

การออกแบบอัลกอริทึมสำหรับแปลภาษามือของตัวอักษรภาษาอังกฤษ
โดยใช้คุณสมบัติที่ไม่แปรผันตามเรขาคณิต

ALGORITHM DESIGN FOR HAND SIGN LANGUAGE OF ENGLISH
CHARACTERS TRANSLATION BY USING GEOMETRIC INVARIANT FEATURE



T139450



วัชรินทร์ ตั้งสุขสันต์

WATCHARIN TANGSUKSANT

วพ.
ว386 ก

2558

เลขหมู่.....

เลขทะเบียน..... 139450

วันเดือนปี..... 3 ม.ย. 2558

.b.....
.i.....

12722133

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมชีวการแพทย์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2558

KMITL-2015-EN-M-045-040

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ALGORITHM DESIGN FOR HAND SIGN LANGUSGUE TRANSLATION OF
ENGLISH CHARACTERS BY USING GEOMETRIC INVARIANT FEATURE



WATCHARIN TANGSUKSANT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ENGINEERING IN BIOMEDICAL ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2015
KMITL-2015-EN-M-045-040

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2015

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์

การออกแบบอัลกอริทึมสำหรับแปลภาษามือของตัวอักษร
ภาษาอังกฤษโดยใช้คุณสมบัติที่ไม่แปรผันตามเรขาคณิต

นักศึกษา

นายวัชรินทร์ ตั้งสุขสันต์

รหัสนักศึกษา

56601191

ปริญญา

วิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชา

วิศวกรรมชีวการแพทย์

พ.ศ.

2558

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รศ.ดร.ชูชาติ ปิ่นทิวรุจน์

บทคัดย่อ

ปัจจุบันมีงานวิจัยมากมายที่ให้ความสำคัญเกี่ยวกับการสื่อสารระหว่างผู้บกพร่องทางการได้ยินกับคนปกติ ซึ่งภาษามือเป็นวิธีหนึ่งที่สำคัญเป็นอย่างมากสำหรับการสื่อสารของผู้บกพร่อง โดยเฉพาะอย่างยิ่งภาษามือที่ใช้แทนพยัญชนะภาษาอังกฤษที่เป็นภาษาสากลทั่วโลก ดังนั้นวิทยานิพนธ์นี้จึงออกแบบถุงมือที่มีมาร์คเกอร์จำนวน 6 จุด และอัลกอริทึมที่ใช้สำหรับแปลภาษามือ ซึ่งประกอบด้วย 4 ขั้นตอนหลัก ดังต่อไปนี้ ขั้นตอนหนึ่งคือการจำแนกประเภทท่าทางจากสัญญาณการเคลื่อนไหว ขั้นตอนที่สองคือการตรวจจับตำแหน่งของมาร์คเกอร์ด้วยวิธีการแปลงวงกลมของฮัฟ ขั้นตอนที่สามคือการคำนวณคุณสมบัติเฉพาะของแต่ละท่าทางด้วยวิธีคำนวณพื้นที่สามเหลี่ยมที่เป็นไปได้ทั้งหมดจากมาร์คเกอร์ทั้งหมด 6 จุดและเรียงค่าจากน้อยไปมาก และขั้นตอนสุดท้ายคือการจำแนกท่าทางด้วยวิธีการของโครงข่ายประสาทเทียม โดยผลการทดลองแสดงร้อยละความแม่นยำของอัลกอริทึมเท่ากับ 95.96 ซึ่งมีประสิทธิภาพการทำงานที่สูงและมีความเป็นไปได้ในการใช้งานจริง

อย่างไรก็ตามการพัฒนาอัลกอริทึมยังคงต้องดำเนินต่อไป เพื่อแก้ไขในจุดบกพร่องบางประการและเพื่อให้มีประสิทธิภาพที่ดียิ่งขึ้น เพื่อเป็นประโยชน์ต่อผู้บกพร่องทางการได้ยินอีกมากมายในอนาคต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Algorithm design for Hand Sign Language of English characters translation by using Geometric Invariant Feature
Student	Mr. Watcharin Tangsuksant
Student ID.	56601191
Degree	Master of Engineering
Program	Biomedical Engineering
Year	2015
Thesis Advisor	Assoc. Prof. Dr. Chuchart Pintavirooj

ABSTRACT

Several researches concern about communication between normal and disabled person. The hand gesture is one of most important communication for the deaf; especially Hand Sign Language (HSL) which is used in order to represent each alphabet (A-Z). In this research, we design the glove with six markers and develop algorithm for recognize and translate each alphabet. The algorithm consists of four main parts. First step is motion classification between static and dynamic posture. Second step is marker detection by using Circle Hough Transform. Third step is feature extraction which is a calculation all feasible triangle areas spanning among all possible triplet of markers. Finally, classification process that use feedforward backpropagation of Artificial Neural Network. The experimental result shows average of accuracy is 95.96 percent that is high performance and feasibility for proposed method.

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงได้ด้วยความกรุณาจาก รศ.ดร. ชูชาติ ปิณฑวิรุจน์ ผู้ซึ่งเป็นอาจารย์ที่ปรึกษา ให้ความรู้ คำแนะนำ อีกทั้งความช่วยเหลือ ในการแก้ไขปัญหาต่างๆ ที่เกิดขึ้น อีกทั้งยังแบ่งปันประสบการณ์ต่างๆ มากมายให้แก่ข้าพเจ้าตลอดเวลาการศึกษา

ขอขอบคุณ นายณัฐนันท์ วันลักษณ์, นายวรพงษ์ ณรงค์วงศ์วัฒนา และนายอนิวัฒน์ จูห้อง นักศึกษาสาขาวิศวกรรมชีวการแพทย์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่เสียสละเวลาส่วนตัว ในการเป็นผู้ทดลองและเก็บข้อมูลภาพ ทั้งที่ใช้ในการฝึกฝนและทดสอบต่างๆ สำหรับวิทยานิพนธ์นี้

ขอขอบพระคุณ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่สนับสนุนการทำวิทยานิพนธ์ฉบับนี้ รวมทั้งเอื้อเฟื้อสถานที่ ห้องแล็บ BIOSIS และ B-204 ในการทำโครงการ ตลอดที่ข้าพเจ้าเป็นนักศึกษาที่สถาบันแห่งนี้

ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคน ที่คอยให้กำลังใจและความช่วยเหลือในการทำวิทยานิพนธ์เล่มนี้ตลอดจนคำปรึกษาที่ทำให้ข้าพเจ้ามีแนวทางที่จะทำวิทยานิพนธ์เล่มนี้ให้เสร็จสมบูรณ์

สุดท้ายนี้ขอขอบพระคุณ บิดา มารดา และสมาชิกในครอบครัว ทุกคน ที่เป็นกำลังใจให้ข้าพเจ้าตลอดจนประสบความสำเร็จ

วัชรินทร์ ตั้งสุขสันต์

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	VIII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	1
1.3 สมมติฐานของการวิจัย.....	2
1.4 ขอบเขตของงานวิจัย.....	2
1.5 ขั้นตอนการศึกษาวิจัย.....	2
1.6 โครงสร้างวิทยานิพนธ์.....	2
บทที่ 2 ทฤษฎีและความรู้พื้นฐานที่เกี่ยวข้อง.....	4
2.1 หูและการได้ยิน.....	4
2.2 ความบกพร่องทางการได้ยิน.....	5
2.3 ภาษามือ.....	6
2.4 งานวิจัยที่เกี่ยวข้องกับการแปลภาษามือ.....	8
2.5 ทฤษฎีที่เกี่ยวข้องกับงานวิจัย.....	10
2.5.1 ทฤษฎีการปรับเทียบกล้อง.....	10
2.5.2 การคำนวณพิกัดสามมิติของจุดจากฉากสองมิติ.....	11
2.5.3 การแปลงวงกลมของฮัฟ.....	13
2.5.4 การคำนวณพื้นที่สามเหลี่ยมใดๆ ของอีรอล.....	15
2.5.5 การจำแนกข้อมูลแบบโครงข่ายประสาทเทียม.....	17
บทที่ 3 ความไม่แปรผันต่อการเปลี่ยนแปลงทางเรขาคณิต.....	23
3.1 การแปลงทางเรขาคณิตเชิงเส้น.....	23

สารบัญ (ต่อ)

	หน้า
3.1.1 การเคลื่อนย้ายภาพ	23
3.1.2 การย่อ/ขยายภาพ.....	24
3.1.3 การหมุนภาพ	26
3.1.4 การเฉือนภาพ	27
3.1.5 ระบบพิกัดแบบโฮโมจีเนียส.....	28
3.2 การแปลงทางเรขาคณิตแบบไม่เป็นเชิงเส้น.....	30
3.3 ประเภทของการแปลงทางเรขาคณิต	33
3.4 ความไม่แปรผันต่อการแปลงทางเรขาคณิต	33
3.4.1 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงแบบคงที่.....	34
3.4.2 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงเหมือน.....	35
3.4.3 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงอัฟไฟน์.....	37
3.4.4 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงเพอสเปคทีฟ.....	41
บทที่ 4 ระเบียบวิธีการวิจัย	44
4.1 การออกแบบการทำงานของระบบ.....	45
4.2 ขั้นตอนการออกแบบอัลกอริทึมสำหรับการแปลภาษามือ	47
4.2.1 กระบวนการรับภาพ.....	47
4.2.2 กระบวนจำแนกชนิดท่าทางของภาษามือ	49
4.2.3 การจำแนกท่าทางแบบ Dynamic.....	51
4.2.4 กระบวนการสกัดคุณลักษณะเฉพาะของท่าทางแบบ Static.....	52
4.2.4.1 การแปลงภาพระดับเทา	52
4.2.4.2 การตั้งค่านับระดับกันของภาพระดับเทา.....	53
4.2.4.3 การตรวจนับมาร์คเกอร์และจุดศูนย์กลางวงกลม	54
4.2.4.4 การตรวจจับสี่ของมาร์คเกอร์.....	55
4.2.4.5 การคำนวณพิกัดสามมิติด้วยวิธีการ Direct Linear Transformation	56
4.2.4.6 การสกัดคุณลักษณะท่าทางด้วยการคำนวณพื้นที่และมุมของ สามเหลี่ยม	57
4.2.5 กระบวนการคัดแยกกลุ่มข้อมูลด้วยวิธีโครงข่ายประสาทเทียม.....	60

สารบัญ (ต่อ)

	หน้า
บทที่ 5 ผลการทดลองและอภิปรายผลการทดลอง	62
5.1 ระบบการจัดเก็บภาพสำหรับแปลภาษามือ	62
5.2 การฝึกฝนโครงข่ายประสาทเทียมสำหรับแปลภาษามือ	63
5.3 ทดสอบอัลกอริทึมสำหรับการแปลภาษามือ	67
5.3.1 ทดสอบความถูกต้องสำหรับท่า A-Z	67
5.3.2 ทดสอบความถูกต้องสำหรับท่า A-Z แบบต่อเนื่อง	72
5.4 ผลการทดสอบความถูกต้องและการคำนวณความผิดพลาดของอัลกอริทึม	79
บทที่ 6 สรุปผลการทดลองและข้อเสนอแนะ	81
เอกสารอ้างอิง	83
ประวัติผู้เขียน	86



สารบัญตาราง

ตารางที่	หน้า
3.1	คุณลักษณะที่มีความไม่แปรผันต่อการแปลงทางเรขาคณิต.....34
4.1	คุณลักษณะของคอมพิวเตอร์ที่ใช้ในการทดลอง49
5.1	ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 1.....67
5.2	ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 2.....68
5.3	ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 3.....69
5.4	ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 4.....70
5.5	สรุปผลการทดลองการทดสอบความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองทั้งหมด .71



สารบัญรูป

รูปที่		หน้า
2.1	องค์ประกอบของหู.....	5
2.2	ท่าทางการสะกดด้วยนิ้วมือสำหรับพยัญชนะภาษาอังกฤษ.....	7
2.3	ถুমือที่ติดตั้งไมโครคอมพิวเตอร์เพื่อแปลภาษามือ	8
2.4	การแปลงภาพสองมิติไปยังพิกัดสามมิติโดยการติดตามข้อนิ้วมือ	9
2.5	การปรับเทียบกล้องด้วยการใช้ตารางหมากรุก.....	12
2.6	แบบจำลองการหาพิกัดสามมิติจากกล้องสองตัว	13
2.7	ส่วนประกอบตัวแปรต่างๆของวงกลม.....	14
2.8	การแปลงวงกลมของฮัฟจากระนาบ x, y ไปเป็นระนาบ a, b, r	14
2.9	รูปสามเหลี่ยมใดๆ	15
2.10	ส่วนประกอบของเซลล์ประสาท	18
2.11	โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม.....	18
2.12 ก.	ฟังก์ชัน Hyperbolic Tangent.....	19
2.12 ข.	ฟังก์ชัน Logistic	19
2.13	โครงข่ายประสาทเทียมสำหรับการฝึกฝนแบบ Backpropagation.....	20
3.1	การเคลื่อนย้ายภาพ.....	24
3.2	การย่อ/ขยายภาพ.....	25
3.3	การหมุนภาพ.....	26
3.4	การเฉือนภาพ.....	28
3.5	การแปลงแบบคงที่ของสี่เหลี่ยม ABCD เป็นสี่เหลี่ยม A'B'C'D'.....	35
3.6	การแปลงเหมือนของสี่เหลี่ยม ABCD เป็นสี่เหลี่ยม A'B'C'D'.....	36
3.7	การแปลงอัฟไฟน์ของสี่เหลี่ยม ABCD เป็นสี่เหลี่ยม A'B'C'D'.....	38
3.8	ตัวอย่างของจุดตัดที่ไม่แปรผันต่อการแปลงอัฟไฟน์	39
3.9	รูปสามเหลี่ยม $a'b'c'$ และ $d'e'f'$ ที่ได้จากการแปลงอัฟไฟน์ของสามเหลี่ยม abc และ def ตามลำดับ	40
3.10	การแปลงเพอสเปคทีฟที่โปรเจกไปยังเส้นตรง	42
3.11	จุดที่อยู่บนระนาบเดียวกัน 5 จุด	43
4.1	กระบวนการโดยสังเขปสำหรับอัลกอริทึม	44
4.2	แผนผังการจำแนกท่าทางภาษามือ	45

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.3	ตำแหน่งของมาร์คเกอร์บนถุงมือ 47
4.4	ภาพอินพุทภาษามือ A-Z จากกล้องเว็บแคม 48
4.5	สัญญาณการเคลื่อนไหวจากภาพ 50
4.6	ตำแหน่งการดึงเฟรมภาพ 50
4.7 ก.	เฟรมเริ่มต้นของ J 51
4.7 ข.	เฟรมสิ้นสุดของ J 51
4.7 ค.	เฟรมเริ่มต้นของ Z 51
4.7 ง.	เฟรมสิ้นสุดของ Z 51
4.8	การแปลงภาพระดับเทา 52
4.9	ฮิสโตแกรมของภาพระดับเทา 53
4.10 ก.	ภาพไบนารีหลังกระบวนการเทรสโฮลด์ดิ้งสำหรับกล้องที่ 1 54
4.10 ข.	ภาพไบนารีหลังกระบวนการเทรสโฮลด์ดิ้งสำหรับกล้องที่ 2 54
4.11 ก.	การตรวจจับมาร์คเกอร์แต่ละจุดสำหรับกล้องที่ 1 54
4.11 ข.	การตรวจจับมาร์คเกอร์แต่ละจุดสำหรับกล้องที่ 2 54
4.12	การสร้างกรอบขนาด 5x5 พิกเซล รอบจุดศูนย์กลางวงกลม 55
4.13	ตัวอย่างพิกัดสามมิติจากการคำนวณแบบ DLT ของท่าทาง Y 56
4.14	พิกัดสามมิติจากการคำนวณแบบ DLT สำหรับทุกท่าทางที่ใช้ในการฝึกฝน 57
4.15	การสร้างสามเหลี่ยมจากมาร์คเกอร์ 6 จุด 58
4.16	ความสัมพันธ์ระหว่างมุมและเวกเตอร์ของสามเหลี่ยม 59
4.17	กราฟความสัมพันธ์ของค่าเฉลี่ยแต่ละท่าทางกับค่าพื้นที่และมุม 59
4.18	คู่มือการฝึกฝนระหว่างอินพุทและเป้าหมาย 60
4.19	การแสดงผลลัพธ์เป็นตัวอักษร 61
5.1	ระบบการจัดเก็บภาพ 63
5.2	ขั้นตอนการฝึกฝนโครงข่ายประสาทเทียมสำหรับ 1 ท่าทาง 63
5.3	การออกแบบกระบวนการฝึกฝนโครงข่ายประสาทเทียม 64
5.4	ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 1) 65
5.5	ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 2) 65
5.6	ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 3) 66

สารบัญรูป (ต่อ)

รูปที่		หน้า
5.7	ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 4).....	66
5.8 ก.	ภาษามือแสดงท่าทางคำว่า KMITL.....	72
5.8 ข.	ภาษามือแสดงท่าทางคำว่า WIFI.....	72
5.8 ค.	ภาษามือแสดงท่าทางคำว่า JANUARY.....	72
5.9	สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 1.....	73
5.10	สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 1.....	73
5.11	สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 1.....	74
5.12	สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 2.....	74
5.13	สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 2.....	75
5.14	สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 2.....	75
5.15	สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 3.....	76
5.16	สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 3.....	76
5.17	สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 3.....	77
5.18	สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 4.....	77
5.19	สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 4.....	78
5.20	สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 4.....	78
5.21 ก.	ตัวอย่างท่าทางตัวอักษร K จากกล้องตัวที่ 1.....	79
5.21 ข.	ตัวอย่างท่าทางตัวอักษร K ที่ถูกบดบัง จากกล้องตัวที่ 2.....	79
5.22	ตัวอย่างตัวอักษร A ที่คัดแยกผิดพลาดเป็นตัวอักษร E.....	80

บทที่ 1

บทนำ

(Introduction)

ในบทนี้จะกล่าวถึงความเป็นมาและความสำคัญของปัญหา วัตถุประสงค์ของงานวิจัย สมมติฐานของการศึกษา ทฤษฎีหรือแนวความคิดที่ใช้ในงานวิจัย ขอบเขตของงานวิจัย และโครงสร้างวิทยานิพนธ์

1.1 ความเป็นมาและความสำคัญของปัญหา

เทคโนโลยีมีความก้าวหน้าอย่างรวดเร็ว โดยเฉพาะอย่างยิ่งเทคโนโลยีทางการแพทย์ที่มีความสำคัญอย่างมากสำหรับการรักษา การวินิจฉัยและการช่วยเหลือหรือทดแทนให้แก่ผู้ป่วยและผู้บกพร่องทางร่างกาย มีการนำเทคโนโลยีหลากหลายแขนง เช่น อิเล็กทรอนิกส์ แมคคานิค คอมพิวเตอร์ ฯลฯ เพื่อประยุกต์ใช้ในการประดิษฐ์คิดค้นเครื่องมือและอุปกรณ์ สำหรับช่วยเหลือผู้ป่วยและผู้บกพร่องทางร่างกายในการที่บุคคลเหล่านั้นสามารถดำรงชีวิตอยู่ได้อย่างดีขึ้น ให้เทียบเท่ากับคนปกติ

งานวิจัยนี้ตระหนักถึงการดำเนินชีวิตของผู้บกพร่องทางการได้ยิน ในการติดต่อสื่อสารกับคนปกติที่ขาดความเข้าใจในท่าทางของภาษามือ ทำให้การสื่อสารระหว่างผู้พิการกับคนปกติเกิดความยุ่งยากและความเข้าใจที่ผิดพลาด ดังนั้นงานวิจัยนี้มุ่งเน้นการประยุกต์ใช้เทคโนโลยีการประมวลผลภาพ และคิดค้นพัฒนาอัลกอริทึมเพื่อนำไปใช้งานสำหรับแปลภาษามือ โดยหวังเป็นอย่างยิ่งว่างานวิจัยนี้สามารถทำให้ผู้บกพร่องทางการได้ยินจะได้รับความสะดวกยิ่งขึ้นในการสื่อสาร และจุดประกายให้กับนักวิจัยในประเทศไทยที่เล็งเห็นความสำคัญในการพัฒนาอุปกรณ์ทางการแพทย์ในประเทศไทยต่อไปในอนาคต

1.2 วัตถุประสงค์ของงานวิจัย

1.2.1 เพื่อนำเสนออัลกอริทึมสำหรับแปลภาษามือไปเป็นตัวอักษรที่มีความถูกต้องสูง

1.2.2 เพื่อทดสอบความแม่นยำของอัลกอริทึมที่นำเสนอ

1.2.3 เพื่อนำความรู้ด้านการประมวลผลภาพมาประยุกต์ใช้ให้เกิดประโยชน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 สมมติฐานของการศึกษาวิจัย

งานวิจัยนี้มุ่งเน้นการออกแบบพัฒนาอัลกอริทึมสำหรับแปลภาษามือ โดยภาษามือนั้นมีท่าทางที่แตกต่างกัน โดยใช้กระบวนการประมวลผลภาพสำหรับการคำนวณหาคุณสมบัติเฉพาะของแต่ละท่าทาง ซึ่งคุณสมบัติเฉพาะเหล่านี้มีความสำคัญอย่างมากสำหรับกระบวนการแยกแยะท่าทาง โดยงานวิจัยนี้ตั้งสมมติฐานเพื่อหาคุณสมบัติเฉพาะที่ไม่แปรผันตามเรขาคณิต ซึ่งวิธีการของการคำนวณพิกัดสามมิติ สามารถนำมาประยุกต์ใช้เป็นอย่างดี โดยตามหลักการการคำนวณหาพื้นที่ใดๆ ซึ่งเกิดจากพิกัดสามมิตินั้น มีความไม่แปรผันตามเรขาคณิต ดังนั้นผู้วิจัยจึงใช้หลักการตามทฤษฎีและสมมติฐานดังกล่าวในการคำนวณหาคุณสมบัติเฉพาะของแต่ละท่าทาง เพื่อแยกแยะและแปลผลเป็นตัวอักษร

1.4 ขอบเขตของงานวิจัย

1.4.1 ออกแบบถุงมือสำหรับผู้ใช้งาน โดยมีมาร์คเกอร์ 6 จุด

1.4.2 ออกแบบและพัฒนาอัลกอริทึมสำหรับแปลภาษามือตัวอักษร A-Z

1.5 ขั้นตอนการศึกษา

1.5.1 กำหนดวัตถุประสงค์ของการศึกษา ความเป็นไปได้และขอบเขตของงานวิจัย

1.5.2 รวบรวมเนื้อหาและงานวิจัยอื่นๆที่เกี่ยวข้องกับงานวิจัย

1.5.3 ออกแบบและพัฒนางานวิจัย

1.5.4 ทำการทดลอง และเก็บผลการทดลอง

1.5.5 สรุปผลงานวิจัย

1.6 โครงสร้างวิทยานิพนธ์

ในวิทยานิพนธ์ฉบับนี้จะแบ่งเนื้อหาออกเป็น 6 บท ในแต่ละบทจะมีเนื้อหาดังต่อไปนี้

บทที่ 1 กล่าวถึงความเป็นมาและความสำคัญของปัญหา สมมติฐานของงานวิจัย วัตถุประสงค์ของการศึกษา และขอบเขตของงานวิจัย

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

บทที่ 3 อธิบายทฤษฎีและรายละเอียดเกี่ยวกับเรื่องความไม่แปรผันต่อการแปลงทางเรขาคณิต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4 ระเบียบวิธีการวิจัย ประกอบไปด้วยขั้นตอนต่าง ๆ ได้แก่ การออกแบบระบบโดยรวม การออกแบบข้อมูลและการออกแบบอัลกอริทึมสำหรับแปลภาษามือ

บทที่ 5 ผลการทดลองและอภิปรายผล

บทที่ 6 บทสรุปและข้อเสนอแนะงานวิจัย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

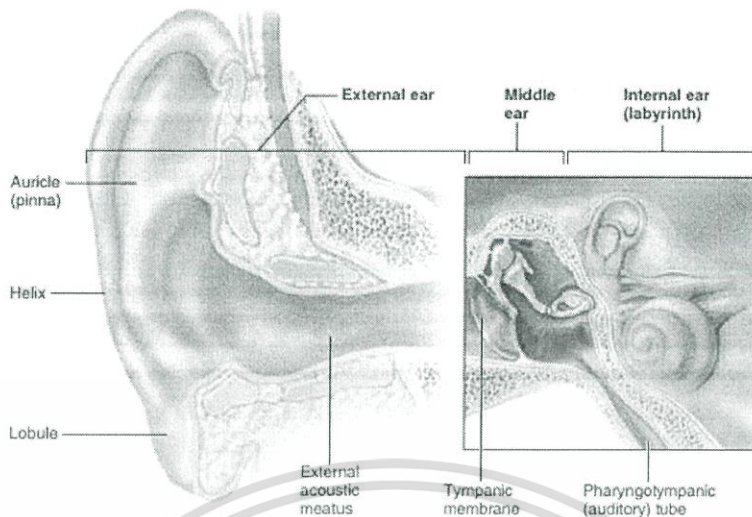
ทฤษฎีและความรู้พื้นฐานที่เกี่ยวข้อง (Theory and Related Basic Knowledge)

2.1 หูและการได้ยิน

หู คือ อวัยวะรับรู้ของร่างกาย ซึ่งหน้าที่ของหูมีหน้าที่เปลี่ยนการสั่นสะเทือนทางกายภาพ (Physical Vibration) ไปเป็นสัญญาณประสาท (Nerve Impulse) โดยมีองค์ประกอบ 3 ส่วนหลักๆ ได้แก่ หูชั้นนอก หูชั้นกลางและหูชั้นใน ดังรูปที่ 2.1

หูชั้นนอก ประกอบด้วย ใบหู (Pinna) และช่องหู (Ear Canal) โดยมีหน้าที่หลักคือ การรับเสียงหรือทางผ่านเสียงเพื่อเข้าสู่ในส่วนถัดไป ในส่วนถัดไป คือ หูชั้นกลาง ประกอบด้วย แก้วหูหรือเยื่อแก้วหู (Eardrum หรือ Tympanic Membrane) ซึ่งมีหน้าที่แปลงเสียงเป็นแรงสั่นสะเทือนไปยังกระดูก 3 ส่วน ในหูชั้นกลาง ได้แก่ กระดูกค้อน (Malleus) ทัง (Incus) และโกลน (Stapes) โดยกระดูกชิ้นเล็กๆ ทั้ง 3 ชิ้นนี้ ทำหน้าที่ส่งการสั่นสะเทือนไปยังหูชั้นในต่อไป หูชั้นในหรือโคเคลีย (Cochlea) ประกอบด้วยน้ำและเซลล์ขน (Hair Cell) ซึ่งมีการรับและการตอบสนองด้วยความไวสูง โครงสร้างเล็กๆ นี้ จะเคลื่อนไหวเมื่อถูกกระตุ้นด้วยความสั่นสะเทือนของเสียงและส่งไปยังเส้นประสาทรับเสียง (Auditory Nerve) ซึ่งสมองจะมีหน้าที่รับสัญญาณประสาทนี้เข้าไปเพื่อแปลความหมายต่อไป นอกจากนี้หูชั้นในจะมีหน้าที่รับรู้เสียงแล้ว แต่มีหน้าที่ในการรักษาสสมดุลร่างกายได้ด้วยเช่นเดียวกัน [1]

ปกติเสียงที่มนุษย์สามารถรับรู้จะมีความถี่อยู่ในช่วง 20 - 20,000 เฮิร์ต (Hertz) ซึ่งความถี่เสียงที่พบได้ปกติทั่วไปในชีวิตประจำวันอยู่ในช่วง 125 - 8,000 เฮิร์ต แต่ช่วงความถี่ของการพูดคุยสื่อสารอยู่ในช่วงความถี่ระหว่าง 500 - 2,000 เฮิร์ต เท่านั้น โดยความถี่ที่หูจะตอบสนองได้ดีที่สุดอยู่ในช่วง 3,000 - 4,000 เฮิร์ต เนื่องจากมี Amplifying Mechanism มากที่สุดภายในช่องหู ดังนั้นเสียงกระตุ้นที่อยู่ในช่วงความถี่ดังกล่าวมีผลทำให้ hair cell ที่ทำงานในช่วงความถี่นี้ถูกทำลายได้มากที่สุด จึงอธิบายได้ว่าเหตุใดการสูญเสียการได้ยินที่เกิดจากสัญญาณรบกวนจึงเริ่มเสียหายที่ความถี่ 3,000 - 4,000 เฮิร์ต [2]



รูปที่ 2.1 องค์ประกอบของหู [3]

2.2 ความบกพร่องทางการได้ยิน

ความบกพร่องทางการได้ยินหรือที่เรียกกันว่าหูหนวก เกิดจากหลายสาเหตุ ซึ่งสามารถแบ่งออกเป็น 2 ประเภทหลัก คือ หูหนวกก่อนคลอด และหูหนวกหลังคลอด โดยประเภทของการหูหนวกก่อนคลอดอาจมีสาเหตุจากพันธุกรรมของพ่อและแม่ เช่น พ่อหรือแม่หูหนวกอาจส่งผลให้ลูกหรือหลานหูหนวก หรืออีกสาเหตุ เกิดจากการกระทบกระเทือนต่อครรภ์ขณะตั้งครรภ์ เช่น การหกล้มของแม่ ทำให้อวัยวะของลูกในครรภ์ถูกกดทับและส่งผลให้เลือดไปเลี้ยงอวัยวะไม่เพียงพอ เมื่อทารกเกิดมาจึงหูหนวก หรือเกิดจากการที่อวัยวะของลูกในครรภ์เติบโตไม่เต็มที่ อาจทำให้ไม่มีรูหูเมื่อคลอดออกมา นอกจากนั้นยังมีโอกาสที่ทารกจะหูหนวกเนื่องจากการใช้ยาบางประเภทที่เป็นอันตรายต่อทารกในครรภ์คาคิวินิน ยาแอสไพริน ยาสเตียรอยด์มัยซิน และยาเพนนิซิลิน ซึ่งส่วนมากเหตุการณ์เหล่านี้จะมีผลมากในช่วง 3 เดือนแรกของการตั้งครรภ์และช่วงใกล้คลอด ดังนั้นมารดาต้องดูแลตัวเองเป็นพิเศษในช่วงนี้ [4]

ส่วนประเภทของการหูหนวกหลังคลอด เกิดได้จากหลากหลายปัจจัย เช่น หูหนวกจากโรคระบบประสาท หูหนวกจากโรคติดต่อ หูหนวกร่วมกับโรคต่อมไร้ท่อ หูหนวกจากยาพิษ เป็นต้น นอกจากนั้น เสียงรบกวน (Noise) หมายถึงเสียงไม่พึงประสงค์ ที่สามารถกระทบกับอารมณ์และความรู้สึกของผู้ฟังได้ หากใช้ความรู้สึกเป็นตัววัดว่าเป็นเสียงรบกวนหรือไม่จึงเกิดความยาก ดังนั้นจึงมีการกำหนดหน่วยวัดระดับความดังของเสียงขึ้นมา ในชื่อว่า เดซิเบล (dB) ซึ่งการสัมผัสเสียงดังเป็นเวลานานส่งผลให้เกิดความผิดปกติทางการได้ยิน ไม่ว่าจะเป็นเสียงดังจากการทำงาน เช่น เครื่องจักร ในโรงงานอุตสาหกรรม การฟังดนตรีเสียงดัง การขับรถจักรยานยนต์ หรือเสียงจากของเล่นเด็กที่มีเสียงดัง ก็สามารถก่อให้เกิดการสูญเสียการได้ยินได้ เนื่องจากการสัมผัสเสียงดังทำให้เกิดการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปลี่ยนแปลงระดับกันของการได้ยิน (Threshold of Hearing) ซึ่งการเปลี่ยนแปลงนี้ แบ่งออกเป็น 2 ประเภท คือ แบบชั่วคราว (Temporary) และแบบถาวร (Permanent) [2]

