Pick-up the Musical Information from Digital Musical Score Based on Mathematical Morphology and Music Notation

Chen Genfang
School of Information Science and Engineering
Hangzhou Normal University
Hangzhou, China
e-mail: cgf@hznu.edu.cn

Zhang Wenjun and Wang Qiuqiu
School of Film & TV Arts and Technology
Shanghai University
Shanghai, China
e-mail: wjzhang@mail.shu.edu.cn

***Abstract*—The basic rule of musical notation for image processing is analyzed, in this paper. Using the structuring elements of musical notation and the basic algorithms of mathematical morphology, a new recognizing for the musical information of digital musical score is presented, and then the musical information is transformed to MIDI file for the communication and restoration of musical score. The results of experiment show that the statistic average value of recognition rate for musical information from digital musical score is 94.4%, and can be satisfied the practical applied demand, and it is a new way for applications of digital library, musical education, musical theory analysis and so on.**

Keywords-Musical score; Image recognition; Mathematical morphology; Music notation; MIDI

I. INTRODUCTION

There are two communication carries for music: paper musical scores (include: Conventional Music Notation-CMN, Numbered Musical Notation, Chinese Tradition Gong-Chi Notation, Chinese Tradition Abbreviated Character Notation and so on), musical acoustic (include: Tape, CD, TV, Perform, Network Play and so on). The two communication carries usually didn't synchronous spread, and a great lot of music works were recorded by paper musical scores; but music is an aural media, the musical acoustic was easily understood by mass, so it is the primary communication carry for music.

How does a great lot of paper musical scores transform into musical acoustic? OMR (Optical Music Recognition) uses synthetically some techniques which include image processing, patten recognition, AI, MIDI and so on, it can transform paper musical scores into musical acoustic, and it is a basic way to apply to digital medium music data, large digital music library, robot reading musical score and perform, computer music education, Chinese tradition music digitalization [1].

In this paper, firstly it was analyzed basic rule of CMN based on image processing, and get some relation of structures of musical symbols, secondly it describes a new method to pick-up musical information based on Mathematical Morphology, and it transforms musical information into MIDI file, the new method had achieved carry transformation from musical score to musical acoustic, and it can restoration musical score from MIDI file by using musical software.

II. KNOWLEDGE OF CMN

Notation is a method for recoding music, there are various music notation in the music history, such as CMN(Common Music Notation), Numbered Musical Notation, Chinese Tradition Abbreviated Character Notation and so on.

The notation was required to recode all musical activity, such as pitch, loudness, duration, timbre, expression symbols and so on. Note is a symbol for recoding tone of different duration, and rest is also a symbol for recoding gap time of some notes, there are some common notes and rests of CMN in Fig. 1.One row of piano score constitutes two staves, a bracket joints them in their left, a bar line is a vertical line which joint several horizontal staff, see Fig. 2.

The basic rule of CMN[2] based on image processing are following:

- Note includes three parts: head, stem and tail, the rate of high and length of head is the minimum in musical symbols, and it is between two neighbor stave or across one a staff.
- Bar line is a vertical line, from top staff line to botton staff line, its length is big then the length of stem, the total duration of all notes and rests in two bar line was fixed by deciding time signature.

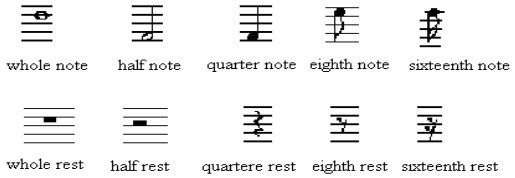

Figure 1. Common notes and rests

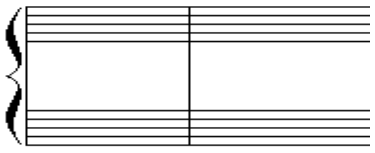

Figure 2. One row of piano score

978-0-7695-3557-9/09 $25.00 © 2009 IEEE
DOI 10.1109/ETCS.2009.261

- The most long lines in the digital musical score image are stave which include five horizontal lines, the length of stave are long then two-thirds of the length of image.
- Clef locates the right of first bar line and the left of key signature.
- Key signature locates the right of clef and the left of first note, meanwhile, the appearance order of sharps or flats keeps to fixed rule.

