การจำแนกกลุ่มเพลงไทยเดิม

โดย

นายโอภาส แก้วต่าย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

ภาควิชาคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2552

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การจำแนกกลุ่มเพลงไทยเดิม

โดย
นายโอภาส แก้วต่าย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
ภาควิชาคอมพิวเตอร์
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552
ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

THAI CLASSICAL MUSIC CLASSIFICATION

By

O-pas Kaewtai

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Computing

Graduate School

SILPAKORN UNIVERSITY

2009

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “ การจำแนกกลุ่มเพลงไทยเดิม ” เสนอโดย นายโอภาส แก้วต่าย เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

……..

(รองศาสตราจารย์ ดร.ศิริชัย ชินะตังกูร)

คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน..... พ.ศ.....

อาจารย์ที่ปรึกษาวิทยานิพนธ์

อาจารย์ ดร.สุนีย์ พงษ์พินิจภิญโญ

คณะกรรมการตรวจสอบวิทยานิพนธ์

..... ประธานกรรมการ

(รองศาสตราจารย์ ดร.จันทนา ผ่องเพ็ญศรี)

...../...../.....

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ณัฐธนนท์ หงส์วริทธิ์ธร)

...../...../.....

..... กรรมการ

(อาจารย์ ดร.สุนีย์ พงษ์พินิจภิญโญ)

...../...../.....

48309335 : สาขาวิชาเทคโนโลยีสารสนเทศ
คำสำคัญ : การจำแนกหมวดหมู่เพลงไทยเดิม / การเลือกค่าคุณลักษณะ
โอภาส แก้วต่าย : การจำแนกกลุ่มเพลงไทยเดิม. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : อ.ดร. สุนีย์ พงษ์พินิจภิญโญ.82 หน้า.

การจัดหมวดหมู่ข้อมูลเสียงเพลงนั้น ต้องใช้ข้อมูลที่อธิบายเพลง หรือเรียกว่า Metadata ที่ได้มาจากการกรอกข้อมูลโดยมนุษย์ เช่น ผู้ที่ทำหน้าที่จัดการข้อมูลในระบบการค้นคืนข้อมูลเสียงเพลง (Music Information Retrieval System-MIRS) เป็นผู้กรอก ดังนั้นหากผู้กรอกข้อมูล Metadata ไม่สามารถอธิบายหรือจำแนกได้ว่าเพลงนั้น ๆ อยู่ในหมวดหมู่อะไร เนื่องจากขาดความรู้ความเข้าใจในบทเพลง จะส่งผลให้การจัดหมวดหมู่ของเสียงเพลงเกิดการผิดพลาด จึงได้มีการศึกษาและวิจัยวิธีการจัดหมวดหมู่เสียงเพลงแบบอัตโนมัติด้วยข้อมูลภายในเสียง (Content-based Music Genre Classification) กันอย่างกว้างขวางในเพลงสากล แต่ยังไม่มากในเพลงไทย โดยเฉพาะเพลงไทยเดิม โดยอาศัยค่าคุณลักษณะหรือค่าที่สกัดได้จากคลื่นเสียงเป็นข้อมูลในการจำแนก

งานวิจัยนี้นำเสนอการจำแนกหมวดหมู่เพลงไทยเดิมด้วยข้อมูลภายในเสียง และนำเสนอการศึกษาเปรียบเทียบวิธีการจำแนกหมวดหมู่เพลงไทยเดิม โดยใช้ค่าคุณลักษณะ Timbrel Feature แบบLinear Predictive Coefficients (LPC), Spectral Centroid, อัตราค่าตัดศูนย์ (Zero Crossing Rate), Mel-frequency cepstral coefficients (MFCC) ซึ่งเป็นค่าคุณลักษณะที่สกัดได้จากความถี่ของเสียงเพลงร่วมกับค่าพลังงานสเปกตรัม (Power Spectrum) ซึ่งเป็นค่าพลังงานของคลื่นเสียง มาจำแนกหมวดหมู่ด้วยวิธีการ k-Nearest Neighbor (k-NN), Support Vector Machine (SVM) และ Artificial Neural Network (ANN) เพื่อทดสอบหาวิธีการที่จะให้ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมได้ดีที่สุด ผลจากการทดลองพบว่าสามารถจำแนกหมวดหมู่เพลงไทยเดิมได้ดีที่สุดด้วยวิธีการ Support Vector Machine (SVM) ด้วยค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับค่าพลังงานสเปกตรัม (Power Spectrum) ที่ขนาดความกว้างของเฟรมเท่ากับ 512 samples และความช่วงยาวของเพลงที่ 0 ถึง 20 วินาที โดยได้ค่าความถูกต้องในการจำแนกสูงที่สุดคือ 86.16% ซึ่งค่าความถูกต้องที่ได้เป็นไปตามเกณฑ์ที่ผู้วิจัยได้กำหนดไว้

ภาควิชาคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2552
ลายมือชื่อนักศึกษา.....
ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ .....

48309335 : MAJOR : INFORMATION TECHNOLOGY
KEY WORD : MUSIC GENRE CLASSIFICATION / FEATURE SELECTION
O-PAS KAEWTAI : THAI CLASSICAL MUSIC CLASSIFICATION. THESIS ADVISOR : SUNEE PONGPINIGPINYO.,Ph.D. 82 pp.

To classify music data requires using descriptive music data called Metadata which is gathered by filling information by human. For example, the one who manage data in Music Information Retrieval System ((MIRS) cannot explain or identify which group that song should be in because of lacking of music understanding. This will take an effect on mistaken classification of music. Therefore, there are a lot of research works on Content-based Music Genre Classification in worldwide in Western music but not many Thai music, especially Thai Classical Music. Most of these research works used data features which were extracted from sound wave to classify the music.

This research presents the methods of content-based genre classification of traditional Thai classical music. The classification methods use Timbrel features such as Linear model, Predictive Coefficients (LPC), Spectral Centroid, Zero Crossing Rate, Mel-Frequency Cepstral Coefficients (MFCC) which were extracted from the sound of music frequency. The power spectrums which are power of sound waves are also used as the features in this research. In order to get the best method and features that result the most classification accuracy, the comparison on various machine learning classification algorithms, including k-Nearest Neighbor, Support Vector Machines and Artificial Neural Networks using with timbral feature vectors and power spectrum features are experimented. The result shows that Support Vector Machine (SVMs) using Mel-Frequency Cepstral Coefficients and power spectrums features with 512 samples of windows size and audio length 0 to 20 second achieve the most accuracy rate of 86.16% for traditional Thai classical music genre classification.

Department of Computing Graduate School, Silpakorn University Academic Year 2009
Student's signature .....
Thesis Advisor's signature .....

กิตติกรรมประกาศ

งานวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ผู้วิจัยต้องขอขอบพระคุณ อาจารย์ ดร. สุนีย์ พงษ์พินิจภิญโญ อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นอย่างสูง ที่ได้คอยให้ความรู้ คำแนะนำ และอุปกรณ์ในการค้นคว้าทดลองจนงานวิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ขอขอบพระคุณ รองศาสตราจารย์ ดร.จันทนา ผ่องเพ็ญศรี ประธานคณะกรรมการตรวจสอบวิทยานิพนธ์ และผู้ช่วยศาสตราจารย์ ดร.ณัฐธนนท์ หงส์วริทธิ์ธร กรรมการตรวจสอบวิทยานิพนธ์ ที่ช่วยทำให้งานวิจัยฉบับนี้สำเร็จสมบูรณ์ไปได้ด้วยดี ขอขอบคุณอาจารย์ประจำภาควิชาคอมพิวเตอร์ที่ให้ความรู้ ขอขอบคุณเจ้าหน้าที่ภาควิชาทุกท่านที่คอยให้ความช่วยเหลือประสานงานต่างๆ เป็นอย่างดี ขอขอบคุณ คุณพิศาล สุขขี ที่ได้ให้ความช่วยเหลือแนะนำการเขียนโปรแกรมรวมไปถึงความช่วยเหลืออื่นๆ จนวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงมาได้ ขอขอบคุณเพื่อนๆที่ได้คอยให้กำลังใจกันตลอดมา

สุดท้ายนี้ต้องขอขอบพระคุณบิดา มารดาและคนในครอบครัวที่เป็นแรงผลักดัน เป็นกำลังใจจนงานวิจัยฉบับนี้สำเร็จลุล่วงด้วยดี

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญตาราง ..... ฎ

สารบัญภาพ ..... ฐ

บทที่

1 บทนำ 1

ความเป็นมาและความสำคัญของปัญหา 1

จุดประสงค์งานวิจัย 2

ขอบเขตการศึกษา 2

ขั้นตอนการศึกษา 3

เครื่องมือและอุปกรณ์ 4

ประโยชน์ที่คาดว่าจะได้รับ 4

2 ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง 5

ทฤษฎีที่เกี่ยวข้อง 5

ดนตรีไทย 5

ประเภทของเครื่องดนตรีไทย 5

วงดนตรีไทย 6

เสียง 6

ระดับเสียง 7

ความเข้มของเสียง 7

สีสันเสียง 7

คุณภาพเสียง 7

ความยาวเสียง 7

การสุ่มหน้าคลื่น 7

ความสูงของคลื่น 8

ค่าคุณลักษณะของเสียง 8

บทที่ หน้า
การประมาณเชิงเส้นหรือกระบวนการแอลพีซี 8
เซปสตรัมที่คำนวณบนแกนความถี่แบบเมล 11
อัตราค่าตัดศูนย์ 12
สเปกทรอลเซนทรอยด์ 12
พลังงานสเปกตรัม 13
การจำแนกประเภทข้อมูล 13
ซัพพอร์ตเวคเตอร์แมชชีนส์ 14
แนวความคิดของซัพพอร์ตเวคเตอร์แมชชีนส์ 14
สมการพื้นฐานของซัพพอร์ตเวคเตอร์แมชชีน 15
ค่าความกว้างเส้นขอบ 16
การแก้ปัญหา Dual Problem 17
ค่าผิดพลาดในการวางคุณลักษณะ..... 18
เคอร์เนล 19
K-Nearest Neighbor 21
เครือข่ายใยประสาทเทียม..... 22
ข้อมูลอินพุต 23
ชั้นซ่อน 23
ค่าถ่วงน้ำหนักและไบแอส..... 24
รูปแบบการเรียนรู้ของระบบเครือข่ายใยประสาท..... 24
Multilayer perceptron 25
Back-propagation algorithm 26
Transfer functions 28
วรรณกรรมที่เกี่ยวข้อง 28
3 วิธีดำเนินการวิจัย 32
การรวบรวมเพลงไทยเดิม 33
การเตรียมข้อมูลเพื่อนำไปใช้ในการจำแนก 33
การแปลงสัญญาณเสียง 33
การตัดแบ่งส่วนเพลง 33
การสกัดค่าคุณลักษณะ 33

บทที่ หน้า
การเลือกค่าคุณลักษณะ 33
การแบ่งข้อมูลออกเป็นเฟรม 36
การสุ่มข้อมูลตัวอย่าง 37
การจำแนกหมวดหมู่เพลงไทยเดิม 37
การจำแนกด้วยวิธีการ k-Nearest Neighbor 37
การจำแนกด้วยวิธีการ Support Vector Machine 37
การจำแนกด้วยวิธีการ Artificial Neural Network..... 38
การเปรียบเทียบค่าความถูกต้องของการจำแนก 39
สรุปผลการทดลองและจัดทำรายงานวิทยานิพนธ์ 39
4 ผลการดำเนินการวิจัย 40
การรวบรวมข้อมูลเพลงไทยเดิมที่ใช้ในงานวิจัย 40
การเตรียมข้อมูล 41
การแปลงรูปแบบ 41
การตัดแบ่งส่วนเพลงไทยเดิม 41
การเลือกค่าคุณลักษณะ 42
ขนาดความกว้างของเฟรม..... 50
ความยาวของเพลงที่ใช้สกัดค่าคุณลักษณะ..... 50
การจำแนกหมวดหมู่เพลงไทยเดิม 51
ผลการจำแนกด้วยวิธีการ k-Nearest Neighbor 52
ผลการจำแนกด้วยวิธีการ Support Vector Machine 54
ผลการจำแนกด้วยวิธีการ Artificial Neural Network 55
เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิม 56
5 สรุปผลการดำเนินงานวิจัย 58
สรุปผลการดำเนินงานวิจัย 58
การเลือกค่าคุณลักษณะ..... 58
ค่าคุณลักษณะที่ใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม..... 58
ขนาดความกว้างของเฟรม..... 58
ความยาวของเพลงที่ใช้สกัดค่าคุณลักษณะ..... 59
วิธีการในการจำแนกหมวดหมู่..... 59

บทที่ หน้า
การจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor.. 59
วิธีการ Support Vector Machine …..... 59
วิธีการ Artificial Neural Network ……... 59
การเปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่
เพลงไทยเดิม…….. 60
วิเคราะห์และอภิปรายผลการทดลอง…..... 60
ข้อเสนอแนะ …..... 66

บรรณานุกรม ……...…… 67

ภาคผนวก ……...…… 70
ภาคผนวก ก ตารางเครจซี่ มอร์แกน สำหรับการกำหนดขนาดตัวอย่าง …..... 71
ภาคผนวก ข ตารางแสดงรายชื่อเพลงที่ใช้ในการทดลอง ..……....……..… 75
ภาคผนวก ค ค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ... 80

ประวัติผู้วิจัย ……. 82

สารบัญตาราง

ตารางที่		หน้า
1	การจำแนกตระกูลและเครื่องดนตรีตามประเภทของวงดนตรี	6
2	สรุปค่าคุณลักษณะแบบต่าง ๆ ที่ใช้ในงานวิจัยที่เกี่ยวข้อง.....	31
3	แสดงตัวอย่างรายชื่อเพลงที่ใช้ในการทดลอง	40
4	แสดงตัวอย่างรายชื่อเพลงที่ตัดแบ่งเป็นส่วนวินาทีต่าง ๆ.....	42
5	ค่าความถูกต้องในการจำแนกของค่าคุณลักษณะแต่ละแบบ.....	43
6	แสดงค่าความถูกต้องในการจำแนกโดยการจับคู่ค่าคุณลักษณะ.....	43
7	แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ3 แบบ	44
8	แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ 4 แบบ	45
9	แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะทั้ง 5 แบบ	46
10	เปรียบเทียบค่าความถูกต้องในการจำแนกที่มากที่สุดโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ของทั้งสามวิธีการ.....	47
11	เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่โดยใช้ค่าคุณลักษณะแบบต่าง ๆ	48
12	ค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum ที่ความกว้างของเฟรมขนาดต่าง ๆ	50
13	ค่าความถูกต้องในการจำแนกหมวดหมู่ที่ช่วงความยาวต่าง ๆ ของเพลง	51
14	แสดงค่าความถูกต้องในการจำแนกหมวดหมู่ด้วยวิธีการ k-NN โดยพิจารณาที่จำนวนของกลุ่มข้อมูลใกล้เคียง (k)	52
15	แสดงผลการทดลองด้วยวิธีการ Support Vector Machine	54
16	ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ ANN ที่จำนวนโหนดในชั้นต่าง ๆ	56
17	ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ ANN ที่ทำการปรับจำนวนโหนดในชั้นซ่อน	56
18	เปรียบเทียบผลการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-NN, SVM และ ANN	57

ตารางที่ หน้า

19 จำนวนของเฟรมเมื่อแบ่งตามขนาดความกว้างของเฟรมขนาดต่าง ๆ ที่ช่วงความยาวของเพลงต่าง ๆ ต่อหนึ่งเพลง ……. 62
20 แสดงเวลาที่ใช้ในการเรียนรู้และทดสอบข้อมูลของวิธีการจำแนกหมวดหมู่แบบ SVM และ k-NN ที่ได้ค่าความถูกต้องในการจำแนกที่มากที่สุด ……... 63
21 แสดงตารางเครจซี่ มอร์แกน สำหรับการกำหนดขนาดตัวอย่าง ……..... 72
22 แสดงรายชื่อเพลงที่ใช้ในการทดลอง …… 76
23 ค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ของทั้งสามวิธีการ..... 81

สารบัญภาพ

ภาพที่		หน้า
1	เปรียบเทียบ sampling rate ที่ 8 kHz และ sampling rate ที่ 44.1 kHz ….....	8
2	วิธีการคำนวนเซปสตรัมที่คำนวณบนแกนความถี่แบบเมล ….....	11
3	ตัวอย่างอัตราค่าตัดศูนย์เท่ากับ 432 ….....	12
4	การขยายตัวของเส้นขอบ ….....	14
5	เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง ….....	15
6	การเกิดเวคเตอร์อนุโลม (Slack Vector) ….....	18
7	รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้ ….....	20
8	ตัวอย่างการจำแนกหมวดหมู่ของวิธีการ k-Nearest Neighbor ….....	22
9	รูปแบบของเซลล์ประสาท (Neuron model) ….....	23
10	โครงสร้างของ Feed-forward multilayer perceptron แบบ 3 ชั้น ….....	26
11	รูปแบบ Back-propagation neural network ….....	27
12	ขั้นตอนการดำเนินการวิจัย ….....	32
13	เปรียบเทียบ Spectrum ของเพลงค้างคาวกินกล้วยที่บรรเลงด้วยวงมโหรีและวงปี่พาทย์……..	34
14	เปรียบเทียบ Spectrum ของเพลง Reggae และเพลง Rock……...	34
15	ตัวอย่างค่าคุณลักษณะแบบต่าง ๆ ที่ใช้ในงานวิจัย……...	35
16	แสดงการแบ่งขนาด Frame ของเพลง ….....	36
17	การจำแนกหมวดหมู่เพื่อหาซัพพอร์ตเวคเตอร์ของแต่ละหมวดหมู่ด้วย LibSVM ...	38
18	ตัวอย่างwaveform ของเพลงค้างคาวกินกล้วยเมื่อแปลงเป็นไฟล์ wav ที่มีค่า Sampling Rate ที่ 44 kHz ….....	41
19	เปรียบเทียบค่าความถูกต้องในการจำแนกโดยเฉลี่ยเมื่อใช้ค่าคุณลักษณะจำนวนต่าง ๆ ……...	46
20	เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่โดยใช้ค่าคุณลักษณะแบบต่าง ๆ ……...	49
21	แสดงค่าความถูกต้องในการจำแนกต่อจำนวนกลุ่มข้อมูลใกล้เคียง (k) ต่าง ๆ ……	54
22	เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมของวิธีการ k-NN, SVM และ ANN ……...	57

ภาพที่ หน้า

23 ค่าความถูกต้องในการจำแนกโดยเฉลี่ยที่ความกว้างของเฟรมขนาดต่าง ๆ 61
24 ส่วนที่เป็นเสียงเงียบในเพลงที่ใช้ในการจำแนกข้อมูล 64

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การค้นคืนข้อมูลเสียงเพลง (Music Information Retrieval (MIR)) เพื่อการรับฟังหรือเพื่อการค้นหาในปัจจุบันนั้น ได้รับความสนใจอย่างกว้างขวางในการพัฒนาวิธีการให้มีการจัดการข้อมูลที่ถูกต้องมากยิ่งขึ้น เนื่องจากในการจัดเก็บข้อมูลเสียงเพลงนั้นต้องใช้ข้อมูลที่อธิบายเพลงนั้น ๆ หรือเรียกว่า Metadata ที่ได้มาจากการกรอกข้อมูลโดยมนุษย์ เช่น ผู้ที่ทำหน้าที่จัดการข้อมูลในระบบ MIRเป็นผู้กรอก ดังนั้นหากผู้กรอกข้อมูล Metadata ไม่สามารถอธิบายหรือจำแนกได้ว่าเพลงนั้น ๆ อยู่ในหมวดหมู่อะไร เนื่องจากขาดความรู้ความเข้าใจในบทเพลง จะส่งผลให้การจัดหมวดหมู่ของเสียงเพลงเกิดการผิดพลาด อีกทั้งยังทำให้การค้นคืนข้อมูลเสียงเพลงไม่ถูกต้องตามที่ผู้ใช้ต้องการ จึงได้มีการศึกษาและวิจัยวิธีการจัดหมวดหมู่เสียงเพลงแบบอัตโนมัติด้วยข้อมูลภายในเสียง (Content-based Music Genre Classification) กันอย่างกว้างขวาง เพื่อนำมาช่วยจัดหมวดหมู่บทเพลงให้มีความถูกต้องและสะดวกมากยิ่งขึ้น เช่น เตาลี่และคณะ (2003) ที่ทำการศึกษาเปรียบเทียบการจำแนกหมวดหมู่โดยอัตโนมัติ โดยทำการศึกษาการจำแนกหมวดหมู่เพลงสากล 10หมวดหมู่ด้วยวิธีการต่าง ๆ หรือ เชน (2009) ที่จำแนกหมวดหมู่เพลงสากลจำนวน 9 หมวดหมู่ด้วยวิธีการSupport Vector Machine (SVMs) หรือ อานนท์ นามสนิท (2549) ที่จำแนกหมวดหมู่เพลงไทย 3ประเภทคือ เพลงไทยลูกทุ่ง เพลงไทยลูกกรุง และเพลงไทยสากล ด้วยวิธีการ Support Vector Machine (SVMs)

ในการจำแนกหมวดหมู่แบบอัตโนมัติด้วยข้อมูลของเสียง สามารถแบ่งขั้นตอนใหญ่ ๆ ได้สองขั้นตอนคือ การสกัดค่าคุณลักษณะของเสียง (Feature Extraction) และขั้นตอนการจำแนกหมวดหมู่ (Classification) ซึ่งในการสกัดค่าคุณลักษณะของเสียงนั้น ค่าคุณลักษณะที่บ่งบอกเอกลักษณ์ของเสียง (Timbral Feature) (Tzanetakis and Cook 2002) เป็นค่าคุณลักษณะที่นิยมสกัดมาใช้ในการจำแนกหมวดหมู่เป็นอย่างมาก (Li et al. 2003 ; Doraisamy et al. 2008 ; Chen and Chen 2009) เช่น ค่า Spectral Centroid, อัตราค่าตัดศูนย์ (Zero Crossing Rate), Mel-frequency cepstral coefficients (MFCC) หรือ Linear Predictive Coefficients (LPC) และค่าคุณลักษณะที่เป็นค่าพลังงานก็เป็นอีกค่าหนึ่งที่นิยมนำมาใช้ในการจำแนกหมวดหมู่ของเพลง เพราะเสียงเพลง

นั้นอยู่รูปของคลื่นเสียงซึ่งสามารถสกัดเอาค่าพลังงานออกมาได้ เช่น ค่าพลังงานสเปกตรัม (Power Spectrum)

จากการศึกษาพบว่างานวิจัยส่วนใหญ่มุ่งเน้นไปที่การจำแนกหมวดหมู่เพลงสากล หรือเพลงที่บรรเลงด้วยเครื่องดนตรีสากล ยังไม่มีงานวิจัยใดที่ทำการศึกษาการจำแนกหมวดหมู่ของเพลงที่บรรเลงด้วยเครื่องดนตรีไทย หรือเพลงไทยเดิมซึ่งนอกจากจะมีการใช้เครื่องดนตรีที่หลากหลายแล้ว ยังสามารถแบ่งออกได้เป็นหลายประเภท โดยหากแบ่งตามวงที่บรรเลงแล้ว เพลงไทยเดิมสามารถแบ่งออกได้เป็น สามวงใหญ่ๆ คือ วงมโหรี วงปี่พาทย์ และวงเครื่องสาย โดยในแต่ละวงนั้นใช้บรรเลงในโอกาสที่แตกต่างกัน ทำให้ได้อรรถรสในการรับฟังที่แตกต่างกันออกไป ทั้งยังเป็นเพลงที่บ่งบองถึงวัฒนธรรมและเอกลักษณ์ของชาติไทยอีกด้วย

ดังนั้นงานวิจัยนี้จึงนำเสนอการศึกษาเปรียบเทียบวิธีการจำแนกหมวดหมู่เพลงไทยเดิมโดยใช้ค่าคุณลักษณะแบบ Timbrel Featureซึ่งเป็นค่าคุณลักษณะที่สกัดได้จากความถี่ของเสียงเพลง ทำให้สามารถระบุเอกลักษณ์ของเพลงได้ และค่าพลังงานสเปกตรัม (Power Spectrum) ซึ่งเป็นค่าพลังงานของเพลง มาจำแนกหมวดหมู่ด้วยวิธีการ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) เพื่อให้ได้วิธีการและค่าคุณลักษณะที่เหมาะสมที่สุดในการจำแนกหมวดหมู่เพลงไทยเดิม

จุดประสงค์งานวิจัย

1. เพื่อศึกษาเปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN)

2. เพื่อศึกษาเปรียบเทียบขนาดของ Windowsize และค่าคุณลักษณะแบบTimbrel Feature ที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิม

ขอบเขตการศึกษา

1. ศึกษาทฤษฎีและวิธีการจำแนกหมวดหมู่บทเพลงโดยอาศัยข้อมูลภายใน (Content-based Music Genre Classification)

2. ศึกษาทฤษฎีและวิธีการทำงานของ k-Nearest Neighbor (k-NN),Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) เพื่อนำไปใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม

3. ศึกษาทฤษฎีและวิธีการในการสกัดค่าคุณลักษณะ (Feature Extraction) จากข้อมูล

เสียงประเภท Timbral Featureได้แก่ Linear Predictive Coefficients (LPC), Spectral Centroid, อัตราค่าตัดศูนย์ (Zero Crossing Rate), Mel-frequency cepstral coefficients (MFCC) และค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum)

4. ศึกษาวิธีการแบ่งไฟล์เสียงออกเป็นเฟรม(Windowsize)โดยใช้จำนวน sample เป็นตัวแบ่ง

5. สกัดค่าคุณลักษณะ (Feature Extraction) จากข้อมูลเสียงประเภท Timbral Feature ได้แก่ Linear Predictive Coefficients (LPC), Spectral Centroid, อัตราค่าตัดศูนย์ (Zero Crossing Rate), Mel-frequency cepstral coefficients (MFCC) และค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum)

6. จำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธี k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN)

7. เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่จากวิธี k-Nearest Neighbor (k-NN), Support Vector Machine(SVMs) และ Artificial Neural Network (ANN)

ขั้นตอนการศึกษา

ในงานวิจัยนี้สามารถแบ่งขั้นตอนในการศึกษาและดำเนินการดังนี้

1. รวบรวมความรู้และข้อมูลจากเอกสารและแหล่งข้อมูลที่เกี่ยวข้องในเรื่องการจำแนกหมวดหมู่บทเพลงโดยอาศัยข้อมูลภายใน (Content-based Music Genre Classification)

2. รวบรวมความรู้และข้อมูลจากเอกสารและแหล่งข้อมูลที่เกี่ยวข้องกับวิธีการทำงานของ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) ในการจำแนกหมวดหมู่

3. รวบรวมความรู้และข้อมูลจากเอกสารและแหล่งข้อมูลที่เกี่ยวข้องกับในการสกัดค่าคุณลักษณะ (Feature Extraction) จากข้อมูลเสียงประเภท Timbral Featureได้แก่ Linear Predictive Coefficients (LPC), Spectral Centroid, อัตราค่าตัดศูนย์ (Zero Crossing Rate), Mel-frequency cepstral coefficients (MFCC) และค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum)

4. รวบรวมบทเพลงที่จะนำมาใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม

5. ทำการทดลองและทดสอบความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิม

6. วิเคราะห์ผลการทดลองเปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN),Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) โดยเลือกพิจารณาที่ค่าความถูกต้องในการจำแนกที่ มากกว่า 85% ขึ้นไป

7. สรุปผลการทดลอง

8. รวบรวมข้อเสนอแนะ

เครื่องมือและอุปกรณ์

1. ซอฟต์แวร์

- ระบบปฏิบัติการ Linux
- Java J2SE version 6.0
- jAudio
- SoX
- WLSVM
- WEKA

2. ฮาร์ดแวร์

- Processor0 Intel(R) Core(TM)2 Duo CPU E7500 @ 2.93GHz
- Processor1 Intel(R) Core(TM)2 Duo CPU E7500 @ 2.93GHz
- RAM : 2 GB
- Hard disk 250 GB

ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย มีดังต่อไปนี้

1. สามารถจำแนกหมวดหมู่เพลงไทยเดิมได้
2. ได้วิธีการจำแนกหมวดหมู่เพลงไทยเดิมที่เหมาะสม
3. ได้ขนาดเฟรม (Windowsize) ที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิม
4. ได้ค่าคุณลักษณะ Timbrel Feature ที่นำมาใช้งานร่วมกับค่าพลังงานสเปกตรัม (Power Spectrum) ที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิม
5. สามารถนำไปประยุกต์ใช้ในการพัฒนาแอพพลิเคชั่น เพื่อเก็บรวบรวมและจัดหมวดหมู่บทเพลงไทยโดยอัตโนมัติต่อไปได้

