

ผลการทดลอง



96.00%

คือความแม่นยำสูงสุดของการจำแนกเสียง
ในสภาพที่ไร้เสียงรบกวน โดยการใช้ k เป็น 1
Euclidean Distance, Flat top windowing

โดยเฉลี่ย Flat top windowing จะให้ความแม่นยำสูงกว่า Window function อื่นๆ ในทุกเงื่อนไขและสภาพแวดล้อม แต่สำหรับ window function อีก 5 ตัวที่เหลือนั้น จะมีความแม่นยำใกล้เคียงกับกรณีที่ไม่ได้ใช้ window function

เมื่อ k=2 และ k=4 จะมีร้อยละความแม่นยำ ต่ำกว่าในกรณี k=1, k=3 และ k=5 ซึ่งหากพิจารณาจาก ร้อยละของกรณีที่ไม่สามารถทำนายได้แล้ว จะพบว่าสาเหตุหลักเกิดจากที่เมื่อ k=2 และ k=4 จะมีร้อยละของกรณีที่ไม่สามารถทำนายได้สูงกว่า

ค่าเฉลี่ยของร้อยละความแม่นยำในแต่ละ Test case เป็นดังนี้

1: ไม่มี Noise	60.56	ซึ่งจะเห็นได้ว่า ค่าของความถูกต้องจะแปรผันกับอัตราส่วนของ Noise
2: Noise 10%	58.90	
3: Noise 25%	47.33	
4: Noise 50%	36.13	
5: Noise 75%	30.70	

หากลองสร้าง confusion matrix โดยนำผลลัพธ์ของ

การทดลองทั้งหมดมารวมกัน ได้ดังนี้

Predict \ Expect	Piano	Guitar	Violin
Piano	21.34 %	1.34 %	6.13 %
Guitar	10.25 %	14.37 %	4.01 %
Violin	15.72 %	0.62 %	11.02 %

ไม่สามารถทำนายได้ : 15.20 %

อภิปรายผลการทดลองและข้อเสนอแนะ

การทดลองนี้นั้น ถูกทดลองขึ้น โดยมีขอบเขตของการทดลองค่อนข้างจำกัด คือเป็นการจำแนกเสียงของเครื่องดนตรีเพียง 3 ชนิด และจำแนกเสียงที่มีลักษณะเป็นโน้ตเดียวต่อ 1 ไฟล์เท่านั้น ดังนั้นในอนาคต จึงสามารถต่อยอดไปยังในกรณีที่เสียงที่นำมาจำแนกมีหลายชนิด และไม่จำกัดประเภทของเสียงที่จำแนกได้ รวมถึงอาจจะเพิ่มจำนวนของ test data ขึ้นได้มากกว่านี้ ซึ่งจะเพิ่มความถูกต้องของอัลกอริทึมได้ และนอกจากนี้ เนื่องจากเมื่อ Noise เพิ่มขึ้น จะทำให้อัตราการตอบถูกลดลงอย่างเห็นได้ชัด การเพิ่มส่วนของขั้นตอนที่ใช้ในการลด Noise จึงเป็นสิ่งที่ควรกระทำ และจากที่กรณีที่ไม่สามารถทำนายได้เป็นหนึ่งในตัวแปรสำคัญต่อความแม่นยำในการทำนาย ดังนั้นจึงควรมีการจัดการกับกรณีเหล่านี้ด้วย

แหล่งอ้างอิง

Adam Głowacz, Witold Głowacz, Andrzej Głowacz. (2553). Sound Recognition of Musical Instruments with Application of FFT and K-NN Classifier with Cosine Distance. Automatyka.

Erik de Castro Lopo. (2545). Introduction to Audio Digital Signal Processing on Linux. (ออนไลน์).

แหล่งที่มา : <http://www mega-nerd.com/Res/IADSPL/>

Robert A. Schilling, Sandra L. Harris. (2556). Introduction to Digital Signal Processing using MATLAB 2nd Edition. พิมพ์ครั้งที่ 2. Cengage Learning. Boston. US.

Toni Heittola, Anssi Klapuri, Tuomas Virtanen. (2552). Musical Instrument Recognition in Polyphonic Audio Using Source-Filter Model for Sound Separation. Kobe International Conference Center. Kobe. Japan.

