RESEARCH ARTICLE OPEN ACCESS

Musical Sheet Segmentation

Vinaya V*, Dhanya M Dhanalakshmy**
*(Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Coimbatore-641112)
** (Department of Computer Science and Engineering, Vishwa Vidyapeetham, Coimbatore-641112)

ABSTRACT
Music has had an important role in human life from time immemorial. Music has been shared in two ways aurally and as written documents known as musical notes or musical scores. Many ancient cultures have used musical symbols to represent melodies and lyrics but none of them is as comprehensive as a written language or document. Thus the knowledge of ancient music is limited to a few fragments which are mostly unpublished. Hence to preserve such music we need to introduce a computerized system to digitalize and to decode the musical symbol images and reconstruct it as new score which will be in machine readable format. Here the machine readable format is MIDI. Prior to converting the musical symbols into MIDI format, musical sheets need to be segmented for isolating the musical symbols. Since synthetically generated musical sheets are used here, the segmentation process can be carried out using recursive graph cut method. Here we discuss the initial few steps such as staff line removal, text removal and segmentation of musical sheet.

Keywords—Digitalization, MIDI(Musical Instrument Digital Interface), Musical scores, Reconstruction, Segmentation, Staff lines.

I. INTRODUCTION

The musical score sheets are the only available artifacts for transmitting musical expressions non - aurally. Over the years musical score sheet has changed dramatically in the way they are presented as well as in the symbolic content. Musical symbols are processed in order to convert them into machine readable electronic format. The machine readable format is known as MIDI (Musical Instrument Digital Interface). For this process isolation of musical symbols is important. Thus we have to make the musical sheet devoid of any staff lines and text in the first step. Then the musical sheet must be subjected to segmentation process using any suitable algorithm. The unique property of musical sheets is staff lines. The five equally spaced lines over which the musical notes are written. These staff lines itself poses a problem when musical sheet segmentation is considered. Since staff lines touches the musical symbols, their removal will leave musical symbols fragmented with corresponding loss of information.

In this paper, a segmentation process for the extraction of musical symbols is described. It is based on extensive use of projection profiles after staff line removal. This paper is organized as follows. Section II presents the general architecture for the musical sheet segmentation process. Section III presents the related works in this field. Section IV describes the algorithms used, and the corresponding results. Section V discusses about the future enhancements and section VI makes the conclusion.

II. GENERAL ARCHITECTURE

The steps for isolation of musical symbols are staff line removal, text removal and the segmentation process. Segmentation here is done using recursive x-y graph cut method.

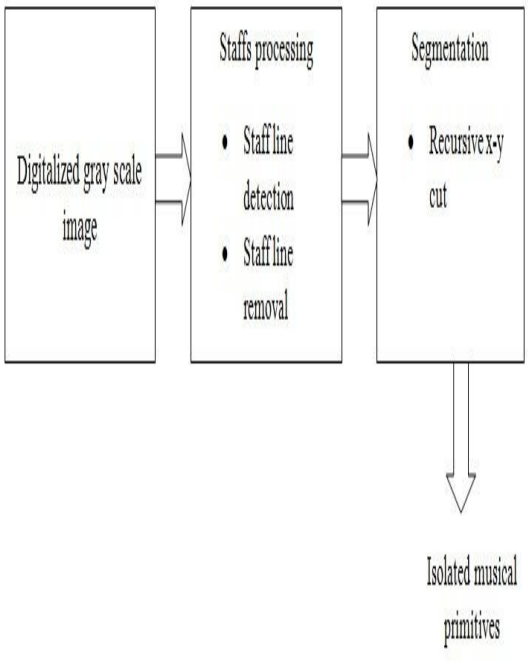

Fig 1: Basic architecture for music sheet segmentation

III. RELATED WORKS

The works on the conversion of musical note to MIDI conversion can be dated back to the early works by

In [1] the input to the system is musical scores send through facsimiles. It will be corrupted by noise, handwriting variations, non standard page layout and notations. The method proposed in this paper handles binarization, staff detection and noise removal by dynamic programming. The main assumptions used here are:

a) the staves (the 5-groups of staff lines) are represented by repetitive line patterns, which are constrained as well as informative , and thus a direct optimization over such patterns is used instead of first spotting single staff lines,

b) the binarization threshold which is optimum also is the one giving the maximum evidence for the presence of staves

c) the noise, or background, is given by the regions where there is insufficient stave pattern evidence.