บทที่ 2
ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

บทนี้จะกล่าวถึงทฤษฎีและวรรณกรรมที่เกี่ยวข้องที่นำมาใช้ในการดำเนินการวิจัย โดยจะกล่าวตามลำดับดังนี้

ทฤษฎีที่เกี่ยวข้อง

ดนตรีไทย

เฉลิมศักดิ์ พิกุลศรี (2530) ได้กล่าวถึงประเภทของเครื่องดนตรีไทยและวงดนตรีไทยไว้ดังต่อไปนี้

1. ประเภทของเครื่องดนตรีไทย

เครื่องดนตรีของแต่ละชาตินั้นมีการแบ่งประเภทที่แตกต่างกันออกไป แต่มีวัตถุประสงค์เดียวกันคือ เพื่อสะดวกในการจะใช้ที่แตกต่างกัน ในคัมภีร์ที่ชื่อว่า นาฏยศาสตร์ (Natyasastra) ซึ่งเป็นคัมภีร์สันสกฤตของพวกพราหมณ์ ได้จำแนกเครื่องดนตรีออกเป็น 4 ประเภท คือ

- ตะตะ คือ เครื่องดนตรีประเภทมีสายสำหรับดีด หรือสีเป็นเสียง
- สุษิระ คือ เครื่องดนตรีประเภทที่เป่าเป็นเสียง
- อะวะนัทธะ คือ เครื่องดนตรีที่หุ้มหนังสีเป็นเสียง
- ฆะนะ คือ เครื่งอดนตรีที่กระทบเป็นเสียง

ดนตรีทางตะวันตกนั้นการจำแนกเครื่องดนตรี จำแนกตามลักษณะของเครื่องดนตรีแบ่งออกเป็น 4 ประเภทดังนี้

- เครื่องสาย (String Instruments)
- เครื่องลมไม้ (Wood wind Instruments)
- เครื่องทองเหลือง (Brass Instruments)
- เครื่องเพอคัชฌัน (Percussion Instruments)

สำหรับเครื่องดนตรีไทยนั้น การจำแนกเราอาศัยจำแนกกิริยาอาการปฏิบัติของผู้เล่น ซึ่งสามารถแบ่งออกได้เป็น 4 ประเภทดังนี้

- เครื่องดนนตรีประเภทดีด (Plucked String Instruments)
- เครื่องดนตรีประเภทสี (Bowed String Instruments)
- เครื่องดนตรีประเภทตี (Percussion Instruments)
- เครื่องดนตรีประเภทเป่า (Wind Instruments)

2. วงดนตรีไทย วงดนตรีไทยในปัจจุบันแบ่งออกเป็น 3 ประเภท คือ

วงเครื่องสาย หมายถึง วงดนตรีที่ประกอบด้วยดนตรีประเภทเครื่องสายเป็นหลัก และมีเครื่องดนตรีประเภทเครื่องเป่าเป็นส่วนประกอบ โดยมีฉิ่ง ฉาบ กรับ เป็นเครื่องประกอบจังหวะ เครื่องดนตรีประเภททำทำนองประกอบด้วย จะเข้ ซออู้ ซอด้วง โทน-รำมะนา และฉิ่ง ฉาบ กรับ โหม่ง ความมาน้อยเครื่องดนตรีขึ้นอยู่กับขนาดของวง

วงปี่พาทย์ หมายถึง วงดนตรีที่เกิดจากการประสมกันระหว่างเครื่องดนตรีประเภทเป่า และเครื่องดนตรีประเภทตี

วงมโหรี เป็นวงดนตรีที่เกิดจากการประสมกันระหว่างวงปี่พาทย์และวงเครื่องสาย โดยตัดเครื่องดนตรีที่มีเสียงดังออก เช่น ปี่ออก วงมโหรีเป็นวงดนตรีที่มีความสมบูรณ์ที่สุด กล่าวคือ มีเครื่องดนตรีที่ประสมอยู่ในวงครบทุกตระกูล คือ ดีด สี ตีและเป่า

ตารางที่ 1 แสดงการจำแนกตระกูลและเครื่องดนตรีตามประเภทของวงดนตรี

	ดีด	สี	ตี	เป่า
เครื่องสาย	จะเข้	ซอด้วง ซออู้	-	ขลุ่ย
ปี่พาทย์	-	ไม้นวม ซออู้	ระนาดเอก ระนาดทุ้ม ระนาดเอกเหล็ก ฆ้องวงใหญ่และเล็ก	ไม้แข็ง ปี่ใน นอก กลาง ขลุ่ยหลีบ, เพียงออ, อู้
มโหรี	จะเข้	ซอสามสาย ซออู้ ซอด้วง	ระนาดเอก-ทุ้ม ระนาดเอกเหล็ก-ทุ้ม ฆ้องวงใหญ่-เล็ก	ขลุ่ยอู้ ขลุ่ยเพียงออ ขลุ่ยหลีบ

เสียง (Sound) (อานนท์ นามสนิท 2549 : 4-5)

เสียงเกิดจากการสั่นของโมเลกุลของวัตถุ แล้วถ่ายเทพลังงานให้กับโมเลกุลของ

อากาศที่วิ่งมาชน ทำให้โมเลกุลของอากาศมีความเร็วสูงขึ้นแล้วเกิดการชนกันของโมเลกุล เมื่อเกิดการชนกันแล้วถ่ายเทพลังงานออกมาเป็นทอดๆ ก็เหมือนคลื่นของความดันที่แผ่ออกไป เมื่อมากระทบแผ่นไดอะแฟรมที่หูของมนุษย์ ก็จะเปลี่ยนให้เป็นสัญญาณไฟฟ้า ส่งไปยังสมอง ทำให้เราสามารถรับรู้และแยกแยะเสียงต่างๆ ได้

1. ระดับเสียง (Pitch) ระดับเสียงเกิดจากการสั่นสะเทือนของวัตถุ วัตถุที่สั่นสะเทือนเร็วจะทำให้เกิดระดับเสียงที่สูงกว่า ในขณะเดียวกันวัตถุที่สั่นสะเทือนช้าก็จะทำให้เกิดระดับเสียงที่ต่ำกว่า โดยระดับเสียงจะมีหน่วยเป็นรอบต่อนาที วัตถุที่สั่นสะเทือนมากกว่าจะมีความถี่มากกว่าทำให้เกิดระดับเสียงที่สูงกว่า ถ้าความถี่มากขึ้นเท่าตัว ระดับเสียงจะสูงขึ้นหนึ่งช่วงคู่แปด (Octave) เช่น เสียงที่มีความถี่ 220 รอบต่อนาที จะมีความถี่เป็นช่วงคู่แปดกับเสียงที่มีความถี่ 110 รอบต่อวินาที

2. ความเข้มของเสียง (Intensity) หมายถึง เสียงเบา เสียงดัง เกิดจากแรงสั่นสะเทือนของวัตถุที่เป็นแหล่งกำเนิดเสียงสะเทือนมากจึงเกิดเสียงดังมาก ในทางตรงกันข้ามหากวัตถุต้นกำเนิดเสียงสั่นสะเทือนน้อย ก็จะเกิดเสียงเบา ดังนั้นความเข้มของเสียงจึงขึ้นอยู่กับความแรงที่ส่งจากแหล่งกำเนิดเสียงไปยังหู โดยสามารถวัดได้จากความสูงของคลื่นเสียง (Amplitude) ปัจจัยที่มีผลต่อความเข้มเสียงก็คือระยะทาง เพราะเมื่อเสียงเดินทางผ่านบรรยากาศความเข้าเสียงจะน้อยลงตามลำดับ

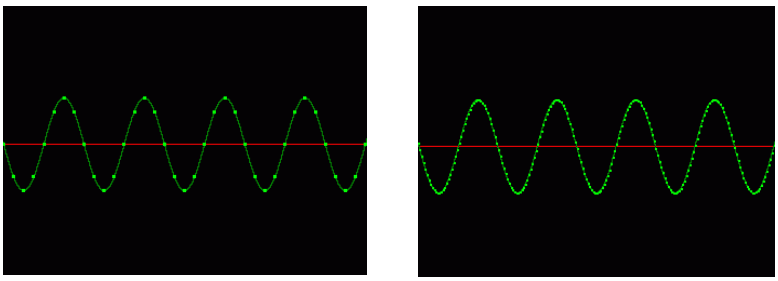
3. สีสันเสียง (Timbre หรือ Tone color) คือลักษณะที่แตกต่างกันของเสียงแต่ละชนิด โดยคลื่นเสียงที่ออกมาจากแหล่งกำเนิดเสียงที่ต่างชนิดกันก็จะมีลักษณะรูปร่างที่แตกต่างกัน เช่น เมื่อเราเล่นโน้ตเดียวกันโดยใช้เครื่องดนตรีที่ต่างกันสองชนิด รูปคลื่นที่ได้จากเครื่องดนตรีทั้งสองก็จะแตกต่างกัน ทำให้เราสามารถแยกแยะได้ว่างเสียงที่ได้ยินเป็นเสียงของเครื่องดนตรีชนิดใดบ้าง

4. คุณภาพเสียง (Tone quality) หมายถึงเสียงที่ได้ออกมาจากแหล่งกำเนิดเสียง มีคุณภาพที่ดีมากหรือดีน้อย เช่น เมื่อเราเล่นโน้ตเดียวกันจากเปียโนคนละตัว คุณภาพเสียงที่ได้ก็จะต่างกัน

5. ความยาวเสียง (Duration) เป็นพื้นฐานของดนตรีซึ่งต้องเกี่ยวข้องกับเวลา เสียงแต่ละเสียงที่เกิดขึ้นต้องมีระยะเวลาซึ่งทำให้เกิดเสียงสั้นเสียงยาว ไม่ว่าเสียงจะมีระดับเสียงที่แน่นอนหรือไม่ก็จะต้องมีระยะเวลาเข้ามาเกี่ยวข้องเสมอ ความยาวของเสียงจึงเป็นที่มาของจังหวะในดนตรี ซึ่งรวมไปถึงความยาวของความเงียบด้วย เนื่องจากดนตรีเป็นผลของกระบวนการเกิดเสียงสลับกับความเงียบซึ่งต่างก็ต้องมีระยะเวลาทั้งนั้น

6. การสุ่มหน้าคลื่น (Sampling) เนื่องจากเสียงนั้นเป็นสัญญาณต่อเนื่อง เมื่อ

ต้องการแปลงให้อยู่ในรูปของสัญญาณดิจิตอลก็จะต้องทำการสุ่มหน้าคลื่นด้วยการดูว่าขณะนี้มี Amplitude เท่าไหร่ โดยอัตราในการสุ่มหน้าคลื่นนี้เรียกว่า Sampling Rate เช่น ถ้าสุ่ม 8000 ครั้งต่อวินาที เรียกว่า 8 kHz sampling หมายความว่าในเวลา 1 วินาที จะได้ข้อมูล amplitude จำนวน 8000 จุด



ภาพที่ 1 เปรียบเทียบ sampling rate ที่ 8 kHz (ซ้าย) และ sampling rate ที่ 44.1 kHz (ขวา)
ที่มา : Sampling [Online], accessed 20 May 2009. Available from http://www.bcae1.com/sampling.htm

7. ความสูงของคลื่น (amplitude) Sampling rate เป็นตัวกำหนดความมากน้อยของจุดที่ใช้แทนหน้าคลื่น โดยจุดที่บันทึก นั่นก็คือ ความสูงของคลื่น ซึ่งจะใช้ตัวเลขแค่ 0 กัย 1 แทน amplitude ของคลื่น เรียกการบันทึกแบบนี้ว่ามี quantization แบบ 2 bit ถ้าเพิ่มจำนวน bit มาเป็น 4 bit จะสามารถกำหนดระดับ amplitude ได้เพิ่มขึ้นเป็น 16 ระดับ

ค่าคุณลักษณะของเสียง (Feature)

ค่าที่ได้จากเสียงนั้นเป็นตัวแทนสำคัญ ที่นำมาแทนข้อมูลเสียงเพื่อนำไปใช้สร้างแบบจำลองและนำไปใช้ในการจำแนกหมวดหมู่ ซึ่งได้มาจากการสกัดค่าคุณสมบัติเสียง (Feature Extraction) งานวิจัยนี้ใช้ค่าคุณลักษณะแบบ Timbrel Feature (Tzanetakis and Cook 2002) ซึ่งเป็นค่าคุณลักษณะที่บ่งบอกเอกลักษณ์ของเสียง โดยจะกล่าวถึงการวิเคราะห์คุณลักษณะต่าง ๆ ที่นำมาใช้ในงานวิจัยดังต่อไปนี้

1. การประมาณเชิงเส้นหรือกระบวนการแอลพีซี (Linear Predictive Coding : LPC) บุญธิดา สุวัชระกุลธร (2547 : 16) กล่าวว่าได้มีการประยุกต์ใช้กระบวนการนี้ในการวิเคราะห์หาพารามิเตอร์ต่างๆ ของสัญญาณเสียง เช่น คาบเวลาพิทช์, ความถี่ฟอร์แมนท์, สเปกตรัม, ฟังก์ชันของพื้นที่โวคัลแทร็ค (Vocal Tract Area Function) และการส่งสัญญาณเสียงด้วยอัตราความเร็วบิตต่ำ (Low Bit Rate Voice) แนวความคิดพื้นฐานของกระบวนการแอลพีซี ก็คือการสุ่ม

ตัวอย่างของสัญญาณเสียง (Speech Sample) สามารถจะสร้างได้จากผลรวมเชิงเส้น (Linear Combination) ของการสุ่มตัวอย่างสัญญาณเสียงหลายๆ ตัวอย่างที่ผ่านมาในอดีต โดยการทำให้ค่าผลรวมกำลังสองของความแตกต่าง หรือค่าผิดพลาดระหว่างค่าตัวอย่างจริงและค่าที่ทำนายได้ บนช่วงที่กำจัดขนาดหนึ่งมีค่าน้อยที่สุด เราสามารถที่จะหาค่าสัมประสิทธิ์ของตัวทำนายออกมาได้ ซึ่งก็คือค่าสัมประสิทธิ์ให้น้ำหนัก (Weighting Coefficient) ที่ใช้ในผลรวมเชิงเส้นนั่นเอง

รูปแบบและวิธีการที่ใช้ในกระบวนการแอลพีซีนั้นมีอยู่หลายแบบด้วยกัน ความแตกต่างของแต่ละรูปแบบอยู่ที่การมองปัญหา หรืออีกนัยหนึ่งก็คือ วิธีการคำนวณเพื่อที่จะให้ได้มาซึ่งสัมประสิทธิ์ของตัวทำนายนั่นเอง รูปแบบต่างๆที่พบเห็นได้มีดังต่อไปนี้

1. วิธีโควาเรียนซ์ (Covariance Method)
2. รูปแบบอัตสหสัมพันธ์ (Autocorrelation Formulation)
3. วิธีการแลตติช (Lattice Method)
4. รูปแบบตัวกรองความถี่กลับ (Inverse Filter Formulation)
5. รูปแบบการประมาณเชิงสเปกตรัม (Spectral Estimation Formulation)
6. รูปแบบความน่าจะเป็นจริงสูงสุด (Maximum Likelihood Formulation)
7. รูปแบบผลคูณภายใน (Inner Product Formulation)
8. รูปแบบของกระบวนการแอลพีซีซึ่งเป็นรูปแบบที่นิยมใช้กันมาก คือ วิธีโควาเรียนซ์และรูปแบบอัตสหสัมพันธ์โดย อานนท์ (2549: 8-10) ได้อธิบายถึงรูปแบบอัตสหสัมพันธ์ไว้ว่า

รูปแบบอัตสหสัมพันธ์ (Autocorrelation Method) ในการวิเคราะห์การทำงานแบบเชิงเส้นช่วงสั้นหาได้โดยการวิเคราะห์แบบหน้าต่างสัญญาณเสียงและอ้างว่าตัวอย่างสัญญาณและอ้างว่าตัวอย่างสัญญาณภายนอกหน้าต่างนี้มีค่าเท่ากับศูนย์ แล้วจึงทำให้ค่าผิดพลาดตามสมการที่ 2.1 มีค่าต่ำสุด ซึ่งสมการที่ 2.1 เมื่อนำมาแตกจะได้ดังสมการที่ 2.2

$$E = \sum_{n=-\infty}^{\infty} e_n^2 \text{ ---------- (2.1)}$$

$$E = \sum_{n=-\infty}^{\infty} s_n^2 - 2\sum_{k=1}^{p} a_k \sum_{n=-\infty}^{\infty} s_n s_{n-1} + \left[\sum_{k=1}^{p} a_k s_{n-k}\right]^2 \text{ ---------- (2.2)}$$

การหาค่า a_k ที่ทำให้ E มีค่าต่ำที่สุดสามารถกระทำได้โดยการหาในสมการที่ 2.2 ของแต่ละค่า $i = 1,2,3,\ldots, p$ เมื่อ $k = 1,2,3,\ldots, p$ ได้เป็นสมการเชิงส้น p

สมการ และแต่ละสมการมี k ค่า ซึ่งก็คือ a_k ที่ไม่ทราบค่า p ตัวดังนี้

$$\frac{\partial E}{\partial a_i} = 0 = -2\sum_{n=-\infty}^{\infty} s_n s_{n-1} + 2\sum_{n=-\infty}^{\infty}\left[\sum_{k=1}^{p} a_k s_{n-k}\right] s_{n-i} \text{ ---------- (2.3)}$$

$$\sum_{n=-\infty}^{\infty} s_{n-1} s_n = \sum_{k=1}^{p} a_k \sum_{n=-\infty}^{\infty} s_{n-i} \cdot s_{n-k} \text{ ---------- (2.4)}$$

เนื่องจากพจน์แรกเป็นอัตสหสัมพันธ์ R_i ของ s_n ที่มีความยาวจำกัด สามารถเขียนสมการที่ 2.4 ใหม่ได้เป็นสมการที่ 2.5

$$\sum_{k=1}^{p} r_{|i=k|} a_k = -r_i \quad \text{เมื่อ } 1 \le i \le p \text{ ---------- (2.5)}$$

เมื่อแก้สมการที่ 2.3 ซึ่งประกอบไปด้วยสมการเชิงเส้น p สมการได้เป็นค่าสัมประสิทธิ์ a_k ที่ต้องการออกมา โดยเขียนสมการที่ 2.3 ให้อยู่ในรูปเมตริกซ์ $R_a = -r$

โดยที่

$$R = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & r_0 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & r_0 & \cdots & r_{p-3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_0 \end{bmatrix} \text{ ---------- (2.6)}$$

$$a = [a_1, a_2, \ldots, a_p]^T \text{ ---------- (2.7)}$$

$$r = [r_1, r_2, \ldots, r_p]^T \text{ ---------- (2.8)}$$

จากสมการที่ 2.6เมตริกซ์ R เรียกว่าเมตริกซ์อัตสหสัมพันธ์ (Autocorrelation Method) ซึ่งมีโครงสร้างแบบโทพสิทธิ์ (Toeplitz) เป็นเมตริกซ์แบบสมมาตร คือมีค่าในแนวทแยงมุมเท่ากันทั้งหมดทุกแถวสำหรับวิธีที่ช่วยหาผลเฉลยของสมการยูล-วอล์กเกอร์ ใน

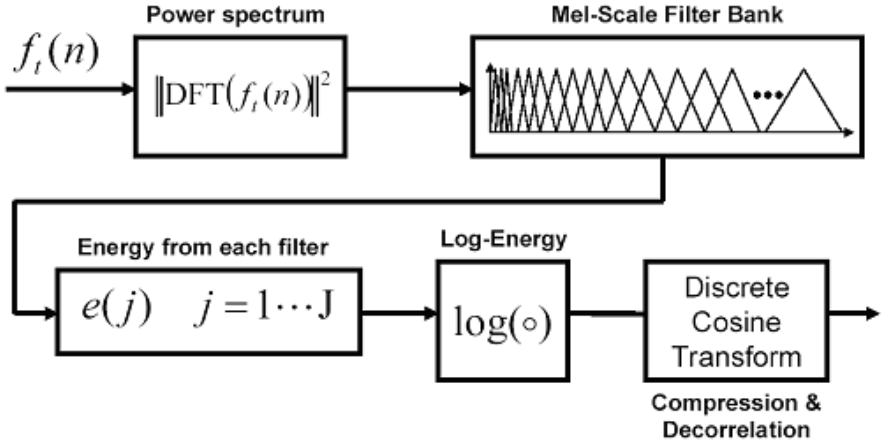
สมการที่ 2.1และหาสัมประสิทธิ์การทำนายแบบเชิงเส้นในสมการที่ 2.2 มีหลายวิธีดังเช่น วิธีเลวินสัน-เดอร์บิน (Levinson-Durbin Algorithm) และวิธีเชอร (Schur Algorithm) เป็นต้น

2. เซปสตรัมที่คำนวณบนแกนความถี่แบบเมล (Mel-Frequency Cepstral Coefficients (MFCC)) Mel Scale เป็นการแบ่งความถี่ออกตามความสามารถในการได้ยินของมนุษย์ โดยที่ช่วงไหนที่คนไม่ได้ยินก็จะถูกกรองทิ้ง ซึ่งจะทำให้เหลือแค่ส่วนที่มนุษย์สามารถได้ยิน ดังสมการ

$$Mel(f) = 2595\log_{10}\left(1+\frac{f}{700}\right) \qquad \text{---------- (2.9)}$$

โดย $Mel(f)$ คือ ความถี่แบบเมล หรือ Mel-Frequency

จากนั้นจึงนำ ความถี่แบบเมล มาใช้ร่วมกับ เซปสตรัม จึงได้ออกมาเป็นเซปสตรัมที่คำนวณบนแกนความถี่แบบเมล ดังภาพที่ 2



ภาพที่ 2 แสดงวิธีการคำนวนเซปสตรัมที่คำนวณบนแกนความถี่แบบเมล

ที่มา : ทรงฤทธิ์ มณีวงศ์วัฒนา และคณะ, Music Search Engine [ออนไลน์], เข้าถึงเมื่อ 23 พฤษภาคม 2552. เข้าถึงจาก http://cpe.kmutt.ac.th/previousproject/2005/26/theory.html

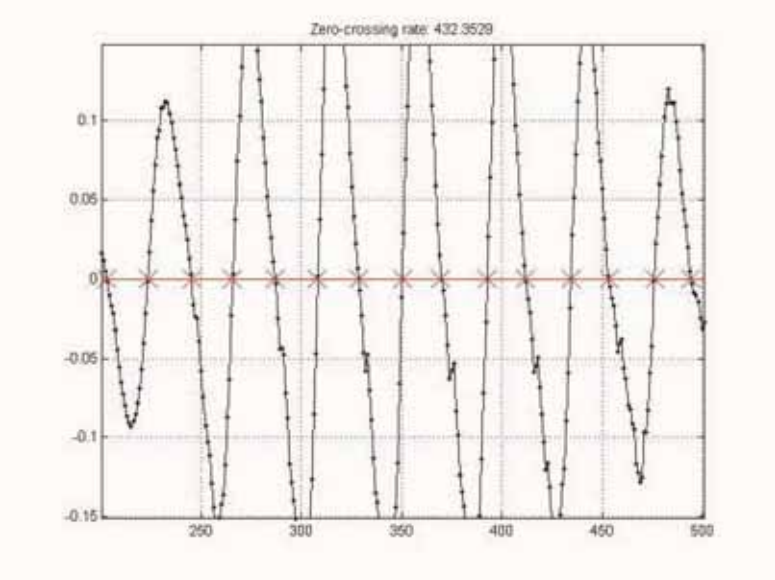
จากภาพที่ 2ในขั้นแรกเป็นการสัญญาณเข้ามาจากภายนอก แล้วมาทำการแปลงฟูเรียร์แบบไม่ต่อเนื่อง (Discrete Fourier Transform) จากนั้นนำแอมพลิจูดที่ได้ไปยกกำลังสอง เพื่อให้กลายเป็น Power Spectrum ซึ่งจะได้สัญญาณในแกนความถี่ ขั้นต่อมา นำไปคูณกับ Mel-Scale Filter Bank เพื่อทำการกรองสัญญาณที่มนุษย์ไม่ได้ยินออกไป หลังจากนั้นจะได้ค่าพลังงานจากตัวกรองแต่ละตัว (สามเหลี่ยมแต่ละลูก) ตั้งแต่ตัวที่ 1 ถึง J แล้วนำค่าพลังงานที่ได้แต่

ละตัว ไปทำการผ่านฟังก์ชันลอการิทึม ซึ่งจะได้เป็นค่าเดซิเบลของแต่ละตัวออกมา จากนั้น นำไปผ่านฟังก์ชันการแปลงโคไซน์แบบไม่ต่อเนื่อง (Discrete Cosine Transform) ดังสมการ2.10

$$MFCC[i][t] = \sqrt{\frac{2}{P}\sum_{j=1}^{p}\left\{\left(\log\left(e[i][t]\right)\right).\cos\left(\frac{\pi i}{P}\left(j-0.5\right)\right)\right\}} \text{ ------- (2.10)}$$

ซึ่งจะได้ค่าคุณลักษณะแบบ MFCC ออกมา

3. อัตราค่าตัดศูนย์ (Zero Crossing Rate) แนวความคิดในการหาค่าอัตราค่าตัดศูนย์ของสัญญาณนั้น คือการนับจำนวนครั้งของสัญญาณที่ผ่านแกนศูนย์นั่นเอง



ภาพที่ 3 ตัวอย่างอัตราค่าตัดศูนย์เท่ากับ 432

ที่มา : Geoffroy Peeters. A Large Set of Audio Features for Sound Description (similarity and classification) in the CUIDADO project [Online], accessed 23 May 2009. Available from http://recherche.ircam.fr/equipes/analysesynthese/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf

4. สเปกทรอลเซนทรอยด์ (Spectral Centriod) คือค่าคุณลักษณะที่แสดงถึงค่ากึ่งกลางของSpectrum โดยหาได้จากสมการที่ 2.11

$$Centriod = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \text{ ---------- (2.11)}$$

เมื่อ x(n) คือ ขนาดของ bin จำนวน n

และ f(n) คือ ค่ากึ่งกลางของความถี่ของ bin

5. พลังงานสเปกตรัม (Power Spectrum) ก่อนที่จะหาค่าพลังงานสเปกตรัมนั้น แต่ละเฟรมจะถูกหาค่าน้ำหนักก่อนด้วย Hamming Window

$$h(n) = \frac{\sqrt{\frac{8}{3}}}{2}\left[1 - \cos\left(2\pi \frac{n}{N}\right)\right] \text{ ---------- (2.12)}$$

เมื่อ n คือ ลำดับที่จุดเสียง (Sample Point)

N คือ จำนวนจุดเสียง (Sample Point) ในแต่ละเฟรม

ค่าพลังงานสเปกตรัมคำนวณได้จากสมการการแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform -FFT)

$$s(k) = 10\log_{10}\left[\frac{1}{N}\left\|\sum_{n=0}^{N-1} s(n)h(n)\exp\left(-j2\pi \frac{nk}{N}\right)\right\|^2\right] \text{-------(2.13)}$$

เมื่อ k คือ จำนวนค่าพลังงานสเปกตรัม

โดยวิธีการ Fast Fourier Transform นี้ มีพื้นฐานมาจากการคำนวณ Discrete Fourier Transform (DFT)

การจำแนกประเภทข้อมูล (Classification)

มีวิธีการในการทำงาน คือ การเรียนรู้ (Learning) เพื่อให้ได้มาซึ่งความรู้ (Knowledge) และใช้ในการสร้างตัวจำแนกข้อมูล (Classifier) โดยสามารถใช้ตัวจำแนกที่ได้มานี้แยกแยะ หรือจำแนกข้อมูลที่เข้ามาใหม่ ให้อยู่ในกลุ่มหรือประเภทที่มีลักษณะเดียวกันอยู่ด้วยกัน กลไกการเรียนรู้ (Machine Learning) สามารถแบ่งออกได้ 2 ลักษณะคือ

อันซุปเปอร์ไวสด์ เลินนิ่ง (Unsupervised Learning) คือ การจัดกลุ่มข้อมูลจากข้อมูลที่มีอยู่โดยไม่ต้องอาศัยข้อมูลเดิมในการสร้างตัวต้นแบบเพื่อการเรียนรู้ก่อน หรือบางครั้งอาจจะเรียกว่าการทำคลัสเตอร์ริ่ง (Clustering)

ซุปเปอร์ไวสด์ เลินนิ่ง (Supervised Learning) คือ การนำข้อมูลจากฐานข้อมูล

ในอดีตที่มีอยู่มาสร้างตัวแบบการเรียนรู้ เพื่อจำแนกประเภทข้อมูลโดยใช้ตัวจำแนกข้อมูล เมื่อมีข้อมูลใหม่เข้ามาจะสามารถจำแนกข้อมูลใหม่เหล่านั้นได้อย่างอัตโนมัติ

การวัดความแม่นยำในการจำแนกประเภท

ปวันรัตน์ มีชัย (2551) กล่าวว่า ความแม่นยำ (Accuracy) แสดงถึงความน่าเชื่อถือของกฎซึ่ง มักถูกแสดงในรูปของอัตราส่วนความถูกต้องของการจำแนกประเภท มีวิธีการเช่น

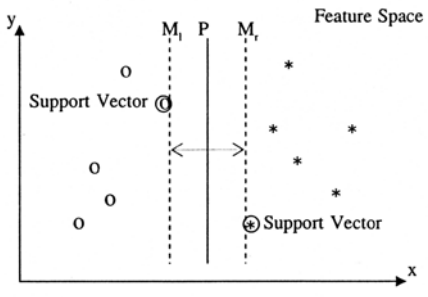
1. แบ่งข้อมูลเป็นสองส่วน ได้แก่ชุดข้อมูลฝึกฝน กับข้อมูลชุดทดสอบ วิธีทั่วไปที่ใช้ในการวัดคือ การให้ข้อมูลโดยสมมุติว่าข้อมูลรู้จักคลาสครบทุกคลาส

2. Cross Validation คือการตรวจวัดความแม่นยำโดยการแบ่งข้อมูลเป็นส่วนๆ สลับให้แต่ละส่วนเป็นข้อมูลทดสอบ แล้วให้ทุกๆส่วนที่เหลือเป็นชุดฝึกฝน แล้ววนจนกระทั่งทุกชุดได้ผ่านการเป็นชุดทดสอบ วิธีนี้จะเหมาะกับชุดข้อมูลที่มีจำนวนเรคคอร์ดน้อยกว่าวิธีข้างต้น และเป็นวิธีการที่งานวิจัยนี้เลือกใช้ในการวัดความแม่นยำในการจำแนกหมวดหมู่เพลงไทยเดิม

สำหรับเทคนิคที่นำมาใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมที่ใช้ในงานวิจัยนี้ ได้ทำการศึกษาและรวบรวมไว้ดังต่อไปนี้

ซัพพอร์ตเวคเตอร์แมชชีนส์ (Support Vector Machines – SVMs)

1 แนวความคิดของซัพพอร์ตเวคเตอร์แมชชีนส์ อานนท์ (2549 : 11-17) กล่าวถึงแนวความคิดของ SVMว่า เป็นสมการที่ใช้ในการจำแนกค่าคุณลักษณะของกลุ่มสองกลุ่มที่วางตัวอยู่ในพื้นที่คุณลักษณะ (Feature Space) ออกจากกัน โดยจะสร้างเส้นแบ่งที่เป็นเส้นตรงขึ้นมา เพื่อให้ทราบว่าเส้นตรงที่แบ่งกลุ่มออกจากกันนั้นเส้นตรงใดเป็นเส้นตรงที่ดีที่สุด โดยเส้นตรงนั้นก็จะถูกเพิ่มเส้นขอบ (Margin) ออกไปทั้งสองข้าง โดยเส้นขอบที่เพิ่มเข้ามานั้นจะขนาดเส้นเดิมเสมอ และจะขยายออกไปจนกว่าจะสัมผัสกับค่าของกลุ่มตัวอย่างที่ใกล้ที่สุด ดังภาพที่ 4



ภาพที่ 4 การขยายตัวของเส้นขอบ

ที่มา : อานนท์ นามสนิท, “การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์” (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549), 12.

