โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ การประเมินเสียงร้องโดยใช้ลักษณะทางสวนศาสตร์

ชื่อนิสิต ปิยภูมิ แสงชัย

ภาควิชา ฟิสิกส์

ปีการศึกษา 2561

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของโครงงานทางวิชาการที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของโครงงานทางวิชาการที่ส่งผ่านทางคณะที่สังกัด
The abstract and full text of senior projects in Chulalongkorn University Intellectual Repository(CUIR)
are the senior project authors' files submitted through the faculty.

โครงงาน

เรื่อง

การประเมินเสียงร้องโดยใช้ลักษณะทางสวนศาสตร์

Singing Evaluation Based on Acoustics Features

ผู้จัดทำโครงงาน

นายปิยภูมิ แสงชัย

อาจารย์ที่ปรึกษาโครงงาน

ผู้ช่วยศาสตราจารย์ ดร.ณัฐกร ทับทอง

ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์

จุฬาลงกรณมหาวิทยาลัย

ชื่อโครงงาน (ภาษาไทย)	การประเมินเสียงร้องโดยใช้ลักษณะทางสวนศาสตร์
ผู้วิจัย	นายปิยภูมิ แสงชัย เลขประจำตัว 5633478823
ชื่ออาจารย์ที่ปรึกษา	ผศ.ดร.ณัฐกร ทับทอง
ภาควิชา/คณะ	ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา	2561

โครงงานนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรรายวิชาปริญญาวิทยาศาสตรบัณฑิต ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

คณะกรรมการได้ตรวจสอบและรับรองรายงานฉบับนี้แล้ว

________________ประธานคณะกรรมการสอบ

(อาจารย์วิสิทธิ์ ลีลาศิริวงศ์)

________________กรรมการสอบ

(อาจารย์ ดร.ยุทธนา รุ่งธรรมสกุล)

________________อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.ณัฐกร ทับทอง)

ชื่อโครงงาน (ภาษาไทย)	การประเมินเสียงร้องโดยใช้ลักษณะทางสวนศาสตร์
ผู้วิจัย	นายปิยภูมิ แสงชัย เลขประจำตัว 5633478823
ชื่ออาจารย์ที่ปรึกษา	ผศ.ดร.ณัฐกร ทับทอง
ภาควิชา/คณะ	ภาควิชาฟิสิกส์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา	2561
คำสำคัญ	การประเมินเสียงร้อง, ลักษณะทางสวนศาสตร์, คาราโอเกะ

บทคัดย่อ

คาราโอเกะ (Karaoke) เป็นกิจกรรมนันทนาการรูปแบบหนึ่งที่ได้ความนิยมในเอเชียฝั่งตะวันออก เครื่องคาราโอเกะประกอบไปด้วยชุดเครื่องเสียงและไมโครโฟนสำหรับร้องเพลงไปตามเนื้อเพลงที่ปรากฏบนจอภาพ โดยเพลงในเครื่องคาราโอเกะจะถูกตัดเสียงร้องหลักของศิลปินออกไป เพื่อให้ผู้ร้องสามารถร้องแทนเสียงร้องหลักและสามารถใช้ฝึกฝนความสามารถในการร้องเพลงได้ โครงงานนี้มีจุดประสงค์เพื่อสร้างการประเมินการร้องเพลงด้วยคอมพิวเตอร์ โดยพิจารณาลักษณะทางสวนศาสตร์ ได้แก่ ระดับเสียง ความยาว และความหนักเบา และความถูกต้องของการร้องตามเนื้อร้อง จากการทดลองพบว่าการประเมินการร้องด้วยคอมพิวเตอร์นี้ สามารถใช้ประเมินการร้องของนักร้องได้จริง แต่ยังให้ผลที่มีความแตกต่างจากผลการประเมินโดยมนุษย์อยู่มาก

Project Name	Singing Evaluation Based on Acoustics Features
Researcher	Mr.Piyapoom Saengchai ID 5633478823
Advisor	Assistant Professor Nuttakorn Thubthong
Department/Faculty	Physics / Science
Academic Year	2018
Keywords	Singing evaluation, acoustic features, karaoke

Abstract

Karaoke is one of the most popular recreational activities in East Asia. A karaoke machine consists of a set of audio devices and a microphone for singing according to the lyrics displayed on a monitor. The song in the karaoke machine will be removed the main vocals of the artist. So that the singer can sing instead of the main vocals and can use it to practice singing ability. This project aims to create an computerized singing evaluation by considering the acoustic characteristics, including pitch, duration and dynamic, and the accuracy of the singing by the lyrics. From the experiment, it was found that this computerized singing evaluation can be used to evaluate the singer's vocals. However, the evaluation results are still many differences from human evaluation results

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณครอบครัวที่สนับสนุนและเป็นกำลังใจในการศึกษาเล่าเรียนและทำงานวิจัย ขอขอบพระคุณอาจารย์ที่ปรึกษาโครงงาน ผศ.ดร.ณัฐกร ทับทอง ที่ได้เป็นธุระในการกำกับดูแล ให้ข้อชี้แนะและให้การสนับสนุนตลอดการทำโครงงาน ขอขอบคุณเพื่อนๆ ที่ให้ข้อมูลเสียงร้องเพลงสำหรับการทำวิจัยและช่วยเหลือในเวลาต่างๆ ขอขอบคุณห้องซ้อมดนตรี สุดท้ายนี้ ขอขอบคุณตัวผู้วิจัยที่ทุ่มเท อดทนฟันผ่าอุปสรรคต่างๆ ในการทำงานจนสำเร็จลุล่วงไปด้วยดีทุกประการ

สารบัญ

	หน้า
บทที่ 1 บทนำ	1
ที่มาของโครงงาน	1
วัตถุประสงค์	1
ขอบเขตการทดลอง	1
ประโยชน์ที่คาดว่าจะได้รับ	1
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	2
การกำเนิดเสียงของมนุษย์	2
ลักษณะทางสวนศาสตร์ของเสียงร้อง	4
บันไดเสียงดนตรี (Musical Scale)	6
บทที่ 3 การทดลอง	7
การบันทึกเสียง	7
การสกัดลักษณะทางสวนศาสตร์	9
การประเมินเสียงร้อง	12
บทที่ 4 ผลการทดลอง	14
อภิปรายผลการทดลอง	19
บทที่ 5 สรุปผลการทดลอง	20
อุปสรรคและข้อเสนอแนะ	20
เอกสารอ้างอิง	21
ภาคผนวก	22
ก. เนื้อร้องของเพลงทดสอบ	22
ข. ระเบียบวิธีการคำนวณคะแนนด้วยโปรแกรม Microsoft Excel	28

บทที่ 1 บทนำ

ที่มาและความสำคัญของงานวิจัย

ในช่วงทศวรรษที่ผ่านมา มีงานประกวดแข่งขันทางดนตรีและการร้องเพลงเกิดขึ้นหลายรายการที่เปิดโอกาสให้นักร้องและนักดนตรีรุ่นใหม่ได้แสดงความสามารถที่ผ่านการฝึกฝนมาอย่างเต็มที่ โดยแต่ละคนต่างมีรูปแบบในการฝึกฝนฝีมือตนเอง วิธีการหนึ่งก็คือการฝึกกับเพลงคาราโอเกะ (Karaoke) ซึ่งเป็นกิจกรรมนันทนาการรูปแบบหนึ่งที่ได้ความนิยมในโซนเอเชียฝั่งตะวันออก ซึ่งประกอบไปด้วยชุดเครื่องเสียงและไมโครโฟนสำหรับร้องเพลงไปตามเนื้อเพลงที่ปรากฏบนจอแสดงผล โดยเพลงในเครื่องคาราโอเกะจะถูกตัดเสียงร้องหลักของศิลปินออกไป เพื่อให้ผู้ร้องสามารถร้องแทนเสียงร้องหลักได้

ในเครื่องคาราโอเกะบางเครื่องสามารถทำการประเมินให้คะแนนการร้องเพลงได้ แต่การประเมินของเครื่องคาราโอเกะมีความแตกต่างจากการประเมินของมนุษย์ ผู้วิจัยจึงมาความสนใจที่จะสร้างโปรแกรมที่สามารถประเมินการร้องเพลงโดยใช้ลักษณะทางสวนศาสตร์ (Acoustic features) ได้แก่ ระดับเสียง (Pitch) ความยาว (Duration) และความหนักเบา (Dynamic) รวมทั้งความถูกต้อง (Correction) ในการร้องเนื้อเพลงเป็นเกณฑ์อ้างอิงในการประเมินการร้องเพลง

วัตถุประสงค์ของโครงงาน

1. เพื่อศึกษาลักษณะทางสวนศาสตร์ของเสียงร้อง
2. เพื่อพัฒนาขั้นตอนวิธีการประเมินการร้องเพลงโดยใช้ลักษณะทางสวนศาสตร์เป็นเกณฑ์อ้างอิง

ขอบเขตของการวิจัย

โครงการนี้ใช้เสียงต้นแบบจากคลิปเสียงคาราโอเกะความยาวประมาณ 5 นาที ที่ตัดมาจากเพลงไทยสากล จำนวน 3 เพลง และใช้เสียงของศิลปินต้นฉบับเป็นเกณฑ์ในประเมิน และใช้เสียงร้องของนักร้องทดสอบจำนวน 6 คน ในการทดสอบ

ประโยชน์ที่คาดว่าจะได้รับ

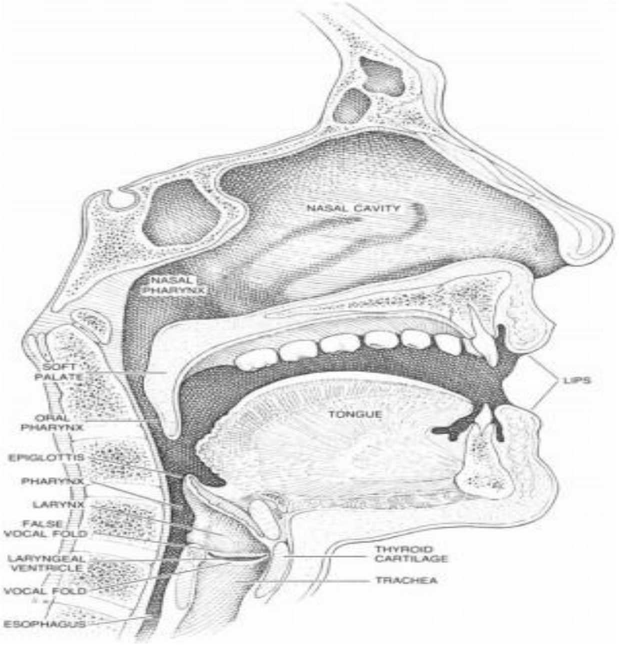
1. สามารถวิเคราะห์ลักษณะทางสวนศาสตร์ได้
2. ได้ขั้นตอนวิธีการประเมินการร้องเพลงโดยใช้ลักษณะทางสวนศาสตร์เป็นเกณฑ์อ้างอิงได้

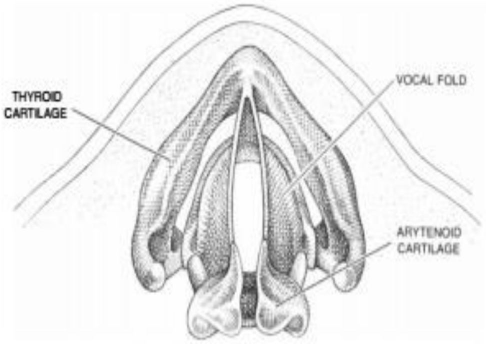
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง

การกำเนิดเสียงของมนุษย์

เสียงพูด (Voice) เป็นเสียงที่สร้างจากอวัยวะภายในช่องทางเดินเสียง (Vocal Tracts) ได้แก่ ปอดทำหน้าที่สร้างกระแสอากาศไหลผ่านเส้นเสียง (Vocal folds) ตามรูปที่ 2.1 ซึ่งจะสั่นและสร้างคลื่นเสียงที่มีความถี่มูลฐาน (Fundamental Frequency) หรือ f_0 จากนั้นคลื่นเสียงที่ได้จะเกิดการสั่นพ้องภายในช่องปากและจมูกเกิดเป็นความถี่ฟอร์แมนท์ (Formant) ซึ่งเป็นฮาร์มอนิกของความถี่มูลฐาน ตามรูปที่ 2.2

