

Received: 14 เม.ย. 2568

Revised: 7 พ.ค. 2568

Accepted: 14 พ.ค. 2568

ขั้นตอนวิธีรู้จำระดับเสียงดนตรีไทยเดิมด้วยวิธีการถดถอยพหุนาม
The Thai Classical Music Pitch Recognition Algorithm Based on
The Polynomial regression Method

ธีราวุฒิ หงษ์ฟ่อน¹, พาสน์ ปราโมกษ์ชน¹, กิตติกร หาญตระกูล¹ และ ปวีณ เชื้อนแก้ว^{1*}

¹สาขาวิชานวัตกรรมเทคโนโลยีดิจิทัล คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้

Teerawut Hongfon¹, Part Pramokchon¹, Kittikorn Hantrakul¹ and Paween Khoenkaw^{1*}

¹Department of Digital Technology Innovation, Faculty of Science, Maejo University

*Corresponding author: paween_k@mju.ac.th

Abstract

Thai Classical Music has its own standards for pitch frequency and frequency range that differ from Western music. Therefore, applying Western note identification algorithms to Thai Classical Music results in inaccurate note identification, especially when used for creating musical signatures for retrieval or similarity comparison. Furthermore, the pitch frequencies of Thai Classical Music vary across different standards. This research proposes a new algorithm for identifying notes in Thai Classical Music based on a regression method. Using a dataset compiled from textbooks and musical pieces from various sources, the researchers derived an equation to identify the pitch frequencies of Thai musical notes. This equation was then used to generate a new frequency table, referencing the sequence of notes on a piano. The signature creation process for retrieval begins by segmenting the audio data into smaller parts, then applying the Fast Fourier Transform (FFT) to convert each segment into frequency data. The peak frequency is extracted from the one-sideband frequency data is analyzed to identify peak frequencies, which are then matched with a proposed frequency table to determine the corresponding musical notes. These notes are then concatenated into a character string, forming the signature of the piece. Experimental results from developing a music retrieval system show that the proposed method

outperforms the traditional Western-based method, achieving an accuracy of 72.53 percent.

Keywords: Thai Classical Music; Fast Fourier Transform; Regression Method

บทคัดย่อ

ดนตรีไทยเดิมมีมาตรฐานของความถี่เสียงและช่วงความถี่ที่แตกต่างจากดนตรีสากล ซึ่งหากนำขั้นตอนวิธีระบุตัวโน้ตแบบตะวันตกมาใช้กับดนตรีไทยเดิมจะส่งผลให้การระบุตัวโน้ตเพื่อใช้สำหรับสร้างลายเซ็นสำหรับการค้นคืนหรือเปรียบเทียบความคล้ายของเพลงไม่แม่นยำ ความถี่ตัวโน้ตของดนตรีไทยเดิมก็มีหลายมาตรฐานแตกต่างกัน งานวิจัยนี้ได้นำเสนอขั้นตอนวิธีการระบุตัวโน้ตดนตรีไทยเดิมโดยใช้วิธีใหม่บนพื้นฐานของการถดถอย จากชุดข้อมูลที่ผู้วิจัยได้รวบรวมมาจากตำราและบทเพลงจากหลายแหล่งจึงได้สมการที่ใช้ระบุความถี่เสียงของโน้ตดนตรีไทยเดิม จากนั้นจึงนำสมการนี้ไปสร้างตารางความถี่ตัวโน้ตขึ้นมาใหม่โดยอ้างอิงลำดับตัวโน้ตของเปียโน ขั้นตอนวิธีการสร้างลายเซ็นสำหรับสืบค้นเริ่มจากการแยกข้อมูลเสียงดนตรีออกเป็นส่วนย่อย จากนั้นจึงนำข้อมูลส่วนย่อยนี้ไปแปลงเป็นความถี่ด้วยวิธีแปลงฟูริเยร์แบบเร็ว จากนั้นจึงนำข้อมูลความถี่เสียง 1 sideband มาหาจุดสูงสุด และนำความถี่สูงสุดนั้นไปเทียบจากตารางความถี่ที่นำเสนอ จะได้เป็นตัวโน้ต จึงนำข้อมูลนี้มาร้อยต่อกันเป็นสายอักขระเพื่อใช้เป็นลายเซ็นของเพลง จากการทดลองสร้างระบบค้นคืนเพลงโดยเปรียบเทียบระบบที่นำเสนอกับวิธีการของดนตรีสากลพบว่าระบบที่นำเสนอมีประสิทธิภาพร้อยละ 72.53

คำสำคัญ: ดนตรีไทยเดิม; แปลงฟูริเยร์แบบเร็ว; การถดถอย

1. บทนำ

พัฒนาการดนตรีไทยซึ่งเกิดขึ้นมาตั้งแต่สมัยยุคประวัติศาสตร์ได้เริ่มต้นจากการเป่าสังข์ใช้ในงานมงคลต่างๆ ตามวัฒนธรรม ความเชื่อของศาสนาพราหมณ์ เมื่อเข้าสู่สมัยทวารวดี เริ่มมีการสร้างเครื่องดนตรีไทยดังหลักฐานปูนปั้นที่จังหวัดราชบุรี ลักษณะคล้ายพิณน้ำเต้า, พิณ 5 สาย, ฉิ่ง, กรับ และระฆังหิน เข้าสู่ยุคสุโขทัย มีหลักฐานปรากฏเด่นชัดคือ วงปี่พาทย์เครื่องห้า และวงขับไม้ หน้าที่หลักๆ คือ บรรเลงประกอบการแสดง, ขับกล่อม, พิธีการ, ทำหน้าที่สัญญาณ ตั้งแต่นั้นเป็นต้นมา ดนตรีไทยได้มีบทเพลงที่หลากหลาย และมีการพัฒนามากขึ้น ช่วงรัชกาลที่ 5 ได้มีการบันทึกเสียงจากกระบอกเสียง เมื่อ พ.ศ.2443 จนกระทั่งมีการศึกษาพบว่าการเก็บข้อมูลเกี่ยวกับเสียงดนตรีไทย

