ระบบรู้จำภาษามือ

Sign Language Recognition System

จุฑาพิชญ์ โสวจัสสตากุล

ชลิดา ธรรมศร

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2556

ระบบรู้จำภาษามือ

Sign Language Recognition System

จุฑาพิชญ์ โสวจัสสตากุล

ชลิดา ธรรมศร

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2556

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2556

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบรู้จำภาษามือ

Sign Language Recognition System

ผู้จัดทำ

1. นางสาวจุฑาพิชญ์	โสวจัสสตากุล	รหัสนักศึกษา	53010242
2. นางสาวชลิดา	ธรรมศร	รหัสนักศึกษา	53010325

..... อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.สมศักดิ์ วลัยรัชต์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบรู้จำภาษามือ

นางสาวจุฑาพิชญ์	โสวจัสสตากุล	53010242
นางสาวชลิดา	ธรรมศร	53010325
ผศ. ดร.สมศักดิ์	วลัยรัชต์	อาจารย์ที่ปรึกษา

บทคัดย่อ

ระบบรู้จำภาษามือเป็นระบบซึ่งใช้ในการแปลภาษามือภาษาไทยโดยอาศัยกล้องไคเนค ตรวจจับการเคลื่อนไหวของผู้พิการทางการได้ยินที่เป็นผู้แสดงท่าทางภาษามือ กล้องไคเนคจะทำการบันทึกตำแหน่งการเคลื่อนที่ของข้อต่อข้อมือซ้าย ข้อศอกซ้าย ข้อมือขวา และข้อศอกขวาของผู้แสดง แล้วนำมาเปรียบเทียบกับท่าทางภาษามือของผู้พิการทางการได้ยิน แสดงผลในรูปแบบคำศัพท์อย่างง่าย นอกจากนั้นระบบยังทำการตรวจจับภาพมือด้านขวา นำภาพมาวิเคราะห์ประมวลผลด้วยเทคนิคการประมวลผลภาพเพื่อค้นหาตำแหน่งปลายนิ้วมือ ตำแหน่งปลายนิ้วมือที่ได้มาจะถูกเปรียบเทียบกับภาพมือที่แสดงโดยผู้พิการทางการได้ยิน แสดงผลเป็นการสะกดพยัญชนะตัวอักษร ทำให้คนปกติสามารถเข้าใจได้ว่าการแสดงภาพมือนั้นหมายถึงพยัญชนะตัวอักษรใด ในส่วนของข้อมูลการเคลื่อนที่ของตำแหน่งข้อต่อต่าง ๆ และข้อมูลตำแหน่งปลายนิ้วที่ประมวลผลได้จะถูกรวบรวมสร้างเป็นฐานข้อมูล เพื่อใช้ในการแปลภาษามือ โครงงานนี้เมื่อพัฒนาเสร็จสมบูรณ์แล้ว จะได้นำไปทดสอบและประเมินประสิทธิภาพการทำงานจริงกับผู้พิการทางการได้ยิน เพื่อนำผลการประเมินมาปรับปรุงระบบให้มีประสิทธิภาพยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Sign Language Recognition System

Ms. Jutapid	Sovajastarkul	53010242
Ms. Chalida	Thammasorn	53010325
Asst.Prof. Dr. Somsak	Walairacht	Advisor

ABSTRACT

This project proposes a Sign Language Recognition System that translates sign language into spoken Thai language using Kinect camera. While the deaf person performs sign language, a Kinect camera is used to record movement of joint and then being translated positions at the left hand, the left elbow, the right hand and the right elbow. The trajectories of joint positions are compared with pose of sign language into simple words which normal person can understand. Moreover, the system is able to detect the depth images of left and right hands of sign language perform and analyze them with image processing techniques for finding fingertip positions with hand images spelling corresponding alphabet by the deaf person, normal person can understand what that pose is. The trajectories of joint positions and fingertip positions are collected and used to construct sign language database. In the final stage of development of this project, the system will be evaluated in real usages for improving system efficiency.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การที่ปริญญานิพนธ์ฉบับนี้สำเร็จลงได้ก็ด้วยความช่วยเหลือจาก ผศ.ดร.สมศักดิ์ วลัยรัชต์ อาจารย์ที่ปรึกษาประจำโครงงานนี้ ที่ได้ให้คำปรึกษา คำแนะนำต่าง ๆ มากมาย รวมทั้งความถูกต้องในการทำงาน ทำให้ข้าพเจ้าได้รับประสบการณ์ในการทำงานมาอย่างมากมาย ข้าพเจ้ารู้สึกซาบซึ้งในความอนุเคราะห์จากท่านอาจารย์ และขอขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณคณาจารย์สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกท่านที่ได้ให้ความรู้ ให้คำปรึกษาตลอดจนความช่วยเหลือ ตลอดระยะเวลาที่ผ่านมา

ขอขอบคุณเพื่อน ๆ รุ่นพี่ และรุ่นน้องคณะวิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ทุกคนที่ให้คำแนะนำ ให้กำลังใจ ให้ความร่วมมือในการทดลอง และช่วยเหลือในด้านต่าง ๆ มาโดยตลอด

ขอกราบขอบพระคุณบิดามารดาและครอบครัวที่คอยเป็นกำลังใจและสนับสนุนในทุก ๆ เรื่องที่ทำให้ข้าพเจ้าสามารถทำปริญญานิพนธ์นี้สำเร็จลุล่วงด้วยดี

สุดท้ายนี้ขอให้คุณค่าและประโยชน์อันพึงมาจากปริญญานิพนธ์ฉบับนี้ มอบแด่ผู้มีพระคุณทุกท่าน

นางสาว จุฑาพิชญ์ โสวจัสสตากุล

นางสาว ชลิดา ธรรมศร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญตาราง VII

สารบัญรูป VIII

บทที่ 1 บทนำ 1

1.1 ความสำคัญและที่มาของโครงงาน 1

1.2 วัตถุประสงค์ของโครงงาน 2

1.3 ขอบเขตของโครงงาน 2

1.4 วิธีการดำเนินการ 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

1.6 ส่วนประกอบของปริญญานิพนธ์ 3

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 4

2.1 ภาษามือ 4

2.2 เครื่องมือที่ใช้ 4

2.2.1 อุปกรณ์ด้าน Hardware ของไคเนค 5

2.2.2 อุปกรณ์ด้าน Software ของไคเนค 6

2.2.3 ขั้นตอนการเตรียมภาพ (Preprocessing) 7

2.2.3.1 มาตรการวัดของไคเนค (Kinect recalibration) 7

2.2.3.2 การกรองข้อมูลความลึก (Depth Data Filtering) 7

2.2.4 หลักการทำงานของไคเนค 8

2.2.4.1 การตรวจจับวัตถุและการติดตาม 8

2.2.4.2 การทำงานของไคเนค 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.3 การประยุกต์ใช้งาน 9
2.4 ทฤษฎีเกี่ยวกับ Skeleton Tracking 9
2.5 ทฤษฎีเกี่ยวกับ Fingertip Tracking 10
2.6 ทฤษฎีเกี่ยวกับ Dynamic Time Warping (DTW) 13
2.6.1 หลักการทำงานของ Dynamic Time Warping (DTW) 13
2.6.1.1 การหา gesture ที่สนใจ 13
2.6.1.2 การคำนวณค่า cost ด้วย matrix 13
2.6.1.3 การคำนวณค่า cost สะสมต่ำสุดด้วย matrix 14
2.6.1.4 การหาระยะทางที่สั้นที่สุด 14
2.7 งานวิจัยที่เกี่ยวข้อง 15

บทที่ 3 การออกแบบระบบ 18
3.1 ภาพรวมของระบบ 18
3.2 การแปลภาษามือ 19
3.2.1 การสะกดพยัญชนะ 19
3.2.2 การแปลคำ 22
3.3 ขั้นตอนการทำงานของระบบ 23
3.4 การดึงคุณลักษะ (Feature extraction) 25
3.4.1 ข้อมูลของข้อต่อที่สนใจ 25
3.4.2 ข้อมูลของมือส่วนที่สนใจ 25
3.4.3 ขบวนการปรับปรุงค่าข้อมูล 26
3.5 กระบวนการแปล 27
3.6 ส่วนติดต่อผู้ใช้ 27

บทที่ 4 การทดลอง และผลการทดลอง 29
4.1 การทดลองการติดตั้งอุปกรณ์ 29

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

4.1.1 เครื่องมือที่ใช้ในการพัฒนา 29
4.1.2 การทดลองการหาระยะระหว่างกล้องและผู้ใช้ 29
4.1.2.1 การหาระยะทางระหว่างกล้องและผู้ใช้งานในโหมดการแปลคำ 30
4.1.2.2 การหาระยะทางระหว่างกล้องและผู้ใช้งานในโหมดการแปลพยัญชนะ 33
4.2 การสร้างฐานข้อมูลสำหรับการแปลคำ 35
4.3 การทดสอบประสิทธิภาพการแปลคำ 36
4.3.1 การทดลองหาค่า BufferSize 36
4.3.2 การทดลองหาค่า Ignore และ MinimumFrames (Match frame) 37
4.3.3 เปอร์เซ็นต์ความคลาดเคลื่อนของการแปลคำ 39
4.4 การทดสอบประสิทธิภาพการแปลพยัญชนะ 41

บทที่ 5 บทสรุปและข้อเสนอแนะ 44
5.1 บทสรุป 44
5.2 ปัญหาอุปสรรคต่าง ๆ 45
5.3 แนวทางในการพัฒนาต่อ 46

บรรณานุกรม 47

ภาคผนวก ก 49
ก1 ท่าทางภาษามือของคำที่ใช้ในการทดลอง 50
ก2 ท่าทางภาษามือของพยัญชนะต่าง ๆ ที่ใช้ในการทดลอง 52
ภาคผนวก ข 54
ข1 การใช้งานโปรแกรม 55
ภาคผนวก ค 57

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

4.1 ผลของการตรวจจับข้อต่อที่ระยะต่าง ๆ ในท่าทางยืน 31
4.2 ผลของการตรวจจับข้อต่อที่ระยะต่าง ๆ ในท่าทางนั่ง 32
4.3 ผลของการตรวจจับจุดปลายนิ้วที่ระยะต่าง ๆ ในท่าทางยืน 33
4.4 ผลของการตรวจจับจุดปลายนิ้วที่ระยะต่าง ๆ ในท่าทางนั่ง 34
4.5 ขนาดข้อมูลการเคลื่อนไหวของท่าทางแต่ละคำ 37
4.6 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 1 ค่า BufferSize เท่ากับ 34 และค่า MinimumFrames ที่ระดับต่าง ๆ 38
4.7 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 2 ค่า BufferSize เท่ากับ 17 และค่า MinimumFrames ที่ระดับต่าง ๆ 38
4.8 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 3 ค่า BufferSize เท่ากับ 11 และค่า MinimumFrames ที่ระดับต่าง ๆ 39
4.9 ผลลัพธ์การแปลคำจากการแสดงท่าทางของผู้ใช้จำนวน 10 คนขณะนั่ง 40
4.10 ผลลัพธ์การแปลพยัญชนะจากการแสดงท่าทางของผู้ใช้จำนวน 10 คน 43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ **หน้า**

2.1 ภาพแสดงแบบสะกดอักษรนิ้วมือประดิษฐ์ของคุณหญิงกมลา ไกรฤกษ์ [1] 4

2.2 ภาพกล้อง Microsoft Kinect XBOX 360™ [5] 5

2.3 ภาพแสดงส่วนประกอบของกล้อง Microsoft Kinect XBOX 360TM [6] 5

2.4 (ก) ภาพสี RGB (ข) ภาพความลึก (Depth image) [7] 5

2.5 ภาพแสดงระบบพิกัดของกล้อง Microsoft Kinect XBOX 360TM 8

2.6 ภาพแสดงระยะห่างระหว่างเซนเซอร์กับคนหรือวัตถุ หน่วยเมตร (m) [10] 10

2.6 ภาพแสดงระยะห่างระหว่างเซนเซอร์กับคนหรือวัตถุ หน่วยเมตร (m) [10] 10

2.7 ภาพแสดงข้อมูลภาพมือ 11

2.8 ภาพแสดงตำแหน่งที่เป็น Convexity Defects [11] 11

2.9 ภาพแสดงผลลัพธ์การหา Convexity Defects 11

2.10 ภาพแสดงจุด Defect [13] 12

2.11 ภาพแสดงตำแหน่งปลายนิ้วตำแหน่งระหว่างนิ้วมือ 12

2.12 ภาพผลลัพธ์การทำ Convexity Defects [14] 13

2.13 ภาพแสดง Matrix ของเฟรมมือ 14

2.14 ภาพแสดง User Interface ของงานวิจัยของ Daniel Martínez Capilla 16

2.15 ภาพแสดงการตรวจจับมือและนิ้วมือด้วยกล้องไคเนค 16

2.16 ภาพแสดงคุณลักษณะต่าง ๆ ของโปรแกรม 17

3.1 ภาพรวมของระบบ 18

3.2 ตัวอย่างการสะกดพยัญชนะภาษามือไทย 20

3.3 ท่าพยัญชนะที่ต่อเนื่องของพยัญชนะ “ก” 20

3.4 พยัญชนะที่สามารถแปลได้ในระบบ 22

3.5 ภาษามือภาษาไทย 23

3.6 แผนผังขั้นตอนการทำงานของระบบ 24

3.7 ตำแหน่งข้อต่อที่สนใจ 25

3.8 ข้อมูลภาพมือ 26

3.9 ค่าคงที่จากการปรับปรุงค่าข้อมูล 27

3.10 User Interface 28

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่ **หน้า**

4.1 ตำแหน่งการติดตั้งกล้องไคเนค 28
4.2 หน้าต่างแสดงรายชื่อคำศัพท์ที่บันทึกลงในฐานข้อมูล 35
4.3 ตัวอย่างข้อมูลที่บันทึกไว้ในรูปของ Textfile 36
4.4 ท่าทางพยัญชนะท่ามือเดียว 41
4.5 ท่าทางการนับสำหรับพยัญชนะสองท่ามือ 41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความสำคัญและที่มาของโครงงาน

ผู้พิการทางการได้ยิน ที่ไม่สามารถสื่อสารด้วยคำพูดเหมือนเช่นคนปกติทั่วไปได้ แต่สายตาของเขาปกติ มองเห็นกิริยาอาการ ท่าทางต่าง ๆ ที่เคลื่อนไหวไปมาได้ ภาพต่าง ๆ ที่เห็นนั้นเป็นสื่อทำให้ผู้พิการทางการได้ยินเรียนรู้ความหมาย แม้จะเข้าใจได้ไม่มาก แต่ก็เป็นส่วนหนึ่งที่มีอิทธิพลให้ผู้พิการทางการได้ยินพยายามใช้ท่าทาง ร่างกาย และสีหน้าเพื่อแสดงความรู้สึกของเขาให้คนอื่นเข้าใจถึงความต้องการได้บ้าง ท่าทางที่แสดงนั้น เราจะสังเกตได้ว่าเป็นท่าทางที่เลียนแบบธรรมชาติมากที่สุด และท่าทางจากธรรมชาตินั้นเอง ก็ได้วิวัฒนาการต่อไปเป็นท่าที่ใช้มือเป็นส่วนใหญ่ เกิดเป็นท่าทางใช้แทนความหมายในคำพูดของคนทั่วไปได้อย่างสะดวก เราเรียกภาษาท่าทางแบบนี้ว่า “ภาษามือ”

ภาษามือ คือ ภาษาสำหรับผู้พิการทางการได้ยิน ใช้มือ สีหน้า และกิริยาท่าทาง ประกอบในการสื่อความหมาย และถ่ายทอดอารมณ์แทนการพูด ภาษามือของแต่ละชาติมีความแตกต่างกันเช่นเดียวกับภาษาพูด ซึ่งแตกต่างกันตามขนบธรรมเนียม ประเพณี วัฒนธรรม และลักษณะภูมิศาสตร์ เช่น ภาษามือจีน ภาษามืออเมริกัน และภาษามือไทย เป็นต้น ภาษามือเป็นภาษาที่นักการศึกษาทางด้านการศึกษาของผู้พิการทางการได้ยินตกลงและยอมรับกันแล้วว่า เป็นภาษาหนึ่งสำหรับติดต่อสื่อความหมาย ระหว่างผู้พิการทางการได้ยินด้วยกัน และระหว่างผู้พิการทางการได้ยินกับคนปกติทั่วไป ภาษาอังกฤษเรียกการสื่อสารด้วยมือนี้ว่า Sign Language หรือ Manual Communication [1]

ระบบภาษามือที่ใช้อยู่ในกลุ่มผู้พิการทางการได้ยินชาติต่าง ๆ มี 2 แบบคือ ภาษามือธรรมชาติ และภาษามือประดิษฐ์ การเลือกระบบที่ใช้ขึ้นอยู่กับความสามารถและความรู้ภาษามือของคู่สนทนาทั้งคู่ ภาษามือธรรมชาติ ผู้พิการทางการได้ยินเป็นผู้สร้างและใช้ร่วมกันในชุมชนหรือในแต่ละชาติ เป็นท่าเลียนแบบธรรมชาติ ส่วนภาษามือประดิษฐ์ เป็นภาษามือที่ครู ผู้ปกครอง หรือญาติมิตรของผู้พิการทางการได้ยินคิดขึ้นแทนคำพูดและภาษาเขียนประจำชาติ โดยโครงสร้างของภาษามือนั้นประกอบด้วย ท่ามือ ตำแหน่งของมือ การเคลื่อนไหวของมือ และทิศทางของฝ่ามือ [2] แต่ก็ยังเกิดความสับสนของการใช้ภาษามืออยู่บ้าง เนื่องจากพบว่าในประเทศไทยแบ่งการใช้ภาษามือออกเป็น 2 แบบ คือ แบบที่ถูกต้อง และแบบที่ประดิษฐ์ขึ้นโดยคนปกติ ซึ่งไม่ใช่แบบที่ถูกต้อง จึงมีการริเริ่มจัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ ดัดแปลง เผยแพร่ หรือแสวงหาประโยชน์ทางการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ลบข้อความนี้ออก

"พจนานุกรมภาษามือไทย" โดยได้รับความร่วมมือจากราชบัณฑิตยสถานร่วมกับวิทยาลัยราชสุดา มหาวิทยาลัยมหิดล แต่ยังอยู่ในช่วงการจัดทำจึงยังไม่มีการนำออกเผยแพร่ต่อสาธารณชน [3]

เนื่องจากภาษามือเป็นภาษาที่ใช้กันเฉพาะกลุ่มผู้พิการทางการได้ยิน ทำให้การขอใช้บริการจากศูนย์บริการต่าง ๆ ซึ่งต้องติดต่อสื่อสารกับบุคคลปกติทั่วไป ยังยากที่จะสื่อสารกันได้สะดวก จำเป็นต้องมีตัวกลางการสื่อสาร เช่น ล่ามภาษามือ เป็นสื่อกลางในการติดต่อ

