Journal of Advanced Computer Science & Technology, 3 (2) (2014) 136-142
©Science Publishing Corporation
www.sciencepubco.com/index.php/JACST
doi: 10.14419/jacst.v3i2.3196
Research Paper

Effective staff line detection, restoration and removal approach for different quality of scanned handwritten music sheets

Fatemeh Alirezazadeh [1]*, Mohammad Reza Ahmadzadeh [2]

[1] *Dept. of Electrical and Computer Engineering, Najaf Abad University, Isfahan, Iran*
[2] *Dept. of Electrical and Computer Engineering, Isfahan University of Technology (IUT), Iran*
**Corresponding author E-mail: alirezazadeh_fa@yahoo.com*

Copyright © 2014 Alirezazadeh, Ahmadzadeh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Musical staff detection and removal is one of the most important preprocessing steps of an Optical Music Recognition (OMR) system. This paper proposes a new method for detecting and restoring staff lines from global information of music sheets. First of all the location of staff lines is determined. Therefore, music staff is sliced. The staff line segments are recognized at each slice and then with adequate knowledge of staff line locations, the deformed, interrupted or partly removed staff lines can be rebuilt. A new approach for staff removal algorithm is suggested in this paper fundamentally based on removing all detected staff lines. At last, the Fourier transform and Gaussian lowpass filter will help to reconstruct the separated and interrupted symbols. It has been tested on the dataset of the musical staff removal competition held under ICDAR 2012. The experimental results show the effectiveness of this method under various kinds of deformations in staff lines.

Keywords: *Fourier Transform, Gaussian Low Pass Filter, Optical Music Recognition, Run Length Coding, Staff Line Removal*.

1. Introduction

Staff detection and removal is a fundamental step for an Optical Music Recognition (OMR) system. Typically, staff lines are approximately parallel lines that cover a large proportion of a music sheet and are interconnected with most of the musical symbols. Superimposed notations make the process of musical symbols recognition more complicated. Furthermore, without adequate knowledge of the staff line location, the symbols cannot be correctly separated, processed, recognized and classified. Also, the staff detection and restoration step is essential for every OMR system in order to find out the pitch of music notes on the staves.
The process of staff removal that aims to remove staff pixels facilitates the stages of symbol recognition.
Although many approaches have been proposed for the process of staff detection and removal, it still represents a challenge due to several reasons. First of all, staff lines are not exactly straight, horizontal or continuous. Additionally, several variations of deformations can be observed for staff lines and musical documents that are illustrated in fig. 1. The problem increases where the music sheet is handwritten. Finally, symbols and notations will be affected during the staff line removal.
Some techniques have been proposed for the staff line detection and removal in the literature. One technique for the first step of staff detection is based on horizontal and vertical projection [2, 4]. This technique assumes that the lines are straight and horizontal. Genfang et al., [5] used Hough Transform to detect staff lines. Graph-theoretic framework is another strategy that treats the image as a graph, where the staff lines are connected paths from the left side to the right side [3], [6]. Thus, in this method, the need to have connected and continuous components of staff line structure is evitable. Other techniques include rule-based classification [7]; skeletonization [9] and stable path [6] are also used to detect staff positions. The concept of strong staff pixels (SSP), proposed in [3] which works with grayscale image rather than binary image, is a set of pixels with a high probability of belonging to a staff line. It is proposed to overcome the results of the binary based staff detection and removal.
Fujinaga proposed a staff removal algorithm which uses the thickness of staff lines and the space between them by applying run-length coding [2].

In spite of many algorithms which have been suggested to detect and remove the staff pixels, they have some limitations and are not effective and robust under deformations. They totally work with local information of the image. Our proposed method improves these problems and has been tested on the dataset of the recent musical staff removal competition.

2. Overview of the algorithm

In a typical piece of music sheet, staff lines cover a large proportion of the image. Finding the location of these staff lines and removing them is an essential task for the most of the OMR systems.
Like most OMR systems, two stages can be identified; staff line identification and staff line removal. The staff line identification stage involves a set of image processing techniques. First, the common method of run-length encoding (RLE) [10] is used to find the thickness of staff lines and the space between the staff lines that are illustrated in fig. 2. Second, thinning algorithm is applied to the whole of the image.
To facilitate the process of staff line detection, it is developed by using the fact that five staff lines are not always equally spaced but partly they are. Therefore, a vertical slicing step is devised to determine and restore the five staff lines. For each area of interest, the staff lines are located, the length and rotation of the lines are measured and extra pixels for amending probable discontinuities and defragmentation may be added.
The second step of staff removal technique is proposed to remove the staff lines from the image completely. Then, 2-D fast Fourier transform and Gaussian lowpass filter are served to compensate the eliminated semantic parts that were removed.

2.1. Staff line identification

Finding the staff lines is a logical precursor to all other recognition tasks to find out the pitch of music notes.
The process of staff detection represents a lot of challenges. For instance, staff lines are not properly straight, horizontal or continuous. Even, music score can be distorted by some types of deformations (see fig. 1) [1].

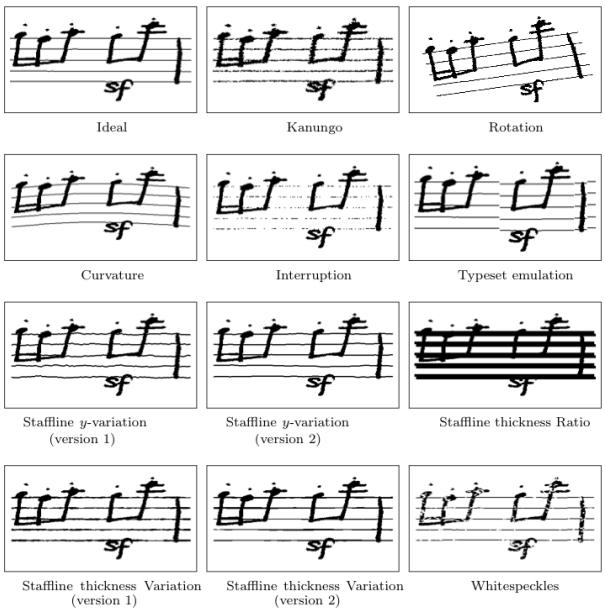


Fig. 1: Examples of Staff Deformations.

2.1.1. Estimation of reference lengths

The first step of the staff line identification relies on the estimation of staff line thickness and the distance between two consecutive lines or the staff space height (see fig. 2). Further processes are based on these estimated parameter values. The input scanned image is converted to a binary image. Then, the document is scanned column by column. The staff line thickness and the staff space height are the most repetition of vertical white-run length and black-run length, respectively.

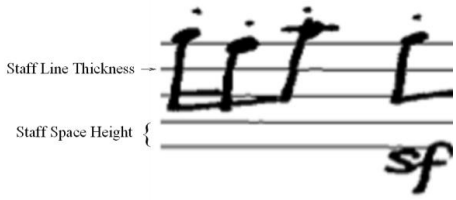


Fig. 2: Common Features of Staff Line.

2.1.2. Thinning

The thinning process provides a convenient representation of image object information. Skeletons of objects can preserve topological properties of the object and reduce storage requirements. It seems to be necessary to facilitate the process of staff lines recognition.
Fig. 3(A) shows the binary musical document. The whole image is thinned as shown in fig. 3(B). Thinning method removes pixels so that an object without holes shrinks to a minimally connected stroke, and an object with holes shrinks to a connected ring halfway between each hole and the outer boundary.

2.1.3. Slicing and finding the location of staff lines

The staff lines are not exactly parallel with each other, but for small segments they are approximately equally-spaced. When a whole staff line cannot be represented by a straight line, the staff line is broken into piecewise linear parts with a small tolerance.
To measure up expectations, the vertical slicing technique is used as shown in fig. 4. So, for simplicity, the staff is divided into several slices so that the length of each slice is chosen as follows:

$$\left|\text{Width of one staff line and its relevant space}\right| \le \left|\text{length}\right| \le \left|\text{Width of five staff lines}\right| \quad (1)$$

$$(\text{staff_space}+\text{staff_line_thickness}) \le \left|\text{length}\right| \le (4*\text{staff_space}+5*\text{staff_line_thickness}) \quad (2)$$

For each slice, line segments are recognized. The line segments are black pixels of the image that make lines. Fig. 3(C) shows the detected line segments of a piece of music document in fig. 3(A).
Some of the line segments cannot be assigned to staff lines or even some discontinuities can be observed at detected segments.
With the lines identified, the next step is to eliminate the line segments that don't belong to staff lines. Hence, the line segments that cannot satisfy the following conditions are removed from the image.