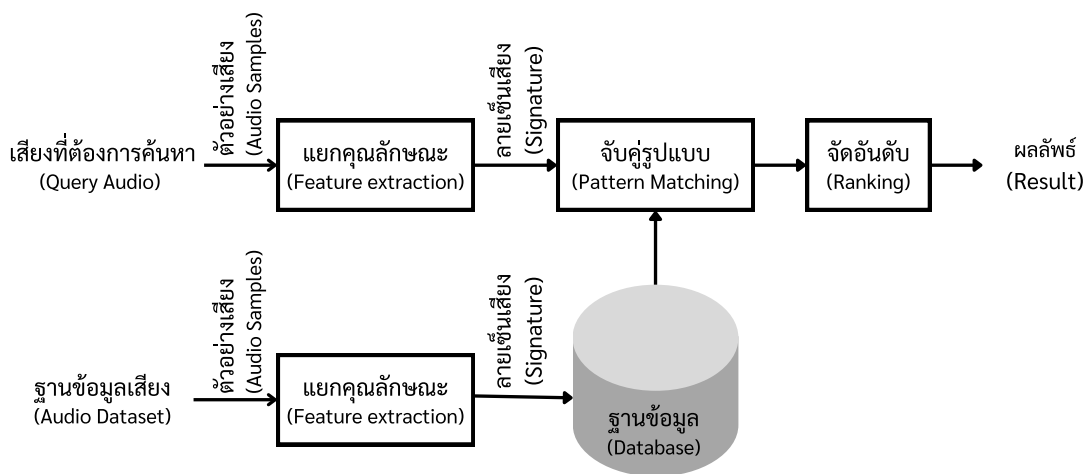
ถึง 10 กว่าครั้ง และทำตารางค่าทฤษฎีบันไดเสียงไดอะทอนิก 7 เสียงเท่ากัน หรือทฤษฎี 7 เสียงเท่ากันของระบบดนตรีไทย ซึ่งไม่ตรงกับค่าที่บันทึกไว้แต่ละสำนักดนตรี

ระบบเสียงดนตรีไทย และดนตรีสากลมีความแตกต่างกันอย่างชัดเจนซึ่งระบบเสียงดนตรีสากลได้ใช้ค่าสมการการบันไดเสียงโครมาติก 12 เสียงเท่ากัน (12 Tone Equal Temperament Tuning System / Equal temperament) คือระดับเสียงถัดไปแต่ละเสียง ได้จากการคูณเสียงก่อนหน้าด้วย รากที่ 12 ของ 2 โดย 1 ระดับเสียงสามารถแบ่งออกเป็น 12 ส่วนเท่า ๆ กัน ซึ่งระบบดนตรีสากลในการเรียกชื่อตัวโน้ต(การระบุระดับเสียง) จะมีทั้ง C, C[#]/D^b, D, E^b/D[#], E, F, F[#]/G^b, G, A^b/G[#], A, B^b/A[#], และ B ซึ่งแตกต่างกับระบบเสียงดนตรีไทย ที่ปัจจุบันมีการอ้างอิงทฤษฎีบันไดเสียงไดอะทอนิก 7 เสียงเท่ากัน (Diatonic scale) หรือ Major scale 7 เสียง โดยตัดครึ่งเสียงของระบบดนตรีสากลออก จะเหลือแค่ C, D, E, F, G, A, และ B แต่ความเป็นจริง ดนตรีไทยตั้งแต่ยุคสมัยที่กล่าวมานั้นไม่มีการบันทึกค่าความถี่เสียงสูง - ต่ำ ที่เป็นมาตรฐาน เนื่องจากเครื่องดนตรีไทยที่มีการประดิษฐ์ออกมา มีการผลิตที่แตกต่างกัน และวัดค่าเสียงโดยใช้หู, เครื่องดนตรีแม่แบบประจำเจ้าสำนัก หรือการเทียบลักษณะเครื่องดนตรี ทำให้ความถี่เสียงที่ออกมา ไม่เหมือนกันและมีความเบี่ยงเบนออกไปไม่มากนักน้อยซึ่งเป็นการยากที่จะจับจุดค่าความถี่ที่คงที่ได้

โดยความท้าทายในการรู้จำระดับเสียงของดนตรีไทย สืบเนื่องจากการเก็บข้อมูลของค่าความถี่เสียงสูง-ต่ำจากในแต่ละชุดข้อมูลและแหล่งอ้างอิง ไม่มีความคงที่ของค่าดังกล่าว ผู้วิจัยได้นำค่าความถี่เสียงสูง - ต่ำ ของแต่ละที่มารวบรวม และคำนวณโดยใช้วิธีการถดถอยแทนการจับจุดเส้นความถี่และระบุระดับเสียงแบบตรงๆ เพื่อให้เกิดความถูกต้องและชัดเจนมากยิ่งขึ้น ซึ่งบทนำข้างต้นนี้ นำไปสู่การศึกษาวิธีรู้จำระดับเสียงดนตรีไทยเดิมด้วยวิธีการถดถอยกำลังสองน้อยที่สุด

2. ขั้นตอนวิธีที่นำเสนอ

ขั้นตอนวิธีการนำเสนอในงานวิจัยนี้ได้ทำการประมวลผลโดยวัดความถูกต้องในการจำแนกเสียงดนตรี โดยชุดข้อมูลที่ใช้ประกอบด้วยเพลง บุหลั่นลอยเลื่อน, โหมโรงโอเอเรศ 3 ชั้น (วงเครื่องสาย), ลาวคำหอม (ขลุ่ยเพียงออ), ลาวดวงเดือน (ซอสามสาย), ซับไม้ (ซอสามสาย), แขมมอญ บางช้าง 3 ชั้น, ลาวเสียงเทียน (เครื่องสายวงใหญ่) และสร้อยเพลง (ขลุ่ยเพียงออ) จากต้นฉบับที่บันทึกในรูปแบบของตลับเทป และแผ่นซีดี ซึ่งถูกแปลงเป็นรูปแบบแฟ้มข้อมูลดิจิทัลโดยการจำแนกโน้ตดนตรีใช้วิธีการดัง *รูปภาพที่ 1* โดยมีขั้นตอนการทำงานดังนี้



