ชื่อเรื่องวิทยานิพนธ์ การถอดเมโลดีของเสียงร้องจากเพลงแบบโพลีโฟนิก
ผู้เขียน นางสาวธนัชพร คันธรส
ปริญญา วิศวกรรมศาสตรมหาบัณฑิต (วิศวกรรมคอมพิวเตอร์)
อาจารย์ที่ปรึกษาวิทยานิพนธ์ อ.ดร. กานต์ ปทานุคม

บทคัดย่อ

ในงานวิจัยนี้ได้ทำการวิจัยเกี่ยวกับการถอดเมโลดีของเสียงร้องจากเพลงแบบโพลีโฟนิก ซึ่งเป็นส่วนหนึ่งของงานวิจัยทางด้านการค้นคืนข้อมูลดนตรี (Music Information Retrieval: MIR) โดยเมโลดีก็ถือได้ว่าเป็นข้อมูลที่มีความสำคัญมากเป็นอันดับต้นๆของเพลง ข้อมูลของเมโลดีจึงถือเป็นข้อมูลที่เป็นประโยชน์อย่างมากสำหรับงานหลายๆอย่าง แต่เนื่องจากเมโลดีในเพลงแบบโพลีโฟนิกซึ่งในแต่ละช่วงเวลาจะประกอบไปด้วยเสียงร้องและเสียงเครื่องดนตรีหลายๆชนิดปนกันอยู่นั้นส่วนใหญ่จะมากจากพิทช์ของเสียงร้อง โดยงานวิจัยนี้มีแนวคิดในการถอดเมโลดีของเสียงร้องโดยการด้วยการทำงานอยู่ 4 อย่างหลักๆคือ การเพิ่มความเด่นชัดให้กับความถี่ที่คาดว่าจะเป็นความถี่มูลฐานของเสียงที่มีลักษณะโครงสร้างเป็นฮาร์มอนิกซึ่งเป็นลักษณะของเสียงร้องและเสียงดนตรี โดยงานวิจัยนี้ได้พัฒนาวิธีการในการสร้างแผนภาพที่ใช้ในการวิเคราะห์เชิงเวลาและความถี่ที่เหมาะสมกับการทำงานที่ต้องการขึ้นมาซึ่งมีชื่อว่า"การปรับปรุงการทำนอร์มอไลซ์ผลรวมของฮาร์มอนิกย่อย (Modified Sub-Harmonic Summation Method: MSHS)"ขั้นตอนที่สองในการแก้ปัญหาคือ การตัดความถี่ที่มาจากเสียงเครื่องดนตรีที่มีลักษณะของการเปลี่ยนแปลงความถี่ที่เรียบ จากนั้นในขั้นตอนถัดไปจะมีการเลือกบริเวณของเมโลดีเนื่องจากพิทช์ของเมโลดีมักจะมีลักษณะที่กระจายตัวอยู่ในออกเตฟ (Octave) เดียวกันหรือใกล้เคียงกัน ดังนั้นหากสามารถเลือกบริเวณของเมโลดีได้ดีก็จะสามารถช่วยลดปัญหาของการตรวจจับพิทช์ได้ความถี่ที่อยู่คนละออกเตฟกับความถี่ที่เกิดขึ้นจริงได้ซึ่งปัญหานี้ถือว่าเป็นปัญหาที่มักจะพบบ่อยในการตรวจจับพิทช์ของเสียงดนตรีหรือเสียงร้อง ในขั้นตอนสุดท้ายนั้นเมโลดีของเสียงร้องจะได้มาจากการถอดพิทช์ออกมาจากบริเวณของเมโลดีที่หาได้จากส่วนก่อนหน้า

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ระบบที่พัฒนาขึ้นนี้ได้ถูกนำไปทดสอบด้วยชุดข้อมูลมาตรฐาน โดยผลการทดลองจากการแยกชุดข้อมูลทดสอบเป็นกลุ่มข้อมูลย่อยตามลักษณะที่คล้ายคลึงกันของไฟล์เสียงทดสอบพบว่าความแม่นยำของระบบจะแตกต่างกันไปตามลักษณะของเสียงที่แตกต่างกัน อย่างไรก็ตามผลการทดลองก็ยังแสดงให้เห็นว่าความแม่นยำของระบบจะเพิ่มขึ้นเรื่อยๆในแต่ละขั้นตอนการทำงานของระบบ นอกจากนั้นระบบยังสามารถลดจำนวนแคนดิเดตได้เป็นจำนวนมากในขณะที่รักษาจำนวนแคนดิเดตที่ถูกต้องให้ถูกตัดทิ้งไปเพียงจำนวนเล็กน้อย

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

Thesis Title Singing Melody Extraction from Polyphonic Songs

Author Miss Thanatphorn Kuntharose

Degree Master of Engineering (Computer Engineering)

Thesis Advisor Dr. Karn Patanukhom

Abstract

With the fact that music has always gained lots of interest from people and the increasing number of audio and acoustic recordings nowadays, researches in the field of Music Information Retrieval (MIR) have received much attention. Melody extraction is one of the problems in this research field that many researchers are trying to solve. It is still considered as an open problem since there is no method that has widely recognized as the best method for solving this problem. Melody is also considered as one of the most significant information of a song. So, it is crucial information to various applications.

In this research, we have developed a system for"Singing Melody Extraction from Polyphonic Songs". In polyphonic songs which comprise of both a singing voice and instrumental sounds, most of melodies come from pitches of the singing voice. Thus, our research tries to extract melodies that come from that singing voice. Our approach of singing melody extraction composes of 4 main steps. The first step is to increase the dominance of frequencies that are likely to be fundamental frequencies of harmonic structured sounds which is one of the characteristics of instrumental sounds and singing voices. We have developed a suitable method for creating a Time-Frequency (T-F) map that is called"Modified Sub-Harmonic Summation (MSHS)"method. The second step of the system is to eliminate frequencies from instrumental sounds. The next step is to estimate the melody region since melody pitches are

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

usually in the same octave or the octave nearby. Thus, estimating the melody region can decrease the problem of octave error which is a common problem in pitch detection. In the last step of the system, singing melody is obtained by extracting most dominant pitches of each time frame from the melody region.

We tested our system with the standard datasets and grouped test files into smaller data groups according to their characteristics. The experimental results show that the system's accuracy varies by input's characteristics. However, the results also show that in most data groups, the system's accuracy gets higher in each process of the system. Lots of candidates can be eliminated while maintaining the decreasing number of the candidates matched accuracy to be low in most cases.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved