

# An Efficient Staff Removal Approach from Printed Musical Documents

Anjan Dutta

Computer Vision Centre  
UAB, 08193 Bellaterra  
Barcelona, Spain  
Email: [adutta@cvc.uab.es](mailto:adutta@cvc.uab.es)

Umapada Pal

CVPR Unit, Indian Statistical Institute  
203 B. T. Road, Kolkata-108  
Kolkata, India  
[umapada@isical.ac.in](mailto:umapada@isical.ac.in)

Alicia Fornés

Computer Vision Centre  
UAB, 08193 Bellaterra  
Barcelona, Spain  
[afornes@cvc.uab.es](mailto:afornes@cvc.uab.es)

Josep Lladós

Computer Vision Centre  
UAB, 08193 Bellaterra  
Barcelona, Spain  
[josep@cvc.uab.es](mailto:josep@cvc.uab.es)

**Abstract**—Staff removal is an important preprocessing step of the Optical Music Recognition (OMR). The process aims to remove the stafflines from a musical document and retain only the musical symbols, later these symbols are used effectively to identify the music information. This paper proposes a simple but robust method to remove stafflines from printed musical scores. In the proposed methodology we have considered a staffline segment as a horizontal linkage of vertical black runs with uniform height. We have used the neighbouring properties of a staffline segment to validate it as a true segment. We have considered the dataset along with the deformations described in [8] for evaluation purpose. From experimentation we have got encouraging results.

**Keywords**—OMR; musical scores; staffline; staffline segments; staffline height; staffspace height; staff removal;

## I. INTRODUCTION

The staff or stave is a set of parallel lines appeared mainly in the western musical documents. The number of lines in each of the staffs varies depending upon the type of that document and each of the lines represents a different musical pitch. The stafflines are used for placing appropriate musical symbols according to their corresponding pitch or function. They usually have uniform thickness i.e. *staffline\_height* and they are vertically separated by uniform *staffspace\_height* (Figure 1). The position of the musical symbols in a staff is very important as they are used to determine the pitch of the music by the human reader. For Optical Music Recognition (OMR) purpose the removal of stafflines is a very important preprocessing step as the stafflines create obstruction for recognizing the symbols.

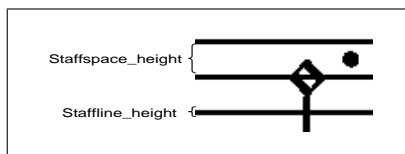


Figure 1. Portion of a staff

The main purpose of staff removal is to isolate the musical symbols for their posterior recognition. Staff removal is a difficult problem because of the overlapping of stafflines with the symbols. It demands exact segmentation of the musical symbols at the junction points, otherwise they couldn't

be recognised correctly. Moreover in real world situation the stafflines don't always appear parallel due to warping, wrinkle, paper deformation which make the problem more difficult. However staff detection can be considered as the first step of staff removal and there are several staff detection algorithms with varying success rate exist in literature. A prior approach involves to roughly detect the staffline pixels and then eliminate the pixels which belong to musical symbols using some criteria and later enhance the initial staff pixels to detect all the staff lines. Different techniques viz. linetracking [2], [4], [5], two-dimensional vector fields [1], skeletonization [8] have been used for detecting these initial staff pixels. Among the other approaches, [3], [7] have worked on a set of staff portions and tried to merge horizontally and vertically the overlapping segments. Diego Nehab [6] has used horizontal projection based technique for detecting stafflines, although he has taken care of the fact of skewness of document in his method. Cardoso et al. [9] have taken a graph based approach to detect the staffline pixels, as they have considered stafflines as the object made from black pixels connected in a stable path. For more information about literature see [8]. In the proposed methodology we have considered a staffline segment as a horizontal linkage of vertical black runs with uniform height. We have used the neighbouring properties of a staffline segment to validate it as a true segment.

This paper is organized as follows: Section II consists of the description of the proposed methodology. In Section III we have described the results of our algorithm in details. Finally in Section IV the future work and concluding statement of the paper is given.

## II. PROPOSED METHODOLOGY

Our staff removal technique consists in: A. Initial identification of staffline segments, B. Removal of false segments, C. Re-addition of wrongly removed staffline segments.