รูปภาพที่ 1 ภาพรวมขั้นตอนการจำแนกเสียงดนตรี

โดยขั้นตอนแรกทำการนำเข้าข้อมูลเสียงเพลงดิจิทัล และทำการประมวลผลด้วยการสกัดคุณลักษณะเด่น (Feature Extraction) จะได้ออกมาเป็นข้อมูลเอกลักษณ์ของเสียงดนตรีที่รับเข้ามา (Signature) เพื่อบันทึกเป็นฐานข้อมูล เพื่อใช้ในการเปรียบเทียบความคล้ายคลึงกับเสียงเพลงที่ต้องการค้นคืนต่อไป [P. Khoenkaw & P. Piamsa-nga. 2014], [Uitdenbogerd, Alexandra & Zobel, Justin. 2002], [W. -j. Liu and J. -l. Feng. 2001], [Asif Ghias, Jonathan Logan, David Chamberlin, & Brian C Smith. 1995]

ปัจจัยสำคัญที่ส่งผลต่อความถูกต้องในการรู้จำระดับเสียงดนตรีไทยเดิมคือเอกลักษณ์ที่สกัดออกมาได้ต้องสามารถเป็นตัวแทนของเพลงหรือทำนองนั้นได้อย่างสมบูรณ์ หากขั้นตอนการสกัดเอกลักษณ์หรือสัดลักษณะเด่นมีความผิดพลาด ย่อมจะส่งผลให้การค้นคืนทำงานผิดพลาดตามไปด้วย วิธีการหนึ่งที่เป็นที่นิยมคือการใช้ โครโมแกรม (Chromogram, Chroma feature) เพื่อระบุระดับเสียงในระบบโน้ตสากล [Shah, Ayush & Kattel, Manasi & Nepal, Araj & Shrestha, D.. 2019] ซึ่งวิธีการนี้จะสกัดระดับเสียงดนตรี ในรูปแบบของดนตรีสากล 12 ระดับแบบระบบการตั้งเสียงในดนตรีที่แบ่งช่วงเสียงออกเป็น 12 ส่วนที่เท่ากัน (Equal Temperament) ดังแสดงใน สมการที่ (1) โดยที่ 88 คีย์มาตรฐาน หรือ 108 คีย์ขยาย โดยใช้เปียโนเป็นตัวหลัก จะอ้างอิงระบบ 12 โทนเท่ากัน ด้วยคีย์ที่ 49 (เสียง A_4) ความถี่ที่ 440 Hz [International Organization for Standardization. 1975] การกระจายความถี่ระดับเสียงถัดไปแต่ละเสียง ได้จากการคูณเสียงก่อนหน้านี้ ด้วย $^{12}\sqrt{2}$

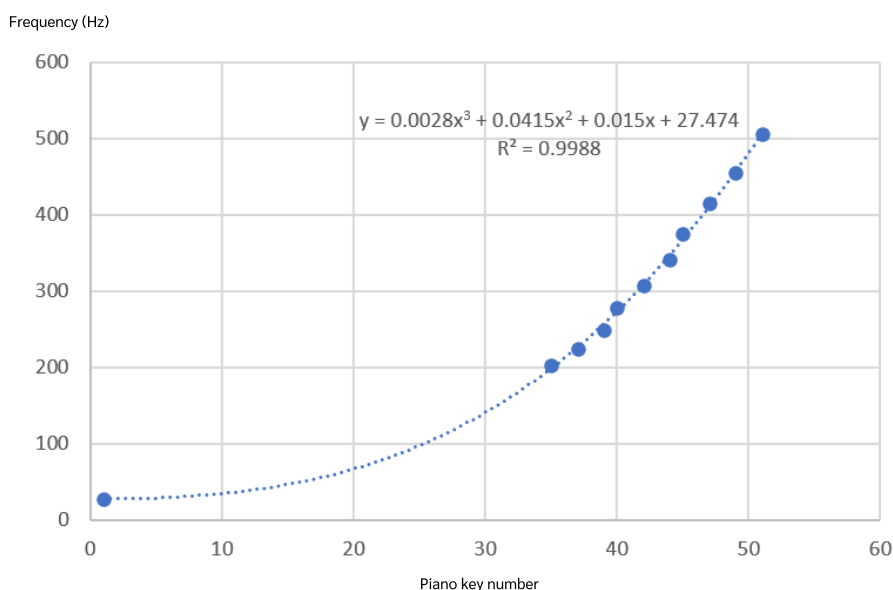
สูตรคำนวณค่าคลื่นความถี่เสียง f

$$f(n) = (\sqrt[12]{2})^{n-49} \times 440\text{Hz} \quad (1)$$

โดยที่ $f(n)$ = ค่าคลื่นความถี่เสียง

n = ลำดับคีย์เสียง

การใช้โครโมแกรม (Chromogram or Chroma feature) นั้นไม่สามารถนำมาประยุกต์ใช้กับดนตรีไทยเดิมได้โดยสมบูรณ์ เนื่องจากดนตรีไทยเดิมไม่ได้กำหนดระดับเสียงแบบระบบการตั้งเสียงในดนตรีที่แบ่งช่วงเสียงออกเป็น 12 ส่วนที่เท่ากัน (Equal Temperament) หากแต่จะมีระบบที่เป็นของไทยเอง ซึ่งไม่ได้แบ่งเป็น 12 ระดับ นอกจากนั้นการคำนวณหาค่าความถี่เสียงสูง-ต่ำของดนตรีไทยนั้น มีแหล่งอ้างอิง เอกสารวิชาการ และหนังสือเรียนที่ไม่สามารถระบุได้ว่าความถี่เสียงสูง-ต่ำเป็นของโน้ตตัวใด ผู้วิจัยได้ดำเนินการสืบค้นและรวบรวมความถี่เสียงสูง-ต่ำของเสียงฆ้องวง, ระนาดเอกเหล็ก และขลุ่ยเพียงออ ดังสรุปได้ใน *รูปภาพที่ 3* มาคำนวณหาระยะห่างและความเบี่ยงเบนของเสียงโดยใช้วิธีการวิเคราะห์การถดถอยเชิงพหุนาม และใช้สมการนั้นในการสกัดข้อมูลระดับเสียงดนตรีดังนี้



