An Adaptive Staff Line Removal in Music Score Images

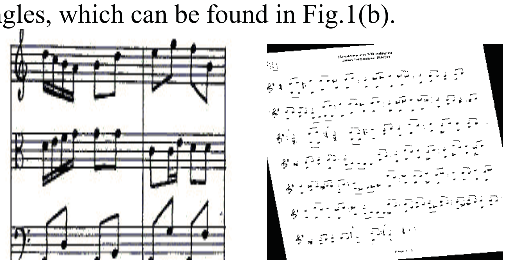
Jiali Cui[1], Huan He[2], Yiding Wang[1]
Department of Electronics
North China University of Technology,
5 Jinyuanzhuang Rd, Beijing 100144, P.R. China
[1]{jialicui, wangyd}@ncut.edu.cn, [2] hehuan1002@gmail.com

Abstract—**This paper proposes an adaptive staff line removal method for music score images. The staff line removal is a key procedure in Optical Music Recognition (OMR) before musical symbol segmentation and classification. The proposed method consists of four main procedures, i.e., ROI(Region of Interest) selection, candidate point extraction, rotation correction and staff line removal. It selects a ROI, extracts candidate points in the selected ROI and estimates rotation angle with these candidate points. After rotation correction, the staff line width is estimated using horizontal projection and the staff lines are removed correspondingly. Extensive experiments show that the proposed method is robust for both printed and scanned music score images.**

Keywords- Music score images, Staff line removal, ROI, Rotation correction, Entropy

I. INTRODUCTION

Music score images are usually the digital form of musical notations. Two examples of music score images are shown in Fig.1, in which the groups of five parallel horizontal lines are called staff lines. Most music score images have no rotation or obvious rotation, which means that the staff lines are completely horizontal, as can be seen in Fig.1(a). However, the scanned music score images usually have slight rotation angles, which can be found in Fig.1(b).



(a)Printed form (b) Scanned form

Figure 1. Two music score images

Optical music recognition (OMR), which aims to recognize musical symbols automatically via computers, is an active topic in image processing and pattern recognition. OMR has many potential applications, such as human-computer-interface and intelligent toys, so it attracts more and more interests from researchers. . OMR has been the research focus since 1970s and some work has been done [1-11, 13, 14].

The target of staff line removal is to remove the staff lines. It is quite important, because the symbol features, such as size and shape, can be extracted after the staff lines are removed. Therefore, it is a key procedure before musical symbol segmentation and recognition. However, staff line removal is still a challenging topic in OMR [7], because the staff line width is variable due to different resolutions and the staff lines are not even horizontal.

Some researchers have proposed methods for staff line removal, for example, references [8,11]. However, these methods depend on the knowledge of the music score images. To overcome the shortcoming, this paper proposes an adaptive staff line removal method. It selects a region of interest (ROI), estimates the rotation angle and does rotation correction so that the staff lines are horizontal. After that, line width is estimated automatically and the staff lines are removed with adaptive parameters.

The remainder of the paper is organized as follows. Section 2 details the proposed method. Experimental results are reported in section 3. Section 4 concludes the whole paper. The references are given at the end of the paper.

II. ALGORITHM DESCRIPTION

A. *Flowchart of the Proposed Method*

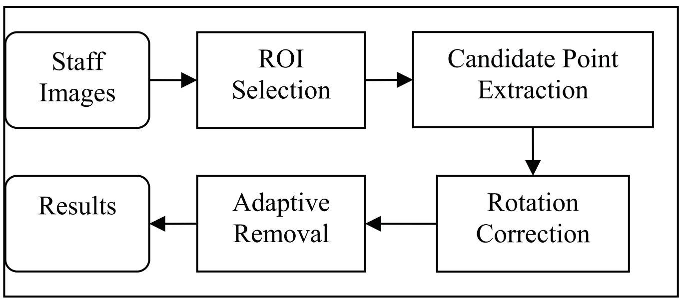


Figure 2. Flowchart of the proposed method

The flowchart of the proposed method is shown in Fig.2. It takes the raw music score images as input and the results are the processed music score images without staff lines. The whole processing consists of four main procedures, i.e., ROI selection, candidate point extraction, rotation correction and adaptive removing.

B. *ROI Selection*

Given a music score image with a resolution of M*N pixels and the maximal possible rotation angle θ_{Max}, we want to select a ROI to locate limited foreground area. The reason of ROI selection is to reduce the number of candidate points in subsequent section. A ROI is a small area which contains at least one group of staff lines and lies in the middle of given music score image. To compute the parameters to locate the ROI, We propose the following model as can be illustrated in Fig.3.