การสูญเสียการได้ยินแบบชั่วคราว (Temporary Threshold Shift) ทำให้การได้ยินความถี่ ในช่วง 3,000 – 6,000 เฮิร์ต ลดลง ปกติจะมีระยะเวลาการฟื้นตัวไม่เกิน 24 ชั่วโมง ยกเว้นบางรายอาจใช้เวลานานเป็นสัปดาห์ แต่หากยังคงรับฟังเสียงดังต่อเนื่องเป็นเวลานานๆ อาจทำให้เกิดการสูญเสียแบบถาวรได้ ส่วนการสูญเสียการได้ยินแบบถาวร (Permanent Threshold Shift) เกิดจากการทำลายอย่างถาวร ในส่วนของ Critical element of Cochlea ซึ่งเกิดได้จาก 2 เหตุการณ์ คือ การได้รับเสียงดังอย่างมากในระยะเวลาสั้นๆ ทำให้เกิดการเปลี่ยนแปลงความดันอย่างรวดเร็วภายในหู ทำให้หูชั้นกลางถูกทำลายโดยตรงอย่างทันที เช่น การได้ยินเสียงระเบิด เสียงปืน ส่วนอีกเหตุการณ์หนึ่ง คือ การได้รับเสียงที่เบากว่าแต่นานกว่า ทำให้เกิดการทำลายในส่วนของ Cochlea ภายในหูชั้นในอย่างถาวร

ผู้บกพร่องทางการได้ยินจะมีปัญหาทางด้านต่างๆตามมา เช่น ด้านอารมณ์ ด้านสังคม ตลอดจนด้านบุคลิกภาพ ฯลฯ โดยเฉพาะอย่างยิ่งปัญหาทางด้านภาษาและการสื่อสาร หากผู้พิการที่ไม่สามารถรักษาให้หายด้วยวิธีการใดๆ ได้ ผู้พิการเหล่านั้นจำเป็นต้องเรียนรู้และฝึกฝนการใช้ชีวิต รวมทั้งภาษามือหรือภาษาใบ้ เพื่อการสื่อสารกับผู้พิการด้วยกัน รวมไปถึงผู้ที่เข้าใจภาษามือ เช่นเดียวกัน

2.3 ภาษามือ

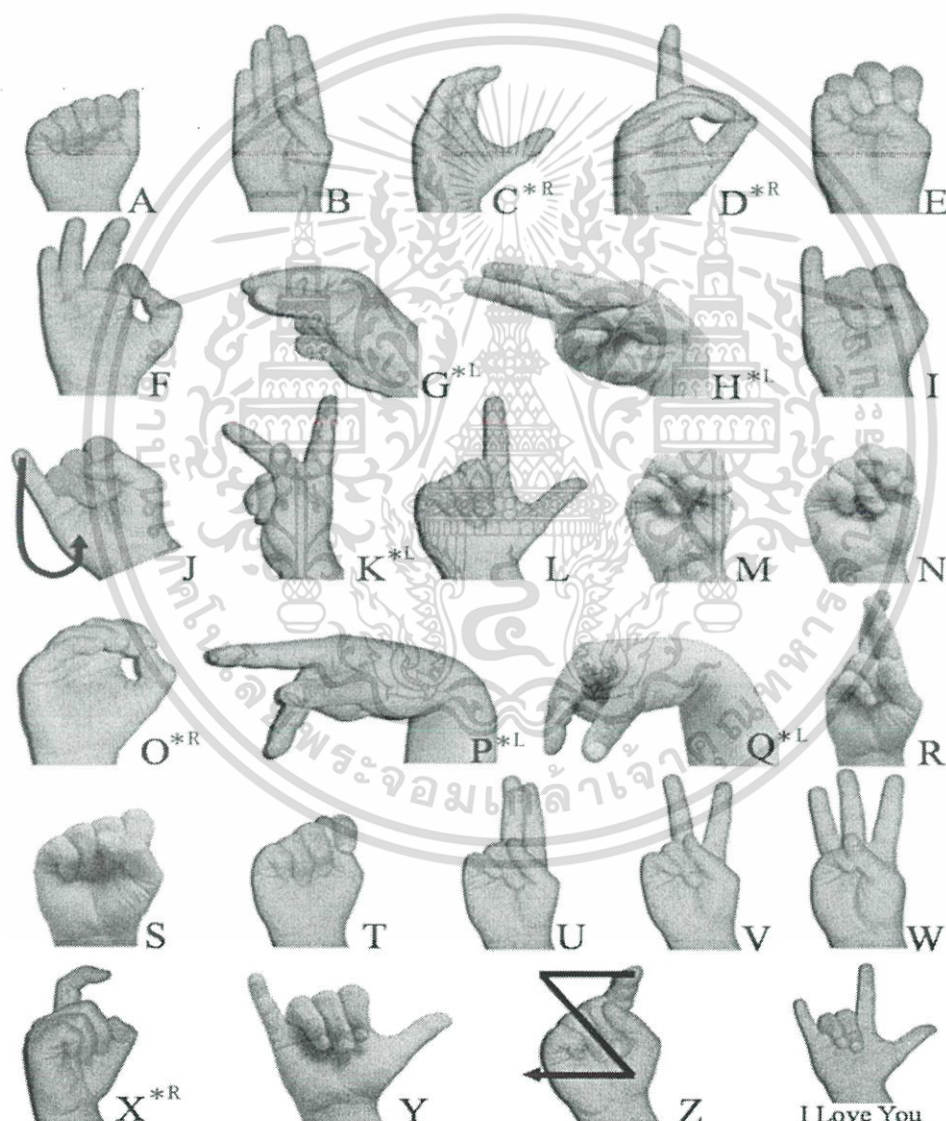
ภาษามือ คือ วจนภาษา ที่สื่อสารด้วยมือ การสื่อสารด้วยร่างกาย และการใช้ริมฝีปากในการสื่อความหมายต่างๆ แทนคำพูด โดยภาษามืออาจถูกนำไปใช้ในด้านต่างๆ เช่น ทางกการทหารเพื่อการปฏิบัติการในพื้นที่ที่ไม่สามารถสื่อสารด้วยเสียงได้ และที่สำคัญอย่างยิ่งคือภาษามือที่ถูกนำไปใช้สำหรับผู้พิการทางหู

ภาษามือสำหรับคนหูหนวกจะใช้มือ ทำท่าทางและสีหน้า สำหรับการแสดงออกเพื่อสื่อถึงความหมายและอารมณ์ โดยภาษามือมีความแตกต่างกันออกไป เช่นเดียวกับภาษาพูด ตามประเพณี วัฒนธรรมและภาษา ซึ่งภาษามือที่ใช้ในกลุ่มคนหูหนวกมี 2 แบบ คือ ภาษามือธรรมชาติ และภาษามือประดิษฐ์ โดยภาษามือธรรมชาตินั้นเกิดจากการเลียนแบบท่าทางตามธรรมชาติของกลุ่มคนหูหนวกในแต่ละชุมชน ซึ่งอาจจะเหมือนหรือแตกต่างกันก็ได้ ส่วนภาษามือประดิษฐ์เป็นภาษามือที่ผู้ปกครอง คุณครูหรือญาติของคนหูหนวกคิดขึ้นมา ใช้แทนคำพูดและภาษาเขียนประจำชาติเพื่อให้เพียงพอต่อการสื่อความหมายและการเรียนรู้ โดยเฉพาะเรื่องนามธรรม

การสะกดด้วยนิ้วมือ เป็นวิธีการสื่อสารชนิดหนึ่งสำหรับผู้สูญเสียการได้ยิน คือการที่บุคคล ทำท่าด้วยนิ้วมือ เพื่อแทนพยัญชนะ สระ วรรณยุกต์ ตลอดจนลักษณะอื่นๆ ของภาษาประจำชาติเพื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสื่อสารที่เป็นพื้นฐานและความเข้าใจที่ตรงกันของผู้พิการ [5] ภาษาศาสตร์สำหรับคนปกตินั้นคือภาษาอังกฤษ เช่นเดียวกับภาษามือที่มีการคิดค้นขึ้นมาโดยชาวอเมริกันชื่อว่า Thomas Hopkins Gallaudet เป็นผู้ก่อตั้งโรงเรียนสำหรับคนหูหนวกคนแรกของโลกในปี ค.ศ. 1817 ร่วมกับ Louis Laurent Marie Clere ซึ่งต่อมาเขาได้ยกย่องให้เป็นบิดาแห่งคนหูหนวกของอเมริกา [6] โดยภาษามือที่สะกดด้วยนิ้วมือที่ถูกคิดค้นนี้เป็นการแสดงท่าทางด้วยนิ้วมือสำหรับพยัญชนะในภาษาอังกฤษทั้ง 26 ตัว เรียกว่า “American Sign Language Alphabet” (ย่อว่า ASL) ดังรูปที่ 2.2 ซึ่งต่อมาถูกนำไปใช้ดัดแปลงไปใช้กับพยัญชนะในแต่ละประเทศ รวมถึงพยัญชนะไทยด้วย



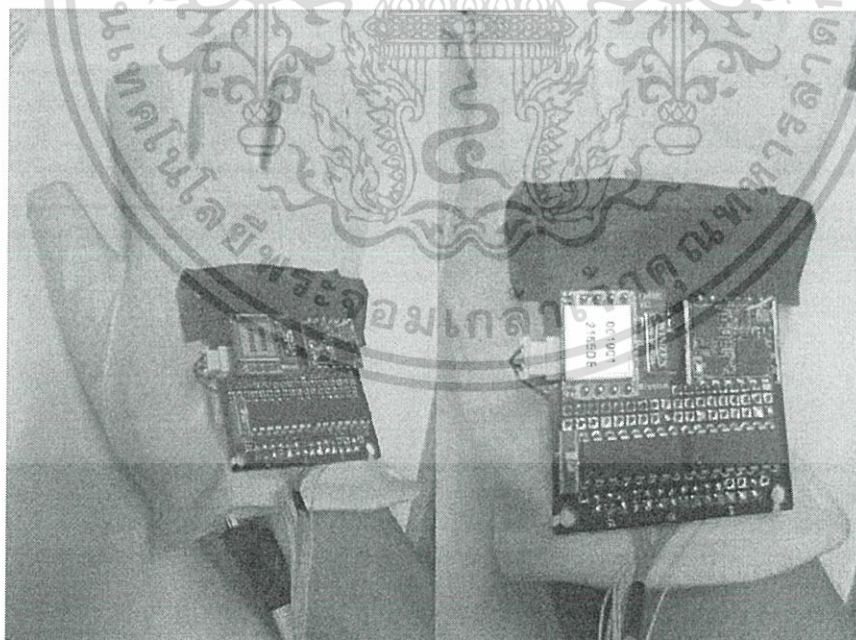
รูปที่ 2.2 ท่าทางของการสะกดด้วยนิ้วมือสำหรับพยัญชนะภาษาอังกฤษ [7]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 งานวิจัยที่เกี่ยวข้องกับการแปลภาษามือ

สำหรับหัวข้อนี้กล่าวถึงงานวิจัยที่เกี่ยวข้องโดยมุ่งเน้นไปยังวิธีการแปลภาษามือ อย่างไรก็ตามที่ได้กล่าวในหัวข้อ 2.3 เกี่ยวกับภาษามือที่มีความคล้ายกับภาษาพูด กล่าวคือมีทั้งภาษาประจำชาตินั้นๆ และภาษาสากล ดังนั้นงานวิจัยบางงานนำเสนอการแปลภาษามือสำหรับภาษาประจำชาติของตนเอง เช่น ภาษาอินเดีย ภาษามาลาเลย์และภาษาไทย [8-10] อย่างไรก็ตามภาษามือแบบ ASL มีความสำคัญอย่างมาก ถูกใช้และดัดแปลงไปใช้กับภาษาอื่นๆ อีกมากมาย ดังนั้นจึงมีผู้ศึกษาวิจัยเพื่อแปลภาษามืออย่างอัตโนมัติ

สำหรับงานวิจัยของ Ensoek Jong และคณะ [11] นำเสนอวิธีการแปลภาษามือด้วยวิธีการประยุกต์ใช้ถุงมือที่มีไมโครคอมพิวเตอร์และบลูทูธ (Bluetooth) ประกอบอยู่บนถุงมือ ซึ่งผู้วิจัยใช้ชื่อว่า “Velostat” ดังรูปที่ 2.3 แต่อย่างไรก็ตามด้วยลักษณะดังกล่าวของถุงมือทำให้เกิดความไม่สะดวกในการเคลื่อนไหวมือเพราะขาดความยืดหยุ่นจากการใช้งาน นอกจากนี้ยังมีงานวิจัยของ N. Tanyawiwat และ S. Thiemjarus [12] ได้ออกแบบถุงมือที่ติดตั้งเซนเซอร์ (Sensor) วัดความงอของแต่ละนิ้วมือ และเซนเซอร์วัดความเร่งแบบสามมิติ (Accelerometer) รวมทั้งติดตั้งโมดูลไวเลส (Module Wireless) ไว้ที่ด้านหลังถุงมือ และทำการวัดความถูกต้องแม่นยำของการแปลภาษาโดยผลความถูกต้องต่ำกว่าร้อยละ 80

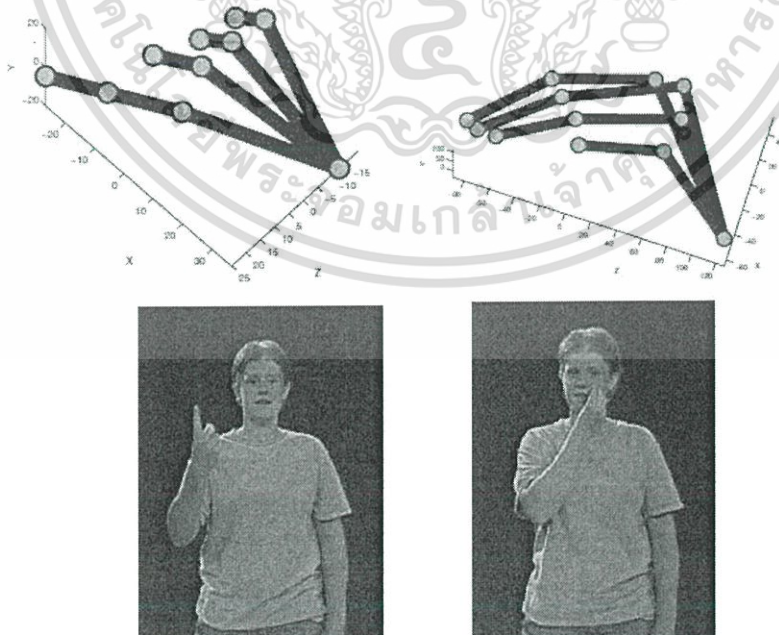


รูปที่ 2.3 ถุงมือที่ติดตั้งไมโครคอมพิวเตอร์เพื่อแปลภาษามือ [8]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในปี ค.ศ. 2012 M.S. Sinith และคณะ [13] นำเสนอวิธีการจดจำภาษามือด้วยวิธีการซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) โดยจำแนกตัวอักษรเพียง 6 ตัวอักษรเท่านั้น คือ A, H, I, O, L และ W นอกจากนั้น M. Ashrafal Amin และ Hong Yan [14] ได้นำเสนอวิธีการจดจำภาษามือในอีกรูปแบบหนึ่ง ด้วยการใช้เทคนิคการคำนวณหาหลักขณะเด่นแบบการวิเคราะห์ส่วนประกอบสำคัญ (Principal Components Analysis) และจำแนกท่าทางด้วยวิธีการของฟัซซีซีมีน (Fuzzy C-Means) นอกจากนั้นแล็บวิว (LabView) ถูกนำมาใช้ในงานวิจัยของ Andreas Domingo และคณะ [15] เพื่อจดจำภาษามือ A-Z ด้วยวิธีการคำนวณหาจุดศูนย์กลางมวลของแต่ละนิ้ว และปี ค.ศ. 2013 Rudy Hartanto และคณะ [16] นำเสนอการแปลภาษามือพยัญชนะอังกฤษ (ยกเว้น J และ Z) โดยใช้เทคนิคคอนทัวร์ (Contour) เพื่อหาขอบเขตของมือ และคำนวณลักษณะสำคัญด้วยวิธีการ SIFT

นอกจากนั้น เทคนิคที่น่าสนใจอีกแบบหนึ่งคือการคำนวณคุณลักษณะสำคัญด้วยวิธีการแบบสามมิติ ซึ่งถูกประยุกต์ใช้ในหลายงานวิจัย [17-21] เช่น งานวิจัยของ Liya Ding และ Aleix M. Martinez [17] นำเสนอการแปลงภาพสองมิติจากวิดีโอไปยังเป็นจุดบนระนาบสามมิติ ตามเส้นทางที่มือเคลื่อนไหว โดยใช้วิธีการติดตามข้อมือ ดังรูปที่ 2.4 และในปี ค.ศ. 2013 Myoung-Kyu Sohn และคณะ [20] ประยุกต์ใช้กล้องวัดระยะลึก เพื่อคำนวณเส้นทางการเคลื่อนไหวของมือในสามมิติและนอมอลไลซ์เพื่อทำให้เกิดคุณลักษณะเฉพาะที่ไม่แปรผันตามเรขาคณิต นำมาแปลผลเป็นตัวอักษรด้วยวิธีการแยกแยะแบบ K-NN



รูปที่ 2.4 การแปลงภาพสองมิติไปยังพิกัดสามมิติโดยการติดตามข้อมือ [17]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับงานวิจัยนี้นำเสนอวิธีการแปลภาษาเมื่อพยัญชนะภาษาอังกฤษ โดยคำนวณลักษณะเฉพาะที่ไม่แปรผันตามเรขาคณิตของแต่ละท่าทางแบบ 3 มิติ ด้วยวิธีการของ Direct Linear Transformation และการแบ่งประเภทของท่าทางด้วยวิธีการของโครงข่ายประสาทเทียม (Artificial Neural Network) ซึ่งทฤษฎีและขั้นตอนจะกล่าวถึงในหัวข้อถัดไป

2.5 ทฤษฎีที่เกี่ยวข้องกับงานวิจัย

2.5.1 การปรับเทียบกล้อง (Camera Calibration)

การปรับเทียบกล้องเป็นการประมาณค่าโปรเจกชันเมทริกซ์ (Projection Matrix : M) จากจุดในฉากสามมิติที่รู้ตำแหน่งและภาพที่เกิดขึ้นของจุด จากนั้นทำการหาพารามิเตอร์ภายใน (Intrinsic) และพารามิเตอร์ภายนอก (Extrinsic) โดยพิจารณาปรับเทียบกล้องทีละตัว [22]

ในการหาค่าของเมทริกซ์ M จะทำการหาจุดในฉากที่เราทราบพิกัด $X=[x,y,z]^T$ และจุดในภาพสองมิติที่สอดคล้องกัน $[u,v]^T$ ดังนี้

$$\begin{bmatrix} \alpha u \\ \alpha v \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.1)$$

$$\begin{bmatrix} \alpha u \\ \alpha v \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11}x + m_{12}y + m_{13}z + m_{14} \\ m_{21}x + m_{22}y + m_{23}z + m_{24} \\ m_{31}x + m_{32}y + m_{33}z + m_{34} \end{bmatrix} \quad (2.2)$$

จากนั้นทำการนอร์มอลไลซ์ (Normalize) ให้สมาชิกที่สามเป็นหนึ่ง จะได้ตามสมการที่ 2.3 และ 2.4

$$u(m_{31}x + m_{32}y + m_{33}z + m_{34}) = m_{11}x + m_{12}y + m_{13}z + m_{14} \quad (2.3)$$

$$v(m_{31}x + m_{32}y + m_{33}z + m_{34}) = m_{21}x + m_{22}y + m_{23}z + m_{24} \quad (2.4)$$

ดังนั้นจะได้สมการเชิงเส้นสองสมการ สำหรับจุดในฉากสามมิติหนึ่งจุดและจุดในภาพสองมิติที่สอดคล้องกัน โดยสามารถนำสมการที่ 2.3 และ 2.4 มาเขียนในรูปแบบเมทริกซ์ได้ดังสมการที่ 2.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -ux & -uy & -uz & -u \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -vx & -vy & -vz & -v \\ \vdots & & & & & & & & & & & \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{34} \end{bmatrix} = 0 \quad (2.5)$$

หรือเขียนเป็นสมการเชิงเส้นได้ดังสมการที่ 2.6 ว่า

$$AM = 0 \quad (2.6)$$

ในการแก้สมการโฮโมจีเนียส (Homogeneous) สามารถใช้วิธี Least Square หรือใช้การคำนวณหาแบบ Singular Value Decomposition (SVD) ของ A เมื่อได้เมทริกซ์ M แล้วขั้นตอนถัดไปทำการแยกพารามิเตอร์ภายในได้ เนื่องจากสมการ M เป็นไปตามสมการที่ 2.7

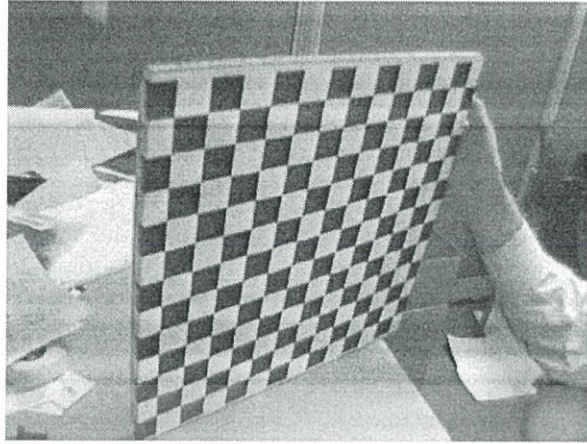
$$M = [KR|-KRT] = [A|b] \quad (2.7)$$

ดังนั้นเมทริกซ์ย่อยภายในขนาด 3×3 แทนด้วย A และเวกเตอร์ด้านขวามือคือ b ซึ่งภายในเวกเตอร์ b แทนด้วย $t = -A^{-1}b$ จะประกอบด้วยค่าพารามิเตอร์ระยะเลื่อน (Translate Parameter แทนด้วย t) โดยสามารถคำนวณค่า ซึ่งค่า t จะบ่งบอกถึงตำแหน่งกึ่งกลางของระนาบภาพ

จากนั้นพิจารณาเมทริกซ์ย่อย $A = KR$ โดย K จะเป็นเมทริกซ์สามเหลี่ยมบน (Upper Triangle) และ R เป็นเมทริกซ์ออร์โธโกนอล โดยสามารถแยก K และ R ออกจากกันด้วยวิธีการ QR composite สำหรับ A

2.5.2 การหาพิกัดสามมิติของจุดในฉากสองมิติ

การหาพิกัดสามมิติจากกล้องสองตัว โดยกล้องทั้งสองตัวต้องผ่านกระบวนการปรับเทียบกล้องตามกระบวนการในหัวข้อที่ 2.5.1 โดยในกระบวนการการปรับเทียบกล้องจะต้องถ่ายภาพฉากสามมิติที่ทราบค่า เช่นตารางหมากรุกที่ทราบขนาดของตารางแต่ละช่อง ดังรูปที่ 2.5 จากนั้นหาพิกัดที่จุดตัดต่างๆของแต่ละช่อง เพื่อทำการหา Perspective Transformation M ของแต่ละกล้อง หลังจากนั้นนำวัตถุเข้าไปแทนที่ตารางหมากรุก และหาจุดสอดคล้องในฉากสามมิติของแต่ละกล้อง



รูปที่ 2.5 การเปรียบเทียบกล้องด้วยการใช้ตารางหมากรุก [23]

กำหนดให้เมทริกซ์ Perspective Transformation ของกล้องที่ 1 และ 2 มีค่า M และ M' ตามลำดับ และแทนค่าแต่ละแถวของเมทริกซ์คือ m_1^T, m_2^T, m_3^T สำหรับกล้องที่ 1 และ $m_1'^T, m_2'^T, m_3'^T$ สำหรับกล้องที่ 2 การคำนวณสำหรับกล้องที่ 1 เป็นไปตามสมการที่ 2.8

$$u = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = MX = \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix} X \quad (2.8)$$

และคำนวณสำหรับกล้องที่ 2 เป็นไปตามสมการที่ 2.9

$$u' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = M'X = \begin{bmatrix} m_1'^T \\ m_2'^T \\ m_3'^T \end{bmatrix} X \quad (2.9)$$

โดย u และ u' แทนพิกัดในภาพของวัตถุจากกล้องที่ 1 และ 2 ตามลำดับ และวิธีการกำจัดค่าสเกลที่ไม่ทราบค่า สามารถทำได้ด้วยการหาสัดส่วนระหว่างสามแถวในเมทริกซ์ ดังสมการที่ 2.10 และ 2.11

$$u : v : w = m_1^T X : m_2^T X : m_3^T X \quad (2.10)$$

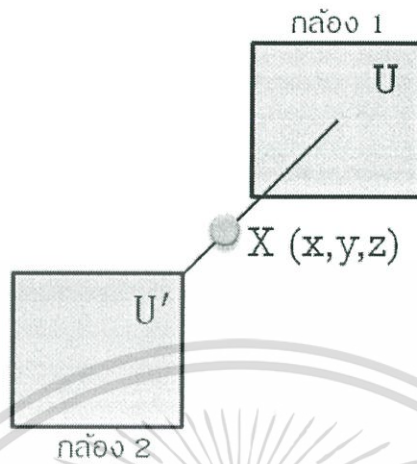
$$u' : v' : w' = m_1'^T X : m_2'^T X : m_3'^T X \quad (2.11)$$

และสามารถเขียนสมการโฮโมจีเนียสได้ดังสมการที่ 2.12

$$\begin{bmatrix} um_3^T - wm_1^T \\ vm_3^T - wm_2^T \\ u'm_3'^T - w'm_1'^T \\ v'm_3'^T - w'm_2'^T \end{bmatrix} X = AX = 0 \quad (2.12)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำการแก้สมการ 2.12 ในรูปแบบโฮโมจีเนียส สำหรับพิกัดสามมิติ เพื่อหาคำคำตอบของ X ซึ่งเป็นคำตอบเฉพาะประเภท Non-Trivial หรือ $\det(A)=0$



รูปที่ 2.6 แบบจำลองการหาพิกัดสามมิติจากกล่องสองตัว

2.5.3 การแปลงวงกลมของฮัฟ (Circle Hough Transform)

การแปลงวงกลมของฮัฟ คือการคำนวณจากระนาบสองมิติ (x,y) ไปยังระบบตัวแปรในระนาบสามมิติ (a, b, r) โดย a, b คือจุดศูนย์กลางของวงกลม และ r คือรัศมีของวงกลม [24] สมการที่ 2.13 แสดงสมการวงกลมของรูปที่ 2.7

$$(x - a)^2 + (y - b)^2 = r^2 \quad (2.13)$$

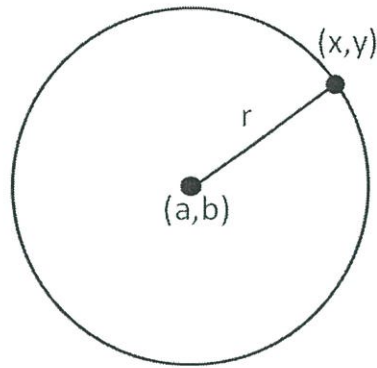
และค่า x, y เป็นไปตามสมการที่ 2.14 และ 2.15 ตามลำดับ

$$x = a + r \cos \theta \quad (2.14)$$

$$y = b + r \sin \theta \quad (2.15)$$

วิธีการแปลงวงกลมของฮัฟ ทำการสร้างวงกลมที่เหมาะสมกับจุดต่างๆ บนขอบวัตถุในภาพ ซึ่งกระบวนการนี้จะใช้วิธีการโหวตหรือหาค่าอาร์เรย์สะสม (Accumulator Array) แทนด้วย $M[a, b]$ ที่มีค่าสูงสุด หากกระบวนการแปลงนี้กำหนดค่ารัศมีวงกลม r [25] จะสามารถอธิบายวิธีการแปลงได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.7 ส่วนประกอบตัวแปรต่างๆของวงกลม

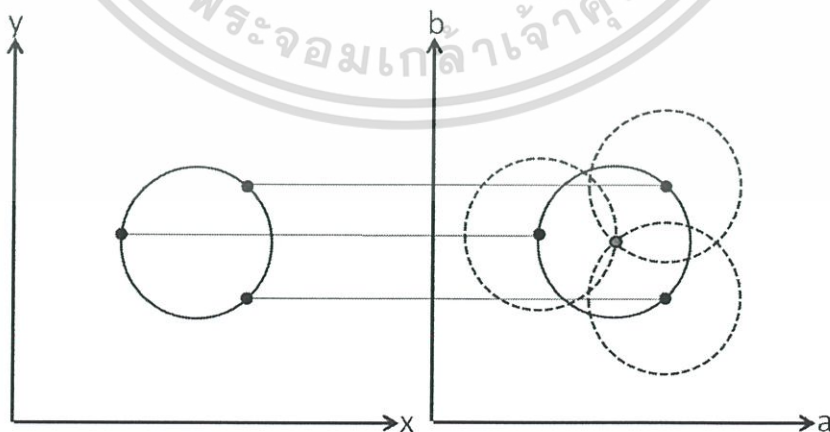
1. กำหนดปริมาณสำหรับพารามิเตอร์ a และ b
2. กำหนดค่าเริ่มต้นให้กับอาร์เรย์สะสม ($M[a, b]$) จากนั้นคำนวณหาขนาด $G[x, y]$ และทิศทางแกรเดียนต์ $Q[x, y]$
3. วาดวงกลมตามสมการ 2.16 และ 2.17 เพื่อคำนวณหาค่า a และ b

$$a = x - r \cos \theta \quad (2.16)$$

$$b = y - r \sin \theta \quad (2.17)$$

โดย x, y คือพิกัดของขอบวัตถุ และ $0 < \theta < 360$

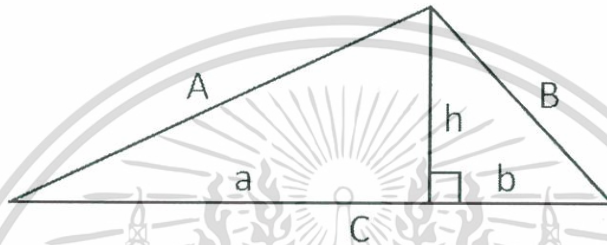
4. วาดวงกลมในพื้นที่สะสม (Accumulator Space) ไปยังทุกจุดบนขอบภาพ (x,y) ดังตัวอย่างในรูปที่ 2.8 โดยจะเกิดการสะสมค่าเพิ่มมากขึ้น ณ ตำแหน่งที่วาดวงกลมที่ $M[a, b]$
5. เลือกค่า $M[a, b]$ ที่มีค่ามากที่สุด และกำหนดให้ a, b เป็นจุดศูนย์กลางของวงกลม โดยมีรัศมีเท่ากับ r

รูปที่ 2.8 การแปลงวงกลมของฮัฟจากระนาบ x, y ไปเป็นระนาบ a, b, r

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.4 ทฤษฎีการคำนวณพื้นที่สามเหลี่ยมใดๆ ของฮีรอน