รูปภาพที่ 2 การถดถอยที่ผู้วิจัยเสนอ

การถดถอยนี้ได้มาจากการรวบรวมบันทึกข้อมูลค่าความถี่เสียงสูง-ต่ำ โดยนำมาจากการเก็บบันทึกค่าความถี่เสียงสูง-ต่ำของดนตรีไทยที่พบ โดยแกน X คือลำดับโน้ตโดยเทียบเคียงกับคีย์เปียโนสากล (Piano key number) และแกน Y คือค่าความถี่เสียงสูง-ต่ำที่พบ (Frequency) จึงได้การถดถอย ดัง *สมการที่ (2)* โดยใช้พหุนามกำลัง 3 ซึ่งให้ผลลัพธ์ใกล้เคียง และดีที่สุดของการประเมิน

ความเหมาะสมของแบบจำลองของวิธีการคำนวณหาค่าความถี่ ซึ่งผู้วิจัยได้ทดลองพหุนามกำลังอื่นๆ พบว่าผลลัพธ์มีความเบี่ยงเบนสูง และไม่สามารถนำไปใช้ได้ เนื่องจากค่าแกน X แกน Y ไม่มีความสัมพันธ์กันนำไปสู่ค่าความถี่ที่สูงกว่าความเป็นจริง

$$A_x = (0.0028x^3) + (0.0415x^2) + (0.015x) + 27.474 \quad (2)$$

โดยที่ A_x = ค่าความถี่ที่นำเสนอ

x = ค่าลำดับโน้ต (Piano key number) จาก *รูปภาพที่ 2*

ทำการหาค่าลำดับเสียงสูง-ต่ำวิธีคล้ายกับการหาค่าความถี่เสียงสูง-ต่ำของดนตรีสากล ซึ่งมีความแตกต่างตรงที่ระยะค่าความถี่เสียงสูง-ต่ำของดนตรีสากลมีความคงที่ หรือที่เรียกว่าเสียงโครมาติก (12 เสียงเท่า) ค่าลำดับเสียงสูง-ต่ำของดนตรีไทยจะไม่มีค่าที่คงที่ และระยะห่างของเสียงไม่เท่ากัน จึงไม่สามารถใช้ทฤษฎี 7 เสียงเท่ากันได้ [อุทิศ นาคสวัสดิ์. 2531], [สราวุฒ สุจิตจร. 2546] เนื่องด้วยทฤษฎี 7 เสียงเท่ากันใช้วิธีล็อกโน้ตตัวโด (C) ให้เท่ากับสากล และโน้ตตัวอื่นกระจายค่าให้เท่ากัน โดยผู้วิจัยได้ทำการคำนวณหาค่าลำดับเสียงที่ 1-88 ใหม่โดยอ้างอิงลำดับจาก 88 คีย์มาตรฐานของเปียโนสากลดัง *สมการที่ (3)* และเทียบระหว่าง 12 เสียงเท่าของดนตรีสากล กับ 7 เสียงของดนตรีไทยที่ผู้วิจัยนำเสนอ ซึ่งได้ผลดัง *รูปภาพที่ 3* นี้ โดย **เส้นสีเขียว** หมายถึงค่าความถี่เสียงสูง-ต่ำที่ใกล้เคียงกัน, ลำดับเสียงสูง-ต่ำ และตัวโน้ตที่เหมือนกัน **เส้นสีแดง** หมายถึงค่าความถี่เสียงสูง-ต่ำที่เบี่ยงเบนกัน, ลำดับเสียงสูง-ต่ำ และตัวโน้ตที่ไม่เหมือนกัน

$$piano\ key = \min_i |f_{peak} - A_x| \quad (3)$$

โดยที่ $piano\ key$ = ลำดับเสียงอ้างอิงคีย์เปียโน

f_{peak} = ค่าความถี่เสียงโดย

วิธีการหาจุดสูงสุด (Peak detection)

A_x = ค่าความถี่ที่นำเสนอ *สมการที่ (2)*

ค่าความถี่เสียงโครมาติก 12 เสียงเท่ากัน			ค่าความถี่ 7 เสียงดนตรีไทยที่ผู้วิจัยเสนอ		
ลำดับ	ความถี่ (Hz)	ตัวโน้ต	ลำดับ	ความถี่ (Hz)	ตัวโน้ต
1	27.500	A	1	27.5333	A
2	29.135	Bb	2		
3	30.868	B	3	27.9681	B
4	32.703	C	4	28.3772	C
5	34.648	C#	5		
6	36.708	D	6	29.6628	D
7	38.891	Eb	7		
8	41.203	E	8	31.6836	E
9	43.654	F	9	33.0117	F
10	46.249	F#	10		
11	48.999	G	11	36.3873	G
12	51.913	Ab	12		
13	55.000	A	13	40.8341	A
14	58.270	Bb	14		
15	61.735	B	15	46.4865	B
16	65.406	C	16	49.8068	C
17	69.296	C#	17		
18	73.416	D	18	57.5196	D
19	77.782	Eb	19		
20	82.407	E	20	66.774	E
21	87.307	F	21	72.0213	F
22	92.499	F#	22		
23	97.999	G	23	83.8401	G
24	103.826	Ab	24		
25	110.000	A	25	97.5365	A
26	116.541	Bb	26		
27	123.471	B	27	113.2449	B
28	130.813	C	28	121.8956	C
29	138.591	C#	29		
30	146.832	D	30	140.874	D
31	155.564	Eb	31		
32	164.814	E	32	162.2004	E
33	174.614	F	33	173.7861	F
34	184.997	F#	34		
35	195.998	G	35	198.8865	G
36	207.652	Ab	36		
37	220.000	A	37	226.6709	A
38	233.082	Bb	38		
39	246.942	B	39	257.2737	B
40	261.626	C	40	273.674	C
41	277.183	C#	41		
42	293.665	D	42	308.7564	D
43	311.127	Eb	43		
44	329.628	E	44	346.9932	E
45	349.228	F	45	367.3365	F
46	369.994	F#	46		
47	391.995	G	47	410.5569	G

