

การรู้จำภาษามือไทยแบบต่อเนื่องโดยใช้แบบจำลองภาษาไบนแกรม

ศิริเศรษฐ์ จิรภัทร์ชาญเดช และ นพพร โชติภักดิ์

คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: vernomcrp@gmail.com, nopporn@it.kmitl.ac.th

บทคัดย่อ

บทความฉบับนี้นำเสนอเทคนิคการรู้จำภาษามือไทยแบบต่อเนื่อง โดยนำเอาแบบจำลองภาษาไบนแกรม (Bigram) มาประยุกต์ใช้งานร่วมกับแบบจำลองฮิดเดนมาร์คอฟ (Hidden Markov Model) เพื่อเพิ่มความถูกต้องในการรู้จำ วิธีที่นำเสนอคือการนำแบบจำลองภาษาไบนแกรมมาใช้ในส่วนของการเปลี่ยนผ่านระหว่างคำ ซึ่งแบบจำลองไบนแกรมนี้จะเป็นตัวเชื่อมระหว่างสถานะ (State) สุดท้ายของคำภาษามือหนึ่งกับสถานะเริ่มต้นของอีกคำหนึ่ง โดยในกระบวนการรู้จำจะใช้อัลกอริทึมการค้นหาเส้นทางของไวเทอบี (Viterbi algorithm) ในการรู้จำ วิธีการที่เสนอได้ถูกนำไปใช้ทดลองกับท่ามือในภาษามือไทยที่ใช้เพื่อการสนทนาในชีวิตประจำวัน ผลการทดลองจะแสดงให้เห็นถึงประสิทธิภาพที่เพิ่มขึ้นของวิธีการที่นำเสนอ

คำสำคัญ— การรู้จำภาษามือ; แบบจำลองฮิดเดนมาร์คอฟ; แบบจำลองไบนแกรม; การค้นหาเส้นทางของไวเทอบี

1. บทนำ

ภาษามือนั้นเป็นรูปแบบของการแลกเปลี่ยนข่าวสารที่มีความเป็นธรรมชาติมากของบุคคลที่มีความพิการทางการได้ยิน ซึ่งประกอบด้วยท่ามือ ลักษณะการเคลื่อนไหวของมือ รวมทั้งลักษณะการเคลื่อนไหวของหน้า คิ้ว และริมฝีปาก ซึ่งการเปลี่ยนสถานะของท่ามือก็คือการเคลื่อนไหวจากท่ามือหนึ่งไปยังอีกท่ามือหนึ่งเพื่อแสดงความหมายของคำในภาษามือนั้น [1] การรู้จำภาษามืออัตโนมัติ เป็นเทคนิคด้านการรู้จำรูปแบบที่สามารถนำไปใช้ประโยชน์ได้หลายทาง เช่น เพื่อการติดต่อสื่อสารระหว่างคนพิการและคนปกติ หรือระหว่างคนพิการด้านการได้ยินกับระบบคอมพิวเตอร์ ซึ่งที่ผ่านมาได้มีการใช้เทคนิคหรือแนวคิดหลายอย่างเพื่อสร้างระบบหรือปรับปรุงให้ระบบการรู้จำภาษามือนั้นมีประสิทธิภาพมากขึ้นอย่างต่อเนื่องทั้งทางด้านการเพิ่มความถูกต้องของการรู้จำ ความเร็วของการรู้จำ การลดความซับซ้อนของระบบการรู้จำ

สำหรับในการรู้จำท่าทางภาษามือในงานวิจัยส่วนใหญ่นิยมใช้แบบจำลองฮิดเดนมาร์คอฟซึ่งคุณสมบัติส่วนใหญ่ของแบบจำลองฮิดเดนมาร์คอฟสามารถจำลองเวลาและความสูงของสัญญาณซึ่งเป็นประโยชน์มากสำหรับการรู้จำเสียง (Speech recognition) ซึ่งมีความคล้ายคลึงกับการรู้จำภาษามือ (Sign Language recognition) [7]

