



รายงานการวิจัยฉบับสมบูรณ์

การรู้จำตัวอักษรภาษาไทยโดยใช้ตัวอธิบายคุณลักษณะ

Printed Thai Character Recognition Using Standard descriptor

ผศ. ดร. กนต์พงษ์ วรรณปัญญา

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2557

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รายงานการวิจัยฉบับสมบูรณ์



T137616

การรู้จำตัวอักษรภาษาไทยโดยใช้ตัวอธิบายคุณลักษณะ

Printed Thai Character Recognition Using Standard descriptor

RCH
ท 3899
8554

ผศ. ดร. กนต์พงษ์ วรรณปัญญา

เลขหมู่.....
เลขทะเบียน 137616
ปี เดือน ปี 13 ก.ค. 2558

b. 12681611
i.....

ได้รับทุนสนับสนุนงานวิจัยจากเงินรายได้ ประจำปีงบประมาณ 2557

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อโครงการ (ภาษาไทย) การรู้จำตัวอักษรภาษาไทยโดยใช้ตัวอธิบายคุณลักษณะ

แหล่งเงิน

เงินรายได้

ประจำปีงบประมาณ

2557

จำนวนเงินที่ได้รับการสนับสนุน 100,000 บาท

ระยะเวลาทำการวิจัย

1 ปี

ตั้งแต่ 1 ตุลาคม 2556 ถึง 30 กันยายน 2557

ชื่อ-สกุล หัวหน้าโครงการ ผศ. ดร. กนต์พงษ์ วรรณปัญญา สังกัด คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

ความคล้ายคลึงกันของตัวอักษรภาษาไทยเป็นความท้าทายที่สำคัญอย่างหนึ่งในการรู้จำตัวอักษร โดยเฉพาะอย่างยิ่งความแตกต่างของชุดแบบของตัวอักษร รูปแบบ และขนาดของตัวอักษร โดยทั่วไปการรู้จำตัวอักษรจะประกอบด้วย 3 กระบวนการหลัก ได้แก่ กระบวนการจัดเตรียมข้อมูล กระบวนการสกัดคุณลักษณะ และกระบวนการแบ่งแยกกลุ่มข้อมูล ในกระบวนการจัดเตรียมข้อมูล ชุดข้อมูลจะถูกแปลงให้อยู่ในรูปแบบมาตรฐานโดยการกลับค่าสีของภาพ การเติมเต็มภาพให้เป็นจัตุรัส และการปรับขนาดของภาพตัวอักษร จากนั้นในกระบวนการสกัดคุณลักษณะ ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (Adaptive Histogram of Oriented Gradient: AHOG) ที่พัฒนาขึ้นจะถูกนำมาใช้เพื่อแก้ไขปัญหาด้านความคล้ายคลึงกันของตัวอักษร AHOG เป็นการพัฒนามาจากฮิสโตแกรมของการไล่ระดับเชิงมุม (Histogram of Oriented Gradient: HOG) ซึ่งได้รับการปรับปรุงใน 2 ส่วน คือ การแบ่งส่วนภาพออกเป็นภาพย่อยแบบปรับเหมาะ (Adaptive Partition) และการกำหนดความกว้างของถังแบบปรับเหมาะ (Adaptive Binning) โดยในส่วนแรกจะเป็นการแบ่งส่วนภาพแบบต้นไม้จตุนิยม (Quadtree Decomposition) ตามความแปรปรวนของข้อมูลเพื่อเพิ่มประสิทธิภาพในการสกัดคุณลักษณะส่วนย่อย และส่วนที่สองจะเป็นการกำหนดความกว้างของแต่ละถังตามหลักการแบบไม่สม่ำเสมอ (Non-Uniform) เพื่อให้ตัวอธิบายคุณลักษณะที่ถูกสร้างจาก AHOG มีข้อผิดพลาดน้อยที่สุด และในกระบวนการสุดท้าย การวัดระยะทางแบบยุคลิดจะถูกนำมาใช้ในการแบ่งแยกคุณลักษณะของ AHOG จากผลการทดลองกับภาพตัวอักษรที่ไม่มีสัญญาณรบกวนทั้งหมด 59,408 ภาพซึ่งถูกแบ่งออกเป็นชุดทดสอบกับชุดฝึกสอนเท่า ๆ กันพบว่าคุณลักษณะที่สกัดจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะมีประสิทธิภาพการรู้จำถูกต้องเฉลี่ย 97.59% ซึ่งสูงกว่าประสิทธิภาพของคุณลักษณะที่ได้จากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบเดิม รวมทั้งฮิสโตแกรมของการไล่ระดับเชิงมุมที่มีการปรับปรุงแล้ว ได้แก่ SSHOG, PHOG, MHOG และ HOGC

Research Title: Printed Thai Character Recognition Using Standard descriptor

Researcher: Asst.Prof. Dr. Kuntpong Woratpanya

Faculty: Information Technology **Department:** -

ABSTRACT

A similarity of printed Thai characters is a grand challenge of optical character recognition (OCR), especially in case of a variety of font-types, -sizes, and -styles. In general, the framework of character recognition is composed of pre-processing, feature extraction, and classification processes. In pre-processing process, a dataset is transformed to a standard form by using image complementation, image zero padding, and image resizing procedures. In feature extraction process, an adaptive histogram of oriented gradient (AHOG) is proposed for overcoming the character similarity. This method improves the conventional histogram of oriented gradient (HOG) in two principal aspects, which are (i) adaptive partition for gradient images and (ii) adaptive binning for oriented histograms. The former is implemented with quadtree partition based on gradient image variance so as to provide an effective local feature extraction. The latter is implemented with non-uniform mapping technique, so that the AHOG descriptor can be constructed with minimal errors. In the last process, Euclidean distance is applied to classify AHOG features. Based on 59,408 single character noise-free images equally divided into training and testing samples, the experimental results show that the AHOG method achieves 97.59% average accuracy rate which outperforms the conventional HOG and state-of-the-art methods, including scale space histogram of oriented gradient (SSHOG), pyramid histogram of oriented gradient (PHOG), multilevel histogram of oriented gradient (MHOG), and HOG column encoding algorithm (HOGC).

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่ให้ความอนุเคราะห์และสนับสนุนทุนวิจัยจากแหล่งทุนเงินรายได้ ประจำปีงบประมาณ พ.ศ. 2557

ขอขอบคุณสมาชิกทุกคนห้องวิจัยการรู้จำรูปแบบและการประมวลผลภาพ (PRIP LAB) คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่ช่วยกันเตรียมการทดลองและประมวลผลข้อมูล และขอขอบคุณผู้ช่วยวิจัย นายธราวิชญ์ ธิติจรูญโรจน์ ที่ได้ดำเนินการทดลองวิจัยตามแผนการวิจัยที่ได้กำหนดไว้ได้อย่างครบถ้วนและสมบูรณ์

สุดท้าย ขอขอบคุณศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ สำหรับการให้ข้อมูลทรัพยากรภาษา ทั้งตัวอักษรภาษาอังกฤษ ตัวอักษรภาษาไทย ตัวเลข และสัญลักษณ์ ซึ่งผู้วิจัยได้นำมาใช้ในงานวิจัยนี้

ผศ.ดร.กัณฑ์พงษ์ วรรัตน์ปัญญา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูปภาพ.....	VIII
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	2
1.3 ขอบเขตการวิจัย.....	3
1.4 กรอบแนวคิดที่นำเสนอ.....	3
1.5 ข้อตกลงเบื้องต้น.....	4
1.6 คำจำกัดความที่ใช้ในการศึกษา.....	5
1.7 ประโยชน์ที่คาดว่าจะได้รับ.....	5
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ทฤษฎีที่เกี่ยวกับตัวอักษรไทย.....	7
2.2 การประมวลผลภาพดิจิทัล (Digital Image Processing).....	8
2.2.1 ภาพดิจิทัล.....	8
2.2.2 การปรับขนาดรูปภาพ.....	9
2.2.3 การแปลงข้อมูลภาพให้เป็นเวกเตอร์.....	10

สารบัญ (ต่อ)

	หน้า
2.3 การรู้จัก	10
2.3.1 การสกัดคุณลักษณะ	10
2.3.2 การแบ่งแยกข้อมูล	16
2.4 การทดสอบความทนทาน	17
2.5 งานวิจัยที่เกี่ยวข้อง	18
บทที่ 3 การดำเนินงานวิจัย	24
3.1 ปัญหาของการรู้จักตัวอักษรภาษาไทย	24
3.2 แนวทางการดำเนินงานวิจัย	24
3.2.1 การแยกองค์ประกอบคำเดียวสำหรับการสกัดคุณลักษณะ และใช้การวัดเชิงมุมด้วยการฉายเมตริกซ์	24
3.2.2 การสกัดคุณลักษณะ โดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ และใช้การวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษร	29
บทที่ 4 ผลการวิจัย	37
4.1 ผลการทดลองที่ 1	38
4.1.1 การเลือกองค์ประกอบเมตริกซ์คุณลักษณะที่เหมาะสม	38
4.1.2 การทดสอบความถูกต้องของตัวอธิบายคุณลักษณะ SD1 และ SD2	41
4.1.3 การทดสอบความทนทานของตัวอธิบายคุณลักษณะ	42
4.2 ผลการทดลองที่ 2	44
4.2.1 การทดสอบความถูกต้องของตัวอธิบายคุณลักษณะ SD-AHOG	44
4.2.2 การเปรียบเทียบขนาดตัวอธิบายคุณลักษณะและเวลาที่ใช้ในการสกัดคุณลักษณะ ..	45
4.2.3 การทดสอบความทนทานของตัวอธิบายคุณลักษณะ	46
4.3 ผลการทดลองที่ 3	47
4.3.1 การเปรียบเทียบความถูกต้องของตัวอธิบายคุณลักษณะ	47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
4.3.2 การเปรียบเทียบขนาดของคุณลักษณะและเวลาที่ใช้ในการสกัดคุณลักษณะ	48
4.3.3 การเปรียบเทียบความทนทานของตัวอธิบายคุณลักษณะ	49
4.4 ผลการทดลองที่ 4.....	49
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	52
5.1 สรุปผลการวิจัย	52
5.1.1 ผลงานวิจัยของการสกัดคุณลักษณะ โดยการแยกองค์ประกอบค่าเดียว	52
5.1.2 ผลงานวิจัยของการสกัดคุณลักษณะ โดยฮิสโตแกรมของการได้ระดับเชิงมุม แบบปรับเหมาะ	53
5.1.3 ผลงานวิจัยของการเปรียบเทียบการสกัดคุณลักษณะ	53
5.1.4 ผลงานวิจัยของการเปรียบเทียบประสิทธิภาพการรู้จำโดยจำแนกตามกรณีต่าง ๆ	54
5.2 ข้อเสนอแนะ	54
บรรณานุกรม	56
ภาคผนวก ก ผลงานวิจัยที่ได้รับการตีพิมพ์	59
ประวัติผู้แต่ง	82

สารบัญตาราง

ตารางที่	หน้า
2.1 สมาชิกตัวอักษรภาษาไทย.....	7
4.1 ผลการทดลองความถูกต้องของการรู้จำของข้อมูลชุดทดสอบ.....	42
4.2 ผลการทดสอบความทนทานในการรู้จำที่มีความแตกต่างของชุดแบบอักษรกับ ขนาดของตัวอักษรด้วยวิธี SD1 และ SD2.....	43
4.3 ผลการเปรียบเทียบประสิทธิภาพการรู้จำ HOG, SSHOG, PHOG, MHOG, HOGC และ SD-AHOG	44
4.4 ผลการเปรียบเทียบเวลาในการสกัดคุณลักษณะและจำนวนคุณลักษณะของ HOG, SSHOG, PHOG, MHOG, HOGC และ SD-AHOG.....	45
4.5 ผลการทดสอบความทนทานในการรู้จำที่มีความแตกต่างของชุดแบบของอักษรกับ ขนาดของตัวอักษรด้วยวิธี SD-AHOG, HOG, SSHOG, PHOG, MHOG และ HOGC.....	46
4.6 การเปรียบเทียบความทนทานต่อขนาดและชุดแบบอักษรระหว่าง SD2 กับ SD-AHOG.....	49
4.7 กลุ่มของตัวอักษรภาษาไทยที่มีความคล้ายคลึงกัน โดยแบ่งแยกตามระดับ	50
4.8 ผลการทดลองเปรียบเทียบกรณีการรู้จำถูกต้อง, การรู้จำผิดพลาดในกลุ่มที่มี ความคล้ายคลึงกัน และการรู้จำผิดพลาดในกลุ่มที่ไม่มีความคล้ายคลึงกัน	50