48	415.305	Ab
49	440.000	A
50	466.164	Bb
51	493.883	B
52	523.251	C
53	554.365	C#
54	587.330	D
55	622.254	Eb
56	659.255	E
57	698.457	F
58	739.989	F#
59	783.991	G
60	830.609	Ab
61	880.000	A
62	932.328	Bb
63	987.767	B
64	1046.502	C
65	1108.731	C#
66	1174.659	D
67	1244.508	Eb
68	1318.510	E
69	1396.913	F
70	1479.978	F#
71	1567.962	G
72	1661.219	Ab
73	1760.000	A
74	1864.655	Bb
75	1975.533	B
76	2093.005	C
77	2217.461	C#
78	2349.318	D
79	2489.016	Eb
80	2637.020	E
81	2793.826	F
82	2959.955	F#
83	3135.963	G
84	3322.438	Ab
85	3520.000	A
86	3729.310	Bb
87	3951.066	B
88	4186.009	C

48		
49	457.2677	A
50		
51	507.6033	B
52	534.1724	C
53		
54	590.1972	D
55		
56	650.1828	E
57	681.7029	F
58		
59	747.8817	G
60		
61	818.3573	A
62		
63	893.2641	B
64	932.4212	C
65		
66	1014.2268	D
67		
68	1100.7996	E
69	1145.9157	F
70		
71	1239.8913	G
72		
73	1338.9701	A
74		
75	1443.2865	B
76	1497.4508	C
77		
78	1609.8756	D
79		
80	1727.874	E
81	1789.0053	F
82		
83	1915.6161	G
84		
85	2048.1365	A
86		
87	2186.7009	B
88	2258.2916	C

รูปภาพที่ 3 แผนผังตารางเทียบค่าความถี่เสียงสูง-ต่ำของดนตรีสากลกับค่าความถี่เสียงสูง-ต่ำของดนตรีไทยที่นำเสนอในการวิจัย

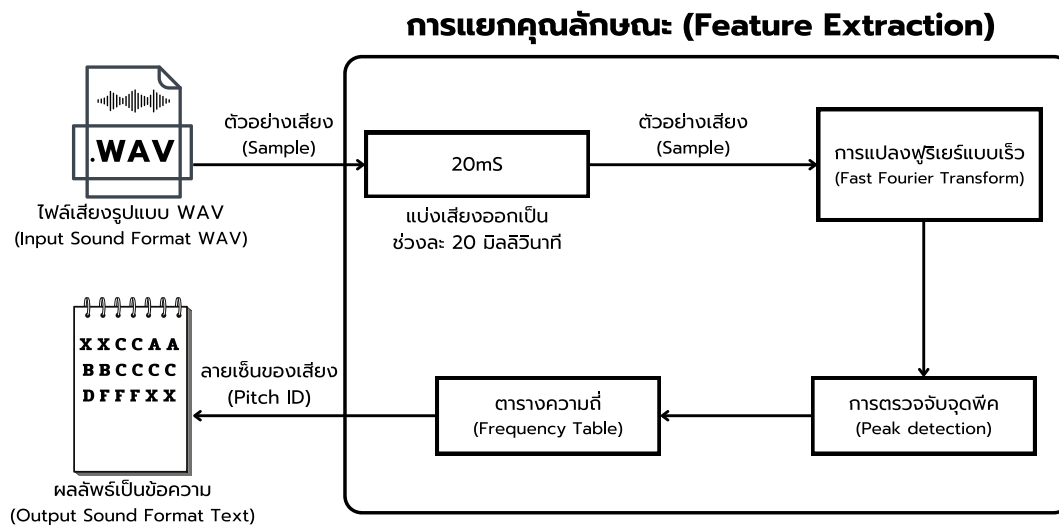
เมื่อทำการเทียบกับค่าสูง-ต่ำของเสียงดนตรีสากลพบว่าค่าลำดับเสียงสูง-ต่ำของดนตรีไทยที่ได้จะมีค่าที่ใกล้เคียงกับค่าลำดับเสียงสูง-ต่ำของดนตรีสากลไม่น้อย และไม่มากเกินไปซึ่งจะอยู่ในช่วงที่สามารถใช้ชื่อของค่าลำดับเสียงสูง-ต่ำของสากลได้ แต่ลำดับที่ 30 - 61 พบการเบี่ยงเบน ตั้งแต่ลำดับที่ 1 - 28 และ 63 - 88

หลังจากที่ผ่านการคำนวณหาค่าความถี่เสียงสูง-ต่ำและได้ตัวโน้ตดนตรีแล้ว ก่อนที่จะนำไปเทียบกับข้อมูลของบทเพลงไทยเดิม ซึ่งผู้วิจัยได้ใช้ตัวโน้ตของดนตรีสากลในการแทนระดับเสียงดนตรีดัง ตารางที่ 1

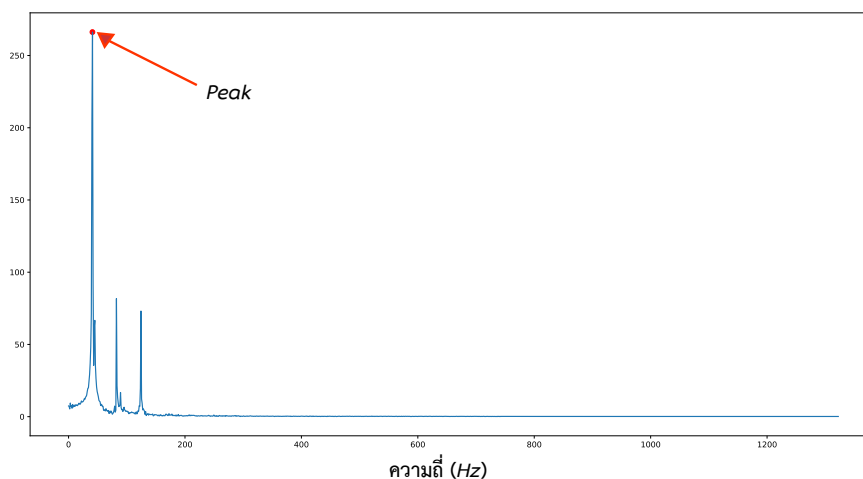
ตารางที่ 1 วิธีการเรียกชื่อตัวโน้ตที่ผู้วิจัยเสนอ

