Air Valve Fuzzy Control Combined with Sheet Music Recognition Techniques Applied to Autoplaying Soprano Recorder Machines

Chun-Chieh Wang[1], Guang-Ming Jhang[2]

[1] Department of Automation Engineering and Institute of Mechatronoptic Systems, Chienkuo Technology University, Taiwan
(Tel: 886-4-7111111ext3350, Fax: 886-4-7111164)
[1]jasonccw@ctu.edu.tw
[2] Department of Electronic Engineering, Chienkuo Technology University, Taiwan

Abstract

In the past research, there are many disadvantages to score recognition and flute performance. Therefore, this paper would improve on these two points. For the part of music score recognition, we use the y-axis projection method to detect the staff position and remove it to replace the erosion and expansion in morphology. It was found that the notes have a specific writing style on the staff. Therefore, this feature is used to distinguish the notes. For the soprano recorder playing part, the finger-shaped electric arm is used to press the blow hole, it will cause the situation that the speed of the score cannot be kept up. Therefore, the motor is changed to a solenoid valve to facilitate the pneumatic cylinder to smoothly press the blow hole. In addition, since the difference in pitch of the soprano recorder requires different air pressure, we increase one valve to three valves. Moreover, the range is divided into bass, midrange, and treble. Not only that, fuzzy control theory is used to control the air valve to greatly improve the sound distortion caused by the original single air valve. Experiments prove that the air valve fuzzy control combined with sheet music recognition techniques can fully realize the functions of autoplaying soprano recorder machines.

Keywords: Autoplaying Soprano Recorder Machines (ASRM), LabVIEW, Arduino, Air Valve Fuzzy Control, Pneumatic Cylinder

1 INTRODUCTION

Nowadays, robots are no longer just used in industrial production, and they can also be seen in medical and artistic fields. The International Robotic Art Competition, which started in 2016, has also been held three times. Many works created by robots have reached a level comparable to human artists. As for performance robots, there have been significant advances and improvements due to artificial intelligence. Among them, the research on music score recognition technology has been proposed in many documents. Although the music notation is limited and there are restrictions on the writing position on the staff. However, there are still many blind spots to be overcome and improved in identifying it with artificial intelligence technology, even if it is only for non-handwritten scores.

In 2018, Yang S.F. [1] proposed sheet music detection techniques applied to fluted recorder robots. The disadvantage is that the mechanism is poorly designed, which often leads to inaccurate beats. In 2004, Tsai Z.W. [2] proposed the recognition system of printed music score. Due to the addition of a variety of staff removal and note recognition methods, the overall score recognition time was extended. Lee H.W. [3] and Peng Z.X. [4] use optical symbol recognition (OMR). Since it must build a symbol database, it is possible to recognize errors as long as there are slight errors. In 2009, Lan S.M. [5] presented the musical notation recognition system for two-wheeled robot. The disadvantage is that the recognized scores are only digital scores. In 2014, Wang B.R. [6] proposed a musical score recognition system for iOS devices, because its recognition process requires human assistance, so there is no way to achieve the effect of automatic recognition.

Based on the shortcomings of the above-mentioned scholars, this article proposes air valve fuzzy control combined with sheet music recognition techniques to improve the recognition of music scores and improve the phenomenon of out of sound. Experiments prove that presented method can fully realize the functions of ASRM.

2 ASRM

To improve the problem that the finger-shaped robotic

© The 2021 International Conference on Artificial Life and Robotics (ICAROB2021), January 21 to 24, 2021

arm cannot keep up with the rhythm of the music, the motor is changed to a solenoid valve and combined with the Arduino UNO board to control the air cylinder to press the blow hole. The finished ASRM is shown in Fig. 1.

