PRIORITIZED CONTIG COMBINING TO SEGREGATE VOICES IN POLYPHONIC MUSIC

Asako Ishigaki, Masaki Matsubara, Hiroaki Saito
Graduate School of Science and Technology
Keio University
asako@nak.ics.keio.ac.jp

ABSTRACT

Polyphonic music is comprised of independent voices sounding synchronously. The task of voice segregation is to assign notes from symbolic representation of a musical score to monophonic voices. Human auditory sense can distinguish these voices. Hence, many previous works utilize perceptual principles. Voice segregation can be applied to music information retrieval and automatic music transcription of polyphonic music. In this paper, we propose to modify the voice segregation algorithm of contig mapping approach by Chew and Wu. This approach consists of 3 steps; segmentation, separation, and combining. We present a modification of "combining" step on the assumption that the accuracy of voice segregation depends on whether the segregation manages to correctly identify which voice is resting. Our algorithm prioritize voice combining at segmentation boundaries with increasing voice counts. We tested our voice segregation algorithm on 78 pieces of polyphonic music by J.S.Bach. The results show that our algorithm attained 92.21% of average voice consistency.

1. INTRODUCTION

Human auditory sense can distinguish plural melodic lines out of music. In musicology, these melodic lines are called voices. The task of voice segregation is to assign notes from symbolic representation of a musical score, such as MIDI, to voices. Unlike audio source separation, voice segregation cannot seek such clues as sound source or timbre, but relies solely upon pitch height, onset time, and duration and employs perceptual principles such as pitch proximity and stream crossing, indicated by Huron [1].

Voice segregation can be applied to music information retrieval and automatic music transcription. In queries by humming and theme finding, for instance, preliminary voice segregation of polyphonic music and homophonic music should be able to facilitate pattern recognition and extraction of monophonic queries and improve hit rates. Furthermore, in automatic music transcription, since the result of multiple pitch estimation by acoustic signal processing is segregated, voice segregation generates easy-to-read scores.

In this paper, we propose a voice segregation algorithm in polyphonic music based on modification of contig mapping approach by Chew and Wu [2]. Research by Chew and Wu of contig mapping approach provided high accuracy of voice segregation in polyphonic music. This approach consists of 3 steps; first, the piece of music is split between rests into units called contig, then voices are separated within each contig, and finally, the contigs are combined. We recognized the possibility of improving the contig combining step of this approach. Pitch proximity determines the voices to be connected across adjacent contigs in the contig combining step. We noticed that accuracy of connection may vary, depending on whether combining is done prior to a rest or after a rest. In this study, therefore, we propose an algorithm to determine the combining priority.

This paper is organized as follows. Section 2 presents a number of recent voice segregation algorithms. Section 3 describes the contig mapping approach and our modification. Section 4 presents our evaluation methods and experimental results. In Section 5, we discuss the results of experiment. Finally, Section 6 concludes this paper and presents future work.

2. RELATED WORK

In recent years, a number of approaches have been proposed for voice segregation [2, 3, 4, 5, 6, 7, 8].

Some of them employ machine learning. "VoiSe" system by Kirlin and Utgoff [3], for instance, learns about features of pairs of notes in the music and checks whether they belong to the same voice, and the learned decision tree is utilized for voice segregation. Some other algorithms which do not rely upon machine learning utilize perceptual principles by Huron [1] or the preference rules by Temperley [9]. Those approaches that group together a number of notes tend to have higher accuracy. Madsen and Widmer [4] group together notes with the same onset time and duration, establish the cost based upon Temperley's well-formedness rules, and then arrive at the shortest path by branch and bound search.

In contig mapping approach advocated by Chew and Wu [2], music piece is split into units known as contigs before and after rests. When a certain voice part is resting, fewer voices are sounding during that period of time. The number of voices vary before and after a rest, but since the music piece is split into contigs at that timing, the number of voices remains constant within each contig.

Then voices are separated within each contig. According to stream crossing principle by Huron [1], “humans have great difficulty of tracking auditory streams that cross with respect to pitch.” Therefore Chew and Wu assumed that voices would not cross. Since there are constant voice counts within a contig, voices can be separated simply by numbering in the order of pitch.

Finally, contigs are combined. Then it would be necessary to determine which voices are to be interconnected across adjacent contigs. Chew and Wu rely upon pitch proximity to determine which voices are to be connected. When all voices sound synchronously within a contig, such a contig is called maximal voice contig. Those contigs that are adjacent to a maximal voice contig are combined to that maximal voice contig. This combining process shall be performed to all of the maximal voice contigs of the entire music piece. Voice segregation of the music is complete when all the contigs are connected to the voices within the maximal voice contigs. Chew and Wu proposed metrics to measure the correctness of voice separation algorithm and achieved high hit ratio of polyphonic music data.

Voice segregation is applied to polyphonic as well as homophonic music, and the term ’voice’ has different meanings between them, as described by Cambouropoulos [6]. Polyphonic music is comprised of independent voices sounding synchronously, and each voice is monophonic. In polyphonic music, voice structure noted by the composer on the score is defined as the ground truth. The proportion of the notes assigned to the correct voice facilitates quantitative evaluation of the voice segregation. In this paper, we aim to segregate polyphonic music into monophonic voices, matching up precisely with the voice structure noted in the score.

In music piece with many voice parts, voice segregation of polyphonic music is prone to errors because music pieces with many voices tend to have also many rests. For example, when 3 voices out of 5 are resting, it would be difficult to infer which 2 voices out of 5 are continuing. This affects the combining step of the contig mapping approach. We propose to enhance the hit rate by improving this step.

3. THE ALGORITHM

3.1 Contig Mapping Approach

We assume that the accuracy of voice segregation depends on whether the segregation manages to correctly identify

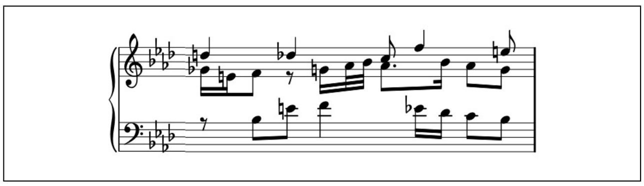

Figure 1. Example of a 3-voice polyphonic score. Measure 8 of Bach’s Sinfonia No.9 (BWV795)

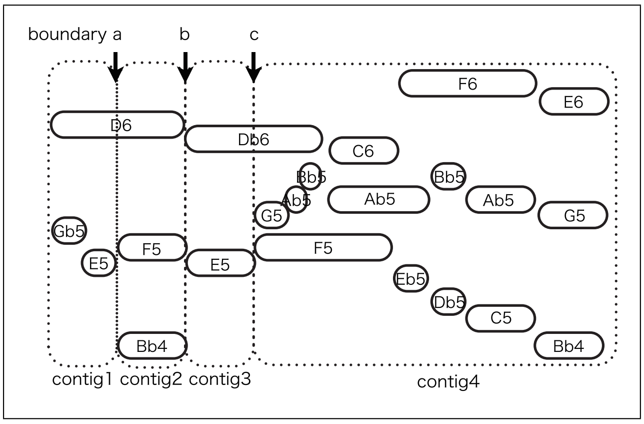

Figure 2. Example of segmentation into contigs

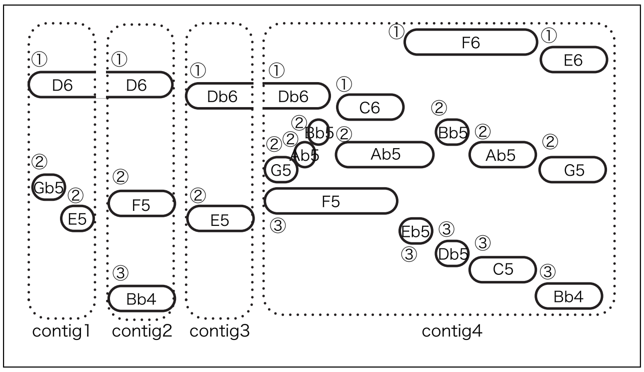

Figure 3. Example of voice segregation within each contig

which voice is resting.

Figure 1 illustrates an example of a 3-voice music. The upward stem notes of the upper staff correspond to the first voice, downward stem notes of the upper staff correspond to the second voice, and the downward stem notes of the lower staff correspond to the third voice. There are 2 eighth rests, one at the first beat, the other at the second. It would be necessary to correlate those rests with the corresponding voices. In contig mapping approach by Chew and Wu, the segregation process is divided into 3 steps to address this issue. The followings describe each step with its underlying concept.

3.1.1 Segmentation

First, music piece is split into units called contigs before and after rests. Figure 1 describes an example of a 3-voice score. This music is split into contigs before and after the eighth rests at the first and the second beat. Figure 2 illustrates how the segmentation is done. Contig 1 has one voice resting and 2 voices sounding, and contig 3 does the same. The number of voices changes at the boundary between contigs, but remains unchanged within a contig. If a note extends across a contig boundary, such as the quarter note on the second beat, a clone for the note is placed in the subsequent contig.

3.1.2 separation

According to the perceptual principles review by Huron, "humans have great difficulty of tracking auditory streams that cross with respect to pitch", which means that streams would hardly cross in polyphonic music, as it is intended by one musical instrument.

On the assumption that voices would not cross within a contig with constant voice counts, notes would be easily assignable to voices following the pitch order. Figure 3 illustrates the numbering of notes within each contig for the example given in Figure 2. Experiments by Chew and Wu demonstrated 99.8 percent accuracy of voice separation within a contig, but this applies only to allocation of notes to streams within a contig. For example, if a contig in a 3-voice music has 2 voices, this phase of separation does not infer as to which 2 out of the 3 voices they corresponded to. Such an inference is performed in the following combining step.

3.1.3 Combining

The final step combines the whole music piece which has been segmented into contigs. In this step, connection of voices across two adjacent contigs is inferred. In Figure 3, for instance, 2 voices in contig 3 have provisionally assigned numbers ① and ② respectively, but in reality, the first and the third voices are sounding. In the combining step, therefore, these two voices should be inferred to be connected to voices ① and ③ out of the 3 voices of adjacent contigs.

Huron reports, " the coherence of an auditory stream is maintained by close pitch proximity in successive tones within the stream. "It is inferred that across adjacent contigs, voices with closer pitch proximity belong to the same original voice.

The following is the method to connect voices with closer pitch proximity by the definition of the cost of voice connection based on pitch proximity.

First, the number of voices in each of the 2 contigs to be combined are counted. Then list all possible combinations of the voice connection. In the case of a 2-voice contig(P_1, P_2) to be combined to a 3-voice contig ($N_1 \dots N_3$), 3 patterns of connection combinations would be possible as indicated in Table 1, assuming that voices would not cross.

3-voice contig	2-voice contig		
	pattern 1	pattern 2	pattern 3
N_1	$\emptyset$	P_1	P_1
N_2	P_1	$\emptyset$	P_2
N_3	P_2	P_2	$\emptyset$

Table 1. Combinations of combining a 2-voice contig and a 3-voice contig

In pattern 2, P_1 is connected to N_1 and P_2 to N_3, and N_2 is on rest. It should be possible to enumerate all the connection combinations each time combining process takes place, as the number of voices is limited. Then arrive at the sum of connection costs for all of the patterns listed above. The cost shall be defined as the absolute difference between the pitches of the two notes; the final note of the preceding voice and the first note of the subsequent voice. If, however, the first note of the subsequent voice is the clone of the last note of the preceding note, the cost should be substituted by a negative value(-1), to ensure connection between these voices. "0" cost shall apply when either of the voices is on rest. Between contig 1 and contig 2 on Figure 3, since the first note of N_1 is the clone of the last note of P_1, the cost shall be -1, the cost for combining of N_2 and a rest shall be 0, and the cost between the first note of N_3 and the last note of P_2 shall be 6 because they are 6 semitones apart. That brings the total sum of cost for pattern 2 will be 5, arrived as $-1 + 0 + 6 = 5$. Following the same procedure, the cost for pattern1 shall be 15 and cost for pattern 3 shall be 0.