ด้วยเทคโนโลยีคอมพิวเตอร์ในปัจจุบันสามารถพัฒนาอุปกรณ์เชื่อมต่อ ที่ช่วยในการจับภาพ ซึ่งนำมาประยุกต์ใช้ในการเก็บข้อมูลในรูปแบบของ ภาพเคลื่อนไหวหรือวีดีโอ โดยเป็นการเก็บข้อมูลการเคลื่อนไหวของผู้ใช้ในทุก ๆ ส่วนตำแหน่งของร่างกาย อีกทั้งสามารถใช้ตรวจจับ ท่ามือ ตำแหน่งของมือ และนำข้อมูลเหล่านั้นมาประมวลผลแปลไปเป็นภาษามือ ผู้จัดทำจึงพัฒนาเป็นโปรแกรมระบบรู้จำภาษามือ เพื่อใช้เป็นสื่อกลางการสื่อสาร โดยการเก็บข้อมูลลักษณะท่าทางภาษามือ แปลเป็นภาษาเขียนหรือพยัญชนะซึ่งคนปกติเข้าใจ ทำให้การสื่อสารกันระหว่างผู้พิการทางการได้ยินกับคนปกติได้ติดต่อสื่อสารกันได้สะดวกและเข้าใจทั้งสองฝ่าย

1.2 วัตถุประสงค์ของโครงงาน

1) เพื่อสร้างระบบ ซึ่งช่วยในการสื่อสาร ระหว่างผู้พิการทางการได้ยิน และบุคคลทั่วไป ให้ติดต่อสื่อสารกันได้สะดวกและเข้าใจทั้งสองฝ่าย
2) เพื่อศึกษาภาษา และหลักการการสื่อสารด้วยภาษามือไทย
3) เพื่อศึกษา และพัฒนาระบบรู้จำภาษามือ จากการประยุกต์การใช้งานเทคโนโลยีการตรวจจับภาพเคลื่อนไหว

1.3 ขอบเขตของโครงงาน

1) ระบบนี้จะทำการวิเคราะห์และประมวลผลภาพจากท่าทางภาษามือ ที่จัดเก็บผ่านกล้องไคเนค แล้วแปลไปเป็นคำหรือพยัญชนะ
2) ระบบนี้เป็นการสื่อสารโต้ตอบกัน โดยผู้พิการทางการได้ยินจะแสดงท่าทางภาษามือ ซึ่งระบบจะแปลไปเป็นคำหรือพยัญชนะ
3) ระบบจะสร้างฐานข้อมูลคำศัพท์ ซึ่งสามารถเพิ่ม และแก้ไข โดยเก็บเฉพาะคำที่ใช้การเคลื่อนไหวของมือด้านซ้ายและขวา และการสะกดพยัญชนะ อ้างอิงจากพจนานุกรมอิเล็กทรอนิกส์ภาษามือไทย โดยสำนักงานคณะกรรมการกิจการโทรคมนาคมแห่งชาติ (กสทช.) [4]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 วิธีการดำเนินการ

1) ศึกษาข้อมูล งานวิจัยที่เกี่ยวข้องกับการใช้เทคโนโลยีคอมพิวเตอร์แปลภาษามือ
2) ศึกษาข้อมูล รูปแบบของภาษามือ โครงสร้างของภาษามือ คำศัพท์ภาษามือที่ใช้ในชีวิตประจำวัน
3) ศึกษาข้อมูล การทำงาน การบันทึกข้อมูลการเคลื่อนไหวในรูปแบบ 3 มิติ จากกล้อง Microsoft Kinect XBOX 360™
4) กำหนดขอบเขตของโครงงาน
5) ออกแบบซอฟต์แวร์ที่ใช้ในระบบ
6) พัฒนาซอฟต์แวร์
7) ทดลองการทำงานของระบบ และประเมินประสิทธิภาพ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1) ได้พัฒนาระบบแปลภาษามือ เพื่อใช้เป็นตัวกลางการสื่อสารระหว่างผู้พิการทางการได้ยินและคนปกติ
2) ได้รับความรู้ ความเข้าใจเกี่ยวกับการวิเคราะห์ข้อมูลการเคลื่อนไหว
3) ได้พัฒนาความสามารถในการออกแบบระบบ และการเขียนโปรแกรมคอมพิวเตอร์
4) ได้รับประสบการณ์ในการพัฒนาตนเอง จากการทำงานร่วมกับผู้อื่น

1.6 ส่วนประกอบของปริญญานิพนธ์

ปริญญานิพนธ์ฉบับนี้ได้แบ่งเนื้อหาออกเป็น 5 บทด้วยกันคือ

บทที่ 1 บทนำ กล่าวถึงความสำคัญและที่มาของโครงงาน วัตถุประสงค์ของโครงงาน ขอบเขตของโครงงาน วิธีการดำเนินการ ประโยชน์ที่คาดว่าจะได้รับ และส่วนประกอบของรายงาน

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง กล่าวถึง ภาษามือ เครื่องมือที่ใช้ หลักการทำงานของกล้อง Microsoft Kinect XBOX 360™ การประยุกต์ใช้งาน ทฤษฎีเกี่ยวกับ skeleton tracking ทฤษฎีเกี่ยวกับ Fingertip Tracking ทฤษฎีเกี่ยวกับ Dynamic Time Warping และงานวิจัยที่เกี่ยวข้อง

บทที่ 3 การออกแบบและพัฒนา กล่าวถึงรายละเอียดของการออกแบบระบบ

บทที่ 4 การทดลองและผลการทดลอง กล่าวถึงการทดลองระบบที่ได้พัฒนาขึ้นมา

บทที่ 5 บทสรุป กล่าวถึงบทสรุปของโครงงาน สรุปสิ่งที่ได้รับจากโครงงาน รวมถึงปัญหาอุปสรรคต่าง ๆ ของโครงงาน และแนวทางในการพัฒนาต่อ

ภาคผนวก กล่าวถึงพจนานุกรมท่าทางภาษามือ คู่มือการใช้งาน และตัวอย่างโค้ดโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

2.1 ภาษามือ

ภาษามือแบ่งออกได้เป็น 2 ประเภทใหญ่คือ

1) Static gesture คือการจดจำท่าทางที่ไม่มีการเคลื่อนไหว เช่น การสะกดพยัญชนะของผู้พิการทางการได้ยิน ดังแผนภาพการสะกดพยัญชนะในรูปที่ 2.1 เป็นต้น
2) Dynamic gesture คือการจดจำท่าทางที่มีการเคลื่อนไหว เช่น การติดตามมือของผู้ใช้งาน เป็นต้น

รูปที่ 2.1 ภาพแสดงแบบสะกดอักษรนิ้วมือประดิษฐ์ของคุณหญิงกมลา ไกรฤกษ์ [1]

2.2 เครื่องมือที่ใช้

เครื่องมือที่ใช้คือ กล้อง Microsoft Kinect XBOX 360™ เนื่องจากเป็นอุปกรณ์ที่สามารถตรวจจับการเคลื่อนไหวทั้งร่างกายได้อย่างสมจริง หรือที่เรียกโดยทั่วไปว่า เทคโนโลยี Motion Controller ที่ใช้ระบบกล้องอินฟราเรดตรวจจับการเคลื่อนไหวได้ทั้งร่างกายและยังเพิ่มประสิทธิภาพในการประมวลผลข้อมูล ทำให้ภาพที่ได้มีความสมจริงมากขึ้น

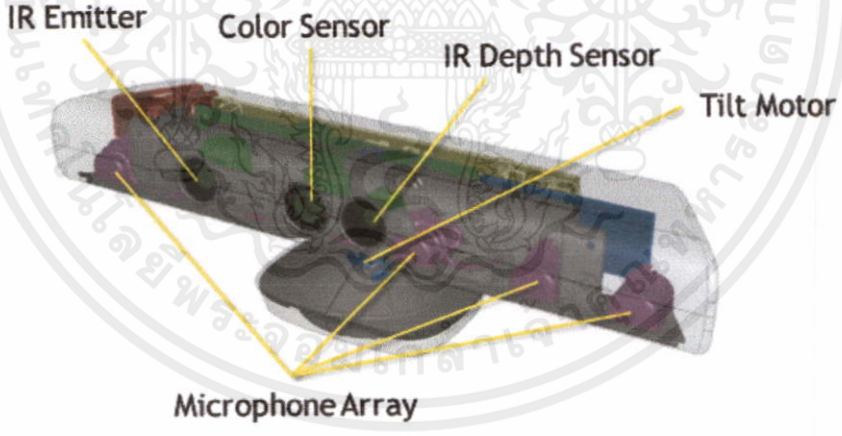
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไคเนค คือ อุปกรณ์ RGB/Depth Sensing และเป็นซอฟต์แวร์เทคโนโลยีที่ประมวลผลออกมาเป็นสัญญาณ RGB/Depth โดยที่อุปกรณ์นี้ประกอบด้วยกล้อง RGB (Normal RGB Camera),

เซนเซอร์ความลึก (Depth Sensor) และไมโครโฟน 4 อัน (Four-Microphone Array) ที่เป็นตัวจัดการสัญญาณความลึก (Depth Signal) รูปภาพ RGB และสัญญาณ Audio ซึ่งซอฟต์แวร์นี้อนุญาตให้ผู้ใช้งานสามารถทำการพัฒนาผลิตภัณฑ์สำหรับแอพลิเคชั่นต่าง ๆ ได้ เครื่องมือเหล่านี้อำนวยความสะดวกในการทำสัญญาณภาพ (Image Signal) การตรวจจับการเคลื่อนไหวของคนในรูปแบบ 3 มิติ การระบุหน้าของคนและการจดจำเสียง เป็นต้น

2.2.1 อุปกรณ์ด้าน Hardware ของไคเนค

รูปที่ 2.2 ภาพกล้อง Microsoft Kinect XBOX 360™ [5]



รูปที่ 2.3 ภาพแสดงส่วนประกอบของกล้อง Microsoft Kinect XBOX 360TM [6]

(ก) (ข)

รูปที่ 2.4 (ก) ภาพสี RGB (ข) ภาพความลึก (Depth Image) [7]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับ[illegible]ท่านั้น[illegible]ะโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อ[illegible]และต้องอ้างอิงถึงเจ้าของเอก[illegible]ทุกครั้งที่มีการนำไปใช้

Microsoft Kinect XBOX 360™ ประกอบด้วยอุปกรณ์ดังที่แสดงในรูปที่ 2.3 คือ IR Emitter, IR Depth Sensor, Color Sensor, Tilt Motor หรือ Motorized Tilt และ Microphone Array โดยที่เซนเซอร์ความลึก (Depth Sensor) ประกอบไปด้วย IR Emitter และ IR Depth Sensor ซึ่ง IR Emitter จะทำการคำนวณจุดของอินฟราเรดบนฉาก 3 มิติ ในขณะที่ IR Depth Sensor ทำการตรวจจับการสะท้อนของจุดอินฟราเรด โดย IR Emitter จะทำการฉายจุดแสงลงบนฉาก 3 มิติ ซึ่งจะเห็นจุดนี้ได้ที่ IR Depth Sensor แต่ไม่สามารถเห็นได้ที่ Color Sensor จากนั้นจะทำการเปรียบเทียบจุดแสงที่ฉายลงไปกับจุดแสงที่เจอในภาพด้วยการวัดขนาดของจุด ความลึกของจุดมาจากความสัมพันธ์ของการแปลจุดจากซ้ายไปขวา ซึ่งการแปลนี้จะขึ้นอยู่กับระยะทางของวัตถุไปยังฉากของกล้อง (Camera-Projector Plane) องค์ประกอบของอุปกรณ์ไคเนค มีดังนี้

1) Color Sensor: มีหน้าที่ส่งข้อมูลสี 3 สี คือ แดง (R) เขียว (G) และน้ำเงิน (B) โดยทำงานที่ 30 Hz และสามารถนำเสนอรูปภาพที่ 640x480 พิกเซลด้วยขนาด 8 บิตต่อช่อง (8-bit Per Channel) อีกทั้งไคเนคยังสามารถผลิตรูปภาพที่มีความละเอียดสูง ทำงาน 10 เฟรมต่อวินาทีที่ความละเอียด 1280x1024 พิกเซลได้
2) 3D Depth Sensor : ประกอบด้วย IR Emitter และ IR Depth Sensor โดยอุปกรณ์ทั้งสองอย่างนี้จะทำการสร้างแผนที่ความลึก ซึ่งจะได้ข้อมูลระยะทางระหว่างวัตถุและกล้องเซนเซอร์มีช่วงระยะทางอยู่ระหว่าง 0.8 เมตรถึง 3.5 เมตร และมีวิดีโอผลลัพธ์มีค่าความละเอียดอยู่ที่ 640x480 พิกเซล ที่อัตรา 30 เฟรมต่อวินาที ทำมุม 57 องศากับแนวนอน และ 43 องศากับแนวตั้ง
3) Motorized Tilt : เป็นตัวปรับระดับเซนเซอร์ ซึ่งสามารถปรับขึ้นลงได้ ±27 องศา

2.2.2 อุปกรณ์ด้าน Software ของไคเนค

ไคเนคทางด้านซอฟต์แวร์หมายถึง Kinect Development Library ซึ่ง Library ที่เราใช้คือ Microsoft Kinect SDK Version 1.7 ของ Microsoft เป็น Open Source Library ที่สามารถใช้ได้กับวินโดว์ (Window) เท่านั้นและในการตรวจจับโครงกระดูก (Skeleton Tracking) ไม่ต้องมีท่าเริ่มที่เฉพาะเจาะจง ทำการตรวจจับท่าทางง่าย ๆ สามารถตรวจจับโครงกระดูกทั้งนั่งและยืน จดจำท่าทางของร่างกาย ตรวจจับหน้าและวิเคราะห์ท่าทางของมือ เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 ขั้นตอนการเตรียมภาพ (Preprocessing)

ข้อมูลที่ได้จากไคเนคเป็นข้อมูลที่ยังไม่เหมาะสมในคอมพิวเตอร์แต่ละรุ่น หรือก็คือข้อมูลความลึกดิบ (Raw Depth Data) ที่มีเสียงรบกวน (Noise) และมีหลายพิกเซลที่ยังไม่มีข้อมูลความลึก เนื่องจากการสะท้อนหลาย ๆ ครั้ง ความโปร่งแสงของวัตถุหรือการกระจัดกระจายของแสงบนพื้นผิว เช่น ผมและเนื้อเยื่อของคน ทำให้ข้อมูลความลึกที่ได้ขาดหายหรือมีการคำนวณที่ไม่ถูกต้องจึงต้องทำการกู้คืนข้อมูลก่อนที่จะนำข้อมูลมาใช้งานต่อไป ดังนั้นไคเนคจึงเริ่มทำงานที่กระบวนการขั้นตอนการเตรียมภาพก่อนเป็นอันดับแรก

2.2.3.1 มาตรการวัดของไคเนค (Kinect Recalibration)

พารามิเตอร์ของกล้องไคเนคจะถูกเก็บลงเมมโมรี่ เพื่อที่จะสามารถออกผลลัพธ์มาเป็นข้อมูลภาพสี (RGB) และข้อมูลความลึก (Depth Data) ซึ่งกล้องไคเนคได้มีการบิดเบือน เพื่อที่จะทำการแก้ปัญหานี้จึงมีการคำนวณค่าความลึกระหว่างค่าความลึกที่คำนวณได้กับค่าความลึกที่ต้องการ นอกจากนี้อุณหภูมิและสภาพแวดล้อมก็มีผลต่อมาตรวัดของไคเนคด้วยเหมือนกัน

2.2.3.2 การกรองข้อมูลความลึก (Depth Data Filtering)

เป็นขั้นตอนที่ทำการลดเสียงรบกวน (Noise) และทำการเติมเต็มส่วนของข้อมูลความลึก (Depth Data) ที่หายไป ข้อมูลความลึกหรือก็คือข้อมูลที่เป็นภาพที่มีสีเดียวแล้วนำภาพนั้นมาผ่านกระบวนการกรอง เช่น Gaussian Filter และ Adaptive Filter เป็นต้น ในการกรองหลุม (Hole) ในแผนภาพความลึก (Depth Map) อันดับแรกที่ต้องทำคือ การหาช่องว่างของความเหมือนกันระหว่าง 2 ช่องทาง โดยวิธีการจับคู่เส้นขอบของวัตถุ หลังจากนั้นวัตถุจะเป็นภาพสีผ่านการแบ่งส่วน (Segmentation) และตำแหน่งที่ได้จากการแบ่งส่วนจะถูกเคลื่อนย้ายไปเป็นภาพความลึก (Depth Image) และวิธีการแบ่งส่วนนี้ยังสามารถทำการเพิ่มข้อมูลความลึกที่หายไปด้วยการสมมติสัญญาณความลึกที่เหมาะสมในแต่ละการแบ่งส่วน แต่วิธีการนี้อาจจะไม่ทำการสร้างสัญญาณความลึกที่สม่ำเสมอ เมื่อพื้นผิววัตถุมีสีสันมาก หรือค่าความลึกที่หายไปสามารถเติมเต็มด้วยวิธีการกรอง Joint-Bilateral Filter ที่พิกเซลความลึก ซึ่งวิธีการนี้คือการกรองโดยเก็บขอบและลดเสียงรบกวน (Noise) คือค่าความลึกในแต่ละพิกเซลจะถูกแทนที่ด้วยค่าเฉลี่ยน้ำหนักค่าความลึกพิกเซลข้างเคียง ในการกรองน้ำหนักนี้จะพิจารณาจากข้อมูล RGB ข้อมูลความลึกและการ Map ที่ตรงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.4 หลักการทำงานของไคเนค

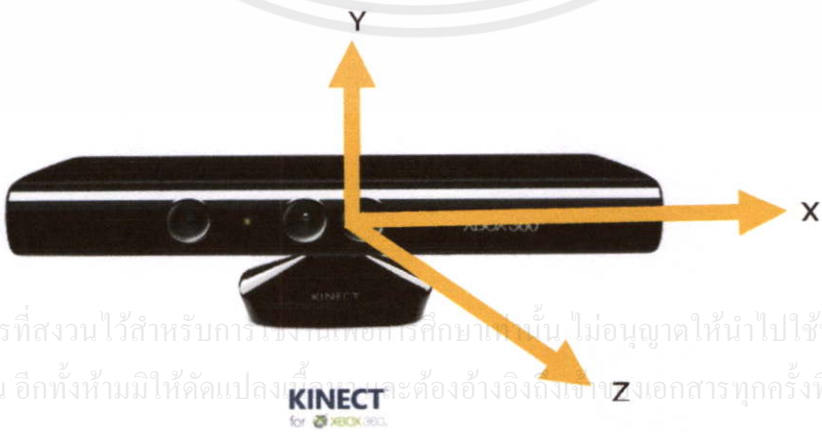
2.2.4.1 การตรวจจับวัตถุและการติดตาม

ในการแยกความแตกต่างระหว่างวัตถุกับพื้นหลังนั้น ไคเนคได้ใช้ทฤษฎี Background Subtraction โดยที่สมมติให้พื้นหลัง (Background) เป็นค่าคงที่ตลอดเวลาและทำการแยกเบื้องหน้า (Foreground) ออกมาด้วยการเอารูปภาพที่ได้รับ (Input Image) ลบด้วยพื้นหลังที่ทำการสมมติไว้ในตอนแรก [7]