กล้ามเนื้อในกล่องเสียง (Larynx) ทำหน้าควบคุมการสั่นของเส้นเสียง มีผลต่อการเปลี่ยนระดับเสียง ขณะที่ลิ้นควบคุมรูปร่างภายในช่องปาก มีผลต่อการสั่นพ้อง ทำให้เสียงพูดของมนุษย์มีความหลากหลาย [1]





รูปที่ 2.1

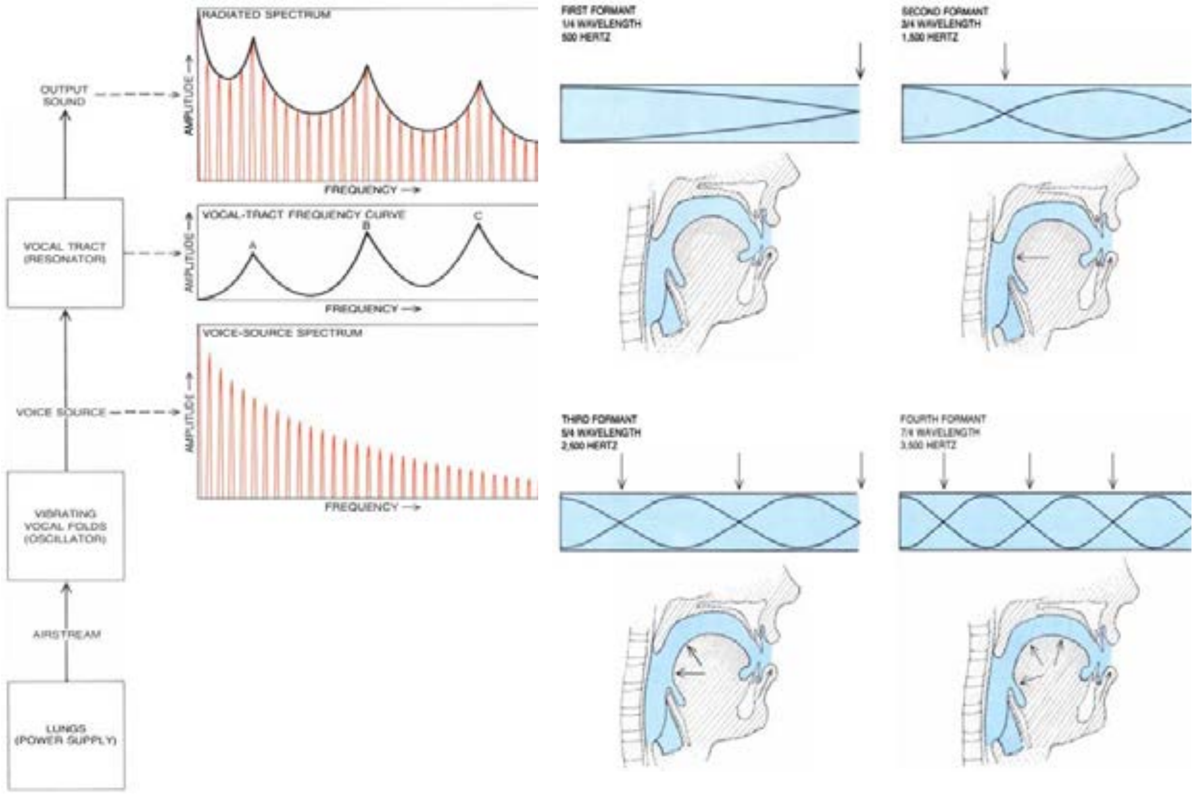
(ซ้าย) อวัยวะในช่องทางเดินเสียง

(บน) องค์ประกอบของกล่องเสียง

(ที่มา Sundberg,1977[1])

---

1 Sundberg, J. (1977) “The acoustics of the singing voice”, *Scientific American*, 236, 82.



รูปที่ 2.2 (ซ้าย) ขั้นตอนการกำเนิดเสียงของมนุษย์ (ขวา) การเกิดความถี่ฟอร์แมนท์

(ที่มา Sundberg,1977[2])

ในการพูดปกติ เพศชายจะมีความถี่ f_0 อยู่ในช่วง 85 ถึง 180 เฮิร์ตซ์ และในเพศหญิงอยู่ในช่วง 165 ถึง 255 เฮิร์ตซ์ [2] แต่ในการขับร้องจะมีการแบ่งเสียงร้องออกเป็นหกช่วงตามระดับเสียงคร่าวๆ [3] ได้แก่

1. โซปราโน (Soprano) เป็นช่วงเสียงที่สูงที่สุดสำหรับเพศหญิง เริ่มจากโน้ต C_4 ไปจนถึง C_6
2. เมสโสโซปราโน (Mezzo-Soprano) เป็นช่วงเสียงกลางสำหรับเพศหญิง เริ่มจากโน้ต A_3 ไปจนถึง A_5
3. คอนทราอัลโต (Contralto) เป็นช่วงเสียงที่ต่ำที่สุดสำหรับเพศหญิง เริ่มจากโน้ต F_3 ไปจนถึง E_5
4. เทเนอร์ (Tenor) เป็นช่วงเสียงสูงที่สุดสำหรับเพศชาย เริ่มจากโน้ต B_2 ไปจนถึง A_4
5. บาริโทน (Baritone) เป็นช่วงเสียงกลางที่สุดสำหรับเพศชาย เริ่มจากโน้ต G_2 ไปจนถึง F_4
6. เบส (Bass) เป็นช่วงเสียงต่ำที่สุดสำหรับเพศชาย เริ่มจากโน้ต E_2 ไปจนถึง E_4

---

[2] Sundberg, J. (1977) “The acoustics of the singing voice”, *Scientific American*, 236, 82.

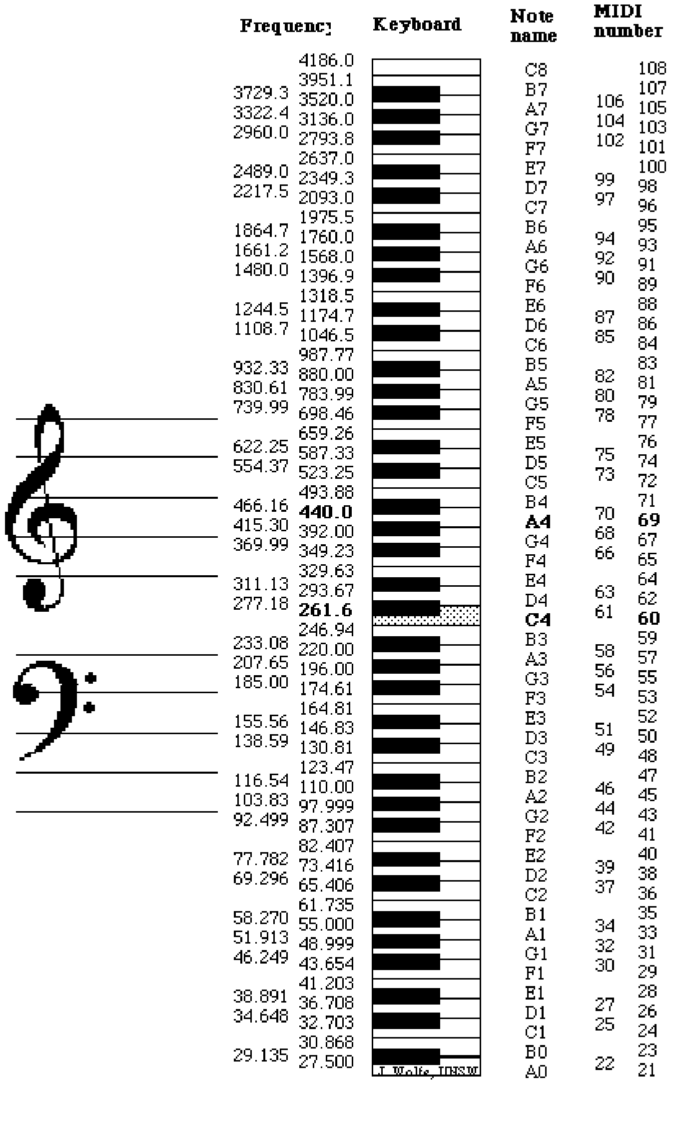
ลักษณะทางสวนศาสตร์ของเสียงร้อง

ระดับเสียงร้องมีความสัมพันธ์โดยตรงกับความถี่มูลฐานของเสียง Tsai และคณะ [4] ได้เสนอการหาระดับเสียง (pitch) โดยการประยุกต์ Fast Fourier Transforms (FFTs) กับคลื่นเสียงที่อยู่ในโดเมนเวลา ซึ่งสามารถเทียบออกมาเป็นโน้ต (Note) ตามมาตรฐาน ISO 16 ที่กำหนดความถี่ของโน้ต A_4 เป็น 440 Hz [5] และสามารถแปลงเป็นรูปแบบ MIDI (Music Instrument Digital Interface) ได้ โดยใช้สมการที่ 2.1

$$m = \left[12 \cdot log_2\left(\frac{f}{440\ Hz}\right) + 69\right] \qquad \text{-(2.1)}$$

เมื่อ f คือ ความถี่มูลฐาน และ m คือ เลข MIDI

ความสัมพันธ์ระหว่างความถี่มูลฐาน โน้ตทางดนตรี และเลข MIDI แสดงดังรูปที่ 2.3



รูปที่ 2.3 การเปรียบเทียบความถี่ของเสียง โน้ตทางดนตรี และเลข MIDI (ที่มา https://newt.phys.unsw.edu.au)

นอกจากนี้ยังได้เสนอแนวทางการประเมินความหนักเบาของเสียง (Dynamic) ในแต่ละพยางค์โดยการพิจารณาความเข้มเสียง (Intensity) เพื่อวิเคราะห์ความแตกต่างของพลังงานระหว่างผู้ทดสอบกับเสียงศิลปิน

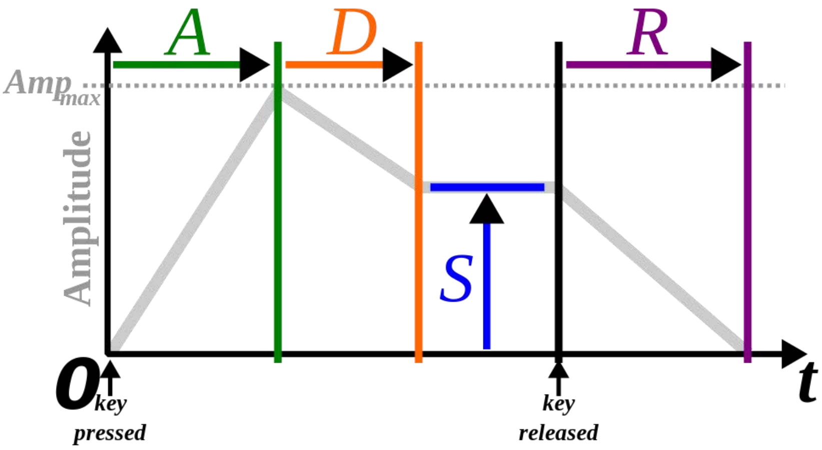
รูปร่างและสเปกตรัมของคลื่นเสียงสามารถนำหาขอบเขตของพยางค์ (Syllable Boundary) ความยาวของพยางค์ (Syllable Duration) และคุณภาพของเสียงร้อง (Timbre) โดยรูปร่างของคลื่นเสียงมีส่วนประกอบสำคัญตามหลัก ADSR [6] 4 ระยะ ที่แสดงดังรูปที่ 2.4 คือ