Then finally, choose the pattern with the minimum cost. In the above-mentioned example, the cost for pattern 3 is 0 and therefore the smallest. Pattern 3 combination is inferred to be optimum to connect contig 1 and contig 2 between closer voices. Thus voices ① and ② are identified as first and second voices, respectively.

However, an issue arises when contig 3 is combined. The above-mentioned optimum pattern to combine contig 2 and contig 3 indicates that voices ① and ② of contig 3 are first and second voices. On the other hand, optimum pattern to combine contig 3 and 4 infers contig 3 voices ① and ② to be first and third voices. In order to overcome this incoherence, we define the priority of combining.

3.2 Proposed Modification

3.2.1 Type of Boundaries between Contigs

At the boundary between contigs, the number of voices changes. There are two kinds of boundaries; one kind is boundary with increasing voice counts, such as boundary a, c on Figure 2 where voices increase from 2 to 3, and the other kind is boundary with decreasing voice counts, such as boundary b where voices decrease from 3 to 2.

Polyphonic music would develop by presenting the theme with each voice. The beginning of a stream is most often accompanied by the theme, so when a new stream begins at a boundary, which means that the voice counts increase with this boundary, the new stream should stand out. Based on the pitch proximity principle, therefore, the very first note of the new stream should have a pitch substantially apart from the notes of the remaining part of the voices that sounded immediately before. It is desirable to focus on boundaries where voices are easily distinguishable, as notes with closer pitch is the most relied-upon clue for voice connection.

We presumed that voice connection would be more accurate when contigs are combined at boundary with increasing voices than at boundary with decreasing voices. We verified our presumption with 78 pieces of music of J.S.Bach, such as invention, sinfonia, and fugue [10, 11, 12]. We made distinction between inter-contig boundaries with increasing voices and decreasing voices for all of the

78 pieces of music. Then we counted the accurate connection of voices across boundaries, following the contig combining procedure as described in 3.1.3. The result is shown in Table 2.

Out of the 2271 boundaries with increasing voice counts, 2004 boundaries achieved correct connection of adjacent contigs; the success ratio is 88.2%. Whereas boundaries with decreasing voice counts recorded 79.1% of connection accuracy. The result indicates voices are combined more accurately across contigs at boundaries where voice number is increasing. We propose an algorithm to prioritize voice combining at boundaries with increasing voice counts to improve the accuracy of voice segregation. The prioritization algorithm is described in the following section.

3.2.2 Connection Algorithm

We propose an algorithm that combines contigs only at those boundaries where the voice counts are increasing and then all the contigs are to be connected to one. The prioritization algorithm is described below.

(1) Specify contigs $C_0 \dots C_M$, and $|C_M|$ represents the number of voices of C_M

(2) Combine C_0 and C_1 if $M = 1$, then go to Step (9)

(3) Go to Step (9) if $M = 0$

(4) Set $N \leftarrow 0$

(5) Go to Step (8) if $N = M$

(6) Combine C_N and C_{N+1} if $|C_N| \leq |C_{N+1}|$

(7) Set $N \leftarrow N + 1$ then go to Step (5)

(8) Specify the combined contigs $C_0 \dots C_I$, set $M \leftarrow I - 1$, then go to Step (1)

(9) All contigs are combined into a contig

The following describes this algorithm taking an example illustrated sequentially in Figure 4, Figure 5, and Figure 6. In Figure 4, dotted lines show 8 boxes corresponding to contigs 1 through 8. It indicates that contigs 3 and 5 have 2 voices each, contigs 1, 4, 6 and 8 have 3 voices each, and contigs 2 and 7 have 4 voices each. The number of voices increase at 4 boundaries out of 7, i.e., boundaries $a, c, e,$ and *f*. In Step (4) through (8), contigs before and after these 4 boundaries are combined following the procedure described in 3.1.3. Then, contigs are combined into 4 broken-lined boxes, contigs $\{1, 2\}, \{3, 4\}, \{5, 6, 7\}$ and $\{8\}$, which are now called contig 1', 3', 5', and 8'. Contig

boundary type	increasing	decreasing
no. of boundaries	2271	1995
successfully combined	2004	1579
success ratio	88.2%	79.1%

