

International Journal of Advance Research in Computer Science and Management Studies

Research Paper

Available online at: www.ijarcsms.com

Practical Issues in the Field of Optical Music Recognition

Apurva Ashokbhai Mehta¹

Dept. of CE
DDU
Nadiad - India

Malay S. Bhatt²

Assoc. Prof.
Dept. of CE
DDU
Nadiad - India

Abstract: 50,000 years and still counting that's history of music. Aural and visual forms are very widely known at the same time written music is also in existence. Music written down on paper known as music score is taken as input for OMR and makes the sheet readable and editable for computers. In last 50 years several authors have applied their idea to make OMR more rich and efficient as OCR. Highly complex and diverse structure of OMR sheet makes it difficult to generate music automatically as it requires extensive knowledge of various field such as music theory, image processing, machine learning. That's why still process of OMR is unstable, less effective and rarely robust. Presence of various techniques creates confusion while selecting a proper one. This paper identifies practical issues presence in the field of OMR and surveys relevant techniques available for the same.

Keywords: Optical Music Recognition, Standard Score, Handwritten Score, Symbol Segmentation, Symbol Recognition With staff line, Without Staff line.

I. INTRODUCTION

Music – This word is part of every living and non living material. It sounds weird but we can have hobby paradox for music as we have birthday paradox in general. Music serves as a therapy as well. We are having lots of music, now days. Based on time and place it gets its flavor. Music is in existence since last 50,000 years [16]. Music was invented in Africa and then it has marked its presence in entire world. Its unanimous fact that span of music is everywhere. Day by day it's increasing, due to its relation with many other fields (i.e. Entertainment, Therapy, Education, etc...). Music is represented in different format. One of them is sheet music representation. Musician use the score sheet to pen down music and even use it to play the music written down on score sheet. To make music available and understandable to every individual even with less amount of prerequisite knowledge optical music recognition (OMR) can help. OMR is summarized as 'A process that recognizes music from any form of score sheet and makes the sheet readable and editable for computers'.

Representing music in score sheet format is well known idea since lots of centuries. Various historical documents suggest the presence of music representation in score sheet format. The following figure 1 describes the taxonomy of OMR representing its evolution.

OMR is basically categorized in two parts. Handwritten score OMR and Standard score OMR [2]. For both category evolution is tremendous. Evolutions of various forms are based on time, place and some other parameters. These forms have somewhat variations in symbols representation for different set of musical instrument available. Research work can be carried out in any of the prescribed domain. As still accuracy, robustness and heterogeneity are open issues in OMR.

OMR is somewhat related to optical character recognition (OCR) [9]. Following figure 2 describes both OCR and OMR representation. Since 1960s OMR is one of the challenging topics among researchers from various streams like computer

science, music theory, mathematics and many others. Sometimes OMR is also referred as a Music OCR a well spanned pattern recognition field. OMR is much more complex compare to its counterpart OCR due to OMR's two dimensional structure, as represented in figure 2. Here, horizontal scale provides the order in which music notes are arranged. In music sense it's the lyrics. The presence of a symbol (i.e. note) at particular vertical position with respect to stave lines or staff space determines the pitch of that symbol. Symbols can be found outside of staff boundary as well.

