Object recognition of handwritten music score sheet using fuzzy logic

**1**

Garol, vincent ram D., Ateneo de Davao University -mondragon, kevin williams j., Ateneo de Davao University

**Abstract.** Music has always been a part of human culture, as such, music evolved together with humanity. From early centuries, music arrangements have always been written down on music score sheets, and like all treasures, is either lost, or destroyed. Researchers like Pruslin and Prerau sought to solve this problem with the help of technology. Optical music recognition, after many years has achieved great improvements, from features like pitch editing to audio playback. Although these features were great, the accuracy of the recognition of these musical symbols are below satisfactory. This study is aimed towards the improvement of the current approach of these commercial systems by using fuzzy logic as the algorithm in the object recognition phase.

General Terms:

OMR – Optical Music Recognition

OCR – Optical Character Recognition

Additional Key Words and Phrases:

Music score sheet recognition, Optical music recognition, and Fuzzy logic.

1. INTRODUCTION
   1. Background of the study

Music score sheets have been used by musicians and composers alike ever since they were invented. Some original compositions of earlier writers have degraded or have been damaged due to time and the elements and others have been lost throughout history. In dealing with the problems mentioned, early researchers have developed an application that could provide a solution. An application called “Optical music recognition (OMR)” was first developed by Pruslin and Prerau [[1](#2et92p0)]. Using this technology, scanned handwritten music score sheets can now be stored in computer systems easily. And with more improvements and features added, the OMR now has audio playback, re-editing, and storage of semantic information (notes, pitch, dynamics, etc.) capabilities [[2](#tyjcwt)], some examples of these commercially available programs are MidiScan, PianoScan, Finale, Sibelius [[3](#3dy6vkm)]. However, even with the development of the OMR, there remain some problems. The recognition of musical notations of handwritten score sheets is still not totally accurate due to the fact the there are many ways to write musical notation [[1](#2et92p0),[4](#1t3h5sf)]. Another problem lies with the quality of the paper and the symbols. Some have erasures; others are smudged. These decrease the accuracy of recognizing the musical symbols of the score sheet.

To achieve better accuracy of digitally reading these musical notations, it is necessary to develop an algorithm that can identify broken symbols. The collaborative model offers a promising solution. By integrating both with and without the staff lines in a way that symbols are segregated depending on which is recognized better (with staff lines or without), each symbol is recognized with less distortions thus boosting its recognition success rate [5]. Another recognition processes that was able to achieve better accuracy in both printed and written music score sheets are the Neural Networks and the Support Vector Machines. Despite achieving good results, both processes were still receiving low percentage recognition in some of the symbols [[4](#1t3h5sf)].

The Fuzzy logic algorithm makes use of its multi-valued decision making capabilities for classifying optical inputs. Since the OMR must classify outputs as specific musical notations, the fuzzy logic algorithm is a prime candidate for this type of implementation [[6](#4d34og8)]. Rossant and Bloch conducted a study utilizing fuzzy logic as one of the algorithms used for recognizing printed score sheets. Their output displayed significant results in recognition of various musical symbols [10]. There are many studies involving the optical music recognition some examples are [[4](#1t3h5sf),5,7,10], but, no study has ever been made for the Fuzzy logic to be used for optical music recognition of handwritten music score sheets. With these reasons, the proponents would like to pursue this study in the hope that it would open possibilities for more development and breakthroughs in OMR.

* 1. Problem Statements

This study primarily seeks to investigate a better object recognition phase in optical music recognition by using fuzzy logic algorithm. More concretely, the study would like to answer the following questions:

1. What are the different factors to be considered in identifying musical notations?
2. How can fuzzy logic be used to identify musical notations written on a music score sheet?
3. How is fuzzy logic different from other approaches?
4. What are the different factors to be considered in using fuzzy logic on OMR?
5. What approach is more suitable for object recognition in optical music recognition (OMR)?
   1. Objectives

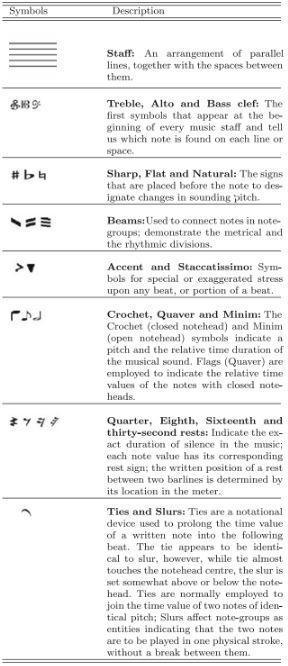
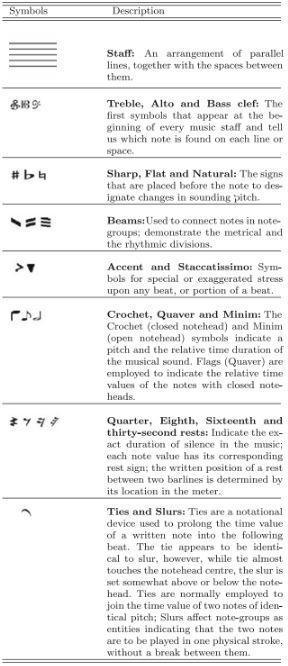
This study intends to investigate how optical music recognition (OMR) will react if used with fuzzy logic in its object recognition. In its various phases, this study hopes to:

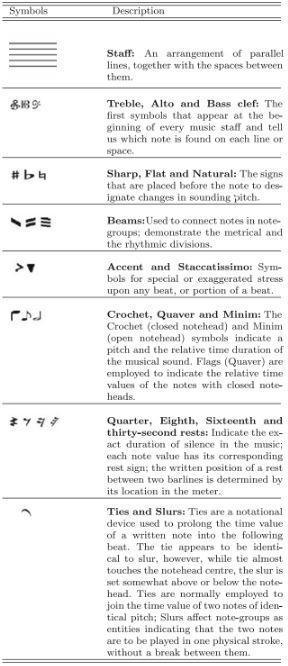
1. Discover the various factors needed in identifying musical notations.
2. Explain how fuzzy logic can be used to identify musical notations that are written on the music score sheet.
3. Find out the differences of fuzzy logic to other approaches.
4. Discover the factors that comes with using fuzzy logic on OMR.
5. Determine which approach is better suited for object recognition in OMR.
   1. Significance of the study

Optical music recognition is a very beneficial application to composers and musicians alike, since handwritten music score sheets can be damaged or destroyed. Composers would then be able to compile their works so that they last for decades to come. Moreover, OMR is a necessary tool for the seamless conversion of these music score sheets to other musical digital formats.

With these reasons, a study regarding the possible improvement of the recognition of objects using fuzzy logic will contribute to the development of the optical music recognition (OMR) approach.

* 1. Scope and limitation

This research is limited to the recognition of handwritten notations of music score sheets with printed staff lines only. This research will only utilize the symbols present on table 1. Furthermore, only the fuzzy logic algorithm will be used during the object recognition phase, all other steps of the optical music recognition will still use the traditional approaches.

**Table 1a** Music Notation taken from [[4](#1t3h5sf)]

**Table 1b** Music Notation taken from [[4](#1t3h5sf)]

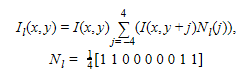
1. REVIEW OF RELATED WORKS
   1. Optical music recognition

The Optical music recognition is the application that interprets music score sheets into computer readable format from scanned images. These outputs can be used for audio playback, re-editing, and many more. The OMR is comparable to the Optical character recognition for automatic reading of text document. Although both systems are similar, the OCR approach cannot be used in dealing with musical notes and symbols [[3](#3dy6vkm)].