การศึกษาผลของ Window function ต่อความแม่นยำในการจำแนกเสียงเครื่องดนตรีด้วย FFT และ K-NN Classifier



Study of effect of window function on musical instrument classification with FFT and K-NN Classifier

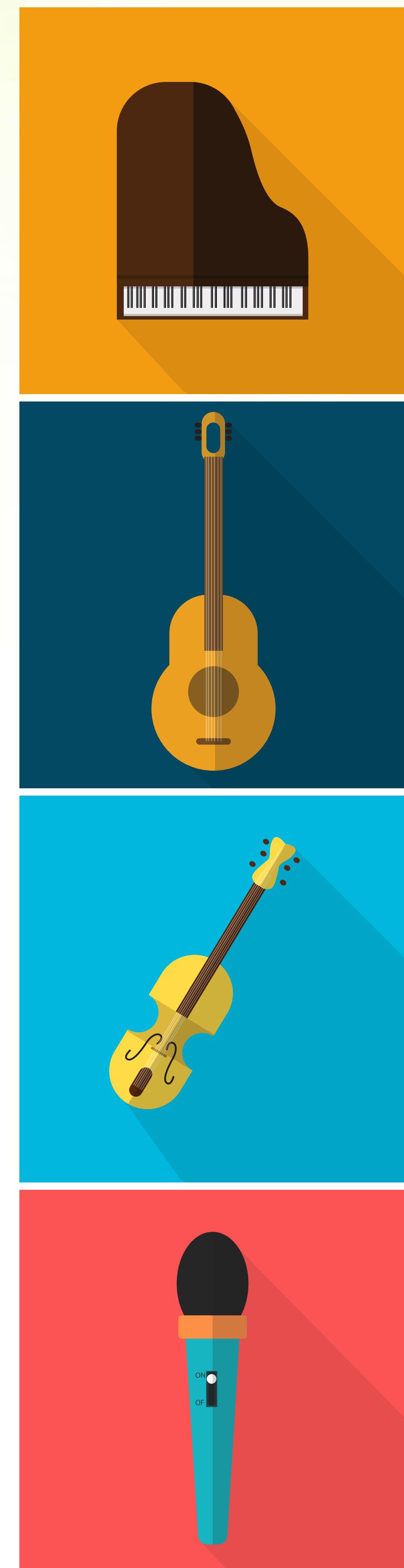
ผู้พัฒนา นายเบนรุจ บุณยะเพล็ก นายจิตรภาณุ อัศวพิชญ์ชิต นายพุฒิพัฒน์ สีริเศรษฐภักดี
ครุฑีปรีกษา นายอภินันท์ เงินมูล Contact : korla.march@gmail.com

