

การสร้างแม่แบบภาษามลายูแบบอักษรยาวีด้วยเทคนิคแบบเทรชโฮลด์

A Template Creation Bahasa Melayu of Jawi Language Using Thresholding

ศิริเรือง พัฒน์ช่วย¹ และ มหศักดิ์ เกตุฉำ¹

Siriruang Phatchuay¹ and Mahasak Ketcham¹

¹คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

¹Faculty of Information Technology, King Mongkut's University of Technology North Bangkok

Received: May 23, 2019

Revised: July 16, 2019

Accepted: July 22, 2019

บทคัดย่อ

ภาษามลายูนิยมใช้กันมากที่สุดใน 3 จังหวัดชายแดนใต้ ภาษามลายูแบบอักษรยาวีมีความสำคัญอย่างยิ่ง เนื่องจกนิยมใช้ในการเรียนการสอนของกลุ่มโรงเรียนในระดับประถมศึกษา การติดต่อสื่อสารทางราชการ รวมถึงเอกสารและป้ายสัญลักษณ์ในสถานที่ราชการ ดังนั้นวิจัยฉบับนี้มีวัตถุประสงค์ในการการสร้างแม่แบบคำแปลจากอักษรยาวีเป็นคำไทยด้วยเทคนิคแบบเทรชโฮลด์ โดยกระบวนการเริ่มจากการนำตัวอักษรแบบเขียนด้วยลายมือมาพิมพ์เป็นคำ และสร้างเป็นไฟล์ภาพ โดยใช้ภาพ จำนวน 100 ภาพ ซึ่งในงานวิจัยได้นำเทคนิคการประมวลผลภาพมาใช้ ได้แก่ การปรับปริภูมิสี การหาค่าเทรชโฮลด์ การหาทิศทางของขอบภาพ และการตัดภาพเฉพาะส่วนที่เป็นข้อความ โดยแม่แบบที่ได้จากการสร้างจะเก็บไว้ใช้เป็นฐานข้อมูลสำหรับกระบวนการแปลภาษาต่อไป โดยผลจากการสร้างแม่แบบอยู่ในระดับที่สามารถนำไปใช้ต่อในขั้นตอนต่อไปของงานวิจัยได้อย่างถูกต้อง

คำสำคัญ: ภาษามลายู, อักษรยาวี, การสร้างแม่แบบ

Abstract

The Bahasa Melayu language is a widely used language of 3 southern border provinces in Thailand. Jawi is an important alphabet of Bahasa Melayu. Handwritten letters are generally used in these areas for studying in elementary school, official communication, documents and use in signs and in government places. Therefore, the purposes of this research were to develop a template

matching translation process for Jawi alphabet into Thai language with thresholding techniques. In the initial process, handwritten letters were transformed to typewritten words and creating an image file using 100 image datasets. This research used image processing techniques including color spaces adjustment, thresholding values, gradient direction and Image segmentation. Thus, in the Create Template Matching was stored in a database for translation processes method/ The results provided in this this study can be continued in the next step of research correctly.

Keywords: Bahasa Melayu, Jawi, template creation



บทนำ

ปัจจุบันการติดต่อสื่อสารของมนุษย์มีความใกล้เคียงกันมากขึ้น ทำให้รู้จักผู้คนที่ใช้ภาษาสำหรับการสื่อสารกันที่หลากหลายภาษา โดยในแต่ละประเทศ แต่ละท้องถิ่น ก็จะมีภาษาถิ่น เป็นของตัวเองเพื่อใช้สำหรับสื่อสารกันโดยรูปแบบภาษาก็จะเหมาะสมกับรูปแบบของการใช้งาน อีกอย่างหนึ่งสำหรับในประเทศไทยเอง ก็มีภาษาที่ใช้สื่อสารกันหลายภาษาเช่นกัน ลักษณะการใช้ภาษาก็ขึ้นอยู่กับแต่ละภูมิภาค อาทิ เช่น ภาคกลางใช้ภาษากลาง ภาคเหนือใช้ภาษาคำเมืองภาคตะวันออกเฉียงเหนือใช้ภาษาลาวอีสาน และภาคใต้ใช้ภาษาไทยถิ่นใต้ ในงานวิจัยมุ่งเน้นศึกษาทางภาษาภาคใต้ โดยลักษณะการใช้ภาษาแบ่งออกเป็นช่วงจังหวัด เช่น ภาษาไทยถิ่นใต้ตะวันออก (สำเนียง นครศรีธรรมราช) บริเวณจังหวัดนครศรีธรรมราช, สุราษฎร์ธานี และภาษาไทยถิ่นใต้ตะวันตก บริเวณพื้นที่จังหวัดกระบี่, พังงา, ภูเก็ต, ระนอง, สุราษฎร์ธานี, และชุมพร ภาษาไทยถิ่นใต้สำเนียงสงขลา บริเวณพื้นที่จังหวัดสงขลา บางส่วนของจังหวัดพัทลุง, จังหวัดปัตตานี และจังหวัดยะลา ภาษาไทยถิ่นใต้สำเนียงตากใบ หรือ ภาษาไทยถิ่นใต้สำเนียงเจ๊ะเห บริเวณพื้นที่จังหวัดนราธิวาส จังหวัดปัตตานี ซึ่งเป็นพื้นที่ 3 จังหวัดชายแดนภาคใต้ที่ใช้ภาษาไทยถิ่นใต้สำเนียงเจ๊ะเห หรือเรียกอีกอย่างว่า ภาษายาวี สำหรับสื่อสาร ทำให้สนใจการใช้ภาษายาวี โดยได้ศึกษาแล้วทำให้รู้ว่า การเขียนภาษามลายู (Abdullah, 2008) ในท้องถิ่นในอดีตนั้นใช้อักขรยาวี ก่อนที่จะเปลี่ยนไปใช้อักขรรูมี ในปัจจุบัน ภาษามลายูเป็นภาษาที่มีบทบาทสำคัญสำหรับ