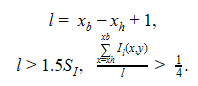
Bellini et al. pointed out that “despite the availability of some commercial OMR’s, none of these is completely satisfactory in terms of precision and reliability” [[3](#3dy6vkm)]. There are also problems regarding the preprocessing phase, and the musical notations as well. In a study conducted by Novotný and Pokorný as well as Harris and Verma, the use of binarization algorithm during the image processing stage, though it is often assessed to be the best and fastest, sometimes produces unnecessary objects and the advantages of this algorithm are not clear in the complete OMR process [[2](#tyjcwt)] [8]. These are just some of the many challenges of the Optical Music Recognition

The OMR has a few steps to reach the final output.

* Image preprocessing, where the scanned image is adjusted by utilizing different methods such as deskewing, enhancements, blurring and morphological operations, binarization, etc. Per Novotny et al. and Rebelo et al. Binarization is the most important step in the image processing phase [1,2].
* Segmentation, in which the musical symbols are isolated by detecting, measuring the staff height and spacing, and the removal of the staff lines [2]. Horizontal projection or Hough transform for the staff line detection [2] and vertical line detection for stem lines of some musical symbols [2,5,]. The study of Rossant et al [10]. they used the following equations to determine the vertical segments.



The equations above deals with geometrical and topological features of the vertical segments to be extracted. The following are what is to be solved using the equation, the label image of the vertical runs , horizontal black run-lengths , and filtered image [10].



The equations above is made to search for the longest vertical run at every y-coordinate, where and are the coordinates of the extremities. The final equation deals with skew problems. Average thickness of the segment is assessed using labelling image by considering only the pixels that are maxima (1.0) in image [10].

* Object recognition, for the classification of the music notations based on its distinct features.
* Semantic reconstruction, where all the classified music notations from the previous step is used to recreate entire music score sheet exactly as it was scanned [[2](#tyjcwt)].

These are only commonly used methods stated in the study of Novotny et al. [2], there are some researchers which use their own method for the different phases [4,5,10].

* 1. Object recognition
     1. Collaborative model

The collaborative model, researched by Pham and Lee, is a study which combines two approaches in identifying musical symbols. Recognition using the staff line and without, typically in an OMR system, the approaches is not used together, it was either one of the two [[2](#tyjcwt)]. The researchers of the collaborative model sought to do just the opposite, by applying both approaches to specific musical notations, the accuracy of the output increased, providing a very satisfactory output in recognizing the musical symbols of printed music score sheets [5].

* 1. Fuzzy logic

Fuzzy logic was created with the purpose of efficiently correcting and handling the imprecise nature of human reasoning. Lotfi A. Zadeh, a computer science professor, also known as the ‘Father of Fuzzy Logic’, created this algorithm with the idea to convert humans’ deduction skills and recreate this through a mathematical and logical way. Fuzzy logic, as a result, became a multi-valued logic which allows the given values to be defined by true or false, yes or no evaluations [[6](#4d34og8)]. But, since fuzzy logic is a multi-valued logic, it can give an output outside of these two decisions (Yes or No, True or False, 0 or 1). Using the fuzzy subset, it allows the decisions to have an ‘uncertain’ output like humans, e.g. the temperature is ‘fairly low or moderately high’ [[9](#2s8eyo1)].

A study conducted by Rossant and Bloch utilized fuzzy logic in an OMR system which recognizes printed music score sheets. Their work aimed to provide a significant increase in the accuracy of the recognition by utilizing Fuzzy model and Template matching. The reason why two recognition methods was because of the problem brought by the variability of the symbols as well as the shape imprecision, in which the fuzzy logic is adept in recognizing and classifying these types of symbols. Utilizing this equation, where is defined as the degree of possibility that S belongs to class k as an increasing function of [10].

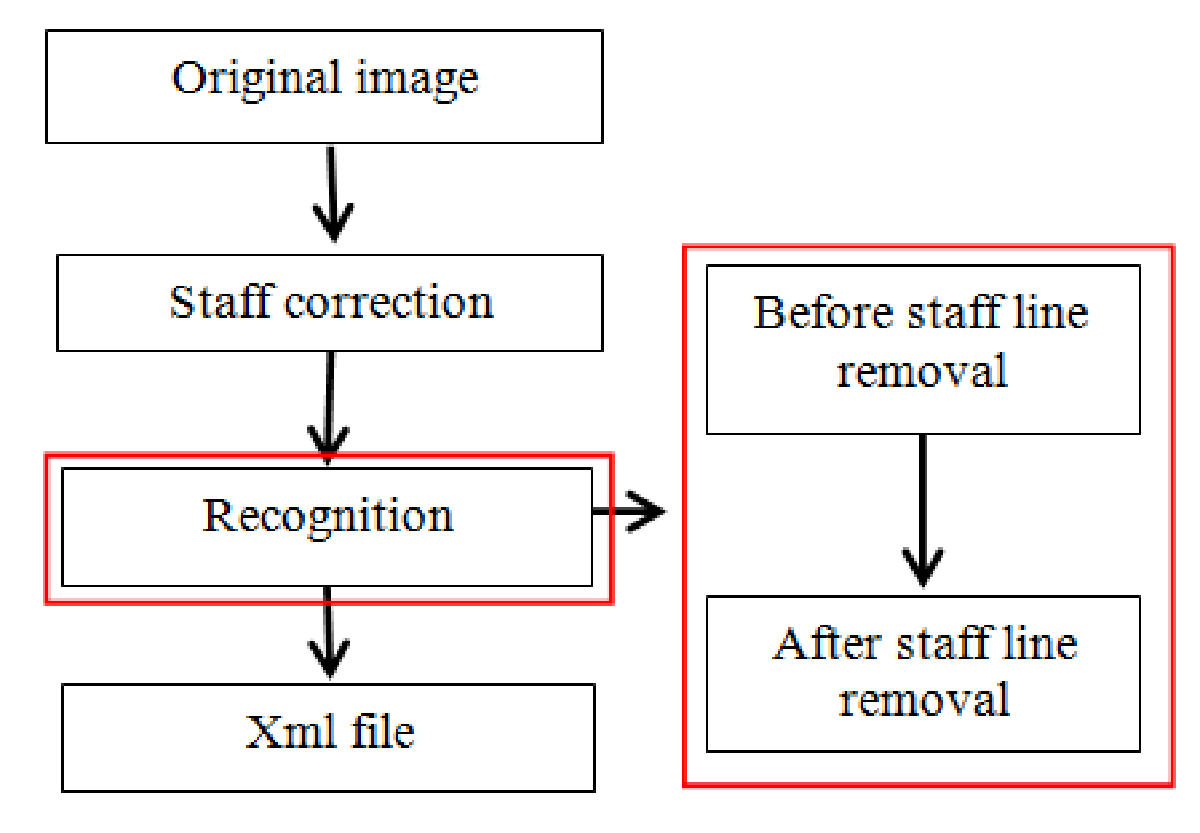


As well as this equation, where the shape of the possibility distribution is defined by and D, n(k) is the number of objects with high correlation score with model , threshold value , average value [10].



Their study using these approaches and methods led to an impressive accuracy count of the recognized music symbols, with much of the results over 100 and 99 percent [[10](#17dp8vu)].

* 1. Theoretical framework
     1. Collaborative model



**Figure 1** Conceptual framework of Collaborative Model[5]

The following framework is based on the research paper of Van Khien Pham and GueeSang Lee (2015). Wherein, the collaborative model was used to identify musical notes from an image source [5].

V. Pham and G. Lee conducted a study that the Collaborative Model is a combination of two approaches in recognizing musical symbols, recognizing with the staff line and without the staff line. The problem with recognizing with the staff line is that it has low accuracy since when the symbols are being detected the staff lines would also be included thus giving a low percentage accuracy rate. As for detecting without the staff line, when the staff line is removed the musical symbols will tend to be distorted thus also having a low percentage accuracy rate [5].

In solving the low percentage accuracy rate, the Collaborative Model uses both detection with and without the staff line. In detecting with the staff line, it uses the vertical line detection to see the left space, right space and the search range information. Basically, it is to get the actual area in where to detect/read [5].