Fig.1. Finished product of the ASRM

2.1 Control Systems

The Arduino UNO board is used as the control core of the ASRM. LINX in LabVIEW Graphical Programming environment is the bridge between LabVIEW program and Arduino UNO board. Like LIFA, LINX must first write a set of communicable commands in Arduino. The difference is that LIFA can provide users with the ability to directly use Arduino connection operation in LabVIEW without writing Arduino code. The LINX kit is an open source kit developed by MakerHub. Its feature is that it cannot only be connected to Arduino, but also supported by many development platforms such as myRIO/Digilent/. Moreover, LINX can directly burn the firmware required for LINX execution on the Arduino through this LINX Firmware wizard. The control system architecture diagram is shown in Fig. 2. Fig. 3 is the Arduino UNO control board.

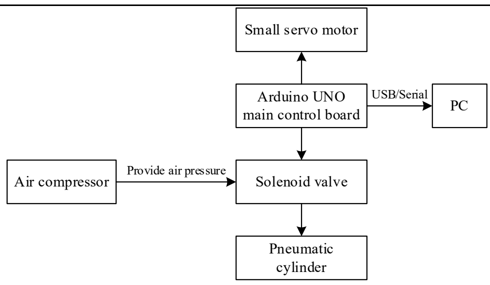


Fig.2. The control system architecture diagram

Fig.3. Arduino UNO

2.2 ASRM Platform

The physical architecture of the ASRM platform is shown in Fig. 4. The main feature is that the part of the original flute that presses the blow hole is pressed by the solenoid valve controlled pneumatic cylinder. Significantly improved the original inability to play music faster than the response speed of the motor. Not only that, the blowing control part has two more solenoid valves. It solves the problem of broken sound when blowing.

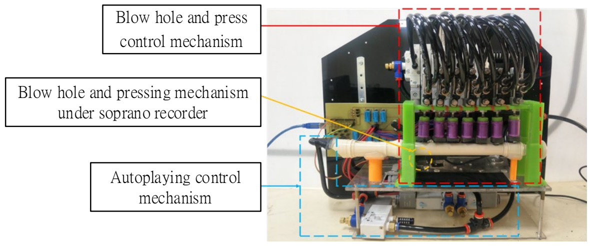


Fig.4. The physical architecture of the ASRM platform

3 MUSIC SCORE RECOGNITION SYSTEM

First, we remove the parts of the score that are not related to performance, as shown in Fig.5. The processed music score is input into the system. After that, binarization and scale recognition are performed through LabVIEW. Fig.6 is its program flow chart.

Fig.5. Score preprocessing

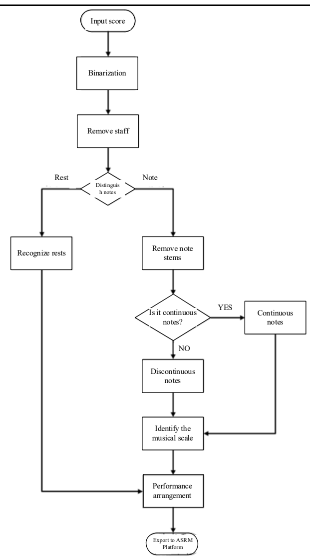


Fig.6. Program flow chart

We use the y-axis projection method to detect the staff position and remove it to replace the erosion and expansion in morphology. It was found that the notes have a specific writing style on the staff. Therefore, this feature is used to distinguish the notes. First, the note stems are distinguished and removed by the x-axis projection detection method. Second, the center point of the head of the note and the position of the staff are identified by the pixel clustering method. Finally, the encoding of the scale is arranged in sequence from bass to treble.

3.1 Binarization

Binarization is to divide the grayscale image into black and white according to the threshold set by the user. When the grayscale value of the pixel is greater than the threshold, it is judged as a white point, otherwise it is a black point. The grayscale image can be converted into a binary image through binarization, as shown in Fig.7.