ประชาชนในภูมิภาคแหลมมลายูและในจังหวัดชายแดนภาคใต้ของไทย และเป็นภาษาที่ถูกนำมาใช้เป็นเครื่องมือในการสื่อสารในการดำเนินชีวิต โดยเฉพาะอย่างยิ่งภาษามลายู อักขรยาวี และการนำแอปพลิเคชันมาช่วยแปลภาษายาวี ยิงมีน้อย ซึ่งใน (Parhizkar, Oteng, Ndaba, Lashkari & Gebiril, 2013) ได้ทำการพัฒนาแอปพลิเคชันสำหรับการแปลภาษามลายูผ่านมือถือแบบเรียลไทม์ แต่ปัญหายังมีเนื่องจากหนังสือแต่ละชนิดแต่ละเล่มมีความแตกต่างกัน มีทั้งการพิมพ์ตัวอักษรด้วยสี ขาวดำ และบางเล่มยังมีรอยเปื้อน ทำให้เกิดปัญหาระหว่างการแปลได้ บางครั้งความหมายอาจผิดเพี้ยนไป ปัจจุบันการใช้ภาษามลายู อักขรยาวี มีความจำเป็นต่อทางระบบราชการเป็นอย่างมาก เนื่องจากเมื่อมีอาเซียนเข้ามาทำให้ภาษามลายู อักขรยาวี มีบทบาทต่อการทำงาน และติดต่อสื่อสารเป็นอย่างมาก และทางหน่วยงานการศึกษามีการจัดทำสื่อวัตกรรมการเรียนการสอนภาษามลายูเพื่อการสื่อสาร ตลอดจนสถาบันการศึกษา ระดับอุดมศึกษา ยังมีการวางแผนยุทธศาสตร์ ระยะ 5 ปี (2560-2564) เพื่อกำหนดแนวทางการพัฒนาการสอน และจัดทำเอกสารประกอบการสอนภาษามลายูให้เป็นไปในทิศทางเดียวกัน เพราะลักษณะของการแปลภาษามลายู อักขรยาวี เป็นภาษารูมี หรือ ภาษารูมี เป็นภาษายาวี ก่อนถึงจะได้ภาษาไทยออกมา อีกทั้งปัญหาของชนิดตัวอักษรที่ใช้ในงานในปัจจุบัน ตลอดจนความไม่เข้าใจระหว่างตัวอักษรของ 2 ภาษาที่ยังไม่สามารถเข้าใจกันได้

จากเหตุผลข้างต้น ผู้วิจัยจึงศึกษาและสร้างแม่แบบภาษามลายูแบบอักษรยาวีด้วยการนำเทคนิคการประมวลผลภาพมาใช้ ได้แก่ การปรับปริภูมิสี การหาค่าเทรชโฮลด์ การหาทิศทางของขอบภาพ และการตัดภาพเฉพาะส่วนที่เป็นข้อความ เพื่อใช้เป็นฐานข้อมูลสำหรับกระบวนการแปลภาษามลายูแบบอักษรยาวีเป็นภาษาไทย

วัตถุประสงค์การวิจัย

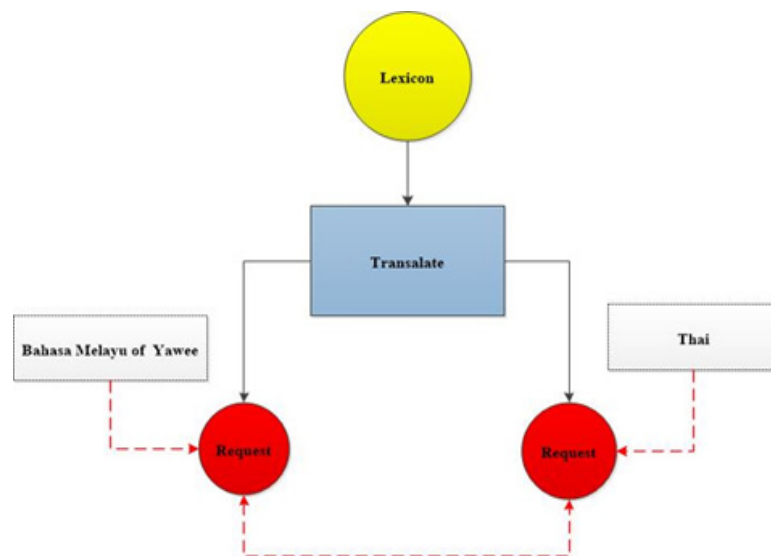
1. เพื่อสร้างต้นแบบของแม่แบบภาษาสำหรับใช้ในการแปลงภาษามลายูแบบอักษรยาวีเป็นภาษาไทย แนวคิดทฤษฎีที่เกี่ยวข้อง