วิธีการคำนวณหาพื้นที่สามเหลี่ยมที่มีความเฉพาะเจาะจง เช่น สามเหลี่ยมด้านเท่า สามเหลี่ยมหน้าจั่ว ฯลฯ อย่างไรก็ตามวิธีการคำนวณพื้นที่สามเหลี่ยมใดๆ มีประโยชน์อย่างมากกับงานวิจัยที่วิเคราะห์พื้นที่สามเหลี่ยมที่มีลักษณะไม่แน่นอน ซึ่งการคำนวณพื้นที่สามเหลี่ยมใดๆ ของฮีรอน (Heron's Formula) ถูกนำมาใช้อย่างแพร่หลาย โดยสมการเป็นไปตามสมการที่ 2.39 และ 2.40 ซึ่งสัมพันธ์กับรูปที่ 2.9 ซึ่งมีการแสดงที่มาของสมการดังสมการที่ 2.18 ถึง 2.40



รูปที่ 2.9 รูปสามเหลี่ยมใดๆ

จากรูปที่ 2.8 สามารถนำมาสร้างสมการพื้นฐานได้ดังนี้

$$A^2 = h^2 + a^2 \quad (2.18)$$

$$B^2 = h^2 + b^2 \quad (2.19)$$

$$C = a + b \quad (2.20)$$

นำสมการที่ 2.18 และ 2.19 มาลบกัน จะได้ว่า

$$A^2 - B^2 = (a - b)(a + b) \quad (2.21)$$

แทนสมการที่ 2.20 ลงใน สมการ 2.21

$$\frac{A^2 - B^2}{C} = a - b \quad (2.22)$$

นำสมการที่ 2.20 บวกกับสมการที่ 2.22 จะได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$C + \frac{A^2 - B^2}{C} = 2a \quad (2.23)$$

จัดรูปสมการที่ 2.23 โดยการนำ C คูณตลอด

$$2aC = A^2 - B^2 + C^2 \quad (2.24)$$

จากนั้น ยกกำลังสองในสมการที่ 2.24 ให้เกิดค่า a^2 คล้ายกับสมการที่ 2.18

$$4a^2C^2 = (A^2 - B^2 + C^2)^2 \quad (2.25)$$

นำสมการที่ 2.18 คูณด้วย $4C^2$ จะได้

$$4A^2C^2 = 4h^2C^2 + 4a^2C^2 \quad (2.26)$$

นำค่าจากสมการที่ 2.25 แทนลงใน 2.26 ได้ว่า

$$4A^2C^2 = 4h^2C^2 + (A^2 - B^2 + C^2)^2 \quad (2.27)$$

เนื่องจากพื้นที่สามเหลี่ยมมีค่าเท่ากับ $\frac{hC}{2}$ ดังนั้นต้องทำการจัดรูปแบบสมการที่ 2.27 ใหม่ ดังนี้

$$4h^2C^2 = 4A^2C^2 - (A^2 - B^2 + C^2)^2 \quad (2.28)$$

$$4h^2C^2 = (2AC + A^2 - B^2 + C^2)(2AC - A^2 + B^2 - C^2) \quad (2.29)$$

$$4h^2C^2 = [(A + C)^2 - B^2][B^2 - (A - C)^2] \quad (2.30)$$

$$4h^2C^2 = (A + C + B)(A + C - B)(B + A - C)(B - A + C) \quad (2.31)$$

$$4h^2C^2 = (A + B + C)(-A + B + C)(A - B + C)(A + B - C) \quad (2.32)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากนั้น ถอดรากสมการที่ 2.32

$$2hC = \sqrt{(A+B+C)(-A+B+C)(A-B+C)(A+B-C)} \quad (2.33)$$

แต่เนื่องจากพื้นที่สามเหลี่ยมเท่ากับ

$$\frac{hC}{2} = \frac{1}{4} \sqrt{(A+B+C)(-A+B+C)(A-B+C)(A+B-C)} \quad (2.34)$$

กำหนดให้ $s = \frac{A+B+C}{2}$ จะได้ว่า

$$2s = A + B + C \quad (2.35)$$

$$2(s - A) = -A + B + C \quad (2.36)$$

$$2(s - B) = A - B + C \quad (2.37)$$

$$2(s - C) = A + B - C \quad (2.38)$$

ดังนั้นจึงได้สมการคำนวณพื้นที่สามเหลี่ยมใดๆ ดังสมการที่ 2.39

$$A = \sqrt{s(s-A)(s-B)(s-C)} \quad (2.39)$$

และ

$$S = \frac{A+B+C}{2} \quad (2.40)$$

2.5.5 การจำแนกข้อมูลแบบโครงข่ายประสาทเทียม

การพัฒนาการจำแนกข้อมูลแบบปัญญาประดิษฐ์นั้น [22] ถูกพัฒนากระบวนการมาอย่างต่อเนื่อง เพื่อนำมาประยุกต์ใช้กับการจำแนกกลุ่มข้อมูล โครงข่ายประสาทเทียมเป็นวิธีหนึ่งของปัญญาประดิษฐ์ ซึ่งลอกเลียนแบบเซลล์ประสาทของสิ่งมีชีวิต ประกอบด้วยส่วนสำคัญ 3 ส่วน ได้แก่ ตัวรับกระแสประสาท (Dendrite) ตัวส่งกระแสประสาท (Axon) และตัวเซลล์ (Soma) ตัวรับกระแสประสาทรับสัญญาณแบบอิมพัลส์ที่ส่งผ่านช่องไซแนป (Synaptic Gap) ระหว่างเซลล์ ด้วยกระบวนการทางเคมี ซึ่งมีความแรงของสัญญาณไม่เท่ากันขึ้นกับกระบวนการทางเคมี โดยตัวเซลล์ทำ

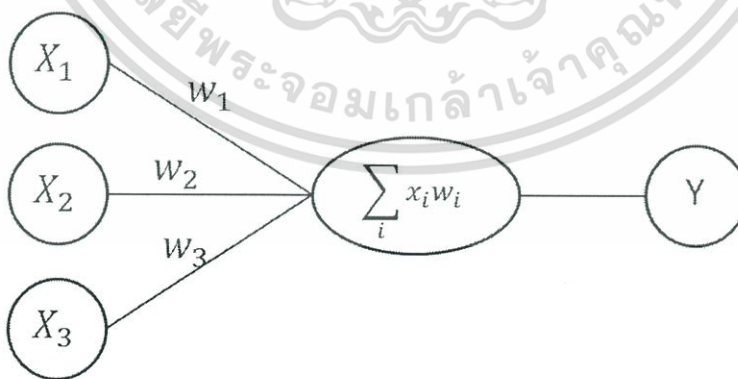
หน้าที่รวมสัญญาณประสาทที่ผ่านเข้ามา หากความแรงมีค่ามากกว่าระดับกั้น ตัวเซลล์จะส่งสัญญาณประสาทต่อไปทางตัวส่งกระแสประสาทดังรูปที่ 2.10 ซึ่งแสดงรูปการทำงานของระบบประสาท



รูปที่ 2.10 ส่วนประกอบของเซลล์ประสาท [3]

องค์ประกอบพื้นฐานของแบบจำลองโครงข่ายประสาทเทียมประกอบด้วยนิวรอน X_1 , X_2 และ X_3 ซึ่งจะส่งสัญญาณ x_1 , x_2 และ x_3 ไปยังนิวรอน Y ด้วยสัดส่วนของน้ำหนักที่ไม่เท่ากัน คือ w_1 , w_2 และ w_3 ถ้ากำหนด y_{input} แทนสัญญาณรวมของนิวรอน Y ซึ่งเป็นไปตามสมการ 2.41 ซึ่งมีความสัมพันธ์กับรูปที่ 2.11

$$y_{input} = \sum_i x_i w_i \quad (2.41)$$

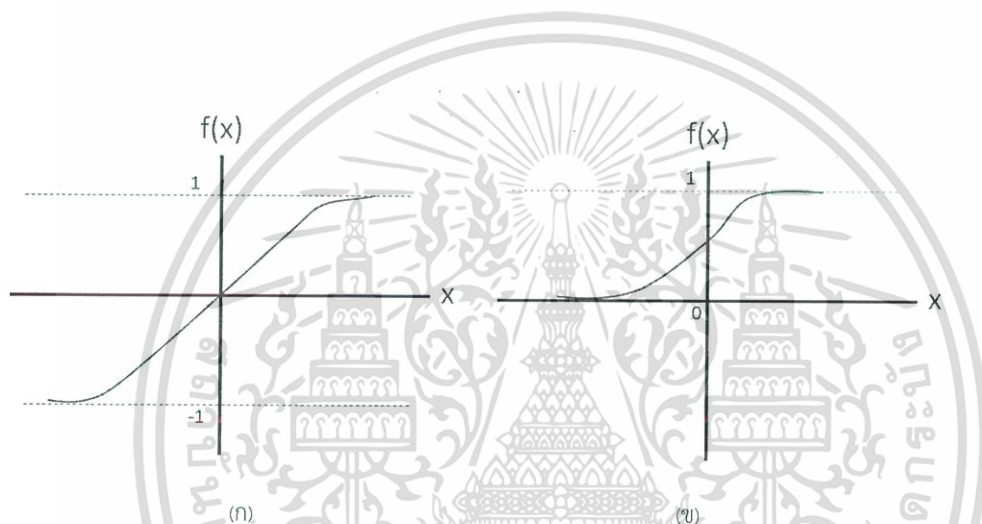


รูปที่ 2.11 โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชันของสัญญาณสุทธิ แทนด้วย $y = f(y_{\text{input}})$ เป็นค่าของอินพุตคูณกับน้ำหนักตามอัตราส่วนและส่งไปยังเอาต์พุต (Output) เรียกว่า ฟังก์ชันกระตุ้น (Activation Function) ซึ่งมีหลากหลายรูปแบบ ซึ่งรูปแบบที่สำคัญชนิดหนึ่งคือ ไฮเพอร์ซิกมอยด์ (Bipolar Sigmoid) ดังรูปที่ 2.12 ได้แก่ ฟังก์ชัน Logistic และ Hyperbolic Tangent ส่วนใหญ่จะใช้ในการฝึกฝนโครงข่ายประสาทเทียมแบบ Backpropagation เนื่องจากให้ค่าความสัมพันธ์อย่างง่ายและลดเวลาในการฝึกฝนลง

ฟังก์ชัน Logistic เป็นฟังก์ชันที่มีค่าระหว่าง 0 และ 1 มักใช้เป็นฟังก์ชันกระตุ้นสำหรับโครงข่ายประสาทเทียมที่ต้องการให้อาไรท์พุทมีค่าเป็นไบนารี ส่วนฟังก์ชัน Hyperbolic Tangent จะมีย่านในช่วง -1 และ 1

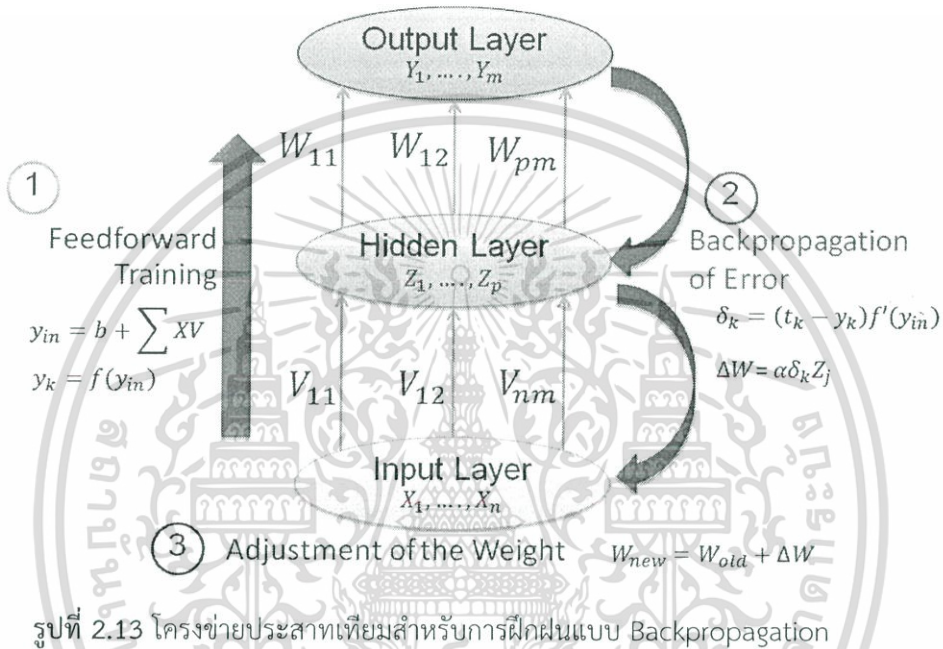


รูปที่ 2.12 ก. ฟังก์ชัน Hyperbolic Tangent ข. ฟังก์ชัน Logistic

การฝึกฝนโครงข่ายประสาทเทียม เป็นวิธีการหาค่าน้ำหนักที่เหมาะสมสำหรับอัลกอริทึม โดยอัลกอริทึมที่มักนำมาใช้ในการฝึกฝนคือ การแพร่แบบย้อนกลับ (Backpropagation Training) ซึ่งเป็นวิธีการหาค่าต่ำสุดและสูงสุดแบบ Gradient Descent คือการหาค่าผิดพลาดต่ำสุดยกกำลังสองของเอาต์พุตที่ถูกคำนวณด้วยการฝึกฝนแบบแพร่ย้อนกลับ ซึ่งการฝึกฝนมี 3 ขั้นตอน คือ การฝึกฝนแบบไปข้างหน้า (Feedforward The Training) การคำนวณค่าผิดพลาดแบบแพร่ย้อนกลับ (Backpropagation of the Error) และการปรับค่าน้ำหนัก (Adjustment of The Weight) ดังรูปที่ 2.13

ขั้นตอนแรกคือการฝึกฝนแบบไปข้างหน้า เริ่มจากชั้นอินพุต (Input layer) โดย X_1, \dots, X_n รับสัญญาณอินพุตและถ่ายทอดไปยังชั้นซ่อนเร้น (Hidden layer) ที่ Z_1, \dots, Z_p ที่ชั้นนี้ ทำการคำนวณค่าการกระตุ้นและส่งสัญญาณ z_j ต่อไปที่แต่ละหน่วยของเอาต์พุต จากนั้นทุกหน่วยของชั้นเอาต์พุตคำนวณค่าการตอบสนอง y_k เพื่อหาค่าการตอบสนองต่อรูปแบบอินพุตระบบที่ได้รับ

สำหรับการฝึกฝนในชั้นเอาวท์พุท ค่าของเอาวท์พุทที่คำนวณได้จะถูกเปรียบเทียบกับเอาวท์พุทเป้าหมาย t_k เพื่อหาค่าผิดพลาดที่เกิดขึ้นของแต่ละหน่วย โดยแฟคเตอร์ของค่าผิดพลาดจะถูกแทนด้วย δ_k โดยค่า $k = 1, \dots, m$ จากนั้นค่าความผิดพลาดในแต่ละหน่วยจะถูกนำไปคำนวณเพื่อปรับแก้น้ำหนักในชั้นซ่อนเร้น ในทำนองเดียวกัน δ_j โดยค่า $j = 1, \dots, p$ ซึ่งเป็นค่าความผิดพลาดในชั้นซ่อนเร้น โดยกระบวนการเหล่านี้จะถูกกระทำซ้ำๆ จนกระทั่งได้ค่าตามเงื่อนไขที่ตั้งไว้



โดยรายละเอียดและขั้นตอนของอัลกอริทึมของโครงข่ายประสาทเทียม ที่ใช้การฝึกฝนแบบ Feedforward Backpropagation มีดังต่อไปนี้

ขั้นที่ 0 กำหนดค่าเริ่มต้นน้ำหนัก ปกติเป็นเลขที่สุ่มขึ้นมา

ขั้นที่ 1 ขณะที่เงื่อนไขการหยุดเป็นเท็จ ให้ทำตามขั้นที่ 2-9

ขั้นที่ 2 สำหรับแต่ละเวกเตอร์การฝึกฝน ให้ทำตามขั้นตอนที่ 3-8

สำหรับขั้นตอน Feedforward เป็นไปตามขั้นตอนที่ 3 - 5

ขั้นที่ 3 ทุกหน่วยของชั้นอินพุท X_i โดย $i = 1, \dots, n$ รับสัญญาณอินพุท x_i ในแต่ละโหนดไปยังชั้นซ่อนเร้น

ขั้นที่ 4 แต่ละหน่วยของชั้นซ่อนเร้น Z_j โดย $j = 1, \dots, p$ คำนวณสัญญาณอินพุตแต่ละหน่วยจากสมการ 2.42

$$z_{input} = v_{0j} + \sum_i x_i v_{ij} \quad (2.42)$$

โดยจะเขียนเป็นอินพุตสุทธิได้ ดังนี้ $z_j = f(z_{input})$

ขั้นที่ 5 แต่ละหน่วยของเอาวาทพุท Y_k โดย $k = 1, \dots, m$ คำนวณสัญญาณอินพุตแต่ละหน่วยจากสมการที่ 2.43

$$y_{input} = w_{0k} + \sum_i z_i w_{jk} \quad (2.43)$$

โดยจะเขียนเป็นอินพุตสุทธิได้ ดังนี้ $y_k = f(y_{input})$

สำหรับขั้นตอน Backpropagation of error เป็นไปตามขั้นตอนที่ 6 - 7

ขั้นที่ 6 คำนวณค่าความผิดพลาดของแต่ละหน่วยเอาวาทพุท Y_k โดย $k = 1, \dots, m$ เทียบกับค่าเป้าหมาย t_k ของแต่ละหน่วยที่กำหนดไว้ ตามสมการ 2.44

$$\delta_k = (t_k - y_k) f'(y_{input}) \quad (2.44)$$

คำนวณเทอมของการแก้ไขน้ำหนักแต่ละหน่วย เพื่อนำไปใช้ปรับค่า w_{jk} ในภายหลัง ด้วยสมการ (α คือค่าอัตราการเรียนรู้ของการฝึกฝน) ซึ่งคำนวณตามสมการที่ 2.45

$$\Delta w_{jk} = \alpha \delta_k z_j \quad (2.45)$$

คำนวณเทอมการแก้ไขของไบอัส (Bias) เพื่อปรับค่าของ w_{0k} ในภายหลัง ด้วยสมการที่ 2.46

$$\Delta w_{0k} = \alpha \delta_k \quad (2.46)$$

ขั้นที่ 7 สำหรับแต่ละหน่วยในชั้นซ่อนเร้น Z_j โดย $j = 1, \dots, p$ ทำการรวมข้อมูลจาก δ_k ที่ได้จากชั้นบน หรือชั้นที่อยู่เหนือขึ้นไป ตามสมการที่ 2.47

$$\delta_{input} = \sum_{k=1}^m \delta_k w_{jk} \quad (2.47)$$

นำค่าที่ได้คูณกับอนุพันธ์ของฟังก์ชันกระตุ้น เพื่อใช้คำนวณค่าผิดพลาด ดังสมการที่ 2.48

$$\delta_j = \delta_{input} f'(z_{input}) \quad (2.48)$$

คำนวณเทอมของการแก้ไขน้ำหนัก เพื่อใช้ปรับค่า v_{ij} ในภายหลัง ตามสมการที่ 2.49

$$\Delta v_{jk} = \alpha \delta_j z_i \quad (2.49)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำนวณเทอมการแก้ไขของไบอัส เพื่อปรับค่าของ v_{0j} ในภายหลัง ด้วยสมการ 2.50

$$\Delta v_{0j} = \alpha \delta_j \quad (2.50)$$

สำหรับขั้นตอนการปรับค่าน้ำหนักและค่าไบอัส เป็นไปตามขั้นตอนที่ 8 - 9

ขั้นที่ 8 แต่ละหน่วยของเอวาทพุท Y_k โดย $k = 1, \dots, m$ ปรับค่าน้ำหนักและไบอัส ดังนี้

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk} \quad (2.51)$$

และแต่ละหน่วยของเอวาทพุท Z_j โดย $j = 1, \dots, p$ ปรับค่าน้ำหนักและไบอัส ดังสมการที่ 2.52

$$v_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij} \quad (2.52)$$

ขั้นที่ 9 ทดสอบเงื่อนไขการหยุด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความไม่แปรผันต่อการแปลงทางเรขาคณิต (Geometric Invariance)

ความไม่แปรผันต่อการแปลงทางเรขาคณิตคือ คุณสมบัติของตัวอธิบายรูปร่าง (Shape Descriptor) ในเชิงเรขาคณิต เช่น คุณสมบัติของเส้นโค้งหรือพื้นผิวบนภาพที่ไม่มีการเปลี่ยนแปลง เมื่ออยู่ภายใต้การแปลงทางเรขาคณิตของภาพ (Geometric Transformation) การแปลงทางเรขาคณิตของภาพเป็นขบวนการพื้นฐานในการประมวลผลภาพดิจิทัล ซึ่งเป็นขบวนการที่เกี่ยวข้องกับการเคลื่อนย้ายภาพ (Translation) การหมุนภาพ (Rotation) การย่อ/ขยายภาพ (Scaling) การเฉือนภาพ (Shear) รวมทั้งการแปลงภาพแบบไม่เป็นเชิงเส้น (Non-Linear Transformation) โดยความไม่แปรผันต่อการแปลงเรขาคณิตสามารถจำแนกได้ตามประเภทของการแปลงทางเรขาคณิต ดังอธิบายในหัวข้อถัดไป

3.1 การแปลงทางเรขาคณิตเชิงเส้น (Geometric Linear Transformation)

การแปลงทางเรขาคณิต โดยทั่วไปมักกระทำในระบบพิกัดคาร์ทีเซียน (Cartesian) นั่นคือมีจุดกำเนิด (0,0) อยู่ที่มุมล่างซ้ายของภาพ โดยกำหนดให้ $g(u,v)$ ที่ $1 \leq u \leq P$ และ $1 \leq v \leq Q$ เป็นภาพเอาร์ทพุทที่ได้จากการแปลงทางเรขาคณิตของภาพอินพุท $f(x,y)$ โดยที่ $1 \leq x \leq M$ และ $1 \leq y \leq N$

3.1.1 การเคลื่อนย้ายภาพ

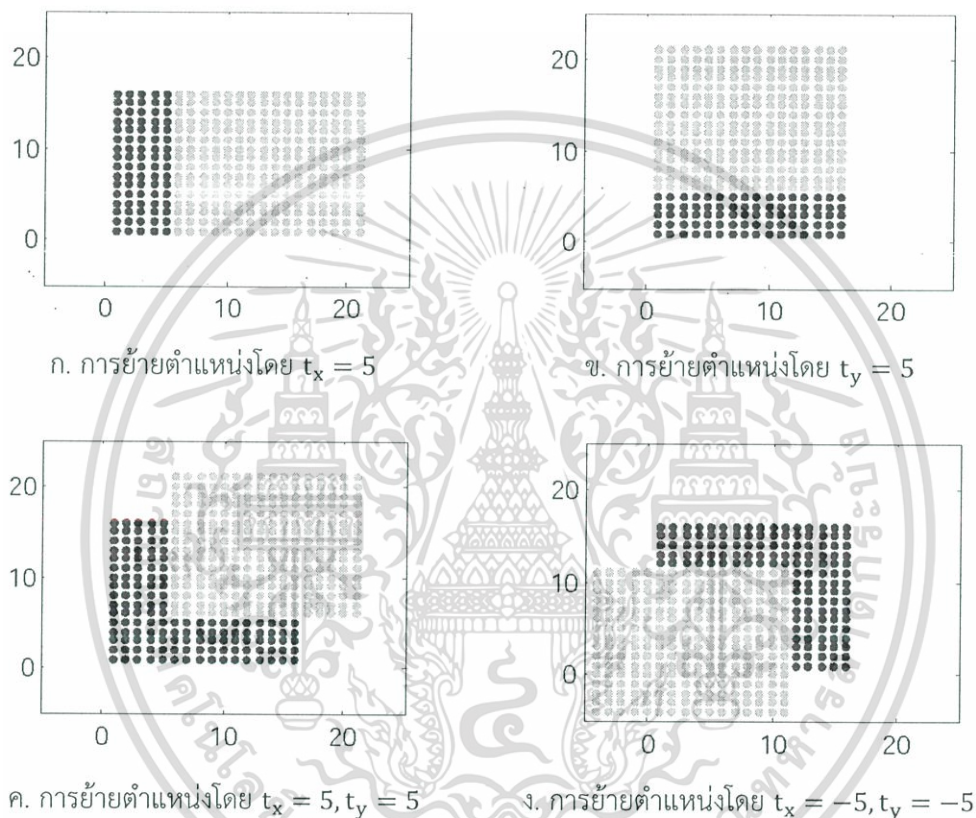
ความสัมพันธ์ระหว่างตำแหน่งอินพุท $f(x,y)$ กับภาพเอาร์ทพุทที่ถูกเคลื่อนย้ายไปยังตำแหน่ง $g(u,v)$ เป็นไปตามสมการที่ 3.1 และ 3.2

$$u = x + t_x \quad (3.1)$$

$$v = y + t_y \quad (3.2)$$

โดยที่ t_x และ t_y คือระยะของการย้ายตำแหน่ง สามารถเขียนอยู่สมการที่ 3.1 และ 3.2 และแทนสมการดังกล่าว ในรูปของเวกเตอร์ได้ ดังสมการที่ 3.3 ซึ่งได้แสดงตัวอย่างการเคลื่อนย้ายภาพแสดงในรูปที่ 3.1 โดยภาพสีดำคือภาพต้นฉบับ และภาพสีเทาคือภาพที่ย้ายตำแหน่ง

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (3.3)$$



รูปที่ 3.1 การเคลื่อนย้ายภาพ [26]

3.1.2 การย่อ/ขยาย ภาพ

การย่อหรือขยายขนาดของภาพ แสดงความสัมพันธ์ได้ดังสมการที่ 3.4 และ 3.5

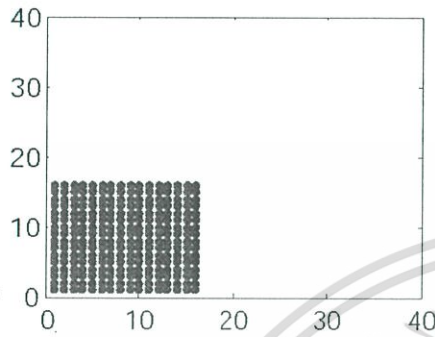
$$u = s_x x \quad (3.4)$$

$$v = s_y y \quad (3.5)$$

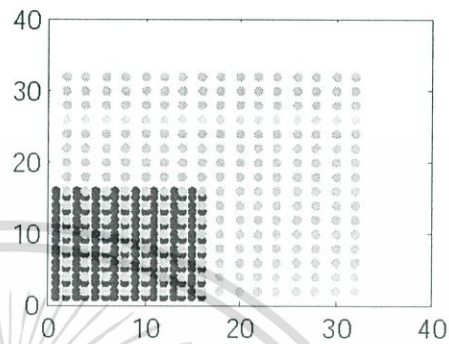
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และสามารถเขียนอยู่ในรูปของเวกเตอร์ได้ดังสมการที่ 3.6

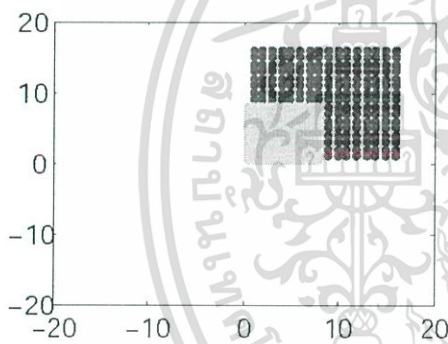
$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.6)$$



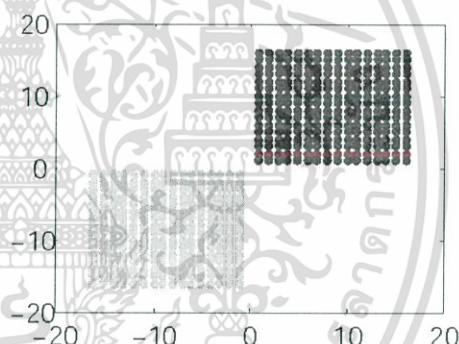
ก. ภาพต้นฉบับ



ข. การขยายขนาดของภาพ $s_x = 2, s_y = 2$



ค. การย่อขนาดของภาพ $s_x = 0.5, s_y = 0.5$



ง. การเกิด Reflection $s_x < 0$ และ $s_y < 0$

รูปที่ 3.2 การย่อ/ขยายภาพ [26]

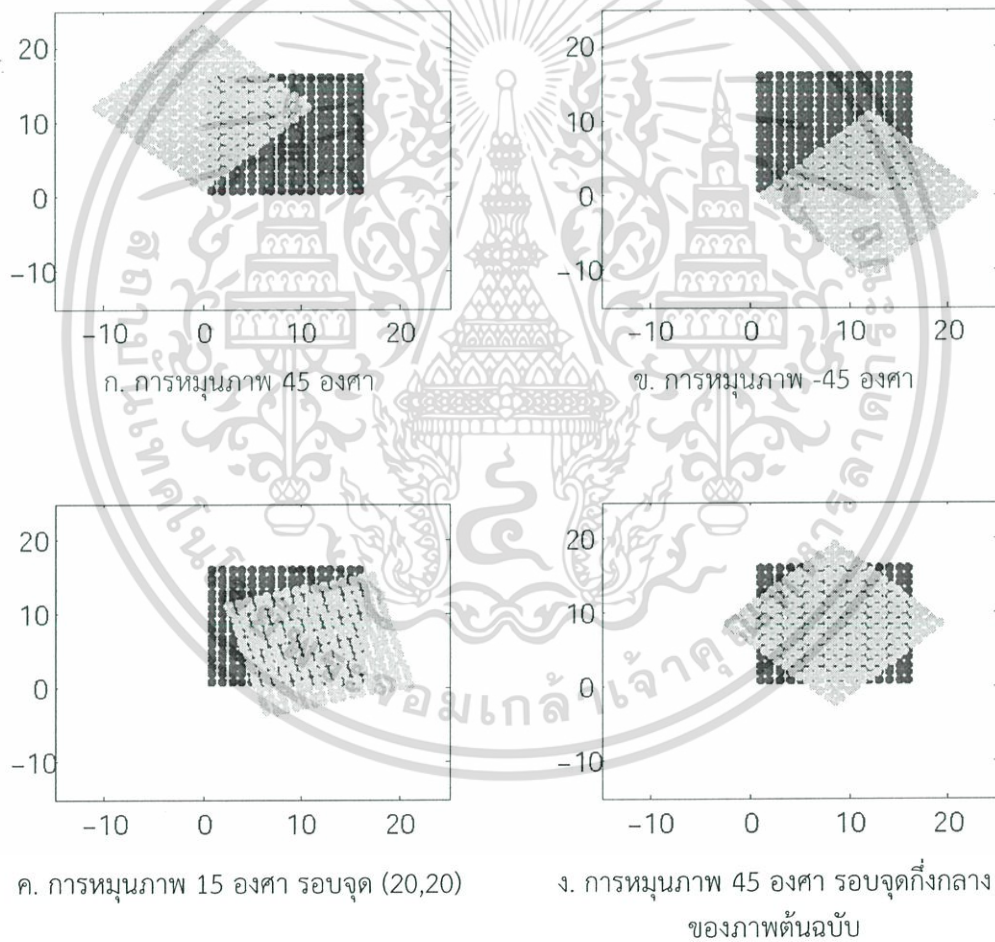
ตัวแปร s_x และ s_y คือ พารามิเตอร์ของการย่อ/ขยาย ภาพ (Scaling Parameter) ซึ่งพารามิเตอร์นี้มีผลต่อการขยายหรือย่อขนาดของภาพในทิศทาง x และ y ตามลำดับ ถ้าพารามิเตอร์ของการย่อ/ขยาย มีค่าอยู่ในช่วง 0 ถึง 1 หมายถึงการย่อภาพ ถ้าพารามิเตอร์ของการย่อ/ขยาย มีค่ามากกว่า 1 หมายถึงการขยายภาพ และถ้าพารามิเตอร์ของการย่อ/ขยายภาพ s_x และ s_y มีค่าน้อยกว่าศูนย์ จะเกิดลักษณะภาพแบบการสะท้อน (Reflection) คือภาพกลับซ้ายเป็นขวา ขวาเป็นซ้าย โดยถ้า $s_x < 0$ ภาพจะเกิด Reflection รอบแกน y และถ้า $s_y < 0$ ภาพจะเกิด Reflection รอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แกน x และถ้า $s_x < 0$ และ $s_y < 0$ ภาพจะเกิด Reflection ทั้งในแกน x และ y ตัวอย่างการย่อ/ขยาย ภาพแสดงดังรูปที่ 3.2 โดยภาพสีดำคือภาพต้นฉบับ และภาพสีเทาคือภาพที่ถูกย่อ/ขยาย