รหัสโครงการ :
Y19TCSC328

โรงเรียนสามเสนวิทยาลัย

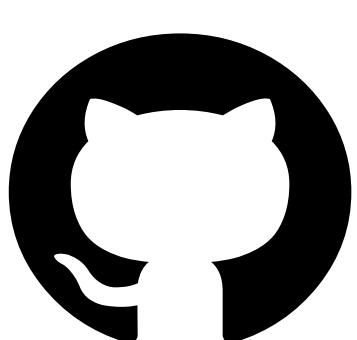
บทคัดย่อ

ปัจจุบัน การจำแนกเสียงเป็นอุปกรณ์ประเภทต่างๆ โดยการใช้คอมพิวเตอร์ได้รับความนิยมสูงมาก
เนื่องจากสามารถนำไปประยุกต์ใช้หลากหลาย ยกตัวอย่างเช่น การแปลเสียงออกมาเป็นข้อความ,
การยืนยันตัวตนด้วยเสียง หรือการนำเสียงของผู้ร้ายไปหาผู้ต้องสงสัยในคดีอาชญากรรม เป็นต้น
ซึ่งสิ่งเหล่านี้ ต่างต้องการอัลกอริทึมที่สามารถจำแนกเสียงออกเป็นกลุ่มต่างๆ ได้อย่างรวดเร็วและ
แม่นยำด้วยกันทั้งสิ้น ดังนั้น พวกเรามีสิ่งนี้มาอัลกอริทึมในการจำแนกเสียงของเครื่องดนตรี
3 ชนิด ได้แก่ กีตาร์ เปียโน และ ไวโอลิน โดยที่จะนำ Window functions ชนิดต่างๆ ได้แก่
Rectangular window, Triangular window, Hanning window, Hamming window,
Blackman-Harris window และ Flat top window มาใช้ร่วมกับการจำแนกเสียงด้วย
Fast Fourier Transform และ k-Nearest Neighbors (k-NN) เพื่อเพิ่มความแม่นยำในการ
จำแนกเสียง ซึ่งสำหรับ Nearest Neighbors เราได้ทดลองแบบห้ามใช้ Euclidean Distance
และ Manhattan Distance ในการหา นอกเหนือนี้ยังได้ทดลองกับค่า k (ใน k-NN) ตั้งแต่ 1-5
โดยได้พัฒนาโปรแกรมในภาษา C++ และ Node.js ขึ้น เพื่อทดลองสมมุติฐาน และทดสอบความ
ถูกต้องของอัลกอริทึม ซึ่งเราได้ใช้เสียงตัวอย่างจากเครื่องดนตรี 3 ชนิด ชนิดละ 50 เสียง รวม
150 เสียง มาทดสอบ โดยการทดสอบจะถูกจัดขึ้นในสภาพแวดล้อมที่ต่างกัน 5 ชุด คือมีเสียงรบกวน
0%, 10%, 25%, 50%, 75% การพยายามหลีกเลี่ยงการทดสอบ เราสามารถสรุปผลการทดลองได้ว่า
การทดลองนั้นเป็นไปตามสมมุติฐานที่ตั้งไว้ นั่นคือการใช้ Window function ช่วยเพิ่มความแม่นยำ
ของการจำแนกเสียงได้จริง และในกรณีที่ใช้ k=1, Euclidean Distance และ Flat top windowing
จะมีอัตราการตอบถูกที่สูงที่สุดในสภาพแวดล้อมที่ไม่มีเสียงรบกวนถึงร้อยละ 96.00 และการใช้ค่า k เป็น
1,3,5 จะมีประสิทธิภาพในการแยกแยะเสียงสูงกว่ากรณีที่ค่า k เป็น 2,4



วัตถุประสงค์

สร้างโปรแกรมที่สามารถ
การจำแนกเสียงเครื่องดนตรี
ได้มีประสิทธิภาพและความแม่นยำ



All source code is available on
github.com/KorlaMarch/sound-identification

กลุ่มเป้าหมาย



ผู้พัฒนาแอ��陌เพลเคชั่นและโปรแกรมที่ต้องใช้การจำแนกเสียง
 เช่น แอพหาชื่อเพลงจากเสียงที่รับเข้ามา, แอพแปลงเสียง
 เป็นข้อความ (speech to text), ระบบ Tag เพลงอัตโนมัติ

ขั้นตอนในการดำเนินงาน



1. ศึกษาทฤษฎีและแนวทางในการพัฒนา
ศึกษาเรื่อง Fast fourier transform
K-nearest neighbor และ Window function