สารบัญรูปภาพ

รูปที่	หน้า
1.1 กรอบแนวคิดของการวิจัยตัวอธิบายคุณลักษณะเพื่อการรู้จำตัวอักษรภาษาไทย.....	4
2.1 โครงสร้างการเขียนประโยคภาษาไทยซึ่งถูกแบ่งออกเป็น 4 ระดับ	8
2.2 การเก็บข้อมูลภาพในระดับจุดภาพ	9
2.3 การปรับขนาดของตัวอักษร	9
2.4 การแปลงเมตริกให้เป็นเวกเตอร์	10
2.5 ผลคูณของเมตริกซ์ U, S และ V	11
2.6 กระบวนการทำงานของฮิสโตแกรมของการไล่ระดับเชิงมุม	13
2.7 กระบวนการ Gradient Computation.....	13
2.8 การกำหนดขนาดของเซลล์และจำนวนถัง	14
2.9 การกำหนดขนาดของบล็อก และอัตราการเคลื่อนที่	15
2.10 การแบ่งข้อมูลสำหรับทดสอบความทนทาน.....	18
3.1 กระบวนการรู้จำด้วยการแยกตัวประกอบค่าเดียวสำหรับการสกัดคุณลักษณะและ การวัดเชิงมุมด้วยการฉายเมตริกซ์	25
3.2 การจัดเตรียมข้อมูล	26
3.3 ขั้นตอนการสกัดคุณลักษณะ	26
3.4 เมตริกซ์ฝึกสอนซึ่งสร้างจากภาพอักษรจากชุดฝึกสอน	27
3.5 การเปรียบเทียบไอเกนเวกเตอร์ลำดับต้นกับลำดับปลาย.....	28
3.6 กระบวนการรู้จำด้วยการสกัดคุณลักษณะ โดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุม แบบปรับเหมาะและการวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษร	30
3.7 การหาขอบภาพ.....	31
3.8 การแบ่งภาพออกเป็นส่วนย่อยโดยอาศัยการแยกองค์ประกอบแบบต้นไม้ตัดสินใจ	32
3.9 ความแตกต่างการแบ่งภาพออกเป็นย่อย.....	32

สารบัญรูปภาพ (ต่อ)

รูปที่	หน้า
3.10 การกำหนดความกว้างของถึง.....	33
3.11 การกำหนดจำนวนถึงและความกว้างแต่ละถึง	33
3.12 เปรียบเทียบความแตกต่างของการรวมขนาดแบบสเกลาร์และเวกเตอร์.....	34
3.13 เปรียบเทียบผลลัพธ์ก่อนและหลังผ่านกระบวนการ จัดรูปแบบฮิสโตแกรมมาตรฐาน.....	35
4.1 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับบน	39
4.2 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับกลาง	39
4.3 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับล่าง.....	40
4.4 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับบน	40
4.5 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับกลาง	41
4.6 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับล่าง.....	41
4.7 ตัวอย่างภาพอินพุทของ “ก-ภ-ถ” ที่มีความคล้ายคลึงกัน	43
4.8 การเปรียบเทียบความถูกต้องของการรู้จำระหว่าง SD2 กับ SD-AHOG ของตัวภาษาไทยในระดับกลาง.....	47
4.9 การเปรียบเทียบความถูกต้องของการรู้จำระหว่าง SD2 กับ SD-AHOG ของตัวภาษาไทยในระดับล่าง	47

สารบัญรูปภาพ (ต่อ)

รูปที่	หน้า
4.10 การเปรียบเทียบขนาดของคุณลักษณะระหว่าง SD2 กับ SD-AHOG	48
4.11 การเปรียบเทียบเวลาที่ใช้ในการสกัดคุณลักษณะระหว่าง SD2 กับ AHOG	48
4.13 ผลลัพธ์ของ AHOG ในกรณีการรู้จำผิดพลาดในกลุ่มที่ไม่มีมีความคล้ายคลึงกัน ในระดับบน	51
4.14 ผลลัพธ์ของ AHOG ในกรณีการรู้จำผิดพลาดในกลุ่มที่ไม่มีมีความคล้ายคลึงกัน ในระดับกลาง	51



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

วิวัฒนาการของชุดแบบอักษร (Font) สำหรับภาษาไทยได้ถูกพัฒนาอย่างต่อเนื่อง เพื่อสร้างสรรค์งานสื่อสิ่งพิมพ์ในรูปแบบต่าง ๆ อาทิเช่น นิตยสาร หนังสือพิมพ์ ตำราเรียน ใบปลิว และอื่น ๆ สื่อเหล่านี้มีความหลากหลายของชุดแบบอักษรจึงกลายมาเป็นความท้าทายที่สำคัญของการรู้จำตัวอักษร นอกจากนี้ชนิดของตัวอักษร (Font Type) และขนาดตัวอักษร (Font Size) ถือเป็นปัจจัยที่ส่งผลโดยตรงต่อประสิทธิภาพการรู้จำตัวอักษรภาษาไทย

ถึงแม้ว่าการรู้จำตัวอักษรภาษาไทยจะถูกพัฒนามาอย่างต่อเนื่องตลอดระยะเวลาที่ล่วงเลยมา (Kimpan, C. et.al. 1987 ; Tanprasert and Koanantakool 1996 ; Tanprasert and Sae-tang 1999 ; Duangphasuk, S. 2001 ; Tangsurakit, N. 2005 ; Kruatrachue, B. et.al. 2007 ; Suttapakti, U. et.al. 2013) แต่การรู้จำตัวอักษรภาษาไทยก็ยังคงต้องได้รับการปรับปรุงให้มีประสิทธิภาพการรู้จำสูงขึ้น ปัญหาหลักของการรู้จำตัวอักษรภาษาไทยสามารถจำแนกได้ 2 ประเด็น คือ ประเด็นแรกความคล้ายคลึงกันของตัวอักษร (ตัวอักษรที่อยู่ในกลุ่มเดียวกัน) เช่น “ข-ช-ซ” หรือ “ญ-ฉ-ฉ-ฉ” เป็นต้น ซึ่งปัญหาที่เกิดขึ้น เมื่อผ่านกระบวนการแบ่งแยกภาพตัวอักษรแล้ว ภาพตัวอักษรที่อยู่คนละคลาส แต่มีความคล้ายคลึงกันจะถูกจัดให้อยู่ในกลุ่มเดียวกัน ประเด็นที่ 2 ในช่วงหลายปีที่ผ่านมาชุดแบบอักษรได้ถูกพัฒนาขึ้นเป็นจำนวนมาก นั้นหมายความว่าประสิทธิภาพของการรู้จำตัวอักษรก็จะลดลงตามไปด้วย เพราะตัวอักษรที่อยู่ในคลาสเดียวกัน แต่มีรูปแบบของตัวอักษรแตกต่างกันมีโอกาสถูกแยกให้อยู่คนละกลุ่ม ดังนั้น เพื่อให้ประสิทธิภาพการรู้จำตัวอักษรภาษาไทยสูงขึ้น จำเป็นต้องพัฒนาตัวอธิบายคุณลักษณะ (Descriptor) ที่มีประสิทธิภาพ ซึ่งเป็นตัวแทนของภาพตัวอักษรที่ได้จากการสกัดคุณลักษณะ (Feature Extraction) ในการสกัดคุณลักษณะเชิงรูปร่างด้วยวิธีฮิสโตแกรมของการไล่ระดับเชิงมุม (Histogram of Oriented Gradient: HOG) ถือเป็นวิธีการที่มีประสิทธิภาพได้ถูกนำเสนอโดย Navneet and Bill (2005: 1-8) หลังจากงานวิจัยนี้ได้รับการเผยแพร่ มีนักวิจัยท่านอื่น ๆ ได้นำแนวคิดของฮิสโตแกรมของการไล่ระดับเชิงมุมมาปรับปรุงเพื่อการรู้จำวัตถุ อาทิเช่น Scale Space Histogram of Oriented Gradient โดยเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Ning, H.et.al. (2008: 167-170), Pyramid Histogram of Oriented Gradient โดย Anna, B.et.al. (2007) และ Multiscale Histogram of Oriented Gradient โดย Subhransu, M.et.al. (2008), แต่มีเพียงงานวิจัยของ Andrew and Lewis (2011: 1085-1089) ที่นำเสนอ Histogram of Oriented Gradient Column สำหรับการรู้จำตัวอักษรในภาพธรรมชาติ

ถึงแม้ว่าฮิสโตแกรมของการไล่ระดับเชิงมุมแบบดั้งเดิม (Conventional HOG) และฮิสโตแกรมของการไล่ระดับเชิงมุมที่ได้รับการปรับปรุง (Modified HOG) จะประสบความสำเร็จในการรู้จำวัตถุ และตัวอักษรภาษาอังกฤษจากภาพธรรมชาติ แต่เมื่อนำมาใช้กับการรู้จำตัวอักษรภาษาไทยพบว่า วิธีการข้างต้นไม่สามารถแยกตัวอักษรภาษาไทยที่มีความคล้ายคลึงกันได้ ดังนั้น ในงานวิจัยนี้ได้นำเสนอวิธีการฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (Adaptive Histogram of Oriented Gradient: AHOG) โดย Woraratpanya and Titjaroonroj (2014: 83-92) ซึ่งได้ทำการปรับปรุงมาจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบดั้งเดิมใน 2 ส่วนหลักคือ ส่วนการแบ่งภาพออกเป็นส่วนย่อยโดยอาศัยการแยกองค์ประกอบแบบต้นไม้จตุรนิยม (Quadtree Decomposition) การแบ่งด้วยวิธีการนี้จะทำให้คุณลักษณะเชิงรูปร่างมีความละเอียดมากขึ้น เนื่องจากสามารถสกัดคุณลักษณะเชิงรูปร่างแบบท้องถิ่น (Local Shape Feature) ได้ดี และในส่วนที่สอง ได้ทำการปรับปรุงการแบ่งขนาดและจำนวนถัง (Bin) โดยอาศัยค่าความแปรปรวน (Variance) ในการกำหนดจำนวนและความกว้างของถัง อีกทั้งปรับเปลี่ยนวิธีการรวมขนาดของแต่ละถัง ซึ่งวิธีการเดิมเป็นการรวมขนาดแบบสเกลาร์ให้เป็นการรวมขนาดแบบเวกเตอร์ เพื่อให้คุณลักษณะเชิงรูปร่างมีความเด่นชัดมากยิ่งขึ้น

ดังนั้น ในงานวิจัยนี้จึงได้นำเสนอกระบวนการการรู้จำตัวอักษร โดยใช้วิธีสกัดคุณลักษณะด้วยฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ เพื่อเพิ่มประสิทธิภาพของการรู้จำตัวอักษรภาษาไทย วิธีการนี้จะสามารถสร้างตัวอธิบายคุณลักษณะ (Descriptor) ที่มีความทนทานต่อชุดแบบและขนาดที่แตกต่างกันของตัวอักษรภาษาไทย ตลอดจนความคล้ายคลึงกันของตัวอักษร (Character Similarity) อีกด้วย

1.2 วัตถุประสงค์

1.2.1 เพื่อสร้างตัวอธิบายคุณลักษณะที่สามารถแบ่งแยกข้อมูลที่มีความคล้ายคลึงกันและทนทานต่อความหลากหลายของชุดแบบตัวอักษรและขนาดตัวอักษร โดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2.2 เพื่อประเมินและเปรียบเทียบประสิทธิภาพของกระบวนการรู้จำตัวอักษรภาษาไทย โดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะกับการอื่น ๆ ในกลุ่มของฮิสโตแกรมของการไล่ระดับเชิงมุม

1.3 ขอบเขตการวิจัย

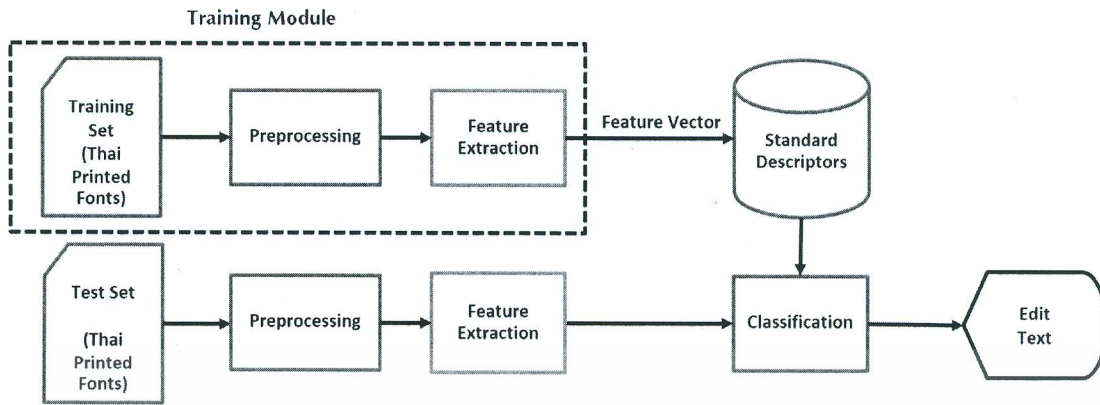
1.3.1 งานวิจัยนี้ มุ่งเน้นเฉพาะการสกัดคุณลักษณะสำหรับการรู้จำตัวอักษรภาษาไทย เท่านั้น

1.3.2 งานวิจัยนี้ทำการศึกษาระบบการรู้จำตัวอักษรแบบออฟไลน์ (Off-line Character Recognition) เฉพาะในส่วนของตัวอักษรภาษาไทย สระ วรรณยุกต์ และตัวเลขไทยเท่านั้น ไม่ครอบคลุมถึงภาษาต่างประเทศ เลขอารบิก และอักขระพิเศษ โดยที่ภาพตัวอักษรที่ใช้ในการวิจัยเป็นตัวอักษรเดี่ยวที่ไม่มีสัญญาณรบกวน (Noise) ข้อมูลที่ใช้ในการวิจัยจะแบ่งเป็น 2 ชุด คือ ชุดฝึกฝน (Training Set) และชุดทดสอบ (Test Set) โดยชุดฝึกฝนมีทั้งหมด 12 ชุดแบบอักษรในแต่ละชุดแบบอักษรมีขนาด 8, 10, 12, 14, 16, 18, 20 และ 22 พอยต์ รูปแบบปกติ หนา เอียง และหนาป็นเอียง