1. ระยะโจมตี Attack เป็นช่วงที่คลื่นเสียงมีการวิวัฒน์แอมพลิจูดจากศูนย์ไปจนถึงค่าสูงสุด

2. ระยะสลาย Decay เป็นช่วงที่แอมพลิจูดลงมาอยู่ในระดับคงตัวในระยะถัดไป

3. ระยะคงตัว Sustain เป็นช่วงที่เสียงมีความหนักเบาคงที่

4. ระยะปลดปล่อย Release เป็นช่วงที่แอมพลิจูดลงไปจนถึงศูนย์



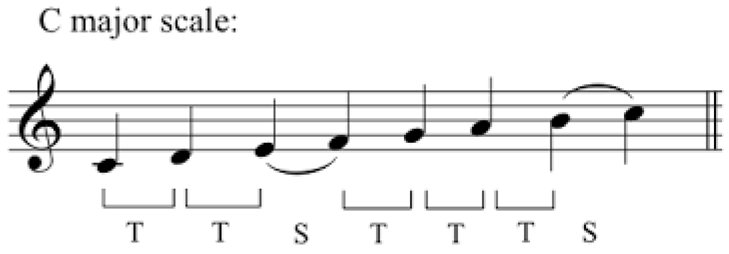
รูปที่ 2.4 แสดงระยะต่าง ๆ ของรูปร่างคลื่น (ที่มา https://www.keithmcmillen.com)

โดยของเขตของพยางค์ นับจากเริ่มต้นระยะโจมตีไปจนสิ้นสุดระยะปลดปล่อย ซึ่งความยาวของพยางค์ คือ ค่าเวลาของโน้ต และรูปร่างของคลื่นเสียงประกอบกับเสียงฮาร์มอนิกในลำดับสูงส่งผลให้เสียงต่างกันมีคุณภาพเสียงที่ต่างกัน ทำให้มนุษย์สามารถแยกเสียงเกิดจากเครื่องดนตรีต่าง ๆ ที่มีระดับเสียงและความหนักเบาเท่ากันได้

บันไดเสียงดนตรี (Musical Scale)

บันไดเสียง หมายถึง การเรียงลำดับของตัวโน้ตโดยใช้ระดับเสียงหรือความถี่มูลฐาน เนื่องจากหลักความสมมูลของช่วงคู่แปด (Octave) บันไดเสียงจะเป็นตัวแทนของการแบ่งช่วงคู่แปดออกเป็นระยะห่างย่อยๆ ภายในช่วงคู่แปดนี้ เช่น บันไดเสียงเมเจอร์ (Major Scale) ที่แบ่งช่วงคู่แปดออกเป็น T-T-S-T-T-T-S โดย T หมายถึงระยะห่างหนึ่งเสียง (Whole Tone) และ S หมายถึงระยะห่างครึ่งเสียง (Semitone)

บันไดเสียงใด ๆ สามารถระบุได้โดยระยะห่างของแต่ละขั้นกับโน้ตโทนิค (Tonic Note) เป็นตัวโน้ตตัวแรกในบันไดเสียงนั้น ซึ่งสามารถทราบได้เครื่องหมายจากกุญแจเสียงในการจดบันทึกโน้ต โดยปกติชื่อเรียกของบันไดเสียงในงานดนตรีจะใช้ชื่อโน้ตโทนิคและชื่อของบันไดเสียง เช่น บันไดเสียงซีเมเจอร์ (C major Scale) จะเริ่มต้นที่โน้ต C และมีระยะห่างกับโน้ตตัวถัดไปตามการแบ่งขั้นในบันไดเมเจอร์ตามที่กล่าวในขั้นต้นตามรูปที่ 2.5 [7]



รูปที่ 2.5 ตัวอย่างของบันไดเสียงซีเมเจอร์ (ที่มา https://musictheorysite.wordpress.com)

บันไดเสียงของเพลงอาจอยู่ในช่วงเสียงที่นักร้องไม่สามารถร้องได้ จึงต้องมีการปรับโน้ตโทนิคจนอยู่ในช่วงเสียงที่นักร้องสามารถร้องได้ โดยการใช้สมการโน้ต MIDI ข้างต้นมาประยุกต์ใช้เป็นสมการ 2.2

$$x = m' - m = 12 \cdot log_2(\frac{f'}{f}) \qquad \text{-(2.2)}$$

โดย x คือ ผลต่างของเลข MIDI ของโน้ตโทนิคใหม่กับโน้ตโทนิคเดิม

f' คือ ความถี่ของโน้ตโทนิคใหม่

f คือ ความถี่ของโน้ตโทนิคเดิม

บทที่ 3 การทดลอง

ในบทนี้จะกล่าวถึงขั้นตอนวิธีการดำเนินการทดลอง ซึ่งแบ่งเป็นสามส่วน ได้แก่ การบันทึกเสียงสำหรับใช้เป็นข้อมูลในการทดลอง การสกัดลักษณะทางสวนศาสตร์ เพื่อใช้สำหรับการประเมินเสียงร้อง และวิธีการประเมินเสียงร้อง

1. การบันทึกเสียง

1.1 เสียงร้องต้นแบบ

โครงงานนี้ใช้เสียงร้องต้นแบบจาก 3 เพลง ที่ร้องโดยศิลปิน 3 คน ดังรายละเอียดในตารางที่ 3.1

โดยทำการคัดแยกเฉพาะเสียงร้องจากต้นฉบับไฟล์เสียงเพลงเต็มที่มีเสียงร้องและเสียงดนตรี ด้วยการนำเสียงนั้นไปทำการหักลบกับไฟล์เสียงดนตรีคาราโอเกะที่มีเสียงดนตรีเท่านั้น ซึ่งเสียงของเครื่องดนตรีที่มีอยู่ในทั้งสองไฟล์จะการแทรกสอดแบบหักล้างกันเนื่องจากมีความถี่และเฟสเดียวกัน เหลือเพียงเสียงร้องเท่านั้น

ตารางที่ 3.1 เสียงร้องต้นแบบที่ใช้ในการทดสอบ

ลำดับ	ชื่อศิลปิน	ชื่อเพลง	จำนวนพยางค์	จังหวะต่อนาที	เครื่องหมายกุญแจเสียง
1	พีรพัฒน์ เถรว่อง	รู้	342	66	Bb
2	วง Bodyslam	แสงสุดท้าย	349	114	C
3	วง Cocktail	คู่ชีวิต	427	72	F

1.2 เสียงร้องทดสอบ

เสียงร้องทดสอบ บันทึกจากนักร้องจำนวน 6 คน โดยมีรายละเอียดดังตาราง 3.2

ตารางที่ 3.2 ข้อมูลนักร้องทดสอบ

รหัส	เพศ	อายุ	อาชีพ	ประสบการณ์ในการร้องเพลง (ปี)
F1	หญิง	25	สอนพิเศษ/นิสิต	1
M1	ชาย	28	ครู/นักดนตรี	5
M2	ชาย	27	ครู	2
M3	ชาย	24	นิสิต	0
M4	ชาย	27	อาจารย์มหาวิทยาลัย	1
M5	ชาย	24	ว่างงาน	0

นักร้องทดสอบทุกคน ร้องเพลง 3 เพลง จำนวน 1 ครั้ง โดยที่นักร้องแต่ละคนต้องฟังเพลงต้นฉบับก่อนทำการบันทึกเสียง ในขณะบันทึกเสียง นักร้องทดสอบสามารถดูเนื้อร้องประกอบ (ดูเนื้อร้องได้ตามภาคผนวก ก.) หากไม่นักร้องทดสอบสามารถร้องในช่วงเสียงของเพลงได้ จะต้องทำการปรับเครื่องหมายกุญแจเสียงโดยใช้สมการที่ 2.2

1.3 การบันทึกเสียง

อุปกรณ์ที่ใช้ในการบันทึกเสียงบันทึกเสียง

- Audio Interface ยี่ห้อ TASCAM รุ่น US-200

- ไมโครโฟน ยี่ห้อ Shure รุ่น PG58

- คอมพิวเตอร์โน้ตบุ๊ค Dell Vostro 5460
- โปรแกรม Adobe Audition CC 2019[3]

---

3 Adobe System Incorporated

การบันทึกเสียงจะทำในสถานที่ที่มีเสียงรบกวนและการสะท้อนของเสียงน้อยเพื่อให้เสียงร้องมีความชัดเจนมากที่สุด สภาวะการบันทึกเสียง แสดงดังรูปที่ 3.1 โดยบันทึกเป็นไฟล์ wav ที่อัตราการชักข้อมูล (sampling rate) 44.1 kHz ความละเอียด (resolution) 16 bits แบบโมโน

รูปที่ 3.1 สภาวะการบันทึกสียง

2. การสกัดลักษณะทางสวนศาสตร์ (Acoustic Feature Extraction)

2.1 การแบ่งพยางค์ (Syllable Segmentation) และกำกับฉลาก (Label)

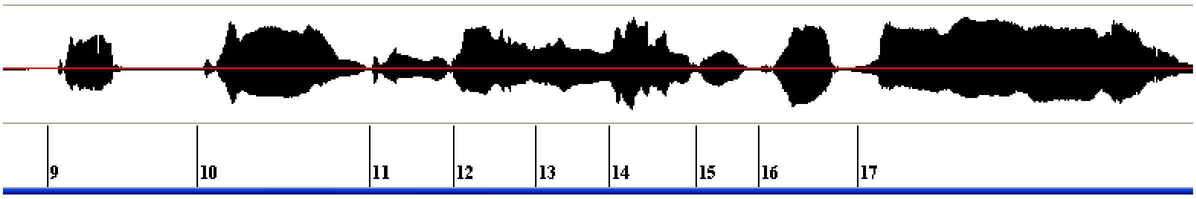
เมื่อได้ข้อมูลเสียงร้องแล้ว จากนั้นจะทำการแบ่งพยางค์ และกำกับฉลาก โดยใช้โปรแกรม Speech Viewer[4] ในการแบ่งพยางค์จะเริ่มตั้งแต่เสียงพยัญชนะต้นไปจนจบพยางค์ ซึ่งระยะเวลาตั้งจุดเริ่มต้นของพยัญชนะต้นไปจนจบพยางค์ คือ **ความยาวของตัวโน้ต**

ส่วนการกำกับฉลาก จะเริ่มจาก “1”, “2”, “3” ไปเรื่อย ๆ จนจบเพลง

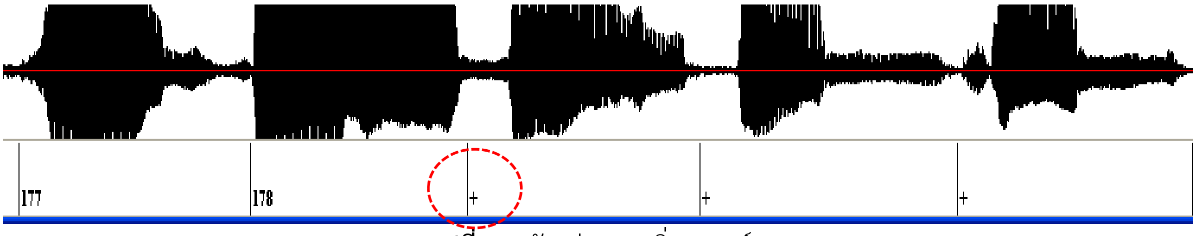
จากนั้นผู้วิจัยได้ตรวจสอบและกำกับการเพิ่มพยางค์ (Addition) การลบพยางค์ (Deletion) และความผิดพลาดของแต่ละพยางค์ โดยใช้สัญลักษณ์ “+”, “-” และ “*” ตามลำดับ แสดงดังรูปที่ 3.2 – 3.5

โปรแกรมจะบันทึกตำแหน่งเวลาเริ่มต้น ตำแหน่งเวลาสุดสิ้นสุด และฉลากของแต่ละพยางค์ ซึ่งฉลากเหล่านี้ จะถูกนำไปคำนวณหา**ค่าความถูกต้องของการร้อง**ต่อไป

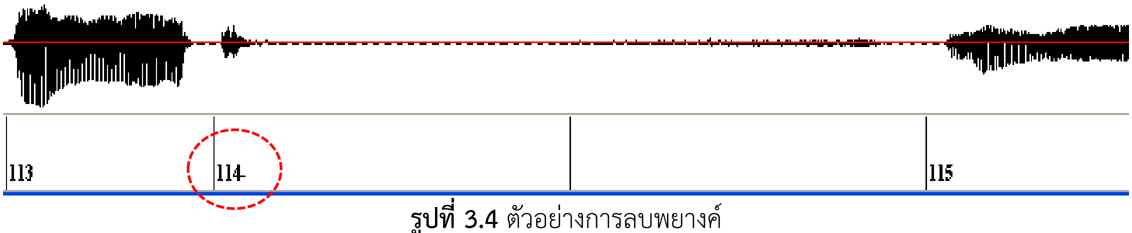
[4] Center for Spoken Language Understanding, Oregon Health & Science University.