III. SOME OPERATIONS OF MATHEMATICAL MORPHOLOGY FOR DIGITAL MUSICAL SCORE

Mathematical morphology[3] is a powerful set of binary image processing operations, it was from a set-theoretical approach, it uses a structuring element to detect an image, and validates whether the structuring element can fill in the image, the selection of structuring element aims to the result of processing and different structuring element may get different analytical result.

Digital score images are binary images, *erosion operation*, *dilation operation* and *hit-or-miss transformation* are some operations of mathematical morphology for digital musical score.

Gerneral *erosion* is defined by:

$$E = B\Theta S = \{x, y \mid S_{xy} \in B\} \tag{1}$$

That is, the binary image E that results from eroding B by S is the set of points (x,y) such that if S is translated so that its origin is located at (x,y), then it is completely contained within B. S is a structuring element.

General *dilation* is defined by:

$$D = B\Theta S = \{x, y \mid S_{xy} \cap B \neq \phi\} \tag{2}$$

That is, the binary image D that results from dilatiting B by S is the set of points (x,y) such that if S is translated so that its origin is located at (x,y), then its intersection with B is not empty.

Hit-miss transformation needs two structuring elements E and F, one detects within image, other detects external image, lets $B=(E, F)$, it is defined by:

$$A * B = (A\Theta E) \cap (A^c \Theta F) \tag{3}$$

IV. PICK-UP MUSICAL INFORMATION BASED ON MATHEMATICAL MORPHOLOGY

There is a flow diagram based on OMR system for digitizing paper scores and picking-up musical information, see Fig. 3.

It supposes the original image as A , the highness of A marks as H, the width of A marks as W.

For musical information recognition, it uses Y-projection to get score staff lines[4], there had counted the amount of

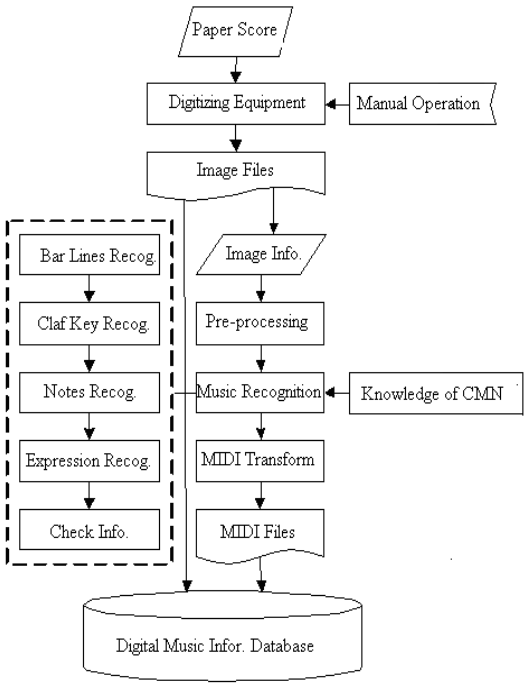

Figure 3. A flow diagram for digitial music processing

black pixels of each line, and had conserved the number to an array which marks as $S[n]$, $1 \leqslant n \leqslant H$, if it looked upon the number of S[n] as gray-level, and had got a gray-level histogram; Apparently, the gray-level histogram have two peaks, so we get the threshold valve to plot the two peaks, the threshold valve marks as f , the valve of some elements of $S[n]$ are large than f , the number of such elements marks as m, and the sequence of the subscript of such elements marks as R i , it suffices:

$$1 < R_i < W, 1 \leq i \leq m, S[R_i] > f \tag{4}$$

It marks the breadth of staff line as k, apparent, $k \geqslant 1$, and it has a property:

If k =1, that $R_{i+1} - R_i > 1, 1 \leqslant i < m$; if k>1, that it exists i, suffices $R_{i+1} - R_i = 1, 1 \leqslant i < m$.