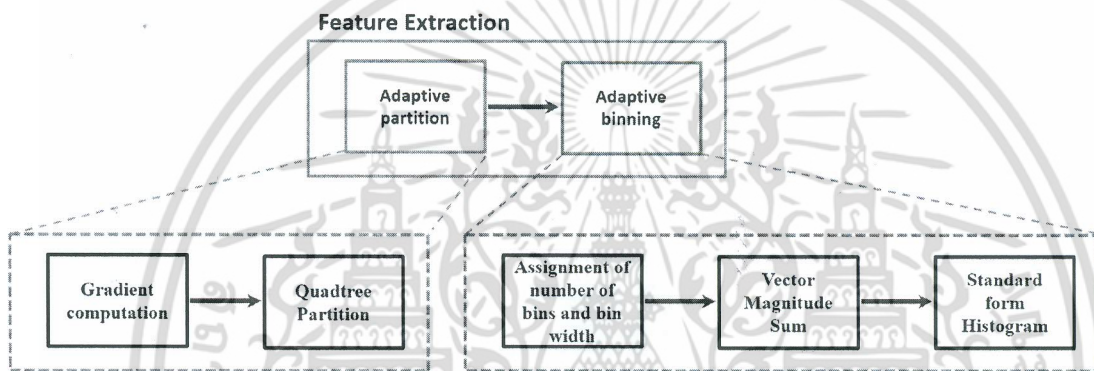
1.4 กรอบแนวความคิดของงานวิจัย

ในงานวิจัยนี้มุ่งเน้นการสกัดคุณลักษณะที่เหมาะสมกับการรู้จำตัวอักษรภาษาไทยตามที่ได้กล่าวไว้ในขอบเขต ภายใต้สมมติฐานที่ว่าถ้าสามารถสกัดคุณลักษณะที่มีประสิทธิภาพ ถึงจะใช้ตัวแบ่งแยกข้อมูลแบบพื้นฐาน ก็สามารถแบ่งแยกกลุ่มข้อมูลที่มีความคล้ายคลึงออกจากกันได้ ดังนั้นกรอบแนวความคิดที่นำเสนอในรูปที่ 1.1(ก) จึงมุ่งเน้นขั้นตอนการสกัดคุณลักษณะ ซึ่งประกอบด้วย 2 ขั้นตอนหลัก คือ ขั้นตอนการแบ่งส่วนภาพแบบปรับเหมาะ (Adaptive Partition) เป็นขั้นตอนการหาขอบจากภาพอักษร ผลลัพธ์ที่ได้ คือ ค่าความเข้มแสงที่ประกอบด้วยขนาด และทิศทางของทุกจุดภาพ หลังจากนั้น จึงทำการแบ่งภาพอักษรเป็นภาพย่อย (Cell) และขั้นตอนการแบ่งถึงแบบปรับเหมาะ (Adaptive Binning) เป็นขั้นตอนการกำหนดความกว้างของถังให้สอดคล้องกับข้อมูลจากนั้นจะดำเนินการรวมขนาดในแต่ละถังโดยแบ่งตามองศา ดังรูปที่ 1.1(ข)) และในการจัดเตรียมข้อมูลเบื้องต้น (Pre-processing) ได้ดำเนินการตามงานวิจัยของ Woraratpanya and Titijaronroj (2013: 165-173) เพื่อทดสอบประสิทธิภาพของตัวอธิบายคุณลักษณะที่สกัดได้จากกระบวนการสกัดคุณลักษณะด้วยฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ ในงานวิจัยนี้เลือกใช้วิธีการวัดระยะทางแบบยูคลิด ซึ่งเป็นวิธีพื้นฐานสำหรับการแบ่งแยกภาพอักษร ซึ่งการปรับปรุงในประเด็นดังกล่าวข้างต้น ส่งผลให้ประสิทธิภาพการรู้จำตัวอักษรสูงขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(ก)



(ข)

รูปที่ 1.1 กรอบแนวคิดของการวิจัยตัวอธิบายคุณลักษณะเพื่อการรู้จำตัวอักษรภาษาไทย (ก) กรอบแนวคิดของการวิจัย (ข) กระบวนการสกัดคุณลักษณะด้วยฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

1.5 ข้อตกลงเบื้องต้น

ภาพตัวอักษรที่ใช้ในข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบจะเป็นหนึ่งภาพต่อหนึ่งตัวอักษรโดยที่พื้นหลังของภาพจะเป็นสีดำ และตัวอักษรเป็นสีขาว ในส่วนของการทดสอบเพื่อวัดประสิทธิภาพของตัวอธิบายคุณลักษณะจะทดสอบ 2 ส่วน ได้แก่ ส่วนแรกการทดสอบค่าความถูกต้องของการรู้จำโดยใช้การเลือกสุ่มข้อมูลแบบเที่ยงตรง K กลุ่ม (K -Fold Cross Validation) และการทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อการเปลี่ยนแปลงขนาด และชุดแบบของตัวอักษรโดยใช้การเลือกสุ่มข้อมูลแบบเที่ยงตรง K กลุ่ม แบบมีเงื่อนไขสำหรับการแบ่งข้อมูล ซึ่งในหัวข้อ 2.4 ซึ่งจะได้อธิบายขั้นตอนการทดสอบความทนทานของตัวอธิบายคุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 คำจำกัดความที่ใช้ในการศึกษา

1.6.1 ตัวอธิบายคุณลักษณะ (Descriptor) หมายถึง คุณลักษณะที่เหมาะสมสำหรับเป็นตัวแทนภาพอักษรภาษาไทย โดยจะถูกสร้างมาจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

1.6.2 SD1 หมายถึง ตัวอธิบายคุณลักษณะที่ถูกสร้างจากข้อมูลชุดฝึกสอนด้วยวิธีการแยกองค์ประกอบค่าเดียวจากการเลือกไอเกนเวกเตอร์ใน “ลำดับปลาย” มาตรฐานเป็นตัวอธิบายคุณลักษณะโดยในงานวิจัยนี้ได้เลือกไอเกนเวกเตอร์ที่ลำดับ 1 ถึง 7

1.6.3 SD2 หมายถึง ตัวอธิบายคุณลักษณะที่ถูกสร้างจากข้อมูลชุดฝึกสอนด้วยวิธีการแยกองค์ประกอบค่าเดียวจากการเลือกไอเกนเวกเตอร์ใน “ลำดับต้น” มาตรฐานเป็นตัวอธิบายคุณลักษณะโดยในงานวิจัยนี้ได้เลือกไอเกนเวกเตอร์ที่ลำดับ 512 ถึง 704

1.6.4 SD-AHOG หมายถึง ตัวอธิบายคุณลักษณะที่ถูกสร้างจากข้อมูลชุดฝึกสอนด้วยกระบวนการฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

1.6.5 ความหลากหลายของชุดแบบตัวอักษร หมายถึง ตัวอักษรภาษาไทยหนึ่งตัวที่ถูกเขียนด้วยชุดแบบอักษร (Font) ที่แตกต่างกัน เช่น “ก-ก-ก-ก-ก” เป็นต้น

1.6.6 ความคล้ายคลึงของตัวอักษร หมายถึง ตัวอักษรภาษาไทยที่มีรูปร่างใกล้เคียงกัน เช่น “ก-ภ-ถ” “ญ-ฒ-ณ-ณ” เป็นต้น

1.6.7 ขนาดตัวอักษร หมายถึง ตัวอักษรภาษาไทยที่มีขนาดแตกต่างกันซึ่งมีขนาดอยู่ในช่วง 8, 10, 12, 14, 16, 18, 20 และ 22 พอยต์ เช่น “ก-ก-ก-ก-ก” เป็นต้น

1.6.8 ชนิดตัวอักษร หมายถึง ตัวอักษรภาษาไทยที่มีลักษณะแตกต่างกันได้แก่ ตัวปกติ ตัวหนา ตัวเอียง และตัวหนาปนเอียง เช่น “ก-ก-ก-ก” เป็นต้น

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 ได้ตัวอธิบายคุณลักษณะที่มีประสิทธิภาพ และทนทานต่อความหลากหลายของชุดตัวอักษรและขนาดของตัวอักษรสำหรับการรู้จำตัวอักษรพิมพ์ภาษาไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.7.2 สามารถนำไปประยุกต์ใช้ในกระบวนการสร้างหนังสือสำหรับผู้พิการทางสายตา (Braille Book) และหนังสือเสียง (Audio Book) ซึ่งกระบวนการสร้างหนังสือสำหรับผู้พิการทางสายตาจะต้องพึ่งพากระบวนการรู้จำตัวอักษรภาษาไทย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงความรู้พื้นฐานในการรู้จำตัวอักษรภาษาไทย โดยเนื้อหาจะประกอบไปด้วย 4 ส่วน คือ ทฤษฎีที่เกี่ยวกับตัวอักษรภาษาไทย การประมวลผลภาพดิจิทัล ทฤษฎีการสกัดคุณลักษณะร่วมกับการแบ่งแยกกลุ่มข้อมูล และวรรณกรรมที่เกี่ยวข้องกับการรู้จำตัวอักษร

2.1 ทฤษฎีที่เกี่ยวกับตัวอักษรไทย

การรู้จำตัวอักษรภาษาไทยจำเป็นต้องเข้าใจลักษณะหรือธรรมชาติของโครงสร้างทางภาษาก่อนเป็นอันดับแรก ซึ่งจะเป็นประโยชน์ในการวิเคราะห์ปัญหาหรือหาความสัมพันธ์ของคุณลักษณะทางภาษา ดังนั้น ในหัวข้อนี้จะกล่าวถึง โครงสร้างตัวอักษรภาษาไทย ดังนี้ ตัวอักษรภาษาไทยประกอบด้วยพยัญชนะ 44 รูป สระ 16 รูป วรรณยุกต์ 4 รูป ตัวอักษรพิเศษ 4 รูป และเลขไทย 10 ตัว ดังแสดงในตารางที่ 2.1

ตารางที่ 2.1 สมาชิกตัวอักษรภาษาไทย

ประเภทตัวอักษร	สมาชิก
พยัญชนะ	ก ข ค ด ต ถ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห พ อ ฮ
สระ	อะ , อา , อิ , อี , อึ , อือ , อำ , อู , อุ , โอ , ไอ , ไอ , เอ , แอ , ฤ , ฌ
วรรณยุกต์	อ่ , ั , ็ , ๋
ตัวอักษรพิเศษ	ั้ , ั๊ , ั๋ , ั๊ , ั
ตัวเลขไทย	๐ , ๑ , ๒ , ๓ , ๔ , ๕ , ๖ , ๗ , ๘ , ๙

โครงสร้างการเขียนภาษาไทยสามารถแบ่งออกได้เป็น 4 ระดับ ได้แก่ ตัวอักษรภาษาไทยในระดับเหนือบน ระดับบน ระดับกลาง และระดับล่างดังแสดงในรูปที่ 2.1 จากโครงสร้างดังกล่าวสามารถใช้ในการจำแนกข้อมูลออกเป็น 3 ระดับ ได้แก่ตัวอักษรภาษาไทยในระดับบน ระดับกลาง และระดับล่าง โดยที่ตัวอักษรภาษาไทยในระดับบนมาจากการรวมกันของระดับเหนือบนกับระดับบน และมีสมาชิกทั้งหมด 12 ตัว ประกอบด้วย อิ อี อึ อือ อ่ ั ็ ๋ ั๊ ั๋ ั๊ ั และ อำ ใน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนตัวอักษรภาษาไทยในระดับล่างจะประกอบด้วย อ อู และสมาชิกนอกเหนือจากที่กล่าวมาจะอยู่ในตัวอักษรภาษาไทยในระดับกลางทั้งหมด

<h1>นิตยสาร</h1> <h1>นิตยสาร</h1> <h1>นิตยสาร</h1> <h1>นิตยสาร</h1>	ระดับเหนือบน
	ระดับบน
	ระดับกลาง
	ระดับล่าง

รูปที่ 2.1 โครงสร้างการเขียนประโยคภาษาไทยซึ่งถูกแบ่งออกเป็น 4 ระดับ

2.2 การประมวลผลภาพดิจิทัล

ในการรู้จำตัวอักษรต้องอาศัยองค์ความรู้พื้นฐานการประมวลผลภาพดิจิทัล (Digital Image Processing) ในหัวข้อนี้จะอธิบายทฤษฎีการประมวลผลภาพเชิงดิจิทัลที่เกี่ยวข้องกับการปรับขนาดรูปภาพ และการแปลงข้อมูลภาพให้อยู่ในรูปแบบเวกเตอร์ โดยเริ่มจากทฤษฎีภาพเชิงดิจิทัล

2.2.1 ภาพดิจิทัล

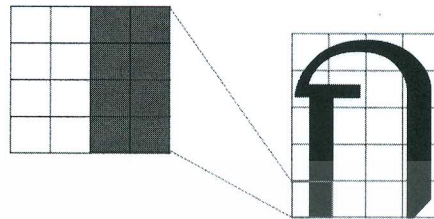
ภาพดิจิทัล (Digital Image) สามารถนิยามเป็นฟังก์ชันสองมิติ $f(x, y)$ โดยที่ x และ y เป็นพิกัดของภาพและแอมพลิจูดของ f ที่พิกัด (x, y) ใด ๆ ภายในภาพ คือ ค่าความเข้มแสง (Intensity) ของภาพที่ตำแหน่งนั้น ๆ และเมื่อ x, y และแอมพลิจูดของ f เป็นค่าจำกัด (Finite Value) จึงเรียกรูปภาพนี้ว่า ภาพดิจิทัล และถ้ากำหนดให้ภาพ $f(x, y)$ มีขนาด M แถวและ N คอลัมน์ และพิกัดของจุดกำเนิด (Origin) ของภาพ คือ ตำแหน่ง $(x, y) = (0, 0)$ แล้ว สามารถแสดงให้อยู่ในรูปแบบเมตริกซ์ได้ ดังสมการที่ 2.1

