

บทที่ 1

บทนำ

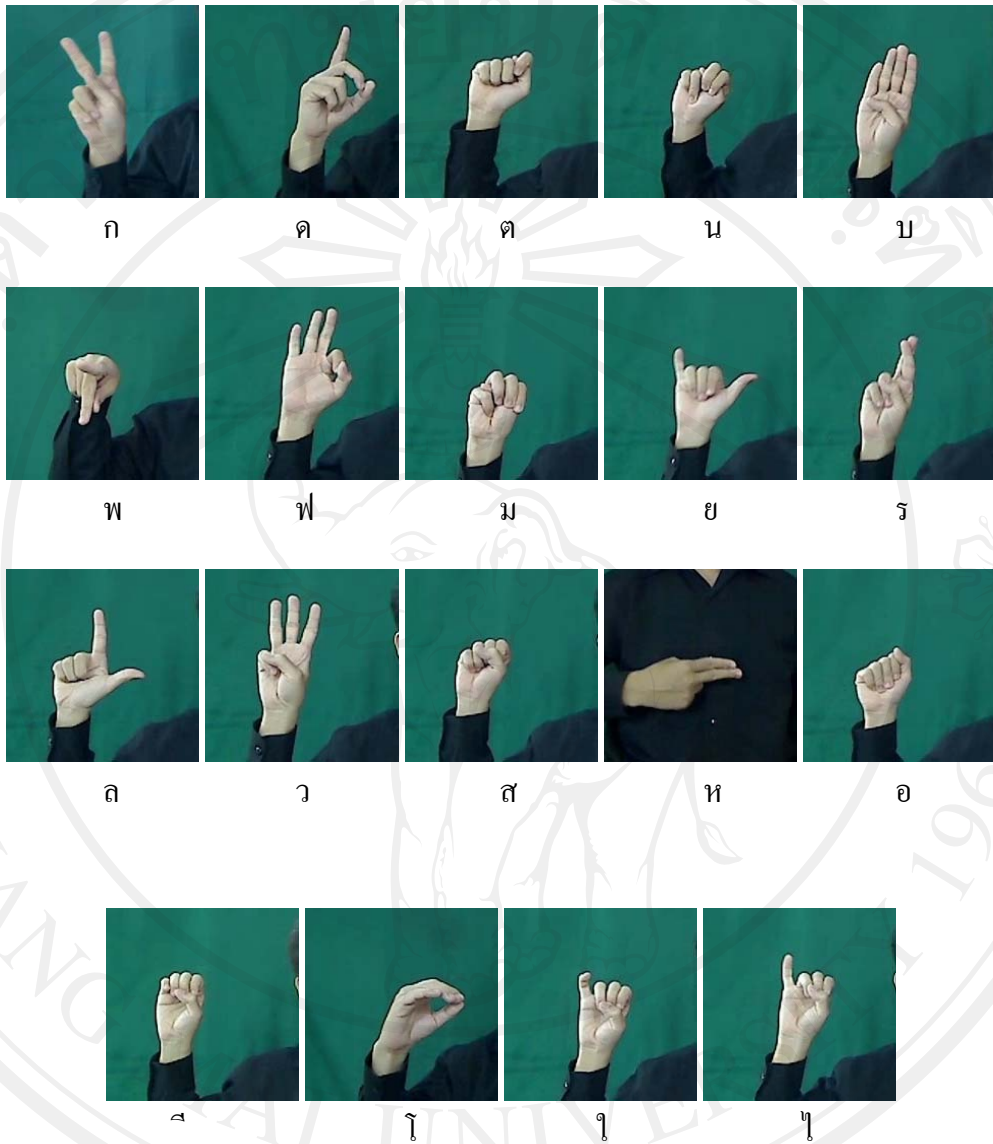
ในบทนี้จะกล่าวถึงที่มาและความสำคัญของปัญหาในการทำวิทยานิพนธ์ เรื่องการประยุกต์ใช้งานซอฟต์แวร์ร่วมกับวิธีการจัดกลุ่มในระบบการแปลภาษามือไทยด้วยวิธีการสะกดนิ้วมือ จากนั้นจะกล่าวถึงสรุปสาระสำคัญของเอกสารที่เกี่ยวข้องกับงานวิจัยนี้ วัตถุประสงค์ของการวิจัย ขอบเขตการวิจัย ประโยชน์ที่ได้รับ และในส่วนท้ายสุดจะเป็นขั้นตอนการทำวิจัย

