

# การเสริมสร้าง Pitch Class Profile ในการเรียนรู้และจดจำคอร์ด ด้วย Convolution Neural Network

## Enhancing Pitch Class Profile for Music Chord Recognition Using Convolution Neural Network

ลิตา วรรณตรง<sup>1</sup>, ณปภัช วรรณตรง<sup>2</sup>

<sup>1</sup>ภาควิชาคอมพิวเตอร์และสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ  
1518 ถนนประชากรราษฎร์ 1 แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพฯ 10800 โทรศัพท์ 02-5552000

E-mail: lita.wanatrong@gmail.com

<sup>2</sup>สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

439 ถ.จระ ต.ในเมือง อ.เมือง จ.บุรีรัมย์ 31000 โทรศัพท์ 044-611221

E-mail: wannatrong@hotmail.com

### บทคัดย่อ

คอร์ด คือเสียงโน้ตที่หลากหลายซึ่งสามารถเล่นให้เกิดเสียงได้ในเวลาเดียวกัน การแกะคอร์ดเป็นงานที่ยุ่งยากที่ต้องอาศัยผู้เชี่ยวชาญ จึงได้มีงานวิจัยจำนวนมากที่ศึกษาการเรียนรู้และจดจำคอร์ด (chord recognition) เพื่อช่วยในการแกะคอร์ดอัตโนมัติ PCP เป็นวิธีหนึ่งที่นิยมใช้ในการเรียนรู้และจำแนกคอร์ด แต่ยังไม่ดีนักเนื่องจากยังไม่สามารถขจัดสัญญาณที่ไม่ใช่ฮาร์โมนิก (non-harmonic) ได้อย่างมีประสิทธิภาพเพียงพอ ส่งผลให้ความถูกต้องในการจำแนกคอร์ดลดลง งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อขยายความสามารถ PCP เพื่อที่จะแก้ไขข้อจำกัด โดยจะนำเสนอวิธี Learning Feature โดยการสกัดคุณลักษณะ (Feature extraction) และจำแนกคอร์ด (Classification Feature) ไปพร้อม ๆ กันด้วย Convolution Neural network โดยงานวิจัยนี้ประกอบด้วย 2 ขั้นตอน ขั้นตอนแรกใช้ PCP ที่เพิ่มส่วนการกรองคอร์ด (chord filter) เข้าไป ขั้นตอนที่สองคือ learning feature ใน PCP ที่มีกระบวนการกรองคอร์ด การเรียนรู้และจดจำคอร์ดโดย PCP ที่มีกระบวนการกรองคอร์ดด้วย Convolution Neural network ให้ค่าความถูกต้องที่มากกว่า PCP ที่ไม่มีกระบวนการกรองคอร์ด 2.1%

คำสำคัญ: Chord recognition, Convolution Neural network, PCP

### Abstract

Chords are various notes sounding at the same time. Chord detection is a quite complicated task which needs skillful persons. So, there are many studies investigating chord recognition in order to detect chords automatically. PCP is a popular method used for chord recognition, but it is not quite good for chord classification since it is unable to remove non-harmonic signal effectively. This study aimed to extend PCP to solve its limits by proposing the methods of learning feature by feature extraction and classification together. This article consists of 2 steps. The first is to use PCP and add the process of chord filter according to chord filter working. The second one is learning feature on PCP with chord. The results revealed that Chord Recognition through PCP with chord filter and Convolution Neural network showed higher accuracy value than PCP with no chord filter 2.1%.

**Keywords:** Chord recognition, Convolution Neural network, PCP

### 1. คำนำ

ดนตรีกับมนุษย์มีความเกี่ยวพันกันมาเป็นเวลานาน โดยเกี่ยวข้องกับมนุษย์ในหลายด้านได้แก่ ประเพณีและวัฒนธรรมให้ความบันเทิง คอร์ด (Chord) คือการเล่นโดยประสานของกลุ่มโน้ต 3 โน้ต (Note) ขึ้นไป [1] การรู้และจดจำคอร์ด (Chord Recognition) คือกระบวนการของความพยายามที่จะแกะคอร์ดและจดจำ

คอร์ดจากเสียงเพลง การแกะคอร์ดด้วยตัวเองเป็นงานที่ยาก ต้องอาศัยความรู้และความสามารถด้านดนตรี

