



ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์คอมพิวเตอร์แมชชีนส์
โดย นายอานนท์ นามสนิท

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ครุศาสตร์อุตสาหกรรมมหาบัณฑิต สาขาวิชาเทคโนโลยีคอมพิวเตอร์

คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร.วิบูลย์ ชื่นแขก)

19 มิถุนายน 2549

คณะกรรมการสอบวิทยานิพนธ์

ประธานกรรมการ

(ร้อยตำรวจตรีหญิง ดร.นิดาพรรณ สุริรัตน์)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.กัณฑ์พงษ์ วรรณปัญญา)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ นิตยสุวัฒน์)

การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์คอมพิวเตอร์แมชชีนส์

นายอานนท์ นามสนิท

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ครุศาสตร์อุตสาหกรรมมหาบัณฑิต
สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษา
บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2549
ISBN 974-19-0765-6
ลิขสิทธิ์ของสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ชื่อ : นายอานนท์ นามสนิท
ชื่อวิทยานิพนธ์ : การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์วิเคราะห์แมชชีนส์
สาขาวิชา : เทคโนโลยีคอมพิวเตอร์
สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ที่ปรึกษาวิทยานิพนธ์ : ร้อยตำรวจตรีหญิง ดร. นิตาพรรณ สุวรรรัตน์
ผู้ช่วยศาสตราจารย์ ดร. กันต์พงษ์ วรรณปัญญา
ปีการศึกษา : 2549

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนอ การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์วิเคราะห์แมชชีนส์ โดยได้ทำการสรรหาวิธีการหาค่าคุณลักษณะที่เหมาะสมกับการจำแนกเพลงไทย ซึ่งประกอบไปด้วย เพลงไทยลูกทุ่ง เพลงไทยลูกกรุง และเพลงไทยเดิม ทำการเรียนรู้และจำแนกด้วยเครื่องมือ ซอฟต์แวร์วิเคราะห์แมชชีนส์ โดยทำการจำแนกครั้งละสองกลุ่มตามความสอดคล้องของค่าคุณลักษณะที่สามารถจำแนกได้ การวิจัยนี้ใช้ข้อมูลเสียงทั้งหมด 75 เพลงแบ่งเป็นกลุ่มเรียนรู้ และกลุ่มทดสอบ 45 และ 30 เพลงตามลำดับ ผลการวิจัยพบว่าวิธีการที่นำเสนอสามารถจำแนกกลุ่มเพลงไทยได้ถูกต้องร้อยละ 86.7

(วิทยานิพนธ์มีจำนวนทั้งสิ้น 53 หน้า)

คำสำคัญ : การจำแนกกลุ่มเพลง, การจำแนกกลุ่มเพลงไทย, ซอฟต์แวร์วิเคราะห์แมชชีนส์, เอสวีเอ็ม



อาจารย์ที่ปรึกษาวิทยานิพนธ์

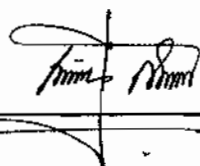
Name : Mr. Arnon Namsanit
Thesis Title : Thai Musical Group Classification Using Support Vector Machines
Major Field : Technology Computer
King Mongkut's Institute of Technology North Bangkok
Thesis Advisors : Pol. Sub. Lt. Dr. Nidapan Sureerattanan
Assistant Professor Dr. Kuntpong Woraratpanya
Academic Year : 2006

Abstract

This thesis presents the Thai musical group classification using support vector machines. The Thai musical features for classification have been investigated. These songs are trained and classified into their groups. The groups of Thai musics including Thai folk, Thai classic and Thai traditional musics are considered. There are 75 samples of musics deviding for train and test sets, 45 and 30 samples respectively. The experimental result illustrates that the proposed method performs 86.7% of classification accuracy.

(Total 53 Pages)

Keywords : Musical Classicfication, Thai Musical Classification, Suport Vector Machines, SVM.



Advisor

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลือจากหลายฝ่าย ขอขอบพระคุณ ร้อยตำรวจตรีหญิง ดร. นิดาพรรณ สุวีรัตน์ ผู้ช่วยศาสตราจารย์ ดร. กันต์พงษ์ วรรณปัญญา และผู้ช่วยศาสตราจารย์ ดร.สุพจน์ นิตยสุวัฒน์ ที่กรุณาให้คำแนะนำ ข้อคิดเห็นและแก้ไขข้อบกพร่องต่างๆ ในการทำวิจัยมาโดยตลอด และขอขอบคุณคุณคุณณัฐนันท์ ทัดพิทักษ์กุลและเพื่อนร่วมงาน เพื่อนร่วมห้องเรียนที่ให้ความช่วยเหลือและให้กำลังใจตลอดเวลา

ขอขอบพระคุณบิดา มารดา พี่และน้อง ที่ให้ความช่วยเหลือในด้านต่างๆ ให้สามารถทำงานวิจัยนี้ให้สำเร็จลุล่วงด้วยดีได้ ณ โอกาสนี้

อานนท์ นามสนิท

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
กิตติกรรมประกาศ	ง
สารบัญตาราง	ช
สารบัญภาพ	ซ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตของการวิจัย	2
1.4 คำจำกัดความในการวิจัย	3
บทที่ 2 ทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้อง	4
2.1 เสียง	4
2.1.1 ระดับเสียง	4
2.1.2 ความเข้มของเสียง	4
2.1.3 ลีสันเสียง	5
2.1.4 คุณภาพเสียง	5
2.1.5 ความยาวเสียง	5
2.2 ค่าคุณลักษณะที่ได้จากเสียง	5
2.2.1 พลังสเปกตรัม	5
2.2.2 อัตราค่าตัดศูนย์	6
2.2.3 การทำนายแบบเชิงเส้น	6
2.2.4 ความถี่เส้นสเปกตรัม	10
2.3 ซัพพอร์ตเวกเตอร์แมชชีนส์	11
2.3.1 แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีนส์	11
2.3.2 สมการพื้นฐานของซัพพอร์ตเวกเตอร์แมชชีนส์	12
2.3.3 ค่าความกว้างเส้นขอบ (Margin)	13
2.3.4 การแก้ปัญหา Dual Problem	14
2.3.5 ค่าผิดพลาดในการวางค่าคุณลักษณะ	15
2.3.6 เคอร์เนล	16
2.4 งานวิจัยที่เกี่ยวข้อง	17

สารบัญ (ต่อ)

	หน้า
บทที่ 3 วิธีการดำเนินการวิจัย	19
3.1 ศึกษาเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้อง	20
3.2 เตรียมข้อมูลและสกัดค่าคุณลักษณะ	20
3.3 เขียนโปรแกรมและทดสอบโปรแกรม	27
3.4 สรุปผลการทดสอบโปรแกรม	27
3.5 เขียนรายงานการศึกษาวิจัย	27
บทที่ 4 ผลการวิจัย	28
4.1 ผลการวิจัย	28
4.1.1 ผลการทดลองหาวิธีการหาค่าคุณลักษณะ	28
4.1.2 ผลการทดลองหาจำนวนชุดการเรียนรู้ที่ดีและเหมาะสม	33
4.1.3 ผลการทดลองหาเคอร์เนล	38
4.2 วิเคราะห์ผลการวิจัย	46
บทที่ 5 บทสรุป และข้อเสนอแนะ	47
5.1 อภิปรายผลการวิจัย	47
5.2 สรุปผล	49
5.3 ปัญหาที่พบ	50
5.4 ข้อเสนอแนะ	50
5.5 ประโยชน์ที่ได้รับ	50
เอกสารอ้างอิง	51
ประวัติผู้วิจัย	53

สารบัญตาราง

ตารางที่		หน้า
4-1	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีบีทสเปกตรัม (Beat Spectrum)	30
4-2	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีแอลพีซีดีโรไฟ์เซปสตรัม (LPC-Derived Cepstrum)	30
4-3	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีอัตราค่าตัดศูนย์ (Zero Crossing)	30
4-4	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีพลังงานสเปกตรัม (Spectrum Power)	31
4-5	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีเมลฟรีเควนซีเซปสตรัลโคเอฟฟิเชียนท์ (Mel Frequency Cepstral Coefficient-MFCC)	31
4-6	ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF)	32
4-7	ผลการทดลองหาจำนวนชุดเรียนรู้ที่ดีและเหมาะสม	34
4-8	ผลการทดลองหาค่าความถูกต้องของขนาดเฟรมต่างๆ	35
4-9	ผลการทดลองหาค่าความถูกต้องของจำนวนเฟรม	37
4-10	ผลของการใช้เคอร์เนลแบบเชิงเส้น (Linear)	38
4-11	ผลของการใช้เคอร์เนลแบบโพลิโนเมียล (Polynomial)	39
4-12	ผลของการใช้เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF)	41
4-13	ผลของการใช้เคอร์เนลแบบซิกมอยด์ (Sigmoid)	42
4-14	เปรียบเทียบค่าที่ดีที่สุดของแต่ละเคอร์เนล	44
4-15	แสดงค่าความถูกต้องระดับเฟรมและเพลงจากกลุ่มทดสอบ	45

สารบัญภาพ

ภาพที่	หน้า
2-1 ตัวอย่างสัญญาณเสียง	6
2-2 การขยายตัวของเส้นขอบ	12
2-3 เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง	13
2-4 การเกิดเวกเตอร์อนุโลม (Slack Vector) ในพื้นที่ค่าคุณลักษณะ	15
2-5 รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้	16
3-1 แผนภาพขั้นตอนการทำวิจัย	19
3-2 แผนภาพขั้นตอนการจำแนกกลุ่มเพลง	21
3-3 แผนภาพขั้นตอนกระบวนการเรียนรู้	22
3-4 แผนภาพขั้นตอนกระบวนการทดสอบ	24
3-5 แผนภาพขั้นตอนกระบวนการจำแนกอย่างละเอียด	25
4-1 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีบีทสเปกตรัม (Beat Spectrum)	28
4-2 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีแอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum)	29
4-3 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีอัตราค่าตัดศูนย์ (Zero Crossing Rate)	29
4-4 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีพลังงานสเปกตรัม (Spectrum Power)	29
4-5 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีเมลฟรีควเอนซีเซปสตรัลโคเอฟฟิเชียนท์ (Mel Frequency Cepstral Coefficient-MFCC)	30
4-6 กราฟแสดงผลการทดลองหาจำนวนเพลงชุดการเรียนรู้	34
4-7 กราฟแสดงผลการทดลองการหาขนาดของเฟรม	36
4-8 กราฟแสดงผลการทดลองการหาจำนวนเฟรมที่ดีที่สุดและเหมาะสม	37
4-9 กราฟแสดงผลการทดลองค่าความถูกต้องของเคอร์เนลแบบเชิงเส้น (Linear)	39
4-10 กราฟแสดงผลการทดลองใช้เคอร์เนลแบบโพลิโนเมียล (Polynomial)	40
4-11 กราฟแสดงผลการทดลองใช้เคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF)	42
4-12 กราฟแสดงผลการทดลองใช้เคอร์เนลแบบซิกมอยด์ (Sigmoid)	43
4-13 การเปรียบเทียบค่าที่ดีที่สุดของแต่ละเคอร์เนล	44

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การวิจัยในด้านเสียงได้รับความสนใจในโลก [1], [2], [3], [4], [5], [6], [7], [8] เนื่องจากเป็นอีกแขนงหนึ่งของสื่อ ที่ได้นำเข้ามาประยุกต์ใช้กับคอมพิวเตอร์ เพื่อให้คอมพิวเตอร์สามารถวิเคราะห์ ประมวลผล หรือสังเคราะห์ได้ ซึ่งจะทำให้คอมพิวเตอร์มีความสามารถใกล้เคียงกับมนุษย์มากยิ่งขึ้นด้วย ความสามารถหนึ่งของมนุษย์ในการได้ยินคือการจดจำและเมื่อได้ฟังอีกครั้งสามารถที่จะระบุหรือจำแนกได้ว่าเป็นเสียงใด ประเภทใด และถือเป็นที่มาของการทำวิจัยด้านเสียงในคอมพิวเตอร์ขึ้น ในอีกด้านหนึ่งที่น่าสนใจเป็นอย่างยิ่งคือเสียงเพลง เมื่อมนุษย์ได้รับฟังเพลงก็สามารถแยกชนิดหรือกลุ่มของเพลงนั้นๆ ได้ และถ้าหากนำความสามารถนี้มาประยุกต์ใช้ในคอมพิวเตอร์ได้ ก็จะทำให้เราสามารถจัดแบ่งหมวดหมู่ของเพลงที่มีอยู่มากมายในระบบคอมพิวเตอร์ได้อย่างเป็นระเบียบและง่ายต่อการเข้าถึงหรือค้นหาต่อไปในอนาคตด้วย

ไทยเป็นชาติหนึ่งที่มีดนตรีเป็นของตนเอง โดยเฉพาะดนตรีไทยเดิม ซึ่งบรรเลงโดยเครื่องดนตรีไทยที่มีเอกลักษณ์เฉพาะตัว ทำให้เสียงและท่วงทำนองที่ได้ไม่เหมือนดนตรีของชาติอื่น และเมื่อกระแสของเครื่องดนตรีทางตะวันตกได้แผ่เข้ามาในประเทศไทย ก็มีการได้นำเครื่องดนตรีบางอย่างเข้ามาประยุกต์เล่นกับเครื่องดนตรีไทย ทำให้เกิดแนวเพลงใหม่ ซึ่งได้แก่เพลงลูกทุ่ง และเพลงลูกกรุง อย่างไรก็ตามเพลงทั้งสองประเภทที่กล่าวมา ก็ยังมีความเป็นไทยสูง เนื่องจากท่วงทำนองมีการยืมทำนองเพลงไทยเดิมนำมาประยุกต์บรรเลง เพลงลูกทุ่งนั้นเป็นการผสมผสานระหว่างเพลงพื้นบ้าน เครื่องดนตรีไทยและเครื่องดนตรีสากล ส่วนเพลงลูกกรุงนั้นเป็นการผสมผสานระหว่างทำนองเพลงไทยเดิมและทำนองเพลงสากล ตัวอย่างเช่น เพลง “น้ำตาแสงใต้” เป็นเพลงประกอบละครเรื่อง “พันท้ายนรสิงห์” ประพันธ์ทำนองโดยสง่า ได้แรงบันดาลใจมาจากเพลง “เขมรไพรโยค” และเพลง “ลาวครวญ” ผู้ขับร้องคนแรกคือ สุรสิทธิ์ [9] แต่นำมาบรรเลงด้วยเครื่องดนตรีสากล จึงมีความเป็นสากลสูงกว่าเพลงกลุ่มใด และยังเป็นต้นแบบของวิวัฒนาการเพลงในเมืองไทยอีกด้วย ปัจจุบันเพลงไทยที่ได้ยินกันทั่วไป มีทำนองของเพลงหรือเครื่องดนตรีบรรเลงนั้นไม่แตกต่างจากเพลงสากล ทำให้ขาดเอกลักษณ์เฉพาะตัวไป ดังนั้นกลุ่มเพลงไทยที่ผู้วิจัยให้ความสนใจในการนำมาจำแนก จึงประกอบด้วย 3 กลุ่มคือ เพลงไทยเดิม เพลงไทยลูกทุ่ง และเพลงไทยลูกกรุง

ในปัจจุบันยังไม่พบว่ามีงานวิจัยด้านการจำแนกกลุ่มเพลงในประเทศไทยแต่อย่างใด ซึ่งหลักฐานงานวิจัยในการจำแนกกลุ่มเพลงนั้นปรากฏในต่างประเทศ [1] โดย ชู ได้ทำการจำแนก

กลุ่มเพลง 4 กลุ่ม ได้แก่ คลาสสิก (Classic) ร็อก (Rock) ป๊อป (Pop) และแจ๊ซ (Jazz) และได้ใช้ค่าคุณลักษณะ 5 ชนิด ได้แก่ บีทสเปกตรัม (Beat Spectrum), แอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum), อัตราค่าตัดศูนย์ (Zero Crossing Rate), พลังงานสเปกตรัม (Spectrum Power) และสัมประสิทธิ์เซปสตรัลบนความถี่เมล (Mel Frequency Cepstral Coefficient-MFCC)

แต่จากการวิจัยในเบื้องต้นพบว่าวิธีการหาค่าคุณลักษณะของ ชู นั้นมีบางค่าคุณลักษณะไม่เหมาะกับกลุ่มเพลงไทย ซึ่งจะแสดงผลการทดลองในบทที่ 4 ต่อไป จึงต้องหาวิธีการหาค่าคุณลักษณะที่เหมาะสมกับกลุ่มเพลงไทยเพิ่มเติม โดยณัฐนันท์ [7] ได้กล่าวถึงการใช้วิธีการความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF) ในการหาค่าคุณลักษณะของสัญญาณเสียงรบกวน แล้วจึงใช้ซัพพอร์ตเวกเตอร์แมชชีนส์จำแนกค่าคุณลักษณะที่ได้ ซึ่งเป็นวิธีหนึ่งที่ไม่ปรากฏในงานของ ชู ผู้วิจัยจึงได้นำวิธีการความถี่เส้นสเปกตรัม เข้ามาใช้ในงานวิจัยนี้ด้วย

ในการนำเสนองานวิจัยนี้ ได้หยิบยกกลุ่มเพลงสามกลุ่มคือ เพลงไทยเดิม เพลงไทยลูกทุ่ง และเพลงไทยลูกกรุง มาเป็นกลุ่มตัวอย่างที่จะทำการจำแนกโดยใช้คอมพิวเตอร์ โดยใช้ค่าคุณลักษณะ 4 ชนิดด้วยกันได้แก่ แอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum), อัตราค่าตัดศูนย์ (Zero Crossing Rate), พลังงานสเปกตรัม (Spectrum Power) และความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF) และใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ (Support Vector Machines) ในการจำแนกกลุ่มเพลงไทย โดยรายละเอียดของค่าคุณลักษณะที่ใช้และซัพพอร์ตเวกเตอร์แมชชีนส์จะกล่าวในบทที่ 2 ต่อไป

1.2 วัตถุประสงค์

เพื่อพัฒนาขั้นตอนวิธีการจำแนกกลุ่มเพลงไทยโดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ (Support Vector Machines – SVM)

1.3 ขอบเขตของการวิจัย

1.3.1 ขั้นตอนวิธีการจำแนกที่พัฒนาขึ้นมาให้คอมพิวเตอร์ทำตามวัตถุประสงค์ จะใช้โปรแกรม MATLAB 6.5 ในการพัฒนา

1.3.2 เพลงไทยที่นำมาเรียนรู้และทดสอบจะเป็นเพลงบรรเลง ที่มีแต่เสียงดนตรี อาจมีเสียงประสาน (Chorus) ได้

1.3.3 เพลงที่นำมาทดสอบจะต้องจัดเก็บอยู่ในรูปแบบไฟล์ .wav และมีความละเอียดที่ 44 กิโลเฮิรท์ซ

1.3.4 กลุ่มเพลงที่จะทำการจำแนกประกอบด้วย 3 กลุ่มได้แก่ เพลงไทยเดิม เพลงไทยลูกทุ่ง และเพลงไทยลูกกรุง

1.4 คำจำกัดความในการวิจัย

1.4.1 การจำแนก (Classification) หมายถึง การจำแนกออกเป็นประเภท หรือการแบ่งกลุ่มข้อมูล

1.4.2 การสกัดคุณลักษณะ (Feature Extraction) หมายถึง การแยกคุณลักษณะเด่น

1.4.3 การทดสอบ (Testing) หมายถึง กระบวนการทดสอบการหาผลลัพธ์

1.4.4 การเรียนรู้ (Training) หมายถึง กระบวนการสอน การฝึกฝนโปรแกรมหรือการเรียนรู้

1.4.5 เฟรม (Frame) หมายถึง หน่วยเสียงที่ตัดแบ่งย่อยมาจากหน่วยเสียงใหญ่

1.4.6 ชุดต้นแบบ (Model) หมายถึง ชุดข้อมูลที่ใช้เป็นแบบในการจำแนกที่ได้มาจากการเรียนรู้

1.4.7 เคอร์เนล (Kernel) หมายถึง สมการที่ใช้แก้ปัญหการวางตัวของข้อมูลให้แบ่งแบบเชิงเส้นได้

1.4.8 พื้นที่คุณลักษณะ (Feature Space) หมายถึง พื้นที่การวางตัวของข้อมูลคุณลักษณะ

บทที่ 2

ทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยเรื่องการจำแนกกลุ่มเพลงไทย มีเอกสารที่เกี่ยวข้องดังนี้

2.1 เสียง

เสียงคือพลังงานอย่างหนึ่ง [13] เกิดจากการสั่นสะเทือนของวัตถุ การสั่นสะเทือนนี้จะแผ่กระจายจากวัตถุที่สั่น แล้วแผ่รัศมีออกไปโดยรอบผ่านตัวกลาง ซึ่งอาจจะเป็น อากาศ น้ำ หรือของแข็งก็ได้ การสั่นสะเทือนทำให้โมเลกุลของตัวกลางนั้น ๆ เคลื่อนไหวไปตามแรงการสั่นสะเทือนนั้น จึงเกิดความเคลื่อนไหวคล้ายละลอกคลื่นและเกิดแรงดันตามมา เกิดเป็นเสียงขึ้น

เรื่องของเสียงเป็นศาสตร์สาขาหนึ่งทางวิทยาศาสตร์เรียกว่า สวณศาสตร์ (อ่านว่า สะ-วะ-นะ-สาด) หมายถึง วิชาที่ว่าด้วยเรื่องเสียง และเนื่องจากดนตรีมีความสัมพันธ์กับเสียงโดยตรง จึงมีความจำเป็นที่จะต้องทำความเข้าใจเรื่องวิทยาศาสตร์เบื้องต้นเกี่ยวกับเสียง

2.1.1 ระดับเสียง

ระดับเสียง (Pitch) หมายถึงเสียงสูงต่ำ เราสามารถเปรียบเทียบเสียง 2 เสียงได้ว่าเสียงไหนมีระดับเสียงสูงกว่า และเสียงไหนมีระดับเสียงต่ำกว่าโดยการฟัง ถ้าระดับเสียงต่างกันมากก็จะแยกแยะได้ง่าย ถ้าต่างกันเล็กน้อยจะแยกแยะได้ยากกว่า ระดับเสียงที่ต่ำกว่าจะเกิดจากแหล่งกำเนิดเสียงที่ใหญ่กว่าหรือแหล่งกำเนิดเสียงที่ตึ่งน้อยกว่า หรือแหล่งกำเนิดเสียงที่มีช่วงสั้นกว่า [13] ยกตัวอย่างเช่นเสียงผู้หญิงและเสียงผู้ชายจะมีค่าระดับเสียงที่ต่างกัน เนื่องจากผู้ชายจะเสียงต่ำกว่าผู้หญิง