Vertical Line Detection Algorithm:

1. Determining the information of the staff line
2. Use vertical histogram for each stave
3. Use local maximum to choose the longest line
4. Give a condition about height for vertical line

After knowing the area of detection, it now uses Musical Notes Detection since musical notes have white notes, multi white notes, black notes and multi black notes thus having an algorithm with 5 steps:

1. Create a template
2. Black note detection
3. White note detection
4. If black notes are detected, then detect multi black note using template matching
5. If white notes are detected, then detect multi white note using template matching

Music notes also have whole notes and dots. In detecting the position of whole notes, they used an algorithm which:

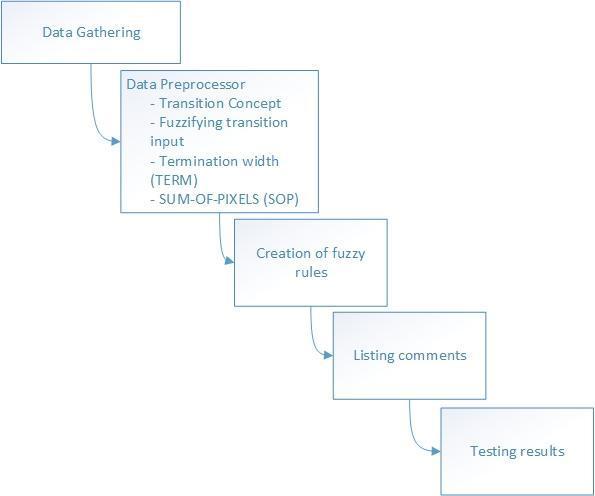
1. Determine bar lines information
2. If there is not stem line between two consecutive bar lines, then Scan horizontal projection.

Since most music notes have tails, there is also an algorithm that was used to detect those:

1. Determine the note head position,
2. If two consecutive note heads are the same bottom or top position then use RLE for detecting beam,
3. If these note heads do not have crossing thick line, then detect tail using RLE for right stem side.

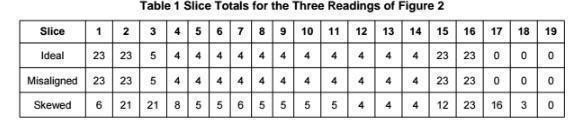
On the other hand, in detecting without the staff line, it only recognizes mainly the Dot, Rest2, Rest4, Rest8 and lastly Rest16 Symbols [5].

* + 1. Optical character recognition



**Figure 2** Visual Representation of OCA framework [[6](#4d34og8)]

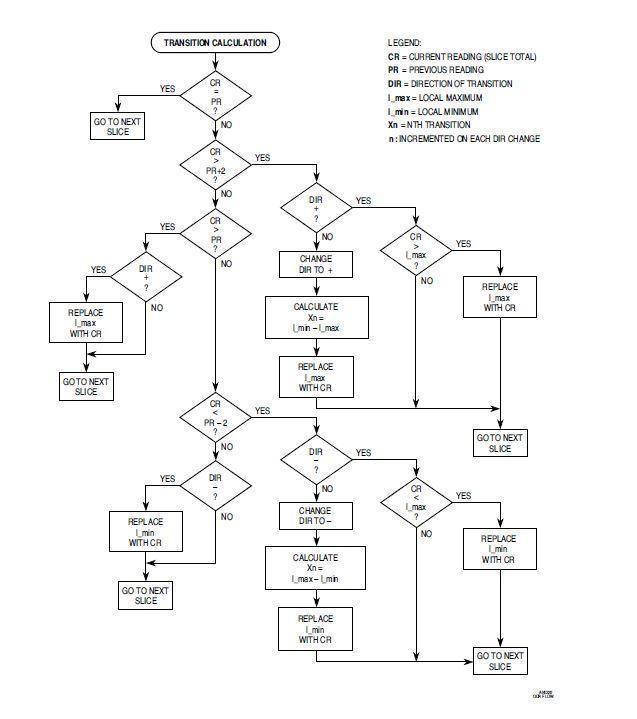
According to W. Gowan, the data is gathered by using an optical sensor called TSL 214, this sensor reads the data and converts it to analog values. Following this step is the data preprocessor phase, this phase handles the preparation of the data to create the fuzzy rules [[6](#4d34og8)].



**Table 2** Slice Totals for the Three Readings of OCR [[6](#4d34og8)]

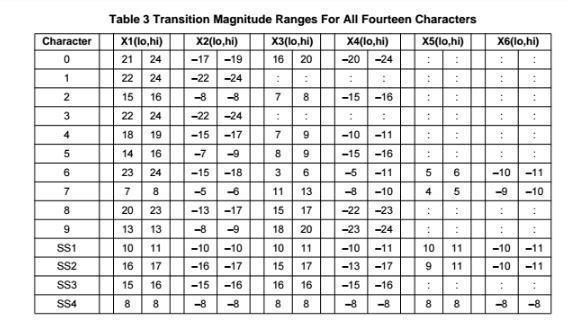
The image shown above takes the total amount of ideal, misaligned, and skewed readings of the characters. This data would then be plotted to determine the transitions or the changes of magnitude of the slice totals as shown image below. The plot shows the difference between the current local maximum (or maximum) and the previous local minimum (or maximum) [[6](#4d34og8)].

Each time the current variables of the data from the plot is updated, the data that is discarded is now fed to the transition calculation algorithm to generate the transition outputs [[6](#4d34og8)].

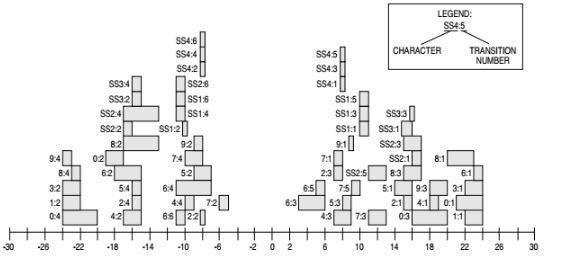


**F**

**Figure 4** The preprocessor for Transition Calculation[[6](#4d34og8)]

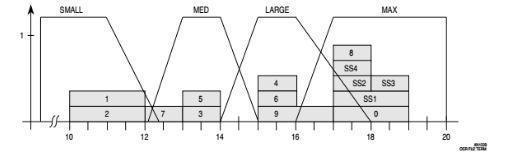
According to the researchers, the transition must be at least three (3) bits or greater magnitude to be recognized. The outputs would then be added and categorized [[6](#4d34og8)].

**Table 3** Transition Magnitude Ranges for All Fourteen Characters [[6](#4d34og8)]

This phase prepares the transition output to provide a clear visualization for the fuzzification of the data. This table was obtained by simulating each character eight (8) times, and with skews up to 5% in each direction. Using this data, the researchers established the range of possible values for the fuzzy input, this is called the universe of discourse, once this is established fuzzy sets can now be created [[6](#4d34og8)].

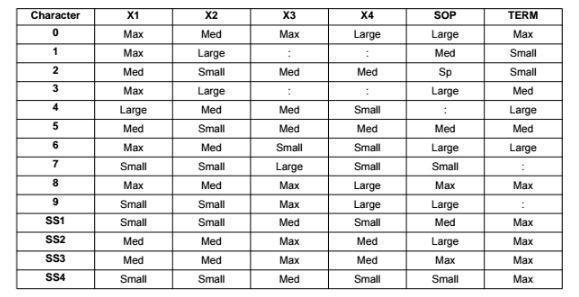
**Figure 5** Distributions of Transitions Across the Universe of Discourse [[6](#4d34og8)]

This graph represents the universe of discourse, derived from the image above, this graph can be used to create the fuzzy sets that is applicable to all six (6) transition inputs. Although, this graph can be further enhanced to increase the precision of the character recognition. Using SUM-OF-PIXELS and TERMINATION WIDTH (TERM), the graph above is now arranged to a clearer table where we can see each character's precise definition [[6](#4d34og8)].



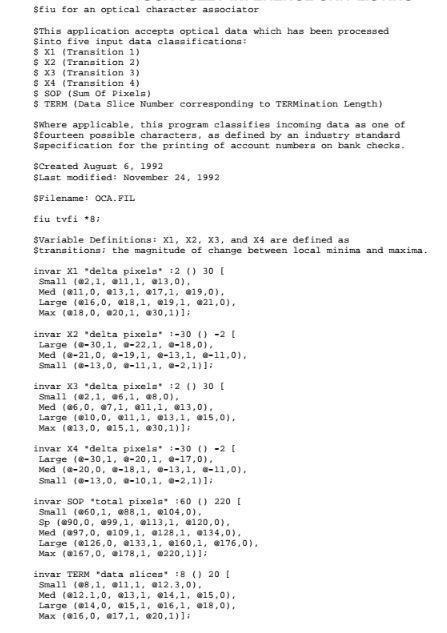
**Figure 6** TERM Fuzzy Set Definitions [[6](#4d34og8)]

From the tables and data that was created by the SOP, TERM, and Transition magnitudes, the researchers summarized all three to one table [[6](#4d34og8)].



**Table 4** OCA Fuzzy Set Associations [[6](#4d34og8)]

Using this, the researchers created the fuzzy rule set with ease. The following images will show the fuzzy rules that was derived from the table above. [[6](#4d34og8)].

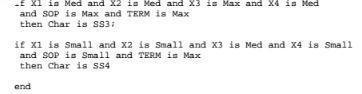


**Figure 7a** Fuzzy Rule of OCA [[6](#4d34og8)]





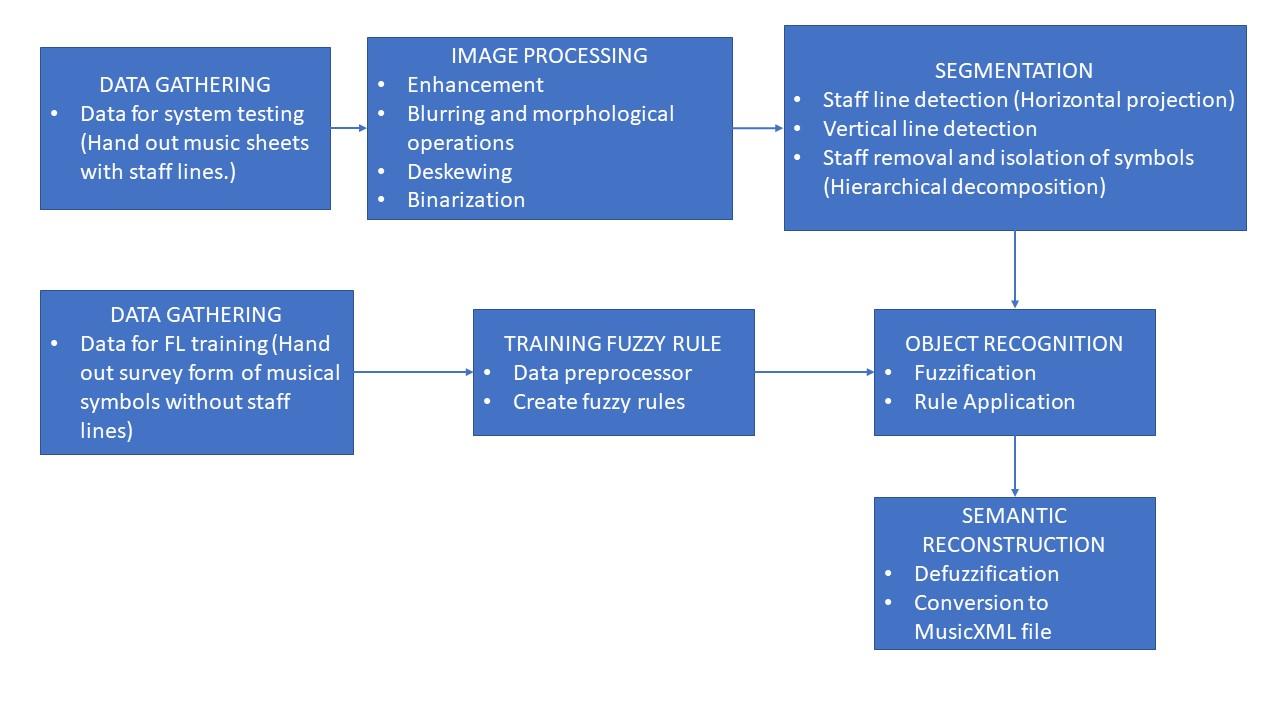
**Figure 7b** Fuzzy Rule of OCA [[6](#4d34og8)]



**Figure 7c** Fuzzy Rule of OCA [[6](#4d34og8)]

Finally, using Fuzzy inference development environment (FIDE), the fuzzy inference unit (FIU) is generated which holds the source code for the range and the input and the output functions, and define the fuzzy rule set [[6](#4d34og8)].

1. RESEARCH DESIGN AND METHODOLOGY
   1. Conceptual framework



**Figure 8** Architecture of Framework

* 1. Data gathering

As this research is aimed to the improvement of handwritten music score sheets, gathering of data will have two (2) parts. First, the researchers will hand out papers, surveys, containing printed music symbols and boxes with staff lines inside for the students to write/copy the musical symbols for the training of data in fuzzy logic and for the testing if the written musical symbols match the printed symbol. The second part will gather data for testing of the system. This will be done by printing staff lines and manually write the musical symbols. Using various penmanship, the aim of making the model/system learn using comparing the handwritten symbol to the printed ones will be possible with high accuracy rate.

* 1. Training Fuzzy rules

The researchers will follow the steps taken from the research of W. Gowan. The data gathered will then be used for the data preprocessing phase and will be converted to its transition values similar to the OCA approach. After the transition, the values would then be used to measure the universe of discourse and be further simplified using the SUM OF PIXELS and the TERM method. Lastly, with output, the researchers will use the values to create the fuzzy rules using Fuzzy Inference System professional (FISpro) to be used during the object recognition phase [[6](#4d34og8)][15].

* 1. Image processing

This step turns the images to gray, preparing it for adaptive threshold transforming the image to binary image. With this the removal of staff lines are done easier. The process used in this staff line removal was the one that Theodore T. made that was based on Morphological operations. It relied on relative ordering of pixel values. This technique uses a template called Structuring element that is a small matrix of pixels, each with a value of either one or zero. The size of the structuring element is being specified by the matrix dimensions, its shape is specified by the pattern of ones and zeros and its origin can be found in one of its pixels.

The Structuring element is said to be “fit” if all of its pixels are set to 1. It would be a “hit” if at least 1 of its pixels contain a 1. Lastly, if none of its pixels contain a 1, it would be ignored.

Morphological operation mainly uses erosion and dilation. The said two are used in order to make the scanned image clear or distinct. Erosion is used in order to shrink the size of the binarized image making it thinner by cutting out layer of pixels, whereas dilation is the opposite of erosion. It adds a layer of pixels to both the inner and outer boundaries of regions making it bigger/thicker.

After the extraction of staff lines by the use of the said operation, the symbols are now enhanced by extracting the edges making it smooth and by the use of blurring in order to achieve even smoother symbols

* 1. Segmentation

The data will then be subjected to segmentation in preparation for its recognition. This phase was done by the use of photoshop and manual cropping with the size of approximately 8x10 cm.