จากภาพที่ 4 เส้น M_1 และ M_2 คือเส้นขอบที่ขยายออกไปทางด้านซ้ายและขวาตามลำดับ และ P คือ เส้นแบ่งกลุ่มข้อมูลทั้งสอง เมื่อเส้น M_1 และ M_2 ขยายออกไปสัมผัสค่าข้อมูลที่ใกล้ที่สุด ซึ่งข้อมูลที่อยู่ใกล้เส้นขอบทั้งสองฝั่งนั้นเรียกว่า เวคเตอร์สนับสนุน (Support Vector) ค่าระยะความห่างหรือขนาดของเส้นขอบก็จะถูกวัด โดยเส้น P จะเปลี่ยนความชันไปเรื่อย ๆ เพื่อที่จะหาความกว้างสูงสุดของเส้นขอบ

กระบวนการโดยรวมของซัพพอร์ตเวคเตอร์แมชชีนนั้น เป็นการหาค่าความชันของเส้น P ที่มีขนาดของเส้นขอบสูงสุดนั่นเอง

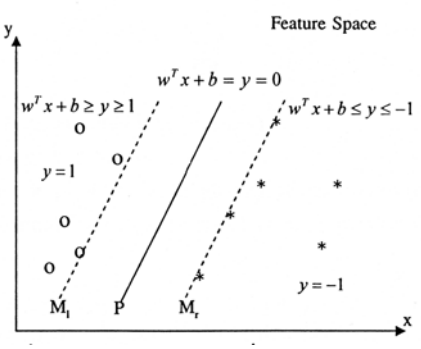
2. สมการพื้นฐานของซัพพอร์ตเวคเตอร์แมชชีน

หากนำแนวคิดของซัพพอร์ตเวคเตอร์แมชชีนที่กล่าวไปแล้ว มาเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยข้อมูลที่นำมาวางบนพื้นที่คุณลักษณะนั้นเป็นกลุ่มข้อมูลที่มีอยู่ในรูปเวคเตอร์

$$X = ((x_1,y_1), \ldots,(x_i,y_i)) \text{ ----------} \quad (2.14)$$

โดย X คือ ค่าคุณลักษณะ

ค่าคุณลักษณะที่วางตัวอยู่ในพื้นที่คุณลักษณะ จะถูกแบ่งด้วยเส้นตรงดังภาพที่ 4 และเมื่อนำเส้นตรงมาแทนค่าด้วยสมการเส้นตรง $y = mx + b$ โดยมีการกำหนดกลุ่มของข้อมูลทั้งสองฝั่งเป็นเพียงสองค่าซึ่งแทนด้วยค่า y เพื่อให้ข้อมูลที่อยู่ในกลุ่มเดียวกันที่มาจากหลายค่ากลายเป็นค่าเดียวดังแสดงในภาพที่ 5



ภาพที่ 5 เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง

ที่มา : อานนท์ นามสนิท, “การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์” (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549), 13.

จากภาพที่ 5 เส้นครง M_1 แทนด้วยสมการ $w^t x \square b \geq y \geq 1$ ซึ่งข้อมูล y ที่มากกว่า 1 ก็จะถูกกำหนดค่าใหม่โดยให้ y เท่ากับ 1 และพจน์ของ w ก็คือความชันเช่นเดียวกันกับเส้นตรง M1 ที่ค่าของ y จะถูกกำหนดใหม่เมื่อ y น้อยกว่า -1 เป็น -1 ดังนั้นสมการที่เกิดใหม่จากสมการเส้นของ (2.15) และ (2.16) สามารถกำหนดได้ดังสมการที่ (2.17)

เมื่อ $w^t x + b \geq y$ กำหนด $y = 1$ ---------- (2.15)

เมื่อ $w^t x + b \leq y$ กำหนด $y = -1$ ----------(2.16)

$y\left(w^t + b\right) - 1 \geq 0$ ----------(2.17)

โดย y คือ ค่ากลุ่มข้อมูล (1,-1)

w คือ ค่าความชัน

x คือ ค่าคุณลักษณะ

b คือ ค่าคงที่ (ค่าตัดแกน y)

3. ค่าความกว้างเส้นขอบ (Margin)

การคำนวณความกว้างของเส้นของ (Margin) ต้องทำการคำนวณพจน์ w ให้อยู่ในรูปปกติ (Normalization) โดยคำนวณจากสมการที่ (2.15) และ (2.16) เมื่อแทนค่า y ลงไปแล้ว

$$w^T x^+ + b = 1$$

$$w^T x^- + b = -1$$

$$w^T \left(x^- - x^+\right) = 2$$

$$M = \left(\frac{w}{|w|}\right)^T \left(x^+ - x^-\right)$$

$$M = \left(\frac{2}{|w|}\right) \text{----------(2.18)}$$

โดย M คือ ความกว้างของเส้นขอบ (Margin)

4. การแก้ปัญหา Dual Problem

หลังจากที่ได้สมการ (2.17) และ (2.18) ของการหาเส้นแบ่งและค่าความกว้างตามลำดับแล้ว จึงทำการแก้สมการนั้นต่อด้วย Lagrangian Dual Problem เพื่อหาค่าในพจน์ของ w ในสมการที่ (2.17)

$$L(w,b,\alpha)=\frac{1}{2}(w.w)-\sum_{i=1}^{N}\alpha_i\left[y_i\left((w_i.w_i)+b\right)-1\right] \text{---------- (2.19)}$$

เมื่อ $\alpha_1 \geq 0; i=1,...N$

สมการที่ (2.19) ถูกนำมาหาอนุพันธ์ (Differential)

$$\frac{\partial L(w,b,\alpha)}{\partial b}=w-\sum_{i=1}^{N}y_i\alpha_i=0 \text{---------- (2.20)}$$

$$\frac{\partial L(w,b,\alpha)}{\partial b}=\sum_{i=1}^{N}y_i\alpha_i=0 \text{---------- (2.21)}$$

หลังจากที่ได้สมการที่ผ่านการแก้ปัญหาด้วย Dual Problem แล้ว ค่า w จะลดรูปและหาได้จาก

$$w=\sum_{i=1}^{N}y_i\alpha_i x_i \text{---------- (2.22)}$$

เมื่อ α คือ ค่าสัมประสิทธิ์คงที่

ค่า b นั้นไม่ปรากฏในการแก้ปัญหาด้วย Dual Problem จึงหาได้ยาก

$$b=-\frac{\max_{y_i=-1}\left(w^T x_i\right)+\min_{y_i=1}\left(w^T x_i\right)}{2} \text{----------(2.23)}$$

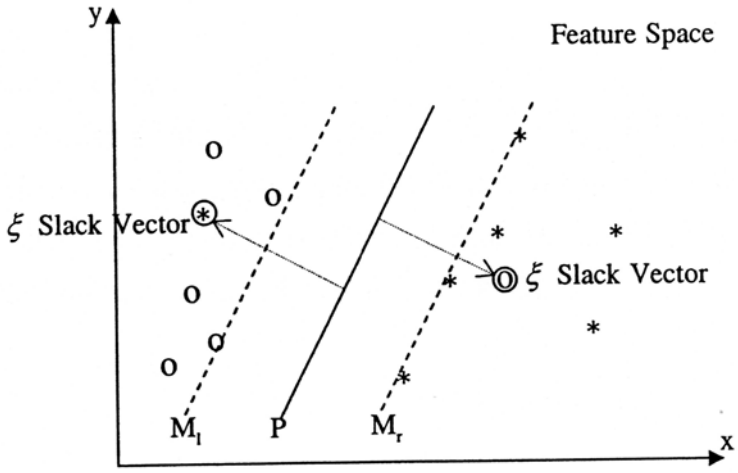
เมื่อนำค่า w ที่แก้ปัญหาไว้ไปใส่ในสมการที่ 2.17 ซึ่งเป็นสมการในการหาเส้นแบ่งจะได้

$$y_i\left(\sum_{i=1}^{N}\alpha_i y_i\left(x^T x_i+1\right)\right)\geq 0 \text{---------- (2.24)}$$

จากสมการที่ 2.24 คือการหาสัมประสิทธิ์ α เพื่อนำมาใช้บอกการวางตัวของเส้นแบ่งจากข้อมูลต่อไป

5. ค่าผิดพลาดในการวางคุณลักษณะ

ในการวางค่าคุณลักษณะในพื้นที่คุณลักษณะ บางครั้งอาจมีคุณลักษณะที่วางตัวอยู่ผิดกลุ่มออกไปซึ่งจะทำให้สมการของเส้นแบ่ง ไม่เป็นความจริง การแก้ปัญหานี้คือการยอมให้มีการเกิดข้อผิดพลาดนี้ได้ในระยะหนึ่งที่ยอมรับได้ ค่านั้นเรียกว่าเวคเตอร์อนุโลม (Slack Vector)



ภาพที่ 6 การเกิดเวคเตอร์อนุโลม (Slack Vector)
ที่มา : อานนท์ นามสนิท, “การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์” (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549), 15.

เมื่อค่าเวคเตอร์อนุโลมเข้ามาในสมการที่ 2.19 ซึ่งเป็นขั้นการแก้ปัญหา Dual Problem ก็จะเกิดตัวแปร Cขึ้นซึ่งเป็นค่าคงที่ค่าหนึ่ง ดังสมการที่ 2.25

$$L(w,b,\xi,\alpha)=\frac{1}{2}(w.w)+\frac{C}{2}\sum_{i=1}^{N}\xi_i^2-\sum_{i=1}^{N}\alpha_i\left[y_i\left((w_i.w_i)+b\right)-1+\xi_i\right] \text{---------- (2.25)}$$

สมการที่ 2.25 ถูกนำมาหาอนุพันธ์ (Differential) เช่นเดียวกับขั้นตอนที่ปรากฏในหัวข้อ 2.1.5.4

$$\frac{\partial L(w,b,\xi,\alpha)}{\partial w} = w - \sum_{i=1}^{N} y_i \alpha_i x_i = 0 \text{ ----------(2.26)}$$

$$\frac{\partial L(w,b,\xi,\alpha)}{\partial \xi} = C\xi - \alpha = 0 \text{ ---------- (2.27)}$$

$$\frac{\partial L(w,b,\xi,\alpha)}{\partial b} = \sum_{i=1}^{N} y_i \alpha_i = 0 \text{ ----------(2.28)}$$

โดยที่ ξ คือ เวคเตอร์อนุโลม

หลังจากได้สมการที่ผ่านการแก้ปัญหาด้วย Dual Problem แบบมีค่าเวคเตอร์อนุโลมแล้ว ค่า w ยังคงเป็นเช่นเดิมแต่จะได้ขอบเขตของ α ขึ้นมา

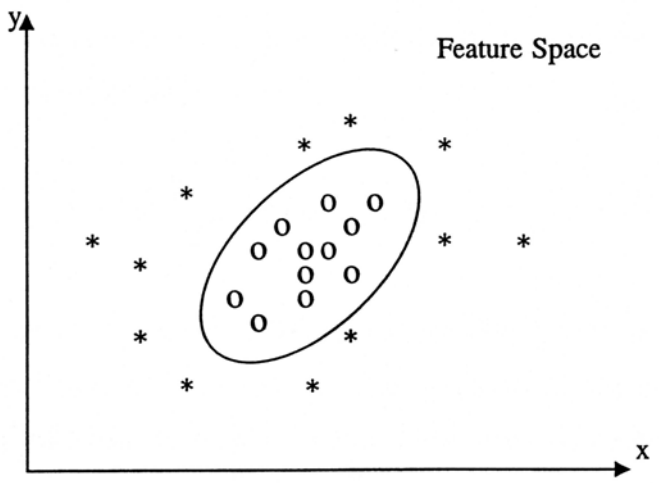
$$w = \sum_{i=1}^{N} y_i \alpha_i x_i \text{ ---------- (2.29)}$$

เมื่อ $C = \alpha = 0$

ค่า C จึงถือเป็นค่าพารามิเตอร์อีกค่าหนึ่งที่จะต้องกำหนดขึ้นมาเพื่อใช้หาค่าสัมประสิทธิ์ α

6. เคอร์เนล (Kernel)

ในโลกความเป็นจริงข้อมูล 2 กลุ่มไม่ได้วางตัวในพื้นที่คุณลักษณะเพียงสองกลุ่มและแบ่งได้โดยเส้นตรง แต่ข้อมูลอาจจะจับกลุ่มกันในตำแหน่งต่าง ๆ ดังนั้นจึงเป็นปัญหาทำให้ไม่สามารถที่จะใช้สมการซัพพอร์ตเวคเตอร์แมทชีนท์แบบเชิงเส้นได้ ดังนั้นจะต้องมีเครื่องมือมาช่วยให้ข้อมูลเหล่านั้นเรียงตัวเสียใหม่ในพื้นที่ที่เรียกว่าพื้นที่หลากมิติ (Higher Dimensional Space)



ภาพที่ 7 รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้
ที่มา : อานนท์ นามสนิท, “การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์” (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549), 16.

ในเคอร์เนลนั้นคือ การคูณกันของชุดเวคเตอร์ของ x ใด ๆ

$$K(x_i, x_j) = x_i^T x_j$$

เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิดด้วยกัน

โพลิโนเมียล (Polynomial)

$$K(x_i, x_j) = (< x_i, x_j > +1)^d \text{ ---------- (2.30)}$$

เมื่อ d คือค่าเลขยกกำลัง

เรเดียลเบลิสฟังก์ชั่น (Radial Basis Function-RBF)

$$K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right) \text{ ----------(2.31)}$$

เมื่อ σ คือ ค่าพารามิเตอร์

ซิกมอยด์ (Sigmoid)

$$K(x_i, x_j) = \tanh(k < x_i, x_j > + \mu) \text{---------- (2.32)}$$

เมื่อ k, μ คือ ค่าพารามิเตอร์

ดังนั้นจากสมการของเคอร์เนลนั้นสารถที่จะแทนลงไปในตำแหน่งของ $X_j^T \ X_j$ ในสมการที่ 2.24 จึงเขียนเป็นสมการใหม่ดังนี้

$$y_i\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x_j) + b\right) - 1 \geq 0 \text{---------- (2.33)}$$

สมการที่ 2.33 เป็นสมการที่ใช้ในขั้นตอนที่จะเรียนรู้ว่าจะวางตำแหน่งเส้นแบ่งไว้ที่ตำแหน่งใดโดยทำงานร่วมกับเคอร์เนล เพื่อแปลงให้ข้อมูลที่ยากต่อการแบ่งแบบเชิงเส้นสามารถแบ่งได้เมื่อถูกทำให้เป็นข้อมูลแบบหลายมิติ (Higher Dimension) ดังนั้นจึงมีอีกสมการหนึ่งที่ใช้ค่า w และ b เดิมมาจัดตำแหน่งของข้อมูลเพื่อให้ทราบว่าข้อมูลนั้นเป็นกลุ่มใด กำหนดได้ดังสมการที่ 2.34

$$f(x) = \text{sgn}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x_j) + b\right) \text{---------- (2.34)}$$

เมื่อ f(x) คือ ค่า y หาในรูปของ x

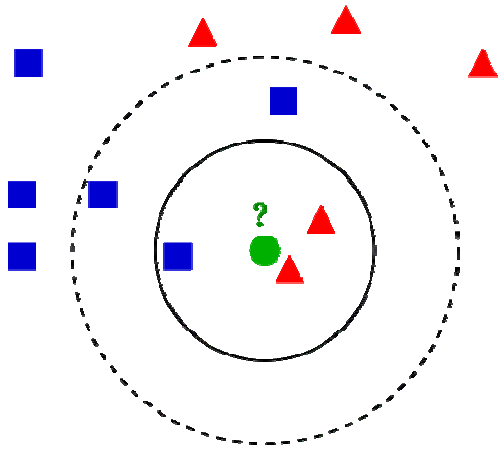
K-Nearest Neighbor (k-NN)

K-Nearest Neighbor หรือ k-NN เป็นวิธีการที่ใช้อย่างแพร่หลายและนิยมนำมาใช้ในการเปรียบเทียบวิธีการในการจำแนกหมวดหมู่เพลงอยู่เสมอ (Mingchun Liu and Chunru Wanp 2001: 247-248 ; Li, T et al. 2003 : 282-289 ; Chen, S., and Chen, S 2009 : 1095-1101) ในการจำแนกข้อมูลเสียงโดยหลักการทำงานก็คือ จะทำการแทนข้อมูลกลุ่มตัวอย่างที่ใช้ในการเรียนรู้ด้วยตัวเลข โดยที่ข้อมูลแต่ละข้อมูลจะถูกกำหนดตำแหน่งบนพื้นที่ และกลุ่มของข้อมูลในการเรียนรู้ (Training samples) แต่ละกลุ่ม จะถูกวางอยู่ในพื้นที่ของตนเอง (pattern space) ซึ่งเมื่อใส่ข้อมูลใหม่ที่ยังไม่ได้มีการระบุว่าอยู่ในกลุ่มข้อมูลใดเข้าไป k-NN จะทำการค้นหาตำแหน่งของ

กลุ่มข้อมูลที่ใกล้เคียงกันกับตำแหน่งของข้อมูลที่ใส่เข้าไปใหม่ โดย k คือจำนวนของกลุ่มข้อมูลที่อยู่ใกล้เคียงกับกลุ่มข้อมูลที่ใส่เข้าไปใหม่ ซึ่งในการหาตำแหน่งที่ใกล้เคียงกันระหว่างข้อมูลสองข้อมูลนั้น สามารถหาได้โดยสมการหาระยะห่าง (Euclidean distance) โดยที่เป็นการหาระยะห่างระหว่างจุดสองจุด คือ $X=(x_1,x_2,...,x_n)$และ $Y=(y_1,y_2,...,y_n)$ ด้วยสมการที่ 2.35

$$d(X,Y)=\sqrt{\sum_{i=1}^{N}(x_i-x_j)^2} \quad \text{---------- (2.35)}$$

ตัวอย่างการทำงานในการจำแนกข้อมูลของ k-Nearest Neighbor กำหนดให้วงกลมคือข้อมูลทดสอบที่จะต้องจำแนกว่าอยู่ในกลุ่มใด ระหว่างกลุ่มสี่เหลี่ยม และกลุ่มสามเหลี่ยม ถ้ากำหนดให้ค่า k เท่ากับ 3 ข้อมูลทดสอบจะจำแนกได้ว่าอยู่ในกลุ่มสามเหลี่ยม เนื่องจากเมื่อวัดระยะทางกับข้อมูล5จุดที่ใกล้ที่สุดพบว่ามีสามเหลี่ยมอยู่ 2 จุดและมีสี่เหลี่ยมอยู่ 1 จุด แต่ถ้ากำหนดค่า k เท่ากับ 5 ข้อมูลทดสอบนั้นจะจำแนกได้ว่าอยู่ในกลุ่มสี่เหลี่ยม เนื่องจากระยะห่างที่ใกล้ที่สุด 5 จุดนั้น มีสีเหลี่ยมอยู่ 3 จุด และมีสามเหลี่ยม 2 จุด ดังภาพที่ 8

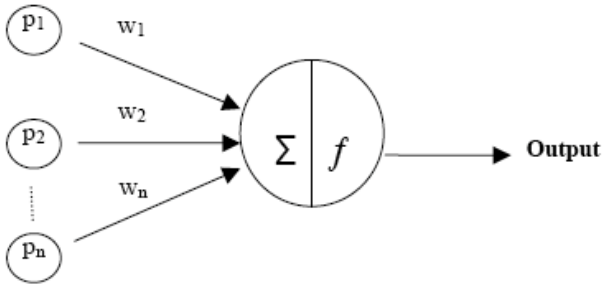


ภาพที่ 8 ตัวอย่างการจำแนกหมวดหมู่ของวิธีการ k-Nearest Neighbor

เครือข่ายใยประสาทเทียม (Artificial neural network – ANN)

เครือข่ายใยประสาทเทียม (Artificial neural network) เป็นการจำลองการทำงานบางส่วนของสมองมนุษย์ ซึ่งประกอบด้วยเซลล์ประสาท (neuron) เป็นจำนวนมาก โดยในแต่ละเซลล์จะประกอบด้วยนิวเคลียส (nucleus) ตัวเซลล์ (cell body) ใยประสาทนำเข้า (dendrite) แกน

ประสาทนำออก (axon) โดยใยประสาทนำเข้าจะมีหน้าที่รับสัญญาณไฟฟ้าเคมีซึ่งส่งมาจากเซลล์ประสาทที่อยู่ใกล้เคียง เมื่อสัญญาณไฟฟ้าเคมีที่ได้รับเข้ามาเกินค่าค่าหนึ่ง เซลล์ประสาทจะถูกกระตุ้นและส่งสัญญาณไปทางแกนประสาทนำออกต่อไป เมื่อปรับโครงสร้างและรูปแบบมาใช้กับเครือข่ายใยประสาทเทียมจะใช้โหนด (nodes) ทำหน้าที่คล้ายกับตัวเซลล์ประสาท (neurons) ดังภาพที่ 9 เซลล์ประสาทแต่ละโหนดจะรับค่าอินพุตได้หลายค่า $(p_1, p_2, ..., p_n)$ แต่ผลการกระตุ้นหรือค่าเอาต์พุตที่ได้มีเพียงหนึ่งค่า ซึ่งคำนวณได้จากการใช้ transfer function (f) กับผลรวมเชิงเส้นแบบถ่วงน้ำหนัก $(w_1, w_2, ..., w_n)$ ของอินพุต



ภาพที่ 9 รูปแบบของเซลล์ประสาท (Neuron model)
ที่มา : จารวี ฉันทสิทธิพร, “การจำแนกชนิดยาเม็ดจากภาพถ่าย โดยใช้เทคนิคเครือข่ายใยประสาท” (วิทยานิพนธ์ปริญามหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร, 2548), 16.