Using dynamic programming the stave evidence can be propagated in one pass from one side of the document to another. The best staves are generated by tracing back the solutions. Another DP formulation solves the problem more optimally by labeling of such tracked solution into background and stave segments. This leads to accurate stave/ background segmentations and thus noise removal. The disadvantage is that is only suited for musical images transmitted through facsimiles.

In [2], the paper discusses about the implementation of a complete optical music recognition system named lemon. The implementation is starting from staff line detection. Then the text regions are segmented out, lines are detected and removed. It is followed by symbol recognition and semantic interpretation of the symbols. The algorithm for staff line processing works by examining the vertical columns of the image. Here also the a priori knowledge that the staff line consists of five set of lines with four equal spacing in between is used. Segmentation is carried out using line adjacency graphs. Since this LAG is found out after staff line removal, there will be one LAG for each symbol. Then the music symbol is recognized. Symbol recognition is split into three parts,

i) line/curve detection,
ii) character-symbol detection,
iii) note head detection.

Thus the individual symbols are obtained. These individual symbols are synthesized to graph grammars which are more complex for further recognition process.

Paper [3] deals with the fundamental problem of staff line detection. The need for the staff line detection and removal lies on the fact that the musical symbols must be isolated for more efficient and correct detection of the same. Staff lines can be considered as the shortest path between the 2 margins.

1. Staff space height is estimated which can be used as reference length
2. Shortest path between left most pixel and right most pixel is found out using Dijikstra's algorithm.
3. Only the strip centered in the row of interest is used.
4. Inorder to discard false staff lines thresholding is applied.
5. Retained path is trimmed.
6. Paths are then validated.

In [4] the proposed algorithm uses images as graphs. Input images are gray scale rather than binarised. Thus the information loss is avoided. Here the concept of strong staff pixel is introduced. It is a pixel with high probability of belonging to a staff line. All pixels in the same run will either be considered as staff pixels or not staff pixels. The decision approach consists in scanning the run-length encoding of each column of the binary image of the music sheet in order to find black runs of *staffline height* pixels followed or preceded by a white run of *staffspace height* pixels. The pixels in the black runs that meet this condition form the set of Strong Staff-pixels. Here the disadvantage is the difficulty in deciding the cost function

In [5], The author proposes an effective staff line detection and removal method that makes use of the global information of the musical document. It models the staff line shape. It first estimates the staff height and space. Then the staff line is modelled by examining the orientation of the staff pixels. At last the estimated model is used to find out the location of staff lines for its removal. The proposed technique is simple and robust and involves only less number of parameters.

A musical sheet has several staff lines. These staff lines are a set of five equally spaced parallel lines on which the musical notes are written. The first note on the staff lines will be the clef which indicates the beginning of the musical notes. Tonality, which gives information about the accidental which is to be applied on the notes and the time signature indicating the number of beats per bar. The specialities of musical notes are that they are two dimensional. The horizontal axis represents the time and the position of musical note head in the vertical direction represents the pitch. The two dimensions of musical notes make the recognition processes difficult.

IV. ALGORITHMS

A.STAFF LINE DETECTION AND REMOVAL

Given a synthetically generated musical sheet our aim is to detect and remove the staff lines first. Staff line removal is one of the challenges faced while producing an electronically readable format of music. For manually reading the notes, the staff lines are very important since the position of the musical notes on the staff lines distinguishes the musical notes from one another. But for machines these staff lines will become an obstacle as it produces difficulties at the time of segmentation and further processing. One major problem faced while removing the staff lines are the broken musical notes. In order to overcome this, efficient ways to reconstruct the symbols have to be devised on a later stage.

Staff line detection

The two important parameters used for the detection of staff lines are the staff line thickness and the vertical space between a pair of staff lines, the staff space height. These are the most frequent black and white run of pixels respectively. To detect staffs we sum up all the binary values for each row and get a histogram which shows the number of non zero values in a row. The lines having the staff will correspond to remarkably high peak. The position of peaks in histogram gives the location of staff lines. Using the location information we can obtain the width or the staff space.