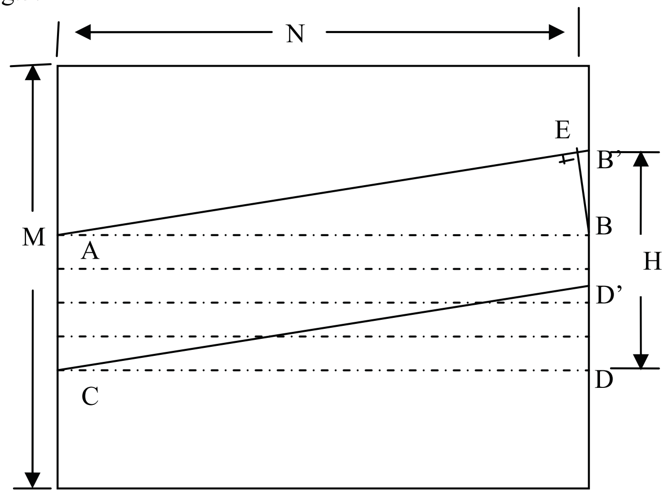


Figure 3. Illustration of ROI selection

A group of horizontal dashed staff lines is shown in Fig.3. The first and last lines are denoted by AB and CD, respectively. If the staff lines are rotated by an angle θ_{Max}, line AB and CD are moved to AB' and CD'. The height of rotated staff lines is the length of line segment DB', so obviously,

$$|DB'| = |DB| + |BB'| > |DB| + |BE|, \qquad (1)$$

where $|\bullet|$ is the length of line segment $\bullet$, and BE is perpendicular to AB'.

To make sure that the ROI contains a complete group of staff lines, we want the height H a little bigger than $|DB'|$, thus,

$$H = N * \sin(\theta_{Max}) + \frac{N}{\alpha}, \qquad (2)$$

where $N * \sin(\theta_{Max})$ equals to $|BE|$, and $\frac{N}{\alpha}$ is bigger than $|DB|$. In our experiments, α is set to 8.

Besides, the ROI is a rectangle that lies in the middle of given staff image. The first and last rows that segment a ROI are represented as r_1 and r_2, which can be computed with Eq(3,4).

$$r_1 = \frac{(N-H)}{2} = \frac{N}{2} - \frac{H}{2} \qquad (3)$$

$$r_2 = r_1 + H \quad = \frac{N}{2} + \frac{H}{2} \qquad (4)$$

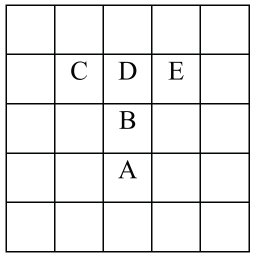
C. Candiate Point Extraction

Although method [8] cannot remove staff lines adaptively, it does remove the points lying on the staff lines. The idea is introduced briefly as follows.

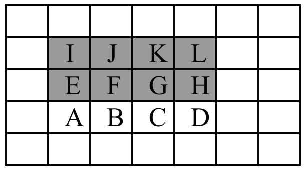
Table 1. Staff line removal method in reference [8]

Given a foreground pixel A (see Fig.4(a)), it is kept if and only if it satisfies the following condition. **Condition: the upper pixel B is a foreground pixel and at least one pixel above B(either C, D or E) is also a foreground pixel.**

It depends on the context to decide whether a pixel is a foreground pixel. In most cases, a pixel is regarded as a foreground pixel if its gray value is zero.



(a) Illustration of idea in reference [8]



(b) Analysis of staff line removal in reference [8]

Figure 4. Illustration of method in reference [8]

It can be concluded that the method removes the staff line points that lie on the top of a line, and it removes two rows each time. For example, if there is a short line segment occupying 3 rows, then the points lying on the upper two rows, i.e., from E to L, will be removed (see Fig.4(b)). This property can be used in section 4.

Therefore, we adopt this method and run it only once. The removed points are candidate points which will be used to estimate the rotation angle.

D. Rotation Angle Estimation

An optimization procedure is used to estimate the rotation angle of a music score image. The idea is that we rotate the music score image by a candidate angle which lies in the pre-selected range and compute a criterion correspondingly. If the criterion reaches its minimum, we obtain the angle by which the image should be rotated.