1. ข้อมูลอินพุต

โครงข่ายใยประสาทประดิษฐ์ ประกอบไปด้วยตัวแปรอิสระ หรือ ข้อมูลอินพุตและตัวแปรตาม หรือผลลัพธ์ โดยหลักการเลือกตัวแปรที่ใช้ในโครงข่ายที่เกี่ยวข้องจะมี 2 แบบ คือ วิธีแรก ข้อมูลจะต้องถูกแปลงรูปให้อยู่ในรูปที่เหมาะสม และวิธีที่สอง คือการเลือกข้อมูลโดยใช้พื้นฐานระหว่าง predictiveness และ covariance โดยปกติแล้ว ตัวแปรอิสระที่ถูกเลือกจะมีความสามารถในการทำนายผลหากตัวแปรที่เลือกมีความสัมพันธ์กันในทางตรงกันข้ามหากตัวแปรอิสระ 2 ตัวมีความสัมพันธ์ต่อกัน จะทำให้แบบจำลองมีความอ่อนไหว (sensitive) และเกิดปัญหาที่เรียกว่า over fitting และ limit generalization ด้วยเหตุผลนี้ การเลือกข้อมูลจะต้องเลือกเฉพาะตัวแปรอิสระที่มีความสามารถทำนายผลเพื่อให้ได้ ผลลัพธ์หรือตัวแปรตาม โดยตัวแปรอิสระที่เลือกมานั้นจะต้องไม่มีความสัมพันธ์กัน แต่อย่างไรก็ดี ก็ขึ้นอยู่กับรูปแบบของโครงข่ายที่ใช้และเพื่อที่จะลดจำนวนตัวอย่างที่ใช้ในการสอนและเวลาที่ใช้ในการเรียนรู้ ก็ควรจะต้องมีการคัดเลือกข้อมูลอินพุตให้เหมาะสม เพราะ

การคัดเลือกข้อมูลเป็นปัจจัยที่สำคัญในการสร้างแบบจำลอง

2. ชั้นซ่อน (Hidden Layer)

ชั้นซ่อนเป็นชั้นประมวลผลที่อยู่ระหว่างชั้นข้อมูลอินพุต และชั้นเอาท์พุต โดยปกติแล้ว ชั้นซ่อนอาจมีมากกว่า 1 ชั้น โดยโครงข่ายจะสามารถประมวลหาฟังก์ชั่นที่เหมาะสม จากปัญหาที่ซับซ้อนได้หากมีชั้นซ่อนที่มากพอ ข้อมูลที่ได้จากชั้นซ่อนจะได้เป็นตัวแปรใหม่ที่จะถูกส่งต่อให้กับชั้นเอาท์พุต หรือชั้นตัวแปรตาม ถ้าโครงข่ายแบบแพร่กลับ (Back-propagation) มีชั้นซ่อนที่น้อยเกินไปแล้วจะทำให้โครงข่ายไม่สามารถที่จะหาทางแก้ปัญหาได้ แต่ถึงอย่างไรก็ตามถ้าเกิดโครงข่ายมีชั้นซ่อนที่มากจนเกินไป จะทำให้โครงข่ายมีระยะเวลาในการเรียนรู้นาน หากมีชั้นซ่อนที่มากก็จะไม่ช่วยให้โครงข่ายมีประสิทธิภาพมากขึ้น หรืออีกนัยหนึ่ง การที่มีโหนดในแต่ละชั้นที่มากเกิน จะทำให้โครงข่ายไม่สามารถที่จะหาจุดสิ้นสุดได้ การที่มีโหนดในชั้นซ่อนมากเกินจะทำให้เกิดปัญหาที่เรียกว่า over fitting โดยโครงข่ายจะจำลองโครงสร้างใหม่เกินความเป็นจริงจาก noise ของข้อมูล แทนที่จะหาฟังก์ชั่นที่เหมาะสมในการวิเคราะห์ปัญหาให้ถูกต้องตามที่ควรจะเป็น

ดังนั้น การที่จะทำให้โครงข่ายเกิดประสิทธิภาพสูงสุด ต้องกำหนดให้มีโหนดในชั้นซ่อนอยู่ให้น้อยเท่าที่จะเป็นไปได้ จำนวนของโหนดในชั้นซ่อนควรอยู่ระหว่างค่าเฉลี่ยของผลรวมของโหนดในชั้นข้อมูลอินพุต และชั้นเอาท์พุต จำนวนโหนดในชั้นซ่อนที่มากที่สุดควรจะเท่ากับผลรวมของโหนดในชั้นข้อมูลอินพุต และชั้นเอาท์พุต และจำนวนโหนดที่น้อยที่สุดควรจะเท่ากับร้อยละ 75 ของโหนดชั้นข้อมูลอินพุต หรือ เท่ากับค่าเฉลี่ยของผลรวมของโหนดในชั้นข้อมูลอินพุตและชั้นเอาท์พุต

3. ค่าถ่วงน้ำหนักและไบแอส (Weights and biases)

ค่าถ่วงน้ำหนักถูกแทนด้วยตัวเลขเพื่อแสดงถึงความแรงในการเชื่อมต่อของโหนดแต่ละโหนดที่ถูกเชื่อมต่อเข้าด้วยกัน ซึ่งผลรวมของค่าถ่วงน้ำหนักที่ป้อนเข้าจะไปปรับปรุงการประมวลผลในแต่ละโหนด ค่าถ่วงน้ำหนักคือค่าความแรงสัมพัทธ์ (ในทางคณิตศาสตร์) ของการเชื่อมต่อซึ่งส่งผลต่อการส่งผ่านข้อมูลจากชั้นหนึ่งไปยังชั้นต่อไป โดยปกติค่าถ่วงน้ำหนักจะถูกกำหนด และเริ่มป้อนเข้าสู่โครงข่ายในขั้นตอนการเรียนรู้ ซึ่งต้องมีหลักการในการกำหนดค่าเพื่อที่จะให้โครงข่ายสามารถแก้โจทย์ปัญหา และลดเวลาการเรียนรู้ได้ สำหรับโครงข่ายใด ๆ ค่าถ่วงน้ำหนักจะมีค่าเท่ากับผลคูณของจำนวนโหนดของทุก ๆ การเชื่อมต่อ และค่าของไบแอสจะเท่ากับผลรวมของจำนวนโหนดของทุก ๆ การเชื่อมต่อ

4. รูปแบบการเรียนรู้ของระบบเครือข่ายใยประสาท (จารวี ฉันทสิทธิพร 2548 : 20 -23)

ลักษณะการเรียนรู้ของระบบเครือข่ายใยประสาทสามารถแบ่งได้เป็น 3

รูปแบบ ดังนี้

การเรียนรู้แบบมีผู้สอน (Supervised learning)

เครือข่ายใยประสาทจะได้รับการฝึกด้วยกลุ่มฝึกที่ประกอบด้วย ข้อมูลอินพุตและค่าเอาต์พุตที่ต้องการ การปรับค่าน้ำหนักและ bias นั้นจะกระทำเพื่อให้ค่าเอาต์พุตที่ได้มีความใกล้เคียงหรือตรงกับเอาต์พุตที่ต้องการมากที่สุด ตัวอย่างเครือข่ายใยประสาทที่ใช้วิธีการเรียนรู้แบบนี้ได้แก่ back-propagation และ multilayer perceptron เป็นต้น

Reinforcement learning

การเรียนรู้แบบนี้จะคล้ายคลึงกับการเรียนรู้แบบมีผู้สอนแต่ต่างกันตรงที่การเรียนรู้แบบนี้จะใช้อินพุตและใช้ระดับ (grade) หรือคะแนน (score) แทนค่าเอาต์พุต ซึ่งวิธีนี้จะมีความซับซ้อนมากกว่าและใช้เวลาในการฝึกมากกว่าทำให้เป็นที่นิยมน้อยกว่าการเรียนรู้แบบมีผู้สอน การเรียนรู้รูปแบบนี้มีความเหมาะสมสำหรับระบบขนาดใหญ่ที่ประกอบด้วยเครือข่ายใยประสาทหลายเครือข่ายทำงานร่วมกัน

การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)

การเรียนรู้เพื่อปรับค่าน้ำหนักและค่า bias จะตอบสนองต่อข้อมูลอินพุตเท่านั้นโดยไม่ได้ใช้ค่าเอาต์พุตร่วมด้วยในการฝึก วิธีการเรียนรู้แบบนี้จะใช้การจัดแบ่งกลุ่มโดยพิจารณาตามรูปแบบของอินพุตตัวอย่างเครือข่ายใยประสาทที่ใช้วิธีการเรียนรู้แบบนี้ได้แก่ counterpropagation และKohonen network เป็นต้น

ในงานวิจัยนี้เลือกใช้ระบบเครือข่ายใยประสาท Multilayer perceptron แบบ Back-propagation ซึ่งจะอธิบายดังต่อไปนี้

5. Multilayer perceptron

เครือข่ายใยประสาทประกอบด้วยหลายชั้นโดยในแต่ละชั้นประกอบด้วย โหนด (nodes) หรือเปรียบได้กับตัวเซลล์ประสาท (neurons) ค่าน้ำหนักของเส้นที่เชื่อมต่อระหว่างโหนดในแต่ละชั้น (เมทริก W), ค่า bias vector (b) และค่า output vector (a) โดย m เป็นตัวเลขบอกลำดับชั้นที่กำกับไว้ด้านบน เมื่อ p เป็น input vector การคำนวณค่าเอาต์พุตสำหรับเครือข่ายใยประสาทที่มีจำนวน M ชั้นดังสมการที่ 3.6 และภาพที่ 10

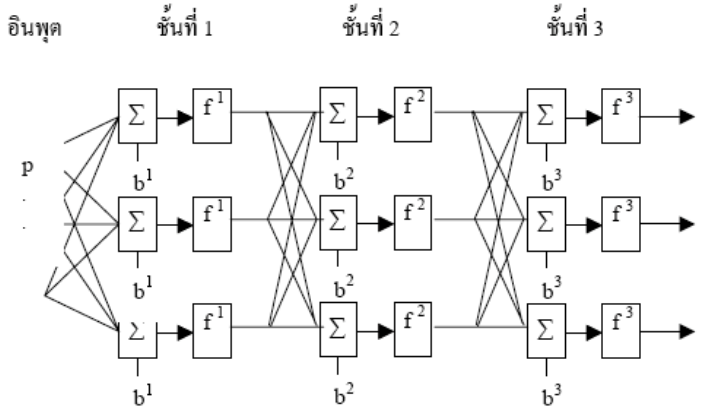
$$a^{m+1} = f^{m+1}\left(W^{m+1}a^{m} + b^{m+1}\right) \text{---------- (2.36)}$$

เมื่อ $m = 0, 2, \ldots, M-1$

$a^{0} = p$

$a = a^m$

และ f เป็น transfer function



$$a^1 = f^1\left(W^1 p + b^1\right) \quad a^2 = f^2\left(W^2 a^1 + b^2\right) a^3 = f^3\left(W^3 a^2 + b^3\right)$$

ภาพที่ 10 โครงสร้างของ Feed-forward multilayer perceptron แบบ 3 ชั้น

ที่มา : จารวี ฉันทสิทธิพร, “การจำแนกชนิดยาเม็ดจากภาพถ่าย โดยใช้เทคนิคเครือข่ายใยประสาท” (วิทยานิพนธ์ปริญามหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร, 2548), 21.

6. Back-propagation algorithm

Back-propagation เป็นอัลกอริทึมที่ใช้ในการเรียนรู้ของเครือข่ายใยประสาทวิธีหนึ่งที่นิยมใช้ใน multilayer perceptron เพื่อปรับค่าน้ำหนักในเส้นเชื่อมต่อระหว่างโหนดให้เหมาะสม โดยการปรับค่านี้จะขึ้นกับความแตกต่างของค่าเอาต์พุตที่คำนวณได้กับค่าเอาต์พุตที่ต้องการ ดังภาพที่ 11 โดยมีขั้นตอนดังต่อไปนี้

1. กำหนดค่าอัตราเร็วในการเรียนรู้ (rate parameter : r)

2. สำหรับแต่ละตัวอย่างอินพุตจะทำตามขั้นตอนต่อไปนี้จนกว่าได้ระดับ performance ที่ต้องการ

- คำนวณหาค่าเอาต์พุตโดยใช้ค่าน้ำหนักเริ่มต้นซึ่งอาจได้จากการสุ่ม
- คำนวณหาค่า β : แทนประโยชน์ที่จะได้รับสำหรับการเปลี่ยนค่าเอาต์พุตของแต่ละโหนดในชั้นเอาต์พุต

$$\beta_z = d_z - o_z \text{ ---------- (2.37)}$$

เมื่อ d_z = ค่าเอาต์พุตที่ต้องการ

o_z = ค่าเอาต์พุตที่คำนวณได้

ในชั้นซ่อน

$$\beta_j = w_{jk} \rightarrow o_k(1-o_k)\beta_k \text{ ---------- (2.38)}$$

เมื่อ $w_{jk} \rightarrow$ = น้ำหนักของเส้นเชื่อมระหว่างชั้นที่ j กับ k

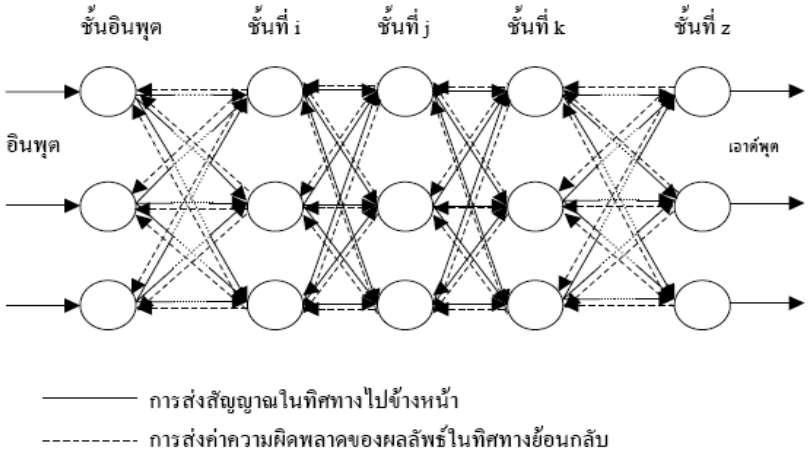
- คำนวณค่าน้ำหนักที่เปลี่ยนแปลงไปสำหรับในทุกน้ำหนัก โดย

สมการที่ 2.39

$$\Delta w_{ij} \rightarrow = ro_i o_j(1-o_j)\beta_j \text{ ---------- (2.39)}$$

- เพิ่มค่าน้ำหนักที่เปลี่ยนแปลง สำหรับตัวอย่างอินพุตทั้งหมด และ

เปลี่ยนค่าน้ำหนัก



ภาพที่ 11 รูปแบบ Back-propagation neural network

ที่มา : จารวี ฉันทสิทธิพร, “การจำแนกชนิดยาเม็ดจากภาพถ่าย โดยใช้เทคนิคเครือข่ายใยประสาท”

(วิทยานิพนธ์ปริญามหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัย

ศิลปากร, 2548), 23.

7. Transfer functions

Transfer functions ที่นิยมใช้ใน Back-propagation neural network มี 3 รูปแบบ ได้แก่ Log-Sigmoid, Tan-Sigmoid และ Linear เนื่องจากงานวิจัยนี้ไม่ได้หาฟังก์ชั่นในลักษณะเส้นตรง ดังนั้นรูปแบบ Transfer functions ที่นำมาใช้สำหรับส่งผ่านข้อมูลไปยังชั้นถัดไปในงานวิจัยนี้ ได้แก่ Log-Sigmoid, Tan-Sigmoid ซึ่งหาได้จากสมการที่ 2.40 และ 2.41

Log-Sigmoid $$y = \frac{1}{1+e^{-a}}$$ ---------- (2.40)

เมื่อ $a = w^T p$

Tan-Sigmoid $$y = \frac{e^{a} - e^{a-1}}{e^{a} + e^{a-1}}$$ ---------- (2.41)

เมื่อ $a = w^T p$

วรรณกรรมที่เกี่ยวข้อง

งานวิจัยเกี่ยวกับการจำแนกหมวดหมู่ด้วยเนื้อหาภายในเสียง (Content-based Music Genre Classification) นั้นจะแบ่งเนื้อหาหรือวิธีการทำเป็นสองส่วน คือส่วนของการเลือกและสกัดค่าคุณลักษณะ และก็ส่วนของการจำแนกหมวดหมู่ งานวิจัยนี้จึงได้ทำการรวบรวมและศึกษางานวิจัยต่าง ๆ ที่เกี่ยวข้อง เพื่อทำการศึกษาวิธีการ รวมถึงเทคนิคในการจำแนกข้อมูลเสียงดังต่อไปนี้

Mingchun Liu และ Chunru Wanp (2001: 247-248) ทำการศึกษาวิธีการเพื่อหาคุณลักษณะที่ดีที่สุดที่จะนำมาใช้ในการจำแนกเสียงเครื่องดนตรีจากสัญญาณเสียง โดยการแบ่งสัญญาณเสียงของแต่ละแฟ้มข้อมูลออกเป็น 58 คุณลักษณะ ซึ่งมีแฟ้มข้อมูลในฐานข้อมูลทั้งหมด 351 แฟ้ม โดยข้อมูลภายในฐานข้อมูลจะแบ่งออกเป็น 5 ประเภท ซึ่งแบ่งตามธรรมชาติของเสียง คือ เสียงที่เกิดจากเครื่องเป่าทองเหลือง เสียงของเปียโน เสียงของเครื่องเคาะ เสียงของเครื่องสาย และเสียงของเครื่องเป่า แล้วทำการเลือกกลุ่มของคุณลักษณะของเสียง (feature set) ด้วยวิธี Sequential Forward Selection (SFS) เพื่อเลือกกลุ่มของคุณลักษณะที่ให้ผลสัมฤทธิ์ในการจำแนกที่ดีที่สุด แล้วใช้เทคนิค Nearest Neighbor (NN), k-Nearest Neighbor (k-NN) และ Gaussian Mixture Model

(GMM) ในการทดสอบเพื่อหาผลสัมฤทธิ์ ซึ่งพบว่าค่าของความแม่นยำในการจำแนกจะเพิ่มขึ้นอย่างรวดเร็วตามคุณลักษณะของเสียงแล้วสม่ำเสมอในช่วงกลาง จากนั้นความแม่นยำจึงลดลงในช่วงท้าย และได้ค่าความแม่นยำในการจำแนกมากที่สุดอยู่ที่ 93% เมื่อทำการทดสอบด้วย k-NN โดยใช้คุณลักษณะ 19 คุณลักษณะคือ 6 temporal, 8 spectral และ 9 coefficient ซึ่งเป็นผลที่น่าพอใจและเป็นประโยชน์สำหรับโปรแกรมแบบ real-time เช่น การค้นคืนข้อมูลเสียงจากฐานข้อมูลขนาดใหญ่

ในการจำแนกกลุ่มเพลงสำหรับการนำไปประยุกต์ใช้ประโยชน์ในการจัดกลุ่ม การค้นคืนของระบบฐานข้อมูลเสียงที่ใช้ในห้องสมุดเสียงนั้น ได้มีผู้วิจัยไว้เช่นกัน คือ Downie, J. S., Ehmann, A. F., และ Tcheng, D (2005 : 377-377) ได้ทำการจำแนกกลุ่มเพลง 14 ประเภทโดยแบ่งตาม Genre ของบทเพลงและใช้ feature 40 feature ซึ่งสกัดจากวิธีการ spectral centroid, spectral roll-off, spectral flux, zero crossing rate จำนวนทั้งสิ้น 20 feature และจาก spectral envelope 16 feature โดยในการจำแนกนั้นใช้วิธีการ Decision tree ระดับ 10 จำนวน 13 tree ซึ่งทำการแบ่งข้อมูลในการเรียนรู้ (Train) และทดสอบ (Test) เท่า ๆ กัน คือ 50% ของข้อมูลทั้งหมดคือ 70,000 ตัวอย่าง ซึ่งมีความยาวของเสียงตัวอย่างละ 10 วินาที ได้ผบความแม่นยำในการจำแนก 72.9% ซึ่งการจำแนกด้วยมนุษย์ได้ผลความแม่นยำอยู่ที่ 70% เท่านั้น

สำหรับงานวิจัยเกี่ยวกับการจำแนกเพลงไทยนั้นได้มีผู้วิจัยเอาไว้คือ อานนท์ (2549) ได้ทำการศึกษาวิจัยการจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีน ไว้ดังนี้ การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีน เป็นการจำแนกเสียงเพลงไทยสามกลุ่มคือ เพลงลูกทุ่ง เพลงลูกกรุง และเพลงไทยเดิม โดยเลือกมาเฉพาะแทร็คดนตรี ประเภทละ 25 เพลง โดยในแต่ละประเภทนั้นจะถูกแบ่งเป็น ชุดข้อมูลในการเรียนรู้ (Train) 15 เพลง และชุดข้อมูลสำหรับทดสอบ 10 เพลง โดยเพลงที่ใช้ทดลองนั้นทำได้ทำการตัดออกเป็นเพลงละ 2000 เฟรม และใช้ค่าคุณลักษณะ (feature) แบบ Beat Spectrum, LPC-Derived Ceptrum, Zero Crossing Rate, Spectrum Power, Mel Frequency Coefficient – MFCC และ Line Spectral Frequency แล้วนำมาทดสอบด้วย Support Vector Machine โดยใช้ เคอร์เนลเรเดียลเบลิสฟังก์ชั่น (Radial Basis Function-RBF) ซึ่งได้ผลการจำแนกถูกต้อง 83%

Li, T., Ogihara, M., and Li, Q. (2003 : 282-289) ได้ทำการศึกษาเปรียบเทียบวิธีการจำแนกหมวดหมู่เพลงโดยอัตโนมัติโดยการนำเสนอค่าคุณลัษณะใหม่ที่เรียกว่า Daubechies Wavelet Coefficient Histograms (DWCHs) โดยนำมาเปรียบเทียบกับค่าคุณลักษณะแบบ Timbrel Feature ซึ่งประกอบด้วยค่า Mel Frequency Coefficient (MFCC), Spectral Centriod, Spectral Rolloff, Spectral Flux, Zero Crossing Rate, Low Energy ค่าคุณลักษณะแบบ Beat ซึ่งสกัดได้จาก

การคำนวณ Beat Spectrum ของบทเพลง และค่าคุณลักษณะแบบ Pitch ซึ่งสกัดจากการคำนวณ FFT จากความถี่ของเพลง โดยทำการจำแนกกลุ่มเพลงสากล 10 กลุ่มซึ่งประกอบไปด้วย Blues, Classical, Country, Disco, Hip-hop, Jazz, Metal, Pop, Reggae และ Rock จำนวน 100เพลงต่อกลุ่ม ด้วยวิธีการ Support Vector Machine (SVM), k-Nearest Neighbor (k-NN) และ Gaussian Mixture Models (GMM) โดยได้ผลการทดลองว่าค่าคุณลักษณะที่นำเสนอคือ DWCHs นั้นได้ค่าความถูกต้องในการจำแนกดีที่สุดเมื่อเปรียบเทียบกับค่าคุณลักษณะแบบอื่น ๆ รวมกัน และวิธีการในการจำแนกทีให้ผลการทดลองที่ดีที่สุดคือวิธีการ Support Vector Machine (SVM)

Chen, S., and Chen, S (2009 : 1095-1101) ได้ทำการจำแนกหมวดหมู่เพลงสากล จำนวน 9 หมวดหมู่คือ Classical, Jazz, Dance, Lullaby, Country, Bossa, Piano, Bluesและ Hip-hop ด้วยค่าคุณลักษณะแบบ Timbrel Feature โดยใช้ค่า Mel Frequency Coefficient (MFCC) ร่วมกับค่า log energy ด้วยวิธีการ Support Vector Machine (SVM) โดยทำการทดลองปรับเปลี่ยนขนาดของเฟรมที่นำมาสกัดค่าคุณลักษณะซึ่งได้ผลว่าขนาดความยาวของเพลงที่ 41 มิลลิวินาที ให้ค่าความถูกต้องที่ดีที่สุดคือ 86%

Shyamala Doraisamy et al. (2008 : 331-336) ได้ทำการศึกษาเพื่อเลือกค่าคุณลักษณะและเทคนิคในการจำแนกหมวดหมู่เพลงพื้นบ้านของประเทศมาเลเซีย ซึ่งแบ่งออกเป็น 10 หมวดหมู่คือ Dikir Bart, Etnik Sabah, Gamelan, Inang, Joget, Keroncong, Tumbuk Kalang, Wayang Kulit และ Zupin โดยเลือกใช้ค่าคุณลักษณะแบบ STFT ,Mel Frequency Coefficient (MFCC) และ Beat ร่วมกัน ในส่วนของการจำแนกหมวดหมู่เพลงพื้นบ้านมาเลเซียนั้น ได้ทำการทดลองเปรียบเทียบวิธีการ 18 วิธีที่มีอยู่ในโปรแกรม Weka ซึ่งเป็นโปรแกรมฟรีสำหรับการทำ Machine Learning และ Data Mining ดังนี้คือAIRS (Artificial Immune System (AIS)), Bagging, Bayesian Network, Cart, Conjunctive rule learner (Conj-Rules), Decision Stump, Decision Table, IB1, J48 (C4.5), Kstar, Logistic, LogitBoost, Multi-layer neural network with back propagation (MLP), Nave Bayesian, Nbtree, PART, RBF Network, และ SMO (Support vector machine) โดยได้ผลว่าวิธีการที่ได้ค่าความถูกต้องในการจำแนกที่ดีที่สุดคือวิธีการ MLP (Multi-layer neural network with back propagation) โดยได้ค่าความถูกต้องในการจำแนก 88.6%

โดยสามารถสรุปวิธีการที่ใช้ในการจำแนกที่ได้ผลดีที่สุดและค่าคุณลักษณะแบบต่าง ๆ ที่ใช้ในงานวิจัยที่เกี่ยวข้องที่ได้กล่าวมาแล้ว ได้ดังตารางที่ 2

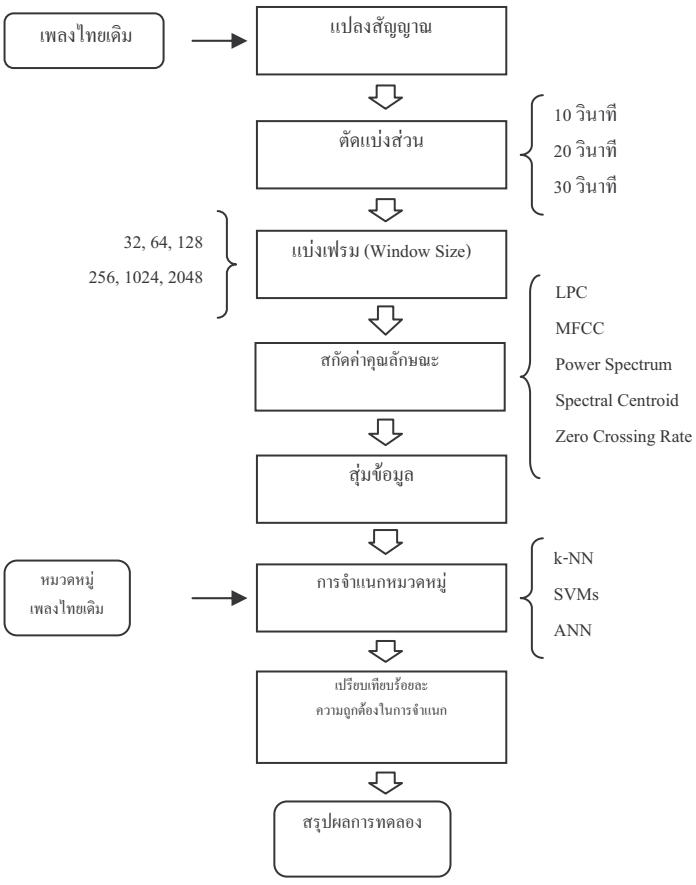
ตารางที่ 2 สรุปค่าคุณลักษณะแบบต่าง ๆ และวิธีการที่ใช้ในการจำแนกที่ได้ผลดีในงานวิจัยที่เกี่ยวข้อง

ชื่อผู้วิจัย	ค่าคุณลักษณะที่ใช้						วิธีการที่ใช้ในการจำแนกที่ได้ผลดี
Chen, S., and Chen, S	Mel Frequency Coefficient (MFCC)	log energy					SVM
Li, T., Ogihara, M., and Li, Q.	Mel Frequency Coefficient (MFCC)	Low Energy	Spectral Centriod	Spectral Rolloff	Spectral Flux	Zero Crossing Rate	SVM
Shyamala Doraisamy et al	Mel Frequency Coefficient (MFCC)	STFT	Beat				MLP (Multi-layer neural network with back propagation)
อานนท์	Mel Frequency Coefficient (MFCC)	LPC-Derived Ceptrum	Zero Crossing Rate	Power Spectrum	Line Spectral Frequency	Beat Spectrum	SVM

จากตารางที่ 2 พบว่าในการวิจัยการจำแนกหมวดหมู่บทเพลงแบบต่าง ๆ นั้นวิธีการ Support Vector Machine (SVM) เป็นวิธีการที่ได้ค่าความถูกต้องในการจำแนกที่ดี ดังนั้นผู้วิจัยจึงตั้งสมมุติฐานว่าวิธีการ Support Vector Machine (SVM) น่าจะเป็นวิธีการที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิม ส่วนค่าคุณลักษณะที่นำมาใช้ในการจำแนกนั้น จะเห็นได้ว่าค่าคุณลักษณะแบบ Mel Frequency Coefficient (MFCC) เป็นค่าคุณลักษณะที่นิยมนำมาใช้ในงานวิจัยต่าง ๆ จึงน่าจะได้ให้ค่าความถูกต้องที่ดีเมื่อนำมาใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม

บทที่ 3
วิธีดำเนินการวิจัย

บทนี้กล่าวถึงวิธีดำเนินการวิจัย ซึ่งในการจำแนกหมวดหมู่ของเพลงไทยเดิมนั้นมีขั้นตอนหลักอยู่ 5 ขั้นตอนคือ การรวบรวมเพลงไทยเดิม การเตรียมข้อมูลเพื่อนำไปใช้ในการจำแนก การสกัดค่าคุณลักษณะของเพลงไทยเดิม และการจำแนกหมวดหมู่เพลงไทยเดิมซึ่งประกอบไปด้วยวิธีการในการจำแนก 3 วิธีการได้แก่ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) และในขั้นตอนสุดท้ายเป็นการเปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมของทั้งสามวิธีการ ซึ่งจะอธิบายขั้นตอนโดยละเอียดต่อไปนี้



ภาพที่ 12 ขั้นตอนการดำเนินการวิจัย

การรวบรวมเพลงไทยเดิม

ทำการรวมรวมเพลงไทยเดิมที่อยู่ในรูปของ Compact DiskหรือCassette Tapeจากห้องสมุดต่าง ๆ และเพลงไทยเดิมที่มีจำหน่ายตามท้องตลาด ซึ่งมีการระบุหมวดหมู่โดยผู้ประพันธ์หรือผู้จัดทำอัลบัมนั้น ๆ ไว้แล้ว โดยคัดเลือกเพลงจากหมวดหมู่ต่าง ๆ คือ วงปี่พาทย์ วงเครื่องสายและวงมโหรีหมวดหมู่ละ 25 เพลง ดังนั้นข้อมูลเพลงไทยเดิมที่นำมาทดลองรวมทั้งสิ้น 75 เพลง

การเตรียมข้อมูลเพื่อนำไปใช้ในการจำแนก

1. การแปลงสัญญาณเสียง

เนื่องจากข้อมูลที่ทำการรวบรวมนั้นมาจากแหล่งต่าง ๆ กันดังนั้นจึงต้องทำการจัดทำข้อมูลให้อยู่ในรูปแบบเดียวกันเสียงก่อน ซึ่งในขั้นตอนนี้เป็นการแปลงสัญญาณเสียงที่ได้มาจาก Compact Disks หรือ Classes Tape ให้อยู่ในรูปของแฟ้มสัญญาณเสียงนามสกุล wav ด้วยโปรแกรมแปลงสัญญาณเสียง (Audio Convert Software) โดยในการแปลงสัญญาณเสียงในงานวิจัยนี้จะใช้ค่า Sampling Rate หรือค่าการสุ่มหน้าคลื่นสัญญาณเพื่อแทนด้วยสัญญาณดิจิตอลที่ 44 kHz ซึ่งเป็นค่าที่ถ่ายทอดเสียงได้ใกล้เคียงเสียงต้นฉบับ ทำให้ได้ข้อมูลไฟล์เสียงนามสกุล wav เพื่อนำไปใช้ในการสกัดค่าคุณลักษณะต่อไป