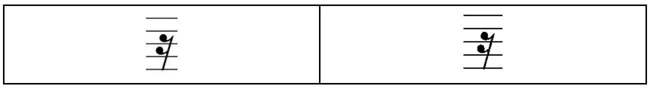
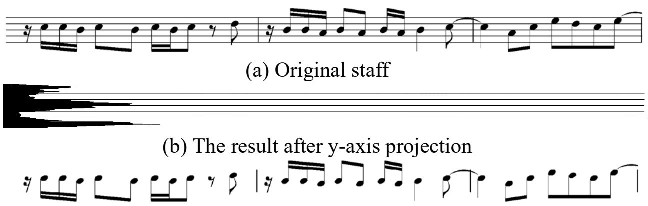


Fig.7. Difference before and after binarization

3.2 Remove staff

We use the y-projection in the orthogonal projection method to project the music score to be identified onto a single axis, which will form a graph called the projection profile, as shown in Fig.8. We can clearly find the position of the staff and remove it. This method greatly improves the original use of erosion and swelling in typology to cause unclear symbols, as shown in Fig.9.

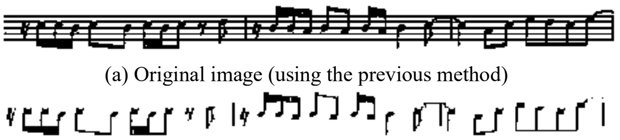


(a) Original staff

(b) The result after y-axis projection

(c) The result after removing the staff

Fig.8. The process of removing the staff



(a) Original image (using the previous method)

(b) The result after removing the staff (using the previous method)

Fig.9. The process of removing the staff(using the previous method)

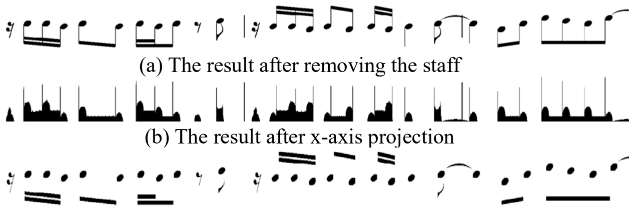
3.3 Symbol distinction

For rests, this article only deals with the commonly used full rests, two-quarter rests, four-quarter rests, eighth rests, and sixteenth rests. General notes are also processed for commonly used whole notes, half notes, quarter notes, eighth notes, and sixteenth notes. Generally speaking, the height of most rests is not greater than the height of the notes. For example, the height of a quarter note on the staff is approximately equal to the height of the staff, as shown in Fig. 10. The height of the quarter rest is only 3/4 of the height of the staff, as shown in Fig.11. Based on this, the notes are preliminarily divided into rest notes and general notes to facilitate subsequent identification.

Fig.10. Quarter note

Fig.11. Quarter rest

To distinguish between continuous notes and discontinuous notes, the image is projected on the x-axis using the x-projection in the orthogonal projection method. The note stems exceeding a certain value are removed, as shown in Fig.12. Moreover, the pixel clustering method will be used to distinguish the following three types: discontinuous notes, continuous notes, and discontinuous notes but with tails, as shown in Fig. 13.

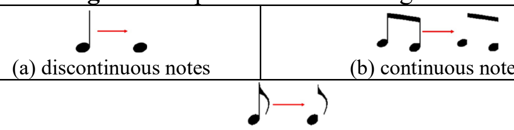


(a) The result after removing the staff

(b) The result after x-axis projection

(c) The result after removing note stems

Fig.12. The process of removing note stems



(a) discontinuous notes

(b) continuous notes

(c) discontinuous notes but with tails

Fig.13. Comparison chart after removing note stems

The pixel clustering method uses the IMAQ Count Objects 2 VI component in the Vision development module in LabVIEW. The function of this component is to cluster the binarized pixels. Let the user set the pixel threshold to distinguish the number of pixel groups. Fig.14 is the result of pixel cluster identification.