ผู้วิจัยได้นำแนวคิดวิธีและกระบวนการสร้างแม่แบบสำหรับใช้ในการเปรียบเทียบแม่แบบ โดยได้แนวคิดจากวิจัยของ (Barni et al., 2010) และ(Rohlfing et al., 2012) สำหรับการสร้างแม่แบบของภาพ และจัดกลุ่มภาพ ได้นำหลักการวิจัยของ (Lin, Davis, Doermann & DeMenthon, 2007) โดยใช้วิธีการแบบเบย์มาช่วยในการตรวจจับและแบ่งส่วนของมนุษย์ โดยมีวิธีการสร้าง 3 ขั้นตอน คือ การจับคู่แบบโครงสร้างต้นไม้ และการดูความน่าจะเป็นของแม่แบบและขั้นตอนสุดท้ายคือ การจับคู่แบบลำดับขั้นเพื่อสร้างชุดแม่แบบไว้สำหรับทดสอบ โดยวิธีการนี้ สามารถตรวจจับรูปร่างของมนุษย์ได้พร้อมกันหลายคน อีกทั้งสามารถตรวจจับในกรณีที่มีผู้คนพลุกพล่านได้ด้วย จุดเด่นของวิธีการนี้ คือ ภาพที่ตรวจจับสามารถทำได้ทั้งที่มีการลบพื้นหลังและไม่มี การลบพื้นหลัง

เทคนิคที่นำมาใช้ในงานวิจัยนี้คือ เทคนิคแบบเทรชโฮลด์ เพราะเป็นขั้นตอนของการเตรียมภาพโดยสร้างภาพไบนารีจากภาพระดับสีเทาโดยเปลี่ยนพิกเซลที่มีค่าตั้งแต่ 0 ถึง 255 ให้เป็นภาพสองระดับเพื่อแบ่งพิกเซลของภาพเป็นสองกลุ่มคือกลุ่มสีขาวมีค่าเป็น 1 และกลุ่มสีดำมีค่าเป็น 0 ซึ่งได้แนวคิดจากงานของ (Raju & Neelima, 2012) ในการแบ่งส่วนภาพ อีกทั้ง (Sujji, Lakshmi & Jiji, 2013) ได้กล่าวถึงวิธีการใช้งานของ Global Thresholding

กรอบแนวคิดการวิจัย

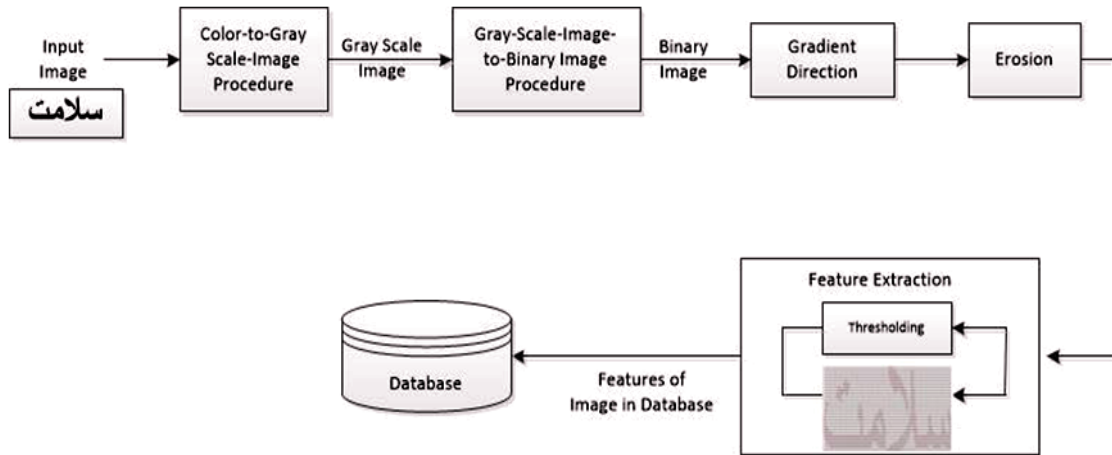
การสร้างแม่แบบภาษามลายูอักษรยาวีนั้น มีกรอบแนวคิดขึ้นเพื่อให้ 2 ภาษาสามารถทำความเข้าใจลักษณะของรูปแบบและข้อความของแต่ละภาษาได้อย่างถูกต้อง โดยแม่แบบที่สร้างขึ้นจะดำเนินการสร้างตามกระบวนการของการประมวลผลภาพด้วยเทคนิคแบบเทรชโฮลด์ โดยแม่แบบที่สร้างทั้งหมดจะถูกจัดเก็บไว้ที่พจนานุกรม (lexicon) เมื่อกระบวนการต้องการแปลข้อความ ก็เข้าสู่การร้องขอของแต่ละภาษา 2 ภาษา คือ ภาษามลายูอักษรยาวี และ ภาษาไทยสามารถทำความเข้าใจความหมายของคำได้อย่างถูกต้อง ดังภาพ 1



