meets the right margin of the image. In order to ignore all of these white pixels, the initial pixels of the path are discarded until a run of at least `blackrun` black pixels are found in the path. In the same way, all pixels of the path after the last occurrence of a run of at least `blackrun` black pixels are discarded. A threshold of $\texttt{blackrun} = 2 \times \texttt{staffspaceheight}$ was used on the experiments. Finally, lines are smoothed with a standard average low-pass filter. A window size of $2 \times \texttt{staffspaceheight}$ was selected on the experiments.

3. RESULTS

This section provides experimental results obtained on a set of scores. Although the assessment of a new staff detection algorithm may be done by visually inspecting the output on a set of scores—as adopted on [1]—, here we support the comparison on quantitative measures. The test set adopted for the qualitative evaluation of the proposed method is the one presented in [14]. The test set consists of ideal scores to which known deformations can be applied. The distortions range from rotation and curvature to typeset emulation and staff line thickness variation—see [14, 15] for more details. In total, 630 images were generated from the originally perfect scores. To conveniently measure the performance of a staff line detection algorithm, we considered two different error metrics: the number of false positive staff lines and missed to detect staff lines.

To evaluate these metrics, we start by computing the average Euclidian distance between each reference staff line and each actually detected staff line; then we solve the matching problem on the resulting bipartite graph by minimizing the assignment cost (= distance). Only pairs with average error-distance bellow the staff line height were considered correctly matched (the other pairs were assumed to originate from a false positive staff line being matched to an undetected true staff line and were therefore unmatched). Now the two metrics result as the number of unmatched detected staff lines (false positive) and unmatched reference staff lines (missed to detect).

The proposed algorithm was compared with the three methods considered in [14] for staff line detection[3]. As Dalitz's algorithm performs significantly better than the two other algorithms evaluated in [14], we have only included Dalitz results in subsequent figures. It is important to state that the comparison reports only to staff line detection algorithms, not to the removal phase. That explains the need to introduce the aforementioned metrics, while not adopting the metrics introduced in [14] for assessing staff line removal.

The effects of the different deformations over the respective parameter ranges are shown in Figure 2. With respect to the distortions considered, our connected path based approach is the most robust and clearly outperforms the Dalitz algorithm. In fact, the performance of our approach is almost independent of intensity of the deformation, for the range of values considered. This performance gain is even more noteworthy as the Dalitz algorithm is receiving as input the correct number of lines per staff, while the proposed approach does not rely on that information. Had not this been the case, the differential between both would have been much larger.

In summary, these experiments show the strengths of the presented algorithm. Despite being based on a simple and intuitive underlying principle, the performance of the proposed algorithm is quite competitive. Moreover, the results are prone to be improved even further by elaborating the stopping criterion of the iterative search or the post-processing rules, while leaving intact the main principle of the method.

4. CONCLUSION

The first challenge faced by an OMR system is staff line detection. This first task dictates the possibility of success for the recognition of the music score. In the case of handwritten music scores, the existing solutions are far from presenting satisfactory results.

In this paper, a new algorithm for the automatic detection of staff lines in music scores was proposed. The connected path approach for staff line detection algorithm is adaptable to a wide range of image conditions, thanks to its intrinsic robustness to skewed images, discontinuities, and curved staff lines. The handwritten staff lines are rarely straight and horizontal, and are not parallel to each other. Some staves may be tilted one way or another on the same page or they may be curved. While current approaches apply a chain of heuristics to correct these undesired imperfections, the connected path algorithm is naturally robust to these challenging conditions. The proposed approach is robust to broken staff lines (due to low-quality digitalization or low-quality originals) or staff lines as thin as one pixel. Missing pieces are automatically 'completed' by the algorithm.