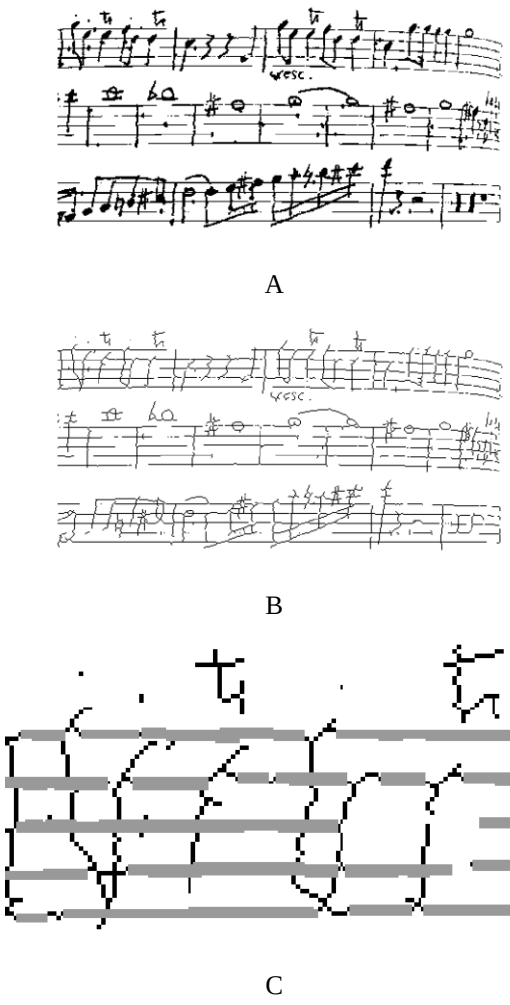
1. The rotation of staff segments is rather close to each other.
2. The distance between staff segments is close to each other (the distance is the same as considering vertical projection of the slice mentioned in [2]).
3. At each slice exist five distinct staff lines.

When staff line segments of one staff line are discontinuous, staff line restoration is necessary to link discontinuities. So, two line segments that are separated by filling gap pixels are merged into a single line.

2.2. Staff line removal

The reason of removing the staff lines lies on the necessity to isolate musical symbols and notations from staff lines in order to have more efficient and suitable detection of musical symbols present in the score.
In spite of the methods available for staff lines removal, they all have some limitations. First, most of these techniques incorporate a set of image processing methods based on the staff line thickness. A common problem of all these techniques is that they lead to remove other meaningful and vulnerable parts of the image. In particular, this problem becomes serious when the score is handwritten. Second, these techniques may be failed to extract the musical notes symbol when the staff line thickness is large [11]-see fig. 1.
In our proposed method, the whole staff lines which were identified in the previous step are removed totally from the image-see fig. 6(C). It causes the symbols to be separated from each other. Then, the next step will aid to bring back the eliminated parts of the symbols to the image.

2.3. 2-D fast Fourier transform and Gaussian low pass filter

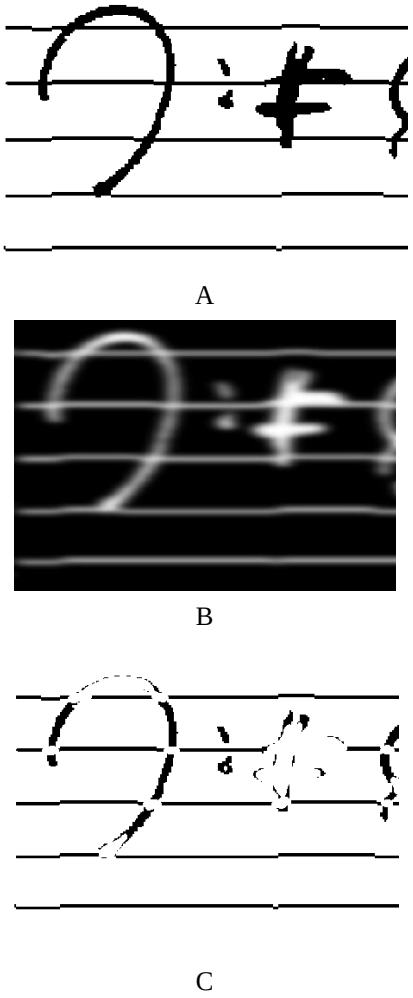


A

B

C

Fig. 3: The Procedure of Staff Line Detection. (A) The Input Musical Document. (B) The Thinned Image. (C) The Detected Staff Line in Gray



A

B

C

Fig. 5: The Procedure of Applying FFT and Filter. (A) The Input Image. (B) Applying FFT and Lowpass Filter. (C) Removing the Extracted Pixels From the Original Image.

