ขั้นตอนวิธีสำหรับการแปลภาษามลายูอักษรยาวีเป็นภาษาไทย

ศิริเรือง พัฒน์ช่วย[1,*], มหศักดิ์ เกตุฉ่ำ[2]
[1,2] คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

Received: 23 May 2019
Revised: 4 September 2019
Accepted: 6 September 2019

บทคัดย่อ

ในการประมวลผลภาพดิจิทัลนั้น กระบวนการปรับปรุงคุณภาพของภาพมีความสำคัญเป็นอย่างยิ่ง เนื่องจากถูกนำไปประยุกต์ใช้ร่วมกับเทคนิคอื่นเพื่อการประยุกต์ในโดเมนต่างๆ อาทิ การแปลภาษาด้วยภาพ งานวิจัยนี้ปรับปรุงคุณภาพของภาพโดยการสร้างแผ่นแบบจากการแปลงภาพเป็นสีเทาร่วมกับแบบรูปฐานสองเชิงขีดเริ่มเปลี่ยน ซึ่งแผ่นแบบที่ถูกสร้างขึ้นจะถูกจัดเก็บไว้ในฐานข้อมูล นอกจากนี้ งานวิจัยนี้ยังเสนอวิธีการเปรียบเทียบแผ่นแบบยูเนียนอินเตอร์เซกชัน เพื่อแปลภาษามาลายูอักษรยาวีเป็นภาษาไทย จากผลการศึกษาพบว่าความถูกต้องแม่นยำมีค่าร้อยละ 98.50

คำสำคัญ: ภาษามลายูอักษรยาวี เปรียบเทียบแผ่นแบบ ยูเนียน อินเตอร์เซกชัน

* ผู้ประสานงานหลัก; อีเมล: S5607011956137@email.kmutnb.ac.th

An Algorithm for Translation Bahasa Jawi into Thai

Siriruang Phatchuay[1,*], Mahasak Ketcham[2]
[1,2] Faculty of Information Technology,
King Mongkut's University of Technology North Bangkok

ABSTRACT

In digital image processing, the process of image enhancement is necessary as it was applied with other techniques for diverse application domains such as image-based language translation. In this research, the image enhancement was carried out by creating image pattern based on greyscale conversion and threshold-based binary method. Created image pattern was stored in a database. Image-based algorithm template matching with union intersection (TMUI) was also performed in this work for translating from Bahasa Jawi into Thai. The results showed that the accuracy was 98.50%.

Keywords: Bahasa Jawi, pattern matching, union, intersection

* Corresponding Author; Email: S5607011956137@email.kmutnb.ac.th

ความเป็นมาและความสำคัญของปัญหา

ภาษาคือสัญลักษณ์สำคัญที่ใช้ในการติดต่อสื่อสารกัน โดยแต่ละภูมิภาคของแต่ละประเทศก็จะมีภาษาถิ่นที่แตกต่างกัน แม้แต่ตัวอักษร พยัญชนะ และการเขียนของภาษาที่ใช้เป็นภาษาประจำชาติของแต่ละประเทศนั้นก็ย่อมมีคุณลักษณะพิเศษที่แตกต่างกันออกไป ดังนั้นตัวแปรสำคัญของการติดต่อสื่อสารกันเพื่อให้ผู้ใช้สามารถตีความหมายและเข้าใจตรงกันได้ คือ การแปลภาษา (Chiang, 2010) รวมถึงการวิเคราะห์ไวยากรณ์ของภาษา (Egashira, 2000) ของแต่ละภาษาเพราะประชากรที่อาศัยอยู่ในแต่ละประเทศต่างประกอบด้วยหลายเชื้อชาติ และภาษาที่ใช้ในการติดต่อสื่อสารกันก็มีหลากหลายเช่นกัน ในงานวิจัยนี้มุ่งเน้นไปที่การศึกษาภาษาทางตอนใต้ของประเทศไทย คือ ภาษามลายูที่เขียนด้วยอักษรยาวี หรือภาษามลายูปัตตานี (Matheson & Hooker, 1988) โดยใช้กันมากใน 3 จังหวัดชายแดนใต้ คือ จังหวัดยะลา นราธิวาส และปัตตานี ซึ่งเปรียบเสมือนภาษาถิ่น

สำหรับปัจจุบันภาษามลายู (Ting, 2010) มีบทบาทสำคัญในการติดต่อสื่อสาร และเมื่อประเทศไทยเข้าสู่ประชาคมอาเซียนยิ่งทำให้ภาษามลายูอักษรยาวีมีบทบาทมากยิ่งขึ้น อาทิ การใช้ติดต่อสื่อสารทางราชการ ป้ายสัญลักษณ์บนท้องถนน และนำไปใช้สำหรับการเรียนการสอนของการศึกษา ในระดับประถมศึกษา มัธยมศึกษาของชุมชนชาวมาเลย์มุสลิม (Herriman, 2005) รวมไปจนถึงระดับมหาวิทยาลัย โดยระบบการแปลภาษาหรือการประมวลผลภาษาธรรมชาติ (Natural Language Processing : NLP) (Cohn & Specia, 2013) ถูกพัฒนาขึ้นมาในลักษณะของแอพพลิเคชันเพื่อช่วยให้ชาวต่างชาติสามารถรับรู้ความหมายและสื่อสารกันได้อย่างถูกต้อง แต่เนื่องจากภาษามลายูแบบอักษรยาวียังไม่ได้พัฒนาขึ้นมาในลักษณะของกระบวนการทางภาษาและพัฒนาในรูปแบบระบบคอมพิวเตอร์อย่างเป็นทางการ