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix} \quad (2.1)$$

จากสมการที่ 2.1 แต่ละค่าที่อยู่ในเมตริกซ์จะเรียกว่า “จุดภาพ” ตำแหน่ง $(0, 0)$ จะอยู่ทางด้านซ้ายมือสุดด้านบนของภาพ การจัดลำดับตำแหน่งของจุดภาพจะเรียงจากซ้ายไปขวาในแต่ละเส้นจุดและการจัดลำดับของเส้นจุดจะเรียงจากบนลงล่าง การเก็บค่าของความเข้มแสงของภาพดิจิทัลลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

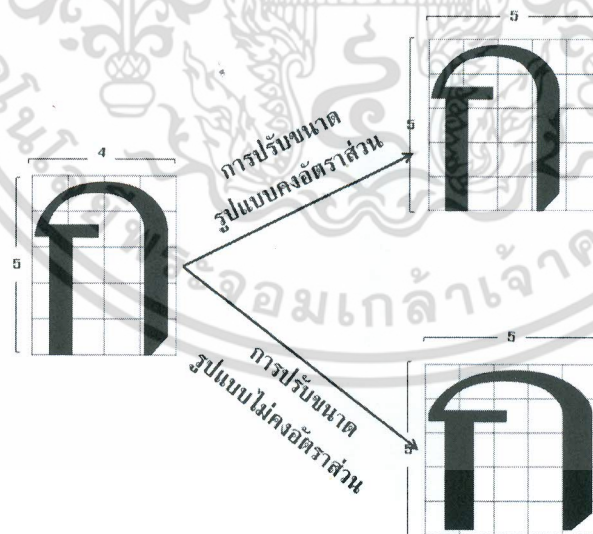
หน่วยความจำในลักษณะเส้นจุด (Raster) นี้ เรียกว่า “ภาพบิตแมป” (Bitmap Image) หรือ “ภาพราสเตอร์” (Raster Image)



รูปที่ 2.2 การเก็บข้อมูลภาพในระดับจุดภาพ

2.2.2 การปรับขนาดรูปภาพ

การปรับขนาดภาพ (Image Resizing) เป็นส่วนที่มีความสำคัญอีกส่วนหนึ่งสำหรับงานวิจัยนี้ เนื่องจากภาพตัวอักษรแต่ละภาพมีขนาดที่แตกต่างกัน ดังนั้น การปรับขนาดภาพจึงมีความจำเป็นที่จะทำก่อนนำภาพไปประมวลผลต่อ วิธีที่นิยมในการปรับขนาดภาพให้มีขนาดเป็นจัตุรัสมีอยู่ 2 วิธีคือ การปรับขนาดภาพโดยคำนึงถึงที่ภาพที่ปรับแล้ว จะมีอัตราส่วนคงเดิม แต่ทำการเติมขอบ (Padding) ของตัวอักษรเพื่อให้ภาพมีขนาดเป็นจัตุรัส ($n \times n$) และการปรับขนาดของภาพโดยไม่คำนึงถึงอัตราส่วนของภาพ โดยยึดให้ภาพมีขนาดเป็นจัตุรัส ($n \times n$) ดังรูปที่ 2.3

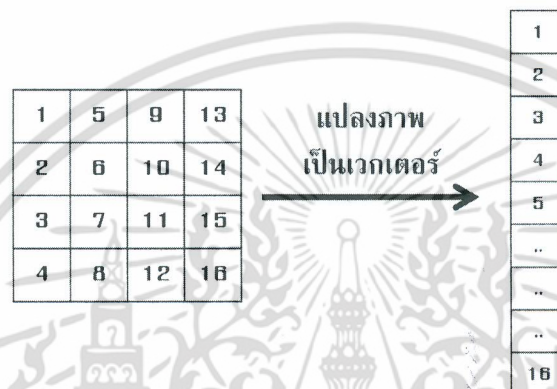


รูปที่ 2.3 การปรับขนาดของตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.3 การแปลงข้อมูลภาพให้เป็นเวกเตอร์

ในการสกัดหาตัวอธิบายคุณลักษณะ จะต้องเตรียมภาพให้อยู่ในรูปแบบเวกเตอร์ เพื่อใช้ในการสกัดหาตัวอธิบายคุณลักษณะ โดยหลักการ คือ การพิจารณาภาพเป็นเมตริกซ์ โดยเมตริกซ์จะมีองค์ประกอบ 2 ส่วนคือ หลัก (สมาชิกตามแนวตั้ง) และแถว (สมาชิกตามแนวนอน) ให้นำหลักต่อหลักจะ ได้ดังรูป 2.4 หรือนำแถวต่อแถวก็ได้ ซึ่งสำหรับการแปลงข้อมูลภาพให้เป็นเวกเตอร์สามารถทำได้ทั้งสองรูปแบบโดยผลลัพธ์ที่ได้ไม่มีความแตกต่างกัน



รูปที่ 2.4 การแปลงเมตริกซ์ให้เป็นเวกเตอร์

2.3 การรู้จำ

การรู้จำ (Recognition) เป็นกระบวนการที่นำข้อมูลจากการสกัดคุณลักษณะ (Feature Extraction) เพื่อให้ได้คุณลักษณะที่มีความสำคัญมาทำการแบ่งแยกข้อมูล (Classification) ซึ่งในกระบวนการนี้ จะนำเอาคุณลักษณะข้างต้นที่ได้มาใช้ในการจำแนกประเภทข้อมูลออกเป็นกลุ่ม ๆ ตามคุณลักษณะ โดยในส่วนแรกจะกล่าวถึงขั้นตอนการสกัดคุณลักษณะหรือกระบวนการหาคุณลักษณะที่มีความสำคัญออกมาจากรูปภาพ

2.3.1 การสกัดคุณลักษณะ

ในงานวิจัยนี้จะกล่าวถึงการสกัดคุณลักษณะเพียง 2 วิธีเท่านั้น คือ การสกัดคุณลักษณะด้วยวิธีการแยกองค์ประกอบค่าเดี่ยว (Singular Value Decomposition: SVD) และการสกัดคุณลักษณะด้วยวิธีการปรับปรุงฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (Adaptive Histogram of Oriented Gradient: AHOG)

2.3.1.1 การแยกองค์ประกอบแบบค่าเดี่ยว

การแยกองค์ประกอบแบบค่าเดี่ยวเป็นการสร้างเมตริกซ์ตัวแทนที่แยกออกมาจากเมตริกซ์เดิม โดยจะทำการแยกเมตริกซ์ A ออกเป็นสามเมตริกซ์ คือ U , S และ V เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

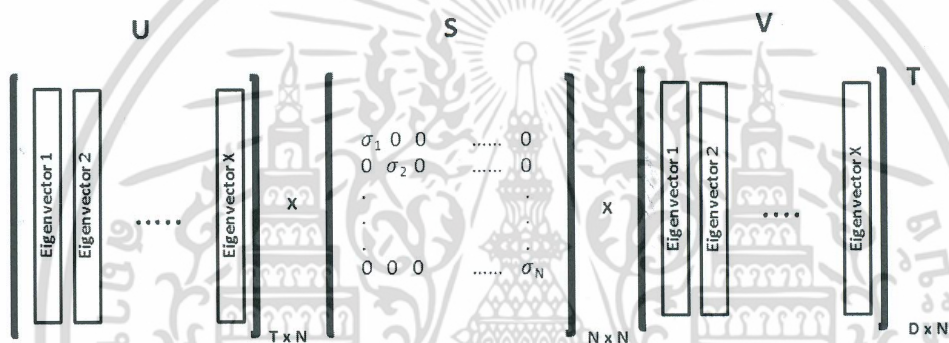
ดังรูปที่ 2.5 ซึ่งเมตริกซ์ S จะเป็นเมตริกซ์ทแยง ส่วนเมตริกซ์ U และ V จะเป็นเมตริกซ์ตั้งฉาก ในการแยกองค์ประกอบแบบค่าเดียวสามารถคำนวณได้ตามสมการที่ 2.2

$$A = U \times S \times V^T \quad (2.2)$$

โดยที่ U คือ เมตริกซ์ของคอลัมน์เป็นไอเจนเวกเตอร์ของ AA^T หรือเรียกว่าไอเจนเวกเตอร์ซ้าย (Left Eigenvector)

V^T คือ เมตริกซ์ของคอลัมน์เป็นไอเจนเวกเตอร์ของ $A^T A$ หรือเรียกว่าไอเจนเวกเตอร์ขวา (Right Eigenvector)

S คือ เมตริกซ์ทแยง (Diagonal) คือค่าเอกฐาน (Singular Values) ของ A



รูปที่ 2.5 ผลคูณของเมตริกซ์ U S และ V

การคำนวณการแยกองค์ประกอบแบบค่าเดียวขึ้นอยู่กับ การคำนวณหาค่าของไอเจนเวกเตอร์ของ AA^T และ $A^T A$ โดยที่ไอเจนเวกเตอร์ของ $A^T A$ จะใช้ในการสร้างคอลัมน์ของ V ส่วนไอเจนเวกเตอร์ของ AA^T จะใช้ในการสร้างคอลัมน์ของ U ส่วน S คำนวณได้จากรากที่สองของ AA^T หรือ $A^T A$ โดยค่าเอกฐานเป็นการเรียงลำดับจากค่ามากไปหาน้อยของค่าเมตริกซ์ทแยง S ซึ่งค่าเอกฐานเป็นจำนวนจริง (Real Number) ถ้า A เป็นเลขจำนวนจริงจะทำให้ U และ V เป็นจำนวนจริงด้วย ซึ่งในการหาค่าไอเจนของ A ได้จากตามสมการที่ 2.3

$$W = A^T A = AA^T \quad (2.3)$$

$$Wx = \lambda x \quad (2.4)$$

จากสมการที่ 2.4 จะเห็นว่า λ คือ ค่าไอเจนของ A และ x คือ ไอเจนเวกเตอร์ของ A

$$(W - \lambda I)x = 0 \quad (2.5)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมการที่ 2.5 โดยที่ I คือ เมตริกซ์เอกลักษณ์ ในการหาค่าสามารถคำนวณ S ได้จาก λ โดยทำการเรียงลำดับจากค่า λ มากไปหาน้อยตามสมการที่ 2.6 และการหารากที่สองของปริมาณ λ ตามสมการที่ 2.7 และสร้างเมตริกซ์ S โดยที่ σ_i เป็นค่าเอกฐานหรือค่าเดี่ยว (Singular Value) เรียงลำดับในตำแหน่งเมตริกซ์ที่แยงจากค่าเอกฐานมากไปหาน้อย โดยผลลัพธ์

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \quad (2.6)$$

$$\sigma_i = \sqrt{|\lambda_i|} \quad (2.7)$$

ในที่นี้เมตริกซ์ A คือ เมตริกซ์ความแปรปรวนร่วม (Covariance Matrix: C_x) ซึ่งหาได้จากการนำภาพตัวอักษรที่มีลักษณะเป็นเมตริกซ์มาทำการแปลงให้เป็นเวกเตอร์ X_1 โดยเมื่อภาพ N ตัวอักษรมีขนาดกว้าง a_1 จุด และความยาว a_2 จุด จะได้เวกเตอร์ X_1 มีขนาดเท่ากับ $(a_1 \times a_2)$ ซึ่งเป็นจำนวนจุดภาพทั้งหมด ในทำนองเดียวกันทำการสร้างเวกเตอร์ $X_2, X_3, X_4, \dots, X_N$ จากนั้น นำเวกเตอร์ดังกล่าวมาเรียงต่อกันในแนวคอลัมน์ตั้งแต่ X_1 ถึง X_N นั่นคือ $[X_1, X_2, \dots, X_N]$ เมื่อ N คือจำนวนเวกเตอร์ของภาพที่ใช้ในการฝึกฝน ค่ารวมเวกเตอร์เฉลี่ย (Mean Vector: M_x) ดังสมการที่ 2.8 และคำนวณเมตริกซ์ความแปรปรวนร่วม ตามสมการที่ 2.9

$$M_x = \frac{1}{N} \sum_{i=1}^N X_i \quad (2.8)$$

$$C_x = \frac{1}{N} \sum_{i=1}^N (X_i - M)(X_i - M)^T \quad (2.9)$$

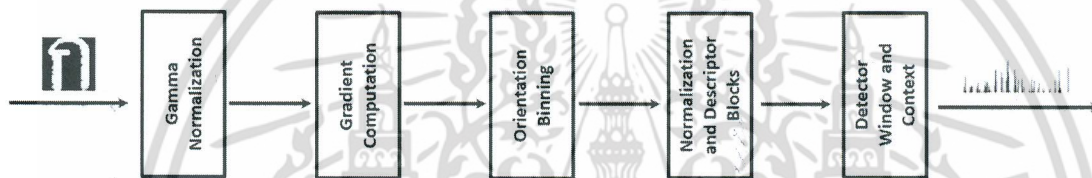
เมตริกซ์ความแปรปรวนร่วมที่ได้จะเป็นเมตริกซ์ขนาด $N \times N$ โดยที่เมตริกซ์ U และ V คือโอแกนเวกเตอร์ซ้ายและโอแกนเวกเตอร์ขวา ส่วนเมตริกซ์ S คือ เมตริกซ์ค่าเอกฐาน (Singular Value Matrix: S) สัมประสิทธิ์ที่ได้จากโอแกนเวกเตอร์ซ้ายเมื่อเรียงตามค่าเอกฐานมากที่สุดจะอยู่ในแถวแรก และเป็นองค์ประกอบที่มีความสำคัญ แกนองค์ประกอบที่สองจะรองลงมาและแกนถัดไปลงตามลำดับ โดยทั่วไปเมตริกซ์ U เป็นโอแกนเวกเตอร์ซ้ายสามารถนำมาใช้เป็นตัวแทนของข้อมูล