ในงานของ Hermann Hienz, Britta Bauer และ Karl-Friedrich Kraiss [5] ได้มีการนำเสนอการใช้แบบจำลองของไวยากรณ์ประเภทที่สร้างโดยวิธีการทางสถิติร่วมกับระบบการรู้จำภาษามือเยอรมันแบบต่อเนื่อง ซึ่งนำข้อมูลที่ได้จากภาพวิดีโอมาสร้างเป็นแบบจำลองฮิดเดนมาร์คอฟ ส่วนแบบจำลองไบนแกรมนั้นเป็นส่วนที่ใช้สำหรับกระบวนการรู้จำของระบบซึ่งจะใช้ในการคำนวณเกี่ยวกับความน่าจะเป็นในการเปลี่ยนสถานะระหว่างการรู้จำแบบต่อเนื่อง ซึ่งงานชิ้นนี้ได้มีการทดลองด้วยการใช้แบบจำลองแบบยูนิแกรมด้วย ผลที่ได้ออกมาแสดงให้เห็นว่าการใช้แบบจำลองทางภาษานั้นให้ความถูกต้องในการรู้จำได้ดีกว่าไม่มีการใช้แบบจำลองทางภาษาโดยในระบบการรู้จำภาษามือเยอรมันแบบต่อเนื่องนั้นการรู้จำที่ไม่ใช่แบบจำลองทางภาษามีความถูกต้อง 92.2% การรู้จำโดยใช้แบบจำลองยูนิแกรม 94.2% และการทดลองด้วยแบบจำลองไบนแกรมได้ความถูกต้อง 95.0% งานของ Rung-Huei Liang และ Ming Ouhyoung [3] ได้นำเสนอวิธีการแก้ปัญหาความซับซ้อนของท่ามือที่เป็นท่าเคลื่อนไหว (dynamic gesture) ในระบบการรู้จำภาษามือได้หวั่น โดยการแยกท่าภาษามือ (gesture) ออกเป็นท่ามือ (postures) ย่อยๆหลายท่าซึ่งจะสามารถรู้จำได้อย่างรวดเร็ว โดยการใช้งานแบบจำลองฮิดเดนมาร์คอฟ หลังจากนั้นจะแบบจำลองไวยากรณ์แบบไบนแกรม (grammar model) เพื่อทำการประมาณค่าความน่าจะเป็นของประโยคที่ประกอบด้วยท่าภาษามือ (gestures) แต่อย่างไรก็ตาม งานของ Rung-Huei Liang และ Ming Ouhyoung ยังไม่ได้พิจารณาในส่วนของคุณค่าของคลังศัพท์ซึ่งในการทดลองยังมีขนาดเล็กอยู่ งานของ Christian Vogler และ Dimistis Metaxas [8] นำเสนอวิธีการรู้จำแบบต่อเนื่องของภาษามืออเมริกันซึ่งใช้ข้อมูลนำเข้าเป็นข้อมูลลักษณะสามมิติของการเคลื่อนไหวของแขน โดยได้มีการสร้างแบบจำลองท่ามือเป็นแบบฮิดเดนมาร์คอฟ และสร้างแบบจำลองฮิดเดนมาร์คอฟแบบไม่ขึ้นกับเนื้อหา (context-dependent HMM model) และแบบจำลองการเคลื่อนไหวระหว่างคำ ในงานนี้ได้มีการทดลองว่าเมื่อใช้แบบจำลองไบนแกรมในส่วนของการรู้จำร่วมกับแบบจำลองทั้งสองแบบนี้จะให้ผลในการรู้จำมากกว่าไม่ได้ใช้แบบจำลองไบนแกรม

ในบทความฉบับนี้ นำเสนอแนวทางการศึกษาการใช้แบบจำลองภาษาไบนแกรมในส่วนของการรู้จำ ซึ่งใช้การคำนวณหาความน่าจะเป็นในการเปลี่ยนสถานะของท่ามือเพื่อหาเส้นทางที่มีความน่าจะเป็นมากที่สุดของการเชื่อมต่อระหว่างแบบจำลองท่ามือฮิดเดนมาร์คอฟที่สถานะสุดท้ายของแบบจำลองแรกกับสถานะแรกของแบบจำลองถัดมานั้นอยู่ติดกัน