ระดับเสียงเกิดจากความถี่ของการสั่นสะเทือนของวัตถุ หรืออาจกล่าวได้ว่า ความถี่ของการสั่นสะเทือนเป็นตัวกำหนดระดับเสียง วัตถุที่สั่นสะเทือนเร็วกว่าทำให้เกิดระดับเสียงสูงกว่า ในขณะที่วัตถุที่สั่นสะเทือนช้ากว่าทำให้เกิดระดับเสียงต่ำกว่า ความถี่ซึ่งเป็นตัวกำหนดระดับเสียงได้มาจากการวัดความถี่ของการสั่นสะเทือนของวัตถุโดยมีหน่วยเป็นรอบต่อวินาที วัตถุที่สั่นสะเทือนมากกว่าจะมีความถี่มากกว่าทำให้เกิดระดับเสียงสูงกว่า ถ้าความถี่มากขึ้นเท่าตัว ระดับเสียงจะสูงขึ้น 1 ช่วงคู่แปด (Octave) เช่น โน้ตที่มีความถี่ 220 รอบต่อวินาทีจะมีระดับเสียงสูงกว่าโน้ตที่มีความถี่ 110 รอบต่อวินาทีอยู่ 1 ช่วงคู่แปด

2.1.2 ความเข้มของเสียง

ความเข้มของเสียง (Volume หรือ Intensity) หมายถึง เสียงเบาเสียงดัง ความเข้มเสียงเกิดจากแรงสั่นสะเทือนของวัตถุที่เป็นแหล่งกำเนิดเสียงสั่นสะเทือนมากเสียงจึงดังมาก และในทางกลับกันถ้าวัตถุสั่นสะเทือนน้อย เสียงก็จะเบา ความเข้มของเสียงขึ้นอยู่กับความแรงที่ส่งจาก

แหล่งกำเนิดเสียงไปยังหู เวลาที่เราตะโกนเสียงจะสั่นสะท้อนมากทำให้เกิดเสียงดัง ความเข้มของเสียงวัดได้จากความสูงคลื่นเสียง (Amplitude) ถ้าคลื่นเสียงยิ่งสูงมากเสียงก็จะยิ่งดัง ในขณะที่เสียงเดินทางผ่านชั้นบรรยากาศนั้น ความเข้มเสียงจะน้อยลงตามลำดับกล่าวคือ เสียงจะเบาลง ดังนั้น ระยะทางที่เสียงต้องเดินทางก็จะมีผลต่อความเข้มเสียงด้วย

2.1.3 สีสันเสียง

สีสันเสียง (Timbre หรือ Tone color) หมายถึงเสียงที่มาจากแหล่งกำเนิดเสียงต่างๆ กัน เช่น เปียโน กีตาร์ มนุษย์ ฯลฯ คลื่นเสียงที่เกิดจากแหล่งที่มีสีสันเสียงต่างกันจะมีรูปร่างลักษณะต่างกัน สื่อดนตรี (Medium) ที่ต่างกันจะทำให้เกิดสีสันเสียงต่างกัน สื่อดนตรีก็คือเครื่องดนตรีหรือเส้นเสียงในกรณีของการขับร้อง เพลงเดียวกันที่บรรเลงด้วยเครื่องดนตรีต่างชนิดกันจะทำให้เกิดรสชาติต่างกันเพราะสีสันเสียงต่างกัน สิ่งสำคัญที่สุดที่เป็นตัวกำหนดสีสันเสียงก็คือ พาร์เชียล (Partials) พาร์เชียลที่น้อยจะทำให้เสียงนั้นใส อย่างเช่นเสียงของฟลูตที่เล่นคีย์สูง จะมีพาร์เชียลน้อย

2.1.4 คุณภาพเสียง

คุณภาพเสียง (Tone Quality) หมายถึง เสียงตีมากติน้อย เช่น เสียงโน้ตตัวเดียวกันที่เกิดจากเปียโน 2 ตัว จะมีคุณภาพเสียงไม่เหมือนกันทีเดียว แม้จะมีระดับเสียง ความเข้มเสียง และสีสันเสียงเหมือนกันทุกประการ เวลาเลือกซื้อเครื่องดนตรีจึงต้องคำนึงถึงคุณภาพเสียง ซึ่งแต่ละบุคคลจะชอบคุณภาพเสียงไม่เหมือนกัน

2.1.5 ความยาวเสียง

ความยาวเสียง (Duration) เป็นพื้นฐานของดนตรีที่เกี่ยวข้องกับระยะเวลา เสียงแต่ละเสียงที่เกิดขึ้นต้องมีระยะเวลาซึ่งทำให้เกิดเสียงสั้นยาว เสียงไม่ว่าจะมีระดับเสียงที่แน่นอนหรือไม่ก็ต้องมีความยาวเข้ามาเกี่ยวข้อง ความยาวเสียงเป็นที่มาของจังหวะซึ่งมีส่วนสำคัญอย่างยิ่งในดนตรีของทุกชาติทุกภาษา ความยาวของเสียงในที่นี้หมายถึง ความเจ็บบของตัวหยุดด้วย เนื่องจากดนตรีเป็นผลของกระบวนการเกิดเสียงสลับกับความเจ็บบ ไม่ว่าจะเป็นเสียงหรือความเจ็บบ ย่อมต้องมีความยาวหรือระยะเวลาทั้งสิ้น

2.2 ค่าคุณลักษณะที่ได้จากเสียง

ค่าคุณลักษณะที่ได้จากเสียงนั้นเป็นตัวแทนสำคัญ ที่เกิดจากการสกัดค่าคุณลักษณะ (Feature Extraction) [11] โดยการวิเคราะห์หาค่าที่นำมาแทนข้อมูลเสียงเพื่อนำไปใช้สร้างแบบจำลอง โดยงานวิจัยนี้ได้เสนอการวิเคราะห์คุณลักษณะ 4 ชนิด ได้แก่ พลังสเปกตรัม (Spectrum Power) อัตราค่าตัดศูนย์ (Zero Crossing Rate) การทำนายแบบเชิงเส้น (Linear Predictive Coding-LPC) และความถี่เส้นสเปกตรัม (Line Spectrum Frequency - LSF)

2.2.1 พลังสเปกตรัม (Spectrum Power)

ก่อนที่จะหาค่าพลังสเปกตรัมนั้น แต่ละเฟรมจะถูกหาค่าน้ำหนักก่อนด้วย Hanning Window [1], $h(n)$:

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

เมื่อ n คือ ลำดับที่จุดเสียง (Sample Point)
 N คือ จำนวนจุดเสียง (Sample Point) ในแต่ละเฟรม

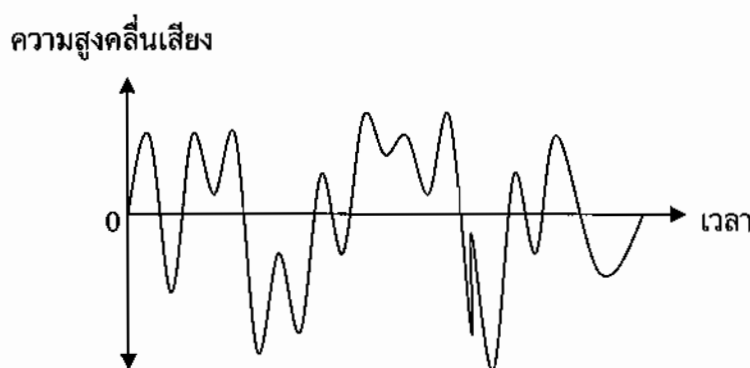
ค่าพลังงานสเปกตรัมคำนวณได้จากสมการการแปลงฟูเรียร์แบบเร็ว (Fast Fourier Transform – FFT)

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

เมื่อ k คือ จำนวนค่าพลังงานสเปกตรัม

2.2.2 อัตราค่าตัดศูนย์ (Zero Crossing Rate)

แนวความคิดของการหาค่าอัตราตัดศูนย์ของสัญญาณนั้น คือการนับจำนวนครั้งของสัญญาณที่ผ่านแกนศูนย์นั่นเอง



ภาพที่ 2-1 ตัวอย่างสัญญาณเสียง

จากภาพที่ 2-1 อัตราค่าตัดศูนย์ของสัญญาณมีค่าเท่ากับ 11

2.2.3 การทำนายแบบเชิงเส้น (Linear Predictive Coding-LPC)

การทำนายแบบเชิงเส้นเป็นเทคนิคที่สำคัญทางด้านการวิเคราะห์เสียงและการสังเคราะห์เสียงเนื่องจากมีความแม่นยำในการประมาณค่าพารามิเตอร์ของเสียงพูดเมื่อเทียบกับความเร็วในการประมวลผล หลักการพื้นฐานของการทำนายแบบเชิงเส้นอาศัยแนวคิดที่ว่าตัวอย่างสัญญาณเสียงสามารถประมาณค่าได้จากผลรวมของตัวอย่างสัญญาณเสียงจากอดีต การวิเคราะห์พารามิเตอร์เพื่อใช้การทำนายโดยทั่วไปเรียกว่า การเข้ารหัสการทำนายแบบเชิงเส้นหรือแอลพีซี (Linear Predictive Coding – LPC) ในทางด้านการประมวลผลสัญญาณเสียง การเข้ารหัสการทำนายแบบเชิงเส้นนี้ถูกนำไปใช้ในสองแนวทางคือ [12]

2.2.3.1 การเข้ารหัสสัญญาณเสียง โดยที่การเข้ารหัสการทำนายแบบเชิงเส้นถูกนำไปใช้เป็นวงจรกรองวิเคราะห์การทำนายแบบเชิงเส้น (LP Analysis Filter) เพื่อแยกส่วนซ้ำซ้อน (Redundancy) ของสัญญาณเสียงออก ส่วนที่เหลือเรียกว่าสัญญาณตกค้าง (Residual Signal)

2.2.3.2 การสังเคราะห์สัญญาณเสียง โดยที่การเข้ารหัสการทำนายแบบเชิงเส้นถูกนำไปใช้เป็นวงจรกรองการทำนายแบบเชิงเส้นผกผัน (Inverse LP Filter) หรือวงจรกรองสังเคราะห์การทำนายแบบเชิงเส้น (LP Synthesis filter) โดยที่ฟังก์ชันการถ่ายโอนของวงจรกรองดังกล่าวแสดงกรอบสเปกตรัมของสัญญาณเสียง วงจรกรองสังเคราะห์การทำนายแบบเชิงเส้นถูกใช้แสดงแทนช่องทางเดินเสียงของมนุษย์และใช้หาสัญญาณกระตุ้นที่เหมาะสมสำหรับวงจรกรองดังกล่าว

ในการวิเคราะห์การเข้ารหัสการทำนายแบบเชิงเส้นเริ่มต้นจากพิจารณากรอบสัญญาณที่มี N ตัวอย่าง คือ $s_1, s_2, \dots, s_N$ โดยอ้างว่าตัวอย่างสัญญาณปัจจุบันสามารถทำนายได้จากผลรวมของตัวอย่างสัญญาณในอดีต p ตัวอย่าง ดังสมการที่ 2-3

$$\tilde{s}_n = -a_1 s_{n-1} - a_2 s_{n-2} - a_3 s_{n-3} - \dots - a_p s_{n-p} = -\sum_{k=1}^p a_k s_{n-k} \quad (2-3)$$

เมื่อ p คืออันดับของการวิเคราะห์การทำนายแบบเชิงเส้นและ $a_1, a_2, \dots, a_p$ คือสัมประสิทธิ์การเข้ารหัสการทำนายแบบเชิงเส้น กำหนด e_n แทนค่าผิดพลาดระหว่างค่าจริงและค่าที่ทำนายได้จะได้ตามสมการที่ 2-4

$$e_n = s_n - \tilde{s}_n \quad (2-4)$$

เมื่อ s_n คือค่าจริง และ $\tilde{s}_n$ คือค่าที่ทำนายได้ เมื่อนำทั้งสองค่ามาหาความต่างก็จะได้เป็นค่าความผิดพลาด e_n จากสมการที่ 2-4 สามารถเขียนใหม่ได้เป็นสมการที่ 2-5

$$e_n = s_n + \sum_{k=1}^p a_k s_{n-k} \quad (2-5)$$

สัญญาณ e_n เรียกว่าสัญญาณตกค้าง เนื่องจากสัญญาณ e_n ได้จากผลการลบสัญญาณ s_n ด้วย $\tilde{s}_n$ เนื่องจากค่าสหสัมพันธ์ช่วงสั้น (Short-term Correlation) ระหว่างตัวอย่างของสัญญาณตกค้างมีค่าต่ำ ดังนั้นประมาณได้ว่ากรอบสเปกตรัมกำลังของสัญญาณตกค้างมีลักษณะเรียบ (Flat) เมื่อทำการแปลงซี (Z Transform) ของสมการที่ 2-5 ได้ดังสมการที่ 2-6

$$E(z) = A(z)S(z) \quad (2-6)$$

โดยที่ $S(z)$ $A(z)$ และ $E(z)$ เป็นผลการแปลงซีของสัญญาณเสียง ค่าสัมประสิทธิ์และสัญญาณตกค้างตามลำดับ

$$A(z) = 1 + \sum_{k=1}^p a_k z^{-k} \quad (2-7)$$

โดยที่ $A(z)$ หรือวงจรกรองไวเทนนิ่ง (Whitening Filter) มีหน้าที่แยกความสัมพันธ์ช่วงสั้นที่ปรากฏในสัญญาณเสียงซึ่งก็คือการทำให้สเปกตรัมเรียบ เนื่องจาก $E(z)$ ประมาณได้ว่ามีสเปกตรัมเรียบดังนั้นสามารถออกแบบกรอบสเปกตรัมของสัญญาณช่วงสั้นได้จากการวิเคราะห์การทำนายแบบเชิงเส้นรูปแบบออลโพล (All-pole Model) หรือรูปแบบออโตเรเกรสซีฟ (Autoregressive Model) ดังสมการที่ 2-8

$$H(z) = \frac{1}{A(z)} \quad (2-8)$$

วงจรกรอง $A(z)$ เรียกได้อีกชื่อหนึ่งว่าวงจรกรองผกผัน (Inverse Filter) เนื่องจากเป็นส่วนผกผันของรูปแบบออลโพล $H(z)$ ของสัญญาณเสียงและรากของ $A(z)$ ทำให้เกิดโพลใน $H(z)$ นั่นคือตำแหน่งฟอร์แมนต์ของเสียงที่ได้จากช่องทางเสียงที่มีฟังก์ชันถ่ายโอน $H(z)$

การหากรอบสเปกตรัมกำลังช่วงสั้นของเสียงด้วยวิธีการวิเคราะห์การทำนายแบบเชิงเส้นคำนวณได้จาก $H(z)$ บนวงกลมหนึ่งหน่วย โดยในขั้นแรกต้องหาสัมประสิทธิ์การทำนายแบบเชิงเส้นของสัญญาณเสียงก่อน โดยปกติหาได้จากการทำให้ค่าผิดพลาดการทำนายแบบเชิงเส้นทั้งหมดยกกำลังสองดังสมการที่ 2-9 มีค่าต่ำที่สุด

$$E = \sum_{n=-\infty}^{\infty} e_n^2 \quad (2-9)$$

วิธีการหาค่าสัมประสิทธิ์ a_k สามารถกระทำได้โดยการทำให้ค่า e_n มีค่าน้อยที่สุดเพราะนั่นหมายถึงค่าผิดพลาดยิ่งน้อยเท่าไร ค่า s_n และ $\tilde{s}_n$ ก็จะมีค่าใกล้เคียงกันมากเท่านั้น และมีสองวิธีที่ใช้ในการหาค่าคือวิธีอัตสหสัมพันธ์ (Autocorrelation Method) และวิธีโคแวลเรียนต์ (Covariance Method) ในงานวิจัยนี้ได้ใช้วิธีอัตสหสัมพันธ์ จึงกล่าวถึงแต่วิธีนี้เท่านั้น

วิธีอัตสหสัมพันธ์ (Autocorrelation Method) ในการวิเคราะห์การทำงานแบบเชิงเส้นช่วงสั้นหาได้โดยการวิเคราะห์แบบหน้าต่างสัญญาณเสียงและอ้างว่าตัวอย่างสัญญาณภายนอกหน้าต่างนี้มีค่าเท่ากับศูนย์ แล้วจึงทำให้ค่าผิดพลาดตามสมการที่ 2-9 มีค่าต่ำสุด ซึ่งสมการที่ 2-9 เมื่อนำมาแตกจะได้ดังสมการที่ 2-12

$$E = \sum_{n=-\infty}^{\infty} \left[s_n - \sum_{k=1}^p a_k s_{n-k} \right]^2 \quad (2-10)$$

$$E = \sum_{n=-\infty}^{\infty} s_n^2 - \sum_{n=-\infty}^{\infty} \left[2s_n \sum_{k=1}^p a_k s_{n-k} \right] + \sum_{n=-\infty}^{\infty} \left[\sum_{k=1}^p a_k s_{n-k} \right]^2 \quad (2-11)$$

$$E = \sum_{n=-\infty}^{\infty} s_n^2 - 2 \sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s_n s_{n-k} + \sum_{n=-\infty}^{\infty} \left[\sum_{k=1}^p a_k s_{n-k} \right]^2 \quad (2-12)$$

การหาค่า a_k ที่ทำให้ E มีค่าต่ำที่สุดสามารถกระทำได้โดยการหา $\partial E / \partial a_i = 0$ ในสมการที่ 2-12 ของแต่ละค่า $i = 1, 2, 3, \dots, p$ เมื่อ $k = 1, 2, 3, \dots, p$ ได้เป็นสมการเชิงเส้น p สมการ และแต่ละสมการมี k ค่า ซึ่งก็คือ a_k ที่ไม่ทราบค่า p ตัวดังนี้

$$\frac{\partial E}{\partial a_i} = 0 = -2 \sum_{n=-\infty}^{\infty} s_n s_{n-i} + 2 \sum_{n=-\infty}^{\infty} \left[\sum_{k=1}^p a_k s_{n-k} \right] s_{n-i} \quad (2-13)$$

$$\sum_{n=-\infty}^{\infty} s_{n-i} s_n = \sum_{k=1}^p a_k \sum_{n=-\infty}^{\infty} s_{n-i} s_{n-k} \quad (2-14)$$

เนื่องจากพจน์แรกเป็นอัตราสัมพันธ์ R_i ของ s_n ที่มีความยาวจำกัด สามารถเขียนสมการที่ 2-14 ใหม่ได้เป็นสมการที่ 2-15

$$\sum_{k=1}^p r_{|i-k|} a_k = -r_i \quad \text{เมื่อ } 1 \leq i \leq p \quad (2-15)$$

เมื่อแก้สมการที่ 2-12 ซึ่งประกอบไปด้วยสมการเชิงเส้น p สมการได้เป็นค่าสัมประสิทธิ์ a_k ที่ต้องการออกมา โดยเขียนสมการที่ 2-12 ให้อยู่ในรูปเมทริกซ์ $Ra = -r$ โดยที่

$$R = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & r_0 & r_1 & \cdots & r_{p-2} \\ r_2 & r_1 & r_0 & \cdots & r_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_0 \end{bmatrix} \quad (2-13)$$

$$a = [a_1, a_2, \dots, a_p]^T \quad (2-14)$$

$$r = [r_1, r_2, \dots, r_p]^T \quad (2-15)$$

เมตริกซ์ R เรียกว่าเมตริกซ์อัตโนมัติสัมพันธ์ (Autocorrelation Method) ซึ่งมีโครงสร้างแบบโทพลิตซ์ (Toeplitz) เป็นเมตริกซ์แบบสมมาตรคือมีค่าในแนวทแยงมุมเท่ากันทั้งหมดทุกแถว สำหรับวิธีที่ช่วยหาผลเฉลยของสมการยูล-วอล์กเกอร์ ในสมการที่ 2-10 และหาสัมประสิทธิ์การทำนายแบบเชิงเส้นในสมการ 2-12 มีหลายวิธีดังเช่น วิธีเลวินสัน-เดอร์บิน (Levinson-Durbin Algorithm) และวิธีเชอร์ (Schur Algorithm) เป็นต้น

2.2.4 ความถี่เส้นสเปกตรัม (Line Spectrum Frequency - LSF)

ความถี่เส้นสเปกตรัม (Line Spectrum Frequency - LSF) เป็นพารามิเตอร์รูปแบบหนึ่งที่เกิดขึ้นมาจากการทำนายแบบเชิงเส้น เนื่องจากพารามิเตอร์การทำนายแบบเชิงเส้นในขั้นตอนการประมาณค่าพารามิเตอร์อาจทำให้เกิดความไม่เสถียรของสัญญาณได้ ส่งผลต่อคุณภาพของเสียง ในขณะที่พารามิเตอร์ความถี่เส้นสเปกตรัมมีคุณสมบัติที่เด่นคือ ค่าพารามิเตอร์อยู่ภายในขอบเขตที่จำกัดมีการเรียงลำดับของค่าพารามิเตอร์และสามารถตรวจสอบเสถียรภาพของวงจรกรองได้ง่าย [12]

ในการคำนวณหาความถี่เส้นสเปกตรัมเริ่มต้นจากพหุนามอันดับ M ของวงจรกรองผกผันในเชิงซี ดังสมการที่ 2-7 โดยทำการแยกส่วนสมการดังกล่าวเป็นพหุนามอันดับ M+1 เป็นจำนวนสองพหุนามดังสมการที่ 2-23 และ 2-24

$$P(z) = A(z) + z^{-(M+1)} A(z^{-1}) \quad (2-23)$$

$$Q(z) = A(z) - z^{-(M+1)} A(z^{-1}) \quad (2-24)$$

โดยพหุนาม $P(z)$ และ $Q(z)$ มีความสัมพันธ์กับ $A(z)$ ตามสมการที่ 2-25

$$A(z) = \frac{P(z) + Q(z)}{2} \quad (2-25)$$

พหุนาม $P(z)$ และ $Q(z)$ สอดคล้องกับรูปแบบช่องทางเดินเสียงที่ไร้การสูญเสียขณะที่ช่องระหว่างเสียงปิดและเปิดตามลำดับ [12] และรากของพหุนาม $P(z)$ และ $Q(z)$ เรียกว่าความถี่เส้นสเปกตรัม โดยพหุนามทั้งสองนี้มีคุณสมบัติดังต่อไปนี้

1. ศูนย์ (Zeros) หรือรากทั้งหมดของพหุนาม $P(z)$ และ $Q(z)$ จะตั้งอยู่บนวงกลมหนึ่งหน่วยเสมอ
2. ศูนย์ของ $P(z)$ และ $Q(z)$ วางเรียงสลับซึ่งกันและกันตามลำดับจากน้อยไปมาก

โดยสามารถแสดงได้ว่าวงจรกรองผกผัน $A(z)$ จะมีเฟสต่ำสุดถ้าเส้นคู่สเปกตรัมของ $A(z)$ สอดคล้องกับเงื่อนไขทั้งสองนี้ ดังนั้นวงจรสังเคราะห์การเข้ารหัสการทำนายแบบเชิงเส้น

สามารถยืนยันได้ว่ามีเสถียรภาพโดยการทำการประมวลพารามิเตอร์การเข้ารหัสการทำนายแบบเชิงเส้นในรูปแบบคู่เส้นสเปกตรัม