2.2.4.2 การทำงานของไคเนค

การทำงานเริ่มจากการฉายแสงอินฟราเรดออกจากตัวไคเนค ซึ่งไม่สามารถมองเห็นได้ด้วยตาเปล่า แสงที่ถูกฉายออกมาจะมีลักษณะเป็นจุด ๆ ตามแนวตั้ง 480 จุด แนวนอน 640 จุด แต่ละจุดห่างกัน 3 มิลลิเมตร (ที่ระยะสองเมตรจากแหล่งกำเนิดแสง) หลังจากนั้นกล้องวัดความลึกจะรับภาพระดับความสว่างของอินฟราเรดที่ตกกระทบลงบนวัตถุส่งไปให้เซนเซอร์ เพื่อทำการวัดความลึกตามแนวแกน Z (Axis-Z) ดังรูปที่ 2.5 ที่แสดงพิกัดต่าง ๆ เมื่อเทียบกับกล้องไคเนค โดยที่จุดกำเนิดแกนอยู่กึ่งกลางกล้อง แกน Z คือแกนที่อยู่ในแนวพุ่งไปหาผู้เล่น ซึ่งสามารถนำความลึกที่วัดได้ในแนวแกน Z มาจำลองสภาพแวดล้อมเป็นสามมิติได้ หากความสว่างมีมากแสดงว่า วัตถุนั้นอยู่ใกล้ แต่ในทางตรงกันข้ามหากมีความสว่างน้อยลงแสดงว่า วัตถุนั้นอยู่ไกลออกไป นอกจากนี้ไคเนคยังทำการบันทึกใบหน้าผู้เล่น และสามารถใช้เสียงในการควบคุมการใช้งานได้อีกด้วย

เมื่อได้ระดับความลึกของภาพแล้ว ทำให้เซนเซอร์ของไคเนคสามารถแยกผู้เล่นออกจากสภาพแวดล้อมภายในห้องได้ เช่น ผนัง ที่นั่งเล่นหรือแม้แต่การจำแนกมือของผู้เล่นอยู่ข้างหน้าหรือข้างหลัง



รูปที่ 2.5 ภาพแสดงระบบพิกัดของกล้อง Microsoft Kinect XBOX 360™

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การประยุกต์ใช้งาน

การออกแบบในระบบรู้จำภาษามือ เป็นการนำไคเนคมาตรวจจับการเคลื่อนไหวของมือ โดยที่นำเทคนิค Skeleton Tracking มาทำการตรวจจับการเคลื่อนไหวของข้อต่อ ในที่นี้ได้ทำการตรวจจับกลางฝ่ามือและข้อศอกทั้งมือขวาและมือซ้าย จะได้ข้อมูลตำแหน่งในแนวแกน X Y และ Z ของแต่ละจุดข้อต่อ ซึ่งจะนำข้อมูลเหล่านี้มาปรับเป็นค่าคงที่และนำไปสร้างเป็นฐานข้อมูล เพื่อใช้ในการเปรียบเทียบกับข้อมูลตำแหน่งข้อต่อต่าง ๆ ที่เข้ามาในขณะที่ผู้พิการทางการได้ยินกำลังแสดงท่าทางภาษามือ [8] โดยที่วิธีนี้อยู่ในขั้นตอนการแปลของระบบ ซึ่งได้ผลลัพธ์คือคำศัพท์ ส่วนการแปลพยัญชนะ จะทำการเปรียบเทียบตำแหน่งของปลายนิ้วมือและระยะห่างระหว่างนิ้วมือ เพราะในแต่ละท่าทางของการสะกดพยัญชนะจะมีระดับปลายนิ้วมือและจำนวนของนิ้วมือไม่เท่ากัน

2.4 ทฤษฎีเกี่ยวกับ Skeleton Tracking

เทคนิค Skeleton Tracking เป็นหนึ่งในวิธีที่ใช้ประโยชน์ของกล้องไคเนคโดยการติดตามข้อต่อหลัก ๆ ของร่างกาย ว่ามีการเคลื่อนที่ไปอยู่ตรงส่วนไหนของเฟรม นอกจากนั้นยังเพิ่มความอิสระของนิ้ว เพื่อใช้ในการสื่อสารและแสดงท่าทางใหม่ ๆ ได้ ซึ่งวิธีการนี้จะนำมาใช้ในการติดตามส่วนของมือทั้ง 2 ข้าง เพื่อดูลักษณะของการเคลื่อนที่ของมือว่าเป็นอย่างไร กำลังบิดมืออยู่หรือไม่ แล้วนำไปเปรียบเทียบกับฐานข้อมูลว่าลักษณะของมือตรงกับข้อมูลใด ถ้าตรงกันก็จะสามารถนำข้อมูลนั้นมาใช้ในการแปลความหมายของลักษณะท่าทางนั้น ๆ แต่ถ้าไม่ตรงกับข้อมูลใดในฐานข้อมูลเลย สามารถเลือกที่จะทำการบันทึกลักษณะท่าทางนั้นไว้ได้ (เหมือนเป็นการอัพเดตฐานข้อมูล)

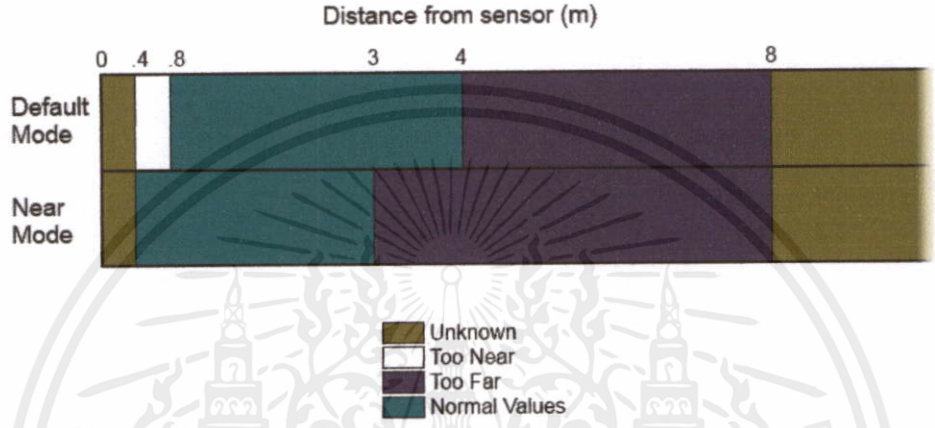
ในการใช้งาน Skeletal Tracking ก่อนอื่นต้องเลือกภาษาที่จะใช้ในการเขียน ซึ่งเลือกใช้เป็นภาษา C# อีกทั้ง Skeletal Tracking อนุญาตให้ไคเนคจดจำผู้ใช้งานและติดตามการกระทำต่าง ๆ โดยใช้กล้องอินฟราเรด ซึ่งไคเนคมีคุณสมบัติ คือ สามารถที่จะจดจำได้มากที่สุด 6 คน ที่อยู่ในขอบเขตของเซนเซอร์ ซึ่งจะแสดงการติดตามได้ 2 คนพร้อม ๆ กัน แอพพลิเคชั่นสามารถที่จะระบุตำแหน่งของข้อต่อ และติดตามการเคลื่อนไหวต่าง ๆ ได้

Skeletal Tracking ถูกทำให้เหมาะสมแก่การจดจำผู้ใช้งานไม่ว่าจะท่ายืน ท่านั่ง กระทั่งใบหน้า ในการจดจำนั้น ผู้ใช้งานต้องยืนอยู่ข้างหน้าเซนเซอร์และให้แน่ใจว่าเซนเซอร์สามารถเห็นส่วนของร่างกายที่เราต้องการให้จดจำได้ เช่น การจดจำใบหน้า เป็นต้น [9]

มุมมองของกล้องไคเนค ถูกกำหนดโดยการตั้งค่ากล้องอินฟราเรด (IR Camera) ซึ่งถูกตั้งค่าด้วย DepthRange Enumeration คือช่วงข้อมูลความลึก ซึ่งถูกกำหนดโดยระยะห่างระหว่างเซนเซอร์กับคน ซึ่งเซนเซอร์วัดระยะ (Depth Sensor) จะแบ่งช่วงความลึกเป็น Default กับ Near ดังรูปที่ 2.6 ที่แสดงการแบ่งระยะทางของ Default Mode กับ Near Mode มีช่วงความลึก ดังนี้

เอกสารนี้ ... ไม่ว่ากรณีใด ...

1) Default : มีช่วงความลึกอยู่ที่ 800 ถึง 4,000 มิลลิเมตร สามารถใช้ได้ใน Kinect for Windows Sensor และ Kinect for Xbox 360 Sensor
2) Near : มีช่วงความลึกอยู่ระหว่าง 400 ถึง 3,000 มิลลิเมตร สามารถใช้ได้ใน Kinect for Windows Sensor เท่านั้น



รูปที่ 2.6 ภาพแสดงระยะห่างระหว่างเซนเซอร์กับคนหรือวัตถุ หน่วยเมตร (m) [10]

ความแตกต่างระหว่าง Kinect for Windows กับ Kinect for Xbox 360 คือ Kinect สำหรับ Windows นั้นสามารถที่จะนำไปพัฒนาเพื่อธุรกิจได้ และที่ระบบ Near Mode นั้นสามารถใช้ได้ใน Windows นั่นเป็นเพราะว่ารุ่น Windows สามารถใช้งานได้ในระยะใกล้สุดคือ 40 เซนติเมตร แต่ Kinect for Xbox 360 ใช้ได้ใกล้สุดที่ 80 เซนติเมตร แต่ถ้าดูในเรื่องของราคาแล้ว Kinect for Windows มีราคาแพงกว่าและในเรื่องของฮาร์ดแวร์มีความแตกต่างกันเล็กน้อย

2.5 ทฤษฎีเกี่ยวกับ Fingertip Tracking

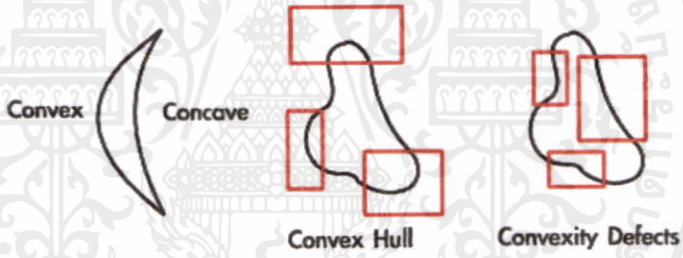
เทคนิค Fingertip Tracking เป็นเทคนิคในการตรวจจับหาตำแหน่งของปลายนิ้วมือ โดยที่ในระบบรู้จำภาษามือจะใช้ Convexity Defects Technique ซึ่งเป็นเทคนิคที่มีขั้นตอนการทำงานดังนี้

1) ทำการตรวจจับมือก่อน เพื่อทำการแยกรูปมือออกจากพื้นหลัง ดังรูปที่ 2.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.7 ภาพแสดงข้อมูลภาพมือ

2) ใช้ Convex Hull ในการหาตำแหน่งปลายนิ้วมือ โดย Convex Hull คือเซตของ Convex พื้นฐานที่ล้อมรอบภาพมือ ผลลัพธ์ดังรูปที่ 2.9



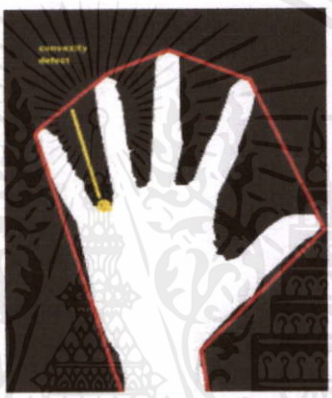
รูปที่ 2.8 ภาพแสดงตำแหน่งที่เป็น Convexity Defects [11]

รูปที่ 2.9 ภาพแสดงผลลัพธ์การหา Convexity Defects

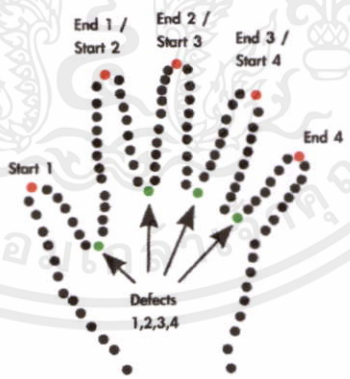
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากภาพมือรูปที่ 2.9 เส้นกรอบนอก (สีแดง) คือ Convex Hull

3) ทำการหา Convexity Defects ดังรูปที่ 2.10 ที่จุดสีเหลืองคือ Defect Point เพื่อให้ Convexity Defects เป็นตัวบอกใบ้ว่าตำแหน่งของปลายนิ้วอยู่ตรงไหน โดยในการหา Convexity Defects จะมีพารามิเตอร์ 4 ชนิด คือ start (จุดที่เริ่มหา Convexity Defects) end (จุดสิ้นสุดการหา Convexity Defects) depth_point (ระยะที่ไกลที่สุดจากจุด Convex Hull มายังส่วนเว้าของมือ) และ depth (ระยะทางระหว่างจุดที่ไกลที่สุดและ Convex Hull) [12] ดังรูปที่ 2.11



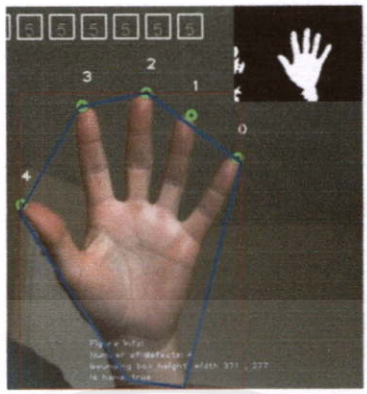
รูปที่ 2.10 ภาพแสดงจุด Defect [13]



รูปที่ 2.11 ภาพแสดงตำแหน่งปลายนิ้วตำแหน่งระหว่างนิ้วมือ

4) เมื่อผ่านกระบวนการการหา Convexity Defects แล้วจะได้ผลลัพธ์ดังรูปที่ 2.12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.12 ภาพผลลัพธ์การทำ Convexity Defects [14]

2.6 ทฤษฎีเกี่ยวกับ Dynamic Time Warping (DTW)

การใช้เทคนิค DTW เพื่อนำมาแก้ปัญหาระยะเวลาในการทำภาษามือ เพราะการแสดงท่าทางอาจจะช้าหรือเร็วกว่าท่าทางภาษามือที่จัดเก็บอยู่ในฐานข้อมูล

2.6.1 หลักการทำงานของ Dynamic Time Warping (DTW)

2.6.1.1 การหา Gesture ที่สนใจ

เริ่มจากการหาค่า Cost ของเฟรมปัจจุบันโดยใช้ Euclidean Distance และทำการเปรียบเทียบเฟรมเฟรมปัจจุบันกับเฟรมสุดท้าย ซึ่งค่า Cost จะเป็นตัวบอกถึงความเหมือนกันของเฟรมสองเฟรม ถ้าค่า Cost = 0 แสดงว่าเฟรมสองเฟรมที่ทำการเปรียบเทียบกันนั้นเหมือนกัน แต่ถ้าค่า Cost > 0 แสดงว่าเฟรมสองเฟรมต่างกัน จะต่างกันมากน้อยแค่ไหนก็ขึ้นอยู่กับค่า Cost และจะให้ค่า Cost ที่ต่ำที่สุดเป็น Gesture ที่สนใจ (Candidate Gesture)

2.6.1.2 การคำนวณค่า Cost ด้วย Matrix

การคำนวณหาค่า Cost ระหว่าง Reference Gesture Frame กับ Input Gesture Frame โดยใช้ Matrix ซึ่งจะใช้ Euclidean Distance เป็นสูตรในการคำนวณ ดังสมการที่ (2.1) และ (2.2)

$$Euclidean\ distance = d(p,q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2 + (p_y - q_z)^2} \quad (2.1)$$

$$Total\ euclidean\ distance = \sum_{i=1}^{5} d(p_i\ ,\ q_i) \quad (2.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย p คือ ตำแหน่งของนิ้วใน Reference Gesture Frame

q คือ ตำแหน่งของนิ้วใน Input Gesture Frame

d คือ ระยะทาง หรือก็คือ ค่า Cost

2.6.1.3 การคำนวณค่า Cost สะสมต่ำสุดด้วย Matrix (Accumulated Distance Matrix)

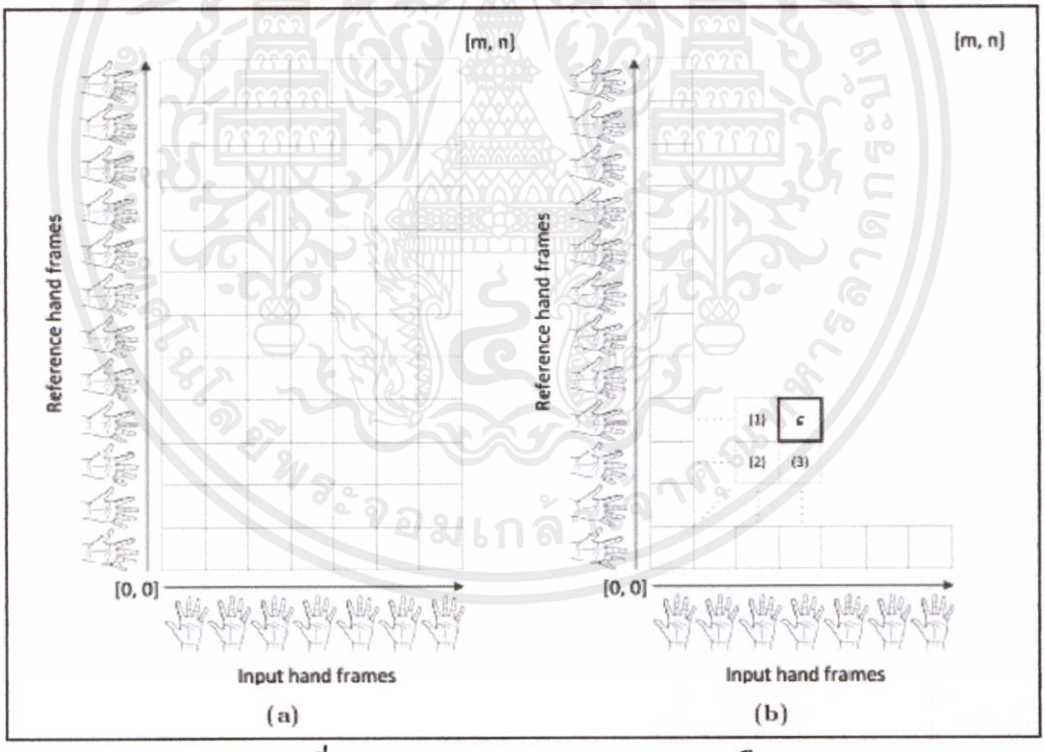
เริ่มจากการคำนวณหาค่า Cost ต่ำสุดที่ไปถึงเซลล์ (c) จากทางซ้าย ด้านล่างหรือด้านทแยงลงมาก็ได้ และทำการคำนวณหาค่า Cost ของเซลล์ (C_{cost}) แล้วนำค่า Cost ที่คำนวณได้ข้างต้นมาบวกเข้ากับค่า Cost ของเซลล์ที่เพิ่งคำนวณก็จะได้ค่า Cost สะสม จากนั้นทำการเริ่มคำนวณใหม่ตั้งแต่ต้นไปเรื่อย ๆ โดยให้เซลล์ [0,0] เป็นเซลล์แรก ดังรูปที่ 2.10 เป็นภาพแสดง Matrix ของเฟรมมือ โดยที่ c คือ cell และ

(1) คือ cell ที่อยู่ด้านซ้ายของเซลล์ c และสามารถไปถึงเซลล์ c ได้

(2) คือ cell ที่อยู่ด้านทแยงลงมาของเซลล์ c และสามารถไปถึงเซลล์ c ได้

(3) คือ cell ที่อยู่ด้านล่างของเซลล์ c และสามารถไปถึงเซลล์ c ได้

เพราะฉะนั้นในขั้นแรกก็ทำการคำนวณหาค่า Cost ต่ำสุดจากเซลล์ (1), (2) และ (3) มาถึงเซลล์ c และหาค่า Cost ที่เซลล์ c แล้วเอามาบวกกัน