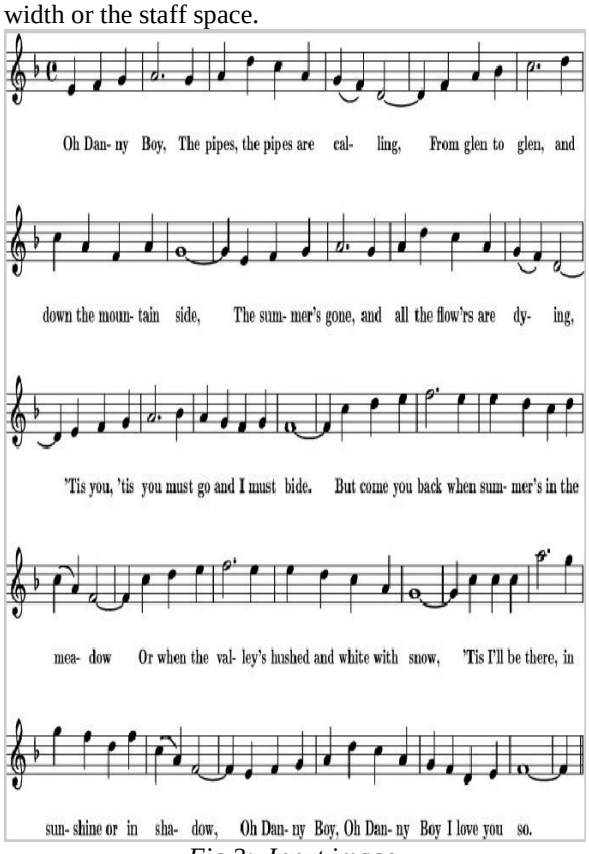

Fig 2: Input image

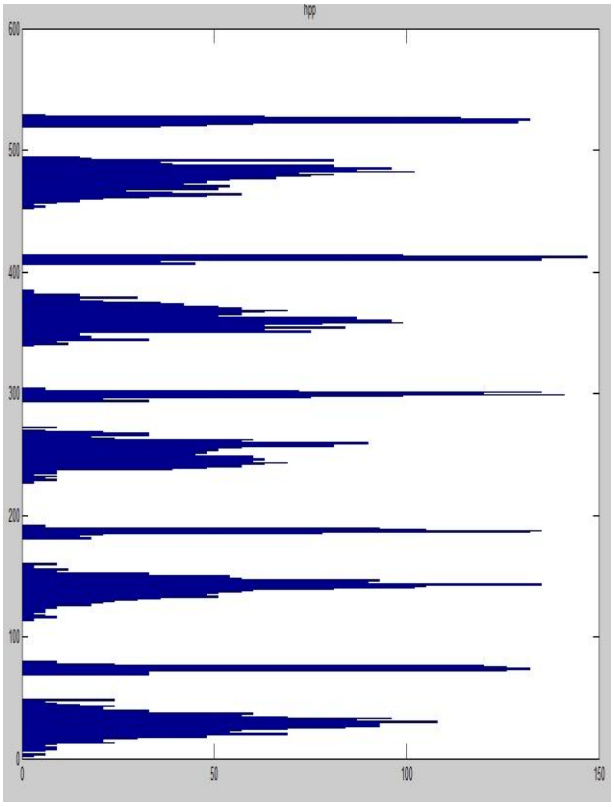

Fig 3: Histogram showing the row index of input image

Staff line removal

Any vertical black run that are more than twice the staff line height are removed as well as any component whose width is less than staff space height is also removed. After this process all the musical notes are removed and the staff lines are isolated. Now the image is having isolated staff lines and is devoid of any other musical symbols. This image is then XORed with the original image to obtain the staff line removed musical sheet.