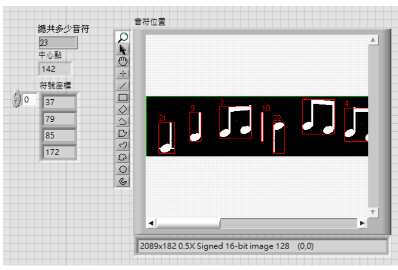


Fig.14 The result of pixel cluster identification

3.4 Scale recognition

While the scale recognition is performed, the note timing is also recognized. Therefore, the head and tail of the note must be distinguished first. First, the discontinuous notes have been separated out after removing the note stem, so there is no need to deal with it. The part with discontinuous notes but with tails is distinguished by the proportion of black pixels in the image extracted by pixel clustering. The pixel ratio calculation is based on the range enclosed by the red frame to calculate the image size and the ratio of black pixels, as shown in Fig.15. Continuous notes cannot be identified by this method because the proportions of black pixels of note stems and note heads are too close. Therefore, we use the aforementioned note stems to account for more than 2/3 of the overall note width to distinguish. The position of the beam will change due to the way the music theory is written. So we divide the continuous note from the center point into the upper and lower parts, as shown in Fig. 16. ((a) Original image with Note head at the bottom, (b) Original image with Note head at the top)

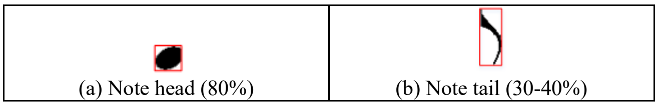


Fig.15. Distinguish of discontinuous notes but with tails

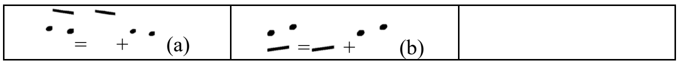


Fig.16. The process of cutting continuous notes

For judgment of discontinuous notes, the difference between discontinuous notes is whether the beam and the head are hollow. We use the ratio of black pixels in the red box to distinguish half notes from quarter notes, as shown in Fig.17. Because we have already recorded the position of the staff while removing the staff. Therefore, we only need to extract the center point of the talisman through the pixel clustering, and then compare it with the previously recorded staff position to know which line or interval the talisman is located on. Put it into the scale table of the recorder to get the scale of the note, as shown in Fig.18 and Fig.19.

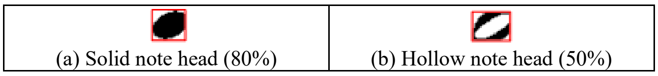


Fig.17. The solid and hollow of the note head

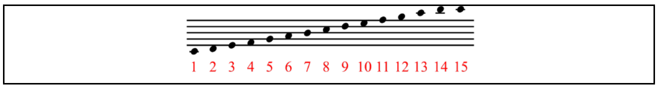


Fig.18. Scale table of the soprano recorder

3.5 Rest identification

The identification of rests is double identification using orthogonal projection method and pixel clustering. According to the appearance of commonly used rests, they are divided into three types: (1) whole rest and half rest, (2) quarter rest, (3) 8th rest and 16th rest. First, we use the pixel clustering method to capture the rest image, and use the black pixel ratio in the red frame to make a preliminary judgment, as shown in Fig.20. Second, make a detailed distinction for the above three types. (1) The image coordinates of the y-axis projection of the full rest and the bipartite rest are captured and compared with the staff position. If the image coordinates are closer to the fourth line of the staff, it is judged as a whole rest. If the coordinates are closer to the third line, it is a half rest. (2) Compare the height value captured after y-axis projection with the highest point value after x-axis projection. If the values are similar, it is judged as a quarter rest, as shown in Fig.21. (3) The 8th rest has only one peak in the y-axis projection image. The 16th rest has two peaks, as shown in Fig.22 and Fig.23.