ภาพ 1 กรอบแนวคิดของของงานวิจัย

วิธีดำเนินการวิจัย

การสร้างแม่แบบภาษามลายูอักษรยาวี ไว้สำหรับ การแปลงข้อความเป็นภาษาไทย ประกอบด้วยขั้นตอน ดังนี้



ภาพ 2 ขั้นตอนการสร้างฐานข้อมูลแม่แบบของภาษามลายูอักษรยาวี

จากภาพ 2 สำหรับขั้นตอนการสร้างฐานข้อมูลแม่แบบของภาษามลายูอักษรยาวี จำนวน 100 คำ ผู้วิจัย

นำคำที่เขียนด้วยลายมือจากผู้ที่มีความรู้ด้านการเขียนและการใช้อักษรยาวี ในหน่วยงานราชการ แสดงดังภาพ 3

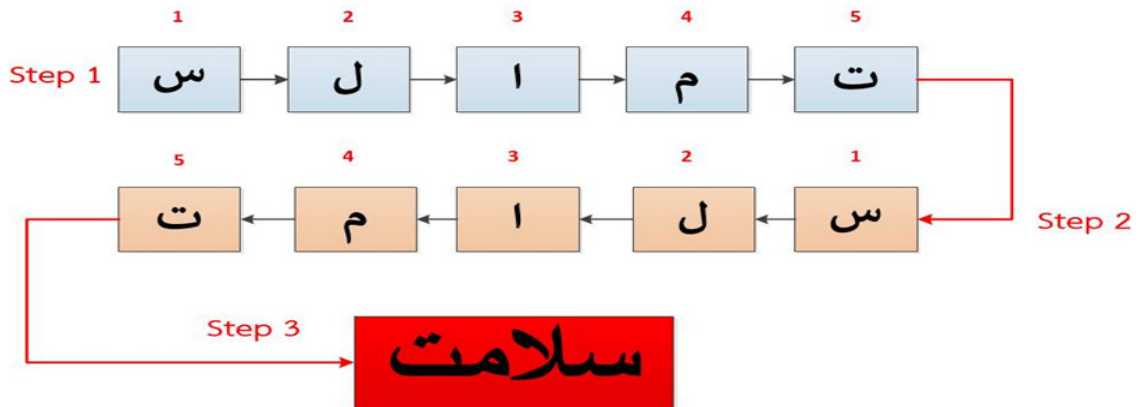
หนังสือ		๑๕ ๑๕ $\langle \text{بولو} \rangle$
สมุด	✓	๑๕ ๑๕ $\langle \text{بوكو} \rangle$
ปากกา	✓	๑๕ ๑๕ $\langle \text{قلم} \rangle$
ยางลบ	✓	๑๕ ๑๕ $\langle \text{محو} \rangle$
ไม้บรรทัด	✓	๑๕ ๑๕ $\langle \text{مسطرة} \rangle$
ดินสอ	✓	๑๕ ๑๕ $\langle \text{قلم} \rangle$
เครื่องคิดเลข	✓	๑๕ ๑๕ $\langle \text{حاسبة} \rangle$
ป้าย	✓	๑๕ ๑๕ $\langle \text{علامة} \rangle$

ภาพ 3 ตัวอย่างคำที่เขียนด้วยลายมือ

เมื่อได้ข้อมูลตัวอย่างคำแล้วกระบวนการต่อไปคือ ขั้นตอนการสร้างกลุ่มคำจากตัวพิมพ์ โดยใช้ชนิดตัวอักษรแบบ Arial ขนาด 48 สามารถ

แสดงขั้นตอนวิธีการดำเนินงานวิจัยได้ ดังนี้

1. ขั้นตอนการสร้างกลุ่มคำตัวอย่างจากการพิมพ์ไว้ในฐานข้อมูล มีขั้นตอนดังภาพ 4 ดังนี้



ภาพ 4 ขั้นตอนการสร้างกลุ่มคำตัวอย่าง

ขั้นตอนที่ 1 คือ การพิมพ์ตัวอักษรในชลัษณะคำ

ขั้นตอนที่ 2 ตัวอักษรที่ถูกพิมพ์เข้าไปจะเรียงจากขวาไปซ้าย

ขั้นตอนที่ 3 ลักษณะคำที่พิมพ์ตามขั้นตอนที่ 1-2 จะติดกันเป็นลักษณะของภาษามลายูอักษรยาวี

นำกลุ่มคำที่สร้างจากคำพิมพ์มาแปลความหมาย และจัดทำเป็นฐานข้อมูล ซึ่งความหมายของคำหรือข้อความแสดงดังตาราง 1

ตาราง 1

ความหมายของคำภาษามลายูแบบอักษรยาวี

อันดับที่	รหัสภาพ	ภาพประกอบ	คำอ่านยาวี	คำแปลเป็นไทย
1	1-2018	نام	นา-มอ	ชื่อ
2	2-2018	سلامت	สือ-ลา-มัด	สวัสดี
3	3-2018	جومف	จุม-ปะ	เจอ
4	4-2018	قيغضل	เมอ-นิง-กัล	จาก
5	5-2018	سلامت تيضكل	สือ-ลา-มัด-ติง-กัล	ลาก่อน

