การรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่

นางสาวสุภาวดี แสงศรี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2555

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)

are the thesis authors' files submitted through the Graduate School.

THAI FINGER-SPELLING SIGN LANGUAGE RECOGNITION FROM DATA GLOVE AND MOTION TRACKING DEVICE

Miss Supawadee Saengsri

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering Program in Computer Engineering
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2012
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูล และอุปกรณ์ติดตามการเคลื่อนที่
โดย	นางสาวสุภาวดี แสงศรี
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

....คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ
(ศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ)

....กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ธนารัตน์ ชลิดาพงศ์)

....กรรมการภายนอกมหาวิทยาลัย
(ดร.ทรงพล องค์วัฒนกุล)

สุภาวดี แสงศรี : การรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่. (THAI FINGER-SPELLING SIGN LANGUAGE RECOGNITION FROM DATA GLOVE AND MOTION TRACKING DEVICE) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ.ดร.โชติรัตน์ รัตนามหัทธนะ, 65 หน้า.

งานวิจัยทางด้านภาษามือไทยยังคงเป็นงานวิจัยที่มีความน่าสนใจเป็นอย่างยิ่ง โดยเฉพาะอย่างยิ่งเมื่อคนทั่วไปยังคงไม่เข้าใจภาษามือทำให้เป็นอุปสรรคของการติดต่อสื่อสารระหว่างคนทั่วไปกับคนหูหนวกหรือเป็นใบ้ และถึงแม้งานวิจัยที่ผ่านมานั้นมีงานวิจัยที่ใช้การประมวลผลภาพจากกล้องมาใช้ในการสร้างอัลกอริทึมสำหรับการรู้จำ แต่อย่างไรก็ตามยังมีประสิทธิภาพไม่ดีพอในการใช้งานจริง ด้วยเหตุผลอันเนื่องมาจากข้อจำกัดทางวิธีการและเทคนิคในการประมวลผลภาพที่มีอยู่ในปัจจุบัน ตัวอย่างเช่น ข้อจำกัดจากการแยกรูปร่างของมือ เมื่อมือมีการซ้อนทับกัน หรือกรณีที่ท่ามือมีการคล้ายคลึงกันอย่างมาก และเนื่องจากการสะกดนิ้วมือของภาษามือไทยนั้นท่ามือที่ใช้มีความใกล้เคียง คล้ายคลึงกันอย่างมาก บางตัวอักษรนั้นก็ใช้มือข้างเดียวทำท่ามือ และบางตัวอักษรก็ใช้ท่ามือที่เกิดจากมือสองข้าง นอกจากนี้ยังมีบางตัวอักษรใช้ท่ามือมากกว่าหนึ่งท่ามาเรียงต่อกันแทนหนึ่งตัวอักษร ในงานวิจัยนี้นำเสนอทางเลือกหนึ่งสำหรับอัลกอริทึมการรู้จำภาษามือไทยโดยใช้ถุงมือรับข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่ ซึ่งทำการรู้จำตัวอักษรทั้งพยัญชนะ สระ วรรณยุกต์ รวมทั้งสิ้น 65 ตัวอักษร ซึ่งที่ผ่านมานั้นยังไม่มีงานวิจัยใดที่ทำการวิจัยครอบคลุมการสะกดนิ้วมือของภาษามือไทยที่ครบทั้ง 65 ตัวอักษร ในงานวิจัยนี้จึงนำเสนอขั้นตอนต่าง ๆ ของอัลกอริทึมที่เหมาะกับการสะกดนิ้วมือของภาษามือไทยทุกตัวอักษร โดยมีลำดับขั้นตอนดังนี้ ขั้นแรกเป็นขั้นตอนการสกัดข้อมูล จากนั้นใช้โครงข่ายประสาทเทียมที่มีการแพร่กระจายแบบย้อนกลับของเอลแมนมาใช้ในสอนอัลกอริทึมให้มีประสิทธิภาพเพิ่มขึ้นในการรู้จำรูปร่างของมือ และขั้นตอนสุดท้ายอัลกอริทึมทำการแปลผลออกมาในรูปของตัวอักษร

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต ____________

สาขาวิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก ____________

ปีการศึกษา 2555

5470426021 : MAJOR COMPUTER ENGINEERING
KEYWORDS : THAI SIGN LANGUAGE / THAI FINGER SPELLING / THAI SIGN LANGUAGE RECOGNITION / SIGN LANGUAGE RECOGNITION / DATA GLOVE / MOTION TRACKING DEVICE

SUPAWADEE SAENGSRI : THAI FINGER-SPELLING SIGN LANGUAGE RECOGNITION FROM DATA GLOVE AND MOTION TRACKING DEVICE. ADVISOR : ASST. PROF. CHOTIRAT RATANAMAHATANA, Ph.D., 65 pp.

Thai Sign Language is considered a research priority requiring some attention since most people still do not understand sign language. It is almost impossible to communicate with people who are deaf or mute. In the past research, there exist many Thai Sign Language Recognition algorithms that use image processing, but have not performed well and still are impractical due to limitations of current image processing techniques and the difficulty of the Thai Sign language itself; many hand signs are very similar, and some signs require both hands. Moreover, some Thai alphabets require multiple sign states to represent only one alphabet. This paper presents an alternative Thai sign language recognition system using the data-gloves and motion tracker devices to recognize 65 Thai alphabets, vowels, and tone marks. We propose a data extraction method, and the Elman Back Propagation Neural Network is applied in the learning step, which greatly improves the accuracy of the system performance.Finally, Outputs are translated into alphabets.

Department : Computer Engineering Student's Signature ....
Field of Study : Computer Engineering Advisor's Signature ....
Academic Year : 2012

กิตติกรรมประกาศ

งานวิจัยนี้จะไม่เกิดขึ้นหากไม่ได้รับความอนุเคราะห์จาก โครงการ “พัฒนาระบบการสื่อสารภาษามือไทยอิเล็กทรอนิกส์” และ สำนักงานคณะกรรมการกิจการโทรคมนาคมแห่งชาติ (กสทช.) ข้าพเจ้าขอขอบพระคุณเป็นอย่างยิ่งที่ให้การสนับสนุนทุนและ อุปกรณ์ต่าง ๆ นอกจากการสนับสนุนทานด้านทุนและอุปกรณ์แล้ว สิ่งสำคัญที่ช่วยให้งานวิจัยในวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วง มิได้มาจากความสามารถของข้าพเจ้าเพียงคนเดียว หากไม่ได้รับการสนับสนุนจากผู้มีพระคุณหลาย ๆ ท่านที่กรุณาช่วยเหลือข้าพเจ้ามาโดยตลอดช่วงระยะเวลาการศึกษา แม้เต็มไปด้วยความลำบากและอุปสรรคนานัปการแต่ก็เป็นบ่อเกิดของประสบการณ์ที่มีค่ายิ่ง ที่ผลักดันให้ข้าพเจ้าสามารถพัฒนาตนเอง ทั้งด้านการคิดอย่างมีเหตุมีผล การรู้จักวิเคราะห์ปัญหา และการอดทนต่ออุปสรรคอย่าย่อท้อ

ขอขอบพระคุณอาจารย์ที่ปรึกษาวิทยานิพนธ์ฉบับนี้ ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ ที่กรุณาสละเวลาอันมีค่าช่วยสนับสนุน ชี้แนะแนวทางในการทางการวิจัย อีกทั้งอบรมและแก้ไขศิษย์คนนี้ด้วยดีเสมอมา

ขอขอบพระคุณ ดร.วิชญ์ เนียรนาทตระกูล ซึ่งเป็นรุ่นพี่ที่ห้องปฏิบัติการที่คอยช่วยเหลือสนับสนุน ชี้แนะแนวทางและให้ข้อคิดต่าง ๆ ในการทำวิจัยตลอดมา

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ฉบับนี้ ซึ่งประกอบไปด้วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล ผู้ช่วยศาสตราจารย์ ดร.ธนารัตน์ ชลิดาพงศ์ และดร.ทรงพล องค์วัฒนกุลที่ให้แนวทางแก้ไขและข้อคิดในงานวิจัยที่ยังมีจุดบกพร่องเพื่อพัฒนาแก้ไขให้ดียิ่ง ๆ ขึ้นไป

ขอขอบพระคุณรุ่นพี่ และเพื่อนทุก ๆ คนทั้งในและนอกห้องปฏิบัติการทุกคนที่คอยช่วยเหลือซึ่งกัน และกันในทุก ๆ ด้านและร่วมทุกข์ร่วมสุขด้วยกันมาตลอดระยะเวลาการศึกษา

สุดท้ายนี้ ขอขอบพระคุณบิดามารดา พี่ชาย และน้องสาวของข้าพเจ้าที่เข้าใจ เป็นกำลังใจให้ และสนับสนุนทุกช่วงเวลาที่ยากลำบากยิ่งเสมอมา

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง

บทคัดย่อภาษาอังกฤษ .... จ

กิตติกรรมประกาศ........ ฉ

สารบัญ .... ช

สารบัญตาราง .... ญ

สารบัญรูป .... ฎ

บทที่ 1 บทนำ........ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์........ 5

1.3 ขอบเขตของการวิจัย........ 5

1.4 ข้อจำกัดของการวิจัย........ 5

1.5 ประโยชน์ที่จะได้รับ........ 5

1.6 วิธีการดำเนินการวิจัย 5

1.7 การเผยแพร่ผลงานวิจัย 6

บทที่ 2 งานวิจัยที่เกี่ยวข้อง 7

2.1 ทฤษฎีที่เกี่ยวข้อง 7

2.1.1 อักษรไทย 7

2.1.2 ภาษามือไทย 8

2.1.3 อุปกรณ์รับข้อมูล 9

2.1.4 ข้อมูลอนุกรมเวลา (Time-Series) 11

2.1.5 โครงข่ายประสาทเทียม (Neural Network)........ 12

2.1.6 การหาลำดับย่อยร่วมที่ยาวสุด (Longest Common Subsequence) [12] 15

2.1.7 การวัดหาค่าความต่างกันของสายอักขระสองชุด (Minimum Edit Distance) [13]........ 16

2.2 งานวิจัยที่เกี่ยวข้อง........ 17

2.2.1 การรู้จำภาษามือโดยใช้การประมวลผลภาพจากกล้อง 17

หน้า

2.2.2 การรู้จำภาษามือโดยใช้การประมวลผลสัญญาณจากถุงมือและอุปกรณ์ที่ใช้ติดตามการเคลื่อนที่ของมือ 20
บทที่ 3 การรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทย จากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่ 25
3.1 ภาพรวมของขั้นตอนการรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่ 25
3.2 การสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ (Data Segmentation) 26
3.2.1 การตรวจสอบตำแหน่งของมือว่าอยู่ในขอบเขตของการทำท่ามือ 27
3.2.2 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่ (Segmentation using data from motion tracking sensors) 28
3.2.3 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยถุงมือรับข้อมูล (Segmentation by using data from the data glove) 30
3.3 การรู้จำท่ามือ (Sign Recognition) 32
3.3.1 โครงข่ายประสาทเทียมสำหรับท่ามือเดียว 33
3.3.1 โครงข่ายประสาทเทียมสำหรับท่ามือสองมือ 33
3.4 การแปลท่ามือ (Translation) 35
3.4.1 การหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence) 35
3.4.2 การวัดหาค่าความแตกต่างกันของสายอักขระสองชุด (Minimum Edit Distance) 36
บทที่ 4 การทดลอง และวิเคราะห์ผลการทดลอง 39
4.1 การเตรียมข้อมูลทดลอง 39
4.2 วิธีวัดผลและประเมินผลการทดลอง 39
4.2.1 การตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ทำท่ามือ 39
4.2.2 การตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ 41
4.2.3 การตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ 42
4.3 ผลการทดลองและวิเคราะห์ผลการทดลอง 43
4.3.1 ผลการตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ 43
4.3.2 ผลการตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ 44

หน้า

4.3.3 ผลการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ 47

4.3.4 การเปรียบเทียบระหว่างอัลกอริทึมที่นำเสนอในงานวิทยานิพนธ์ และ Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform [15].. 50

4.3.5 ผลการทดลองเพิ่มเติมเพื่อทำการรู้จำการสะกดนิ้วมือในระดับคำเบื้องต้น .. 50

4.4 สรุปผลการทดลอง 54

บทที่ 5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ 55

5.1 สรุปและอภิปรายผลการวิจัย 55

5.2 ข้อจำกัดและข้อเสนอแนะ 56

รายการอ้างอิง 57

ภาคผนวก 59

ภาคผนวก ก การติดตั้ง และการใช้งานถุงมือรับข้อมูลรุ่น 5DT 14 Data Glove Ultra โดยใช้ Microsoft Visual Studio C++ 60

ภาคผนวก ข การติดตั้ง และการใช้งานอุปกรณ์ติดตามการเคลื่อนที่รุ่น 3D Guidance trakSTAR™โดยใช้ Microsoft Visual Studio C++ 63

ประวัติผู้เขียนวิทยานิพนธ์ 65

สารบัญตาราง

หน้า

ตารางที่ 2.1 การคำนวณหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence) 15
ตารางที่ 2.2 ผลการ Trace back เพื่อหาสายอักระย่อยที่ยาวที่สุด 16
ตารางที่ 2.3 ตารางการหาค่าความแตกต่างของ Minimum Edit Distance 17
ตารางที่ 3.1 ตารางแสดงไอดี และท่ามือที่ใช้ในการทำการวิจัย [8] 34
ตารางที่ 3.2 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้
Longest Common Subsequence 35
ตารางที่ 3.3 โค้ดตัวอย่างการเขียนฟังก์ชั่น Longest Common Subsequence 36
ตารางที่ 3.4 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้ Minimum Edit Distance 36
ตารางที่ 3.5 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้ Minimum Edit Distance 37
ตารางที่ 4.1 ผลการทดลองของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ 43
ตารางที่ 4.2 ผลการวัดค่า Accuracy ของโมเดลโครงข่ายประสาทเทียมสำหรับการรู้จำ
ท่ามือเดียว และท่ามือสองมือ 44
ตารางที่ 4.3 ตัวอย่างผลการทำนายที่ผิดพลาดของโมเดลที่ 1 45
ตารางที่ 4.4 ผลการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือเป็นตัวอักษร 48
ตารางที่ 4.5 ตัวอย่างผลการแปลตัวอักษร ‘ล’...49
ตารางที่ 4.6 ผลการเปรียบเทียบ Accuracy ของอัลกอริทึมที่นำเสนอกับงานวิจัย [15] 50
ตารางที่ 4.7 ผลการทดลองเพิ่มเติมเพื่อตรวจสอบความถูกต้องของการแปลท่ามือในระดับคำ...51
ตารางที่ 4.8 ตัวอย่างผลการแปลคำของ Longest common subsequences 52
ตารางที่ 4.9 ตัวอย่างผลการแปลคำของ Minimun Edit Distance รูปแบบที่ 1 53
ตารางที่ 4.10 ตัวอย่างผลการแปลคำของ Minimun Edit Distance รูปแบบที่ 2 53
ตารางที่ ก.1 โค้ดตัวอย่างการเชื่อมต่อถุงมือ 62

สารบัญรูป

หน้า

รูปที่ 1.1 เปรียบเทียบการสะกดนิ้วมือของประเทศที่ใช้ภาษาอังกฤษ 2
รูปที่ 1.2 แบบสะกดนิ้วมือไทย ประดิษฐ์โดยคุณหญิงกมลา ไกรฤกษ์ [8] 3
รูปที่ 1.3 การสะกดนิ้วมือของตัวอักษร ข ต้องใช้ 2 ท่ามือเพื่อประกอบกันแสดงถึง ตัวอักษร ข โดยประกอบจากท่ามือ ก ตามด้วยท่ามือเลข 1 3
รูปที่ 1.4 รูปร่างของท่ามือซึ่งมีการซ้อนทับกันของนิ้วมือ และการซ้อนทับกันของมือกับมือ 4
รูปที่ 1.5 ตัวอย่างค่าที่ได้จากตัวรับรู้แต่ละตัวเมื่อทำท่ามือ “ก” 4
รูปที่ 2.1 แบบสะกดนิ้วมือไทย และ ตัวอย่างการสะกดนิ้วมือ [8] 8
รูปที่ 2.2 อุปกรณ์รับข้อมูลที่ใช้ในงานวิจัย 9
รูปที่ 2.3 ถุงมือรับข้อมูล 5DT 14 Data Glove Ultra [9] 9
รูปที่ 2.4 รายละเอียดตัวรับรู้ที่ติดบนถุงมือแต่ละข้าง [9] 10
รูปที่ 2.5 อุปกรณ์ติดตามการเคลื่อนที่ 11
รูปที่ 2.6 การวัดค่าเทียบจากแนวแกน x y z ของตัวรับรู้ และตัวรับสัญญาณ [10] 11
รูปที่ 2.7 เซลล์ประสาท(Natural Neural) [11] 12
รูปที่ 2.8 เพอร์เซปตรอน (Perceptron) [11] 13
รูปที่ 2.9 ตัวอย่างการเรียนรู้ฟังก์ชัน AND และ XOR ของเพอร์เซปตรอนตัวเดียว [11] 14
รูปที่ 2.10 การแก้ปัญหา XOR ด้วยข่ายงานประสาทหลายชั้น [11] 14
รูปที่ 2.11 ระบบการรู้จำของงานวิจัย TSTMT [14] 19
รูปที่ 2.12 ท่ามือการสะกดการรู้จำที่ใช้ในงานวิจัย [15] 20
รูปที่ 2.13 ท่ามือที่ใช้ทำการรู้จำในงานวิจัย [2, 3] 21
รูปที่ 2.14 ขั้นตอนการรู้จำของท่ามือในงานวิจัย [2, 3] 22
รูปที่ 2.15 ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่ไม่มีการเคลื่อนที่ [2] 22
รูปที่ 2.16 ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่มีการเคลื่อนที่ [2] 23
รูปที่ 2.17 ภาพรวมแสดงขั้นตอนการดึงข้อมูลและขั้นตอนการรู้จำ [1] 24
รูปที่ 3.1 ภาพรวมการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ การติดตามการเคลื่อนที่ 25
รูปที่ 3.2 ช่วงเวลา หรือสถานะต่าง ๆ ของผู้ทำท่ามือขณะเริ่มทำท่ามือ ฎ จนสิ้นสุด 26
รูปที่ 3.3 ตำแหน่งการวางตัวส่งสัญญาณ (Transmitter) 28

หน้า

รูปที่ 3.4 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่..............................29

รูปที่ 3.5 การสกัดหาจุดต่ำสุดสัมพัทธ์ด้วยถุงมือรับข้อมูล..............................30

รูปที่ 3.6 การสกัดหาเฟรมสำคัญ หรือคีย์เฟรม..............................31

รูปที่ 3.7 รูปแบบทั้งหมดของท่ามือที่ต้องใช้สำหรับการสะกดนิ้วมือทั้ง 65 ตัวอักษร..............................32

รูปที่ 3.8 โครงข่ายประสาทเทียม..............................32

รูปที่ 4.1 ขั้นตอนการฝึกสอน (Train) และทดสอบ (Test) โครงข่ายประสาทเทียม..............................41

รูปที่ 4.2 การสุ่มเลือกข้อมูลแบบ Leave-one-out Cross Validation..............................42

รูปที่ 4.3 Confusion metrix of Model 1..............................45

รูปที่ 4.4 Confusion metrix of Model 2..............................46

รูปที่ 4.5 Confusion metrix of Model 3..............................46

รูปที่ ก.1 เลือกเข้าเมนู Property..............................60

รูปที่ ก.2 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.h ได้..............................60

รูปที่ ก.3 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.dll ได้..............................61

รูปที่ ก.4 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.lib ได้..............................61

รูปที่ ข.1 การเรียกหน้าต่าง Properties ของ Project..............................63

รูปที่ ข.2 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.h ได้..............................63

รูปที่ ข.3 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.dll ได้..............................64

รูปที่ ข.4 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.lib ได้..............................64

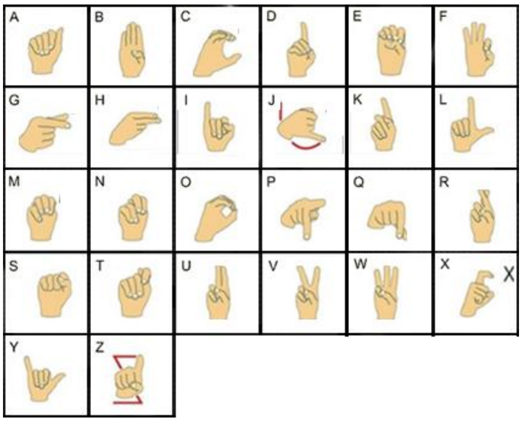
บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

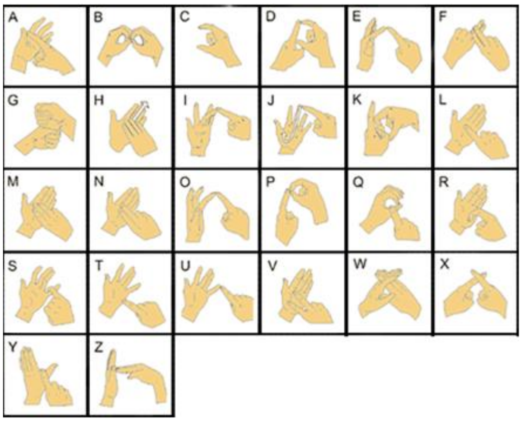
การสื่อสารมีความสำคัญอย่างยิ่งสำหรับมนุษย์ที่มีการอยู่ร่วมกันเป็นสังคม โดยรูปแบบการสื่อสารที่มีความสะดวกที่สุดคือการสื่อสารที่ใช้เสียงจากการพูดเป็นสื่อกลาง เพราะสามารถใช้ได้ทุกขณะ ทั้งในขณะเดินหรือนั่ง อีกทั้งไม่จำเป็นต้องมีอุปกรณ์ช่วยใด ๆ และยังสามารถสื่ออารมณ์ของผู้พูดได้ดี แต่สำหรับผู้ที่มีความบกพร่องทางการได้ยินหรือการพูดเช่น คนหูหนวก หรือเป็นใบ้นั้นไม่สามารถที่จะใช้เสียงจากการพูดในการสื่อสารเหมือนคนทั่วไปทำให้ช่องทางในการสื่อสารมีลดน้อยลง ภาษามือจึงถูกคิดค้นให้เข้ามีบทบาทสำคัญเพื่อทำหน้าที่เป็นสื่อกลางในการสื่อสารภายในกลุ่มผู้ที่มีความบกพร่องทางการได้ยินหรือการพูด เพราะภาษามือสำหรับผู้มีความบกพร่องทางการได้ยินหรือการพูดแล้ว ก็เปรียบได้กับการสื่อสารด้วยการพูดของคนทั่วไป ถึงแม้ภาษามือจะเข้ามามีบทบาทสำคัญในการแก้ปัญหาด้านการสื่อสารในกลุ่มของผู้ที่มีความพิการทางการได้ยินหรือการพูดได้ดี แต่อย่างไรก็ตามก็ยังไม่เพียงพอที่จะใช้เป็นสื่อกลางในการสื่อสารกับคนทั่วไปได้ เพราะในกลุ่มคนทั่วไปนั้นมีคนเป็นจำนวนมากที่ไม่เข้าใจภาษามือ และเป็นไปไม่ได้ที่จะสอนให้ทุกคนเข้าใจภาษามือกันได้ และในการใช้ชีวิตประจำวันของผู้มีความบกพร่องทางการได้ยินหรือการพูดก็ไม่มีทางหลีกเลี่ยงที่จะไม่สื่อสารกับคนทั่วไปเหล่านี้ได้ เนื่องจากความต้องการทางด้านปัจจัยต่าง ๆ ในชีวิตจำเป็นต้องมีการพึ่งพากัน เช่น ในการทำงาน การซื้อของ การรักษา พยาบาล เป็นต้น ซึ่งการแก้ปัญหาด้านการสื่อสารที่ดีที่สุดในขณะนี้คือการใช้ล่ามผู้ที่สามารถแปลภาษามือออกมาเป็นภาษาพูด และแปลภาษาพูดออกเป็นภาษามือ แต่การจ้างล่ามมานั้นก็ย่อมมีค่าใช้จ่ายที่สูง และในบางกรณีการสื่อสารอาจจะต้องการความเป็นส่วนตัวการใช้ล่ามจึงอาจจะเป็นทางออกที่ไม่เหมาะสม ดังนั้น ในวิทยานิพนธ์นี้มีจุดประสงค์เพื่อสร้างอัลกอริทึมที่สามารถทำการรู้จำการสะกดนิ้วมือของภาษามือไทย เพื่อเพิ่มช่องทางในการสื่อสารของผู้ที่มีความบกพร่องทางการได้ยินหรือการพูดโดยการใช้คอมพิวเตอร์มาเป็นสื่อกลางเพื่อเป็นทางเลือกที่ใช้ในการสื่อสารระหว่างผู้ที่มีความบกพร่องทางการได้ยินหรือการพูดกับบุคคลทั่วไป

ภาษามือทั่วโลกมีเป็นจำนวนมากโดยภาษามือแต่ละภาษาก็มีความแตกต่างกันตามชาติพันธุ์ และพื้นที่ของแหล่งกำเนิด ซึ่งภาษามือนั้นแตกต่างจากภาษาพูดที่ไม่มีภาษากลางที่สามารถสื่อสารกันได้ทั่วโลกเช่น ภาษาพูดมีภาษาอังกฤษเป็นภาษาสากลแต่ภาษามือนั้นไม่มีเพราะแม้แต่ในกลุ่มประเทศที่ใช้ภาษาอังกฤษเหมือนกันเช่น ประเทศสหรัฐอเมริกา และประเทศอังกฤษ ยังมีรูปแบบของการสะกดนิ้วมือที่แตกต่างกัน ตัวอย่างเช่น การสะกดนิ้วมือภาษามืออเมริกัน

(American Sign Language) และการสะกดนิ้วมือภาษามือบริติช (British Sign Language) ดังรูปที่ 1.1 โดยการสะกดนิ้วมือของภาษามืออเมริกันใช้เพียงมือข้างเดียวในการทำท่ามือ ในขณะที่การสะกดนิ้วมือของภาษามือบริติชส่วนมากใช้สองมือในการทำท่ามือ มีเพียงตัวอักษร C ที่ใช้ท่ามือเดียวกัน