It marks the number of i (sufficing k>1 and $R_{i+1} - R_i = 1$) as t , the distance of neighbor two staff lines marks as d, and d is defined by:

$$d = \frac{\left(\sum_{i=1}^{m-1}(R_{i+1} - R_i)\right) - t}{m - t} \tag{5}$$

We choose the first sheet of "Turkey March"(Beethoven) to processing, the result sees Fig. 4, and the left is the result of Y-Projection, the top right corner is a gray-level histogram for . the result of Y-Projection

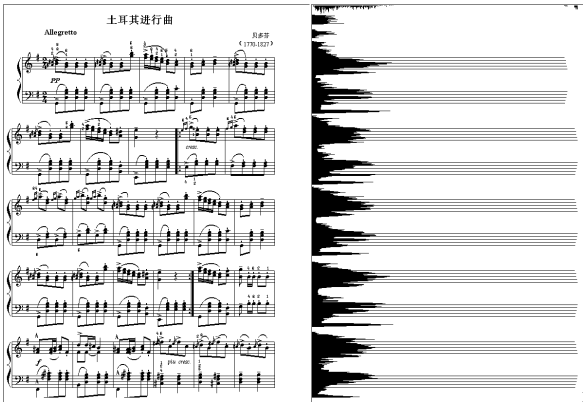

Figure 4. The result of *Y*-projection and gray-level histogram

It uses *hit-miss transformation* to recognise the clef and key signature in score by the location of stave and d .

It uses *erosion operation* to recognise notes. It marks the structuring element as $B1$ which be consist $(d - k)*(d - k)$ rectangle black pixel, It masks a originality image as A, and uses $A\Theta B1$ operation to recognise the head of notes in A, that it gets the location of whole heads. An experiment of $A\Theta B1$ for "Dance of Little Four Swans" (Tchailovsky), see Fig. 5. Meanwhile, it masrks the structuring element as $B2$ which be consist $1*[4*d+3*k]$ rectangle black pixel, and uses $A\Theta B2$ operation to recognise the stem of notes and bar line, an experiment of $A\Theta B2$ for "Für Elise"(Beethoven), see Fig. 6.

It gets two structuring element E and F by the location of head and stem of notes, and get the duration information of notes by *hit-miss transform*.

Now, the pitch and duration information of notes were got, that it uses the location of bar lines to plot out all notes, and amount to the duration of notes each bar, if the total duration of notes in each bar are equality, then the recognition procceing had ended, otherwise, it recognises all notes of the bar again.

Finally, we use the MIDI1.0 protocol to translate the musical information to a MIDI file, and the file will be save in the music database.

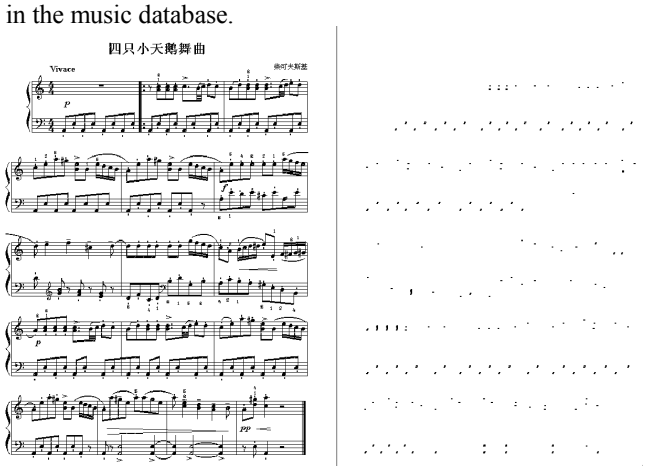

Figure 5. The score of "Dance of Little Four Swans" (Left) and the result of $A\Theta B_1$(Right)