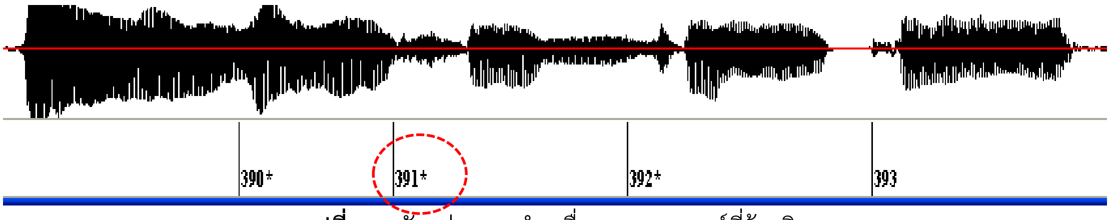
รูปที่ 3.2 ตัวอย่างการแบ่งพยางค์



รูปที่ 3.3 ตัวอย่างการเพิ่มพยางค์



รูปที่ 3.4 ตัวอย่างการลบพยางค์

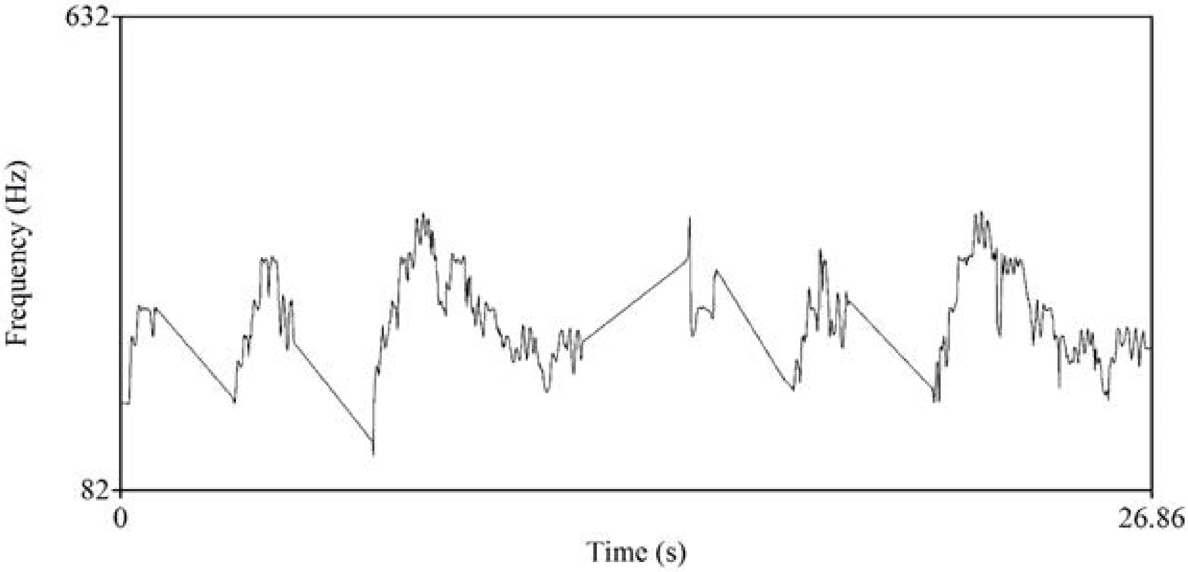


รูปที่ 3.5 ตัวอย่างการทำเครื่องหมายพยางค์ที่ร้องผิด

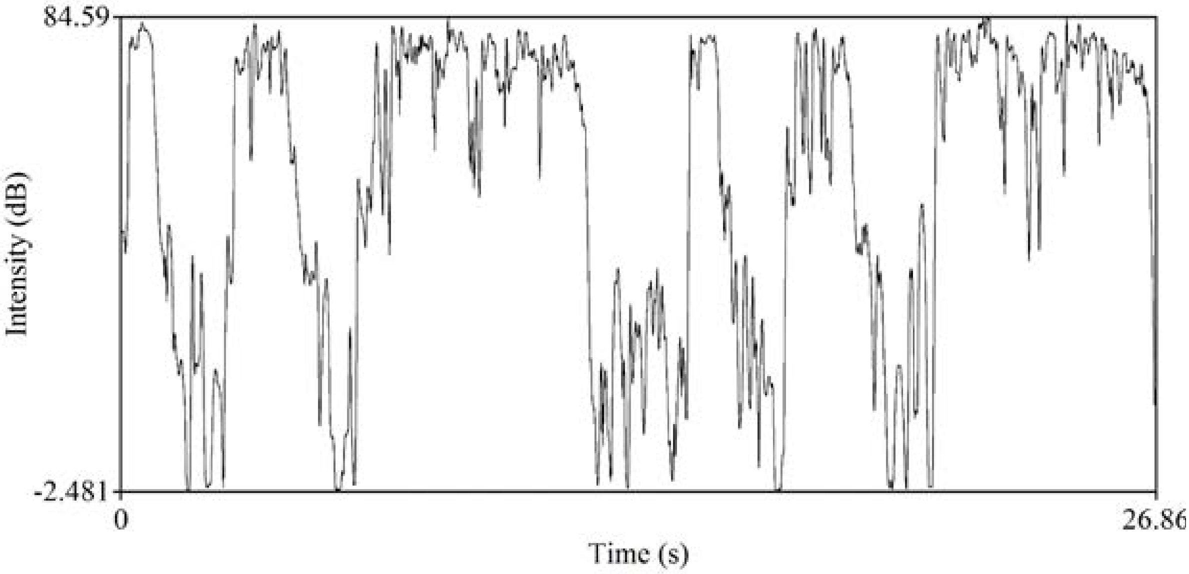
2.2 การสกัดความถี่มูลฐาน และความเข้มเสียง

โปรแกรม Praat[5] ถูกใช้หาความถี่ f_0 ในหน่วยเฮิร์ตซ์ และความเข้มเสียงในหน่วยเดซิเบลของแต่ละพยางค์ โดยกำหนดช่วงความถี่ในช่วงความถี่ 82 ถึง 632 เฮิร์ตซ์ สำหรับนักร้องชาย และ 174 ถึง 1568 เฮิร์ตซ์ สำหรับนักร้องหญิง เพื่อนำไปคำนวณระดับเสียงในรูปของ**เลข MIDI** และ**ความหนักเบา** ตัวอย่างกราฟความถี่มูลฐานและความเข้มเสียง แสดงดังรูปที่ 3.6 และ 3.7

[5] http://www.fon.hum.uva.nl/praat/



รูปที่ 3.6 กราฟความถี่ f_0 ที่ได้จากโปรแกรม Praat



รูปที่ 3.7 กราฟความเข้มเสียงที่ได้จากโปรแกรม Praat

จากนั้นนำตำแหน่งของพยางค์ มาคำนวณค่า**ความยาวของโน้ต** ค่าความถี่ f_0 เฉลี่ยในแต่ละพยางค์ นำไปคำนวณหาค่า**ระดับเสียงของโน้ต**ในรูปแบบตัวเลข MIDI และนำค่าความเข้มเสียงมาปรับบรรทัดฐาน (Normalization) และนำมาหาค่าเฉลี่ยของแต่ละโน้ต เพื่อแทน**ค่าความหนักเบาของโน้ต**

3. การประเมินเสียงร้อง

ข้อมูลความถูกต้อง ความยาวของโน้ต ระดับเสียงของโน้ต และความหนักเบาของตัวโน้ต ของเสียงศิลปินต้นแบบและนักร้องทดสอบ จะนำมาคำนวณหาคะแนนต่าง ๆ สำหรับการประเมินผลซึ่งได้แสดงรายละเอียดไว้ในภาคผนวก ข ดังต่อไปนี้

3.1 คะแนนความถูกต้อง (Correction Score : S_c)

คะแนนความถูกต้องจะให้ในกรณีที่พยางค์นั้นมีการออกเสียงถูกต้องตามเนื้อร้องของเพลง และไม่ให้คะแนนในกรณีของการเพิ่มพยางค์ การลบพยางค์ และการร้องผิด

$$c_i = \begin{cases} 1, & Correct \\ 0, & Addition \\ 0, & Deletion \\ 0, & Wrong \end{cases} \qquad \text{-(3.1)}$$

$$S_c = \frac{1}{N}\sum_{i=1}^{N} c_i \times 100\% \qquad \text{-(3.2)}$$

เมื่อ N คือ จำนวนโน้ตของทั้งเพลง

3.2 คะแนนระดับเสียง (Pitch Score : S_p)

คะแนนระดับเสียงจะได้คะแนนเมื่อโน้ตเสียงนักร้องทดสอบอยู่ในระดับเสียงเดียวกันหรือเป็นจำนวนเท่าของช่วงคู่แปดกับโน้ตเสียงศิลปิน โดยใช้สมการดังนี้

$$p_i = \begin{cases} 1, & when\ m_i\ mod\ 12 - M_i\ mod\ 12 = 0 \\ 0, & when\ m_i\ mod\ 12 - M_i\ mod\ 12 \neq 0 \end{cases} \qquad \text{-(3.3)}$$

เมื่อ m_i และ M_i คือ เลข MIDI พยางค์ที่ i ของนักร้องทดสอบและศิลปินต้นแบบ ตามลำดับ

$$S_p = \frac{1}{N}\sum_{i=1}^{N} p_i \times 100\% \qquad \text{-(3.4)}$$

3.3 คะแนนความยาว (Duration Score : S_t)

คะแนนความยาวจะคิดจากอัตราส่วนของค่าสัมบูรณ์ของผลต่างของความยาวของโน้ตเสียงนักร้องกับเสียงศิลปินต้นแบบกับความยาวของโน้ตเสียงศิลปินต้นแบบ ดังสมการที่ 3.5

$$S_t = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \left|\frac{t_i - T_i}{T_i}\right|\right) \times 100\% \qquad \text{-(3.5)}$$

เมื่อ t_i และ T_i คือ ความยาวของพยางค์ที่ i ของนักร้องทดสอบและศิลปินต้นแบบ ตามลำดับ

3.4 คะแนนความหนักเบา (Dynamic Score : S_d)

คะแนนหนักเบาจะคิดจากอัตราส่วนของค่าสัมบูรณ์ของผลต่างของความหนักเบาของโน้ตเสียงนักร้องกับเสียงศิลปินต้นแบบ กับความหนักเบาของโน้ตเสียงศิลปิน ดังสมการที่ 3.6

$$S_d = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \left|\frac{d_j - D_j}{D_j}\right|\right) \times 100\% \qquad -(3.6)$$

เมื่อ d_i และ D_i คือ ความหนักเบาของพยางค์ที่ i ของนักร้องทดสอบและศิลปิน ตามลำดับ

4. ผลคะแนนรวม

การคำนวณผลคะแนนรวมจะใช้คะแนนทั้งสี่ เพื่อหาค่าเฉลี่ยแบบถ่วงน้ำหนัก (Weighted Average)

$$X_c = \frac{w_c S_c + w_p S_p + w_t S_t + w_d S_d}{w_c + w_p + w_t + w_d} \qquad -(3.7)$$