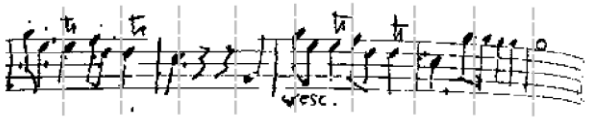


Fig. 4: Vertical Slicing Technique

The foundation for linear filtering is the convolution theorem [12],

$$f(x,y) * h(x,y) \Leftrightarrow H(u,v)F(u,v) \tag{3}$$

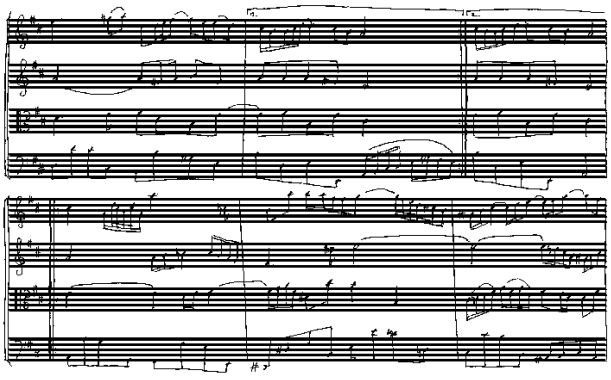
Here, the symbol "*" indicates convolution operator of two functions. Filtering in the spatial domain consists of convolving an image f(x, y) with a filter mask h(x, y). "The transfer function of a Gaussian lowpass filter is given by

$$H(u,v) = e^{-D^2(u,v)/2\sigma^2} \tag{4}$$

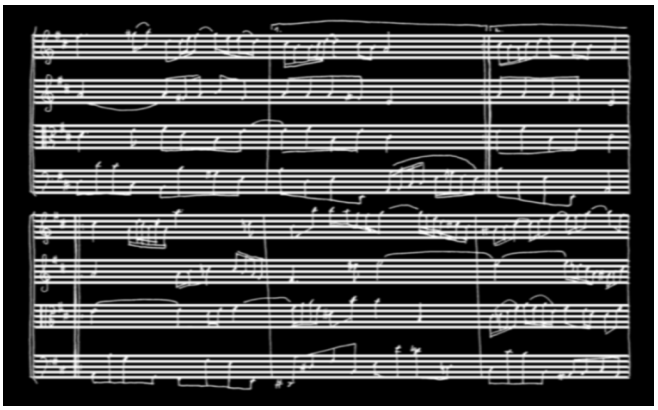
Where σ is the standard deviation".

Fig. 5(B) shows the filtered image of fig. 5(A) carried out in the frequency domain via the Fourier transform. The output image is in gray. It can be concluded from this figure that the Gaussian lowpass filter makes the image blurred and the interconnections and thicker parts of the image are brighter than other parts. So, by simple adaptive thresholding the gray image, the interconnections can be filtered out from the image. The extracted pixels are sum of thicker parts

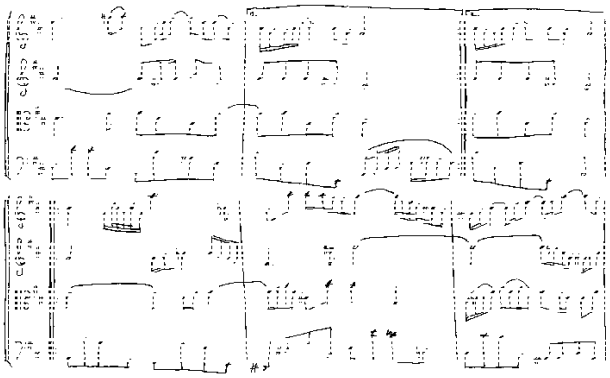
and common pixels between the staff lines and symbols. In the other words, they are exactly meaningful parts of the image that have been removed in the previous step. To clarify this operation, the extracted pixels are intentionally removed from the original image of fig. 5(A), which is shown in fig. 5(C).



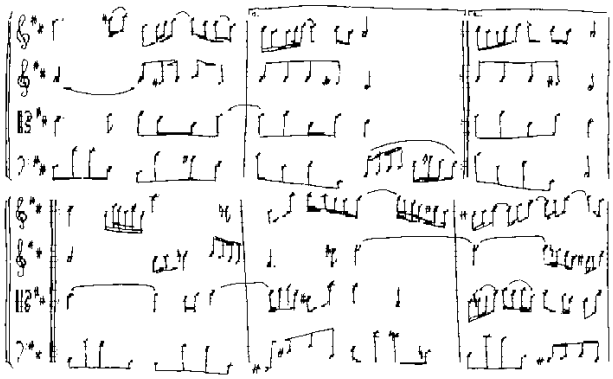
A



B



C



D

Fig. 6: The Procedure of Staff Removal. (A) The Input Image. (B) Applying FFT and Lowpass Filter. (C) The Image Obtained after Removing Staff Lines Completely. (D) The Output Image.