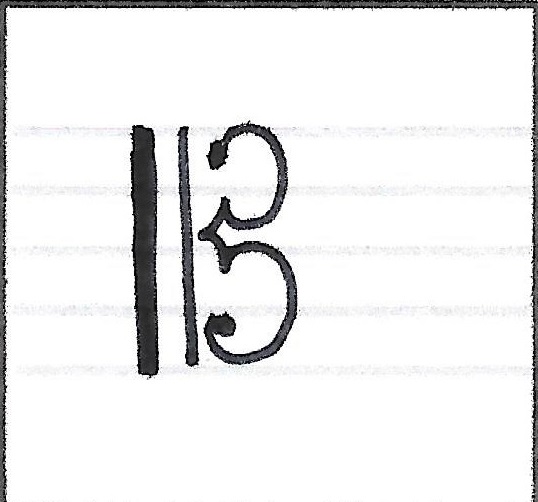
* 1. Object recognition

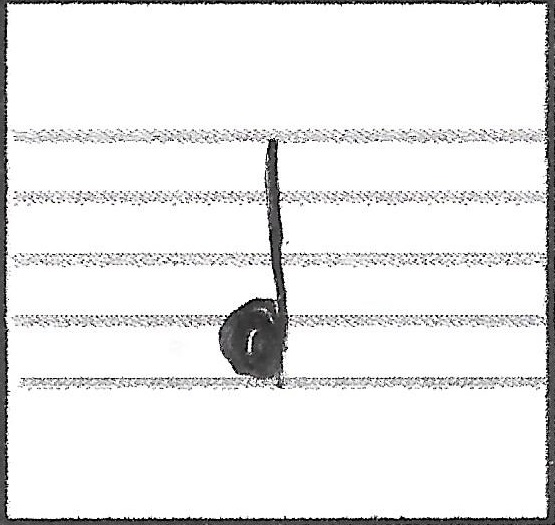
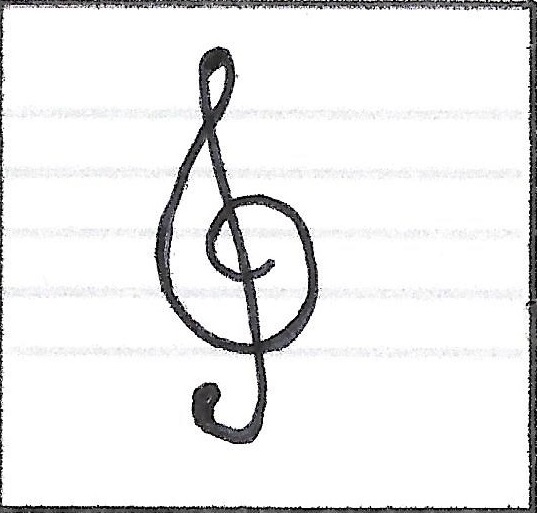
Following W. Gowan’s study. The processed data from the previous step will then be converted from conventional values to linguistic values and is assigned to a fuzzy set, this method is known as fuzzification. Following this step is the rule application, this is when the fuzzified data are evaluated against the fuzzy rules written for this system. The last step of the object recognition phase will be the generation of specific values based on the rule strengths that came from the previous step [[6](#4d34og8)].

1. RESULTS AND DISCUSSION
   1. Data Gathering of musical symbols

Data gathering was done in two parts. The first data gathering was conducted to be used for the creation of fuzzy rules. The survey was done within the Ateneo de Davao University Jacinto campus, with over ten (10) respondents both with and without musical background. This was to ensure data on beginners were included, and taken into account. The survey form consists of twenty (20) different symbols with five (5) separate boxes with staff lines. The second survey was done by handing out blank sheets of paper with staff lines, and another sheet which is a printed musical arrangement containing the 20 symbols with which a handful of respondents will fill out.

The survey was then scanned using an HP Deskjet Ink Advantage 3545 with an image resolution of 600 DPI and saved as JPEG format.



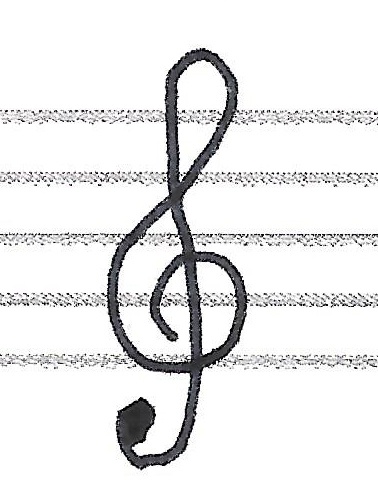
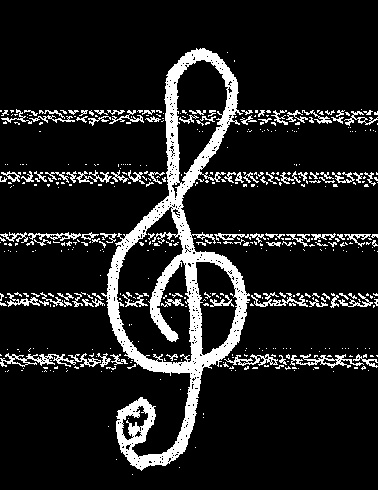
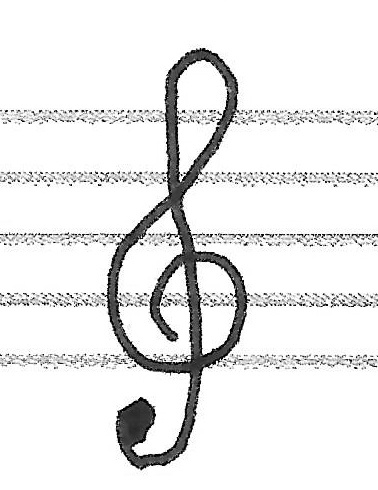


**Figure 8a** Treble **Figure 8b** Alto **Figure 8c** Crochet

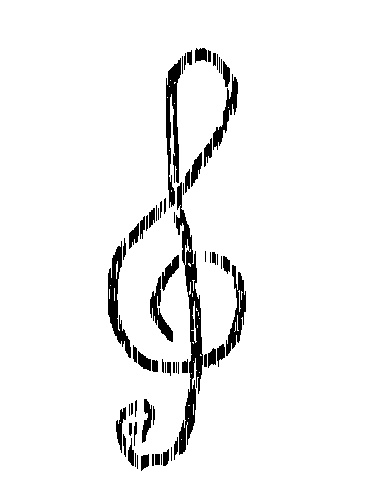
* 1. Image preprocessing
     1. Symbol Segmentation

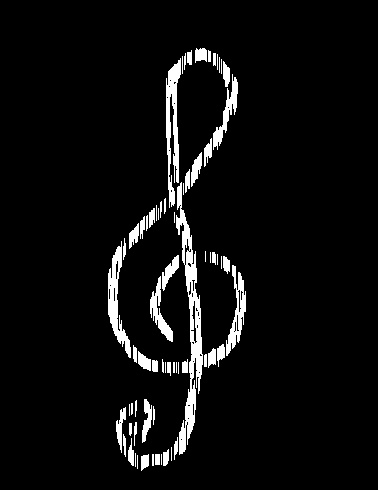
The survey sheets that were scanned, both the ones for the creation of fuzzy rules and for the testing, have staff lines. Removal of the said lines is deemed advisable since doing so, would make the recognizing of the musical notes easier. However, before the removal of staff lines, the image would first be transformed into gray in preparation for the application of adaptive threshold transforming the gray image to binary image making morphological approach in staff line removal possible.

The staff line removal approach that the researchers used was the one founded by Theodore T. which was based on morphological operations which relied on the relative ordering of pixel values. This technique examines an image with a small template called Structuring element and is positioned in all possible locations in the image. This is done to enhance image quality after the staff line removal. After the said process, the image would now be inverted making it look similar to its original form but without the staff line. In order to make the symbol look smoother, adaptive threshold was reapplied, dilation of edges was done, edges were extracted, blurs were applied.