โดย w_c, w_p, w_t และ w_d เป็นค่าถ่วงน้ำหนักของคะแนนความถูกต้อง ระดับเสียง ความยาวโน้ต และความหนักเบา ตามลำดับ ในการทดลองจะกำหนดให้ w_c, w_p, w_t และ w_d มีค่าเท่ากันเพื่อเปรียบเทียบคะแนนแต่ละส่วนโดยตรง

ผลคะแนนที่ได้จะนำมาเปรียบเทียบกับผลการประเมินโดยผู้เชี่ยวชาญ ซึ่งมีประสบการณ์ทางดนตรีไม่ต่ำกว่า 5 ปี โดยการประเมินทั้ง 4 ด้าน คือ ความถูกต้อง ระดับเสียง ความยาวโน้ต และความหนักเบา

จากนั้นจะคำนวณหาความคลาดเคลื่อนสัมพัทธ์ (Relative Error) ระหว่างการประเมินด้วยคอมพิวเตอร์และการประเมินโดยผู้เชี่ยวชาญ ตามสมการที่ 3.8

$$\eta = \left|\frac{X_h - X_c}{X_h}\right| \times 100\% \qquad -(3.8)$$

โดย X_h และ X_c เป็นคะแนนการประเมินโดยผู้เชี่ยวชาญและคอมพิวเตอร์ ตามลำดับ

บทที่ 4 ผลการทดลอง

ในบทนี้นำเสนอผลการทดลองของนักร้องทดสอบ 6 คน โดยแยกตามเพลงทดสอบ 3 เพลง ซึ่งแสดงค่าคะแนน 4 แบบ คือ คะแนนความถูกต้อง คะแนนระดับเสียง คะแนนความยาว และคะแนนความหนักเบา และคะแนนรวม ที่คำนวณด้วยโปรแกรม Microsoft Excel โดยใช้สมการในบทที่สาม (ขั้นตอนการคำนวณอย่างละเอียดแสดงไว้ในภาคผนวก ข.) โดยเครื่องหมาย “+” , “-” และ “*” แสดงถึงการเพิ่มพยางค์ การลบพยางค์ และการร้องผิดในแบบอื่น ๆ และจะเปรียบเทียบผลการประเมินดังกล่าว (คอมพิวเตอร์) กับผลการประเมินด้วยผู้เชี่ยวชาญ

1. เพลงที่ 1 “รู้”

ผลการประเมินด้วยคอมพิวเตอร์และด้วยผู้เชี่ยวชาญ แสดงดังตารางที่ 4.1 และตารางที่ 4.2ตามลำดับ

จากการประเมินด้วยคอมพิวเตอร์ของเพลงที่ 1 ซึ่งเป็นเพลงที่มีจังหวะช้าและการร้องที่มีเวลาของเสียงที่สั้นแต่เร็วพบว่า ได้คะแนนรวมเฉลี่ย 64.63±6.30 โดยมีคะแนนความถูกต้องเฉลี่ย 97.86±1.33 คะแนนระดับเสียงเฉลี่ย 40.01±5.54 คะแนนความยาวเฉลี่ย 56.01±14.27 และคะแนนความหนักเบาเฉลี่ย 97.96±2.33

จากการประเมินโดยผู้เชี่ยวชาญพบว่า ได้คะแนนรวมเฉลี่ย 46.25±17.06 โดยมีคะแนนความถูกต้องเฉลี่ย 45±22.17 คะแนนระดับเสียงเฉลี่ย 53.33±13.74 คะแนนความยาวเฉลี่ย 51.67±16.75 และมีคะแนนความหนักเบาเฉลี่ย 35±18.03

ตารางที่ 4.1 ผลประเมินโดยคอมพิวเตอร์ของนักร้องในเพลงที่ 1 “รู้” ของ พีรพัฒน์ เถรว่อง

	+	-	*	Correction Score	Pitch Score	Duration Score	Dynamic Score	Total
F1	0	8	6	95.91	30.99	37.89	99.80	54.93
M1	1	0	3	98.83	37.43	34.00	92.88	56.75
M2	0	2	9	96.78	44.74	67.65	99.07	69.72
M3	0	1	4	99.12	39.18	64.20	99.12	67.50
M4	0	4	6	97.08	48.54	65.78	98.04	70.46
M5	0	1	1	99.42	39.18	66.54	98.84	68.38
Mean				**97.86**	**40.01**	**56.01**	**97.96**	**64.63**

ตารางที่ 4.2 ผลประเมินโดยผู้เชี่ยวชาญของนักร้องในเพลงที่ 1 “รู้” ของ พีรพัฒน์ เถรว่อง

	Correction Score	Pitch Score	Duration Score	Dynamic Score	**Total**
F1	40	50	60	20	**42.5**
M1	90	80	80	70	**80**
M2	50	60	60	40	**52.5**
M3	40	50	40	40	**42.5**
M4	20	40	30	20	**27.5**
M5	30	40	40	20	**32.5**
Mean	**45.00**	**53.33**	**51.67**	**35.00**	**46.25**

เมื่อเปรียบเทียบผลการประเมินระหว่างคอมพิวเตอร์และผู้เชี่ยวชาญ พบว่ามีความคลาดเคลื่อนสัมพัทธ์ของคะแนนรวม 46.32% ของคะแนนระดับเสียง 28.53% คะแนนความยาว 17.61% ของคะแนนความถูกต้อง 116.97% และของคะแนนความหนักเบา 161.11%

2. เพลงที่ 2 “แสงสุดท้าย”

ผลการประเมินด้วยคอมพิวเตอร์และด้วยผู้เชี่ยวชาญ แสดงดังตารางที่ 4.3 และตารางที่ 4.5 ตามลำดับ

จากการประเมินด้วยคอมพิวเตอร์ของเพลงที่ 2 ซึ่งเป็นเพลงที่มีการร้องด้วยเสียงที่มีความยาวและใช้ความหนักเบาค่อนข้างคงที่พบว่า ได้คะแนนรวมเฉลี่ย 71.70±5.53 โดยมีคะแนนความถูกต้องเฉลี่ย 96.71±10.46 คะแนนระดับเสียงเฉลี่ย 43.74±1.21 คะแนนความยาวเฉลี่ย 74.65±6.39 และมีคะแนนความหนักเบาเฉลี่ย 94.64±2.65

จากการประเมินโดยผู้เชี่ยวชาญพบว่า ได้คะแนนรวมเฉลี่ย 46.67±17.00 โดยมีคะแนนความถูกต้องเฉลี่ย 45.00±21.17 คะแนนระดับเสียงเฉลี่ย 55.00±13.84 คะแนนความยาวเฉลี่ย 51.67±16.75 และมีคะแนนความหนักเบาเฉลี่ย 35.00±18.03

เมื่อเปรียบเทียบผลการประเมินระหว่างคอมพิวเตอร์และผู้เชี่ยวชาญ พบว่ามีความคลาดเคลื่อนสัมพัทธ์ของคะแนนรวม 55.70% ของคะแนนความถูกต้อง 112.99% ของคะแนนระดับเสียง 24.96% คะแนนความยาว 46.75% และของคะแนนความหนักเบา 154.36%

ตารางที่ 4.3 ผลการประเมินโดยคอมพิวเตอร์ของนักร้องในเพลงที่ 2 “แสงสุดท้าย” ของ วง Bodyslam

	+	-	*	Correction Score	Pitch Score	Duration Score	Dynamic Score	Total
F1	0	8	12	94.27	53.01	83.70	96.90	76.99
M1	1	1	1	99.14	39.43	56.36	95.27	64.98
M2	1	16	10	92.29	34.00	66.46	94.41	64.25
M3	0	1	2	99.71	48.14	82.07	94.48	76.64
M4	0	3	4	97.99	47.56	85.47	93.58	77.01
M5	1	7	3	96.86	40.29	73.84	93.20	70.33
Mean				**96.71**	**43.74**	**74.65**	**94.64**	**71.70**

ตารางที่ 4.4 ผลการประเมินโดยผู้เชี่ยวชาญของนักร้องในเพลงที่ 2 “แสงสุดท้าย” ของ วง Bodyslam

	Correction Score	Pitch Score	Duration Score	Dynamic Score	**Total**
F1	40	60	60	20	**45**
M1	90	80	80	70	**80**
M2	50	60	60	40	**52.5**
M3	40	50	40	40	**42.5**
M4	20	40	30	20	**27.5**
M5	30	40	40	20	**32.5**
Mean	**45.00**	**55.00**	**51.67**	**35.00**	**46.67**

3. เพลงที่ 3 “คู่ชีวิต”

ผลการประเมินด้วยคอมพิวเตอร์และด้วยผู้เชี่ยวชาญ แสดงดังตารางที่ 4.5 และตารางที่ 4.6 ตามลำดับ

จากการประเมินด้วยคอมพิวเตอร์ของเพลงที่ 3 ซึ่งเป็นเพลงที่มีอัตราจังหวะแตกต่างจากเพลงอื่นและมีการเปลี่ยนเครื่องหมายกุญแจเสียงในบางท่อน ซึ่งในเพลงนี้ปรากฏว่านักร้องหญิงไม่สามารถร้องได้ จึงต้องมีการปรับระดับเสียงของนักร้องลง 4 ครึ่งเสียง (ลดความถี่ลง 1.26 เท่า) ก่อนนำมาคิดคะแนนพบว่า ได้คะแนนรวมเฉลี่ย 74.78±7.94 โดยมีคะแนนความถูกต้องเฉลี่ย 98.13±13.24 คะแนนระดับเสียงเฉลี่ย 46.73±1.26 คะแนนความยาวเฉลี่ย 79.48±11.41 และมีคะแนนความหนักเบาเฉลี่ย 99.05±1.16

จากการประเมินโดยผู้เชี่ยวชาญพบว่า ได้คะแนนรวมเฉลี่ย 34.58±21.48 โดยมีคะแนนความถูกต้องเฉลี่ย 28.33±20.34 คะแนนระดับเสียงเฉลี่ย 33.33±22.85 คะแนนความยาวเฉลี่ย 40.00±23.10 และมีคะแนนความหนักเบาเฉลี่ย 34.58±21.34

เมื่อเปรียบเทียบผลการประเมินระหว่างคอมพิวเตอร์และผู้เชี่ยวชาญ พบว่ามีความคลาดเคลื่อนสัมพัทธ์ของคะแนนรวม 126.84% ของคะแนนความถูกต้อง 254.32% ของคะแนนระดับเสียง 62.18% คะแนนความยาว 109.45% และของคะแนนความหนักเบา 163.61%

ตารางที่ 4.5 ผลการประเมินโดยคอมพิวเตอร์ของนักร้องในเพลงที่ 3 “คู่ชีวิต” ของ วง Cocktail

	+	-	*	Correction Score	Pitch Score	Duration Score	Dynamic Score	Total
F1	1	5	9	96.50	26.87	50.13	96.54	**57.83**
M1	1	0	14	96.50	65.42	86.64	99.95	**82.85**
M2	0	1	1	99.30	42.06	83.27	99.93	**74.87**
M3	0	0	2	99.77	49.30	85.46	99.80	**78.17**
M4	0	1	5	98.36	48.60	84.69	98.28	**77.22**
M5	0	1	5	98.36	48.13	86.67	99.77	**77.72**
Mean				**98.13**	**46.73**	**79.48**	**99.05**	**74.78**

ตารางที่ 4.6 ผลการประเมินโดยผู้เชี่ยวชาญของนักร้องในเพลงที่ 3 “คู่ชีวิต” ของ วง Cocktail

	Correction Score	Pitch Score	Duration Score	Dynamic Score	**Total**
F1	30	40	40	40	**37.5**
M1	70	80	80	80	**77.5**
M2	30	30	60	40	**40**
M3	20	20	20	20	**20**
M4	10	20	20	20	**17.5**
M5	10	10	20	20	**15**
Mean	**28.33**	**33.33**	**40**	**36.67**	**34.58**

4. รวมสามเพลง

ผลการประเมินรวมด้วยคอมพิวเตอร์และด้วยผู้เชี่ยวชาญของทั้งสามเพลง แสดงดังตารางที่ 4.7 และตารางที่ 4.6 ตามลำดับ