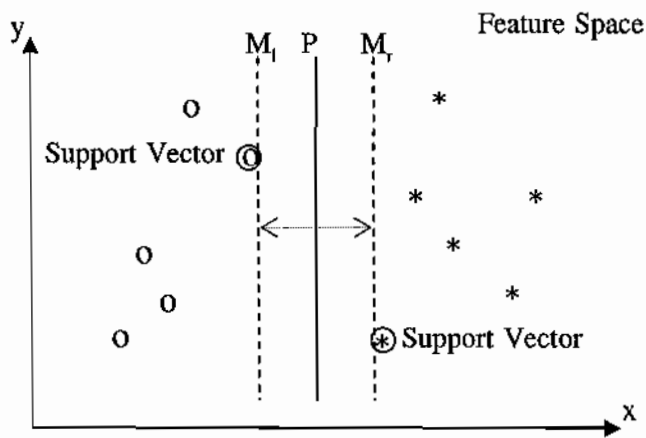
เมื่อพิจารณาจากของพหุนามทั้งสองพบว่าพหุนาม $P(z)$ และ $Q(z)$ มีรากจริงอยู่ที่ -1 และ 1 ตามลำดับสำหรับรากอื่นๆ อยู่บนวงกลมหนึ่งหน่วยโดยวางเรียงสลับกันตามคุณสมบัติของพหุนามทั้งสอง และรากของทั้งสองพหุนามมีลักษณะเป็นคู่เชิงซ้อนสังยุค ดังนั้นในการเก็บรากของพหุนามเพื่อใช้เป็นพารามิเตอร์จึงเก็บเพียง M ค่า

เนื่องจากรากของพหุนาม $A(z)$ แสดงตำแหน่งฟอร์แมนต์ของเสียงและพหุนาม $P(z)$ และ $Q(z)$ สัมพันธ์กับ $A(z)$ ตามสมการที่ 2-23 และ 2-24 ดังนั้นรากของพหุนามทั้งสองจึงสัมพันธ์กับฟอร์แมนต์ด้วย โดยรากของพหุนาม $A(z)$ แต่ละอันจับคู่กับรากของพหุนาม $P(z)$ และ $Q(z)$ อย่างละหนึ่งราก การเกาะกลุ่มของรากหรือความถี่เส้นสเปกตรัมจำนวน 2 หรือ 3 ราก จะกำหนดลักษณะความถี่ฟอร์แมนต์และขนาดแบนด์วิดท์ฟอร์แมนต์โดยขึ้นกับความชิดของกลุ่มรากดังกล่าว ถ้ารากชิดกันมากจะทำให้เกิดฟอร์แมนต์และมีแบนด์วิดท์ ในทางตรงกันข้ามกลุ่มรากจะแสดงถึงสเปกตรัมที่มีแบนด์วิดท์กว้างคือไม่เกิดฟอร์แมนต์ นอกจากนี้คุณสมบัติอีกอย่างหนึ่งของคู่เส้นสเปกตรัมคือความไวทางสเปกตรัมมีลักษณะเฉพาะก็คือเมื่อมีการเปลี่ยนแปลงคู่เส้นสเปกตรัมใดๆ จะทำให้เกิดการเปลี่ยนแปลงของสเปกตรัมกำลังของการเข้ารหัสการทำนายแบบเชิงเส้นเฉพาะบริเวณรอบๆ เท่านั้น ทำให้สามารถทำการประเมินค่า (Quantization) ได้อย่างอิสระโดยไม่มีผลกระทบจากการลดทอนเนื่องจากการประเมินค่า จากสเปกตรัมหนึ่งไปสู่สเปกตรัมอื่นๆ

2.3 ซัพพอร์ตเวกเตอร์แมชชีนส์ (Support Vector Machines – SVM)

2.3.1 แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีนส์

ซัพพอร์ตเวกเตอร์แมชชีนส์เป็นสมการที่ใช้ในการจำแนกค่าคุณลักษณะของสองกลุ่มที่วางตัวอยู่ในพื้นที่คุณลักษณะ (Feature Space) ออกจากกันโดยจะสร้างเส้นแบ่ง (Plane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด โดยเส้นตรงนั้นก็จะถูกเพิ่มเส้นขอบ (Margin) ออกไปทั้งสองข้าง โดยเส้นขอบที่เพิ่มนั้นจะขนานกับเส้นเดิมเสมอ เส้นขอบที่เพิ่มขึ้นมานี้จะขยายออกไปจนกว่าจะสัมผัสกับค่าของกลุ่มตัวอย่างที่ใกล้ที่สุด [13] ดังภาพที่ 2-2



ภาพที่ 2-2 การขยายตัวของเส้นขอบ

จากภาพที่ 2-2 เส้น M_1 และ M_2 คือเส้นขอบที่ขยายออกไปทางด้านซ้ายและขวาตามลำดับ และ P คือเส้นแบ่งข้อมูลทั้งสองกลุ่ม เมื่อเส้น M_1 และ M_2 ขยายออกจนไปสัมผัสค่าข้อมูลที่ใกล้ที่สุดซึ่งข้อมูลที่อยู่บนเส้นขอบของทั้งสองฝั่งนั้นเรียกว่าเวกเตอร์สนับสนุน (Support Vector) [13], [14] ค่าระยะความห่างหรือขนาดของเส้นขอบก็จะถูกวัด โดยเส้น P จะเปลี่ยนความชันไปเรื่อย ๆ เพื่อที่จะหาความกว้างสูงสุดของเส้นขอบ

กระบวนการโดยรวมของซัพพอร์ตเวกเตอร์แมชชีนนั้น เป็นการหาค่าความชันของเส้น P ที่มีขนาดของเส้นขอบสูงที่สุดนั่นเอง

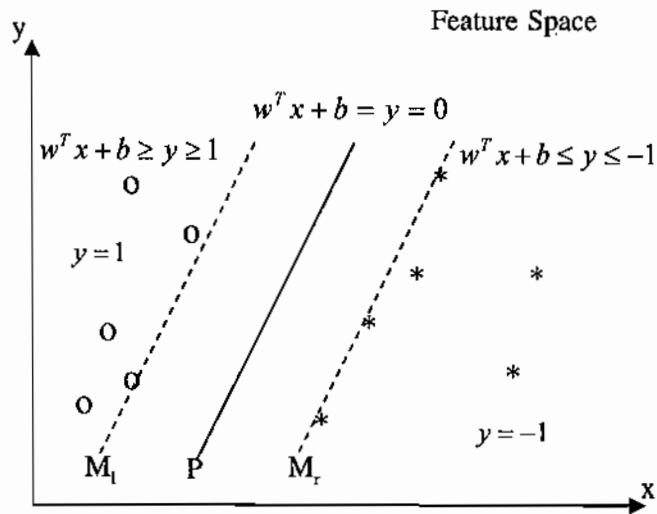
2.3.2 สมการพื้นฐานของซัพพอร์ตเวกเตอร์แมชชีนส์

หากนำแนวความคิดของซัพพอร์ตเวกเตอร์แมชชีนส์ที่กล่าวไปในข้อ 2.3.1 มาเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยข้อมูลที่นำมาวางลงในพื้นที่คุณลักษณะนั้นเป็นกลุ่มข้อมูลที่อยู่ในรูปของเวกเตอร์

$$X = ((x_1, y_1), \dots, (x_i, y_i)) \quad (2-26)$$

โดย X คือ ชุดค่าคุณลักษณะ

ค่าคุณลักษณะที่วางตัวอยู่ในพื้นที่คุณลักษณะ จะถูกแบ่งด้วยเส้นตรงดังภาพที่ 2-2 และเมื่อนำเส้นตรงมาแทนค่าด้วยสมการเส้นตรง $y = mx + b$ โดยมีการกำหนดกลุ่มของข้อมูลทั้งสองฝั่งเป็นเพียงสองค่าซึ่งแทนด้วยค่า y เพื่อให้ข้อมูลที่อยู่ในกลุ่มเดียวกันที่มาจากหลายค่ากลายเป็นค่าเดียว ดังแสดงในภาพที่ 2-3



ภาพที่ 2-3 เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง

จากภาพที่ 2-3 เส้นตรง M_l แทนด้วยสมการ $w^T x + b \geq y \geq 1$ ซึ่งข้อมูล y ที่มากกว่า 1 ก็จะถูกกำหนดค่าใหม่โดยให้ y เท่ากับ 1 และพจน์ของ w ก็คือค่าความชัน เช่นเดียวกันกับเส้นตรง M_r ที่ค่าของ y จะถูกกำหนดใหม่เมื่อ y น้อยกว่า -1 เป็น -1 ดังนั้นสมการที่เกิดขึ้นใหม่จากสมการเส้นขอบ 2-27 และ 2-28 สามารถกำหนดได้ดังสมการที่ 2-29

$$\text{เมื่อ } w^T x + b \geq y \quad \text{กำหนด } y = 1 \quad (2-27)$$

$$\text{เมื่อ } w^T x + b \leq y \quad \text{กำหนด } y = -1 \quad (2-28)$$

$$y(w^T x + b) - 1 \geq 0 \quad (2-29)$$

โดย y คือ ค่ากลุ่มข้อมูล (1, -1)

w คือ ค่าความชัน

x คือ ค่าคุณลักษณะ

b คือ ค่าคงที่ (ค่าตัดแกน y)

2.3.3 ค่าความกว้างเส้นขอบ (Margin)

การคำนวณความกว้างของเส้นขอบ (Margin) ต้องทำการคำนวณพจน์ w ให้อยู่ในรูปปกติ (Normalization) [13] โดยคำนวณจากสมการที่ 2-27 และ 2-28 เมื่อแทนค่า y ลงไปแล้ว

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

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

$$w^T (x^+ - x^-) = 2$$

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

$$M = \frac{2}{\|w\|} \quad (2-30)$$

โดยที่ M คือ ความกว้างของเส้นขอบ (Margin)

2.3.4 การแก้ปัญห Dual Problem

หลังจากที่ได้สมการที่ 2-29 และ 2-30 ของการหาเส้นแบ่งและค่าความกว้างตามลำดับแล้ว จึงทำการแก้สมการนั้นต่อด้วย Lagrangian Dual Problem [13] เพื่อหาค่าในพจน์ของ w ในสมการที่ 2-29

$$L(w, b, \alpha) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^N \alpha_i [y_i((w_i \cdot w_i) + b) - 1] \quad (2-31)$$

เมื่อ $\alpha_i \geq 0; i = 1, \dots, N$

สมการที่ 2-31 ถูกนำมาหาอนุพันธ์ (Differential)

$$\frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^N y_i \alpha_i x_i = 0 \quad (2-32)$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^N y_i \alpha_i = 0 \quad (2-33)$$

หลังจากได้สมการที่ผ่านแก้ปัญหด้วย Dual Problem แล้วค่า w จะลดรูป และหาได้จาก

$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (2-34)$$

เมื่อ α คือ สัมประสิทธิ์คงที่

ค่า b นั้นไม่ปรากฏในการแก้ปัญหด้วย Dual Problem จึงหาได้จาก

$$b = -\frac{\max_{y_i=-1} (w^T x_i) + \min_{y_i=1} (w^T x_i)}{2} \quad (2-35)$$

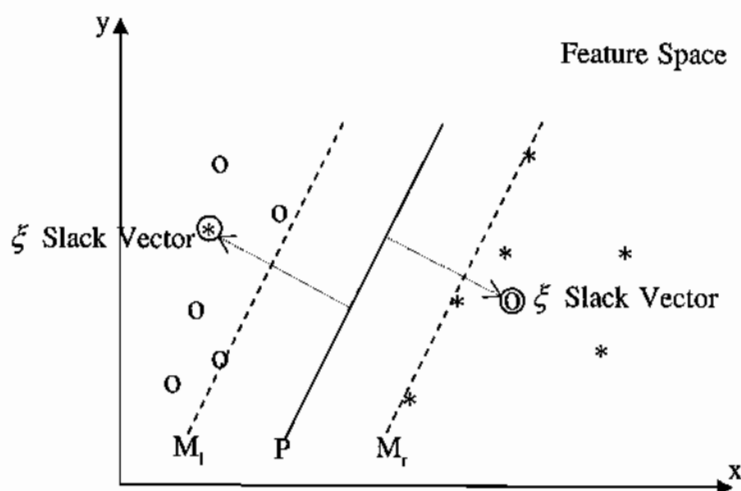
เมื่อนำค่า w ที่แก้ปัญหามาไว้ ไปใส่ในสมการที่ 2-29 ซึ่งเป็นสมการในการหาเส้นแบ่งจะได้

$$y_i \left(\sum_{i=1}^N \alpha_i y_i (x^T x_i) + b \right) - 1 \geq 0 \quad (2-36)$$

จากสมการที่ 2-36 คือการหาค่าสัมประสิทธิ์ α เพื่อนำมาใช้บอกการวางตัวของเส้นแบ่งข้อมูลต่อไป

2.3.5 ค่าผิดพลาดในการวางค่าคุณลักษณะ

ในการวางค่าคุณลักษณะในพื้นที่คุณลักษณะ บางครั้งอาจมีบางคุณลักษณะที่วางตัวอยู่ผิดกลุ่มออกไปซึ่งจะทำให้สมการของเส้นแบ่ง ไม่เป็นความจริง การแก้ปัญหานี้คือการยอมให้มีการเกิดข้อผิดพลาดนี้ได้ในระยะหนึ่งที่ยอมรับได้ คำนันเรียกว่าเวกเตอร์อนุโลม (Slack Vector)



ภาพที่ 2-4 การเกิดเวกเตอร์อนุโลม (Slack Vector) ในพื้นที่ค่าคุณลักษณะ

เมื่อนำค่าเวกเตอร์อนุโลมเข้ามาในสมการที่ 2-31 ซึ่งเป็นขั้นตอนการแก้ปัญหา Dual Problem ก็จะทำให้เกิดตัวแปร C ขึ้นซึ่งก็เป็นค่าคงที่ค่าหนึ่ง ดังสมการที่ 2-37

$$L(w, b, \xi, \alpha) = \frac{1}{2}(w \cdot w) + \frac{C}{2} \sum_{i=1}^N \xi_i^2 - \sum_{i=1}^N \alpha_i [y_i ((w_i \cdot w_i) + b) - 1 + \xi_i] \quad (2-37)$$

สมการที่ 2-37 ถูกนำมาหาอนุพันธ์ (Differential) เช่นเดียวกับขั้นตอนที่ปรากฏในหัวข้อ

2.3.4

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial w} = w - \sum_{i=1}^N y_i \alpha_i x_i = 0 \quad (2-38)$$

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial \xi} = C \xi - \alpha = 0 \quad (2-39)$$

$$\frac{\partial L(w, b, \xi, \alpha)}{\partial b} = \sum_{i=1}^N y_i \alpha_i = 0 \quad (2-40)$$

โดยที่ ξ คือ เวกเตอร์อนุโลม

หลังจากได้สมการที่ผ่านแก้ปัญหาด้วย Dual Problem แบบมีค่าเวกเตอร์อนุโลมแล้วค่า w ยังคงเป็นเช่นเดิมแต่จะได้ขอบเขตของ α ขึ้นมา

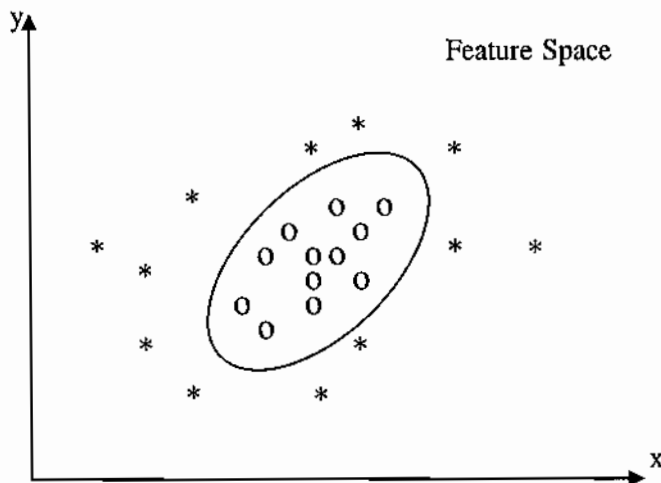
$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (2-41)$$

เมื่อ $C \geq \alpha \geq 0$

ค่า C จึงถือเป็นค่าพารามิเตอร์อีกค่าหนึ่งที่จะต้องกำหนดขึ้นมาเพื่อใช้หาค่าสัมประสิทธิ์ α

2.3.6 เคอร์เนล (Kernel)

ในโลกความเป็นจริงนั้นข้อมูลสองกลุ่มไม่ได้วางตัวในพื้นที่คุณลักษณะ เป็นเพียงสองกลุ่ม และแบ่งได้โดยเส้นตรง แต่ข้อมูลอาจจะจับกลุ่มกันในตำแหน่งต่างๆ ดังนั้นจึงเป็นปัญหาทำให้ไม่สามารถที่จะใช้สมการซัพพอร์ตเวกเตอร์แมชชีนส์แบบเชิงเส้นได้ ดังนั้นจะต้องมีเครื่องมือมาช่วยให้ข้อมูลเหล่านั้นเรียงตัวเสียใหม่ในพื้นที่ที่เรียกว่าพื้นที่หลายมิติ (Higher Dimensional Space)



ภาพที่ 2-5 รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้

ในเคอร์เนลนั้นคือการคูณกันของชุดเวกเตอร์ของ x ใดๆ

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

เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิดด้วยกัน

โพลิโนเมียล (Polynomial) :

$$K(x_i, x_j) = (< x_i, x_j > + 1)^d \quad (2-42)$$

เมื่อ d คือ ค่าเลขยกกำลัง

เรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF):

$$K(x_i, x_j) = \exp \left[-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right] \quad (2-43)$$

เมื่อ σ คือ ค่าพารามิเตอร์

ซิกมอยด์ (Sigmoid):

$$K(x_i, x_j) = \tanh(k \langle x_i, x_j \rangle + \mu) \quad (2-44)$$

เมื่อ k, μ คือ ค่าพารามิเตอร์

ดังนั้นจากสมการของเคอร์เนลนั้นสามารถที่จะแทนลงไปในตำแหน่งของ $x_i^T x_j$ ในสมการที่ 2-36 จึงเขียนเป็นสมการใหม่ดังนี้

$$y_i \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) - 1 \geq 0 \quad (2-45)$$

สมการที่ 2-45 เป็นสมการที่ใช้ในขั้นตอนที่จะเรียนรู้ว่าจะวางตำแหน่งเส้นแบ่งไว้ที่ตำแหน่งใดโดยทำงานร่วมกับเคอร์เนล เพื่อแปลงให้ข้อมูลที่ยากต่อการแบ่งแบบเชิงเส้นสามารถแบ่งได้เมื่อถูกทำให้เป็นข้อมูลแบบหลายมิติ (Higher Dimension) ดังนั้นจึงมีอีกสมการหนึ่งที่ใช้ค่า w และ b เดิมมาจัดตำแหน่งของข้อมูลเพื่อให้ทราบว่าข้อมูลนั้นเป็นกลุ่มใด กำหนดได้ดังสมการที่ 2-46

$$f(x) = \text{sgn} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) \quad (2-46)$$

เมื่อ $f(x)$ คือค่า y หาในรูปของ x

2.4 งานวิจัยที่เกี่ยวข้อง

งานวิจัยด้านเสียงที่เกี่ยวข้องและเป็นแรงบันดาลใจของงานวิจัยนี้มีดังนี้

ชู [1] ทำการวิจัยโดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์จำแนกกลุ่มเพลง โดยการจำแนกกลุ่มเพลงออกเป็น 4 กลุ่มได้แก่ ร็อก (Rock) ป๊อป (Pop) คลาสสิก (Classic) และแจ๊ซ (Jazz)

แล้วจึงทำการดึงค่าคุณลักษณะโดยใช้ บีทสเปกตรัม (Beat Spectrum) พลังสเปกตรัม (Power Spectrum) การทำนายแบบเชิงเส้น (Linear Predictive Coding-LPC) อัตราค่าตัดศูนย์ (Zero Crossing Rate) และสัมประสิทธิ์เซปสตัลบนความถี่เมล (Mel Frequency Cepstral Coefficient) แล้วจึงนำเอาค่าคุณลักษณะเหล่านี้ไปทำการเรียนรู้ด้วยซัพพอร์ตเวกเตอร์แมชชีนส์ ก็จะได้แบบจำลองหรือชุดต้นแบบ ซึ่งจะใช้ร่วมกับการจำแนกด้วยซัพพอร์ตเวกเตอร์แมชชีนส์

สำหรับงานวิจัยด้านการวิเคราะห์ค่าคุณลักษณะพลังสเปกตรัม ดังปรากฏในงานวิจัยของชู และ พุด [1], [2] โดยทำการดึงค่าพลังสเปกตรัมออกจากเสียง แต่พุดนำค่าสเปกตรัมไปหาค่าของบีทสเปกตรัมต่อ เพื่อหาคาบเวลาของจังหวะในเพลง

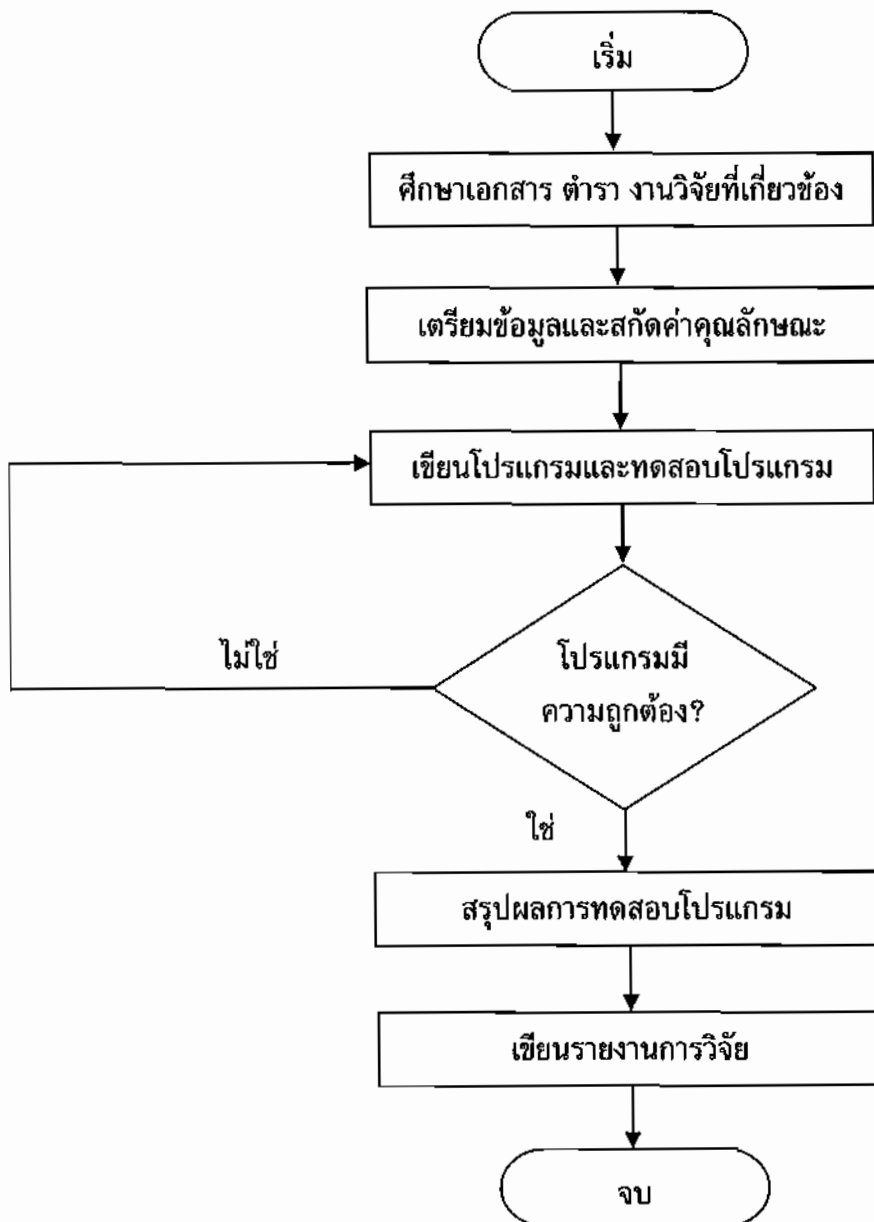
สำหรับงานวิจัยด้านการวิเคราะห์ค่าคุณลักษณะการทำนายแบบเชิงเส้น (Linear Predictive Coding-LPC) ดังปรากฏในงานวิจัยของ ศรีวิต [11] โดยนำไปใช้ในการบ่งชี้ตัวผู้พูดจากเสียงพูด