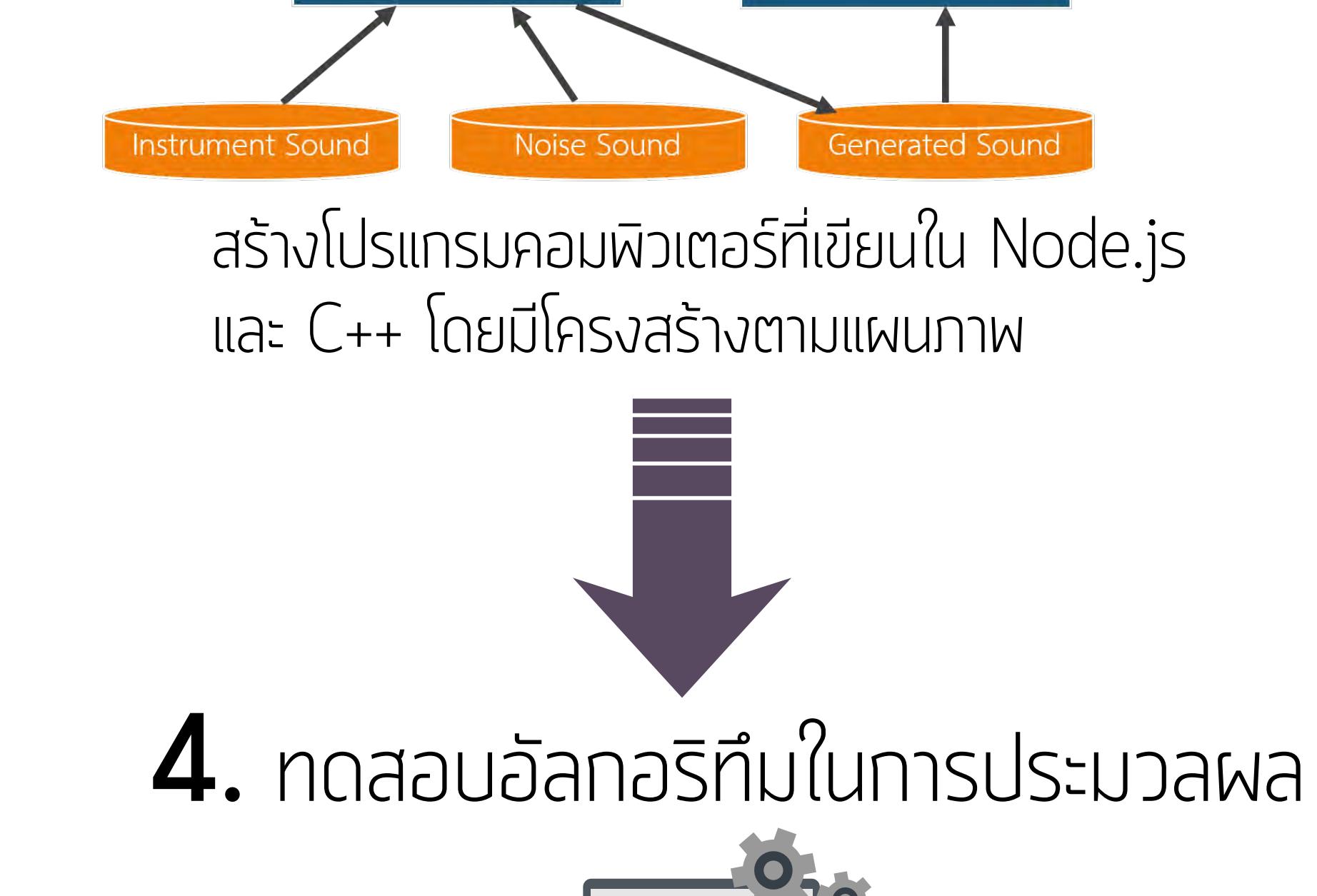
2. เก็บรวมข้อมูลเสียง

โดยเก็บข้อมูลเสียงจากเครื่องดนตรี 3 ชนิด
คือ เปียโน กีตาร์ และ ไวโอลิน รวมถึงหา
เสียงรบกวน (Noise) อีกจำนวน 10 เสียง

โดยเก็บข้อมูลเสียงจากเครื่องดนตรี 3 ชนิด
คือ เปียโน กีตาร์ และ ไวโอลิน รวมถึงหา
เสียงรบกวน (Noise) อีกจำนวน 10 เสียง



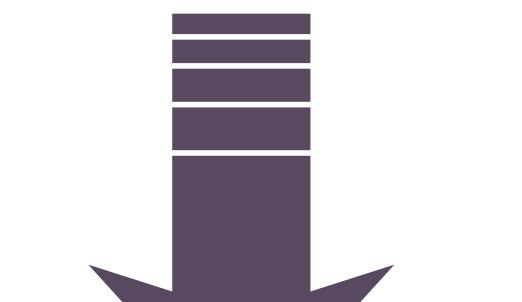
3. พัฒนาโปรแกรมคอมพิวเตอร์



สร้างโปรแกรมคอมพิวเตอร์ที่เขียนใน Node.js
และ C++ โดยมีโครงสร้างตามแผนภาพ



4. ทดสอบอัลกอริทึมในการประมวลผล



ทดสอบโปรแกรมโดยทดลองกับตัวแปรตันหลายๆตัว
คือ ชนิดของ Window function, Distance
function, ค่าคงที่ K และปริมาณ Noise ในชุดเสียง
ทดสอบ โดยมีการทดสอบทั้งสิ้น 180,000 ครั้ง



5. สรุปและอภิปรายผล

วิเคราะห์ผลลัพธ์ที่ได้จากขั้นตอนที่ 4 เพื่อหา
ความถูกต้องของอัลกอริทึม โดยการคำนวณ
ร้อยละความถูกต้องและสร้าง confusion matrix