จากการประเมินด้วยคอมพิวเตอร์ของทั้งสามเพลงพบว่า ได้คะแนนรวมเฉลี่ย 70.37±3.95 โดยมีคะแนนความถูกต้องเฉลี่ย 97.57±1.3416 คะแนนระดับเสียงเฉลี่ย 43.49±4.00 คะแนนความยาวเฉลี่ย 70.05±8.65 และมีคะแนนความหนักเบาเฉลี่ย 97.22±0.67

ตารางที่ 4.7 ผลการประเมินเฉลี่ยรวมทั้งสามเพลงโดยคอมพิวเตอร์

	Correction Score	Pitch Score	Duration Score	Dynamic Score	**Total**
F1	95.56	36.96	57.24	97.75	**63.25**
M1	98.16	47.43	59.00	96.04	**68.19**
M2	96.12	40.26	72.46	97.80	**69.62**
M3	99.53	45.54	77.24	97.80	**74.11**
M4	97.81	48.23	78.65	96.63	**74.90**
M5	98.21	42.53	75.68	97.27	**72.14**
Mean	**97.57**	**43.49**	**70.05**	**97.22**	70.37

ตารางที่ 4.8 ผลการประเมินเฉลี่ยรวมทั้งสามเพลงโดยผู้เชี่ยวชาญ

	Correction Score	Pitch Score	Duration Score	Dynamic Score	**Total**
F1	36.67	50.00	53.33	26.67	**41.67**
M1	83.33	80.00	80.00	73.33	**79.17**
M2	43.33	50.00	60.00	40.00	**48.33**
M3	33.33	40.00	33.33	33.33	**35.00**
M4	16.67	33.33	26.66	20.00	**24.17**
M5	23.33	30.00	33.33	20.00	**26.67**
Mean	**39.44**	**47.22**	**47.78**	**35.56**	**42.5**

จากการประเมินโดยผู้เชี่ยวชาญพบว่า ได้คะแนนรวมเฉลี่ย 42.5±18.36 โดยมีคะแนนความถูกต้องเฉลี่ย 39.44±21.47 คะแนนระดับเสียงเฉลี่ย 47.22±16.49 คะแนนความยาวเฉลี่ย 47.78±18.63 และมีคะแนนความหนักเบาเฉลี่ย 35.56±18.32

เมื่อเปรียบเทียบผลการประเมินระหว่างคอมพิวเตอร์และผู้เชี่ยวชาญ พบว่ามีความคลาดเคลื่อนสัมพัทธ์ของคะแนนรวม 65.58% ของคะแนนความถูกต้อง 147.36% ของคะแนนระดับเสียง 7.90% คะแนนความยาว 46.62% และของคะแนนความหนักเบา 173.43%

อภิปรายผลการทดลอง

จากการทดลองที่อยู่ในภาวะเหมาะสมใกล้เคียงอุดมคติ คือ มีเสียงรบกวนน้อยทำให้มีอัตราส่วนสัญญาณต่อสัญญาณรบกวน (Signal to Noise Ratio) ที่สูง ซึ่งในภาวะปกติอาจมีเสียงรบกวนต่าง ๆ แทรกเข้ามาทำให้การสกัดคุณลักษณ์ทางสวนศาสตร์มีคุณภาพลดลง ส่งผลให้คะแนนในส่วนต่าง ๆ เปลี่ยนไปได้

จากการประเมินโดยผู้เชี่ยวชาญ คะแนนจากการประเมินมีความสัมพันธ์กับประสบการณ์ในการร้องเพลง โดยนักร้องที่มีประสบการณ์มากจะมีระดับคะแนนที่สูงกว่าอย่างเห็นได้ชัด เมื่อเปรียบเทียบกับผลการประเมินจากคอมพิวเตอร์พบว่า คอมพิวเตอร์สามารถคำนวณคะแนนระดับเสียงได้ใกล้เคียงกับการประเมินด้วยมนุษย์ในระดับความผิดพลาดร้อยละ 7.90 แต่การประเมินคะแนนส่วนอื่นยังมีความคลาดเคลื่อนอยู่มาก

คะแนนระดับเสียงเฉลี่ยของนักร้องทดสอบทั้งหมดมีค่าต่ำที่สุดในทุกเพลงที่นำมาทดสอบ อาจเกิดจากอัตราการเปลี่ยนแปลงของความถี่ f_0 ในพยางค์เดียวกันของนักร้องทดสอบแต่ละคนมีความแตกต่างกันประกอบกับกระบวนการที่ใช้ในการสกัดเสียงศิลปินต้นฉบับไม่สามารถสกัดเสียงนักร้องออกมาได้อย่างสมบูรณ์จึงทำให้ค่าระดับเสียงที่ได้จากการคำนวณที่ใช้ฟังก์ชันในการปัดเศษทศนิยมเป็นจำนวนเต็มมีความคลาดเคลื่อนไป

คะแนนความยาวเฉลี่ยของนักร้องทดสอบทั้งหมดมีคะแนนต่ำ อาจเกิดจากนักร้องมีความสามารถในการกักเก็บลมหายใจไม่เท่ากันทำให้มีช่วงเว้นในการหายใจไม่เท่ากัน ประกอบกับการร้องผิดอาจทำให้สัดส่วนจังหวะในการร้องผิดแผกไปจากเดิม

เมื่อพิจารณาคะแนนความถูกต้องและความหนักเบาเฉลี่ยของนักร้องทดสอบทั้งหมดพบว่ามีคะแนนสูง โดยคะแนนความถูกต้องมักจะเสียคะแนนจากการร้องผิดมากกว่าจากการลบพยางค์ ซึ่งเป็นได้ทั้งการแทนพยางค์ การเปลี่ยนพยางค์ หรือการกลืนพยางค์ และส่วนน้อยที่มีการเพิ่มพยางค์ โดยมักเกิดในท่อนที่มีเสียงพยัญชนะต้นยาวหรือร้องผิดท่อนของเพลงที่มีเนื้อร้องคล้ายๆกัน และการร้องผิดอาจส่งผิดให้ความหนักเบาในพยางค์นั้นๆ มีความผิดเพี้ยนไป

บทที่ 5 สรุปผลโครงงาน

โครงงานนี้มีจุดประสงค์เพื่อสร้างระเบียบวิธีในการประเมินการร้องเพลงที่ใช้ลักษณะทางสวนศาสตร์และใช้ข้อมูลเสียงร้องของศิลปินต้นฉบับ 3 เพลง เพื่อใช้เป็นเกณฑ์อ้างอิงในการให้คะแนนการร้องเพลง ได้แก่ (1) เพลงรู้ ของ บี พีรพัฒน์ เถรว่อง (2) เพลงแสงสุดท้าย ของวง Bodyslam และ (3) เพลงคู่ชีวิต ของวง Cocktail โดยใช้ข้อมูลเสียงจากนักร้องทดสอบจำนวน 6 คน นำมาสกัดลักษณะทางสวนศาสตร์ ได้แก่ ระดับเสียง (Pitch) ความหนักเบา (Dynamic) และความหนักเบา (Dynamic) จากนั้นนำลักษณะเหล่านี้ พร้อมทั้งความถูกต้องของการร้อง มาใช้ประเมินการร้องด้วยระเบียบวิธีที่นำเสนอ

จากการทดลองพบว่า ระเบียบวิธีการประเมินเสียงร้องโดยใช้ลักษณะทางสวนศาสตร์สามารถใช้ประเมินเสียงร้องของนักร้องได้ โดยมีคะแนนรวมเฉลี่ย 70.37±3.95 แบ่งเป็นค่าคะแนนความถูกต้อง 97.57±1.34 คะแนนระดับเสียง 43.49±4.00 ค่าคะแนนความยาว 70.05±8.65 และค่าคะแนนความหนักเบา 97.22±0.67

เมื่อเทียบกับการประเมินโดยผู้เชี่ยวชาญซึ่งมีผลการประเมินคะแนนรวมเฉลี่ย 42.5±18.36 โดยมีคะแนนความถูกต้องเฉลี่ย 39.44±21.47 คะแนนระดับเสียงเฉลี่ย 47.22±16.49 คะแนนความยาวเฉลี่ย 47.78±18.63 และคะแนนความหนักเบาเฉลี่ย 35.56±18.32 พบว่ามีความคลาดเคลื่อนสัมพัทธ์ของคะแนนรวม 65.58% และมีความแม่นยำในการประเมินคะแนนระดับเสียงโดยมีความคลาดเคลื่อนสัมพัทธ์ 7.90%

อุปสรรคและข้อเสนอแนะ

1. การแบ่งพยางค์ที่มีการลบอาจทำให้เวลาสิ้นสุดของพยางค์ก่อนหน้าและเวลาเริ่มต้นของพยางค์ถัดไปมีความคลาดเคลื่อน
2. การคำนวณการเพิ่มของพยางค์ด้วยโปรแกรม Microsoft Excel มีความยุ่งยาก ควรใช้การเขียนโปรแกรมในภาษาอื่นเพื่อช่วยในการคำนวณคะแนน
3. ในขั้นตอนการสกัดลักษณะทางสวนศาสตร์ด้วยระเบียบวิธีนี้ควรพัฒนาให้สามารถวิเคราะห์เทคนิคในการร้องเพลงได้ เช่น การสั่นเสียง (Vibrato) เป็นต้น

เอกสารอ้างอิง

[1] Juan, G. Roederer (2008), *The Physics and Psychophysics of Music: An Introduction,* (4th ed.) Springer, ISBN 978-0-387-09470-0.

[2] Titze, I. R. (1994), *Principles of Voice Production*, Prentice Hall (currently published by NCVS.org), ISBN 978-0-13-717893-3.

[3] Peckham, A. (2005), *Vocal Workouts for the Contemporary Singer*. Berklee Press Publications, ISBN 978-0-87639-047-4.

[4] Tsai, W., and Lee, H. (2012), “Automatic Evaluation of Karaoke Singing Based on Pitch, Dynamic, and Rhythm Features”, *IEEE Transactions on Audio, Speech, and Language Processing*, 20(4), 1233-1243.

[5] ISO 16:1975 - Acoustics – “Standard tuning frequency (n.d.)”. Retrieved from https://www.iso.org/standard/3601.html

[6] Vail, M. (2014), *The Synthesizer: A Comprehensive Guide to Understanding, Programming, Playing, and Recording the Ultimate Electronic Music Instrument*. OUP USA. ISBN 978-0-19539-489-4.

[7] Benward, B. and Marilyn S. (2003), *Music: In Theory and Practice*, (7th edition), vol. 1, Boston: McGraw-Hill. ISBN 978-0-07-294262-0.