สำหรับงานวิจัยด้านการวิเคราะห์ค่าคุณลักษณะความถี่เส้นสเปกตรัม (Line Spectrum Frequency – LSF) ดังปรากฏในงานวิจัยของ ณัฐนันท์ และชัชวาลย์ [7,12] โดยณัฐนันท์ ใช้ในการจำแนกเสียงรบกวน

บทที่ 3

วิธีการดำเนินการวิจัย

ในการทำวิจัยครั้งนี้ ผู้วิจัยได้แบ่งขั้นตอนการทำงานของงานของการจำแนกกลุ่มเพลงไทยออกเป็นขั้นตอนต่าง ๆ 5 ขั้นตอน ดังภาพที่ 3-1



ภาพที่ 3-1 แผนภาพขั้นตอนการทำวิจัย

3.1 ศึกษาเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้อง

ขั้นตอนนี้เป็นขั้นตอนที่ผู้วิจัยได้ทำการศึกษาข้อมูลจากเอกสาร ตำรา งานวิจัยที่เกี่ยวข้อง และแหล่งข้อมูลอื่น ๆ โดยได้ศึกษาถึงวิธีการ เทคนิคต่าง ๆ ที่ใช้ และทำการทดลองในเบื้องต้น แล้วนำความรู้ต่าง ๆ ที่ได้มาวิเคราะห์ เพื่อใช้เป็นพื้นฐานและแนวทางในการทำวิจัย ซึ่งจะนำไปสู่การออกแบบงานวิจัยให้ได้ผลลัพธ์ที่มีประสิทธิภาพ

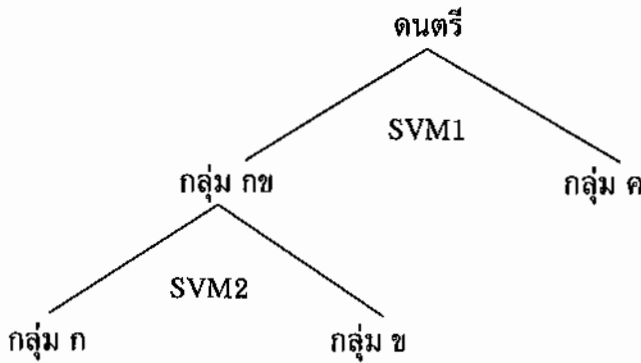
3.2 เตรียมข้อมูลและสกัดค่าคุณลักษณะ

3.2.1 จัดหาและรวบรวมเสียงเพลงที่จะนำมาใช้เป็นข้อมูลในการวิจัยอันได้แก่ เพลง 3 กลุ่มคือ เพลงไทยเดิม เพลงไทยลูกทุ่ง และเพลงไทยลูกกรุง โดยหาและสะสมข้อมูลจากที่อยู่ในระบบอินเทอร์เน็ตและซีดีคาราโอเกะเพื่อดึงเอาเฉพาะแทร็คดนตรีเท่านั้น กลุ่มละ 25 เพลง

3.2.2 จากนั้นจึงนำเสียงเพลงเหล่านั้นมาตัดเป็นเฟรมหรือวินโดว์ โดยแต่ละเพลงจะตัดออกเป็น 2,000 เฟรม เฟรมละ 1,024 จุด รวม 75 เพลง ดังนั้นจะได้จำนวนเฟรมรวมทั้งสิ้น 150,000 เฟรม

3.2.3 ข้อมูลที่ได้สะสมในแต่ละกลุ่มเพลง กลุ่มละ 25 เพลง จะถูกใช้เป็นชุดการเรียนรู้ 15 เพลง และชุดทดสอบ 10 เพลง

3.2.4 นำเฟรมที่ได้ตัดไว้แล้วนั้นมาหาคุณลักษณะของแต่ละกลุ่มเพลงในวิธีที่ได้ศึกษาไว้ จากนั้นนำค่าคุณลักษณะที่ได้หาไว้ของแต่ละกลุ่มเพลงมาวาดกราฟตามแต่ละกลุ่มของตน เพื่อใช้เป็นเครื่องมือในการตัดสินใจเลือกวิธีการหาคุณลักษณะที่เหมาะสมกับกลุ่มเพลงไทยมากที่สุด เพื่อให้เป็นไปตามลำดับการแบ่งกลุ่มที่ได้ออกแบบไว้ดังแสดงในภาพที่ 3-2 โดยจะนำกลุ่มเพลงทั้งสามกลุ่มมาแบ่งเป็นสองสาย หรืออาจมองเป็นการจำแนกที่ละสองกลุ่ม ดังนั้นจึงมีความจำเป็นที่จะต้องจับสองกลุ่มใด ๆ มารวมเป็นกลุ่มเดียวกัน โดยจะใช้หลักเกณฑ์ของกราฟที่นำมาวาดจากค่าคุณลักษณะของแต่ละกลุ่มเพลงว่าเพลงกลุ่มใดมีกราฟที่แตกต่างมากที่สุด อีกสองกลุ่มที่เหลือจะถูกจับเป็นกลุ่มเดียวกัน กลุ่มเพลงที่นำมารวมกันจะใช้ตัวแทนกลุ่มละครึ่งหนึ่งของชุดการเรียนรู้ที่ได้กำหนดไว้ คือกลุ่มละ 7.5 เพลง (15,000 เฟรม) รวมกันคือ 15 เพลง ซึ่งกลุ่มแรกที่เกิดจากการรวมกันของกลุ่มเพลงสองกลุ่มกำหนดแทนด้วยกลุ่ม กข และอีกกลุ่มเพลงที่สามารถถูกแยกได้จากกราฟค่าคุณลักษณะกำหนดแทนด้วยกลุ่ม ค และแต่ละขั้นของการจำแนกจะใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ (SVM) แต่จะใช้ค่าคุณลักษณะที่ต่างกันออกไป ซึ่งขึ้นอยู่กับประเภทของเพลงที่จะทำการจำแนกในขณะนั้น ดังที่กล่าวไปแล้วคือดูจากกราฟของค่าคุณลักษณะ



ภาพที่ 3-2 แผนภาพขั้นตอนการจำแนกกลุ่มเพลง

3.2.5 เลือกวิธีการหาค่าคุณลักษณะ ในการจำแนกของกลุ่มเพลงที่ต่างกันนั้นจะไม่สามารถใช้วิธีการหาค่าคุณลักษณะวิธีเดียวกันได้ สำหรับขั้นแรกในการเลือกวิธีที่จะนำมาใช้ในการหาค่าคุณลักษณะของเสียงคือการใช้วิธีการหรืออัลกอริธึมที่มีอยู่ในวงการของการวิจัยด้านเสียง ด้วยการหาเอกสารงานวิจัยที่เกี่ยวข้องมาศึกษา แล้วทำการลงมือพัฒนาตามวิธีการหาค่าคุณลักษณะเหล่านั้นของแต่ละกลุ่มเพลง จากนั้นนำค่าคุณลักษณะที่ได้ของแต่ละกลุ่มเพลงมาวาดกราฟแยกกันตามวิธีการและกลุ่มเพลง แล้วนำกราฟที่ได้มาวิเคราะห์ว่าวิธีการใดสามารถจำแนกเพลงกลุ่มใดได้บ้าง และเพื่อให้เป็นไปตามขั้นตอนการจำแนกที่ได้ออกแบบไว้ดังแสดงในภาพที่ 3-2 การวิเคราะห์กราฟค่าคุณลักษณะทำให้รู้ว่าจะรวมเพลงกลุ่มใดเข้าด้วยกันเป็นกลุ่ม กข

หลังจากได้ข้อมูลเสียงที่มีความยาวเท่ากันแล้ว จะนำข้อมูลนั้นมาสกัดค่าคุณลักษณะโดยวิธีการหาค่าคุณลักษณะที่ได้เลือกสรรไว้ที่สอดคล้องกับกลุ่มเพลงที่จะทำการจำแนกสองกลุ่ม ดังนั้นตามขั้นตอนที่ได้ออกแบบไว้จะต้องมีการหาค่าคุณลักษณะสองชุด ชุดแรกใช้สำหรับการจำแนกกลุ่ม กข และกลุ่ม ค และอีกชุดสำหรับกลุ่ม ก และกลุ่ม ข

3.2.5.1 ค่าคุณลักษณะที่ใช้สำหรับจำแนกกลุ่ม กข และกลุ่ม ค

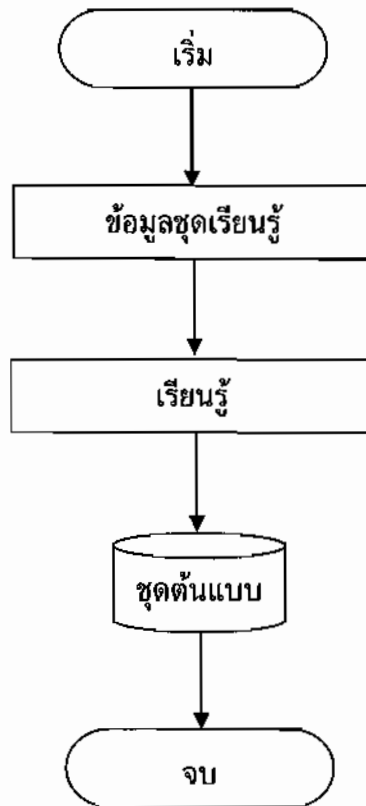
ขั้นตอนนี้ทำให้รู้ว่ากลุ่มเพลง กข คือการรวมกันของกลุ่มเพลงไทยลูกกรุงและเพลงไทยลูกทุ่ง และกลุ่ม ค คือกลุ่มเพลงไทยเดิม และค่าคุณลักษณะที่เหมาะสมใช้กับการจำแนกในครั้งแรกคือ พลังงานสเปกตรัม (Spectrum Power) และการเข้ารหัสการทำนายแบบเชิงเส้น (Linear Prediction Coding – LPC) (แสดงการได้มาของวิธีหาค่าคุณลักษณะที่เหมาะสมในบทที่ 4)

3.2.5.2 ค่าคุณลักษณะที่ใช้สำหรับจำแนกกลุ่ม ก และกลุ่ม ข

วิธีการหาค่าคุณลักษณะที่ใช้ในการจำแนกขั้นนี้ได้แก่วิธีการเข้ารหัสการทำนายแบบเชิงเส้น (Linear Prediction Coding – LPC) และอัตราค่าตัดศูนย์ (Zero Crossing Rate) และเส้นความถี่สเปกตรัม (Line Spectral Frequency-LSF) (แสดงการได้มาของวิธีหาค่าคุณลักษณะที่เหมาะสมในบทที่ 4)

3.2.6 กระบวนการเรียนรู้ (Training Process)

กระบวนการเรียนรู้คือกระบวนการที่ทำให้ได้มาซึ่งชุดต้นแบบ (Model) ซึ่งจะใช้ชุดต้นแบบนี้เป็นมาตรฐานในการจำแนกต่อไป กระบวนการเรียนรู้สามารถอธิบายได้ภาพที่ 3-3



ภาพที่ 3-3 แผนภาพขั้นตอนกระบวนการเรียนรู้

จากแผนภาพดังกล่าว สามารถอธิบายรายละเอียดของขั้นตอนกระบวนการต่างๆ ได้ดังต่อไปนี้

3.2.6.1 ขั้นตอนการเตรียมข้อมูลชุดเรียนรู้ (Training Data Set)

นำข้อมูลเสียงที่ตัดเป็นเฟรมและหาค่าคุณลักษณะแล้วในขั้นตอนที่ 3.2.3 ถึง 3.2.5 ของทั้งสามกลุ่ม ที่เป็นเฟรมของเพลงที่คัดเลือกไว้ใช้ในการเรียนรู้เข้าทำการเรียนรู้ด้วยซัพพอร์ตเวกเตอร์แมชชีนส์

3.2.6.2 ขั้นตอนการเรียนรู้ (Training)

เมื่อหาค่าคุณลักษณะของแต่ละกลุ่มเพลงที่จะทำการจำแนกแล้ว ค่าของคุณลักษณะเหล่านั้นจะถูกนำเข้าไปทำการเรียนรู้ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ (SVM) ในการเรียนรู้โดยสมการทางคณิตศาสตร์ที่ใช้สำหรับการเรียนรู้นั้นดังสมการที่ 2-45 ในบทที่ 2

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b$$

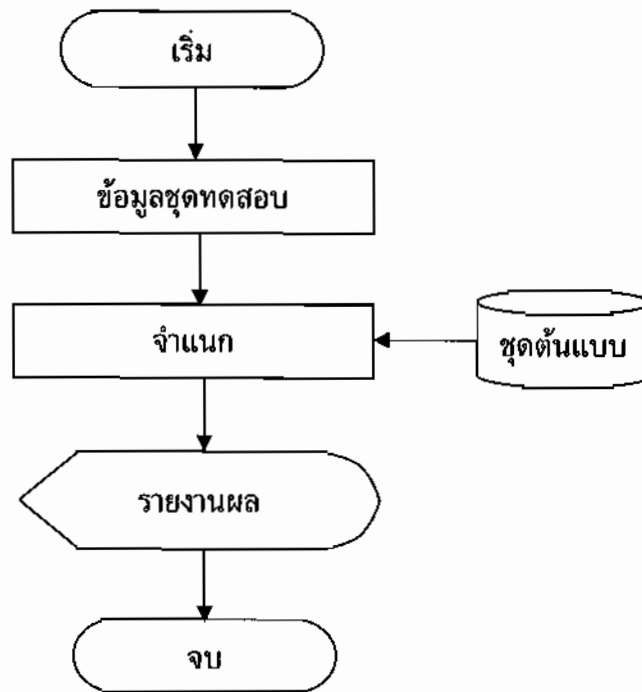
เมื่อ α	คือ สัมประสิทธิ์ความชัน
b	คือ ค่าคงที่
x	คือ ค่าคุณลักษณะ
y	คือ กลุ่มของค่าคุณลักษณะ
i, j	คือ ลำดับค่าคุณลักษณะ
K	คือ เคอร์เนล
N	คือ จำนวนค่าคุณลักษณะทั้งหมด

จากสมการที่ 2-45 เป็นสมการที่ใช้ในการเรียนรู้ แต่ในบางครั้งค่าคุณลักษณะที่ถูกจัดวางในพื้นที่คุณลักษณะ (Feature Space) ไม่สามารถแบ่งแยกได้ด้วยสมการแบบซัพพอร์ตเวกเตอร์แมชชีนส์แบบเชิงเส้น (Linear Support Vector Machines – LSVM) จึงจะต้องเอาเคอร์เนล (Kernel) เข้ามาช่วย ในการทำให้ข้อมูลสองกลุ่มวางตัวเรียงกันเสียใหม่ จนสามารถหาแผ่นแบ่ง (Hyperplane) ได้ สมการเคอร์เนลที่ใช้มี 3 ชนิดด้วยกันได้แก่ โพลีโนเมียล (Polynomial) เรเดียลเบสิสฟังก์ชัน (Radial Basis Function–RBF) และซิกมอยด์ (Sigmoid)

สำหรับงานวิจัยนี้เลือกใช้เคอร์เนลเรเดียลเบสิสฟังก์ชัน เนื่องจากเป็นเคอร์เนลที่เหมาะสมที่สุดในงานวิจัยนี้ ซึ่งได้ทำการทดลองและเสนอผลไว้ในบทที่ 4 ผลลัพธ์ที่ได้ออกมาคือชุดต้นแบบ (Model) ซึ่งจะใช้เป็นข้อมูลมาตรฐานที่ใช้ในการจำแนกร่วมกับสมการซัพพอร์ตเวกเตอร์แมชชีนส์ชุดที่ใช้ในการจำแนกต่อไป

3.2.7 กระบวนการทดสอบ (Testing)

กระบวนการทดสอบนั้นเป็นกระบวนการที่มีลำดับงานคล้ายคลึงกับการเรียนรู้ จะต่างกันเพียงอัลกอริธึมของซัพพอร์ตเวกเตอร์แมชชีนส์ที่ใช้ในการจำแนก ซึ่งอธิบายได้ดังภาพที่ 3-4



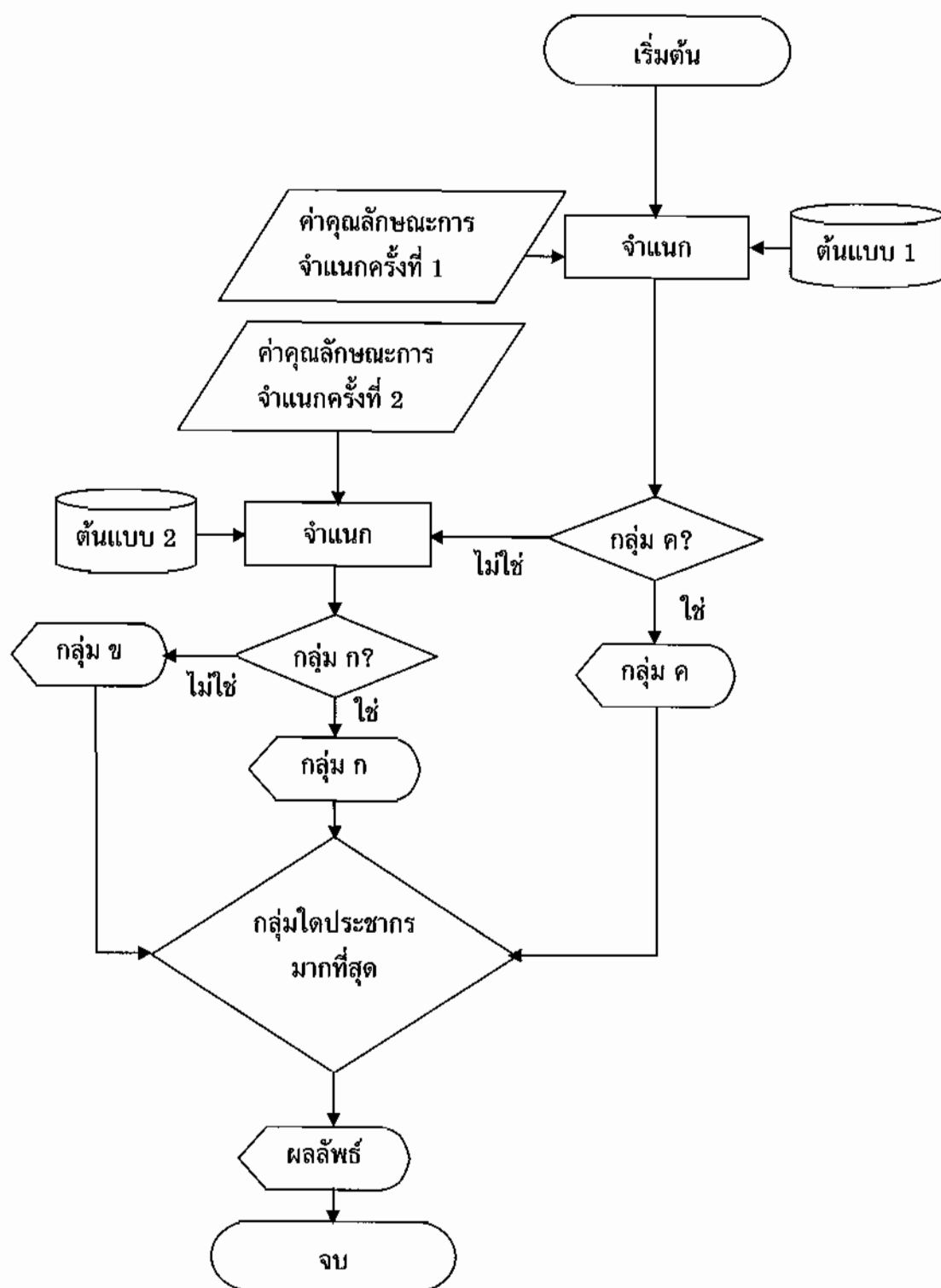
ภาพที่ 3-4 แผนภาพขั้นตอนกระบวนการทดสอบ

3.2.7.1 ขั้นตอนการเตรียมข้อมูลชุดทดสอบ (Testing Data Set)

นำข้อมูลเสียงที่ตัดเป็นเฟรมและหาค่าคุณลักษณะแล้วในขั้นตอนที่ 3.2.3 ถึง 3.2.5 ของทั้งสามกลุ่ม ซึ่งเป็นเฟรมของเพลงที่คัดเลือกไว้ใช้ในการทดสอบเข้าทำการจำแนกด้วยซอฟต์แวร์เวคเตอร์แมชชีนส์

3.2.7.2 ขั้นตอนการจำแนก (Classification)

หลังจากได้หาค่าคุณลักษณะของเฟรมที่ต้องการจำแนกแล้ว ค่าคุณลักษณะเหล่านั้นก็จะถูกส่งมายังระบบการจำแนกคือซอฟต์แวร์เวคเตอร์แมชชีนส์ กระบวนการทดสอบหรือจำแนกนั้นเมื่อนำมาเขียนเป็นแผนผังงานตามการออกแบบในภาพที่ 3-2 จะได้ดังภาพที่ 3-5



ภาพที่ 3-5 แผนภาพขั้นตอนกระบวนการจำแนกอย่างละเอียด

ตามภาพที่ 3-5 เมื่อค่าคุณลักษณะในชั้นแรกของเฟรมใดๆ เข้ามา ซึ่งก็คือค่าคุณลักษณะที่ได้จากวิธีสเปกตรัมพาวเวอร์ (Spectrum Power) และลิเนียร์พรีดิคชันโคัดดิง (LPC) จะถูกนำเข้ามาจำแนกด้วยซัพพอร์ตเวกเตอร์แมชชีนส์ โดยใช้ต้นแบบที่ 1 (Model 1) ซึ่ง