รูปที่ 2.13 ภาพแสดง Matrix ของเฟรมมือ

2.6.1.4 การหาระยะทางที่สั้นที่สุด

ทำการหาระยะทางสะสมที่สั้นที่สุด (The Lowest Accumulated Cost) โดยเริ่มจากเซลล์สุดท้ายไปยังเซลล์แรกและทำการเลือกเซลล์ที่มีค่า Cost ต่ำที่สุดใน 3 เซลล์ (เซลล์ (1) (2) และ (3))

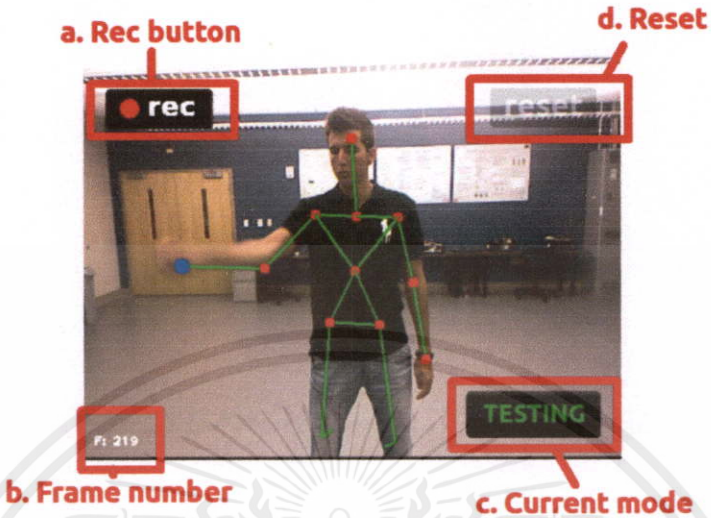
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แล้วทำการเลือกเซลล์ต่อไปจากด้านซ้าย ด้านล่างหรือด้านทแยงลงมาก็ได้ ระหว่างที่ทำการเลือกเซลล์นั้นก็ทำการบวกค่า Cost ไปเรื่อย ๆ ตามระยะทางที่เกิดขึ้น [15]

2.7 งานวิจัยที่เกี่ยวข้อง

งานวิจัยของ Daniel Martínez Capilla [16] เป็นงานวิจัยที่พัฒนาการแปลภาษามือโดยอัตโนมัติซึ่งใช้ข้อมูลที่ได้จากกล้อง Microsoft Kinect XBOX 360™ ข้อมูลภาษามือที่รับเข้ามา (Input Data) มาจากการบันทึกภาพภาษามือที่กระทำโดยคนที่พิการทางการได้ยิน ภาพข้อมูลที่ได้มา (Raw Image) นำไปทำการประมวลผล ตัวแปลภาษา (Translator) จะแปลออกมาเป็น คำหรือข้อความที่เป็นภาษาพูด การที่จะทำงานระบบนี้ใช้ Ubuntu เป็นระบบปฏิบัติการ ใช้ภาษา C++ และ Library OpenCV, OpenNI/NITE Middleware และ Qt (ใช้ Qt GUI ในการออกแบบตัวแปลภาษามือในหน้า User Interface) การแปลภาษามือนี้สิ่งที่จำเป็นก็คือ การตรวจจับข้อต่อ (Joint Tracking) โดยใช้ PriimeSense ในการเข้าถึงข้อมูลภาพสีและภาพความลึกที่ได้มาจากเซนเซอร์ สำหรับ Gesture Recognition (เทคโนโลยีตรวจจับภาษากาย) Feature Detection (คุณลักษณะในการตรวจจับ) และ Joint Tracking (การตรวจจับข้อต่อ) ใช้ซอฟต์แวร์ OpenNI (Open Natural Interaction) เขียนโปรแกรมตั้งแต่การโหลดโปรแกรมขึ้นมา การรับข้อมูลเข้า การแปลงภาพ การสร้างข้อต่อที่ต้องการตรวจจับ การสร้างเส้นเชื่อมต่อระหว่างข้อต่อ (กระดูก : Bone) การสร้างพื้นหลังของภาพกระดูก (Skeleton Image) การใส่สีของข้อต่อ (Joint) และกระดูก (Bone) ไปจนถึงจบการทำงานต่าง ๆ กระบวนกาทำงานของโปรแกรมก็คือ คนที่พิการทางการได้ยินยืนหน้ากล้องทำภาษามือ กล้องก็จะได้รับภาพมาจากเซนเซอร์เป็นเฟรม (Frame) ทำการแสดงเฟรมบนจอภาพซึ่งเป็นภาพสี (RGB Image) โดยภาพที่นำไปแสดงจะมีการตรวจสอบว่าจะทำการบันทึกหรือไม่ ถ้าต้องการบันทึกก็จะไปเข้าสู่กระบวนการตรวจจับข้อต่อ (Joint Tracking) คือ ได้รับข้อมูลข้อต่อที่สนใจ (Joint of Interest : JoI) มาจากกล้องและนำไปทำการลดข้อมูล เพื่อที่จะสามารถนำมาวิเคราะห์ได้ง่าย (Normalized Data) จากนั้นก็สร้าง Frame Descriptor (เฟรมที่เก็บข้อมูลตำแหน่งของมือ) แต่ถ้าไม่ต้องการบันทึกก็ต้องไปเลือกว่าต้องการทำงานโหมดไหน มี 2 โหมดคือ TESTING และ TRAINING ดังรูปที่ 2.14 ถ้าโหมด TRAINING คือโหมดที่ใช้ในการบันทึกภาษามือ เป็นการเพิ่ม Frame Descriptor ไปยังฐานข้อมูล Gesture Dictionary ส่วนโหมด TESTING คือโหมดในการแปลภาษามือ โดยการเพิ่ม Frame Descriptor ไปยังไฟล์ test.data หลังจากนั้นระบบทำการตรวจสอบว่าเฟรมปัจจุบันที่รับเข้ามาเป็นเฟรมสุดท้ายหรือยัง ถ้ายังให้ทำการรับเฟรมต่อ แต่ถ้าเป็นเฟรมสุดท้ายแล้วให้หยุดรับเฟรม ทำการแปลภาษามือโดยการนำภาพในเฟรมที่ผ่านการ Normalized Data แล้วมาเปรียบเทียบกับภาพในฐานข้อมูล Gesture Dictionary ถ้าตรงกันก็แสดงคำซึ่งเป็นผลลัพธ์ออกทางจอแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้ง



รูปที่ 2.14 ภาพแสดง User Interface ของงานวิจัยของ Daniel Martínez Capilla

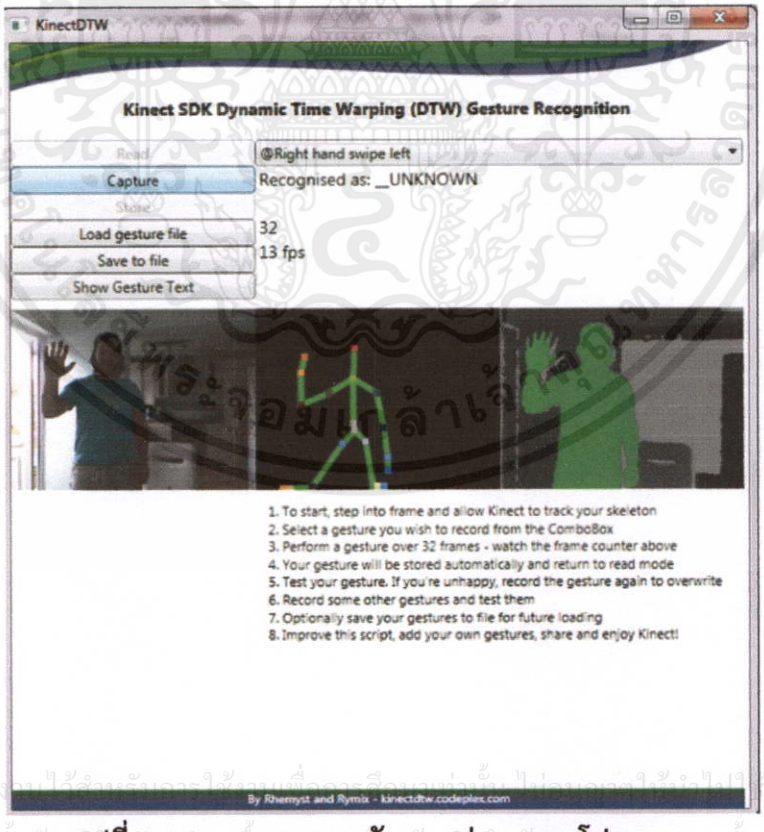
งานวิจัยของ Stefan Stegmueller [17] “Candescent NUI” [18] เป็นโปรแกรมที่ทำการตรวจจับและติดตามมือและนิ้วมือของกล้องไคเนคจากข้อมูลภาพความลึก (Depth Data) โดยพัฒนาจาก OpenNI, Microsoft Kinect SDK และ C# ซึ่งโปรแกรมสามารถทำการตรวจจับมือ และนิ้วมือดังรูปที่ 2.15 นอกจากนี้ยังสามารถทำการเลื่อนตำแหน่ง ย่อ ขยายภาพได้ตามที่ต้องการและสามารถแปลจำนวนนิ้วมือออกมาเป็นตัวเลขได้อีกด้วย

รูปที่ 2.15 ภาพแสดงการตรวจจับมือและนิ้วมือด้วยกล้องไคเนค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

งานวิจัยของ Rymix [19] “Kinect SDK Dynamic Time Warping (DTW) Gesture Recognition” เป็นโปรแกรมที่ใช้กับ Window 7 ส่วนซอฟต์แวร์ใช้ Microsoft Kinect SDK โปรแกรมที่ใช้คือ Microsoft Visual Studio 2010

โปรแกรมนี้สามารถทำการตรวจจับร่างกาย โดยใช้วิธี Skeleton Tracking แล้วสามารถทำการตรวจจับท่าทางต่าง ๆ จากนั้นสามารถเลือกได้ว่าทำการตรวจจับท่าทางชื่ออะไรและเก็บข้อมูลที่ได้จากการตรวจจับ ซึ่งก็คือตำแหน่งของข้อต่อ (ตำแหน่ง X, Y และ Z) หรือจะทำการตรวจจับท่าทางนั้น ๆ ใหม่อีกครั้ง เมื่อทำการเก็บข้อมูลตำแหน่งของข้อต่อลงไฟล์ ไฟล์ที่ได้จะเป็นไฟล์อักษร (Text File) โดยจะกำหนดให้เก็บข้อมูลข้อต่อ 32 เฟรมต่อ 1 วินาที หรือสามารถเก็บข้อมูล 32 เฟรมแบบหนึ่งเฟรมเว้นหนึ่งเฟรมก็ได้ นอกจากนี้ยังสามารถนำข้อมูลที่เก็บลงไฟล์โหลดขึ้นมาเปรียบเทียบกับข้อมูลข้อต่อที่กำลังตรวจจับอยู่ได้ โดยอาจจะกำหนดให้ทำการจับคู่เฟรมทั้ง 32 เฟรมหรือจะ 10 เฟรมก็ได้ตามที่กำหนดไว้ เพื่อทำการจับคู่เฟรมที่ตรงกัน ถ้ากำหนดให้จับคู่เฟรม 32 เฟรมแล้วให้ขนาดของการเก็บข้อมูลข้อต่อเท่ากับ 32 เฟรมต่อวินาที โปรแกรมจะทำการตรวจดูทั้ง 32 เฟรมว่า 32 เฟรมที่โหลดขึ้นมากับ 32 เฟรมที่กำลังบันทึกอยู่ตรงกันหรือไม่ ถ้าตรงกันก็จะออกผลลัพธ์คำนั้นออกมา ส่วนการทำงานของปุ่ม Show Gesture Text เป็นฟังก์ชันการทำงานที่เมื่อมีการเลือกชื่อและตรวจจับท่าทางนั้น ๆ ไม่ว่าจะบันทึกลงไฟล์แล้วหรือไม่ก็ตาม เมื่อกดปุ่มนี้ จะแสดงชื่อท่าทางและตำแหน่งของข้อต่อต่าง ๆ ทั้งหมดที่จดจำอยู่ ณ ขณะนั้นทั้งหมด



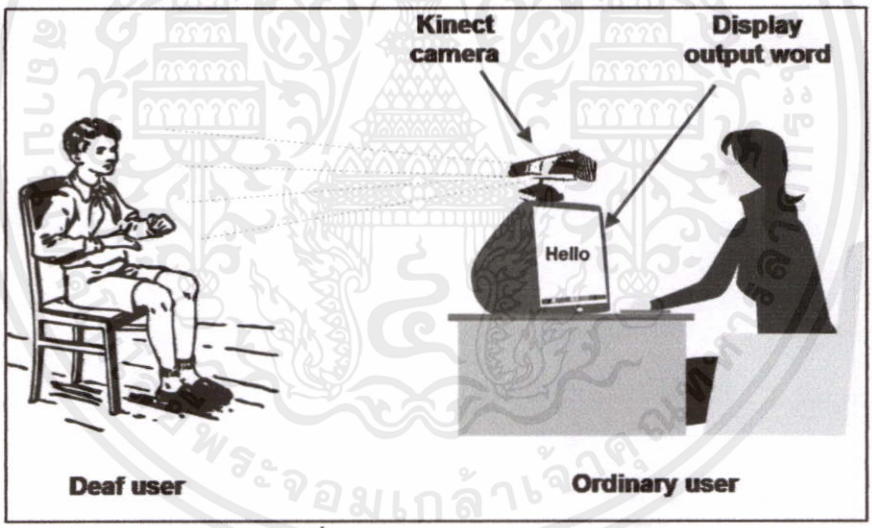
รูปที่ 2.16 ภาพแสดงคุณลักษณะต่าง ๆของโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การออกแบบระบบ

3.1 ภาพรวมของระบบ

จากการศึกษาข้อมูล ระบบภาษามือ และหลักการทำงานของกล้องไคเนค ผู้จัดทำโครงงานได้ทำการออกแบบการทำงานของระบบโดยเริ่มจาก ผู้พิการทางการได้ยินแสดงท่าทางภาษามือ ที่หน้ากล้อง เพื่อให้กล้องรับข้อมูลภาพเคลื่อนไหว ระบบจะทำการตรวจจับไปที่ตำแหน่งของมือทั้งสองข้าง จากนั้นจะทำการประมวลผลภาพท่ามือทั้งสองข้าง และตำแหน่งของข้อมือ เทียบกับฐานข้อมูล ซึ่งเป็นตัวอย่างคำในภาษามืออย่างง่าย หลังจากการประมวลผลจะแสดงผลลัพธ์ในรูปของคำหรือพยัญชนะ ซึ่งรูปที่ 3.1 จะแสดงภาพรวมของระบบที่ผู้จัดทำโครงงานได้ออกแบบไว้



รูปที่ 3.1 ภาพรวมของระบบ

ระบบรู้จำภาษามือนี้ เป็นระบบในการแปลภาษามือแบบทางเดียว กล่าวคือ จากท่าทางซึ่งเป็นภาษามือ ผ่านการประมวลผลแปลเป็นพยัญชนะ หรือคำ ไม่สามารถแปลงกลับจากคำหรือพยัญชนะเป็นท่าทางภาษามือได้ เนื่องจากผู้ที่พิการทางการได้ยินนั้น อาจสามารถสื่อสารจากการอ่านข้อความจากบุคคลปกติได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 การแปลภาษามือ

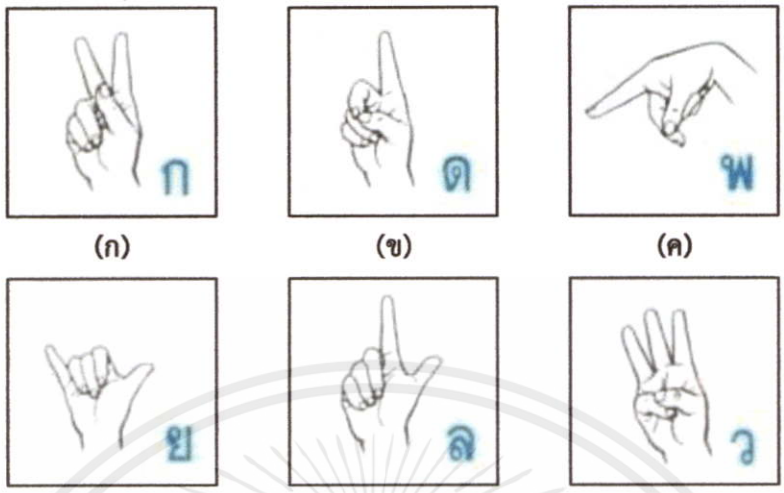
3.2.1 การสะกดพยัญชนะ

ในการสะกดพยัญชนะจะใช้มือเพียงข้างเดียว ทำท่าด้วยนิ้วมือเป็นรูปต่าง ๆ แทนตัวพยัญชนะ แสดงให้เห็นว่าข้อมูลที่จำเป็นต้องใช้คือ ท่าทางของนิ้วมือ ยกตัวอย่างการสะกดพยัญชนะแบบที่เป็นท่ามือเดียวและแสดงตำแหน่งนิ้วมือชัดเจน ดังต่อไปนี้

ตัวอย่างการสะกดพยัญชนะ

1) “ก” ท่ามือที่แสดงคือ ชูนิ้วชี้และนิ้วกลางขึ้น นิ้วหัวแม่มือแตะที่ข้อล่างนิ้วกลาง ส่วนนิ้วนางและนิ้วก้อยหุบลง ดังรูปที่ 3.2(ก)
2) “ด” ท่ามือที่แสดงคือ ชูนิ้วชี้ขึ้น นิ้วหัวแม่มือแตะที่ข้อสองของนิ้วกลาง ดังรูปที่ 3.2(ข)
3) “พ” ท่ามือที่แสดงคือ ใช้นิ้วชี้ชี้ลง นิ้วหัวแม่มือแตะที่ข้อที่หนึ่งของนิ้วกลาง ดังรูปที่ 3.2(ค)
4) “ย” ท่ามือที่แสดงคือ ชูนิ้วหัวแม่มือและนิ้วก้อย ส่วนสามนิ้วที่เหลือหุบลง ดังรูปที่ 3.2(ง)
5) “ล” ท่ามือที่แสดงคือ ชูนิ้วหัวแม่มือและนิ้วชี้ขึ้น ส่วนสามนิ้วที่เหลือหุบลง ดังรูปที่ 3.2(จ)
6) “ว” ท่ามือที่แสดงคือ ชูนิ้วชี้ นิ้วกลาง และนิ้วนางขึ้น อีกสองนิ้วที่เหลือหุบลงแตะกัน ดังรูปที่ 3.2(ฉ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก) (ข) (ค)

(ง) (จ) (ฉ)

รูปที่ 3.2 ตัวอย่างการสะกดพยัญชนะภาษามือไทย

(ก) ท่ามือพยัญชนะ “ก”

(ข) ท่ามือพยัญชนะ “ด”

(ค) ท่ามือพยัญชนะ “พ”

(ก) ท่ามือพยัญชนะ “ย”

(ก) ท่ามือพยัญชนะ “ล”

(ก) ท่ามือพยัญชนะ “ว”

นอกจากการสะกดพยัญชนะแบบที่เป็นท่ามือเดียว ยังมีบางพยัญชนะที่ใช้ท่ามือสองท่าในการระบุพยัญชนะ เช่น “ข ค ฆ” ซึ่งเป็นท่าต่อเนื่องจากท่ามือพยัญชนะ “ก” โดยทำการนับเพิ่มจากท่าต้น (ท่ามือพยัญชนะ “ก”) ต่อไป ซึ่งการนับหนึ่งจะได้ท่าพยัญชนะ “ข” การนับสองจะได้ท่าพยัญชนะ “ค” และการนับสามจะได้ท่าพยัญชนะ “ฆ” ดังรูปที่ 3.3