(Boundary state) การสร้างแบบจำลองภาษาโปรแกรมสามารถสร้างได้จากข้อมูลทางสถิติของคลังศัพท์ภาษามือ[1]หรือบทสนทนาทั่วไป[2]โดยไม่เกี่ยวข้องกับข้อมูลที่น่ามาสร้างแบบจำลองฮิดเดนมาร์คอฟ

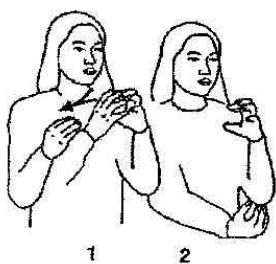
ส่วนของเนื้อหาของบทความนี้ หัวข้อที่ 2 กล่าวถึงภาพรวมของระบบการรู้จำภาษามือไทยและอุปกรณ์ที่เกี่ยวข้องในระบบการรู้จำ ในหัวข้อ 3 จะได้กล่าวถึงแนวทางในการประยุกต์ใช้งานแบบจำลองภาษาโปรแกรมในระบบการรู้จำภาษามือแบบต่อเนื่อง หัวข้อที่ 4 จะเป็นเนื้อหาที่เกี่ยวกับการทดลอง รายละเอียดการสร้างและใช้งานแบบจำลองทางภาษาแบบโปรแกรมและบทสรุปในหัวข้อที่ 5 โดยจะสรุปแนวคิดและผลของการทดลองที่ได้รวมถึงข้อเสนอแนะที่ได้จากการทดลอง

2. ภาพรวมของระบบการรู้จำภาษามือไทย

ระบบภาษามือไทยนั้นถูกพัฒนาขึ้นมาเพื่อให้การสื่อสารระหว่างคนทั่วไปกับคนที่มีความพิการทางการได้ยินสามารถทำได้อย่างมีประสิทธิภาพมากขึ้น โดยระบบการรู้จำภาษามือไทยจะถูกฝึกฝนจากข้อมูลที่ได้มาจากแหล่งต่างๆ ตัวอย่างเช่น หนังสือภาษามือ[2] วิดีโอสอนภาษามือ แล้วนำมาสร้างแบบจำลองทางสถิติในรูปแบบต่างๆ เช่น แบบจำลองทางสถิติแบบ ฮิดเดนมาร์คอฟ โดยใช้งานร่วมกับอุปกรณ์เครื่องต่างๆ เช่น ถุงมืออิเล็กทรอนิกส์ อุปกรณ์ติดตามการเคลื่อนไหว

2.1. ภาษามือไทย

ภาษามือไทยเป็นรูปแบบของภาษามือที่ใช้สื่อสารกันส่วนมากภายในประเทศไทย เป็นลักษณะของภาษาที่ได้จากการรวมกันระหว่างการเคลื่อนไหวต่างๆ ที่แสดงออกตามธรรมชาติของมนุษย์ ในภาษามือไทยนั้นแต่ละคำภาษามือจะแสดงออกด้วยท่าเคลื่อนไหว, ท่านิ่ง, การเคลื่อนไหวของมือและการแสดงออกทางสีหน้า

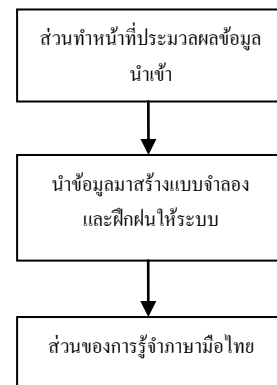


รูปที่ 1. ตัวอย่างการทำท่าภาษามือไทยของคำว่า "มันเทศ"