การรู้และจดจำคอร์ดอัตโนมัติ คือ การวิเคราะห์สัญญาณเพลงและระบบการรู้จำคอร์ดอัตโนมัติ ซึ่งประกอบด้วย 2 กระบวนการได้แก่การสกัดคุณลักษณะ (Features Extraction) และกระบวนการจำแนกคอร์ด (Classification Feature) [2]

วิธีหนึ่งที่นิยมใช้ในการสกัดคุณลักษณะ คือ PCP [3-5] แต่วิธีนี้ยังมีข้อจำกัดคือ ยังไม่สามารถขจัดสัญญาณที่ไม่ใช่ฮาร์โมนิก (non-harmonic) เช่น สัญญาณรบกวนชั่วคราว (transient noise) และฮาร์โมนิกโอเวอร์โทน (harmonic overtones) ได้ดีนัก ซึ่งสัญญาณเหล่านี้อาจจะสร้างพลังงานความถี่ใส่ลงในโครมาเวกเตอร์ (chroma vector) นอกเหนือจากโน้ตที่แท้จริงของคอร์ดที่เกี่ยวข้อง และการจำแนกคอร์ดในรูปแบบเดิมนั้นไม่สามารถ learning feature โดยการสกัดคุณลักษณะไปพร้อมกับการจำแนกคอร์ดได้ ดังงานวิจัยในหัวข้อ Chord Recognition for Symbolic Music Using Autoencoder ที่ใช้วิธี Autoencoder เปรียบเทียบกับการใช้ PCP ในการสกัดคุณลักษณะและใช้ Neural Networks ในกระบวนการจำแนกคอร์ดเพื่อจำแนกคอร์ดออกมา [6] ซึ่งปัจจุบันแนวโน้มการรู้และจดจำคอร์ดอัตโนมัตินั้นนิยมวิธีที่สามารถ learning feature โดยการสกัดคุณลักษณะไปพร้อมกับการจำแนกคอร์ดได้ เช่น Convolution Neural Network [7-10]

ดังนั้นงานวิจัยนี้จึงนำเสนอวิธี PCP ที่เพิ่มกระบวนการกรองคอร์ดซึ่งจะช่วยแก้ปัญหา PCP ที่มีอยู่เดิม และใช้ Convolution Neural network ที่สามารถ learning feature ใน PCP ไปพร้อมกับการจำแนกคอร์ดไปพร้อมกันได้ โดยหลักการทำงานของกรองคอร์ดคือจะทำการกรอง PCP โดยใช้ตัวกรองของแต่ละเมเจอร์คอร์ด (Major Chord) ให้ได้ลักษณะความเป็นคอร์ดที่แท้จริงที่อยู่ในมุมมองของเมเจอร์คอร์ดนั้น ๆ ออกมา

โดยในงานวิจัยนี้จะทำการเปรียบเทียบประสิทธิภาพในการจำแนกคอร์ดระหว่างวิธีที่นำเสนอคือ PCP ที่มีกระบวนการกรองคอร์ดด้วย Convolution Neural network และ PCP

ในส่วนต่อไปจะนำเสนอเนื้อหา ได้แก่ ข้อมูลที่ใช้ในงานวิจัย ขั้นตอนการวิจัย และวิธีการวัดประสิทธิภาพ ดังนำเสนอในหัวข้อถัดไป

## 2. วิธีที่นำเสนอ

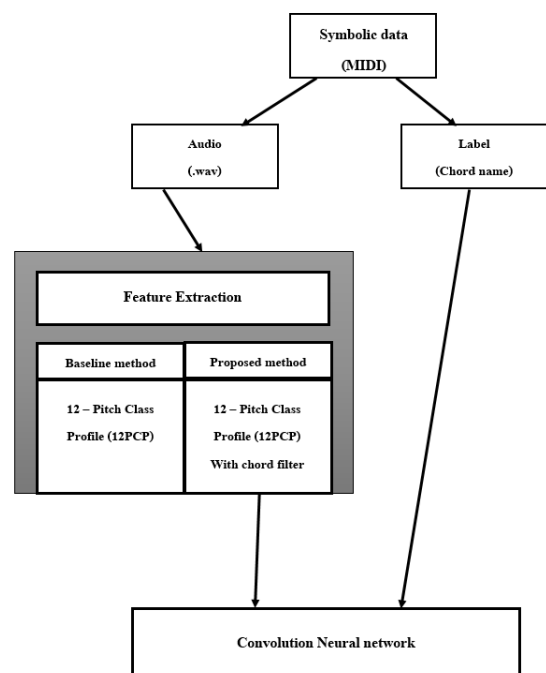
ในงานวิจัยนี้ได้ทำการเปรียบเทียบการรู้และจดจำคอร์ดระหว่างวิธี PCP และวิธีที่นำเสนอคือวิธี PCP ที่มีกระบวนการกรองคอร์ดด้วย Convolution Neural network ซึ่งสามารถแสดงข้อมูลที่ใช้งานวิจัย และขั้นตอนการวิจัย ได้ดังนี้