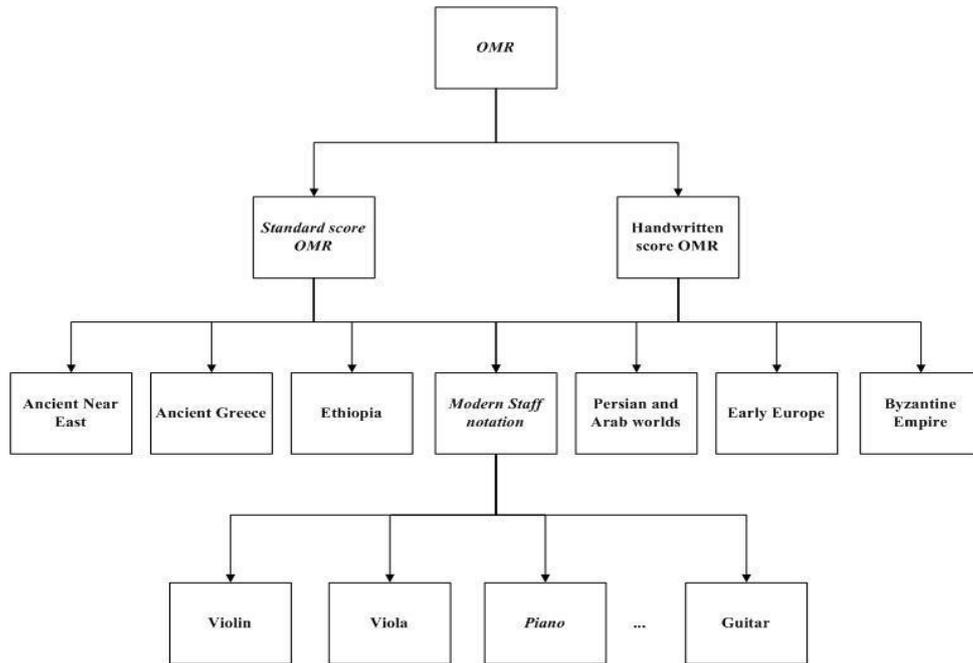


Figure 1 OMR Taxonomy

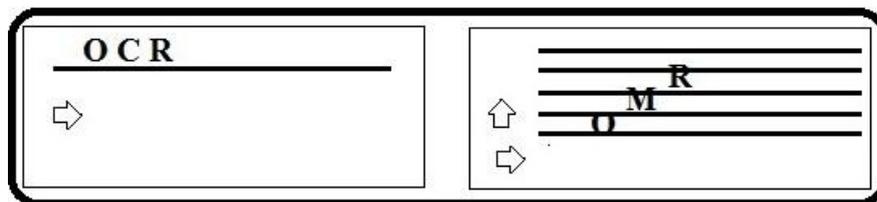


Figure 2 OCR OMR Representations

Application area of OMR is wide spread. It helps in educational and training field as well in recent digital era OMR helps us in preparing digital database of old score sheets. It can serves as Sur recognizer and music player and mixer.

Paper is prepared in following manner. Section II discusses basic architecture of OMR. Section III showcases practical issues in the field of OMR. Section IV concludes the paper. Last two sections give clue about score sheet terminology and acknowledge the references.

II. OMR ARCHITECTURE

Entire OMR can be divided into various phase based on the task to be perform to transform music score written on page into music. OMR process starts with preparing an image for OMR. Once image is preprocessed it is segmented in form of basic and complex symbols. Symbols are recognized in the later stage phase of OMR. Recognized symbols are required to represent in the 2-D form. Spatial relation exists between various symbols based on their position in score sheet. Represented 2-D form notes are transformed into music during final stage of OMR. Each stage requires support of various techniques to perform OMR. Each stage of OMR architecture is discussed at length in the following section.

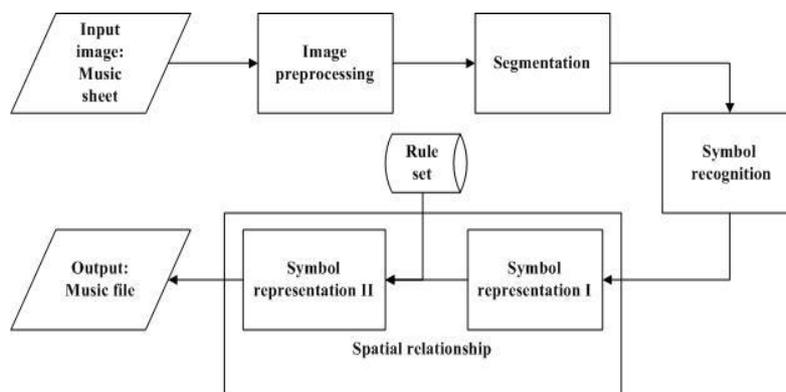


Figure 3 OMR Architecture

III. PRACTICAL ISSUE IN OMR

Practical issues and techniques that assistance in solving corresponding issues is discussed stage wise.

A. Input

Input to the system is image of any widely used format. Pixel relationship is very important in music score sheet. So, it is advisable to fix the scale of input image. Again this limits scale of system. Solution is to rescale every input image to specific value so algorithms and techniques present in following stage can work properly without any further modifications.

B. Preprocessing

Image pre processing is required in most of the pattern recognition application. The main idea is to increase the quality of being reliable in OMR. Initial input to the system is an RGB image. However the color detail is never of interest to OMR systems [3]. In effect we can optimize the OMR process by reducing color detail as much as possible. It is preferable to convert given RGB input image into grayscale image. Grayscale conversion is merely based on method used by tool in use. Various authors perform binarization on grayscale image so that image foreground details are clearly separated from background. To perform binarization threshold value is required. Various local and global algorithms can help here. Otsu's thresholding based on clustering with global threshold is very famous technique. Even Niblack's thresholding based on mean and standard deviation in the local neighborhood is also popular techniques for same task.