2.2. ระบบการรู้จำภาษามือไทยโดยใช้แบบจำลองฮิดเดนมาร์คอฟ

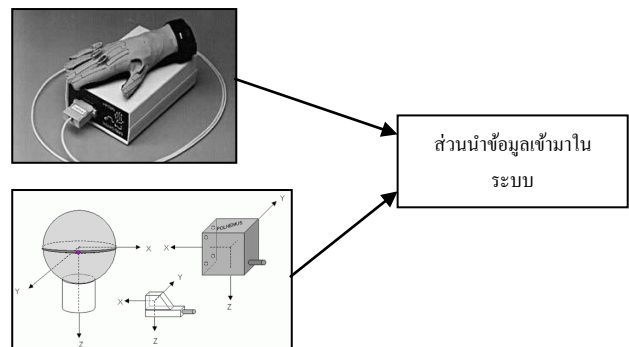
ระบบการรู้จำภาษามือไทยจะเริ่มจากการรับสัญญาณที่เป็นข้อมูลของลักษณะต่างๆ ที่ตรวจจับได้เช่น ตำแหน่งของมือ ทิศทางหรือแกนหมุนของข้อมือ ความโค้งงอของนิ้วมือและข้อนิ้วต่างๆ แล้วนำมาประมวลผลโดยใช้

วิธีของการสร้างแบบจำลองฮิดเดนมาร์คอฟ หลังจากนั้นจะทำการรู้จำโดยใช้วิธีการของไวเทอบี ซึ่งเป็นการสร้างเส้นทางที่เป็นไปได้ของสถานะของท่าภาษามือในเวลาต่างๆ ดังรูปต่อไป



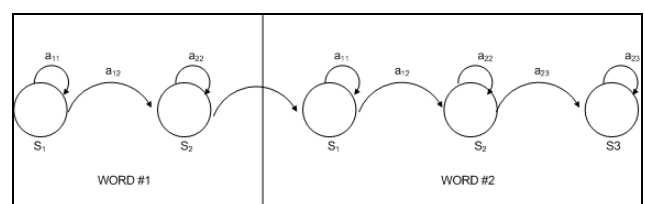
รูปที่ 2. ภาพรวมของระบบการรู้จำภาษามือไทย

ส่วนแรกเป็นส่วนที่ทำหน้าที่รับข้อมูลจากผู้ที่ทำท่ามือโดยค่าต่างๆ ที่ได้จะมาจากเซ็นเซอร์ที่ถุงมืออิเล็กทรอนิกส์ (Cyberglove) และการเคลื่อนไหว, มุมต่างๆ ของมือจะถูกตรวจจับโดยอุปกรณ์ตรวจจับตำแหน่งของมือ (Tracker)



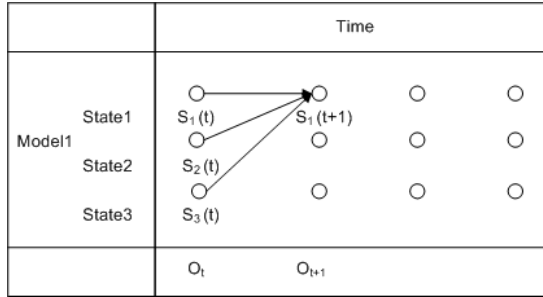
รูปที่ 3. ส่วนรับข้อมูล

ส่วนที่สองเป็นส่วนของการสอนและการสร้างแบบจำลองการเคลื่อนไหวระหว่างมือโดยนำข้อมูลที่ได้จากส่วนแรกมาผ่านขั้นตอนของการแบ่งกลุ่มของข้อมูลโดยใช้วิธีแบบเคมีน (K-Mean) จำนวนของกลุ่มที่กำหนดนั้นแบ่งตามจำนวนท่าหนึ่งที่ประกอบอยู่ในลำดับของท่ามือที่ทำให้เกิดคำภาษามือนั้น หลังจากนั้นใช้อัลกอริทึมแบบ Baum-Walsh เพื่อปรับค่าพารามิเตอร์ของแบบจำลองให้เหมาะสม



รูปที่ 4. ส่วนของแบบจำลองฮิดเดนมาร์คอฟ

ส่วนที่สามเป็นส่วนของการรู้จำโดยใช้อัลกอริทึมการค้นหาเส้นทางแบบไวเทอบี ร่วมกับการใช้แบบจำลองทางภาษาแบบไบแกรม