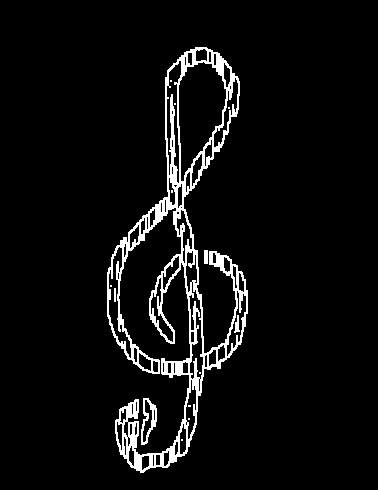
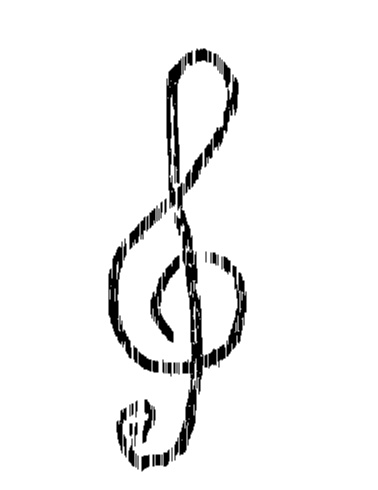
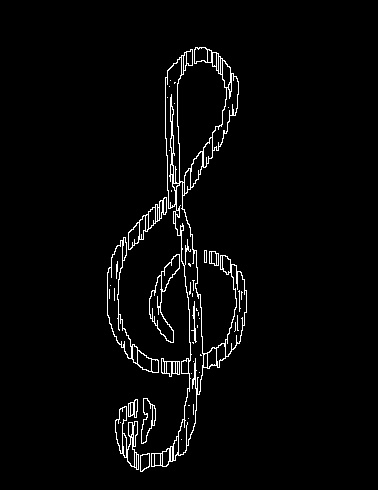


**Figure 9a** w/ Staff line **Figure 9b** Gray Image  **Figure 9c** Adaptive Threshold





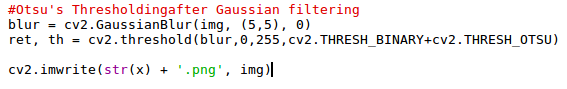
**Figure 9d** Staff line removal **Figure 9e** Inverse Image



**Figure 9f** Adaptive Threshold **Figure 9g** Dilation of Edge **Figure 9h** Smooth Final Result

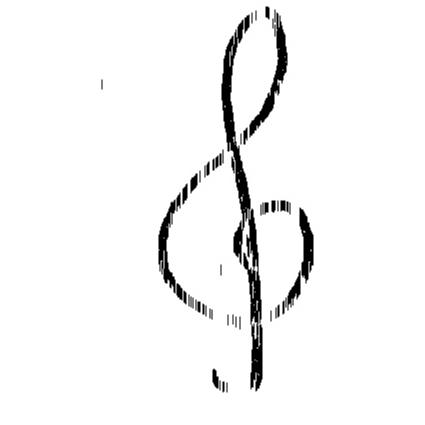
Following the staff line segmentation, instead of seeking out for a certain framework that’s supposedly auto-detects and segments the symbol, manual cropping was done in place of the latter by the use of photoshop. Each symbol was cropped by approximately 8x10 cm.

* + 1. Image Binarization

After the symbols were segmented and cropped out, each individual images of the symbols were binarized and saved as PNG format. This was done by using OTSU’s method of binarization utilizing the OpenCV package on python. Using the advantage of OTSU’s method on bimodal images, the result is an accurate threshold.

**Figure 10** Binarization Code

The code snippet was taken from the OpenCV website [16]. The code has two parts, the blurring of the image which removed the background noise, then the OTSU’s thresholding with a minimum and maximum of 0 and 255 respectively for the global thresholding. After the Binarization was applied it was then saved as a PNG image. The images below show the symbol before and after the process was done.



**Figure 11a** Original Treble **Figure 11b** Binarized Treble

* + 1. Canny edge detector

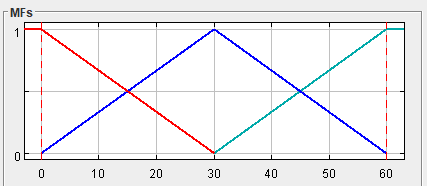
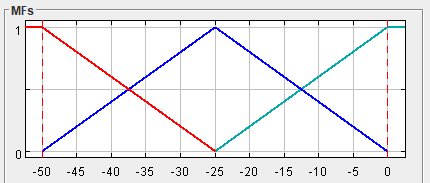
Following the image preprocessing, the next step is the feature extraction. The Proponents chose this feature extractor due to the accurate detection of edges within an image. The symbols that were fed to the Edge detector and its binary string was extracted from it. Utilizing the transition calculation found in W. Gowan’s research [6], the following results were acquired.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Symbol name** | **X1 (LO & HIGH)** | | **X2 (LO & HIGH)** | | **X3 (LO & HIGH)** | | **X4 (LO & HIGH)** | | **X5 (LO & HIGH)** | |
| **8th Rest** | -15 | -3 | 10 | 24 | -13 | -5 | 15 | 25 | -11 | 4 |
| **16th Rest** | -23 | -12 | 17 | 28 | -13 | -4 | 17 | 28 | -13 | -6 |
| **32nd Rest** | -16 | -11 | 14 | 27 | -14 | -6 | 11 | 25 | -10 | -3 |
| **Accent** | -12 | -7 | 10 | 18 | -8 | -5 | 11 | 21 | -10 | -7 |
| **Alto** | -29 | -20 | 32 | 42 | -14 | -5 | 11 | 21 | -10 | -7 |
| **Bass** | -24 | -11 | 16 | 26 | -15 | -9 | 18 | 29 | -12 | -3 |
| **Beam** | -40 | -21 | 41 | 2 | -11 | -4 | 41 | 59 | -21 | -11 |
| **Beam 2** | -14 | -8 | 14 | 23 | -13 | -5 | 20 | 30 | -11 | -3 |
| **Beam 3** | -12 | -6 | 6 | 13 | -16 | -9 | 15 | 27 | -9 | -5 |
| **Crochet** | -13 | -5 | 7 | 13 | -11 | -2 | 15 | 22 | -10 | -5 |
| **Crochet 2** | -14 | -8 | 15 | 24 | -8 | 3 | 14 | 25 | -9 | -3 |
| **Flat** | -5 | -3 | 3 | 19 | -12 | -3 | 18 | 26 | -13 | -4 |
| **Minim** | -13 | -5 | 10 | 17 | -11 | 3 | 17 | 24 | -7 | -4 |
| **Natural** | -9 | -3 | 3 | 13 | -17 | -8 | 22 | 36 | -13 | -3 |
| **Quarter Rest** | -15 | -5 | 12 | 24 | -15 | -5 | 11 | 22 | -18 | -7 |
| **Quaver** | -12 | -7 | 11 | 23 | -7 | -3 | 12 | 20 | -7 | -3 |
| **Sharp** | -13 | -6 | 7 | 15 | -15 | -9 | 19 | 33 | -16 | -8 |
| **Slur** | -10 | -6 | 3 | 11 | -6 | -3 | 3 | 17 | -8 | -5 |
| **Staccatissimo** | -11 | -5 | 10 | 18 | -8 | -4 | 11 | 20 | -9 | -3 |
| **Tie** | -12 | -7 | 5 | 11 | -10 | -4 | 7 | 13 | -19 | -11 |
| **Treble** | -13 | -8 | 12 | 20 | -12 | -3 | 17 | 22 | -13 | -4 |

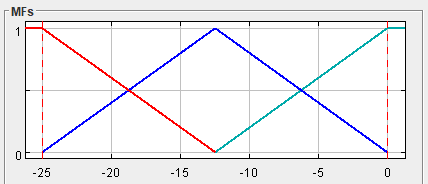
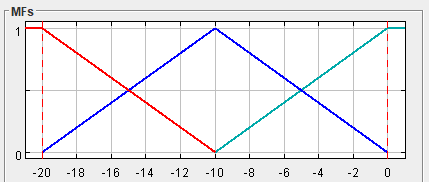
**Table 5** Transitions

* 1. Creation of membership function

The data acquired from the transition calculations was plotted and used as basis for the fuzzy membership function which is also called “the Universe of discourse” (the range of possible values for fuzzy input) [6]. This research utilized a Fuzzy Inference System named “Fuzzy Inference System Profession” (FISpro), this is an open source software created by Serge Guillaume and Brigitte Charnomordic [15]. FISpro allows the users to create fuzzy inference systems with the help of a Graphical User Interface.