3.1.3 การหมุนภาพ

การหมุนภาพรอบจุดกำเนิดเป็นการหมุนจุด (x,y) รอบจุดกำเนิดไปที่จุด (u,v) ซึ่งมีความสัมพันธ์ดังสมการที่ 3.7 และ 3.8 และตัวอย่างการหมุนภาพแสดงดังรูปที่ 3.3 โดยภาพสีดำคือภาพต้นฉบับ และภาพสีเทาคือภาพที่หมุนไป



รูปที่ 3.3 การหมุนภาพ [26]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$u = x \cos \theta - y \sin \theta \quad (3.7)$$

$$v = x \sin \theta + y \cos \theta \quad (3.8)$$

โดย θ คือมุมของการหมุนในทิศทวนเข็มนาฬิกาเมื่อเทียบกับแกนนอน และสามารถเขียนอยู่ในรูปของเวกเตอร์ได้ ดังสมการที่ 3.9

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.9)$$

3.1.4 การเลื่อนภาพ

การเลื่อนภาพ หมายถึง การทำให้ภาพเปลี่ยนจากรูปร่างสี่เหลี่ยมผืนผ้า หรือสี่เหลี่ยมจัตุรัสไปเป็นสี่เหลี่ยมขนมเปียกปูน โดยกระทำดังสมการที่ 3.10 ถึง 3.13 โดยตัวอย่างการเลื่อนภาพแสดงดังรูปที่ 3.4 โดยภาพสีดำคือภาพต้นฉบับ และภาพสีเทาคือภาพที่ได้จากการเลื่อน

สมการของการเลื่อนไปข้างหน้าในทิศทาง x คือ

$$u = x + sh_x y, \quad v = y \quad (3.10)$$

สมการของการเลื่อนไปข้างหน้าในทิศทาง y คือ

$$v = y + sh_y x, \quad u = x \quad (3.11)$$

สมการของการเลื่อนไปข้างหน้าในทิศทาง x และ y

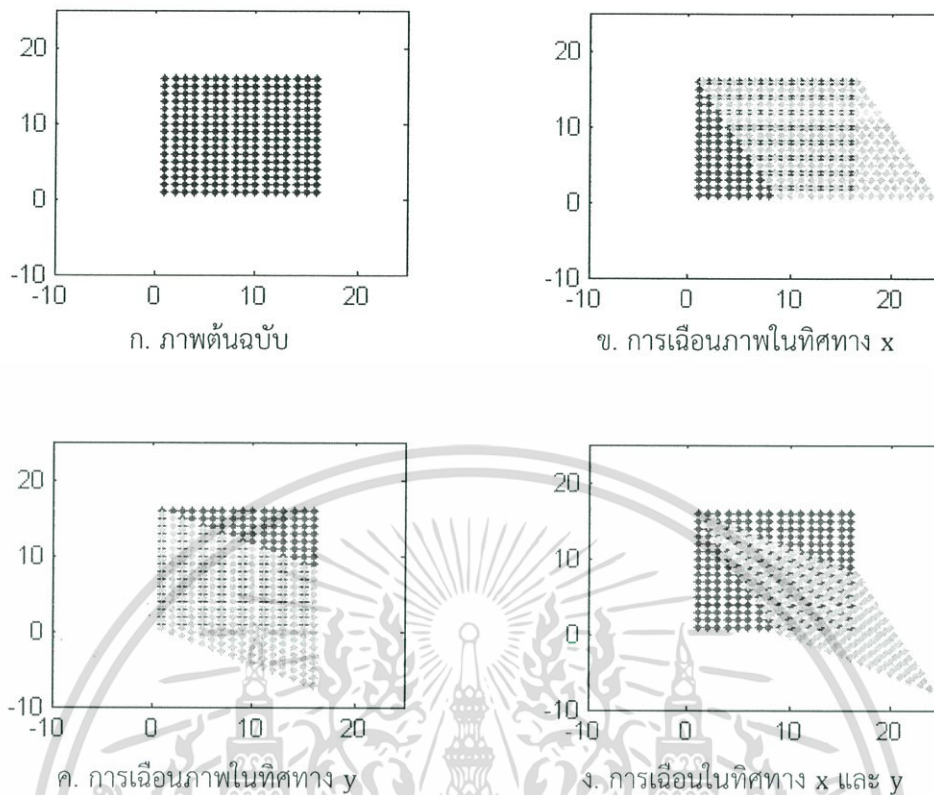
$$u = x + sh_x y \quad (3.12)$$

$$v = y + sh_y x \quad (3.13)$$

และสามารถเขียนอยู่ในรูปของเวกเตอร์ได้ดังสมการที่ 3.14

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & sh_x \\ sh_y & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.14)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 การเลื่อนภาพ [26]

3.1.5 ระบบพิกัดโฮโมจีเนียส (Homogeneous Coordinate)

ระบบพิกัดโฮโมจีเนียส คือ การกำหนดเวกเตอร์ขนาด 3 มิติ $(x_k, y_k, 1)$ ขึ้นมา แทนเวกเตอร์ตำแหน่งขนาด 2 มิติ โดยการแปลงเรขาคณิตในระบบพิกัดโฮโมจีเนียสเขียนแทนได้ดังต่อไปนี้

การแปลงเรขาคณิตของการเคลื่อนตำแหน่งภาพ ในระบบพิกัดโฮโมจีเนียส แทนด้วยสมการที่ 3.15

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3.15}$$

การแปลงเรขาคณิตของการย่อ/ขยายภาพ ในระบบพิกัดโฮโมจีเนียส เขียนได้ดังสมการที่

3.16

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.16)$$

การแปลงเรขาคณิตของการหมุน ในระบบพิกัดโฮโมจีเนียส เขียนได้ตามสมการที่ 3.17

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.17)$$

การแปลงเรขาคณิตของการเอียง ในระบบพิกัดโฮโมจีเนียส เขียนได้ดังสมการที่ 3.18

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & sh_x & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.18)$$

การเขียนเมทริกซ์การแปลงให้อยู่ในระบบโฮโมจีเนียส ทำให้สามารถหาเมทริกซ์การแปลงรวมได้ง่ายขึ้น โดยการแปลงทางเรขาคณิตรวม เป็นไปตามสมการที่ 3.19 ซึ่งเป็นการนำตัวแปรต่างๆ ทั้งหมด ได้แก่ การเคลื่อนตำแหน่ง การย่อ/ขยายภาพ การหมุนและการเอียงภาพ มาคูณกัน

$$T = S \cdot Sh \cdot R \cdot Tr \quad (3.19)$$

โดยที่ S คือ การย่อ/ขยาย ในระบบพิกัดโฮโมจีเนียส

Sh คือ การเอียงในระบบพิกัดโฮโมจีเนียส

R คือ การหมุนในระบบพิกัดโฮโมจีเนียสรอบจุดหมุน (x_c, y_c) ตามสมการที่ 3.20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$R = \begin{bmatrix} 1 & 0 & x_c \\ 0 & 1 & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{bmatrix} \quad (3.20)$$

Tr คือ การเคลื่อนย้ายในระบบพิกัดโฮโมจีเนียส

ดังนั้น การแปลงสำหรับตัวแปร T สามารถลดรูปให้อยู่ในรูปของเมทริกซ์เดี่ยวขนาด 3x3 ได้ ดังสมการที่ 3.21

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.21)$$

โดยที่ a_0, a_1, a_2 และ b_0, b_1, b_2 คือพารามิเตอร์ของการแปลง

3.2 การแปลงทางเรขาคณิตแบบไม่เป็นเชิงเส้น (Geometrical Non-Linear Transformation)

การแปลงทางเรขาคณิตแบบไม่เป็นเชิงเส้น ทำให้ภาพมีรูปร่างผิดเพี้ยนหรือบิดเบือนไปจากเดิม ซึ่งมีลักษณะคล้ายกับการยืดหรือหดแผ่นยาง ตัวอย่างเช่น การแปลงโพลีโนเมียล (Polynomial Warping) โดยการแปลงโพลีโนเมียลอันดับที่สอง (Second Order Polynomial Mapping) ของ พิกัดภาพถูกนิยามดังสมการที่ 3.22 และ 3.23

$$u = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 \quad (3.22)$$

$$v = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 \quad (3.23)$$

และสามารถเขียนอยู่ในรูปของเวกเตอร์ได้สมการที่ 3.24

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ b_0 & b_1 & b_2 & b_3 & b_4 & b_5 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix} \quad (3.24)$$

การแปลงโพลีโนเมียลอันดับที่สองหรือสูงกว่า มีจุดประสงค์เพื่อการแก้ไขการบิดเบือนไป (Spatial Distortion) ของระบบรับภาพ สมมติให้การแปลงพิกัดจากภาพในอุดมคติไปเป็นภาพที่บันทึกได้ เป็นไปตามสมการที่ 3.25 และ 3.26

$$u = O_u\{x, y\} \quad (3.25)$$

$$v = O_v\{x, y\} \quad (3.26)$$

โดย $O_u\{x, y\}$ และ $O_v\{x, y\}$ คือ ฟังก์ชันการแปลงพิกัดแบบไม่เป็นเชิงเส้น ฟังก์ชันนี้สามารถนำไปทำการแปลงย้อนกลับเพื่อแก้ไขการบิดเบือนไปของภาพได้ ซึ่งทำได้โดยการประมาณค่าฟังก์ชัน

กำหนดให้ ฟังก์ชันโพลีโนเมียลอันดับที่สอง ในสมการ 3.22 และ 3.23 เป็นฟังก์ชันสำหรับการแปลงพิกัด จากนั้นต้องทำการคำนวณค่าสัมประสิทธิ์ (a_i, b_i) ที่ทำให้ค่าผลลัพธ์ของความผิดพลาดเฉลี่ยยกกำลังสอง (Mean-Square Error) ระหว่างพิกัดในภาพที่บันทึกได้ (u_m, v_m), $1 \leq m \leq M$ และพิกัดที่ได้จากการประมาณค่าโพลีโนเมียล (u, v) มีค่าน้อยที่สุด สำหรับเซตของจุด (x_m, y_m), $1 \leq m \leq M$ พิกัดของภาพที่บันทึกได้สามารถเขียนให้อยู่ในรูปแบบของเวกเตอร์ได้ดังสมการที่ 3.27 และ 3.28

$$\mathbf{u}^T = [u_1, u_2, \dots, u_M] \quad (3.27)$$

$$\mathbf{v}^T = [v_1, v_2, \dots, v_M] \quad (3.28)$$

และสัมประสิทธิ์ของฟังก์ชันโพลีโนเมียลอันดับที่สอง เขียนแทนด้วยสมการที่ 3.29 และ 3.30

$$\mathbf{a}^T = [a_0, a_1, \dots, a_5] \quad (3.29)$$

$$\mathbf{b}^T = [b_0, b_1, \dots, b_5] \quad (3.30)$$

โดยการคำนวณค่าผิดพลาดเฉลี่ยยกกำลังสองเป็นไปตามสมการที่ 3.31

$$\varepsilon^2 = (\mathbf{u} - \mathbf{Aa})^T(\mathbf{u} - \mathbf{Aa}) + (\mathbf{v} - \mathbf{Ab})^T(\mathbf{v} - \mathbf{Ab}) \quad (3.31)$$

โดยที่ค่าของตัวแปร A เป็นไปตามสมการที่ 3.32

$$\mathbf{A} = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_M & y_M & x_M^2 & x_M y_M & y_M^2 \end{bmatrix} \quad (3.32)$$

การหาค่าสัมประสิทธิ์ที่ทำให้ค่าผิดพลาดเฉลี่ยยกกำลังสองมีค่าน้อยที่สุด ทำได้โดยการหาอนุพันธ์ของสมการที่ 3.31 เทียบกับสัมประสิทธิ์ a และ b แล้วทำให้ได้ผลลัพธ์เท่ากับศูนย์ดังสมการที่ 3.33 และ 3.34 ตามลำดับ

$$\frac{d\varepsilon^2}{da} = -2\mathbf{A}^T(\mathbf{u} - \mathbf{Aa}) = 0 \quad (3.33)$$

$$\frac{d\varepsilon^2}{db} = -2\mathbf{A}^T(\mathbf{v} - \mathbf{Ab}) = 0 \quad (3.34)$$

ดังนั้น จะได้สัมประสิทธิ์ a และ b เป็นไปตามสมการที่ 3.35 และ 3.36 ตามลำดับ

$$\mathbf{a} = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{u}) \quad (3.35)$$

$$\mathbf{b} = (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{v}) \quad (3.36)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ประเภทของการแปลงทางเรขาคณิต

การแยกเป็นประเภทการแปลงทางเรขาคณิตของพิกัดภาพ สามารถอาศัยความสอดคล้องกับความไม่แปรผันต่อการแปลงทางเรขาคณิตได้ 4 ประเภทดังต่อไปนี้

1. การแปลงแบบคงที่ (Rigid Transformation) เป็นการแปลงที่รูปร่างของภาพไม่เปลี่ยนแปลง ประกอบด้วย การเคลื่อนตำแหน่ง และการหมุน
2. การแปลงเหมือน (Similarity Transformation) ประกอบด้วย การเคลื่อนตำแหน่ง การหมุน และการสเกล
3. การแปลงอัฟไฟน์ (Affine Transformation) ประกอบด้วย การเคลื่อนตำแหน่ง การหมุน การสเกล และการเอียง
4. การแปลงเพอสเปกทิฟ (Perspective Transformation) สามารถใช้ในการจำลองขบวนการเกิดภาพ เมื่อมีการมองวัตถุสามมิติผ่านเลนส์ของตาหรือของกล้องแล้วเห็นเป็นภาพสองมิติ

3.4 ความไม่แปรผันต่อการแปลงทางเรขาคณิต

ความไม่แปรผันต่อการแปลงทางเรขาคณิต คือ คุณสมบัติของตัวอธิบายรูปร่าง (Shape Descriptor) ในเชิงเรขาคณิต เช่น คุณสมบัติของเส้นโค้งหรือพื้นผิวบนภาพที่ไม่มีการเปลี่ยนแปลงเมื่ออยู่ภายใต้การแปลงทางเรขาคณิตของภาพ โดยความไม่แปรผันต่อการแปลงเรขาคณิตสามารถแบ่งได้เป็น 2 แบบ คือ

1. ความไม่แปรผันต่อการแปลงทางเรขาคณิตสัมบูรณ์ (Absolute Invariance) คือคุณสมบัติที่ไม่มีการเปลี่ยนแปลงทั้งก่อนและหลังการแปลงทางเรขาคณิตของภาพ
2. ความไม่แปรผันต่อการแปลงทางเรขาคณิตสัมพัทธ์ (Relative Invariance) คือคุณสมบัติที่ก่อนและหลังการแปลงมีความสัมพันธ์กันด้วยฟังก์ชันใดฟังก์ชันหนึ่ง

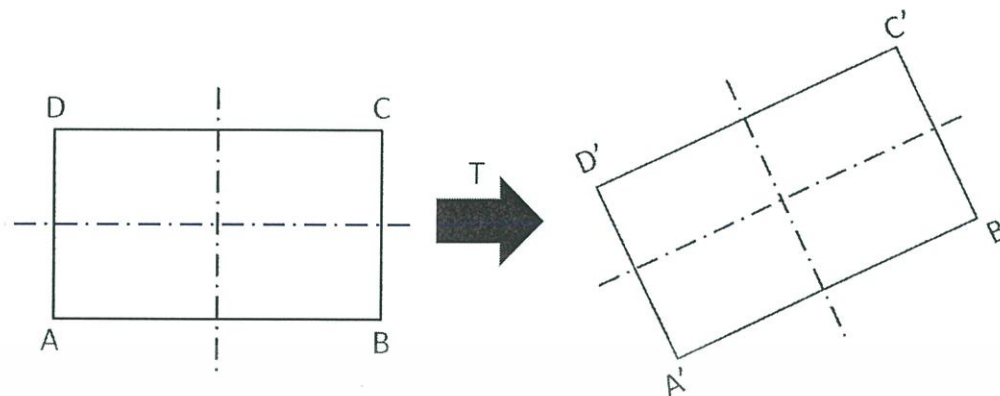
คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงทางเรขาคณิตประเภทต่างๆ แสดงดังตารางที่ 3.1

ตารางที่ 3.1 คุณลักษณะที่มีความไม่แปรผันต่อการแปลงทางเรขาคณิต

Geometric Transformation	Geometric Invariance	
	Absolute Invariance	Relative Invariance
Rigid transformation	- length, angle, area - principal axis	-
Similarity transformation	- angle - ratio of length - ratio of area	length and area (related by determinant of transformation)
Affine transformation	- parallelism - ratio of length - ratio of area	- length, area - intersection point - centroid - inflection point (zero-curvature) - convex hull - median axis
Perspective transformation	- cross ratio (ratio of ratio of length) - five coplanar points	

3.4.1 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงแบบคงที่

คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสมบูรณ์ภายใต้การแปลงแบบคงที่ หรือ Rigid Transformation ได้แก่ ด้านหรือความยาวมุม พื้นที่ และแกนหลัก กำหนดให้สี่เหลี่ยม $A'B'C'D'$ ได้จากการแปลงแบบคงที่ของสี่เหลี่ยม $ABCD$ โดย T คือเมทริกซ์ของการแปลงแบบคงที่ แสดงดังรูปที่ 3.5 และเป็นไปตามสมการที่ 3.37 ถึง 3.39



รูปที่ 3.5 การแปลงแบบคงที่ของสี่เหลี่ยม ABCD เป็นสี่เหลี่ยม A'B'C'D'

$$\angle ABC = \angle A'B'C', \angle BCD = \angle B'C'D', \angle CDA = \angle C'D'A', \angle DAB = \angle D'A'B' \quad (3.37)$$

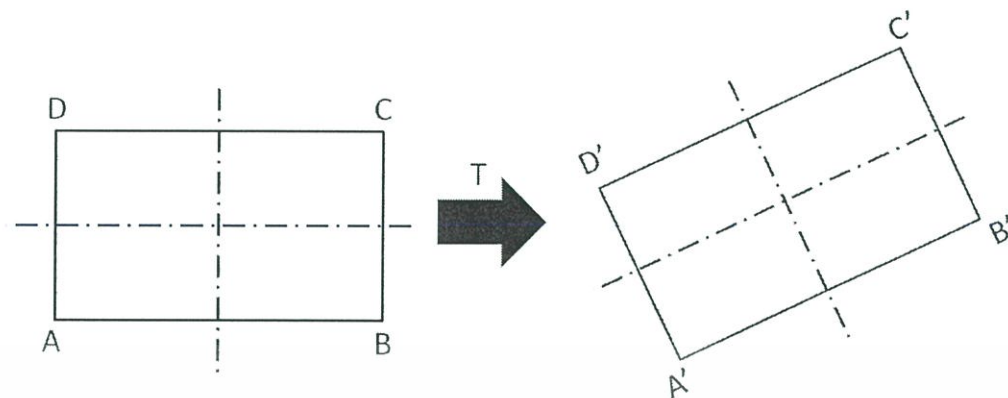
$$\|AB\| = \|A'B'\|, \|BC\| = \|B'C'\|, \|CD\| = \|C'D'\|, \|DA\| = \|D'A'\| \quad (3.38)$$

$$\text{Area}(ABCD) = \text{Area}(A'B'C'D') \quad (3.39)$$

โดย $\text{Area}(ABCD)$ และ $\text{Area}(A'B'C'D')$ คือ พื้นที่ของสี่เหลี่ยม ABCD และ A'B'C'D' ตามลำดับ

3.4.2 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงเหมือน

คุณสมบัติที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสมบูรณ์ภายใต้การแปลงเหมือน ได้แก่ มุม อัตราส่วนของความยาว (Ratio of Length) และอัตราส่วนของพื้นที่ (Ratio of Area) คุณสมบัติที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสัมพันธ์ภายใต้การแปลงเหมือน ได้แก่ ความยาวและพื้นที่ โดยคุณสมบัติก่อนและหลังการแปลงจะสัมพันธ์กันด้วยค่า determinant ของเมทริกซ์การแปลง T



รูปที่ 3.6 การแปลงเหมือนของสี่เหลี่ยม ABCD เป็นสี่เหลี่ยม A'B'C'D'

กำหนดให้สี่เหลี่ยม A'B'C'D' ได้จากการแปลงเหมือนของสี่เหลี่ยม ABCD โดย T คือเมทริกซ์ของการแปลงเหมือน แสดงดังรูปที่ 3.6 จะได้ตามสมการที่ 3.40 ถึง 3.40

$$\text{Area}(A'B'C'D') = |T| \cdot \text{Area}(ABCD) \quad (3.40)$$

โดย Area(ABCD) และ Area(A'B'C'D') คือพื้นที่ของรูปสี่เหลี่ยม ABCD และ A'B'C'D' ตามลำดับ และ

$$\text{Area}(\Delta A'B'C') = |T| \cdot \text{Area}(\Delta ABC) \quad (3.41)$$

โดย Area(ΔABC) และ Area($\Delta A'B'C'$) คือพื้นที่ของรูปสามเหลี่ยม ABC และ A'B'C' ตามลำดับ

$$\text{Area}(\Delta B'C'D') = |T| \cdot \text{Area}(\Delta BCD) \quad (3.42)$$

โดย Area(ΔBCD) และ Area($\Delta B'C'D'$) คือพื้นที่ของรูปสามเหลี่ยม BCD และ B'C'D' ตามลำดับ

สมการ 3.41 ทารด้วยสมการ 3.42 จะได้เป็นสมการที่ 3.43

$$\frac{\text{Area}(\Delta A'B'C')}{\text{Area}(\Delta B'C'D')} = \frac{\text{Area}(\Delta ABC)}{\text{Area}(\Delta BCD)} \quad (3.43)$$

3.4.3 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงอัฟไฟน์

คุณสมบัติที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสมบูรณ์ ภายใต้การแปลงอัฟไฟน์ ได้แก่ การขนานกัน (Parallelism) อัตราส่วนของความยาวและอัตราส่วนของพื้นที่ คุณสมบัติที่ไม่แปรผันต่อการแปลงเรขาคณิตสัมพันธ์ภายใต้การแปลงอัฟไฟน์ ได้แก่ ความยาว, พื้นที่ จุดตัด (Intersection Point), จุดเซนทรอยด์ (Centroid), จุดหักเห (Inflection Point), Convex Hull และแกนกลาง (Median Axis) การแปลงอัฟไฟน์ของระนาบ ($T: \mathcal{R}^2 \rightarrow \mathcal{R}^2$) สามารถแสดงอยู่ในรูปเมทริกซ์ nonsingular ขนาด 2×2 (A) และเวกเตอร์ของการย้ายตำแหน่ง b ได้ ดังสมการที่ 3.44

$$x' = Ax + b, \text{ สำหรับ } x \in \mathcal{R}^2 \quad (3.44)$$

คุณสมบัติต่างๆ ของการแปลงอัฟไฟน์ มีดังต่อไปนี้

คุณสมบัติที่ 1

การแปลงอัฟไฟน์ของระนาบถูกนิยามอย่างจำเพาะด้วยจุดที่เป็นคู่กันสามคู่ นั่นคือ ถ้า a , b และ c เป็นจุดที่ไม่ได้อยู่บนแนวเส้นตรงเดียวกัน (Non-Collinear Points) และ a' , b' และ c' เป็นจุดที่สอดคล้องกัน จะได้ว่ามีการแปลงอัฟไฟน์เพียงการแปลงเดียว ($T: \mathcal{R}^2 \rightarrow \mathcal{R}^2$) ที่แปลงจุดทั้งสามไปยังจุดที่สอดคล้องกัน

คุณสมบัติที่ 2

อัตราส่วนของระยะทาง (Ratio of Distances) บนเส้นไม่แปรผันต่อการแปลงอัฟไฟน์ ถ้า a , b และ c เป็นจุดที่อยู่บนแนวเส้นตรงเดียวกัน (Collinear Points) และ T คือการแปลงอัฟไฟน์ เขียนแทนด้วยสมการที่ 3.45

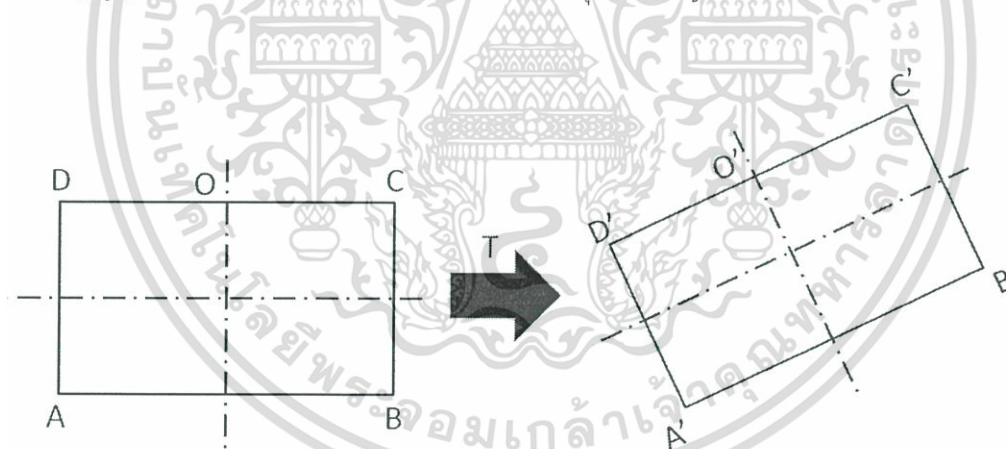
$$\frac{|a-b|}{|b-c|} = \frac{|T(a)-T(b)|}{|T(b)-T(c)|} \quad (3.45)$$

โดยที่ $|x|$ คือ ความยาวของส่วนของเส้นตรง x

ตัวอย่างเช่นกำหนดรูปสี่เหลี่ยม $ABCD$ โดยส่วนของเส้นตรง DC ประกอบด้วย \overline{DO} กับ \overline{OC} และรูปสี่เหลี่ยม $A'B'C'D'$ โดยส่วนของเส้นตรง $D'O'C'$ ประกอบด้วย $\overline{D'O'}$ กับ $\overline{O'C'}$ แสดงดังรูปที่ 3.7 และเขียนได้ตามสมการที่ 3.46

$$\frac{\|DO\|}{\|OC\|} = \frac{\|D'O'\|}{\|O'C'\|} \quad (3.46)$$

โดยที่ $\|xy\|$ คือ ความยาวของส่วนของเส้นตรงระหว่างจุด x และ y



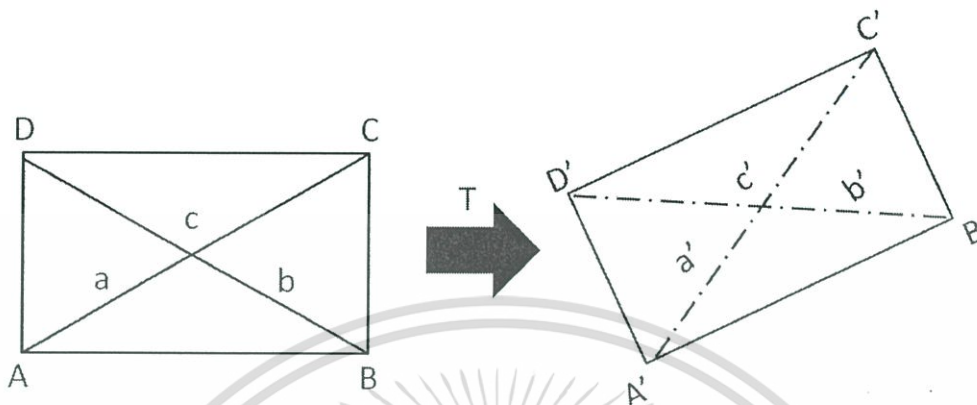
รูปที่ 3.7 การแปลงอัฟไฟน์ของสี่เหลี่ยม $ABCD$ เป็นสี่เหลี่ยม $A'B'C'D'$

คุณสมบัติที่ 3

จุดตัด (Intersection Points) ไม่แปรผันต่อการแปลงอัฟไฟน์ ถ้า c คือจุดตัดของเส้น a และเส้น b และ T คือการแปลงอัฟไฟน์ จะได้ว่า $T(c)$ เป็นจุดตัดของเส้น $T(a)$ และ $T(b)$ ตัวอย่างของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จุดตัดที่ไม่แปรผันต่อการแปลงอัฟไฟน์แสดงดังรูปที่ 3.8 โดย c' เป็นจุดตัดของเส้นทแยงมุม a' กับ b' ของรูปสี่เหลี่ยม $A'B'C'D'$ ที่ได้จากการแปลงอัฟไฟน์ของสี่เหลี่ยม $ABCD$



รูปที่ 3.8 ตัวอย่างของจุดตัดที่ไม่แปรผันต่อการแปลงอัฟไฟน์

คุณสมบัติที่ 4

จุด centroid ของเซตของจุดไม่แปรผันต่อการแปลงอัฟไฟน์ ถ้า c คือจุด centroid ของเซตของจุด (X) และ T คือการแปลงอัฟไฟน์ จะได้ว่า $T(c)$ เป็นจุด centroid ของเซตของจุด $X' = \{x' | x' = A(x), x \in X\}$

คุณสมบัติที่ 5

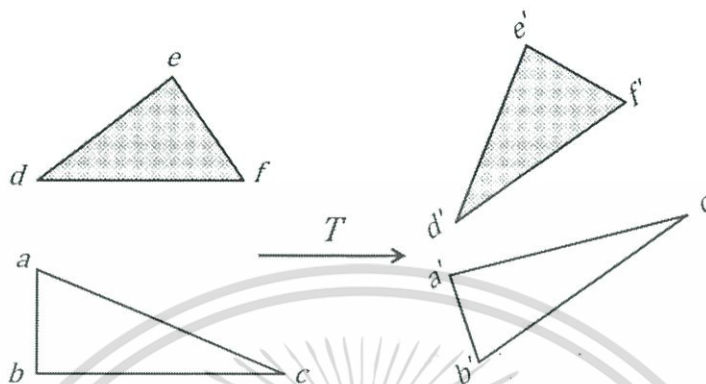
อัตราส่วนของพื้นที่ของสามเหลี่ยมไม่แปรผันต่อการแปลงอัฟไฟน์ ถ้า a, b และ c เป็นจุดที่ไม่ได้อยู่บนแนวเส้นตรงเดียวกัน (Non-Collinear Points) และ d, e และ f เป็นจุดที่ไม่ได้อยู่บนแนวเส้นตรงเดียวกัน (โดยอาจจะไม่แตกต่างจาก a, b และ c) และ T คือการแปลงอัฟไฟน์ ดังสมการที่

$$\frac{|\Delta abc|}{|\Delta def|} = \frac{|\Delta A(a)A(b)A(c)|}{|\Delta A(d)A(e)A(f)|} \tag{3.47}$$