รูปที่ 5. ส่วนการรู้จำโดยใช้อัลกอริทึมไวเทอบี

3. การใช้แบบจำลองภาษาไบแกรมสำหรับการรู้จำภาษาไทย

ในการรู้จำภาษามือ ระบบต้องมีการสร้างแบบจำลองการเคลื่อนไหวของท่ามือ ซึ่งแบบจำลองฮิดเดนมาร์คอฟเป็นแบบจำลองที่สามารถจำลองการเคลื่อนย้ายสถานะการเคลื่อนไหวของท่ามือ แนวทางการใช้งานแบบจำลองไบแกรมจะเป็นการใช้แบบจำลองเพื่อใช้ในการคำนวณหาเส้นทางที่มีความเป็นไปได้มากที่สุดในช่วงเวลาที่ต่อเนื่องกันโดยการใช้อัลกอริทึมในการหาเส้นทางของไวเทอบี โดยกำหนดให้สัญลักษณ์ O เป็นเซตของความน่าจะเป็นของค่าปรากฏ

$$O = \{O_1, O_2, \dots, O_n\} \quad (1)$$

และให้ A เป็นความน่าจะเป็นในการเคลื่อนย้ายสถานะภายในแบบจำลองฮิดเดนมาร์คอฟ

$$A = \{a_{ij}\} \quad (2)$$

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i) \quad (3)$$

โดยที่ $1 \leq i, j \leq N$

3.1. การคำนวณความน่าจะเป็นของค่าปรากฏ

ให้ O เป็นเซตของความน่าจะเป็นของค่าปรากฏ ดังนั้นค่าความน่าจะเป็นของค่าปรากฏ ที่สถานะ j

$$b_j(k) = P(v_k(t) | q_t = S_j) \quad (4)$$

โดยที่ $1 \leq j \leq N$ และ $1 \leq k \leq M$

3.2. การคำนวณการเปลี่ยนผ่านท่ามือโดยใช้แบบจำลองไบแกรม

ในการสร้างแบบจำลองไบแกรมในงานชิ้นนี้จะใช้สมการ (5) ในการสร้างแบบจำลองเป็นหลัก โดยสร้างจากบทสนทนาในชีวิตประจำวัน

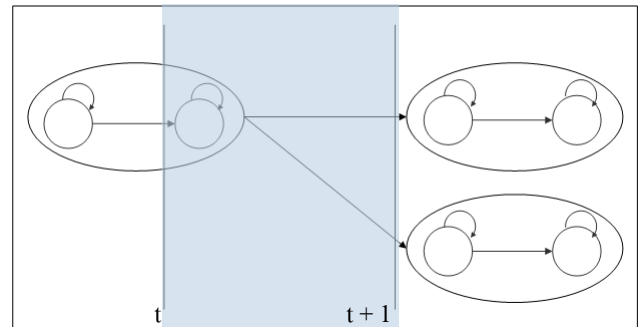
$$P(W_n | W_{n-1}) = \frac{P(W_{n-1}, W_n)}{P(W_{n-1})} \quad (5)$$

เพื่อนำมาใช้ร่วมกับอัลกอริทึมไวเทอบีซึ่งจะใช้ในการคำนวณหาเส้นทางที่มีความน่าจะเป็นมากที่สุด โดยที่ δ จะเป็นค่าความน่าจะเป็นที่มากที่สุด และ ϕ เป็นสถานะที่มีความน่าจะเป็นมากที่สุด

$$\delta_t(i) = \max_j (\delta_{t-1}(j) a_{ji} b_{i k_t}) \quad (6)$$

$$\phi_t(i) = \arg \max_j (\delta_{t-1}(j) a_{ji}) \quad (7)$$

รูปที่ 6 ส่วนของพื้นที่สีทึบแสดงถึงการใช้งานความน่าจะเป็นของแบบจำลองไบแกรมในส่วนของการคำนวณเส้นทางที่มีความเป็นไปได้ของแบบจำลองฮิดเดนมาร์คอฟในสถานะต่างๆในแต่ละช่วงเวลา



รูปที่ 6. การคำนวณเส้นทางของไวเทอบีร่วมกับการใช้งานความน่าจะเป็นของแบบจำลองไบแกรม