รูปที่ 3.3 ท่าพยัญชนะที่ต่อเนื่องของพยัญชนะ “ก”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น

จากลักษณะท่ามือที่ได้ยกตัวอย่างไปนั้นล้วนมีความแตกต่างกันทั้งในเรื่องตำแหน่งของนิ้วมือและจำนวนนิ้วมือ จึงนำไปกำหนดเงื่อนไขสำหรับการแปลแต่ละพยัญชนะ ดังนี้

1) นับจำนวนนิ้วที่ตรวจจับได้ และกำหนดให้ระหว่างนิ้วห่างกัน k หน่วย

(1) 1 นิ้ว

- พยัญชนะ “ด” ; ตำแหน่งปลายนิ้วสูงกว่าตำแหน่งมือ ที่ค่าความแตกต่างมากกว่าความสูงนิ้วชี้

(2) 2 นิ้ว

- พยัญชนะ “ก” ; ตำแหน่งปลายนิ้วพิกัดแกน Y อยู่ในระดับเดียวกัน และค่าความต่างในแนวแกน X น้อยกว่าค่าคงที่ k
- พยัญชนะ “พ” ; ตำแหน่งปลายนิ้วพิกัดแกน Y อยู่ต่ำกว่าระดับของมือ และค่าความต่างในแนวแกน X น้อยกว่าค่าคงที่ k
- พยัญชนะ “ย” ; ตำแหน่งปลายนิ้วพิกัดแกน Y อยู่ต่ำกว่าระดับของมือ และค่าความต่างในแนวแกน X มากกว่าค่าคงที่ 4k
- พยัญชนะ “ล” ; ค่าความต่างพิกัดแกน Y มากกว่าความสูงนิ้วชี้ และค่าความต่างพิกัดแกน X น้อยกว่าค่าคงที่ k

(3) 3 นิ้ว

- พยัญชนะ “ว” ; ตำแหน่งปลายนิ้วพิกัดแกน Y อยู่ในระดับเดียวกัน

2) กำหนดให้ พยัญชนะแรก (firstLetter) ได้มาจากผลลัพธ์ข้อ 1)

3) ตรวจสอบพยัญชนะแรก และจำนวนนิ้วมือ

(1) 1 นิ้ว

- พยัญชนะ “ข” ; ถ้าพยัญชนะแรกคือ “ก”
- พยัญชนะ “ป” ; ถ้าพยัญชนะแรกคือ “พ”
- พยัญชนะ “ฏ” ; ถ้าพยัญชนะแรกคือ “ด”
- พยัญชนะ “ญ” ; ถ้าพยัญชนะแรกคือ “ย”
- พยัญชนะ “ฬ” ; ถ้าพยัญชนะแรกคือ “ล”

(2) 2 นิ้ว

- พยัญชนะ “ค” ; ถ้าพยัญชนะแรกคือ “ก”
- พยัญชนะ “ผ” ; ถ้าพยัญชนะแรกคือ “พ”

(3) 3 นิ้ว

- พยัญชนะ “ฆ” ; ถ้าพยัญชนะแรกคือ “ก”
- พยัญชนะ “ภ” ; ถ้าพยัญชนะแรกคือ “พ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้วยระบบสามารถแปลพยัญชนะได้จากการตรวจจับตำแหน่งปลายนิ้วมือ จึงไม่สามารถแปลพยัญชนะที่มีท่าทางกำมือ หรือท่าทางที่นิ้วมือแนบชิดกัน ดังนั้นพยัญชนะที่สามารถแปลได้ในระบบมีจำนวนทั้งหมด 15 ตัวอักษร ดังรูปที่ 3.4 ซึ่งในสี่กรอบสี่เหลี่ยมเส้นทึบ คือ พยัญชนะท่ามือเดียว และในกรอบสี่เหลี่ยมเส้นประ คือ พยัญชนะสองท่ามือ ซึ่งเป็นพยัญชนะที่มีท่ามือต่อเนื่องมาจากพยัญชนะท่ามือเดียว

รูปที่ 3.4 พยัญชนะที่สามารถแปลได้ในระบบ

3.2.2 การแปลคำ

คำศัพท์ในภาษามือนั้นจะมีท่าทางลักษณะเฉพาะของแต่ละคำ ยกตัวอย่างคำที่มีท่าทางอย่างง่าย เช่น คำว่า “ฉัน” “ไป” “โรงเรียน” ซึ่งมีการแสดงท่าทาง คือ

1) “ฉัน” ท่าทางที่แสดงคือ ใช้มือข้างขวาชี้เข้าหาตัวเอง อีกมือหนึ่งที่เหลือไม่แสดงท่าทางอะไร ดังรูปที่ 3.5(ก)
2) “ไป” ท่าทางที่แสดงคือ วางมือซ้ายในท่าชี้ไปทางขวาอยู่ที่ระดับอก แล้วใช้มือขวาซึ่งอยู่ในท่าชี้ไปทางซ้ายเคลื่อนไหววนรอบมือซ้าย ดังรูปที่ 3.5(ข)
3) “โรงเรียน” ท่าทางที่แสดงคือ มือซ้ายแบมือขึ้น มือขวาคว่ำมือลง ทำการประกบกันแบบขึ้นลงซ้ำสองครั้ง ดังรูปที่ 3.5(ค)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(ก) (ข) (ค)

รูปที่ 3.5 ภาษามือภาษาไทย

(ก) ท่าทางแสดงคำว่า "ฉัน"

(ข) ท่าทางแสดงคำว่า "ไป"

(ค) ท่าทางแสดงคำว่า "โรงเรียน"

จากตัวอย่างคำภาษามือข้างต้น จะพบว่าการเคลื่อนไหวอวัยวะที่สำคัญคือ มือทั้งสองข้าง ผู้จัดทำจึงทำการเลือกส่วนของข้อต่อที่สำคัญ (Joint of Interest: JOI) คือ ข้อต่อของมือทั้งสองข้าง และเลือกข้อต่อข้อศอกซ้ายและขวา เพื่อช่วยเพิ่มความแม่นยำของตำแหน่งการเคลื่อนไหว รวมทั้งหมดเป็น 4 จุด ได้แก่

1) ข้อต่อมือซ้าย (HandLeft)
2) ข้อต่อข้อศอกซ้าย (ElbowLeft)
3) ข้อต่อมือขวา (HandRight)
4) ข้อต่อข้อศอกขวา (ElbowRight)

3.3 ขั้นตอนการทำงานของระบบ

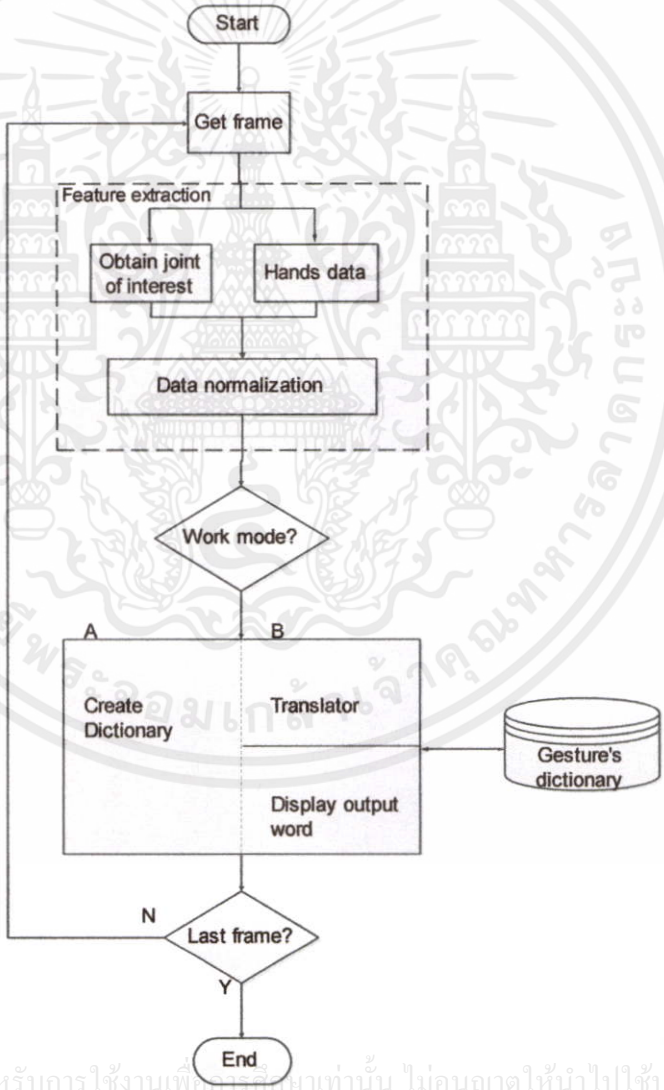
เริ่มแรกผู้พิการทางการได้ยินมายืนอยู่ข้างหน้ากล้องและแสดงท่าทางภาษามือ เราจะได้รับเฟรมใหม่ควบคู่ไปกับแสดงผลภาพวีดีโอ ระบบจะทำการอัพเดตข้อมูล Skeleton Frame และ Depth Frame ของผู้ใช้งาน และทำการดึงคุณลักษณะ (Feature Extraction) ซึ่งจะมี 3 ขบวนการประมวลผล ได้แก่

1) ขบวนการเก็บข้อมูลของข้อต่อที่สนใจ (Joint of Interest : JoI) ทำการดึงข้อมูลจาก Skeleton Frame ข้อมูลที่ได้คือพิกัดแนวแกน X Y และ Z ของจุดข้อต่อต่าง ๆ ที่สนใจ ซึ่งมีทั้งหมด 4 จุด ได้แก่ มือซ้าย มือขวา ข้อศอกซ้าย และข้อศอกขวา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ขบวนการเก็บข้อมูลของมือ (Hand Data) จากการดึงข้อมูลจาก Depth Frame ผ่านกระบวนการ Fingertip Detection เพื่อระบุตำแหน่งปลายนิ้วมือ
3) ขบวนการปรับปรุงค่าข้อมูล (Data Normalization) เพื่อปรับข้อมูลจากขบวนการที่ 1 เป็นค่าคงที่

หลังจากการดึงคุณลักษณะ จะมีการประมวลผลระหว่างข้อมูลและฐานข้อมูล ซึ่งแบ่งออกเป็น 2 ระบบย่อย คือ ขั้นตอนการสร้างฐานข้อมูลภาษามือ และขั้นตอนการแปลภาษามือ ดังรูปที่ 3.6



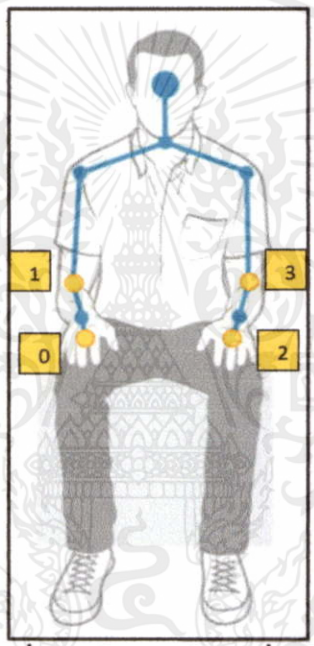
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.6 แผนผังขั้นตอนการทำงานของระบบ

3.4 การดึงคุณลักษะ (Feature Extraction)

3.4.1 ข้อมูลของข้อต่อที่สนใจ

ข้อมูลที่ได้รับจากการตรวจจับตำแหน่งข้อต่อที่สนใจ ซึ่งมีทั้งหมด 4 จุด แต่ละข้อจะประกอบไปด้วย ค่าพิกัด X Y และ Z โดยเรียงต่อกันเป็น Array แล้วส่งต่อไปที่ขบวนการปรับปรุงค่าข้อมูลและบันทึกค่าให้อยู่ในรูป String

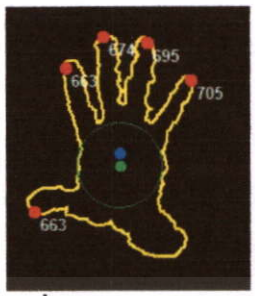


รูปที่ 3.7 ตำแหน่งข้อต่อที่สนใจ

3.4.2 ข้อมูลของมือส่วนที่สนใจ

การดึงข้อมูลของมือมาดำเนินการโดยใช้เทคนิค Fingertip Recognition สำหรับการตรวจจับตำแหน่งปลายนิ้วมือ ซึ่งจะทำโดยการตัดภาพมือออกจากพื้นหลัง ผ่าน Algorithm ในการหา Convexity Defects เพื่อระบุตำแหน่งของปลายนิ้ว ดังรูปที่ 3.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 ข้อมูลภาพมือ

3.4.3 ขบวนการปรับปรุงค่าข้อมูล

ขบวนการปรับปรุงค่าข้อมูล (Data Normalization) เป็นการปรับค่าข้อมูลจากขบวนการที่ 1 เพื่อให้เป็นค่าคงที่ โดยทำการออกแบบเป็นค่าคงที่ของอัตราส่วน ซึ่งมาจากความกว้างของตัวผู้ใช้งาน (shoulderDist) ต่อระยะห่าง (Distance) ซึ่งเป็นระยะจากกึ่งกลางลำตัว (center) ถึงตำแหน่งพิกัดของข้อต่อที่สนใจ ($Point[i]_{(x,y,z)}$) ดังสมการ 3.1 โดยค่า shoulderDist และ Distance(x,y,z) คำนวณจากสมการ 3.2 3.3 และ 3.4

$$Normalize = \frac{Distance_{(x,y,z)}}{shoulderDist} \tag{3.1}$$

$$shoulderDist = \left| \overrightarrow{shoulderLeft_{(x,y,z)} shoulderRigth_{(x,y,z)}} \right| \tag{3.2}$$

$$Distance_{(x,y,z)} = Point[i]_{(x,y,z)} - center_{(x,y,z)} \tag{3.3}$$

$$center_{(x,y,z)} = \frac{shoulderLeft_{(x,y,z)} + shoulderRigth_{(x,y,z)}}{2} \tag{3.4}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.9 ค่าคงที่จากการปรับปรุงค่าข้อมูล

3.5 กระบวนการแปล

หลังจากการเก็บข้อมูลที่ต้องการได้แล้ว ระบบจะทำการแปลโดยการเปรียบเทียบข้อมูลที่รับเข้ามา เทียบกับฐานข้อมูล ด้วยเทคนิค Dynamic Time Warping: DTW ซึ่งเป็นการเปรียบความคล้ายของข้อมูลที่มีความแตกต่างกัน ในด้านเวลา

3.6 ส่วนติดต่อผู้ใช้

ส่วนติดต่อผู้ใช้ (User Interface) แบ่งเป็น 7 ส่วน ได้แก่

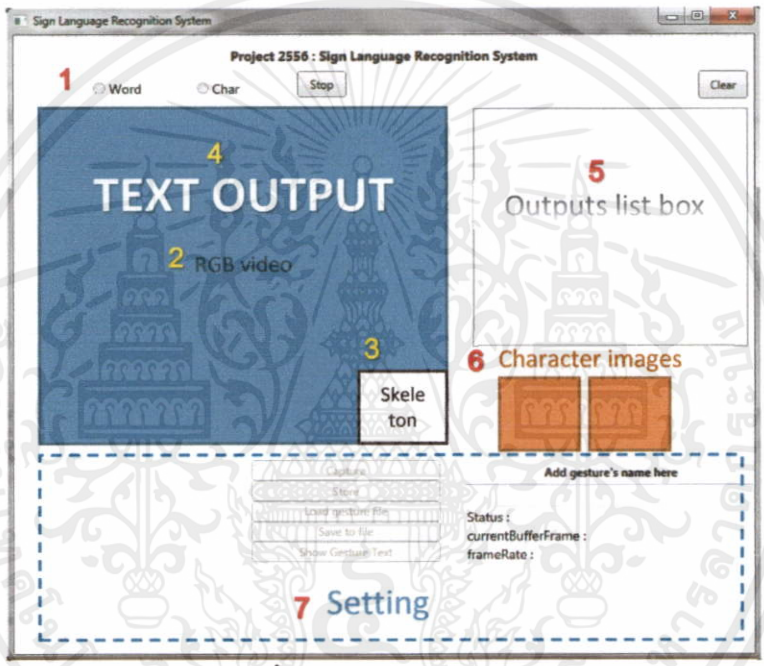
1) Mode คือ ส่วนสำหรับเลือกโหมดการทำงานระหว่าง การแปลคำ และการแปลพยัญชนะ
2) RGB Video คือ ส่วนสำหรับแสดงผลภาพวีดีโอการเคลื่อนไหวซึ่งเป็นภาพสี RGB 24 bits
3) Skeleton Frame คือ ส่วนสำหรับแสดงภาพการเคลื่อนที่ของตำแหน่งข้อต่อต่าง ๆ
4) Text Output คือ คือ ส่วนสำหรับแสดงคำ หรือพยัญชนะ ซึ่งเป็นผลลัพธ์จาการแปลของระบบ
5) Output Box List คือ ส่วนสำหรับแสดงข้อความจากระบบ และผลลัพธ์การแปลที่ผ่านมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6) Character Images คือ ส่วนสำหรับแสดงภาพท่ามือของพยัญชนะซึ่งเป็นผลลัพธ์และผลลัพธ์ที่ใกล้เคียง จากการแปลของระบบ ซึ่งการแสดงภาพประกอบ ใช้เพื่อเป็นทางเลือกให้กับผู้ใช้สำหรับบางพยัญชนะที่มีความใกล้เคียงกัน

7) Setting คือ ส่วนการดำเนินการต่าง ๆ ได้แก่ สร้างฐานข้อมูล การนำเข้าข้อมูลจากฐานข้อมูล และการบันทึกฐานข้อมูล

ตัวอย่างหน้าจอติดต่อผู้ใช้ดังแสดงในรูปที่ 3.10



รูปที่ 3.10 User Interface

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทดลอง และผลการทดลอง

บทนี้จะกล่าวถึงการทดลองติดตั้งอุปกรณ์ ผลการทดลองระบบการแปลคำและพยัญชนะ และผลการทดสอบประสิทธิภาพการทำงานของระบบ

4.1 การทดลองการติดตั้งอุปกรณ์

4.1.1 เครื่องมือที่ใช้ในการพัฒนา

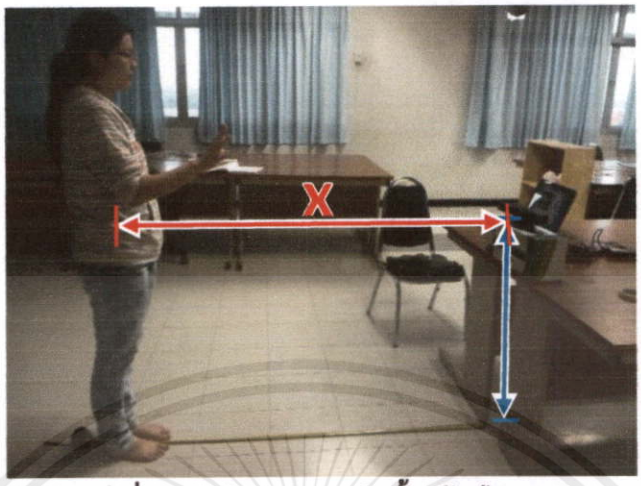
ระบบรู้จำภาษามือพัฒนาขึ้นด้วยภาษา C# โดยโปรแกรม Microsoft Visual Studio 2012 บนระบบปฏิบัติการ Windows 7 ติดตั้ง Microsoft Kinect SDKs เวอร์ชัน 1.7 เพื่อใช้ Package "Microsoft.Kinect.dll" และใช้ Package จาก Candescent NUI [15]