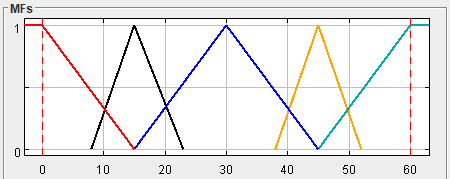
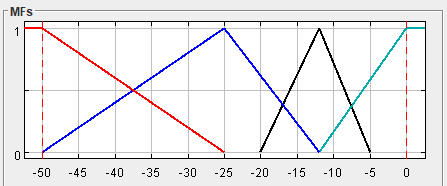
After defining the inputs and their respective ranges, the proponents used the automatic membership function (automf) available within the FISpro program to create the membership labels.

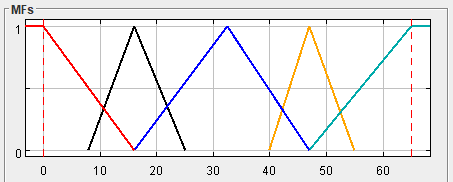
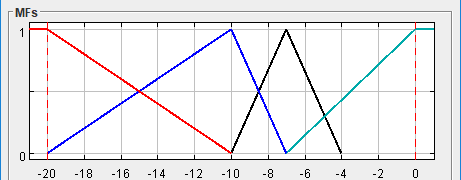
**Figure 12a** X1 Universe of discourse  **Figure 12b** X2 & X4 Universe of discourse

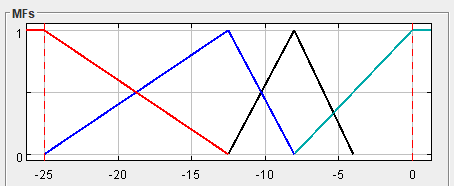


**Figure 12c** X3 Universe of discourse **Figure 12d** X5 Universe of discourse

The membership functions that were generated had some symbols that overlaps each other due to their values being near or is within the values of other symbols. Additionally, there were values found throughout the transitions that could not be recognized due to it being in between two membership function. To resolve the latter problem, the proponents manually added a new function to the existing functions, this was also done by W. Gowan in his research to resolve this particular problem [6].

**Figure 13a** X1 Revised Membership Function **Figure 13b** X2 Revised Membership Function

**Figure 13c** X3 Revised Membership Function **Figure 13d** X4 Revised Membership Function



**Figure 10e** X1 Revised Membership Function

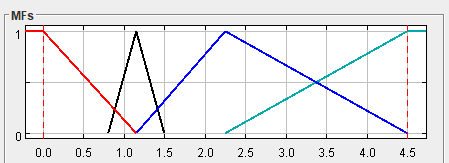
Based on the data from the table #, the proponents systematically added new membership functions to satisfy the values that could not be read. By adding these new membership functions, all values now has a specific membership function in which it is included, as well as decreased some of the conflicts found previously.

* 1. Sum of Pixels (SOP)

With the conflicts found in table #, the researchers found it necessary to include Sum of Pixels (SOP) [6] to resolve the conflicts that were found. The SOP reads the total number of white pixels of an image by counting through the binary string which was acquired during the Canny feature detector stage.

|  |  |  |
| --- | --- | --- |
| **Symbol Name** | **SOP (LO)** | **SOP (HIGH)** |
| **8th Rest** | 400 | 700 |
| **16th Rest** | 950 | 1200 |
| **32nd Rest** | 1350 | 1750 |
| **Accent** | 440 | 800 |
| **Alto** | 3400 | 4100 |
| **Bass** | 1100 | 1600 |
| **Beam** | 800 | 1100 |
| **Beam 2** | 1400 | 2000 |
| **Beam 3** | 2100 | 2500 |
| **Crochet** | 800 | 1200 |
| **Crochet 2** | 900 | 1100 |
| **Flat** | 500 | 900 |
| **Minim** | 700 | 1000 |
| **Natural** | 800 | 1000 |
| **Quarter Rest** | 1500 | 2100 |
| **Quaver** | 1300 | 1700 |
| **Sharp** | 950 | 1300 |
| **Slur** | 850 | 1050 |
| **Staccatissimo** | 100 | 500 |
| **Tie** | 900 | 1200 |
| **Treble** | 2700 | 3100 |

**Table 6** SOP



**Figure 14** SOP Universe of discourse

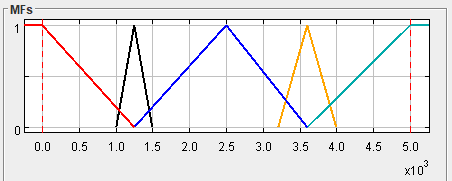
With the SOP data acquired, it was then plotted to the universe of discourse and generated its corresponding membership function using Automf and created an additional membership function to resolve the values that could not be read. SOP helped resolve some of the conflicts that was found, but not all.

* 1. Termination Width (TERM)

To resolve the remaining conflicts, the proponents used TERM [6] which calculated for the width of a symbol by determining which binary string has the most white pixels. This process is called Termination width (TERM) [6].

|  |  |  |
| --- | --- | --- |
| **Symbol Name** | **TERM (LO)** | **TERM (HIGH)** |
| **8th Rest** | 1300 | 1394 |
| **16th Rest** | 1540 | 1540 |
| **32nd Rest** | 1975 | 2149 |
| **Accent** | 701 | 783 |
| **Alto** | 4170 | 4742 |
| **Bass** | 2169 | 2595 |
| **Beam** | 1812 | 1915 |
| **Beam 2** | 2286 | 2371 |
| **Beam 3** | 2173 | 2443 |
| **Crochet** | 1250 | 1360 |
| **Crochet 2** | 1061 | 1175 |
| **Flat** | 860 | 1039 |
| **Minim** | 1589 | 1589 |
| **Natural** | 746 | 1066 |
| **Quarter** | 1500 | 1900 |
| **Quaver** | 1709 | 1982 |
| **Sharp** | 1252 | 1766 |
| **Slur** | 1040 | 1241 |
| **Staccatissimo** | 1930 | 1930 |
| **Tie** | 1111 | 1348 |
| **Treble** | 3452 | 3858 |

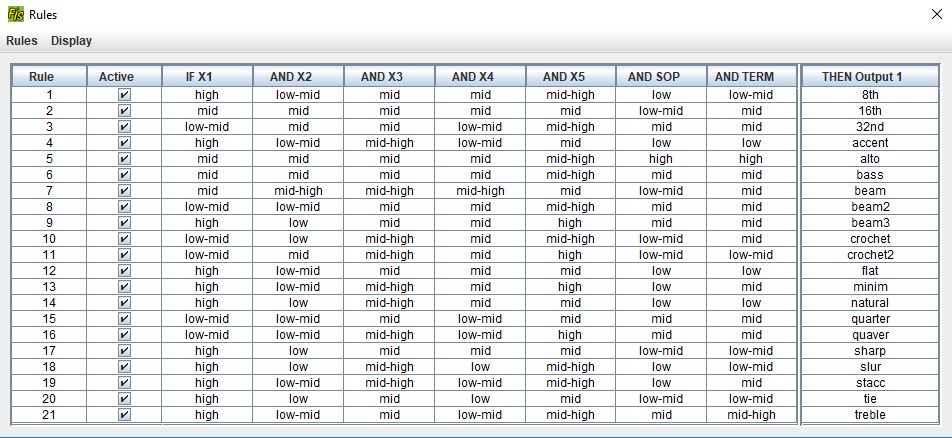
**Table 7** TERM

****

**Figure 15** TERM Revised Membership Function

With the conflicts resolved, and the addition of new membership functions to satisfy the values that could not be read. The data was then used for the creation of rules.

* 1. Fuzzy rule creation

After all the conflicts were resolved, the proponents created the rules of each twenty-one (21) symbols according to the data gathered. Creating the rules was very simple and efficient due to the FISpro environment.

**Figure 16** Rules

These rules can also be written with the basic IF-ELSE statements using any Fuzzy Inference Language (FIL) [6].