Suppose I represents an image which contains the candidate points extracted in section C, and I_θ represents I rotated by an angle θ, then the horizontal projection of I_θ can be obtained with Eq.(5).

$$P_\theta(m) = \sum_n I_\theta(m,n) \qquad (5)$$

The best angle can be estimated using an entropy minimization procedure which is described with the following equations. First, the projection is normalized with Eq.(6).

$$\bar{P}_\theta(m) = \frac{P_\theta(m)}{\sum_m P_\theta(m)} \tag{6}$$

Then, the entropy of the normalized projection is used to compute the Shannon entropy [12].

$$En(\bar{P}_\theta(m)) = -\sum_m \bar{P}_\theta(m)\log(\bar{P}_\theta(m)), \tag{7}$$

where $\bar{P}_\theta(m)\log(\bar{P}_\theta(m))$=0 when $\bar{P}_\theta(m)$=0.

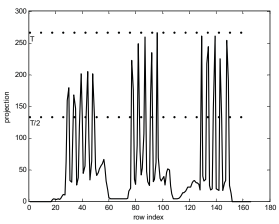
Finally, the rotation angle is estimated using the following optimization procedure.

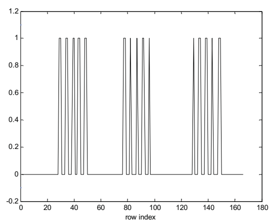
$$\theta_{opt} = \arg\min_{\theta\in(-30^\circ,30^\circ)} En\left(\bar{P}_\theta(m)\right) \tag{8}$$

As we know, entropy is a measure of uncertainty. $En(\bar{P}_\theta(m))$ reaches its maximum when $\bar{P}_\theta(m)$ is equilibrium and it arrives its minimum when one component is 1 and the others are zeros. It can be induced that when one group of staff lines is projected to only five or a little more bins, Eq.(7) will arrive its minimum. The conclusion can be shown in Fig.6(d).

E. Adaptive Parameter Estimation

(a) Original music score image





(b) Horizontal projection (c)Thresholded projection

Figure 5. Illustration of line width estimation

After the best rotation angle is obtained, the music score image is corrected. The gray values of the image are first reversed, i.e., bigger value means foreground pixel. The horizontal projection, $P_{\theta_{opt}}(m)$, is obtained again and the maximum of the projection is computed correspondingly.

$$T = \max_m(P_{\theta_{opt}}(m)) \tag{9}$$

Then the projection is binarized using a threshold of the half of the maximum.

$$V(m) = \begin{cases} 1, & P_{\theta_{opt}}(m) > T/2 \\ 0, & otherwise \end{cases} \tag{10}$$

The staff line width, W, can be estimated by computing the mean number of consecutive ones in $V(m)$. The whole staff line width estimation can be illustrated in Fig.5.

To completely remove the staff lines, we need to run method [8] Z times and Z can be computed with Eq(11).

$$Z = ceil(\frac{W}{2}), \tag{11}$$

where $ceil(\bullet)$ is the smallest integer no less than $\bullet$.

Sections B to E propose an adaptive method which estimates the rotation angle automatically and remove the staff lines with an adaptive parameter Z.

III. Experimental Results

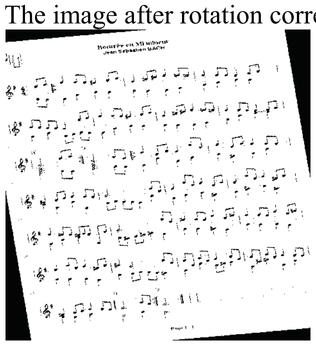
There are two datasets. The first one contains of 430 music score images downloaded on Internet and the second one contains of 125 scanned images. The scanned music score images have different rotation angle. Two examples can be found in Fig.1. The proposed method works well for both datasets. The processing results are shown as follow. Since our focus is the pre-processing of music score images and the classification is not done yet, all the results are visually checked. Comparison to other methods will be done in the future.

A. Scanned Images

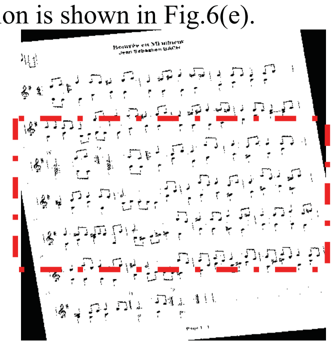
For a scanned music score image, the processing is shown in Fig.6. A ROI is first selected to do rotation correction. The selected ROI is marked with a dashed rectangle in Fig.6(b).

The candidate points are extracted according to section II.C and the results are shown in Fig.6(c). We assume that the biggest possible rotation angle is 30 degree, so the ROI contains 4 groups of staff lines. Actually, few images have a rotation angle bigger than 10 degree, so the ROI can be still reduced.