2. การตัดแบ่งส่วนเพลง

ก่อนที่จะนำข้อมูลไปสกัดค่าคุณลักษณะจะต้องทำการตัดแบ่งเสียงเพลงเสียก่อนเพื่อศึกษาว่าการตัดแบ่งเพลงกี่วินาทีที่จะให้ผลการจำแนกที่ถูกต้องมากที่สุด โดยในการทดลองนี้ผู้วิจัยได้ทำการตัดแบ่งเพลงตั้งแต่วินาทีที่ 0 ถึง 10, วินาทีที่ 0 ถึง 20 และวินาทีที่ 0 ถึง 30

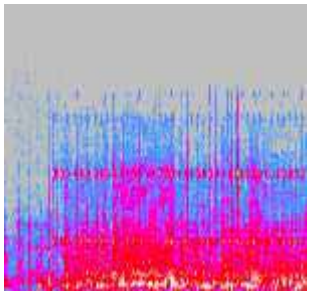
การสกัดค่าคุณลักษณะ

ในขั้นตอนนี้เป็นขั้นตอนของการสกัดค่าคุณลักษณะของเพลงไทยเดิม โดยนำเพลงที่ตัดแบ่งออกเป็นความยาวต่าง ๆ มาหาค่าคุณลักษณะของแต่ละเพลง เพื่อนำไปใช้ในการจำแนกหมวดหมู่ของเพลงไทยเดิม โดยผู้วิจัยได้ใช้ jAudio (McEnnis et al. 2005) ซึ่งเป็นส่วนหนึ่งของชุดโปรแกรม jMIRที่พัฒนาขึ้นเพื่องานวิจัยด้านการค้นคืนข้อมูลดนตรี (Music Information Retrieval) และการจัดหมวดหมู่ดนตรีโดยอัตโนมัติ (Automatic Music Classification) ในการสกัดค่าคุณลักษณะต่าง ๆ ซึ่งมีขั้นตอนดังต่อไปนี้

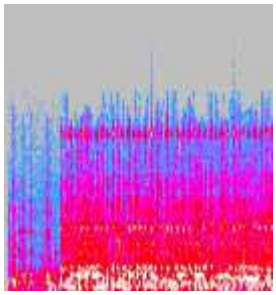
1. การเลือกค่าคุณลักษณะ (Feature Extraction)

เมื่อได้เพลงที่แบ่งออกเป็นเฟรมแล้ว จึงเข้าสู่กระบวนการสกัดค่าคุณลักษณะซึ่งเป็นการแปลงข้อมูลจากคลื่นเสียงออกเป็นชุดตัวเลข เพื่อนำไปใช้ในการจำแนกหมวดหมู่ของดนตรี

ไทย โดยค่าคุณลักษณะต่าง ๆ ก็จะให้ค่าชุดตัวเลขที่แตกต่างกันออกไป โดยในเพลงไทยเดิมนั้นจะมีจังหวะและท่วงทำนองที่คล้ายคลึงกัน การประพันธ์เพลงหนึ่ง ๆ อาจนำมาใช้ในการบรรเลงด้วยประเภทวงที่แตกต่างกันได้เช่นในภาพที่ 13



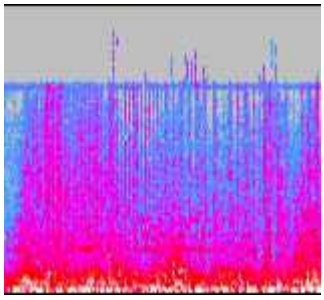
วงมโหรี



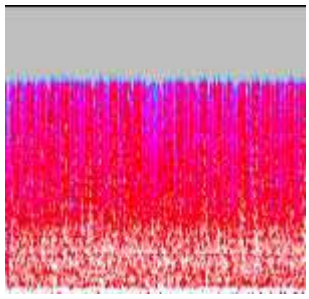
วงปี่พาทย์

ภาพที่ 13 เปรียบเทียบ Spectrum ของเพลงค้างคาวกินกล้วยที่บรรเลงด้วยวงมโหรีและวงปี่พาทย์

จากภาพที่ 13 จะเห็นได้ว่า Spectrum ที่ได้จากจังหวะและท่วงทำนองของเพลงไทยเดิมนั้นมีความใกล้เคียงกันมากซึ่งต่างจากเพลงสากลดังภาพที่ 14



Reggae



Rock

ภาพที่ 14 เปรียบเทียบ Spectrum ของเพลง Reggae และเพลง Rock

ดังนั้นเพลงไทยเดิมจึงอาจจะไม่สามารถใช้ค่าคุณลักษณะแบบจังหวะ (Beat) มาทำการจำแนกได้ ดังนั้นในงานวิจัยนี้ได้ใช้ค่าคุณลักษณะแบบ Timbrel Feature ซึ่ง เตาลี่และคณะ (2003) กล่าวไว้ว่าเป็นค่าคุณลักษณะที่นิยมนำใช้ในงานวิจัยเกี่ยวกับการรู้จำเสียงพูด (Speech Recognition) เพราะสามารถบ่งบอกเอกลักษณ์ของเสียงได้ โดยเลือกใช้ค่าคุณลักษณะแบบ Timbrel จากงานวิจัยที่เกี่ยวข้องดังต่อไปนี้

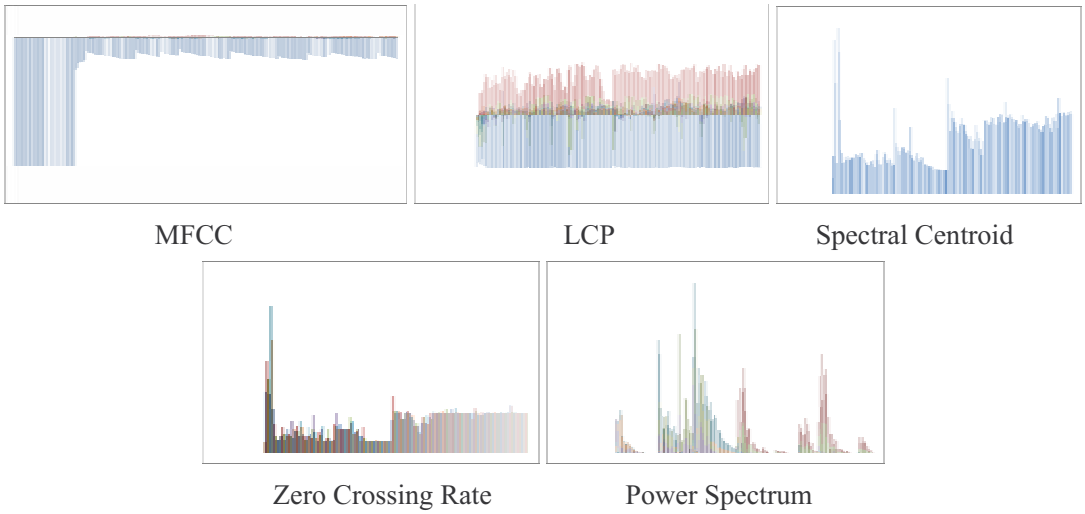
Linear Predictive Coefficients (LPC) เป็นค่าคุณลักษณะที่ได้การสุ่มตัวอย่างของสัญญาณเสียง ซึ่งสามารถสร้างได้จากผลรวมเชิงเส้น (Linear Combination) ของการสุ่มตัวอย่างสัญญาณเสียงหลาย ๆ ตัวอย่างที่ผ่านมาในอดีต

Mel-frequency cepstral coefficients (MFCC) เป็นค่าคุณลักษณะที่ได้มาจากการปรับปรุงสัมประสิทธิ์เซปสตรัม ด้วยการปรับสเกลของสเปกตรัมให้อยู่บนสเกลที่เหมาะสมสำหรับการรับฟังของมนุษย์ จึงมีการออกแบบสเกลของสเปกตรัมให้สามารถเก็บรายละเอียดของสัญญาณเสียงช่วงความถี่ต่ำได้มากกว่า เรียกว่า สเกลเมล (Mel scale)

Spectral Centroid คือค่าคุณลักษณะที่แสดงถึงค่ากึ่งกลางของ Power Spectrum โดยคำนวณจากค่าเฉลี่ยของค่า Power Spectrum

Zero Crossing Rate คือค่าคุณลักษณะที่ได้จากการนับจำนวนครั้งของสัญญาณที่ผ่านแกนศูนย์

โดยงานวิจัยนี้จะใช้นำค่าคุณลักษณะแบบ Timbrel Feature และค่าคุณลักษณะแบบพลังงานสเปกตรัม Power Spectrum ซึ่ง **พลังงานสเปกตรัม (Power Spectrum)** คือค่าคุณลักษณะที่คำนวณได้จากสมการการแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform - FFT) โดยจะถูกหาค่าน้ำหนักก่อนด้วย Hamming Window



ภาพที่ 15 ตัวอย่างค่าคุณลักษณะแบบต่าง ๆ ที่ใช้ในงานวิจัย

โดยค่าคุณลักษณะที่จะนำไปใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมนั้น จะถูกคัดเลือกซึ่งมีขั้นตอนดังนี้

1.นำค่าคุณลักษณะแต่ละแบบไปทำการจำแนกหมวดหมู่ เพื่อศึกษาค่าคุณลักษณะที่ได้ค่าความถูกต้องในการจำแนกที่ดีที่สุด แล้วนำมาเรียงลำดับตามค่าความถูกต้องที่ได้

2.นำค่าคุณลักษณะที่ได้ลำดับที่ 1 จากการเรียงลำดับในข้อ 1 มาเป็นค่าคุณลักษณะหลักแล้วนำเอาค่าคุณลักษณะลำดับต่าง ๆ มาทำการสกัดค่าคุณลักษณะร่วมกับค่าคุณลักษณะลำดับที่ 1 แล้วนำไปจำแนกหมวดหมู่เพื่อศึกษาค่าความถูกต้องในการจำแนกที่ได้

3.นำคู่ของค่าคุณลักษณะที่ได้ค่าความถูกต้องที่มากที่สุดในข้อ 2 เป็นหลักแล้วนำเอาค่าคุณลักษณะที่เหลือมาสกัดค่าคุณลักษณะร่วมด้วย แล้วนำไปจำแนกหมวดหมู่

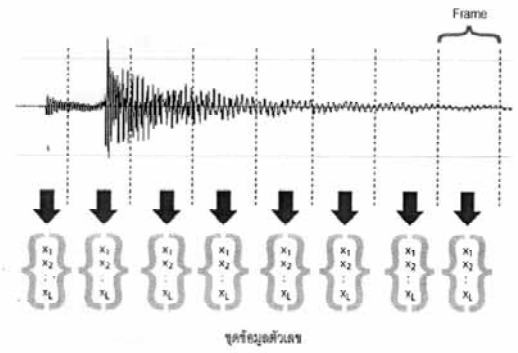
4.นำค่าคุณลักษณะที่ใช้ร่วมกัน 3 แบบที่ได้จากข้อ 3 มาเป็นค่าคุณลักษณะหลักแล้วนำค่าคุณลักษณะที่เหลือมาสกัดค่าคุณลักษณะร่วม แล้วนำไปจำแนกหมวดหมู่เพื่อศึกษาค่าความถูกต้องในการจำแนกที่ได้

5.นำค่าคุณลักษณะทั้ง 5 แบบมาสกัดค่าคุณลักษณะร่วมกันแล้วนำไปจำแนกหมวดหมู่เพื่อศึกษาค่าความถูกต้องในการจำแนก

จากนั้นทำการเปรียบเทียบค่าความถูกต้องของการใช้ค่าคุณลักษณะทั้ง 5 แบบโดยใช้ค่าเฉลี่ยของค่าความถูกต้องในการจำแนกของการสกัดค่าคุณลักษณะ เพื่อเลือกรูปแบบการใช้ค่าคุณลักษณะและเลือกค่าคุณลักษณะชุดที่ให้ค่าความถูกต้องมากที่สุดมาเป็นค่าคุณลักษณะที่จะนำไปใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม

2. การแบ่งข้อมูลออกเป็นเฟรม

งานวิจัยนี้ต้องการศึกษาว่าขนาดของเฟรมที่แบ่งตามขนาดของ Sample นั้น มีผลต่อค่าความถูกต้องของการจำแนกหรือไม่ ดังนั้นจึงต้องทำการแบ่งข้อมูลออกเป็นเฟรมอีกครั้งหนึ่ง โดยนำเพลงแต่ละเพลงมาทำการแบ่งส่วนข้อมูลออกเป็นเฟรมที่ขนาด 32, 64, 128, 256, 512, 1024 และ 2048 Sample เพื่อนำไปหาค่าคุณลักษณะของแต่ละเฟรมดังแสดงตามภาพที่ 16



ภาพที่ 16 แสดงการแบ่งขนาด Frame ของเพลง

3. การสุ่มข้อมูลตัวอย่าง

เมื่อทำการสกัดค่าคุณลักษณะแล้ว ค่าคุณลักษณะที่ได้จะเป็นกลุ่มของชุดตัวเลข ซึ่งอาจมีจำนวนของข้อมูลมากทำให้ข้อมูลที่จะนำไปใช้ในการจำแนกนั้นมีจำนวนมากอาจทำให้ใช้เวลาในการจำแนกนานกว่าที่จำเป็น และยังอาจเกิดความผิดพลาดเนื่องมาจากข้อมูลที่ไม่สามารถจำแนกได้ ผู้วิจัยจึงใช้การสุ่มข้อมูลตัวอย่าง โดยสุ่มจากจำนวนค่าคุณลักษณะที่สกัดมาแล้ว ด้วยวิธีการสุ่มกลุ่มตัวอย่าง (Sampling) จากตารางเครจซี่ มอร์แกน (Krejcie & Morgan 1970 : 607-610) โดยเลือกใช้ที่ช่วงความเชื่อมั่น 99% และที่ความคลาดเคลื่อน (error) เท่ากับ 0.01 (ภาคผนวก ก) ซึ่งนำมาเทียบเป็นจำนวนของข้อมูลทั้งหมด และจำนวนของข้อมูลตัวอย่างออกมาเป็นเปอร์เซนต์แล้วนำมาใช้ในการสุ่มข้อมูล

การจำแนกหมวดหมู่เพลงไทยเดิม

ในการจำแนกหมวดหมู่เพลงไทยเดิมนั้น จะการทดสอบความถูกต้องของการจำแนกด้วยวิธีการ 10-fold cross-validation คือ นำข้อมูลที่ผ่านการสกัดค่าคุณลักษณะมาแล้ว มาแบ่งเป็นส่วนเท่า ๆ กัน 10 ส่วน นำส่วนที่ 9 ถึง 10 ไปเป็นชุดข้อมูลเรียนรู้ (Training Set) แล้วนำส่วนที่ 1 มาเป็นชุดข้อมูลทดสอบ (Test Set) จากนั้นไขว้พับเปลี่ยนชุดข้อมูลเรียนรู้และทดสอบจนครบ 10 ครั้ง แล้วจึงหาค่าเฉลี่ยในการจำแนกทั้ง 10 ครั้ง ซึ่งวิธีการจำแนกหมวดหมู่เพลงไทยเดิมนั้นได้แบ่งออกเป็น 3 วิธีการ คือ การจำแนกด้วย k-Nearest Neighbor (k-NN), การจำแนกด้วย Support Vector Machine (SVM) และการจำแนกด้วย Artificial Neural Network (ANN) โดยสามารถอธิบายได้ดังนี้

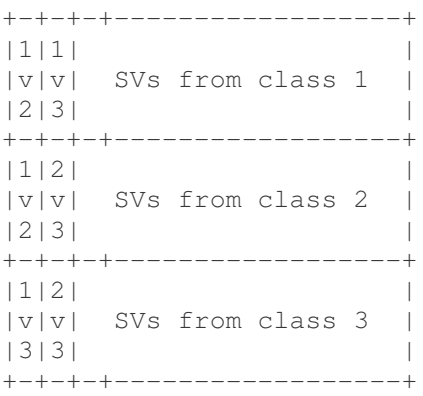
1. การจำแนกด้วยวิธีการ k-Nearest Neighbor (k-NN)

k-Nearest Neighbor (k-NN) เป็นวิธีการที่ทำงานโดยการหาระยะห่างของแต่ละชุดข้อมูล ด้วยสมการ Euclidean distance โดยในงานวิจัยนี้จะทำการกำหนดค่า k ตั้งแต่ 1 ถึง 50 เพื่อศึกษาว่าค่า k หรือค่ากลุ่มของข้อมูลที่ดีที่สุด ในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN) นั้นเป็นเท่าใด

2. การจำแนกด้วยวิธีการ Support Vector Machine (SVMs)

Support Vector Machine (SVMs) นั้น ทำงานโดยอาศัยการสร้างเส้นแบ่ง (Plane) เพื่อแยกกลุ่มชุดข้อมูลตัวเลขที่เป็นข้อมูลสองหมวดหมู่ออกจากกัน และสามารถแบ่งกลุ่มข้อมูลได้ครั้งละ 2 หมวดหมู่เท่านั้น แต่เนื่องจากหมวดหมู่ของเพลงไทยเดิมที่ต้องการจำแนกมีทั้งหมด 3 หมวดหมู่ ผู้วิจัยจึงใช้ WLSVM (Weka LibSVM) (Yasser and Vasant 2005) ที่พัฒนา

มาจาก LIBSVM (Chang and Lin 2001) โดยWLSVMนั้น เป็นส่วนเสริมของ Weka ที่สามารถจำแนกข้อมูลด้วยวิธีการ Support Vector Machine แบบหลายหมวดหมู่ได้ ในงานวิจัยนี้กำหนดหมวดหมู่ (Class) เท่ากับ 3 WLSVM ก็จะทำการจำแนกโดย จำแนกหมวดหมู่ที่ 1 กับหมวดหมู่ที่ 2 และ 3 จำแนกหมวดหมู่ที่ 2 กับหมวดหมู่ที่ 1 และ 3 และสุดท้ายจะจำแนกหมวดหมู่ที่ 3 กับหมวดหมู่ที่ 1 และ 2 เพื่อหาซัพพอร์ตเวคเตอร์ของแต่ละหมวดหมู่ ซึ่งจะได้ซัพพอร์ตเวคเตอร์ของทั้ง 3หมวดหมู่ เพื่อใช้เป็น Modelในการทดสอบการจำแนกต่อไป (Chang and Lin 2001) ดังภาพที่ 17



ภาพที่ 17 การจำแนกหมวดหมู่เพื่อหาซัพพอร์ตเวคเตอร์ของแต่ละหมวดหมู่ด้วย LibSVM
ที่มา : Chih-Chung Chang and Chih-Jen Lin. LIBSVM FAQ [Online], accessed 20 November 2009. Available from http://www.csie.ntu.edu.tw/~cjlin/libsvm/faq.html

การจำแนกหมวดหมู่ด้วย Support Vector Machine นั้น หากข้อมูลไม่วางตัวอยู่ในพื้นที่คุณลักษณะ จะทำให้ไม่สามารถแบ่งหมวดหมู่ของข้อมูลออกจากกันได้ ซึ่งจะมีการใช้ Kernel เข้าไปช่วยในการวางกลุ่มข้อมูลใหม่ โดยจะทำการทดลองเพื่อเลือก Kernel ที่ดีที่สุดในการจำแนกหมวดหมู่เพลงไทยเดิมจาก Kernel แบบ Linear, Polynomial, Radial basis function และ Sigmoid

3. การจำแนกด้วยวิธีการ Artificial Neural Network (ANN)

งานวิจัยนี้ใช้ Artificial Neural Network (ANN) แบบ Back-Propagation ซึ่งเป็นที่นิยมใช้ใน Multilayer Perceptron เพื่อปรับค่าน้ำหนักในเส้นเชื่อมต่อระหว่างโหนดให้เหมาะสม โดยมีชั้น (Layer) ทั้งหมด 3 ชั้นคือ ชั้นอินพุต (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นเอาต์พุต (Output Layer) ซึ่งจำนวนโหนดของชั้นเอาต์พุตคือ 3 ตามจำนวนของหมวดหมู่เพลงไทยเดิม ชั้นซ่อนจะทำการทดลองเพื่อเลือกจำนวนของชั้นซ่อนที่ดีที่สุดจาก ค่าเฉลี่ยของ

ผลรวมระหว่างจำนวนของค่าคุณลักณะและจำนวนของหมวดหมู่ และจำนวนของหมวดหมู่เพลงไทยเดิมส่วนโหนดในชั้นอินพุตนั้นกำหนดจากจำนวนกำหนดโหนดเท่ากับจำนวนของค่าคุณลักษณะที่สกัดได้ และTransfer Function ที่ใช้คือSigmoid ทั้งหมด

การเปรียบเทียบค่าความถูกต้องของการจำแนก

เมื่อทำการจำแนกหมวดหมู่เพลงไทยเดิมทั้งสามวิธีการโดยทำการปรับค่าพารามิเตอร์เพื่อหาค่าความถูกต้องที่มากที่สุดในแต่ละวิธีการแล้ว งานวิจัยนี้จะนำค่าความถูกต้องในการจำแนกที่มากที่สุดของทั้งสามวิธีการ มาทำการเปรียบเทียบ เพื่อศึกษาว่าวิธีการใดให้ค่าความถูกต้องในการจำแนกมากที่สุด และใช้คุณลักษณะแบบใดจึงจะเหมาะสมที่สุด

สรุปผลการทดลองและจัดทำรายงานวิทยานิพนธ์

เมื่อได้ผลการทดลองตรงตามวัตถุประสงค์ที่กำหนดแล้ว จึงจัดทำรายงานวิทยานิพนธ์

บทที่ 4
ผลการดำเนินการวิจัย

หลังจากที่ได้ดำเนินการวิจัยมาแล้วนั้น ในส่วนนี้จะกล่าวถึงผลการดำเนินการวิจัย โดยประกอบไปด้วยส่วนของการรวบรวมข้อมูลเพลงไทยเดิม การเตรียมข้อมูล ผลของการเลือกค่าคุณลักษณะ และผลการจำแนกหมวดหมู่เพลงไทยเดิมทั้ง 3 วิธีการ คือ การจำแนกด้วย k-Nearest Neighbor (k-NN), การจำแนกด้วย Support Vector Machine (SVM) และการจำแนกด้วย Artificial Neural Network (ANN) ซึ่งจะอธิบายโดยละเอียดในแต่ละข้อดังนี้

การรวบรวมข้อมูลเพลงไทยเดิมที่ใช้ในงานวิจัย

ผู้วิจัยได้รวบรวมเพลงไทยเดิมจาก Compact Disk ที่มีอยู่ในห้องสมุดและจากที่จัดจำหน่ายตามต้องตลาด โดยคัดเลือกเพลงจากหมวดหมู่ต่าง ๆ ที่มีการแบ่งหมวดหมู่โดยผู้ประพันธ์หรือผู้จัดทำอัลบัมไว้แล้วคือ วงปี่พาทย์ วงเครื่องสาย และวงมโหรี หมวดหมู่ละ 25 เพลง รวมทั้งสิ้น 75เพลง (ภาคผนวก ข) โดยได้แสดงตัวอย่างรายชื่อเพลงที่ใช้ในการทดลองไว้ตามตารางที่ 3

ตารางที่ 3 แสดงตัวอย่างรายชื่อเพลงที่ใช้ในการทดลอง

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
1	มโหรีย์	ชงโค	1 ค้างคาวกินกล้วย
2	มโหรีย์	ชงโค	2 แขกอาหวัง
3	มโหรีย์	ชงโค	3 แขกหนัง
4	มโหรีย์	ชงโค	4 แขกสาหร่าย
5	มโหรีย์	ชงโค	5 พัดชา
6	มโหรีย์	ชงโค	6 มฤคระเริง
7	มโหรีย์	ชงโค	7 เต่าเห่
8	มโหรีย์	ชงโค	8 เขมรพายเรือ

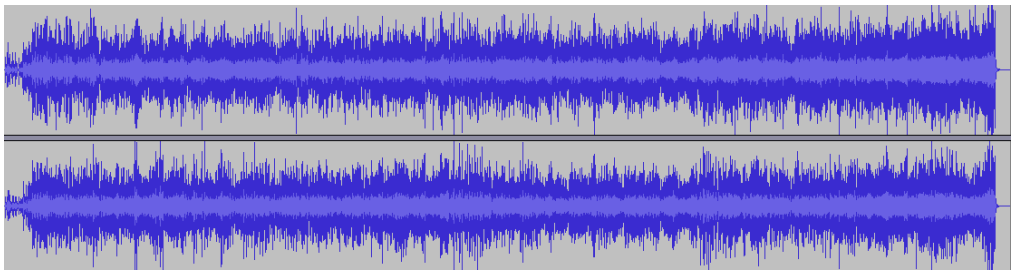
ตารางที่ 3 (ต่อ)

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
9	มโหรีย์	มโหรีย์	1 เขมรโพธิสัตว์
10	มโหรีย์	มโหรีย์	2 แขกสาหร่าย
11	มโหรีย์	มโหรีย์	3 ลาวดวงเดือน
12	มโหรีย์	มโหรีย์	4 ลาวเสี่ยงเทียน
13	มโหรีย์	มโหรีย์	5 ยอเร
...	...	...	...