### 2.1 ข้อมูลที่ใช้ในงานวิจัย

เพลงที่นำมาใช้ในการวิจัยนี้ ประกอบด้วย เพลงเดี่ยวเปียโน (solo piano music) ประพันธ์โดย Beethoven โดยนำมาจากฐานข้อมูล CCARH (Computer Assisted Research in Humanities at Stanford University) ถูกบันทึกเป็นไฟล์ข้อมูลเชิงสัญลักษณ์ (Symbolic File) เก็บในรูปแบบไฟล์มิดิ (MIDI file) ที่อัตราการสุ่ม (Sampling Rate) เท่ากับ 44100 เฮิร์ตซ์ (Hertz: Hz) โดยข้อมูลเพลงที่ใช้ในการทำวิจัยทั้งหมดมีจำนวน 135,334 เรคอร์ด (Records)

### 2.2 ขั้นตอนการวิจัย

ได้ทำการเปรียบเทียบการรู้และจดจำคอร์ดระหว่างวิธีที่ 1 PCP และวิธีที่ 2 PCP ที่มีกระบวนการกรองคอร์ด โดยทั้งสองวิธีใช้การจำแนกคอร์ดด้วย Convolution Neural Network โดยมีขั้นตอนการวิจัยวิธีที่ 1 และ 2 ดังภาพ 1



ภาพที่ 1 : ขั้นตอนการวิจัย

จากภาพที่ 1 วิธีที่ 1 PCP และวิธีที่ 2 PCP ที่มีกระบวนการกรองคอร์ด มีรายละเอียดขั้นตอนดังนี้  
 2.2.1 PCP เป็นวิธีพื้นฐานที่ใช้ในงานวิจัยนี้

1) นำข้อมูลสัญญาณเสียงที่เป็นไฟล์ข้อมูลเชิงสัญลักษณ์ (symbolic file) จากฐานข้อมูล CCARH แบ่งเป็น 2 ส่วน

ส่วนแรก ทำการแปลงไฟล์เสียงที่อยู่ในรูปแบบไฟล์มิดิ (MIDI file) ให้อยู่ในรูปแบบไฟล์เสียงที่เก็บในรูปแบบไฟล์เวฟ (WAV file) โดยใช้เครื่องมือคือ โปรแกรม Timidity++ [11]

ส่วนที่สอง นำไฟล์เสียงที่อยู่ในรูปแบบไฟล์มิดิ (MIDI file) มาทำการสร้างไฟล์เลเบล (label files) โดยใช้เครื่องมือคือ โปรแกรม Melisma Music Analyzer [12]

2) แปลงจาก .wav เป็น PCP โดยใช้ Chromagram Toolbox

3) นำผลลัพธ์จากขั้นตอน 2) และไฟล์เลเบล จากขั้นตอน 1) มาทำการกระบวนการ Convolution Neural Network เพื่อหาคอร์ดที่ถูกตัด