งานวิจัยนี้นำเสนออัลกอริทึมสำหรับการแปลภาษามลายูอักษรยาวีเป็นภาษาไทย โดยมีกระบวนการสำคัญลำดับแรกเป็นการปรับปรุงคุณภาพของภาพประกอบด้วยการสร้างแผ่นแบบเก็บไว้ในฐานข้อมูลโดยใช้คำสำคัญในการค้นหารูปภาพ (Zakraoui, Elloumi, Alja'am, & Yahia, 2019) โดยคำนวณจากความหนาแน่นของพิกเซลทั้งหมดของภาพ และเข้าสู่กระบวนการการแปลงภาพเป็นสีเทา และการทำภาพแบบไบนารีโดยเทคนิคการแบ่งภาพ งานวิจัยนี้เลือกการแบ่งภาพแบบเทรสโช (Al-Amri & Kalyankar, 2010) เพื่อให้ได้ภาพที่สัญญาณรบกวนน้อยที่สุดเก็บไว้ในฐานข้อมูล และเข้าสู่ขั้นตอนการเปรียบเทียบภาพกับแผ่นแบบที่สร้างไว้ ด้วยหลักการและทฤษฎีทางคณิตศาสตร์แบบยูเนียน (Union) $A \cup B$ และอินเตอร์เซกชัน (Intersection) $A \cap B$ เพื่อทำการแปลภาษาได้อย่างถูกต้องแม่นยำ โดยรายละเอียดกระบวนการและขั้นตอนการสร้างแผ่นแบบและอัลกอริทึมสำหรับแปลข้อความอธิบายไว้ในส่วนของวิธีดำเนินการวิจัยในครั้งนี้

วัตถุประสงค์ของการวิจัย

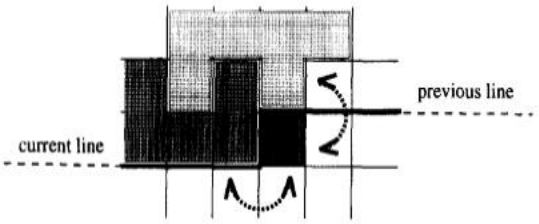
เพื่อพัฒนาอัลกอริทึมสำหรับการแปลภาษามลายูอักษรยาวีเป็นภาษาไทยเพื่อใช้เป็นต้นแบบสำหรับการเรียนการสอนเกี่ยวกับความรู้เบื้องต้นของภาษามลายูอักษรยาวีสำหรับนักเรียนระดับประถมศึกษา

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในส่วนของทฤษฎีและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยได้ศึกษาหลักการของการทำงานการเปรียบเทียบแผ่นแบบของภาพและการจับคู่ความเหมือนของแผ่นแบบ (Briechle & Hanebeck, 2001) และประยุกต์ใช้หลักการและทฤษฎีทางคณิตศาสตร์ 2 อย่าง คือ Set Union and Set Intersection เพื่อประยุกต์ใช้สำหรับงานวิจัย ดังนี้

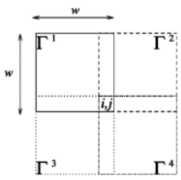
1. การเปรียบเทียบแผ่นแบบ (Pattern Matching) สามารถทำได้โดยการหาตำแหน่งของภาพที่เหมือนกับภาพแผ่นแบบหลักโดยทำการเปรียบเทียบแบบพิกเซลต่อพิกเซล โดย (Khalil, 2010) นำเสนออัลกอริทึมสำหรับการจดจำป้ายทะเบียนภาษาอาหรับแบบการจับคู่กับแผ่นแบบ แต่องค์ประกอบภายในภาพแต่ละภาพก็มีความแตกต่างกัน ซึ่ง (Ahuja & Tuli, 2013) ใช้อัลกอริทึม Normalized Cross-correlation (NCC) ในการวิเคราะห์ความเหมือนของภาพ แต่วิธีการยังไม่สามารถรู้จำองค์ประกอบของภาพแผ่นแบบได้ทั้งหมด ซึ่งทำให้ (Dawoud, Samir, & Janier, 2011) นำเสนออัลกอริทึมใหม่ที่ดีกว่าการจับคู่ภาพแบบ Normalized Cross-correlation (NCC) โดยนำไปใช้ในการตรวจจับวิเคราะห์ความเหมือนของใบหน้า ทำให้ผลการทดลองของเวลาในการจับคู่ภาพทำได้เร็วขึ้นต่อมา (Sedaghat, Mokhtarzade, & Ebadi, 2011) ยังพัฒนาอัลกอริทึมเพิ่มเติมจาก Scale-Invariant Feature Transform (SIFT) เพราะลักษณะของภาพที่ใช้ในการเปรียบเทียบกับแผ่นแบบนั้นไม่ได้มีขนาดที่เท่ากันทุกภาพทำให้เกิดปัญหาในขั้นตอนการเปรียบเทียบกับแผ่นแบบ อีกทั้ง (Zheng, Yin, Lu, Xu, & Sole, 2016) ได้ประยุกต์ใช้หลักการเปรียบเทียบแผ่นแบบมาช่วยลดความซ้ำซ้อนในการเข้ารหัสและถอดรหัสของวีดีโอ ดังนั้นผู้วิจัยจึงพัฒนาอัลกอริทึม TMUI ขึ้นเนื่องจากใช้วิธีการลดความผิดพลาดตามแนวขอบของภาพที่เกิดขึ้นตอนแปลงภาพเป็นบิต 0,1 ทำให้เพิ่มความถูกต้องในการเปรียบเทียบค่าพิกเซลของภาพได้