จากที่กล่าวมาข้างต้น จะพบว่า การแยกองค์ประกอบค่าเดี่ยวเป็นกระบวนการทางคณิตศาสตร์ที่จัดการเรียงค่าโอแกนเวกเตอร์ตามค่าเอกฐาน โดยทั่วไปแล้ว ค่าโอแกนเวกเตอร์ได้ถูกนำไปใช้เป็นตัวแทนภาพอักษรหรือคุณลักษณะ ซึ่งถ้าเลือกโอแกนเวกเตอร์ในลำดับแรกจะได้เอกสารที่เป็นเอกสารที่ส่งวนเวสสำหรับใช้ในงานเพื่อการศึกษาค้นคว้า เมืออนุญาตเห็นไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะในรูปแบบความถี่ต่ำ แต่ถ้าเลือกไอเกนเวกเตอร์ในลำดับหลังจะได้คุณลักษณะในรูปแบบความถี่สูงหรือรูปร่างของภาพอักษร ดังนั้น จึงสามารถกล่าวได้ว่ากระบวนการแยกองค์ประกอบค่าเดี่ยวสามารถนำมาประยุกต์ใช้ในการสกัดคุณลักษณะจากภาพอักษรภาษาไทยได้

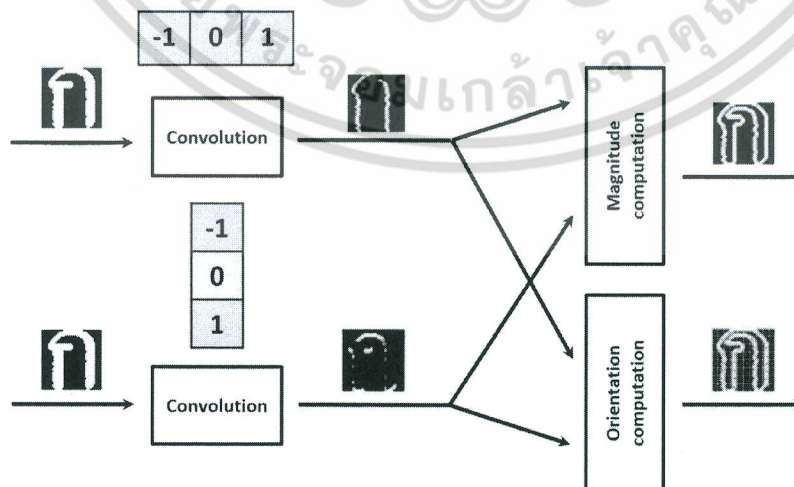
2.3.1.2 ฮิสโตแกรมของการไล่ระดับเชิงมุม

ฮิสโตแกรมของการไล่ระดับเชิงมุม (Histogram of Oriented Gradient: HOG) เป็นกระบวนการที่ใช้ในการสกัดคุณลักษณะเชิงรูปร่างที่มีประสิทธิภาพ ซึ่งวิธีการนี้ถูกนำเสนอครั้งแรกโดย Navneet and Bill (2005: 1-8) เพื่อใช้ในงานทางด้านการระบุภาพว่าเป็นภาพบุคคลหรือไม่ โดยกระบวนการฮิสโตแกรมของการไล่ระดับเชิงมุมจะประกอบไปด้วย 5 ขั้นตอน ดังรูปที่ 2.6



รูปที่ 2.6 กระบวนการทำงานของฮิสโตแกรมของการไล่ระดับเชิงมุม

ขั้นตอนที่ 1 คือ Gamma and Color Normalization เป็นการนอร์มัลไลเซชันภาพอินพุท กล่าวคือ เป็นการทำแกมมาอีควอลไลเซชัน (Gamma Equalization) เพื่อปรับให้ภาพอยู่ในบรรทัดฐานเดียวกัน ซึ่งจากรายงานวิจัยของ Navneet D. และ Bill T. แสดงให้เห็นว่าการปรับมาตรฐานของภาพให้เป็น Gray Space จะให้ผลลัพธ์ที่ดีกว่า RGB Space



รูปที่ 2.7 กระบวนการ Gradient Computation

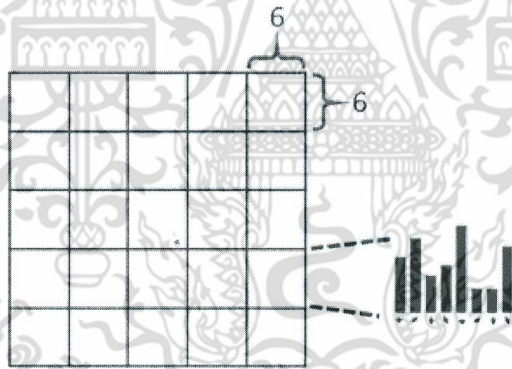
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 2 คือ Gradient Computation เป็นการคำนวณขนาด (Magnitude) และทิศทาง (Orientation: θ) ของเวกเตอร์ในทุกจุดภาพของขอบภาพโดยอาศัยอนุพันธ์อันดับ 1 ของตัวกรองเกาส์เซียน (Gaussian Filter) คือ $[-1 \ 0 \ 1]$ และ $[-1 \ 0 \ 1]^T$ จากนั้นนำตัวกรองข้างต้นมาดำเนินการคอนวอลูชัน (Convolution) กับภาพอินพุท ซึ่งผลลัพธ์ที่ได้คือ I_x และ I_y ขนาดและทิศทางของเวกเตอร์คำนวณตามสมการที่ 2.10 และสมการที่ 2.11 ตามลำดับ

$$\text{Magnitude} = \sqrt{I_x^2 + I_y^2} \quad (2.10)$$

$$\theta = \tan^{-1}\left(\frac{I_y}{I_x}\right); -\pi \leq \theta \leq \pi \quad (2.11)$$

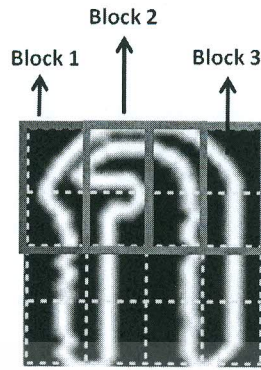
ขั้นตอนที่ 3 คือ Spatial and Orientation Binning ทำการแบ่งภาพอินพุทออกเป็นหลายส่วน แต่ละส่วนจะมีขนาดเป็นจัตุรัส โดยเริ่มจากการกำหนดจำนวนพิกเซลต่อเซลล์ (Cells) และกำหนดจำนวนถังของทิศทาง (Bins) ตัวอย่างเช่นจากรูปที่ 2.8 กำหนดให้แต่ละเซลล์มีขนาด 6×6 พิกเซล และจำนวนถังเท่ากับ 8



รูปที่ 2.8 การกำหนดขนาดของเซลล์และจำนวนถัง

ขั้นตอนที่ 4 คือ Normalization and Descriptor Blocks จะสร้างเวกเตอร์คุณลักษณะ (Feature Vector: w) โดยการรวมขนาดตามถังแบบสเกลาร์ในแต่ละบล็อก (Block) ซึ่งบล็อกหนึ่งจะประกอบไปด้วยหลายเซลล์ขึ้นกับการกำหนด จากนั้นบล็อกจะขยับตามค่า (Overlapping Rate) ที่กำหนด ดังแสดงในรูปที่ 2.9 จะเห็นได้ว่าในหนึ่งบล็อกมี 2×2 เซลล์ และเคลื่อนที่ละครึ่งของขนาดบล็อก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 การกำหนดขนาดของบล็อก และอัตราการเคลื่อนที่

จากนั้นทำการนอร์มัลไลซ์เวกเตอร์คุณลักษณะของแต่ละบล็อก (V) ซึ่งทำได้หลายวิธี แต่วิธีที่ได้รับการนิยมนิยม คือ L1-norm และ L2-norm ตามสมการที่ 2.12 และ 2.13

$$V = \frac{V'}{(\|V\| + \epsilon)} \quad (2.12)$$

$$V = \frac{V'}{\sqrt{\|V\|^2 + \epsilon}} \quad (2.13)$$

โดยที่ V' คือ เวกเตอร์คุณลักษณะของแต่ละบล็อกที่ไม่ผ่านการนอร์มัลไลเซชัน

V คือ เวกเตอร์คุณลักษณะของแต่ละบล็อกที่ผ่านการนอร์มัลไลเซชัน

ϵ คือ ค่าคงที่ ๆ มีค่าน้อย

จากนั้น V ของทุกบล็อกจะถูกนำมาเรียงต่อกันเป็นเวกเตอร์คุณลักษณะหรือสามารถเรียกอีกอย่างว่าค่าฮิสโตแกรมของการไล่ระดับเชิงมุม

ในส่วนขั้นตอนสุดท้าย คือ Detector Window and Context เป็นขั้นตอนการปรับเปลี่ยนค่าพารามิเตอร์ต่าง ๆ อาทิเช่น จำนวนจุดภาพต่อเซลล์ จำนวนเซลล์ต่อบล็อก และอัตราการเคลื่อนที่เป็นต้น เพื่อให้ได้เวกเตอร์คุณลักษณะที่มีประสิทธิภาพ

จากที่กล่าวมาข้างต้น จะพบว่า HOG เป็นกระบวนการสกัดคุณลักษณะเชิงรูปร่างแบบหนึ่งที่มีประสิทธิภาพ อีกทั้งมีขนาดของคุณลักษณะที่เล็ก ถือได้ว่าเป็นข้อดีประการหนึ่งเมื่อเปรียบเทียบกับขนาดของคุณลักษณะที่ได้จากการสกัดคุณลักษณะด้วยกระบวนการแยกองค์ประกอบค่าเดียว ซึ่งในงานวิจัยนี้ได้นำ HOG มาปรับปรุงเพื่อให้ได้คุณลักษณะที่เหมาะสมกับภาพอักษรภาษาไทยที่มีความคล้ายคลึงกันมาก

2.3.2 การแบ่งแยกข้อมูล

การแบ่งแยกข้อมูล (Classification) คือกระบวนการที่นำคุณลักษณะที่สกัดได้มาใช้ในการแบ่งแยกข้อมูลออกเป็นกลุ่ม โดยข้อมูลที่อยู่ในกลุ่มเดียวกันจะมีคุณลักษณะที่ใกล้เคียงกัน และข้อมูลที่อยู่คนละกลุ่มกันจะมีคุณลักษณะที่แตกต่างกัน โดยการแบ่งแยกข้อมูลนั้นจะใช้ค่าความถูกต้องในการรู้จำในการชี้วัดตามสมการที่ 2.14

$$\text{ความถูกต้องของการรู้จำ} = \frac{\text{จำนวนตัวอักษรที่รู้จำถูกต้อง}}{\text{จำนวนตัวอักษรทั้งหมด}} \times 100 \quad (2.14)$$

โดยการแบ่งแยกกลุ่มข้อมูลนั้นสามารถทำได้หลายวิธี แต่ในงานวิจัยนี้จะเลือกใช้เฉพาะวิธีพื้นฐานในการแบ่งแยกกลุ่ม ซึ่งได้แก่ การฉายภาพเมตริกซ์ และระยะทางแบบยุคลิด

2.3.2.1 การฉายภาพเมตริกซ์

การวัดเชิงมุม คือหลักการวัดความคล้ายตามมุม ถ้ากำหนดให้ \vec{n} และ \vec{y} เป็นเวกเตอร์ใด ๆ แล้ว ถ้า \vec{n} และ \vec{y} เป็นเวกเตอร์ที่มีความคล้ายคลึงกันจะส่งผลให้มุมระหว่าง \vec{n} และ \vec{y} แคบหรือค่าโคไซน์จะมาก แต่ในทางกลับกันถ้า \vec{n} และ \vec{y} เป็นเวกเตอร์ที่มีความแตกต่างกันจะส่งผลให้มุมระหว่าง \vec{n} และ \vec{y} กว้างหรือส่งผลให้ค่าโคไซน์จะน้อยตาม โดยอาศัยการฉายภาพเมตริกซ์ (Projection Matrix) เป็นคณิตศาสตร์ที่ใช้ในการวัดเชิงมุมตามสมการที่ 2.15

$$S = (\vec{n}^T \vec{y})^2 \quad (2.15)$$

ผลที่ได้จากสมการที่ 2.15 เทียบเคียงได้ตามสมการที่ 2.16

$$\sqrt{S} = \cos \theta \quad (2.16)$$

เมื่อ \vec{n} , \vec{y} คือเวกเตอร์และ S คือ ค่าความคล้ายเชิงมุม หรือค่าโคไซน์

2.3.2.2 ระยะทางแบบยุคลิด

ระยะทางแบบยุคลิด (Euclidean Distance) เป็นการวัดระยะห่างระหว่าง \vec{x}_1 และ \vec{x}_2 ถ้า \vec{x}_1 และ \vec{x}_2 มีความคล้ายกัน ระยะทาง (d) ที่ได้จะมีค่าน้อย แต่ถ้า \vec{x}_1 และ \vec{x}_2 มีความแตกต่างกัน ระยะทางที่ได้จะมาก ซึ่งสามารถคำนวณได้จากสมการที่ 2.17