เป็นต้นแบบที่ได้จากการเรียนรู้จากค่าคุณลักษณะเดียวกันกับการหาค่าคุณลักษณะในขั้นแรก เมื่อทำการตัดสินใจได้ว่าเฟรมที่จำแนกนั้นเป็นเพลงกลุ่มใดระหว่างกลุ่ม กข หรือกลุ่ม ค หากเป็นกลุ่ม ค ก็จะแสดงผลและสรุปได้ทันทีว่าเฟรมนั้นเป็นเพลงกลุ่ม ค เนื่องจากในการจำแนกขั้นแรกนี้ใช้ค่าคุณลักษณะที่มีอำนาจจำแนกเพลงกลุ่ม ค ได้แต่ยังไม่สามารถจำแนกเพลงกลุ่ม ก และ ข ออกจากกันได้ แต่ถ้าหากการจำแนกในขั้นแรกไม่ปรากฏว่าเป็นกลุ่ม ค เฟรมนั้นก็จะถูกนำมาหาค่าคุณลักษณะใหม่อีกครั้งด้วยวิธีแอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum) และอัตราค่าตัดศูนย์ (Zero Crossing Rate) และเส้นความถี่สเปกตรัม (Line Spectral Frequency-LSF) ซึ่งเป็นวิธีที่ใช้จำแนกเพลงกลุ่ม ก และ ข ออกจากกัน แล้วทำการจำแนกอีกครั้งด้วยซัพพอร์ตเวกเตอร์แมชชีนส์ โดยใช้ต้นแบบที่ 2 ร่วมในการจำแนก สมการที่ใช้ในการจำแนก

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

เมื่อ α	คือ สัมประสิทธิ์ความชัน
b	คือ ค่าคงที่
x	คือ ค่าคุณลักษณะ
y	คือ กลุ่มของค่าคุณลักษณะ
i, j	คือ ลำดับค่าคุณลักษณะ
K	คือ เคอร์เนล
N	คือ จำนวนค่าคุณลักษณะทั้งหมด

3.2.7.3 ขั้นแสดงผล (Result)

เมื่อผ่านขั้นตอนของการทดสอบหรือการจำแนกมาแล้วนั้น เฟรมแต่ละเฟรมก็จะถูกจำแนกว่าเป็นเพลงในกลุ่มใด แต่นั่นเป็นเพียงเฟรมเดียวของทั้งหมด 2,000 เฟรมจากหนึ่งเพลง ดังนั้นการจำแนกได้ว่าเพลงนั้นเป็นเพลงในกลุ่มใดจะต้องมีการจำแนกในสองระดับด้วยกัน ได้แก่ ในระดับเฟรม และในระดับเพลง

ในระดับเฟรมนั้น เฟรมแต่ละเฟรมจะถูกจำแนกว่าเป็นกลุ่มใดจนครบ 2,000 เฟรม จากนั้นจึงหากลุ่มเพลงที่เฟรมทั้ง 2,000 เฟรมถูกจำแนกให้เป็นมากที่สุด หรือกล่าวอีกนัยหนึ่งคือกลุ่มเพลงใดมีเปอร์เซ็นต์ของการจำแนกสูงสุด เพลงนั้นก็จะถูกจำแนกให้เป็นกลุ่มเพลงดังกล่าวที่มีสมาชิกเฟรมมากที่สุด

ในการรายงานผลนั้น ค่าความแม่นยำ (Accuracy) สามารถที่จะรายงานได้เป็นสองระดับเช่นกัน ซึ่งจะมีผลต่างกัน ยกตัวอย่างเช่น หากมีเพลงที่ต้องการจำแนก 1 เพลง สมมติให้เป็นเพลงกลุ่ม ก เมื่อทำตามกระบวนการจำแนกแล้ว ปรากฏว่าจำนวนเฟรมทั้ง 2,000 เฟรมตกเป็นกลุ่ม ก ทั้งสิ้น 1,750 เฟรม คิดเป็นค่าความแม่นยำร้อยละ 87.5 แต่อย่างไรในระดับเพลง เพลง

นั้นก็ถูกจำแนกให้เป็นเพลงกลุ่ม ก ซึ่งคิดเป็นค่าความแม่นยำร้อยละ 100 (ถูก 1 เพลงจาก 1 เพลง)

3.3 เขียนโปรแกรมและทดสอบโปรแกรม

เป็นขั้นตอนการนำอัลกอริธึมที่สร้างขึ้นไปเขียนโปรแกรมด้วยโปรแกรม MATLAB 6.5 แล้วทำการทดสอบว่าอัลกอริธึมที่สร้างขึ้น สามารถใช้ในการจำแนกกลุ่มเพลงไทยได้อย่างมีประสิทธิภาพมากน้อยเพียงใด ถ้าประสิทธิภาพที่ได้ยังไม่ดีเท่าที่ควร ก็กลับไปทำการแก้ไขปรับปรุงอัลกอริธึมที่พัฒนาไว้ในขั้นตอนที่ 3.2 ต่อไปให้มีประสิทธิภาพมากขึ้น หลังจากที่ได้ทดสอบโปรแกรมที่สร้างขึ้นแล้วว่าทำงานได้อย่างมีประสิทธิภาพแล้ว จึงสร้างส่วนติดต่อผู้ใช้เพื่อการสะดวกและใช้งานง่ายของผู้ใช้ทั่วไป

3.4 สรุปผลการทดสอบโปรแกรม

ขั้นตอนนี้เป็นขั้นตอนหลังจากทำการทดสอบแล้วว่าโปรแกรมที่พัฒนาขึ้นมีประสิทธิภาพเพียงพอ จึงทำการสรุปผลการทดสอบโปรแกรม

3.5 เขียนรายงานการศึกษาวิจัย

ขั้นตอนนี้เป็นขั้นตอนการสร้างเอกสารงานวิจัย รวมข้อมูลทั้งประวัติความเป็นมา ทฤษฎีที่เกี่ยวข้อง ขั้นตอนการทำงาน ผลการทดลอง และข้อเสนอแนะ เพื่อไว้เป็นข้อมูลงานวิจัยต่อไป

บทที่ 4

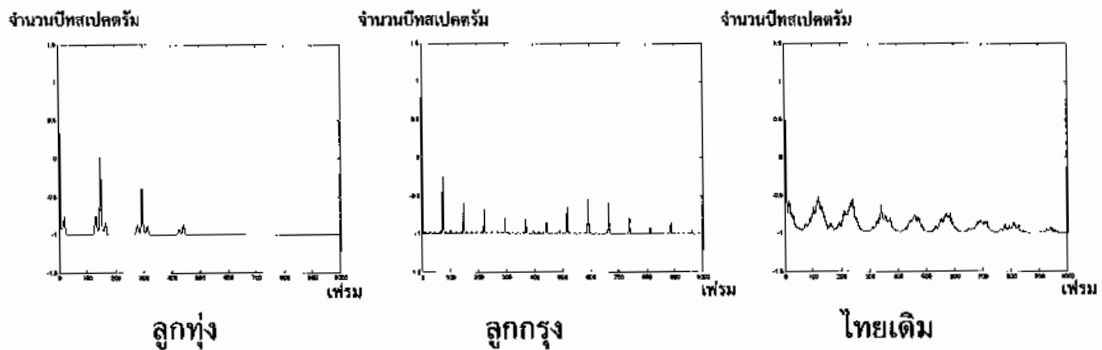
ผลการวิจัย

4.1 ผลการวิจัย

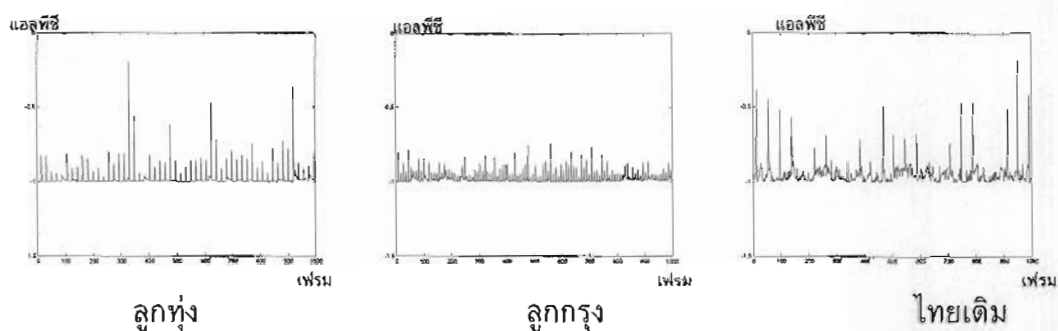
ในการวิจัยครั้งนี้มีขั้นตอนการหาผลของการวิจัยที่สำคัญ ประกอบด้วย 3 ขั้นตอนดังนี้

4.1.1 ผลการทดลองหาค่าคุณลักษณะที่เหมาะสม

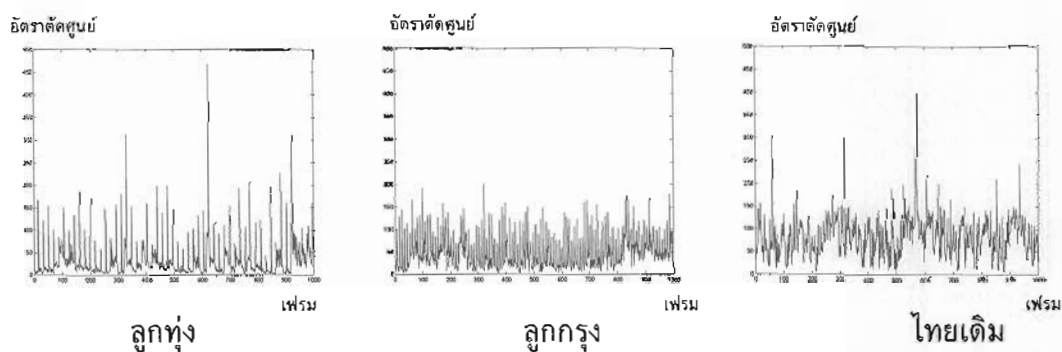
ในการหาค่าคุณลักษณะ (Feature) ชู [1] ได้ใช้วิธีการในการหาค่าคุณลักษณะทั้งสิ้น 5 วิธี ในการจำแนกกลุ่มเพลง 4 กลุ่ม คือ คลาสสิก (Classic) ร็อก (Rock) ป๊อป (Pop) และแจ๊ซ (Jazz) และวิธีที่ใช้หาค่าเหล่านั้นได้แก่ บีทสเปกตรัม (Beat Spectrum), แอลพีซีดีโรฟด์เซปสตรัม (LPC-Derived Cepstrum), อัตราค่าตัดศูนย์ (Zero Crossing Rate), พลังงานสเปกตรัม (Spectrum Power) และสัมประสิทธิ์เซปสตรัลบนความถี่เมล (Mel Frequency Cepstral Coefficient-MFCC) ผู้วิจัยจึงได้นำเอาวิธีการหาค่าคุณลักษณะเหล่านี้มาทดลองหาค่าความถูกต้องในการจำแนกและส่งผลถึงการจัดกลุ่มที่เหมาะสมเพื่อการจำแนกแบบคู่เป็นลำดับ ของกลุ่มเพลงไทยทั้งสามกลุ่ม ได้แก่ เพลงไทยลูกทุ่ง เพลงไทยลูกกรุง และเพลงไทยเดิม ดังภาพที่ 3-2 ในบทที่ 3 ได้นำค่าคุณลักษณะที่ได้จากวิธีการต่างๆ มาวาดกราฟเพื่อดูความแตกต่างของแต่ละวิธีในแต่ละกลุ่มเพลงโดยมีผลดังภาพที่ 4-1 ถึง 4-5



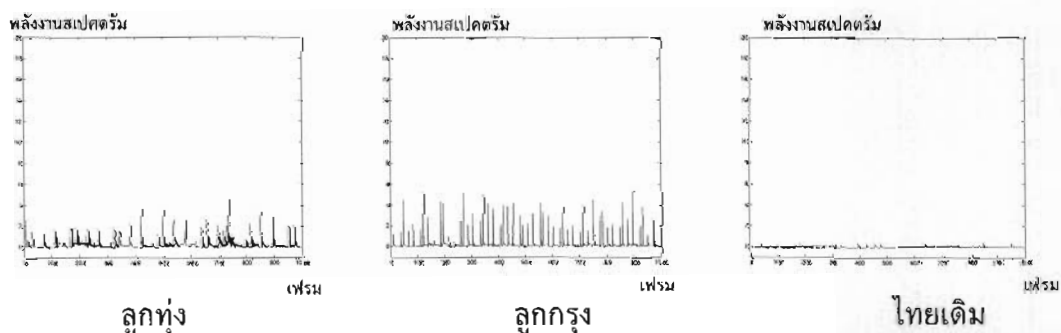
ภาพที่ 4-1 ผลตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีบีทสเปกตรัม (Beat Spectrum)



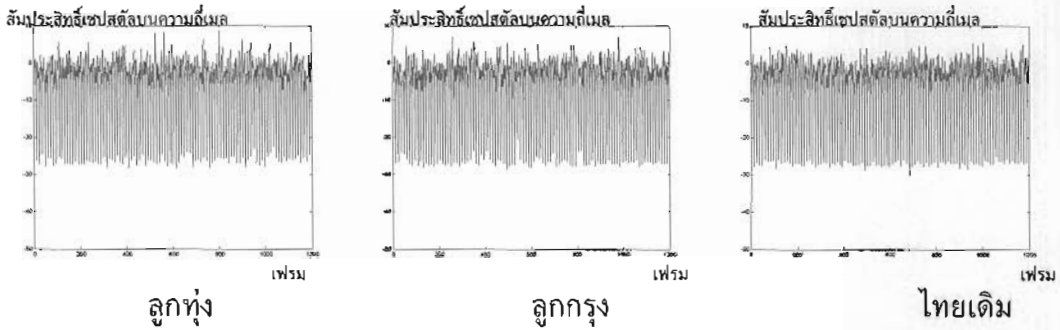
ภาพที่ 4-2 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีแอสพิซีดีไรฟด์ เชปสตรัม (LPC-Derived Cepstrum)



ภาพที่ 4-3 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีอัตราค่าตัดศูนย์ (Zero Crossing Rate)



ภาพที่ 4-4 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีพลังงานสเปกตรัม (Spectrum Power)



ภาพที่ 4-5 ตัวอย่างกราฟค่าคุณลักษณะของทั้งสามกลุ่มเพลงไทยด้วยวิธีสัมประสิทธิ์เซปสตัลบนความถี่เมล (Mel Frequency Cepstral Coefficient-MFCC)

จากกราฟค่าคุณลักษณะของแต่ละวิธีนั้น ได้นำไปหาค่าความถูกต้องของการจำแนกระหว่างกลุ่มเพลงต่าง ๆ โดยใช้เพลงสำหรับชุดเรียนรู้กลุ่มละ 10 เพลง เพลงละ 1,000 เฟรมและกลุ่มทดสอบ 5 เพลง เพลงละ 1,000 เฟรมเช่นเดียวกัน ซึ่งมีผลดังตารางที่ 4-1 ถึง 4-5

ตารางที่ 4-1 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีบีทสเปกตรัม (Beat Spectrum)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	49.1
2	ลูกทุ่งและไทยเดิม	49.7
3	ลูกกรุงและไทยเดิม	49.8

ตารางที่ 4-2 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีแอลพีซีดีไรฟด์เซปสตรัม (LPC-Derived Cepstrum)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	52.4
2	ลูกทุ่งและไทยเดิม	61.2
3	ลูกกรุงและไทยเดิม	59.7

ตารางที่ 4-3 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีอัตราค่าตัดศูนย์ (Zero Crossing Rate)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	51.8

ตารางที่ 4-3 (ต่อ)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
2	ลูกทุ่งและไทยเดิม	52.6
3	ลูกกรุงและไทยเดิม	52.8

ตารางที่ 4-4 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีพลังงานสเปกตรัม (Spectrum Power)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	31.2
2	ลูกทุ่งและไทยเดิม	75.6
3	ลูกกรุงและไทยเดิม	78.3

ตารางที่ 4-5 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีสัมประสิทธิ์เซปสตัลบนความถี่ เมล (Mel Frequency Cepstral Coefficient-MFCC)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	45.3
2	ลูกทุ่งและไทยเดิม	48.4
3	ลูกกรุงและไทยเดิม	47.8

จากตารางที่ 4-1 และภาพที่ 4-1 พบว่าวิธีบีทสเปกตรัม (Beat Spectrum) ไม่มีความสามารถแยกเพลงทั้งสามกลุ่มเพลงได้ดีพอ เนื่องจากในแต่ละกลุ่มมีค่าความถูกต้องที่ต่ำกว่าร้อยละ 50 ทั้งสิ้น โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 49.1 ลูกทุ่งและไทยเดิมได้ร้อยละ 49.7 ลูกกรุงและไทยเดิมได้ร้อยละ 49.8

จากตารางที่ 4-2 และภาพที่ 4-2 พบว่า วิธีแอลพีซีดีไรฟต์เซปสตรัม (LPC-Derived Cepstrum) มีความสามารถแยกเพลงทั้งสามกลุ่มเพลงได้ โดยแต่ละกลุ่มมีค่าความถูกต้องที่เกินกว่าร้อยละ 50 ทั้งสิ้น โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 52.4 ลูกทุ่งและไทยเดิมได้ร้อยละ 61.2 ลูกกรุงและไทยเดิมได้ร้อยละ 59.7

จากตารางที่ 4-3 และภาพที่ 4-3 พบว่า วิธีอัตราค่าตัดศูนย์ (Zero Crossing Rate) มีความสามารถแยกเพลงทั้งสามกลุ่มเพลงได้ โดยแต่ละกลุ่มมีค่าความถูกต้องที่เกินกว่าร้อยละ 50 ทั้งสิ้น โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 51.8 ลูกทุ่งและไทยเดิมได้ร้อยละ 52.6 ลูกกรุงและไทยเดิมได้ร้อยละ 52.8

จากตารางที่ 4-4 และภาพที่ 4-4 พบว่า วิธีพลังงานสเปกตรัม (Spectrum Power) ไม่สามารถแยกเพลงกลุ่มลูกทุ่งและลูกกรุงได้ โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 31.2 ลูกทุ่งและไทยเดิมได้ร้อยละ 75.6 ลูกกรุงและไทยเดิมได้ร้อยละ 78.3

จากตารางที่ 4-5 และภาพที่ 4-5 พบว่า วิธีเมลฟรีควেনซีเซปสตรัลโคเอฟฟิเชียนท์ (Mel Frequency Cepstral Coefficient-MFCC) ไม่มีความสามารถแยกเพลงทั้งสามกลุ่มเพลงได้ดีพอ เนื่องจากในแต่ละกลุ่มมีค่าความถูกต้องที่ต่ำกว่าร้อยละ 50 ทั้งสิ้น โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 45.3 ลูกทุ่งและไทยเดิมได้ร้อยละ 48.4 ลูกกรุงและไทยเดิมได้ร้อยละ 47.8

จากค่าความถูกต้องของการจำแนกทั้งหมด ค่าที่โดดเด่นที่สุดคือค่าคุณลักษณะพลังงานสเปกตรัม (Spectrum Power) ซึ่งสามารถที่จะจำแนกเพลงไทยเดิมออกจากกลุ่มเพลงอื่นได้ โดยมีความถูกต้องเมื่อจำแนกกับเพลงลูกทุ่งและลูกกรุงที่ร้อยละ 75.5 และ 78.3 ตามลำดับ ดังนั้นในขั้นการจำแนกแรกจะเป็นการจำแนกหากกลุ่มเพลงไทยเดิม โดยที่นำกลุ่มเพลงลูกทุ่งและลูกกรุงรวมเป็นกลุ่มเดียวกัน โดยใช้ค่าคุณลักษณะพลังงานสเปกตรัม (Spectrum Power) เป็นตัวหลักในการจำแนก และใช้ค่าคุณลักษณะแอลพีซีดีโรไฟด์เซปสตรัม (LPC-Derived Cepstrum) มาร่วมด้วยเพื่อให้ได้ค่าความถูกต้องที่สูงขึ้น

ในขั้นต่อไปคือการจำแนกกลุ่มเพลงลูกทุ่งและลูกกรุงออกจากกัน โดยเลือกใช้ค่าคุณลักษณะแอลพีซีดีโรไฟด์เซปสตรัม (LPC-Derived Cepstrum) เนื่องจากมีค่าการจำแนกเพลงกลุ่มนี้สูงที่สุดคือร้อยละ 52.4 และอัตราค่าตัดศูนย์ (Zero Crossing Rate) ที่มีค่าความถูกต้องร้อยละ 51.8

เนื่องค่าการจำแนกในขั้นที่สองระหว่างเพลงลูกทุ่งและเพลงลูกกรุงนั้นมีค่าความถูกต้องที่ค่อนข้างต่ำ ผู้วิจัยจึงได้แก้ปัญหาด้วยการหาวิธีหาค่าคุณลักษณะวิธีอื่นที่อยู่นอกเหนือจากวิธีอ้างอิงจากงานวิจัยของ ชู [1] นั่นคือวิธีความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF) ซึ่งมีผลการทดลองดังตารางที่ 4-6

ตารางที่ 4-6 ผลการทดลองการจำแนกกลุ่มเพลงไทยด้วยวิธีความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF)

ครั้งที่	กลุ่มเพลงที่จำแนก	ค่าความถูกต้อง (ร้อยละ)
1	ลูกทุ่งและลูกกรุง	62.1
2	ลูกทุ่งและไทยเดิม	52.3
3	ลูกกรุงและไทยเดิม	51.6

จากตารางที่ 4-6 พบว่า วิธีความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF) มีความสามารถแยกเพลงทั้งสามกลุ่มเพลงได้ดีพอ เนื่องจากในแต่ละกลุ่มมีค่าความถูกต้องที่สูงกว่า

ร้อยละ 50 ทั้งสิ้น โดยค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 62.1 ลูกทุ่งและไทยเดิมได้ร้อยละ 52.3 ลูกกรุงและไทยเดิมได้ร้อยละ 51.6

จากการทดลองการหาค่าคุณลักษณะที่เหมาะสมเพื่อนำมาใช้ในการจำแนก สามารถสรุปได้ว่าการจำแนกชั้นแรกนั้น จะเป็นการจำแนกระหว่างกลุ่มเพลงไทยเดิมกับกลุ่มที่รวมกันของเพลงลูกทุ่งและลูกกรุง โดยใช้ค่าคุณลักษณะพลังงานสเปกตรัม (Spectrum Power) และใช้ค่าคุณลักษณะแอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum) ในขั้นที่สองเป็นการจำแนกระหว่างกลุ่มเพลงลูกทุ่งและกลุ่มเพลงลูกกรุง โดยใช้ค่าคุณลักษณะแอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum) และอัตราค่าตัดศูนย์ (Zero Crossing Rate) และความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF)

4.1.2 ผลการทดลองหาจำนวนชุดการเรียนรู้ที่ดีและเหมาะสม

ในขั้นแรกของการทำวิจัยนั้นจำเป็นจะต้องมีกลุ่มประชากรที่จะใช้ในการวิจัยเสียก่อน ซึ่งงานวิจัยนี้เป็นงานวิจัยการจำแนกกลุ่มเพลงไทยอันได้แก่ กลุ่มเพลงไทยลูกทุ่ง กลุ่มเพลงไทยลูกกรุง และกลุ่มเพลงไทยเดิม โดยข้อมูลทั้งหมดที่มีอยู่ในฐานข้อมูลมีทั้งสิ้น 75 เพลง แบ่งเป็นกลุ่มเพลงไทยลูกทุ่ง 25 เพลง กลุ่มเพลงไทยลูกกรุง 25 เพลงและกลุ่มเพลงไทยเดิม 25 เพลง