4.1.2 การทดลองการหาระยะระหว่างกล้องและผู้ใช้

เป็นการทดลองเพื่อหาตำแหน่งที่เหมาะสมระหว่างกล้องไคเนคและผู้ใช้งาน ซึ่งเป็นตำแหน่งที่กล้องสามารถตรวจจับข้อต่อที่ต้องการได้ครบทุกจุดขณะที่มีการเคลื่อนไหวร่างกาย และสามารถตรวจจับตำแหน่งปลายนิ้วมือ โดยจำลองการวางตำแหน่งของกล้องตามรูปที่ 4.1 โดยกำหนดค่าคงที่ดังนี้

1) ค่าความสูงของกล้อง ที่ระยะ 0.90 เมตร
2) ค่ามุมเงยของกล้อง ที่ 0 องศา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 ตำแหน่งการติดตั้งกล้องไคเนค

จากนั้นทำการหาระยะทางระหว่างกล้องและผู้ใช้งาน (X) โดยที่กำหนดระยะการทดลองตั้งแต่ 1.60 ถึง 0.70 เมตร ทำการทดลองในระบบการแปลคำ และการแปลพยัญชนะ ซึ่งโหมดการแปลคำจะแสดงผลลัพธ์ภาพจุดข้อต่อโครงกระดูก (Skeleton Image) ที่ระยะต่าง ๆ ในท่ายืนและนั่ง ส่วนโหมดการแปลพยัญชนะ จะแสดงผลลัพธ์จุดปลายนิ้วที่ตรวจจับได้บนภาพสี RGB

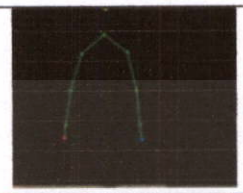
4.1.2.1 การหาระยะทางระหว่างกล้องและผู้ใช้งานในโหมดการแปลคำ

กำหนดให้ข้อต่อที่เป็นจุดสีเขียว หมายถึง สามารถตรวจจับได้ (Tracked) ข้อต่อที่เป็นจุดสีเหลือง หมายถึง ค่าที่ได้จากการประมาณ (Inferred) ข้อมูลตำแหน่งของข้อต่อในส่วนนั้นจะเป็นค่าประมาณ ในบางครั้งอาจเป็นค่าที่ตรวจจับได้ล่าสุดก่อนที่จะตรวจจับไม่ได้ในเวลาต่อมา จุดสีแดง หมายถึง ข้อต่อมือซ้าย และจุดสีน้ำเงิน หมายถึงข้อต่อมือขวา

เมื่อผู้ใช้แสดงท่าทางยืน และท่าทางนั่ง ที่ระยะทาง 1.60 ถึง 0.70 เมตร โดยห่างกันช่วงละ 0.10 เมตร ทำการบันทึกภาพโครงกระดูกในแต่ละระยะ ได้ผลการทดลอง ดังแสดงตารางที่ 4.1 และ 4.2

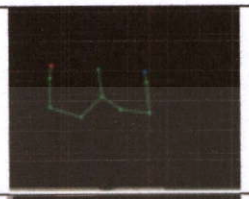
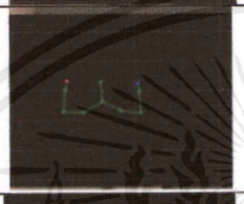
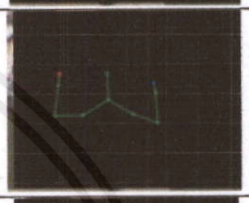
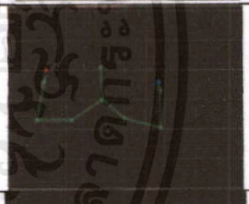
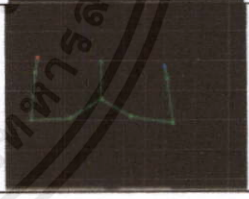
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ผลของการตรวจจับข้อต่อที่ระยะต่าง ๆ ในท่าทางยืน

No.	Distance (m)	Skeleton Image	No.	Distance (m)	Skeleton Image
1	1.60		6	1.10	
2	1.50		7	1.00	
3	1.40		8	0.90	
4	1.30		9	0.80	
5	1.20		10	0.70	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 ผลของการตรวจจับข้อต่อที่ระยะต่างๆ ในท่าทางนั่ง

No.	Distance (m)	Skeleton Image	No.	Distance (m)	Skeleton Image
1	1.60		6	1.10	
2	1.50		7	1.00	
3	1.40		8	0.90	
4	1.30		9	0.80	
5	1.20		10	0.70	

จากการทดลองระยะระหว่างกล้องและผู้ใช้ ที่ระยะทางต่าง ๆ พบว่าในท่าทางยืนที่ระยะทาง 1.60 ถึง 1.20 เมตร กล้องสามารถตรวจจับตำแหน่งข้อต่อได้ครบทุกจุด ที่ระยะทาง 1.10 ถึง 0.80 เมตร กล้องตรวจจับข้อต่อศีรษะไม่ได้ และที่ระยะทาง 0.70 เมตร กล้องไม่สามารถตรวจจับตำแหน่งผู้ใช้แล้วสร้างภาพโครงกระดูกได้ และในท่าทางนั่งที่ระยะทาง 1.70 ถึง 0.70 เมตร หรือทุกระยะการทดลอง กล้องสามารถตรวจจับตำแหน่งข้อต่อได้ครบทุกข้อต่อ แต่หากระยะทางยิ่งไกลก็จะทำให้ ภาพโครงกระดูกนั้นเล็กลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.2.2 การหาระยะทางระหว่างกล้องและผู้ใช้งานในโหมดการแปลพยัญชนะ

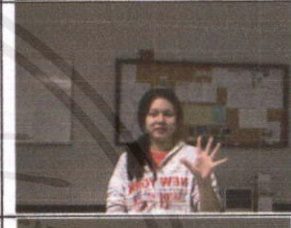
เมื่อผู้ใช้เลือกโหมดการทำงานเป็นการแปลพยัญชนะ จากนั้นยกมือข้างหนึ่งขึ้น ให้ฝ่ามืออยู่ระดับอก แสดงท่าทางยืน และท่าทางนั่ง ที่ระยะทาง 1.60 ถึง 0.70 เมตร โดยห่างกันช่วงละ 0.10 เมตร ทำการบันทึกภาพปลายนิ้วมือในแต่ละระยะ ได้ผลการทดลอง ดังแสดงตารางที่ 4.3 และ 4.4

ตารางที่ 4.3 ผลของการตรวจจับจุดปลายนิ้วที่ระยะต่าง ๆ ในท่าทางยืน

No.	Distance (m)	Hand Image	No.	Distance (m)	Hand Image
1	1.60		6	1.10	
2	1.50		7	1.00	
3	1.40		8	0.90	
4	1.30		9	0.80	
5	1.20		10	0.70	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 ผลของการตรวจจับจุดปลายนิ้วที่ระยะต่าง ๆ ในท่าทางนั่ง

No.	Distance (m)	Hand Image	No.	Distance (m)	Hand Image
1	1.60		6	1.10	
2	1.50		7	1.00	
3	1.40		8	0.90	
4	1.30		9	0.80	
5	1.20		10	0.70	

จากการทดลองในท่าทางยืน และนั่งที่ระยะทางต่าง ๆ พบว่าในท่าทางยืนมีเพียงระยะทาง 1.00 ถึง 1.10 เมตรเท่านั้น ที่กล้องสามารถตรวจจับตำแหน่งปลายนิ้วมือได้ และในท่าทางนั่งมีเพียงระยะทาง 0.80 ถึง 0.90 เท่านั้น ที่กล้องสามารถตรวจจับตำแหน่งปลายนิ้วมือได้

จากการทดลองการหาระยะทางระหว่างกล้องและผู้ใช้ พบว่าท่าทางนั่งที่ระยะทาง 0.80 ถึง 0.90 เมตร เป็นระยะทางที่สามารถตรวจจับได้ทั้งข้อต่อโครงกระดูก และตำแหน่งจุดปลายนิ้ว

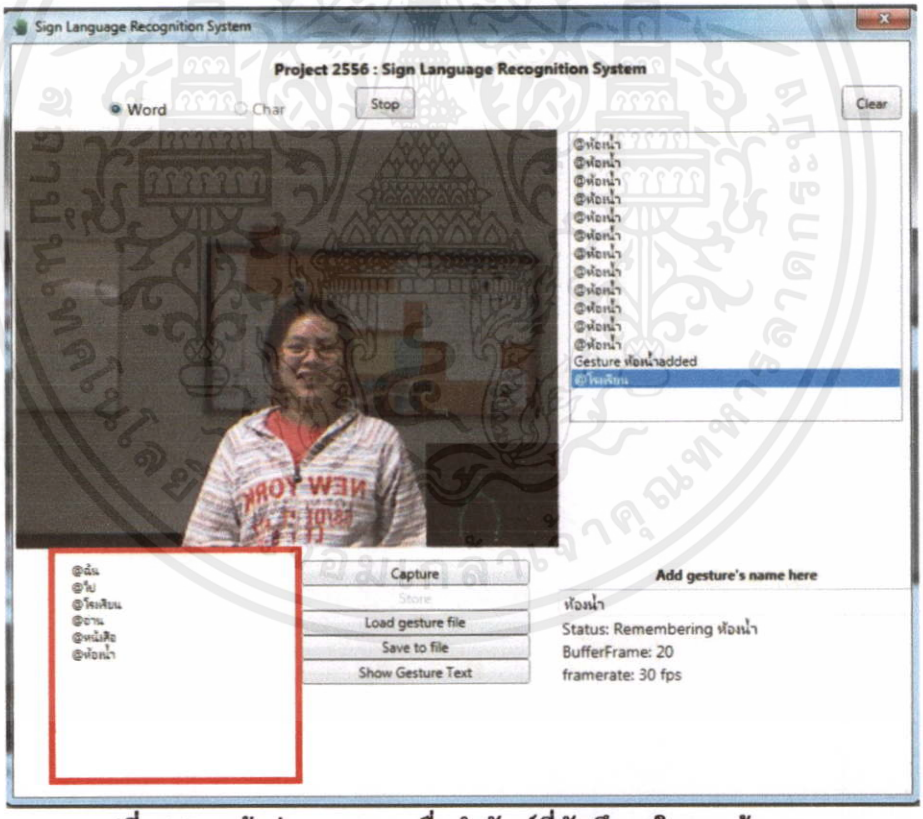
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ ดัดแปลง เผยแพร่ หรือแสวงหาประโยชน์ทางการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้นำไฟล์นี้ไปเผยแพร่ต่อ ไม่ว่าจะมีค่าตอบแทนหรือไม่ก็ตาม

4.2 การสร้างฐานข้อมูลสำหรับการแปลคำ

การแปลคำ เป็นการแปลโดยใช้ข้อมูลตำแหน่งข้อต่อมือซ้าย มือขวา ข้อศอกซ้ายและข้อศอกขวา จำนวนทั้งสิ้น 4 จุด ผ่านขบวนการปรับปรุงค่าข้อมูล ซึ่งได้อธิบายขั้นตอนไว้ในบทที่ 3

จากการทดลองที่ 4.1.2 ผู้จัดทำโครงงานเลือกใช้ระยะทาง 0.80 ถึง 0.90 เมตรในท่าทางนั่งมาทำการทดลองตั้งแต่ขั้นตอนนี้เป็นต้นไป และเริ่มต้นการสร้างฐานข้อมูลโดยการป้อนข้อมูลชื่อของท่าทางภาษามือ แล้วกดปุ่ม Capture จากนั้นแสดงท่าทาง โดยคำศัพท์ที่ทำการเพิ่มลงในฐานข้อมูลได้แก่ ฉัน, ไป, โรงเรียน, อ่าน, หนังสือ และห้องน้ำ

หลังจากที่บันทึกครบทุกท่าทางแล้ว กดปุ่ม Show Gesture Text จะแสดงรายชื่อคำศัพท์ภาษามือที่ได้ทำการบันทึกไว้ (กรอบสี่เหลี่ยมแดง) ดังรูป 4.2 และตัวอย่างข้อมูลที่บันทึกไว้ในรูปของ Text File ดังรูป 4.3



รูปที่ 4.2 หน้าต่างแสดงรายชื่อคำศัพท์ที่บันทึกลงในฐานข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.3 ตัวอย่างข้อมูลที่บันทึกไว้ในรูปของ Textfile

4.3 การทดสอบประสิทธิภาพการแปลคำ

การแปลคำจะมาจากการเปรียบเทียบท่าทางในฐานข้อมูลกับท่าทางตัวอย่าง ด้วยเทคนิคไดนามิกไทม์วอร์ปปิง และในระบบจะต้องทำการกำหนดตัวแปรค่าคงที่ 3 ชนิด คือ

1) ค่าคงที่ BufferSize สำหรับกำหนดจำนวนบัฟเฟอร์ที่ใช้ในการบันทึกข้อมูล
2) ค่าคงที่ Ignore สำหรับกำหนดค่าความถี่ในการบันทึกข้อมูล
3) ค่าคงที่ MinimumFrames (Match Frame) สำหรับกำหนดจำนวนเฟรมที่น้อยที่สุดที่เปรียบเทียบแล้วตรงกับฐานข้อมูล

4.3.1 การทดลองหาค่า BufferSize

จัดทำการทดลองโดยบันทึกข้อมูลการเคลื่อนไหวของท่าทางแต่ละคำ เพื่อหาขนาดของข้อมูลจริง โดยมีความถี่ในการเก็บข้อมูลเป็น 30 เฟรม/วินาที ทำการทดลองแสดงท่าทาง 10 ครั้งต่อหนึ่งคำ และทำการหาค่าเฉลี่ย จากท่าทางที่ใช้ทดลองจำนวน 6 ท่า ซึ่งเฉลี่ยแล้วแต่ละท่ามีจำนวนข้อมูล 33.08 เฟรม คิดเป็นค่าประมาณคือ 34 เฟรม ตามตารางที่ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารนี้ทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 ขนาดข้อมูลการเคลื่อนไหวของท่าทางแต่ละคำ

ที่	ชื่อท่าทาง	ค่าเฉลี่ยจำนวนข้อมูลที่บันทึก (เฟรม)
1	ฉัน	18.4
2	ไป	31.2
3	โรงเรียน	33.5
4	อ่าน	26.3
5	หนังสือ	43.0
6	ห้องน้ำ	46.1

4.3.2 การทดลองหาค่า Ignore และ MinimumFrames (Match Frame)

จากการทดลองในหัวข้อ 4.3.1 เป็นการเก็บข้อมูลแบบทั้งหมด ซึ่งตำแหน่งจะมีการเปลี่ยนแปลงไม่มากในลำดับที่ติดกัน และเพื่อเป็นการลดจำนวนข้อมูลที่ซ้ำซ้อนลง จึงต้องทดลองหาค่าตัวแปร Ignore ที่จะระบุความถี่ในการบันทึกข้อมูล โดยทำการกำหนดค่า BufferSize เท่ากับ 34 แล้วทำการทดลอง ดังนี้

1) กำหนดค่า Ignore เท่ากับ 1 หมายความว่า บันทึกข้อมูลทุกๆ เฟรม ที่ตรวจจับได้ ส่งผลให้ BufferSize มีค่าเท่ากับ 34 เฟรม
2) กำหนดค่า Ignore เท่ากับ 2 หมายความว่า บันทึกข้อมูลทุกๆ 2 เฟรม ที่ตรวจจับได้ ส่งผลให้ BufferSize มีค่าเท่ากับ 17 เฟรม
3) กำหนดค่า Ignore เท่ากับ 3 หมายความว่า บันทึกข้อมูลทุกๆ 3 เฟรม ที่ตรวจจับได้ ส่งผลให้ BufferSize มีค่าเท่ากับ 11 เฟรม

สำหรับการทดลองหาค่า MinimumFrames (Match Frame) นั้นจะกำหนดให้ทดลองที่ระดับความตรงกันต่างๆ ได้แก่ 100% 80% 60% และ 40%

จากการทดลอง เมื่อผู้ใช้แสดงท่าทางภาษามือจำนวน 10 รอบต่อหนึ่งคำ แล้วนับผลลัพธ์การแปลที่ถูกต้องบันทึกผลการทดลอง ดังแสดงในตารางที่ 4.6 4.7 และ 4.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 1 ค่า BufferSize เท่ากับ 34 และค่า MinimumFrames ที่ระดับต่าง ๆ

ชื่อท่าทาง	จำนวนครั้งที่ถูกต้องที่ค่า MinimumFrames ต่างๆ			
	34 f. (100%)	27 f. (80%)	20 f. (60%)	14 f. (40%)
ฉัน	10	10	10	10
ไป	7	10	8	10
โรงเรียน	8	10	6	8
อ่าน	8	9	10	10
หนังสือ	9	10	8	10
ห้องน้ำ	10	6	8	3
เฉลี่ย	**8.67**	**9.16**	**8.33**	**8.50**

ตารางที่ 4.7 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 2 ค่า BufferSize เท่ากับ 17 และค่า MinimumFrames ที่ระดับต่าง ๆ

ชื่อท่าทาง	จำนวนครั้งที่ถูกต้องที่ค่า MinimumFrames ต่างๆ			
	17 f. (100%)	14 f. (80%)	11 f. (60%)	7 f. (40%)
ฉัน	10	10	8	10
ไป	9	7	9	10
โรงเรียน	5	7	7	10
อ่าน	10	10	9	10
หนังสือ	7	9	8	7
ห้องน้ำ	5	5	9	8
เฉลี่ย	**7.67**	**8.00**	**8.33**	**9.16**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.8 จำนวนผลลัพธ์การแปลที่ถูกต้องเมื่อกำหนดค่า Ignore เท่ากับ 3 ค่า BufferSize เท่ากับ 11 และค่า MinimumFrames ที่ระดับต่าง ๆ

ชื่อท่าทาง	จำนวนครั้งที่ถูกต้องที่ค่า MinimumFrames ต่างๆ			
	11 f. (100%)	9 f. (80%)	7 f. (60%)	5 f. (40%)
ฉัน	10	8	6	6
ไป	3	5	6	9
โรงเรียน	1	5	7	7
อ่าน	2	6	10	10
หนังสือ	7	7	8	10
ห้องน้ำ	1	0	4	8
เฉลี่ย	**4.00**	**5.16**	**6.83**	**8.33**

จากการผลการทดลองตารางที่ 4.6, 4.7 และ 4.8 พบว่า ที่ระดับ MinimumFrames 100% 80% 60% และ 40% ส่งผลให้

1) ค่า BufferSize เท่ากับ 34 มีค่าเฉลี่ยจำนวนครั้งที่ถูกต้อง 8.67 9.16 8.33 และ 8.50 ตามลำดับระดับ MinimumFrames
2) ค่า BufferSize เท่ากับ 17 มีค่าเฉลี่ยจำนวนครั้งที่ถูกต้อง 7.67 8.00 8.33 และ 9.16 ตามลำดับระดับ MinimumFrames
3) ค่า BufferSize เท่ากับ 11 มีค่าเฉลี่ยจำนวนครั้งที่ถูกต้อง 4.00 5.16 6.83 และ 8.33 ตามลำดับระดับ MinimumFrames

แสดงให้เห็นว่าจำนวน BufferSize ที่มากกว่าให้ผลความถูกต้องมากกว่า จำนวน BufferSize ที่น้อยกว่า