2. ยูเนียน (Union) เป็นหลักการและทฤษฎีทางคณิตศาสตร์ที่ใช้วิเคราะห์ค่าสมาชิกทั้งหมด แทนด้วยสัญลักษณ์ $A \cup B$ ต่อมาได้มีการประยุกต์ใช้ในงานด้านการประมวลผลภาพ เช่น การจำแนกความเหมือนของพิกเซลภาพ โดยดูลักษณะของพิกเซลภาพด้านซ้ายและด้านบนด้วยวิธีการสแกนค่าความเหมือนของพิกเซลในแต่ละบรรทัด แสดงดังภาพที่ 1



ภาพที่ 1 การตรวจสอบจำนวนพิกเซลภาพ (Fiorio & Gustedt, 1996)

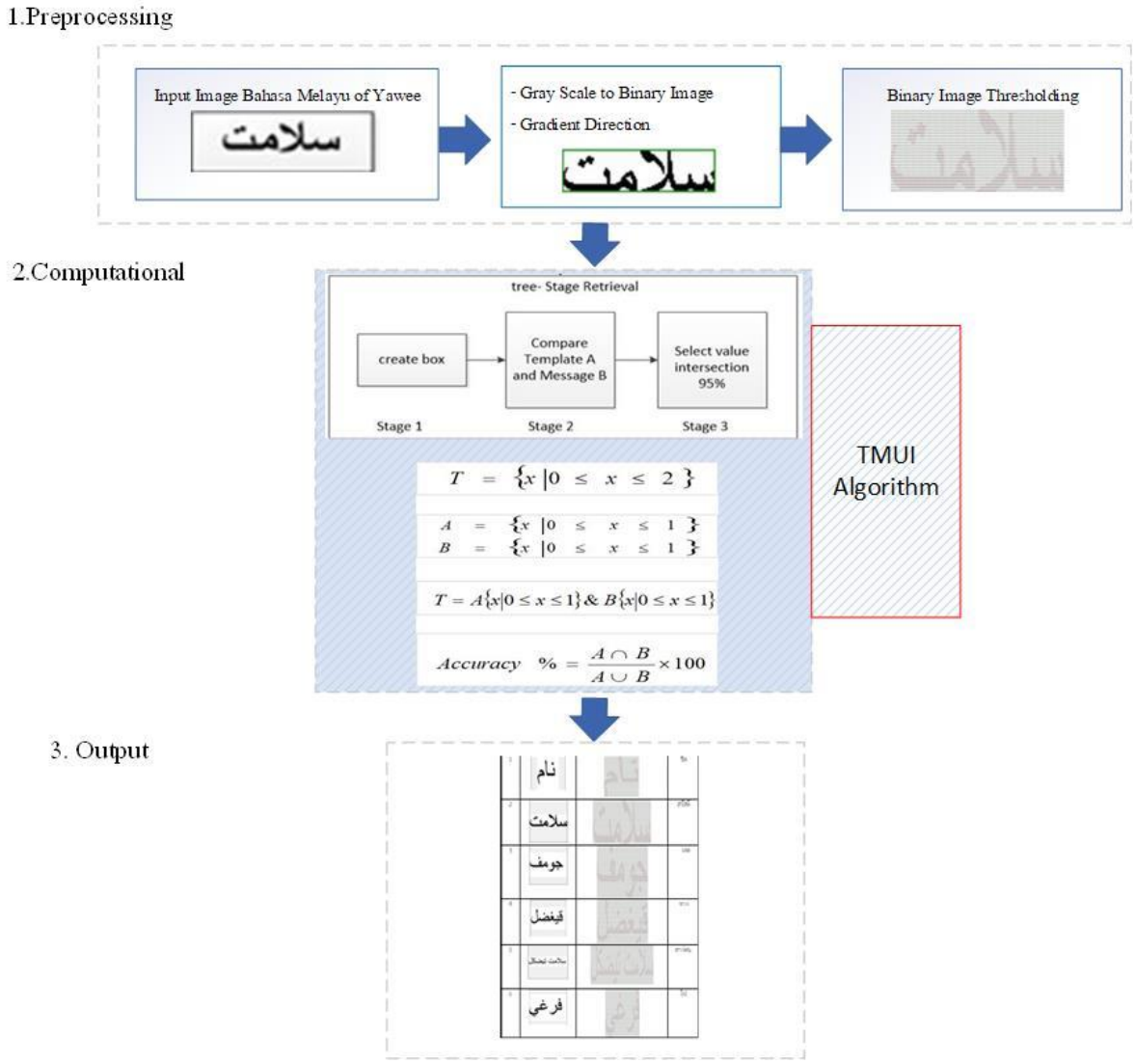
3. อินเตอร์เซกชัน (Intersection) เป็นอีกหนึ่งหลักการและทฤษฎีทางคณิตศาสตร์ที่ใช้วิเคราะห์สำหรับการหาจำนวนสมาชิกทั้งหมดที่มีความเหมือนกันกับเซตแผ่นแบบที่สร้างขึ้น แทนด้วยสัญลักษณ์ $A \cap B$ โดยคำนวณจากความเหมือนของพิกเซลแต่ละภาพ โดยทำงานในลักษณะทำซ้ำ (Loop) จนครบจำนวนของพิกเซล แสดงดังภาพที่ 2