1.1. ที่มาและความสำคัญของปัญหา

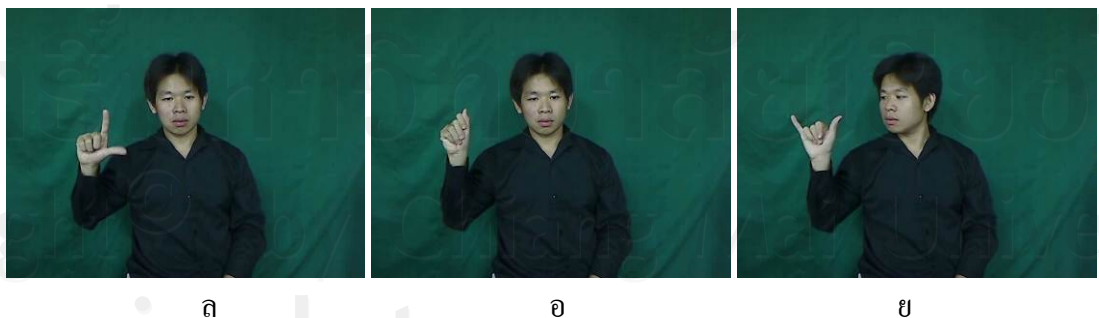
ในการสื่อสารกับผู้ที่มีความบกพร่องทางการได้ยิน และระหว่างผู้ที่มีความบกพร่องทางการได้ยินเองนั้น จำเป็นจะต้องมีการศึกษาเรียนรู้ลักษณะท่าทางของมือที่ได้มีการกำหนดไว้เป็นมาตรฐานภาษามือ ซึ่งจำเป็นต้องใช้ระยะเวลาในการศึกษาเรียนรู้ถึงจะสามารถใช้ภาษามือในการสื่อสารได้ ทำให้เกิดความยุ่งยากซับซ้อนในการที่จะศึกษาเรียนรู้ภาษามือ จึงเป็นปัญหาในการสื่อสารกับผู้ที่มีความบกพร่องทางการได้ยิน

ดังนั้น จึงได้มีแนวทางการพัฒนาคอมพิวเตอร์ให้สามารถจดจำท่าทางของมือ เพื่อช่วยอำนวยความสะดวกในการสื่อสารกับผู้บกพร่องทางการได้ยิน โดยคอมพิวเตอร์จะสามารถแปลท่าทางของมือที่เป็นรูปแบบของไฟล์วิดีโอออกมาเป็นคำได้ โดยปัจจุบันนี้ได้มีการค้นคว้าและพัฒนาให้คอมพิวเตอร์สามารถจดจำและแปลท่าทางของมือ[1] โดยถูกพัฒนาด้วยหลักการความฉลาดทางการคำนวณ (computational intelligence: CI)

ในการพัฒนาคอมพิวเตอร์ให้สามารถจดจำท่าทางของมือในรูปแบบสะกดนิ้วมือ เริ่มต้นจากการแบ่งเฟรมวิดีโอของคำที่ต้องการรู้ความหมายออกเป็นช่วงเพื่อนำไปหาลักษณะเฉพาะด้วยวิธีการของซิปต์ [2] แล้วนำไปหาค่าความเหมือนเปรียบเทียบกับรูปท่าสะกดนิ้วมือของตัวอักษร โดยตัวอย่างตัวอักษรที่ใช้ในการสะกดนิ้วมือถูกแสดงไว้ดังรูปที่ 1.1 และตัวอย่างของเฟรมตัวอักษรที่ได้จากการแบ่งเฟรมวิดีโอถูกแสดงไว้ดังรูปที่ 1.2