2.2.2 PCP ที่เพิ่มกระบวนการกรองคอร์ดเป็นวิธีการที่นำเสนอในงานวิจัยนี้

1) และ 2) ทำกระบวนการเดียวกันกับขั้นตอน 1) และ 2) ของวิธี PCP พื้นฐาน

3) นำ PCP จาก ขั้นตอน 2) มาผ่านกระบวนการกรองคอร์ดเพื่อหาความปั่นคอร์ดที่แท้จริงออกมา โดยใช้ฟังก์ชัน min (PCP, Note) โดยการกรองผ่านกระบวนการกรองคอร์ดซึ่งมีอัลกอริทึมตามภาพที่ 2

```

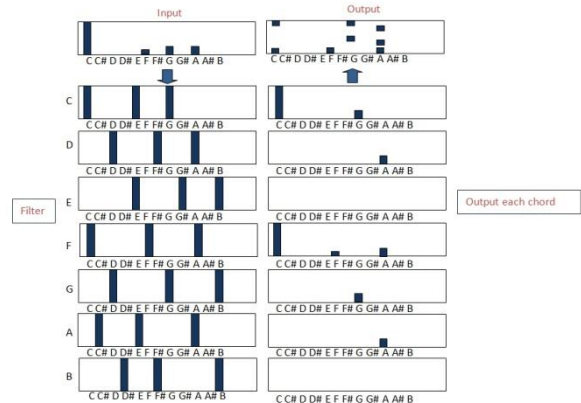
Algorithm Separating PCP with 7 major chord filters
Input : Pitch class profile PCP (1x12)
Output : 7x12 matrix that called 2D-PCP

% C C# D D# E F F# G G# A A# B
Ideal_PCP=[1 0 0 1 0 0 1 0 0 0 0; %C
           0 0 1 0 0 0 1 0 0 1 0 0; %D
           0 0 0 0 1 0 0 0 1 0 0 1; %E
           1 0 0 0 0 1 0 0 0 1 0 0; %F
           0 0 1 0 0 0 0 1 0 0 0 1; %G
           0 1 0 0 1 0 0 0 0 1 0 0; %A
           0 0 0 1 0 0 1 0 0 0 0 1; %B
           ];

For each Ideal_PCPi ∈ Nchord do
    PCPmin = min(PCP, Ideal_PCPi)
    2D-PCP.append(PCPmin)
End for
    
```

ภาพที่ 2 : อัลกอริทึมของกระบวนการกรองคอร์ด

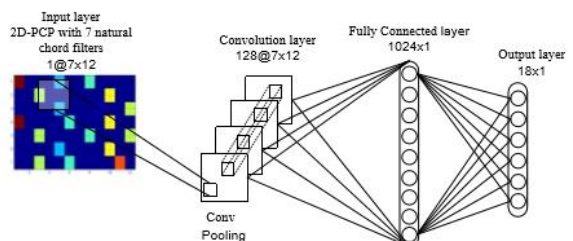
จากภาพที่ 2 กระบวนการกรองคอร์ดที่นำมาใช้ในการวิจัยนี้เกิดจากการที่คอร์ดประกอบด้วยโน้ตธรรมชาติ 3 โน้ต เช่น คอร์ด C จะประกอบไปด้วยโน้ต C, E, G เราก็จะกำหนดตัวกรอง (filter) มีค่าเป็น 1 ตามตำแหน่งโน้ตตามลำดับดังภาพ โดยตัวกรองที่เราใช้นี้มีทั้งหมด 7 คอร์ด ได้แก่ C, D, E, F, G, A และ B จากนั้นนำมาผ่านกระบวนการกรองคอร์ด เพื่อหาสัญญาณเสียงที่แท้จริง โดยใช้ฟังก์ชัน min ดังภาพที่ 2



ภาพที่ 3 : ตัวกรองคอร์ด

จากภาพที่ 3 นำคลื่นสัญญาณเสียงที่เป็นข้อมูลเข้า (input) มาผ่านการกรองซึ่งประกอบด้วยเมเจอร์คอร์ดจำนวน 7 คอร์ด ได้แก่ C, D, E, G, A, B, F ดังภาพฝั่งซ้ายมือ จะได้ผลลัพธ์ (output) คือสัญญาณที่ผ่านการกรองดังภาพฝั่งขวามือ และนำผลลัพธ์ทั้ง 7 คอร์ดมารวมกันอยู่ในรูปแบบของ 2D-PCP ตามภาพฝั่งขวามือบนสุดโดยผลลัพธ์ที่ได้จะออกมาเป็นลักษณะของคอร์ดที่แท้จริงตามมุมมองของคอร์ดนั้น ๆ และนำไปเป็นข้อมูลเข้าให้กับ Convolution Neural Network ต่อไป

4) ใช้ไฟล์เลเบลจากขั้นตอน 1) และใช้ PCP ที่ผ่านกระบวนการกรองคอร์ดจากขั้นตอน 3) และทำการสกัดคุณลักษณะ และจำแนกคลาสโดย Convolution Neural Network ดังภาพที่ 4