ตัวโน้ตดนตรี	ตัวโน้ตดนตรีที่ผู้วิจัย	การเรียกชื่อตัวโน้ต	
	เสนอ	สากล	ไทย
(ไม่มี)	X	ไม่มีเสียง	ไม่มีเสียง
A	A	ลา	ลา
Ab	Ab	ลา-แฟลต	-
B	B	ที	ที
Bb	Bb	ที-แฟลต	-
C	C	โด	โด
C#	C#	โด-ชาร์พ	-
D	D	เร	เร
Eb	Eb	มี-แฟลต	-
E	E	มี	มี
F	F	ฟา	ฟา
F#	F#	ฟา-ชาร์พ	-
G	G	ซอล	ซอล
G#	G#	ซอล-ชาร์พ	-

การสกัดระดับทำนองเพลงมีขั้นตอนการทำงานดัง *รูปภาพที่ 4* โดยเริ่มจากการนำแฟ้มข้อมูลเสียงเพลงในรูปแบบ .wav ทำการแบ่งช่วงเพลง ช่วงละ 20mS จากนั้นจะทำการแปลงข้อมูลเสียงเป็นโดเมนความถี่ ด้วยวิธี การแปลงฟูริเยร์แบบไม่ต่อเนื่อง (Fast Fourier transform) ขั้นตอนถัดไป จะทำการระบุความถี่เสียงที่เป็นทำนองหลักด้วยวิธีการหาจุดสูงสุด (Peak detection) ดัง *รูปภาพที่ 5* โดยตำแหน่งที่ค้นพบว่าเป็นจุดสูงสุดได้แสดงเป็นจุดสีแดงในภาพ จากนั้นทำการแปลงความถี่ที่ตรวจจับมาได้จากการหาจุดสูงสุดให้เป็นระดับเสียงหรือ ตัวโน้ตด้วยวิธีตารางค้นหา (Lookup table) ด้วยการหาตำแหน่งที่ใกล้ที่สุดเมื่อเทียบตำแหน่งระดับเสียงที่ได้จาก *สมการที่ (2)* จะได้เป็นทำนองดนตรีในรูปแบบสายอักขระ (String of Melody) ดังแสดงตัวอย่างใน *รูปภาพที่ 6* และ *รูปภาพที่ 7* เพื่อใช้เป็นเอกลักษณ์สำหรับใช้ค้นคืนเพลงไทยเดิมต่อไป



รูปภาพที่ 4 ขั้นตอนการสกัดทำนองเพลง



รูปภาพที่ 5 การระบุความถี่เสียงที่เป็นทำนองหลักด้วยวิธีการหาจุดสูงสุด (Peak detection)

ตารางที่ 2 ชุดข้อมูลที่ทำกรวิจัย

บทเพลง	แหล่งข้อมูล
บุหลันลอยเลื่อน	ดลั้บเทป
โหมโรงไอยเรศ 3 ชั้น	แผ่นซีดี
ลาวคำหอม (ขลุ่ย)	แผ่นซีดี
ลาวดวงเดือน	ดลั้บเทป
ขับไม้ (ซอสามสาย)	ดลั้บเทป
แขกมอญบางช้าง 3 ชั้น	ออนไลน์
ลาวเสียงเทียน	ออนไลน์
สร้อยเพลง (ขลุ่ย)	แผ่นซีดี

แต่ละเพลง(.wav) ถูกตัดออกมาประมวลผลเพียง 30 วินาทีแรก ด้วยการสกัดออกมาเป็นช่วงละ 20 mSec เพื่อแปลงเป็นตัวโน้ต (ดังอธิบายในขั้นตอนใน *รูปภาพที่ 4*) มีผลทำให้แต่ละเพลงมีจำนวน 1500 ตัวโน้ต เพื่อใช้เป็นเอกลักษณ์ในการค้นหา โดยค่าที่สกัดได้จากขั้นตอนวิธีที่นำเสนอจะถูกนำมาเปรียบเทียบความถูกต้องกับผลเฉลยที่สร้างโดยผู้เชี่ยวชาญ การวัดความถูกต้องของการสกัดทำนองเพลงใช้วิธีการเปรียบเทียบความเหมือนของสายอักขระที่สกัดได้ในสองรูปแบบคือ

1. การคำนวณหาค่าความถูกต้องโดยนับโน้ตเทียบตำแหน่งต่อตำแหน่ง วิธีการนี้จะทำการเปรียบเทียบสายอักขระที่ละตำแหน่งว่าเหมือนกันหรือไม่ หากไม่เหมือนกันก็จะถือว่าตำแหน่งนั้นมีความผิดพลาด และนำค่าที่ผิดพลาดนั้นมาเฉลี่ยและคำนวณเป็นความถูกต้องดัง *สมการที่ (4)*

$$E_i = \begin{cases} 0 & \text{when } A_i = B_i \\ 1 & \text{others} \end{cases}$$

โดยที่ A = ผลเฉลย

B = ผลลัพธ์

E_i = ระดับเสียงที่ผิดพลาด

$$accuracy = 100 - \left(\frac{1}{N} \sum_{i=0}^n E_i \right) \times 100 \quad (4)$$

2. การคำนวณหาค่าความถูกต้องโดยนับโน้ตเทียบตำแหน่งต่อตำแหน่งอิง Interval วิธีการนี้จะไม่ทำการเปรียบเทียบระดับเสียงแบบต้องเหมือนกันทุกประการ แต่จะมีการหาระดับการ

เปลี่ยนแปลง เนื่องจากในการบรรเลงดนตรี อาจไม่ได้มีการตั้งระดับเสียงเครื่องดนตรี (Instrument tuning) ให้เป็นมาตรฐานเดียวกันก่อนการบรรเลง จึงส่งผลให้เพลงเดียวกันโน้ตดนตรีมีค่าที่ไม่เหมือนกัน หรือ การบรรเลงคนละระดับเสียง (Scientific pitch notation) หรือ ช่วงคู่แปด (octave) ดังนั้นการเปรียบเทียบความถูกต้องจึงจะพิจารณาเฉพาะระดับ ช่วงความถี่ (Interval) ที่เปลี่ยนไปของเสียงที่สูงขึ้น หรือ ต่ำลงเท่านั้น จากนั้นจึงนำค่านี้มีเปรียบเทียบกับหาความผิดพลาด และนำมาเฉลี่ยเป็นค่าความถูกต้องดัง สมการที่ (5)

$$E_i = \begin{cases} 0 & \text{when } \Delta A = \Delta B \\ 1 & \text{others} \end{cases}$$