4.3.3 เปอร์เซ็นต์ความคลาดเคลื่อนของการแปลคำ

การทดสอบหาเปอร์เซ็นต์ความคลาดเคลื่อน กำหนดให้ค่าคงที่ BufferSize เท่ากับ 17 ค่าคงที่ Ignore เท่ากับ 2 และค่าคงที่ MinimumFrames เท่ากับ 14 ทำการทดลองโดยผู้ใช้จำนวน 10 คน ทำการแสดงท่าทางภาษามือคำละ 10 รอบ แล้วบันทึกจำนวนรอบของการแปลที่ถูกต้องลงในตาราง ซึ่งได้ผลการทดลองดังตารางที่ 4.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 ผลลัพธ์การแปลคำจากการแสดงท่าทางของผู้ใช้จำนวน 10 คน

ผู้ใช้ที่ / ชื่อท่าทาง	1	2	3	4	5	6	7	8	9	10	%Error
ฉัน	8	10	10	10	10	10	10	10	10	9	3
ไป	4	5	3	6	10	2	9	9	4	6	42
โรงเรียน	9	9	3	3	10	5	6	9	4	3	39
อ่าน	7	3	4	2	10	6	7	8	2	4	47
หนังสือ	9	10	8	5	2	2	7	7	6	6	38
ห้องน้ำ	8	7	7	1	1	3	3	6	2	3	59
										ค่าเฉลี่ย	38.00

จากการทดลอง เมื่อผู้ใช้จำนวน 10 คน แสดงท่าทางภาษามือที่กำหนดอย่างละ 10 รอบ พบว่าท่าทางที่มีความคลาดเคลื่อนน้อยที่สุดเท่ากับ 3 เปอร์เซ็นต์ คือ ท่าทาง “ฉัน” ซึ่งเป็นท่าทางที่มีความยาวน้อยหรือมีการเคลื่อนไหวน้อยที่สุด และท่าทางที่มีความคลาดเคลื่อนที่มากที่สุดเท่ากับ 59 เปอร์เซ็นต์ คือ ท่าทาง “ห้องน้ำ” ซึ่งเป็นท่าทางที่มีความซับซ้อนมาก ส่วนท่าทางอื่น ๆ มีค่าความคลาดเคลื่อนระหว่าง 39 ถึง 47 เปอร์เซ็นต์ และจากการเฉลี่ยค่าความคลาดเคลื่อนของทุกท่าทางมีความคลาดเคลื่อนเท่ากับ 38.00 เปอร์เซ็นต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การทดสอบประสิทธิภาพการแปลพยัญชนะ

พยัญชนะที่กำหนดไว้ในระบบมีทั้งสิ้น 15 ตัวอักษรแบ่งเป็นพยัญชนะท่ามือเดียว 6 ตัว ได้แก่ ก ด พ ย ล และ ว ดังแสดงในรูปที่ 4.4 และพยัญชนะสองท่ามืออีก 9 ตัว ได้แก่ ข ค ฆ ป ผ ภ ฏ ญ และ ฬ ซึ่งจะต้องทำการท่าทางการนับต่อจากท่าแรก โดยการชูหนึ่งนิ้ว สองนิ้ว หรือสามนิ้ว ดังแสดงในรูปที่ 4.5

พยัญชนะ “ก” พยัญชนะ “ด” พยัญชนะ “พ”

พยัญชนะ “ย” พยัญชนะ “ล” พยัญชนะ “ว”

รูปที่ 4.4 ท่าทางพยัญชนะท่ามือเดียว

รูปที่ 4.5 ท่าทางการนับสำหรับพยัญชนะสองท่ามือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จัดทำการทดลองโดยให้ผู้ใช้จำนวน 10 คน แสดงท่าทางของแต่ละพยัญชนะด้วยมือขวา พยัญชนะละ 10 รอบ ซึ่งมีวิธี ดังนี้

1) แสดงท่ามือของพยัญชนะที่มีท่ามือเดียว ได้แก่ ก ด พ ย ล และ ว ดังรูปที่ 4.4 โดยแต่ละท่าจะต้องทำค้างไว้ประมาณ 1 วินาที แล้วกำมือเพื่อไม่ให้ระบบทำการแปลท่ามือถัดไป
2) แสดงท่ามือของพยัญชนะที่เป็นสองท่ามือ ได้แก่
   2.1) พยัญชนะ “ข” จะต้องแสดงท่า “ก” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 1 นิ้ว
   2.2) พยัญชนะ “ค” จะต้องแสดงท่า “ก” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 2 นิ้ว
   2.3) พยัญชนะ “ฆ” จะต้องแสดงท่า “ก” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 3 นิ้ว
   2.4) พยัญชนะ “ฎ” จะต้องแสดงท่า “ด” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 1 นิ้ว
   2.5) พยัญชนะ “ป” จะต้องแสดงท่า “พ” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 1 นิ้ว
   2.6) พยัญชนะ “ผ” จะต้องแสดงท่า “พ” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 2 นิ้ว
   2.7) พยัญชนะ “ภ” จะต้องแสดงท่า “พ” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 3 นิ้ว
   2.8) พยัญชนะ “ญ” จะต้องแสดงท่า “ย” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 1 นิ้ว
   2.9) พยัญชนะ “ฬ” จะต้องแสดงท่า “ล” ค้างไว้ 1 วินาที แล้วนับหนึ่งโดยชู 1 นิ้ว

แล้วทำการบันทึกจำนวนรอบที่ระบบสามารถแปลได้ถูกต้อง ลงในผลการทดลอง ซึ่งแสดงดังตารางที่ 4.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.10 ผลลัพธ์การแปลพยัญชนะจากการแสดงท่าทางของผู้ใช้จำนวน 10 คน

ผู้ใช้ที่ / พยัญชนะ	1	2	3	4	5	6	7	8	9	10	%Error
ก	10	8	7	4	6	5	6	7	9	8	30
ข	9	7	8	6	8	6	8	8	6	7	27
ค	10	10	10	8	9	10	10	10	10	10	3
ฆ	9	10	9	3	7	8	8	5	3	10	28
พ	10	8	6	6	7	8	10	6	8	9	22
ป	9	6	7	3	6	3	5	8	3	8	42
ผ	10	10	5	9	8	10	9	7	8	8	16
ภ	5	6	5	2	2	5	6	3	3	4	59
ด	7	9	8	8	6	4	3	6	5	8	36
ฏ	10	10	9	10	7	10	9	8	10	10	7
ว	6	8	10	6	4	9	5	7	6	7	32
ย	10	10	9	9	8	10	10	8	9	10	7
ญ	5	9	7	8	6	5	7	6	5	7	35
ล	10	10	8	9	8	10	9	9	9	8	10
ฬ	10	10	8	8	8	9	8	9	9	10	11
										ค่าเฉลี่ย	24.33

จากการทดลองเมื่อผู้ใช้จำนวน 10 คน แสดงท่าทางพยัญชนะภาษามือที่กำหนดอย่างละ 10 รอบ นำผลลัพธ์มาเรียงลำดับเปอร์เซ็นต์ความคลาดเคลื่อนจากน้อยไปมากจะได้ พยัญชนะ ค ฏ ย ล ฬ ผ พ ข ฆ ก ว ญ ด ป และ ภ ซึ่งท่าทางพยัญชนะที่มีความคลาดเคลื่อนน้อยที่สุดเท่ากับ 3 เปอร์เซ็นต์ คือ พยัญชนะ “ค” ส่วนท่าทางที่มีความคลาดเคลื่อนที่มากที่สุดเท่ากับ 59 เปอร์เซ็นต์ คือ ท่าทาง “ภ” และจากการเฉลี่ยค่าความคลาดเคลื่อนของทุกท่าทางพยัญชนะมีความคลาดเคลื่อนเท่ากับ 24.33 เปอร์เซ็นต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
บทสรุปและข้อเสนอแนะ

5.1 บทสรุป

ระบบรู้จำภาษามือมีการออกแบบให้ทำการแปลจากท่าทางการเคลื่อนไหวของภาษามือแปลไปเป็นคำหรือพยัญชนะ และเป็นระบบทิศทางเดียวไม่มีการแปลย้อนกลับไปเป็นภาษามือ โดยระบบรู้จำภาษามือใช้เทคโนโลยีของกล้องไคเนคเป็นเครื่องมือในการตรวจจับการเคลื่อนไหวของผู้พิการทางการได้ยิน

ระบบการทำงานแบ่งออกเป็น 2 ระบบย่อย คือ การแปลคำ และการแปลพยัญชนะ โดยการแปลคำจะมีการแสดงผลภาพเคลื่อนไหว 2 รูปแบบคือ ภาพสี และภาพตำแหน่งข้อต่อ ซึ่งทั้งหมดแสดงที่ความละเอียด 640x480 พิกเซล 30 เฟรม/วินาที การบันทึกฐานข้อมูลจะใช้ข้อมูลการเคลื่อนไหวของข้อต่อที่สนใจจำนวน 4 จุด ได้แก่ ข้อต่อมือซ้าย มือขวา ข้อศอกซ้ายและข้อศอกขวา ส่วนการแปลพยัญชนะจะทำการแสดงผลภาพสี ภาพตำแหน่งจุดปลายนิ้วมือ และภาพพยัญชนะภาษามือ โดยใช้ตำแหน่งพิกัดระบบสองมิติของจุดปลายนิ้วมือในการแปลพยัญชนะ

จากการทดลองการติดตั้งอุปกรณ์ จะต้องวางกล้องที่ตำแหน่งความสูง 0.90 เมตร ค่ามุมเงยกล้องอยู่ที่ 0 องศา จะสามารถบันทึกตำแหน่งข้อต่อได้ครบทุกจุดเมื่อผู้ใช้อยู่ที่ตำแหน่งห่างจากกล้องไคเนคที่ระยะทาง 0.80 ถึง 0.90 เมตร ในท่าทางนั่ง ซึ่งเป็นระยะทางที่ผู้ทำโครงงานใช้เป็นตำแหน่งในการเก็บข้อมูลเพื่อสร้างฐานข้อมูล

จากการทดลองหาประสิทธิภาพการแปลคำ จากคำที่ใช้ทดลองทั้งหมด 6 คำ โดยผู้ใช้จำนวน 10 คน แสดงท่าทางคำละ 10 รอบ มีค่าความคลาดเคลื่อนเฉลี่ยอยู่ที่ 38.00 เปอร์เซ็นต์ ซึ่งคำที่มีความคลาดเคลื่อนน้อยที่สุดอยู่ที่ 3 เปอร์เซ็นต์ คือคำว่า “ฉัน” และคำที่มีความคลาดเคลื่อนมากที่สุดอยู่ที่ 59 เปอร์เซ็นต์ คือคำว่า “ห้องน้ำ”

จากการทดลองการแปลพยัญชนะ ซึ่งมีในระบบทั้งสิ้น 15 ตัว ทำการทดลองโดยผู้ใช้จำนวน 10 คน แสดงท่าทางพยัญชนะละ 10 รอบ มีค่าความคลาดเคลื่อนเฉลี่ยอยู่ที่ 24.33 เปอร์เซ็นต์ ซึ่งคำที่มีความคลาดเคลื่อนน้อยที่สุดอยู่ที่ 3 เปอร์เซ็นต์ คือพยัญชนะ “ค” และคำที่มีความคลาดเคลื่อนมากที่สุดอยู่ที่ 59 เปอร์เซ็นต์ คือพยัญชนะ “ภ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ปัญหาอุปสรรคต่างๆ

ในการทำระบบรู้จำภาษามือนี้ ปัญหาที่พบคือ

1) การศึกษาค้นคว้า มีงานวิจัยต่าง ๆ ให้ได้ศึกษามากมาย ทำให้ต้องทำการศึกษางานต่าง ๆ ทีละเรื่อง เพื่อดูว่างานชิ้นไหนที่เหมาะสำหรับนำมาใช้ และเมื่อได้ตัวอย่างงานในลักษณะที่ต้องการแล้วก็ต้องทำการศึกษา และทดลองการใช้โปรแกรมต่าง ๆ ในการควบคุมการทำงานของกล้องไคเนค ซึ่งก็สามารถที่จะเขียนได้หลายภาษา จึงต้องทำการศึกษาวิธีการเขียนโปรแกรม โดยที่ในเริ่มแรกจะใช้โปรแกรมเหมือนกันงานวิจัยที่ใช้เป็นแบบอย่างก่อน ซึ่งก็มีซอฟต์แวร์ในการใช้งานเกี่ยวกับกล้องไคเนคมากกว่าหนึ่งชนิด เมื่อได้ทำการติดตั้งซอฟต์แวร์ต่าง ๆ เรียบร้อยแล้ว ก็ต้องมาศึกษาวิธีการใช้งานโปรแกรมก่อนว่าใช้งานอย่างไร ซึ่งใช้เวลาพอสมควร
2) ขั้นตอนการเขียนโปรแกรมจะใช้เวลานาน เพราะต้องศึกษาเรื่องการเขียนโปรแกรมไปพร้อมกับการทำความเข้าใจในโค้ดต่าง ๆ ไปด้วย ในส่วนของการเขียนโปรแกรมนี้มีปัญหาและอุปสรรคต่าง ๆ มากมาย เช่น
    (1) การเขียนโปรแกรมแล้วพบข้อผิดพลาด
    (2) การแก้โปรแกรมให้ผลลัพธ์ออกมาตามที่ต้องการซึ่งกว่าจะได้ผลลัพธ์ที่ต้องการฟังก์ชันการทำงานบางส่วนก็ต้องใช้เวลาในการศึกษาให้สามารถเขียนโปรแกรมได้
    (3) ในขั้นตอนการแปลภาษามือออกมาเป็นคำ เพราะคำแต่ละคำในภาษามือจะมีท่าทางภาษามือที่ไม่เหมือนกัน ระยะเวลาในการทำภาษามือแต่ละจึงแตกต่างกันไป ทำให้เป็นอุปสรรคในการกำหนดค่า BufferSize (Buffer ที่เก็บตำแหน่งของข้อต่อที่สนใจในจำนวนเฟรมตามที่ระบุ) ถ้ากำหนดค่าของ BufferSize น้อยเกินไปท่าทางภาษามือของคำบางคำอาจไม่สามารถทำการแปลออกมาได้ เช่น ภาษามือคำว่า “ห้องน้ำ” เป็นท่าทางภาษามือที่ต้องทำมือ 3 ขั้นตอน ซึ่งจากการทดลองจะใช้ BufferSize อยู่ที่ 47 แต่ถ้ากำหนดให้ BufferSize อยู่ที่ 34 ทำให้เฟรมที่ใช้ในการจับคู่น้อยลงไปด้วย เมื่อจำนวนเฟรมที่จับคู่น้อยลง ค่าความถูกต้องของผลลัพธ์ก็น้อยลงด้วย ทำให้ตรวจจับคำว่า “ห้องน้ำ” ได้ยาก อีกทั้งยังมีปัญหาที่โครงกระดูกทับซ้อนกัน เช่น การไขว้มือ การเอาแขนมาทับกัน เป็นต้น ทำให้ค่าตำแหน่งที่ได้มีความผิดเพี้ยน
    (4) การแปลพยัญชนะจะมีปัญหาในเรื่องของพยัญชนะที่มีท่าต่อเนื่องแต่ท่าทางคล้ายกันกับท่าแรก เช่น “ค” ที่มีท่าทางคล้ายท่าพยัญชนะ“ก” คือชูนิ้วชี้กับนิ้วกลาง หากทำท่า “ก” ค้างไว้ระบบจะแปลเป็นพยัญชนะ “ค” ทำให้ในการแปลพยัญชนะต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพิ่มการแนะนำพยัญชนะตัวที่มีท่าทางภาษามือที่ใกล้เคียงกัน และส่วนการแสดงผลภาพสีกับภาพตำแหน่งจุดปลายนิ้วมือ ภาพตำแหน่งปลายนิ้วมือจะอยู่เยื้องจากนิ้วมือไปเล็กน้อย เพราะเราเอาภาพตำแหน่งปลายนิ้วมือที่ทำการคำนวณมาจากภาพความลึก (Depth Image) มาซ้อนทับกับภาพสี (RGB Image) ซึ่งภาพทั้งสองภาพนี้มาจากกล้องคนละตัว ทำให้ค่าตำแหน่งที่ได้ไม่เท่ากัน เมื่อนำมาซ้อนทับกันภาพจุดตำแหน่งปลายนิ้วมือจึงไม่อยู่ตรงตำแหน่งปลายนิ้วมือพอดี และพยัญชนะบางตัวจะไม่สามารถทำการตรวจจับได้ เช่น “บ” “ส” เป็นต้น เพราะเป็นท่าทางภาษามือที่มีการกำมือ แบบมือโดยนิ้วชิดกัน ซึ่งวิธีการที่เราใช้ไม่สามารถจะทำการตรวจจับตำแหน่งปลายนิ้วมือได้

3) ในการประมวลผลโปรแกรม การแปลออกมาเป็นพยัญชนะตำแหน่งของการตรวจจับมือจากกล้องมายังผู้ใช้งานมีการจำกัดระยะในแนวแกน (Z) ซึ่งมีในช่วงแคบ ดังนั้นถ้าอยู่นอกระยะที่กำหนดจะไม่สามารถทำการตรวจจับมือ และตำแหน่งปลายนิ้วมือได้

5.3 แนวทางในการพัฒนาต่อ

แนวทางในการพัฒนาระบบรู้จำภาษามือด้วยกล้องไคเนคต่อในอนาคต คือ ทำการศึกษาและแก้ไขปัญหาต่างๆ ดังนี้

1) กำหนดขนาด BufferSize ให้มีประสิทธิภาพมากยิ่งขึ้น โดยอาจจะทำการกำหนดขนาดของ BufferSize ให้แต่ละคำ เพื่อประหยัดขนาดของ Buffer เพราะถ้าใช้ Buffer ไม่หมด ส่วนที่เหลือของ Buffer จะกลายเป็นขยะ
2) ในการแปลภาษามือมาเป็นคำ ต้องหาวิธีการในการไม่ให้ค่าตำแหน่งของการทับกันของโครงกระดูกออกมาผิดเพี้ยนมากเกินไป ทางที่ดีหาวิธีที่ทำให้ระบบมีความคงที่มากที่สุดจะดีกว่า
3) ในการแปลภาษามือมาเป็นพยัญชนะ ต้องหาวิธีการในการตรวจจับนิ้วมือได้แม้ว่าจะกำมือหรือแบมือโดยนิ้วชิดติดกันก็ตาม
4) ในการแปลภาษามือมาเป็นพยัญชนะ อาจจะเพิ่มช่วงระยะห่างระหว่างกล้องกับผู้ใช้งานให้มีช่วงที่กว้างขึ้น เพื่อให้สามารถตรวจจับมือและตำแหน่งปลายนิ้วมือได้มากขึ้น
5) การทำให้ผลลัพธ์ออกมาอยู่ในรูปของประโยค โดยทางเราได้มีแนวคิดการทำผลลัพธ์ออกมาในรูปของประโยคนั้นอาจจะต้องเพิ่มตัวแปรเพื่อระบุคุณลักษณะ (Attribute) ประธาน กริยา หรือกรรม ซึ่งทำการระบุคุณลักษณะในขั้นตอนสร้างฐานข้อมูลคำศัพท์ แล้วทำการแสดงผลลัพธ์ โดยการตรวจสอบคุณลักษณะว่าเป็นประธาน หรือกริยา หรือเป็นกรรม เพื่อทำการเรียงลำดับการแสดงผลลัพธ์ แล้วสร้างเป็นประโยคอย่างง่ายขึ้นมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] มหาวิทยาลัยราชภัฏสวนดุสิต. **"ภาษามือเป็นอย่างไร?"** [Online]. Available : http://www.dssdusit.com/aural/sign. 2013.