(ก)



(ข)

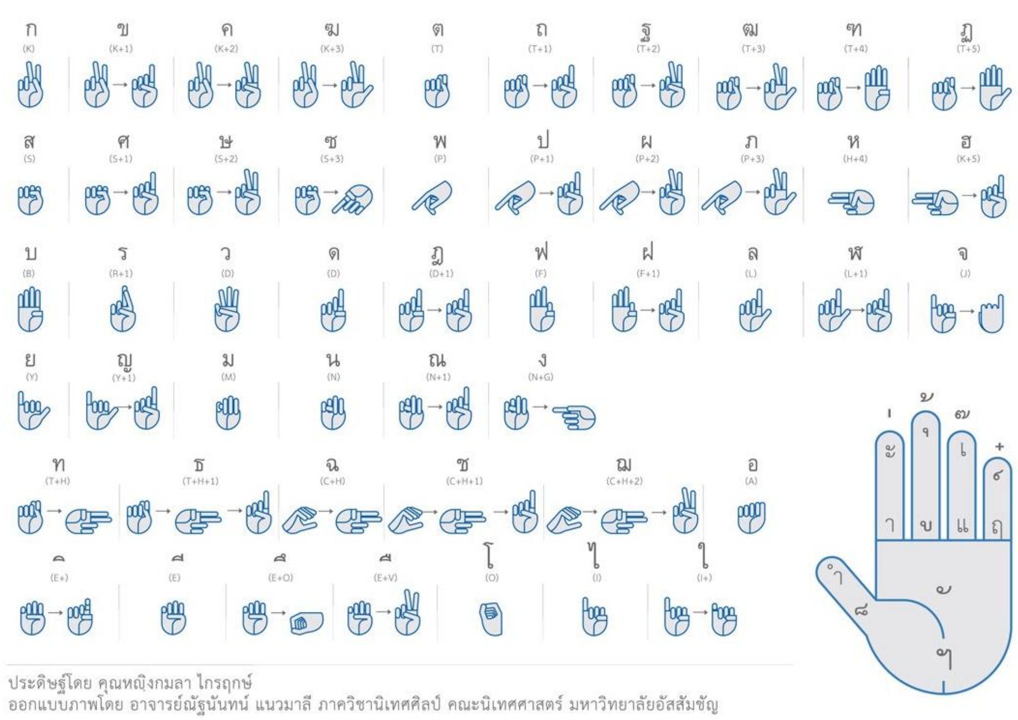
รูปที่ 1.1 เปรียบเทียบการสะกดนิ้วมือของประเทศที่ใช้ภาษาอังกฤษ

(ก) ภาษามืออเมริกัน (American Sign Language)

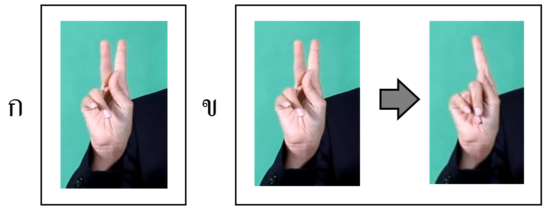
(ข) ภาษามือบริติซ (British Sign Language)

ทำให้งานวิจัยทั่วโลกที่มีการวิจัยเกี่ยวกับภาษามือนั้นมีความแตกต่างกันตามไปด้วย เช่น ภาษามืออเมริกัน [1, 2, 3] ภาษาออสเตรเลีย[4] ภาษามือบริติส[5] ภาษามือปากีสถาน[6] และภาษามือญี่ปุ่น[7]

วิทยานิพนธ์นี้ให้ความสนใจกับการรู้จำการสะกดนิ้วมือของภาษามือไทยที่มีการพัฒนามาจากการสะกดนิ้วมือของภาษามืออเมริกัน ซึ่งการสะกดนิ้วมือของภาษามือไทยรูปที่ 1.2 ถูกประดิษฐ์โดยคุณหญิงกมลา ไกรฤกษ์ เนื่องจากภาษาไทยมีพยัญชนะ สระ วรรณยุกต์ ที่มีจำนวนมากกว่าภาษามืออเมริกันมาก จึงจำเป็นต้องการคิดค้นเพิ่มโดยสร้างท่ามือขึ้นมาเพิ่มเติมซึ่งทำให้หลายท่ามือนั้นมีความคล้ายคลึงกันมาก และบางตัวอักษรใช้ท่ามือมากกว่าหนึ่งท่ามือประกอบกันเพื่อแสดงแทนตัวอักษร เช่น ตัวอักษร ข ดังรูปที่ 1.3 ใช้ท่ามือหลายท่าเพื่อประกอบเป็น 1 ตัวอักษร ซึ่งทำให้การรู้จำของการสะกดนิ้วมือไทยมีความยากกว่าการรู้จำการสะกดนิ้วมืออเมริกัน หรือบริติช



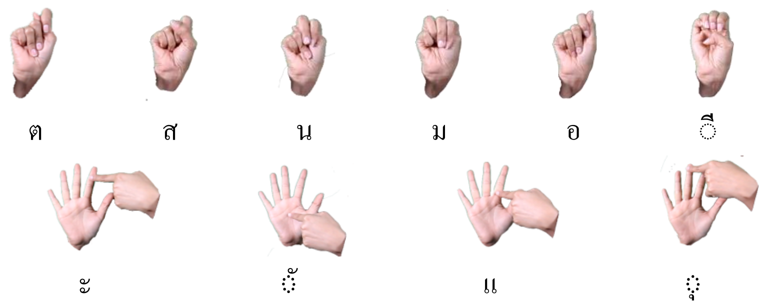
รูปที่ 1.2 แบบสะกดนิ้วมือไทย ประดิษฐ์โดยคุณหญิงกมลา ไกรฤกษ์ [8]



รูปที่ 1.3 การสะกดนิ้วมือของตัวอักษร ข ต้องใช้ 2 ท่ามือเพื่อประกอบกันแสดงถึงตัวอักษร ข โดยประกอบจากท่ามือ ก ตามด้วยท่ามือเลข 1

ในวิทยานิพนธ์นี้ได้เลือกใช้อุปกรณ์พิเศษในการรู้จำได้แก่ ถุงมือ (Data Glove) และอุปกรณ์ติดตามการเคลื่อนที่ (Motion Tracking Sensor) เพื่อเป็นตัวรับข้อมูลจากผู้ทำท่าภาษามือ เพราะสามารถให้ข้อมูลที่มีความแม่นยำมากกว่ากล้องเว็บแคมทั่วไปเนื่องจากเป็นการรับค่าจากตัวรับรู้ (Sensor) ที่ติดอยู่บนข้อต่อต่าง ๆ ของนิ้วมือโดยตรง ไม่ต้องขึ้นกับสภาวะแวดล้อมใด เช่น คุณภาพแสง ความคมชัดของภาพ นอกจากนั้นอุปกรณ์นี้สามารถตรวจพบการซ้อนทับกันของนิ้วมือ หรือมือ กับมือ ดังรูปที่ 1.4ได้ดีกว่าการใช้กล้อง และความเร็วในการประมวลผลก็เร็วกว่าเพราะสามารถนำข้อมูลที่ได้มาใช้ได้โดยตรงไม่ต้องผ่านขั้นตอนทางการประมวลผลทางภาพ เพื่อสกัด

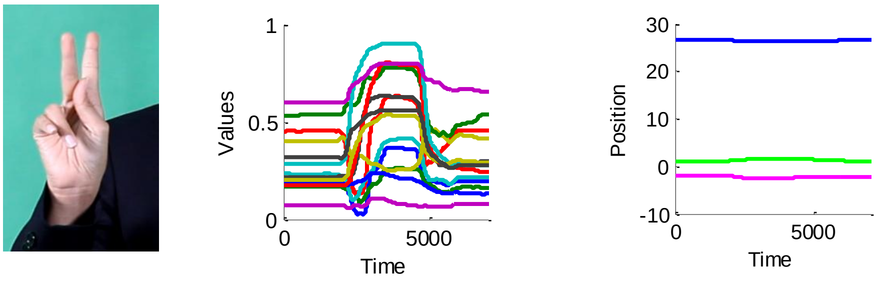
คุณสมบัติของภาพ ดังนั้นระบบการรู้จำที่ได้รับข้อมูลจากอุปกรณ์พิเศษจึงมีแนวโน้มที่ให้ประสิทธิผลที่สูงกว่า



รูปที่ 1.4 รูปร่างของท่ามือซึ่งมีการซ้อนทับกันของนิ้วมือ และการซ้อนทับกันของมือกับมือ

ซึ่งความยากของงานวิทยานิพนธ์นี้คือการออกแบบขั้นตอนวิธีการนำข้อมูลจากตัวรับรู้ (Sensor) จำนวนมากดังรูปที่ 1.5 ซึ่งให้ข้อมูลที่เป็นข้อมูลอนุกรมเวลา (Time-Series) หลายช่องสัญญาณ มามาเข้าสู่ขั้นตอนการรู้จำจึงถือว่าเป็นการยากที่จะจำแนกช่วงเวลาที่มีการทำท่ามือซึ่งสื่อความหมายออกจากช่วงเวลาอื่น ๆ ที่เป็นท่ามือไม่สื่อความหมาย เพราะการทำท่ามือต้องประกอบจากรูปแบบการงอของข้อต่อบนนิ้วมือทุกนิ้ว ไม่ใช่เพียงนิ้วใดนิ้วหนึ่ง ทำให้ต้องมีการวิเคราะห์จากหลาย ๆ ตัวรับรู้ที่ติดบนแต่ละนิ้วพร้อม ๆ กันจึงจะสามารถบอกถึงรูปร่างของท่ามือได้ อีกทั้งในงานวิจัยนี้ต้องการทำการรู้จำการสะกดนิ้วมือของพยัญชนะ สระ วรรณยุกต์ ทั้งหมดซึ่งมีจำนวนคลาสมากกว่างานวิจัยที่การรู้จำการสะกดนิ้วมือภาษามือไทยที่มีก่อนหน้า

ท่ามือ “ก”



(ก) (ข)

รูปที่ 1.5 ตัวอย่างค่าที่ได้จากตัวรับรู้แต่ละตัวเมื่อทำท่ามือ “ก”

(ก) ค่าที่ได้จากตัวรับรู้บนถุงมือข้างขวา จำนวน 14 ตัวรับรู้

(ข) ค่าที่ได้จากอุปกรณ์ติดตามการเคลื่อนที่ ตำแหน่งของมือในแกน x y z

1.2 วัตถุประสงค์

พัฒนาอัลกอริทึมที่ใช้ในการรู้จำการสะกดนิ้วมือไทยครอบคลุม พยัญชนะ สระ วรรณยุกต์ และสัญลักษณ์รวมทั้งสิ้นจำนวน 65 ตัวอักษร

1.3 ขอบเขตของการวิจัย

สามารถพัฒนาอัลกอริทึมที่สามารถทำการรู้จำคำโดยการสะกดนิ้วมือไทยจำนวน 65 ตัวอักษรโดยใช้อุปกรณ์เฉพาะทางได้แก่ถุงมือรับข้อมูล และอุปกรณ์รับข้อมูล

1.4 ข้อจำกัดของการวิจัย

อุปกรณ์ที่ใช้คือถุงมือรับข้อมูล 5DT 14 Ultra Data Glove และอุปกรณ์ติดตามการเคลื่อนที่ 3D Guidance Trakstar™

1.5 ประโยชน์ที่จะได้รับ

ได้แบบจำลองที่ใช้ในการรู้จำการสะกดภาษามือไทยโดยแปลออกมาเป็นตัวอักษร

1.6 วิธีการดำเนินการวิจัย

1. ศึกษางานวิจัยก่อนหน้า(Related Work)เพื่อหาแนวทางในการทำงานวิจัย
2. เก็บข้อมูล (Data Collection) และทาการเตรียมข้อมูล (Data Preprocessing)
3. พัฒนาวิธีการในการหาช่วงเวลาที่มีการทำท่าการสะกดนิ้วมือ (Segmentation)
4. สร้างแบบจำลอง
5. ประเมินผล และ ปรับแบบจำลอง
6. เรียบเรียงวิทยานิพนธ์

1.7 การเผยแพร่ผลงานวิจัย

ส่วนหนึ่งของงานวิทยานิพนธ์นี้ ได้รับการตีพิมพ์เป็นบทความวิชาการจำนวน 1 เรื่องคือ

TFRS: Thai Sign Language Recognition System. โดย นางสาวสุภาวดี แสงศรี ดร.วิชญ์ เนียรนาทตระกูล และ ผู้ช่วยศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัธนะ ในงานประชุมวิชาการ "DICTAP 2012 : The Second International Conference on Digital Information and Communication Technology and its Applications " ซึ่งจัดขึ้น ณ กรุงเทพฯ ประเทศไทย ระหว่างวันที่ 16 ถึง 18 พฤษภาคม 2555

บทที่ 2
งานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 อักษรไทย

ระบบอักษรไทยประกอบด้วยเสียงและรูปอักษรไทย 3 ชนิด คือ สระ พยัญชนะ วรรณยุกต์ และเครื่องหมายต่าง ๆ ที่จำเป็นในการเขียน อ่าน สะกดคำในภาษาไทย ซึ่งภาษาเขียนของภาษาไทยเป็นการใช้สัญลักษณ์หรือตัวอักษรแทนเสียงในภาษาซึ่งประกอบด้วยเสียงสระ พยัญชนะ และวรรณยุกต์ ดังต่อไปนี้

พยัญชนะ

พยัญชนะในภาษาไทยแต่เดิมนั้นมีจำนวนทั้งหมด 44 ตัว แต่มีพยัญชนะที่ไม่ได้ใช้ตั้งแต่ปี พ.ศ. 2445 อยู่ 2 ตัวได้แก่ ฃ ฅ ทำให้ในปัจจุบันเหลือพยัญชนะที่ใช้อยู่จำนวน 42 ตัว ได้แก่

ก	ข	ค	ฆ	ง	จ	ฉ	ช	ซ	ฌ
ญ	ฎ	ฏ	ฐ	ฑ	ฒ	ณ	ด	ต	ถ
ท	ธ	น	บ	ป	ผ	ฝ	พ	ฟ	ภ
ม	ย	ร	ล	ว	ศ	ษ	ส	ห	ฬ
อ	ฮ								

สระ

สระในภาษาไทยถ้านับตามรูปสระที่ใช้แทนเสียงจำนวน 17 รูป ดังนี้

1. ะ	2. ◌ั	3. า	4. ◌ำ	5. ◌ิ
6. ◌ี	7. ◌ึ	8. ◌ื	9. ◌ุ	10. ◌ู
11. ◌็	12. เ	13. แ	14. โ	15. ใ
16. ไ	17. ฤ			

วรรณยุกต์ และ สัญลักษณ์

วรรณยุกต์ในภาษาไทยนั้นใช้สำหรับการผันเสียงของคำเช่น ปา ป่า ป้า ป๊า ป๋า เป็นต้น ซึ่งวรรณยุกต์ในภาษาไทยมี 4 รูป 5เสียงได้แก่ เสียงสามัญ (ไม่มีรูป) เสียงเอก (◌่) เสียงโท (◌้) เสียงตรี (◌๊) และเสียงจัตวา (◌๋) และ สัญลักษณ์อื่น ๆ ซึ่งมีการสร้างไว้โดยคุณหญิงกมลา ไกรฤกษ์ ได้แก่ ◌์ และ ฯ

2.1.2 ภาษามือไทย

ภาษามือเป็นภาษาสัญลักษณ์ซึ่งถือได้ว่าภาษามือเป็นอวัจนภาษาอย่างหนึ่ง ซึ่งประกอบไปด้วย การสื่อสารด้วยมือ, การสื่อสารด้วยร่างกาย และการใช้ริมฝีปากในการสื่อความหมายแทนการใช้เสียงพูด การสื่อสารจะใช้ลักษณะของมือที่ทำเป็นสัญลักษณ์ การเคลื่อนไหวของมือ แขนร่างกาย และการแสดงความรู้สึกทางใบหน้าเพื่อช่วยในการสื่อสารความคิดเห็นของผู้สื่อสาร ซึ่งภาษามือนั้นไม่ได้มีการใช้อยู่แต่เพียงกลุ่มของผู้มีความบกพร่องทางการได้ยินหรือการพูดเท่านั้น แม้แต่ในทางทหารเองก็มีการนำภาษามือไปใช้

ภาษามือที่งานวิจัยนี้ให้ความสนใจคือการสะกดนิ้วมือไทยรูปที่ 2.1 ซึ่งการสะกดนิ้วมือไทยในภาษามือไทย คือการใช้นิ้วมือทำท่าภาษามือแทนตัวอักษร สระ วรรณยุกต์ในภาษาไทย รวมทั้งตัวเลขเพื่อสะกดเป็นคำที่มีความหมายต่าง ๆ ผู้บกพร่องทางการได้ยินหรือการพูดใช้การสะกดนิ้วมือ เมื่อกล่าวถึงบุคคล ชื่อเฉพาะ ชื่อย่อ และสถานที่ที่ไม่คุ้นเคย เพื่อสื่อสารความเข้าใจที่ถูกต้องชัดเจนและเพื่อการศึกษาหาความรู้ การสะกดนิ้วมือไทยมีเป็นครั้งแรกเมื่อคุณหญิงกมลา ไกรฤกษ์ ซึ่งจบการศึกษา อ.บ. ป.ม.และ M.A. การสอนคนหูหนวกจากมหาวิทยาลัยกาเลาเด็ท (Gallaudet) กรุงวอชิงตัน สหรัฐอเมริกา และเป็นอดีตอาจารย์ใหญ่โรงเรียนสอนคนหูหนวก (ปัจจุบันคือ โรงเรียนเศรษฐเสถียร) ประดิษฐ์แบบสะกดนิ้วมือไทยขึ้นโดยดัดแปลงจากการสะกดนิ้วมืออเมริกันรูปที่ 1.1 (ก) ในปี พ.ศ. 2499 โดยผู้วาดภาพนิ้วมือไทยคราวนั้นคือ นายกุลน้อย ทองน้อย ต่อ กรมวิชาการ กระทรวงศึกษาธิการประกาศรับรองการสะกดนิ้วมือไทยในปี พ.ศ. 2542 ทำให้มีการใช้แบบสะกดนิ้วมือไทยดังกล่าวอย่างแพร่หลายและเป็นส่วนหนึ่งของภาษามือไทยจนถึงปัจจุบัน

รูปที่ 2.1 แบบสะกดนิ้วมือไทย และ ตัวอย่างการสะกดนิ้วมือ [8]

2.1.3 อุปกรณ์รับข้อมูล

ในวิทยานิพนธ์นี้ได้เลือกใช้อุปกรณ์รับข้อมูลจากการทำท่าภาษามือไทยคือ ถุงมือรับข้อมูล และอุปกรณ์ติดตามการเคลื่อนที่ เนื่องจากถุงมือสามารถตรวจจับการซ้อนทับกันของนิ้วมือ กับนิ้วมือและมือกับมือได้ดีกว่าการรับค่าจากกล้อง แต่การใช้ถุงมือเพียงอย่างเดียวไม่สามารถระบุทิศทางการเคลื่อนที่และตำแหน่งของมือได้ ดังนั้นจึงมีการใช้อุปกรณ์ติดตามการเคลื่อนที่ร่วมกับถุงมือ ดังรูปที่ 2.2

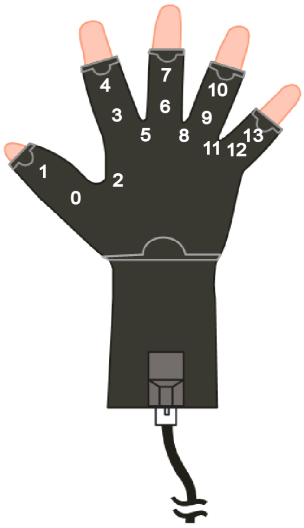
รูปที่ 2.2 อุปกรณ์รับข้อมูลที่ใช้ในงานวิจัย

โดยอุปกรณ์รับข้อมูลที่ใช้ในงานวิจัยได้แก่ ถุงมือรับข้อมูล และอุปกรณ์ติดตามการเคลื่อนที่ใช้ในงานวิจัยนี้มีรายละเอียดดังนี้

ถุงมือรับข้อมูล

ถุงมือรับข้อมูลที่ใช้ในงานวิจัยนี้คือ 5DT 14 Data Glove Ultra ดังรูปที่ 2.3 มีหน้าที่ในการตรวจจับการงอของนิ้ว และการกางของนิ้วมือโดยการให้ค่าที่อยู่ในช่วง 0 – 4095 โดยถุงมือแต่ละข้างมีตัวรับรู้ (Sensor) ทั้งหมด 14 ตัว โดยจำนวนตัวรับรู้บนแต่ละนิ้วจะมีตัวรับรู้ (Sensor) ที่ใช้วัดการงอของนิ้วอยู่อย่างละ 2 ตัว (ตำแหน่งปลายนิ้ว และตำแหน่งโคนนิ้ว) และมีตัวรับรู้ระหว่างนิ้วซึ่งใช้วัดการกางออกของนิ้วโดยมีรายละเอียดตามรูปที่ 2.4

รูปที่ 2.3 ถุงมือรับข้อมูล 5DT 14 Data Glove Ultra [9]



ตัวรับรู้ที่	รายละเอียด
0	ตัวรับรู้วัดการงอของนิ้วโป้งในตำแหน่งโคนนิ้วโป้ง
1	ตัวรับรู้วัดการงอของนิ้วโป้งในตำแหน่งปลายนิ้วโป้ง
2	ตัวรับรู้วัดการกางออกระหว่างนิ้วโป้ง และนิ้วชี้
3	ตัวรับรู้วัดการงอของนิ้วชี้ในตำแหน่งโคนนิ้วชี้
4	ตัวรับรู้วัดการงอของนิ้วชี้ในตำแหน่งปลายนิ้วชี้
5	ตัวรับรู้วัดการกางออกระหว่างนิ้วชี้ และนิ้วกลาง
6	ตัวรับรู้วัดการงอของนิ้วกลางในตำแหน่งโคนนิ้วกลาง
7	ตัวรับรู้วัดการงอของนิ้วกลางในตำแหน่งปลายนิ้วกลาง
8	ตัวรับรู้วัดการกางออกระหว่างนิ้วกลาง และนิ้วนาง
9	ตัวรับรู้วัดการงอของนิ้วนางในตำแหน่งโคนนิ้วนาง
10	ตัวรับรู้วัดการงอของนิ้วกนางในตำแหน่งปลายนิ้วนาง
11	ตัวรับรู้วัดการกางออกระหว่างนิ้วนาง และนิ้วก้อย
12	ตัวรับรู้วัดการงอของนิ้วก้อยในตำแหน่งโคนนิ้วก้อย
13	ตัวรับรู้วัดการงอของนิ้วก้อยในตำแหน่งปลายนิ้วก้อย

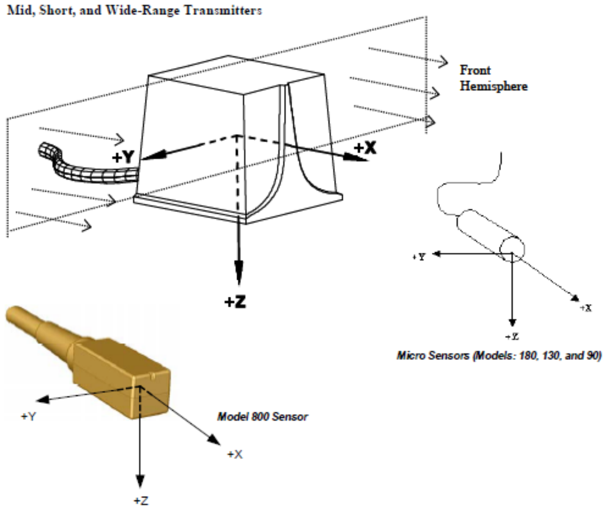
รูปที่ 2.4 รายละเอียดตัวรับรู้ที่ติดบนถุงมือแต่ละข้าง [9]

อุปกรณ์ติดตามการเคลื่อนที่

อุปกรณ์ติดตามการเคลื่อนที่ที่ใช้เพื่อติดตามการเคลื่อนที่ของมือในงานวิทยานิพนธ์นี้คือ 3D Guidance trakSTAR™ ดังรูปที่ 2.5 โดยประกอบไปด้วยแผงรับข้อมูล (รูปที่ 2.5 ก) ตัวส่งสัญญาณ (รูปที่ 2.5 ข) และตัวรับรู้ (sensor) โดยค่าที่ได้จากอุปกรณ์นี้มีทั้งหมด 6 ค่าต่อ 1 ตัวรับรู้ ได้แก่ พิกัดในแนวแกน x, y, z และมุมในแนวแกน x, y, z อ้างอิงจากการเทียบระยะจากตัวส่งสัญญาณถึงตัวรับรู้ ดังรูปที่ 2.6 ซึ่งทั้งตัวส่งสัญญาณและตัวรับรู้นั้นถูกต่อเข้ากับแผงรับข้อมูล และเพื่อให้สามารถตรวจจับตำแหน่งการเคลื่อนที่ของมือได้ จึงนำตัวรับรู้มาติดที่ด้านหลังมือและที่ข้อมือของทั้งสองข้าง

(ก) (ข) (ค)

รูปที่ 2.5 อุปกรณ์ติดตามการเคลื่อนที่ประกอบด้วยอุปกรณ์ 3 ชนิด
(ก) แผงรับข้อมูล (ข) ตัวส่งสัญญาณ (ค) ตัวรับรู้ [10]



รูปที่ 2.6 การวัดค่าเทียบจากแนวแกน x y z ของตัวรับรู้ และตัวรับสัญญาณ [10]