4.1.2.1 จำนวนเพลงสำหรับชุดการเรียนรู้ที่ดีและเหมาะสม

เมื่อได้ทำการสะสมข้อมูลไว้ในฐานข้อมูลแล้ว จากนั้นจึงได้ทำการทดลองเพื่อหาจำนวนเพลงที่จะนำมาใช้เป็นชุดการเรียนรู้ที่ดีและเหมาะสมที่สุด โดยทำการทดลองทั้งหมด 4 ครั้งดังนี้

การทดลองครั้งที่ 1 แบ่งเพลงทั้งสิ้นออกเป็น เพลงกลุ่มเรียนรู้ 15 เพลง (กลุ่มเพลงไทยลูกทุ่ง 5 เพลง กลุ่มเพลงไทยลูกกรุง 5 เพลงและกลุ่มเพลงไทยเดิม 5 เพลง) และเพลงกลุ่มทดสอบ 30 เพลง (กลุ่มเพลงไทยลูกทุ่ง 10 เพลง กลุ่มเพลงไทยลูกกรุง 10 เพลงและกลุ่มเพลงไทยเดิม 10 เพลง)

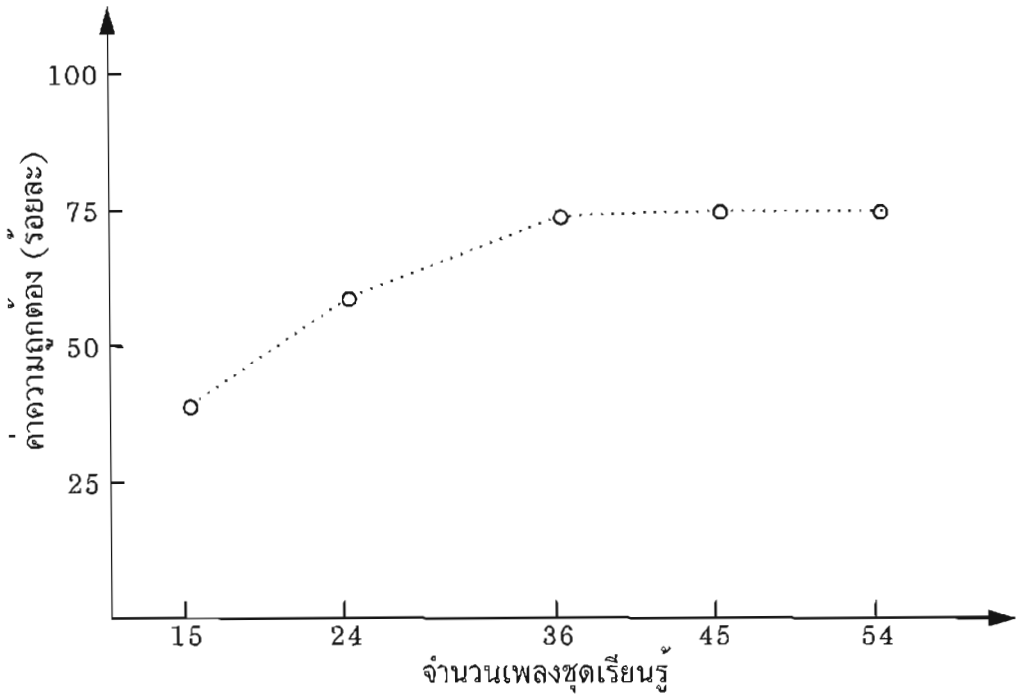
การทดลองครั้งที่ 2 แบ่งเพลงทั้งสิ้นออกเป็น เพลงกลุ่มเรียนรู้ 24 เพลง (กลุ่มเพลงไทยลูกทุ่ง 8 เพลง กลุ่มเพลงไทยลูกกรุง 8 เพลงและกลุ่มเพลงไทยเดิม 8 เพลง) และเพลงกลุ่มทดสอบ 30 เพลง (กลุ่มเพลงไทยลูกทุ่ง 10 เพลง กลุ่มเพลงไทยลูกกรุง 10 เพลงและกลุ่มเพลงไทยเดิม 10 เพลง)

การทดลองครั้งที่ 3 แบ่งเพลงทั้งสิ้นออกเป็น เพลงกลุ่มเรียนรู้ 36 เพลง (กลุ่มเพลงไทยลูกทุ่ง 12 เพลง กลุ่มเพลงไทยลูกกรุง 12 เพลงและกลุ่มเพลงไทยเดิม 12 เพลง) และเพลงกลุ่มทดสอบ 30 เพลง (กลุ่มเพลงไทยลูกทุ่ง 10 เพลง กลุ่มเพลงไทยลูกกรุง 10 เพลงและกลุ่มเพลงไทยเดิม 10 เพลง)

การทดลองครั้งที่ 4 แบ่งเพลงทั้งสิ้นออกเป็น เพลงกลุ่มเรียนรู้ 45 เพลง (กลุ่มเพลงไทยลูกทุ่ง 15 เพลง กลุ่มเพลงไทยลูกกรุง 15 เพลงและกลุ่มเพลงไทยเดิม 15 เพลง) และเพลงกลุ่มทดสอบ 30 เพลง (กลุ่มเพลงไทยลูกทุ่ง 10 เพลง กลุ่มเพลงไทยลูกกรุง 10 เพลงและกลุ่มเพลงไทยเดิม 10 เพลง)

ตารางที่ 4-7 ผลการทดลองหาจำนวนชุดเรียนรู้ที่ดีและเหมาะสม

ครั้งที่	ชุดเรียนรู้				ชุดทดสอบ				ความถูกต้อง (ร้อยละ)
	ลูกทุ่ง	ลูกกรุง	ไทยเดิม	รวม	ลูกทุ่ง	ลูกกรุง	ไทยเดิม	รวม	
1	5	5	5	15	10	10	10	30	40.44
2	8	8	8	24	10	10	10	30	60.2
3	12	12	12	36	10	10	10	30	75.23
4	15	15	15	45	10	10	10	30	76.06
5	18	18	18	54	10	10	10	30	76



ภาพที่ 4-6 กราฟแสดงผลการทดลองหาจำนวนเพลงชุดการเรียนรู้

จากตารางที่ 4-7 พบว่าในการทดลองครั้งที่ 1 ได้ใช้เพลงในการเรียนรู้ 15 เพลงและใช้ 30 เพลงในการทดสอบ และได้ค่าความถูกต้องที่ร้อยละ 40.44 และการทดลองครั้งที่ 2 ได้ใช้เพลงในการเรียนรู้ 24 เพลงและใช้ 30 เพลงในการทดสอบ และได้ค่าความถูกต้องที่ร้อยละ 60.2 ในการทดลองครั้งที่ 3 ได้ใช้เพลงในการเรียนรู้ 36 เพลงและใช้ 30 เพลงในการทดสอบ และได้ค่าความถูกต้องที่ร้อยละ 75.23 ในการทดลองครั้งที่ 4 ได้ใช้เพลงในการเรียนรู้ 45 เพลงและใช้ 30 เพลงในการทดสอบ และได้ค่าความถูกต้องที่ร้อยละ 76.06 และในการทดลองครั้งที่ 5 ได้ใช้เพลงในการเรียนรู้ 54 เพลงและใช้ 30 เพลงในการทดสอบ และได้ค่าความถูกต้องที่ร้อยละ 76 และเมื่อนำค่าความถูกต้องจากตารางที่ 4-7 ไปวาดกราฟจะได้กราฟดังภาพที่ 4-6

เมื่อสังเกตจากภาพที่ 4-6 ก็จะพบว่าค่าความถูกต้องเริ่มชะลอตัวที่จำนวนเพลงระหว่าง 36 เพลง และ 54 เพลง จำนวนเพลงที่ให้ค่าสูงสุดคือ 45 เพลง แสดงว่าถ้าหากทำการเพิ่มจำนวนเพลงของชุดการเรียนรู้ก็อาจจะไม่ได้ผลที่ต่างไปจากนี้ หรืออาจเพิ่มขึ้นแต่ค่าที่ได้ไม่มีนัยสำคัญมากพอเมื่อเทียบกับจำนวนเพลงที่เพิ่มขึ้น

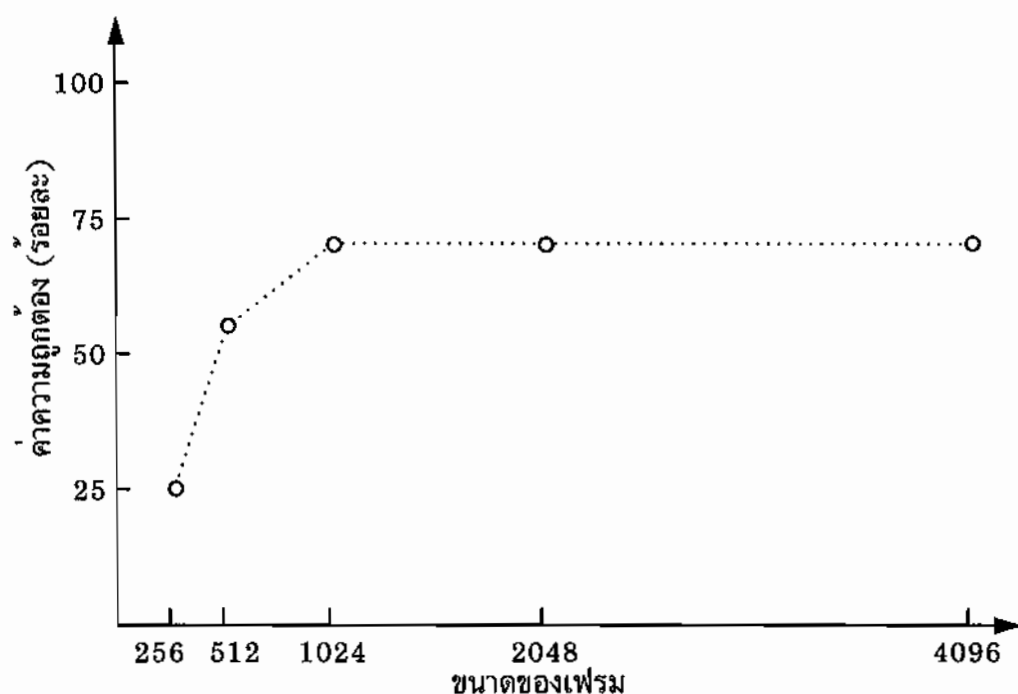
ดังนั้นในการทำวิจัยในครั้งนี้จึงเลือกใช้จำนวนเพลง 45 เพลงสำหรับใช้เป็นชุดการเรียนรู้ และ 30 เพลงสำหรับใช้เป็นชุดทดสอบ

4.1.2.2 จำนวนเฟรมและขนาดเฟรมที่เหมาะสม

เมื่อได้จำนวนเพลงของชุดเรียนรู้ที่เหมาะสมแล้ว ขั้นตอนต่อไปจะต้องนำเพลงเหล่านั้นมาตัดเป็นเฟรม โดยแต่ละเฟรมนั้นมีขนาดที่เท่ากัน ดังขั้นตอนของการเตรียมข้อมูลที่ได้กล่าวไปแล้ว ใน บทที่ 3 ในงานวิจัยนี้ได้ยึดค่าอัตราส่วนระหว่างความกว้างของเฟรมต่อระดับความถี่ของเสียง โดย ฟุต [2] ได้ใช้ที่ความกว้าง 256 จุดต่อหนึ่งเฟรมและที่ระดับความถี่ 22 กิโลเฮิรท์ ดังนั้นความกว้างของเฟรมคิดเป็นเวลาประมาณ 11 มิลลิวินาที (11.63 มิลลิวินาที) และเมื่อนำค่าเวลาความกว้างของเฟรมตามงานวิจัยดังกล่าวมาเป็นหลักของการเริ่มต้นหาค่าความกว้างเฟรมในงานวิจัยนี้ ซึ่งข้อมูลเพลงมีระดับความถี่ที่ 44 กิโลเฮิรท์ ก็จะได้ที่ความกว้าง 512 จุด และได้ทำการทดลองหาความกว้างที่เหมาะสมของเฟรมทั้งหมด 5 ระดับดังตารางที่ 4-8 โดยในแต่ละครั้งของการทดลองมีจำนวนทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง

ตารางที่ 4-8 ผลการทดลองหาค่าความถูกต้องของขนาดเฟรมต่าง ๆ

ครั้งที่	ความกว้างของเฟรม	ค่าความถูกต้อง (ร้อยละ)
1	256	26.3
2	512	56.4
3	1,024	71.5
4	2,048	71.6
5	4,096	71.8



ภาพที่ 4-7 กราฟแสดงผลการทดลองการหาขนาดของเฟรม

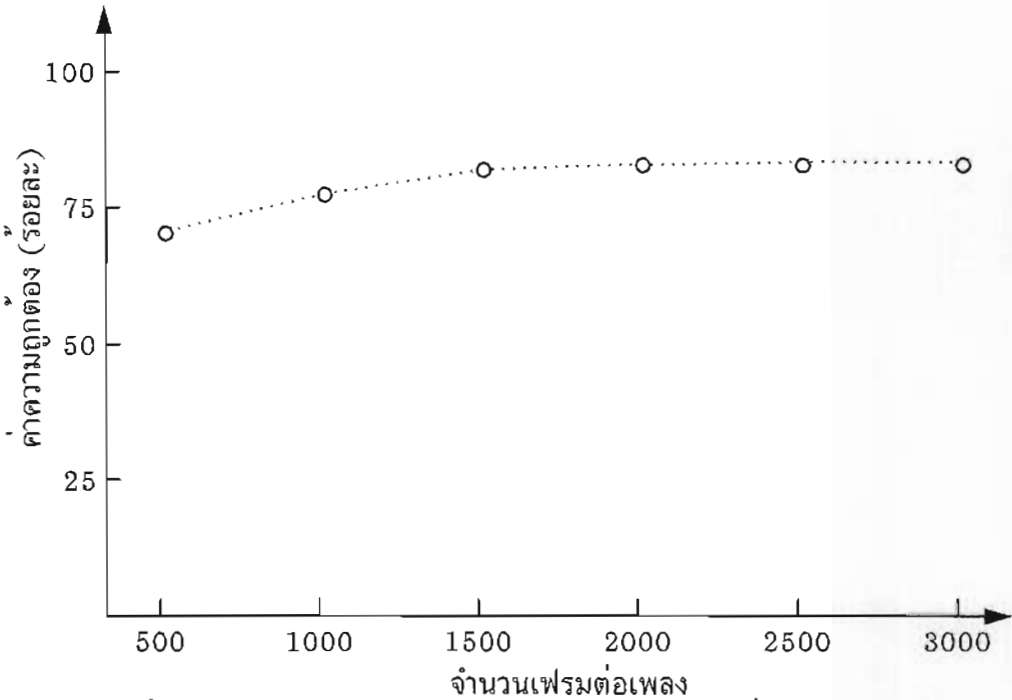
จากตารางที่ 4-8 พบว่า การจำแนกเพลงไทยที่ระดับความถี่ที่ 44 กิโลเฮิรตซ์ ที่ขนาดเฟรม 256 จุดให้ค่าความถูกต้องที่ร้อยละ 26.3 ที่ขนาดเฟรม 512 จุดให้ค่าความถูกต้องที่ร้อยละ 56.4 ที่ขนาดเฟรม 1,024 จุดให้ค่าความถูกต้องที่ร้อยละ 71.5 ที่ขนาดเฟรม 2,048 จุดให้ค่าความถูกต้องที่ร้อยละ 71.6 และที่ขนาดเฟรม 4,096 จุดให้ค่าความถูกต้องที่ร้อยละ 71.8 และเมื่อนำค่าจากตารางที่ 4-8 มาเขียนกราฟ ก็จะได้ดังภาพที่ 4-7 และสังเกตได้ว่าค่าความถูกต้องเริ่มชะลอตัวที่ค่าความกว้างของเฟรมตั้งแต่ 1,024 จุดเป็นต้นไป ถึงแม้ว่าที่ขนาดเฟรมที่สูงกว่าขนาด 1,024 จุด จะมีค่าความถูกต้องที่สูงกว่า แต่เมื่อเทียบกับขนาดของเฟรมที่เพิ่มขึ้นแบบเท่าตัวกับค่าความถูกต้องที่เพิ่มขึ้นนั้น ค่าที่ได้กลับไม่มากขึ้นเป็นเท่าตัวตามขนาดของเฟรม

ดังนั้นในงานวิจัยนี้ขนาดของเฟรมที่เลือกใช้คือ 1,024 จุด เนื่องจากให้ค่าความถูกต้องอยู่ในช่วงสูง และจะเป็นขนาดที่ประหยัดเวลาและทรัพยากรในการประมวลผลที่สุด

เมื่อได้ค่าขนาดของเฟรมที่เหมาะสมแล้ว จึงนำค่ามาหาจำนวนของเฟรมต่อหนึ่งเพลง โดยชู [1] ใช้เฟรมเป็นจำนวน 2,000 เฟรม ที่ขนาด 882 จุด ดังนั้นจึงใช้เป็นข้อมูลอ้างอิงข้อมูลหนึ่งถึงจำนวนเฟรมที่จะนำมาทำการทดลอง โดยทำการทดลอง 6 ครั้งด้วยกัน ดังตารางที่ 4-9

ตารางที่ 4-9 ผลการทดลองหาค่าความถูกต้องของจำนวนเฟรม

ครั้งที่	จำนวนเฟรม	ค่าความถูกต้อง (ร้อยละ)
1	500	71.5
2	1,000	78.6
3	1,500	83.2
4	2,000	84.1
5	2,500	84.1
6	3,000	84.2



ภาพที่ 4-8 กราฟแสดงผลการทดลองการหาจำนวนเฟรมที่ดีและเหมาะสม

จากตารางที่ 4-9 พบว่า การจำแนกที่จำนวนเฟรม 500 เฟรมให้ค่าความถูกต้องที่ร้อยละ 71.5 การจำแนกที่จำนวนเฟรม 1,000 เฟรมให้ค่าความถูกต้องที่ร้อยละ 78.6 การจำแนกที่จำนวนเฟรม 1,500 เฟรมให้ค่าความถูกต้องที่ร้อยละ 83.2 การจำแนกที่จำนวนเฟรม 2,000 เฟรมและ 2,500 เฟรมให้ค่าความถูกต้องที่ร้อยละ 84.1 และการจำแนกที่จำนวนเฟรม 3,000 เฟรมให้ค่าความถูกต้องที่ร้อยละ 84.2 จากนั้นจึงนำค่าจากตารางที่ 4-9 มาวาดกราฟ ได้ดังภาพที่ 4-8 ค่าความถูกต้องได้ชะลอตัวตั้งแต่จำนวนเฟรมที่ 2,000 เป็นต้นไป และเพิ่มขึ้นอีกร้อยละ 0.1 ที่จำนวนเฟรม 3,000 เฟรม

ในการเลือกจำนวนเฟรมที่ดีที่สุดนั้น ได้คำนึงถึงระยะเวลาและทรัพยากรระบบที่จะใช้ในการประมวลผลในการจำแนกด้วย เมื่อเปรียบเทียบค่าความถูกต้องระหว่างจำนวนเฟรมที่ 2,000

เฟรมและ 3,000 เฟรมแล้ว ค่าที่เพิ่มขึ้นร้อยละ 0.1 นั้นมีประโยชน์น้อยกว่าจำนวนเฟรมที่เพิ่มขึ้นจำนวนมากและต้องใช้หน่วยความจำเพิ่มขึ้นอย่างมากในการประมวลผลด้วยเช่นกัน ดังนั้นจำนวนเฟรมที่ดีและเหมาะสมที่สุดที่เลือกใช้คือ 2,000 เฟรม

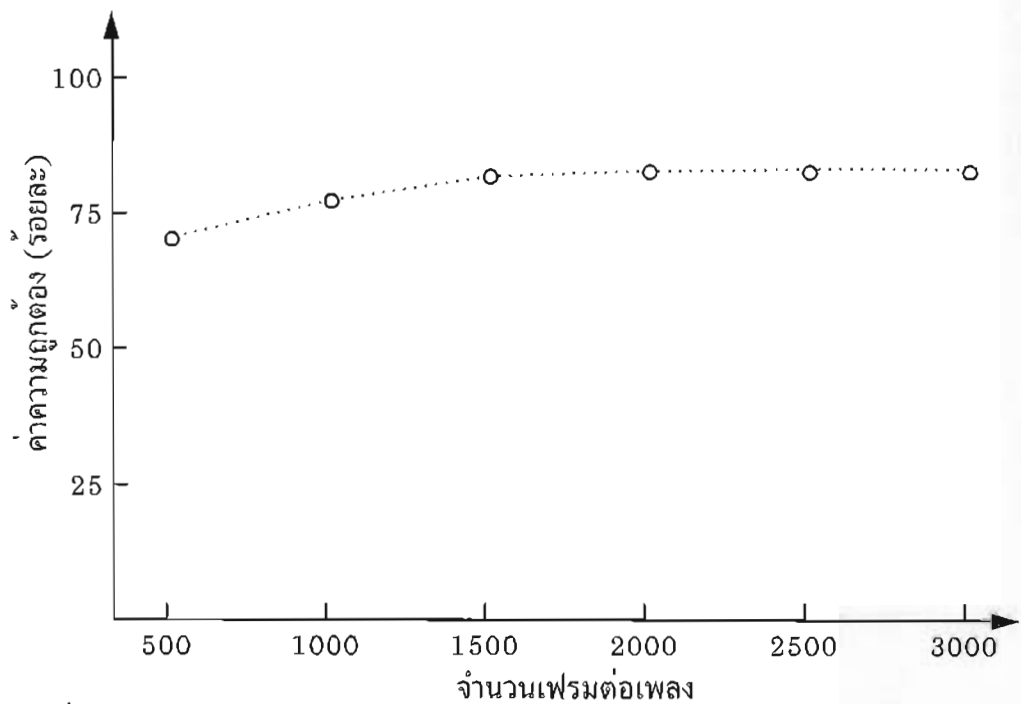
เมื่อนำข้อมูลจากการทดลองทั้งหมดของการหาค่าตัวแปรต่างๆ ของชุดการเรียนรู้นั้น สรุปได้ว่า จำนวนเพลงของชุดเรียนรู้ที่ดีและเหมาะสมที่สุดคือ 45 เพลง หรือกลุ่มละ 15 เพลง ประกอบด้วยกลุ่มเพลงไทยลูกทุ่ง 15 เพลง กลุ่มเพลงไทยลูกกรุง 15 เพลง และกลุ่มเพลงไทยเดิม 15 เพลง โดยแต่ละเพลงจะนำมาแบ่งเป็นเฟรมจำนวน 2,000 เฟรม ซึ่งมีความยาวเฟรมละ 1,024 จุด