Input of the system can be considered from various sources. Provided source is a camera image then due to illumination and lens adjustment image might have several noise effects [4]. Source is a scanner image then due to scanner adjustment variation image might have several noise effects [8]. Normally image might get blurred due to improper adjustment parameter used in aforementioned sources. Sometimes image content is transformed to skewed dimension again due to improper adjustment parameter. For De-skewing Connected Component Analysis based on centroid values is smart choice. It is also possible that by projecting image with various projections based on variance values we can correct skew present in score sheet.

C. Segmentation

In any pattern recognition problem segmentation is one essential phase. Segmentation step makes recognition and representation easier for OMR. During segmentation we require to separate each music symbol with maximal difference between boundaries of next and previous music symbol. Entire music score sheet can be segmented into various parts. In few OMR system [3][4][5][8][10] authors have preferred to remove stave line from score sheet. This is obvious as once staff lines are removes sheet is left with score sheet symbols only. It is easy to segment symbols, now. For this task we can use morphological open operation with proper structuring element. If this approach is followed during segmentation then it becomes complex to restore the 2nd dimension of symbols recognized from score sheet. Few authors [1][6][7][9][11][12][13] that's why prefer to segment sheet along with stave lines. Even while we remove staff lines it is possible that this operation might disturb usual shape of symbols, introduce blurring and other noise in score sheet.

Hierarchical decomposition is seems to most suitable one for segmentation process. It is natural to hierarchy in which sheet is prepared. This method is used by several authors [4][9]. Hierarchical decomposition can also be achieved in couple with thresholding methods [1][8]. Morphological operation can also considered as another option. This option turns out to costly and complex as we require designing various structuring element with respect to varied set of OMR symbols. Connected Component Analysis [5][7] and Template Matching [4] are also usable approach for segmentation. Formal requires support of complex data set and later requires having so many template for large set of music symbols set. So, again both approaches turn out to complex and time consuming.

D. Recognition

Symbol segmentation process outputs dozens of symbols. It requires recognizing all those symbols. Recognition process gives identity to a mere symbolic structure and help system to convert a normal symbol into a musical symbol. Recognition of symbols can be accomplished with the help of various methodologies. Recognition using MusBNF grammar a completely linguistic approach is one such recognition way [10]. Again this approach requires strong domain knowledge and requires building large rule set. So, it turns out a time consuming process but at end it serve with strong results. Few systems have followed morphology based approach [3][12]. Problem with morphology approach is already discussed in previous stage. Statistical and probabilistic approach is gaining lots of attention now days in particularly OMR after its successful implementation in various products. Classification trees is also one more such approach for recognition [11]. Previous work suggests that it gives poor result. Even classification trees over feet very often. Frequently it becomes unstable because some small variation in data might generate an entirely new tree. Fuzzy system is one viable option for recognition [13]. This requires broad knowledge for subject under study. Many experts together form various rules and that make a strong fuzzy system. This is hard to achieve as rules are so many and they frequently change based on mind set of rule creator. Artificial neural network is widely used approach in recognition for OMR [1][6][8][9]. Neural network is considerable weapon in the area of pattern recognition. Lots of reason support decision in favor of neural network. Neural network can easily adapt to any compound function due to characteristic of being non linear. Problem solving can be made efficient by introducing hidden layers in between input layer and output layers. Neural network proves its effectiveness by helping us solving and moderating curse of dimensionality issue. Neural network learns by examples. This feature provides ease of use with respect to knowledge require for input domain. Neural network is parametric model. This fact is beneficial while input is of large training sample. Neural network helps us in concluding multiple patterns. That's how neural network has proved its importance in the field of pattern recognition by its heterogeneous applicability. Neural network allows supervised learning, unsupervised learning and reinforcement learning with the support of various algorithms available to support each learning method. There exist many methods to recognize music notes. One is to use multi-layer back propagation neural network. It takes advantages of non linear approximation with greater flexibility. However, it is critical to select the number of layers and the threshold of error functions for a reasonable convergence of weights. We can use probabilistic based and density based neural networks as well. With SVM problem of over fitting makes the correct thing incorrect or vice versa but if have big and diverse dataset SVM can turns out a big achiever in terms of recognition rate.

E. Representation

Recognized symbols given 1st dimension of symbols present in score sheet, still to construct 2nd dimension we need to locate exact vertical position of a recognized symbol in score sheet with respect to stave line or stave space. Once both dimensions are recognized it is more essential to express spatial relationship of a symbol with neighboring symbol. This can be achieved by building rule set that handles all complex spatial relation cases. We can deliver this representation in text file or through NIFF or via XML notations.