ภาพที่ 2 การตรวจสอบความเหมือนพิกเซลของภาพ (Petrosino & Salvi, 2006)

วิธีดำเนินการวิจัย

สำหรับเนื้อหาส่วนนี้นำเสนอกรอบแนวคิดอัลกอริทึม Template Matching Union Intersection (TMUI) โดยวิธีการดำเนินการวิจัยการเปรียบเทียบแผ่นแบบของอัลกอริทึม TMUI ประกอบด้วยขั้นตอนการทำงานแสดงดังภาพที่ 3



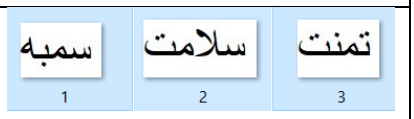
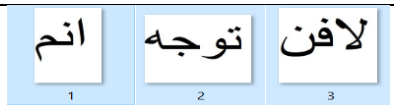
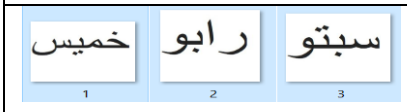
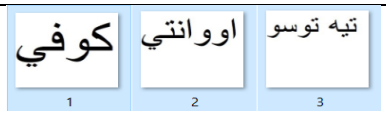
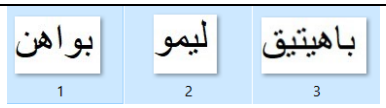
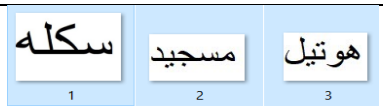
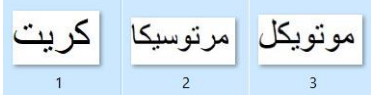
ภาพที่ 3 กรอบแนวคิดการวิจัย

จากภาพที่ 3 ขั้นตอนการทำงานของอัลกอริทึม TMUI ของการแปลภาษามลายูอักษรยาวีเป็นภาษาไทย สามารถแสดงขั้นตอนวิธีการดำเนินงานวิจัยดังนี้

1. กระบวนการก่อนการประมวลผล (Preprocessing) ของขั้นตอนการทำงานของอัลกอริทึม จะนำคำที่แปลงเป็นไฟล์รูปภาพเข้าสู่กระบวนการประมวลผลภาพโดยทำการปรับไฟล์ภาพให้อยู่ในลักษณะของภาพที่เหมาะสมกับการนำไปใช้งาน และเข้าสู่กระบวนการการทำงานของอัลกอริทึม TMUI ของการแปลภาษามลายูอักษรยาวีเป็นภาษาไทย ประกอบด้วย 3 ขั้นตอนดังนี้

1.1 การโหลดไฟล์ภาพเป็นขั้นตอนแรกของการนำตัวอักษรมลายูที่เก็บเป็นไฟล์ภาพไว้เข้าสู่ระบบโดยผ่านอุปกรณ์เพื่อเป็นการแปลงภาพให้เป็นภาพดิจิทัล โดยภาพที่ใช้เป็นตัวอย่างมาจากผู้เชี่ยวชาญที่มีความรู้ด้านอักษรยาวี โดยเกิดจากการสร้างคำให้เป็นไฟล์ภาพ โดยภาพที่ใช้ทดลองจำนวน 500 ภาพ มีการแบ่งกลุ่มออกเป็น 5 กลุ่ม คือ คำทั่วไป ตัวเลข/วัน/เดือน เครื่องดื่ม/ผลไม้ สถานที่ ยานพาหนะ แสดงตัวอย่างดังตารางที่ 1