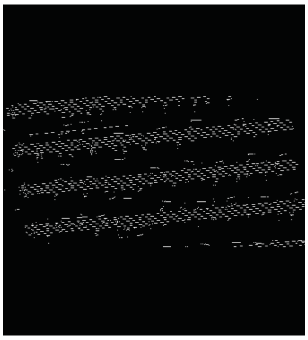
Then we estimate the rotation angle with Eq.(7). For the given image, we rotate the image by an angle and compute the corresponding entropy. We can see that when the rotation angle is -7 degree, entropy reaches it minimum. This indicates that the image will be corrected if it is rotated by -7 degree. The image after rotation correction is shown in Fig.6(e).



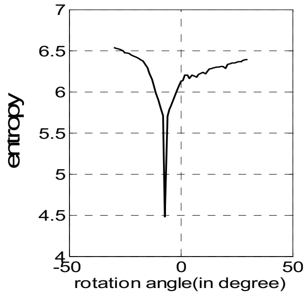
(a) Original image



(b) ROI selection



(c) Candidate points



(d) entropy vs. rotation angle

(e)After rotation correction

(f) Staff lines removed

Figure 6. A scanned images and processing results (line width is 1 pixel)

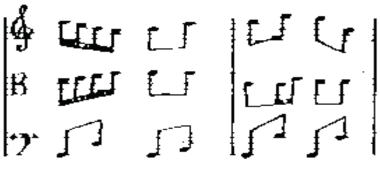
Now, the staff lines are horizontal so we can compute horizontal projection according to Eq.(5). The line width is estimated and the estimated width is one pixel, so we need to run method in Table 1 only one time to remove the staff lines. The results are shown in Fig.6(f).

B. Downloaded Images

For a downloaded music score image, there is no obvious rotation, so the rotation correction procedure may be omitted. This can further improve the efficiency of the proposed method.

C. Image with Thick Staff Lines

(a) Original image



(b) Staff lines removed

Figure 7. Processing results for images with thick staff lines (line width is 5 pixels)

To evaluate the performance of the line width estimation, we select some images of different line width and the proposed method works well.

A music score image is shown in Fig.7 and its staff line width is 5 pixels, so we need to repeat the method in Table 1 three times to remove the staff lines.

IV. CONCLUSION

This paper proposed an adaptive method for staff line removal in music score image pre-processing. The method corrects the rotation and removes the staff lines automatically. Experimental results showed that the method is quite robust. The originality can be described in two aspects.

(1) The rotation angle estimation. It uses a statistical modal and is quite robust.
(2) The adaptive staff line removal.

In the future, we will make a comparison to other staff line removal methods.

REFERENCES

[1] B. Alphonce, Optical Music Recognition: A Progress Report, Proceedings of the small Computers in the Arts, 8-12, 1988.,

[2] D. Blostein, H. Baird, A Critical Survey of Music Image Analysis, in Proceedings of Structured Document Image Analysis, Berlin: Springer-Verlag, 1992:405-434.

[3] Ichiro Fujinaga, Adaptive Optical Music Recognition, Doctoral Thesis, McGill University, Montreal Canada, 1996.

[4] H. Miyao, et al, Symbol Recognition for Printed Piano Score Based on the Music Knowledge, IEICE Trans, 1992:J75-D-II:1848-1855

[5] B. MOdayur, et al, MUSER-A Prototype Musical Recognition System Using Mathematical Morphology, Machine Vision and Applications, 1993:6(2-3):140-150.

[6] Florence Rossant, A Global Method for Music Symbol Recognition in typeset Music Sheets, Pattern Recognition Letters, 2003:1129-1141.

[7] D. Bainbridge, et al, The challenge of Optical Music Recognition, Computers and the Humanities, 35:95-121, 2001.

[8] A.T.Clarke, et al, Inexpensive Optical Character Recognition of Music Notation: A new alternative for publishers, Proceedings of the Computers in Music Research Conference, Lancaster, UK, April 1988, pp.84-87.

[9] Etienne Sicard, An Efficient Method for the Recognition of Printed Music, Proceedings of 11th International Conference on Pattern Recognition (IAPR).

[10] S. Marinai, et al, Projection Based Segmentation of Musical Sheets, Proceedings of the Fifth International Conference on Document Analysis and Recognition, 1999.

[11] Xiaoxiang Liu, et al, A fast and Robust Approach to Detect and Remove Staff Lines from Music Score Image, in Computer Engineering and Applications, pp.210-212 & 217, 2004.

[12] C.E. Shannon, A Mathematical Theory of Communication, Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948.

[13] L. Pugin, "Optical Music Recognition of Early Typographic Prints using Hidden Markov Models," *Proc. Seventh Int'l Conf. Music Information Retrieval,* pp. 53-56, 2006.

[14] J. Lorenzo, Optical Music Recognition for Scores Written in White Mensural Notation, EURASIP Journal on Image and Video Processing, Volume 2009.