This method does not hurt the vulnerable parts of the musical document. The net result of lowpass filtering is image blurring (smoothing). Sensitive symbols such as whole notes, half notes, rest symbols, accidentals and key signatures are smoothed via this way. Thus, this method unlike most of the available methods of staff lines removal does not destroy or even does not lead to fragmentation of symbols. It is, therefore, advantageous to remove staff lines, attempting not to distort any intersecting symbols.
As an illustration, fig. 6(A) shows an example taken from the competition dataset with large staff line thickness. A Gaussian lowpass filter is applied to the image that is shown in fig. 6(B). Then the staff lines are removed completely as shown in fig. 6(C). The final output of staff lines removal is shown in fig. 6(D).

3. Experimental results

The proposed method was tested using on the dataset of the musical staff removal competition held under ICDAR 2012. This database consists of 1,000 handwritten music score images written by 50 different musicians. 11 distortion models were applied to the 1,000 original images: Kanungo, Rotation, Curvature, Staff line interruption, Typeset emulation, Staff line y-variation (two models), Staff line thickness ratio, Staff line thickness variation (two models) and white speckles. As a result, this dataset consists of 12,000 images consist of 11,000 distorted images and 1,000 original images. The first 6000 of the images were used as training data and the other 6000 images were used as testing data.
The performance of the algorithm was evaluated based on pixel based metric. The error rate as described in [9] was computed as:

$$E.R. = 100.\frac{\#\text{missclassified(sp)} + \#\text{missclassified(non - sp)}}{\#\text{all(sp)} + \#\text{(non - sp)}} \quad (5)$$

Where # means “number of” and sp means “staff pixels”.
Several types of methods suggested in music staff removal competition are explained in [1]. It is important to note that some distortions such as interruption and thickness made the detection process very difficult in most of these methods and most images are rejected because of the lack of identification. The one which gets better results in most cases and also without rejecting any image is ISI01-HA and its total algorithm is mentioned in [1]. The comparison between our proposed method and ISI01-HA method is shown in table 1.

Table 1: Error Rate of Staff Removal in % for Each One of 12 Distortions.

Distortion	ISI01-HA	Proposed method
Ideal	1.50	1.30
Curvature	1.66	2.11
Interruption	0.91	0.82
Kanungo	2.84	1.84
Rotation	1.76	1.53
Staff line thickness v1	2.17	1.76
Staff line thickness v2	2.15	1.79
Staff line y-variation v1	1.89	2.07
Staff line y-variation v2	1.83	2.04
Thickness Ratio	2.86	1.10
Typeset emulation	1.60	1.55
White speckles	1.48	1.20
Overall Error Rate	1.89	1.59

As can be observed from this table, the worst result for ISI01-HA method is when the thickness ratio increases. By increasing the thickness of staff lines in an arbitrary image, the symbols fade away slightly. The conventional methods of staff removal try to remove the staff lines according to their thickness. A slight difference in thickness may cause to eliminate the other symbols parts. Thanks to the Gaussian lowpass filter makes the interconnections and thicker parts of the image to be brighter than other parts. Therefore, the proposed method works well encountering with thickness ratio.
In the table can be seen that the proposed method is better in interruption, kanungo and white speckle distortions. In the proposed algorithm, a line segment is black pixels of the image making a line. It can be a combination of several dots that lie in a line. In spite of the fact that misidentification happens during staff detection algorithm, the line segments that cannot satisfy the conditions mentioned in sec. 2.1.3 are removed from the image. Hence, its operation is robust enough to be compared with the other conventional algorithms.
The drawback of this algorithm is when the severe curvature distortion [1] takes place-see fig. 7. It causes the process of staff detection more complicated. On one hand, the algorithm is enforced to deal with shorter line segments; On the

other hand, the misidentification of staff lines increases. Hence, there should be a trade-off between them. But, this trade- off cannot satisfy the results.



Fig. 7: Severe Distortion.

Totally, the evaluation shows that the proposed method is more powerful for interrupted or thickened staff lines and also it works well for white speckles. The overall Error Rate of 1.59 shows the efficiency of this method compare to the other available methods.