โดยที่ $|\Delta x|$ คือ พื้นที่ของรูปสามเหลี่ยม

ตัวอย่างเช่นกำหนดรูปสามเหลี่ยม $a'b'c'$ และ $d'e'f'$ ได้จากการแปลงอัฟไฟน์ของสามเหลี่ยม abc และ def ตามลำดับ แสดงดังรูปที่ 3.9 และสามารถเขียนได้ดังสมการที่ 3.48

$$\frac{\text{Area}(\triangle abc)}{\text{Area}(\triangle def)} = \frac{\text{Area}(\triangle a'b'c')}{\text{Area}(\triangle d'e'f')} \quad (3.48)$$



รูปที่ 3.9 รูปสามเหลี่ยม $a'b'c'$ และ $d'e'f'$ ที่ได้จากการแปลงอัฟไฟน์ของสามเหลี่ยม abc และ def ตามลำดับ [26]

คุณสมบัติที่ 6

จุดหักเห (Inflection Point) หรือจุดที่ความโค้งเป็นศูนย์ (Zero-Curvature) ไม่แปรผันต่อการแปลงอัฟไฟน์ สำหรับเส้นโค้ง (Curve) ที่ถูกกำหนดโดยพารามิเตอร์ t โดย t คือความยาวคอร์ด (Chord Length) ความโค้ง (Curvature) ถูกนิยามตามสมการที่ 3.49

$$k(t) = \frac{|r^{(1)t} \times r^{(2)t}|}{|r^{(1)t}|^3} \quad (3.49)$$

โดยที่ $r^{(1)t}$ คือเส้นสัมผัสส่วนโค้ง (Tangent Vector) และ $r^{(2)t}$ คือ เวกเตอร์ความโค้ง (Curvature Vector)

ที่จุดหักเหใดๆ ความโค้งมีค่าเป็นศูนย์และ $r^{(1)t}$ และ $r^{(2)t}$ ขนานกัน เนื่องจากการแปลงอัฟไฟน์มีคุณสมบัติของการขนานกัน จะได้ว่า $r_a^{(1)t}$ และ $r_a^{(2)t}$ ขนานกันด้วย โดย $r_a^{(1)t}$ และ $r_a^{(2)t}$ คือเส้นสัมผัสส่วนโค้งและเวกเตอร์ความโค้งของเส้นโค้งหลังจากการแปลงอัฟไฟน์ตามลำดับ ดังนั้นจุด

หักของเส้นโค้งหลังการแปลงอัฟไฟน์ ก็คือจุดหักเห (ของเส้นโค้งเดิม) ที่ถูกแปลงด้วยการแปลงอัฟไฟน์ นั่นคือจุดหักเหมีคุณสมบัติที่ไม่แปรผันต่อการแปลงอัฟไฟน์

คุณสมบัติที่ 7

convex hull มีคุณสมบัติที่ไม่แปรผันต่อการแปลงอัฟไฟน์นั่นคือ convex hull ของเซตของจุดข้อมูลหลังการแปลงอัฟไฟน์คือ convex hull (ของจุดข้อมูลก่อนการแปลง) ที่ถูกแปลงด้วยการแปลงอัฟไฟน์ โดย convex hull คือรูปทรงเรขาคณิตที่เกิดจากการเชื่อมต่อส่วนนูน (Convex) ของวัตถุ ซึ่งมีคุณสมบัติ ดังนี้

1. มีลักษณะเฉพาะเป็นหนึ่งเดียว (Uniqueness)

2. มีประสิทธิภาพในการคำนวณ (Computational Efficiency) โดยในการหา convex hull ของจุดข้อมูลจำนวน N จุด ความซับซ้อนในการคำนวณมากที่สุดที่เป็นไปได้คือ $O(N \log_2 N)$ และถ้าจุดข้อมูลมีการกระจายที่สม่ำเสมอ ความซับซ้อนในการคำนวณที่เป็นไปได้คือ $O(N)$ โดยที่ $O(f(n))$ คือจำนวนของการดำเนินการ (Number of Operations) ซึ่งคำนวณได้จากฟังก์ชัน $f(n)$ โดยที่ N คือจำนวนจุดข้อมูลที่จะนำมาหา convex hull

3. convex hull มีคุณสมบัติที่ไม่แปรผันต่อการแปลงอัฟไฟน์

4. มีความสามารถในการควบคุมเฉพาะที่ (Local Controllability) นั่นคือเมื่อมีการเพิ่มหรือลดจุดข้อมูลจากจุดข้อมูลเดิมที่มีอยู่ จะมีผลต่อการหา convex hull เพียงบางตำแหน่งของข้อมูลเท่านั้น

3.4.4 คุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงเพอสเปคทีฟ

คุณสมบัติที่ 1

cross ratio ของจุด 4 จุดบนเส้นตรงไม่แปรผันต่อการแปลงเพอสเปคทีฟ (Perspective Transformation) กำหนดให้ A, B, C และ D เป็นจุด 4 จุดบนเส้นตรง μ และกำหนดให้จุดเหล่านี้โปรเจกต์ไปยังจุด a, b, c และ d ตามลำดับบนเส้นตรง μ_1 แสดงดังรูปที่ 3.10 และสมการที่ 3.50

$$\frac{|a-c||b-d|}{|a-b||c-d|} = \frac{|A-C||B-D|}{|A-B||C-D|} \quad (3.50)$$

โดยที่ $|x_1 - x_2|$ คือระยะทางระหว่างจุด x_1 และ x_2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

cross ratio สามารถสร้างจากจุดใดบนเส้นตรงก็ได้ ที่ทำให้จำนวนจุดทั้งในตัวเศษและตัวส่วนของ cross ratio มีจำนวนจุดเท่ากัน เพื่อที่จะกำจัด determinant ของเมทริกซ์การแปลงในระบบพิกัดโฮโมจีเนียส (Homogeneous Transformation Matrix) และ projective scale factors

คุณสมบัติที่ 2

จุดที่อยู่บนระนาบเดียวกัน 5 จุด (Five Coplanar Points) ไม่แปรผันต่อการแปลงเพอสเปคทีฟ พิจารณาจุด 5 จุด (P_1, \dots, P_5) ที่อยู่บนระนาบเดียวกันและไม่ได้อยู่บนเส้นตรงเดียวกัน แสดงดังรูปที่ 3.11 จะได้ว่าอัตราส่วน 2 อัตราส่วนที่เป็นอิสระต่อกันของ determinant ของเมทริกซ์ $m_{i,j,k}$ ของจุด 5 จุดมีความไม่แปรผันต่อการแปลงเพอสเปคทีฟ ดังนั้นสามารถสร้างคุณลักษณะที่ไม่แปรผัน (I_1) ของจุด 5 จุดได้ ตัวอย่างเช่น



รูปที่ 3.10 การแปลงเพอสเปคทีฟที่โปรเจคไปยังเส้นตรง [26]

$$I_1 = \frac{|m_{431}| |m_{521}|}{|m_{421}| |m_{531}|} \quad (4.41)$$

โดยที่ $m_{i,j,k} = (P_i, P_j, P_k)$ โดย $P_i = (x_i, y_i, 1)^t$ และ

$$|m_{i,j,k}| \text{ คือ determinant ของ } m_{i,j,k} \text{ เช่น } |m_{521}| \text{ คือ } \begin{vmatrix} x_5 & x_2 & x_1 \\ y_5 & y_2 & y_1 \\ 1 & 1 & 1 \end{vmatrix}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กำหนดให้ $P_1 = (x_1, y_1), \dots, P_5 = (x_5, y_5)$ เป็นจุดที่อยู่บนระนาบเดียวกันและไม่ได้อยู่บนเส้นตรงเดียวกันและ $P'_1 = (x'_1, y'_1), \dots, P'_5 = (x'_5, y'_5)$ ได้จากการแปลงเพอสเปคทีฟของ P_1, \dots, P_5 สามารถสร้างความสัมพันธ์ที่ไม่แปรผันต่อการแปลงเพอสเปคทีฟได้ ดังสมการที่ 4.42

$$\frac{A_{12}A_{34}}{A_{14}A_{23}} = \frac{A'_{12}A'_{34}}{A'_{14}A'_{23}} \quad (4.42)$$

โดย A_{ij} คือพื้นที่ของรูปสามเหลี่ยมที่เกิดจากการเชื่อมต่อเวกเตอร์ $\overrightarrow{P_5P_i}$ กับ $\overrightarrow{P_5P_j}$, $i = 1, \dots, 4$, $j = 1, \dots, 4$ เป็นไปตามสมการที่ 4.43

$$\frac{\sin a_{12} \sin a_{34}}{\sin a_{14} \sin a_{23}} = \frac{\sin a'_{12} \sin a'_{34}}{\sin a'_{14} \sin a'_{23}} \quad (4.43)$$

โดย a_{ij} คือมุมระหว่างเส้น P_5P_i กับ P_5P_j โดย $i = 1, \dots, 4$ และ $j = 1, \dots, 4$ คือจุดที่เหลื่อมอยู่



รูปที่ 3.11 จุดที่อยู่บนระนาบเดียวกัน 5 จุด [26]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ระเบียบวิธีการวิจัย (Research Methodology)

ในบทนี้กล่าวถึงขั้นตอนในการศึกษาวิจัย เพื่อออกแบบอัลกอริทึม (Algorithm) สำหรับแปลภาษามือแบบสะกดคำด้วยนิ้วมือ โดยใช้กระบวนการประมวลผลภาพ ซึ่งภาพรวมของอัลกอริทึมสามารถแบ่งออกเป็น 3 ขั้นตอนหลักๆ ดังภาพที่ 4.1 กระบวนการเริ่มต้นสำหรับอัลกอริทึมคือการปรับปรุงภาพเบื้องต้น (Preprocessing Process) ได้แก่ การปรับภาพด้วยกระบวนการเทรชโฮลดิ้ง (Thresholding) การกำจัดสัญญาณรบกวนภาพ ส่วนถัดมาคือการสกัดคุณลักษณะเฉพาะ (Feature Extraction) จากแต่ละท่าทางของภาษามือด้วยเทคนิคการคำนวณพื้นที่สามเหลี่ยม และขั้นตอนสุดท้ายคือการคัดแยกประเภทข้อมูลที่สกัดได้ด้วยวิธีการของโครงข่ายประสาทเทียม

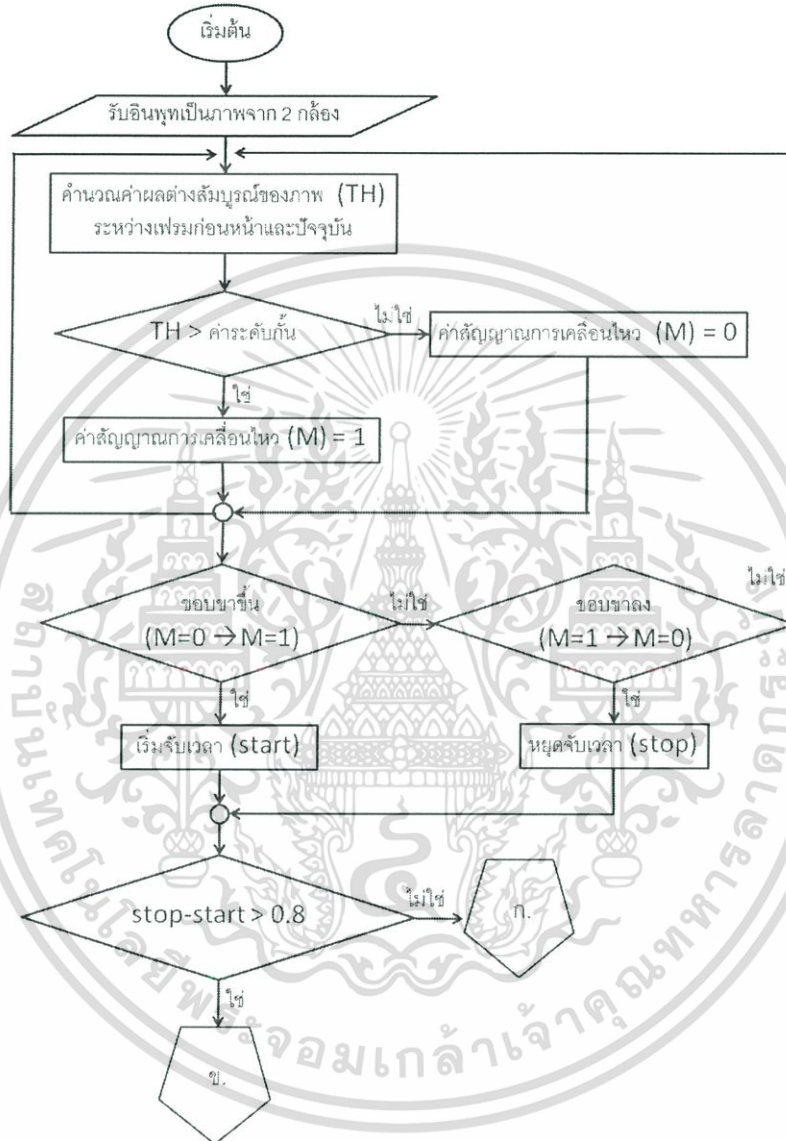


รูปที่ 4.1 กระบวนการโดยสังเขปสำหรับอัลกอริทึม

นอกจากนั้นผู้วิจัยได้ออกแบบถุงมือที่มีลักษณะดังรูปที่ 4.3 โดยใช้ถุงมือสีดำและมีมาร์คเกอร์ (Marker) ที่มีสีแตกต่างกันออกไป ติดอยู่บริเวณส่วนปลายของถุงมือทั้ง 5 นิ้ว และบริเวณตรงกลางข้อนิ้วด้านใน 1 จุด ทั้งนี้เพื่อให้ลดความยุ่งยากสำหรับกระบวนการประมวลผลภาพเบื้องต้น สามารถคัดแยกมาร์คเกอร์ออกจากบริเวณฝ่ามือได้อย่างง่ายดาย นอกจากนี้ยังมีการควบคุมปริมาณแสงของระบบและใช้พื้นหลังเป็นสีดำ

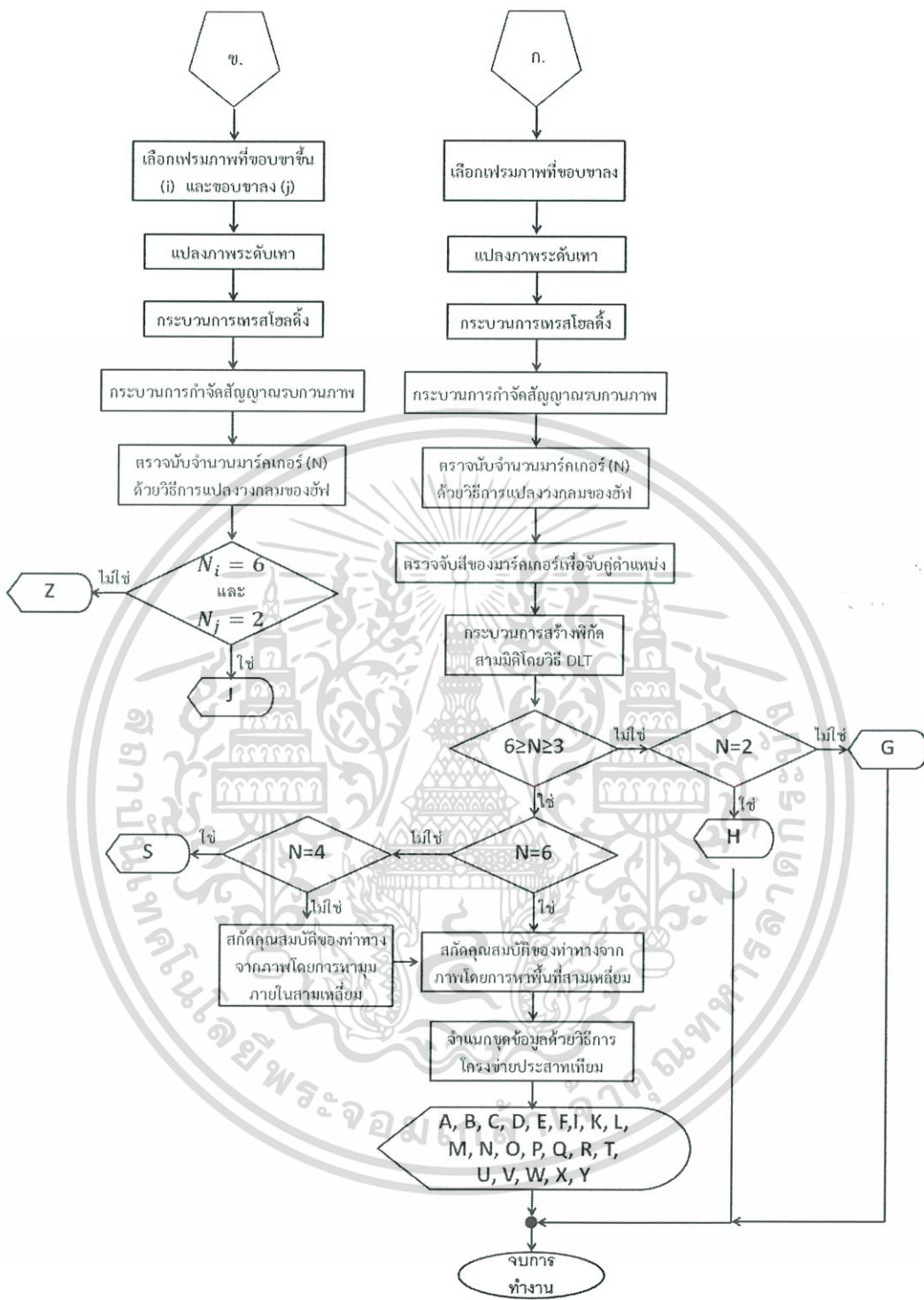
4.1 การออกแบบการทำงานของระบบ

4.1.1 ขั้นตอนการจำแนกท่าทางของภาษามือ แสดงในรูปที่ 4.2



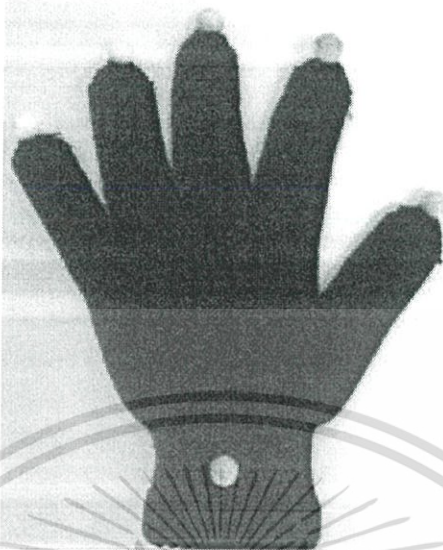
รูปที่ 4.2 แผนผังการจำแนกท่าทางภาษามือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 ตำแหน่งของมาร์คเกอร์บนถุงมือ

4.2 ขั้นตอนการออกแบบอัลกอริทึมสำหรับแปลภาษามือ

กระบวนการการออกแบบอัลกอริทึมสำหรับการแปลภาษามือ แบ่งออกเป็น 3 ขั้นตอนหลัก โดยเริ่มตั้งแต่กระบวนการรับภาพ (Image Acquisition) เข้าสู่ระบบเพื่อนำไปประมวลผลในขั้นถัดไป ซึ่งขั้นตอนที่มีความสำคัญถัดมาคือ กระบวนการสกัดคุณลักษณะเฉพาะ (Feature Extraction) ของแต่ละท่าทาง โดยประกอบด้วยหลายๆ กระบวนการแยกย่อย เช่น การแปลงชนิดภาพ การตรวจจับวงกลม การหาพื้นที่สามเหลี่ยมใดๆ ฯลฯ ซึ่งจะอธิบายในส่วนถัดไป นอกจากนี้กระบวนการของการแบ่งกลุ่มข้อมูล (Classification) ด้วยวิธีการของโครงข่ายประสาทเทียม (Artificial Neural Network) เป็นวิธีที่มีประสิทธิภาพที่ดีในการคัดแยกกลุ่มข้อมูลที่มีจำนวนมากและซับซ้อน

4.2.1 กระบวนการรับภาพ

กระบวนการรับภาพเป็นขั้นตอนแรกสำหรับการนำภาพไปประมวลผล ซึ่งภาพต้นฉบับที่นำเข้าสู่กระบวนการอยู่ในรูปแบบของโมเดลสีอาร์จีบี (RGB Color Model) ดังรูปที่ 4.4 โดยใช้วิธีการรับภาพจากกล้อง 2 ตัว โดยการใช้คอมพิวเตอร์เพียง 1 เครื่อง ซึ่งใช้วิธีการถ่ายภาพสลับกันอย่างรวดเร็วผ่าน Port USB ของคอมพิวเตอร์ โดยระยะเวลาเฉลี่ยของการรับภาพระหว่างสอง Port มีค่าเฉลี่ย 0.08 วินาที หลังจากรับภาพจากกล้อง ภาพดังกล่าวถูกแปลงเป็นภาพระดับเทา (Gray Scale Image) และเข้าสู่กระบวนการประมวลผลในส่วนถัดไป ดังรูปที่ 4.2



รูปที่ 4.4 ภาพอินพุตภาษามือ A-Z จากกล้องเว็บแคม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทำงานของระบบได้รับอินพุตมาจากกล้องเว็บแคมจำนวน 2 ตัว ยี่ห้อ Microsoft รุ่น LifeCam HD-3000 ความละเอียดภาพสูงสุด 720p HD ขนาดของภาพที่ใช้ในการประมวลผลภาพเท่ากับ 320x240 พิกเซล (Pixel) และอัตราเร็วต่อเฟรมเป็น 30 เฟรมต่อวินาที โดยระบบคอมพิวเตอร์ที่ใช้สำหรับการทดลองมีคุณลักษณะเป็นไปตามตารางที่ 4.1

ตารางที่ 4.1 คุณลักษณะของคอมพิวเตอร์ที่ใช้ในการทดลอง

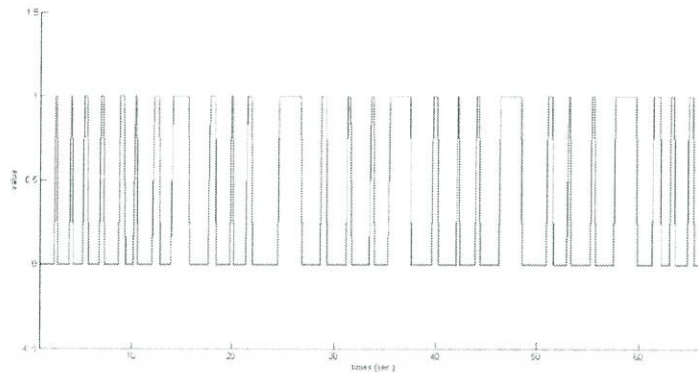
รายการ	คุณลักษณะ
CPU	Core i-7 – 4500U CPU @ 1.80 GHz
Memory	8.00 GB RAM
Operation System	Windows 7 Ultimate รุ่น 64-Bit

4.2.2 กระบวนการจำแนกชนิดของท่าทางจากการเคลื่อนไหว

จากท่าทางของการสะกดคำด้วยภาษามือแบบ American Sign Language สามารถจำแนกออกเป็น 2 ประเภท ตามการเคลื่อนไหว คือ ท่าที่มีการเคลื่อนไหวแบบสถิตย์ (Static Posture) ได้แก่ A, B, C, D, E, F, G, H, I, K, L, M, N, O, P, Q, R, S, T, U, V, W, X และ Y ส่วนท่าที่มีการเคลื่อนไหวแบบพลวัต (Dynamic Posture) ได้แก่ J และ Z

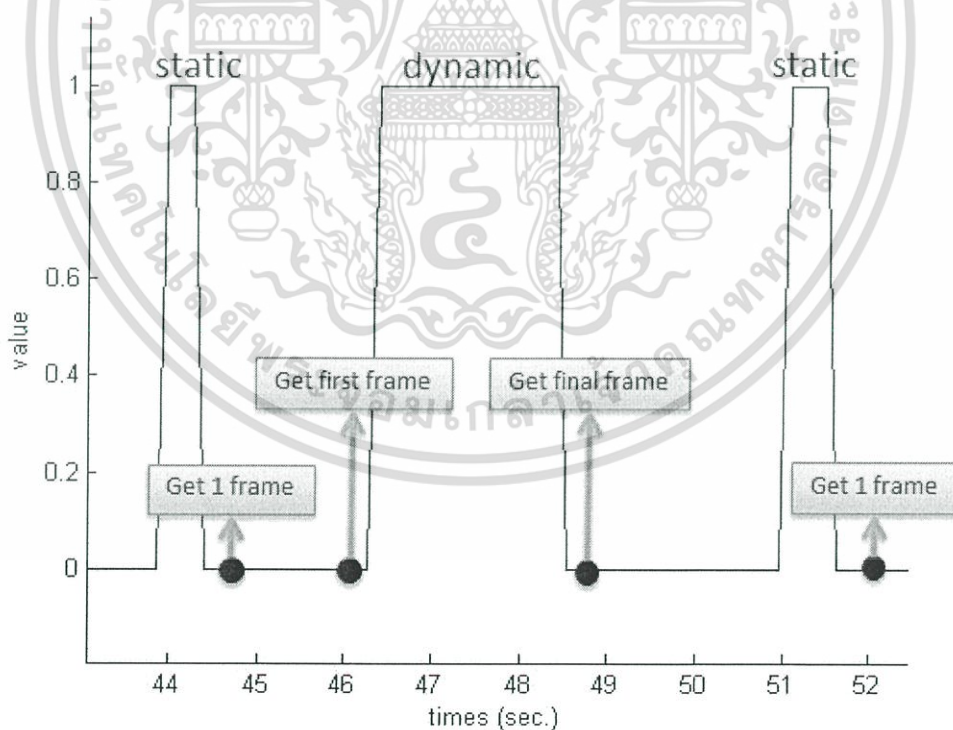
ดังนั้นจำเป็นต้องมีกระบวนการจำแนกชนิดของท่าทางจากการเคลื่อนไหว (Motion Classification) โดยวิธีการของการหาผลต่างสัมบูรณ์ (Absolute Difference) ระหว่างเฟรมภาพปัจจุบันกับเฟรมภาพก่อนหน้า โดยสามารถนำผลต่างสัมบูรณ์ดังกล่าว จะถูกแปลงให้มีค่า 0 และ 1 ด้วยค่าระดับกัน (Threshold) เรียกว่าสัญญาณการเคลื่อนไหว (Motion Signal) ซึ่งความหมายของค่า 0 คือ ไม่มีการเคลื่อนไหว และ 1 คือ มีการเคลื่อนไหว ดังแสดงในรูปที่ 4.5

การจำแนกชนิดของท่าทางภาษามือสำหรับโครงการนี้ใช้เวลาของการเคลื่อนไหว เพื่อจำแนกท่าทางระหว่าง Static กับ Dynamic จากรูปที่ 4.5 สังเกตได้ว่า มีสัญญาณ 2 รูปแบบ กล่าวคือ สัญญาณที่มีค่าเป็น 1 ในช่วงที่กว้างและแคบ ถ้าหากเวลาการเคลื่อนไหวมีค่ามากกว่าระดับกัน กำหนดให้เป็นท่าแบบ Dynamic แต่หากเวลาการเคลื่อนไหวมีค่าน้อยกว่าระดับกัน กำหนดให้เป็นท่าแบบ Static ซึ่งวิธีการดึงเฟรมภาพเพื่อเอาไปพิจารณามีความแตกต่างกัน



รูปที่ 4.5 สัญญาณการเคลื่อนไหวจากภาพ

ตำแหน่งการดึงเฟรมภาพสำหรับท่าทางแบบ Static เพื่อนำไปพิจารณา คือตำแหน่งหลังขอบขาลงสัญญาณและนำภาพไปเข้าสู่กระบวนการสกัดคุณลักษณะเฉพาะตามหัวข้อ 4.2.3 ถัดไป ส่วนตำแหน่งการดึงเฟรมสำหรับท่าทางแบบ Dynamic มี 2 ตำแหน่ง คือ ตำแหน่งก่อนขอบขาขึ้นและหลังขอบขาลง ดังรูปที่ 4.6

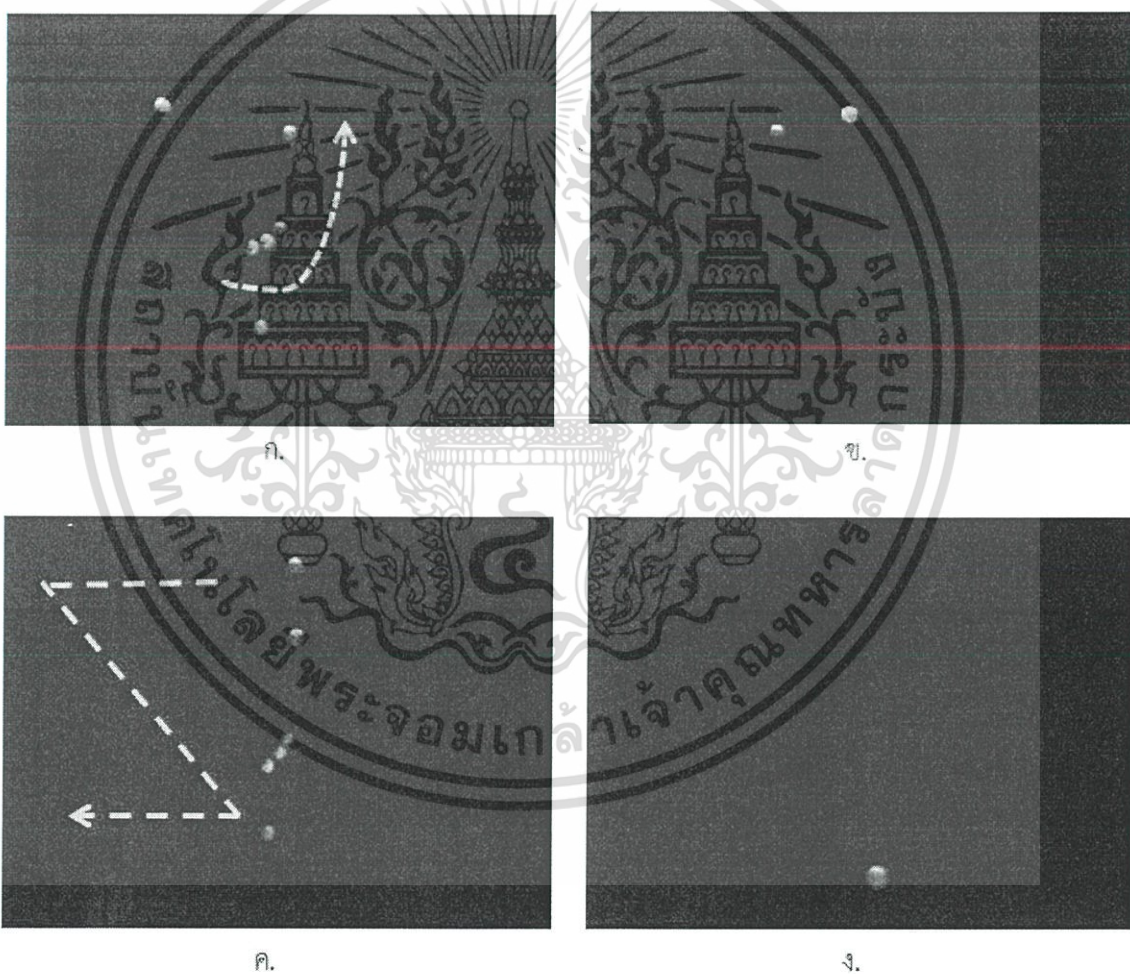


รูปที่ 4.6 ตำแหน่งการดึงเฟรมภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3 กระบวนการจำแนกชนิดของท่าทางแบบ Dynamic

