Recognizing Chord Symbols in Printed Korean Musical Images Using Lexicon-Driven Approach

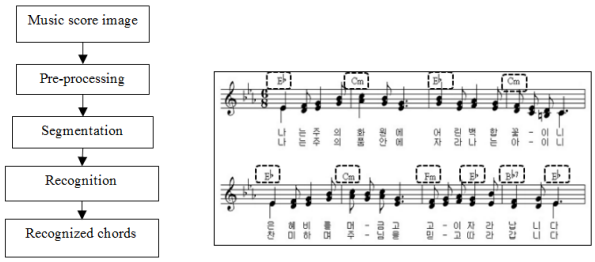
Minh Dinh*, Hyung-Jeong Yang**, Guee-Sang Lee,
Soo-Hyung Kim, In-Seop Na
Chonnam National University

ABSTRACT

Optical music recognition (OMR) systems have been developed in recent years. However, chord symbols that play a role in a music sheet have been still disregarded. Therefore, we aimed to develop a proper approach to recognize these chord symbols. First, we divide the image of chord symbol into small segments in horizontal by a method based on vertical projection. Then, the optimal combination of these segments is found by using a lexicon-driven word scoring technique and a nearest neighbor classifier. The word that corresponds to the optimal combination is the result of recognition. The experiment gives an impressive result with accuracy 97.32%.

I. Introduction

Recent years optical music recognition (OMR) [1] system gains great interest among researchers. The idea of this system is that printed or handwritten music score images can be played by user without any knowledge of music primitives or musical instruments by machine. There have been many such systems that have strongly developed. The recent survey of OMR systems and related publications can be found in [2, 3]. However, most of them have focused on musical symbols, such as notes, rests, sharps or flats, and disregarded chord symbols that give harmony information in a music sheet. A chord is important because it is a harmonic unit with at least three different tones sounding simultaneously. A combination of three or more pitches sounding at the same time. That is, two or more notes sounding simultaneously are known as a chord [4]. The most frequently encountered chords are triads, because they consist of three distinct notes: seventh chords, extended chords, or added tone chords. The most common chords are the major and minor triads and then the augmented and diminished triads. The descriptions "major", "minor", "augmented" and "diminished" are sometimes referred to collectively as chordal "quality". Chords are also commonly classed by their root note. For instance, the chord C Major may be described as a triad of major quality built upon the note C.]



(a) (b)

▶▶ Figure 1. (a) Flowchart of our method (b) Chord symbol locations

Ⅱ. Proposed Method

Our method consists of three stages: (1) Preprocessing, (2) Segmentation, and (3) Recognition as flowchart in Figure 1a. It starts taking the original music score image and results the chord symbols.

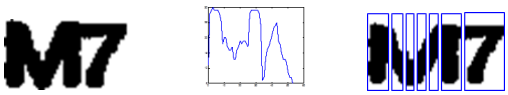
- Pre-processing: In this stage, we detect the areas of chords.
- Segmentation: After obtaining the area of a chord, we divide it into small segments in horizontal. The output of this stage is vertical slices.
- Recognition: At this stage, we apply a lexicon-driven approach proposed in [5]. For each chord, this approach find the optimal combination of segments from the second stage and best matching word in the lexicon simultaneously.

1. Preprocessing

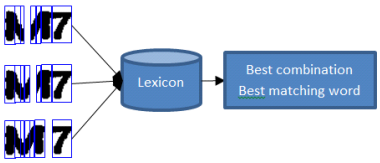
In order to detect the area of chords from a music score image, we perform some steps as following. Step 1: Input a music score image. Step 2: Binarization. Step 3: Detecting and removing stave-lines. Step 4: Finding connected components. Step 5: Detecting the components of chords by using stave-line information. Step 6: Output the locations of chords. The result is shown in Figure 1b.

2. Segmentation

After the locations of chords are gained by above stage, we perform the vertical projection only for the region of each chord. Then, the local minima of the projection histogram is detected. Two continuous local minima creates one vertical segment. The output of this stage is the set of vertical segments, as shown in Figure 2c.



▶▶Figure 2. Chord segmentation: (a) Chord symbol iamge (b) Vertical projection histogram (c) Vertical slices



▶▶ Figure 3. Chord recognition

3. Recognition

In this stage, we apply the approach proposed in [7] to find the optimal combination of segments that are divided by previous stage, and to simultaneously search the word in the lexicon that best matches to the chord. In this approach, as shown in Figure 3, the lexicon of chord symbols is defined before and templates of characters are available. Many combinations are generated from the list of segments. Then, a lexicon-driven word scoring technique and a nearest neighbor classifier are applied to find the best combination. The best matching word that corresponds to the best combination is the result of recognition.

Ⅲ. Experiment

We implemented the proposed method in C++. The sample images are captured in controlled environment with average resolution is 1223 by 1458. We collected about 30 sample music score images with over 200 chord symbols. Our method gives an impressive result with accuracy 97.32%.

Ⅳ. Conclusion

In this paper we have proposed a new method for the chord recognition. The main idea is that we simultaneously perform the segmentation and the matching to overcome touching cases. The results we obtained prove that our method is effective and provides high performance.

▮ References ▮

[1] Pugin L., “Optical music recognition of early typographic prints using hidden Markov models,” Proceedings of the International Conference on Music Information Retrieval, Vol. 7, pp. 53–6, 2006.

[2] Rebelo A., Fujinaga I., Paszkiewicz F., Andre M., Guedes C., Jaime C.S., “Optical music recognition: state-of-theart and open issues, International Journal of Multimedia,” Information Retrieval, Vol. 1, issue 3, pp 173-190, 2012.

[3] Rebelo A.. Capela G., and Cardoso S.C., “Optical recognition of music symbols: a comparative study,” International Journal on Document Analysis and Recognition, Vol. 13, Issue 1, 2010.

[4] Gorow R., “Hearing and Writing Music: Professional Training for Today’s Musician,” 2nd edition, 2000.

[5] Kim S. H., Jeong S., Suen C. Y., “A lexicon-driven approach for optimal segment combination in off-line recognition of unconstrained handwritten Korean words,” Pattern Recognition, Vol. 34, pp. 1437-1447, 2001.