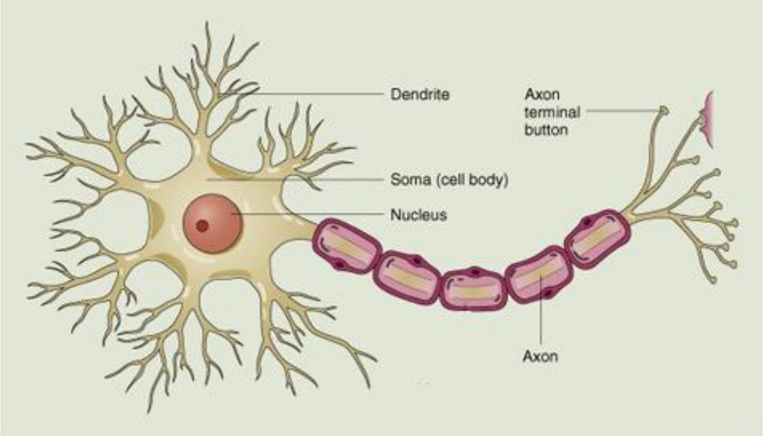
2.1.4 ข้อมูลอนุกรมเวลา (Time-Series)

ข้อมูลอนุกรมเวลา (Time Series data) คือ ข้อมูลมีการเก็บรวบรวมตามระยะเวลาโดยมีระยะห่างคงที่ เช่น ข้อมูลจำนวนเด็กแรกเกิดในแต่ละวัน ข้อมูลยอดขายสินค้าที่มีการเก็บรวบรวมทุก ๆ เดือน ข้อมูลรายได้ประชาชาติที่มีการเก็บทุก ๆ ปี ซึ่งในงานวิจัยนี้ข้อมูลที่ได้จากตัวรับรู้ (Sensor) ถือว่าเป็นข้อมูลอนุกรมเวลาเนื่องจากระยะห่างในการเก็บค่าจากตัวรับรู้ (Sensor) ในแต่ละครั้งมีระยะห่างคงที่ หรือความถี่ในการเก็บคงที่ ซึ่งข้อมูลที่ได้จากตัวรับรู้แต่ละตัวก็คือข้อมูลอนุกรมเวลา 1 ตัวและข้อมูลที่ได้เป็นข้อมูลที่มาจากตัวรับรู้หลายตัว ดังนั้นข้อมูลที่ใช้ในการทดลองนี้จึงเป็นข้อมูลอนุกรมเวลาที่มีหลายตัว

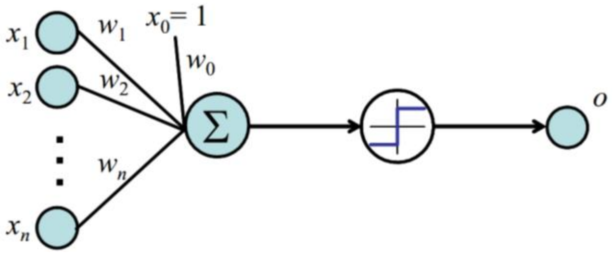
2.1.5 โครงข่ายประสาทเทียม (Neural Network)

ในงานวิทยานิพนธ์นี้พัฒนาระบบการรู้จำโดยเลือกใช้โครงข่ายประสาทเทียม (Artificial Neural Network) หรือที่มักจะเรียกสั้น ๆ ว่าข่ายงานประสาท (Neural Network หรือ Neural Net) มาเป็นส่วนประกอบในการพัฒนาอัลกอริทึมเนื่องจากโครงข่ายประสาทเทียม มีความสามารถในการจดจำรูปแบบที่มีความเหมือน หรือคล้ายคลึงกันได้ (Pattern Recognition) อีกทั้งยังเป็นอัลกอริทึมที่มีความยืดหยุ่นในเรื่องจำนวนของอินพุท และเหมาะกับงานวิทยานิพนธ์นี้ที่มีการใช้ตัวรับรู้จำนวนมากจึงมีอินพุทจำนวนมากด้วย

โครงข่ายประสาทเทียม [11] มีจุดเริ่มต้นมาจากแนวคิดที่ต้องการจำลองให้คอมพิวเตอร์สามารถคิดคำนวณได้เหมือนสมองของมนุษย์ที่มีความสามารถในการจดจำ และเรียนรู้ประสบการณ์ที่มีความซับซ้อนได้ ซึ่งสิ่งที่ทำให้สมองของคนเรานั้นสามารถมีความสามารถในการจดจำ และเรียนรู้ประสบการณ์นั้นเนื่องจากสมองประกอบไปด้วยเซลล์ประสาทดังรูปที่ 2.7 หลาย ๆ เซลล์ทำงานร่วมกันเพื่อประมวลผลด้วยการคำนวณแบบคอนเนคชันนิสต์ (connectionist)ซึ่งมีขั้นตอนการทำงานคือเซลล์ประสาททำการรับอินพุทที่เป็นสัญญาณไฟฟ้าจากเซลล์ประสาทอื่น ๆ เข้ามาทางใยประสาทนำเข้า (Dendrites) แล้วส่งต่อให้ตัวเซลล์ (Cell body) ประมวลผลเพื่อตัดสินใจส่งสัญญาณ หรือ output ออกมาทางแกนประสาทนำออก (Axon) ที่จะส่งต่อสัญญาณไฟฟ้าหรือผลลัพธ์ไปยังเซลล์ประสาทอื่น ๆ ให้ทำงานเช่นเดิม ซึ่งที่กล่าวมาข้างต้นนี้เป็นต้นแบบของการสร้างโครงข่ายประสาทเทียม โดยโครงข่ายประสาทเทียมมีการสร้างเพอร์เซปตรอน(Perceptron) รูปที่ 2.8 เพื่อเลียนแบบการทำงานของเซลล์ประสาท และยังสามารถเชื่อมกับเพอร์เซปตรอนตัวอื่น ๆ ได้เป็นเครือข่าย



รูปที่ 2.7 เซลล์ประสาท(Natural Neural) [11]



รูปที่ 2.8 เพอร์เซปตรอน (Perceptron) [11]

โดยเพอร์เซปตรอนหนึ่งตัวจะเป็นข่ายงานประสาทเทียมแบบง่าย ที่ประกอบไปด้วยตัวรับอินพุทซึ่งสามารถรับอินพุทเป็นเวกเตอร์จำนวนจริง $(x_1, x_2, \ldots, x_n)$ จากนั้นนำอินพุทมาคำนวณผลลัพธ์ด้วยการคำนวณหาผลรวมเชิงเส้น (Linear Combination) ที่มีการถ่วงน้ำหนักให้แก่อินพุทแต่ละตัว ได้เป็น $(w_1x_1 + w_2x_2 + \cdots + w_nx_n)$ โดยที่ถ้าผลรวมที่ได้มีค่าต่ำกว่า หรือสูงกว่าค่าขีดแบ่ง (θ) แล้วจะส่งผลต่อค่าเอาท์พุทที่ออกมาจากเพอร์เซปตรอน ที่มีการกำหนดด้วยฟังก์ชันกระตุ้นที่เลือกใช้เช่น ฟังก์ชันสองขั้ว (Bipolar function) จะสามารถแสดงเอาท์พุทได้เป็น -1 และ 1 เป็นต้น ซึ่งหากเขียนผลลัพธ์เป็นความสัมพันธ์ของค่าเอาท์พุท(o) ในรูปฟังก์ชันของอินพุท $(x_1, x_2, \ldots, x_n)$ ได้ดังนี้

$$o(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & if\ w_1x_1 + w_2x_2 + \cdots + w_nx_n > \theta \\ -1 & if\ w_1x_1 + w_2x_2 + \cdots + w_nx_n < \theta \end{cases} \quad (2.1)$$

ซึ่งสามารถทำการเขียนให้อยู่ในรูปใหม่ได้โดยย้ายข้างของค่าขีดแบ่ง (θ) ให้อยู่ในรูปผลรวมเชิงเส้นแบบถ่วงน้ำหนัก w_0x_0 ซึ่ง x_0 มีค่าเท่ากับ 1 เสมอ ดังนั้นเมื่อจัดฟอร์มใหม่จะทำให้อยู่ในรูป

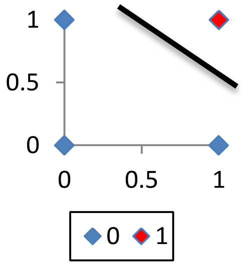
$$o(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & if\ w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n > 0 \\ -1 & if\ w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n < 0 \end{cases} \quad (2.2)$$

การนำเพอร์เซปตรอนไปแก้ปัญหาจะต้องมีการสอนให้เพอร์เซปตรอนมีการเรียนรู้ก่อน โดยการให้อินพุทแก่เพอร์เซปตรอน และต้องกำหนดเอาท์พุทให้แก่เพอร์เซปตรอนด้วย ซึ่งเพอร์เซปตรอนจะมีการเรียนรู้เพื่อปรับน้ำหนักให้เหมาะสมกับตัวอย่างที่ได้รับ ซึ่งเพอร์เซปตรอนนั้นจะมีลักษณะเป็นเส้นตรงในกรณีที่มีข้อมูลอินพุทเวกเตอร์ขนาด 2 มิติ และถ้ามากกว่า 2 มิติจะมีลักษณะเป็นระนาบตัดสินใจ ในที่นี้จะขอยกตัวอย่างการใช้เพอร์เซปตรอนเรียนรู้ฟังก์ชัน AND ดังรูปที่ 2.9 (ก) เป็นฟังก์ชันที่เพอร์เซปตรอนสามารถแบ่งแยกได้เรียกว่า ฟังก์ชันแยกเชิงเส้นได้ (linearly separable function) แต่อย่างไรก็ตามก็มีบางฟังก์ชันที่เพอร์เซปตรอนไม่สามารถแบ่งแยก

ได้ ดังรูปที่ 2.9 (ข) จากตัวอย่าง จะเห็นได้ว่าเพอร์เซปตรอนไม่สามารถเรียนรู้เพื่อสร้างเส้นตรงหรือระนาบที่ใช้แบ่งแยกประเภทได้ ซึ่งเป็นข้อจำกัดของเพอร์เซปตรอนที่ไม่สามารถเรียนรู้บางฟังก์ชันที่เป็นฟังก์ชันแยกเชิงเส้นไม่ได้ (linearly non-separable function) ดังนั้นจึงมีการนำเพอร์เซปตรอนมาต่อกันเป็นข่ายงานประสาทหลายชั้น (multilayer neural network)

AND

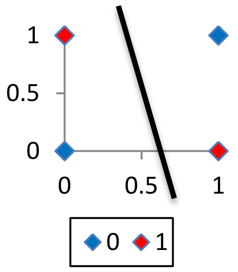
x_1	x_2	$output$
0	0	0
0	1	0
1	0	0
1	1	1



(ก)

XOR

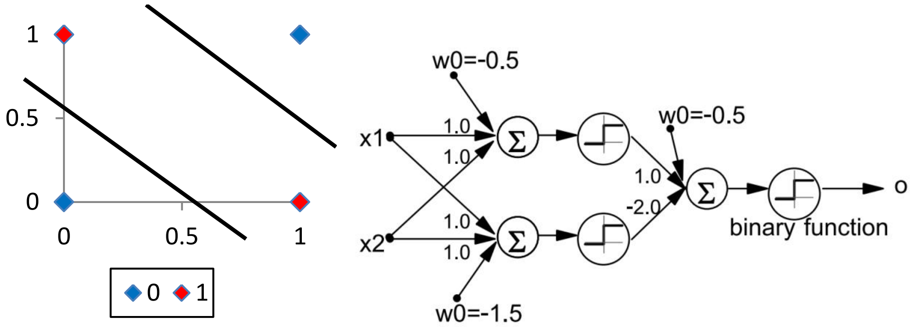
x_1	x_2	$output$
0	0	0
0	1	1
1	0	1
1	1	0



(ข)

รูปที่ 2.9 ตัวอย่างการเรียนรู้ฟังก์ชัน AND และ XOR ของเพอร์เซปตรอนตัวเดียว [11]
(ก) ตัวอย่างฟังก์ชัน AND ซึ่งเพอร์เซปตรอนสามารถเรียนรู้ได้
(ข) ตัวอย่างฟังก์ชัน XOR ซึ่งเพอร์เซปตรอนไม่สามารถเรียนรู้ได้

ข่ายงานประสาทหลายชั้น (Multilayer neural network) สามารถแก้ปัญหาที่เพอร์เซปตรอนเพียงตัวเดียวไม่สามารถแก้ไขได้เช่น ปัญหาของฟังก์ชัน XOR ดังรูปที่ 2.10



รูปที่ 2.10 การแก้ปัญหา XOR ด้วยข่ายงานประสาทหลายชั้น [11]

นอกจากจะช่วยแก้ปัญหาที่เป็นฟังก์ชันแยกเชิงเส้นไม่ได้แล้ว ข่ายงานประสาทหลายชั้นยังสามารถใช้ในกรณีที่ข้อมูลมีจำนวนมากกว่า 2 กลุ่ม (class) หรือเป็น multi-class ยกตัวอย่างเช่นงานทางด้านภาษามือเองก็มีลักษณะข้อมูลเป็น multiclass ก็มีงานวิจัยที่ใช้ ข่ายงานประสาทหลาย มาแก้ปัญหา เช่น การรู้จำการสะกดนิ้วมือของภาษาอเมริกัน [1] และของภาษาญี่ปุ่น [7] เป็นต้น ซึ่งใน

งานวิจัยนี้ต้องการจำแนกถึง 65 คลาสจึงเป็นอีกหนึ่งเหตุผลที่ใช้โครงข่ายประสาทเทียมเพราะประมวลผลลัพธ์ได้เร็ว

2.1.6 การหาลำดับย่อยร่วมที่ยาวสุด (Longest Common Subsequence) [12]

การหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence) เป็นวิธีการหาสายอักขระย่อยที่ยาวที่สุดและเหมือนกัน ของสายอักขระ 2 สาย โดยแต่เดิมนั้นนิยมใช้เพื่อเปรียบเทียบความเหมือนของสาย DNA การของการหาลำดับย่อยร่วมที่ยาวที่สุดมีหลักการหาคล้ายกับการหา Dynamic Time Warping โดยมีการคำนวณแต่ละเซลล์ด้วยสมการนี้

$$c[i,j] = \begin{cases} 0; & if\ i = 0\ or\ j = 0 \\ c[i-1,j-1]+1; & if\ i,j > 0\ and\ x_i = y_i \\ \max(c[i,j-1], c[i-1,j]); & if\ i,j > 0\ and\ x_i = y_i \end{cases} \quad (2.3)$$

เมื่อ $c[i,j]$ คือ คะแนนความเหมือนของในแต่ละช่องบนเมทริกซ์ที่มีขนาด MxN

M คือ ขนาดความยาวของสายอักขระที่ 1

N คือ ขนาดความยาวของสายอักขระที่ 2

x_i คือ อักขระตำแหน่งที่ i บนสายอักขระที่ 1

y_i คือ อักขระตำแหน่งที่ i บนสายอักขระที่ 2

ตัวอย่าง

ตารางที่ 2.1 การคำนวณหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence)

สายอักขระที่ 1: ก 1 ว า ง				**สายอักขระที่ 2 : ก 2 ไม้โท ว า ง**		
		ก	**1**	**ว**	**า**	**ง**
	0	0	0	0	0	0
ก	0	0	1	1	1	1
2	0	1	1	2	2	2
ไม้โท	0	1	2	2	3	3
ว	0	1	2	2	3	4
า	0	1	2	3	2	3
ง	0	1	2	3	3	2

ผลลัพธ์ที่บอกความยาวของสายอักขระย่อยที่ยาวที่สุดของการคำนวณจะเท่ากับค่าที่ $c[M, N]$ และหากต้องการหาสายอักขระที่ยาวที่สุดสามารถหาได้โดยการ trace back จากตารางที่ 2.1 ซึ่งจะได้ผลลัพธ์ดังตารางที่ 2.2 ซึ่งผลลัพธ์ที่ได้คือ สายอักขระย่อย กวาง

ตารางที่ 2.2 ผลการ Trace back เพื่อหาสายอักระย่อยที่ยาวที่สุด

		ก	า	ว	า	ง
	0	0	0	0	0	0
ก	0	ก	ก	ก	ก	ก
2	0	ก	ก	ก	ก	ก
ไม้โท	0	ก	ก	ก	ก	ก
ว	0	ก	ก	กว	กว	กว
า	0	ก	ก	กว	กวา	กวา
ง	0	ก	ก	กว	กวา	กวาง

ซึ่งในงานวิจัยชิ้นนี้ได้ทำการทดสอบอัลกอริทึมนี้มาใช้ประยุกต์เพื่อหาความเหมือนกันของลำดับท่ามืออินพุทที่เข้ามาสู่ขั้นตอนการรู้จำ กับลำดับท่ามือที่มีเก็บอยู่ในพจนานุกรมซึ่งจะกล่าวถึงอย่างละเอียดในบทที่ 3

2.1.7 การวัดหาค่าความต่างกันของสายอักขระสองชุด(Minimum Edit Distance) [13]

การวัดหาค่าความต่างกันของสายอักขระสองชุด(Minimum Edit Distance) เป็นขั้นตอนวิธีการวัดหาค่าความต่างกันของสายอักขระสองชุด ระหว่างชุดแรกที่เป็นต้นแบบ และ ชุดที่สองที่เป็นชุดเปรียบเทียบ โดยค่าความต่างกันจะวัดจากจำนวนของขั้นตอนการที่จะต้องทำการตัดออก (Delete) แทรก (Insert) และแทนที่ (Substitution) อักขระในชุดที่นำมาเปรียบเทียบจนกระทั่งมีลักษณะเหมือนชุดอักขระที่เป็นต้นแบบทุกประการ ซึ่งหลักการของขั้นตอนวิธีการนี้จะเป็นการนำชุดอักขระ 2 ชุด มาเปรียบเทียบจำนวนความแตกต่างกัน จะพิจารณาด้วยวิธีการดังนี้

1. แทรก (Insert) เป็นการนำเอาอักขระตัวใดๆมา เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งใน เช่น *ยา* ⟶ *ยาว* จะแทรกตัว *ว* ให้กับ *ยา* เพื่อให้ *ยา* กลายเป็น *ยาว* เป็นต้น
2. ตัดออก (Delete) เป็นการตัดอักขระออกครั้งละ 1 ตัว จากชุดอักขระตัวหนึ่ง เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งในภายหลัง เช่น *ยาว* ⟶ *ยา* จะตัดตัว *ว* ออก เพื่อให้ *ยาว* กลายเป็น *ยา* เป็นต้น

3. แทนที่ (Substitution) เป็นการนำอักขระของชุดอักขระหนึ่งไปแทนอักขระของอีกชุดอักขระหนึ่ง เพื่อให้ชุดอักขระชุดนั้นเหมือนกับอีกชุดอักขระหนึ่งในภายหลัง เช่น *ขาว* ⟶ *ยาว* จะแทนที่ *ข* ด้วย ย เพื่อให้ *ขาว* กลายเป็น *ยาว* หรือ อาจมองในทางกลับกันก็ได้ เป็นต้น

ตัวอย่าง เมื่อต้องการเปรียบเทียบสายอักขระ ก, 1 ว, า, ง และ ก, 2, ไม้โท, ว, า, ง

ตารางที่ 2.3 ตารางการหาค่าความแตกต่างของ Minimum Edit Distance

ก	*	1	*	ว	า	ง
|	|	|	|	|	|	|
ก	2	*	ไม้โท	ว	า	ง
-	Del	Ins	Del	-	-	-
0	**1**	**1**	**1**	0	0	0

กำหนดให้การตัดออก (Delete) แทรก (Insert) และแทนที่ (Substitution) มีค่าความแตกต่างของสายอักขระเท่ากับ 1 ทุกวิธีการแล้วเมื่อทำการหาค่าความแตกต่างระหว่าง ก, 1 ว, า, ง และ ก, 2, ไม้โท, ว, า, ง แล้วจะได้คะแนนค่าความแตกต่างระหว่างสองสายอักขระเท่ากับ 3

2.2 งานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะเน้นกล่าวถึงงานวิจัยต่าง ๆ ที่เกี่ยวข้องกับการรู้จำภาษามือไทย และงานวิจัยที่เกี่ยวข้องกับการรู้จำภาษามือต่างประเทศทั่วโลกบางส่วนที่น่าสนใจ เนื่องจากงานวิจัยเกี่ยวกับการรู้จำภาษามือทั่วโลกนั้นมีมามากกว่า 10 ปี จึงทำให้มีงานวิจัยการรู้จำภาษามือต่าง ๆ เป็นจำนวนมากที่ถูกเผยแพร่ขึ้นมา เช่น ภาษามืออเมริกัน [1] ภาษามือบริติช [5] ภาษามือออสเตรเลีย [4] ภาษามือญี่ปุ่น [3] และภาษามือปากีสถาน [6] เป็นต้น ซึ่งถ้าสรุปโดยภาพรวมแล้วสามารถแบ่งออกได้เป็น 2 กลุ่มใหญ่ ๆ ตามลักษณะของการรับข้อมูลเข้าสู่ระบบดังนี้

2.2.1 การรู้จำภาษามือโดยใช้การประมวลผลภาพจากกล้อง

งานวิจัยในด้านนี้จะใช้อุปกรณ์นำข้อมูลเข้าคือกล้องวิดีโอ โดยระบบจะสกัดคุณลักษณะเด่นของมือออกมา เช่น ระยะห่างระหว่างนิ้วแต่ละนิ้ว โมเมนต์ของภาพ (Image moment) พื้นที่ (Area) และเส้นขอบ (Contour) เป็นต้น ดังนั้นในขั้นตอนแรกของการรู้จำภาษามือโดยใช้การ

ประมวลผลภาพจากกล้องจะต้องมีขั้นตอนการสกัดคุณลักษณะต่าง ๆ ที่สำคัญของมือจากรูปภาพที่เข้ามาให้ได้เสียก่อน จากนั้นจึงจะใช้แบบจำลองเพื่อทำการรู้จำ

ข้อดีของการรู้จำภาษามือโดยใช้การประมวลผลภาพจากกล้องคือ สามารถหาซื้ออุปกรณ์ได้ง่าย และมีราคาถูกเพราะสามารถใช้กล้องวิดีโอที่ใช้กับเครื่องคอมพิวเตอร์ทั่วไปได้ แต่ก็มีข้อเสียคือคุณภาพของคุณลักษณะที่สกัดออกมาจากภาพนั้นยังไม่เพียงพอต่อการนำมาใช้ในการรู้จำ เนื่องจากข้อจำกัดของกล้องปกติทั่วไปทำให้การระบุส่วนต่าง ๆ ของมือนั้นยังทำได้ยากมีความผิดพลาดมาก ซึ่งส่งผลไปถึงการรู้จำภาษามือที่จะเกิดผิดพลาดตามไปด้วย นอกจากนี้กล้องยังมีขีดจำกัดในการตรวจจับการงอของนิ้วเมื่อมีการซ้อนทับกันของนิ้วหรือมือ ซึ่งในความเป็นจริงแล้วมีท่าในภาษามือเป็นจำนวนมากที่มีการซ้อนทับกันของนิ้วมือหรือฝ่ามือจึงทำให้ประสิทธิผลของระบบการรู้จำโดยใช้กล้องเป็นอุปกรณ์นำเข้าค่อนข้างต่ำ

งานวิจัยที่เป็นการรู้จำภาษามือไทยโดยใช้การประมวลผลภาพจากกล้องมี 2 งานวิจัยที่น่าสนใจได้แก่

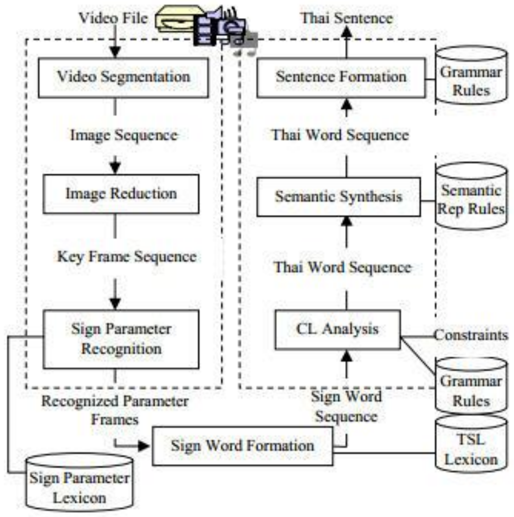
- TSTMT: Step towards an Accurate Thai Sign Translation [14]
- Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform [15]

TSTMT: Step towards an Accurate Thai Sign Translation [14]

งานวิจัยนี้ใช้การรู้จำภาษามือโดยใช้การประมวลผลภาพจากกล้อง โดยมีแบบจำลองที่ใช้ในการรู้จำคือแบบจำลองมาคอฟ (Hidden Markov Models : HMMs) ซึ่งเป็นแบบจำลองที่ใช้ในการจำลองความน่าจะเป็นของลำดับของเหตุการณ์ที่สนใจ ซึ่ง TSTMT นอกจากจะมีการใช้แบบจำลองมาคอฟทำการรู้จำแล้ว ยังมีการนำวิธีประมวลผลทางภาษาธรรมชาติ (NLP: Natural Language Processing) มาใช้เพื่อเพิ่มความแม่นยำในการรู้จำโดยอาศัยการวิเคราะห์ในการสร้างประโยคโดยใช้ไวยากรณ์ที่เป็นสิ่งที่น่าสนใจ

โดยขั้นตอนในการทำงานของงานวิจัยนี้แสดงดังรูปที่ 2.11 ซึ่งระบบจะทำการเก็บวิดีโอมาเป็นภาพถ่ายแต่ละเฟรมเรียงต่อ ๆ กันเป็นลำดับและทำการลดจำนวนเฟรมที่ไม่มีความหมายซึ่งหมายถึงเฟรมที่ไม่มีการทำท่ามือที่มีความหมายอยู่ โดยจะนำเฟรมดังกล่าวออกไป เหลือไว้แต่เฟรมที่มีท่ามือสื่อความหมายเท่านั้น จากนั้นจะนำเฟรมที่เหลือไปทำการรู้จำ และทำการเรียบเรียงตามหลักไวยากรณ์ เพื่อจัดเรียงรูปแบบการเรียงของคำภายในประโยคให้ถูกต้อง เพราะรูปแบบการ

สื่อสารด้วยภาษามือนั้นมีลำดับการเรียงของคำในประโยคภาษามือที่แตกต่างจากการเรียงคำในภาษาไทยปกติ



รูปที่ 2.11 ระบบการรู้จำของงานวิจัย TSTMT [14]

Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform [15]

เป็นงานวิจัยที่ทำการทดลองการรู้จำการสะกดนิ้วมือของภาษามือไทยโดยใช้การประมวลผลภาพจากกล้อง ซึ่งในงานวิจัยนี้ใช้ Fuzzy C-Means (FCM) เพื่อสร้างแบบจำลอง และ Scale Invariant Feature Transform (SIFT) เป็นวิธีในการสกัดคุณลักษณะของมือออกมา และทำการรู้จำโดยข้อมูลตัวอักษรที่นำมาใช้ในการรู้จำมีเพียง 15 ตัวอักษรดังรูปที่ 2.12 แต่ในความเป็นจริงมีตัวอักษรจำนวนมากถึง 65 ตัวอักษร ซึ่งหลักการในการสกัดคุณสมบัติของภาพจากงานวิจัยนี้ไม่สามารถใช้ได้กับท่าที่มีความคล้ายคลึงกันมาก ๆ หรือท่ามือสะกดตัวอักษรบางตัวซึ่งใช้สองมือในการทำท่าดังรูปที่ 1.4 ซึ่งจะมีหลายรูปแบบท่ามือที่มีความซ้อนทับกันของนิ้วมือซึ่งมีความคล้ายคลึงกันมาก อีกทั้งในบางท่ามือมีการซ้อนทับกันของนิ้วมือ หรือการซ้อนทับกันของมือ ซึ่งงานวิจัยนี้ไม่ได้ทำการรู้จำที่ครอบคลุมท่ามือทั้งหมด รวมไปถึงไม่รองรับตัวอักษรที่เกิดจากการประกอบกันของลำดับท่ามือที่มากกว่า 1 ท่ามือ ถึงแม้ผลการทดลองจะให้ Accuracy ที่ดีที่สุดถึง 80.57% ด้วยการ manual เลือก key frame ที่สูงที่สุด และหากใช้การเลือก key frame แบบอัติโนมัติ ระบบจะให้ค่า Accuracy 79.45% แต่อย่างไรก็ตามเป็นเพียงการรู้จำตัวอักษรเพียงส่วนน้อยของตัวอักษรไทย

รูปที่ 2.12 ท่ามือการสะกดการรู้จำที่ใช้ในงานวิจัย [15]

2.2.2 การรู้จำภาษามือโดยใช้การประมวลผลสัญญาณจากถุงมือและอุปกรณ์ที่ใช้ติดตามการเคลื่อนที่ของมือ

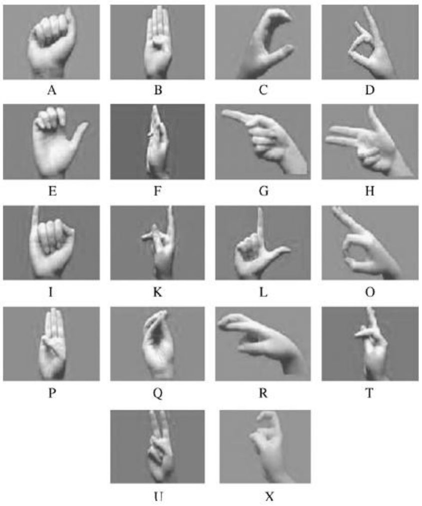
งานวิจัยในด้านนี้จะเน้นใช้อุปกรณ์นำข้อมูลเข้าคือถุงมือและอุปกรณ์การติดตามการเคลื่อนที่ของมือ ซึ่งข้อมูลที่ได้จากตัวรับรู้บนถุงมือโดยส่วนใหญ่แม้จะต่างรุ่นกันแต่จะให้ข้อมูลของการงอของข้อนิ้ว และการกางออกระหว่างนิ้วได้ ส่วนข้อมูลจากสัญญาณอุปกรณ์ติดตามมือจะให้ตำแหน่งของมือ และมุมที่มือทำมุมกันแนวแกน x, y, z ซึ่งเมื่อทำการรับข้อมูลจากอุปกรณ์ทั้ง 2 อย่างมาแล้วนั้นสามารถนำมาใช้ทำการรู้จำได้โดยตรงทำให้ไม่ต้องผ่านขั้นตอนการสกัดคุณลักษณะจากภาพก่อน

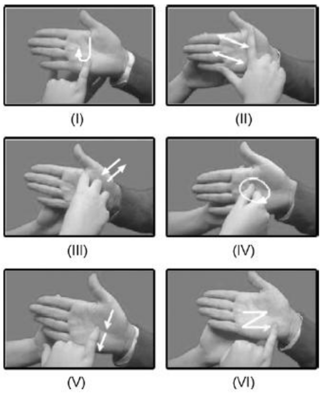
ข้อดีของการรู้จำแบบนี้คือการรู้จำมีความแม่นยำเนื่องจากได้ข้อมูลของลักษณะมือโดยตรงทำให้การรู้จำมีประสิทธิผล และสามารถลดขั้นตอนของความผิดพลาดในการได้มาซึ่งข้อมูลสำหรับสร้างแบบจำลอง และข้อเสียของการรู้จำแบบนี้คือความไม่สะดวกในการใช้งานเนื่องจากต้องสวมถุงมือทุกครั้งที่จะใช้งาน และอุปกรณ์ที่ใช้มีราคาค่อนข้างสูง ซึ่งงานวิจัยที่ใช้อุปกรณ์ชนิดนี้ มีเป็นจำนวนมากซึ่งในที่นี้จะขอยกตัวอย่างงานวิจัยเพียง 2 งานที่มีความน่าสนใจได้แก่

- Layered Architecture for Real-Time Sign Recognition [2, 3]
- American Sign Language word recognition with a sensory glove using artificial Neural networks [1]

Layered Architecture for Real-Time Sign Recognition [2, 3]

งานวิจัยนี้มีการแบ่งอักขระในการสะกดนิ้วมือในภาษาอังกฤษออกเป็น 2 กลุ่ม คือ ท่ามือที่ไม่มีการเคลื่อนที่ของมือ (Hand Gesture) ดังรูปที่ 2.13 (ก) และท่ามือที่มีการเคลื่อนที่ของมือ (Hand Movement) ดังรูปที่ 2.13 (ข)





(ก) (ข)

รูปที่ 2.13 ท่ามือที่ใช้ทำการรู้จำในงานวิจัย [2, 3]

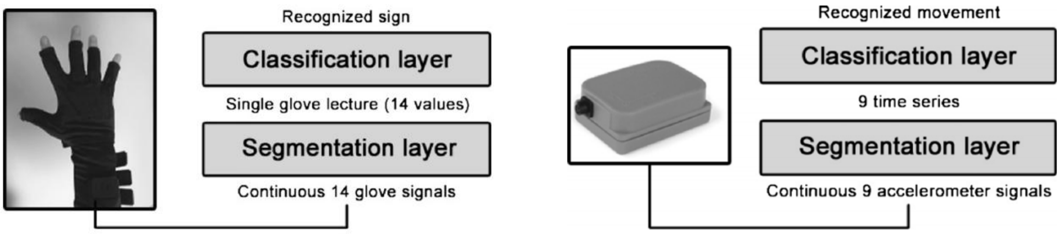
(ก) ท่ามือที่ไม่มีการเคลื่อนที่ของมือ (Hand Gesture)

(ข) ท่ามือที่มีการเคลื่อนที่ของมือ (Hand Movement)

ซึ่งทั้ง 2 ลักษณะมีคุณสมบัติที่แตกต่างกันคือการเคลื่อนที่ โดยท่ามือที่ไม่มีการเคลื่อนที่จะมีลักษณะเป็นเหมือนการทำท่ามือค้างไว้แล้วเกิดความหมาย โดยไม่ต้องอาศัยการเคลื่อนที่ หรือทิศทางของมือในการสื่อความหมายใด ๆ ซึ่งแตกต่างจากการทำท่ามือที่มีการเคลื่อนที่อาศัยการเคลื่อนที่โดยอย่างมีทิศทางของมือเป็นส่วนประกอบในการสื่อความหมาย

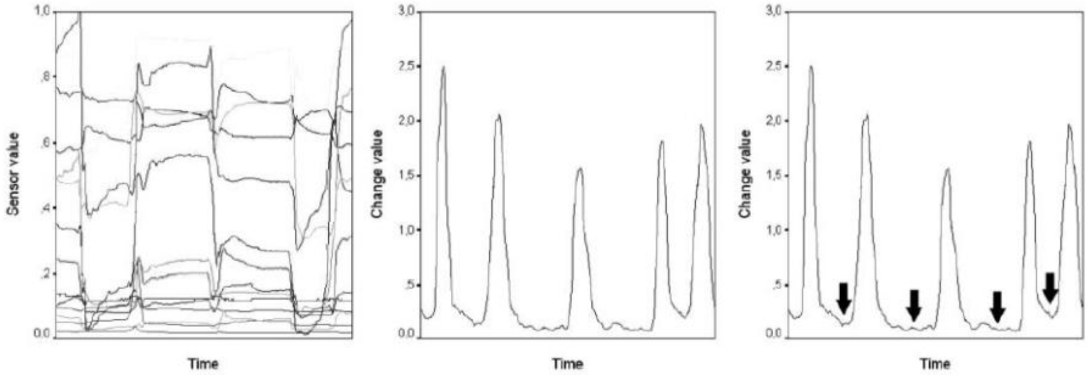
เนื่องจากมีการแบ่งท่ามือออกเป็น 2 ประเภท ทำให้ในขั้นตอนการรู้จำในงานวิจัยนี้มีการแบ่งออกเป็น 2 ส่วนใหญ่ ๆ โดยแบ่งตามลักษณะของท่ามือ ซึ่งทั้ง 2 ส่วนก็มีการแบ่งเป็นขั้นตอนย่อย ๆ ในการรู้จำเหมือน ๆ กันคือ การตัดแบ่งสัญญาณ (Segmentation) และการจำแนกประเภท (Classification) ดังรูปที่ 2.14 ซึ่งในขั้นตอนการรู้จำของท่ามือที่ไม่มีการเคลื่อนที่จะใช้อินพุทเป็นข้อมูลถุงมือดังรูปที่ 2.14 (ก) เพราะจะให้ความสนใจกับรูปร่างของมือเป็นหลัก แตกต่างจาก

ขั้นตอนการรู้จำที่จะใช้ข้อมูลจากอุปกรณ์ติดตามการเคลื่อนที่ของมือเป็นหลักเพื่อต้องการดูการเคลื่อนที่ของมือว่ามีความหมายอย่างไร ดังรูปที่ 2.14 (ข)



(ก) (ข)

รูปที่ 2.14 ขั้นตอนการรู้จำของท่ามือในงานวิจัย [2, 3] แบ่งออกเป็น 2 ขั้นตอนคือ
การตัดแบ่งสัญญาณ (Segmentation) และ การจำแนกประเภท (Classification)
(ก) ท่ามือที่ไม่มีการเคลื่อนที่ (Hand Gesture) (ข)ท่ามือที่มีการเคลื่อนที่ของมือ (Hand Movement)

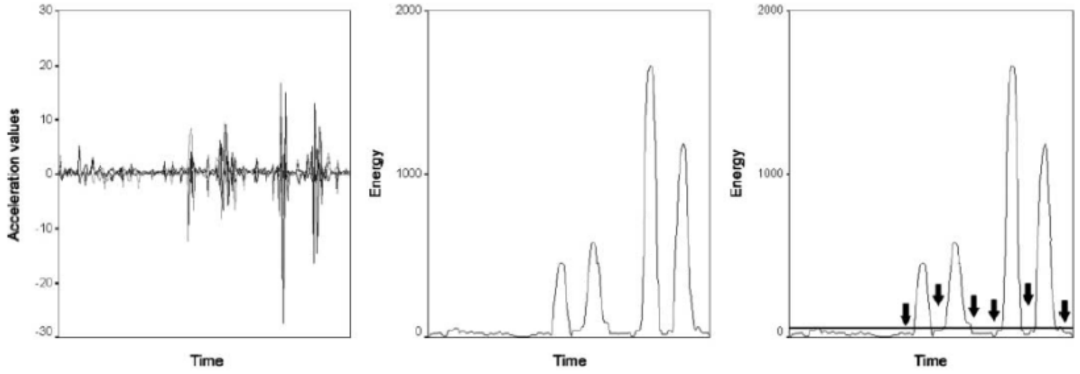


(ก) (ข) (ค)

รูปที่ 2.15 ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่ไม่มีการเคลื่อนที่ [2]
(ก) ค่าที่ได้จากตัวรับรู้บนถุงมือรับข้อมูลทั้ง 14 ตัว
(ข) การหาการเปลี่ยนแปลงของค่าที่ได้จาตัวรับรู้ทั้ง 14 ตัว
(ค) การหาจุดที่เป็น local minimum

ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่ไม่มีการเคลื่อนที่นั้นใช้หลักการในการหาจุดเวลาที่มีอัตราการเปลี่ยนแปลงของตัวรับรู้ที่อยู่บนถุงมือที่มีค่าน้อยที่สุดหรือจุดที่เรียกว่า Local Minimum และเมื่อหาจุดเวลาดังกล่าวได้ดังรูปที่ 2.15 (ค) จะนำค่าของตัวรับรู้มาทำการรู้จำและอัลกอริทึม K-NN มาทำการรู้จำ

ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่มีการเคลื่อนที่ดังรูปที่ 2.16 นั้นใช้หลักการเดียวกับการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่ไม่มีการเคลื่อนที่แต่เปลี่ยนจากการหาจุดเวลาที่มีอัตราการเปลี่ยนแปลงของตัวรับรู้ที่อยู่บนถุงมือที่มีค่าน้อยที่สุด เป็นการหาจุดเวลาที่มีค่าน้อยที่สุด (Local Minimum) ของค่าพลังงาน (Energy) ที่เกิดจากการเคลื่อนไหวมือ จากนั้นจะนำช่วงเวลาที่สกัดได้นี้ไปหาค่าคุณสมบัติต่าง ๆ เช่น จำนวนจุดยอด ,จำนวนจุดยอดที่มีค่าบวก ,จำนวนจุดยอดที่มีค่าลบ และแอมพลิจูด เป็นต้น จากนั้นจะนำคุณสมบัติเหล่านี้มาเข้าสู่ขั้นตอนการรู้จำซึ่งมีการทดลองใช้ 3 อัลกอริทึม ได้แก่ C4.5 ,1-NN และNaïve-Bayes ซึ่งผลที่ดีที่สุดคือ Naïve-Bayes

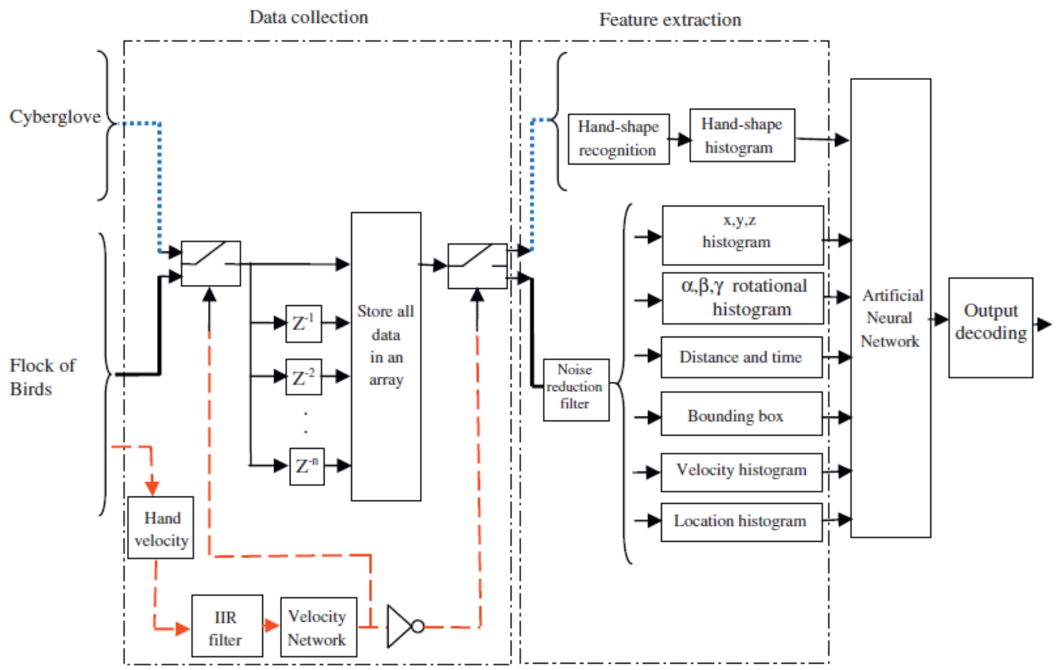


รูปที่ 2.16 ขั้นตอนการตัดแบ่งสัญญาณ (Segmentation) ของท่ามือที่มีการเคลื่อนที่ [2]
(ก) ค่าที่ได้จากอุปกรณ์ติดตามการเคลื่อนที่
(ข) พลังงานที่คำนวณจากการเคลื่อนที่มือ
(ค) การหาจุดที่เป็น local minimum

แม้ว่าหลักการในการทำ Segmentation ในงานวิจัยนี้เป็นหลักการที่ดี แต่จากการที่ระบบไม่สามารถนำขั้นตอนการรู้จำของท่ามือทั้ง 2 แบบได้แก่ ท่ามือที่ไม่มีการเคลื่อนที่ และท่ามือที่มีการเคลื่อนที่มารวมกันได้ ทำให้ระบบยังไม่สมบูรณ์เพราะต้องมีการจำแนกประเภทของท่ามือให้ได้ก่อนที่จะเข้าสู่ขั้นตอนการรู้จำ

American Sign Language word recognition with a sensory glove using artificial Neural networks [1]

งานวิจัยนี้มีความเป็นระบบ และมีความอัตโนมัติมากกว่างานวิจัย [2, 3] เพราะระบบนี้ใช้ Artificial Neural Network ในการพัฒนาโดยระบบนี้มีหลักการทำงานตามรูปที่ 2.17 กล่าวคือ ในขั้นตอนการเก็บข้อมูลจะมีการใช้ Neural Network ในการตัดแบ่งข้อมูล (Segmentation) ไปพร้อมกัน และหลังจากที่ทำการตัดแบ่งข้อมูลแล้วนั้น จะนำเอาข้อมูลเหล่านั้นมาผ่านขั้นตอน Feature Extraction เพื่อสกัดคุณลักษณะต่าง ๆ จากนั้นจึงนำคุณลักษณะดังกล่าวไปเป็นข้อมูลเข้า ให้แก่ Artificial Neural Network ตัวสุดท้ายเพื่อทำการรู้จำและแปลออกมาเป็นคำศัพท์ แต่อย่างไรก็ตามก็ยังไม่สามารถนำมาใช้กับการสะกดนิ้วมือไทยได้ อาจจะต้องมีการพัฒนาเพิ่มเติมให้สามารถใช้กับการสะกดนิ้วมือที่ใช้มือทั้งสองข้างทำท่ามือ



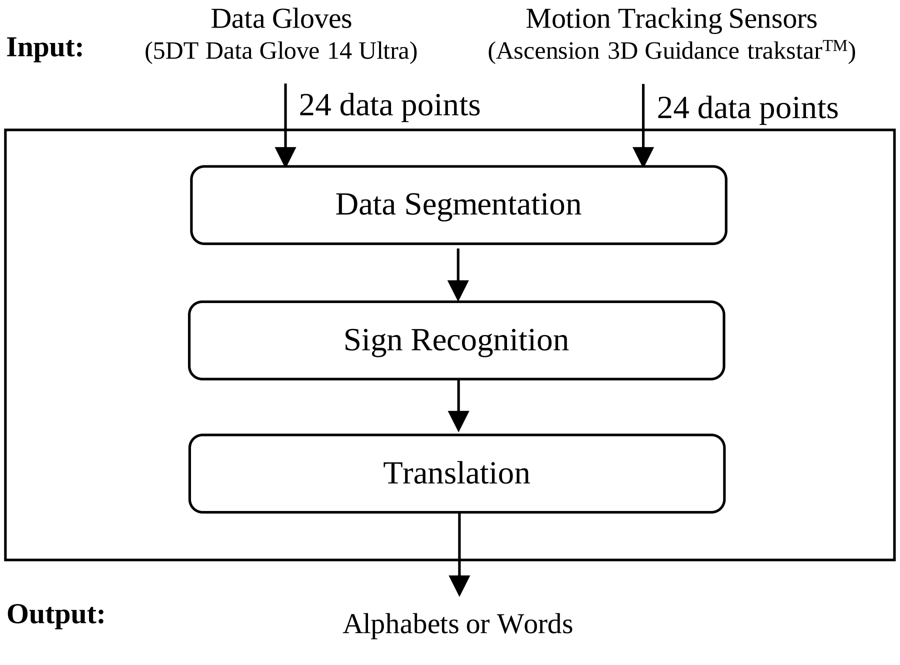
รูปที่ 2.17 ภาพรวมแสดงขั้นตอนการดึงข้อมูลและขั้นตอนการรู้จำ [1]

บทที่ 3
การรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทย
จากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่

ในบทนี้จะเป็นการกล่าวถึงรายละเอียดต่าง ๆ เกี่ยวกับขั้นตอน และอัลกอริทึมที่ใช้ในการพัฒนาการรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์การติดตามเคลื่อนที่ โดยเนื้อหาส่วนแรกจะกล่าวถึงภาพรวมของขั้นตอนต่าง ๆ ในการพัฒนาการรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่ซึ่งสามารถแบ่งออกเป็น 3 ส่วนหลัก ๆ ซึ่งจะถูกบรรยายในหัวข้อต่อ ๆ ไปตามลำดับ

3.1 ภาพรวมของขั้นตอนการรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่

ขั้นตอนการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูล และอุปกรณ์การติดตามเคลื่อนที่ในวิทยานิพนธ์นี้สามารถแบ่งได้เป็น 3 ขั้นตอนดังรูปที่ 3.1



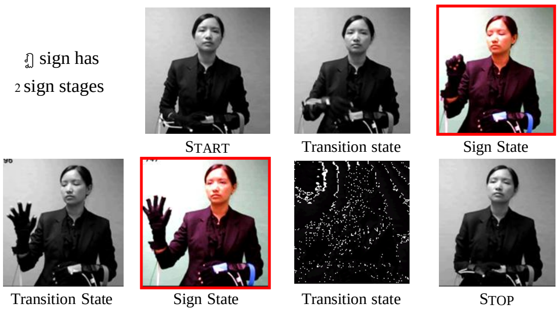
รูปที่ 3.1 ภาพรวมการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์การติดตามการเคลื่อนที่

โดยในขั้นตอนแรกนี้คือ การสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ (Data Segmentation) หลังจากสามารถหาช่วงเวลาที่ทำท่ามือได้แล้วจะนำค่าที่ได้จากอุปกรณ์ทั้งสองอย่าง คือทั้งถุงมือ และอุปกรณ์ติดตามการเคลื่อนที่ บนช่วงเวลาที่ได้จากขั้นตอนที่แรก มาเข้าสู่ขั้นตอนที่สองซึ่งเรียกว่าการรู้จำท่ามือ (Sign Recognition) ในขั้นตอนนี้จะทำการแปลงจากข้อมูลที่ได้จากอุปกรณ์ทั้งสองได้ผลลัพธ์เป็นท่ามือ ซึ่งจากท่ามือในขั้นตอนที่สองนี้จะถูกเป็นไว้เป็นลำดับจนกว่าจะหมดช่วงเวลาที่ทำท่ามือ หรือเสร็จสิ้นการสะกดคำ ลำดับท่ามือที่ถูกเก็บไว้จะถูกนำมาเข้าสู่ขั้นตอนสุดท้ายเรียกว่าการแปลท่ามือ (Translation) จะทำการแปลจากลำดับท่ามือออกมาเป็นตัวอักษร หรือคำซึ่งเป็นผลลัพธ์สุดท้าย

หลังจากที่ได้ทราบภาพรวมของขั้นตอนการรู้จำการรู้จำการสะกดนิ้วมือภาษามือไทยจากถุงมือข้อมูลและอุปกรณ์ติดตามการเคลื่อนที่แล้ว ในหัวข้อถัดไปของบทนี้จะกล่าวถึงรายละเอียดของแต่ละขั้นตอนเริ่มจาก 3.2 การสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ (Data Segmentation) 3.3 การรู้จำท่ามือ (Sign Recognition) และสุดท้าย 3.4 การแปลท่ามือ (Translation)

3.2 การสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ (Data Segmentation)

แนวคิดในการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือนี้มาจากการสังเกตลักษณะการเคลื่อนไหวของมือ และนิ้วมือขณะทำท่ามือในการสะกดนิ้วมือ ประกอบกับการสอบถามล่ามผู้มีความเชี่ยวชาญในการทำท่ามือ โดยการทำการสะกดนิ้วมือนั้นจะต้องมีช่วงเวลาที่การทำท่ามือเมื่อทำท่ามือเสร็จ ตำแหน่งของมือที่เคลื่อนที่ และการขยับงอหรือยืดออกของนิ้วมือจะมีการหยุดนิ่ง แต่ก่อนที่จะเริ่มอธิบายถึงวิธีการที่ใช้ในการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือนั้น ผู้เขียนวิทยานิพนธ์ขอกล่าวถึงช่วงเวลา หรือสถานะของการทำท่ามือที่จำเป็นต้องกล่าวถึงในขั้นตอนนี้ โดยยกตัวอย่างภาพประกอบเพื่อความเข้าใจดังรูปที่ 3.2 เป็นสถานะต่างๆตั้งแต่มีการเริ่มทำท่ามือจนจบการทำท่ามือของท่ามือ ฎ



รูปที่ 3.2 ช่วงเวลา หรือสถานะต่าง ๆ ของผู้ทำท่ามือขณะเริ่มทำท่ามือ ฎ จนสิ้นสุด

สถานะเริ่มต้นทำท่ามือ หรือช่วงเวลาเริ่มต้นทำท่ามือ (Start State) และสถานะสิ้นสุดการทำท่ามือ (End State) ทั้งสองสถานะนี้ผู้ทำท่ามือจะมีการทำท่ามืออยู่ในท่าพักมือคือนำฝ่ามือทั้งสองคว่ำลง และวางลงบนหน้าตักของผู้ทำท่ามือ