ตารางที่ 1 ตัวอย่างภาพที่ใช้ในงานวิจัย

ชนิดกลุ่ม	ไฟล์ภาพ
คำทั่วไป	
ตัวเลข/วัน/เดือน	
	
เครื่องดื่ม/ผลไม้	
	
สถานที่	
ยานพาหนะ	

1.2 การแปลงภาพเป็นสีเทา (Shu, Zhang, Chen, Haigron, & Luo, 2010) โดยนำภาพที่รับเข้าจากขั้นตอนที่ 1 มาเข้าสู่สมการเพื่อคำนวณหาค่าสีเทาและนำค่าที่ได้ไปแทนจุดพิกเซลเดิม และทำการหาทิศทางของข้อความโดยหาได้จากสมการ (1) และ (2) ดังนี้

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

โดยกำหนดให้

G' คือ ค่าระดับสีเทา

R คือ ค่าระดับสีแดง

G คือ ค่าระดับสีเขียว

B คือ ค่าระดับสีน้ำเงิน

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

โดยกำหนดให้

GD (x, y) ทิศทางของขอบภาพที่ตำแหน่ง (x, y)

1.3 การทำ Thresholding (Zhu, Jing, & Bi, 2010) ภาพที่ได้จากขั้นตอนที่ 2 คือ การแปลงจากภาพสีเทาที่มีค่าตั้งแต่ 0 ถึง 255 ให้เป็นภาพสองระดับเพื่อแบ่งพิกเซลของภาพเป็นสองกลุ่มคือกลุ่มสีขาวมีค่าเป็น 1 และกลุ่มสีดำมีค่าเป็น 0 โดยหาได้จากสมการ (3) ดังนี้

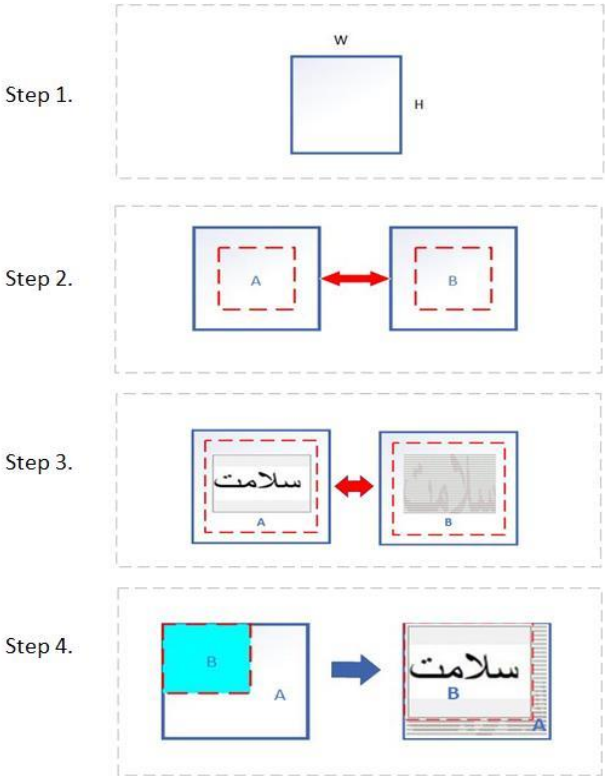
$$g(x,y) = \begin{cases} 1 \ if \ \ f(x,y) > \mathrm{T} \\ 0 \ if \ \ f(x,y) \leq \ \mathrm{T} \end{cases} \quad (3)$$

โดยกำหนดให้

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

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

2. กระบวนการแก้ปัญหา (Computational) เป็นขั้นตอนการทำงานของอัลกอริทึม TMUI ในส่วนการประมวลผลของการเปรียบเทียบกับแผ่นแบบที่สร้างเก็บไว้ในฐานข้อมูล โดยวิธีการนี้เป็นวิธีการสามารถที่จะแก้ไขและลดความพลาดของการวิเคราะห์หาค่าพิกเซลของภาพได้ โดยการทำงานของอัลกอริทึม TMUI ประกอบด้วย 4 ขั้นตอนดังต่อไปนี้



ภาพที่ 4 การทำงานของอัลกอริทึม TMUI

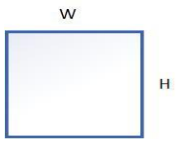
ขั้นตอนที่ 1 การสร้างพื้นที่กล่องสี่เหลี่ยมเปล่าขึ้นมาเพื่อใช้สำหรับเก็บค่าตัวแปร A และ B ซึ่งสร้างจากความกว้าง x ยาว ของแบบเปล่า ดังสมการ 4 แสดงการทำงานดังภาพที่ 5

$$T = \{x|0 \leq x \leq 2\} \tag{4}$$

โดยกำหนดให้

T คือ กระดาษเปล่าที่สร้างขึ้นสำหรับเก็บค่าตัวแปร A และ B

X คือ สมาชิกที่อยู่ในเซต



ภาพที่ 5 การสร้างกล่องสี่เหลี่ยมเปล่า

ขั้นตอนที่ 2 ระบบทำการสร้างตัวแปร A และตัวแปร B ขึ้นมา ดังสมการ 5 แสดงการทำงานดังภาพที่ 6

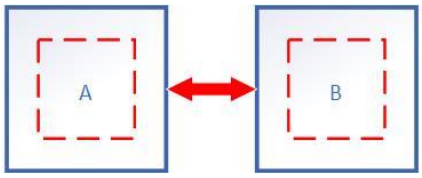
$$A = \{x|0 \leq x \leq 1\}$$
$$B = \{x|0 \leq x \leq 1\} \tag{5}$$