Table 2. Success ratio in two boundary types

3' and 8' contain 3 voices each and contig 1' and 5' contain 4 voices each. It is still unknown which 3 voices out of the 4 are contained in contig 3' and 8'. As we repeat combining of these 4 contigs only at the boundaries where the voice counts increase, contig 3' and 5' are combined and we have 3 contigs; contig 1", 3" and 8"(indicated by the solid lines in Figure 5).

Now we still have a 3-voice contig, that is contig 8". Final combining is between contig 1" and 3", and then Step (2) is applied to combine contig 3" and 8" at a declining boundary as an exceptional measure. As shown in Figure 6, out of the iterations described above, all of the 8 contigs in Figure 4 have all their voices connected as indicated by gray lines, and the combining step is complete.

4. EXPERIMENTS AND RESULTS

4.1 Evaluation Methods

In polyphonic music, voices are recorded onto the score as is intended by the composer. The correspondence between notes and voices are established as the ground truth.

We tested the voice segregation algorithm on polyphonic music dataset. We chose the average voice consistency (AVC) out of the three metrics advocated by Chew and Wu [2] to quantify the algorithm performance. The voice segregation algorithm assigns all the notes in the music to voices. The assigned result is compared to the ground truth, and AVC is obtained as the percentage of the number of notes assigned to the correct voice.

4.2 Experiments

We obtained AVC from MIDI data of 78 music pieces by J.S. Bach [10, 11, 12]; 15 Inventions (2-voice), 15 Sinfonias (3-voice), 48 Fugues (2 to 5 voices). Since J.S. Bach is the representative composer of polyphonic music, a large number of related researches utilize this dataset for their experiments. Each track of the MIDI data contains a voice segregated in accordance with the ground truth. In this study, we iterated combining of contigs as described in 3.1.3 . After iteration of combining up to the phase described in Figure 6 is referred to as Full Experiment. The part of experiment up to the phase described in Figure 4, i.e. combining just once before iteration is referred to as Subset. We tested another algorithm which combines contigs only at the boundaries with decreasing voices and iterates the combining, referred to as Reverse, to contrast with

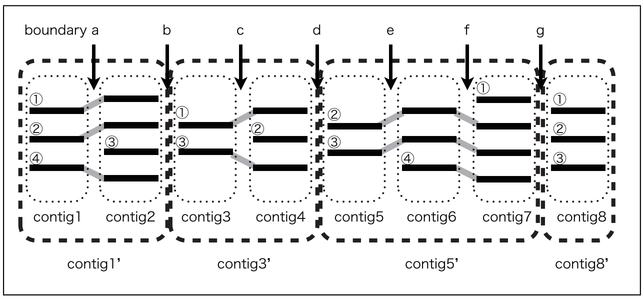

Figure 4. First phase of combining contigs

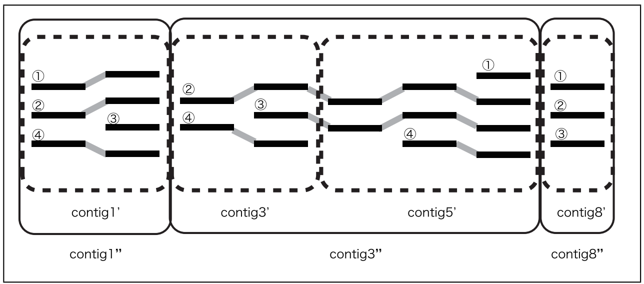

Figure 5. Second phase of combining contigs

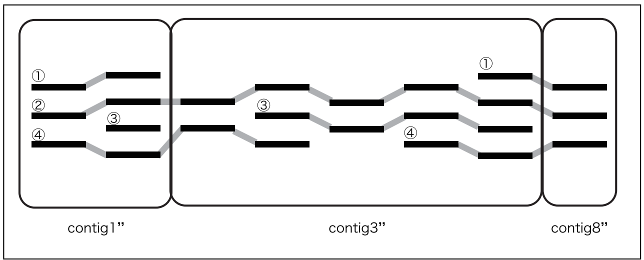

Figure 6. Last phase of combining contigs

Full Experiment.