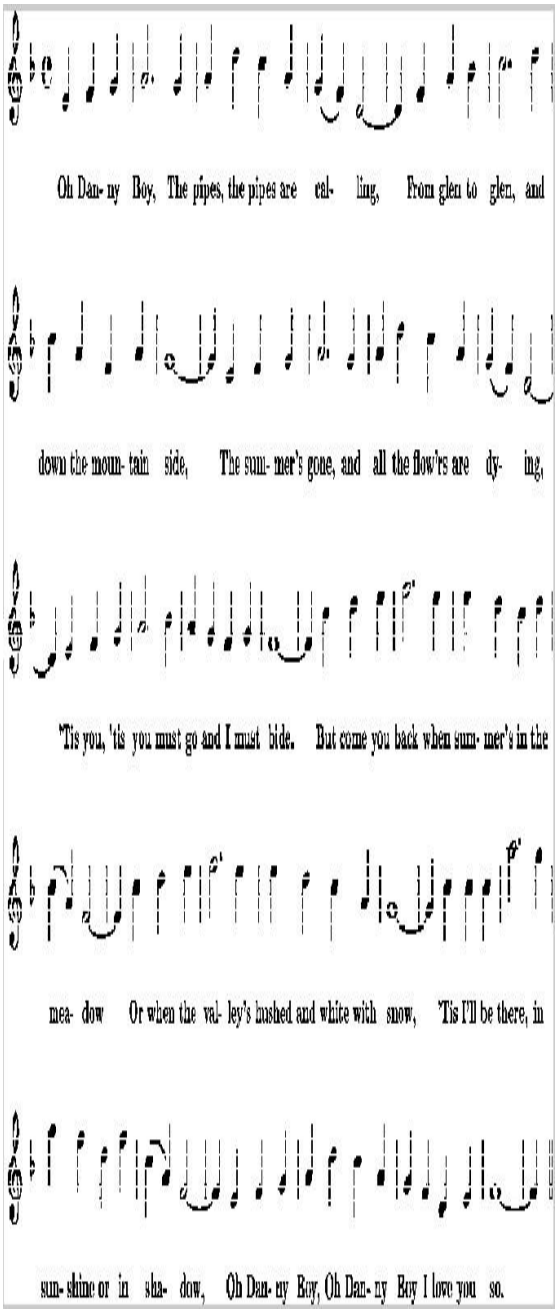

Fig 3: Staff line removed image

B.TEXT REMOVAL

Text such as lyrics should be removed as much as possible so as to reduce the burden of the classifier. Text can be differentiated from musical symbols by its characteristics. Text symbols have basically same height and are placed side to side. For this a connected component analysis is done. Those connected components which have a value greater than a threshold and aspect ratio are removed. The musical notes have larger aspect ratio and is composed of bigger connected components. Thus all the musical notes are removed leaving the lyrics alone in the sheet music.

Again the output image with only lyrics is XORed with the staff line removed image. Now we obtain the musical sheet having neither staff lines nor lyrics but musical symbols only.