ภาคผนวก

ก. เนื้อเพลงที่ใช้ในการทดสอบ

เพลงลำดับที่ 1 : รู้ – บี พีรพัฒน์ เถรว่อง

(1) รู้ แต่มันทำไม่ได้
อยากจะทำแค่ไหน ทำอย่างไร ฉันก็ทำ ทำอย่างนั้น ทำไม่ไหว
รู้ ต้องปล่อยเธอไปซะที
แต่ก็ยังไม่รู้ ว่าจะทำ ฉันจะทำ ทำอย่างนั้น ได้อย่างไร

(2) พยายามทำความเข้าใจ อยู่ในสมอง
คำบางคำมันลอยวนเวียน อยู่เป็นคำถาม
ย้ำซ้ำๆ ว่าทำไม ทำไม เหตุผลคืออะไรบอกเลิกกันง่ายๆ
ทางที่เราเคยเดินด้วยกันยังมีความหมาย
และทุกภาพทรงจำของฉันและเธอ
ก็ยังเป็นความทรงจำที่มันยังตราตรึงอยู่ในใจ

(3) รู้ แต่มันทำไม่ได้
อยากจะทำแค่ไหน ทำอย่างไร คิดจะทำ ทำอย่างนั้น ทำไม่ไหว
รู้ ต้องปล่อยเธอไปซะที
แต่ก็ยังไม่รู้ ว่าจะทำ ฉันจะทำ ทำอย่างนั้นได้อย่างไร

(4) ในนาทีที่เธอ กำลังบอกลากับฉัน
ภายในใจมันยังคิดว่าเป็นแค่ฝัน
ฉันไม่รู้ว่าทำไม ทำไม และฉันไม่เข้าใจยังรับมันไม่ได้
ดาวที่เราเคยมองด้วยกันก็ยังสดใส
รูปของฉันและเธอดูแล้วก็ยังซึ้งใจ
แต่ละเวลานาทีความรักฉันไม่เคยเปลี่ยนไป

(5) รู้ แต่มันทำไม่ได้

อยากจะทำแค่ไหน ทำอย่างไร ฉันก็ทำ ทำอย่างนั้น ทำไม่ไหว

รู้ ต้องปล่อยเธอไปซะที

แต่ก็ยังไม่รู้ว่าจะทำ ฉันจะทำ ทำอย่างนั้นได้อย่างไร

(6) ก็ฉันไม่รู้และไม่เข้าใจ อะไรกับเหตุผล

และไม่รู้จะทำยังไงให้เธออยู่กับฉัน

รู้ว่าต้องทำใจแต่ทำไม่ได้

ฉันไม่รู้จะห้ามยังไงให้เธอไม่ไปไหน

และไม่รู้จะทำยังไงกับใจที่สลาย

และมันคือความเป็นจริง

ที่เลวร้ายเกินใจจะรับไหว

(7) รู้ แต่มันทำไม่ได้

อยากจะทำแค่ไหน ทำอย่างไร ฉันก็ทำ ทำอย่างนั้น ทำไม่ไหว

รู้ ต้องปล่อยเธอไปซะที

แต่ก็ยังไม่รู้ ว่าจะทำ ฉันจะทำ ทำอย่างนั้นได้อย่างไร

เมื่อหัวใจมันไม่ยอม

เพลงลำดับที่ 2 : แสงสุดท้าย - วง Bodyslam

(1) รอนแรมมาเนิ่นนาน เพียงหนึ่งใจ
กับทางที่โรยเอาไว้ด้วยขวากหนาม
ถูกแหลมคมทิ่มแทง
จนมันแทบจะทนไม่ไหว
ชีวิต ทำไมยากเย็นขนาดนั้น
สองมือจะมีเรี่ยวแรงขนาดไหน
แต่หัวใจของคน
ยังยืนยันจะไม่ถอดใจ

(2) ในค่ำคืนที่ฟ้านั้นไม่มีดาวอยู่ตรงนี้
ฉันยังคงก้าวไป
ยังคงมีรักแท้เป็นแสงนำไป
ในคืนที่หลงทาง

(3) วันเวลาไม่เคยจะหยุดเดิน
อย่างไรเราคงต้องเดินไปกับมัน
เก็บทุกความผิดพลั้ง
เป็นคำเตือนให้เราเข้าใจ
ชีวิตเริ่มตรงที่คำว่าฝ่าฟัน
ขอเพียงใจเราเท่านั้นไม่หวั่นไหว
บทชีวิตของเรา
เราจะทำให้มีความหมาย

(4) ในค่ำคืนที่ฟ้านั้นไม่มีดาวอยู่ตรงนี้
ฉันยังคงก้าวไป
ยังคงมีรักแท้เป็นแสงนำไป
ในคืนที่หลงทาง

(5) นาทีที่ความฝันนั้นพร้อมเป็นเพื่อนตาย
เส้นทางนี้ฉันยังมีจุดหมาย ตราบใดที่ปลายท้องฟ้ามีแสงรำไร
จะไปจนถึงแสงสุดท้าย...

(6) ความเดียวดายในคืนเหน็บหนาว แหงนมองฟ้ายังนึกถึงวันเก่า
มันคงจริงที่ทางยาวไกล กร่อนหัวใจ
ภาวนากับความมืดมิด ขอให้รักยังคุ้มครองเราอยู่
เติมพลังให้ใจดวงนี้ ไม่ยอมแพ้

(7) ในค่ำคืนที่ฟ้าท้าทายใจคนอยู่ตรงนี้ และฉันยังคงก้าวไป
ยังคงมีรักแท้เป็นแสงนำไปในคืนที่หลงทาง

(8) นาทีที่ความฝันนั้นพร้อมเป็นเพื่อนตาย
เส้นทางนี้ฉันยังมีจุดหมาย ตราบใดที่ปลายท้องฟ้ามีแสงรำไร
จะไปจนถึงแสงสุดท้าย... จนถึงแสงสุดท้าย...

(9) ตราบใดที่ปลายท้องฟ้า... ตราบใดที่ปลายท้องฟ้า...

เพลงลำดับที่ 3 : คู่ชีวิต - วง Cocktail

(1) เธอคือทุกสิ่ง ในความจริงในความฝัน
คือทุกอย่าง เหมือนใจต้องการ
เธอเป็นนิทาน ที่ฉันอ่าน ก่อนหลับตาและนอนฝัน

(2) เธอคือหัวใจ ไม่ว่าใครไม่อาจเทียมเทียบเท่าเธอ
ช่างโชคดีที่เจอ ได้ตกหลุมรักเธอ ได้มีเธอ เคียงข้างกัน

(3) คงจะมีเพียงเธอทำให้โลกนั้นหยุดหมุน เพียงเธอสบตาฉัน
คงจะมีเพียงเธอที่หยุดหัวใจของฉันไว้ตรงนี้ ตรงที่เธอ

(4) เธอเพียงคนเดียวและเพียงเธอที่ต้องการ
ฉันจะทำทุกๆ ทางด้วยวิญญาณและหัวใจ
นั่นคือฉันจะรักเธอไม่ว่าเป็นเมื่อไรสถานใด
ทั้งหัวใจฉันมีเธอเพียงคนเดียว

(5) เธอคือรักจริง ฉันยอมทิ้งทุกๆ อย่างเพียงเพื่อเธอ
ดั่งฟ้าให้มาเจอ ให้เธอคู่กับฉัน ให้เราได้เดินเคียงข้างกันนับจากนี้

(6) คงจะมีเพียงเธอทำให้โลกนั้นหยุดหมุน เพียงเธอสบตาฉัน
คงจะมีเพียงเธอที่หยุดหัวใจของฉันไว้ตรงนี้ ตรงที่เธอ
เธอเพียงคนเดียวและเพียงเธอที่ต้องการ
ฉันจะทำทุกๆ ทางด้วยวิญญาณและหัวใจ
นั่นคือฉันจะรักเธอไม่ว่าเป็นเมื่อไรสถานใด
ทั้งหัวใจฉันมีเธอเพียง

(7) เธอเพียงคนเดียวและเพียงเธอที่เฝ้ารอ
ฉันจะขอภาวนา ต่อหน้าฟ้าอันแสนไกล
นั่นคือฉันจะรักเธอไม่ว่าเป็นเมื่อไรสถานใด
ทั้งหัวใจฉันมีเธอเพียงคนเดียว

(8) จะทุกข์หรือยามที่เธอนั้นสุขใจ
ยามป่วยไข้หรือสุขกายสบายดี
ฉันอยู่ตรงนี้และจะมีเพียงเธอทุกวินาที
จะอยู่ใกล้ไม่ห่างไกล จะเคียงชิดไม่ห่างไป ไม่ไปไหน

(9) เธอเพียงคนเดียวและเพียงเธอที่ต้องการ
ฉันจะทำทุกๆ ทางด้วยวิญญาณและหัวใจ
นั่นคือฉันจะรักเธอไม่ว่าเป็นเมื่อไรสถานใด
ทั้งหัวใจฉันมีเธอเพียง

(10)เธอเพียงคนเดียวและเพียงเธอ เพียงเธอที่รอ
ฉันขอภาวนาต่อหน้าฟ้าอันแสนไกล
นั่นคือฉันจะรักเธอไม่ว่าเป็นเมื่อไรสถานใด
เกิดชาติไหนฉันมีเธอ มีเธอเพียง คนเดียว

ข. ระเบียบวิธีการการคำนวณ

ในภาคผนวก ข จะแสดงระเบียบวิธีที่ใช้ในการประเมินคะแนนเสียงร้องของนักร้องด้วยโปรแกรม Microsoft Excel โดยละเอียด โดยเครื่องหมาย “+” “-” และ “*” หลังหมายเลขพยางค์หมายถึงการเพิ่ม การตัว และการผิด ตามลำดับ

ตารางที่ ข.1 ตัวอย่างผลการสกัดลักษณะทางสวนศาสตร์ 50 พยางค์แรก ด้วยโปรแกรม Speech Viewer และ Praat ของศิลปินต้นฉบับ

O				
label	T_start (ms)	T_stop (ms)	Freq (Hz)	Intensity (dB)
1	44.22735	45.47589	170.9342	61.31191
2	45.47589	46.24761	190.6625	64.09704
3	46.24761	46.64978	218.7511	51.00336
4	46.64978	47.09557	176.8936	67.94952
5	47.09557	47.47927	125.356	57.99494
6	47.47927	47.92657	145.3487	56.53483
7	47.92657	48.33594	174.4628	65.06402
8	48.33594	48.71941	194.0649	72.40613
9	48.71941	49.15611	214.5054	67.29223
10	49.15611	50.52556	202.4809	64.18727
11	50.52556	51.35444	193.1139	67.98534
12	51.35444	51.71906	211.1705	60.12363
13	51.71906	52.71959	171.3543	61.29567
14	52.71959	53.3747	218.2207	68.89985
15	53.3747	53.7584	268.4217	71.16416
16	53.7584	54.18238	206.3177	71.32302
17	54.18238	55.12029	185.1331	69.79307
18	55.42653	55.84995	128.6725	61.97152
19	55.84995	56.2408	175.7785	71.26128
20	56.2408	56.68169	187.2082	65.73771
21	56.68169	57.93639	127.2709	60.9097
22	58.35463	58.76294	263.662	55.77998

23	58.76294	59.16368	173.0952	68.24489
24	59.16368	60.4439	129.653	65.42515
25	60.87064	61.21846	113.4987	73.02204
26	61.21846	61.61165	139.4182	64.9786
27	61.61165	62.08081	173.4359	70.57924
28	62.08081	62.4018	144.6458	70.11943
29	62.4018	62.91561	168.1583	67.91486
30	62.91561	63.89145	181.3399	72.50507
31	64.23255	65.49409	171.3102	63.92036
32	65.49409	66.26038	188.5405	74.82191
33	66.26038	66.70947	202.7375	70.27989
34	66.70947	67.0928	175.0152	75.08968
35	67.0928	67.51756	127.6694	76.12439
36	67.51756	67.93338	141.276	70.50272
37	67.93338	68.356	167.7141	77.25904
38	68.356	68.74563	138.6485	75.14772
39	68.74563	69.24417	128.4886	63.85019
40	69.24417	70.48763	173.5032	63.86575
41	70.48763	71.29601	190.6855	67.66871
42	71.29601	71.73305	200.8951	69.78213
43	71.73305	72.69041	171.589	65.52713
44	72.69041	72.92743	183.4019	60.69739
45	72.92743	73.36927	275.6341	75.83305
46	73.36927	73.79165	260.984	63.36176
47	73.79165	74.07981	197.335	56.28744
48	74.07981	75.26997	183.6466	69.08615
49	75.26997	75.48197	163.5629	65.11125
50	75.48197	75.8271	259.8155	66.19685

ตารางที่ ข.2 ตัวอย่างผลการสกัดลักษณะทางสวนศาสตร์ 50 พยางค์ แรกด้วยโปรแกรม Speech Viewer และ Praat ของนักร้อง