### A. Initial identification of staffline segments

We have considered a staffline segment as a horizontal linkage of vertical runs of uniform height. At the very beginning we have done a horizontal run-length smoothing of  $t_1$  pixels (here we have experimentally set  $t_1$  equal to

*staffline\_height*) to join the broken portions of a staffline that might appear in some noisy environment. Hereafter we will consider this smoothed image as the operating image.

Let  $V = \{v_1, v_2, v_3, \dots\}$  be the set of all vertical runs of black pixels in a document. Then a particular vertical black run  $v_n \in V$  may be a part of the stafflines if and only if  $|staffline\_height - \delta| \leq run\_length(v_n) \leq |staffline\_height + \delta|$ , here we have estimated the *staffline\_height* by the most frequent vertical black run in the document and we have taken  $\delta = 2$ , to consider slight fluctuation in the thickness of a staffline that may occur due to noise.

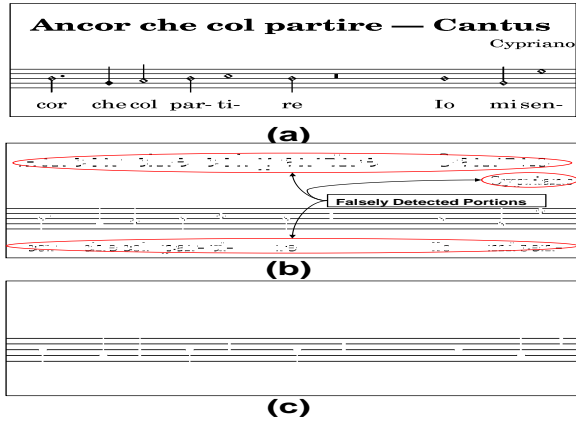


Figure 2. (a) Stafflines (original) (b) Selected vertical runs from the original stafflines (c) Detected stafflines

Thus we get a set  $V_S \subseteq V$  of vertical black runs having height near to the *staffline\_height*, the adjacent vertical runs form the probable stafflines segments. All the segments of stafflines of the image in Figure 2(a) are shown in Figure 2(b).

### B. Removal of false segments

First of all we have eliminated all the small components having width less than  $t_2$  pixels (here we have taken  $t_2$  equal to  $2 \times staffline\_height$ , and it is set by experiment), as most of the staff portions are expected to be wider than that and the smaller staffline portions appeared in place of frequent overlapping with the musical symbols. We have taken care of the problem of eliminating true smaller staffline portions in the next step. It is discussed earlier that the stafflines are always grouped in staves, the vertical gap between two true staff component is less than  $(staffline\_height + staffspace\_height)$ , where the *staffline\_height*, *staffspace\_height* can be easily estimated by the most frequent vertical run length of black and white pixels respectively. Further the horizontal gaps between the components belonging to the same staffline is also very small and they are joined by some musical symbols in the original documents (Figure 3(a), (b)). We have used

these observations to remove the false portions of stafflines.

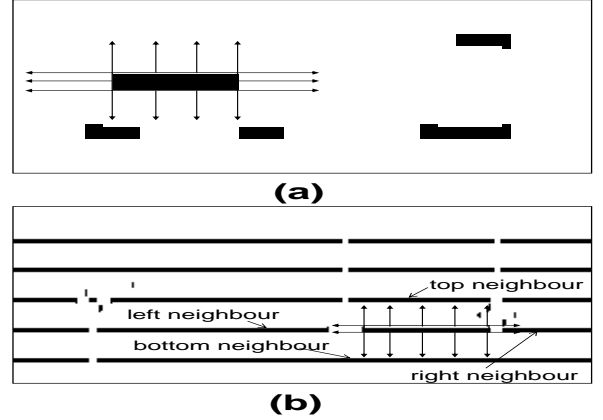


Figure 3. (a) Non-staff components (b) Staff components

Let  $C = \{c_1, c_2, c_3, \dots\}$  be the set of all components having width more than  $t_2$  pixels. Here we define the four neighbouring components  $c_i^l$  (left neighbour),  $c_i^r$  (right neighbour),  $c_i^t$  (top neighbour),  $c_i^b$  (bottom neighbour) of a component  $c_i \in C$  as follows:

- 1) A component  $c_i^l \in C$  is called the left neighbour of  $c_i$  if  $c_i^l$  is connected to the left of  $c_i$  in the original document.
- 2) A component  $c_i^r \in C$  is called the right neighbour of  $c_i$  if  $c_i^r$  is connected to the right of  $c_i$  in the original document.
- 3) A component  $c_i^t \in C$  is called the top neighbour of  $c_i$  if  $c_i^t$  is on the vertically upper side of  $c_i$  and  $vertical\_distance(c_i, c_i^t) < (staffline\_height + staffspace\_height)$ .
- 4) A component  $c_i^b \in C$  is called the bottom neighbour of  $c_i$  if  $c_i^b$  is on the vertically lower side of  $c_i$  and  $vertical\_distance(c_i, c_i^b) < (staffline\_height + staffspace\_height)$ .

Here  $vertical\_distance(s, d)$  denotes the minimum vertical run length between the component  $s$  and  $d$ . Now any arbitrary component  $c_n \in C$  with the list of neighbours  $\langle c_n^l, c_n^r, c_n^t, c_n^b \rangle$  is a part of the staffline if it has at least one of left and right neighbours and at least one of top and bottom neighbours. This eliminates most of the false stafflines portions.

### C. Re-addition of wrongly removed staffline segments

This is actually the post-processing part of our method. We have observed that due to noise and removal of small components (described in sub-section II-B) some small portions of a staffline might get eliminated. This is the reason we have added this post-processing module in our method. Here we re-add some wrongly eliminated portions of staffline which will satisfy some criteria to be a part of

valid staffline portion. Let  $S = \{s_1, s_2, s_3, \dots\}$  be the set of all valid staff component and  $E = \{e_1, e_2, e_3, \dots\}$  be the set of all components we have eliminated yet. We will add an eliminated component  $e_n \in E$  to  $S$  if there exist  $s_k \in S$  such that  $s_k$  is the left or right neighbour of  $e_n$  and also there exist  $s_l \in S$  such that  $s_l$  is the top or bottom neighbour of  $e_n$ . Thus we have got the new set  $S' = S \cup \{x\}, \forall x \in E$  that are satisfying the above condition. The complete set of stafflines detected by our method for the image in Figure 2(a) are shown in Figure 2(c).

### III. RESULTS AND DISCUSSION

#### A. Dataset

For experimental purpose we have used the dataset described in [8]. The dataset contains 32 artificially generated ideal images along with the ground truth. It covers a wide range of music notations (historic, modern and tablature). Further the authors have generated 8 types of deformations (Kanungo, Staffline thickness variation (version-1), Staffline thickness variation (version-2), Staffline y variation (version-1), Staffline y variation (version-2), Curvature, White speckles and Typeset Emulation) on the above images, which simulate the dataset with real world situation. The sample deformed images of an original ideal image (Figure 4(a)) are shown in Figure 4(b)-(i) in the above mentioned deformation order. The parameters for generating the deformed images are set as explained in [8].

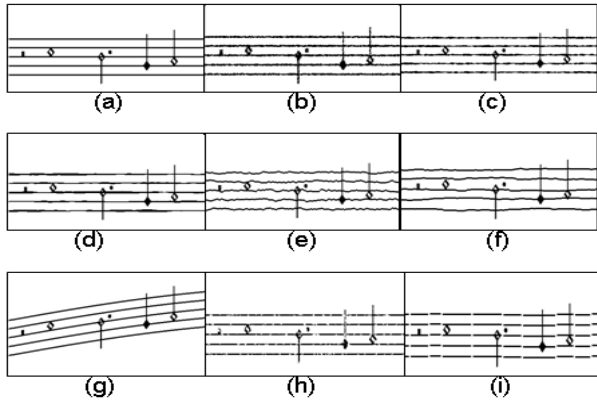


Figure 4. A sample image with its eight deformations (a) Original (b)-(i) Deformed images