[2] ศูนย์การศึกษาพิเศษประจำจังหวัดสุราษฎร์ธานี. **"เนื้อหาทางวิชาการเรื่องความหมาย-ความสำคัญของภาษามือ"** [Online]. Available : http://web.specialsurat.go.th/PDF/hand.pdf. 2013.

[3] สำนักงานส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการแห่งชาติ(พก.). **"พจนานุกรมภาษามือไทย"** [Online]. Available : http://www.nep.go.th/index.php?mod=newsmedia_detail&id=254. 2013

[4] สำนักงานคณะกรรมการกิจการโทรคมนาคมแห่งชาติ (กสทช.). **"พจนานุกรมอิเล็กทรอนิกส์ภาษามือไทย"** [Online]. Available : http://e-tsl.com/indexx.php. 2013.

[5] WIKIPEDIA. **"Kinect"** [Online]. Available : http://en.wikipedia.org/wiki/Kinect. 2014.

[6] Microsoft Developer Network. **"Kinect for Windows Sensor Components and Sepcifications"** [Online]. Available : http://msdn.microsoft.com/en-us/library/jj131033.aspx. 2014.

[7] Jungong Han, Ling Shao, Dong Xu and Jamie Shotton. **"Enhanced Computer Vision with Microsoft Kinect Sensor: A Review"** [Online]. Available : https://www.yumpu.com/en/document/view/15127932/kinectreview-tc2013. 2013.

[8] WordPress.com. **"หลักการทำงานและประยุกต์ใช้งานของไคเนค(Kinect)"** [Online]. Available : http://langisser.wordpress.com/2011/03/07/หลักการทำงานและประย/. 2013.

[9] Microsoft. **"Skeleton Tracking"** [Online]. Available : http://msdn.microsoft.com/en-us/library/hh973074.aspx. 2013.

[10] Microsoft. **"Coordinate Spaces"** [Online]. Available : http://msdn.microsoft.com/en-us/library/hh973078.aspx#Depth_Ranges. 2013.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[11] Thomas Royal. **"A Highlevel Description of two fingertip tracking technique: k-curvature and convexity defects"** [Online]. Available : http://tmroyal.com/a-high-level-description-of-two-fingertip-tracking-techniques-k-curvature-and-convexity-defects/. 2014.

[12] OpenCV. **"Structural Analysis and Shape Descriptors"** [Online]. Available : http://docs.opencv.org/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html#void. 2014.

[13] Aniket Tatipamula. **"Hand gesture using opencv"** [Online]. Available : http://anikettatipamula.blogspot.com/2012/02/hand-gesture-using-opencv.html. 2014.

[14] Simene Andresen. **"Hand Tracking and Recognition With OpenCV"** [Online]. Available : http://simena86.github.io/blog/2013/08/12/hand-tracking-and-recognition-with-opencv/. 2014.

[15] Daniel James Ryan. **"Finger and gesture recognition with Microsoft Kinect"** [Online]. Available : http://brage.bibsys.no/uis/bitstream/URN:NBN:no-bibsys_brage_33088/1/Ryan,%20Daniel%20James.pdf. 2014.

[16] Daniel Martínez Capilla. **MSc Thesis VIBOT Sign Language Translator using Microsoft Kinect XBOX 360™**. 2012.

[17] Stefan Stegmueller. **"Candescent NUI"** [Online]. Available : http://www.candescent.ch/CandescentNui.aspx. 2014.

[18] CodePlex. **"Candescent NUI"** [Online]. Available : http://candescentnui.codeplex.com/. 2014.

[19] CodePlex. **"Kinect SDK Dynamic Time Warping (DTW) Gesture Recognition"** [Online]. Available : http://kinectdtw.codeplex.com/. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

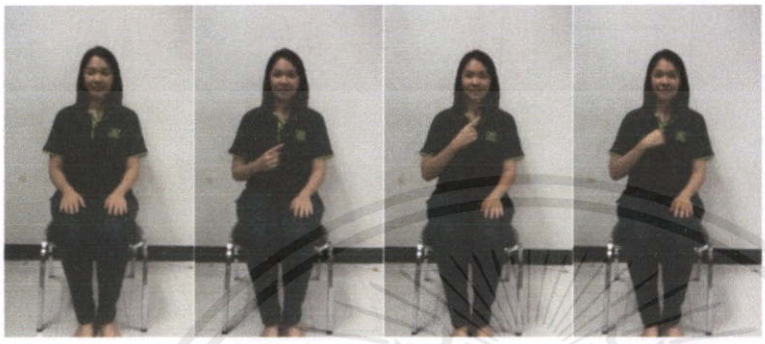
ภาคผนวก ก

ท่าทางภาษามือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก1 ท่าทางภาษามือของคำที่ใช้ในการทดลอง

1) ท่าทาง “ฉัน”



2) ท่าทาง “ไป”

3) ท่าทาง “โรงเรียน”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

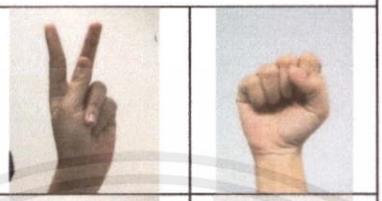
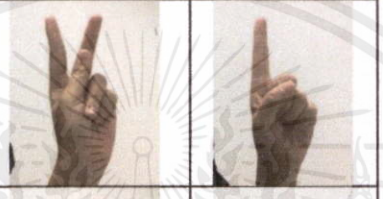
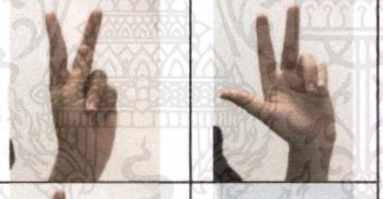
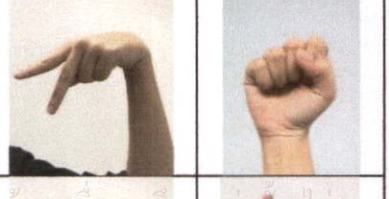
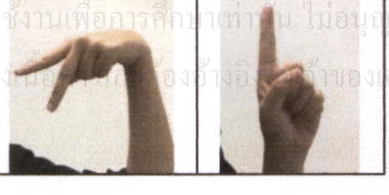
4) ท่าทาง “อ่าน”

5) ท่าทาง “หนังสือ”

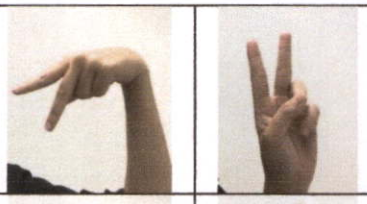
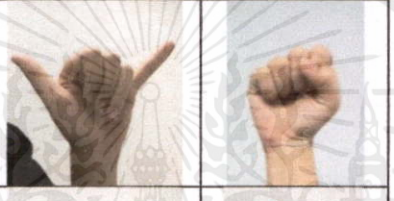
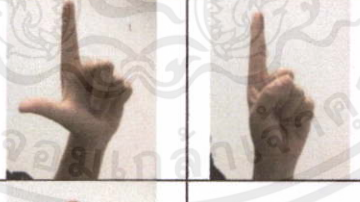
6) ท่าทาง “ห้องน้ำ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ก2 ท่าทางภาษามือของพยัญชนะต่างๆ ที่ใช้ในการทดลอง

ลำดับที่	พยัญชนะ	ท่ามือ
1	ก	
2	ข	
3	ค	
4	ฆ	
5	ด	
6	ฎ	
7	พ	
8	ป	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับที่	พยัญชนะ	ท่ามือ
9	ผ	
10	ภ	
11	ย	
12	ญ	
13	ล	
14	ฬ	
15	ว	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

คู่มือการใช้งานโปรแกรม

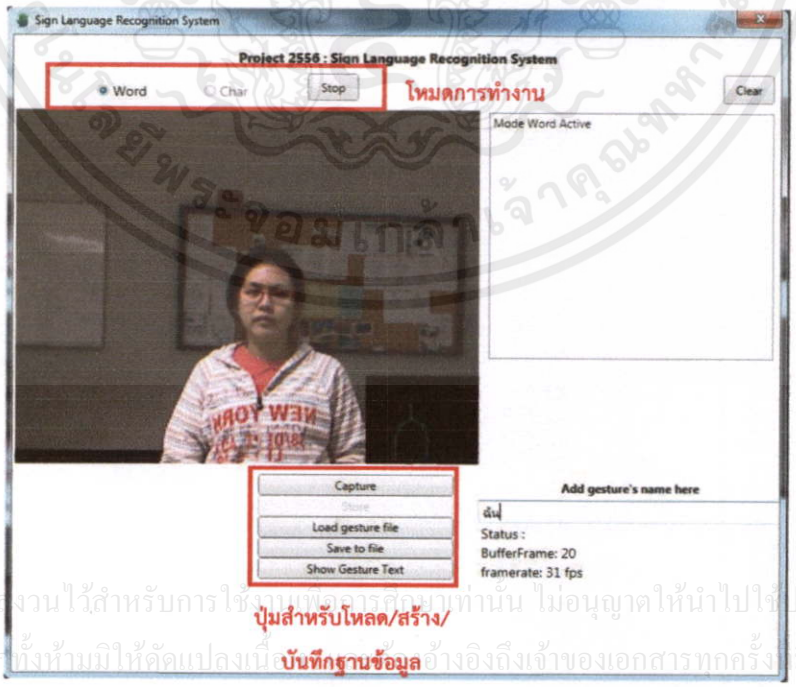
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข1 การใช้งานโปรแกรม

เริ่มจากเลือกโหมดการทำงานระหว่าง การแปลคำ (Word) หรือการแปลพยัญชนะ (Char) หลังจากนั้นจึงเริ่มแสดงท่าทาง

โหมดการแปลคำ

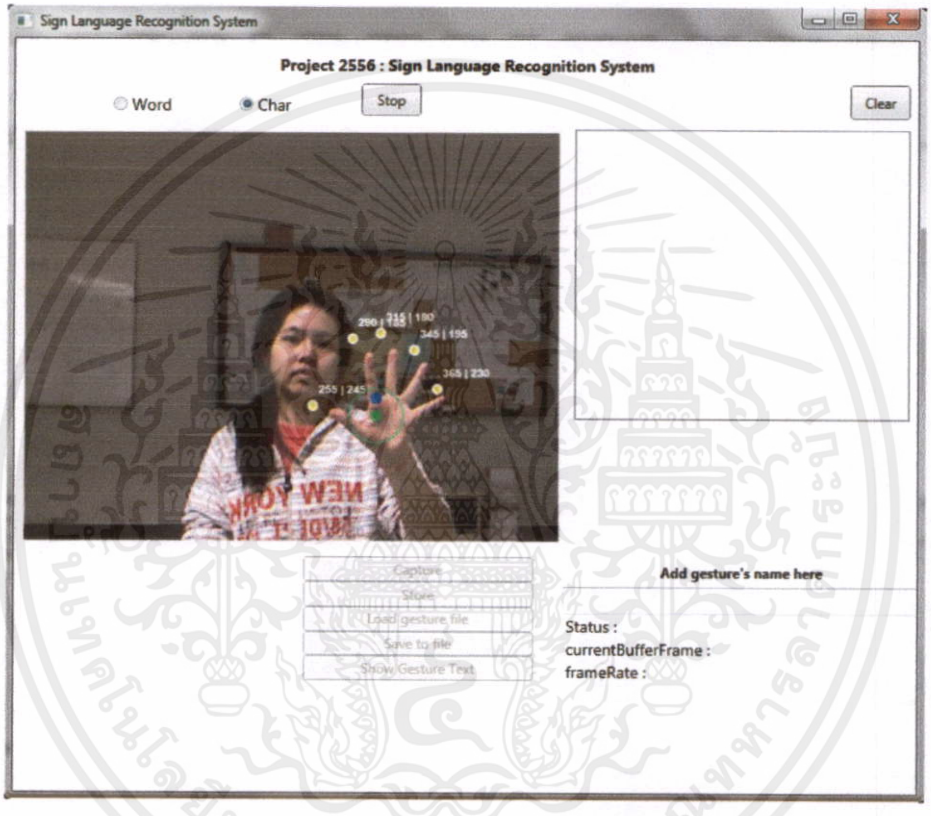
1) รอจนภาพเคลื่อนไหวโครงกระดูกขึ้น และตรวจดูให้แน่ใจว่าตรวจจับข้อต่อได้ครบทุกจุดแล้ว
2) คลิกปุ่ม Load Gesture file เพื่อโหลดฐานข้อมูลคำศัพท์เข้ามาในระบบ
3) หากไม่มีฐานข้อมูลคำศัพท์ ให้ทำการสร้างโดยการใส่ชื่อคำศัพท์ในช่องว่าง Add Gesture's name here จากนั้นกดปุ่ม Capture
4) รอเวลานับถอยหลัง 3 วินาที จากนั้นเริ่มแสดงท่าทางภาษามือของคำศัพท์ที่ระบุไว้ เมื่อระบบบันทึกเสร็จจะกลับเข้าสู่โหมดการแปลโดยอัตโนมัติ หากต้องการเพิ่มคำศัพท์อื่นๆ ให้ทำซ้ำข้อที่ 3) และ 4) อีกครั้ง
5) ผู้ใช้สามารถเรียกดูรายชื่อคำศัพท์ที่ระบบมีอยู่ในปัจจุบันได้โดยการคลิกที่ปุ่ม Show Gesture Text
6) เมื่อทำการสร้างฐานข้อมูลคำศัพท์เสร็จเรียบร้อย และต้องการบันทึกให้คลิกที่ปุ่ม Save to file
7) หากต้องการเปลี่ยนโหมดการทำงานให้คลิกที่ปุ่ม Stop จากนั้นทำการเลือกโหมด Char



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โหมดการแปลพยัญชนะ

สำหรับโหมดการแปลพยัญชนะ ผู้ใช้จะต้องอยู่ในระยะห่างจากกล้อง 0.80 ถึง 0.90 เมตร แล้วทำการยกมือข้างหนึ่งขึ้นแสดงที่หน้ากล้อง เมื่อกล้องตรวจจับมือและปลายนิ้วมือของผู้ใช้พบจะมีการแสดงผลจุดปลายนิ้วมือขึ้นที่ RGB video จากนั้นผู้ใช้จึงเริ่มแสดงท่าทางของพยัญชนะได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

ตัวอย่างโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรมที่ 1 ตัวอย่างโค้ดขบวนการปรับปรุงค่าข้อมูล

```
public static void ProcessData(Skeleton data) {
      // Extract the coordinates of the points.
      var p = new Point3D[4];
     Point3D shoulderRight = new Point3D(),shoulderLeft = new
Point3D();
      foreach (Joint j in data.Joints){
            switch (j.JointType) {
                  case JointType.HandLeft:
                        p[0] = new Point3D(j.Position.X,j.Position.Y,
                              j.Position.Z);
                        break;
                  case JointType.ElbowLeft:
                        p[1] = new Point3D(j.Position.X,j.Position.Y,
                              j.Position.Z);
                        break;
                  case JointType.ElbowRight:
                        p[2] = new Point3D(j.Position.X,j.Position.Y,
                              j.Position.Z);
                        break;
                  case JointType.HandRight:
                        p[3] = new Point3D(j.Position.X,j.Position.Y,
                              j.Position.Z);
                        break;
                  case JointType.ShoulderLeft:
                        shoulderLeft = new Point3D(j.Position.X,
                                    j.Position.Y, j.Position.Z);
                        break;
                  case JointType.ShoulderRight:
                        shoulderRight = new Point3D(j.Position.X,
                                    j.Position.Y, j.Position.Z);
                        break;
            }
      }
      // Center the data
      var center = new Point3D((shoulderLeft.X+shoulderRight.X)/2,
                  (shoulderLeft.Y + shoulderRight.Y) / 2,
                  (shoulderLeft.Z + shoulderRight.Z) / 2);
      for (int i = 0; i < 4; i++){
            p[i].X -= center.X;
            p[i].Y -= center.Y;
            p[i].Z -= center.Z;
      }
      // Normalization of the coordinates
      double shoulderDist = Math.Sqrt(
                  Math.Pow((shoulderLeft.X - shoulderRight.X),2) +
                  Math.Pow((shoulderLeft.Y - shoulderRight.Y), 2) +
                  Math.Pow((shoulderLeft.Z - shoulderRight.Z), 2));
      for (int i = 0; i < 4; i++){
            p[i].X /= shoulderDist;
            p[i].Y /= shoulderDist;
            p[i].Z /= shoulderDist;
      }
      // Launch the event!
  Skeleton3DdataCoordReady(null, new Skeleton3DdataCoordEventArgs(p));
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรมที่ 2 ตัวอย่างโค้ดการแปลพยัญชนะท่ามือเดียว

```
private String findFirstLetter(HandData handData){
      string result = "Unknow";
      if (handData.FingerCount == 3) {
            float diffY1, diffY2;
            diffY1 = handData.FingerPoints[0].Y -
                     handData.FingerPoints[1].Y;
            diffY2 = handData.FingerPoints[0].Y -
                     handData.FingerPoints[2].Y;
            //check ง
            if (Math.Abs(diffY1 - diffY2) < 15)result = "ง";
      }
      if (handData.FingerCount == 2){
            float diffY, diffX;
            diffY = handData.FingerPoints[0].Y -
                    handData.FingerPoints[1].Y;
            diffX = handData.FingerPoints[0].X -
                    handData.FingerPoints[1].X;

            //check ก
            if ( Math.Abs(diffY)<15 && (Math.Abs(diffX)<50)) {
                  result = "ก";
                  firstLetter = "ก";
            }
            //check ล
            if ((Math.Abs(diffY)>70) && (Math.Abs(diffX)<60)) {
                  result = "ล";
                  firstLetter = "ล";
            }
            //check บ
            if ((Math.Abs(diffY)<15) && (Math.Abs(diffX)>90)) {
                  result = "บ";
                  firstLetter = "บ";
            }
            //check พ
            if ((Math.Abs(diffX) < 40) &&
            (handData.FingerPoints[0].Y > handData.Location.Y) &&
            (handData.FingerPoints[1].Y > handData.Location.Y)){
                  result = "พ";
                  firstLetter = "พ";
            }
      }
      if (handData.FingerCount == 1) {
            float diffY = handData.FingerPoints[0].Y -
                          handData.Location.Y;
            //check ด
            if ((diffY < 0) && (Math.Abs(diffY) > 50)) {
                  result = "ด";
                  firstLetter = "ด";
            }
      }
      return result;
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรมที่ 3 ตัวอย่างโค้ดการแปลพยัญชนะสองท่ามือ

```
private String findSecondLetter(HandData handData){
      string result = "";
      if (handData.FingerCount == 1) {
            if (firstLetter == "ก") result = "จ";
            if (firstLetter == "พ") result = "ป";
            if (firstLetter == "ล") result = "ฟ";
            if (firstLetter == "ย") result = "ญ";
            if (firstLetter == "ด") result = "ฎ";
      }
      if (handData.FingerCount == 2) {
            if (firstLetter == "ก") result = "ค";
            if (firstLetter == "พ") result = "ผ";
      }
      if (handData.FingerCount == 3) {
            if (firstLetter == "ก") result = "ฆ";
            if (firstLetter == "พ") result = "ภ";
      }

      return result;
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้