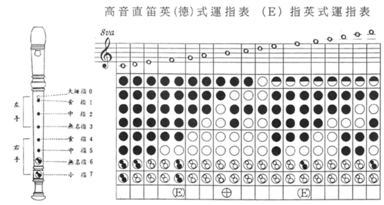


Fig.19. Scale press fingering table of the soprano recorder

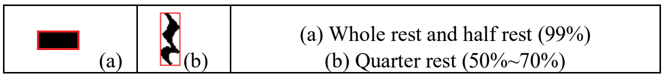


Fig.20. Distinguish between rests

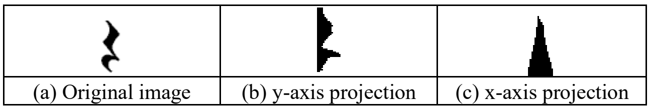


Fig.21. Quarter rest

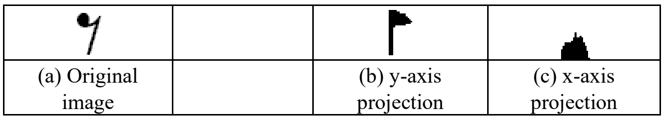


Fig.22. 8th rest

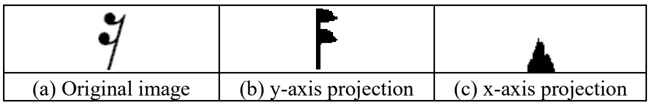


Fig.23. 16th rest

4 AIR VALVE FUZZY CONTROL DESIGN

To provide a smoother air pressure with different sound ranges, this article increases the number of solenoid valves to 3. In addition, we use the fuzzy control law to write the valve control program. The range (R) is divided into bass, midrange, and treble, as shown in Fig.24.

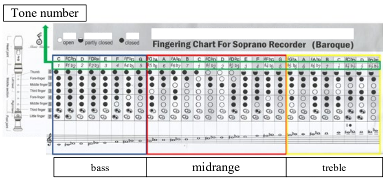
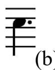


Fig. 24. The range

Define the opening degree of air pressure valve be V. TS1, TS2, and TS3 belong to Takagi-Sugeno type. Fuzzy IF-THEN rules are expressed as follows.
IF R is bass, THEN V is TS1.
IF R is midrange, THEN V is TS2.
IF R is treble, THEN V is TS3.

5 EXPERIMENTAL RESULTS

In view of the large differences in the way of notation for tuplets (Fig.25(a)) and dotted notes (Fig.25(b)) by musicians around the world, it is impossible to effectively identify all these two symbols with a general rule. Therefore, this article will ignore it during the experiment. In addition, the score used for testing is taken from a web site made by netizens and provided free of charge [7, 8]. The scores we used to test included 5 Mandarin pop songs, 2 Japanese pop songs, and 3 movie theme songs. Please refer to the following URL directly for the experimental results. https://www.youtube.com/watch?v=1Z-kPIPQG-U

 (a)	 (b)	**Fig. 25(a).** Tuplets **Fig. 25(b).** Dotted notes

6 CONCLUSION

This paper uses the pixel clustering method, the x & y axis projection method to successfully improve the score recognition results. Moreover, the finger-shaped electric arms are changed to solenoid valves to facilitate the pneumatic cylinder to smoothly press the blow hole. This method has also successfully improved the phenomenon of air leakage. Not only that, fuzzy control theory is used to control the three air valves to greatly improve the sound distortion caused by the original single air valve.

REFERENCES

[1] Yang S.F. (2018), Sheet Music Detection Techniques Applied to Fluted Recorder Robots, Master Thesis of Chienkuo Technology University.
[2] Tsai Z.W. (2004), Recognition System of Printed Music Score, Master Thesis of National Sun Yat-sen University.
[3] Lee H.W. (2016), Automatic Music Score Recognition, Master Thesis of National Tsing Hua University.
[4] Peng Z.X. (2014), Automatic Music Score Recognition System Using Digital Image Processing, Master Thesis of Chung Yuan Christian University.
[5] Lan S.M. (2009), Musical Notation Recognition System for Two-Wheeled Robot, Master Thesis of National Taiwan University of Science and Technology
[6] Wang B.R.(2014), A Musical Score Recognition System for iOS Devices, Master Thesis of National Kaohsiung First University of Science and Technology
[7] Egg Head Liu https://www.youtube.com/channel/UCC6UAv477s0YiQIfioIj0Vw (Search date: 12/15/2020)
[8] TSAI,JU-HSIN https://www.youtube.com/channel/UChE9nxA6DKQUvsJ-HBSgUJQ (Search date: 12/15/2020)