รูปที่ 1.1 ท่าสะกดนิ้วมือของ 19 ตัวอักษร



รูปที่ 1.2 รูปที่ได้จากการแบ่งเฟรมวิดีโอของคำว่า “ลอย”

ในการหาค่าความเหมือนของรูปทดสอบว่าเป็นตัวอักษรใด เวลาในการเปรียบเทียบจะขึ้นอยู่กับจำนวนรูปของท่าสะกดตัวอักษร ซึ่งข้อดีของการที่จำนวนรูปของท่าสะกดตัวอักษรมากก็จะทำให้ความถูกต้องในการหาความเหมือนของรูปทดสอบถูกต้องสูงขึ้น แต่ข้อเสียคือเวลาในการเปรียบเทียบจะเพิ่มมากขึ้นตามจำนวนรูปของท่าสะกดตัวอักษรดังรูปที่ 1.3

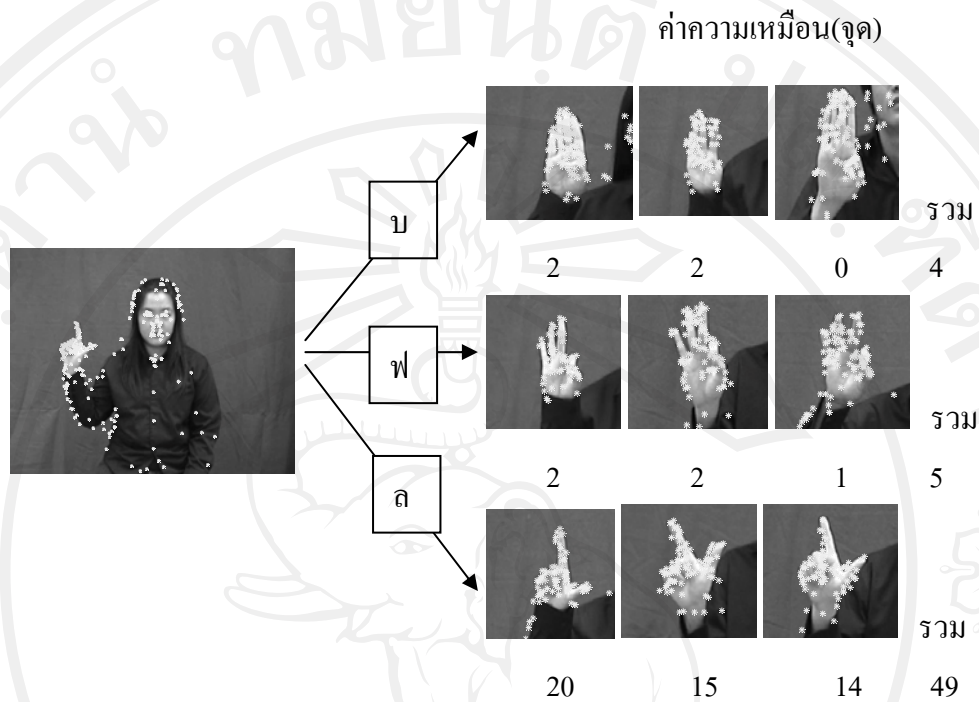
สมมติให้จำนวนรูปท่าสะกดตัวอักษรมีทั้งหมด n รูป โดยสมมติให้แต่ละรูปประกอบไปด้วยจำนวนลักษณะเฉพาะ $k_1, k_2, \dots, k_n$ จุดและสมมติให้ $k = \min_{i=1}^n k_i$ เป็นจำนวนจุดที่น้อยที่สุดจากทั้งหมด n รูป ดังนั้นค่าสัญกรณ์โอใหญ่ (BigO) จะมีค่าเท่ากับ nk ซึ่งแสดงให้เห็นว่ายิ่งจำนวนของรูปท่าสะกดตัวอักษรมีมากขึ้นเท่าไร ก็ยิ่งใช้เวลาในการประมวลผลมากขึ้นเรื่อยๆ