การเตรียมข้อมูล

1. การแปลงรูปแบบ

เนื่องจากเพลงที่ผู้วิจัยได้ทำการรวบรวมนั้น เป็นเพลงที่อยู่ในรูปแบบของ Compact Disk ดังนั้นจึงต้องทำการแปลงข้อมูลทั้ง 75 เพลงให้อยู่ในรูปแบบเดียวกันเสียก่อน และเพื่อเป็นการหลีกเลี่ยงความผิดพลาดที่จะเกิดจากการบีบอัดข้อมูล งานวิจัยนี้จึงเลือกแปลงข้อมูลให้อยู่ในรูปของไฟล์นามสกุล wav ที่มีค่า Sampling Rate ที่ 44 kHz เพื่อพร้อมจะนำไปดำเนินการวิจัยในขั้นต่อไป ซึ่งเมื่อแปลงแล้วจะมีค่า waveform ดังภาพที่ 18



ภาพที่ 18 แสดงตัวอย่างwaveform ของเพลงค้างคาวกินกล้วยเมื่อแปลงเป็นไฟล์ wav ที่มีค่า Sampling Rate ที่ 44 kHz

2. การตัดแบ่งส่วนเพลงไทยเดิม

ในงานวิจัยนี้ต้องการศึกษาว่าการตัดแบ่งเพลงกี่วินาทีที่จะให้ผลการจำแนกที่ถูกต้องมากที่สุด โดยในการทดลองนี้ผู้วิจัยได้ทำการตัดแบ่งเพลงทั้ง 75 เพลง ตั้งแต่วินาทีที่ 0 ถึง

10, วินาทีที่ 0 ถึง 20 และวินาทีที่ 0 ถึง 30 ด้วย SoX ซึ่งเป็นโปรแกรมอัตถประโยชน์ในการประมวลผลทางด้านเสียงแบบโอเพ่นซอร์ส (Chris Bagwell et al.) ดังแสดงตัวอย่างในตารางที่ 4

ตารางที่ 4 แสดงตัวอย่างรายชื่อเพลงที่ตัดแบ่งเป็นส่วนวินาทีต่าง ๆ

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง		
			10 วินาที	20 วินาที	30 วินาที
1	mahori	Chongco	1 Kangkao Kin Kluay.wav	1 Kangkao Kin Kluay-20.wav	1 Kangkao Kin Kluay-30.wav
2	mahori	Chongco	2 Kag Ar Wang.wav	2 Kag Ar Wang-20.wav	2 Kag Ar Wang-30.wav
3	mahori	Chongco	3 Kag Nang.wav	3 Kag Nang-20.wav	3 Kag Nang-30.wav
...	...	...	...	...	...
75	string	03Thai Strings	9 Sarika-Kaew.wav	9 Sarika-Kaew-20.wav	9 Sarika-Kaew-30.wav

การเลือกค่าคุณลักษณะ (Feature Selection)

ในงานวิจัยนี้ได้เลือกใช้ค่าคุณลักษณะแบบ Timbrel Feature จำนวน 5 คุณลักษณะ คือ Linear Predictive Coefficients (LPC), Mel-frequency cepstral coefficients (MFCC), Spectral Centroid, Zero Crossings Rate และค่าคุณลักษณะพลังงานสเปกตรัม (Power Spectrum) ในทดสอบหาค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมที่ดีที่สุด

โดยได้ทำการทดลองเลือกค่าคุณลักษณะเพื่อที่จะหาค่าคุณลักษณะที่ดีที่สุดที่จะนำมาใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม ด้วยการหาค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิม โดยเลือกใช้ข้อมูลในการทดสอบคือเพลงไทยเดิม 3 หมวดหมู่ หมวดหมู่ละ 25 เพลง รวมทั้งสิ้นจำนวน 75 เพลง ที่ความยาวของแต่ละเพลงตั้งแต่ 0 ถึง 20 วินาที ที่ความกว้างของเฟรม 512 sample ที่กำหนดโดยการสุ่ม และในแต่และวิธีการได้ทำการปรับพารามิเตอร์ดังนี้

- k-NN กำหนดให้ใช้ค่า k เท่ากับจำนวนของหมวดหมู่คือ 3
- SVM กำหนดให้ใช้ kernel แบบ radius basis function
- ANN กำหนดให้จำนวน Hidden Node เท่ากับจำนวนของหมวดหมู่คือ 3

ซึ่งปรากฏผลการทดลองดังต่อนี้

1. ค่าคุณลักษณะแบบเดี่ยว

ทำการสกัดค่าคุณลักษณะแต่ละแบบแล้วนำไปทดลองจำแนกหมวดหมู่เพลงไทยเดิม โดยได้ผลการทดลองโดยเรียงลำดับตามค่าความถูกต้องในการจำแนกหมวดหมู่ดังตารางที่ 5

ตารางที่ 5 ค่าความถูกต้องในการจำแนกของค่าคุณลักษณะแต่ละแบบ

ลำดับที่	ค่าคุณลักษณะ	ค่าความถูกต้องในการจำแนกด้วยวิธีการต่าง ๆ (%)			
		k-NN	SVM	ANN	ค่าเฉลี่ย
1	Mel-frequency cepstral coefficients (MFCC)	80.45	79.69	79.63	**79.92**
2	Power Spectrum	70.30	59.49	41.36	**57.05**
3	Linear Predictive Coefficients (LPC)	61.10	49.54	53.41	**54.68**
4	Zero Crossings Rate	44.27	47.25	42.72	**44.75**
3	Spectral Centroid	41.36	49.09	41.61	**44.02**
		ค่าเฉลี่ยรวม			**56.08**

จากตารางที่ 5 จะเห็นได้ว่าค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ให้ค่าความถูกต้องในการจำแนกหมวดหมู่โดยเฉลี่ยดีที่สุดคือ 79.92% และค่าความถูกต้องในการจำแนกเฉลี่ยโดยรวมเมื่อใช้ค่าคุณลักษณะแบบเดี่ยวเท่ากับ 56.08%

2. ค่าคุณลักษณะแบบคู่

เมื่อได้ค่าคุณลักษณะที่ส่งผลให้ได้ค่าความถูกต้องในการจำแนกที่ดีเป็นลำดับที่ 1 แล้ว ซึ่งก็คือค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) จึงทำการทดสอบการใช้ค่าคุณลักษณะอื่น ๆ ร่วมกัน โดยนำเอาค่าคุณลักษณะลำดับที่ 2, 3, 4 และ 5 มาทำการจำแนกหมวดหมู่ร่วมกับค่าคุณลักษณะในลำดับที่ 1 และนำไปทดลองจำแนกหมวดหมู่ เพื่อหาค่าคุณลักษณะคู่ที่ได้ค่าความถูกต้องในการจำแนกที่ดีที่สุด โดยได้ผลการทดลองดังตารางที่ 6

ตารางที่ 6 แสดงค่าความถูกต้องในการจำแนกโดยการจับคู่ค่าคุณลักษณะ

	ค่าคุณลักษณะที่เลือกใช้					ค่าความถูกต้องในการจำแนก (%)			
กลุ่มที่	M	P	L	Z	S	k-NN	SVM	ANN	ค่าเฉลี่ย
1	✓	✓				82.39	86.16	68.26	**78.94**

ตารางที่ 6 (ต่อ)

	ค่าคุณลักษณะที่เลือกใช้					ค่าความถูกต้องในการจำแนก (%)			
กลุ่มที่	**M**	**P**	**L**	**Z**	**S**	**k-NN**	**SVM**	**ANN**	**ค่าเฉลี่ย**
2	✓		✓			79.68	75.71	79.82	**78.40**
3	✓			✓		80.36	66.28	79.83	**75.49**
4	✓				✓	80.41	68.26	79.44	**76.04**
						ค่าเฉลี่ยรวม			**77.22**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

จากตารางที่ 6 พบว่าเมื่อใช้ค่าคุณลักษณะในกลุ่มที่ 1 คือ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum จะให้ค่าความถูกต้องในการจำแนกเฉลี่ยมากที่สุด คือ 78.94% และค่าความถูกต้องในการจำแนกเฉลี่ยของการใช้ค่าคุณลักษณะแบบคู่เท่ากับ 77.22%

3. ค่าคุณลักษณะร่วมกัน 3 แบบ

นำค่าคุณลักษณะแบบคู่ที่ให้ค่าความถูกต้องที่มากที่สุดคือ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum มาใช้ร่วมกับค่าคุณลักษณะอื่น ๆ แล้วนำไปทดลองจำแนกหมวดหมู่ เพื่อหาค่าคุณลักษณะ 3 แบบ ที่ใช้ร่วมกันแล้วได้ค่าความถูกต้องในการจำแนกที่ดีที่สุด โดยได้ผลตามตารางที่ 7

ตารางที่ 7 แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ3 แบบ

	ค่าคุณลักษณะที่เลือกใช้				ค่าความถูกต้องในการจำแนก (%)			
กลุ่มที่	**M+P**	**L**	**Z**	**S**	**k-NN**	**SVM**	**ANN**	**ค่าเฉลี่ย**
1	✓	✓			80.80	86.16	66.62	**77.86**
2	✓		✓		81.96	85.49	69.62	**79.02**
3	✓			✓	82.20	85.73	63.43	**77.12**
					ค่าเฉลี่ยรวม			**78.00**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

จากตารางพบว่าเมื่อใช้ค่าคุณลักษณะ Mel-frequency cepstral coefficients (MFCC) และ Power Spectrum ร่วมกับ ค่าคุณลักษณะแบบ Zero Crossings Rate ให้ค่าความถูกต้องในการจำแนกโดยเฉลี่ยมากที่สุดคือ 79.02% และค่าความถูกต้องในการจำแนกโดยเฉลี่ยของการใช้ค่าคุณลักษณะร่วมกัน 3 แบบเท่ากับ 78.00%

4. ค่าคุณลักษณะร่วมกัน 4 แบบ

นำค่าคุณลักษณะทั้งสามแบบที่ให้ค่าความถูกต้องที่มากที่สุดที่ได้จากการทดลอง หาค่าคุณลักษณะร่วมกัน 3 แบบ คือ Mel-frequency cepstral coefficients (MFCC), Power Spectrum และ Zero Crossings Rate มาใช้ร่วมกับค่าคุณลักษณะในลำดับต่อไป แล้วนำไปทดลองจำแนกหมวดหมู่ได้ผลตามตารางที่ 8

ตารางที่ 8 แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ 4 แบบ

	ค่าคุณลักษณะที่เลือกใช้			**ค่าความถูกต้องในการจำแนก (%)**			
กลุ่มที่	**M+P+Z**	**L**	**S**	**k-NN**	**SVM**	**ANN**	**ค่าเฉลี่ย**
1	✓	✓		80.36	85.39	65.75	**77.17**
2	✓		✓	81.08	85.63	70.83	**79.18**
				ค่าเฉลี่ยรวม			**78.18**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

จากตารางพบว่าเมื่อใช้ค่าคุณลักษณะ Mel-frequency cepstral coefficients (MFCC), Power Spectrum และ Zero Crossings Rate ร่วมกับ ค่าคุณลักษณะแบบ Spectral Centroid ให้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 79.18% และได้ค่าความถูกต้องโดยเฉลี่ยของการใช้ค่าคุณลักษณะร่วมกัน 4 แบบเท่ากับ 78.18%

5. ใช้ค่าคุณลักษณะร่วมกัน 5 แบบ

นำค่าคุณลักษณะทั้งทุกแบบคือ Mel-frequency cepstral coefficients (MFCC), Power Spectrum, Linear Predictive Coefficients (LPC), Zero Crossings Rate และ Spectral Centroid มาใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมร่วมกัน ได้ผลตามตารางที่ 9

ตารางที่ 9 แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะทั้ง 5 แบบ

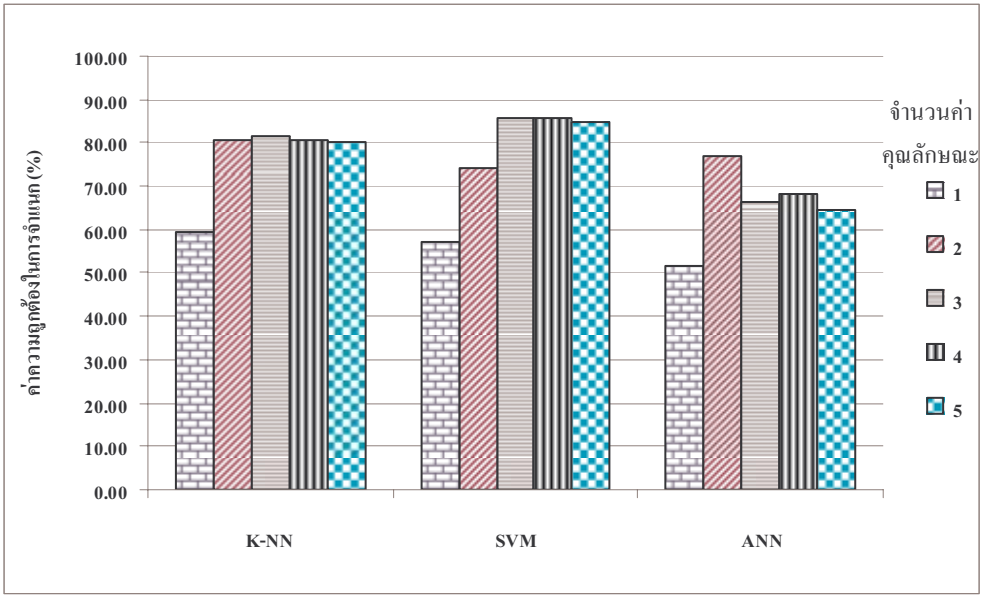
ค่าคุณลักษณะที่เลือกใช้	ค่าความถูกต้องในการจำแนก (%)			
	k-NN	**SVM**	**ANN**	**ค่าเฉลี่ย**
M+P+Z+S+L	80.41	84.95	64.59	**76.65**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

จากตารางที่ 9 พบว่าค่าเมื่อใช้ค่าคุณลักษณะทั้ง 5 แบบร่วมกันได้ค่าความถูกต้องในการจำแนกโดยเฉลี่ยเท่ากับ 76.65%

เปรียบเทียบค่าคุณลักษณะที่ใช้ในการจำแนกหมวดหมู่

เมื่อทำการทดลองใช้ค่าคุณลักษณะแบบต่าง ๆ ในการจำแนกหมวดหมู่ดังที่กล่าวมาแล้วนั้น ได้ผลค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ของทั้งสามวิธีการดังภาคผนวก ค และสามารถสร้างเป็นแผนภูมิเปรียบเทียบค่าความถูกต้องในการจำแนกโดยเฉลี่ยเมื่อใช้ค่าคุณลักษณะจำนวนต่าง ๆ ของแต่ละวิธีการได้ดังภาพที่ 19



ภาพที่ 19 เปรียบเทียบค่าความถูกต้องในการจำแนกโดยเฉลี่ยเมื่อใช้ค่าคุณลักษณะจำนวนต่าง ๆ

โดยสามารถแสดงค่าความถูกต้องในการจำแนกที่มากที่สุดของแต่และวิธีการได้ดังตารางที่ 10

ตารางที่ 10 เปรียบเทียบค่าความถูกต้องในการจำแนกที่มากที่สุดโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ของทั้งสามวิธีการ

วิธีการ	ค่าคุณลักษณะที่ใช้	ค่าความถูกต้องในการจำแนกที่มากที่สุด (%)
k-NN	M	80.45
	M+P	82.39
	M+P+S	82.20
	M+P+Z+S	81.08
	M+P+Z+S+L	80.41
SVM	M	79.69
	M+P	86.16
	M+P+L	86.16
	M+P+Z+S	85.63
	M+P+Z+S+L	84.95
ANN	M	79.63
	M+Z	79.83
	M+P+Z	69.62
	M+P+Z+S	70.83
	M+P+Z+S+L	64.59

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

จากตารางที่ 10 พบว่าวิธีการ k-NN นั้นได้ค่าความถูกต้องมากที่สุดเมื่อใช้ค่าคุณลักษณะแบบคู่ คือ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum โดยได้ค่าความถูกต้อง 82.39% ส่วนวิธีการ SVM นั้นได้ค่าความถูกต้องมากที่สุดที่ 86.16% เมื่อใช้ค่าคุณลักษณะร่วมกันสามแบบ คือ Mel-frequency cepstral coefficients (MFCC), Power Spectrum และ Linear Predictive Coefficients ส่วนวิธีการ ANN ได้ค่าความถูกต้องมากที่สุดที่ 79.83% เมื่อใช้ค่าคุณลักษณะแบบคู่คือ Mel-frequency cepstral coefficients (MFCC) และ Zero Crossings Rate

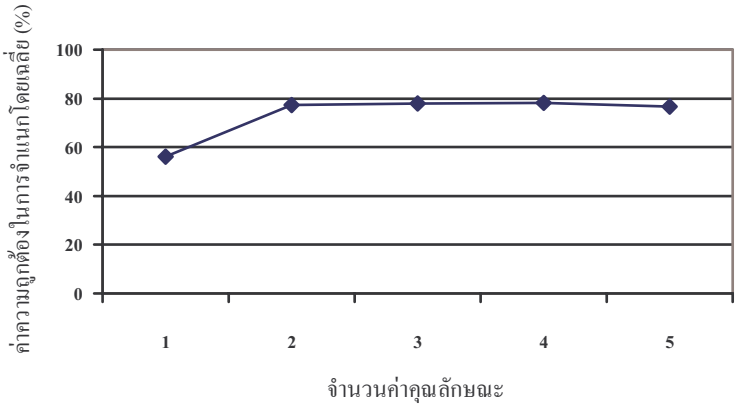
จะเห็นได้ว่าค่าความถูกต้องในการจำแนกด้วยค่าคุณลักษณะแบบต่าง ๆ ด้วยวิธีการ k-NN นั้นมีค่าที่ใกล้เคียงกันมากจึงยังไม่สามารถเลือกได้ว่าค่าคุณลักษณะแบบใดเหมาะสมที่สุด ส่วนวิธีการ SVM นั้นค่าความถูกต้องจะมีการเปลี่ยนแปลงไม่มากเมื่อใช้ค่าความถูกต้องสามแบบขึ้นไป ดังนั้นจึงน่าจะเป็นรูปแบบของค่าคุณลักษณะที่เหมาะสมสำหรับวิธีการจำแนกด้วย SVM และวิธีการ ANN นั้นค่าความถูกต้องในการจำแนกมีการเปลี่ยนแปลงลดต่ำลงเมื่อใช้ค่าคุณลักษณะสามแบบขึ้นไป ดังนั้นการใช้ค่าคุณลักษณะแบบจับคู่จึงเป็นรูปแบบของค่าคุณลักษณะที่เหมาะสมในการจำแนกเพลงไทยเดิมด้วยวิธีการ ANN

เมื่อพิจารณาที่แต่ละวิธีการแล้ว จะเห็นว่ามีการใช้รูปแบบค่าคุณลักษณะที่แตกต่างกัน แต่เนื่องจากในงานวิจัยนี้ต้องการศึกษาถึงวิธีการในการจำแนกว่าวิธีการใดจะสามารถจำแนกหมวดหมู่เพลงไทยเดิมได้ดีที่สุด รูปแบบของค่าคุณลักษณะที่นำมาใช้ในการทดลองจึงควรที่จะเป็นรูปแบบเดียวกัน ดังนั้นจึงต้องพิจารณาที่ค่าความถูกต้องในการจำแนกโดยเฉลี่ยของการใช้ค่าคุณลักษณะแต่ละรูปแบบ โดยสามารถสรุปเป็นตารางเปรียบเทียบค่าความถูกต้องโดยเฉลี่ยในการจำแนกหมวดหมู่โดยการใช้ค่าคุณลักษณะแบบต่าง ๆ ได้ดังตารางที่ 11 และสามารถสร้างเป็นแผนภูมิได้ดังภาพที่ 20

ตารางที่ 11 เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่โดยใช้ค่าคุณลักษณะแบบต่าง ๆ

	ค่าคุณลักษณะที่ให้ค่าความถูกต้องในการจำแนกมากที่สุด	**ค่าความถูกต้องในการจำแนกที่มากที่สุด (%)**	**ค่าความถูกต้องในการจำแนก โดยเฉลี่ย (%)**	**เวลาในการสกัดค่าคุณลักษณะโดยเฉลี่ย (วินาที)**
ค่าคุณลักษณะแบบเดี่ยว	**M**	**79.92**	**56.08**	**45**
ค่าคุณลักษณะแบบคู่	**M+P**	**78.94**	**77.22**	**90**
ค่าคุณลักษณะร่วมกัน 3 แบบ	**M+P+Z**	**79.02**	**78.00**	**780**
ค่าคุณลักษณะร่วมกัน 4 แบบ	**M+P+Z+S**	**79.18**	**78.18**	**798**
ค่าคุณลักษณะร่วมกัน 5 แบบ	**M+P+Z+S+L**	**76.65**	**76.65**	**797.4**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid



ภาพที่ 20 เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่โดยใช้ค่าคุณลักษณะแบบต่าง ๆ

พิจารณาภาพที่ 20 พบว่าเมื่อใช้ค่าคุณลักษณะแบบเดี่ยวจะให้ค่าความถูกต้องในการจำแนกเฉลี่ยที่ 56.08% และเพิ่มขึ้นเป็น 77.22% เมื่อใช้ค่าคุณลักษณะแบบคู่ จากนั้นเมื่อเพิ่มค่าคุณลักษณะเป็น 3 แบบ, 4 แบบ และ 5 แบบ กราฟค่าความถูกต้องในการจำแนกมีการเปลี่ยนแปลงไม่มาก และมีลักษณะเป็นเส้นตรง อีกทั้งเวลาที่ใช้ในการสกัดค่าคุณลักษณะยังเพิ่มขึ้นเป็นเท่าตัว โดยใช้เวลาเฉลี่ยในการสกัดค่าคุณลักษณะต่อหนึ่งเพลงที่ 45 วินาที เมื่อใช้สกัดค่าคุณลักษณะแบบเดี่ยว และ 90 วินาทีเมื่อใช้ค่าคุณลักษณะแบบคู่ แต่เมื่อเพิ่มค่าคุณลักษณะในการสกัดค่าเป็นสาม สี่ และห้าค่าคุณลักษณะจะใช้เวลาเฉลี่ยในการสกัดค่าที่ 780 วินาที 798 วินาที และ 797.4 วินาที ตามลำดับ ดังนั้นเมื่อพิจารณาจากเวลาที่ต้องใช้ในการสกัดค่าคุณลักษณะกับค่าความถูกต้องในการจำแนกที่เพิ่มขึ้นเพียงเล็กน้อยแล้วและไม่ได้เป็นการเพิ่มขึ้นอย่างมีนัยสำคัญ การใช้ค่าคุณลักษณะแบบคู่จึงมีความเหมาะสมในการนำมาใช้จำแนกหมวดหมู่เพลงไทยเดิมมากกว่า และเมื่อพิจารณาค่าคุณลักษณะแบบจับคู่จากตารางที่ 11 จะพบว่า การใช้ค่าคุณลักษณะ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum นั้นให้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 78.94%

ดังนั้นจึงสรุปได้ว่า การใช้ค่าคุณลักษณะแบบจับคู่ โดยใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum เป็นรูปแบบและค่าคุณลักษณะที่เหมาะสมในการนำไปใช้จำแนกหมวดหมู่เพลงไทยเดิม

ขนาดความกว้างของเฟรม (windowsize)

เมื่อได้คู่ของค่าคุณลักษณะคือ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum จากการคัดเลือกคุณลักษณะแล้ว ในขั้นตอนนี้จะทำการทดลองเพื่อหาขนาดความกว้างของเฟรม (windowsize) ในการสกัดค่าคุณลักษณะ เพื่อที่จะนำไปใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมต่อไป โดยทำการสกัดค่าคุณลักษณะคู่ที่ได้กล่าวมาข้างต้นด้วยขนาดความกว้างของเฟรมที่ 32, 64, 128, 256, 512, 1024 และ 2048 samples ตามลำดับ แล้วนำชุดข้อมูลที่ได้จากการสกัดค่า ไปทำการจำแนกหมวดหมู่ ซึ่งได้ผลตามตารางที่ 12

ตารางที่ 12 ค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum ที่ความกว้างของเฟรมขนาดต่าง ๆ

ขนาดความกว้างของเฟรม (samples)	ค่าความถูกต้องในการจำแนก (%)			
	k-NN	SVM	ANN	ค่าเฉลี่ย
32	62.86	69.05	62.84	**64.92**
64	67.23	77.10	68.63	**70.99**
128	73.70	82.19	69.23	**75.04**
256	78.36	85.37	70.79	**78.17**
512	82.39	86.16	68.26	**78.94**
1024	82.98	85.69	34.23	**67.63**
2048	75.63	78.72	33.66	**62.67**

จากตารางที่ 12 พบว่าค่าความถูกต้องในการจำแนกหมวดหมู่ค่าเฉลี่ยมากที่สุดเมื่อใช้ขนาดความกว้างของเฟรมที่ 512 samples และมีค่าลดลงเมื่อความกว้างของเฟรมมีขนาดมากขึ้น ดังนั้นขนาดความกว้างของเฟรมที่เหมาะสมในการนำไปใช้ในการสกัดค่าคุณลักษณะ เพื่อนำไปใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมคือ 512 samples

ความยาวของเพลงที่ใช้สกัดค่าคุณลักษณะ

เมื่อได้คู่ของค่าคุณลักษณะ และขนาดความกว้างของเฟรมที่จะนำไปใช้ในการจำแนกหมวดหมู่แล้ว ปัจจัยอีกอย่างที่ทำการศึกษาคือ ขนาดความยาวของเพลงที่จะนำไปในการสกัดค่าคุณลักษณะ จึงทำการทดลองหาค่าความถูกต้องในการจำแนก โดยใช้ข้อมูลเพลงที่ช่วงความยาวต่าง ๆ คือ 0 ถึง 10 วินาที 0 ถึง 20 วินาที และ 0 ถึง 30 วินาที และสกัดค่าคุณลักษณะ Mel-frequency

cepstral coefficients (MFCC) ร่วมกับ Power Spectrum ที่ขนาดความกว้างของเฟรมที่ 512 sample ซึ่งเป็นค่าที่จากการทดลองที่ได้กล่าวมาแล้วข้างต้น แล้วนำมาทดลองจำแนกหมวดหมู่โดยวิธีการและการปรับค่าพารามิเตอร์เช่นเดียวกับการเลือกค่าคุณลักษณะ ซึ่งได้ผลตามตารางที่ 13

ตารางที่ 13 ค่าความถูกต้องในการจำแนกหมวดหมู่ที่ช่วงความยาวต่าง ๆ ของเพลง

ช่วงความยาวของเพลง	ค่าความถูกต้องในการจำแนก (%)			
	k-NN	SVM	ANN	ค่าเฉลี่ย
0 ถึง 10 วินาที	82.59	86.36	41.20	**70.05**
0 ถึง 20 วินาที	82.39	86.16	68.26	**78.94**
0 ถึง 30 วินาที	81.61	86.97	62.19	**76.92**

จากตารางที่ 13 พบว่าช่วงความยาวของเพลงที่ให้ค่าความถูกต้องในการจำแนกหมวดหมู่เฉลี่ยที่มากที่สุดคือ 0 ถึง 20 วินาที

การจำแนกหมวดหมู่เพลงไทยเดิม

การทดสอบความถูกต้องของการจำแนกหมวดหมู่เพลงไทยเดิมในงานวิจัยนี้ ทำการทดลองโดยใช้เพลงไทยเดิมที่ได้ทำการรวบรวมหมวดหมู่และ 25 เพลง 3 หมวดหมู่รวม 75 เพลง และนำมาสกัดค่าคุณลักษณะ โดยใช้ค่าที่ดีที่สุดที่ได้จากการคัดเลือกค่าคุณลักษณะดังที่ได้กล่าวมาแล้วข้างต้น คือใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum ที่ขนาดความกว้างของเฟรมที่ 512 samples และที่ความยาวของเพลงตั้งแต่ 0 ถึง 20 วินาที แล้วนำมาจำแนกหมวดหมู่โดยแบ่งข้อมูลด้วยวิธีการ 10-fold cross-validation ซึ่งเป็นการแบ่งข้อมูลที่ได้จากการสกัดค่าคุณลักษณะที่ใช้ในการทดสอบ ออกมา 10 กลุ่ม แล้วนำข้อมูลชุดที่ 2-10 มาทำการเรียนรู้ (Train) จากนั้นนำชุดข้อมูลที่ 1 มาทดสอบ (Test) เพื่อหาค่าความถูกต้องในการจำแนก แล้วทำการสลับกลุ่มข้อมูลสำหรับเรียนรู้และทดสอบจนครบทั้ง 10 กลุ่ม จากนั้นจึงหาค่าเฉลี่ยของค่าความถูกต้องในการจำแนกออกมาเป็นเปอร์เซ็นต์

วิธีการจำแนกหมวดหมู่เพลงไทยเดิมในงานวิจัยนี้ได้แบ่งออกเป็น 3 วิธีการ คือ การจำแนกด้วย k-Nearest Neighbor (k-NN), การจำแนกด้วย Support Vector Machine (SVM) และการจำแนกด้วย Artificial Neural Network (ANN) ในทุกวิธีการนั้นจะมีการปรับเปลี่ยนค่าพารามิเตอร์เพื่อหาค่าความถูกต้องในการจำแนก ซึ่งได้ผลดังต่อไปนี้

1. ผลการจำแนกด้วยวิธีการ k-Nearest Neighbor (k-NN)

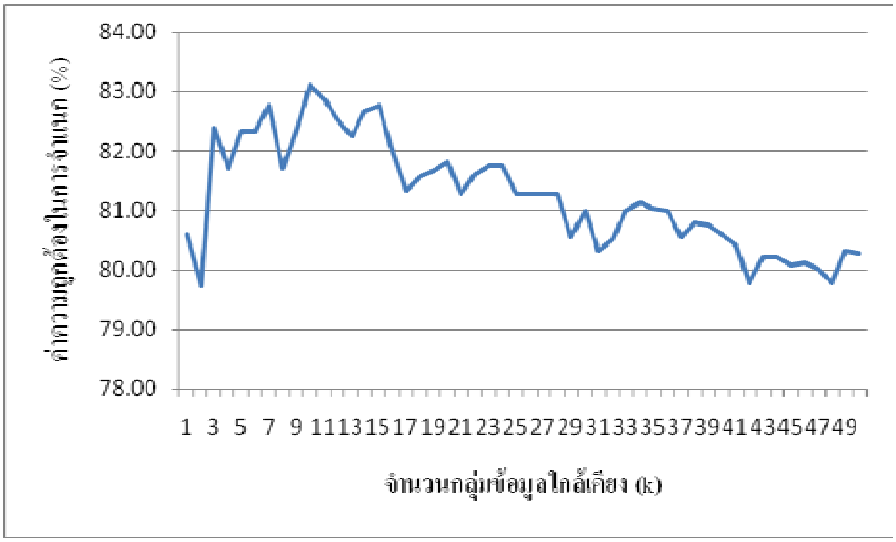
การจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN) นั้น ผู้วิจัยได้ทำการกำหนดค่า k หรือค่ากลุ่มข้อมูลใกล้เคียงของวิธีการจำแนกแบบ k-NN ตั้งแต่ 1-50 ซึ่งได้ผลการทดลองดังนี้ดังตารางที่ 14 และสามารถสรุปเป็นแผนภูมิดังภาพที่ 21

ตารางที่ 14 แสดงค่าความถูกต้องในการจำแนกหมวดหมู่ด้วยวิธีการ k-NN โดยพิจารณาที่จำนวนของกลุ่มข้อมูลใกล้เคียง (k)

กลุ่มข้อมูลใกล้เคียง (k)	ค่าความถูกต้องในการจำแนก (%)
1	80.60
2	79.73
3	82.39
4	81.71
5	82.34
6	82.34
7	82.78
8	81.71
9	82.39
10	83.12
11	82.87
12	82.54
13	82.24
14	82.68
15	82.78
16	81.95
17	81.33
18	81.57
19	81.66
20	81.81
21	81.28
22	81.62
23	81.76
24	81.76

ตารางที่ 14 (ต่อ)

กลุ่มข้อมูลใกล้เคียง (k)	ค่าความถูกต้องในการจำแนก (%)
25	81.28
26	81.28
27	81.28
28	81.28
29	80.55
30	80.99
31	80.31
32	80.50
33	80.99
34	81.13
35	81.04
36	80.99
37	80.55
38	80.79
39	80.75
40	80.60
41	80.41
42	79.78
43	80.21
44	80.21
45	80.07
46	80.12
47	80.02
48	79.78
49	80.31
50	80.26



ภาพที่ 21 แสดงค่าความถูกต้องในการจำแนกต่อจำนวนกลุ่มข้อมูลใกล้เคียง (k) ต่าง ๆ

จากการทดลองพบว่าเมื่อทำการปรับจำนวนกลุ่มข้อมูลใกล้เคียงหรือค่า k เท่ากับ 10 จะให้ค่าความถูกต้องในการจำแนกที่มากที่สุดคือ 83.12%

จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN) ที่กล่าวมาทั้งหมดสามารถสรุปได้ว่า กลุ่มข้อมูลใกล้เคียง (k) ที่ดีที่สุดคือ 10 และได้ค่าความถูกต้องในการจำแนกที่ดีที่สุดคือ 83.12%

2. ผลการจำแนกด้วยวิธีการ Support Vector Machine (SVM)

การทำงานของวิธีการ Support Vector Machine (SVM) นั้นอาศัยการสร้างเส้นแบ่ง (Plane) เพื่อแยกกลุ่มข้อมูลสองกลุ่มออกจากกัน แต่เนื่องจากหมวดหมู่ของเพลงไทยเดิมที่ต้องการจำแนกมีทั้งหมด 3 กลุ่ม งานวิจัยนี้ใช้ WLSVM ที่เป็นส่วนเสริมของ Weka ที่สามารถจำแนกข้อมูลด้วยวิธีการ Support Vector Machine แบบหลายกลุ่มได้ และทดลองใช้ Kernel แบบ Linear, Polynomial, Radial basis function และ Sigmoid ซึ่งได้ผลค่าความถูกต้องในการจำแนกและเวลาที่ใช้ในการทดสอบดังตารางที่ 15

ตารางที่ 15 แสดงผลการทดลองด้วยวิธีการ Support Vector Machine

Kernel	**เวลาในการเรียนรู้ (วินาที)**	**เวลาในการทดสอบ (วินาที)**	**ค่าความถูกต้องในการจำแนก (%)**
Linear	37.42	3.16	78.04

ตารางที่ 15 (ต่อ)

Kernel	เวลาในการเรียนรู้ (วินาที)	เวลาในการทดสอบ (วินาที)	ค่าความถูกต้องในการจำแนก (%)
Polynomial	3003.05	2.58	83.45
Radial basis function	7.57	4.63	86.16
Sigmoid	11.09	7.01	35.17

จากตารางที่ 15 พบว่าการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ Support Vector Machine (SVM) โดยใช้ Kernel แบบ Polynomial และ Radial basis function ได้ค่าความถูกต้องในการจำแนกใกล้เคียงกันคือ Kernel แบบ Polynomial ได้ค่าความถูกต้องที่ 83.45% ส่วน Kernel แบบ Radial basis function ได้ค่าความถูกต้องเท่ากับ 86.16% แต่เมื่อพิจารณาที่เวลาในการเรียนรู้ข้อมูล Radial basis function นั้นใช้เวลาน้อยกว่า

ดังนั้น Radial basis function จึงเป็น Kernel แบบที่เหมาะสมสำหรับการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ SVM โดยได้ค่าความถูกต้องในการจำแนกที่ 86.16% ซึ่งเกินกว่าเกณฑ์ที่ผู้วิจัยได้ตั้งไว้คือ 85%

3. ผลการจำแนกด้วยวิธีการ Artificial Neural Network (ANN)

งานวิจัยนี้ใช้ Artificial Neural Network (ANN) แบบ Back-Propagation โดยมีชั้น (Layer) ทั้งหมด 3 ชั้นคือ ชั้นอินพุต (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นเอาต์พุต (Output Layer) โดยได้กำหนดจำนวนโหนดในชั้นต่าง ๆ ดังนี้

ชั้นอินพุต กำหนดโหนดเท่ากับจำนวนของค่าคุณลักษณะที่สกัดได้ที่ใช้ซึ่งในงานวิจัยนี้ใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) จำนวน 13 ค่าร่วมกับ Power Spectrum จำนวน 256 ค่า ดังนั้นจำนวนโหนดของชั้นอินพุตจึงเท่ากับ 269

ชั้นซ่อน กำหนดจำนวนโหนดในชั้นซ่อนโดยปรับค่าต่าง ๆ ดังนี้

1. ค่าเฉลี่ยของผลรวมระหว่างจำนวนของค่าคุณลักณะและจำนวนของหมวดหมู่ คือ (269+3)/2 เท่ากับ 136
2. จำนวนของหมวดหมู่เพลงไทยเดิม คือ 3

ชั้นเอาต์พุต กำหนดจำนวนโหนดเท่ากับจำนวนหมวดหมู่ของเพลงไทยเดิมคือ 3 ซึ่งได้ผลการทดลองดังตารางที่ 16

ตารางที่ 16 ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ ANN ที่จำนวนโหนดในชั้นต่าง ๆ

จำนวนโหนด			ค่าความถูกต้องในการจำแนก (%)
ชั้นอินพุต	ชั้นซ่อน	ชั้นเอาต์พุต	
269	136	3	36.57
	3		68.26

จากตารางที่ 16 พบว่าเมื่อใช้จำนวนโหนดในชั้นซ่อนเท่ากับจำนวนของหมวดหมู่เพลงไทยเดิมคือ 3 นั้น จะส่งผลให้ค่าความถูกต้องในการจำแนกได้ผลดีกว่าการใช้จำนวนโหนดในชั้นซ่อนเท่ากับค่าเฉลี่ยของผลรวมระหว่าง จำนวนของค่าคุณลักณะและจำนวนของหมวดหมู่ แต่เนื่องจากค่าความถูกต้องในการจำแนกนั้นไม่ผ่านเกณฑ์ที่ตั้งไว้คือ 85% ดังนั้นผู้วิจัยจึงทำการทดลองปรับจำนวนโหนดในชั้นซ่อนอีกครั้งเพื่อให้ได้ค่าความถูกต้องในการจำแนกที่ดีขึ้น โดยเริ่มต้นที่ 3 ตามจำนวนหมวดหมู่ของเพลงไทยเดิมที่ใช้ในการทดลอง ได้ผลการทดลองดังตารางที่ 17

ตารางที่ 17 ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ ANN ที่ทำการปรับจำนวนโหนดในชั้นซ่อน

ชั้นซ่อน	ค่าความถูกต้องในการจำแนก (%)
3	68.26
4	66.33
5	59.75

จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ Artificial Neural Network (ANN) แบบ Back-Propagation สามารถสรุปได้ว่า เมื่อกำหนดจำนวนโหนดในชั้นซ่อนเท่ากับจำนวนหมวดหมู่เพลงไทยเดิมที่ใช้ในการทดลองคือ 3 โหนด จะให้ค่าความถูกต้องในการจำแนกที่ดีที่สุดเท่ากับ 68.26%

เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิม

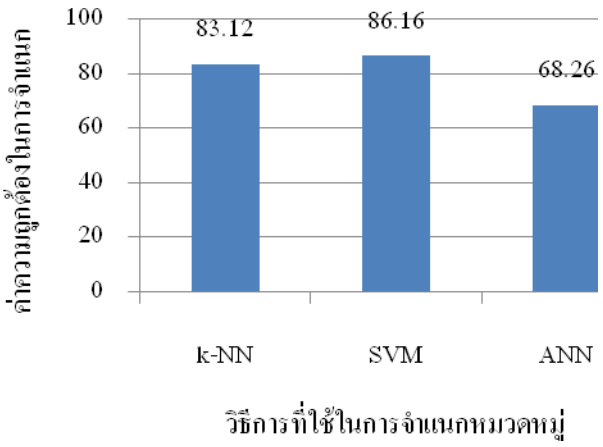
จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN),

Support Vector Machine (SVM) และ Artificial Neural Network (ANN) สามารถแสดงผลการทดลองได้ดังตารางที่ 18

ตารางที่ 18 เปรียบเทียบผลการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-NN, SVM และ ANN

วิธีการ	ค่าคุณลักษณะ	ความกว้างของเฟรม	ค่าพารามิเตอร์ที่ใช้แล้วให้ค่าความถูกต้องที่มากที่สุด	ค่าความถูกต้องในการจำแนก (%)
k-NN	Mel-frequency cepstral coefficients (MFCC) + Power Spectrum	512 Samples	จำนวนกลุ่มข้อมูลใกล้เคียง (k) เท่ากับ 10	83.12
SVM			Kernel แบบ Radial basis function	86.16
ANN			ชั้นอินพุต = 269 โหนด, ชั้นซ่อน = 3 โหนด, ชั้นเอาต์พุต = 3 โหนด	68.26

จากตารางที่ 18 พบว่าเมื่อจำแนกหมวดหมู่ด้วยวิธีการ Support Vector Machine (SVM) จะให้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 86.16% รองลงมาคือวิธีการ k-Nearest Neighbor (k-NN) ซึ่งให้ค่าความถูกต้องในการจำแนกที่ 83.12% และน้อยที่สุดคือวิธีการ Artificial Neural Network (ANN) ที่มีค่าความถูกต้องในการจำแนกที่ 68.26% ดังแสดงตามภาพที่ 22



ภาพที่ 22 เปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมของวิธีการ k-NN, SVM และ ANN

บทที่ 5
สรุปผลและอภิปรายผลการดำเนินงานวิจัย

ในบทนี้จะเป็นการสรุปผลที่ได้จากการดำเนินงานวิจัย การจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) นอกจากนี้จะได้นำเสนอความยาวของเพลงที่นำมาสกัดค่าคุณลักษณะที่เหมาะสม ขนาดความกว้างของเฟรม (Windowsize) ที่เหมาะสม และค่าคุณลักษณะที่ทำให้ได้ค่าความถูกต้องในการจำแนกที่ดีในการจำแนกหมวดหมู่เพลงไทยเดิม และกล่าวถึงปัญหาและแนวทางการแก้ไขรวมถึงข้อเสนอแนะในงานวิจัยนี้ โดยอธิบายดังต่อไปนี้

สรุปผลการดำเนินงานวิจัย

จากผลการดำเนินงานวิจัยในบทที่ 4 สามารถสรุปผลการดำเนินงานวิจัยได้ดังต่อไปนี้

การเลือกค่าคุณลักษณะ

1. ค่าคุณลักษณะที่ใช้ในการจำแนกหมวดหมู่เพลงไทยเดิม

จากการทดลองพบว่าเมื่อใช้ค่าคุณลักษณะแบบเดี่ยวจะให้ค่าความถูกต้องในการจำแนกเฉลี่ยที่ 56.08% และเพิ่มขึ้นเป็น 77.22% เมื่อใช้ค่าคุณลักษณะแบบคู่ และค่าความถูกต้องในการจำแนกมีการเปลี่ยนแปลงเพียงเล็กน้อย เมื่อเพิ่มค่าคุณลักษณะที่ใช้ในการจำแนกเป็น 3 แบบ, 4 แบบ และ 5 แบบ ดังนั้นการใช้ค่าคุณลักษณะแบบคู่จึงเพียงพอต่อการจำแนกหมวดหมู่เพลงไทยเดิม โดยค่าคุณลักษณะคู่ที่ให้ความถูกต้องในการจำแนกหมวดหมู่มากที่สุดคือ ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum โดยเมื่อนำไปใช้จำแนกหมวดหมู่เพลงไทยเดิมมีค่าความถูกต้องเฉลี่ยในการจำแนกเท่ากับ 78.94%

2. ขนาดความกว้างของเฟรม (windowsize)

จากการทดลองนำเพลงไทยเดิมมาค่าคุณลักษณะแบบคู่คือ ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum มาแบ่งออกเป็นเฟรมที่ขนาดความกว้างต่าง ๆ เพื่อทำการสกัดค่าคุณลักษณะ พบว่าเมื่อใช้ขนาดความกว้างของเฟรมที่ 512 samples ค่าความถูกต้องในการจำแนกหมวดหมู่โดยเฉลี่ยจะมีค่ามากที่สุดคือ 78.94% และเมื่อขนาดความกว้างมากกว่า 512 sample ค่าความถูกต้องในการจำแนกหมวดหมู่กลับมีค่าเฉลี่ยที่ลดลง

3. ความยาวของเพลงที่ใช้สกัดค่าคุณลักษณะ

เมื่อทดลองนำค่าคุณลักษณะแบบคู่คือ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum มาสกัดค่าคุณลักษณะโดยใช้ความกว้างของเฟรมที่ 512 samples ที่ช่วงความยาวต่าง ๆ ของเพลง พบว่าช่วงความยาวของเพลงที่ค่าความถูกต้องในการจำแนกหมวดหมู่เฉลี่ยมากที่สุดคือ 0 ถึง 20 วินาที คือเท่ากับ 78.94%

วิธีการในการจำแนกหมวดหมู่

การจำแนกหมวดหมู่ของทั้ง 3 วิธีการในงานวิจัยนี้คือ k-Nearest Neighbor (k-NN), Support Vector Machine (SVMs) และ Artificial Neural Network (ANN) นั้น จะใช้ค่าคุณลักษณะที่ได้ทำการคัดเลือกมาแล้วจากการเลือกค่าคุณลักษณะคือ ใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum มาทำการสกัดค่าคุณลักษณะโดยกำหนดความกว้างของเฟรมที่ 512 samples ซึ่งสกัดจากช่วงความยาวของเพลงที่ 0 ถึง 20 วินาทีของเพลงไทยเดิมหมวดหมู่ละ 25 เพลง รวมทั้งสิ้น 75 เพลง และใช้วิธีการ 10-fold cross-validation ในการแบ่งข้อมูลเพื่อทำการเรียนรู้และทดสอบ และแต่ละวิธีการจะทำการปรับพารามิเตอร์เพื่อให้ได้ค่าคุณลักษระที่ดีที่สุด ซึ่งได้ผลการจำแนกหมวดหมู่เพลงไทยเดิมดังต่อไปนี้

1. การจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ k-Nearest Neighbor (k-NN)

จากการทดลองปรับจำนวนของกลุ่มข้อมูลใกล้เคียงของวิธีการ k-NN ตั้งแต่ 1 ถึง 50 เพื่อทำการจำแนกหมวดหมู่เพลงไทยเดิมทั้ง 3 กลุ่ม พบว่าเมื่อทำการปรับจำนวนกลุ่มข้อมูลใกล้เคียงเท่ากับ 10 ค่าความถูกต้องในการจำแนกจะมากที่สุดคือ 83.12%

2. วิธีการ Support Vector Machine (SVMs)

จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ Support Vector Machine (SVM) โดยใช้ Kernel ทั้ง 4 แบบคือ Linear, Polynomial, Radial basis function และ Sigmoid พบว่า Kernel แบบ Polynomial และ Radial basis function ได้ค่าความถูกต้องในการจำแนกใกล้เคียงกันคือ Kernel แบบ Polynomial ได้ค่าความถูกต้องที่ 83.45% ส่วน Kernel แบบ Radial basis function ได้ค่าความถูกต้องเท่ากับ 86.16% แต่เมื่อพิจารณาที่เวลาในการเรียนรู้ข้อมูล Radial basis function นั้นใช้เวลาน้อยกว่า จึงเป็น Kernel ที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ SVM โดยได้ค่าความถูกต้องในการจำแนกมากที่สุดเท่ากับ 86.16%

3. วิธีการ Artificial Neural Network (ANN)

จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมด้วยวิธีการ Artificial Neural Network (ANN) แบบ Back-Propagation โดยกำหนดจำนวนโหนดในชั้นซ่อนจำนวนต่าง ๆ พบว่าจำนวนโหนดในชั้นซ่อนที่ได้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 3 โหนด โดยค่าความ

ถูกต้องในการจำแนกเท่ากับ 68.26%

การเปรียบเทียบค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิม

จากการทดลองจำแนกหมวดหมู่เพลงไทยเดิมทั้ง 3 วิธีการ พบว่า วิธีการ Support Vector Machine (SVM) เป็นวิธีการที่ได้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 86.16% รองลงมาคือวิธีการ k-Nearest Neighbor (k-NN) ได้ค่าความถูกต้องในการจำแนกที่ 83.12% และน้อยที่สุดคือวิธีการ Artificial Neural Network (ANN) โดยมีค่าความถูกต้องในการจำแนกที่ 68.26%

วิเคราะห์และอภิปรายผลการทดลอง

จากการผลทดลองการจำแนกหมวดหมู่เพลงไทยเดิม โดยใช้ค่าคุณลักษณะแบบ Timbrel ร่วมกับค่าพลังงานสเปกตรัม (Power Spectrum) ด้วยวิธีการ k-Nearest Neighbor (k-NN), Support Vector Machine (SVM) และ Artificial Neural Network (ANN) สามารถวิเคราะห์และอภิปรายผลการทดลองได้ดังต่อไปนี้

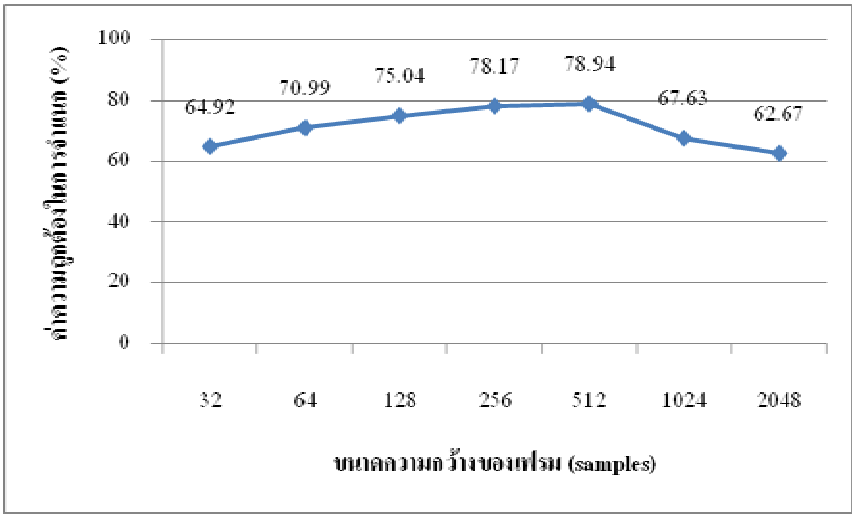
1. ค่าคุณลักษณะที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิม

จากผลการทดลองพบว่าเมื่อใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ Power Spectrum ซึ่งเป็นการใช้ค่าคุณลักษณะแบบคู่ ซึ่งได้ค่าความถูกต้องในการจำแนกมากที่สุด คือ 78.94% ซึ่งสอดคล้องกับสมมุติฐานที่ตั้งไว้ว่าค่าคุณลักณษะแบบ Mel-frequency cepstral coefficients (MFCC) เป็นค่าคุณลักษณะให้ค่าความถูกต้องในการจำแนกที่ดี ทั้งนี้เนื่องมาจาก ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) สามารถเก็บรายละเอียดของสัญญาณเสียงในช่วงความถี่ต่ำได้มากกว่าค่าคุณลักษณะแบบอื่น ๆ เพราะ MFCC ต้องปรับสเกลของสเปกตรัมให้อยู่บนสเกลที่เหมาะสำหรับการรับฟังของมนุษย์หรือที่เรียกว่า Mel Scale จึงทำให้สามารถนำมาเป็นข้อมูลในการสร้างตัวแบบ (model) เพื่อใช้ใช้ในการจำแนกหมวดหมู่เพลงไทยเดิมได้ดี ซึ่งสอดคล้องกับงานวิจัยเรื่อง ระบบรู้จำเสียงพูดภาษาไทยต่อเนื่องแบบเฉพาะบุคคลสำหรับการเข้าถึงอีเมล์ (ฐนียา สัตยพานิช และ อัศนีย์ ก่อตระกูล 2546) ที่กล่าวว่าลักษณะเด่น MFCC สามารถแทนลักษณะของหน่วยเสียงของภาษาได้ เช่น สระและพยัญชนะ โดยมีแนวคิดจากกระบวนการเกิดเสียงพูดของมนุษย์

2. ขนาดความกว้างของเฟรม (windowsize) และช่วงความยาวของเพลงที่ใช้สกัดค่าคุณลักษณะ

จากผลการทดลองพบว่าขนาดความกว้างของเฟรมและช่วงความยาวของเพลงที่นำไปใช้ในการสกัดค่าคุณลักษณะที่ดีที่สุดคือ ความกว้าง 512 samples ได้ค่าความถูกต้องใน

การจำแนกโดยเฉลี่ยที่ 78.94% เพราะเมื่อเพิ่มขนาดความกว้างที่ 1024 samples และ 2048 samples แล้วค่าความถูกต้องในการจำแนกมีค่าลดลงคือ 67.63% และ 62.67 % ตามลำดับ ดังภาพที่ 22



ภาพที่ 22 ค่าความถูกต้องในการจำแนกโดยเฉลี่ยที่ความกว้างของเฟรมขนาดต่าง ๆ

จากภาพที่ 22 จะเห็นได้ว่าค่าความถูกต้องในการจำแนกมีค่ามากที่สุดที่ 78.94% เมื่อใช้ขนาดความกว้างของเฟรมเท่ากับ 512 samples และมีค่าลดลงเมื่อใช้ขนาดความกว้างของเฟรมที่ 1024 samples และ 2048 samples

และเมื่อพิจารณาที่ช่วงความยาวของเพลงที่นำมาใช้ในการสกัดค่าคุณลักษณะ ช่วงความยาวที่ 0 ถึง 20 วินาทีเป็นช่วงเวลาที่ได้ค่าความถูกต้องในการจำแนกมากที่สุด เนื่องมาจากขนาดความกว้างและช่วงความยาวของเพลงมีผลต่อจำนวนข้อมูลที่นำไปใช้ในการจำแนกหมวดหมู่ กล่าวคือขั้นตอนของการเตรียมข้อมูลนั้นผู้วิจัยได้ทำการแปลงรูปแบบของเสียงจาก Compact Disk ให้เป็นรูปแบบของแฟ้มข้อมูลแบบ wav โดยกำหนด sampling rate ซึ่งหมายถึงค่าการสุ่มหน้าคลื่นสัญญาณเพื่อแทนด้วยสัญญาณดิจิตอลในการแปลงเสียงต่อหนึ่งวินาที ที่ 44 kHz ซึ่งเป็นค่ามาตรฐานในการบันทึกเสียงของ Compact Disk (Ballou 2002 : 1058) และใช้ช่วงความยาวของเพลงที่ 0 ถึง 20 วินาที ดังนั้น เมื่อแบ่งข้อมูลตามขนาดความกว้างที่ 512 samples จะทำให้ได้ข้อมูลที่ใช้ในการนำไปจำแนกหมวดหมู่ 1,723 เฟรมต่อหนึ่งเพลง ซึ่งมีจำนวนข้อมูลมากกว่าการแบ่งตามขนาดความกว้างที่ 1024 samples และ 2048 samples ที่มีจำนวนเฟรมเพียง 862 และ431 เฟรมต่อหนึ่งเพลง ดังตารางที่ 19

ตารางที่ 19 จำนวนของเฟรมเมื่อแบ่งตามขนาดความกว้างของเฟรมขนาดต่าง ๆ ที่ช่วงความยาวของเพลงต่าง ๆ ต่อหนึ่งเพลง

ช่วงความยาวของเพลง	จำนวนของเฟรมตามขนาดความกว้างต่าง ๆ ต่อหนึ่งเพลง						
	32 samples	**64 samples**	**128 samples**	**256 samples**	**512 samples**	**1024 samples**	**2048 samples**
0 ถึง 10 วินาที	13,782	6,891	3,446	1,723	862	431	216
0 ถึง 20 วินาที	27,563	13,782	6,891	3,446	1,723	862	431
0 ถึง 30 วินาที	41,344	20,672	10,336	5,168	2,584	1,292	646

ดังนั้นจำนวนเฟรมของข้อมูลเพลงทั้ง 3 หมวดหมู่ (75 เพลง) ที่นำไปใช้ในการจำแนกเมื่อแบ่งตามความกว้างที่ 512 samples คือ 1723 x 75 เท่ากับ 129,225 เฟรม ส่วนการแบ่งเฟรมที่ความกว้าง 1024 samples และ 2048 samples มีจำนวนเฟรมทั้งหมดเพียง 61,950 เฟรม และ 32,325 เฟรมตามลำดับ จะเห็นได้ว่าเมื่อขนาดของเฟรมมีความกว้างมากขึ้นจะมีผลทำให้จำนวนเฟรมที่จะนำไปสกัดค่าคุณลักษณะลดลง เมื่อนำไปเป็นข้อมูลในการจำแนกหมวดหมู่ทำให้ตัวแบบ (model) ที่สร้างจากขั้นตอนการเรียนรู้ไม่ดีเท่าที่ควร จึงทำให้ค่าความถูกต้องในการจำแนกมีค่าลดลง ซึ่งสอดคล้องกับงานวิจัยเรื่อง การใช้เทคนิคดาต้าไมน์นิงเพื่อพัฒนาคุณภาพการศึกษาคณะวิศวกรรมศาสตร์ (กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และธนาวินท์ รักธรรมานนท์ 2544 : 134-142) ที่กล่าวไว้ว่า จำนวนข้อมูลในบางสาขาวิชามีปริมาณค่อนข้างน้อยทำให้โมเดลที่ได้ไม่แม่นยำเท่าที่ควร หากต้องการกำจัดความผิดพลาดที่เกิดจากปริมาณข้อมูลน้อยเกินไปจำเป็นต้องใช้ข้อมูลอย่างน้อยพันคนในแต่ละสาขาวิชาที่ต้องการทำนาย และเมื่อพิจารณาที่ช่วงความยาวที่มากกว่า คือ 0 ถึง 30 วินาที นั้นมีจำนวนของเฟรมที่มากกว่า แต่ค่าความถูกต้องในการจำแนกกลับน้อยกว่าเนื่องมาจากจำนวนข้อมูลมีมาก ซึ่งเป็นไปได้ว่ามีข้อมูลที่ไม่สมบูรณ์ ไม่สามารถระบุได้ว่าอยู่ในหมวดหมู่ใดอยู่ด้วย หรือเป็นข้อมูลส่วนที่ซ้ำกันมากเกินไป จึงทำให้การสร้างตัวแบบไม่แม่นยำและส่งผลให้ค่าความถูกต้องในการจำแนกมีค่าไม่ดีเท่าที่ควร

จึงสรุปได้ว่าขนาดความกว้างของเฟรมที่นำไปใช้ในการสกัดค่าคุณลักษณะที่เหมาะสมในการนำไปจำแนกหมวดหมู่เพลงไทยเดิมคือที่ขนาดความกว้าง 512 samples ที่ช่วงความยาวของเพลงที่ 0 ถึง 20 วินาที

3. วิธีการในการจำแนกหมวดหมู่เพลงไทยเดิม

จากผลการทดลองจำแนกหมวดหมู่เพลงไทยเดิมพบว่าวิธีการ Support

Vector Machine (SVM) เป็นวิธีการที่ได้ค่าความถูกต้องในการจำแนกมากที่สุดคือ 86.16% รองลงมาคือวิธีการ k-Nearest Neighbor (k-NN) ได้ค่าความถูกต้องในการจำแนกที่ 83.12% แต่เนื่องจากค่าความถูกต้องในการจำแนกของทั้งสองวิธีการมีความใกล้เคียงกันมาก จึงต้องนำปัจจัยอื่น ๆ มาพิจารณาร่วมด้วยโดยเลือกพิจารณาที่เวลาที่ใช้ในการเรียนรู้และทดสอบข้อมูลดังแสดงตามตารางที่ 20

ตารางที่ 20 แสดงเวลาที่ใช้ในการเรียนรู้และทดสอบข้อมูลของวิธีการจำแนกหมวดหมู่แบบ SVM และ k-NN ที่ได้ค่าความถูกต้องในการจำแนกที่มากที่สุด

วิธีการ	เวลาในการเรียนรู้	เวลาในการทดสอบ	เวลาที่ใช้โดยเฉลี่ย
Support Vector Machine	7.57 วินาที	4.63 วินาที	6.10 วินาที
k-Nearest Neighbor	0.02 วินาที	22.32 วินาที	11.17 วินาที

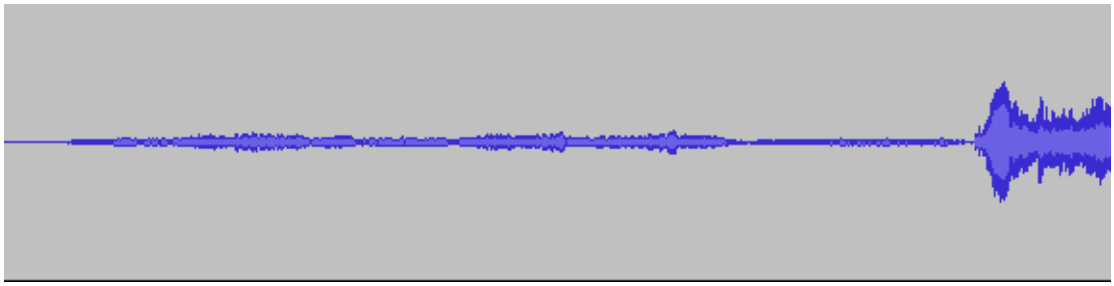
จากตารางที่ 20 พบว่าวิธีการ Support Vector Machine (SVM) ใช้เวลาโดยเฉลี่ยในการจำแนกหมวดหมู่น้อยกว่าวิธีการ k-Nearest Neighbor (k-NN) เนื่องมาจากวิธีการจำแนกข้อมูลด้วย k-NN ไม่มีการสร้างตัวแบบเพื่อทำการจำแนกและจะทำการคำนวนระยะห่างของข้อมูลเพื่อจำแนกกลุ่มเมื่อมีข้อมูลเข้ามาใหม่ทุกครั้ง (Witten and Frank 2005 : 78 – 80) จึงทำให้ใช้เวลาในการทดสอบข้อมูลมาก อีกทั้งหากนำวิธีการ k-NN ไปใช้ในการจำแนกหมวดหมู่บทเพลงที่มีจำนวนเพลงหรือจำนวนข้อมูลที่มีปริมาณมาก จะใช้เวลาในการทำงานนานจนเกินความจำเป็น จึงไม่เหมาะสมที่จะนำมาเป็นวิธีการในการจำแนกหมวดหมู่เพลงไทยเดิม