โดยที่ ΔA = ผลเฉลี่ย

ΔB = ผลลัพธ์

E_i = ระดับเสียงที่ผิดพลาด

$$accuracy2 = 100 - \left(\frac{1}{N} \sum_{i=0}^n E_i \right) \times 100 \quad (5)$$

4. ผลการทดลอง

งานวิจัยนี้ได้ทำการประเมินความถูกต้องของขั้นตอนวิธีโดยเปรียบเทียบผลการทำงานโดยใช้วิธีการสกัดแบบโครมาแกรมซึ่งมีพื้นฐานมาจากระบบ 12 เสียงเท่ากันเป็นขั้นตอนวิธีอ้างอิง กับระบบที่นำเสนอที่มีพื้นฐานจากระบบ 7 เสียงไม่เท่ากัน พบว่ามีความถูกต้องดังแสดงใน ตารางที่ 3 ซึ่งแสดงให้เห็นว่าขั้นตอนวิธีการที่นำเสนอสามารถระบุระดับเสียงดนตรีไทยเดิมได้ดีกว่าขั้นตอนวิธีอ้างอิงแบบเทียบระดับเสียงแบบตำแหน่งต่อตำแหน่งถึงร้อยละ 72.5313 ± 9.1113 และมีประสิทธิภาพดีกว่าขั้นตอนวิธีอ้างอิงเมื่อเปรียบเทียบกับวิธีช่วงระดับคิดเป็นร้อยละ 70.1928 ± 8.1842

ตารางที่ 3 ประสิทธิภาพการจำแนกโน้ตเพลง

บทเพลง	เปอร์เซ็นต์ ความถูกต้อง ของวิธีการ อ้างอิง โดยวิธีเทียบ ตำแหน่งต่อ ตำแหน่ง	เปอร์เซ็นต์ ความถูกต้อง ของวิธีการ นำเสนอ โดยวิธีเทียบ ตำแหน่งต่อ ตำแหน่ง	เปอร์เซ็นต์ ความถูกต้อง ของวิธีการ อ้างอิง โดยวิธี Interval	เปอร์เซ็นต์ ความถูกต้อง ของวิธีการ นำเสนอ โดยวิธี Interval
บุหลันลอยเลื่อน	32.3873	68.1135	57.0952	74.4574
โหมโรงไอยเรศ 3 ชั้น*	37.0000	40.5714	<u>67.2857</u>	<u>64.5714</u>
ลาวคำหอม (ขลุ่ย)	51.6452	53.5050	67.8112	68.2403
ลาวดวงเดือน	36.8948	37.3957	78.1302	79.4658
ขับไม้ (ซอสามสาย)	70.9586	82.9757	84.1202	84.4063
แขกมอญบางช้าง 3 ชั้น	65.7143	64.5714	75.2857	78.5714
ลาวเสียงเทียน*	33.1429	39.0000	<u>62.0000</u>	<u>54.1429</u>
สร้อยเพลง (ขลุ่ย)	54.9356	54.6495	69.8140	76.3948
ค่าเฉลี่ย	47.8348 ± 14.1901	55.0978 ± 15.0938	70.1928 ± 8.1842	72.5313 ± 9.1113

*ดนตรีที่บรรเลง เป็นเสียงโครมาติก (12 เสียงเท่า) สำหรับโทนเสียงดนตรีสากล

จากข้อมูลตัวอย่าง 8 บทเพลง พบว่ามี 2 บทเพลงตัวอย่าง (ลาวเสียงเทียน, โหมโรงไอยเรศ 3 ชั้น) ร้อยละความถูกต้องของวิธีการอ้างอิงสูงกว่า ความถูกต้องของวิธีการที่นำเสนอ เนื่องจากดนตรีที่ใช้บันทึกเสียงเป็นดนตรีที่บรรเลงเป็นเสียงโครมาติก (12 เสียงเท่า) สำหรับโทนเสียงดนตรีสากล ซึ่งไม่ใช่คีย์ไทยเดิมตามที่ผู้ทำวิจัยตั้งขอบเขตการวิจัย นอกจากนั้นเพลงนี้ถูกประพันธ์ด้วยบันไดเสียงแบบเพนทาโทนิค (Pentatonic Scale) [Tyler Connaghan. 2023], [Phra Chen Duriyanga. 2015] ซึ่งมีเพียง 5 ระดับเสียงและระดับเสียงทั้ง 5 ระดับนี้ก็ใกล้เคียงกับระดับเสียงของระบบโครมาติกจึงส่งผลให้ผลการดำเนินงานของขั้นตอนวิธีที่นำเสนอใกล้เคียงกับขั้นตอนวิธีอ้างอิง

ซึ่งวิจัยในครั้งนี้ทำให้การรู้จำระดับเสียงของดนตรีไทยมีประสิทธิภาพมากยิ่งขึ้นพร้อมกับการสร้างอัลกอริทึมในการสร้างเอกลักษณ์เพลงการคำนวณหาค่าความถี่เสียงสูง-ต่ำ ในอนาคตสามารถต่อยอดในการพัฒนาระบบ หรือเป็นแนวทางของวิธีการรู้จำระดับเสียงดนตรี อาจจะใช้อ้างอิงฐานข้อมูลค่าความถี่เสียงของภูมิภาคเอเชียทั้งประเทศจีน, อินเดีย, พม่า หรืออื่นๆ ที่ใช้ทฤษฎีบันไดเสียงไดอะทอนิก 7 เสียงเท่ากัน หรือทฤษฎี 7 เสียงเท่ากันเช่นเดียวกันกับประเทศไทย