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$d = \sqrt{\sum_{i=1}^n (x_{2i} - x_{1i})^2} \quad (2.17)$$

เมื่อ x_{1i} คือ สมาชิกของ \bar{x}_1

x_{2i} คือ สมาชิกของ \bar{x}_2

n คือ ขนาดของเวกเตอร์ x

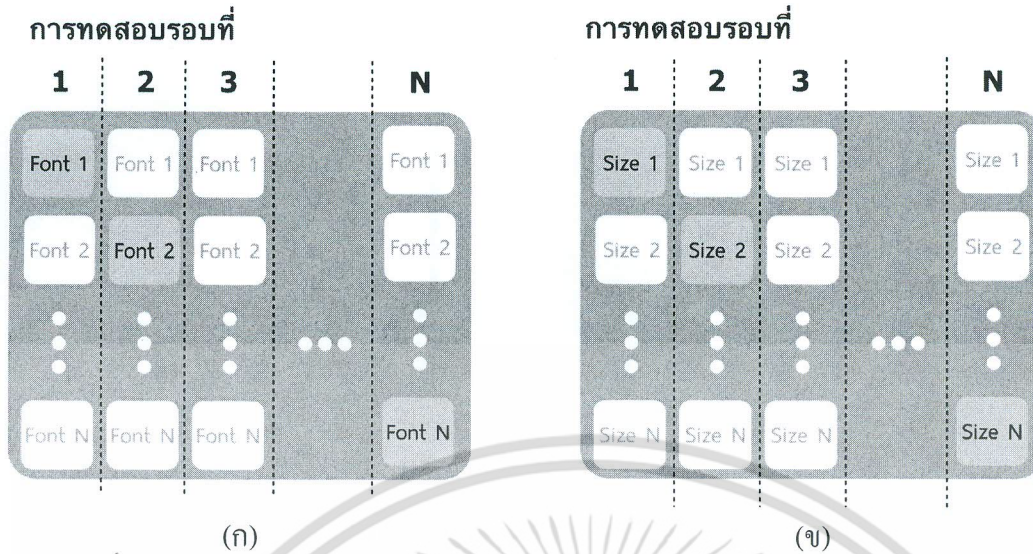
d คือ ค่าระยะทางแบบยุคลิด

การวัดระยะทางเชิงมุมด้วยการฉายภาพเมตริกซ์ และการวัดระยะทางแบบยุคลิดจะทำการวัดระยะห่างระหว่างตัวอธิบายคุณลักษณะทุกตัวเทียบกับคุณลักษณะที่ทดสอบ ถ้าการเทียบคุณลักษณะใดมีระยะห่างน้อยสุดถือว่าคุณลักษณะตัวทดสอบอยู่กลุ่มเดียวกับตัวอธิบายคุณลักษณะ ซึ่งทั้ง 2 กระบวนการถือว่าเป็นการแบ่งแยกข้อมูลแบบพื้นฐาน โดยงานวิจัยนี้เชื่อว่าถ้าคุณลักษณะที่สกัดได้จากภาพอักษรมีประสิทธิภาพ ถึงจะใช้การแบ่งแยกข้อมูลแบบพื้นฐานก็สามารถแบ่งแยกข้อมูลที่มีความคล้ายคลึงกันออกจากกัน

2.4 การทดสอบความทนทาน

การทดสอบความทนทาน (Robustness) คือ การทดสอบความคงทนของวิธีการทดสอบที่สามารถให้ผลการทดสอบที่มีความคลาดเคลื่อนเล็กน้อย โดยการเปลี่ยนแปลงเฉพาะปัจจัยที่ให้ความสนใจและสังเกตผลการเปลี่ยนแปลงที่เกิดขึ้น ในกลุ่มงานวิจัยการรู้จำวัตถุการทดสอบความคงทน หมายถึงการทดสอบความคงทนต่อสัญญาณรบกวนว่าวิธีการทดสอบสามารถจะรู้จำวัตถุได้ถูกต้องหรือไม่ แต่ในงานวิจัยนี้จะเป็นการทดสอบความคงทนต่อชุดแบบอักษร และขนาดอักษรที่เปลี่ยนแปลง เช่น รูปที่ 2.10 (ก) แสดงการแบ่งข้อมูลออกเป็น N ชุด ตามชุดแบบอักษร ได้แก่ ชุดแรกมีอักษรในชุดแบบอักษร AngsanaUPC เท่านั้น ชุดที่สองมีอักษรในชุดแบบอักษร BrowalliaUPC เท่านั้น เป็นต้น จากนั้นในขั้นตอนการทดสอบในรอบที่ 1 จะนำข้อมูลชุดที่ 2 ถึง N ไปสกัดคุณลักษณะเพื่อสร้างตัวอธิบายคุณลักษณะในขณะที่ข้อมูลชุดที่ 1 จะถูกใช้เป็นชุดทดสอบหรือสามารถกล่าวได้ว่าในรอบที่ x ชุดที่ x จะเป็นชุดทดสอบ และชุดที่เหลือจะนำมาสร้างตัวอธิบายคุณลักษณะ ในส่วนของการทดสอบความทนทานต่อขนาดของตัวอักษรก็เช่นเดียวกับการทดสอบความทนทานของชุดตัวอักษร ดังรูปที่ 2.10 (ข)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.10 การแบ่งข้อมูลสำหรับทดสอบความทนทาน (ก) การแบ่งข้อมูลออกตามชุดแบบอักษร และ (ข) การแบ่งข้อมูลออกตามขนาดอักษร

2.5 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้องของการรู้จำตัวอักษรภาษาไทยด้วยวิธีการต่าง ๆ ที่ใช้ในการสกัดคุณลักษณะ ซึ่งในแต่ละงานวิจัยได้มุ่งเน้นการแก้ปัญหาที่แตกต่างกัน เช่น การแก้ปัญหาชุดแบบอักษรที่ไม่มีหัว การแก้ไขปัญหาการขาดหายของภาพ การแก้ปัญหาคัญญารบกวนในภาพอักษร และการแก้ปัญหากล่อนภาพอักษร ซึ่งพอสรุปได้ดังนี้

Kimpan C. et.al. (1987: 257-264) ได้ศึกษาการรู้จำตัวอักษรภาษาไทยโดยอาศัย KL-expansion ในการสกัดคุณลักษณะและใช้ Linear Decision Function ในการจำแนกประเภท ซึ่งในงานวิจัยนี้มุ่งเน้นในการเพิ่มประสิทธิภาพของการรู้จำตัวอักษรภาษาไทย วิธีการนี้มีความถูกต้องถึง 100% แต่ต้องทำการจำแนกประเภทถึง 2 รอบ และในการจำแนกประเภทในรอบที่ 2 ได้มีการเลือกเฉพาะส่วนของตัวอักษรมาพิจารณา ซึ่งขึ้นกับกลุ่มของข้อมูล อีกทั้งในปี ค.ศ. 1987 ยังไม่มีความหลากหลายของชุดแบบอักษรเท่ากับในปัจจุบันนี้

Duangphasuk S. (2001) ได้ศึกษาการรู้จำตัวอักษรภาษาไทยโดยในการทดลองนี้มีการทดลอง 2 รอบ ในรอบแรกใช้คุณลักษณะต่อไปนี้ ขาเดี่ยว สองขา สามขา ขอบล่างปิด ขอบล่างเปิด ขอบบนเปิดและขอบบนปิด ส่วนรอบที่สองใช้ส่วนของขาหน้าคล้ายกัน ส่วนของขาหลังคล้ายกัน ส่วนบนของตัวอักษรที่คล้ายกัน และส่วนล่างของตัวอักษรที่คล้ายกัน ซึ่งทั้งสองรอบอาศัยโครงข่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสาทเทียมแบบ ART I ในการรู้จำตัวอักษร ซึ่งผลการทดลองมีความถูกต้อง 100% แต่ในงานวิจัยนี้มุ่งเน้นการแก้ปัญหาในตัวอักษร และสระภาษาไทยในรูปแบบปกติเท่านั้น ไม่มีการทดลองในกรณีของความหลากหลายของชุดแบบอักษร

Tangsurakit N. (2005) ได้ศึกษาการรู้จำตัวอักษรภาษาไทยโดยทำการจำแนกประเภทโดยอาศัยภาษาโนบิสติสแทนด์ ในการจำแนกตัวอักษรด้วยคุณลักษณะที่นำมาใช้ ได้แก่ ความหนาแน่น โดยการแบ่งตัวอักษรตามแนวนอน ความหนาแน่นโดยการแบ่งตัวอักษรตามแนวตั้ง ความหนาแน่นโดยการแบ่งตัวอักษรตามแนวทแยง ลักษณะเส้นแถบ และอัตราส่วนความกว้างต่อความสูง ซึ่งผลการทดลองด้วยวิธีนี้มีความถูกต้อง 99.62 % เมื่อทดสอบการรู้จำกับ 5 ชุดแบบอักษรพื้นฐาน ซึ่งประกอบด้วย AngsanaUPC , BrowalliaUPC , Cordia New , Dillenia UPC และ Eucrosia UPC มีขนาด 14 , 16 , 18 , 20 , 22 พอยต์ อีกทั้งมีขอบเขตการวิจัยกับเพียงพยัญชนะภาษาไทยเท่านั้น

Kruatrachue B. et.al. (2007: 153-158) ได้นำเสนอการรู้จำตัวอักษรภาษาไทยโดยการสร้างแบบจำลองในรูปแบบเครื่องสถานะจำกัดแบบอัตโนมัติโดยอาศัย Deterministic Finite Automata (DFA) และรับภาพอักษรเข้ามาผ่านกระบวนการรหัสโซ่ (Chain Code) โดยมีประสิทธิภาพการรู้จำในภาพรวมค่อนข้างสูง แต่ถ้าพิจารณาในรายละเอียดจะพบว่า มีบางตัวอักษรที่มีความคล้ายคลึงกัน จะมีประสิทธิภาพการรู้จำค่อนข้างต่ำ เช่น “ก-ถ”, “ข-ช”, “ค-ศ” เป็นต้น ดังนั้น จะเห็นว่าชุดแบบอักษรที่มีความคล้ายคลึงกันนั้นจะส่งผลกระทบต่อประสิทธิภาพการรู้จำของงานวิจัย ซึ่งในงานวิจัยนี้มุ่งเน้นแก้ปัญหาในตัวอักษรภาษาไทย สระ วรรณยุกต์ และเลขอารบิก

Thammano and Duangphasuk (2005: 695-698) ได้นำเสนอการจำแนกประเภทโดยอาศัย Hierarchical Cross-correlation ARTMAP โดยสกัดคุณลักษณะจาก Eight Directional Codes ซึ่งในงานวิจัยนี้มุ่งเน้นในการแก้ไขปัญหาประสิทธิภาพการรู้จำของชุดแบบอักษรที่ไม่มีหัวทำให้มีความถูกต้องเฉลี่ย 83.77% โดยมุ่งเน้นในการแก้ปัญหาคู่ชุดแบบอักษรที่มีความหลากหลาย ทั้งมีหัว และไม่มีหัว ซึ่งชุดแบบอักษรที่ไม่มีหัวก็เป็นหนึ่งในความหลากหลายของชุดแบบอักษรเช่นกัน

Tanprasert and Sae-Tang (1999: 336-339) ได้ศึกษาการรู้จำตัวอักษรภาษาไทยโดยมุ่งเน้นการแก้ปัญหาที่เกิดจากการสแกนภาพอักษรเข้าสู่คอมพิวเตอร์ ได้แก่ การขาดหายของภาพอักษร การเลื่อนตำแหน่งของอักษร และสัญญาณรบกวนในภาพอักษร เป็นต้น โดยได้นำเสนอวิธีการแก้ปัญหาดังนี้ นำภาพเอกสารมาผ่านกระบวนการ Preprocessing ประกอบด้วย Page Alignment,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Segmentation, Normalization, Thinning จากนั้น Artificial Neural Networks (ANNs) จะถูกใช้สำหรับการจำแนกประเภท และ Post processing ซึ่งประกอบด้วย Thai Dictionary และ Rules of Combining Thai Words ซึ่งผลการทดลองมีความถูกต้องโดยเฉลี่ย 90 - 95%

จากงานวิจัยข้างต้นผู้วิจัยพบว่ามิงงานวิจัยของ Tangsurakit N. (2005) ที่ได้กล่าวถึงความคล้ายคลึงของตัวอักษรภาษาไทย แต่ไม่มีงานวิจัยใดที่ให้ความสำคัญกับทั้ง 2 ปัจจัยข้างต้น กล่าวคือ ความหลากหลายของขนาด ชุดแบบอักษร และชนิดของอักษร ดังนั้น ผู้วิจัยจึงได้นำเสนองานวิจัย Printed Thai Character Recognition Using Standard Descriptor โดย Woraratpanya and Titijaroonroj (2013: 165-173) เพื่อแก้ปัญหาข้างต้น ซึ่งรายละเอียดจะกล่าวไว้ในบทที่ 3 จากผลลัพธ์ที่ได้จากงานวิจัย Woraratpanya and Titijaroonroj (2013: 165-173) นั้นได้แสดงให้เห็นถึงประสิทธิภาพที่ดีในกรณีที่ไม่มีสัญญาณรบกวนในชุดข้อมูล แต่ก็ยังมีส่วนที่ต้องปรับปรุง คือ ขนาดคุณลักษณะที่มีขนาดใหญ่ ซึ่งงานวิจัย Woraratpanya and Titijaroonroj (2013: 165-173) ได้สกัดคุณลักษณะโดยเลือกไอเจนเวกเตอร์ในลำดับปลาย หรืออีกนัยหนึ่งคือการพิจารณาความถี่สูงหรือเป็นคุณลักษณะของขอบตัวอักษรหรือรูปร่าง (Shape) ซึ่งผู้วิจัยได้พบว่า คุณลักษณะเชิงรูปร่างเป็นคุณลักษณะที่เหมาะสมกับงานทางด้านความรู้จำตัวอักษรภาษาไทย หรือสามารถกล่าวได้อีกนัยหนึ่งว่าเป็นคุณลักษณะที่สามารถแบ่งแยกภาพอักษรที่มีความคล้ายคลึงกันได้ ดังนั้น ผู้วิจัยจึงได้ทำการทบทวนวรรณกรรมที่เกี่ยวกับกระบวนการสกัดคุณลักษณะเชิงรูปร่างดังนี้