การคำนวณความน่าจะเป็นของเส้นทางในช่วงเวลา t ใดๆ จะเป็นการใช้ความน่าจะเป็นของแบบจำลองไบแกรมที่สถานะ j ที่เวลา t มาที่สถานะ i ใดๆที่เวลา $t+1$ ซึ่งที่สถานะ j ในเวลา t จะไม่มีค่าความน่าจะเป็นในการเคลื่อนย้ายจากสถานะสุดท้ายของแบบจำลองฮิดเดนมาร์คอฟในเวลา t ไปที่สถานะเริ่มต้นของแบบจำลองฮิดเดนมาร์คอฟอื่นๆที่มีความเป็นไปได้ทั้งหมด มีแต่ค่าความน่าจะเป็นในการเคลื่อนย้ายกลับมายังสถานะเดิมที่เวลา t ถัดไปเท่านั้น ดังนั้นจะใช้ค่าความน่าจะเป็นของการเคลื่อนย้ายสถานะกลับมาที่ตัวมันเองมาทำการเฉลี่ยกับความน่าจะเป็นในการเคลื่อนย้ายไปท่ามืออื่นๆที่ได้จากแบบจำลองไบแกรมโดยการให้ความน่าจะเป็นในอยู่ที่สถานะเดิมมีน้ำหนักเป็น $1/n$ โดยที่ n เป็นจำนวนของแบบจำลองท่ามือฮิดเดนมาร์คอฟทั้งหมดที่ใช้ในการรู้จำ หลังจากนั้นจะให้โปรแกรมทำงานในช่วงขอบเขตเวลาที่กำหนดจึงจะทำการรู้จำและ

เลือกเส้นทางจำนวนหนึ่งหลังจากนั้นจะทำการเริ่มต้นการค้นหาเส้นทางด้วยอัลกอริทึมไวเทอบีต่อไป

4. การทดลอง

ในการทดลองจะมีการใช้ถุงมืออิเล็กทรอนิกส์ (CyberGlove) และตัวติดตามตำแหน่ง (Pohelms 3 SPACE Position tracker) เป็นอุปกรณ์รับข้อมูล ซึ่งตัวติดตามตำแหน่งจะติดตั้งอยู่ที่ข้อมือของผู้ใช้ ให้ถุงมืออิเล็กทรอนิกส์ทำหน้าที่เก็บข้อมูลการเปลี่ยนแปลงตำแหน่งต่างๆของรูปทรงของมือและตัวติดตามตำแหน่งเก็บข้อมูลการเปลี่ยนแปลงต่างๆของตำแหน่งการหมุนและการเคลื่อนที่ของมือเข้ามาที่เครื่องคอมพิวเตอร์ตั้งโต๊ะจำนวนหนึ่งชุด ขณะเก็บท่ามือผู้ที่ทำการเก็บท่ามือจะอยู่ในลักษณะยืนหลังตรง โดยที่ข้อมูลที่ได้นั้นจะอยู่ในรูปของเวกเตอร์ของข้อมูล (Row vector) และทำการปรับแต่งให้เป็นเร็คคอร์ดด้วยโปรแกรมภายนอก การสร้างแบบจำลองไบแกรมจะใช้ข้อมูลจากบทสนทนาในชีวิตประจำวันจำนวน 50 บทสนทนา โดยการนับจำนวนความถี่ในแต่ละคำศัพท์ทั้งหมดของทุกบทสนทนาไม่รวมส่วนที่ไม่ได้ประกอบอยู่ในการสนทนา เช่น ชื่อของผู้สนทนา หัวข้อเรื่องของการสนทนา หลังจากนั้นใช้โปรแกรมภายนอกเพื่อคำนวณหาความถี่ของคำที่เกิดขึ้นและนำมาคำนวณหาความน่าจะเป็นแบบไบแกรมเพื่อสร้างแบบจำลองต่อไป