The paper presents a novel technique for detecting musical symbols, based on recognizing the staff lines and completely removing them, in order to isolate the desired symbols, and further using the fast Fourier transform and Gaussian lowpass filter for restoring the separated and interrupted musical notes. The contribution stands in the reconstruction stage that manages to detect the meaningful parts of the image that are removed in the first step of the algorithm and thus obtaining the unaltered musical symbols. The techniques and algorithms implied in the different stages of the process are clearly explained. The robustness of the method is proved in the experiments run on the dataset of the musical staff removal competition held under ICDAR 2012. The extraction of the musical symbols is accurate even when various deformations of the staff lines occur. The authors show that their detector brings a plus to the existing systems for staff line detection and removal, as they manage to decrease the overall error rate with 0.3% compared to the state of the art, which is a promising result.

4. Conclusion

Musical staff detection and removal is a technique to detect staff line positions and remove them from the music score. It is an important step of OMR tasks and is usually considered as the first step of an OMR system. The wider variability of the symbols in size and shape, found on handwritten music score, makes the operation of symbol extraction more complex. This novel method is proposed to detect and restore staff lines from global information of music sheet. At first, the location of staff lines is determined and the music staff is sliced. At each slice, the staff line segments are recognized and then interrupted staff lines can be rebuilt by global information of five staff lines. This new approach is fundamentally based on removing all detected staff lines properly. Then, the fast Fourier transform and Gaussian lowpass filter will aid to reconstruct the separated and interrupted symbols. It has been tested on the database of the first music staff removal competition held in ICDAR. This database consists of 1,000 handwritten music score images written by 50 different musicians. The experimental results show the robustness of this method for various kinds of deformations in staff lines.

References

[1] A. Fornes, A. Dutta, A. Gordo and J. Llados, "The ICDAR 2012 music scores competition: Staff removal and writer identification," *ICDAR, 2013*.

[2] I. Fujinaga, "Staff detection and removal," *In: George S (ed) Visual perception of music notation*, on-line and off-line recognition, 2004. http://dx.doi.org/10.4018/978-1-59140-298-5.ch001.

[3] A. Rebelo and J. S. Cardoso, "Staff Line Detection and Removal in the Grayscale Domain," *12th International Conference on Document Analysis and Recognition ICDAR*, 2013, pp. 57-61.

[4] D. Bainbridge and T. Bell, "A music notation construction engine for optical music recognition," *Softw Pract Exp* vol. 33, no. 2, pp. 173–200, 2003. http://dx.doi.org/10.1002/spe.502.

[5] C. Genfang, Z. Liyin, Z. Wenjun and W. Qiuqiu, "Detecting the staff-lines of musical score with hough transform and mathematical morphology," *International Conference on Multimedia Technology (ICMT)*, 2010, pp. 1-4.

[6] J.S. Cardoso, A. Capela, A. Rebelo, C. Guedes and J.P. da Costa, "Staff detection with stable paths," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, pp. 1134-1139, 2009. http://dx.doi.org/10.1109/TPAMI.2009.34.

[7] A. Rebelo, F. Paszkiewics and C. Guedes, "A method for music symbols extraction based on musical rules," *Bridges: Mathematics, Music, Art, Architecture, Culture*, pp. 81-87, 2011.

[8] A. Fornes, A. Dutta, A. Gordo, and J. Llados, "CVC-MUSCIMA: a ground truth of handwritten music score images for writer identification and staff removal," *International Journal on Document Analysis and Recognition*, vol. 15, 2012, pp. 243–251. http://dx.doi.org/10.1007/s10032-011-0168-2.

[9] C. Dalitz, M. Droettboom, B. Pranzas, and I. Fujinaga, "A comparative study of staff removal algorithms," *IEEE Transactions on PAMI*, vol. 30, no. 5, pp. 753–766, May 2008. http://dx.doi.org/10.1109/TPAMI.2007.70749.

[10] T. Pinto, A. Rebelo, G. Giraldi and J.S. Cardoso, "Music score binarization based on domain knowledge," *pattern recognition and image analysis.* Lecture notes in computer science, vol. 6669. Springer, Heidelberg, pp. 700–708, 2011. http://dx.doi.org/10.1007/978-3-642-21257-4_87.

[11] B. Su, S. Lu, U. Pal and C. L. Tan, "An effective staff detection and removal technique for musical documents," *10th International Workshop on Document Analysis Systems*, Queensland, Australia, March 2012, pp. 160-164.

[12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2008.