ส่วนวิธีการ Artificial Neural Network (ANN) สามารถจำแนกหมวดหมู่เพลงไทยเดิมได้ค่าความถูกต้องเพียง 68.26% เนื่องมาจาก ค่าคุณลักษณะที่นำมาใช้ในการจำแนกกล่าวคือนอกจากจะเป็นค่าคุณลักษณะแบบคู่แล้ว ค่าคุณลักษณะแต่ละแบบจะมีจำนวนของค่าคุณลักษณะที่แตกต่างออกไป เช่นในงานวิจัยนี้ใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) จำนวน 13 ค่า ร่วมกับค่า Power Spectrum จำนวน 256 ค่า โดยได้นำจำนวนของค่าคุณลักษณะที่ได้มาใช้เป็นจำนวนโหนดในชั้นอินพุต ซึ่งข้อมูลอินพุตบางส่วนอาจไม่สามารถจำแนกหมวดหมู่ได้เช่นข้อมูลส่วนที่เป็นเสียงเงียบที่มีค่าเป็น 0 ดังนั้นหากจะนำวิธีการ ANN มาใช้ในการจำแนกหมวดหมู่บทเพลงต่าง ๆ จึงจำเป็นที่จะต้องทำการคัดเลือกข้อมูลอินพุตเสียก่อนเช่นไม่นำข้อมูลส่วนที่เป็นเสียงเงียบเข้ามารวมด้วย ซึ่งสอดคล้องกับ งานวิจัยเรื่อง Prediction of maximum scour depth at spur dikes with adaptive neural networks. (Wu and Lim

1993: 61-66) ที่กล่าวไว้ว่า การเลือกข้อมูลจะต้องเลือกเฉพาะตัวแปรอิสระที่มีความสามารถทำนายผลเพื่อให้ได้ ผลลัพธ์หรือตัวแปรตาม โดยตัวแปรอิสระที่เลือกมานั้นจะต้องไม่มีความสัมพันธ์กัน แต่อย่างไรก็ดี ก็ขึ้นอยู่กับรูปแบบของโครงข่ายที่ใช้และเพื่อที่จะลดจำนวนตัวอย่างที่ใช้ในการสอนและเวลาที่ใช้ในการเรียนรู้ ก็ควรจะต้องมีการคัดเลือกข้อมูลป้อนเข้าให้เหมาะสม เพราะการคัดเลือกข้อมูลเป็นปัจจัยที่สำคัญในการสร้างแบบจำลอง

ซึ่งแตกต่างจากวิธีการ Support Vector Machine (SVM) ที่ใช้การสร้างเส้นแบ่ง (Plan) เพื่อแบ่งข้อมูลสองกลุ่มออกจากกัน อีกทั้งยังมี Kernel มาช่วยคำนวนการจัดวางค่าคุณลักษณะใหม่ หากไม่สามารถแบ่งกลุ่มข้อมูลออกจากกันได้ ดังนั้นจึงสามารถจำแนกข้อมูลเกี่ยวกับเสียงรวมถึงข้อมูลที่มีปริมาณมาก ๆ ได้ดี และยังไม่ต้องคำนึงถึงจำนวนข้อมูลนำเข้าเหมือนวิธีการ ANN จึงทำให้ได้ค่าความถูกต้องในการจำแนกหมวดหมู่เพลงไทยเดิมที่ดีกว่าทั้งสองวิธีการ ซึ่งสอดคล้องกับสมมุติฐานที่ผู้วิจัยได้ตั้งไว้ว่าวิธีการ Support Vector Machine (SVM) เป็นวิธีการที่ได้ผลดีที่สุดในการจำแนกหมวดหมู่เพลงไทยเดิม รวมถึงยังสอดคล้องกับงานวิจัยที่เกี่ยวกับการจำแนกข้อมูลประเภทเสียงต่าง ๆ เช่น งานวิจัยเรื่อง Classification of acoustic events using SVM-based clustering schemes (Temko and Nadeu 2005 : 682-694) ที่ทำการจำแนกเสียงต่าง ๆ จากสภาพแวดล้อมเช่น เสียงเคาะแป้นคีย์บอร์ด เสียงฉีกกระดาษ หรือเสียงปิดประตู เป็นต้น โดยทำการเปรียบเทียบการจำแนกเสียงระหว่างวิธีการ Support Vector Machine (SVM) และ Gaussian Mixture Models (GMM) โดย SVM ให้ค่าความถูกต้องในการจำแนกที่มากกว่า

ดังนั้น Support Vector Machine (SVM) จึงเป็นวิธีการที่เหมาะสมในการจำแนกหมวดหมู่เพลงไทยเดิมโดยได้ค่าความถูกต้องในการจำแนกมากที่สุดเท่ากับ 86.16% ซึ่งเกินเกณฑ์ 85% ที่ผู้วิจัยได้ตั้งไว้ ซึ่งข้อมูลส่วนที่เกิดการผิดพลาดหรือไม่สามารถจำแนกได้นั้น อาจเกิดจากการเลือกใช้ข้อมูลเสียงเพลงในช่วง 0 ถึง 20 วินาที ซึ่งเพลงส่วนใหญ่ จะทำการเว้นเสียงหรือมีส่วนที่เงียบในส่วนต้นของเพลง ดังภาพที่ 23



ภาพที่ 23 ส่วนที่เป็นเสียงเงียบในเพลงที่ใช้ในการจำแนกข้อมูล

ดังนั้นส่วนที่เป็นเสียงเงียบของเพลงจึงมีผลให้ข้อมูลค่าคุณลักษณะที่สกัดได้จากส่วนนี้เป็นข้อมูลที่ไม่สามารถจำแนกได้หรือไม่มีข้อมูล จึงส่งผลให้ค่าความถูกต้องในการจำแนกในส่วนนั้นมีค่าที่ซ้ำกันจำนวนมากทำให้การสร้างตัวแบบในการจำแนกไม่ดีพอ จึงส่งผลให้เกิดการผิดพลาดได้

สรุป

เพลงไทยเดิมมีจังหวะและท่วงทำนองที่คล้ายคลึงกัน การประพันธ์เพลงหนึ่ง ๆ อาจนำมาใช้ในการบรรเลงด้วยประเภทวงที่แตกต่างกันได้ ทำให้ผู้ที่ทำการจัดเก็บข้อมูลไม่ทราบแน่ชัดว่าเพลงนั้นอยู่ในหมวดหมู่ใด อาจทำให้การจัดเก็บหมวดหมู่มีความผิดพลาดได้ ดังนั้นการจำแนกหมวดหมู่เพลงไทยเดิมในงานวิจัยชิ้นนี้จึงเลือกใช้ค่าคุณลักษณะแบบ Timbrel ซึ่งเป็นค่าคุณลักษณะที่ไม่ได้สกัดจากท่วงทำนองหรือจังหวะของเพลง แต่สกัดจากข้อมูลอันเป็นเอกลักษณ์ของบทเพลงนั้น ๆ โดยทำการทดลองเลือกใช้ค่าคุณลักษณะ Timbrel แบบต่าง ๆ คือ Linear Predictive Coefficients (LPC), Mel-frequency cepstral coefficients (MFCC), Spectral Centroid, Zero Crossings Rate และค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum) ซึ่งจากการทดลองพบว่าเมื่อใช้ค่าคุณลักษณะร่วมกันตั้งแต่ 2 แบบขึ้นไป โดยใช้ความกว้างของเฟรม (windowsize) ที่ 512 samples และที่ช่วงความยาวของเพลงที่ 0 ถึง 20 วินาที ในการจำแนกหมวดหมู่จะให้ค่าความถูกต้องที่ใกล้เคียงกัน ดังนั้นการใช้ค่าคุณลักษณะร่วมกัน 2 แบบคือ ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ ค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum) นอกจากจะเป็นการประหยัดเวลาที่ใช้ในการสกัดค่าคุณลักษณะแล้วข้อมูลที่ได้ยังเพียงพอต่อการสกัดค่าคุณลักษณะเพื่อจำแนกหมวดหมู่เพลงไทยเดิม ซึ่งต่างจากการจำแนกหมวดหมู่เพลงไทยอื่น ๆ เช่นงานวิจัยเรื่อง การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์ (อานนท์ นามสนิท 2549) ที่ต้องทำการสกัดค่าคุณลักษณะหลาย ๆ แบบเพื่อนำมาใช้ในการจำแนกกลุ่มของเพลงไทย และในส่วนของวิธีการจำแนกนั้นวิธีการ Support Vector Machine (SVM) เป็นวิธีการที่เหมาะสมที่สุดในการจำแนกหมดวหมู่เพลงไทยเดิม เพราะมีการสร้างตัวต้นแบบเพื่อใช้ในการจำแนก ทำให้ใช้เวลาในการจำแนกหมวดหมู่น้อยกว่า วิธีการ k-Nearest Neighbor (k-NN) อีกทั้งยังไม่ต้องทำการคัดเลือกค่าคุณลักษณะที่จะทำการป้อนเข้าเช่นวิธีการ Artificial Neural Network (ANN) อีกด้วย

จากการทดลองจึงสามารถสรุปได้ว่าการใช้วิธีการ Support Vector Machine (SVM) เป็นวิธีการในการจำแนกหมวดหมู่เพลงไทยเดิม โดยใช้ค่าคุณลักษณะแบบ Mel-frequency cepstral coefficients (MFCC) ร่วมกับ ค่าคุณลักษณะแบบพลังงานสเปกตรัม (Power Spectrum) ที่ความ

กว้างของเฟรม (windowsize) ที่ 512 samples และช่วงความยาวของเพลงที่ 0 ถึง 20 วินาทีนั้น เป็นวิธีการที่เหมาะสมที่สุด โดยได้ค่าความถูกต้องในการจำแนกที่มากที่สุดที่ 86.16% ซึ่งในการทดลองนั้น หากมีการคัดเลือกข้อมูลส่วนที่เป็นเสียงเงียบออกไปอาจทำให้ได้ค่าความถูกต้องในการจำแนกที่มากขึ้น

ข้อเสนอแนะ

1. นำวิธีการที่ได้จากงานวิจัยไปจำแนกหมวดหมู่ของเพลงไทยอื่น ๆ เช่น เพลงพื้นบ้าน เพลงฉ่อย เพลงอีแซว ฯลฯ

2. พัฒนาวิธีการเพื่อให้ได้ค่าความถูกต้องในการจำแนกมากยิ่งขึ้น โดยนำค่าคุณลักษณะอื่น ๆ มาใช้

3. นำวิธีการที่ได้จากงานวิจัยไปประยุกต์ใช้ในการพัฒนาแอพพลิเคชั่น เพื่อเก็บรวบรวมและช่วยตัดสินใจในการจัดหมวดหมู่บทเพลงไทยเดิม

4. ในขั้นตอนการเตรียมข้อมูลเสียงเพลง ควรทำการคัดเลือกข้อมูลโดยตัดข้อมูลเสียงส่วนที่เงียบออกหรือเลือกแต่เฉพาะส่วนที่มีข้อมูลเสียงมาการสกัดค่าคุณลักษณะ ซึ่งอาจจะทำให้ได้ค่าความถูกต้องในการจำแนกหมวดหมู่เพิ่มมากยิ่งขึ้น

บรรณานุกรม

ภาษาไทย

กฤษณะ ไวยมัย, ชิดชนก ส่งศิริ และธนาวินท์ รักธรรมานนท์. **“**การใช้เทคนิคดาต้าไมน์นิ่งเพื่อพัฒนาคุณภาพการศึกษาคณะวิศวกรรมศาสตร์.**”** NECTEC Technical Journal 3, 11 (2544) : 134 -142.

จารวี ฉันทสิทธิพร. “การจำแนกชนิดยาเม็ดจากภาพถ่าย โดยใช้เทคนิคเครือข่ายใยประสาท.” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร, 2548.

เฉลิมศักดิ์ พิกุลศรี. สังคีตนิยมว่าด้วยดนตรีไทย. กรุงเทพมหานคร : สำนักพิมพ์โอเดียนสโตร์, 2530.

บุญธิดา สุวัชระกุลธร. “แบบจำลองการแปลงเสียง.” ปริญญานิพนธ์วิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2547.

ปวันรัตน์ มีชัย. “เครื่องมือจำแนกประเภทโดยวิธีวัดความใกล้เคียงจากข้อมูลของกลุ่ม.” ปริญญานิพนธ์ ปริญญาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาตร์ มหาวิทยาลัยเกษตรศาสตร์, 2551.

อานนท์ นามสนิท. “การจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวคเตอร์แมชชีนส์.” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2549.

ฐนียา สัตยพานิช และอัศนีย์ ก่อตระกูล. “ระบบรู้จำเสียงพูดภาษาไทยต่อเนื่องแบบเฉพาะบุคคลสำหรับการเข้าถึงอีเมล์” National Conference on Computer Science and Computer Engineering, Thailand (2549)

ภาษาต่างประเทศ

Chen, Shi-Huang and Shih-Hao Chen. “Content-based music genre classification using timbral feature vectors and support vector machine.” ICIS '09: Proceedings of the 2nd International Conference on Interaction Sciences 403 (November 2009) : 1095-1101.

Chih-Chung, Chang and Chih-Jen Lin. LIBSVM : a library for support vector machines[Online]. accessed 20 November 2009. Available from http://www.csie.ntu.edu.tw/~cjlin /libsvm

Bagwell, Chris et al. SoX - Sound eXchange[Online]. accessed 3 May 2009. Available from http://sox.sourceforge.net/

Downie, Stephen J. , Andreas F. Ehmann and David Tcheng. “Real-time genre classification for music digital libraries.” ACM/IEEE-CS Joint Conference on Digital Libraries 5 (June 2005) : 377-377.

Glen M. Ballou. Handbook for sound engineers. Boston, Mass. : Focal Press, 2002

Krejcie, Robert V. and Daryle W. Morgan. “Determining sample size for research activities.” Educational and Psychological Measurement 30 (1970) : 607-610.

Li, Tao, Ogihara Mitsunori and Qi Li. “A comparative study on content-based music genre classification.” Annual international ACM SIGIR Conference on Research and Development in informaion Retrieval 26 (2003) : 282-289.

Liu, Mingchun and Chunru Wan. “Feature selection for automatic classification of musical instrument sounds.” ACM/IEEE-CS Joint Conference on Digital Libraries 1 (2001) : 247-248.

McEnnis, Daniel et al. “jAudio: A feature extraction library.” International Conference on Music Information Retrieval (2005) : 600–603.

Shyamala, Doraisamy et al. “A Study on Feature Selection and Classification Techniques for Automatic Genre Classification of Traditional Malay Music.” International Conference on Music Information Retrieval 9 (2008) : 331-336.

Temko, Andrey and Climent Nadeu. "Classification of Acoustic events using SVM-based clustering schemes." Pattern recognition. (2006)

Tzanetakis, George and Perry Cook. “Musical Genre Classification of Audio Signals” IEEE Transactions on Speech and Audio Processing 10 (July 2002) : 293 – 302.

Manzalawy, Yasser EL and Vasant Honavar. WLSVM : Integrating LibSVM into Weka Environment[Online]. accessed 15 May 2009. Available from http://www.cs.Iastate.edu/~yasser/wlsvm

Witten, Ian H, and Eibe Frank. DATA MINING Practical Machine Learning Tools and Techniques. United States of America : Morgan Kaufmann Publishiers, 2005.

Wu, X., and Lim S. Yong. "Prediction of maximum scour depth at spur dikes with adaptive neural networks." In Neural Networks and Combinatorial Optimization in Civil and Structural Engineering, 61-66. Edited by Civil-Comp Press. Edinburgh : Civil-Comp Press,1993.

ภาคผนวก

ภาคผนวก ก

ตารางเครจซี่ มอร์แกน สำหรับการกำหนดขนาดตัวอย่าง

ตารางที่ 21 แสดงตารางเครจซี่ มอร์แกน สำหรับการกำหนดขนาดตัวอย่าง

	Confidence = 95.00%				Confidence = 99.0%			
	Degree of Accuracy/Margin of Error				Degree of Accuracy/Margin of Error			
Population Size	0.05	0.04	0.03	0.01	0.05	0.04	0.03	0.01
10	10	10	10	10	10	10	10	10
20	19	20	20	20	19	20	20	20
30	28	29	29	30	29	29	30	30
50	44	47	48	50	47	48	49	50
75	63	69	72	74	67	71	73	75
100	80	89	94	99	87	93	96	99
150	108	126	137	148	122	135	142	149
200	132	160	177	196	154	174	186	198
250	152	190	215	244	182	211	229	246
300	169	217	251	291	207	246	270	295
400	196	265	318	384	250	309	348	391
500	217	306	377	475	285	365	421	485
600	234	340	432	565	315	416	490	579
700	248	370	481	653	341	462	554	672
800	260	396	526	739	363	503	615	763
900	269	419	568	823	382	541	672	854
1,000	278	440	606	906	399	575	727	943

ตารางที่ 20 (ต่อ)

	Confidence = 95.00%				Confidence = 99.0%			
	Degree of Accuracy/Margin of Error				Degree of Accuracy/Margin of Error			
Population Size	0.05	0.04	0.03	0.01	0.05	0.04	0.03	0.01
1,200	291	474	674	1067	427	636	827	1119
1,500	306	515	759	1297	460	712	959	1376
2,000	322	563	869	1655	498	808	1141	1785
2,500	333	597	952	1984	524	879	1288	2173
3,500	346	641	1068	2565	558	977	1510	2890
5,000	357	678	1176	3288	586	1066	1734	3842
7,500	365	710	1275	4211	610	1147	1960	5165
10,000	370	727	1332	4899	622	1193	2098	6239
25000	378	760	1448	6939	646	1285	2399	9972
50,000	381	772	1491	8056	655	1318	2520	12455
75,000	382	776	1506	8514	658	1330	2563	13583
100,000	383	778	1513	8762	659	1336	2585	14227
250,000	384	782	1527	9248	662	1347	2626	15555
500,000	384	783	1532	9423	663	1350	2640	16055
1,000,000	384	783	1534	9512	663	1352	2647	16317
2,500,000	384	784	1536	9567	663	1353	2651	16478
10,000,000	384	784	1536	9594	663	1354	2653	16560

ตารางที่ 21 (ต่อ)

	Confidence = 95.00%					Confidence = 99.0%			
	Degree of Accuracy/Margin of Error					Degree of Accuracy/Margin of Error			
Population Size	0.05	0.04	0.03	0.01		0.05	0.04	0.03	0.01
100,000,000	384	784	1537	9603		663	1354	2654	16584
264,000,000	384	784	1537	9603		663	1354	2654	16586

ภาคผนวก ข

ตารางแสดงรายชื่อเพลงที่ใช้ในการทดลอง

ตารางที่ 22 แสดงรายชื่อเพลงที่ใช้ในการทดลองทดลอง

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
1	มโหรีย์	ชงโค	1 ค้างคาวกินกล้วย
2	มโหรีย์	ชงโค	2 แขกอาหวัง
3	มโหรีย์	ชงโค	3 แขกหนัง
4	มโหรีย์	ชงโค	4 แขกสาหร่าย
5	มโหรีย์	ชงโค	5 พัดชา
6	มโหรีย์	ชงโค	6 มฤคระเริง
7	มโหรีย์	ชงโค	7 เต่าเห่
8	มโหรีย์	ชงโค	8 เขมรพายเรือ
9	มโหรีย์	มโหรีย์	1 เขมรโพธิสัตว์
10	มโหรีย์	มโหรีย์	2 แขกสาหร่าย
11	มโหรีย์	มโหรีย์	3 ลาวดวงเดือน
12	มโหรีย์	มโหรีย์	4 ลาวเสี่ยงเทียน
13	มโหรีย์	มโหรีย์	5 ยอเร
14	มโหรีย์	มโหรีย์	6 มอญอ้อยอิ่ง
15	มโหรีย์	มโหรีย์	7 ญี่ปุ่นรำพึง- ญี่ปุ่นชะอ้อน
16	มโหรีย์	มโหรีย์	8 ครวญหา
17	มโหรีย์	The small Maho-Ri Ensemble	1 ค้างคาวกินกล้วย
18	มโหรีย์	The small Maho-Ri Ensemble	2 ต้นวรเชษฐ์

ตารางที่ 22 (ต่อ)

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
19	มโหรีย์	The small Maho-Ri Ensemble	3 แสนคำนึง
20	มโหรีย์	The small Maho-Ri Ensemble	4 สิบสองจุไทย
21	มโหรีย์	The small Maho-Ri Ensemble	5 นางครวญ
22	มโหรีย์	The small Maho-Ri Ensemble	6 มอญอ้อยอิ่ง
23	มโหรีย์	The small Maho-Ri Ensemble	7 ม่านมงคล
24	มโหรีย์	The small Maho-Ri Ensemble	8 โหมโรงขับไม้บัณเฑาะว์
25	มโหรีย์	The small Maho-Ri Ensemble	9 เขมรพายเรือ
26	ปี่พาทย์	โหมโรง	01- โหมโรง
27	ปี่พาทย์	โหมโรง	02-ศรีนวล
28	ปี่พาทย์	โหมโรง	03-เขมรไทรโยค
29	ปี่พาทย์	โหมโรง	04-อาหนัง
30	ปี่พาทย์	โหมโรง	05-ปาน้อย
31	ปี่พาทย์	โหมโรง	06-ลำปางใหญ่
32	ปี่พาทย์	โหมโรง	07-เขมรไล่ควาย

ตารางที่ 22 (ต่อ)

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
33	ปี่พาทย์	โหมโรง	08-ลาวเสี่ยงเทียน
34	ปี่พาทย์	ปี่พาทย์ไม้แข็ง	01-ราตรีประดับดาว
35	ปี่พาทย์	ปี่พาทย์ไม้แข็ง	02-การเกศ
36	ปี่พาทย์	ปี่พาทย์ไม้แข็ง	03-นาคบริพัตร
37	ปี่พาทย์	ปี่พาทย์ไม้แข็ง	04-ม้าย่อง
45	ปี่พาทย์	ปี่พาทย์ไม้นวม	04-พยายาม
46	ปี่พาทย์	ปี่พาทย์ไม้นวม	05-ไทยครอง
47	ปี่พาทย์	ปี่พาทย์ไม้นวม	06-ล้านนา ๑
48	ปี่พาทย์	ปี่พาทย์ไม้นวม	07-ล้านนา ๒
49	ปี่พาทย์	ปี่พาทย์ไม้นวม	08-ไทยใหญ่
50	ปี่พาทย์	ปี่พาทย์ไม้นวม	09-ไทยมุง
51	เครื่องสาย	01คำหวาน	1 คำหวาน
52	เครื่องสาย	01คำหวาน	2 ลาวจ้อย
53	เครื่องสาย	01คำหวาน	3 ลาวครวญ
54	เครื่องสาย	01คำหวาน	4 ลาวคำหอม
55	เครื่องสาย	01คำหวาน	5 ลาวดวงเดือน
56	เครื่องสาย	01คำหวาน	6 นางครวญ
57	เครื่องสาย	01คำหวาน	7 ลาวกระแต
58	เครื่องสาย	01คำหวาน	8 ต้นวรเชษฐ์
59	เครื่องสาย	02คลื่นกระทบฝั่ง	1 คลื่นกระทบฝั่ง

ตารางที่ 22 (ต่อ)

ลำดับที่	หมวดหมู่	ชื่ออัลบัม	ชื่อเพลง
60	เครื่องสาย	02คลื่นกระทบฝั่ง	2 บังใบ
61	เครื่องสาย	02คลื่นกระทบฝั่ง	3 แขกสาหร่าย (ทางเดี่ยว)
62	เครื่องสาย	02คลื่นกระทบฝั่ง	4 ลาวจ้อย
63	เครื่องสาย	02คลื่นกระทบฝั่ง	5 ลาวเสี่ยงเทียน
64	เครื่องสาย	02คลื่นกระทบฝั่ง	6 ทยอยญวน
65	เครื่องสาย	02คลื่นกระทบฝั่ง	7 ตารีกีปัส
66	เครื่องสาย	02คลื่นกระทบฝั่ง	8 ลาวลำปางใหญ่
67	เครื่องสาย	03แขกต่อยหม้อ	1 แขกต่อยหม้อ
68	เครื่องสาย	03แขกต่อยหม้อ	2 ดาวทอง
69	เครื่องสาย	03แขกต่อยหม้อ	3 เขมรพวง
70	เครื่องสาย	03แขกต่อยหม้อ	4 จีนหลวง
71	เครื่องสาย	03แขกต่อยหม้อ	5 ตลุ่มโปง
72	เครื่องสาย	03แขกต่อยหม้อ	6 แขกบรเทศ
73	เครื่องสาย	03แขกต่อยหม้อ	7 กราวกลาง
74	เครื่องสาย	03แขกต่อยหม้อ	8 ชมแสงจันทร์
75	เครื่องสาย	03แขกต่อยหม้อ	9 สาริกาแก้ว

ภาคผนวก ค

ค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะแบบต่าง ๆ

ตารางที่ 23 แสดงค่าความถูกต้องในการจำแนกโดยใช้ค่าคุณลักษณะแบบต่าง ๆ ของทั้งสามวิธีการ

วิธีการ	ค่าความถูกต้องในการจำแนกหมวดหมู่ (%)																	
	ค่าคุณลักษณะเดี่ยว						ค่าคุณลักษณะคู่					ค่าคุณลักษณะร่วมสามแบบ				ค่าคุณลักษณะร่วมสี่แบบ			ค่าคุณลักษณะร่วมห้าแบบ
	L	**M**	**P**	**S**	**Z**	**เฉลี่ย**	**M+L**	**M+P**	**M+S**	**M+Z**	**เฉลี่ย**	**M+P+L**	**M+P+S**	**M+P+Z**	**เฉลี่ย**	**M+P+Z+L**	**M+P+Z+S**	**เฉลี่ย**	**M+P+L+C+Z**
K-NN	61.10	80.45	70.30	41.36	44.27	**59.50**	79.68	82.39	80.41	80.36	**80.71**	80.79	82.20	81.95	**81.65**	81.08	80.36	**80.72**	80.41
SVM	49.54	79.69	59.49	49.09	47.25	**57.01**	75.71	86.16	68.26	66.28	**74.10**	86.16	85.73	85.49	**85.79**	85.39	85.63	**85.51**	84.95
ANN	53.41	79.63	41.36	41.61	42.72	**51.75**	79.82	68.26	79.44	79.83	**76.84**	66.62	63.43	69.62	**66.56**	65.75	70.83	**68.29**	64.59
เฉลี่ย	**54.68**	**79.92**	**57.05**	**44.02**	**44.75**		**78.40**	**78.94**	**76.04**	**75.49**		**77.86**	**77.12**	**79.02**		**77.17**	**79.18**		**76.65**
	56.08						**77.22**					**78.00**				**78.18**			**76.65**

หมายเหตุ : M = Mel-frequency cepstral coefficients (MFCC), P = Power Spectrum, L = Linear Predictive Coefficients (LPC), Z = Zero Crossings Rate, S = Spectral Centroid

ประวัติผู้วิจัย

ชื่อ - สกุล	นายโอภาส แก้วต่าย
วันเดือนปีเกิด	14 เมษายน พ.ศ. 2524
ที่อยู่	44/9 หมู่ 3 ต. กะเฉด อ.เมือง จ.ระยอง

ประวัติการศึกษา

พ.ศ. 2541	สำเร็จการศึกษาชั้นมัธยมศึกษาตอนปลาย (ม.6) โรงเรียนระยองวิทยาคม
พ.ศ. 2545	สำเร็จการศึกษาปริญญาศึกษาศาสตร์บัณฑิต วิชาเอกเทคโนโลยีการศึกษา วิชาโททัศนศิลป์ (ถ่ายภาพ) เกียรตินิยมอันดับ 2 มหาวิทยาลัยศิลปากร
พ.ศ. 2548	ศึกษาต่อระดับปริญญามหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ประวัติการทำงาน

พ.ศ. 2546- 2550	ตำแหน่ง นักคอมพิวเตอร์กราฟฟิค ศูนย์คอมพิวเตอร์ มหาวิทยาลัยศิลปากร