Navneet and Bill (2005: 1-8) ได้ศึกษา Histogram of Oriented Gradient for Human Detection เพื่อใช้ในงานทางด้านการระบุภาพว่าเป็นภาพบุคคลหรือไม่ โดยกระบวนการ Histogram of Oriented Gradient หรือ HOG เป็นกระบวนการหนึ่งในกลุ่มของการสกัดคุณลักษณะเชิงรูปร่างที่มีประสิทธิภาพดี ซึ่ง HOG นั้นจะประกอบไปด้วย 5 ขั้นตอน ได้แก่ Gamma an Color Normalization, Gradient Computation, Spatial and Orientation Binning, Normalization and Descriptor Blocks และ Detector Window and Context โดยที่คุณลักษณะที่ได้ คือค่าฮิสโตแกรมของการไล่ระดับเชิงมุมและใช้เวกเตอร์สนับสนุนการตัดสินใจสำหรับการแบ่งแยกข้อมูล จากผลการทดลองพบว่า HOG นั้นมีประสิทธิภาพที่ดีเมื่อเทียบกับ Scale-Invariant Feature Transform (SIFT) แต่เมื่อนำ HOG มาสกัดคุณลักษณะจากภาพอักษรภาษาไทยพบว่าคุณลักษณะที่ได้นั้น ไม่สามารถแบ่งแยกประเภทของกลุ่มข้อมูลที่มีความคล้ายคลึงกันได้ ซึ่งเหตุผลหลักมาจาก HOG ถูกออกแบบมาให้เหมาะสมการรู้จำโดยอาศัยรูปร่างแบบหยาบ เพราะเนื่องจากภาพบุคคลนั้นมีความยืดหยุ่นและเป็นวัตถุที่มีข้อต่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในปี ค.ศ. 2008 ได้มีการนำเสนอ Scale space Histogram of Oriented Gradient for Human Detection โดย Ning, H.et.al. (2008: 167-170) Scale Space Histogram of Oriented Gradient : SSHOG ถูกใช้สำหรับการสกัดคุณลักษณะและถูกพัฒนามาจาก HOG ในการระบุว่าภาพบุคคลหรือไม่นั้นเป็นการระบุที่ยากกว่ารถยนต์ และใบหน้าคน เพราะร่างกายคนเรานั้นมีส่วนที่เชื่อมต่อกันค่อนข้างมาก อีกทั้งยังมีความยืดหยุ่นของรูปร่างค่อนข้างสูง ดังนั้น SSHOG จึงมีการปรับปรุง HOG โดยการนำ Gaussian kernel เข้ามาใช้งาน โดยมีความคิดว่าถ้าภาพยิ่งถูกทำให้เกลี่ยให้เรียบ (smooth) หรือค่าสเกลของซิกม่า (σ) จะทำให้ส่วนต่าง ๆ ของคนเรา เช่น แขน, ขา เป็นต้น ถูกมองรวมกันเป็นวัตถุชิ้นเดียวกันได้ และใช้เวกเตอร์สนับสนุนการตัดสินใจสำหรับการแบ่งกลุ่มข้อมูล ซึ่งผลการทดลองแสดงว่า Miss Rate และ FFPW ของ SSHOG น้อยกว่า HOG ซึ่งบ่งบอกถึง SSHOG นั้นประสิทธิภาพที่ดีกว่า HOG แต่การเลือกใช้สเกลของซิกม่าที่หลากหลายนั้นก็ยังเป็นแบบตายตัว คือทุกภาพต้องใช้สเกลของซิกม่าในช่วงที่กำหนด ซึ่งถ้าอยากได้คุณลักษณะที่ดีควรจะใช้สเกลของซิกม่าให้เหมาะสมกับภาพที่ต้องการสกัดคุณลักษณะมากกว่าการใช้ทุกสเกลของซิกม่า หรือควรเลือกใช้เฉพาะสเกลของซิกม่าที่ส่งผลดีกับการสกัดคุณลักษณะของภาพนั้น อีกทั้งในสเกลของซิกม่าที่สูงขึ้นสำหรับภาพบางประเภทอาจจะส่งผลทำให้รูปร่างที่ได้มีความไม่ชัดเจน หรือผิดเพี้ยนไปจากความเป็นจริงได้ซึ่งส่งผลต่อคุณลักษณะที่สกัดได้เช่นกัน

Representing Shape with a Spatial Pyramid Kernel ได้ถูกนำเสนอ โดย Anna, B.et.al. (2007) อาศัย Pyramid Histogram of Oriented Gradient : PHOG เป็นกระบวนการสำหรับการสกัดคุณลักษณะ และถูกพัฒนามาจาก HOG โดยที่ PHOG มีแนวความคิดเรื่องขนาดของเซลล์ หรือเรียกว่า “ความคมชัด” (Resolution) เซลล์ที่มีขนาดแตกต่างกันก็คือความคมชัดที่แตกต่างกัน ใน HOG ได้เลือกใช้เซลล์หนึ่งขนาดเท่านั้น หรือเพียงความคมชัดเดียว ซึ่ง PHOG นั้นคิดว่าในความคมชัดที่ต่างกันจะพบจุดที่สำคัญนั้นแตกต่างกัน และใช้เวกเตอร์สนับสนุนการตัดสินใจในการแบ่งแยกกลุ่มของวัตถุ PHOG ถูกนำประยุกต์ใช้งานกับการรู้จำวัตถุ จากผลการทดลองเมื่อเปรียบเทียบกับวิธีการอื่นพบว่า PHOG มีประสิทธิภาพที่ดีกว่าวิธีการอื่นโดยเฉลี่ย 10 – 20% หรือได้ค่าความถูกต้อง 77.8% ซึ่งถือว่าคุณลักษณะที่สกัดจาก PHOG นั้นมีประสิทธิภาพในระดับหนึ่ง อีกทั้งจำนวนคุณลักษณะที่ได้น้อยกว่า HOG มากโดยในประเด็นนี้ถือว่าเป็นข้อดีเพราะใน PHOG นั้นไม่มีการคาบเกี่ยวหรือซ้อนทับ (Overlap) ระหว่างบล็อกเหมือน HOG แต่การใช้ความคมชัดก็ยังเป็นแบบตายตัว เพราะทุกภาพต้องใช้ความคมชัดในช่วงที่กำหนดเหมือนกัน โดยการที่จะได้เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะที่ดีนั้นควรเลือกใช้ความคมชัดที่เหมาะสมกับภาพเท่านั้น ไม่ควรเลือกใช้ทุกความคมชัด เพราะจะส่งผลต่อคุณลักษณะที่สกัดได้

Classification using Intersection Kernel Support Vector Machines is Efficient ได้ถูกนำเสนอโดย Subhransu, M.et.al. (2008) ซึ่งใช้ Multilevel Histogram of Oriented Gradient : MHOG ในกระบวนการสกัดคุณลักษณะ MHOG ได้นำกาบอร์ฟิลเตอร์ (Gabor Filter) มาใช้งานร่วมกับ PHOG โดยที่ MHOG นั้นมีแนวความคิดว่าในแต่ละทิศทางจะมีตำแหน่งที่สำคัญแตกต่างกันไปตามทิศทางและความคมชัดเงิน โดยทั่วไปตัวกรองกาบอร์เป็นกระบวนการที่พัฒนามาจาก Gaussian Kernel โดยการปรับปรุงให้ภาพที่ผ่านตัวกรองกาบอร์ได้ขอบที่มีความชัดเจนในทิศทางที่กำหนด ส่วนทิศทางที่ไม่กำหนดจะมีลักษณะที่เบลอ และใช้เวกเตอร์สนับสนุนการตัดสินใจมาใช้ในการแบ่งแยกข้อมูลออกเป็นกลุ่ม จุดที่น่าสนใจของ MHOG จะคล้ายคลึงกับ PHOG เพราะ MHOG ได้หยิบ PHOG มาใช้งานต่อจากกระบวนการทำงานของตัวกรองกาบอร์

ในปี 2011 ได้มีการนำเสนอ Multiscale Histogram of Oriented Gradient Descriptors for Robust Character Recognition โดย Andrew and Lewis (2011: 1085-1089) ซึ่งงานวิจัยนี้มุ่งเน้นการรู้จำตัวอักษรภาษาอังกฤษและเลขเอริกจากป้ายตามธรรมชาติ Histogram of Oriented Gradient Column หรือ HOGC ถูกใช้สำหรับกระบวนการสกัดคุณลักษณะ และถูกพัฒนามาจาก SSHOG โดยมีแนวความคิดว่า HOG ดั้งเดิมได้ใช้สเกลเพียงค่าเดียว หรือ $\sigma = 0$ และ SSHOG ก็ได้มีการปรับปรุงให้ใช้งานหลากหลายสเกลของซิกม่า แต่ HOGC นั้นได้มีการปรับปรุงแนวความคิดเพิ่มเติมอีกว่าสเกลของภาพนั้นควรมีอยู่ 2 แบบ คือ 1) สเกลหยาบ และ 2) สเกลละเอียดซึ่งสเกลละเอียด (Base Scale) คือ σ_{Base} และสเกลหยาบ (Coarser Scale) คือ $R\sigma_{Base}$ โดยที่ R คือ ค่าของอัตราเปรียบเทียบ (Scale Ratio) มีค่าเริ่มจาก 1,2,3,...,n และใช้ Bhattacharyya Distance สำหรับแบ่งแยกกลุ่มตัวอักษร จากผลการทดลองพบว่า HOGC นั้นมีค่าความถูกต้องมากกว่า HOG อยู่โดยเฉลี่ย 10% และประเด็นที่น่าสนใจจึงไม่ต่างจาก SSHOG คือ สเกลของซิกม่าที่สูงขึ้นสำหรับภาพบางประเภทอาจจะส่งผลทำให้รูปร่างที่ได้มีความไม่ชัดเจน หรือผิดเพี้ยน เป็นต้น จะเห็นว่าข้อจำกัดด้านจำนวนสมาชิกกลุ่มของข้อมูล (ตัวอักษรภาษาไทย สระ วรรณยุกต์ และตัวเลขไทย) จำนวนชุดแบบอักษร ขนาด และรูปแบบที่แตกต่างกันของตัวอักษรส่งผลต่อประสิทธิภาพการรู้จำตัวอักษรภาษาไทย หรือกระบวนการที่มีความซับซ้อนค่อนข้างสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากงานวิจัยในกลุ่มของฮิสโตแกรมของการไล่ระดับเชิงมุม (Navneet and Bill 2005 ; Ning, H.et.al. 2008 ; Anna, B.et.al. 2007 ; Subhransu, M.et.al. 2008 ; Andrew and Lewis 2011) จะพบว่า HOG เป็นกระบวนการสกัดคุณลักษณะที่เป็นที่นิยม ซึ่งจะพบว่า HOG นั้นได้ถูกพัฒนาอย่างต่อเนื่องตั้งแต่ ค.ศ. 2005 จนถึงปัจจุบัน ผู้วิจัยพบว่าคุณลักษณะเชิงรูปร่างเป็นคุณลักษณะที่สามารถแบ่งแยกภาพอักษรที่มีความคล้ายคลึงกันได้แต่เมื่อนำคุณลักษณะเชิงรูปร่างที่ได้รับจาก HOG มาใช้งานกับตัวอักษรภาษาไทยพบว่า ยังไม่สามารถแบ่งแยกภาพตัวอักษรที่มีความคล้ายคลึงกันได้ ดังนั้น ในงานวิจัยนี้จึงเสนอแนวทางการแก้ปัญหาโดยการสร้างตัวกลางหรือตัวอธิบายคุณลักษณะของภาพตัวอักษรที่ไม่ขึ้นอยู่กับขนาดและชุดแบบอักษรที่แตกต่างกัน โดยการสกัดคุณลักษณะจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ หรือ AHOG จากงานวิจัย Woraratpanya and Titijaroonroj (2014: 83-92) จะกล่าวถึงในบทที่ 3