* 1. Results

The accuracy results were acquired by using 20 images per symbol and feeding it to the FISpro Inference module.

|  |  |
| --- | --- |
| **Symbol Name** | **Accuracy (%)** |
| **8th Rest** | 30 |
| **16th Rest** | 5 |
| **32nd Rest** | 20 |
| **Accent** | 58.3 |
| **Alto** | 35 |
| **Bass** | 25 |
| **Beam** | 40 |
| **Beam 2** | 0 |
| **Beam 3** | 86.7 |
| **Crochet** | 10 |
| **Crochet 2** | 25 |
| **Flat** | 15 |
| **Minim** | 30 |
| **Natural** | 15 |
| **Quarter Rest** | 5 |
| **Quaver** | 30 |
| **Sharp** | 5 |
| **Slur** | 60 |
| **Staccatissimo** | 35 |
| **Tie** | 20 |
| **Treble** | 35 |
| **Unknown** | 71.4 |
| **TOTAL AVERAGE PER SYMBOL** | 27.86 |

**Table 8** Results and accuracy

Based on the table shown above, the accuracy of the majority of the symbols is not satisfactory. The factors that were found were the raw data itself, since the style of writing differ from person to person, some of the values’ range were too large thus creating a huge gap where some could not be read or had values from two membership functions.

Following the first factor is the staff line segmentation, with the removal of the staff lines, the symbols were left destroyed and almost unreadable. Additionally, there were still signs of the remains of the staff lines present in some images, this was due to the reason of keeping the symbols as intact.

Lastly, the fuzzy membership functions. With the revision of the generated membership functions, the universe of discourse was segregated into parts, although this proved to be effective in resolving the rule conflicts, this was not the case with how the rules were applied. One of the main factors has to do with how these rules were set up.

1. CONCLUSIONS AND RECOMMENDATIONS
   1. Conclusions

The researchers conclude that with the present result, using fuzzy logic as a means to recognize handwritten musical symbols is still far from being a passable algorithm. The factors involving the accuracy of the symbol recognition namely:

* The raw numeric data extracted from the image.
* The Segmentation of the staff lines.
* The fuzzy membership functions.

These three (3) factors must be further explored and researched before Fuzzy logic can become a mainstream algorithm for recognizing handwritten musical symbols. The accuracy of the results is still quite far behind compared to machine learning algorithm found in [4].

Regardless of the less than satisfactory results, this research is helpful to future researchers whose topic include object recognition in either alpha-numeric characters or symbols or for researchers whose objective will be optimizations and enhancements. The researchers also conclude that fuzzy logic can be used for recognition of objects and characters. This has been shown by the symbols “beam 3, Slur, and Accent”.

* 1. Recommendations

The researchers recommend using the Support Vector Machine (SVM) algorithm which was researched by A. Rebelo et.al. [4]. Their results and findings are noteworthy and highly recommended due to their near perfect recognition rate. Additionally, a research on which feature extractor is more suitable with extracting the key points of a musical symbol, regardless of size, and orientation, as well as a research on the best staff line segmentation.

# References

|  |  |
| --- | --- |
| 1 | Rebelo, Ana, Fujinaga, Ichiro, Paszkiewiez, Filipe, Marcal, Andre R.S., Guedes, Carlos, and Cardoso, Jaime S. *Optical Music Recognition - State-of-the-Art and Open Issues*. Int J Multimed info Ret, 2012. |
| 2 | Novotný, Jiří and Pokorný, Jaroslav. *Introduction to Optical Music Recognition: Overview and Practical Challenges*. Charles University, Prague, 2015. |
| 3 | Bellini, P., Bruno, I., and Nesi, P. An Off-line Optical Music Sheet Recognition. In *Visual Perception of Music Notation: On-line and Off-line Recognition*. Department of Systems and Informatics, University of Florence, Florence. |
| 4 | Rebelo, A., Capela, G., and Cardoso, Jaime S. *Optical Recognition of music symbols*. Springer-Verlag, 2009. |
| 5 | Pham, Van Khien and Lee, GueeSang. Music Score Recognition Based on a Collaborative Model. *International Journal of Multimedia and Ubiquitous Engineering*, X (2015), 379-390. |
| 6 | Gowan, William A. *Optical Character Recognition using Fuzzy logic*. Freescale Semiconductor, Inc, Retrieved on December 16, 2016. |
| 7 | Kang, Tim and Perry, Michaels. *Optical Music Recognition without staffline removal*. University of California, San Diego, Retrieved on December 8, 2016. |
| 8 | Harris, Sevy and Verma, Prateek. *Sheet Music Reader*. Department of Electrical Engineering, Stanford University, Stanford, CA, Retrieved on December 08, 2016. |
| 9 | Hellmann, M. *Fuzzy Logic Introduction*. Universite de Rennes, Rennes Cedex, Retrieved on February 7, 2017. |
| 10 | Rossant, Florence and Bloch, Isabelle. Robust and Adaptive OMR System Including Fuzzy Modeling, Fusion of musical rules, and Possible error detection. *EURASIP Journal on Advances in Signal Processing* (2007), 25. |
| 11 | Nehab, Diego. *Staff line Detection by Skewed Projection*. 2003. |
| 12 | Calvo-Zaragoza, Jorge, Vigliensoni, Gabriel, and Fujinaga, Ichiro. *Staff-Line Detection on Greyscale images with pixel classification*. Retrieved on December 20, 2016. |
| 13 | Garrido, Angel. *A brief history of Fuzzy logic*. BRAINotes, Madrid, Retrieved on January 12, 2017. |
| 14  15  16 | Jäkel , Jens and Bretthauer, Georg. Fuzzy system application. In *Control systems, Robotics and automation*. Encylopedia of Life Support Systems, Forschungszentrum Karlsruhe, Retrieved on February 7, 2017.  Serge Guillaume and Brigitte Charnomordic, Learning interpretable Fuzzy Inference Systems with FisPro, International Journal of Information Sciences, [doi:10.1016/j.ins.2011.03.025](http://dx.doi.org/10.1016/j.ins.2011.03.025) , 181(20), 4409-4427, 2011.  Doxygen. Image Thresholding (2016, December). Retrieved from <http://docs.opencv.org/3.2.0/d7/d4d/tutorial_py_thresholding.html> |

# 

# Notes from Proposal Defense

1. Fuzzy Logic on OMR to increase accuracy - finished (background of the study)
2. Have you looked into other algorithms? – finished ( found on background of the study and RRL)  
   SVM was used previously--not 100% accurate on handwritten – finished ( found on B.Study)
3. Why not compare different algorithms?  
   Why did you conclude that FL is the best way? - Finished ( found on B.study)  
   Why not compare FL and ANN? Because FL hasn’t been tested for handwritten music score sheet.

FL promise--user intervention to make it learn better

======================= SAME SA COMMENTS NG PANELISTS ===============================

1. Include in Methodology, comparing the printed with the students’ copy Finished (methodology section 3.7)
2. Expand data gathering  
   --include the steps in data gathering in the figure FINISHED (methodology 3.1)
3. Conceptual Framework  
   --cite sources and which parts? FINISHED (methodology 3.3 , 3.5, 3.6)

E-mail from Chair Panelist

- Include in the RRL the flowchart of the Music score sheet OMR as base on previous studies. FINISHED (found in the theoretical framework)

- Clearly state inside the box of Data Gathering under conceptual framework the data gathering techniques you will use. FINISHED (Found in section 3)

- Indicate where data gathering - > image proc -> segmentation came from in RRL (c/o Sir Ogs) FINISHED (RRL section 2.4.1)

- Identify/state in RRL what was the exact equation used by related researches to further 'expand' their segmentation. (c/o Sir Ogs) - FINISHED (RRL section 2.1)

- Part of your analysis phase is to compare the before and after the correction of the results based on the manual intervention.

Patrick Angelo P. Paasa

Faculty - Computer Studies

Ateneo de Davao University