4.1.3 ผลการทดลองหาเคอร์เนล

เมื่อได้วิธีหาคูณลักษณะและจำนวนและขนาดของเฟรมที่เหมาะสมจะนำมาจำแนกแล้ว ใน การจำแนกนั้น จะมีการใช้เคอร์เนลเป็นตัวจัดการกับค่าคุณลักษณะเหล่านั้น เพื่อให้การจำแนกมี ประสิทธิภาพมากขึ้น โดยเคอร์เนลที่นำมาใช้นั้นมี 4 วิธีด้วยกันคือ เชิงเส้น (Linear), โพลีโนเมียล (Polynomial), เรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) และ ซิกมอยด์ (Sigmoid) โดยแต่ละวิธีมีค่าพารามิเตอร์ที่ต่างกัน จึงได้ทำการทดลองเพื่อหาเคอร์เนลที่ เหมาะสมกับการจำแนกที่สุด

ตารางที่ 4-10 ผลของการใช้เคอร์เนลแบบเชิงเส้น (Linear)

พารามิเตอร์ C	จำแนกขั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
0.1	39.6	22.3
1	44.3	26.5
10	45.2	27
100	46.2	27.3
1,000	46.1	27.3



ภาพที่ 4-9 กราฟแสดงผลการทดลองค่าความถูกต้องของเคอร์เนลแบบเชิงเส้น (Linear)

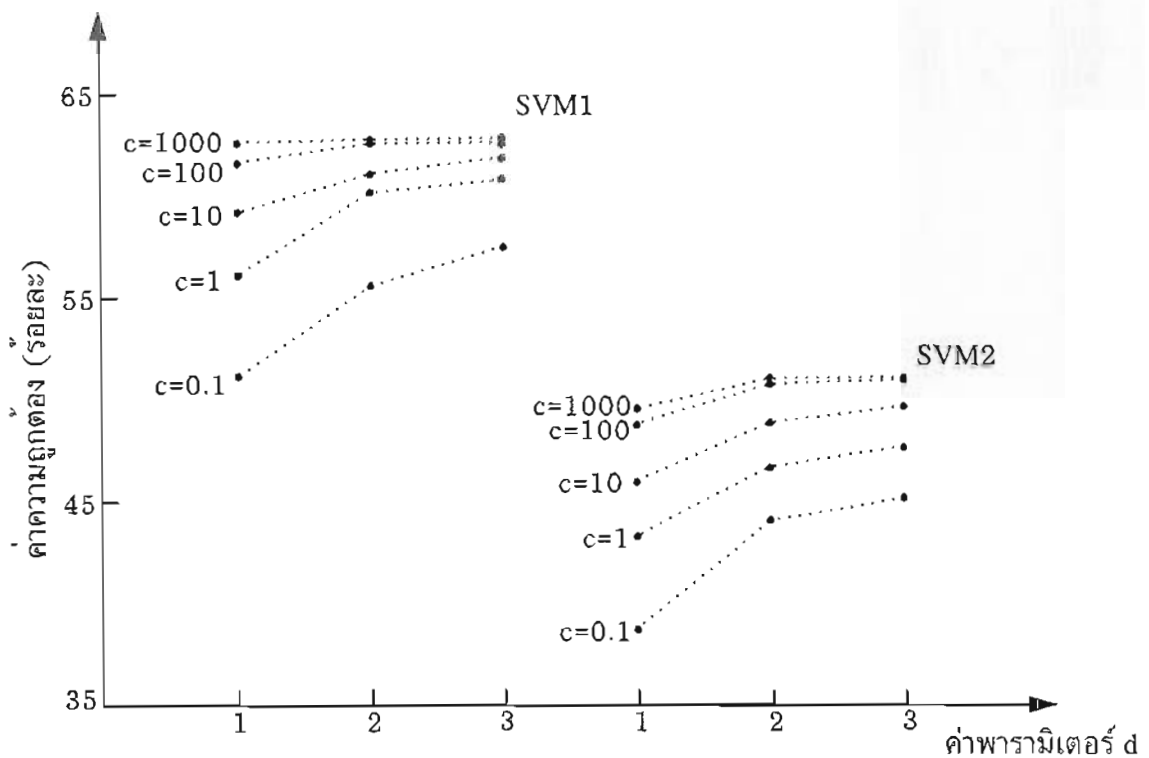
ค่าจากตารางที่ 4-10 ได้นำมาเขียนกราฟได้ดังภาพที่ 4-9 พบว่า ในการจำแนกชั้นแรก (SVM1) ค่าพารามิเตอร์ C ที่ 100 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 46.2 และในการจำแนกชั้นที่ 2 (SVM2) ค่าพารามิเตอร์ C ที่ 100 และ 1,000 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 46.2

ตารางที่ 4-11 ผลของการใช้เคอร์เนลแบบโพลิโนเมียล (Polynomial)

พารามิเตอร์ C	D (Degree)	จำแนกชั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
0.1	1	51.4	38.9
	2	55.8	44.3
	3	57.7	45.4
1	1	56.3	43.5
	2	60.4	46.9
	3	61.1	47.9
10	1	59.4	46.2
	2	61.3	49.1

ตารางที่ 4-11 (ต่อ)

พารามิเตอร์ C	D (Degree)	จำแนกชั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
10	3	62.1	49.9
100	1	61.8	49
	2	62.8	51
	3	62.8	51.2
1,000	1	62.8	49.8
	2	63	51.3
	3	63.1	51.4

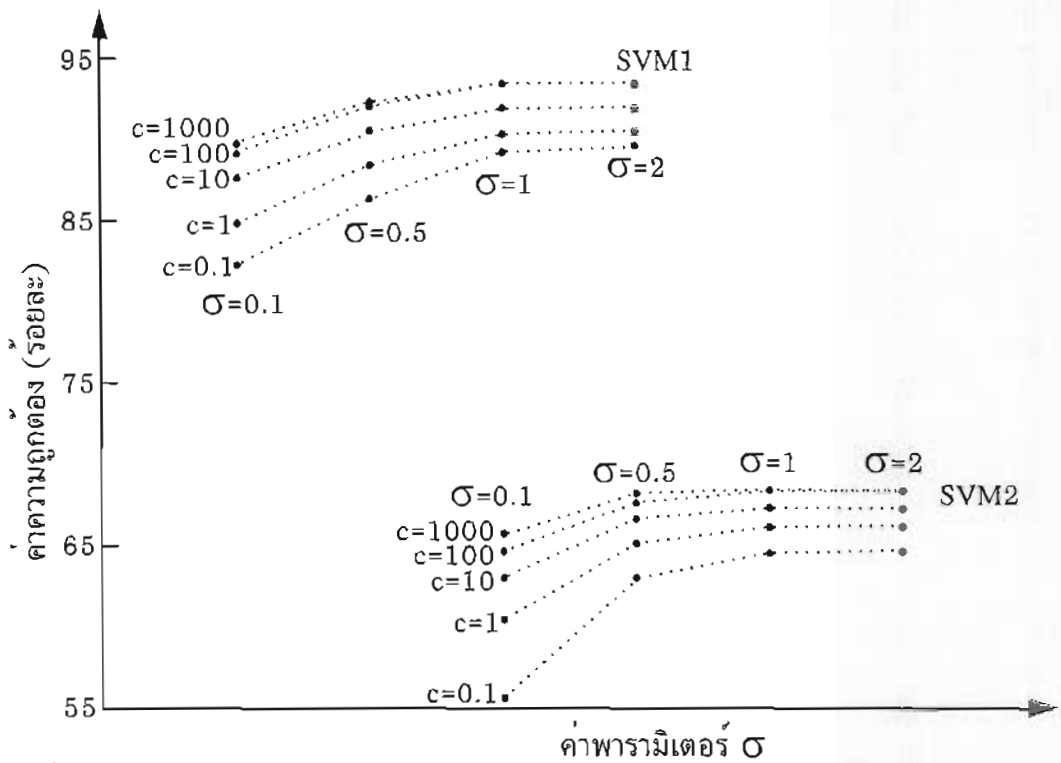


ภาพที่ 4-10 กราฟแสดงผลการทดลองใช้เคอร์เนลแบบโพลิโนเมียล (Polynomial)

ค่าจากตารางที่ 4-11 ได้นำมาเขียนกราฟได้ดังภาพที่ 4-10 พบว่า ในการจำแนกชั้นแรก (SVM1) ค่าพารามิเตอร์ C ที่ 1,000 และค่าดีกรี (D-Degree) ที่ 3 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 63.1 และในการจำแนกชั้นที่สอง (SVM2) ค่าพารามิเตอร์ C ที่ 1,000 และค่าดีกรี (D-Degree) ที่ 3 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 51.4

ตารางที่ 4-12 ผลของการใช้เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF)

พารามิเตอร์ C	σ (Sigma)	จำแนกชั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
0.1	0.1	82.5	55.9
	0.5	86.6	63.3
	1	89.5	64.8
	2	89.9	65
1	0.1	85.1	60.7
	0.5	88.7	65.4
	1	90.6	66.4
	2	90.8	66.5
10	0.1	87.9	63.8
	0.5	90.8	66.9
	1	92.3	67.6
	2	92.3	67.6
100	0.1	89.4	64.9
	0.5	92.3	67.9
	1	93.7	68.7
	2	93.7	68.7
1,000	0.1	90	66
	0.5	92.6	68.5
	1	93.7	68.7
	2	93.7	68.7



ภาพที่ 4-11 กราฟแสดงผลการทดลองใช้เคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF)

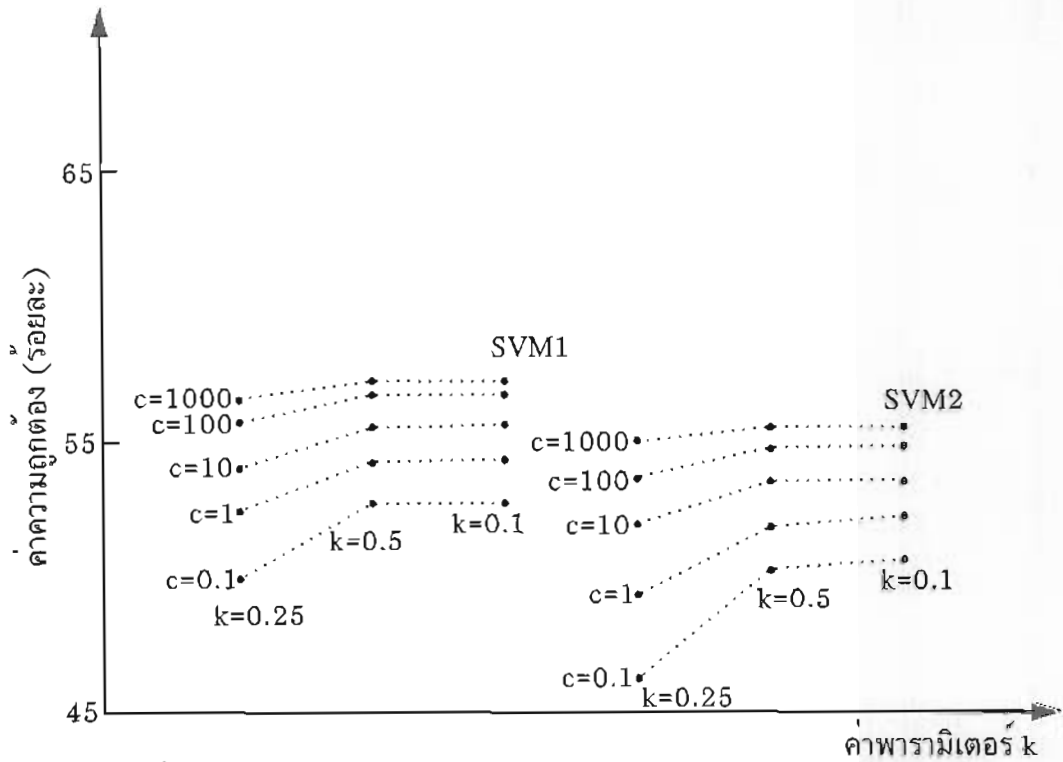
ค่าจากตารางที่ 4-12 ได้นำมาเขียนกราฟได้ดังภาพที่ 4-11 พบว่า ในการจำแนกชั้นแรก (SVM1) ค่าพารามิเตอร์ C ที่ 100 และ 1,000 และค่า σ (Sigma) ที่ 1 และ 2 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 93.7 และในการจำแนกชั้นที่สอง (SVM2) ค่าพารามิเตอร์ C ที่ 100 และ 1,000 และค่า σ (Sigma) ที่ 1 และ 2 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 68.7

ตารางที่ 4-13 ผลของ การใช้เคอร์เนลแบบซิกมอยด์ (Sigmoid)

พารามิเตอร์ C	k	จำแนกชั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
0.1	0.25	50.1	46.4
	0.5	52.9	50.4
	1	52.9	50.8
1	0.25	52.6	49.5
	0.5	54.4	52.1
	1	54.5	52.4

ตารางที่ 4-13 (ต่อ)

พารามิเตอร์ C	k	จำแนกชั้นที่ 1 (SVM1)	จำแนกครั้งที่ 2 (SVM2)
10	0.25	54.2	52.1
	0.5	55.7	53.7
	1	55.8	54.2
100	0.25	55.9	53.8
	0.5	56.2	54.9
	1	56.2	55
1,000	0.25	56.7	55.2
	0.5	57.4	55.7
	1	57.4	55.7



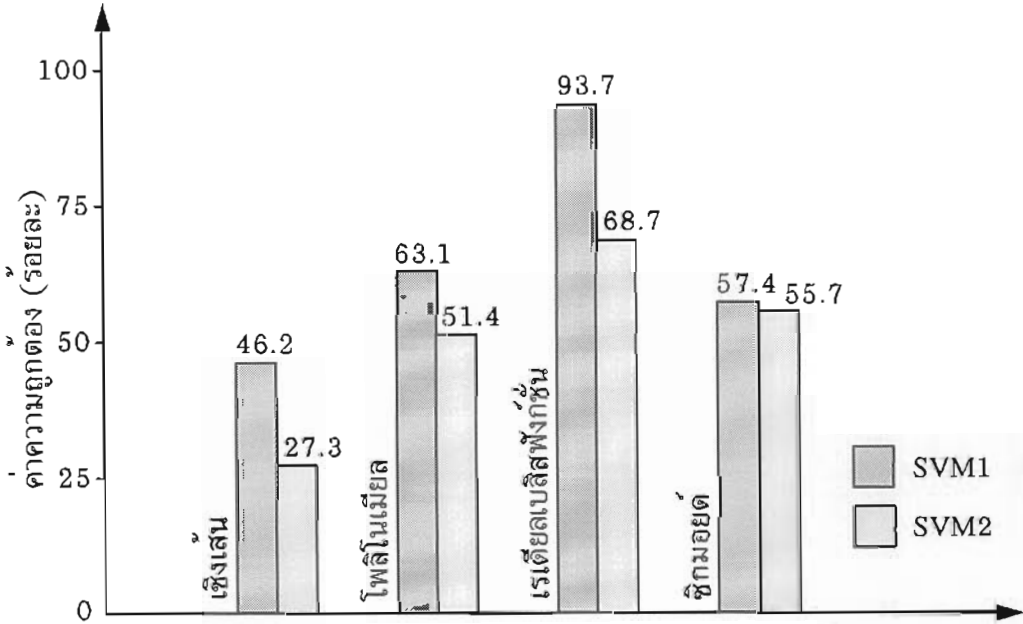
ภาพที่ 4-12 กราฟแสดงผลการทดลองใช้เคอร์เนลแบบซิกมอยด์ (Sigmoid)

ค่าจากตารางที่ 4-13 ได้นำมาเขียนกราฟได้ดังภาพที่ 4-12 พบว่า ในการจำแนกชั้นแรก (SVM1) ค่าพารามิเตอร์ C ที่ 1,000 และค่า k ที่ 0.5 และ 1 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 57.4 และในการจำแนกชั้นที่สอง (SVM2) ค่าพารามิเตอร์ C ที่ 1,000 และค่า k ที่ 0.5 และ 1 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 55.7

เพื่อให้เกิดการเปรียบเทียบที่ชัดเจนของค่าความถูกต้องของแต่ละเคอร์เนล ผู้วิจัยจึงได้ทำการเลือกเอาเฉพาะค่าที่เลือกใช้ที่ให้ค่าความถูกต้องสูงที่สุดของแต่ละเคอร์เนลมาเขียนเป็นตารางและภาพใหม่ โดยค่าที่เลือกหากในแต่ละเคอร์เนลมีพารามิเตอร์ที่ให้ค่าความถูกต้องซ้ำกันค่าพารามิเตอร์ที่น้อยกว่าจะถูกเลือก เนื่องจากเป็นค่าที่ใช้เวลาประมวลผลและทรัพยากรน้อยและดีที่สุด ซึ่งเปรียบเทียบได้ดังตารางที่ 4-14

ตารางที่ 4-14 เปรียบเทียบค่าที่ดีที่สุดของแต่ละเคอร์เนล

เคอร์เนล	จำแนกชั้นที่ 1 (SVM1)		จำแนกชั้นที่ 2 (SVM2)	
	พารามิเตอร์	ความถูกต้อง	พารามิเตอร์	ความถูกต้อง
เชิงเส้น	C=100	46.2	C=100	27.3
โพลีโนเมียล	C=1,000, D=3	63.1	C=1,000, D=3	51.4
เรเดียลเบสิสฟังก์ชัน	C=100, σ=1	93.7	C=100, σ=1	68.7
ซิกมอยด์	C=1,000, k=0.5	57.4	C=1,000, k=0.5	55.7



ภาพที่ 4-13 การเปรียบเทียบค่าที่ดีที่สุดของแต่ละเคอร์เนล

จากค่าที่อ่านได้จากตารางที่ 4-14 นำมาเขียนแผนภูมิแท่งเปรียบเทียบได้ดังภาพที่ 4-13 หากพิจารณาแล้วค่าความถูกต้องสูงสุดในบางเคอร์เนลนั้นมีจำนวนที่ซ้ำกัน ในค่าพารามิเตอร์อื่น ดังนั้นจึงต้องมีการเลือกเพียงเคอร์เนลและค่าพารามิเตอร์เดียว ซึ่งจะพิจารณาจากค่าพารามิเตอร์ที่น้อยกว่าเนื่องจากจะใช้เวลาในการประมวลผลที่สั้นกว่าค่าพารามิเตอร์ที่สูงกว่า ดังนั้นค่าความถูกต้องที่สูงที่สุดของการจำแนกครั้งแรก (SVM1) คือเคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial

Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และ 1,000 และค่า σ (Sigma) ที่ 1 และ 2 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 93.7 เพราะฉะนั้นเพื่อให้เป็นไปตามดังที่กล่าวไปแล้วคือเพื่อการประหยัดเวลาในการประมวลผล ค่าที่เลือกใช้คือ พารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1

ในการจำแนกครั้งที่สอง (SVM2) ค่าความถูกต้องที่สูงที่สุดคือ เคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และ 1,000 และค่า σ (Sigma) ที่ 1 และ 2 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 68.7 เพราะฉะนั้นเพื่อให้เป็นไปตามดังที่กล่าวไปแล้วคือเพื่อการประหยัดเวลาในการประมวลผล ค่าที่เลือกใช้คือ พารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1

ในการวัดค่าความถูกต้องนั้นมีการวัดกันใน 2 ระดับคือ ระดับเฟรม และระดับเพลง ในระดับเฟรมนั้นจะคิดจากจำนวนเฟรมทั้งหมดของชุดทดสอบ ซึ่งในการรายงานผลข้างต้นเป็นการรายงานผลในระดับเฟรมทั้งสิ้น และเพื่อให้เป็นไปตามวัตถุประสงค์ของการวิจัยนี้ คือการจำแนกกลุ่มเพลงไทย ดังนั้นจึงต้องมีการสรุปค่าความถูกต้องในระดับเพลง โดยตัดสินจากจำนวนเฟรม 2,000 เฟรมที่ได้ตัดออกมาจากหนึ่งเพลงนั้นมีประชากรเฟรมข้างมากถูกจำแนกว่าเป็นกลุ่มเพลงใด เพลงนั้นก็จะถูก ตัดสินให้เป็นกลุ่มเพลงนั้นทันที และจากการทดสอบระดับเพลงมีค่าความถูกต้องดังต่อไปนี้

ตารางที่ 4-15 ค่าความถูกต้องระดับเฟรมและเพลงจากกลุ่มทดสอบ

กลุ่มเพลง	ระดับเฟรม			ระดับเพลง		
	ทดสอบ/ เฟรม	ถูกต้อง	ร้อยละ	ทดสอบ/เพลง	ถูกต้อง	ร้อยละ
ไทยลูกทุ่ง	20,000	11,484	57.42	10	7	70
ไทยลูกกรุง	20,000	14,515	72.575	10	9	90
ไทยเดิม	20,000	19,295	96.475	10	10	100
รวม	60,000	45,294	75.49	30	26	86.7

จากตารางที่ 4-15 เพลงไทยลูกทุ่งมีค่าความถูกต้องระดับเฟรมที่ร้อยละ 57.42 และระดับเพลงที่ร้อยละ 70 เพลงไทยลูกกรุงมีค่าความถูกต้องระดับเฟรมที่ร้อยละ 72.575 และระดับเพลงที่ร้อยละ 90 และเพลงไทยเดิมมีค่าความถูกต้องระดับเฟรมที่ร้อยละ 96.475 และระดับเพลงที่ร้อยละ 100

และเมื่อมองค่าความถูกต้องแบบรวมทั้งระบบซึ่งเป็นดัชนีบอกความสามารถของโปรแกรมหรือระบบพบว่า มีค่าความถูกต้องระดับเฟรมที่ร้อยละ 75.49 และระดับเพลงที่ร้อยละ 86.7

4.2 วิเคราะห์ผลการวิจัย

ในขั้นตอนแรกคือการหาวิธีการหาค่าคุณลักษณะที่เหมาะสม และการหาวิธีการที่เหมาะสมนั้นยังช่วยให้สามารถรู้ได้ว่าการจำแนกชั้นที่แรกนั้นจะนำเพลงกลุ่มใดมารวมกันเพื่อให้เหลือเพลงเพียงสองกลุ่มในการจำแนก โดยเมื่อวิเคราะห์จากภาพที่ 4-4 และตารางที่ 4-4 แล้วพบว่าวิธีการหาค่าคุณลักษณะพลังงานสเปกตรัม (Spectrum Power) สามารถแยกกลุ่มเพลงไทยเดิมออกจากกลุ่มเพลงอื่นได้เป็นอย่างดี และอีกวิธีที่มีผลลงลงมาคือวิธีการแอลพีซีดีโรฟด์เซปสตรัม (LPC-Derived Cepstrum) ซึ่งนำมาใช้ร่วมในการจำแนกครั้งที่หนึ่งระหว่างกลุ่มเพลงไทยเดิมและกลุ่มการรวมกันของไทยลูกทุ่งและลูกกรุง

การจำแนกในขั้นที่สอง ระหว่างกลุ่มเพลงไทยลูกทุ่งและเพลงไทยลูกกรุง เลือกใช้วิธีการแอลพีซีดีโรฟด์เซปสตรัม (LPC-Derived Cepstrum) และ ความถี่เส้นสเปกตรัม (Line Spectral Frequency-LSF) เนื่องจากเป็นสองวิธีแรกที่ทำให้ค่าความถูกต้องในการจำแนกกลุ่มเพลงไทยลูกทุ่งและเพลงไทยลูกกรุงได้สูงสุด