สถานะเปลี่ยนแปลงท่ามือ หรือช่วงเวลาเปลี่ยนแปลง (Transition State) สถานะนี้เป็นสถานะที่ผู้ทำท่ามือกำลังเปลี่ยนแปลงรูปร่างของมือ เช่นอาจจะกำลังขยับนิ้วเพื่อทำท่ามือแต่อาจจะทำยังไม่เสร็จสมบูรณ์ หรืออาจจะอยู่ในช่วงขยับมือเปลี่ยนตำแหน่งจากสถานะเริ่มต้นไปเป็นสถานะทำท่ามือ (Sign State)

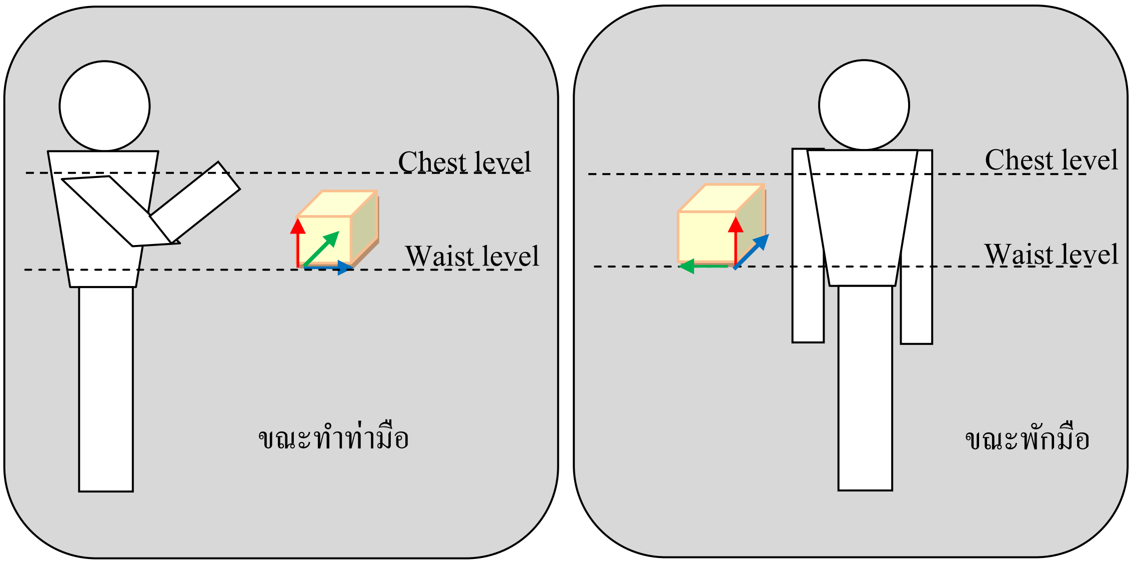
สถานะทำท่ามือ หรือช่วงเวลาขณะทำท่ามือ (Sign State) สถานะนี้เป็นสถานะที่ต้องการทราบเวลาที่ชัดเจนเนื่องจากเป็นช่วงเวลาที่ผู้ทำท่ามือทำท่ามือแล้วท่ามือดังกล่าวสามารถเป็นท่ามือที่สื่อความหมาย หรือมีรูปแบบของการงอนิ้วตรงตามแบบสะกดนิ้วมือของภาษามือไทย

จากสถานะ หรือช่วงเวลาที่ได้กล่าวมาข้างต้นนั้น สถานะหรือช่วงเวลาที่ขั้นตอนของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือให้ความสนใจมากที่สุดคือเวลาขณะที่ผู้ทำท่ามืออยู่ในสถานะทำท่ามือ เพื่อหาเฟรมสำคัญ (Key Frame) จากช่วงเวลาขณะทำท่ามือ โดยในกระบวนการในขั้นตอนนี้มีการแบ่งออกเป็น 3 ส่วนคือ

1. การตรวจสอบตำแหน่งของมือว่าอยู่ในขอบเขตตำแหน่งของการทำท่ามือ
2. การสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่ (Segmentation using data from motion tracking sensors)
3. การสกัดหาช่วงเวลาขณะทำท่ามือด้วยถุงมือรับข้อมูล (Segmentation by using data from the data gloves)

3.2.1 การตรวจสอบตำแหน่งของมือว่าอยู่ในขอบเขตของการทำท่ามือ

ในขั้นตอนนี้เป็นการตรวจสอบเบื้องต้นว่ามีมือข้างใดบ้างกำลังอยู่ในสถานะที่มีแนวโน้มจะทำท่ามือโดยมีการกำหนดบริเวณหรือขอบเขตที่จะทำท่ามือได้โดยกำหนดตำแหน่งความสูงของที่ตั้งอุปกรณ์ตัวส่งสัญญาณ(Transmitter) ให้อยู่ระดับต่ำกว่าระดับหน้าอก แต่สูงกว่าเอวของผู้ทำท่ามือ หรือเมื่ออยู่ในท่าพักมือตำแหน่งของมือจะอยู่ต่ำกว่าตัวส่งสัญญาณ (Transmitter) แต่เมื่อทำท่ามือตำแหน่งของมือควรอยู่สูงกว่าตัวส่งสัญญาณ (Transmitter) ซึ่งขอบเขตนี้สามารถกำหนดได้โดยผู้ใช้เองโดยกำหนดค่าขอบเขต z เพื่อเป็นตัวตัดสินว่ามือข้างใดอยู่ในบริเวณที่กำหนด



รูปที่ 3.3 ตำแหน่งการวางตัวส่งสัญญาณ (Transmitter)

หากมือข้างใดมีค่า z จากอุปกรณ์ติดตามการเคลื่อนที่น้อยกว่า ค่าขอบเขต z ที่กำหนดจะถือว่ามีแนวโน้มทำท่ามือ ในขั้นตอนนี้ใช้ประกอบกับขั้นตอนถัดไปเพื่อใช้ตัดสินว่าควรตัดสินท่ามือด้วยอัลกอริทึมใด

3.2.2 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่ (Segmentation using data from motion tracking sensors)

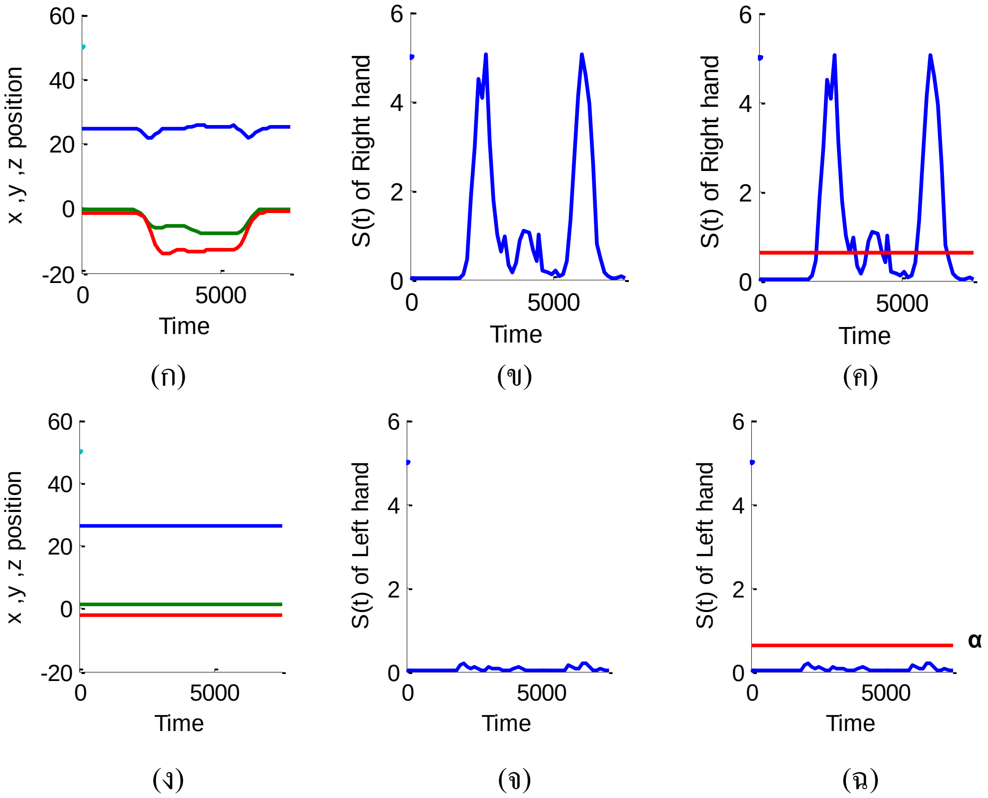
ขั้นตอนการสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่เป็นกระบวนการหาช่วงเวลาที่ตำแหน่งมือไม่มีการเปลี่ยนแปลง หรือมือไม่มีการเคลื่อนที่ โดยการหาอัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือจากค่าที่ได้จากอุปกรณติดตามการเคลื่อนที่ซึ่งถูกติดอยู่บนฝ่ามือของผู้ทำท่ามือ โดยจะใช้ค่าพิกัด ในแนวแกน X Y Z มาคำนวณหาอัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือโดยใช้สมการที่ (3.1) จากนั้นจะทำการเลือกช่วงเวลาที่มีอัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือมีค่าเข้าใกล้ศูนย์ แต่เนื่องจากการเคลื่อนที่โดยธรรมชาติของมนุษย์นั้นไม่สามารถทำให้การเปลี่ยนแปลงตำแหน่งได้ค่าศูนย์ หรือทำให้มือหยุดนิ่งได้โดยสมบูรณ์ ดังนั้นจึงนำค่า α เป็นตัวช่วยตัดสินว่าควรเลือกเวลาใด

$$S(t) = \begin{cases} 1, & t = 0 \\ |x_t - x_{t-1}| + |y_t - y_{t-1}| + |z_t - z_{t-1}|, & t \geq 1 \end{cases} \quad (3.1)$$

เมื่อ t คือ เวลา

$S(t)$ คือ อัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือ ณ เวลา t

x_t, y_t, z_t คือ ค่าตำแหน่งของมือบนแกน x, y, z ณ เวลา t ตามลำดับ



รูปที่ 3.4 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่

(ก) ตำแหน่งของมือขวาในแนวแกน X Y Z

(ข) อัตราการเปลี่ยนแปลงการเคลื่อนที่ของมือขวา

(ค) การตัดสินใจเลือกช่วงเวลาผลลัพธ์โดยใช้ค่า α ของมือขวา

(ง) ตำแหน่งของมือซ้ายในแนวแกน X Y Z

(จ) อัตราการเปลี่ยนแปลงการเคลื่อนที่ของมือซ้าย และ

(ฉ) การตัดสินใจเลือกช่วงเวลาผลลัพธ์โดยใช้ค่า α ของมือซ้าย

จากรูปที่ 3.5 เป็นตัวอย่างของการหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่ โดย (ก) และ (ง) เป็นค่าที่ได้จากอุปกรณ์ติดตามการเคลื่อนที่ซึ่งมีความสัมพันธ์กับเวลา โดยถูกติดบนมือขวาและมือซ้ายตามลำดับ ซึ่งค่าที่ได้นี้เมื่อนำมาทำการคำนวณคำนวณหาอัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือดังรูป (ข) และ (จ) แล้วจะถูกนำมาหาช่วงเวลาขณะทำท่ามือโดยการหาช่วงเวลาที่อัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือมีค่าน้อยกว่า α ดังรูปที่ (ค) และ (ฉ)

3.2.3 การสกัดหาช่วงเวลาขณะทำท่ามือด้วยถุงมือรับข้อมูล (Segmentation by using data from the data glove)

ขั้นตอนการสกัดหาช่วงเวลาขณะทำท่ามือด้วยถุงมือรับข้อมูลเป็นกระบวนการหาช่วงเวลาที่นิ้วมือไม่มีการเปลี่ยนแปลงไม่มีการยืด งอ หรือการการออกของนิ้วเพิ่มเติม ซึ่งเป็นช่วงเวลาที่นิ้วมือหยุดนิ่ง โดยใช้หลักการคล้ายกับขั้นตอนการสกัดหาช่วงเวลาขณะทำท่ามือด้วยอุปกรณ์ติดตามการเคลื่อนที่ซึ่งถูกบรรยายไปก่อนหน้า ในขั้นตอนการสกัดหาช่วงเวลาขณะทำท่ามือด้วยถุงมือรับข้อมูลนี้จะใช้สมการที่ (3.2) ในการหาอัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือ เพื่อหาจุดเวลาที่มีค่าอัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือเป็นจุดที่มีค่าต่ำสุดสัมพัทธ์ (Local Minimum)

$$V(t) = \begin{cases} 0, & t < 1 \\ \sum_{i=0}^{13} \left| x_{i,t} - x_{i,t-1} \right|, & t \geq 1 \end{cases} \tag{3.2}$$

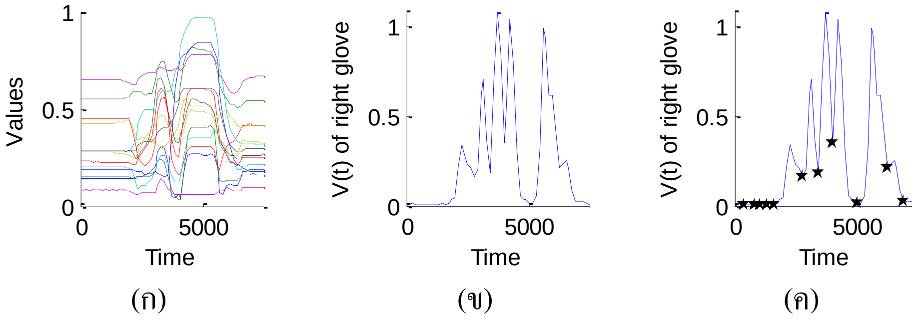
เมื่อ t คือ เวลา

$V(t)$ คือ อัตราการเปลี่ยนแปลงการเคลื่อนที่ของตำแหน่งของมือ ณ เวลา t

i คือ ลำดับของตัวรับรู้บนถุงมือ

$x_{i,t}$ คือ ค่าของตัวรับรู้ลำดับที่ i ณ เวลา t

จากรูปที่ 26 เป็นการนำค่าจากเซ็นเซอร์ทั้ง 14 ตัวบนถุงมือขวา (ก) ที่ผ่านการปรับให้มีค่าอยู่ในช่วง 0 – 1 มาทำการหาอัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือ (ข) และทำการหาจุดต่ำสุดสัมพัทธ์ของอัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือ (ค)

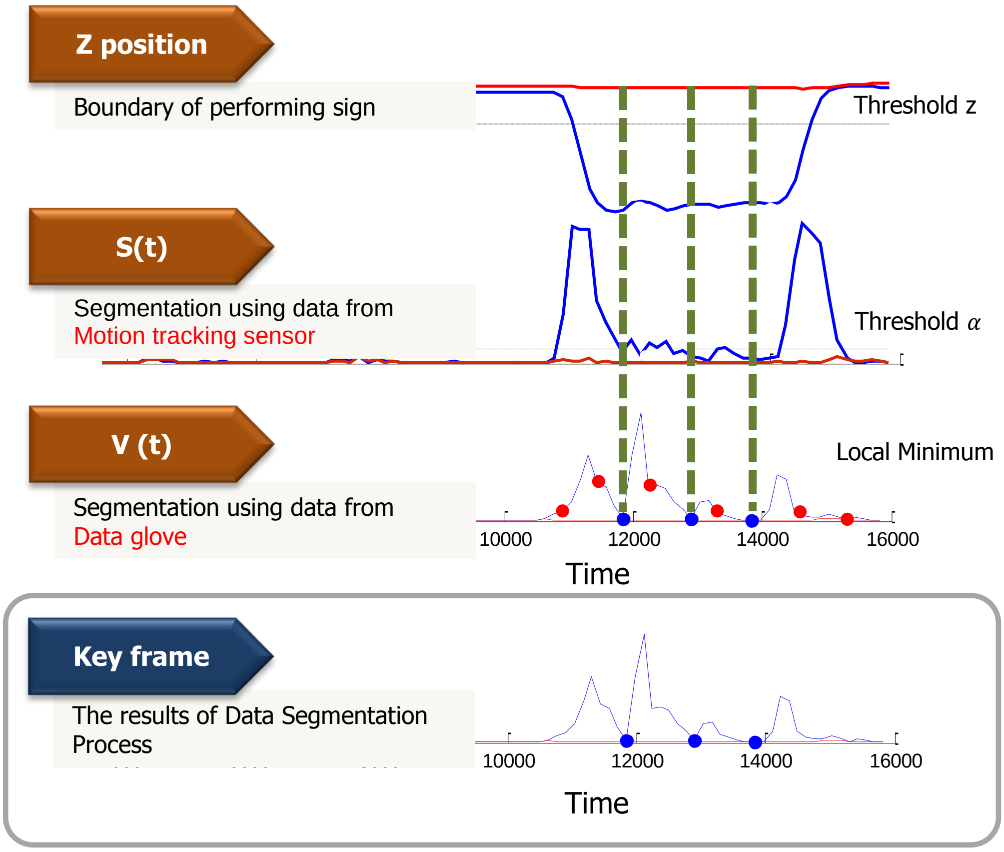


รูปที่ 3.5 การสกัดหาจุดต่ำสุดสัมพัทธ์ด้วยถุงมือรับข้อมูล

(ก) ข้อมูลที่ได้จากตัวรับรู้บนถุงมือขวาทั้ง 14 ตัวซึ่งถูกปรับค่าให้อยู่ในช่วง 0-1

(ข) อัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือ

(ค) จุดต่ำสุดสัมพัทธ์ของอัตราการเปลี่ยนแปลงของการเคลื่อนที่ของนิ้วมือ



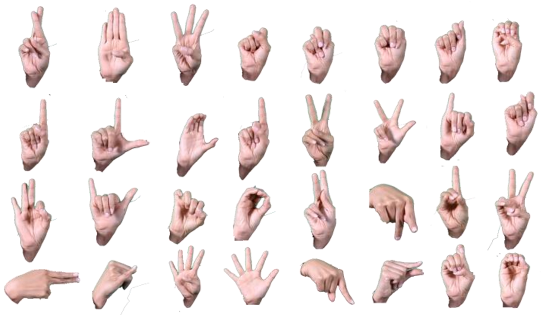
รูปที่ 3.6 การสกัดหาเฟรมสำคัญ หรือคีย์เฟรม

จากรูปที่ 3.6 หลังจากทำการคำนวณหา $S(t)$, $V(t)$ และรู้ค่า Z position แล้วนั้น จะทำการพิจารณาจุดที่เป็นจุดต่ำสุดสัมพัทธ์ (Local Minimum) ซึ่งอยู่บนค่า V(t)จากนั้นทำการพิจารณา ณ เวลาเดียวกัน หาก S(t) และค่า Z position มีค่าน้อยกว่า Threashold α และThreashold Z ตามลำดับ จะถือว่าเวลา ณ จุดนั้นเป็น เฟรมสำคัญ หรือ คีย์เฟรม ซึ่งจากรูปที่ 3.6 จะเหลือเพียง 3 จุดเท่านั้นที่กลายเป็นเฟรมสำคัญ หรือ คีย์เฟรม

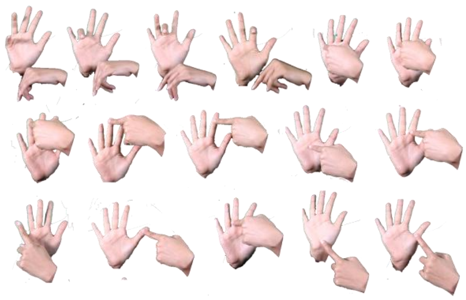
หลังจากที่สามารถหาเฟรมสำคัญได้แล้วนั้นจะทำการส่งค่าของอุปกรณ์ติดตามการเคลื่อนที่ และถุงมือรับข้อมูล ณ จุดเวลาที่ถูกตัดสินใจว่าเป็นเฟรมสำคัญไปในขั้นตอนถัดไปเพื่อทำการรู้จำประกอบกับ สถานะของมือทั้งสองข้างว่ามีตำแหน่งของมืออยู่ในบริเวณทำท่ามือหรือไม่ เพื่อการตัดสินใจเลือกใช้โมเดลการรู้จำที่จะใช้ในขั้นตอนถัดไป

3.3 การรู้จำท่ามือ (Sign Recognition)

ในขั้นตอนการรู้จำท่ามือ (Sign Recognition) นี้มีเพื่อทำการแปลงจากข้อมูลบนจุดเวลาที่เป็นเฟรมสำคัญของอุปกรณ์ทั้งสองชนิด ไปเป็นท่ามือ ท่ามือที่กล่าวถึงนี้คือรูปร่างของท่ามือพื้นฐานที่เมื่อนำมาแต่ละท่ามือมาประกอบกันอย่างมีลำดับแล้วจะสามารถแปลออกมาเป็นตัวอักษรหรือคำได้ ซึ่งท่ามือที่ใช้ในการสะกดนิ้วมือ สามารถแบ่งออกเป็นสองกลุ่มได้แก่ ท่ามือที่เกิดจากการทำท่ามือโดยใช้มือเพียงข้างเดียว ในที่นี้ผู้จัดทำวิทยานิพธ์ขอเรียกว่าท่ามือเดียว และนอกจากท่ามือเดียวแล้วท่ามืออีกกลุ่มคือท่ามือที่ต้องเกิดจากการใช้มือทั้งสองข้างทำเป็นสัญลักษณ์ ซึ่งในที่นี้ผู้วิจัยขอเรียกว่าท่าสองมือ ทำให้อัลกอริทึมที่ใช้ในการรู้จำถูกแบ่งออกเป็นสองส่วนคืออัลกอริทึมสำหรับท่ามือเดียว และอัลกอริทึมสำหรับท่าสองมือ

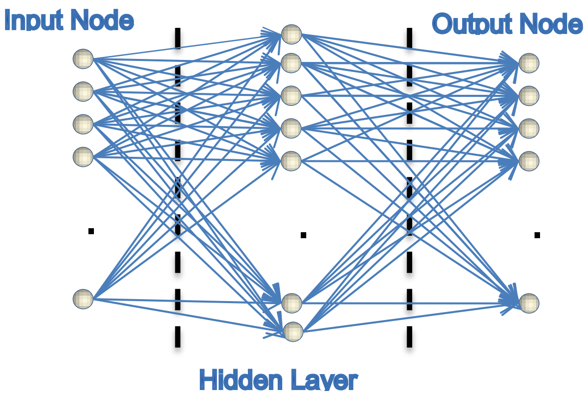


(ก)



(ข)

รูปที่ 3.7 รูปแบบทั้งหมดของท่ามือที่ต้องใช้สำหรับการสะกดนิ้วมือทั้ง 65 ตัวอักษร (ก) ท่ามือเดียวที่ใช้ในการสะกดนิ้วมือ และ (ข) ท่ามือสองมือที่ใช้ในการสะกดนิ้วมือ



รูปที่ 3.8 โครงข่ายประสาทเทียม

ในการรู้จำท่ามือของทั้งสองกลุ่มนี้ใช้ Elman Back Propagation Neural Network (ENN) ในการทำการรู้จำโดยทั้งสองกลุ่มจะมีจำนวนโหนดข้อมูลนำเข้า (Input Node) และโหนดผลลัพธ์

(Output Node) ที่แตกต่างกันตามคุณลักษณะของแต่ละกลุ่มท่ามือ และจำนวนรูปแบบที่มีของท่ามือภายในกลุ่มตามลำดับ และชั้นซ่อน (Hidden Layer) จะมีการคำนวณจำนวนของโหนดจาก จำนวนโหนดข้อมูลนำเข้ารวมกับจำนวนโหนดผลลัพธ์หารด้วยสอง

3.3.1 โครงข่ายประสาทเทียมสำหรับท่ามือเดียว

ในกลุ่มของท่ามือเดียวนี้สามารถแบ่งได้เป็น 32 รูปแบบ ทำให้จำนวนโหนดผลลัพธ์ของโครงข่ายประสาทเทียมมีจำนวนทั้งสิ้น 32 โหนดซึ่งการตัดสินว่าเป็นท่ามือเดียวรูปแบบใดจะตัดสินจากโหนดไหนมีค่าผลลัพธ์ออกมามากที่สุดจะเป็นผู้ชนะเช่น

โหนดลำดับที่	ผลลัพธ์	การจัดอันดับ
1	0.814723686393179	17
2	0.905791937075619	12
3	0.126986816293506	
…	...	…
30	0.546881519204984	
31	0.957506835434298	2
32	***0.964888535199277***	***1***

โหนดลำดับที่ 32 มีคะแนนมากที่สุดก็จะเป็นผู้ชนะ ดังนั้นผลลัพธ์ที่ได้จากอัลกอริทึมนี้คือรหัสของท่ามือ เป็นรหัสของท่ามือหมายเลข 32 และข้อมูลนำเข้าของโครงข่ายประสาทเทียมสำหรับท่ามือเดียวนี้มีจำนวนทั้งหมด 14 ค่าซึ่งได้จากตัวรับรู้บนถุงมือข้อมูล 14 ตัว

3.3.1 โครงข่ายประสาทเทียมสำหรับท่ามือสองมือ

ในกลุ่มของท่ามือสองมือนี้มีจำนวนโหนดผลลัพธ์เท่ากับจำนวนรูปแบบท่ามือสองมือ โดยมีจำนวนทั้งหมด 16 รูปแบบ ซึ่งการตัดสินว่าเป็นท่ามือรูปแบบใดใช้รูปแบบเดียวกับ โครงข่ายประสาทเทียมสำหรับท่ามือเดียว เพียงแต่รหัสของท่ามือของผลลัพธ์จะเริ่มจาก 33 และมีจำนวนโหนดอินพุทที่แตกต่างกันโดยมีการรับค่าจากตัวรับรู้จากทั้งบนถุงมือทั้งสองข้างจำนวน 28 ค่า และระยะห่างระหว่างตัวรับรู้ของอุปกรณ์ติดตามการเคลื่อนที่ 4 ตัวรับรู้ซึ่งค่าระยะห่างมีทั้งหมด 6 ค่า รวมทั้งหมดแล้วจึงมีโหนดนำเข้าทั้งหมด 34 โหนด