ท่าทางแบบ Dynamic สำหรับภาษามือแบบ American Sign Language ได้แก่ J และ Z เป็นท่าที่มีการเคลื่อนไหวอย่างต่อเนื่องในขณะที่ทำท่านั้นๆ การพิจารณาเฟรมภาพที่สนใจ คือ เฟรมที่อยู่ก่อนขอบขาขึ้น (ตำแหน่งเริ่มต้นของท่าทาง) และหลังขอบขาลง (ตำแหน่งสิ้นสุดของท่าทาง) ดังรูปที่ 4.5 ดังนั้นการจำแนกท่าทางระหว่าง J และ Z สามารถพิจารณาจากจำนวนมาร์คเกอร์ที่เฟรมแรก และเฟรมสุดท้ายของท่าทาง ดังรูปที่ 4.7 สำหรับตัว J มีจำนวนมาร์คเกอร์ 6 และ 2 จุด ของเฟรมแรกและเฟรมสุดท้าย ตามลำดับ และ Z มีจำนวนมาร์คเกอร์ 6 และ 1 จุด ของเฟรมแรกและเฟรมสุดท้าย ตามลำดับ เช่นเดียวกัน ซึ่งกระบวนการตรวจจับวงกลมได้อธิบายรวมอยู่ในหัวข้อที่ 4.2.4.3



รูปที่ 4.7 ก. เฟรมเริ่มต้นของ J ข. เฟรมสิ้นสุดของ J
 ค. เฟรมเริ่มต้นของ Z ง. เฟรมสิ้นสุดของ Z

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

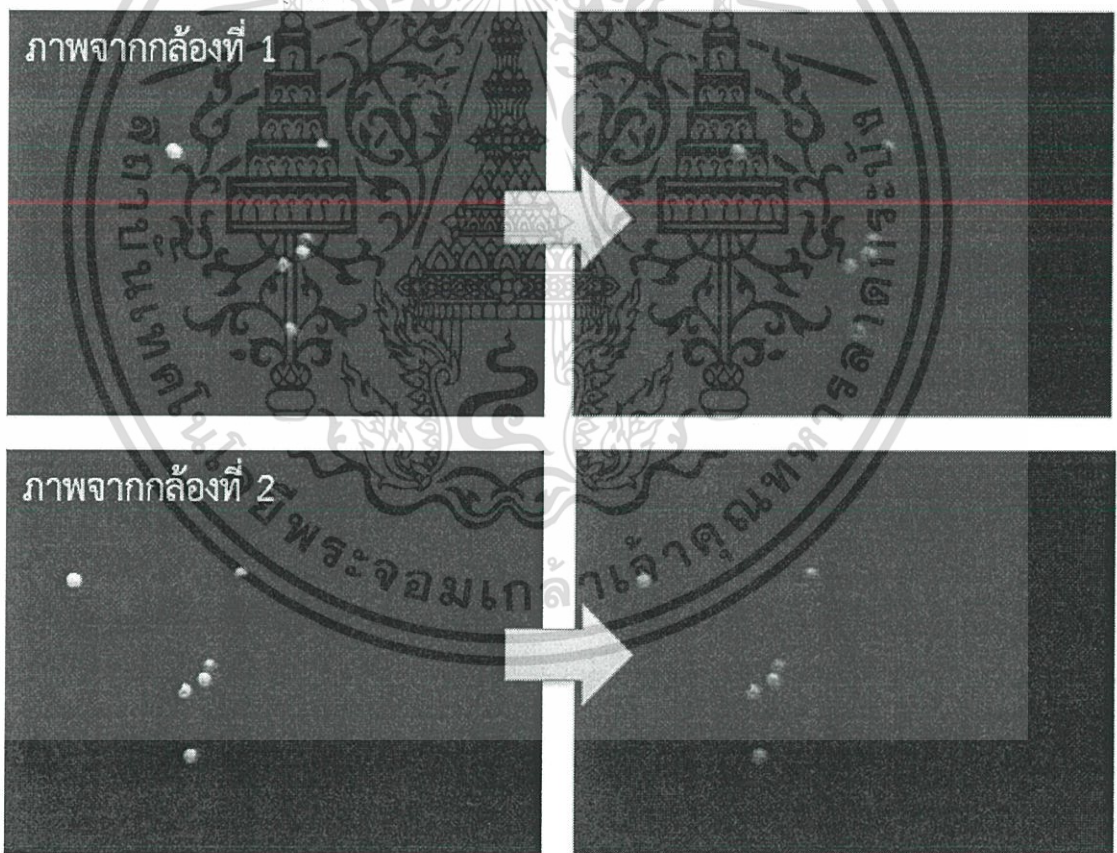
4.2.4 กระบวนการสกัดคุณลักษณะเฉพาะสำหรับทำทางแบบ Static

กระบวนการสกัดคุณลักษณะเฉพาะของแต่ละทำทางของภาษามือ เพื่อแปลงข้อมูลภาพที่รับเข้ามาในแต่ละเฟรมให้อยู่ในรูปแบบข้อมูลตัวเลข ซึ่งวิธีการสกัดคุณลักษณะเฉพาะในโครงการนี้ประกอบไปด้วยขั้นตอนย่อยๆ โดยการอธิบายถึงขั้นตอนเหล่านี้จะยกตัวอย่างภาพเพียง 1 ทำทาง เนื่องจากทุกๆ ทำทาง กระทำด้วยขั้นตอนเหมือนกันทั้งหมด

4.2.4.1 การแปลงภาพระดับเทา (Gray Scale Image Converting)

หลังจากรับภาพจากกล้องทั้งสองตัวแล้ว กระบวนการถัดมา ได้แก่ การแปลงภาพระดับเทาดังรูปที่ 4.8 ซึ่งใช้การเฉลี่ยค่าในโมเดลสี RGB ด้วยน้ำหนักที่แตกต่างกันในแต่ละพิกเซล ดังสมการที่ 4.1 โดย i และ j แทนตำแหน่งพิกเซลในภาพ R G และ B แทนค่าองค์ประกอบสีในภาพต้นฉบับ

$$\text{gray}(i,j) = 0.2989 \times R(i,j) + 0.5870 \times G(i,j) + 0.1140 \times B(i,j) \quad (4.1)$$

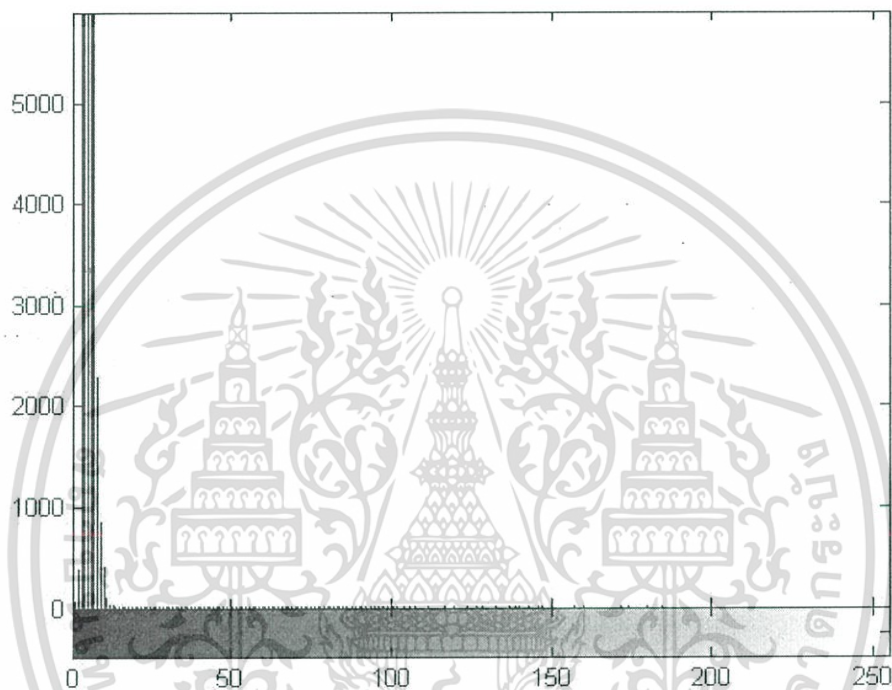


รูปที่ 4.8 การแปลงภาพระดับเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.4.2 การตั้งค่าระดับกันของภาพ (Thresholding Image)

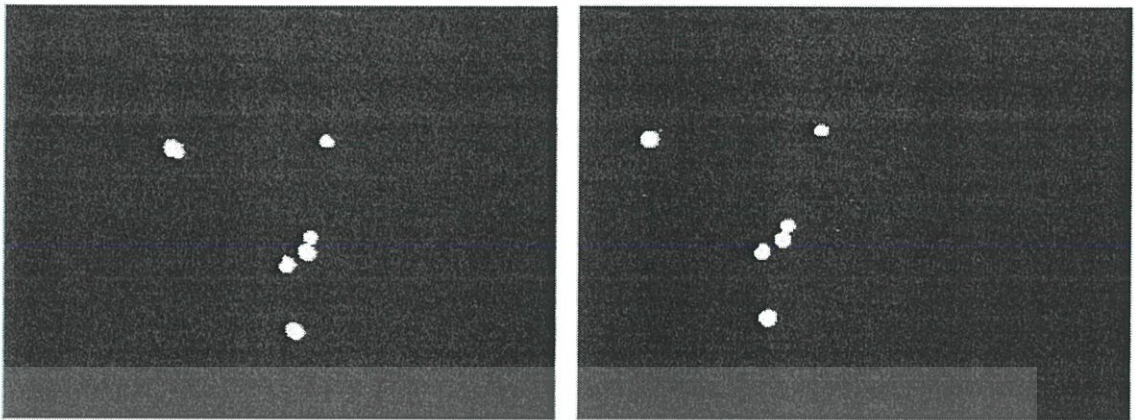
ระบบที่ถูกออกแบบมาเป็นการควบคุมปริมาณแสงที่พอเหมาะ พื้นหลังและถูงมือเป็นสีดำ ทำให้กระบวนการเทรสโฮลดิ้ง มีความคงที่และสามารถแยกบริเวณระหว่างมาร์คเกอร์และพื้นหลังได้อย่างชัดเจน โดยสามารถวิเคราะห์จำนวนพิกเซลและความเข้มแสงเพื่อตั้งค่าระดับกันได้จากฮิสโตแกรม ดังรูปที่ 4.9



รูปที่ 4.9 ฮิสโตแกรมของภาพพระดับเทา

วิธีการนี้เป็นวิธีการแปลงภาพพระดับเทาจากกระบวนการในหัวข้อที่ 4.2.2.1 เป็นภาพไบนารี (Binary Image) โดยบริเวณมาร์คเกอร์คือส่วนที่สนใจ ถูกกำหนดให้มีค่าเป็น 1 (หมายถึงสีขาว) ส่วนบริเวณถูงมือและพื้นหลังถูกกำหนดให้มีค่าเป็น 0 (หมายถึงสีดำ) ดังสมการที่ 4.2 ซึ่งผลลัพธ์เป็นดังรูปที่ 4.10

$$\text{gray}(i, j) = \begin{cases} 1, & \text{gray}(i, j) < \text{threshold} \\ 0, & \text{gray}(i, j) \geq \text{threshold} \end{cases} \quad (4.2)$$



ก.

ข.

รูปที่ 4.10 ก. ภาพไบนารีหลังกระบวนการเทรสโพลด์ดิ่งสำหรับกล่องที่ 1
 ข. ภาพไบนารีหลังกระบวนการเทรสโพลด์ดิ่งสำหรับกล่องที่ 2

4.2.4.3 การตรวจนับจำนวนมาร์คเกอร์และหาศูนย์กลางวงกลม



ก.

ข.

รูปที่ 4.11 ก. การตรวจจับมาร์คเกอร์แต่ละจุดสำหรับกล่องที่ 1
 ข. การตรวจจับมาร์คเกอร์แต่ละจุดสำหรับกล่องที่ 2

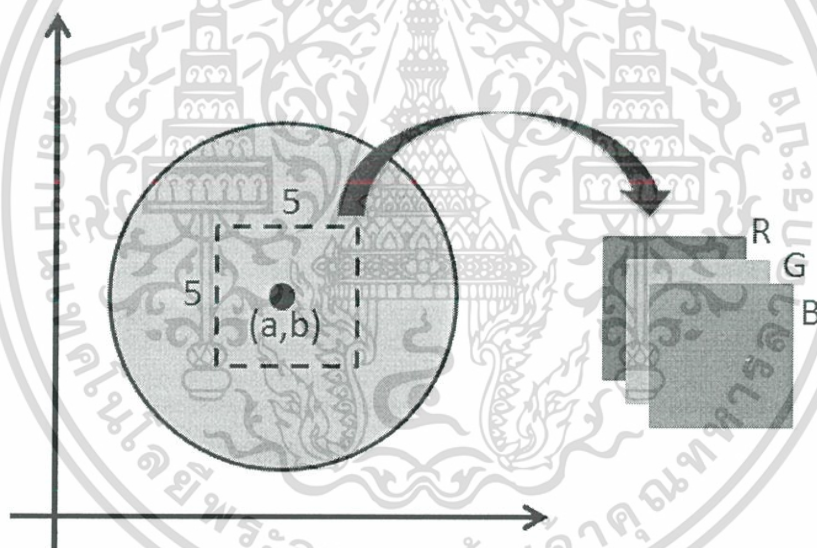
หลังจากกระบวนการแปลงภาพไบนารี ภาพแบ่งออกเป็นสองส่วนอย่างชัดเจน โดยบริเวณสีขาวเป็นส่วนของมาร์คเกอร์ จากภาพดังกล่าวสามารถใช้เทคนิคการแปลงวงกลมของฮัฟ (Circle Hough Transformation) ซึ่งได้กล่าวทฤษฎีตามบทที่ 2 มาประยุกต์ใช้ในการตรวจจับมาร์คเกอร์ในภาพที่มีลักษณะเป็นวงกลมได้ ดังปรากฏในรูปที่ 4.11 โดยมีวงกลมสีแดงล้อมรอบบริเวณขอบมาร์ค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เกอร์บนภาพ จากกระบวนการดังกล่าวนี้สามารถทราบค่าจุดศูนย์กลางของมาร์คเกอร์แต่ละจุดในภาพได้ ซึ่งเป็นประโยชน์ต่อกระบวนการตรวจจับสีต่อไป

4.2.4.4 การตรวจจับสีมาร์คเกอร์

เนื่องจากการวิจัยนี้ใช้เทคนิคของ Direct Linear Transformation เพื่อหาพิกัดสามมิติของมาร์คเกอร์ในแต่ละจุด ดังนั้นจำเป็นต้องทำการจับคู่ตำแหน่งมาร์คเกอร์จากภาพที่ได้จากกล้องแต่ละตัว ด้วยวิธีการตรวจจับสีของมาร์คเกอร์แต่ละจุดบนภาพ ซึ่งกระบวนการตรวจจับสีใช้วิธีการสร้างกรอบสี่เหลี่ยมขนาด 5×5 พิกเซล โดยใช้จุดศูนย์กลางของมาร์คเกอร์แต่ละจุดเป็นหลัก ดังตัวอย่างรูปที่ 4.12 จากนั้นหาค่าเฉลี่ยของความเข้มแสงในแต่ละองค์ประกอบ R G และ B ของพื้นที่กรอบสี่เหลี่ยมสำหรับทุกมาร์คเกอร์ เพื่อใช้ค่าเหล่านี้ในการจับคู่ตำแหน่งมาร์คเกอร์จากภาพที่ได้จากกล้องสองตัว



รูปที่ 4.12 การสร้างกรอบขนาด 5×5 พิกเซล รอบจุดศูนย์กลางวงกลม

วิธีการคำนวณเพื่อให้ได้มาซึ่งตำแหน่งมาร์คเกอร์ที่คู่กันด้วยวิธีการตรวจจับสีนี้ ใช้ผลรวมที่น้อยที่สุดของค่าผลต่างสัมบูรณ์ (Sum of Absolute Different) ตามสมการที่ 4.3 โดยต้องใช้ค่าเฉลี่ยของพื้นที่สี่เหลี่ยมจากภาพทั้งสองกล้องในการเทียบค่า โดย R G และ B แทนค่าเฉลี่ยขององค์ประกอบสีจากกล้องตัวที่ 1 และ R' G' และ B' แทนค่าเฉลี่ยขององค์ประกอบสีจากกล้องตัวที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

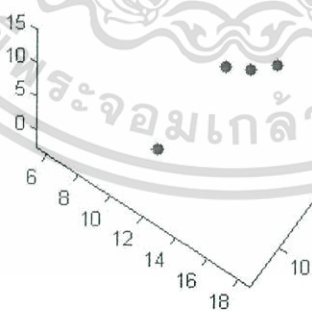
$$\text{pair} = \min(\sum_{i=1}^n (|R - R'_i| + |G - G'_i| + |B - B'_i|)) \tag{4.3}$$

4.2.4.5 การคำนวณพิกัดสามมิติแบบ Direct Linear Transformation

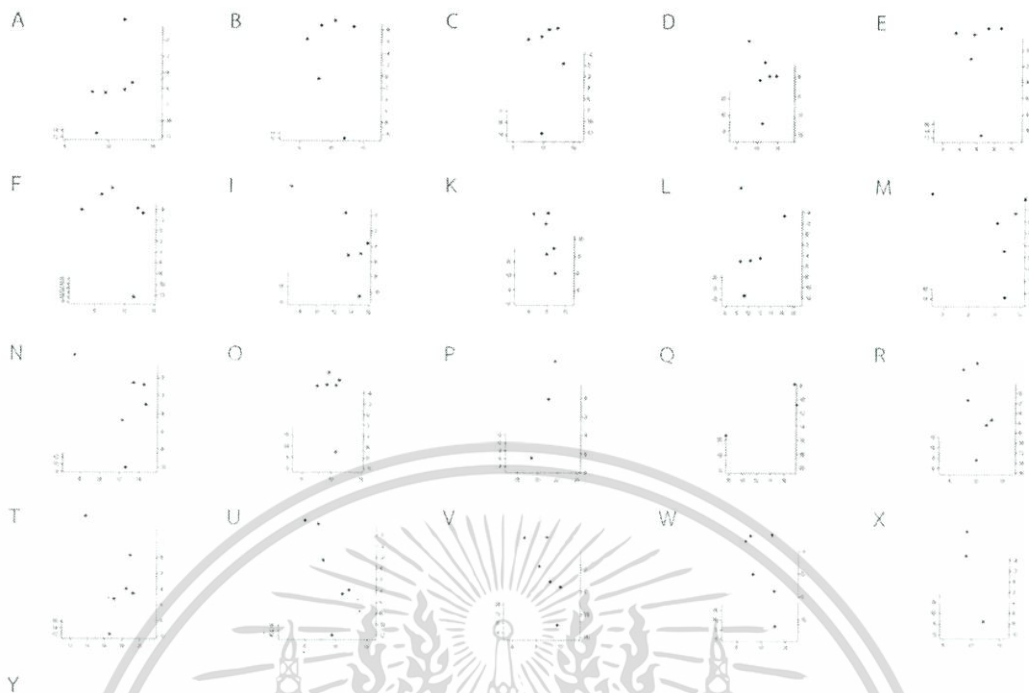
เมื่อสามารถจับคู่ตำแหน่งมาร์คเกอร์บนภาพจากกล้องสองตัว หลังจากนั้นเข้าสู่กระบวนการคำนวณเพื่อหาพิกัดสามมิติตามวิธีการแบบ Direct Linear Transformation ดังที่ได้อธิบายในบทที่ 2 โดยเมื่อทราบจุดร่วมของทั้งสองภาพ สามารถนำไปแทนค่าในสมการที่ 2.9 ได้ดังนี้

$$\begin{bmatrix} um_3^T - wm_1^T \\ vm_3^T - wm_2^T \\ u'm_3^T - w'm_1^T \\ v'm_3^T - w'm_2^T \end{bmatrix} X = AX = 0 \tag{4.4}$$

โดยค่า u และ v กำหนดเป็นค่าจุดศูนย์กลางมาร์คเกอร์บนภาพของกล้องที่ 1 ในทำนองเดียวกัน u' และ v' กำหนดเป็นค่าจุดศูนย์กลางมาร์คเกอร์บนภาพของกล้องที่ 2 โดยจุดศูนย์กลางของทั้งสองกล้องต้องเป็นจุดศูนย์กลางที่เป็นคู่เดียวกัน ที่คำนวณได้จากสมการที่ 4.3 ในส่วนของ w และ w' มีค่าเท่ากับ 1 เนื่องจากระบบพิกัดภาพเป็นแบบสองมิติ หลังจากนั้นทำการแก้สมการแบบโฮโมจีเนียส เพื่อคำนวณค่า X ต่อไป ซึ่งได้ผลลัพธ์จากการคำนวณด้วยวิธีดังกล่าวดังรูปที่ 4.13



รูปที่ 4.13 ตัวอย่างพิกัดสามมิติจากการคำนวณแบบ DLT ของท่าทาง Y



รูปที่ 4.14 พิกัดสามมิติจากการคำนวณแบบ DLT สำหรับทุกท่าทางที่ใช้ในการฝึกฝน

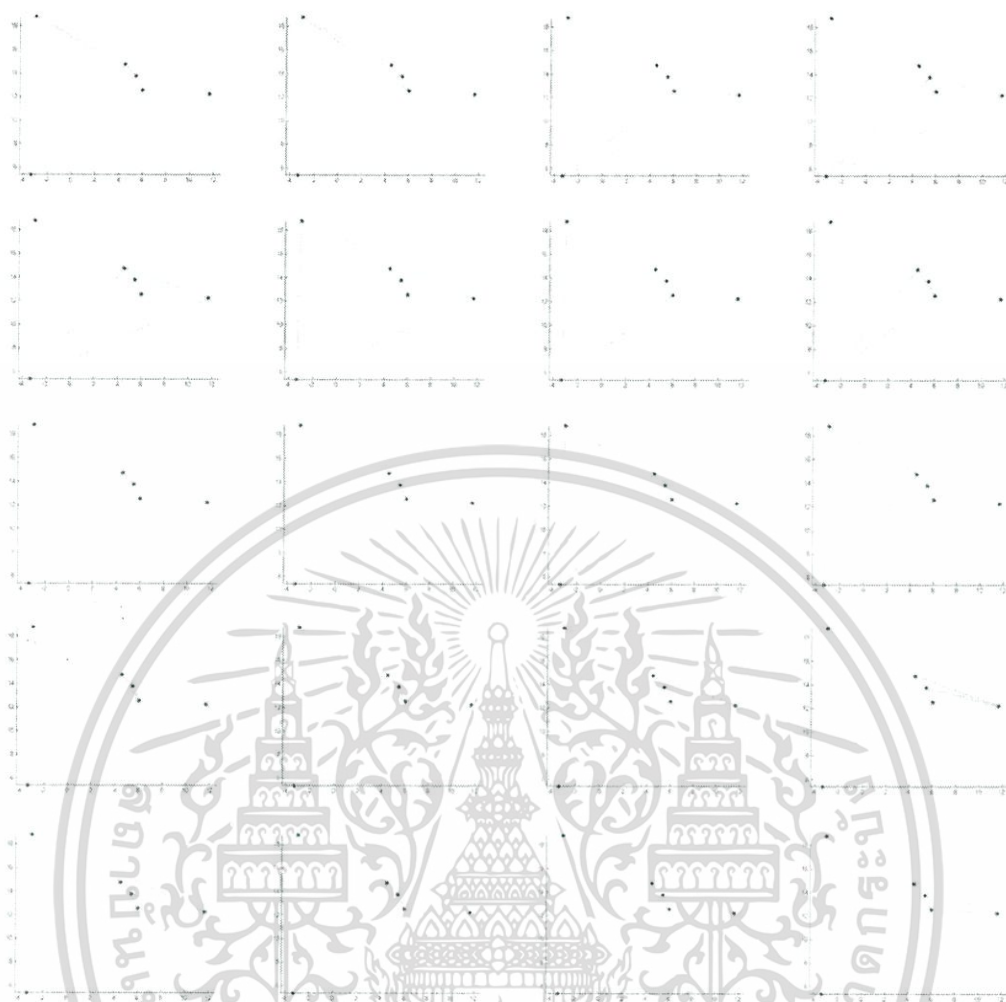
4.2.4.6 การสกัดคุณลักษณะท่าทางด้วยวิธีการคำนวณพื้นที่สามเหลี่ยมและมุม

หัวข้อนี้อธิบายถึงวิธีการสกัดคุณลักษณะสำหรับแต่ละท่าทางของภาษามือ จากแผนผังรูปที่ 4.2 สังเกตว่าวิธีการตรวจนับจำนวนมาร์คเกอร์เป็นหนึ่งในวิธีการสกัดคุณลักษณะของภาษามือได้ กล่าวคือถ้าหากอัลกอริทึมนับจำนวนมาร์คเกอร์ได้ 1 จุด (N=1), 2 จุด (N=2) และ 4 จุด (N=4) หมายถึงเป็นท่าของตัวอักษร ‘G’ ‘H’ และ ‘S’ ตามลำดับ

นอกจากนั้นยังมีท่าทางที่มีจำนวนมาร์คเกอร์เท่ากับ 3 และ 6 จุด ในหลายท่าทาง ดังนั้นจำเป็นต้องมีการคำนวณเพื่อสกัดคุณลักษณะสำหรับท่าทางเหล่านี้เป็นพิเศษและเป็นคุณลักษณะที่ไม่แปรผันตามเลขาคณิต โดยงานวิจัยนี้นำเสนอวิธีการคำนวณโดยใช้พื้นที่ของสามเหลี่ยมทั้งหมดที่เกิดขึ้นสำหรับท่าทางที่มีจำนวนมาร์คเกอร์ 6 จุด และการคำนวณพื้นที่และมุมภายในสามเหลี่ยมสำหรับท่าทางที่มีจำนวนมาร์คเกอร์ 3 จุด ซึ่งจำนวนสามเหลี่ยมที่สามารถสร้างได้จากจำนวนจุดมาร์คเกอร์ที่ตรวจจับได้ เป็นไปตามสมการที่ 4.5

$$C_{n,3} = \frac{n!}{(n-3)!3!} \tag{4.5}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



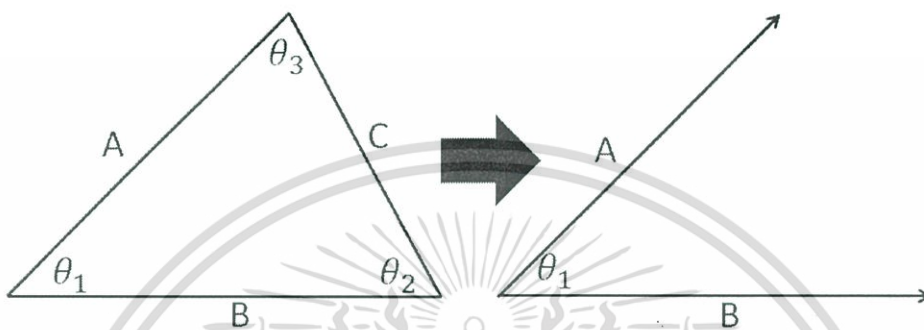
รูปที่ 4.15 การสร้างสามเหลี่ยมจากมาร์คเกอร์ 6 จุด

สมการที่ 4.5 เรียกว่าสมการการจัดหมู่ สามารถคำนวณจำนวนสามเหลี่ยมทั้งหมดที่สร้างได้จากจำนวนจุดมาร์คเกอร์ที่ตรวจจับได้ หากจำนวนมาร์คเกอร์เท่ากับ 6 และ 3 จุด จำนวนสามเหลี่ยมที่เกิดขึ้นได้จำนวน 20 และ 1 สามเหลี่ยม ตามลำดับ รูปที่ 4.15 เป็นการสร้างสามเหลี่ยมจำนวน 20 สามเหลี่ยม เพื่อทำการหาพื้นที่สามเหลี่ยมใดๆ ด้วยสมการของฮีรอน (Heron's Equation)

สำหรับกรณีที่ตรวจจับจำนวนมาร์คเกอร์ได้ 3 จุด ได้แก่ ท่าของตัวอักษร 'P' 'Q' และ 'X' โดยสามารถสร้างสามเหลี่ยมและหาพื้นที่ได้เพียง 1 สามเหลี่ยม ทำให้การสกัดคุณลักษณะของท่าทางมีความใกล้เคียงกัน ดังนั้นจำเป็นต้องเพิ่มคุณลักษณะสำหรับตัวอักษรเหล่านี้ด้วยค่าของมุมภายในสามเหลี่ยม สามารถคำนวณด้วยสมการที่ 4.6 โดยมีความสัมพันธ์ดังรูปที่ 4.16 โดยมุมที่คำนวณได้

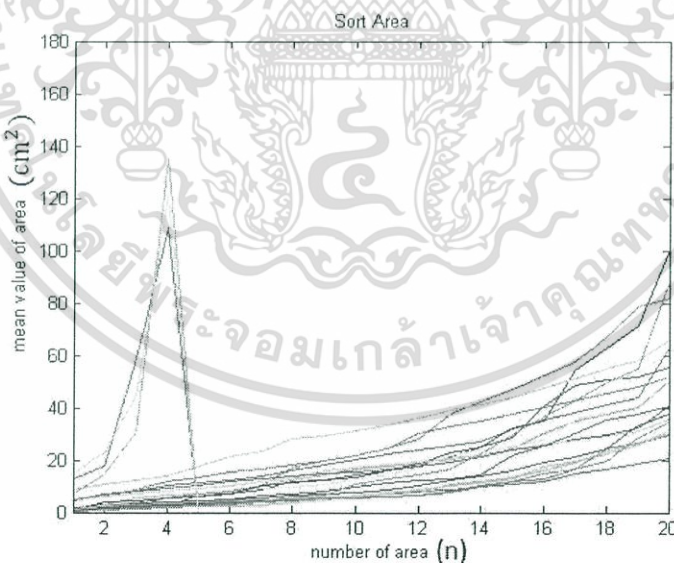
เป็นมุมของทางด้านทางเวกเตอร์กระทำกัน สามารถใช้หลักการนี้คำนวณหามุมทั้งสามด้านของสามเหลี่ยมได้

$$\theta = \cos^{-1} \frac{A \cdot B}{|A||B|} \tag{4.6}$$



รูปที่ 4.16 ความสัมพันธ์ระหว่างมุมและเวกเตอร์ของสามเหลี่ยม

หลังจากคำนวณพื้นที่และมุมด้วยวิธีการดังกล่าวแล้ว จากนั้นนำค่าที่ได้มาจัดเรียงจากน้อยไปมาก รูปที่ 4.17 แสดงกราฟความสัมพันธ์ระหว่างค่าที่คำนวณได้กับท่าทางภาษามือที่แตกต่างกันไป



รูปที่ 4.17 กราฟความสัมพันธ์ของค่าเฉลี่ยแต่ละท่าทางกับค่าพื้นที่และมุม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5 กระบวนการคัดแยกกลุ่มข้อมูลด้วยโครงข่ายประสาทเทียม

กระบวนการคัดแยกกลุ่มข้อมูลด้วยโครงข่ายประสาทเทียม แบ่งออกเป็น 2 ขั้นตอนหลัก ได้แก่ ขั้นตอนการฝึกฝน และขั้นตอนการใช้งานโครงข่ายประสาทเทียม