ภาพที่ 4 : กระบวนการ Convolution Neural Network

จากภาพที่ 4 ในกระบวนการ Convolution Neural Network ได้มีการกำหนดค่าพารามิเตอร์ต่าง ๆ ดังนี้ กำหนดโมเดล (Model) ให้มีจำนวนชั้นในส่วน Convolution layer จำนวน 1 ชั้น ใช้ตัวกรองขนาด 5x5 จำนวน 128 จำนวน strides เท่ากับ 1 จากนั้นผ่านชั้น Max Pooling layer จำนวน 1 ชั้นมีจำนวน strides เท่ากับ 1 และผ่านชั้น Fully connected layer เพื่อจำแนกคลาสจำนวน 1 ชั้น ในขั้นตอนการ train data ใช้วิธีการ train แบบ AdamOptimizer และได้กำหนดพารามิเตอร์ต่างๆ ได้แก่ Learning rate เท่ากับ 0.01, Train iteration ที่ 10,000,000 ใช้ Batch size ขนาด 100, Display step เท่ากับ 20, Training epochs เท่ากับ 15 และ Dropout เท่ากับ 0.75

### 2.3 การวัดประสิทธิภาพ

ใช้การวัดอัตราความถูกต้องด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่ารีคอล (Recall) และค่าวัดประสิทธิภาพ (F-measure) เพื่อการบ่งชี้ความถูกต้องของการจำแนกข้อมูลหรือความเชื่อถือได้ของวิธีการจำแนกแต่ละวิธีตามสมการข้างล่างนี้ [13]

$$accuracy = \frac{TN+TP}{TN+FP+TP+FN} \times 100\% \quad (1)$$

$$precision = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

### 3. ผลการทดลอง

ในการทดลองนี้ ผู้วิจัยได้ใช้การจำแนกคอร์ค 2 วิธี เพื่อที่จะเปรียบเทียบความสามารถในการจำแนกคอร์ค ได้แก่ โครงข่ายประสาทเทียม และ Convolution Neural Network โดยโครงข่ายประสาทเทียมถูกใช้เพื่อทดสอบว่าวิธีที่นำเสนอสามารถเพิ่มประสิทธิภาพในการจำแนกคอร์ค และ Convolution Neural Network ถูกใช้เพื่อแสดงให้เห็นถึงประโยชน์ของการ learning feature และจำแนกคอร์คไปพร้อมกัน

ผลการทดลองมาจากการทดลอง 2 ส่วน ได้แก่ การทดลองที่ 1 เป็นการทดสอบความสามารถในการจำแนกคลาสรหว่าง PCP ที่มีกระบวนการกรองคอร์ค กับ PCP ด้วยการโครงข่ายประสาทเทียม การทดลองที่ 2 เป็นการทดสอบ learning feature ด้วย Convolution Neural Network

ระหว่าง PCP ที่มีกระบวนการกรองคอร์คกับ PCP ได้ผลการทดลองดังตารางที่ 1

ตารางที่ 1: ผลการทดลองเปรียบเทียบประสิทธิภาพระหว่างวิธี PCP และ PCP ที่มีกระบวนการกรองคอร์ค

feature extraction	NN (Accuracy)	CNN			
		Accuracy	Precision	Recall	F-measure
PCP	65.96	94.18	83.05	74.4	78.48
PCP with filter	79.2	96.28	88.84	84.4	86.69

จากตารางที่ 1 แสดงให้เห็นว่าวิธี PCP ที่ไม่มีกระบวนการกรองคอร์ค เมื่อจำแนกคอร์คด้วยโครงข่ายประสาทเทียมได้ค่า accuracy 65.96% และเมื่อจำแนกคอร์คด้วย Convolution Neural network ได้ค่า accuracy 94.18%, ค่า precision 83.05%, ค่า Recall 74.4% และค่า F-measure 78.48%

และวิธี PCP ที่มีกระบวนการกรองคอร์ค เมื่อจำแนกคอร์คด้วยโครงข่ายประสาทเทียมได้ค่า accuracy 79.2% และเมื่อจำแนกคลาสรด้วย Convolution Neural network ได้ค่า accuracy 96.28%, ค่า precision 88.84%, ค่า Recall 84.4% และค่า F-measure 86.69%