ตาราง 1 (ต่อ)

อันดับที่	รหัสภาพ	ภาพประกอบ	คำอ่านยาวี	คำแปลเป็นไทย
6	6-2018	فرغي	ปี-ร-กี	ไป
7	7-2018	جو تي	จู-ตี	ลา
8	8-2018	سمبه	เสิม-บอฮ	ไหว้
9	9-2018	اكت تاغف	อา-กัต-ตา-จัน	พนมมือ
10	10-2018	بيمبغ	บิม-บ้ง	ห้วง

2. นำคำที่ได้แปลงเป็นลักษณะรูปภาพเพื่อนำเข้าสู่กระบวนการ Image processing โดยมีขั้นตอน ดังนี้

2.1 Input Stage เป็นขั้นตอนแรกของการสร้างแม่แบบภาษามลายูอักษรยาวี ซึ่งเป็นกรนำตัวอักษรมาลายูที่เก็บเป็นไฟล์ภาพไว้ในฐานข้อมูล โดยนำไฟล์ภาพเข้าสู่ระบบโดยผ่านอุปกรณ์เพื่อเป็นการแปลงภาพให้เป็นภาพดิจิทัล

2.2 Pre-Processing Stage สำหรับกระบวนการ Pre-Processing Stage นำวิธีการประมวลผลภาพ (image processing) มาใช้ปรับไฟล์ภาพให้อยู่ในลักษณะของภาพที่เหมาะสมกับการนำไปใช้งาน ซึ่งมีรายละเอียด ดังนี้

2.2.1 Standard size กำหนดขนาดมาตรฐานของภาพต้นฉบับทุกภาพโดยอิงพุดที่ w=108 และ h=55 พิกเซล และกำหนดขอบเขตของ Region of Interest-ROI ที่ w=100-100



ภาพ 5 ตัวอย่างการกำหนด ROI

2.2.2 Binary Transform กระบวนการเตรียมภาพ เป็นกระบวนการที่มีจุดมุ่งหมายเพื่อกำจัดส่วนที่ไม่ต้องการออกไปและเน้นเฉพาะส่วนที่ต้องการเท่านั้น โดยเริ่มต้นจากการนำเข้าภาพต้นฉบับ w=108 และ h=55 พิกเซล นามสกุล .jpg แล้วทำการแปลงภาพเป็นสีเทาและเป็นภาพสองระดับตามลำดับ ทำการกำจัดสัญญาณรบกวน ซึ่งมีขั้นตอนดังต่อไปนี้

(1) การแปลงภาพเป็นสีเทา โดยการแยกระดับสีแต่ละพิกเซลออกจากกันในรูปแบบสี RGB เข้าสู่สมการเพื่อคำนวณหาค่าสีเทาและนำค่าที่ได้ไปแทนจุดพิกเซลเดิมโดยหาได้จากสมการ(1) ดังนี้

$$G' = 0.3 R + 0.59G + 0.11 B \quad (1)$$

โดยกำหนดให้

- G' คือ ค่าระดับสีเทา
- R คือ ค่าระดับสีแดง
- G คือ ค่าระดับสีเขียว
- B คือ ค่าระดับสีน้ำเงิน