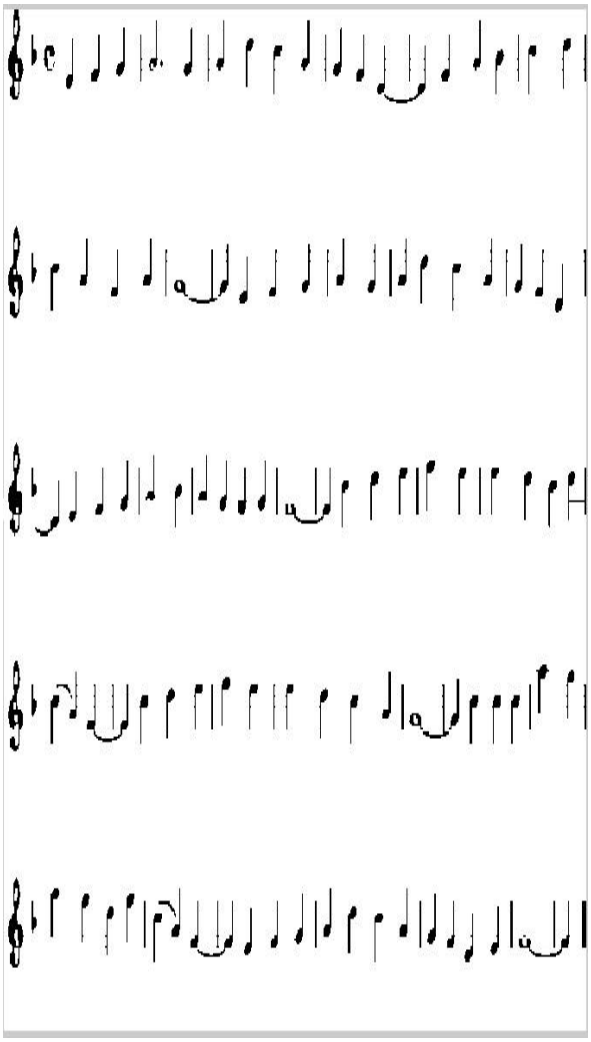

Fig 4: Text removed image

C.SEGMENTATION

Segmentation process here is used to isolate the symbol primitives for identification process. Since the musical sheet is devoid of staff lines and lyrics we can use recursive graph cut for segmentation process. It is a structured top-down method that also uses projection profiles splits the document into successively smaller rectangular blocks by alternately making horizontal and vertical "cuts" along white space, starting with a full page, and continuing with each sub-block. The locations of these cuts are found from minima in the horizontal and vertical projection profiles of each block.

METHOD

-Horizontal and vertical projection profiles of each node are computed in every step of recursion.

-Then, the valleys along the horizontal direction and vertical direction is compared to corresponding predefined thresholds .
- If the valley is larger than the threshold, the node is split at the mid-point
-The process continues until no node can be split again

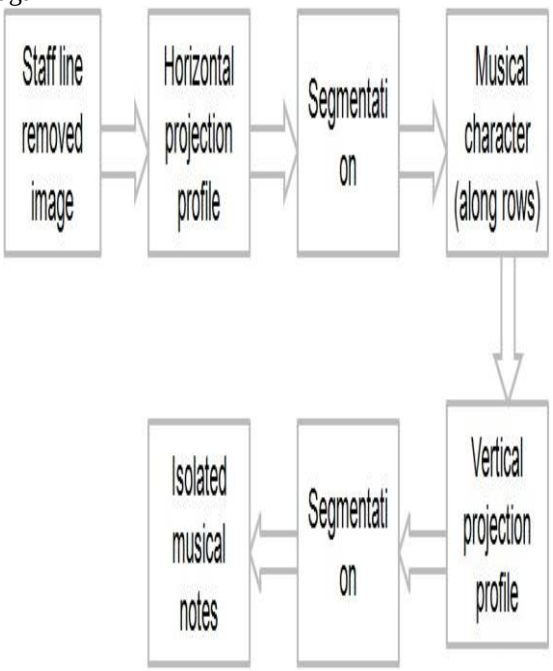

Fig 5: Segmentation using recursive x-y cut

Staff segmentation
The image is first subjected to horizontal (x direction) cut. Here the horizontal projection profile is considered for this process. E Horizontal projection profile represents the number of continuous black pixels graphically. In the graph we obtain peaks, which indicates the presence of pixels and valleys, where there are no pixels. The cut is made on the middle portion of the valley. Thus we obtain individual staff segments(five staff segments for each musical sheet).