4.3 Results

Table 3 shows the comparison of experimental results(AVC) of Full Experiment, Subset and Reverse as well as preceding studies based on the dataset of 78 music pieces, which is the same dataset this study is based upon.

	Invention	Sinfonia	Fugue	Avg.
Full Exp.	**98.73%**	**95.27%**	**89.21%**	**92.21%**
Subset	99.02%	95.30%	86.43%	90.56%
Reverse	98.87%	92.33%	80.75%	86.46%
Chew	99.29%	93.35%	84.39%	88.98%
Madsen	-	-	-	70.11%

Table 3. Experimental results of voice segregation

Fugues vary in voice counts; from 2- up to 5-voice. Table 4 derives from the same dataset as Table 3, but specifically classified by voice counts, which is unique to this particular study.

In the researches to-date, the algorithm by Chew and Wu recorded by far the highest hit rate. The result of our Subset alone surpasses the percentage achieved by Chew and Wu, as well as the result of Reverse fall below them. Furthermore, with the iteration in Full Experiment, the voice segregation accuracy improved even more than Subset. AVC higher than 85% was obtained for relatively more complex and difficult 5-voice music in the dataset.

	2-voice	3-voice	4-voice	5-voice
Full Exp.	**98.81%**	**93.73%**	**84.05%**	**85.50%**
Subset	99.11%	93.65%	79.30%	67.68%
Reverse	98.86%	88.51%	73.37%	69.50%

Table 4. Accuracy in 2-5 voices

5. DISCUSSION OF RESULTS

5.1 The result of prioritization of combine process

Table 3 shows that the Subset gave better results than the experiment by Chew and Wu and our Reverse, which indicates that the prioritization of boundaries with increasing voice counts was effective in the "combining" step of contigs.

For example, the result on Figure 7 is the correct voice segregation of the score of Figure 1. On this colored score, the first, the second, and the third voices are represented by pink, sky blue, and orange, respectively. If contigs are combined at the boundaries with decreasing voice counts, it would result in an error as indicated on Figure 8, where the note E on the second beat is represented by sky blue, although it really should be orange. Whereas the algorithm we propose combines contigs at the boundaries with increasing voice counts and generates correct results as indicated on Figure 7 for this example.

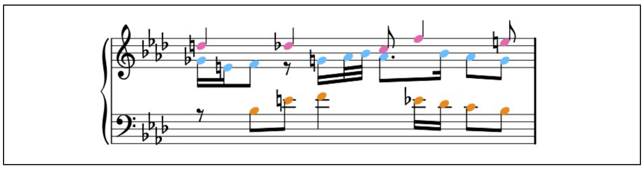

Figure 7. Example of the correctly combined contigs. Measure 8 of Bach's Sinfonia No.9 (BWV795)

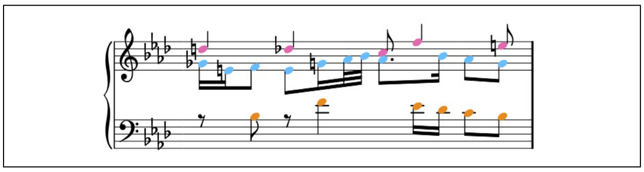

Figure 8. Example of the incorrectly combined contigs. Measure 8 of Bach's Sinfonia No.9 (BWV795)

5.2 The Iteration Effects

Four-voice or 5-voice music tend to have relatively longer duration of rests for 1 or more voices, and it is difficult to presume which of the voices are resting. Furthermore, if there is an error in presuming the voices on rest, that affects many notes and assign them to wrong voices. The success of voice segregation for 4-voice or 5-voice music depends upon the ability to correctly presume which voices out of the 4 or 5 corresponds to the 2 or 3 voices that are sounding.