โดยเมื่อทำการแปลงภาพภาษามลายูอักษรยาวีให้เป็นสีเทาจะได้ดังภาพ 5



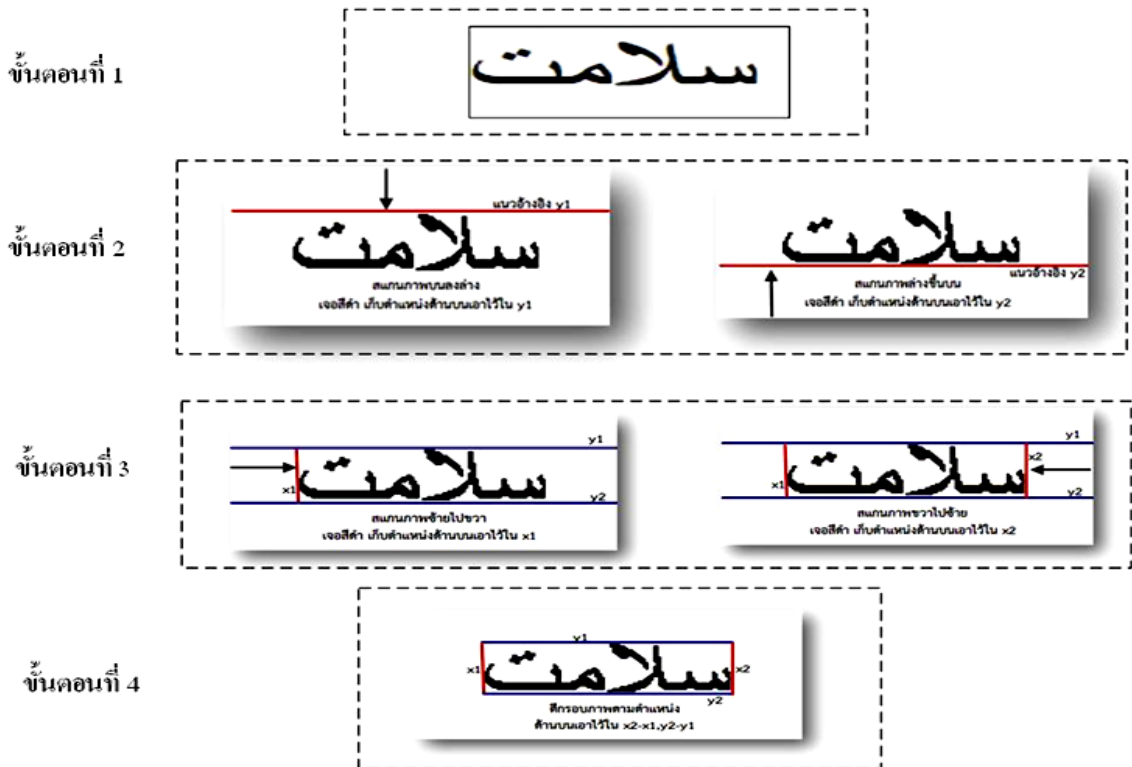
ภาพ 6 การแปลงภาพเป็นสีเทา

(2) การหาทิศทางของขอบภาพ (gradient direction) โดยหาได้จากสมการ (2) ดังนี้

$$GD(x,y) = \tan^{-1} \left\{ \frac{Y_{diff}(x,y)}{X_{diff}(x,y)} \right\} \quad (2)$$

โดย $GD(x,y)$ ทิศทางของขอบภาพที่ตำแหน่ง (x,y) โดยการหาทิศทางขอบภาพ ประกอบด้วย 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 คือ การรับไฟล์ภาพ เข้าสู่ระบบ
 ขั้นตอนที่ 2 เพื่อให้ระบบทำการตีกรอบ (bounding box) โดยมีการทำงานแทนค่าตัวแปร คือ x_1, x_2, y_1, y_2 โดยระบบตรวจสอบพิกเซลว่าเป็นสีดำหรือไม่ ถ้าใช่ให้เก็บตำแหน่งด้านบนไว้ข้างอิงในตัวแปร y_1 และด้านล่างข้างอิงในตัวแปร y_2
 ขั้นตอนที่ 3 ระบบตรวจสอบพิกเซลว่าเป็นสีดำหรือไม่ ถ้าใช่ให้เก็บตำแหน่งด้านซ้ายไว้ข้างอิงในตัวแปร x_1 และด้านขวาข้างอิงในตัวแปร x_2 ซึ่งระบบจะสแกนภาพจากบนลงล่าง โดยเริ่มจากจุดข้างอิงจาก y_1 สแกนถึงระยะ y_2 เท่านั้น
 ขั้นตอนที่ 4 ระบบตีกรอบภาพในส่วนเฉพาะ x_2-x_1, y_2-y_1 โดยกระบวนการทำงานสามารถแสดงได้ดังภาพ 7



ภาพ 7 กระบวนการหาทิศทางของขอบภาพ

(3) การตัดข้อความจะทำการตัดออกมาเฉพาะข้อความที่เป็นลักษณะสี่เหลี่ยม ตามความกว้าง* ความยาวของทิศทางของขอบภาพ ด้วยวิธีการกร่อนภาพ (erosion) คือ การตัดภาพ ขยายภาพให้เล็กลง โดยหาได้จากสมการ (3) ดังนี้

$$A \ominus B^t = \{p | B_p CA\}; B \text{ is denoted } A \ominus B^t \quad (3)$$

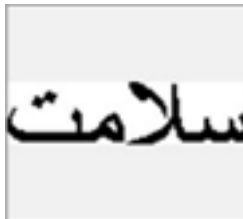
โดยที่

A คือ ภาพที่ต้องการขยายภาพ (dilation)

B^t คือ ส่วนกลับ (transposed) ของ B

B_p คือ ค่ากลาง ณ จุด p

โดยเมื่อทำการทำการตัดออกมาเฉพาะข้อความภาษามลายูอักขรยารวีจะได้ดังภาพ 8



ภาพ 8 การตัดข้อความ

(4) การทำเทรสโฮลด์ เป็นวิธีการแปลงภาพเป็นสองระดับ โดยการแปลงจากภาพสีเทาที่มีค่าตั้งแต่ 0 ถึง 255 ให้เป็นภาพสองระดับเพื่อแบ่งพิกเซลของภาพเป็นสองกลุ่มคือกลุ่มสีขาวมีค่าเป็น 1 และกลุ่มสีดำมีค่าเป็น 0 โดยหาได้จากสมการ (4) ดังนี้