5. สรุป

วิจัยนี้นำเสนอขั้นตอนวิธีการสร้างเอกลักษณ์สำหรับใช้ในการค้นคืนสารสนเทศเสียงดนตรีไทยเดิม โดยอาศัยหลักการถดถอยเพื่อสร้างสมการเทียบความถี่ตัวโน้ตดนตรีไทยเดิมจากข้อมูลความถี่เสียงดนตรีไทยเดิมที่ผู้วิจัยได้ค้นคว้ารวบรวมมา จากการทดลองกับเพลงไทยเดิม 8 บทเพลงพบว่า สามารถระบุโน้ตเพลงได้ในระดับที่น่าพอใจและสามารถใช้สร้างเอกลักษณ์ที่ใช้ในการค้นคืนฐานข้อมูลเสียงดนตรีเพลงไทยเดิมได้อย่างมีประสิทธิภาพ สามารถสรุปได้ว่า โครมาแกรมมีประสิทธิภาพในการระบุตัวโน้ตดนตรีไทยเดิมได้ถูกต้อง อย่างไรก็ตามในงานวิจัยนี้ผู้วิจัยได้เลือกใช้ในการเทียบความถูกต้องแบบตำแหน่งต่อตำแหน่ง โดยไม่ได้พิจารณาถึงการบรรเลงเพลงที่มีความช้า-เร็ว (tempo) แตกต่างกันจึงส่งผลให้ค่าความถูกต้องในภาพรวมยังไม่ดีนัก แต่อย่างไรก็ตามขั้นตอนวิธีที่นำเสนอนี้สามารถพัฒนาต่อยอดเพื่อให้ได้ค่าความถูกต้องในการค้นคืนเพิ่มขึ้นได้อีก

6. เอกสารอ้างอิง

- จุฬาลงกรณ์มหาวิทยาลัย. (2542). การศึกษาค้นคว้าวิจัยเรื่อง **ความถี่ของเสียงดนตรีไทย โดย พระราชดำริในพระบาทสมเด็จพระเจ้าอยู่หัว**. กรุงเทพฯ: โรงพิมพ์เรือนแก้วการพิมพ์.
- ณัชชา พันธุ์เจริญ. (2556). **ทฤษฎีดนตรี**(พิมพ์ครั้งที่ 12). กรุงเทพฯ : สำนักพิมพ์เกษกรัต.
- ชนะรัชต์ อนุกุล. 2563. **ระบบเสียงดนตรีไทยกับการขับร้องเพลงไทยเดิม**. วารสารที่ทัศน์วัฒนธรรม. 19(1). 195-214.
- สงบศึก ธรรมวิหาร. (2567). **ดุริยางค์ไทย**(พิมพ์ครั้งที่ 4). กรุงเทพฯ : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- สรวิชัย สุจิตจร. (2546). **การวิเคราะห์เสียงดนตรีไทย**. นครราชสีมา : โรงพิมพ์เลิศศิลป์ (1994).
- สุกรี เจริญสุข. (2546). **เสียงและระบบดนตรีไทย**. กรุงเทพฯ : วิทยาลัยดุริยางศิลป์ มหาวิทยาลัยมหิดล.
- อุทิศ นาคสวัสดิ์. (2531). **ทฤษฎีและปฏิบัติดนตรีไทย ภาค ๑ ว่าด้วยหลักและทฤษฎีดนตรีไทย**. กรุงเทพฯ : สถาบันฝึกดนตรีและนาฏศิลป์ไทย.
- Asif Ghias, Jonathan Logan, David Chamberlin & Brian C Smith. (1995). **Query by humming: Musical information retrieval in an audio database**. In *Proceedings of the third ACM international conference on Multimedia*. pp.231–236.
- Boonmatham, Pheerasut & Pongpinigpinyo, Sunee & Soonklang, Tasanawan. (2013). **Musical-scale characteristics for Thai Classical Music genre classification**. 2013 International Computer Science and Engineering Conference, ICSEC 2013. pp 227-232. 10.1109/ICSEC.2013.6694784.
- Boonmatham, Pheerasut & Pongpinigpinyo, S. & Soonklang, Tasanawan. (2012). **A comparison of audio features of Thai Classical Music Instrument**. pp 213-218.
- International Organization for Standardization. (1975). **ISO 16:1975 Acoustics — Standard tuning frequency (Standard musical pitch)**. Switzerland.
- Uitdenbogerd, Alexandra & Zobel, Justin. (1999). **Matching Techniques for Large Music Databases**. *ACM MULTIMEDIA '99: Proceedings of the seventh ACM international conference on Multimedia (Part 1)*. 1(978-1-58113-151-2): 57-66.
- P. Khoenkaw & P. Piamsa-nga. (2014). **Video Similarity Measurement Using Spectrogram**, in *Proceeding of 2014 International Computer Science and*

- Engineering Conference (ICSEC 2014), pp 412-417, 2014, Khon Kaen, THAILAND.
DOI:10.1109/ICSEC.2014.6978241.
- P. Khoenkaw & P. Piamsa-nga. (2014). **N-Gram Signature for Video Copy Detection.**
in Recent Advances in Information and Communication Technology. vol. 265,
S. Boonkrong, H. Unger, and P. Meesad, Eds., ed: Springer International Publishing,
2014, pp. 335-344. DOI: 10.1007/978-3-319-06538-0_33
- Phra Chen Duriyanga. (2015). **THAI MUSIC (Thai Culture, New Series No.15) (7th ed.).**
Bangkok : Rung Silp Printing Company Limited.
- Rumsey, F., and McCr. Mick, T. (2006). **Sound & recording: An introduction.** Boston :
Elsevier/Focal Press.
- Shah, Ayush & Kattel, Manasi & Nepal, Araj & Shrestha, D.. (2019). **Chroma Feature
Extraction.**
- Tyler Connaghan. (2023). **Pentatonic Scale: The Ultimate Beginner's Guide.** Available
from: <https://emastered.com/blog/pentatonic-scale>. [22 March 2024]
- W. -j. Liu and J. -l. Feng. (2011). **Qualitative mapping model of humming melody
extraction.** *Proceedings of 2011 International Conference on Computer Science
and Network Technology*, Harbin, China, 2011, pp. 1650-1653. doi:
10.1109/ICCSNT.2011.6182283.