ข้อมูลที่ใช้ในการทดลองจะประกอบไปด้วยท่าภาษามือไทย 20 ท่ามือ แต่ละท่ามือจะมีจำนวนชุดข้อมูล 200 ชุดข้อมูล คำศัพท์ที่นำมาทดลองจะนำมาจาก หนังสือการสอนภาษามือไทยเล่มที่ 1 – 6 [2] และพจนานุกรมภาษามือไทย[1] ประโยคที่ได้จะนำมาจากแบบสนทนาที่ใช้ในสถานการณ์ทั่วไป ตัวอย่างเช่น บทสนทนาทั่วไป ท่ามือที่ได้จะมาจากผู้ที่ทำท่ามือคนเดียวกัน ข้อมูลที่ทำการทดสอบมีจำนวน 5 ชุดแต่ละชุดเป็นบทสนทนาที่ใช้ในชีวิตประจำวัน ซึ่งเกี่ยวข้องกับการแสดงอารมณ์จำนวน 3 ประโยคโดยประกอบด้วยคำทั้งหมดจำนวน 8 ถึง 10 คำ จากนั้นทำการวัดความถูกต้องของการแปลความหมายในแต่ละคำด้วยการตรวจสอบโดยผู้วิจัยซึ่งอ้างอิงจากหนังสือภาษามือไทยที่กล่าวถึงในข้างต้น และทำการทดลองซ้ำจำนวน 20 ครั้งข้อมูลทดสอบแต่ละชุด จากนั้นหาจำนวนครั้งในการแปลความหมายที่ถูกต้อง เปรียบเทียบระหว่างวิธีการที่ใช้แบบจำลองไบแกรมและไม่ใช้แบบจำลองไบแกรม ผลการทดลองแสดงในตารางที่ (1)

ตาราง 1. ผลการทดลอง

วิธีการทดสอบ	จำนวนครั้งที่แปลความหมายคำถูกต้อง	ร้อยละความถูกต้อง
ใช้แบบจำลอง	657	73
ไม่ใช้แบบจำลอง	621	69

5. ข้อเสนอแนะและสรุปผลการทดลอง

จากการทดลองพบว่าการใช้แบบจำลองไบแกรมให้ผลลัพธ์ที่ดีกว่าไม่ได้ใช้้อยู่ร้อยละ 4 เนื่องจากว่าความน่าจะเป็นที่ได้จากแบบจำลองไบแกรมมีส่วนช่วยให้ประโยคที่ได้จากการรู้จำนั้นมีความน่าจะเป็นเพิ่มมากขึ้นตามค่าสถิติจากในแบบจำลอง อย่างไรก็ตามความถูกต้องของการรู้จำที่ได้นั้น อยู่ในเกณฑ์ที่ไม่สูงมากนัก ซึ่งเราสามารถเพิ่มเติมการทดลองในอนาคตได้โดยการใช้ชุดของข้อมูลที่มีความหลากหลายและมีจำนวนมากขึ้น นอกจากนี้แล้วการปรับเปลี่ยนค่าช่วงในการเลือกเส้นทางของอัลกอริทึมไวเทอบีเป็นค่าต่างๆ ก็อาจจะมีผลต่อความถูกต้องของการรู้จำ

เอกสารอ้างอิง

- [1] สมาคมคนหูหนวกแห่งประเทศไทย. **พจนานุกรมหนังสือภาษามือไทย ฉบับปรับปรุงและขยายเพิ่มเติม** พิมพ์ที่ กรุงเทพฯ : สำนักพิมพ์ ไทยวัฒนาพานิช
- [2] สมาคมคนหูหนวกแห่งประเทศไทย. **หนังสือภาษามือไทย เล่ม 1-6** พิมพ์ครั้งที่ 1. กรุงเทพฯ : สำนักพิมพ์พิริการพิมพ์, 2544
- [3] Rung-Huei Liang and Ming Ouhyoung, "A Sign Language Recognition System Using Hidden Markov Model and Context Sensitive Search", Proceeding of ACM VRST'96, pp.59-66, July, Hong Kong, 1996
- [4] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", IEEE, 1989
- [5] Hermann Hienz, Britta Bauer, Karl-Friedrich Kraiss, "HMM-Based Continuous Sign Language Recognition Using Stochastic Grammars"
- [6] Jiyoung Ma, Wen Gao, Jiangin Wu, Chunli Wang, "A Continuous Chinese Sign Language Recognition System"
- [7] V.N. Pashaloudi and K.G. Margaritis, "Hidden Markov Models for Sign Language Recognition: a Review
- [8] Christian Vogler, Dimitris Metaxas. "Adapting Hidden Markov Models for ASL Recognition by Using Three-dimension Computer Vision Method", IEEE, 1997