ดังนั้น ในงานวิจัยนี้จึงมุ่งเน้นที่จะแก้ปัญหาในด้านเวลาของการวัดความเหมือนของรูปทดสอบ โดยที่ยังคงความถูกต้องให้ใกล้เคียงหรือมากกว่าวิธีการเดิม

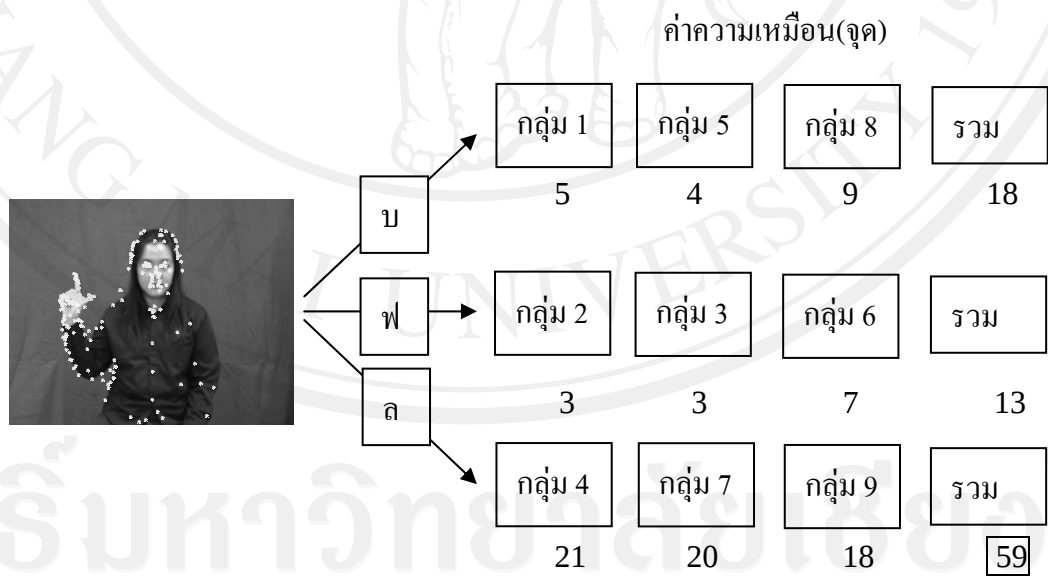
แนวทางในการแก้ปัญหา จะทำการรวมลักษณะเฉพาะของรูปท่าสะกดตัวอักษรของทุกรูปเข้าด้วยกัน แล้วทำการจัดกลุ่ม (clustering) ของลักษณะเฉพาะทั้งหมด

สมมติให้จำนวนของลักษณะเฉพาะที่ได้จากการจัดกลุ่มมีทั้งหมด c จุด ดังนั้นค่าสัญกรณ์โอใหญ่ จะมีค่าเท่ากับ c จะเห็นว่าจากวิธีการที่ได้กล่าวมาแล้ว เปรียบเทียบกับวิธีนี้ ค่าสัญกรณ์โอใหญ่มีค่าลดลงจาก nk เหลือ c นั่นแสดงถึงเวลาที่ใช้ในการประมวลผลน้อยลงจากดัชนีกำลังสองเหลือแค่ค่าของจำนวนเต็ม ซึ่งข้อดีอีกประเด็นของวิธีการจัดกลุ่มคือ ไม่ว่าจะเพิ่มจำนวนรูปท่าสะกดตัวอักษรเพิ่มขึ้นเท่าไร ค่าสัญกรณ์โอใหญ่ก็ยังคงเท่ากับ c