ตารางที่ 3.1ตารางแสดงไอดี และท่ามือที่ใช้ในการทำการวิจัย [8]

Class	**Sign**	**Class**	**Sign**	**Class**	**Sign**	**Class**	**Sign**
1		2		3		4	
5		6		7		8	
9		10		11		12	
13		14		15		16	
17		18		19		20	
21		22		23		24	
25		26		27		28	
29		30		31		32	
33		34		35		36	
37		38		39		40	
41		42		43		44	
45		46		47		48	

หลังจากได้ผลลัพธ์จากการทำการรู้จำท่ามือทั้งสองกลุ่มเป็นไอดีของท่ามือแล้ว จะมีการเก็บไอดีของท่ามือเข้าสู่ลิสต์ท่ามือ โดยมีการเริ่มเก็บตั้งแต่เริ่มจนจบการทำท่ามือของตัวอักษรหรือคำ จากนั้นจะนำลิสต์ ไอดีรูปแบบท่ามือส่งไปยังขั้นตอนถัดไปเพื่อทำการแปลออกมาเป็นตัวอักษรหรือคำในขั้นตอนการแปลท่ามือ (Translation) ซึ่งไอดีของท่ามือถูกกำหนดดังตารางที่ 3.1

3.4 การแปลท่ามือ (Translation)

หลักการของขั้นตอนนี้คือการนำลำดับท่ามือที่ได้จากขั้นตอนก่อนหน้ามาเปรียบเทียบกับทุกตัวอักษรที่เก็บไว้ในพจนานุกรม ซึ่งพจนานุกรมนี้ถูกจัดทำขึ้นโดยการแปลงตัวอักษรในภาษาไทยจำนวน 65 ตัวอักษรเป็นลำดับท่ามือและทำการเก็บไว้เพื่อใช้ในการเปรียบเทียบกับลิสต์ลำดับท่ามือของข้อมูลใหม่ที่เข้ามาใหม่ ซึ่งในการเปรียบเทียบนี้ทำการทดสอบใช้ 2 อัลกอริทึมได้แก่

3.4.1 การหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence)

การหาลำดับย่อยร่วมที่ยาวที่สุด (Longest Common Subsequence) อัลกอริทึมนี้จะทำการหาคำใด ๆ ในพจนานุกรมที่มีความยาวของสายอักขระย่อยที่ยาวที่สุดเป็นคำตอบ

ตารางที่ 3.2 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้ Longest Common Subsequence

FUNCTION TRANSLATION_LCSS(new_handshape_list)

```
set max to – infinity
loop i = 1 to number of word in dictionnary
        cost = LCSS(new_handshape_list,word_list[i]);
        if cost>max
                max = cost;
                clear list_of_answer_word;
                add word[i] to list_of_answer_word;
        else if cost = max
                add word[i] to list_of_answer_word;
        end if
end loop
return list_of_answer_word
```

ตารางที่ 3.3 โค้ดตัวอย่างการเขียนฟังก์ชั่น Longest Common Subsequence

FUNCTION LCSLENGTH(new_handshape_list, word_list[i])

```
Set m to length of new_handshape_list
Set n to length of word_list[i]
C = array(0..m, 0..n)
for i := 0..m
   C[i,0] = 0
for j := 0..n
   C[0,j] = 0
for i := 1..m
   for j := 1..n
      if X[i] = Y[j]
         C[i,j] := C[i-1,j-1] + 1
      else
         C[i,j] := max(C[i,j-1], C[i-1,j])
return C[m,n]
```

3.4.2 การวัดหาค่าความแตกต่างกันของสายอักขระสองชุด (Minimum Edit Distance)

การวัดหาค่าความแตกต่างกันของสายอักขระสองชุด (Minimum Edit Distance) เปรียบเทียบกันเพื่อเลือกคำที่มีค่า cost ในการเปลี่ยนน้อยที่สุด โดยมีการปรับค่า cost ของแต่ละวิธีการ(แทรก, ตัดออก, แทนที่)

ตารางที่ 3.4 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้ Minimum Edit Distance

FUNCTION TRANSLATION_MED(new_handshape_list)

```
set min to infinity
loop i = 1 to number of word in dictionnary
        cost = MED(new_handshape_list,word_list[i]);
        if cost<min
                max = cost;
                clear list_of_answer_word;
                add word[i] to list_of_answer_word;
        else if cost = min
                add word[i] to list_of_answer_word;
        end if
end loop
return list_of_answer_word
```

ตารางที่ 3.5 โค้ดตัวอย่างการเขียนฟังก์ชั่น translation โดยใช้ Minimum Edit Distance

FUNCTION MED(new_handshape_list,word_list[i])

```
set m to length of new_handshape_list
set n to length of word_list[i]
C = array(0..m, 0..n)
set C[0,0...n] to 0
set C[0...m,0] to 0

loop i = 1 to m
    for j = 1 to n
        set cost[0..2] to infinity
        //Match
      if X[i] = Y[j]
        cost[0] = 0
        //Substitution
      else if X[i] != Y[j]
        cost[0] = C[i-1,j-1] +Sub(X[i],Y[j])
        end if
        //Delete
        cost[1] = C[i-1,j-1] + del(X[i])
        //Insert
        cost[1] = C[i-1,j-1] + 1
        C[i,j] = min(cost[0],cost[1],cost[2]);
return C[m,n]
```

โดยการให้คะแนนค่า Cost ของแต่ละวิธีการได้มาจากการทดลองและการสังเกตลักษณะของผลลัพธ์จากขั้นตอนก่อนหน้า

1. การตัดออก(Delete)

   Cost = 0 ถ้า X [i-1] = X[i]

   Cost = 1 ถ้า X [i-1] ไม่เท่ากับ X[i]

2. การแทรก(Insert)

   Cost = 1 เสมอ

3. การแทนที่(Substitution)

   Cost = infinity ถ้า X[i] และ Y[j] อยู่ต่างกลุ่มท่ามือ เช่น

   X[i] เป็นสมาชิกในกลุ่มท่ามือที่ใช้ท่ามือเดียวในการทำท่ามือ แต่ Y[j] เป็นสมาชิกใกลุ่มท่ามือที่ใช้สองมือในการทำท่ามือ

   Cost = คะแนนการจัดอันดับของระยะห่างแบบยูคลิด (Euclidean Diatance) ระหว่าง X[i] และ Y[j] ซึ่งเกิดจากการนำข้อมูลของท่ามือในกลุ่มเดียวกันทั้งหมด

มาหาระยะห่างแบบยูคลิด แล้วทำการจัดเรียงอันดับความเหมือนของรูปร่างท่ามือจากมากไปน้อย โดยหากมีระยะห่างยิ่งน้อยหมายความว่าท่ามือมีความคล้ายคลึงกัน คะแนนที่ได้จะเข้าใกล้ 0 แต่หากมีระยะห่างยิ่งมากแสดงว่าท่ามือมีความคล้ายคลึงกันน้อย คะแนนที่ได้จะเข้าใกล้ 1

สรุปผลลัพธ์สุดท้ายที่ได้จากทั้งสองอัลกอริทึมนี้จะมีลักษณะเป็นเซตของคำ ซึ่งได้คะแนนที่ดีที่สุด ซึ่งเป็นผลลัพธ์ท้ายสุดของการรู้จำการสะกดนิ้วมือภาษามือไทย

บทที่ 4

การทดลอง และวิเคราะห์ผลการทดลอง

4.1 การเตรียมข้อมูลทดลอง

ในงานวิทยานิพนธ์นี้ได้ทำการทดลองโดยมีการเก็บข้อมูลจากล่ามผู้เชี่ยวชาญในการทำภาษามือโดยมีการเก็บการสะกดนิ้วมือระดับตัวอักขระจำนวน 4 เซต โดยแต่ละเซตประกอบไปด้วยการสะกดนิ้วมือของตัวอักษร สระ วรรณยุกต์ และสัญลักษณ์แต่ละตัว ซึ่งผู้จัดทำวิทยานิพนธ์ได้ทำการเขียนพัฒนาโปรแกรมด้วย Microsft Visual C++ และ OpenCV โดยนำ OpenCV มาเพื่อใช้ในการบันทึกภาพ เพื่อใช้ในการตรวจสอบความถูกต้อง และหาช่วงเวลาทำท่ามือที่แท้จริง

หลังจากที่ทำการเก็บข้อมูลแล้ว ขั้นตอนถัดมาก่อนนำข้อมูลมาใช้จะต้องมีการ Calibration ข้อมูลจากถุงมือให้อยู่ในช่วง 0 - 1 จากนั้นมีการตรวจสอบหาเฟรมสำคัญจากการสังเกตวิดีโอจากการจัดเก็บข้อมูลทั้งหมด

4.2 วิธีวัดผลและประเมินผลการทดลอง

วิธีวัดผลและประเมินผลการทดลองถูกแบ่งตามขั้นตอนการรู้จำการสะกดนิ้วมือไทย 3 ขั้นตอนหลัก คือ การตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ ขั้นตอนที่สองคือการตรวจสอบความถูกต้องแม่นยำในการรู้จำท่ามือ และขั้นที่สามคือความถูกต้องในการแปลท่ามือ โดยรายละเอียดของการวัดผลและประเมินผลการทดลองจะกล่าวในหัวข้อย่อยต่อไป

4.2.1 การตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ทำท่ามือ

ในการตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือนั้นผู้วิจัยได้ทำการตรวจสอบช่วงเวลาที่ผู้ทำท่ามือได้ทำท่ามือจากวิดีโอที่ได้ทำการบันทึกในขั้นตอนเก็บข้อมูล ซึ่งการตรวจสอบความถูกต้องใช้การวัดจากค่า Precision Recall และF-measure ในการตรวจสอบความถูกต้อง โดยมีการหาค่าต่าง ๆ ดังนี้

- *True positive (key frame)* คือ จำนวนเฟรมที่สกัดได้จากอัลกอริทึมแล้วเฟรมนั้นอยู่ในช่วงเวลาที่กำลังทำท่ามืออย่างถูกต้อง
- *False positive (unknown frame)* คือ จำนวนเฟรมที่สกัดได้จากอัลกอริทึมแล้วเฟรมนั้นไม่อยู่ในช่วงเวลาที่กำลังทำท่ามือ หรือ สกัดเฟรมผิด

- *False negative (miss)* คือ จำนวนเฟรมที่ถูกต้องที่ควรสกัดได้จากอัลกอริทึม แต่ไม่ได้ถูกสกัดออกมา

		Actual class	
		Key frame	Not Key frame
Predict class	Key frame	**True positive key frame**	**False positive unknown frame**
	Not Key frame	**False negative Miss**	**True negative Don't detect**

หลังจากทำการหาค่า True positive false positive และ false negative จะนำมาทำการคำนวณหาค่า Precision Recall และ F-measure ซึ่งค่าแต่ละค่าเป็นตัวบ่งบอกคความถูกต้องดังนี้

- ***Precision*** จะเป็นตัวบ่งบอกชี้ถึงคุณภาพในการกรองเฟรมที่ไม่ใช่เฟรมสำคัญออกของอัลกอริทึม ว่ามีความถูกต้องแม่นยำได้มากน้อยเพียงใดหากค่าที่ได้ยิ่งมากเข้าใกล้ 1 หรือมีค่าเท่ากับ 1 จะหมายถึงอัลกอริทึมสามารถกรองเฟรมที่ไม่ใช่เฟรมสำคัญออกไปได้ดี ซึ่ง Precision สามารถคำนวณได้โดยใช้สมการนี้

$$Precision = \frac{true\ positive}{true\ positive + false\ positive}$$

- ***Recall*** จะเป็นตัวบ่งชี้ถึงคุณภาพในการสกัดหาเฟรมสำคัญว่าสามารถสกัดเฟรมสำคัญได้ครบถ้วนสมบูรณ์หรือไม่ ค่าที่ได้ยิ่งมากเข้าใกล้ 1 หรือมีค่าเท่ากับ 1 จะหมายถึงอัลกอริทึมสามารถสกัดหาเฟรมสำคัญได้ครบถ้วน ซึ่ง Recall Precision สามารถคำนวณได้โดยใช้สมการนี้

$$Recall = \frac{true\ positive}{true\ positive + false\ negative}$$

- ***F-measure*** เป็นตัวบ่งชี้คุณภาพโดยรวมของอัลกอริทึมทั้งความสามารถในการกรองส่วนที่ไม่ใช่เฟรมสำคัญ และ ความสามารถในการสกัดหาเฉพาะส่วนที่เป็นเฟรมสำคัญ ซึ่งยิ่งค่า F-measure เข้าใกล้ 1 ยิ่งดี

$$F - measure = 2\ x\ \frac{Precision\ x\ Recall}{Precision + Recall}$$

เนื่องจากมีการใช้ Threshold α ดังนั้นจึงมีการหาค่า Threshold α ที่ดีที่สุดโดยการเลือกจาก Threshold α ที่ให้ค่า F-measure ที่เข้าใกล้ 1 มากที่สุด

4.2.2 การตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ

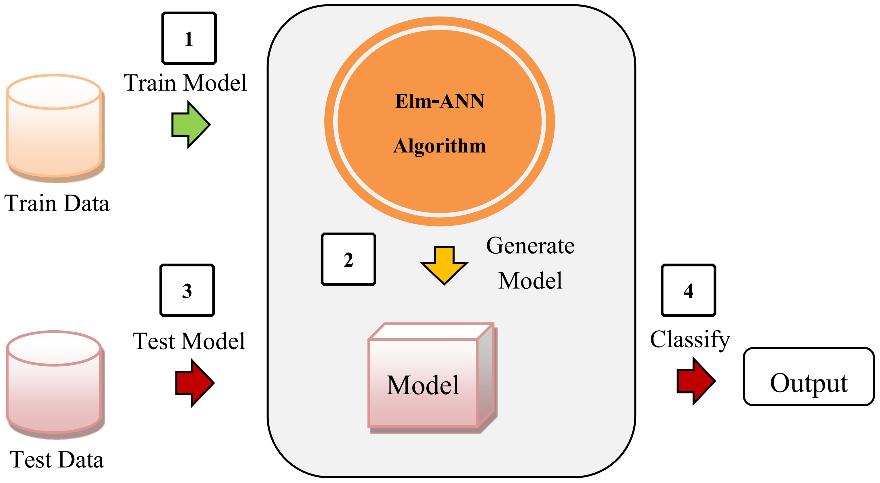
ในการตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ เป็นการตรวจสอบความถูกต้อง โดยใช้ตัววัด Accuracy เพื่อวัดประสิทธิภาพของการทำนายรูปแบบท่ามือของโครงข่ายประสาทเทียมทั้งสำหรับท่ามือเดียว และท่ามือสองมือ โดยมีวิธีการคำนวณ Accuracy คือ

		Actual class	
		Real Class	Other class
Predict class	Real Class	**True positive**	**False positive**
	Other class	**False negative**	**True negative**

โดยสูตรการคำนวณค่า Accuracy คือ

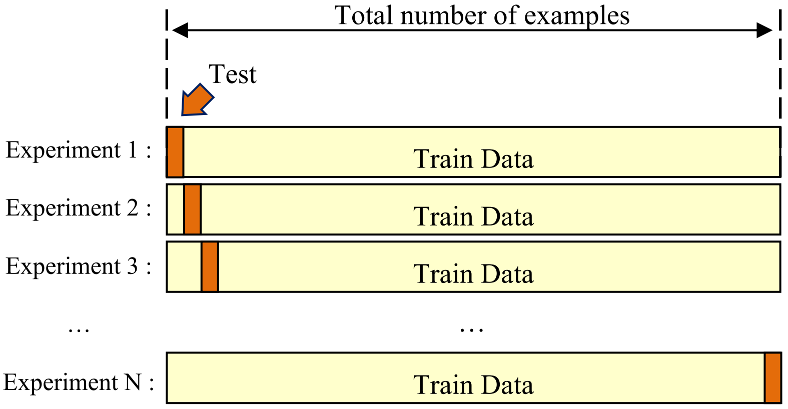
$$Accuracy = \frac{true\ positive + false\ positive}{true\ positive + true\ negative + false\ positive + false\ negative}$$

แต่ก่อนที่จะทำการวัดค่า Accuracy นั้น ข้อมูลที่เป็นเฟรมสำคัญหรือคีย์เฟรมทั้งหมดต้องทำการเตรียมเพื่อใช้สำหรับ การฝึกสอน (train) และทดสอบ (test) โครงข่ายประสาทเทียม โดยเริ่มจากนำข้อมูลฝึกสอนมาใช้สอนโมเดลการรู้จำ จากนั้นจะใช้ข้อมูลทดสอบเข้ามาทดสอบโมเดลที่ถูกสอนเพื่อหาผลลัพธ์ แล้วจึงนำผลลัพธ์ที่ได้ซึ่งเป็นการทำนายคลาสของรูปแบบท่ามือมาวัดค่า Accuracy ดังรูปที่ 4.1



รูปที่ 4.1 ขั้นตอนการฝึกสอน (Train) และทดสอบ (Test) โครงข่ายประสาทเทียม

การแบ่งข้อมูลเพื่อทดสอบนี้ใช้อัลกอริทึม Leave-one-out Cross Validation ซึ่งเป็นการสุ่มเลือกข้อมูลแบบความเที่ยงตรง K กลุ่ม เมื่อกำหนดให้ K มีค่าเท่ากับจำนวนแถวข้อมูลทั้งหมด (N) ดังรูปที่ 4.2



รูปที่ 4.2 การสุ่มเลือกข้อมูลแบบ Leave-one-out Cross Validation

การแบ่งข้อมูลเพื่อทดสอบความถูกต้องด้วย การสุ่มเลือกข้อมูลแบบ Leave-one-out Cross Validation เป็นรูปแบบที่เหมาะสมกับงานวิจัยนี้มากที่สุดเพราะจำนวนรูปร่างของมือในแต่ละกลุ่มมีจำนวนไม่เท่ากัน และปริมาณข้อมูลมีจำกัด ซึ่งการสุ่มเลือกข้อมูลแบบ Leave-one-out Cross Validation เหมาะสำหรับการวัดผลที่มีข้อมูลไม่มากนักเพราะการประมวลผลใช้เวลานาน แต่มีข้อดีที่การวัดผลทำได้แม่นยำเนื่องจากมีการทดสอบข้อมูลทุกจึงไม่มีการเบี่ยงเบนเกิดขึ้น

4.2.3 การตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ

การตรวจสอบของขั้นตอนนี้ทำการเน้นที่การหาวิธีการเลือกคำที่ถูกต้องจากพจนานุกรมซึ่งมีการเลือกทดสอบ 2 อัลกอริทึมเปรียบเทียบกัน คือ Longest Common Subsequence และ Minimum Edit Distance โดยทำการสุ่มเลือกตัวอย่างการสะกดนิ้วมือที่เก็บไว้ทั้งระดับคำ และตัวอักขระมาทำการทดสอบ ว่าสามารถแปลออกมาได้ถูกต้องหรือไม่ แต่เนื่องจากในขั้นตอนนี้การวัดผลค่อนข้างมีความเกี่ยวเนื่องกับขั้นตอนก่อนหน้าทั้งสองขั้นตอน จึงทำให้ผลที่ได้อาจจะมีประสิทธิภาพขึ้นอยู่กับความถูกต้องของสองขั้นตอนก่อนหน้าด้วย

ซึ่งสิ่งที่ใช้วัดประสิทธิภาพในครั้งนี้จะใช้ค่า Accuracy ในการวัดผล แต่เนื่องจากผลลัพธ์ที่ออกมามีลักษณะเป็นเซตของคำ ดังนั้นการตัดสินว่าผลลัพธ์นั้นถูกหรือผิด จะถูกตัดสินโดย ถ้าในเซตของคำตอบมีสมาชิกตัวใดตัวหนึ่งเป็นคำตอบที่ถูก จะถือว่าอัลกอริทึมสามารถทำนายถูก จัดให้

ค่าที่ได้เป็น true แต่หากเซตของคำตอบไม่มีสมาชิกตัวใดตรงกับคำตอบที่ถูกแล้วจะถือว่าอัลกอริทึมทำนายผิดจัดให้ค่าที่ได้เป็นเป็น false

$$Accuracy = \frac{true}{true + false}$$

ในลำดับถัดไปจะเป็นผลการทดลองทั้งหมด โดยใช้วิธีการวัดผลและประเมินผลที่ได้นำเสนอในหัวข้อนี้

4.3 ผลการทดลองและวิเคราะห์ผลการทดลอง

ในหัวข้อนี้ จะนำเสนอผลการทดลองตามลำดับตั้งแต่การตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ ต่อด้วยการตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ และสุดท้ายคือการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ รวมไปถึงการวิเคราะห์ผลการทดลองทั้งหมดของแต่ละขั้นตอน โดยผลการทดลองทั้งหมดมีดังต่อไปนี้

4.3.1 ผลการตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ

ผลการทดลองของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือเป็นดังตารางที่ 4.1

ตารางที่ 4.1 ผลการทดลองของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ

Threshold α	**Precision**	**Recall**	**F – measure**
0.3	0.67	0.84	0.75
0.4	0.8	0.83	0.82
0.5	0.79	0.84	0.82
***0.6**	***0.88***	***0.85***	***0.86***
0.7	0.85	0.84	0.85

จากตารางผลการทดลอง ตารางที่ 4.1 เมื่อค่า α เท่ากับ 0.6 เป็นค่าที่ดีที่สุดที่ส่งผลให้อัลกอริทึมให้ค่า Recall ซึ่งเป็นสิ่งที่บ่งบอกได้ว่าอัลกอริทึมสามารถสกัดเฟรมสำคัญได้ถึง 85 เปอร์เซนต์ จากจำนวนเฟรมสำคัญทั้งหมดที่ควรสกัดได้ และค่า Precision ซึ่งเป็นสิ่งที่บ่งบอกได้ว่าอัลกอริทึมการสกัดเฟรมสำคัญมีความสามารถในการกรองเฟรมที่ไม่ใช่เฟรมสำคัญได้สูงถึง 88

เปอร์เซนต์โดยเทียบจากจำนวนเฟรมทั้งหมดที่ถูกอัลกอริทึมนี้สกัดมา ซึ่งการเลือก ค่า α ที่ดีที่สุดนั้นถูกเลือกโดยตัดสินใจจากค่า F-measure ซึ่งเป็นค่าที่คิดจากค่า Precision และ Recall

ลำดับถัดไปจะนำเสนอผลการวัดความสามารถในการตรวจสอบความถูกต้องของการรู้จำท่ามือ ซึ่งเป็นขั้นตอนที่สองของการรู้จำท่ามือ

4.3.2 ผลการตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ

ผลการทดลองของการวัดความสามารถในการรู้จำของโมเดลโครงข่ายประสาทเทียมสำหรับการรู้จำท่ามือเดียว และการรู้จำท่ามือสองมือโดยการวัดค่า Accuracy ซึ่งเป็นเครื่องมือวัดที่สามารถบ่งบอกได้ว่าโมเดลนั้นทำนายผลลัพธ์ได้ถูกต้องมากน้อยเพียงใด ซึ่งในการทดลองนี้มีการสุ่มข้อมูลสำหรับฝึกสอนโมเดล และ ทดสอบโมเดล ด้วยการสุ่มเลือกข้อมูลแบบ Leave-one-out Cross Validation โดยให้ผลการทดลองดังตารางที่ 4.2

ตารางที่ 4.2 ผลการวัดค่า Accuracy ของโมเดลโครงข่ายประสาทเทียมสำหรับการรู้จำท่ามือเดียว และท่ามือสองมือ

Model	Type of Hand	Input Data	Accuracy
1	One hand	- 14 data points from data glove	0.79
2	Two hands	- 28 data points from 2 data gloves	0.44
3	Two hands	- 28 data points from 2 data gloves - 6 distance values from a pair of 4 motion sensor’s positions.	0.72

จากผลการทดลองดังตารางที่ 4.2 มีการทดลองสร้างโมเดลทั้งหมด 3 โมเดล ดังนี้

โมเดลที่ 1 เป็นการทดสอบโมเดลสำหรับท่ามือเดียว ที่มีข้อมูลนำเข้าเป็นค่าจากตัวรับรู้บนถุงมือข้างเดียว 14 ค่า ซึ่งให้ค่า Accuracy ซึ่งเป็นตัวบ่งความความแม่นยำในการทำนายของโมเดลนี้โดยเฉลี่ยได้ 0.79 ซึ่งแสดง Confusion Metrix ดังรูปที่ รูปที่ 4.3 Confusion metrix of Model 1

true class \ Predict Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	29	1	0	0	1	0	0	0	0	0	0	0	0	0
5	0	0	0	2	0	0	0	0	0	0	0	5	0	0	0	0	0	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	8	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
7	0	0	0	33	0	2	0	1	5	0	0	0	0	0	0	0	0	0	1	0	0	23	5	0	0	0	6	0	0	0	0	0
8	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	84	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	2	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	130	0	0	1	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
13	0	0	0	0	0	0	0	0	3	0	0	0	54	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	1	0	0	0	0	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	1	0	0	0	0	0	33	0	0	0	0	0	62	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	49	0	0	0	0	0	5	0	0	0	7	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	6	0	0	9	0	0	0	0	0	0	0	0	87	0	1	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	48	1	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	208	0	0	0	0	0	0	0	0	9	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	1	10	0	0	0	0	0	0	0	0	15	3	15
22	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	258	20	0	0	0	0	0	0	2	0	0
23	0	2	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	127	0	0	0	0	18	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	23	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	15	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	27	0	3	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0	15	10	0	0	12	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	19