F1 (+4)				
label	T_start (ms)	T_stop (ms)	Freq (Hz)	Intensity (dB)
1	44.14395	45.41279	216.3725	71.419
2	45.41279	46.32451	245.6653	68.15175
3	46.32451	46.62403	263.886	65.80072
4	46.62403	47.17864	204.1711	69.14106
5	47.17864	47.55221	163.6102	61.92287
6	47.55221	48.02124	183.8784	68.05357
7	48.02124	48.46393	208.425	69.5145
8	48.46393	48.80788	185.7144	66.07219
9	48.80788	49.18475	165.7794	63.85513
10	49.18475	50.47663	216.8281	60.98317
11	50.47663	51.26986	241.339	71.33533
12	51.26986	51.60394	260.0157	70.96024
13	51.60394	52.82177	218.4787	61.79478
14	52.82177	53.41916	282.9505	69.74013
15	53.41916	53.86514	342.2139	69.93687
16	53.86514	54.25024	251.5818	69.97477
17	54.25024	55.48617	224.3965	64.47256
18	55.48617	55.94696	162.1798	64.13139
19	55.94696	56.35181	187.8483	70.84247
20	56.35181	56.68169	218.1109	62.42499
21	56.68169	57.75724	163.9662	53.32753
22	58.35463	58.76294	184.9073	67.85012
23	58.76294	59.22357	214.3985	65.53351
24	59.22357	60.47102	160.0648	62.3301
25	60.87064	61.29223	171.9856	66.02986
26	61.29223	61.68061	194.527	47.11219
27	61.68061	62.12989	214.8422	72.49645

28	62.12989	62.49359	187.4197	72.68254
29	62.49359	62.91785	222.8889	73.77381
30	62.91785	64.14423	233.9169	70.99243
31	64.16071	65.42511	217.7535	63.07391
32	65.42511	66.27359	244.9264	73.83259
33	66.27359	66.62598	262.1405	73.92966
34	66.62598	67.0928	215.213	70.66701
35	67.0928	67.55466	170.2948	69.05915
36	67.55466	67.93954	192.8897	67.20108
37	67.93954	68.40504	219.4277	72.46869
38	68.40504	68.7672	176.2332	70.70836
39	68.7672	69.19291	162.4061	60.71786
40	69.19291	70.42555	217.831	64.29815
41	70.42555	71.27377	243.1102	73.86866
42	71.27377	71.6963	263.1477	75.7986
43	71.6963	72.58583	219.8701	61.11571
44	72.58583	72.79551	346.361	62.31808
45	72.79551	73.38959	347.1223	72.8893
46	73.38959	73.74222	331.1978	61.99415
47	73.74222	74.14251	258.6529	66.14141
48	74.14251	75.26997	225.2093	68.35005
49-	75.26997	75.50858	151.9252	43.60507
50	75.50858	75.83262	165.3712	64.92814

ตรวจสอบความถูกต้องโดยใช้เงื่อนไขตรวจสอบว่าหมายเลขพยางค์เป็นตัวเลขหรือไม่

$$C = if(isnumber(label) = true, 1,0)$$

โดย *isnumber* (*x*) เป็นฟังก์ชันตรวจสอบความเป็นตัวเลข

นำค่าความถี่ในตารางมาคำนวณหาระดับเสียงของโน้ตในรูปแบบ MIDI ในกรณีที่มีการปรับกุญแจเสียงขึ้นไป x ครึ่งเสียงให้นำไปหักลบโดยตรงได้ทันที และกรณีที่เกิดการลบพยางค์จะทำให้ระดับเสียงพยางค์นั้นๆมีค่าเป็นศูนย์

$$pitch = 12 * \log\left(\frac{freq}{440}\right) + 69 - x$$

หาค่าเวลาของพยางค์โดยใช้ผลต่างของเวลาสุดท้ายและเวลาเริ่มต้น ในกรณีที่เกิดการลบพยางค์จะทำให้ค่าเวลาพยางค์นั้น ๆ มีค่าเป็นศูนย์

$$duration = if(right(label, 1) = \text{"-"}, 0, (t_{end} - t_{start}))$$

โดย right(text,n) เป็นฟังก์ชันที่อ่านค่าตัวอักษรจากทางด้านขวา n ตัว

คำนวณหาค่าความหนักเบาโดยใช้ค่าค่าเข้มเสียงหารด้วยความเข้มเสียงสูงสุด

$$loudness = \frac{I}{\max(I)}$$

โดย max เป็นฟังก์ชันในการหาค่าสูงสุด

ตารางที่ ข.3 ผลการคำนวณของเสียงศิลปินต้นฉบับโดยใช้สมการดังกล่าว

O				
label	C	pitch	Duration (s)	Dynamic
1	100	52.63125	1.248539	0.123487
2	100	54.52222	0.771721	0.129096
3	100	56.90144	0.402168	0.102725
4	100	53.22454	0.44579	0.136856
5	100	47.26232	0.383701	0.116806
6	100	49.82417	0.4473	0.113866
7	100	52.98499	0.409371	0.131044
8	100	54.82843	0.383468	0.145832
9	100	56.56212	0.436699	0.135532
10	100	55.56338	1.369453	0.129278
11	100	54.74339	0.828877	0.136928
12	100	56.29086	0.364624	0.121094
13	100	52.67374	1.000526	0.123454
14	100	56.85941	0.65511	0.13877
15	100	60.44398	0.383701	0.14333
16	100	55.88837	0.423981	0.14365

17	100	54.01271	0.937916	0.140569
18	100	47.7144	0.423428	0.124815
19	100	53.11506	0.390842	0.143526
20	100	54.20569	0.440899	0.132401
21	100	47.52479	1.2547	0.122677
22	100	60.13423	0.40831	0.112345
23	100	52.84874	0.400745	0.137451
24	100	47.84583	1.28022	0.131771
25	100	45.54207	0.34782	0.147072
26	100	49.10298	0.393185	0.130872
27	100	52.88279	0.469169	0.142152
28	100	49.74025	0.320984	0.141226
29	100	52.3478	0.513809	0.136786
30	100	53.65431	0.975846	0.146031
31	100	52.66929	1.261535	0.128741
32	100	54.32845	0.766297	0.150697
33	100	55.58531	0.449089	0.141549
34	100	53.03972	0.383331	0.151236
35	100	47.57891	0.424759	0.15332
36	100	49.33215	0.415817	0.141998
37	100	52.30201	0.422615	0.155606
38	100	49.00714	0.389633	0.151353
39	100	47.68964	0.498543	0.128599
40	100	52.88951	1.243454	0.128631
41	100	54.5243	0.808388	0.13629
42	100	55.42726	0.437034	0.140547
43	100	52.69744	0.95736	0.131977
44	100	53.85006	0.237022	0.122249
45	100	60.90301	0.441841	0.152734
46	100	59.9575	0.422379	0.127615
47	100	55.11772	0.288162	0.113367

48	100	53.87315	1.190155	0.139145
49	100	51.8681	0.212006	0.131139
50	100	59.87981	0.345131	0.133326

ตารางที่ ข.4 ผลการคำนวณของเสียงนักร้องที่ทดสอบโดยใช้สมการดังกล่าว

F1 (+4)				
label	C	pitch	Duration (s)	Dynamic
1	100	52.71216	1.26884	0.122008
2	100	54.91029	0.911724	0.116427
3	100	56.14894	0.299519	0.11241
4	100	51.7073	0.554604	0.118117
5	100	47.87311	0.373573	0.105786
6	100	49.89498	0.469028	0.116259
7	100	52.0643	0.442696	0.118755
8	100	50.06699	0.343952	0.112874
9	100	48.10113	0.376866	0.109087
10	100	52.74858	1.291881	0.10418
11	100	54.60269	0.793232	0.121865
12	100	55.89314	0.33408	0.121225
13	100	52.87987	1.217823	0.105567
14	100	57.35656	0.597393	0.11914
15	100	60.64874	0.445984	0.119476
16	100	55.32229	0.385097	0.119541
17	100	53.34256	1.235928	0.110141
18	100	47.72109	0.460796	0.109559
19	100	50.26477	0.404846	0.121023
20	100	52.8507	0.329884	0.106643
21	100	47.91074	1.07555	0.091102
22	100	49.99159	0.40831	0.115911
23	100	52.55349	0.460636	0.111954
24	100	47.49383	1.247444	0.106481

25	100	48.73741	0.421585	0.112802
26	100	50.8696	0.388386	0.080484
27	100	52.58929	0.449279	0.123849
28	100	50.22523	0.363701	0.124167
29	100	53.22586	0.424263	0.126031
30	100	54.06191	1.22638	0.12128
31	100	52.82231	1.264404	0.107752
32	100	54.85814	0.84848	0.126132
33	100	56.03404	0.352387	0.126297
34	100	52.61914	0.466827	0.120724
35	100	48.56637	0.461853	0.117977
36	100	50.72327	0.38488	0.114803
37	100	52.9549	0.4655	0.123802
38	100	49.15979	0.362167	0.120794
39	100	47.74522	0.425705	0.103727
40	100	52.82847	1.232636	0.109843
41	100	54.72928	0.848228	0.126193
42	100	56.10043	0.422531	0.12949
43	100	52.98977	0.889527	0.104407
44	100	60.85728	0.209678	0.106461
45	100	60.89528	0.594078	0.12452
46	100	60.08227	0.352631	0.105907
47	100	55.80216	0.400291	0.112992
48	100	53.40515	1.127457	0.116765
49-	0	0	0	0
50	100	48.05846	0.324044	0.11092

คำนวณคะแนนระดับเสียงจากค่าเฉลี่ยของคะแนนระดับเสียงย่อย p ในแต่ละพยางค์ ที่คำนวณจากผลต่างระหว่างมอดูโลฐานสิบสองของโน้ตของนักร้องกับต้นฉบับ และตรวจสอบความเป็นตัวเลขสำหรับกรณีที่ได้ความถี่จากโปรแกรม Praat เป็นศูนย์

$$p = iferror((if(mod(round(O_{pitch}), 12) = mod(round(F1_{pitch}, 0), 12), 1, 0), 0)$$

$$S_p = average(p)$$

โดย iferror(A,B) เป็นฟังก์ชันตรวจสอบข้อผิดพลาดของสูตร A หากพบให้คืนค่า B

และ round(x) เป็นฟังก์ชันที่ปัดเลขทศนิยมเป็นเลขจำนวนเต็ม

คำนวณคะแนนความยาวจากค่าเฉลี่ยของคะแนนความยาวย่อย T ในแต่ละพยางค์ ที่คำนวณจากค่าสัมบูรณ์ผลต่างของความยาวของนักร้องกับต้นฉบับหารด้วยความยาวของต้นฉบับ

$$\mathrm{T} = \mathrm{if}\left(\mathrm{right}(label, 1) = "-", 0, \mathrm{iferror}\left(1 - \frac{\mathrm{ABS}(\mathrm{F1}_{\mathrm{duration}} - \mathrm{O}_{\mathrm{duration}})}{\mathrm{O}_{\mathrm{duration}}}, 0\right), 0\right)$$

$$S_t = average(T)$$

คำนวณคะแนนความถูกต้องจากค่าเฉลี่ยของคะแนนความถูกต้องย่อย C ในแต่ละพยางค์

$$S_c = average(C)$$

คำนวณคะแนนความหนักเบาจากค่าเฉลี่ยของคะแนนเความหนักเบาย่อย V ในแต่ละพยางค์ ที่คำนวณจากค่าสัมบูรณ์ผลต่างของความหนักเบาของนักร้องกับต้นฉบับหารด้วยความหนักเบาของต้นฉบับ

$$\mathrm{D} = \mathrm{if}\left(\mathrm{right}(label, 1) = "-", 0, \mathrm{iferror}\left(1 - \frac{\mathrm{ABS}\left(\mathrm{F1}_{\mathrm{Dynamic}} - \mathrm{O}_{\mathrm{Dynamic}}\right)}{\mathrm{O}_{\mathrm{Dynamic}}}, 0\right), 0\right)$$

$$S_d = average(D)$$

คำนวณคะแนนรวมจากคะแนนแต่ละส่วนโดยใช้ค่าเฉลี่ยแบบถ่วงน้ำหนัก

$$X_c = \frac{w_c S_c + w_p S_p + w_t S_t + w_d S_d}{w_c + w_p + w_t + w_d}$$