จากรูปที่ 1.4 แสดงวิธีการหารูปทดสอบว่าอยู่ในกลุ่มของท่าตัวอักษรใดโดยใช้วิธีการจัดกลุ่มลักษณะเฉพาะ กำหนดให้จำนวนจุดศูนย์กลางมี 9 กลุ่ม โดยสมมติให้กลุ่มที่ 1 5 และ 8 เป็นตัวแทนของตัวอักษร “บ” กลุ่มที่ 2 3 และ 6 เป็นตัวแทนของตัวอักษร “ฟ” และ กลุ่มที่ 4 7 และ 9 เป็นตัวแทนของตัวอักษร “ล” ซึ่งพบว่าที่ผลรวมของกลุ่มของตัวอักษร “ล” มีจำนวนจุดรวมมากที่สุด ดังนั้นรูปทดสอบเป็นตัวอักษร “ล”



รูปที่ 1.3 ตัวอย่างการวัดค่าความเหมือนของรูปทดสอบกับรูปทำสะกดตัวอักษร



รูปที่ 1.4 ตัวอย่างการวัดค่าความเหมือนของรูปทดสอบกับกลุ่มข้อมูลที่ได้จากการจัดกลุ่ม 9 กลุ่ม

1.2. สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง

Lowe [2] ได้ทำการคิดค้นชิฟต์ขึ้นมา ซึ่งในขณะเดียวกันได้นำเสนอวิธีการเปรียบเทียบด้วยการหาระยะห่างแบบยูคลิดระหว่างคีย์เฟรม ของรูปที่ใช้ในการทดสอบกับคีย์เฟรมของรูปที่ใช้ในการฝึกสอน ถ้าอัตราส่วนของระยะห่างแบบยูคลิดค่าน้อยที่สุด กับค่าน้อยที่สุดเป็นอันดับสอง มีค่า น้อยกว่า 0.8 จะถือว่า คีย์เฟรม ของรูปที่ใช้ในการทดสอบ นั้นเข้ากับ รูปที่ใช้ในการฝึกสอน

Lowe ได้แสดงตัวอย่างงานที่ใช้ในการเปรียบเทียบ โดยใช้โปรแกรม Matlab ซึ่งในตัวอย่าง วิธีการเปรียบเทียบใช้ วิธีหาระยะห่างแบบเชิงมุม (angle distance) ซึ่งให้ผลที่เหมือนกับ วิธีหาระยะห่างแบบยูคลิด แต่จะใช้เวลาน้อยลงเมื่อทำงานกับโปรแกรม Matlab

Batra [3] ได้เสนอวิธีการลดจำนวนของคีย์พอยต์ของรูปที่เหมือนกันให้เหลือกลุ่มเดียว โดยนำคีย์พอยต์ของหลายรูปที่ใช้ในการฝึกสอนมารวมกัน จากนั้นทำการกำจัดคีย์พอยต์ด้วยการใช้ตัวกรองความแปรปรวน (variance based filtering) เพื่อเก็บไว้ใช้เป็นฐานข้อมูลไว้เปรียบเทียบกับรูปที่ทดสอบต่อไป ซึ่งปรากฏว่าให้ผลความแม่นยำ (precision) 82% จากรูปทดสอบ 32 รูป ข้อดีของวิธีการนี้คือการรวมคีย์พอยต์ของทุกรูปที่คล้ายกัน และกำจัดคีย์พอยต์ด้วยวิธีการ (variance based filtering) ข้อเสียของวิธีการนี้คือสามารถทำได้เฉพาะรูปที่มีมุมมองในรูปแบบเดียว คือขาดคุณสมบัติของความทนทานต่อการเปลี่ยนแปลงรูปร่าง เปลี่ยนแปลงมุม เพราะรูปที่นำมาใช้ในการทดลองมีการเปลี่ยนแปลงเฉพาะวัตถุที่ต้องการเปรียบเทียบเท่านั้น