#### B. Results

We have chosen the pixel based evaluation metric to get the quantitative measurement of the performance of our algorithm on different images. We have computed precision, recall and error rate [8] at the pixel level for every images and shown their average for every deformation type in Table I. To get the idea about the results of our algorithm for every type of deformed images, in Figure 5 we have

Table I  
PERFORMANCE OF THE STAFF REMOVAL ALGORITHM

Deformation Type	Precision	Recall	Error Rate
Ideal	96.26%	97.54%	0.029
Kanungo	98.25%	93.80%	0.038
Line Thickness Variation-v1	96.46%	96.30%	0.039
Line Thickness Variation-v2	96.52%	96.61%	0.043
Line y Variation-v1	96.21%	97.14%	0.028
Line y Variation-v2	96.37%	97.22%	0.026
Curvature	93.96%	94.82%	0.053
White Speckles	94.03%	98.17%	0.040
Typeset Emulation	95.74%	94.35%	0.043
<b>Average</b>	<b>95.98%</b>	<b>96.22%</b>	<b>0.037</b>

provided the results along with the respective ground truth images in the third column.

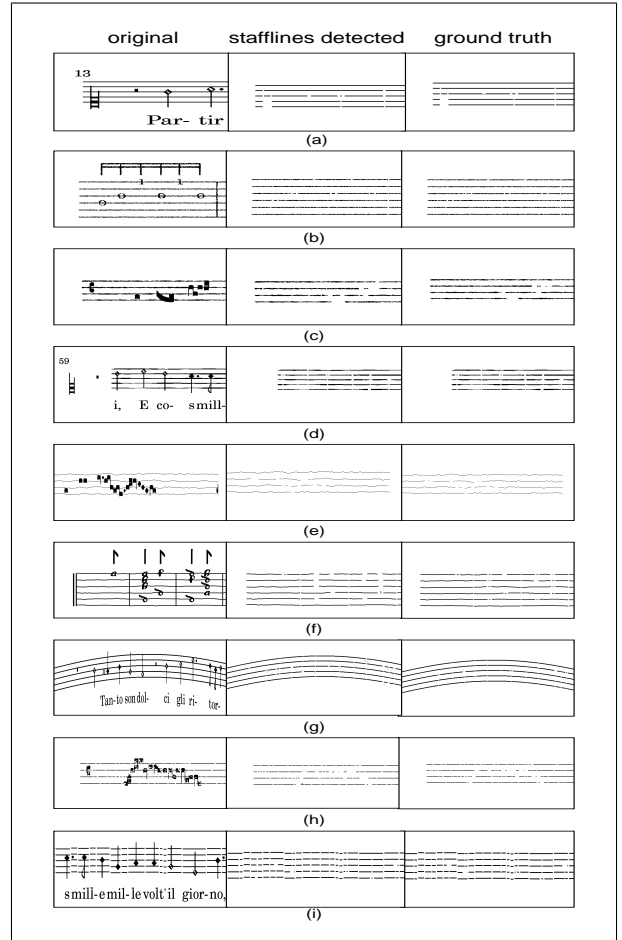


Figure 5. Staffline detection results on various kind of deformation ((a) Ideal Image (b) Kanungo (c) Staffline Thickness Variation-v1 (d) Staffline Thickness Variation-v2 (e) Staffline y Variation-v1 (f) Staffline y Variation-v2 (g) Curvature (h) White Speckles (i) Typeset Emulation)

#### C. Comparison with existing methods

We have compared our algorithm with the two existing methods<sup>1</sup> described in [7], [8] using the same dataset. We

<sup>1</sup>The scripts are made available by Dalitz et al. [8] at <http://lionel.kr.hs-niederrhein.de/~dalitz/data/projekte/stafflines/>

Table II  
COMPARISON WITH THE EXISTING METHODS

Methods	Precision	Recall	Error Rate	Exec. Time
Fujinaga [7]	88.95%	92.40%	0.099	3.75 secs
Dalitz et al [8]	<b>98.77%</b>	87.59%	0.064	7.95 secs
Our method	95.98%	<b>96.22%</b>	<b>0.037</b>	<b>1.72 secs</b>

have considered the average precision, recall, error rate and execution time of the methods as the parameters for the comparison (Table II). Though Dalitz's method [8] has given the highest precision, our method has provided the highest recall, lowest error rate and has taken lower execution time. Our method has taken less than  $\frac{1}{4}$  of the execution time of [8]. Also the recall of our method is 8.63% more than [8], moreover the error rate of our method is about half of [8] and  $\frac{1}{3}$  of [7]. This indicates the overall performance of our method is better than the existing pieces of recent work in this area.

#### D. Discussion

Our method has failed to exactly segment the symbol parts from the staffline portion whenever only some part of a staffline is occluded by a symbol. We have shown one example in Figure 6 where the two musical symbols occupy (compare our output shown in Figure 6(a) with the ground truth in Figure 6(c)) a part of third and fourth stafflines (counted from bottom). Here our method has eliminated all the pixels of the stafflines (Figure 6(b)) belonging to the same vertical runs with the music symbols. This is because of the fact that we have calculated the vertical black run to count the thickness. However, we have considered this fact as a drawback of our method which further adds some errors in the results.

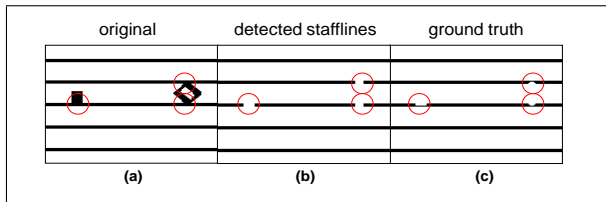


Figure 6. Example of erroneous results

#### IV. CONCLUSION AND FUTURE WORK

In our proposed method an approach to stafflines detection and removal from printed musical documents has been proposed. The method is based on the criteria of neighbouring staff components. We have tested our algorithm with various deformed images and obtained good precision and recall and very less execution time. Our method is able to correctly detect the position of the symbols. We have tested our method on some real data obtained from old musical scores for which also we have got encouraging results but since we do not have any ground truth, we can not provide any

quantitative results for this dataset, instead we have just shown sample result in Figure 7. The main problem to deal with this kind of dataset is the paper degradation, distortion which can not be handled by simple binarization technique and need further investigation which we have decided to be our future work. The encouraging results will also motivate us to work on handwritten musical scores.



Figure 7. Experiments on real dataset

#### Acknowledgement

This work has been partially supported by the Spanish project TIN2009-14633-C03-03. We are also grateful to Dr. Christoph Dalitz for providing us the script for generating various type of deformed stafflines and corresponding ground truths.

#### REFERENCES

- [1] 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, pp. 33-44, 1988.
- [2] P. Martin and C. Bellisant, "Low-Level Analysis of Music Drawing Images", *Proc. First Int'l Conf. Document Analysis and Recognition*, pp. 417-425, 1991.
- [3] N. P. Carter and R. A. Bacon, "Automatic Recognition of Printed Music", *H. S. Baird, H. Bunke, K. Yamamoto, editors, Structured Document Image Analysis*, pp. 454-465, Springer, 1992.
- [4] R. Randriamahefa, J. P. Cocquerez, F. Pépin and S. Philipp, "Printed Music Recognition", *Proc. Second Int'l Conf. Document Analysis and Recognition*, pp. 898-901, 1993.
- [5] D. Bainbridge and T. C. Bell, "Dealing with Superimposed Objects in Optical Music Recognition", *Proc. Sixth Int'l Conf. Image Processing and Its Applications*, pp. 756-760, 1997.
- [6] Diego Naheb, "Staff Line Detection by Skewed Projection", *PhD Thesis*, pp. 417-425, 2003.
- [7] I. Fujinaga, "Staff Detection and Removal", *Visual Perception of Music Notation*, S. George, ed., pp. 1-39, 2004.
- [8] C. Dalitz, M. Droettboom, B. Pranzasnd and I. Fujinaga, "A Comparative Study of Staff Removal Algorithms", *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 753-766, Vol. 30 no. 5.
- [9] J. S. Cardoso, A. Capela, A. Rebelo, C. Guedes and J. P. Costa, "Staff Detection with Stable Paths", *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 1134-1139, Vol. 31 no. 6.