จากผลการทดลองแสดงให้เห็นว่าวิธี PCP ที่ไม่มีกระบวนการกรองคอร์คให้ค่าความถูกต้องมากกว่า PCP ที่ไม่มีกระบวนการกรองคอร์คเมื่อใช้โครงข่ายประสาทเทียม และเมื่อร่วมกับการใช้ Convolution Neural network ยิ่งให้ค่าความถูกต้องที่สูงเพิ่มมากขึ้น ซึ่งแสดงว่า PCP ที่มีกระบวนการกรองคอร์คด้วย Convolution Neural network สามารถช่วยลดสัญญาณที่ไม่ใช่ฮาร์โมนิคได้ดีขึ้นและให้ค่าความถูกต้องที่มากขึ้น

### 4. สรุปผลการทดลอง

งานวิจัยนี้นำเสนอการสกัดคุณลักษณะแบบ PCP ที่มีกระบวนการกรองคอร์คเพื่อให้ได้ลักษณะความเป็นคอร์คที่แท้จริงที่อยู่ในมุมมองของเมเจอร์คอร์คนั้น ๆ เพื่อช่วยเพิ่มประสิทธิภาพในการรู้จำคอร์คจากเสียงเครื่องดนตรี และใช้ Convolution Neural network ในการจำแนกคลาสร และสกัดคุณลักษณะไปพร้อม ๆ กัน เพื่อการรู้และจดจำคอร์ค

ในงานวิจัยนี้คาดหวังว่าสร้างคุณลักษณะใหม่ที่มีประสิทธิภาพ และจากผลการทดลองที่ได้แสดงให้เห็นว่า

วิธีการที่นำเสนอเป็นวิธีการที่มีประสิทธิภาพที่สูงกว่าวิธีการ PCP

นอกจากนี้วิธีการที่นำเสนอจะมีประโยชน์ต่อการเรียนรู้ในการรู้และจดจำคอร์ด วิธีการที่นำเสนอจะช่วยในการพัฒนาแอปพลิเคชันเพื่อการเรียนรู้และจดจำคอร์ดอัตโนมัติ ซึ่งแอปพลิเคชันดังกล่าวจะช่วยประหยัดเวลาและเพิ่มความถูกต้องในการแกะคอร์ดสำหรับนักดนตรีมือสมัครเล่น หรือผู้สนใจทางดนตรีที่ค่อนข้างและเรียนรู้ด้วยตนเอง

## 5. เอกสารอ้างอิง

- [1] S. V. Vaseghi, "Multimedia Signal Processing Theory and Applications in Speech, Music and Communications", England: John Wiley & Sons, 2007 : pp 447-448.
- [2] Müller and Meinard, "Fundamentals of Music Processing", Springer International Publishing, 2015.
- [3] K. Lee, "Automatic chord recognition from audio using enhanced pitch class profile", Proc. International Computer Music Conference, (n.p.) : 306-313, 2006.
- [4] K. Muludi, Aristoteles, and A.F. S. Loupatty, "Chord Identification Using Pitch Class Profile Method With Fast Fourier Transform Feature Extraction", IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 3, No 1, May 2014
- [5] T. Fujishima, "Realtime Chord Recognition of Musical Sound a System Using Common Lisp Music", Proc. of International Computer Music Conference. Beijing, China, 1999.
- [6] V. Phongthongloa, S. Kamonsantiroj and L. Pipanmaekaporn, "Learning high-level features for chord recognition using Autoencoder Neural Network", Master Degree Thesis. King Mongkut's University of Technology North Bangkok Bangkok, Thailand, 2015.
- [7] E. J. Humphrey and J. P. Bello, "Rethinking Automatic Chord Recognition with Convolution Neural Networks", 11th International Conference on Machine Learning and Application, 2012.
- [8] F. Korzeniowski and G. Widmer, "Feature Learning For Chord Recognition: The Deep Chroma Extractor", Proceeding Of the 17th ISMIR Conference New York City, USA. August 7-11, 2016.
- [9] F. Korzeniowski and G. Widmer, "A fully Convolutional Deep Auditory Model For Musical Chord

Recognition", IEEE International Workshop of Machine Learning For Signal Processing 13-16, Salerno, Italy, 2016.

[10] J. Pons, T. Lidy and X. Serra, "Experimenting with Musically Motivated Convolutional Neural Network", IEEE, 2016.

[11] M. Izumo et al., "TiMidity++", Available : <http://timidity.sourceforge.net/> March, 2017.

[12] D. Sleator and D. Temperley, "The Melisma Music Analyzer", Available : <http://www.link.cs.cmu.edu/music-analysis/> March, 2017.

[13] A. Pacharawongsakda, "Introduction to Data Mining Techniques", Data Cute, 2014.