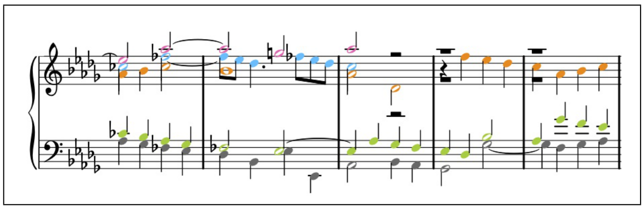

Figure 9. Example of correctly connected voices. Measures 35-39 of Bach's Fugue No.22 (BWV867)

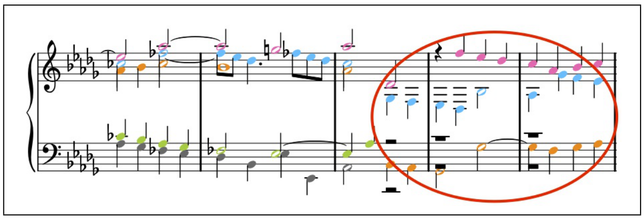

Figure 10. Example of disregarded potentiality of 4th or 5th voices. Measures 35-39 of Bach's Fugue No.22 (BWV867)

Table 3 shows further improvement of success ratio of Full Experiment over Subset. In Full Experiment, combining of contigs was iterated at boundaries with increasing voice counts. Furthermore, the improvements are specifically conspicuous for 4-voice and 5-voice music pieces, as shown in Table 4. It is therefore evident that the iteration by the proposed algorithm assigns sounding notes to the appropriate voices in those contigs with a number of synchronous rests.

Figure 9 illustrates the result of a partially correct voice segregation of a 5-voice music. Color-coding for the first, second and the third voices is the same as in Figure 7, and green and gray represent the fourth and the fifth voices, respectively. As is illustrated in Figure 4, to the extent of Subset, in 2-voice or 3-voice contigs that correspond to the red circle over the score on Figure 10, the potentiality of fourth or fifth voices sounding are disregarded; those 3 voices in reality must be represented by orange, green, and gray colors end up pink, sky blue, and orange. As the experiment advances to Full Experiment, however, the 3 voices are connected to the correct voices out of the five voices, and the result conforms to Figure 9, which is right.

Although this algorithm still does not guarantee perfect connections for all of the cases, our study demonstrated that the ratio of success can be enhanced by prioritizing which contigs are to be combined first.

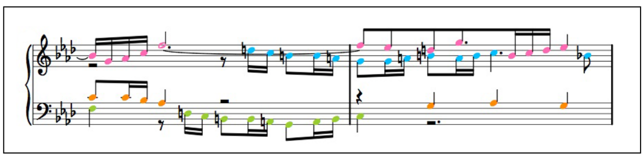

Figure 11. Example of voice permutation. Measures 18 and 19 of Bach's Fugue No.12 (BWV857)

Figure 12. Example of error caused by voice permutation. Measures 18 and 19 of Bach's Fugue No.12 (BWV857)

5.3 Error Analysis

Exceptionally, some music pieces do not end with monophonic voice. Chew and Wu excluded such exceptional contigs from their experiments beforehand; our study did not apply any such exceptions and had some errors. On Table 3, the success ratio of the proposed method for 2-voice was slightly below the experimental result by Chew and Wu. But we consider that this is not a major issue.

There are, however, certain errors that cannot be overcome by contig mapping approach based on pitch proximity principles and voice number variations. Those errors are caused by permutation and crossing of voices.

Figure 11 illustrates the music piece with voice permutation. On the second beat of the second bar, the fourth voice in green has a rest, and at the same time, a new stream begins for the third voice in orange. Contig mapping approach segments contigs by checking the voice number variations. So, if a certain stream ends at the same time with the beginning of another stream, segmentation does not take place. Therefore, the C sound in green and G sound in orange are regarded as one continuing voice. In the red-circled area of Figure 12, the green voice of the first bar is indicated as orange, and the orange voice is indicated as sky blue.