Karddouchi [4] เปลี่ยนการเปรียบเทียบจากการหาระยะห่างแบบยูคลิด มาเป็นใช้ fuzzy similarity measure โดยทำการนำเวกเตอร์ลักษณะเฉพาะของ 2 รูปที่จะทำการเปรียบเทียบ มาสร้าง fuzzy triangular number จากนั้นทำการสร้างกฎขึ้นมา แล้วทำการเปรียบเทียบค่าที่ได้จากกฎว่า เป็นรูปที่เข้าคู่กันหรือไม่ โดยผลที่ได้จากการทดลองให้ความถูกต้องเกือบทุกรูป ข้อดีของวิธีการนี้คือเป็นวิธีการที่ใหม่เนื่องจากการเริ่มต้นใช้ fuzzy similarity มาหาความเหมือนของรูป แต่ข้อเสียคือมีความผิดพลาดเมื่อรูปที่นำมาทดสอบมีคีย์พอยต์มากเกินไปหรือน้อยเกินไป ก็สามารถใช้ได้กับกลุ่มของรูปที่มีคีย์พอยต์สม่ำเสมอทุกๆ รูปเท่านั้น

Delponte [5] ใช้วิธี Singular Value Decomposition เพื่อทำการเปรียบเทียบลักษณะเฉพาะที่ได้จากชิฟต์โดยเปรียบเทียบกับวิธีหาระยะห่างแบบยูคลิดระหว่างลักษณะเฉพาะของชิฟต์ข้อดีคือจำนวนของจุดที่เข้าคู่กันระหว่างรูปที่ใช้ในการเปรียบเทียบมีจำนวนที่เยอะขึ้น แต่ทำให้เกิดข้อเสียคือจำนวนของจุดที่เข้าคู่กันผิดมีจำนวนเยอะขึ้นตาม

Zhou [6] เสนอวิธีการติดตามการเคลื่อนที่ของวัตถุ โดยพัฒนาชิฟต์ร่วมกับ mean-shift โดยในการทดลองจะทำการเปรียบเทียบผลของวิธีที่นำเสนอกับชิฟต์และ mean-shift ด้วยวิธีการวัดค่าความผิดพลาดจากการติดตามวัตถุที่อยู่ในสถานที่ที่แตกต่างกัน ผลที่ได้จากการทดลองพบว่าวิธีการ

ที่นำเสนอ ให้ผลที่ดีกว่าวิธีการซิปต์และ mean-shift แต่ข้อเสียของวิธีการทดลองนี้คือผลที่ได้ยังไม่มีความน่าเชื่อถือมากนัก เพราะรูปที่นำมาใช้ในการทดลองมีเพียงไม่กี่รูปทำให้ไม่สามารถสรุปได้ว่าวิธีการที่นำเสนอจะดีกว่าวิธีการของซิปต์หรือ mean-shift ในกรณีอื่นๆ

Wang [7] ทำการฝึกสอนหุ่นยนต์ให้สามารถจดจำคำสั่งด้วยการสั่งงานด้วยมือ โดยใช้ adaboost algorithm ร่วมกับการใช้ลักษณะเฉพาะของซิปต์วิธีการนี้เริ่มจากการตรวจหารูปที่เป็นมือ โดยที่รูปที่ใช้ฝึกสอนจะมีทั้งรูปที่เป็นมือและรูปที่ไม่มีมือ ขั้นตอนต่อไปเมื่อได้รูปที่เป็นมือจะทำการจำแนกว่ามือเป็นสัญลักษณ์ใด โดยใช้วิธีการของ Lowe จับคู่เพื่อเปรียบเทียบว่ารูปทดสอบเป็นสัญลักษณ์ใด ข้อดีของวิธีการนี้คือสามารถแยกสัญลักษณ์ของมือได้ว่าเป็นสัญลักษณ์ใดและสามารถบอกได้ว่ารูปทดสอบมีมืออยู่หรือไม่ ข้อเสียคือช้าเพราะต้องมีกระบวนการที่เพิ่มขึ้นจากวิธีการที่ Lowe ได้นำเสนอไว้