โดยกำหนดให้

A คือ ไฟล์ภาพ

B คือ แผ่นแบบ

X คือ สมาชิกที่อยู่ในเซต



ภาพที่ 6 พื้นที่เก็บตัวแปร A และ ตัวแปร B

ขั้นตอนที่ 3 ระบบทำการโหลดภาพแผ่นแบบตัวแปรโดยอ่านจากไฟล์ โดยในงานวิจัยนี้แทนค่าตัวแปรของภาพแผ่นแบบที่อ่านเข้ามาเป็น A และทำการเปิดไฟล์ภาพที่ต้องการแปลข้อความเข้าระบบ โดยในงานวิจัยนี้แทนค่าตัวแปรของไฟล์ภาพที่อ่านเข้ามาเป็น B ทำให้ในกระดาษเปล่าที่สร้างขึ้นมีทั้งตัวแปร A และ B ดังสมการ 6 แสดงการทำงานดังภาพที่ 7

$$T = A\{x|0 \le x \le 1\} \,\&\, B\{x|0 \le x \le 1\} \tag{6}$$

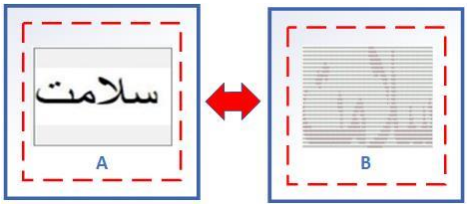
โดยกำหนดให้

T คือ กระดาษเปล่า

A คือ ไฟล์ภาพ

B คือ แผ่นแบบ

X คือ สมาชิกที่อยู่ในเซต



ภาพที่ 7 การโหลดภาพแผ่นแบบและไฟล์ภาพ

ขั้นตอนที่ 4 เมื่อระบบมีข้อมูลเข้าสำหรับใช้ในการเปรียบเทียบแผ่นแบบของอัลกอริทึม TMUI เรียบร้อยแล้ว นั่นคือ ภายในกล่องข้อความเปล่าที่ถูกสร้างขึ้นประกอบด้วยตัวแปร A และ B และดูการ Intersection $A \cap B$ ของพื้นที่ที่ข้อความตรงกับแผ่นแบบ ดังสมการที่ 7 โดยตั้งค่าความเหมือนของข้อความกำหนดไว้ที่ ร้อยละ 95 แสดงการทำงานดังภาพที่ 8 โดยกำหนดค่าดังนี้

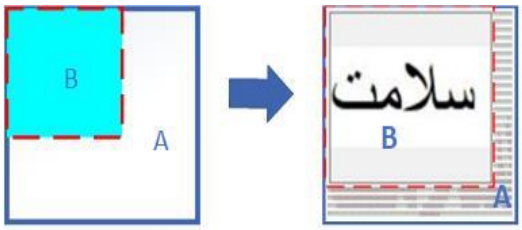
$$Accuracy\,\% = \frac{A \cap B}{A \cup B} \times 100 \qquad (7)$$

โดยกำหนดให้

Accuracy% คือ ค่าความเหมือนของตัวแปร A และ B

$A \cap B$ คือ จำนวนพิกเซลที่ค่าต่างกัน

$A \cup B$ คือ จำนวนพกเซลที่ใช้เปรียบเทียบทั้งหมด



ภาพที่ 8 การเปรียบเทียบแผ่นแบบของอัลกอริทึม TMUI

เครื่องมือที่ใช้ในการวิจัย

ในการวิจัยนี้เป็นการพัฒนาอัลกอริทึมสำหรับการแปลภาษามลายูอักษรยาวีเป็นภาษาไทย ซึ่งมีซอฟต์แวร์และฮาร์ดแวร์ ที่ใช้ในการวิจัยในการออกแบบและทดสอบระบบดังนี้

- เครื่องคอมพิวเตอร์ CPU core TM i7-6500
- โปรแกรม Visual Studio 2017
- Open CV ภาษา C#
- Library System.Drawing;
  - Class System.Drawing.Bitmap
  - Class System.Drawing.Graphics

การวิเคราะห์ข้อมูล

สำหรับการวิเคราะห์ข้อมูล ที่ผู้วิจัยนำเสนอในส่วนวิธีการดำเนินการวิจัย ขอนำเสนอตัวอย่างซอร์ซโค้ด (Source Code) ในส่วนอัลกอริทึมที่พัฒนาด้วยภาษา C# ในส่วนของการตรวจสอบและเปรียบเทียบความเหมือนพิกเซลของภาพของอัลกอริทึม TMUI

อัลกอริทึม TMUI (การตรวจสอบและเปรียบเทียบความเหมือนพิกเซลของภาพ)