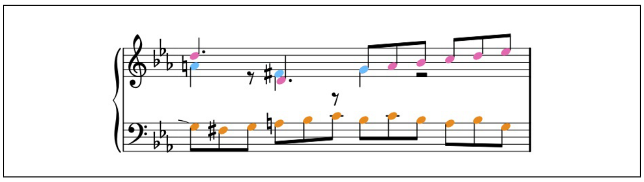

Figure 13. Example of voice crossing. Measure 12 of Bach's Sinfonia No.2 (BWV788)

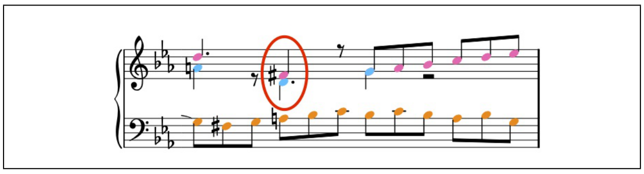

Figure 14. Example of error caused by voice crossing. Measure 12 of Bach's Sinfonia No.2 (BWV788)

Figure 13 illustrates the example of music piece with crossing voices. On the second beat, the D sound in pink has lower pitch than F# sound in sky-blue, and the first and the second voices cross. Many methods proposed to-date including our algorithm rely upon pitch proximity as the clue for voice segregation. The red-circled area on

Figure 14 shows the assignment of F# sound to first voice. This is a difficult problem to overcome in the current circumstances.

As voice crossing and permutations rarely occur, 92% of success ratio on the average was obtained despite of these errors.

6. CONCLUSION AND FUTURE WORK

This study proposes a voice segregation algorithm that has modified the contig mapping approach by Chew and Wu. The "combining" step of contigs has been improved by prioritization of boundaries with increasing voice counts, as higher success ratio was observed than at the boundaries with decreasing voice counts. The proposed algorithm was tested on music data of 78 pieces by J.S. Bach. It was evident from the test results that the success ratio improved and that the prioritization of combining of contigs is effective.

Future work of this study would be to test this algorithm with polyphonic music by other composers as well as other genre of music to broaden its application. We presume that theme finding would be clue for improving voice segregation algorithm because notes within a theme phrase should belong to a same voice. Assembling theme finding and voice segregation may help to overcome voice-crossing and permutation problems.

7. REFERENCES

[1] D. Huron, "Tone and voice: A derivation of the rules of voice-leading from perceptual principles," *Music Perception*, vol. 19, no. 1, pp. 1–64, 2001.

[2] E. Chew and X. Wu, "Separating voices in polyphonic music: A contig mapping approach," *Computer Music Modeling and Retrieval(CMMR)*, pp. 1–20, 2005.

[3] P. Kirlin and P. Utgoff, "Voise: Learning to segregate voices in explicit and implicit polyphony," *International Conference on Music Information Retrieval(ISMIR)*, 2005.

[4] S. Madsen and G. Widmer, "Separating voices in midi," *International Conference on Music Information Retrieval(ISMIR)*, 2006.

[5] A. Jordanous, "Voice separation in polyphonic music: A data-driven approach," *Proceedings of the International Computer Music Conference(ICMC)*, 2008.

[6] E. Cambouropoulos, "Voice and stream: Perceptual and computational modeling of voice separation," *Music Perception*, vol. 26, no. 1, pp. 75–94, 2008.

[7] I. Karydis, A. Nanopoulos, A. Papadopoulos, and E. Cambouropoulos, "Visa: The voice integration/segregation algorithm," *Proceedings of the International Conference on Music Information Retrieval(ISMIR)*, 2007.

[8] J.Kilian and H. Hoos, "Voice separation - a local optimisation approach," *Proceedings of the Fourth Annual International Symposium on Music Information Retrieval(ISMIR)*, 2003.

[9] D. Temperley, "The cognition of basic musical structures," *The MIT Press*, pp. 85–114, 2001.

[10] J.S.Bach, *Inventions and Sinfonias BWV772-801*. Henle, 1979.

[11] ——, *Well-Tempered Clavier - Part I BWV846-869*. Henle, 2007.

[12] ——, *Well-Tempered Clavier - Part II BWV870-893*. Henle, 2007.