รูปที่ 4.3 Confusion metrix of Model 1

จาก Confusion metrix ในรูปที่ 4.3 เนื่องจากมีรูปแบบท่ามือจำนวนมากอีกทั้งท่ามือมีความคล้ายคลึงกันอย่างมาก ดังนั้นจึงส่งผลให้ต่อความแม่นยำในการทำนายรูปแบบท่ามือ โดยท่ามือที่มีความคล้ายคลึงกันอย่างมากที่และส่งผลต่อการรู้จำดังตารางที่ 4.3 กลุ่มแรกรหัสท่ามือหมายเลข 23 มีการทำนายผิดเป็นรหัสท่ามือหมายเลข 28 ซึ่งทั้งสองท่ามือมีความคล้ายคลึงกัน ดังแสดงในตารางที่ 4.3 แสดงกลุ่มรหัสท่ามือที่มีความคล้ายคลึงกันตัวอย่างที่ 1 และ กลุ่มรหัสท่ามือหมายเลข 4, 9, 19 และ20 ซึ่งมีความคล้ายคลึงกันอย่างมากเนื่องจากมีการงอนิ้วทั้ง 5 นิ้วในลักษณะที่แตกต่างกันเพียงเล็กน้อย แม้กระทั่งมนุษย์เองสามารถสับสนได้แสดงในกลุ่มรหัสท่ามือมีความคล้ายคลึงกันตัวอย่างที่ 2 แต่โมเดลนี้ก็สามารถแยกท่ามือที่มีความแตกต่างอย่างชัดเจนเช่น รหัสท่ามือหมายเลข 1 และ 25 ซึ่งรหัสหมายเลขท่ามือต่าง ๆ นี้สามารถอ้างอิงได้จากตารางที่ 3.1 ในหน้าที่ 34

ตารางที่ 4.3 ตัวอย่างผลการทำนายที่ผิดพลาดของโมเดลที่ 1

<table>
<tr><td>
23</td><td>
28</td><td>
4</td><td>
9</td><td>
19</td><td>
20</td></tr>
<tr><td colspan="2">กลุ่มรหัสท่ามือมีความคล้ายคลึงกันตัวอย่างที่ 1</td><td colspan="4">กลุ่มรหัสท่ามือมีความคล้ายคลึงกันตัวอย่างที่ 2</td></tr>
</table>

โมเดลที่ 2 เป็นการทดสอบโมเดลสำหรับท่ามือสองมือ ที่มีข้อมูลนำเข้าเป็น ค่าจากตัวรับรู้บนถุงมือทั้งสองข้างจำนวน 28 ค่า ซึ่งให้ค่า Accuracy ซึ่งเป็นตัวบ่งความความแม่นยำในการทำนายของโมเดลนี้โดยเฉลี่ยได้ 0.44

true class \ Predict Class	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
33	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	1	52	0	0	7	0	4	0	0	0	4	2	0	0	8	0
35	0	0	4	17	5	0	0	0	0	0	16	0	0	7	0	5
36	0	1	0	0	0	0	0	0	0	0	5	0	0	0	0	0
37	0	0	0	0	76	0	10	0	0	0	39	0	0	0	0	0
38	0	6	3	0	8	5	0	0	3	0	12	8	0	0	0	0
39	0	0	0	9	8	0	3	0	0	0	0	0	0	3	6	0
40	0	0	0	0	0	1	0	0	0	0	5	0	0	0	0	0
41	0	0	0	0	0	0	6	0	5	0	13	0	3	0	0	0
42	0	1	1	0	5	0	0	0	0	0	3	0	0	0	0	0
43	0	6	5	0	6	0	1	15	0	0	143	0	0	4	0	5
44	0	6	0	0	0	0	5	0	0	0	8	22	0	0	2	3
45	0	0	0	0	11	0	0	0	2	0	0	0	2	0	7	0
46	0	0	1	0	0	0	0	0	0	0	18	0	0	0	0	0
47	0	3	14	0	0	0	0	0	0	0	7	0	0	0	0	0
48	0	0	1	0	9	3	0	0	0	0	4	1	0	0	0	0

รูปที่ 4.4 Confusion metrix of Model 2

โมเดลที่ 3 เป็นการทดสอบโมเดลสำหรับท่ามือสองมือ ที่มีข้อมูลนำเข้าประกอบด้วย ค่าจากตัวรับรู้บนถุงมือทั้งสองข้างจำนวน 28 ค่า, ค่าระยะห่างระหว่างตำแหน่งหลังมือของมือซ้ายไปถึงตำแหน่งหลังมือของมือขวา, ค่าระยะห่างระหว่างตำแหน่งหลังมือของมือซ้ายไปถึงตำแหน่งข้อมือของมือขวา, ค่าระยะห่างระหว่างตำแหน่งหลังมือของมือขวาไปถึงตำแหน่งข้อมือของมือซ้าย และระยะห่างระหว่างตำแหน่งของข้อมือขวาไปถึงตำแหน่งข้อมือของมือซ้าย โดยที่ค่าตำแหน่งต่าง ๆ ได้มาจากตัวรับรู้ 4 ตัวของอุปกรณ์ติดตามการเคลื่อนที่ โดยผลที่ได้จากโมเดลนี้ให้ค่า Accuracy เท่ากับ 0.72

true class \ Predict Class	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
33	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	48	0	0	0	0	0	0	0	0	30	0	0	0	0	0
35	0	0	27	7	20	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
37	0	1	0	0	95	0	0	1	8	0	17	0	0	3	0	0
38	0	0	0	2	5	31	0	0	0	0	6	0	0	0	1	0
39	0	0	0	0	7	0	17	0	8	0	0	0	0	0	0	0
40	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0
41	0	0	2	0	0	3	0	0	22	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0
43	0	1	0	0	30	0	0	11	0	1	140	0	2	0	0	0
44	0	0	0	0	0	0	0	0	0	0	3	43	0	0	0	0
45	0	0	0	0	0	6	0	0	0	0	0	0	5	0	0	0
46	0	0	0	5	0	0	0	0	0	0	0	0	0	14	0	0
47	0	0	0	0	0	0	0	0	0	0	8	0	0	0	16	0
48	0	0	0	0	0	0	3	0	5	0	0	0	0	0	0	10

รูปที่ 4.5 Confusion metrix of Model 3

จาก Confusion metrix ของโมเดลที่ 2 และ 3 ดังแสดงดังรูปที่ 4.4 และ 4.5 จะสังเกตได้ว่าจำวนท่ามือที่ทำนายถูกต้องมีมากขึ้นเมื่อพิจารณาจากผลรวมในแนวเส้นทะแยงมุม แต่อย่างไรก็ตามเนื่องจากท่ามือมีความใกล้เคียงซึ่งทำให้ความยากต่อการสร้างโมเดลการรู้จำ ซึ่งรหัสหมายเลขท่ามือต่าง ๆ นี้สามารถอ้างอิงได้จากตารางที่ 3.1 ในหน้าที่ 34 แต่เมื่อพิจารณาที่หมายเลข 48 ซึ่งเป็นท่ามือที่มีลักษณะการชี้ที่บริเวณข้อมือ จึงมีความแตกต่างจากท่ามืออื่น ๆ อย่างชัดเจน และโมเดลที่ 2 สามารถทำนายได้ถูกต้องทั้งหมด

จากผลการทดลองทำให้งานวิจัยนี้ตัดสินใจเลือกใช้โมเดลที่ 1 และโมเดลที่ 3 เพื่อใช้ทำการรู้จำท่ามือเดียว และท่ามือสองมือตามลำดับ โดยโมเดลทั้งสองโมเดลนี้จะถูกใช้ในขั้นตอนถัดไป ซึ่งเป็นการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ

4.3.3 ผลการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือ

ในขั้นตอนการตรวจสอบการแปลท่ามือนี้ได้ทำการทดลองโดย ทดสอบในระดับตัวอักษรจำนวน 65 ตัว ซึ่งมีข้อมูลจำนวน 4 ชุด ซึ่งจะถูกนำไปเข้าสู่ 2 กระบวนการก่อนหน้าได้แก่ ขั้นตอนการสกัดหาเฟรมสำคัญซึ่งเป็นช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือที่มีความหมาย และนำค่าจากอุปกรณ์ต่าง ๆ ได้แก่ ถุงมือ และอุปกรณ์ติดตามการเคลื่อนที่มาเข้าสู่ขั้นตอนการรู้จำท่ามือ เพื่อแปลงเป็นท่ามือ จากนั้นนำผลลัพธ์ซึ่งเป็นลำดับของท่ามือเข้าสู่กระบวนการแปลท่ามือออกมาเป็นตัวอักษร โดยมีการวัดผลด้วยค่า Accuracy ดังที่กล่าวไว้ในหัวข้อ 4.2.3 โดยที่ผลลัพธ์ที่ได้จากขั้นตอนนี้จะมีลักษณะเป็นเซตของคำตอบ หากสมาชิกตัวใด ๆ ของเซตคำตอบตรงกับคำตอบที่ถูกต้องแล้ว จะถือว่าผลลัพธ์ที่ได้ให้ผลถูกต้อง ในทางตรงกันข้ามหากไม่มีสมาชิกตัวใดของเซตคำตอบตรงกับคำตอบที่ถูกต้องแล้วนั้น จะถือว่าผลลัพธ์ที่ได้ให้ผลไม่ถูกต้อง ซึ่งในขั้นตอนการทดลองนี้ทำการทดลองเพื่อหาอัลกอริทึม และพารามิเตอร์ ที่ดี โดยมีการทดสอบอัลกอริทึมทั้งหมด 3 รูปแบบได้แก่

1. การใช้อัลกอริทึม Longest common subsequence
2. การใช้อัลกอริทึม Minimum edit distance โดยมีการกำหนดค่าพารามิเตอร์ทุกวิธีการ(การแทรก,การแทนที่,การลบ)มีค่าเท่ากัน คือ 1

   การใช้อัลกอริทึม Minimum edit distance โดยมีการกำหนดค่าพารามิเตอร์ โดยใช้ฟังค์ชั่นของการให้คะแนน

ผลการวัดความถูกต้องแม่นยำของขั้นตอนนี้เป็นดังตารางที่ 4.4 ซึ่งมีการเปรียบเทียบผลลัพธ์จากทั้ง 3 รูปแบบด้วยค่า Accuracy

ตารางที่ 4.4 ผลการตรวจสอบความถูกต้องของขั้นตอนการแปลท่ามือเป็นตัวอักษร

	Algorithm	**Configuration**	**Accuracy of Alphabets**
1	LCSS	- none	0.69
2	MED	- Cost delete = 1 - Cost Insert = 1 - Cost Substitution = 1	0.65
3	MED with weight function	- Cost delete = 0 if X[i] = X[i-1] = 1 if X[i] != X[i-1] - Cost Insert = 1 - Cost Substitution = infinity if Set of X is not subset of Set of Y = weight calculate from Euclidean distance	0.65

จากตารางที่ 4.4 จะเป็นได้ว่าอัลกอริทึมที่นำเสนอในงานวิทยานิพธ์ชิ้นนี้สามารถทำการรู้จำได้ครอบคลุมถึง 65 ตัวอักษรซึ่งถือได้ว่าเป็นงานแรกของไทยที่ทำการรู้จำตัวอักษรได้ครอบคลุมทั้งหมด และยังสามารถให้ค่า Accuracy ที่ดีที่สุดได้สูงเกือบ 70 เปอร์เซ็นต์ ซึ่งถือได้ว่าสูงมากสำหรับงานวิจัยที่มีการทำการรู้จำตัวอักษรจำนวนมากที่มีอยู่ในปัจจุบันทั่วโลก เพราะแม้แต่การรู้จำการสะกดนิ้วมือของภาษามืออเมริกันซึ่งมีจำนวนตัวอักษรที่น้อยกว่าก็ยังไม่มีงานวิจัยไหนที่ทำจนครบทุกตัวอักษร หลังจากผลการทดลองเบื้องต้นด้วยค่า Accuracy ของแต่ละอัลกอริทึมในลำดับถัดไปจะพิจารณาและวิเคราะห์ผลลัพธ์ หรือเซตของคำตอบที่ได้จากแต่ละอัลกอริทึม ซึ่งให้ผลลัพธ์ดังตารางที่ 4.5 ซึ่งยกตัวอย่างมาเพียงการรู้จำตัวอักษร ‘ล’ ซึ่งหมายเลขในคอลัมน์ Sign ในทุกคอลัมน์คือ รหัสท่ามือที่ถูกกำหดตามตารางที่ 3.1

จากตารางที่ 4.5 ตัวอักษร ‘ล’ นั้นเกิดจากการทำท่ารหัสท่ามือหมายเลข 17 เพียงท่ามือเดียวแต่อย่างไรก็ตามเนื่องจากการสกัดข้อมูลช่วงเฟรมสำคัญ การทำให้คอมพิวเตอร์ตัดสินใจได้ว่าช่วงเวลาใดเป็นช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือทำได้ยากดังนั้นแล้วการสกัดเฟรมสำคัญจึงมีการสกัดเฟรมสำคัญ ซ้ำออกมามากกว่า 1 เฟรม และเมื่อลำดับท่ามือมาเข้าสู่ขั้นตอนการรู้จำท่ามือจะได้ลำดับท่ามือคือ 22, 17, 17, 17 ซึ่งจะถูกส่งเข้าสู่อัลกอริทึม LCSS และ MED

ตารางที่ 4.5 ตัวอย่างผลการแปลตัวอักษร ‘ล’

Input		Segment and recognition	OUTPUT					
			LCSS		MED		MED with weight function	
Alphabet	Sign	Sign	Alphabet	Sign	Alphabet	Sign	Alphabet	Sign
ล	17	22,17,17,17	ข	10,22	ล	17	ล	17
			ช	15,22,21	พ	17,22		
			ญ	18,22				
			ฎ	14,22				
			ฌ	19,22				
			ถ	11,22				
			ธ	11,15,22				
			ป	13,22				
			ฝ	16,22				
			ล	17				
			พ	17,22				
			ศ	12,22				
			ฮ	15,22				

จากการวัดในระดับตัวอักษรดังตารางที่ 4.4 นั้นแสดงให้เห็นว่าถึงแม้อัลกอริทึม LCSS จะให้ผลการทดสอบที่มีความแม่นยำมากกว่า แต่ก็เกิดจากการที่ LCSS มีการเลือกทุกตัวอักษรที่มีท่ามือตรงกับท่ามือที่ผ่านขั้นตอนการรู้จำมาจำทำให้เซตของทำตอบมีขนาดใหญ่ แต่อัลกอริทึม MED จะทำการเลือกเฉพาะบางตัวอักษรที่มีความใกล้เคียงโดยมีการพิจารณาลำดับของท่ามือด้วย ดังนั้นจึงมีผลลัพธ์ที่ใกล้เคียงมากที่สุด

ดังนั้น จากผลการทดลองเมื่อพิจารณาโดยภาพรวมแล้ว หากเลือกอัลกอริทึมที่เหมาะสมที่สุดแล้ว Longest common subsequence ถึงแม้จะให้ค่า Accuracy ดีกว่า แต่ผลลัพธ์จากอัลกอริทึม Minimum Edit Distance มีความใกล้เคียงสิ่งที่ถูกมากกว่า โดยโมเดลที่ดีที่สุดของ Minimum edit distance ที่ให้ผลลัพธ์ที่เหมาะสมใกล้เคียงคำตอบที่ถูกต้องมากที่สุดคือ Minimum edit distance with weight function ดังนั้นงานวิจัยนี้จึงเลือกใช้ อัลกอริทึม Minimum edit distance with weight function

4.3.4 การเปรียบเทียบระหว่างอัลกอริทึมที่นำเสนอในงานวิทยานิพนธ์ และ Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform [15]

เพื่อทำการทดสอบเปรียบเทียบกับงานวิจัย Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform (Phitakwinai, Auephanwiriyakul, และ Theera-Umpon, 2008) ซึ่งทำการรู้จำท่ามือจำนวนเพียง 15 ท่ามือดังรูปที่ 2.12 ซึ่งเป็นการเลือกตัวอักษรเพียงบางส่วนที่เป็นท่ามือซึ่งใช้มือเพียงข้างเดียว และไม่มีการใช้ท่ามือมากกว่าหนึ่งท่ามือในการแสดงหนึ่งตัวอักษร โดยมีผลลัพธ์ดังตารางที่ 4.6 ซึ่งจะเห็นได้ว่า อัลกอริทึมของงานวิจัยนี้สามารถให้ผลลัพธ์ที่ดีกว่าทั้ง ๆ ที่โมเดลที่นำมาทดสอบถูกสร้างสำหรับจำนวนคลาสของท่ามือที่มีมากกว่า ซึ่งส่งผลให้สามารถทำนายคลาสอื่นนอกเหนือจาก 15 ท่ามือที่ใช้เปรียบเทียบ

ตารางที่ 4.6 ผลการเปรียบเทียบ Accuracy ของอัลกอริทึมที่นำเสนอกับงานวิจัย [15]

Our Propose Algorithm	**Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform [15]**	
automatic translation system	automatic translation system	manual translation system
80.58%	79.45%	85.67%

สรุปผลการเปรียบเทียบงานวิจัยจากตารางที่ 4.6 จะเห็นได้ว่า อัลกอริทึมของงานวิจัยนี้สามารถให้ผลลัพธ์ที่ดีกว่า Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform โดยที่อัลกอริทึมนี้ทำการเลือกคีย์เฟรมโดยอัตโนมัติ เมื่อเทียบกับผลลัพธ์ที่ดีที่สุดของงานวิจัย Thai Sign Language Translation Using Fuzzy C-Means and Scale Invariant Feature Transform ซึ่งผลลัพธ์ที่ดีที่สุดจะต้องมาจากการเลือกคีย์เฟรมแบบ manual และในลำดับถัดไปจะเป็นการทดลองเพิ่มเติมซึ่งถือได้ว่าอยู่นอกเหนือขอบเขตของการวิจัย ซึ่งเป็นการทดสอบอัลกอริทึมในการแปลระดับคำ

4.3.5 ผลการทดลองเพิ่มเติมเพื่อทำการรู้จำการสะกดนิ้วมือในระดับคำเบื้องต้น

นอกจากในการทดสอบระดับตัวอักษรแล้ว ผู้ทำวิทยานิพนธ์ได้ทำการทดลองเพิ่มเติมเพื่อทดสอบหากนำอัลกอริทึม Longest common subsequence และ Minimum edit distance ไปใช้ทดสอบในระดับคำ โดยทำการเปลี่ยนจากเปรียบเทียบกับตัวอักษรเป็นการเปรียบเทียบกับคำใน

พจนานุกรมซึ่งเป็นจำนวน 1962 คำซึ่งมากจากตัวอย่างการสะกดนิ้วมือจากปธานุกรมภาษามือ แต่อย่างไรก็ตาม การกำหนดค่าใช้จ่ายสำหรับวิธีการ Substitution แล้วนั้น การทดลองนี้อาจจะไม่สามารถให้ค่าคะแนนที่ใกล้เคียงได้เนื่องจากปริมาณข้อมูลในการทดลองอาจจะมีจำนวนไม่เพียงพอเพราะมีข้อจำกัดทางงบประมาณและเวลา ดังนั้นแล้วหากต้องการปรับให้อัลกอริทึมมีความแม่นยำขึ้นนั้น อาจจะมีการทดลองเพิ่มเติมในงานวิจัยต่อไปได้ ซึ่งในตารางที่ 4.7 จะเป็นผลการวัดค่า Accuracy ที่ได้ทำการทดสอบเพิ่มเติมในเบื้องต้น และ ตารางที่ 4.8, ตารางที่ 4.9 และตารางที่ 4.10 จะเป็นการแสดงตัวอย่างผลลัพธ์บางส่วนของแต่ละอัลกอริทึมได้แก่ Longest common subsequences, Minimun Edit Distance และMinimun Edit Distance with weight function ตามลำดับ โดยรหัสหมายเลขท่ามือซึ่งถูกแสดงในคอลัมน์ Sequence of sign จากทั้งสามตารางนั้นมากจากรหัสหมายเลขท่ามือซึ่งถูกกำหนดตามตารางที่ 3.1 ในหน้าที่ 34

ตารางที่ 4.7 ผลการทดลองเพิ่มเติมเพื่อตรวจสอบความถูกต้องของการแปลท่ามือในระดับคำ

	Algorithm	**Configuration**	**Accuracy of Words**
1	LCSS	- none	0
2	MED	- Cost delete = 1 - Cost Insert = 1 - Cost Substitution = 1	0.14
3	MED with weight function	- Cost delete = 0 if X[i] = X[i-1] = 1 if X[i] != X[i-1] - Cost Insert = 1 - Cost Substitution = infinity if Set of X is not subset of Set of Y = weight calculate from Euclidean distance	0.32

ตารางที่ 4.8 ตัวอย่างผลการแปลคำของ Longest common subsequences

Input		Segment and recognition	OUTPUT(LCSS)	
Word	Sequence of sign	Sequence of sign	Word	Sequence of sign
ฝัน	16,22,45,19	16,14,38,15,19	รถโดยสารประจำทาง	20,2,11,22,9,14,18,12,39,2,13,22,2,38,5,34,11,15,39,19,30
			วัดพระเชตุพนวิมลมังคลาราม	30,3,45,14,13,2,38,44,15,21,22,11,42,13,19,3,20,31,…,4
			วัดพระธาตุพนม	15,3,45,14,13,2,38,11,15,22,39,11,42,13,19,4
ปี	13,12,20	11,20,20,20	เจ้าหน้าที่บันทึกวิดีโอ	27,44,5,41,39,15,19,41,39,11,15,20,32,10,3,20,31,14,…,6
			สาธารณรัฐฟิลิปปินส์	28,12,39,11,15,22,39,2,19,22,39,2,19,22,2,45,11,…,6
			สาธารณรัฐอินโดนีเซีย	26,12,39,11,15,22,39,2,19,22,2,45,11,23,6,20,31,…,18
ผี	13,23,20	13,20,20,14	เครื่องพิมพ์ดีดไฟฟ้า	24,44,10,23,2,20,29,37,6,19,30,13,30,31,4,13,47,…,39
			พริกชี้ฟ้าแดง	17,13,2,20,31,10,15,21,22,20,41,16,41,39,36,14,19,30
หวี	15,17,45,19,30	10,13,10,29,29,12,10	เครื่องซักผ้า	18,44,10,23,2,20,29,37,6,19,30,12,27,45,10,13,23,41,39
			เครื่องปรับอากาศ	21,44,10,23,2,20,29,37,6,19,30,13,22,2,45,1,6,39,…,22
			ธนาคารกสิกรไทย	19,11,15,22,19,39,10,23,39,2,10,12,20,31,10,2,7,…,18
			บัตรประกันสังคม	1,45,11,2,13,22,2,38,10, 45,19,12,45,19,30,10,23,4
			ลูกกระเดือก	17,48,10,10,2,38,44,14,20,29,6,10
			ศูนย์การค้าซีคอนแสควร์	12,22,48,19,18,47,10,39,2,10,23,41,39,12,27,20,…,47

จากตารางที่ 4.8 สังเกตได้ว่า Longest common subsequence มีการตัดสินโดยการวัดค่าความเหมือนซึ่งเป็นผลทำให้คำที่แปลออกมาถูกจับคู่กับคำที่มีความยาวของลำดับท่ามือที่มาก เนื่องจากคำที่มีลำดับท่ามือที่มากนั้นประกอบด้วยรูปแบบท่ามือจำนวนมากกว่าจึงส่งผลให้มีแนวโน้มสูงที่จะถูกจับคู่อีกทั้งลำดับท่ามือที่ได้จากการสกัดเฟรมสำคัญ และการแปลท่ามือมีความผิดพลาดยิ่งส่งผลต่อการจับคู่ที่ผิดเมื่อตรวจสอบจากแค่ความเหมือนของสายลำดับท่ามือ

ตารางที่ 4.9 ตัวอย่างผลการแปลคำของ Minimun Edit Distance รูปแบบที่ 1

Input		Segment and recognition	OUTPUT(MED model 1)	
Word	Sequence of sign	Sequence of sign	Word	Sequence of sign
ฝืน	16,22,45,19	16,14,38,15,19	อดทน	6,14,11,15,19
ปี	13,12,20	11,20,20,20	ตี	11,20
			ตึก	11,20,32,10
ผี	13,23,20	13,20,20,14	เร็ว	5,20,29,14
			อูฐ	13,22,6,14
			ไอศครีม	13,22,39,14
			ปิด	13,22,20,31,14
			ปี	13,22,20
			ผิด	13,23,20,31,14
			ผี	13,23,20
			พี่	13,20,37
			พูด	13,48,14
			มีด	4,20,14
หวี	15,17,45,19,30	10,13,10,29,29,12,10	เผือก	44,13,23,20,29,6

จากตัวอย่างดังตารางที่ 4.9 คำที่เป็นผลลัพธ์นั้นมีความใกล้เคียงกับลำดับท่ามือที่ได้จากการสกัดข้อมูล และการแปลท่ามือ มากกว่าจากอัลกอริทึมแรกทั้งความยาว และความคล้ายของลำดับท่ามือแต่อย่างไรก็ตามอัลกอริทึมนี้ยังไม่สามารถช่วยแก้ปัญหาเช่น การใช้วิธีการสับเปลี่ยนนั้นไม่ควรมีการสับเปลี่ยนระหว่างท่ามือมือเดียวกับท่ามือสองมือ

ตารางที่ 4.10 ตัวอย่างผลการแปลคำของ Minimun Edit Distance รูปแบบที่ 2

Input		Segment and recognition	OUTPUT(MED model 2)	
Word	Sequence of sign	Sequence of sign	Word	Sequence of sign
ฝืน	16,22,45,19	16,14,38,15,19	ด่วน	14,37,3,19
			ฝน	16,22,19
			ฝืน	16,22,45,19
			ฟัน	16,45,19
			อดทน	7,15,19
			ที่นี่	11,15,20,37,19,20, 37
ผี	13,23,20	13,20,20,14	ปิด	13,22,20,31 ,14
หวี	15,17,45,19,30	10,13,10,29,29,12,10	เป็นลม	44,13,22,35,19,17 ,4

จากตัวอย่างดังตารางที่ 4.10 ผลลัพธ์ที่ได้มีความใกล้เคียงกับลำดับท่ามือที่ได้จากการสกัดข้อมูล นอกจากนี้อัลกอริทึมสามารถช่วยแก้ปัญหาเช่น การใช้วิธีการสับเปลี่ยนนั้นไม่ควรมีการ

สับเปลี่ยนระหว่างท่ามือมือเดียวกับท่ามือสองมือ และการเลือกสับเปลี่ยนท่ามือที่ผิดด้วยท่ามือที่มีความใกล้เคียงแทน ส่งผลให้ประสิทธิภาพในการแปลมีความถูกต้องมากกว่าสองอัลกอริทึมแรก ในลำดับถัดไปจะกล่าวถึงการสรุปผลการทดลองทั้งหมดของทุกขั้นตอนโดยภาพรวม