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (4)$$

โดยกำหนดให้

$g(x, y)$ คือ ค่าพิกเซลในแต่ละตำแหน่ง (x, y) ของภาพใหม่

$f(x, y)$ คือ ค่าพิกเซลในแต่ละตำแหน่ง (x, y) ของภาพเดิม

ค่าพิกเซล เท่ากับ 1 แทนพื้นหลังสีขาว

ค่าพิกเซล เท่ากับ 0 แทนตัวอักษรสีดำ

โดยงานวิจัยดำเนินการสร้างแม่แบบด้วยหลักการทำเทรสโฮลด์ด้วยการแปลงเป็นภาพสองระดับ คือ การตั้งกฎที่จะแบ่งเป็นจุดดำและจุดขาวแบบตายตัว โดยมีระบบจะทำการสร้างแม่แบบประกอบด้วย 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 อ่านไฟล์ภาพคำทั้งหมด จากในโฟลเดอร์ภาพที่ได้จากการสร้างคำจากการพิมพ์ที่กำหนดไว้

ขั้นตอนที่ 2 ตัดภาพให้เหลือเฉพาะส่วนของคำที่จะใช้

ขั้นตอนที่ 3 อ่านค่าพิกเซลของภาพแล้วแปลงภาพให้เป็นรูปแบบ 0,1 โดยมีเงื่อนไข คือ ถ้าค่าเฉลี่ย RGB (แม่สีของภาพ) < 200 ให้เป็น 0 (ส่วนของตัวอักษร) และถ้าค่าเฉลี่ย RGB (แม่สีของภาพ) > = 200 ให้เป็น 1 (ส่วนของพื้นหลัง)

ขั้นตอนที่ 4 สร้างไฟล์ที่เก็บข้อมูลแม่แบบไว้ในฐานข้อมูลโดยเมื่อทำเทรสโฮลด์ จะได้ดังภาพ 9



ภาพ 9 การทำเทรสโฮลด์

ผลการวิจัย

จากผลการวิจัยเพื่อสร้างต้นแบบภาษามลายูอักขรยารวี โดยงานวิจัยทำการสร้างตัวอย่างต้นแบบสำหรับใช้ในการวิจัยทั้งสิ้น 100 คำ เพื่อนำจัดเก็บไว้ในพจนานุกรมฐานข้อมูล

ตาราง 2

ตัวอย่างคำต้นแบบ

ลำดับที่	ไทย	มลายูอักษรยาวี
1	ชื่อ	نام
2	สวัสดิ์	سلامت
3	เจอ	جومف
4	พบ	جومف
5	จาก	قيغضل
6	ลาก่อน	سلامت. تيضكل
7	ไป	فرغي
8	ลา	جوتي
9	ไหว้	سمبه
10	พนมมือ	اكت. تاغف
11	คิดถึง	ايشط
12	ห้วง	بيمبغ
13	หึ่ง	جمورو
14	เป็น	صيدوف
15	ฉัน	ساي
16	เขา	ديا
17	คุณ	افدا
18	เธอ	ديا
19	สามี	سوامي
20	รัก	كاسيه
21	มา	ماري
22	ภรรยา	استري
23	พ่อ	باف
24	แม่	اييو
25	พี่ชาย	اباغ

ตาราง 2 (ต่อ)

ลำดับที่	ไทย	มลายูอักษรยาวี
26	ยาย	نڤيڤيت
27	หลาน	جو جو
28	เพื่อน	كاون
29	ตำรวจ	فوليس
30	ทหาร	سكاو
31	แม่ชี	نني
32	บวช	ننتايبسد
33	สมุด	بوکو
34	ยางลบ	فبافو
35	เดิน	جالن
36	นั่ง	دودوت
37	ป่วย	نوستو
38	กระโดด	لومفت
39	ร้อง	ฟาغيส
40	ร้องไห้	مناغيส
41	ยิ้ม	سبوم
42	เศร้า	سدیه
43	ดีใจ	سوكاهاتي
44	เสียใจ	سويه
45	ความสุข	كمبوا
46	ความทุกข์	كسوسهن
47	มี	ادا
48	รัก	كاسيه
49	ความรัก	كاسيهن
50	โลก	كتامبکن

จากคำตัวอย่างภาษามลายูอักษรยาวีดังตาราง 1 เมื่อนำไปเข้าสู่ขั้นตอนกระบวนการของวิธีดำเนินการวิจัย จะได้ผลการทดลองเป็นไฟล์ภาพแม่แบบเพื่อเก็บไว้ใน

พจนานุกรม โดยนำเสนอทั้งหมด 20 ภาพตามกระบวนการการทำเทรสโฮลด์ ดังตาราง 2

ตาราง 3

ผลการสร้างแม่แบบด้วยเทคนิคเทรสโฮลด์

ไฟล์ภาพภาษามลายูอักษรยาวี	การสร้างแม่แบบด้วยเทคนิคเทรสโฮลด์	จำนวนพิกเซล	คำภาษาไทย
		1435	ชื่อ
		2496	สวัสดี
		3036	เจอ
		3648	จาก
		7752	ลาก่อน
		3200	ไป
		2926	ลา

ตาราง 3 (ต่อ)

ไฟล์ภาพภาษามลายูอักษรยาวี	การสร้างแม่แบบด้วยเทคนิคเทอร์สโสด์	จำนวนพิกเซล	คำภาษาไทย
		1943	ไหว้
		3718	พนมมือ
		2479	ห้วง
		2442	คิดถึง
		2231	หึ่ง
		3729	เป็น
		2419	ฉั่น
		1190	เขา