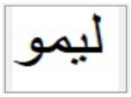
```
private bool CompText(char[,] a, char[,] b, int i, int j
    {
    bool ch = false;
    int c1 = 0, c2 = 0, c3 = 0;
     for (int x = 0; x < i - 1; x++)
       {
        for (int y = 0; y < j; y++)
          {
            c1 = c1 + 1;
            if ((a[x, y] == b[x, y]) && a[x, y] == '2') c1 = c1 - 1;
            else if ((a[x, y] == b[x, y]) && a[x, y] != '2') c2 = c2 + 1;
            else c3 = c3 + 1;
          }
       }
       tbC1.Text = c1.ToString();
       tbC2.Text = c2.ToString();
       tbC3.Text = c3.ToString();
       Double f1 = Convert.ToDouble(c1);
       Double f2 = Convert.ToDouble(c2);
       Double f3 = f2 / f1;
       tbPc.Text = (f3*100).ToString("#.00") +" %";
       if ( f3 > 0.95) ch = true;
       return ch;
     }
```

ภาพที่ 9 อัลกอริทึมการตรวจสอบและเปรียบเทียบความเหมือนพิกเซลของภาพ

ผลการวิจัย

การทดลองของระบบ การทดลองสำหรับงานวิจัยมาจากกลุ่มคำของภาษามลายูอักษรยาวี โดยมีขนาด w= 108 และ h = 55 พิกเซล และกำหนดขอบเขตของ ROI (Region Of Interest) ที่ w = 100-100 ภาพถูกเปลี่ยนจาก RGB เป็นสีเทา โดยมีความละเอียด 300 dpi นามสกุล .jpg และเข้าสู่ระบบการแปลงภาพเป็นไบนารีด้วยเทคนิคการทำ Thresholding ในการทดลองใช้ภาพ 500 ภาพเพื่อดูผลการทำงานของการเปรียบเทียบแผ่นแบบของอัลกอริทึม TMUI

ตารางที่ 2 ผลการทดลอง

ไฟล์ภาพ	เปรียบเทียบกับแผ่นแบบ	จำนวนพิกเซล	ความถูกต้อง	คำแปลภาษาไทย
		2765	100%	ผลไม้
		3360	100%	มะละกอ
		2106	100%	ส้ม
		3036	99.67%	เจอ,พบ
		2640	89.50%	ป้าย

อภิปรายผล

งานวิจัยนี้นำเสนออัลกอริทึมใหม่สำหรับการแปลภาษามลายูอักษรยาวีเป็นภาษาไทยที่มีชื่อว่า Template Matching Union Intersection (TMUI) โดยข้อดีของการทำงานของอัลกอริทึม TMUI สามารถสร้างกล่องสี่เหลี่ยมสำหรับรับภาพ (T) ประกอบด้วยภาพแผ่นแบบ (A) และไฟล์ภาพ (B) ที่ต้องการเข้าสู่กระบวนการแปลความหมายแบบไม่จำกัดขนาดของภาพ โดยทำการวิเคราะห์รูปร่างและความหมายของภาพ ซึ่งระบบจะดูภาพที่รับเข้ามาทั้งภาพ (A) และ (B) ภาพใดมีขนาดใหญ่จะใช้ภาพนั้นเป็นภาพเริ่มต้นตามหลักการทำงานทางคณิตศาสตร์แบบยูเนียน (Union) $A \cup B$ และอินเตอร์เซกชัน (Intersection) $A \cap B$ และระบบจะดูการ Intersection ของพิกเซลภาพที่เหมือนกันมากที่สุดโดยระบบตั้งค่าไว้ที่ร้อยละ 95 เป็นค่าที่คำนึงถึงการผิดผลาดตามแนวของภาพที่เกิดขึ้นตอนแปลงภาพเป็นบิต 0,1 คือ เมื่อค่าความถูกต้องมากกว่าหรือเท่ากับ 95 ระบบจะทำการแปลข้อความ แต่ถ้าน้อยกว่าหรือเท่ากับ 94 ระบบจะไม่แปลข้อความ ซึ่งการทดสอบอัลกอริทึม TMUI ที่พัฒนาขึ้นมีประสิทธิภาพความถูกต้องโดยคิดเป็นค่าเฉลี่ยของความถูกต้องในการแปลร้อยละ 98.50 จากจำนวนภาพ 500 ภาพที่ได้ทำการทดลอง โดยอัลกอริทึมที่พัฒนาขึ้นสามารถนำไปใช้งานได้จริง โดยพัฒนาเป็นระบบสำหรับการแปลภาษา อีกทั้งสามารถเปลี่ยนลักษณะของกลุ่มตัวอย่างไฟล์ภาพเป็นภาษาอื่นที่ต้องการแปลได้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 จากการพัฒนาระบบโดยเลือกใช้ software library Class System.Drawing.Bitmap สำหรับใช้ในการอ่านข้อมูลพิกเซลของไฟล์ภาพนั้น และ ใช้ software library Class System.Drawing.Graphics สำหรับใช้ตีกรอบข้อความเพื่อใช้ในการตัดข้อความออกมาจากพื้นหลังสีขาวของภาพที่ไม่ได้ใช้ โดยดูระดับค่าที่คำนึงถึงการผิด

ผลาดตามแนวของภาพที่เกิดขึ้นตอนแปลงภาพเป็นบิต 0,1 ทำให้กระบวนการการปรับปรุงคุณภาพของภาพสามารถนำไปใช้ในขั้นตอนการสร้างฐานข้อมูลแผ่นแบบของภาษามลายูอักษรยาวีได้อย่างมีประสิทธิภาพ