4.4 สรุปผลการทดลอง

จากการวัดผลการทดลองทั้ง 3 ขั้นตอนที่ได้นำเสนอมาทั้งหมดจะสรุปได้ว่าอัลกอริทึมที่ใช้ในการรู้จำการสะกดนิ้วมือของวิทยานิพนธ์เล่มนี้สามารถสกัดหาเฟรมสำคัญ และทำการกรองเฟรมที่ไม่อยู่ในขณะทำท่ามือออกทิ้งด้วยคุณภาพที่สูงถึง 80 กว่าเปอร์เซนต์ ถัดจากผ่านขั้นตอนการสกัดหาเฟรมสำคัญได้แล้วนั้นงานวิจัยนี้สามารถทำการแปลงข้อมูลที่ได้จากอุปกรณ์รับข้อมูลทั้งสองชนิดเป็นท่ามือ โดยสามารถครอบคลุมท่ามือทั้งหมดที่ใช้ในการสะกดนิ้วมือของภาษามือไทยได้มีความแม่นยำสูง และสุดท้ายงานวิจัยนี้ทดลองเพิ่มเติมเพื่อทำการแปลงลำดับท่ามือที่ได้จากกระบวนการรู้จำให้เป็นตัวอักษร หรือ คำที่ถูกต้องได้ถึงแม้ในขั้นตอนการสกัดเฟรมสำคัญจะสกัดเฟรมสำคัญเกิน หรือขาด และเมื่อท่ามือถูกแปลงผิดพลาดจากขั้นตอนการแปลงข้อมูลจากอุปกรณ์เป็นท่ามือ แต่อย่างไรก็ตามการแปลงจากลำดับของท่ามือไปเป็นคำนั้นสามารถนำไปพัฒนาต่อได้โดยการเพิ่มจำนวนของข้อมูลในการฝึกสอน แต่เนื่องจากในงานวิจัยนี้มีข้อจำกัดทางด้านงบประมาณในการจ้างล่ามทางด้านภาษามือ และเวลาของล่ามทางด้านภาษามือจึงทำให้ปริมาณของข้อมูลนั้นมีจำนวนไม่เพียงพอต่อขั้นตอนการแปลง

สรุปในบทที่ 4 นี้ได้นำเสนอลำดับขั้นตอนที่ใช้ทำการทดลอง และวิธีการในการประเมินผลการทดลองซึ่งประกอบไปด้วย 3 ขั้นตอนคือ การตรวจสอบความถูกต้องของการสกัดหาช่วงเวลาที่ผู้ทำท่ามือทำท่ามือ การตรวจสอบความถูกต้องของขั้นตอนการรู้จำท่ามือ สุดท้ายคือการตรวจสอบความถูกต้องของการแปลท่ามือ

บทที่ 5
สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

งานวิจัยนี้นำเสนอการสร้างอัลกอริทึมสำหรับทำการรู้จำการสะกดนิ้วมือของภาษามือไทย ซึ่งเป็นงานวิจัยแรกที่มีการทำการรู้จำการสะกดนิ้วมือของภาษามือไทยที่ครอบคลุมทั้ง พยัญชนะ สระ วรรณยุกต์ และสัญลักษณ์ ในภาษาไทยตามแบบสะกดนิ้วมือจนครบทุกตัวอักขระ

5.1 สรุปและอภิปรายผลการวิจัย

ในหัวข้อนี้จะสรุปและอภิปรายงานวิจัยทั้งหมดตั้งแต่จุดเริ่มต้นของงานวิจัยตามด้วยอัลกอริทึมที่นำเสนอไปจนถึงสรุปผลการทดลองดังนี้

งานวิจัยเกี่ยวกับการรู้จำการสะกดนิ้วมือของภาษามือนั้นมีอยู่ในหลายประเทศโดยที่แต่ละประเทศมีความแตกต่างกันของอัลกอริทึมโดยแปรผันตามภาษาถิ่นของแต่ละประเทศ และประเทศไทยนั้นมีการสะกดนิ้วมือตามอักษรของภาษาไทย ซึ่งมีจำนวนมากกว่าภาษามือต้นแบบอย่างเช่น การสะกดนิ้วมือของภาษามืออเมริกันที่เป็นต้นแบบในการพัฒนาการสะกดนิ้วมือไทย ในปัจจุบันยังไม่มีงานวิจัยก่อนหน้าที่ทำการรู้จำการสะกดนิ้วมือครบทุกตัวอักษร ตามแบบสะกดนิ้วมือ และหากทำการแบ่งกลุ่มของอัลกอริทึมการรู้จำการสะกดนิ้วมือ หรือการรู้จำภาษามือในปัจจุบันสามารถแบ่งได้เป็นสองกลุ่มตามอุปกรณ์รับข้อมูล กลุ่มแรกคือ งานวิจัยที่มีกล้องเป็นอุปกรณ์นำเข้าข้อมูลซึ่งมีข้อด้อยในการสกัดคุณสมบัติของข้อมูลเพื่อมาทำการแปล จึงทำให้ในงานวิจัยนี้ทำการทดลองด้วยอุปกรณ์นำเข้ากลุ่มที่สองที่ใช้อุปกรณ์พิเศษ เช่น ถุงมือรับข้อมูล และอุปกรณ์ติดตามการเคลื่อนที่ เป็นต้น ซึ่งอุปกรณ์กลุ่มนี้สามารถให้ข้อมูลที่มีคุณภาพมากกว่า แต่อย่างไรก็ตามยังมีความยากในการสกัดหาช่วงเวลาการทำการรู้จำการสะกดนิ้วมือรวมถึงความยากในการสร้างโมเดลสำหรับแปลงข้อมูลเป็นท่ามือ และแปลงจากลำดับท่ามือเป็นอักขระ หรือ คำ

ในงานวิจัยชิ้นนี้ให้ความสนใจในการพัฒนาอัลกอริทึมสำหรับทำการรู้จำการสะกดนิ้วมือของภาษามือไทยที่ครอบคลุมทั้ง พยัญชนะ สระ วรรณยุกต์ และสัญลักษณ์ ในภาษาไทยตามแบบสะกดนิ้วมือจนครบทุกตัวอักขระ โดยมีลำดับขั้นตอนหรือกระบวนการของอัลกอริทึมทั้งสิ้น 3 ขั้นตอนหลัก ๆ ได้แก่

- ขั้นตอนการสกัดหาช่วงเวลาที่ผู้ทำท่ามือกำลังทำท่ามือ เพื่อสกัดหาเฟรมสำคัญ และกรองเฟรมที่ไม่ได้อยู่ในช่วงเวลาที่ผู้ทำท่ามืออยู่ในสถานะทำท่ามือ

- ขั้นตอนการรู้จำท่า เพื่อแปลงข้อมูลได้รับจากอุปกรณ์ ณ จุดเวลาที่เป็นเฟรมสำคัญ ซึ่งถูกสกัดได้จากขั้นตอนแรก ซึ่งจะมีการเก็บท่ามือไว้เป็นลำดับตั้งแต่เริ่มทำการสะกดอักขระ หรือคำ
- ขั้นตอนการแปลท่ามือ เพื่อแปลงจากลำดับท่ามือไปเป็นตัวอักษร หรือคำ

5.2 ข้อจำกัดและข้อเสนอแนะ

แม้ว่าจากการทดลอง เมื่อทำการทดสอบความสามารถในการสกัดหาเฟรมสำคัญ และการแปลข้อมูลจากอุปกรณ์ถุงมือ และอุปกรณ์ติดตามการเคลื่อนที่ จะให้ผลความถูกต้องที่อยู่ในเกณฑ์ที่ดี แต่อย่างไรก็ตามยังมีบางส่วนที่ทั้งสองขั้นตอนมีความผิดพลาด ส่งผลให้การแปลเป็นผลเป็นอักษร หรือคำ ยังไม่ถูกต้องในบางครั้ง และเนื่องจากงานนี้เป็นงานวิจัยชิ้นแรกที่มีการทำการรู้จำตัวอักษรไทยครอบคลุมมากที่สุด ดังนั้นแล้วจำนวนรูปแบบของท่ามือจึงมีมากซึ่งส่งผลทำให้ความถูกต้องในขั้นตอนการแปลงจากข้อมูล เป็นรูปแบบท่ามือมีความถูกต้องลดลง

งานวิจัยนี้เน้นที่การสร้างอัลกอริทึมที่สามารถทำการรู้จำการสะกดนิ้วมือของภาษามือไทย ซึ่งประกอบด้วยการสกัดเฟรมสำคัญ การรู้จำรูปแบบท่ามือ และการแปลลำดับท่ามือเป็นตัวอักษร หรือคำ แต่ยังขาดอัลกอริทึมในการตัดคำซึ่งสามารถนำไปพัฒนาในงานวิจัยอื่น ๆ ได้

รายการอ้างอิง

[1] Oz, C. and Leu, M. C. American Sign Language word recognition with a sensory glove using artificial neural networks. Engineering Applications of Artificial Intelligence 24 (October 2011) : 1204-1213.

[2] Ibarguren, A., Maurtua, I., and Sierra, B. Layered architecture for real-time sign recognition. Computer Journal 53 (October 2010) : 1169-1183.

[3] Ibarguren, A., Maurtua, I., and Sierra, B. Layered architecture for real time sign recognition: Hand gesture and movement. Engineering Applications of Artificial Intelligence 23 (October 2010) : 1216-1228.

[4] Holden, E. J., Lee, G. E., and Owens, R. A., Australian sign language recognition. Machine Vision and Applications 16 (December 2005) : 312-320.

[5] Liwicki, S. and Everingham, M. R. Automatic recognition of fingerspelled words in british sign language In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 50-57. Miami : IEEE, 2009

[6] Aleem, K. A., Yousuf, B. A., Mehmood, U., Suleman, M., Sameer, R. U., Razi, R., and Israr, A., Pakistan Sign Language Recognition Using Statistical Template Matching. International Journal of Information Technology during (ICIT 2004) 1 (Dec 2004) in Istanbul, Turkey : 1-12.

[7] Machacon, H. T. C. and Shiga, S., Recognition of Japanese finger spelling gestures using neural networks. Journal of Medical Engineering and Technology 34 (May 2010) : 254-260.

[8] Setsatian School for the Deaf. Thai finger spelling Chart[Online]. 2013. Available from: http://www.setsatian.ac.th/index.html[2013,May 2]

[9] 5DT Inc. 5DT Products[Online]. 2012. Available from: http://www.5dt.com/products/pdataglove14.html. [2012,August 20].

[10] A. T. Corporation. Ascension Technology Corporation[Online]. Available from: http://www.ascension-tech.com/medical/trakstar.php. [2012,August 20].

[11] Kijsirikul, B. Artificial Intelligence. Bangkok : Chulalongkorn University, 2003.

[12] Hirschberg, D. S. Algorithms for the Longest Common Subsequence Problem. Journal of the ACM 24 (October 1977) : 664-675.

[13] Manning, C. The Longest Common Subsequence[Online]. Available from: http://www.stanford.edu/class/cs124/lec/med.pdf. [2013,January 12].

[14] Ditcharoen, T., Gereone, N., Naruedomkul, K., and Tipakorn, B., TSTMT: Step towards an accurate thai sign translation. In (ICMLA 2005) Fourth International Conference on Machine Learning and Applications, pp. 305-310. Los Angeles: IEEE, 2005.

[15] Phitakwinai, S., Auephanwiriyakul, S. and Theera-Umpon, N. Thai sign language translation using Fuzzy C-Means and scale invariant feature transform. In International Conference on Computational Science and Its Applications (ICCSA 2008), pp. 1107-1119. Perugia: Springer, 2008.

[16] Buehler, P., Everingham, M.R., Huttenlocher, D.P. and Zisserman, A.P. Upper body detection and tracking in extended signing sequences. International Journal of Computer Vision 95 (November 2011) : 180-197.

[17] Hruz, M. and others. Automatic fingersign-to-speech translation system. Journal on Multimodal User Interfaces 4 (July 2011) : 61-79.

[18] Saengsri, S., Niennattrakul, V. and Ratanamahatana, C.A. TFRS : Thai Sign Language Recognition System. In 2nd International Conference on Digital Information and Communication Technology and its Applications (DICTAP2012), pp. 457-462. Thailand: IEEE, 2012.

ภาคผนวก

ภาคผนวก ก

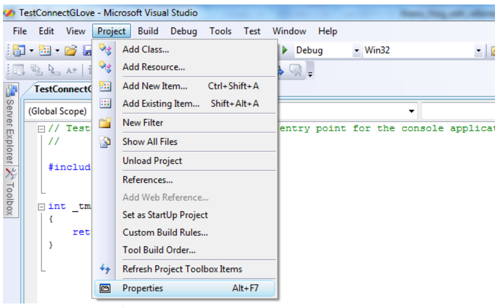
การติดตั้ง และการใช้งานถุงมือรับข้อมูลรุ่น 5DT 14 Data Glove Ultra

โดยใช้ Microsoft Visual Studio C++

การติดตั้ง และใช้งาน SDK ของถุงมือ

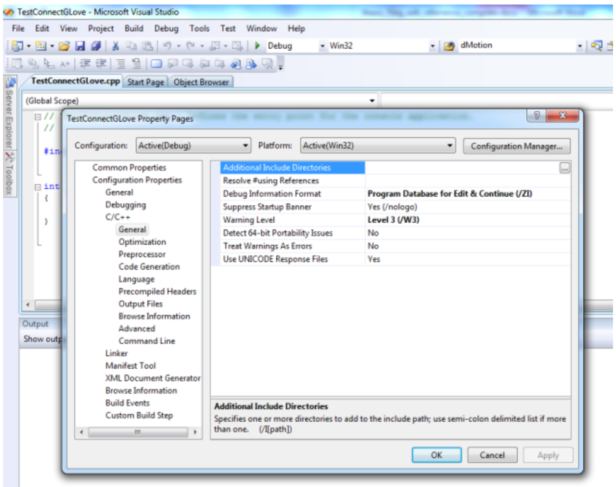
การติดตั้งถุงมือรับข้อมูลรุ่น 5DT 14 Data Glove Ultra นั้นสามารถเข้าไปดาวน์โหลดไดร์ฟเวอร์จาก http://www.5dt.com/ ซึ่งเป็นผู้ผลิต และจัดจำหน่าย จากนั้นทำการคัดลอก fglove.h, fglove.lib และfglove.dll ไปไว้ยัง path ที่ระบบสามารถมองเห็นได้จากนั้นทำการ configuration ค่าต่าง ๆ โดยในภาคผนวกนี้จะเป็นการสาธิตวิธีการนำ dll มาใช้ซึ่งเหมาะสำหรับผู้ที่ไม่เคยใช้งานให้สามารถทำตามได้อย่างง่าย ๆ ซึ่งมีขั้นตอนดังนี้

1. หลังจากสร้าง Project แล้วให้เลือกเมนู Project จากนั้นเลือกเมนู Property



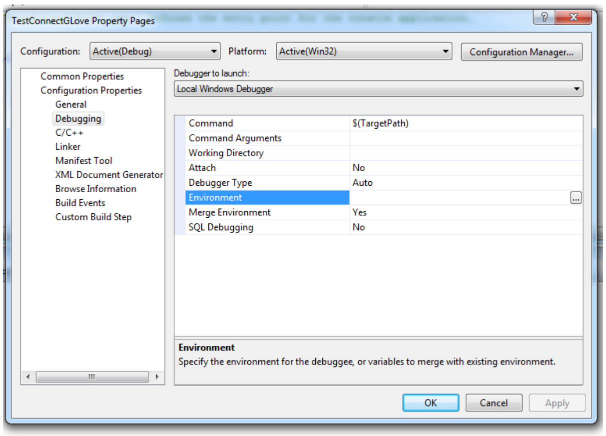
รูปที่ ก.1 เลือกเข้าเมนู Property

2. จากนั้นทำการเลือก Configuration Property >> C/C++ >> General ให้ทำการกรอก Path ของ fglove.h ลงใน Additional Include Directories



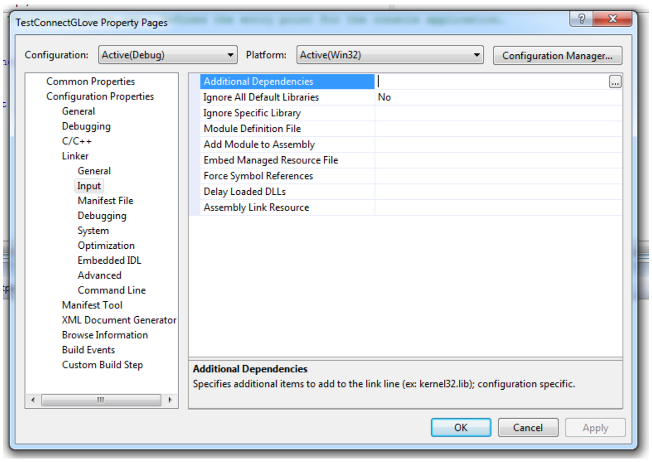
รูปที่ ก.2 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.h ได้

3. จากนั้นให้ทำการเลือก Configuration Property >> Debugging และทำการใส่ Path ของ fglove.dll ในช่องของ Environment



รูปที่ ก.3 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.dll ได้

4. ทำการเลือก Configuration Property >> Linker และทำการพิมพ์ fglove.lib ลงในช่อง Additional Dependencies



รูปที่ ก.4 การตั้ง path ให้ project เข้าถึงไฟล์ fglove.lib ได้

5. สุดท้ายใส่ #include <fglove.h> ลงในโปรเจคซึ่งคำอธิบายคำสั่งในการใช้งานตัวแปรและฟังก์ชั่นอย่างละเอียดจะถูกเก็บอยู่ใน 5DT\Data Glove Ultra\Documents

ตัวอย่างการติดต่อ และคำสั่งเรียกใช้ถุงมือเบื้องต้น

ตารางที่ ก.1 โค้ดตัวอย่างการเชื่อมต่อถุงมือ

File : GloveSample.cpp

```
1 #include <iostream>
2 #include "fglove.h"
3  int main (int argc, char * const argv[])
4 {
5      fdGlove *pGlove = NULL;
6      pGlove = fdOpen(""); // connects to first glove found
7       if (pGlove)
8      {
9           cout<<"found glove \n";
10          cout<<"Glove Type:”<<fdGetGloveType(pGlove);
11     }
12      fdClose(pGlove);
13      return 0;
14 }
```

ตัวแปรชนิด **fdGlove** ในบรรทัดที่ 5 เป็นตัวแปรสำคัญซึ่งใช้ทั้งในการเชื่อมต่อกับถุงมือ รับค่าจากตัวรับรู้บนถุงมือ และการยกเลิกการเชื่อมต่อกับถุงมือซึ่งต้องการประกาศแบบพ็อยต์เตอร์เท่านั้น ในบรรทัดที่ 6 เป็นการเชื่อมต่อถุงมือโดยไม่ระบุช่องสัญญาณ ซึ่งถ้าหากสามารถต่อได้สำเร็จ โปรแกรมนี้จะทำการรีเทิร์นชนิดของถุงมือโดยใช้คำสั่ง fdGetGloveType(pGlove) ในบรรทัดที่ 10 และสุดท้ายทำการยกเลิกการเชื่อมต่อกับถุงมือในบรรทัดที่ 12 ด้วยคำสั่ง fdClose(pGlove) ซึ่งคำสั่งที่จำเป็นต้องใช้นอกเหนือจากคำสั่งที่กล่าวไปข้างต้นมีดังนี้

fdGlove *fdOpen(char *pPort) ในบรรทัดที่ 6 เป็นการเริ่มเชื่อมต่อกับถุงมือโดยที่ไม่ระบุช่องสัญญาณ(Port)

int fdScanUSB(unsigend short *aPID, int &nNumMax) ใช้เมื่อต้องการทราบช่องสัญญาณที่ถุงมือเชื่อมต่ออยู่ เพื่อใช้ร่วมกับคำสั่ง **fdOpen**

int fdGetGloveHand(fdGlove *pFG) ใช้เมื่อต้องการทราบว่าถุงมือเป็นมือซ้ายหรือขวา โดยจะมีการรีเทิร์น FD_HAND_LEFT หรือ FD_HAND_RIGHT

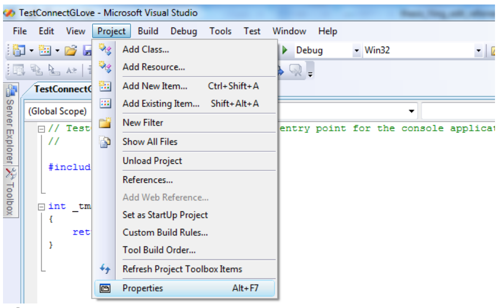
void fdGetSensorRawAll(fdGlove *pFG, unsigned short *pData) เป็นคำสั่งเก็บค่า raw data จากตัวรับรู้ทุกตัวบนถุงมือ

ภาคผนวก ข
การติดตั้ง และการใช้งานอุปกรณ์ติดตามการเคลื่อนที่รุ่น 3D Guidance trakSTAR™ โดยใช้ Microsoft Visual Studio C++

การติดตั้ง และใช้งาน SDK ของอุปกรณ์ติดตามการเคลื่อนที่

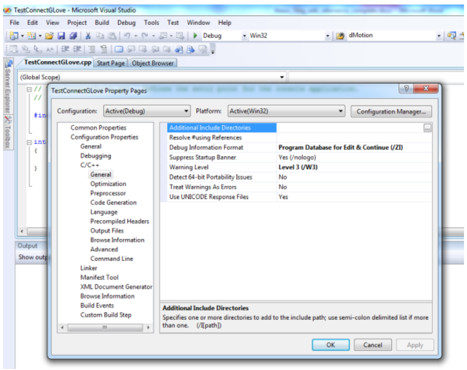
การติดตั้งอุปกรณ์ติดตามการเคลื่อนที่รุ่น 3D Guidance trakSTAR™ นั้นสามารถเข้าไปดาวน์โหลดไดร์ฟเวอร์จาก http://www.ascension-tech.com/ และให้ทำการติดตั้งไดร์ฟเวอร์ โดยที่ระหว่างติดตั้งควรให้คอมพิวเตอร์มีการเชื่อมต่อกับระบบอินเทอร์เน็ต จากนั้นทำการคัดลอก AT3DG.h, AT3DG.lib และ AT3DG.dll ไปไว้ยัง path ที่ระบบสามารถมองเห็นได้จากนั้น และทำการ configuration ค่าต่าง ๆ คล้ายกับภาคผนวก ก โดยมีขั้นตอนดังนี้

1. หลังจากสร้าง Project แล้วให้เลือกเมนู Project จากนั้นเลือกเมนู Property



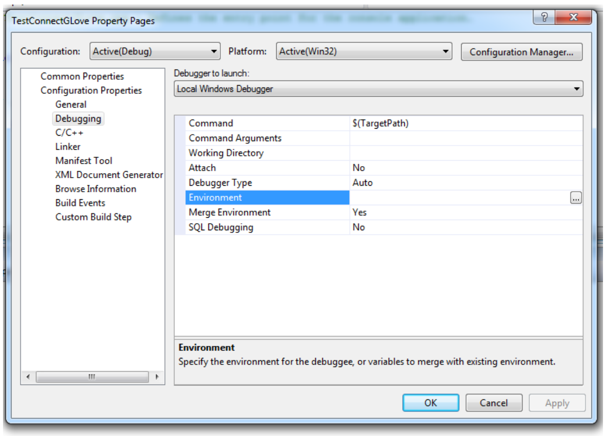
รูปที่ ข.1 การเรียกหน้าต่าง Properties ของ Project

2. จากนั้นทำการเลือก Configuration Property >> C/C++ >> General ให้ทำการกรอก Path ของ AT3DG.h ลงใน Additional Include Directories



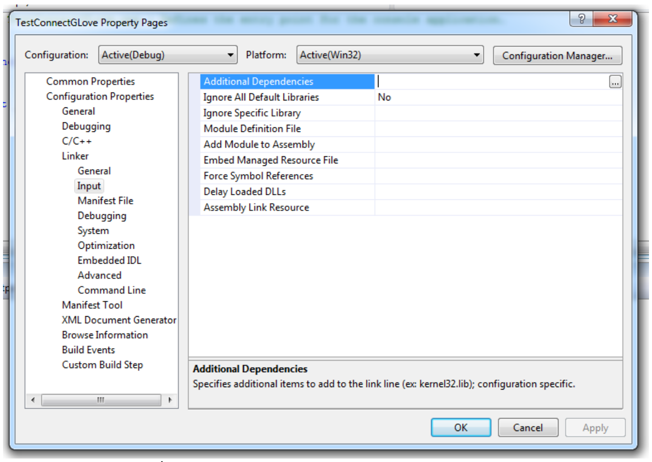
รูปที่ ข.2 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.h ได้

3. จากนั้นให้ทำการเลือก Configuration Property >> Debugging และทำการใส่ Path ของ AT3DG.dll ในช่องของ Environment



รูปที่ ข.3 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.dll ได้

4. ทำการเลือก Configuration Property >> Linker และทำการพิมพ์ AT3DG.lib ลงในช่อง Additional Dependencies



รูปที่ ข.4 การตั้ง path ให้ project เข้าถึงไฟล์ AT3DG.lib ได้

5. สุดท้ายใส่ #include < AT3DG.h> ลงในโปรเจคซึ่งคำอธิบายคำสั่งในการใช้งานตัวแปรและฟังก์ชั่นอย่างละเอียดจะถูกเก็บอยู่ใน path: ...\Ascension\3D Guidance (Rev D) และมีตัวอย่างโค้ดอยู่ใน ...\Ascension\3D Guidance (Rev D)\Samples

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวสุภาวดี แสงศรี เกิดเมื่อวันที่ 4 ธันวาคม 2531 สำเร็จการศึกษาในระดับมัธยมศึกษา จากโรงเรียนสมุทรสาครบูรณะ จังหวัดสมุทรสาคร เข้าศึกษาต่อในระดับอุดมศึกษาที่คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยมหิดล ในปีการศึกษา 2550 และสำเร็จการศึกษาในปีการศึกษา 2553 หลังสำเร็จการศึกษาระดับปริญญาตรี ได้เข้าศึกษาต่อในระดับปริญญามหาบัณฑิตในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ ที่ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2554