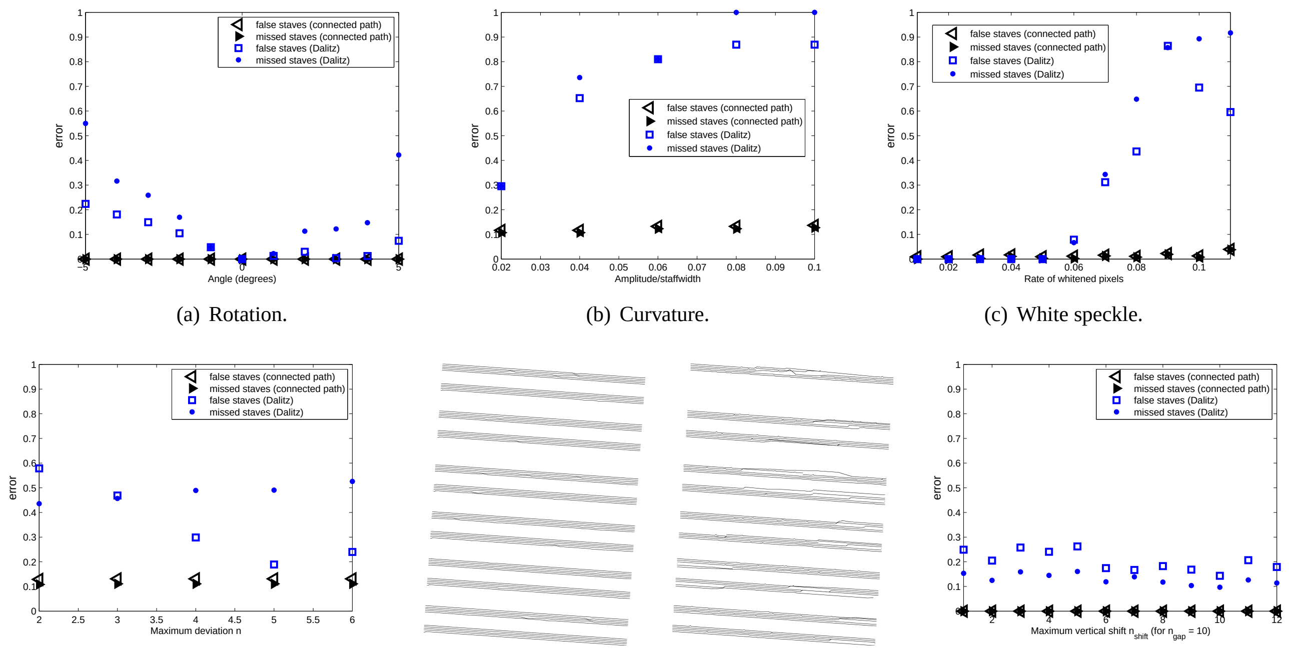
Acknowledgments

This work was partially funded by Fundação para a Ciência e a Tecnologia (FCT) - Portugal through project PTDC/EIA/71225/2006.

5. REFERENCES

[1] Ana Rebelo, Artur Capela, Joaquim F. Pinto da Costa, Carlos Guedes, Eurico Carrapatoso, and Jaime S. Cardoso, "A shortest path approach for staff line detection," in *Third International Conference on Automated Production of Cross Media Content for Multi-channel Distribution (AXMEDIS 2007)*, 2007.

[3] The source code is available upon request to the authors.



(a) Rotation. (b) Curvature. (c) White speckle.

(d) Line y-variation. (e) Shortest path (left) and Dalitz (right) results for rotation (angle=-4°). (f) Typeset emulation.

Fig. 2. Effect of different deformations on the overall error rates. See [14] for parameter details.

[2] Hidetoshi Miyao and Masayuki Okamoto, "Stave extraction for printed music scores using dp matching," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 8, pp. 208–215, 2007.

[3] Mariusz Szwoch, "A robust detector for distorted music staves," in *Computer Analysis of Images and Patterns*, pp. 701–708. Springer-Verlag, Heidelberg, 2005.

[4] Dorothea Blostein and Henry S. Baird, "A critical survey of music image analysis," in *Structured Document Image Analysis*, Baird, Bunke, and Yamamoto (Eds.), Eds., pp. 405–434. Springer-Verlag, Heidelberg, 1992.

[5] N. P. Carter, *Automatic Recognition of Printed Music in the Context of Electronic Publishing*, Ph.D. thesis, Departments of Physics and Music, University of Surrey, 1989.

[6] D. Bainbridge, *Extensible Optical Music Recognition*, Ph.D. thesis, Department of Computer Science, University of Canterbury, Christchurch, NZ, 1997.

[7] Ichiro Fujinaga, "Staff detection and removal," in *Visual Perception of Music Notation: On-Line and Off-Line Recognition*, Susan George, Ed., pp. 1–39. Idea Group Inc., 2004.

[8] Ignacy Gawedzki, "Optical music scores recognition," Tech. Rep., 2002.

[9] J. V. Mahoney, "Automatic analysis of music score images. B.Sc thesis," 1982.

[10] D. Prerau, *Computer pattern recognition of standard engraved music notation*, Ph.D. thesis, Department of Computer Science and Engineering, MIT, 1970.

[11] J. W. Roach and J. E. Tatem, "Using domain knowledge in low-level visual processing to interpret handwritten music: an experiment," *Pattern Recognition*, vol. 21, no. 1, pp. 33–44, 1988.

[12] I. Leplumey, J. Camillerapp, and G. Lorette, "A robust detector for music staves," in *Proceedings of the International Conference on Document Analysis and Recognition*, 1993, pp. 902–905.

[13] Shai Avidan and Ariel Shamir, "Seam carving for content-aware image resizing," in *ACM Transactions on Graphics (SIGGRAPH 2007)*, 2007, vol. 26.

[14] Christoph Dalitz, Michael Droettboom, Bastian Czerwinski, and Ichiro Fujigana, "A comparative study of staff removal algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004.

[15] Christoph Dalitz, Michael Droettboom, Bastian Czerwinski, and Ichiro Fujigana, "Staff removal toolkit for gamera(2005-2007," .