1.2 ผลการจากวิจัยสามารถนำระบบที่พัฒนาขึ้นไปใช้เป็นระบบต้นแบบสำหรับการแปลภาษามลายูอักษรยาวีแบบลักษณะคำและลักษณะรูปแบบประโยค เพื่อใช้สอนกลุ่มนักเรียนในโรงเรียนสอนศาสนาอิสลามในจังหวัดชายแดนภาคใต้สำหรับใช้ในการสื่อสารในชีวิตประจำวัน

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 การศึกษางานวิจัยนี้เป็นเพียงการสร้างกลุ่มคำภาษามลายูอักษรยาวีสำหรับนำไปใช้ในการแปลเป็นภาษาไทยเท่านั้น ควรมีการพัฒนาระบบเพื่อรองรับในลักษณะของการแปลแบบการเขียนด้วยลายมือเพื่อให้ระบบสามารถทำงานได้ทั้ง 2 รูปแบบ คือ การแปลจากการเขียนด้วยลายมือ และการแปลที่เกิดจากการสร้างกลุ่มคำ

2.2 ระบบที่พัฒนาขึ้นสามารถนำไปต่อยอดเพื่อพัฒนาการใช้งานรูปแบบแอปพลิเคชันบนสมาร์ทโฟน ทั้งนี้เพื่อเพิ่มประสิทธิภาพการใช้งานและอำนวยความสะดวกในการใช้งานได้ในทุกรูปแบบ

เอกสารอ้างอิง

Ahuja, K., & Tuli, P. (2013). Object recognition by template matching using correlations and phase angle method. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(3), 1368-1373.

Al-Amri, S. S., & Kalyankar, N. V. (2010). Image segmentation by using threshold techniques. arXiv preprint arXiv:1005.4020, *Journal of Computing*, 2(5),83-86.

Briechle, K., & Hanebeck, U. D. (2001). Template matching using fast normalized cross correlation. *The Proceedings of the Optical Pattern Recognition XII.* Proc. SPIE 4387, 20 March 2001 at Orlando, FL, United States.

Chiang, D. (2010). Learning to translate with source and target syntax. *The Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 11-16 July 2010, Uppsala, Sweden, 1443-1452.

Cohn, T., & Specia, L. (2013). Modelling annotator bias with multi-task gaussian processes: An application to machine translation quality estimation. *The Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 4-9 August 2013, Sofia, Bulgaria, 32–42.

Dawoud, N. N., Samir, B. B., & Janier, J. (2011). Fast template matching method based optimized sum of absolute difference algorithm for face localization. *International Journal of Computer Applications,* 18(8), 0975-8887.

Egashira, R. (2000).Language processing unit and a language processing method to translate a source program and generate an object module file. In: *Google Patents. U.S. Patent* No. 6,014,519. 11 January 2000.

Fiorio, C., & Gustedt, J. (1996). Two linear time union-find strategies for image processing. *Theoretical Computer Science*, 154(2), 165-181.

Herriman, M. (2005). Language, Ethnicity, Culture and the Conflict in Southern Thailand. *NUCB journal of language culture and communication*, 7(1), 15-30.

Khalil, M. (2010). Car plate recognition using the template matching method. *International Journal of Computer Theory and Engineering*, 2(5), 683.

Matheson, V., & Hooker, M. B. (1988). Jawi literature in Patani: The maintenance of an Islamic tradition. *Journal of the Malaysian Branch of the Royal Asiatic Society,* 61(1 (254), 1-86.

Petrosino, A., & Salvi, G. (2006). Rough fuzzy set based scale space transforms and their use in image analysis. *International Journal of Approximate Reasoning*, 41(2), 212-228.

Sedaghat, A., Mokhtarzade, M., & Ebadi, H. (2011). Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(11), 4516-4527.

Shu, H., Zhang, H., Chen, B., Haigron, P., & Luo, L. (2010). Fast computation of Tchebichef moments for binary and grayscale images. *IEEE Transactions on Image Processing*, 19(12), 3171-3180.

Ting, S.-H. (2010).Impact of language planning on language choice in friendship and transaction domains in Sarawak, Malaysia. *Current Issues in Language Planning*, 11(4), 397-412.

Zakraoui, J., Elloumi, S., Alja'am, J. M., & Yahia, S. B. (2019). Improving Arabic Text to Image Mapping Using a Robust Machine Learning Technique. *IEEE Access*, 7, 18772-18782.

Zheng, Y., Yin, P., Lu, X., Xu, Q., & Sole, J. (2016). Methods and apparatus for reduced complexity template matching prediction for video encoding and decoding. *Google Patents. U.S. Patent* No. 9,516,341. 6 December 2016.

Zhu, Q., Jing, L., & Bi, R. (2010). Exploration and improvement of Ostu threshold segmentation algorithm. *The Proceedings of the 2010 8th World Congress on Intelligent Control and Automation*, 7-9 July 2010, Jinan, China, 6183-6188.