Slot and Kim [8] เสนอวิธีการวัดความเหมือนของรูปโดยการจัดกลุ่มคีย์พอยต์ของหนึ่งรูป ด้วยวิธีการสร้างฮิสโตแกรมจากบริเวณพื้นที่ที่ถูกเลือก โดยเลือกเฉพาะบริเวณที่มีคีย์พอยต์หนาแน่นเท่านั้น จากนั้นทำการจัดกลุ่มของฮิสโตแกรมที่ใกล้เคียงกันให้เป็นกลุ่มเดียวกัน แล้วทำการหาค่ามัธยฐานของแต่ละกลุ่มแล้วทำการกำจัดฮิสโตแกรมบริเวณที่อยู่นอกเหนือจากค่ามัธยฐานของแต่ละกลุ่ม วิธีการนี้จะให้ผลการจับคู่ที่ถูกต้องสูงกว่าวิธีการของ Lowe เมื่อมีจำนวนของคีย์พอยต์ที่นำมาเปรียบเทียบจำนวนเยอะ ข้อเสียของวิธีการนี้คือระดับค่าขีดแบ่งที่ใช้ในการทดลองสามารถใช้ได้เฉพาะกลุ่มของข้อมูลที่ใช้ในการทดลองนี้เท่านั้นซึ่งเมื่อต้องการนำไปใช้กับข้อมูลกลุ่มใหม่จำเป็นต้องทดลองเพื่อหาค่าระดับขีดแบ่งที่เหมาะสมใหม่

Daliri และ Torre [9] ได้เสนอวิธีการจัดกลุ่มของ texture และ shape โดยการใช้ HCM กับลักษณะเฉพาะที่ได้จากซิปต์ของรูป texture และ shape โดยจะทำการประมาณค่าจำนวนของกลุ่มที่เหมาะสม ข้อเสียของงานวิจัยนี้คือไม่ได้แสดงผลความถูกต้องของการจำแนกกลุ่ม

1.3. วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาระบบที่ใช้ตรวจจับลักษณะการเคลื่อนไหวของมือโดยใช้ซิปต์และขั้นตอนวิธีในการจัดกลุ่มข้อมูล ให้สามารถทำงานได้เร็วขึ้นและยังคงประสิทธิภาพในแง่ของความถูกต้องให้ใกล้เคียงกับระบบที่ใช้ซิปต์เพียงอย่างเดียว

1.4. ขอบเขตการทำวิจัย

- 1.4.1 เป็นระบบตรวจจับการสะกดนิ้วมือภาษาไทย โดยเกิดจากการใช้มือเดียว
- 1.4.2 ภาพที่ใช้เป็นภาพเคลื่อนไหวของมือโดยผู้ทำท่าทางมือจะใส่เสื้อแขนยาวสีเข้ม
- 1.4.3 การทำลักษณะของมือจะต้องอยู่ในลักษณะหน้าตรง

1.5. ประโยชน์ที่ได้รับจากการวิจัย

ได้ระบบแปลภาษามือภาษาไทยด้วยวิธีการสะกดนิ้วมือที่สามารถนำมาใช้ในการสื่อสารกับผู้พิการทางหูได้จริง

1.6. ขั้นตอนการทำวิจัย

1.6.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง

1.6.2 เก็บข้อมูลลักษณะการเคลื่อนไหวของมือในลักษณะท่าทางต่างๆกัน โดยแบ่งส่วนหนึ่งเป็นข้อมูลที่ใช้สำหรับการฝึกสอนและอีกส่วนเป็นข้อมูลที่ใช้สำหรับทดสอบโปรแกรม

1.6.3 ออกแบบและพัฒนาระบบสำหรับจับการเคลื่อนไหวของนิ้วมือโดยใช้ชิพด์และอัลกอริทึมในการจัดกลุ่มข้อมูล

1.6.4 ทดสอบในแต่ละวิธีการทดลอง

1.6.5 วิเคราะห์และสรุปผลการทดลอง

1.6.6 จัดทำวิทยานิพนธ์ฉบับสมบูรณ์