ตาราง 3 (ต่อ)

ไฟล์ภาพภาษามลายูอักษรยาวี	การสร้างแม่แบบด้วยเทคนิคเทรลโฮลด์	จำนวนพิกเซล	คำภาษาไทย
		1000	คุณ
		1190	เธอ
		3720	สามี
		2660	รัก
		2960	มา

การอภิปรายผล

ผู้วิจัยได้ใช้ตัวอย่างภาพภาษามลายูแบบอักษรยาวีจำนวน 100 ภาพ เพื่อเป็นภาพตัวอย่างสำหรับสร้างต้นแบบภาษามลายูแบบอักษรยาวี และทำการสร้างภาพต้นแบบเป็นไฟล์ ไฟล์ภาพ RGB นามสกุล .jpg แล้วทำการประมวลผลภาพเบื้องต้นด้วยการแปลงภาพที่ได้ให้เป็นภาพระดับเทา และทำการหาทิศทางของขอบภาพ (gradient direction) ด้วยการตีกรอบข้อความ และนำภาพที่ได้ไปแปลงข้อมูลแบบไบนารี และทำเทรลโฮลด์ โดยการเปลี่ยนค่าของจุดภาพให้

มีเพียงสองค่า (binary) คือ ขาวและดำ (0,1) เพื่อเปรียบเทียบแม่แบบ ซึ่งสอดคล้องกับงานวิจัยของ (Vala & Baxi, 2013) และ (Kaur & Kaur, 2011) เพื่อทำการรู้จำภาษามลายูแบบอักษรยาวี โดยระบบสามารถสร้างต้นแบบภาษาสำหรับการแปรภาษามลายูแบบอักษรยาวีเป็นภาษาไทยอยู่ในระดับที่น่าพอใจ ซึ่งจากต้นแบบของแม่แบบที่ได้สร้างเก็บไว้ในฐานข้อมูลจากงานวิจัยจะนำไปใช้สำหรับกระบวนการแปลในงานวิจัยต่อไป

ข้อเสนอแนะ

ค่าที่นำมาสร้างภาพเพื่อใช้ในสร้างแม่แบบมีข้อจำกัด เพราะลักษณะของตัวอักษรที่เขียนนั้นมีขนาดไม่เท่ากัน คือ ยาวสูงด้านบน และยาวลงด้านล่าง ทำให้การหาทิศทางของ ขอบภาพ (gradient direction) ด้วยการตีกรอบข้อความ

ที่ได้ออกมาคุณภาพความละเอียดภาพแตกต่างกัน ทำให้ ค่าในส่วนของการทำเทรสโฮลด์ การรู้แม่แบบแตกต่างกัน ดังนั้นจึงควรหาแนวทางในการศึกษาเพิ่มเติมและการแก้ไข ปัญหาเพื่อให้การหาทิศทางของขอบภาพเกิดข้อผิดพลาด น้อยสุด



References

- Abdullah, A. A. (2008). Southern Thailand: Some grievances of the Patani Malays. *The Journal of International Studies*, 4(1), 102-111.
- Barni, M., Bianchi, T., Catalano, D., Di Raimondo, M., Labati, R. D., Failla, P., . . . & Piva, A. (2010). A privacy-compliant fingerprint recognition system based on homomorphic encryption and fingercode templates. *Paper presented at the 2010 fourth IEEE international conference on Biometrics: Theory, Applications and Systems--BTAS* (pp. 1-7). Washington DC, USA.: IEEE.
- Kaur, N., & Kaur, R. (2011). A review on various methods of image thresholding. *International Journal on Computer Science and Engineering*, 3(10), 3441-3443.
- Lin, Z., Davis, L. S., Doermann, D., & DeMenthon, D. (2007). Hierarchical part-template matching for human detection and segmentation. *Paper presented at the 2007 IEEE 11th international conference on Computer Vision* (pp. 1-8). Rio De Janeiro, Brazil: IEEE.
- Parhizkar, B., Oteng, K., Ndaba, O., Lashkari, A. H., & Gebril, Z. M. (2013). Ubiquitous mobile real time visual translator using augmented reality for bahasa language. *International Journal of Information and Education Technology*, 3(2), 124-128.
- Raju, P. D. R., & Neelima, G. (2012). Image segmentation by using histogram thresholding. *International Journal of Computer Science Engineering and Technology*, 2(1), 776-779.
- Rohlfing, T., Kroenke, C. D., Sullivan, E. V., Dubach, M. F., Bowden, D. M., Grant, K., & Pfefferbaum, A. (2012). The INIA19 template and NeuroMaps atlas for primate brain image parcellation and spatial normalization. *International Journal Frontiers in neuroinformatics*, 6(27), 1-15.
- Sujji, G. E., Lakshmi, Y., & Jiji, G. W. (2013). MRI brain image segmentation based on thresholding. *International Journal of Advanced Computer Research*, 3(1), 97.
- Vala, H. J., & Baxi, A. (2013). A review on Otsu image segmentation algorithm. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 2(2), 387-389.