ขั้นตอนการฝึกฝนโครงข่ายประสาทเทียม เป็นขั้นตอนที่สำคัญก่อนนำโครงข่ายประสาทเทียมไปใช้งาน ซึ่งจำเป็นต้องมีข้อมูลสำหรับการฝึกฝนสำหรับแต่ละท่าทางจำนวนหนึ่ง ยังมีข้อมูลสำหรับฝึกฝนมาก โครงข่ายประสาทเทียมก็ยิ่งมีประสิทธิภาพมากยิ่งขึ้น ขั้นตอนการฝึกฝนนี้อาศัยคู่การฝึกฝน (Training pair) เป็นไปตามสมการที่ 4.7 เมื่อ p คือ อินพุตของข้อมูลที่คำนวณจากการสกัดคุณลักษณะเฉพาะของแต่ละท่าทาง และ t คือ ค่าของเป้าหมายของแต่ละท่าทาง ดังแสดงในรูปที่ 4.18 การออกแบบการฝึกฝนใช้วิธีแบบ Backpropagation ประกอบด้วยชั้นอินพุตจำนวน 21 โหนด และชั้นซ่อนเร้นจำนวน 21 โหนด และใช้ฟังก์ชันกระตุ้นชนิด Log-Sigmoid ที่ให้ค่าเอาต์พุตในช่วง 0 ถึง 1 ซึ่งงานวิจัยนี้เตรียมข้อมูลสำหรับการฝึกฝนทั้งหมด 2,100 ภาพ (ท่าทางละ 100 ภาพ) ต่อผู้ทดลอง 1 คน

$$\text{training pair} = \{(p_1, t_1), (p_2, t_2), \dots, (p_{21}, t_{21})\} \quad (4.7)$$

อินพุต (p)	เป้าหมาย (t)
A	1 0
B	0 1 0
C	0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
D	0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
E	0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
F	0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
I	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
K	0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
L	0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
M	0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
N	0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
O	0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
P	0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
Q	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
R	0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
T	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
U	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
V	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0
W	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
X	0 1 0
Y	0 1

รูปที่ 4.18 คู่การฝึกฝนระหว่างอินพุตและเป้าหมาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากที่ฝึกฝนโครงข่ายประสาทเทียมเรียบร้อยแล้ว ขั้นตอนต่อไปคือการนำโครงข่ายประสาทเทียมที่ฝึกฝนนี้ไปใช้ เมื่อมีข้อมูลภาพใดๆ เข้าสู่ระบบ ถูกดำเนินการตามขั้นตอนดังรูปที่ 4.2 และถูกนำไปประมวลผลกับโครงข่ายประสาทเทียมที่เตรียมไว้ ซึ่งสามารถคืนค่าผลลัพธ์ในรูปแบบเดียวกันกับค่าเป้าหมาย (t) โดยค่าเหล่านั้นถูกแปลงกลับไปเป็นตัวอักษรแสดงบนหน้าจอคอมพิวเตอร์ ดังรูปที่ 4.19



รูปที่ 4.19 การแสดงผลเป็นตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

ผลการทดลองและอภิปรายผลการทดลอง

(Result and Discussion)

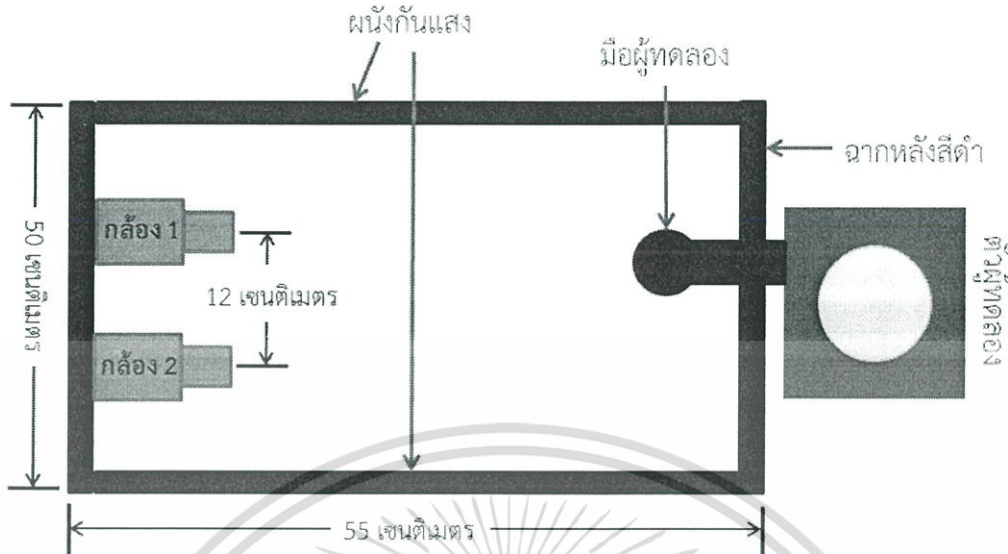
ในบทนี้อธิบายถึงขั้นตอนการทดลอง เพื่อนำไปสู่การแสดงผลการทดลองของอัลกอริทึมสำหรับแปลภาษามือในรูปแบบการสะกดคำด้วยนิ้วแบบ American Sign Language (ASL) ซึ่งการทดลองนี้จะแบ่งออกเป็น 2 ส่วน ได้แก่ การทดลองเพื่อทดสอบความถูกต้องของอัลกอริทึมของการแปลภาษามือแบบพยัญชนะ A-Z และการทดลองเพื่อทดสอบการแปลภาษามือแบบต่อเนื่อง โดยเนื้อหาส่วนถัดไปอธิบายดังหัวข้อต่อไปนี้

1. การออกแบบระบบการจับภาพเพื่อการแปลภาษามือ
2. การจับเก็บข้อมูลภาพเพื่อนำไปเข้ากระบวนการฝึกฝนของโครงข่ายประสาทเทียม
3. การทดสอบความถูกต้องของอัลกอริทึมของการแปลภาษามือแบบพยัญชนะ A-Z
4. การทดสอบการแปลภาษามือแบบต่อเนื่อง

5.1 การออกแบบระบบการจับภาพเพื่อแปลภาษามือ

การออกแบบระบบการจับภาพมีความจำเป็นอย่างยิ่งในลำดับแรกๆ เนื่องจากปริมาณแสงและสิ่งแวดล้อม มีผลกระทบต่อกระบวนการประมวลผลภาพ แต่หากมีการออกแบบระบบการจับภาพที่ดี การควบคุมปริมาณแสง การกำหนดสีพื้นหลังภาพ ทำให้การประมวลผลภาพในขั้นตอนแรก (Preprocessing Process) มีความสะดวกสบาย และลดระยะเวลาการประมวลผลภาพในส่วนถัดไป

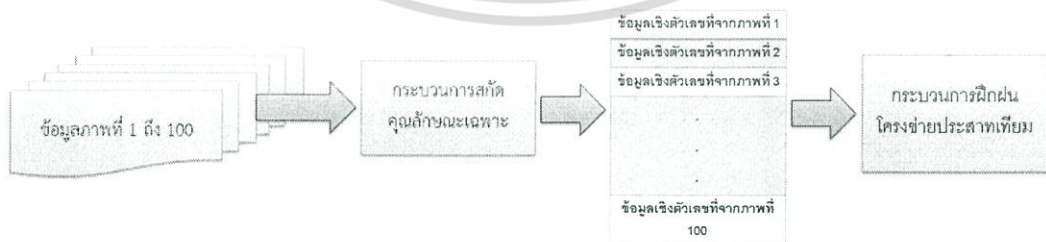
การออกแบบระบบเพื่อกำหนดสิ่งแวดล้อมในการจับภาพข้อมูลเพื่อแปลภาษามือของโครงการนี้ กำหนดให้ผู้ทดสอบใส่ถุงมือสีดำที่มีลูกบิดสีแตกต่างกัน ที่ทำหน้าที่เป็นมาร์คเกอร์ (Marker) อยู่ที่ปลายนิ้วและข้อมือ นอกจากนั้นพื้นหลังของมือและแขนเสื้อถูกกำหนดให้เป็นสีดำทั้งหมด เพื่อง่ายต่อกระบวนการประมวลผลในขั้นแรก และกล้องจำนวน 2 ตัว ถูกยึดเข้ากับอุปกรณ์อย่างแข็งแรง ไม่สามารถขยับได้หลังจากที่ปรับเทียบกล้องทั้งสองเรียบร้อยแล้ว โดยกระบวนการจับเก็บภาพจะกระทำพร้อมกันทั้งสองกล้อง ผ่านทาง Port USB ของคอมพิวเตอร์ ซึ่งผู้ทดลองนั่งหันหน้าเข้าหากกล้องและยื่นแขนเข้าไปภายในระบบการจับภาพที่ถูกออกแบบ ดังรูปที่ 5.1



รูปที่ 5.1 แบบจำลองระบบการจับภาพของโครงการ

5.2 การจับเก็บข้อมูลภาพเพื่อนำไปเข้ากระบวนการฝึกฝนของโครงข่ายประสาทเทียม

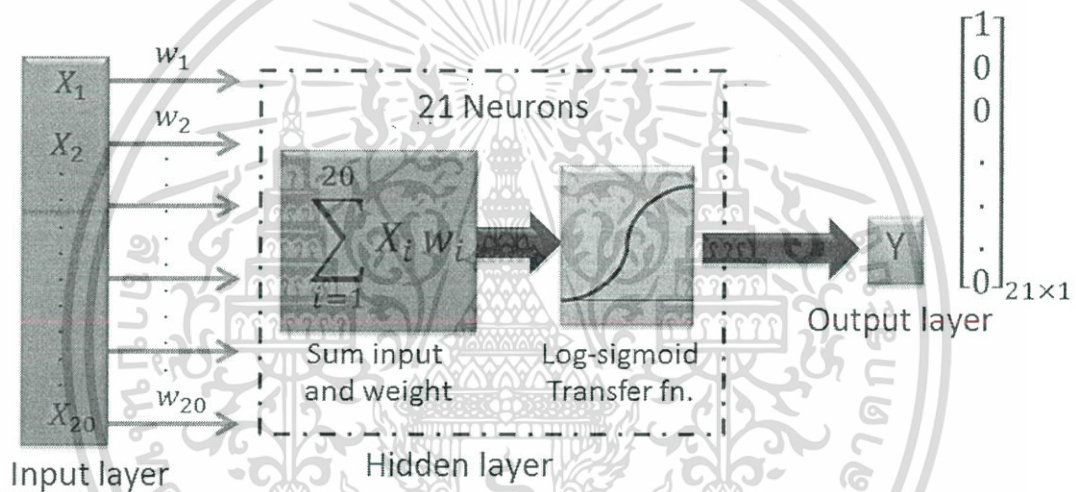
โครงการนี้ใช้กระบวนการฝึกฝนของโครงข่ายประสาทเทียมแบบ Feedforward Backpropagation Training โดยมีความจำเป็นที่ต้องจัดเก็บข้อมูลภาพของผู้ทดลอง เพื่อนำไปแปลเป็นข้อมูลตัวเลข ดังที่อธิบายในหัวข้อวิธีการการสกัดคุณลักษณะเฉพาะของแต่ละท่าทาง (Feature Extraction) การฝึกฝนโครงข่ายประสาทเทียมในโครงการนี้ใช้ภาพ 4,200 ภาพต่อผู้ทดสอบ 1 คน (ท่าทางละ 100 ภาพ) สำหรับกล่องจำนวน 2 ตัว โดยตัวอักษรที่ต้องการการฝึกฝนได้แก่ A, B, C, D, E, F, I, K, L, M, N, O, P, Q, R, T, U, V, W, X, Y เนื่องจากเป็นท่าทางที่ปรากฏมาร์คเกอร์จำนวน 6 และ 3 จุด ขณะแสดงท่าทาง ดังนั้นจำเป็นต้องมีการฝึกฝนข้อมูลดังกล่าว รูปที่ 5.2 แสดงรูปแบบการนำข้อมูลภาพไปฝึกฝนในระบบของโครงข่ายประสาทเทียมที่ออกแบบ



รูปที่ 5.2 ขั้นตอนการฝึกฝนโครงข่ายประสาทเทียมสำหรับ 1 ท่าทาง

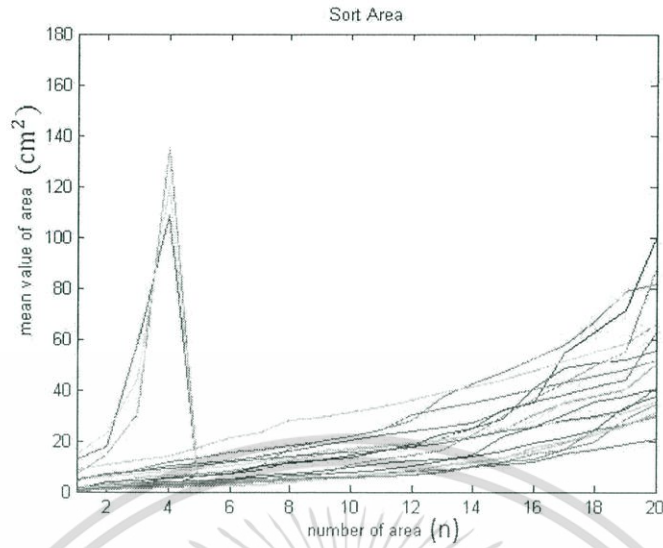
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากจัดเก็บภาพสำหรับฝึกฝนท่าทางละ 100 ภาพ ครบถ้วนแล้ว ขั้นตอนของกระบวนการฝึกฝนข้อมูลเริ่มจากการรับภาพแต่ละท่าทางเข้าไป เพื่อสกัดคุณลักษณะเฉพาะ (Feature Extraction) ดังที่อธิบายในหัวข้อ 4.2.4 ในบทที่ 4 เพื่อแปลงข้อมูลภาพเป็นข้อมูลตัวเลข ซึ่งแต่ละภาพจะมีข้อมูลตัวเลขเป็นเมทริกซ์ขนาด 1×20 ที่ประกอบด้วยขนาดของพื้นที่และมุมของสามเหลี่ยมใดๆ ซึ่งเรียงจากน้อยไปมาก จากนั้นข้อมูลตัวเลขเหล่านี้ ถูกนำไปเข้ากระบวนการฝึกฝนโครงข่ายประสาทเทียม โดยออกแบบให้ชั้นอินพุท (Input Layer) ของโครงข่ายประสาทเทียมประกอบด้วย 20 นิวรัล (Neural) ซึ่งจะรับค่าอินพุท (Input) X_1, X_2, \dots, X_{20} ที่ได้จากการสกัดคุณลักษณะเฉพาะของแต่ละท่าทางเข้าไป โดยแต่ละอินพุทจะถูกคูณกับค่าน้ำหนัก W_i และส่งเข้าสู่ส่วนถัดไป คือ ชั้นซ่อนเร้น (Hidden Layer) ประกอบด้วย 21 นิวรัล และชั้นเอาต์พุท (Output Layer) จะคืนค่าคำตอบในรูปแบบเมทริกซ์ 21×1 ที่ภายในมีค่าเพียง 0 และ 1 เท่านั้น ดังแสดงในรูปแบบที่ 5.3

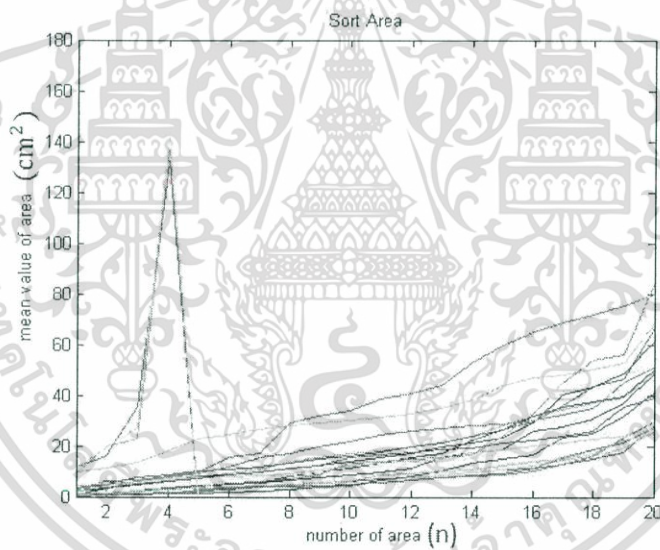


รูปที่ 5.3 การออกแบบกระบวนการฝึกฝนโครงข่ายประสาทเทียม

สำหรับการทดลองในโครงการนี้ใช้ผู้ทดสอบจำนวน 4 คน โดยผู้ทดสอบเป็นคนปกติทั่วไปที่นำมาฝึกฝนการทำท่าทางแบบ ASL สำหรับการทดสอบความแม่นยำของอัลกอริทึม ในการฝึกฝนข้อมูลและทดสอบความถูกต้องของอัลกอริทึมของการแปลภาษามือแบบพยัญชนะ โดยจำนวนชุดข้อมูลการฝึกฝนโครงข่ายประสาทเทียมนั้นถูกแยกเป็นรายบุคคล เมื่อฝึกฝนเสร็จเรียบร้อยแล้ว จึงได้ค่าพารามิเตอร์ของน้ำหนักที่เหมาะสมของผู้ทดสอบแต่ละคนเก็บไว้ เพื่อนำไปใช้ในการทดสอบข้อมูลในกระบวนการการนำไปใช้ (Application Process) รูปที่ 5.4 ถึง 5.7 แสดงค่าเฉลี่ยของข้อมูลที่นำไปฝึกฝนของแต่ละบุคคล ซึ่งเส้นของกราฟแต่ละเส้น หมายถึง ค่าเฉลี่ยของท่าทางแต่ละตัว ที่แตกต่างกันออกไป

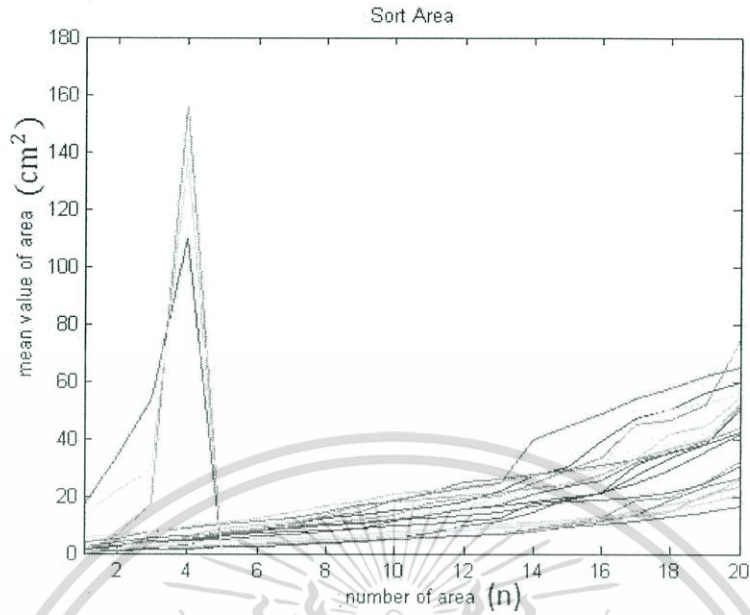


รูปที่ 5.4 ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 1)

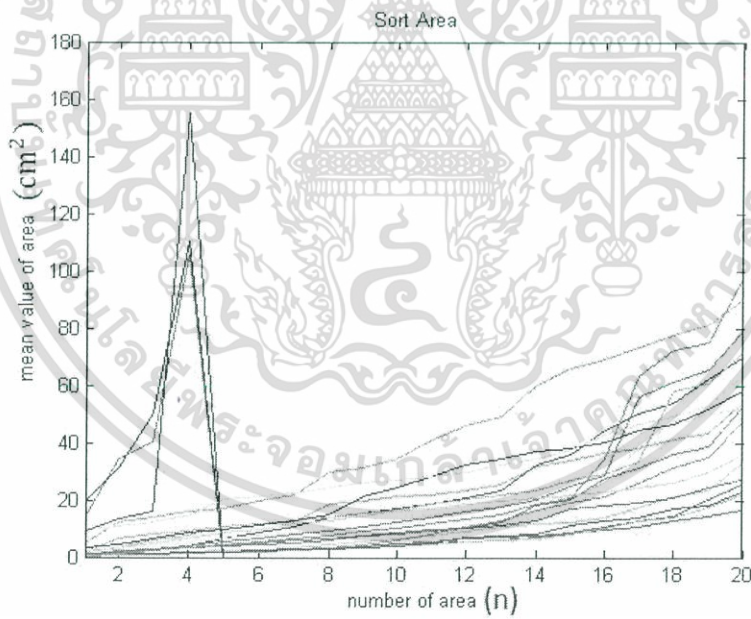


รูปที่ 5.5 ค่าเฉลี่ยข้อมูลที่นำไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.6 ค่าเฉลี่ยข้อมูลที่น่าไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 3)



รูปที่ 5.7 ค่าเฉลี่ยข้อมูลที่น่าไปฝึกฝนของแต่ละท่าทาง (ผู้ทดลองคนที่ 4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ทดสอบอัลกอริทึมสำหรับการแปลภาษามือ

5.3.1 ทดสอบความถูกต้องสำหรับท่า A-Z

หลังจากการฝึกฝนโครงข่ายประสาทเทียมด้วยข้อมูลภาพจำนวนท่าละ 100 ภาพแล้ว ขั้นตอนถัดไปเป็นการทดสอบความถูกต้องของอัลกอริทึมของการแปลภาษามือแบบพหุชนะ A-Z โดยข้อมูลภาพใดๆ (Unknown Data) ที่นำเข้าสู่ระบบเพื่อประมวลผลด้วยอัลกอริทึมที่พัฒนาขึ้นมา จะแปลงข้อมูลภาพใดๆ เป็นตัวอักษร ซึ่งการทดสอบความแม่นยำในโครงข่ายนี้จะทดสอบท่าทางละ 20 ภาพ กับผู้ทดสอบ 4 คน โดยผลการทดลองเป็นไปตามตารางที่ 5.1 ถึง 5.4

ตารางที่ 5.1 ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 1

ท่าทางที่ทดสอบ	จำนวน (ครั้ง)		ร้อยละความถูกต้อง
	ถูกต้อง	ผิดพลาด	
A	17	3	85.00
B	18	2	90.00
C	20	0	100.00
D	18	2	90.00
E	18	2	90.00
F	20	0	100.00
G	20	0	100.00
H	20	0	100.00
I	19	1	95.00
J	19	1	95.00
K	20	0	100.00
L	20	0	100.00
M	19	1	95.00
N	20	0	100.00
O	18	2	90.00
P	20	0	100.00
Q	20	0	100.00
R	20	0	100.00
S	19	1	95.00
T	19	1	95.00
U	18	2	90.00
V	20	0	100.00
W	18	2	90.00
X	20	0	100.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.1 (ต่อ)

ท่าทางที่ทดสอบ	จำนวน (ครั้ง)		ร้อยละความถูกต้อง
	ถูกต้อง	ผิดพลาด	
Y	18	2	90.00
Z	19	1	95.00
ร้อยละความถูกต้องเฉลี่ยของทุกท่าทาง			95.58

ตารางที่ 5.2 ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 2

ท่าทางที่ทดสอบ	จำนวน (ครั้ง)		ร้อยละความถูกต้อง
	ถูกต้อง	ผิดพลาด	
A	20	0	100.00
B	19	1	95.00
C	19	1	95.00
D	20	0	100.00
E	19	1	95.00
F	19	1	95.00
G	20	0	100.00
H	20	0	100.00
I	20	0	100.00
J	19	1	95.00
K	20	0	100.00
L	20	0	100.00
M	20	0	100.00
N	20	0	100.00
O	18	2	90.00
P	20	0	100.00
Q	19	1	95.00
R	17	3	85.00
S	20	0	100.00
T	18	2	90.00
U	20	0	100.00
V	20	0	100.00
W	20	0	100.00
X	18	2	90.00
Y	20	0	100.00
Z	19	1	95.00
ร้อยละความถูกต้องเฉลี่ยของทุกท่าทาง			96.92

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.3 ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 3

ท่าทางที่ทดสอบ	จำนวน (ครั้ง)		ร้อยละความถูกต้อง
	ถูกต้อง	ผิดพลาด	
A	19	1	95.00
B	19	1	95.00
C	19	1	95.00
D	20	0	100.00
E	19	1	95.00
F	20	0	100.00
G	20	0	100.00
H	20	0	100.00
I	18	2	90.00
J	20	0	100.00
K	17	3	85.00
L	20	0	100.00
M	20	0	100.00
N	18	2	90.00
O	19	1	95.00
P	20	0	100.00
Q	19	1	95.00
R	18	2	90.00
S	19	1	95.00
T	18	2	90.00
U	18	2	90.00
V	19	1	95.00
W	20	0	100.00
X	20	0	100.00
Y	20	0	100.00
Z	20	0	100.00
ร้อยละความถูกต้องเฉลี่ยของทุกท่าทาง			95.96

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.4 ผลการทดลองความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองคนที่ 4

ท่าทางที่ทดสอบ	จำนวน (ครั้ง)		ร้อยละความถูกต้อง
	ถูกต้อง	ผิดพลาด	
A	20	0	100.00
B	20	0	100.00
C	20	0	100.00
D	18	2	90.00
E	17	3	85.00
F	19	1	95.00
G	20	0	100.00
H	20	0	100.00
I	20	0	100.00
J	20	0	100.00
K	18	2	90.00
L	20	0	100.00
M	20	0	100.00
N	18	2	90.00
O	17	3	85.00
P	20	0	100.00
Q	20	0	100.00
R	17	3	85.00
S	20	0	100.00
T	17	3	85.00
U	18	2	90.00
V	19	1	95.00
W	19	1	95.00
X	20	0	100.00
Y	19	1	95.00
Z	20	0	100.00
ร้อยละความถูกต้องเฉลี่ยของทุกท่าทาง			95.38

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.5 สรุปผลการทดลองการทดสอบความแม่นยำในการแปลภาษามือ A-Z ของผู้ทดลองทั้งหมด

ท่าทางที่ทดสอบ	ร้อยละความถูกต้อง				ค่าเบี่ยงเบนมาตรฐาน (S.D.)	ร้อยละความถูกต้องเฉลี่ย
	ผู้ทดลองที่ 1	ผู้ทดลองที่ 2	ผู้ทดลองที่ 3	ผู้ทดลองที่ 4		
A	85.00	100.00	95.00	100.00	7.07	95.00
B	90.00	95.00	95.00	100.00	4.08	95.00
C	100.00	95.00	95.00	100.00	2.89	97.50
D	90.00	100.00	100.00	90.00	5.77	95.00
E	90.00	95.00	95.00	85.00	4.79	91.25
F	100.00	95.00	100.00	95.00	2.89	97.50
G	100.00	100.00	100.00	100.00	0.00	100.00
H	100.00	100.00	100.00	100.00	0.00	100.00
I	95.00	100.00	90.00	100.00	4.79	96.25
J	95.00	95.00	100.00	100.00	2.89	97.50
K	100.00	100.00	85.00	90.00	7.50	93.75
L	100.00	100.00	100.00	100.00	0.00	100.00
M	95.00	100.00	100.00	100.00	2.50	98.75
N	100.00	100.00	90.00	90.00	5.77	100.00
O	90.00	90.00	95.00	85.00	4.08	97.50
P	100.00	100.00	100.00	100.00	0.00	90.00
Q	100.00	95.00	95.00	100.00	2.89	97.50
R	100.00	85.00	90.00	85.00	7.07	90.00
S	95.00	100.00	95.00	100.00	2.89	97.50
T	95.00	90.00	90.00	85.00	4.08	90.00
U	90.00	100.00	90.00	90.00	5.00	92.50
V	100.00	100.00	95.00	95.00	2.89	90.00
W	90.00	100.00	100.00	95.00	4.79	97.50
X	100.00	90.00	100.00	100.00	5.00	97.50
Y	90.00	100.00	100.00	95.00	4.79	96.25
Z	95.00	95.00	100.00	100.00	2.89	97.50
ร้อยละความถูกต้องเฉลี่ยของทุกท่าทาง						95.96

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

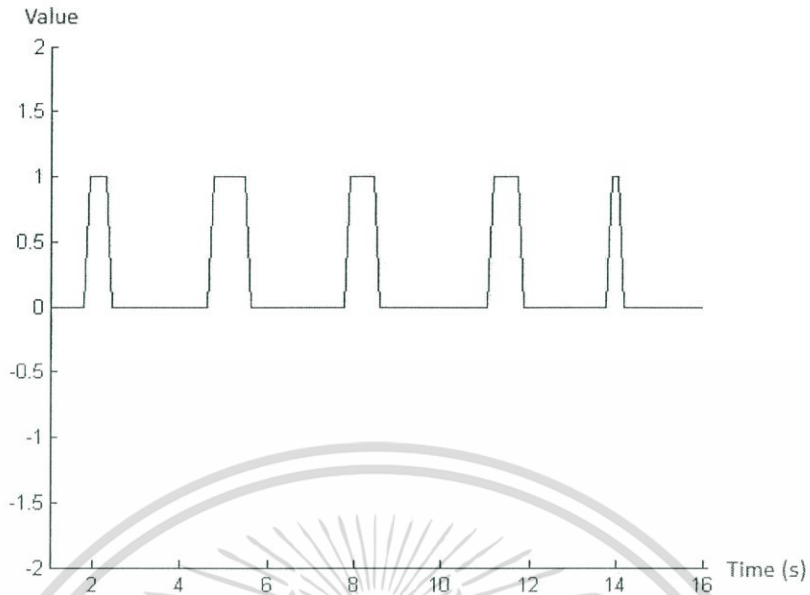
5.3.2 ทดสอบความถูกต้องสำหรับท่า A-Z แบบต่อเนื่อง

การทดสอบการแปลภาษามือแบบต่อเนื่อง โดยใช้วิธีการคำนวณความแตกต่างแบบสัมบูรณ์ (Absolute Difference) ของเฟรมภาพปัจจุบันและภาพก่อนหน้า เพื่อแปลงเป็นสัญญาณการเคลื่อนที่ (Motion Signal) ของมือ ซึ่งวิธีการการแยกแยะท่าทางและขั้นตอนการคำนวณเป็นไปตามที่อธิบายในบทที่ 4 ในหัวข้อนี้ให้ผู้ทดลองทำท่าภาษามือแบบต่อเนื่องจำนวน 3 คำ ได้แก่ KMITL, WIFI และ JANUARY ดังแสดงในรูปที่ 5.8

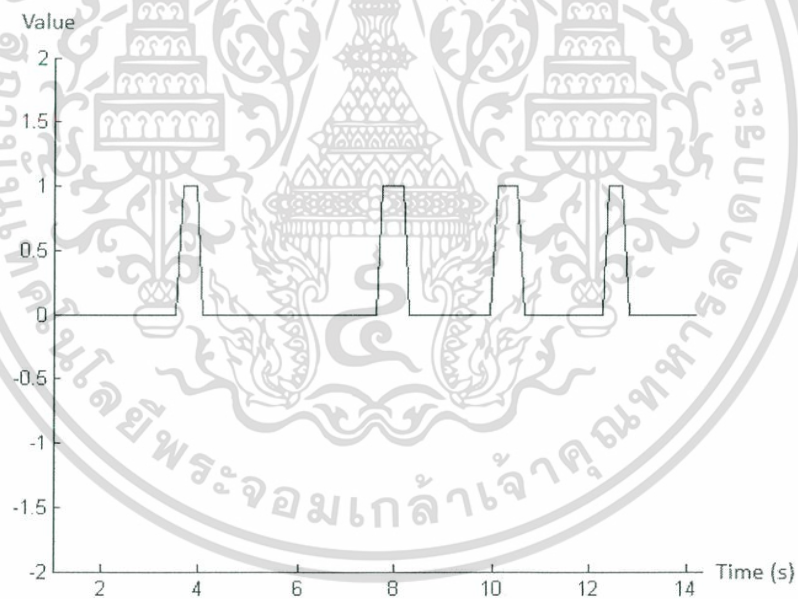


รูปที่ 5.8 ก. ภาษามือแสดงท่าทางคำว่า KMITL ข. ภาษามือแสดงท่าทางคำว่า WIFI
ค. ภาษามือแสดงท่าทางคำว่า JANUARY