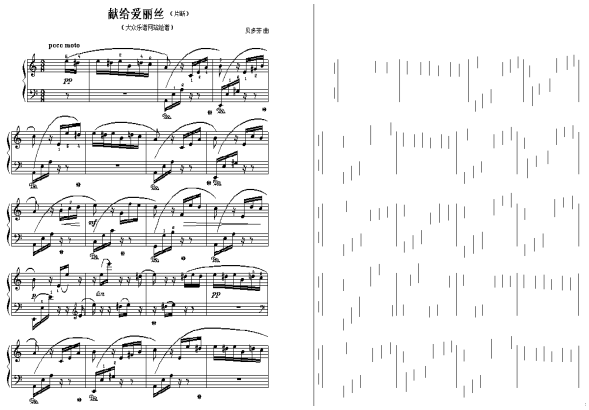

Figure 6. The score of "Für Elise" (Left) and the result of AΘB2 (Right)

V. THE RESULT OF EXPERIMENT

We had designed a prototype system by using some basic operation of mathematical morphology and some rules of CMN, and had downloaded 36 piece piano scores from Internet, the music all are counterpoint music, and include some works of famous composer and some folk song. We get the processing result of these score by the system. See Table 1.

From Table 1, the highest accuracy is 100%, and the lowest accuracy is 77.3%, the average recognition accuracy is 94.4%. By analyzing scores, the low recognition accuracy of score has more noises and expression marks than the high recognition accuracy of score.

VI. CONCLUSION

By mathematical morphology combines with the rule of music notation, we can digitizing the paper music score, and get an acoustics from the them, especially some scores with standard notation rule for example CMN, then the method in the paper can apply to all music score of CMN.

Now, a great of music works had existed based on paper and music notation, then the method in the paper is a key technique for large music database, digital library, music teaching, music analysis and so on.

REFERENCES

[1] Kia Ng, Jerome Barthelemy, Bee Ong, Ivan Bruno, Paolo Nesi, "CIMS: Coding Images of Music Sheets".The Interactive-Music Network,2004.

[2] Chongguang LI, *Basic Knowledge of Music*, Beijing: People Music Press,Beijing,1980.

[3] Qi CUI, Image Procceing and Analysis-Mathematical Morpholog, Beijing:Science Press,Beijing,2000.

[4] Etienne Sicard." An Efficient Method for the Recognition of Printed Music",0-8186-2920-7/92,IEEE,1992.

[5] http://www.myscore.org/.

TABLE I. A RECOGNITION RESULT OF 36 SCORES

Serial Number	Music Name	Total of notes	The amount of mistake recognition of duration	The amount of mistake recognition of pitch	Recognition Accuracy
1	Long Time Ago	205	0	93	0.773171
2	France folk song	170	24	35	0.826471
3	Little sister	200	24	22	0.885
4	March 2	356	27	44	0.900281
5	Italy Song	242	14	29	0.911157
6	Turkey March	447	4	75	0.911633
7	Doll March	218	11	23	0.922018
8	Symphonie No.5 Op.67 1st movement	244	5	31	0.92623
9	Polka	183	7	19	0.928962
10	Mazurka	191	14	12	0.931937
11	Starlight Waltz	379	18	32	0.934037
12	The Dream of Children	284	4	32	0.93662
13	Rhythm in F Major	253	4	25	0.942688
14	Little Cuckoo	100	0	10	0.95
15	Lovely May	170	6	10	0.952941
16	Fair	231	5	16	0.954545
17	Für Elise	129	3	8	0.957364
18	Dance of Little Four Swans	264	5	17	0.958333
19	Minuet in G Major	267	6	16	0.958801
20	Rhythm	256	5	16	0.958984
21	Iec-skating Waltz	265	0	21	0.960377
22	Wedding March	234	14	2	0.965812
23	Minuet	130	0	7	0.973077
24	Trout	163	2	5	0.978528
25	Romeo and Juliet	72	2	1	0.979167
26	March 1	339	6	5	0.983776
27	Ode To Joy	203	0	3	0.992611
28	Andante	152	0	2	0.993421
29	Below Moonlight	108	0	1	0.99537
30	No Title 2	111	0	1	0.995495
31	Beside Spring	149	0	1	0.996644
32	The Blue Danube	255	0	1	0.998039
33	No Title 1	95	0	0	1
34	May	167	0	0	1
35	Yankee Doodle	29	0	0	1
36	Lullaby	129	0	0	1
Statistic		7390	210	615	0.94418