ในขั้นต่อมาคือการหาจำนวนของชุดเรียนรู้ที่เหมาะสม ซึ่งได้แก่จำนวนเพลง และจำนวนเฟรมและความกว้างของเฟรม โดยได้ใช้วิธีการหาค่าคุณลักษณะตามได้ทดลองไปแล้ว โดยค่าของจำนวนเพลงในแต่ละกลุ่มที่ให้ค่าความถูกต้องสูงที่สุดคือ 15 เพลงรวมทั้งสามกลุ่มคือ 45 เพลง และใช้จำนวนเฟรมทั้งสิ้น 2,000 เฟรมที่ความกว้าง 1,024 จุด ในการทดลองหาจำนวนเพลงที่เหมาะสม พบว่าเมื่อเพิ่มจำนวนชุดการเรียนรู้ ค่าที่ได้จะสูงขึ้นเรื่อยๆ และพบว่าอัตราเร่งของค่าความถูกต้องนั้นเริ่มชะลอตัวเมื่อเพิ่มจำนวนขึ้นไป

จากนั้นจึงนำค่าที่เหมาะสมที่ได้ทำการทดลองไปแล้ว นำไปใช้กับค่าเคอร์เนลต่างๆ เพื่อหาเคอร์เนลและค่าพารามิเตอร์ภายในที่ให้ค่าความถูกต้องสูงที่สุด โดยการจำแนกครั้งแรกใช้เคอร์เนล เรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1 ในการจำแนกครั้งที่สอง (SVM2) ค่าความถูกต้องที่สูงที่สุดคือ เคอร์เนล เรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1 เช่นเดียวกัน

บทที่ 5

บทสรุปและข้อเสนอแนะ

5.1 อภิปรายผลการวิจัย

จากผลการวิจัย ผลที่ได้ถือว่าอยู่ในค่าที่สูง แต่ถ้าหากดูตามผังของการจำแนกแล้ว จะมีการจำแนกสองชั้นด้วยกัน คือ ชั้นแรก เป็นการจำแนกระหว่างกลุ่มเพลงไทยเดิมและกลุ่มการรวมกันของเพลงลูกทุ่งและเพลงลูกกรุง ซึ่งค่าความถูกต้องที่ได้้นั้นค่อนข้างสูง ซึ่งเป็นปัจจัยสืบเนื่องมาจากกลุ่มเพลงที่จำแนกนั้นมีความต่างกันทางกายภาพ เพราะเพลงไทยเดิมนั้นมีท่วงทำนองและเครื่องดนตรีที่บรรเลงเป็นเอกลักษณ์โดดเด่น ทำให้การวิเคราะห์ค่าคุณลักษณะที่ได้ออกมาั้นแตกต่างและโดดเด่นจากอีกกลุ่มอย่างเห็นได้ชัด ด้วยเหตุนี้จึงส่งผลให้การจำแนกทำได้ง่ายและให้ค่าความถูกต้องที่สูง

ในขณะเดียวกัน การจำแนกเพลงลูกทุ่งและเพลงลูกกรุงออกจากกันนั้นให้ผลที่ไม่สูงมากนัก ซึ่งสาเหตุมาจากความคล้ายคลึงกันของดนตรี ด้วยเหตุนี้การวิเคราะห์ค่าคุณลักษณะที่นำมาใช้จำแนกเพลงสองกลุ่มนี้ จึงต้องใช้หลายชนิดมากกว่าการจำแนกเพลงไทยเดิม

จากผลการวิจัยการจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์เคเตอร์แมชชีนส์ โดยจำแนกเป็น 3 กลุ่มเพลงด้วยกันได้แก่ กลุ่มเพลงไทยลูกทุ่ง เพลงไทยลูกกรุงและเพลงไทย โดยในชั้นแรกผู้วิจัยได้ทำการทดลองเพื่อหาวิธีหาค่าคุณลักษณะที่เหมาะสม โดยใช้เพลงสำหรับชุดเรียนรู้กลุ่มละ 10 เพลง เพลงละ 1,000 เฟรมและกลุ่มทดสอบ 5 เพลง เพลงละ 1,000 เฟรมเช่นเดียวกัน ซึ่งวิธีบีทสเปกตรัม (Beat Spectrum) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 49.1 ลูกทุ่งและไทยเดิมได้ร้อยละ 49.7 ลูกกรุงและไทยเดิมได้ร้อยละ 49.8

เมื่อใช้คุณลักษณะแอลพีซีดีไฟฟ์เซปสตรัม (LPC-Derived Cepstrum) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 52.4 ลูกทุ่งและไทยเดิมได้ร้อยละ 61.2 ลูกกรุงและไทยเดิมได้ร้อยละ 59.7

เมื่อใช้คุณลักษณะอัตราค่าตัดศูนย์ (Zero Crossing Rate) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 51.8 ลูกทุ่งและไทยเดิมได้ร้อยละ 52.6 ลูกกรุงและไทยเดิมได้ร้อยละ 52.8

เมื่อใช้คุณลักษณะพลังงานสเปกตรัม (Spectrum Power) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 31.2 ลูกทุ่งและไทยเดิมได้ร้อยละ 75.6 ลูกกรุงและไทยเดิมได้ร้อยละ 78.3

เมื่อใช้คุณลักษณะสัมประสิทธิ์เซปสตรัมบนความถี่เมล (Mel Frequency Cepstral Coefficient-MFCC) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 45.3 ลูกทุ่งและไทยเดิมได้ร้อยละ 48.4 ลูกกรุงและไทยเดิมได้ร้อยละ 47.8

เมื่อใช้คุณลักษณะเส้นความถี่สเปกตรัม (Line Spectral Frequency-LSF) หาค่าความถูกต้องการจำแนกระหว่างกลุ่มลูกทุ่งและลูกกรุงได้ร้อยละ 62.1 ลูกทุ่งและไทยเดิมได้ร้อยละ 52.3 ลูกกรุงและไทยเดิมได้ร้อยละ 51.6

ผู้วิจัยจึงเลือกที่จะจำแนกเพลงไทยเดิมในขั้นแรกออกจากกลุ่มอื่น โดยใช้ซอฟต์แวร์เตอร์แมชชีนส์ซึ่งจะต้องมีสองกลุ่มในการจำแนก ดังนั้นอีกสองกลุ่มที่เหลือคือกลุ่มเพลงไทยลูกทุ่งและไทยลูกกรุง จึงถูกรวมให้เป็นกลุ่มเดียวกัน โดยใช้วิธีพลังงานสเปกตรัม (Spectrum Power) และวิธีแอลพีซีดีไรฟต์เซปสตรัม (LPC-Derived Cepstrum) ในการหาค่าคุณลักษณะของกลุ่มนี้ เนื่องจากทั้งสองวิธีให้ค่าความถูกต้องในการจำแนกกลุ่มเพลงไทยเดิมออกจากกลุ่มอื่นได้สูงสุดและในขั้นที่สอง เป็นการจำแนกระหว่างกลุ่มเพลงไทยลูกทุ่งและไทยลูกกรุง โดยใช้ค่าคุณลักษณะแอลพีซีดีไรฟต์เซปสตรัม (LPC-Derived Cepstrum) อัตราค่าตัดศูนย์ (Zero Crossing Rate) และเส้นความถี่สเปกตรัม (Line Spectral Frequency-LSF) เนื่องจากให้ค่าความถูกต้องการจำแนกระหว่างสองกลุ่มสูงสุด

ในขั้นต่อมาผู้วิจัยจึงนำวิธีการหาค่าคุณลักษณะที่ได้เลือกไว้มาทำการทดลองหาค่าจำนวนเพลง เฟรม และขนาดของเฟรม ที่ทำให้ค่าความถูกต้องสูงที่สุด โดยได้ขั้นแรกได้หาจำนวนเพลงที่เหมาะสมโดยได้ทำการทดลอง 4 ครั้งได้แก่ 1) จำนวนเพลงชุดเรียนรู้ 15 เพลง ชุดทดสอบ 30 เพลง 2) จำนวนเพลงชุดเรียนรู้ 24 เพลง ชุดทดสอบ 30 เพลง 3) จำนวนเพลงชุดเรียนรู้ 36 เพลง ชุดทดสอบ 30 เพลง 4) จำนวนเพลงชุดเรียนรู้ 45 เพลง ชุดทดสอบ 30 เพลง จากผลการทดสอบผู้วิจัยเลือกใช้ที่จำนวนชุดเรียนรู้ 45 เพลง ชุดทดสอบ 30 เพลง เพราะให้ค่าความถูกต้องสูงสุด ต่อมาจึงทำการทดลองหาขนาดของเฟรมที่เหมาะสม โดยทำการทดลอง 5 ครั้งด้วยกันได้แก่ 1) ตัดเฟรมที่ความกว้าง 256 จุด รวมทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง 2) ตัดเฟรมที่ความกว้าง 512 จุด รวมทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง 3) ตัดเฟรมที่ความกว้าง 1,024 จุด รวมทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง 4) ตัดเฟรมที่ความกว้าง 2,048 จุด รวมทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง 5) ตัดเฟรมที่ความกว้าง 4,096 จุด รวมทั้งสิ้น 500 เฟรมต่อหนึ่งเพลง โดยผู้วิจัยเลือกใช้ที่ขนาดความกว้าง 1,024 จุด เนื่องจากให้ค่าความถูกต้องที่ร้อยละ 71.5 ถึงแม้จะไม่ใช่ว่าค่าที่สูงที่สุด แต่เมื่อเปรียบเทียบกับค่าความกว้างที่ให้ค่าความถูกต้องสูงกว่านั้น จะเป็นค่าที่ดีที่สุดและประหยัดเวลาและทรัพยากรการประมวลผลที่สุด

ในขั้นต่อมานำค่าความกว้างที่เลือกใช้ไปหาจำนวนเฟรมที่เหมาะสมต่อ โดยได้ทดลอง 6 ครั้ง โดยมีจำนวนเฟรมดังนี้ 1) 500 เฟรม 2) 1,000 เฟรม 3) 1,500 เฟรม 4) 2,000 เฟรม 5) 2,500 เฟรม 6) 3,000 เฟรม ผู้วิจัยได้เลือกที่ 2,000 เฟรม ต่อเพลง โดยมีค่าความถูกต้องที่ร้อยละ 84.1 ซึ่งไม่ใช่ค่าความถูกต้องที่สูงที่สุด แต่เมื่อเปรียบเทียบกับจำนวนเฟรมของค่าที่สูงที่สุด มีความต่างกันมาก คือ 3,000 เฟรม แต่ค่าความถูกต้องที่ได้ คือร้อยละ 84.2 ซึ่งหากเลือกใช้ค่าที่สูงที่สุดจะใช้เวลาในการประมวลผลนานและทรัพยากรมากซึ่งไม่คุ้มค่างับค่าความถูกต้องที่ได้มา ดังนั้นจำนวนเฟรมที่ดีที่สุดคือที่ 2,000 เฟรม

เมื่อเราได้ค่าในการเตรียมข้อมูลแล้ว ผลลัพธ์ที่ได้มาคือค่าคุณลักษณะของชุดการเรียนรู้ เมื่อนำเข้ามาทำการเรียนรู้ในซัพพอร์ตเวกเตอร์แมชชีนส์ ก็จะมีการใช้เคอร์เนลเพื่อช่วยในการหาค่าการแบ่งกลุ่ม โดยผู้วิจัยได้เลือกเคอร์เนลสำหรับการจำแนกครั้งที่ 1 คือ เคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 93.7 และการจำแนกในครั้งที่ 2 คือ เคอร์เนลเรเดียลเบสิสฟังก์ชัน (Radial Basis Function-RBF) ที่ค่าพารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1 ให้ค่าความถูกต้องสูงที่สุดที่ร้อยละ 68.7

ในการจำแนกขั้นที่สองนั้นได้ใช้ค่าคุณลักษณะแอลพีซีดีโรฟต์เซปสตรัม (LPC-Derived Cepstrum) อัตราค่าตัดศูนย์ (Zero Crossing Rate) และวิธีเส้นความถี่สเปกตรัม (Line Spectral Frequency-LSF) โดยใช้จำแนกกลุ่มเพลงไทยลูกทุ่งและไทยลูกกรุงออกจากกัน จากนั้นจึงนำค่าคุณลักษณะที่ได้เลือกไว้ไปหาค่าคุณลักษณะของกลุ่มเพลงสองกลุ่มที่จะจำแนกได้ โดยค่าคุณลักษณะเหล่านั้นหาจากข้อมูลชุดเรียนรู้ของสองกลุ่มที่จะจำแนก ในขั้นแรกเป็นการเป็นการจำแนกระหว่างเพลงไทยเดิม และกลุ่มการรวมกันของเพลงไทยลูกทุ่งและลูกกรุง โดยใช้เพลงไทยเดิม 15 เพลง และตัดเป็นเฟรม เพลงละ 2,000 เฟรม โดยมีความกว้างเฟรมละ 1,024 จุด จากนั้นจึงนำชุดข้อมูลค่าคุณลักษณะของทั้งสองกลุ่มที่ได้ ไปทำการเรียนรู้โดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ โดยใช้ค่าพารามิเตอร์ภายในที่ค่า C เท่ากับ 100 และใช้เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน โดยค่า σ (Sigma) เท่ากับ 1 ก็จะได้ชุดต้นแบบสำหรับการจำแนกครั้งที่ 1

ในการจำแนกครั้งที่ 2 เป็นการจำแนกต่อจากผลลัพธ์ที่ได้จากขั้นที่ 1 ซึ่งหากคำตอบคือเพลงนั้นมีความเป็นกลุ่มการรวมกันของเพลงไทยลูกทุ่งและลูกกรุง เพลงนั้นจะถูกนำมาจำแนกต่อไปว่าเป็นเพลงไทยลูกทุ่งหรือไทยลูกกรุง โดยชุดการเรียนรู้ของการจำแนกขั้นนี้จะใช้เพลงไทยลูกทุ่ง 15 เพลง เพลงละ 2,000 เฟรม มีความกว้างเฟรมละ 1,024 จุด รวมทั้งสิ้น 30,000 เฟรม เช่นเดียวกันกับกลุ่มเพลงลูกกรุง จากนั้นจึงนำค่าคุณลักษณะที่ได้ในขั้นนี้เข้าไปทำการเรียนรู้ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีนส์ โดยใช้ค่าพารามิเตอร์ภายในที่ค่า C เท่ากับ 100 และใช้เคอร์เนลแบบเรเดียลเบสิสฟังก์ชัน โดยค่า σ (Sigma) เท่ากับ 1 ก็จะได้ชุดต้นแบบสำหรับการจำแนกครั้งที่ 2

5.2 สรุปผล

งานวิจัยที่เกี่ยวข้องกับเสียง ได้อำนวยประโยชน์ให้กับการใช้สื่อในคอมพิวเตอร์ให้มีประสิทธิภาพเพิ่มขึ้น งานวิจัยด้านเสียงหนึ่งที่ที่น่าสนใจคือการจำแนกกลุ่มเพลง ในกลุ่มเพลงไทยนั้นยังไม่พบการจำแนกหรือวิธีการจำแนกด้วยคอมพิวเตอร์ ดังนั้นการทำงานวิจัยนี้จึงอาจก่อให้เกิดการใช้สื่อด้านเสียงในคอมพิวเตอร์โดยเฉพาะในไทยได้อย่างมีประสิทธิภาพ

ในงานวิจัยนี้ ได้ใช้กลุ่มเพลง 3 กลุ่มเป็นกลุ่มตัวอย่างในการจำแนกกลุ่มเพลงไทย ได้แก่ กลุ่มเพลงไทยลูกทุ่ง กลุ่มเพลงไทยลูกกรุง และกลุ่มเพลงไทยเดิม โดยใช้กลุ่มละ 25 เพลงรวมทั้งสามกลุ่มเป็น 75 เพลง และใช้จำนวนเพลงกลุ่มละ 15 และ 10 เพลงเป็นกลุ่มเรียนรู้และกลุ่มทดสอบตามลำดับ ในขั้นตอนการตัดเฟรมของเพลงเพื่อใช้ในการจำแนกนั้น จำนวนเฟรมต่อเพลงคือ 2,000 เฟรม ซึ่งแต่ละเฟรมมีความกว้าง 1,024 จุด จากนั้นจึงใช้ซอฟต์แวร์เตอร์แมชชีนส์ทำการจำแนก ในการจำแนกนั้นแบ่งเป็นสองชั้น โดยแต่ละชั้นนั้นใช้เคอร์เนลเรเดียลเบสิสฟังก์ชัน ที่ค่าพารามิเตอร์ C ที่ 100 และค่า σ (Sigma) ที่ 1

วิธีการที่นำเสนอสามารถจำแนกกลุ่มเพลงไทยได้ถูกต้องร้อยละ 86.7 การจำแนกเพลงลูกทุ่งและเพลงลูกกรุงออกจากกันนั้นให้ผลที่ไม่สูงมากนัก ซึ่งสาเหตุมาจากความคล้ายคลึงกันของดนตรี ด้วยเหตุนี้การวิเคราะห์ค่าคุณลักษณะที่นำมาใช้จำแนกเพลงสองกลุ่มนี้ จึงต้องใช้หลายชนิดมากกว่าการจำแนกเพลงไทยเดิม

5.3 ปัญหาที่พบ

ในการวิจัยพบปัญหาดังนี้

5.3.1 วิธีการที่ใช้จำแนกข้อมูลได้เพียง 3 กลุ่มเพลงคือ กลุ่มเพลงไทยลูกทุ่ง เพลงไทยลูกกรุง และไทยเดิม

5.3.2 ค่าความถูกต้องการจำแนกกลุ่มเพลงไทยลูกทุ่งและลูกกรุงยังไม่สูงมาก

5.4 ข้อเสนอแนะ

จากการอภิปรายผลข้อสรุป จะเห็นว่างานวิจัยนี้ยังมีประเด็นที่จะนำไปพัฒนาต่อเพิ่มเติม เพื่อให้การจำแนกกลุ่มเพลงไทยประสิทธิภาพดีขึ้นโดยพัฒนาให้สามารถ

5.4.1 จำแนกกลุ่มเพลงไทยอื่นๆ ได้ เช่น เพลงเพื่อชีวิต เพลงไทยพื้นบ้าน เป็นต้น

5.4.2 จำแนกกลุ่มเพลงไทยจากกลุ่มเพลงสากลได้

5.4.3 จำแนกเพลงไทยที่มีเสียงขับร้องได้

5.4.4 จำแนกกลุ่มเพลงไทยได้ถูกต้องมากกว่าร้อยละ 86.7 ทั้งนี้อาจหาวิธีการคัดแยกคุณลักษณะที่เหมาะสมมากยิ่งขึ้น

5.5 ประโยชน์ที่ได้รับ

5.5.1 ได้วิธีการที่สามารถจำแนกกลุ่มเพลงไทยเดิม ลูกทุ่ง และลูกกรุง

5.5.2 ได้วิธีหาค่าคุณลักษณะที่เหมาะสมสำหรับการจำแนกกลุ่มเพลงไทย

5.5.3 เป็นแนวทางให้ผู้สนใจได้ค้นคว้าและศึกษาเกี่ยวกับการจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์เตอร์แมชชีนส์

เอกสารอ้างอิง

1. C. Xu, et al. Musical Genre Classification Using Support Vector Machines. [serial online] 2003 [cited 2005 May 15]. Available from : URL : <http://www.comp.nus.edu.sg/~shaoxi/papers/ICASSP03.pdf>
2. J. Foote and S. Uchihashi. The Beat Spectrum: A New Approach to Rhythm Analysis. [serial online] 2003 [cited 2005 June 12]. Available from : URL : <http://www.fxpal.com/publications/FXPAL-PR-01-022.pdf>
3. S. Das and G.P. Biswas. Design of Adaptive Speaker Identification System Using LPC Coefficients. [serial online] 2002 [cited 2005 August 8]. Available from : URL : <http://www.cmriindia.nic.in>
4. L. Lu, H. Jiang and H. J. Zhang. A Robust Audio Classification and Segmentation Method. [serial online] 2003 [cited 2005 November 28]. Available from : URL : http://research.microsoft.com/users/llu/Publications/ACMMM01_AudioAnalysis.pdf
5. G. Tzanetakis, G. Essl and P. Cook. Human Perception and Computer Extraction of Musical Beat Strength. [serial online] 2002 [cited 2005 June 24]. Available from : URL : <http://www.scf.usc.edu/~ise575/a/students/zali/zali-4.4.pdf>
6. J. Foote, M. Cooper and U. Nam. Audio Retrieval by Rhythmic Similarity. [serial online] 2003 [cited 2005 June 13]. Available from : URL : <http://www.fxpal.com/people/foote/papers/ISMIR02rhythm.pdf>
7. ณัฐนันท์ ทัดพิทักษ์กุล, บุญธีร์ เครือตราชู และชัย วุฒิวิวัฒน์ชัย. “การจำแนกเสียงรบกวนด้วยวิธีการวิเคราะห์องค์ประกอบหลักแบบเคอร์เนล.” การประชุมวิชาการทางวิศวกรรมไฟฟ้าครั้งที่ 28 เล่ม 2. 28 (ตุลาคม 2548) : 961-964.
8. วีระ ทองไพบูลย์, ไซยนต์ สุวรรณชีวะศิริ และวิทวัส ผ่องญาติ. “อัลกอริทึมการรู้จำเสียงและโน้ตดนตรีไทยสำหรับเครื่องดนตรีชิ้นเดียว.” The 9th National Computer Science and Engineering Conference NCSEC 2005. 9 (ตุลาคม 2548) : 459-466.
9. น้ำตาแสงใต้ [ออนไลน์] 2545. [สืบค้นวันที่ 4 มกราคม 2549]. จาก <http://th.wikipedia.org/wiki/น้ำตาแสงใต้>

10. ณัฏชา โสคติยานุรักษ์. ทฤษฎีดนตรี. พิมพ์ครั้งที่ 1. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2543.
11. ศวิต กาสุริยะ. ระบบการรู้จำผู้พูด: การนำซึ่งผู้พูดแบบขึ้นกับบทคำพูด. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2542.
12. ชัชวาล หาญสกุลบันเทิง. การสังเคราะห์พยางค์ภาษาไทยด้วยวิธีการสังเคราะห์แบบวิเคราะห์โดยใช้คู่เส้นสเปกตรัม. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2542.
13. N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge : Cambridge University press, 2000.

ประวัติผู้วิจัย

ชื่อ : นายอานนท์ นามสนิท

ชื่อวิทยานิพนธ์ : การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์เคเตอร์แมชชีนส์

สาขาวิชา : เทคโนโลยีคอมพิวเตอร์

ประวัติ

เกิดวันที่ 15 สิงหาคม 2521 ที่จังหวัดระยอง สำเร็จการศึกษาระดับปริญญาตรี สาขาวิชาเอก วิศวกรรมคอมพิวเตอร์ จากสถาบันเทคโนโลยีราชมงคล วิทยาเขตพระนครเหนือ