สัญญาณการเคลื่อนไหวมือของผู้ทดลองทั้งหมด 4 คน เป็นไปดังรูปที่ 5.9-5.20 โดยกราฟของสัญญาณทางแกน X เป็นแกนของเวลา ส่วนแกน Y เป็นแกนที่แสดงค่าระหว่าง 0 และ 1 ซึ่งแสดงค่าที่ได้จากการคำนวณ Absolute Difference ของภาพแต่ละเฟรม และการตั้งค่าระดับกัน (Threshold) ของสัญญาณการเคลื่อนไหว

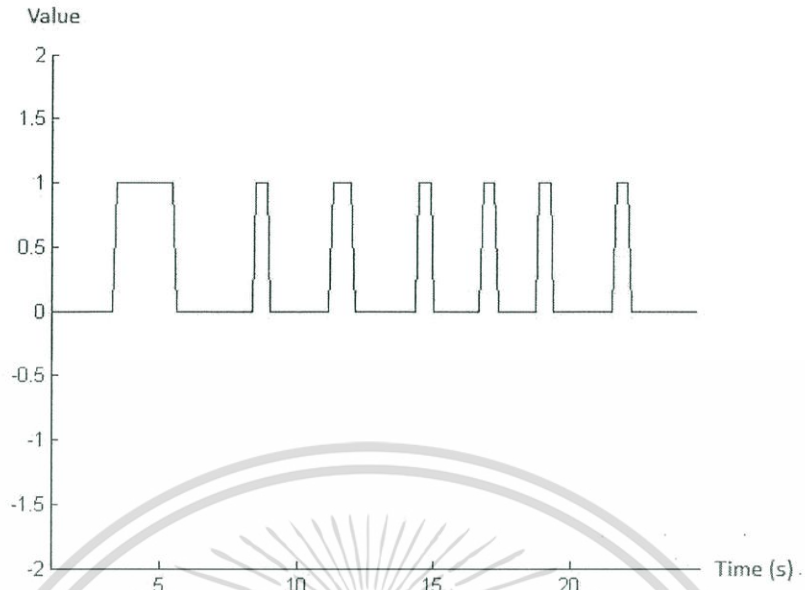


รูปที่ 5.9 สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 1



รูปที่ 5.10 สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

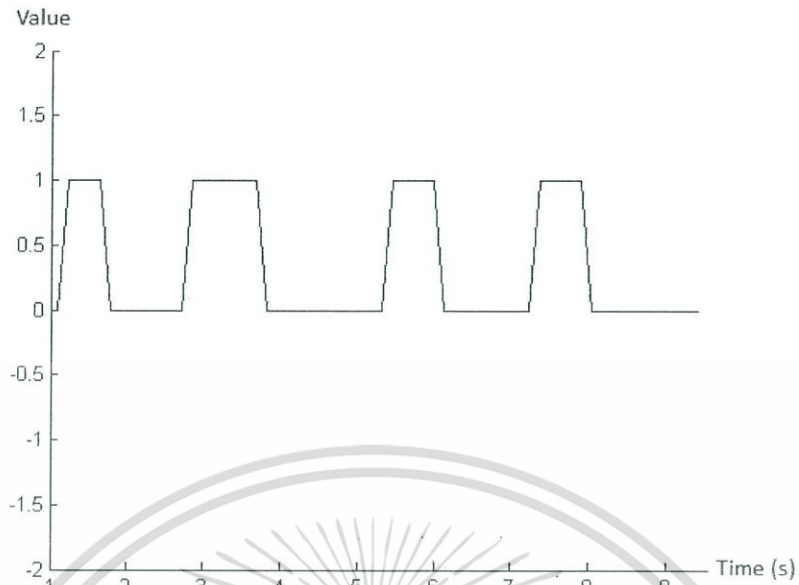


รูปที่ 5.11 สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 1

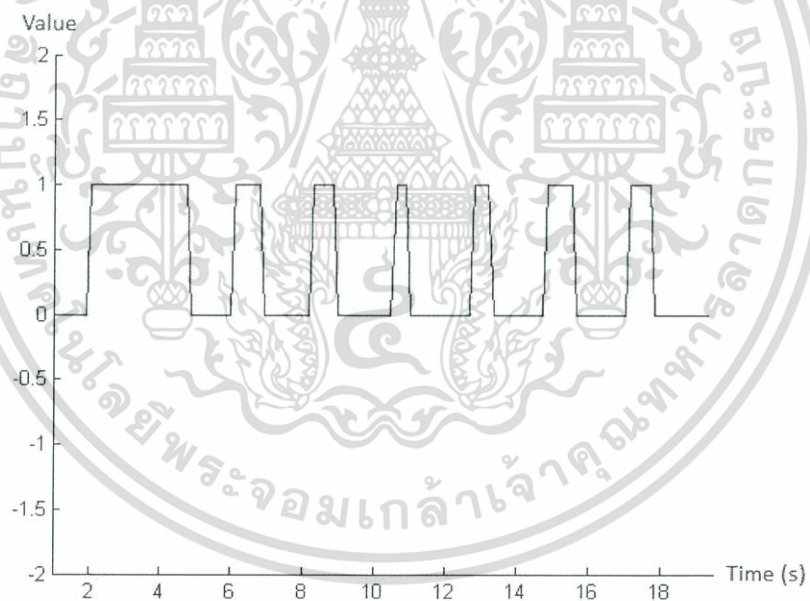


รูปที่ 5.12 สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

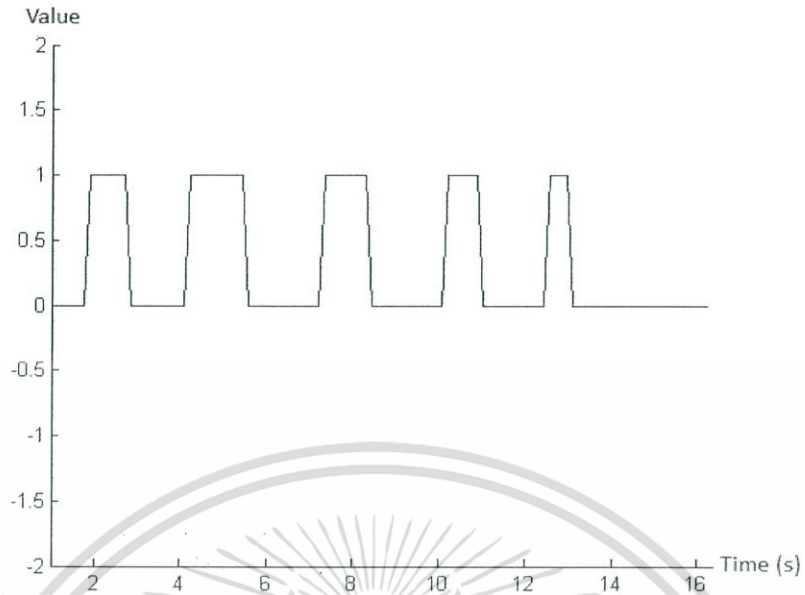


รูปที่ 5.13 สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 2



รูปที่ 5.14 สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

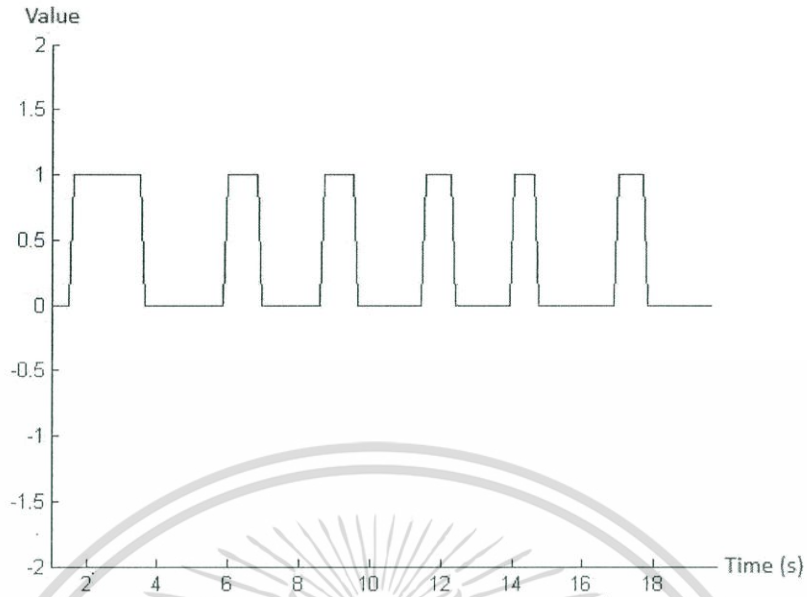


รูปที่ 5.15 สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 3

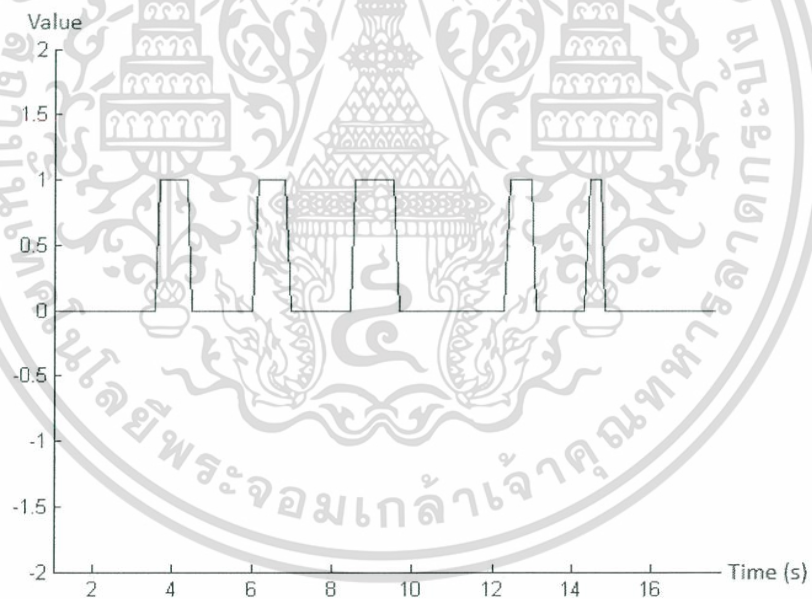


รูปที่ 5.16 สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

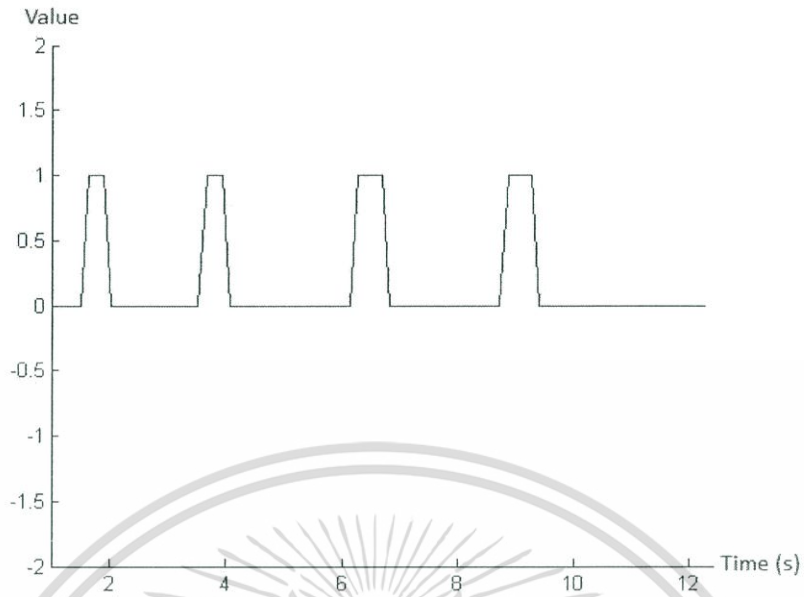


รูปที่ 5.17 สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 3

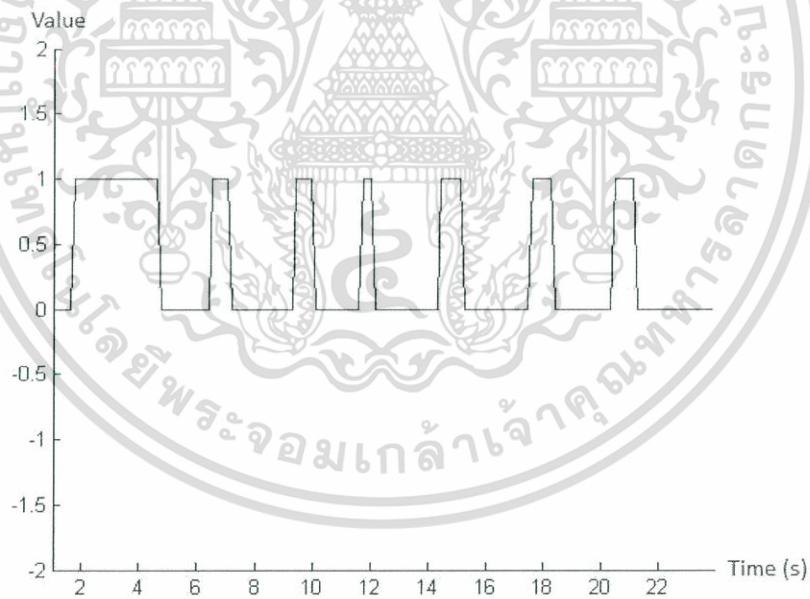


รูปที่ 5.18 สัญญาณการเคลื่อนไหวมือคำว่า KMITL ของผู้ทดลองที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.19 สัญญาณการเคลื่อนไหวมือคำว่า WIFI ของผู้ทดลองที่ 4



รูปที่ 5.20 สัญญาณการเคลื่อนไหวมือคำว่า JANUARY ของผู้ทดลองที่ 4

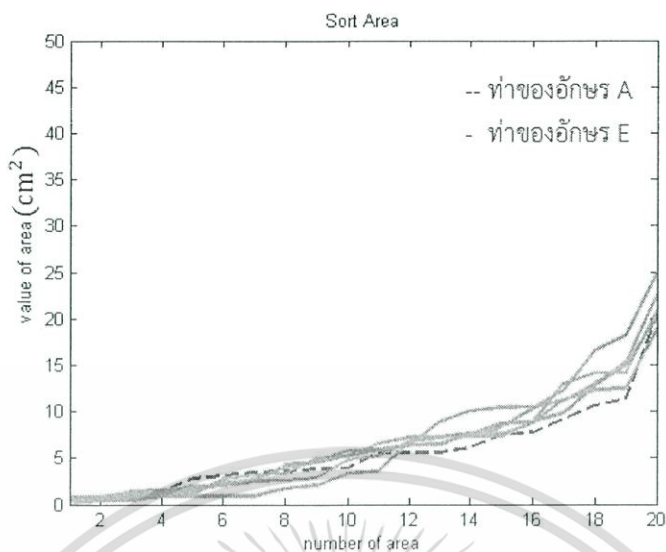
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.4 ผลการทดสอบความถูกต้องของอัลกอริทึม

จากตารางผลการทดลองของผู้ทดลองทั้ง 4 คน ที่ทำการทดสอบความถูกต้องทั้งทำที่เป็น Static และ Dynamic ผลการทดลองแสดงร้อยละความถูกต้องเท่ากับ 95.58, 96.92, 95.96 และ 95.38 ตามลำดับ ซึ่งคำนวณเป็นร้อยละของค่าเฉลี่ยได้เท่ากับ 95.96 ซึ่งความผิดพลาดที่เกิดขึ้นสำหรับการแปลภาษามือ ส่วนหนึ่งเกิดจากการถูกบดบังของมาร์คเกอร์ขณะที่เปลี่ยนจากท่าหนึ่งไปยังอีกท่าหนึ่ง ทำให้การตรวจจับตำแหน่งมาร์คเกอร์จากสองกล้องไม่ถูกต้อง ทำให้การตรวจจับจำนวนมาร์คเกอร์ของทั้งสองกล้องไม่เท่ากัน ทำให้ไม่สามารถแสดงผลการแปลภาษาออกมาได้ ดังรูปที่ 5.21 นอกจากนี้บางท่าทางมีการแปลที่ผิดไป ซึ่งเกิดจากขั้นตอนของการแบ่งประเภทของข้อมูลในการแปลภาษามือ เช่น ท่าทางของตัวอักษร A แปลผิดเป็นตัวอักษร E เนื่องจากชุดข้อมูลจากการสกัดคุณลักษณะเฉพาะของท่าทาง A สำหรับบางตำแหน่งที่ผู้ทดสอบทำ มีความใกล้เคียงกับชุดข้อมูลที่นำมาใช้ฝึกฝนนิรอรอลของตัวอักษร E ดังรูปที่ 5.22 โดยเหตุผลดังกล่าวนี้ทำให้เกิดความผิดพลาดเกิดขึ้น ในทางกลับกันบางชุดข้อมูลของตัวอักษร E มีความใกล้เคียงกับชุดข้อมูลที่นำมาใช้ฝึกฝนนิรอรอลสำหรับตัวอักษร A เช่นกัน



รูปที่ 5.21 ก. ตัวอย่างท่าทางตัวอักษร K จากกล้องตัวที่ 1
ข. ตัวอย่างท่าทางตัวอักษร K ที่ถูกบดบัง จากกล้องตัวที่ 2



รูปที่ 5.22 ตัวอย่างตัวอักษร A ที่คัดแยกผิดพลาดเป็นตัวอักษร E

นอกจากนั้น การทดลองการแปลภาษาเมื่อแบบต่อเนื่องจากผู้ทดลอง 4 คน สามารถคำนวณเป็นสัญญาณการเคลื่อนไหวของท่าทางระหว่างท่าแบบ Static และ Dynamic ได้ และเฟรมภาพที่ถูกนำไปพิจารณามีความถูกต้องสูง โดยอาจมีบางเฟรมภาพที่มีลักษณะเบลอ เนื่องมาจากการเคลื่อนไหวที่ติดต่อกันเร็วเกินไป ทำให้การกระบวนการดึงเฟรมภาพนั้นผิดพลาดไป ดังนั้นควรมีการหยุดการเคลื่อนไหวมือสักช่วงเวลาหนึ่ง หลังจากมีการเปลี่ยนท่าทาง เพื่อให้กระบวนการดึงเฟรมภาพไปพิจารณามีความแม่นยำยิ่งขึ้น ซึ่งจากการทดลองช่วงเวลาการนำเฟรมภาพไปพิจารณาด้วยอัลกอริทึมที่พัฒนาขึ้นมีเวลาเฉลี่ยที่ 1.05 วินาที ด้วยคุณลักษณะของคอมพิวเตอร์ดังตารางที่ 4.1

จากผลการทดลองด้วยวิธีการประมวลผลโดยใช้เทคนิคของ Direct Linear Transformation ในการคำนวณพิกัดของมาร์คเกอร์จากภาพ 2 มิติ เป็น 3 มิติ และด้วยกระบวนการหาคุณลักษณะเฉพาะดังกล่าวนี้ ซึ่งเป็นคุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสัมพันธ์ ทำให้ผลการทดลองมีความถูกต้องมากกว่า 95 เปอร์เซ็นต์ และมีระยะเวลาการประมวลผลภาพในแต่ละท่าทางมีความรวดเร็ว สามารถนำไปใช้งานได้มีประสิทธิภาพ

สรุปผลการทดลองและข้อเสนอแนะ (Conclusion and Suggestion)

วิทยานิพนธ์นี้นำเสนอการแปลภาษามือแบบการสะกดพยัญชนะภาษาอังกฤษ A-Z หรือที่เรียกว่า American Sign Language ซึ่งทำการออกแบบถุงมือสีดำ ที่มีมาร์คเกอร์ที่มีสีแตกต่างกัน จำนวน 6 จุด ติดที่ตำแหน่งปลายนิ้วทั้ง 5 นิ้ว และที่ตำแหน่งข้อมืออีก 1 จุด ใช้กล้องจำนวน 2 ตัว ในการถ่ายภาพวิดีโอ และใช้เทคนิคของ Direct Linear Transformation เพื่อแปลงข้อมูลจุดของ มาร์คเกอร์จากภาพ 2 มิติ เป็น 3 มิติ โดยกระบวนการของอัลกอริทึมที่นำเสนอประกอบด้วย กระบวนการรับภาพ (Image Acquisition) การคำนวณสัญญาณการเคลื่อนไหวของมือ (Motion Signal) การสกัดคุณลักษณะเฉพาะของท่าทาง (Feature Extraction) และการจำแนกท่าทางของ ภาษามือ (Classification)

การทดลองสำหรับวิทยานิพนธ์นี้ เป็นการทดสอบความถูกต้องของอัลกอริทึมในการจำแนก ท่าทางภาษามือและกระบวนการดึงเฟรมภาพเพื่อไปพิจารณาสำหรับการทำท่าทางแบบต่อเนื่อง จาก การทดลองคาร์้อยละความถูกต้องเฉลี่ยของทุกท่าทางเท่ากับ 95.96 ซึ่งมีความแม่นยำสูง เนื่องจาก วิธีการที่ใช้นั้นเป็นการนำข้อมูลที่ถูกละเป็น 3 มิติ มาวิเคราะห์ข้อมูลสำหรับแต่ละท่าทาง ทำให้ ข้อมูลมีความละเอียดและเฉพาะเจาะจงมากกว่าข้อมูล 2 มิติ และไม่มีความแปรผันทางเรขาคณิต ซึ่ง มีคุณสมบัติทางเรขาคณิตที่ไม่แปรผันต่อการแปลงทางเรขาคณิตสมบูรณ์ภายใต้การแปลงแบบคงที่ หรือ Rigid Transformation ได้แก่ ด้านหรือความยาวมุม พื้นที่ และแกนหลัก

นอกจากนี้กระบวนการดึงเฟรมภาพไปพิจารณามีความแม่นยำ โดยค่าเฉลี่ยของขั้นตอนการ ประมวลผลภาพเพื่อแปลภาษามือมีค่าเท่ากับ 1.05 วินาที ซึ่งมีความรวดเร็วในระดับหนึ่ง นอกจากนั้นเงื่อนไขหนึ่งที่ทำให้การดึงเฟรมภาพมีความถูกต้องแม่นยำ คือ ผู้ทดลองต้องหยุดนิ่งสัก ระยะเวลาหนึ่ง เพื่อให้การดึงเฟรมภาพไปพิจารณามีความถูกต้อง ไม่เบลอนั่นเนื่องจากการเคลื่อนไหว มืออย่างต่อเนื่อง

อย่างไรก็ตามยังคงมีข้อเสนอแนะบางประเด็นที่ต้องนำไปพัฒนาและแก้ไข เช่น การถูกบดบัง ของมาร์คเกอร์ในบางท่าทาง ทำให้กระบวนการตรวจจับมาร์คเกอร์ผิดพลาดไป นอกจากนั้นการจำกัด ปริมาณแสงมีความสำคัญเช่นกัน หากมีปริมาณแสงที่มากเกินไปจะเกิดการสะท้อนแสงจากมาร์คเกอร์ ทำให้กล้องเห็นมาร์คเกอร์เป็นสีขาวทั้งหมด จะทำให้กระบวนการตรวจจับสีมาร์คเกอร์ เพื่อจับคู่สี มาร์คเกอร์จากภาพทั้ง 2 กล้อง ผิดพลาดไปเช่นเดียวกัน อีกทั้งกระบวนการการตรวจจับตำแหน่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มาร์คเกอร์โดยวิธีการของ Circle Hough Transformation มีระยะเวลาการประมวลผลที่ค่อนข้างนาน และโดยปกติวิธีการนี้เหมาะกับการตรวจจับวัตถุทรงกลมบนพื้นหลังที่มีลวดลายแตกต่างกัน ดังนั้นจึงมีข้อเสนอแนะให้ใช้วิธีการคำนวณหาจุดศูนย์กลางมวล (Centroid of Mass) ของมาร์คเกอร์ดังกล่าวเพื่อลดเวลาการประมวลผลลง โดยข้อเสนอแนะและความผิดพลาดดังกล่าวนี้ควรต้องได้รับการพัฒนาต่อไป เพื่อให้อัลกอริทึมนี้มีประสิทธิภาพมากยิ่งขึ้น เพื่อที่จะเป็นประโยชน์ต่อไปในภายภาคหน้า



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

- [1] เว็บไซต์ MED-EL international ประเทศไทย, “โครงสร้างของหู” [Online]. Available: <http://www.medel.com/th/anatomy-of-the-ear/>
- [2] พ.ญ. เพ็ญมาศ ธรรมศรีณู, “Noise-Induced Hearing Loss”, วันที่ 24 พฤษภาคม 2554.
- [3] Elaine N. Marieb and Katja Hoehn. “Human Anatomy & Physiology” Ninth Edition, Page 394, 571.
- [4] Written by Chatlada: “สาเหตุความบกพร่องทางการได้ยิน” [Online]. Available: <http://anusan-deaf.ac.th/attachments/article/106/know7.pdf>
- [5] วรุตม์ สุปุทธยางกูร. “เนื้อหาทางวิชาการเรื่องความหมาย-ความสำคัญของภาษามือ” [Online]. <http://web.specialsurat.go.th/PDF/hand.pdf>
- [6] Shielonda Mikle. “Thomas Hopkins Gallaudet” [Online]. Available: <http://www.lifeprint.com/asl101/pages-layout/gallaudet-thomas-hopkins3.htm>, 16, March, 2003.
- [7] The American manual alphabet in photographs [Online]. Available: http://upload.wikimedia.org/wikipedia/commons/5/56/ABC_pict.png
- [8] J. Rekba, J. Bhattacharya and s. Majumder. “Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition” Trendz in Information Sciences and Computing (TISC), 2011 3rd International Conference, 8-9 Dec. 2011.
- [9] Rini Akmelawati, Melanie Po-Leen Ooi and Ye Chow Kuang. “Real-Time Malaysian Sign Language Translation using Colour Segmentation and Neural Network” IMTC 2007 - Instrumentation and Measurement Technology Conference Warsaw, Poland, 1-3 May 2007.
- [10] Phonkrit Chanda, Sansanee Auephanwiriyaikul and Nipon Theera-Umpon. “Thai Sign Language Translation System Using Upright Speed-Up Robust Feature and C-Means Clustering” Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference. 10-15 June 2012.
- [11] Eunseok Jeong, Jaehong Lee and DaeEun Kim. “Finger-gesture Recognition Glove using Velostat” 11th International Conference on Control, Automation and Systems Oct. 26-29, 2011.

- [12] Netchanok Tanyawiwat and Surapa Thiemjarus. "Design of an Assistive Communication Glove using Combined Sensory Channels" Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conferencem, 9-12 May 2012.
- [13] M.S. Sinith, Soorej G Kamal, Nisha B.,Nayana S, Kiran Surendran and Jith P S. "Sign Language Recognition Using Support Vector Machine" Advances in Computing and Communications (ICACC), 2012 International Conference, 9-11 Aug. 2012.
- [14] M. ASHRAFUL AMIN and HONG YAN. "SIGN LANGUAGE FINGER ALPHABET RECOGNITION FROM GABOR-PCA REPRESENTATION OF HAND GESTURES" the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007.
- [15] Andreas Domingo, Rini Akmeliawati and Kuang Ye Chow. "Pattern Matching for Automatic Sign Language Translation System using LabVIEW" Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference, 25-28 Nov. 2007.
- [16] Rudy Hartanto, Adhi Susanto, and P. Insap Santosa. "Preliminary Design of Static Indonesian Sign Language Recognition System" Information Technology and Electrical Engineering (ICITEE), 2013 International Conference, 7-8 Oct. 2013.
- [17] Liya Ding and Aleix M. Martinez. "Three-Dimensional Shape and Motion Reconstruction for the Analysis of American Sign Language" Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06), 17-22 June 2006.
- [18] Mahmoud Elmezain, Ayoub Al-Hamadi, Saira Saleem Pathan and Bernd Michaelis. "Spatio-Temporal Feature Extraction-Based Hand Gesture Recognition for Isolated American Sign Language and Arabic Numbers" Image and Signal Processing and Analysis, 2009. ISPA 2009, 16-18 Sept. 2009.
- [19] Kin Fun Li, Kylee Lothrop, Ethan Gill, and Stephen Lau. "A Web-Based Sign Language Translator Using 3D Video Processing" Network-Based Information Systems (NBIS), 2011 14th International Conference, 7-9 Sept. 2011.

- [20] Myoung-Kyu Sohn, Sang-Heon Lee, Dong-Ju Kim , Byungmin Kim, Hyunduk Kim. “3D Hand Gesture Recognition from One Example” Consumer Electronics (ICCE), 2013 IEEE International Conference, 11-14 Jan. 2013.
- [21] Chenyang Zhang, Xiaodong Yang, and YingLi Tian. “Histogram of 3D Facets: A Characteristic Descriptor for Hand Gesture Recognition” Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops, 22-26 April 2013.
- [22] ชูชาติ ปิณฑวิรุจน์. การประมวลผลภาพดิจิทัลขั้นสูงด้วย Matlab. กรุงเทพมหานคร: หจก.วี.เจ.พรินติ้ง. 2550.
- [23] Camera Calibration Toolbox for Matlab. “First calibration example”[online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html
- [24] Harvey Rho dy. “Hough Circle Transform” [Online]. https://www.cis.rut.edu/class/simg782/lectures/lecture_10/lec782_05_10.pdf, October 11, 2005.
- [25] มหาวิทยาลัยเชียงใหม่, “ทฤษฎีที่ใช้ในการแก้ปัญหา” [Online]. Available: http://archive.lib.cmu.ac.th/full/T/2551/enco1051os_ch2.pdf
- [26] นางสาวรณช เอี่ยมปา, คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, วิทยานิพนธ์ เรื่อง “การสกัดคุณลักษณะภาพลายนิ้วมือโดยอาศัยความไม่แปรผันทางเรขาคณิตซึ่งสร้างจากตำแหน่งของ minutiae ร่วมกับการใช้คุณลักษณะที่ได้จากการฟิลเตอร์ภาพแบบหลายทิศทาง”, หน้าที่ 20-36.

ประวัติผู้เขียน

ชื่อ-นามสกุล นายวัชรินทร์ ตั้งสุขสันต์
วัน เดือน ปีเกิด 14 สิงหาคม 2533
ที่อยู่ 79 ซอยบางแวก 76 แขวงบางแวก เขตภาษีเจริญ กรุงเทพฯ
10160
เบอร์โทรศัพท์ต่อ 02-8652129, 088-008-1457
E-mail w.tangsuksant.m@hotmail.com

ประวัติการศึกษา ปี พ.ศ. 2555 สำเร็จการศึกษา วิศวกรรมศาสตรบัณฑิต
สาขาวิศวกรรมชีวการแพทย์ คณะวิศวกรรมศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ

ผลงานทางวิชาการที่ได้รับการยอมรับ

- ปี ค.ศ. 2010: Topic “Basic Numerical Skill from Electronics Piggy Bank for Learning Disabilities” in The Thai Biomedical Medical Engineering Conference, Pattaya, Thailand.
- ปี ค.ศ. 2012: Topic “Directional Eye Movement Detection System for Virtual Keyboard Controller” in The 5th Biomedical Medical International Conference, Udonthani, Thailand.
- ปี ค.ศ. 2013: Topic “Development Algorithm to Count Blood Cells in Urine Sediment using ANN and Hough Transform” in The 6th Biomedical Medical International Conference, Krabi, Thailand.
- ปี ค.ศ. 2014: Topic “American Sign Language Recognition by Using 3D Geometric Invariant Feature and ANN Classification” in The 7th Biomedical Medical International Conference, Fukuoka, Japan.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้