F. Music Generation

Last stage of OMR is to play represented notes. This can be achieved via various ways. Two ways are most popular. One is MIDI table lookup and another is generating discrete sinusoids of one octave of musical notes.

IV. CONCLUSION

This paper followup entire OMR process in general by considering various practical issues. It also suggests possible solution for corresponding practical issues with broader view. For standard score sheet it seems that community will definitely achieve good results due to availability of various datasets like IMSLP [14] for validation and testing. An effectual OMR system can help in various fields like education, entertainment, search optimization, etc... It is possible to inherit ideas and act on solution given in paper for various practical issues to recognize music from score sheets belongs to several instrument like piano, guitar, drum and many other.

Appendix A

This section defines terms frequently used in Report. For detailed reference [15] is good resource to follow.

Staff: Music is written down on staff. Normally staff contains five horizontal parallel equal space lines. Four spaces are created due to five lines and notes can be placed on this space, too.

Grand staff: Piano music is represented by grand staff. It contains two staff. Each grand staff shows notes to be played with left and right hand on piano. Upper half represents notes to be played by right hand and lower half represents notes to be played by left hand.

Measure: Staff is divided by bar lines. Purpose is simple - making it easy to read. Even it helps in identifying time scale, too. Span between two bar line is known as measure.

Clefs: In piano music two clefs are important: treble clef and bass clef. Treble clef shows notes to be played with right hand and bass clef shows notes to be played with left hand.

Note: Beat of music is represented by notes in music score sheet. These notes are whole, half, quarter, eighth, etc..

Rest: Silence in music is indicated by rest symbols. These rests are whole, half, quarter and so on.

Accidental: Pitch of the note can be altered with the help of accidental. There are three types of accidental are available: Flat, sharp and natural. A flat note lowers the pitch, a sharp note raises the pitch and a normal note cancels the effect.

Octave: An octave is a large musical interval that separates two of the same notes

References

1. Tatiana Tambouratzis, "Identifications of key music symbols for optical music recognition and on-screen presentation", IEEE, 2011.
2. Ana Rebelo, Ichiro Fujinaga, Filipe Paszkiewicz, Andre R.S. Marcal, Carlos Guedes and Jaime S. Cardoso, "Optical music recognition state of the art and open issues", Springer, 2012.
3. Dingyi Li, Bing Han and Jia Ji, "Music score reader based on morphological image processing", Stanford, 2012.
4. Steve Dai, Ye Tian and Chun-Wei Lee, "Optical music recognition and playback on mobile devices", Stanford, 2012.
5. Albert Ng and Asifkhan, "Automated instructional aid for reading sheet music", Stanford, 2012.
6. Toyokazu hori, Shinichiro wada, Howzan Tai and S. Y.Kung, "Automatic music score recognition/play system based on decision based neural network", IEEE, 1999.
7. Christopher Raphael and Jingya Wang, "New approaches to optical music recognition", ISMIR 2011.
8. Tomas Beran and Tomas Macek, "Recognition of printed music score", Springer, 1999.
9. P. Bellini, I. Bruce and P. Nesi, "Optical Music Sheet Segmentation", IEEE, 2001.
10. Mariusz Szwoch, "Guido: A Musical Score Recognition System", IEEE, 2007.
11. Wladyslaw Homonda and Marcin Luckner, "Automatic knowledge acquisition: Recognizing music notation with methods of centroids and classification trees", IEEE, 2006.

12. Bharath Modayur, "Music score recognition-A selective attention approach using mathematical morphology", University of Washington, 1996.
13. Florence Rossant and Isabelle Bloch, "A fuzzy model for optical recognition of musical scores", Elsevier, 2004.
14. International Music Score Library Project (IMSLP), Online Image dataset (<http://imslp.org/>).
15. Jonathan Harnum, "Basic music theory", SolUt Press, 2001
16. http://en.wikipedia.org/wiki/Music_history
17. <http://en.wikipedia.org/wiki/Pian>

AUTHOR(S) PROFILE



Apurva Ashobhai Mehta is working as Asst. Prof. in Computer Engineering at Dharmsinh Desai University, Nadiad, India. His fields of interests are Image Processing and Graph Theory. He has attended various STTPs and seminars. He is active member of professional societies like ISTE and CSI.



Malay S Bhatt is working an Assoc. Prof. in Department of Computer Engineering at Dharmsinh Desai University, Nadiad, India. His fields of interests are Digital Image Processing and Data Mining. He has published more than 35 papers in various international journals and conferences.