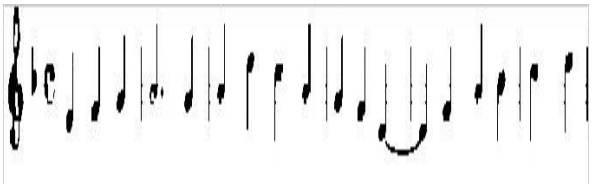

Fig 6: Segment 1 obtained after x cut

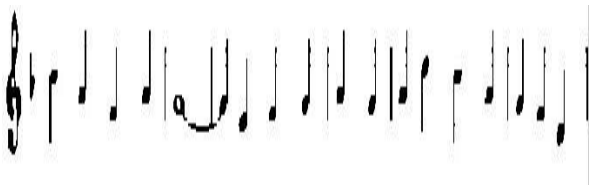

Fig 7: Segment 2 obtained after x cut

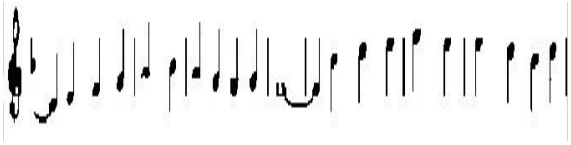

Fig 8: Segment 3 obtained after x cut

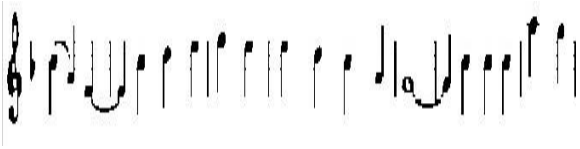

Fig 9: Segment 4 obtained after x cut

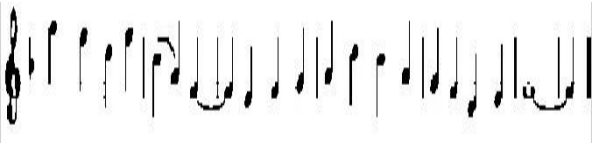

Fig 10: segment 5 obtained after x cut

Isolation of individual notes
1) Each staff segments are then taken as input for vertical graph cut in y direction. The vertical projection profiles of each of the staff segments are taken and then the position to cut is deduced from the resulting graph. Thus individual musical notes are obtained.
2) We can also do a connected component analysis on each of the staff segments and extract them. It also results in the isolation of musical primitive.

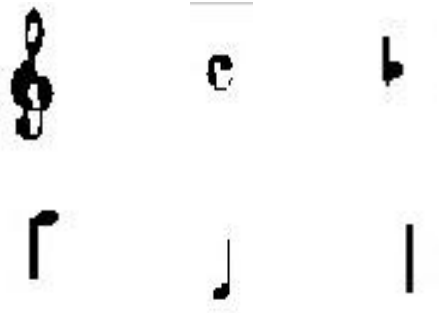

Fig 11: Different symbols obtained after y cut

V. FUTURE WORKS

This work can be extended further to the classification process, musical note reconstruction and representation processes. The final music sheet can be then converted to graphic publishing file using softwares such as photoscore, Vivaldi score etc. The work presented here works for early music notations. This can be further improvised so as to work for scanned handwritten documents, camera captured documents.

VI. CONCLUSION

For the conversion of a musical sheet into its electronic or machine readable format, a proper staff

line removal and segmentation process are required initially. The staff line removal and the segmentation process described in this paper gives a good result for isolation of musical symbols for further classification and recognition. Once the musical symbols are classified and recognized, the musical sheet should be reconstructed. Then the reconstructed musical sheet can be given to softwares like photoscore, vivaldiscan etc. for converting into MIDI format which can be played using any custom music player.

REFERENCES

Journal Papers:

[1] Radu Timofte and Luc Van Gool," **Automatic Stave Discovery for Musical Facsimiles"** , *ACCV 2012, Springer-Verlag Berlin Heidelberg 2013, pp. 510–523*

[2] J.W. Roach, J.E. Tatem "**Using Domain Knowledge in low-level visualProcessing to interpret handwritten Music: an Experiment**." *PatternRecognition 21, pp. 33-44 (1988)*

Proceedings Papers:

[1] K Todd Reed, J R Parker**, "Automatic Computer Recognition of Printed Music***" IEEE proceedings of ICPR-1996,pp: 803 – 807, volume 3*

[2] Alicia Fornés a;1, Josep Lladós a and Gemma Sánchez **"Staff and graphical primitive segmentation in old handwritten music scores"** *6th International Workshop, GREC 2005, Hong Kong, China, August 25-26, 2005, pp:279-290*

[3] Ana Rebelo , Artur Capela , Joaquim F. Pinto da Costa, Guedes, Jaime S. Cardoso**,"A Shortest Path Approach for Staff Line Detection"** *Automated Production of Cross Media Content for Multi-Channel Distribution, 2007. AXMEDIS '07. Third International Conference*

[4] D Bainbridge and T C Bell **"Dealing With Superimposed Objects In Optical Music Recognition"** *Sixth International Conference On Image Processing And Its Applications, pp: 756 - 760 vol.2*

[5] Ana Rebelo and Jaime S.Cardoso, Senior Member, IEEE **"Staff line Detection and Removal in the Grayscale Domain"** *2013; 12th International Conference on Document Analysis and Recognition*