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การดำเนินงานวิจัย

3.1 ปัญหาของการรู้จำตัวอักษรภาษาไทย

ความหลากหลายของชุดแบบตัวอักษร ชนิดของตัวอักษร และขนาดของอักษร อีกทั้งความคล้ายคลึงกันของอักษรภาษาไทย ล้วนเป็นปัจจัยที่ส่งผลกระทบต่อประสิทธิภาพการรู้จำตัวอักษรภาษาไทย ถ้าประสิทธิภาพของการรู้จำตัวอักษรถูกพัฒนาให้ดีขึ้น ย่อมส่งผลให้การประยุกต์ใช้งานของนำการรู้จำตัวอักษรภาษาไทยมีคุณภาพมากยิ่งขึ้นด้วย จากสมมติฐานของผู้วิจัยถ้าคุณลักษณะที่สกัดจากภาพอักษรมีประสิทธิภาพแล้ว ถึงจะใช้กระบวนการแบ่งแยกข้อมูลพื้นฐานก็สามารถแบ่งแยกข้อมูลที่มีความคล้ายคลึงกันได้ อีกทั้งจากการวิจัยพบว่าคุณลักษณะที่เหมาะสมกับภาพตัวอักษรคือคุณลักษณะเชิงรูปร่างหรือขอบของตัวอักษร ซึ่งเป็นแนวทางในการดำเนินงานวิจัยนี้

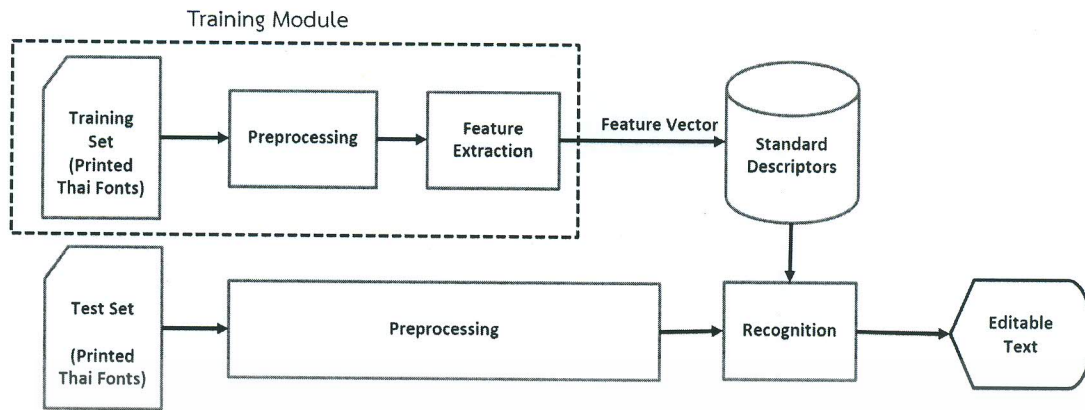
3.2 แนวทางการดำเนินการวิจัย

ในบทนี้จะกล่าวถึงขั้นตอนการรู้จำตัวอักษรภาษาไทยตั้งแต่การเตรียมข้อมูลเบื้องต้นเพื่อปรับข้อมูลให้อยู่ในมาตรฐานเดียวกัน การสกัดคุณลักษณะ และการแบ่งแยกกลุ่มของตัวอักษร ซึ่งในงานวิจัยนี้ได้นำเสนอ 2 แนวทาง คือ การแยกองค์ประกอบค่าเดียวสำหรับการสกัดคุณลักษณะ และใช้การวัดเชิงมุมด้วยการฉายภาพเมตริกซ์ และการสกัดคุณลักษณะโดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและใช้การวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษร

3.2.1 การแยกองค์ประกอบค่าเดียวสำหรับการสกัดคุณลักษณะและใช้การวัดเชิงมุมด้วยการฉายภาพเมตริกซ์

การแยกองค์ประกอบค่าเดียวได้ถูกประยุกต์ใช้สำหรับการสกัดคุณลักษณะ โดยเริ่มจากการนำภาพที่ผ่านขั้นตอนการเตรียมข้อมูลแล้วมาสกัดคุณลักษณะและผลลัพธ์ที่ได้คือ ไอเกนเวกเตอร์ โดย Suttapakti, U. et.al. (2013: 45-48) ได้กล่าวว่า “การเลือกไอเกนเวกเตอร์จะส่งผลกระทบต่อประสิทธิภาพการรู้จำ” ซึ่งกระบวนการรู้จำด้วยวิธีการดังกล่าวสามารถแสดงได้ดังรูปที่ 3.1

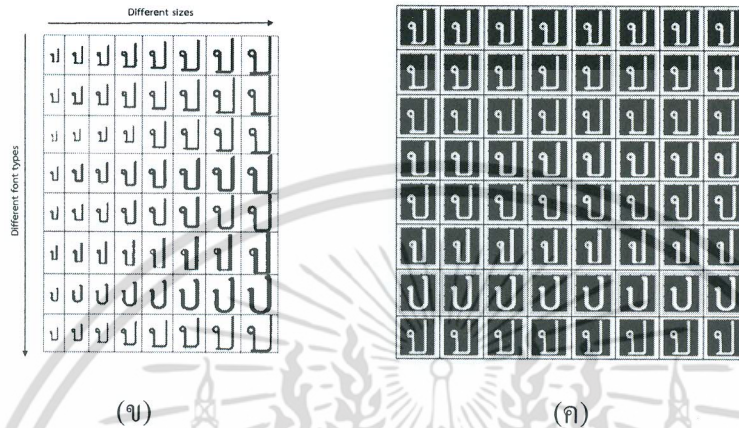
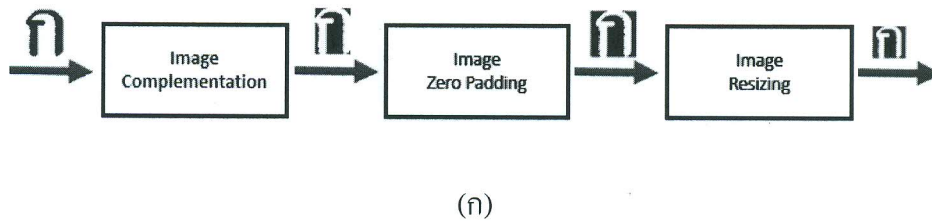
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 กระบวนการรู้จำด้วยการแยกองค์ประกอบค่าเดียวสำหรับการสกัดคุณลักษณะและการวัดเชิงมุมด้วยการฉายเมตริกซ์

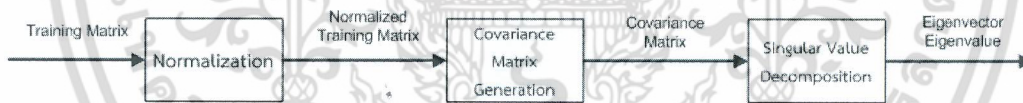
3.2.1.1 ขั้นตอนการเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูล (Preprocessing Procedure) เป็นการเปลี่ยนภาพอักษรให้อยู่ในมาตรฐานเดียวกันดังรูปที่ 3.2(ก) โดยเริ่มจากการนำภาพตัวอักษรที่มีลักษณะดังรูปที่ 3.2(ข) ที่มีขนาดและชุดแบบของตัวอักษร แตกต่างกันมาทำการกลับสีภาพ (Image Complementation) เนื่องจากภาพตัวอักษรเป็นสีดำซึ่งมีค่าเป็น “0” จึงทำการกลับให้เป็น “1” เพื่อให้บริเวณตัวอักษรมีค่าเป็น “1” ซึ่งสามารถนำมาใช้ในการคำนวณ จากนั้นทำการเติมศูนย์ในบริเวณภาพโดยรอบของอักษร (Zero Padding) เพื่อปรับให้ภาพมีอัตราส่วนเป็นจัตุรัส เช่น ตัวอักษร “ป” ส่วนสูงจะมีค่ามากกว่าความกว้าง จะเติมศูนย์ที่ด้านซ้าย และด้านขวาของ “ป” ให้เท่ากัน แต่ในกรณี “ญ” ความกว้างจะมีค่ามากกว่าความสูง ดังนั้น จะเติมศูนย์ในด้านบน และด้านล่างของ “ญ” ให้เท่ากัน ในกรณีที่ไม่สามารถเติมศูนย์ให้เท่ากันทั้งสองด้านได้ จะยึดด้านซ้ายหรือด้านบนเป็นหลักคือให้เติมศูนย์มากกว่าอีกด้านหนึ่งแถวหรือหนึ่งคอลัมน์ จากนั้นปรับขนาดของภาพอักษรให้มีขนาดเป็นมาตรฐานเดียวกัน แต่ยังคงอัตราส่วนไว้ เพื่อช่วยในการแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกัน โดยนำไปใช้กับทุกภาพอักษร ซึ่งผลลัพธ์ที่ได้ คือ ภาพตัวอักษรที่มีขนาด 32 x 32 พิกเซล และพื้นหลังสีดำ (Zero Background) ดังรูปที่ 3.2(ค) ซึ่งแสดงภาพตัวอักษรที่ผ่านขั้นตอนการเตรียมข้อมูล



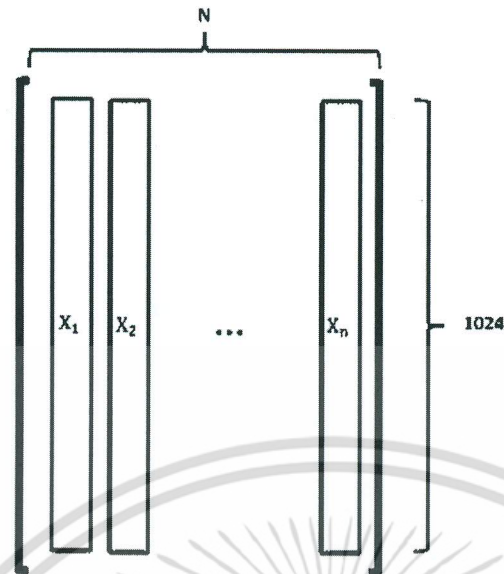
รูปที่ 3.2 การเตรียมข้อมูล (ก) ขั้นตอนการทำงานของขั้นตอนการเตรียมข้อมูล (ข) ภาพตัวอักษรที่มีชุดแบบอักษรแตกต่างกัน และขนาด (ค) ภาพอักษรที่อยู่ในมาตรฐานเดียวกัน

3.2.1.2 ขั้นตอนสกัดตัวอธิบายคุณลักษณะ



รูปที่ 3.3 ขั้นตอนการสกัดคุณลักษณะ

ภาพรวมของการสกัดตัวอธิบายคุณลักษณะ (Feature Extraction) ดังแสดงในรูปที่ 3.3 โดยเริ่มจากการนำภาพตัวอักษรที่อยู่ในรูปแบบมาตรฐานดังรูปที่ 3.2(ค) มาแปลงให้อยู่ในรูปแบบเวกเตอร์ X ดังรูปที่ 3.4 นำเวกเตอร์ของภาพตัวอักษรที่อยู่ในกลุ่มเดียวกัน แต่มีขนาด ชุดแบบอักษร และชนิดที่แตกต่างกันมาเรียงต่อกันตามแนวนอน ซึ่งผลลัพธ์ที่ได้จะเรียกว่า เมตริกซ์ฝึกสอน (Training Matrix)



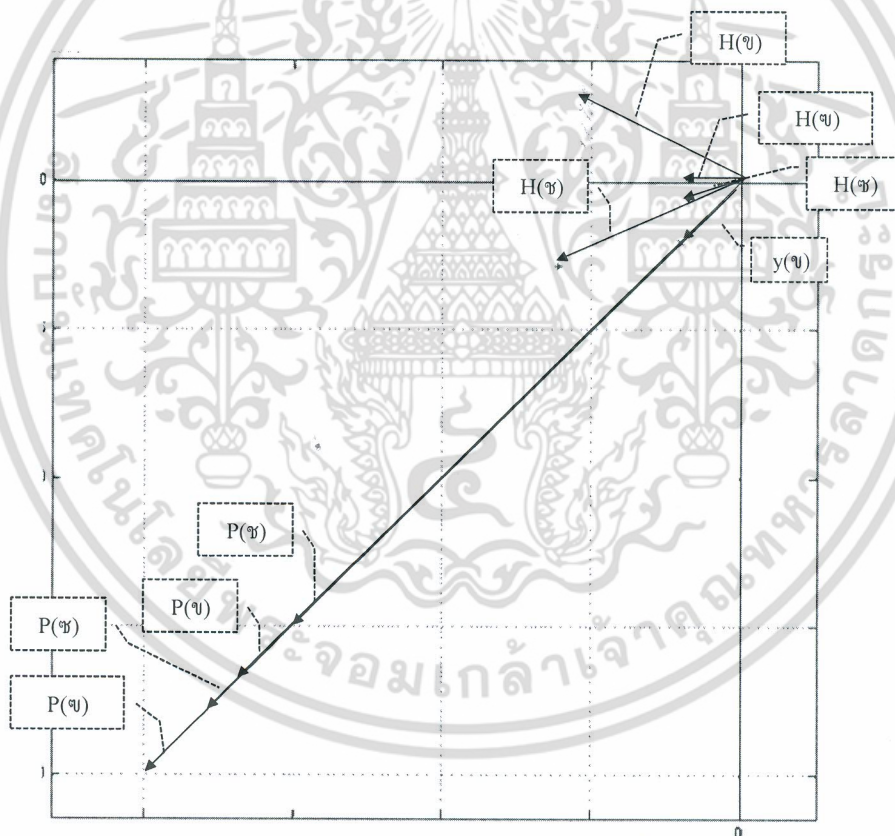
รูปที่ 3.4 เมตริกซ์ฝึกสอนซึ่งสร้างจากภาพตัวอักษรจากชุดฝึกสอน

โดยขั้นตอนการสร้างตัวอธิบายคุณลักษณะนั้น เริ่มจากการนำเมตริกซ์ฝึกสอนมาปรับให้อยู่ในมาตรฐานเดียวกัน (Normalized) ตามสมการที่ 2.8 จากนั้นสร้างเมตริกซ์ความแปรปรวน (Covariance Matrix) โดยการนำเมตริกซ์ฝึกสอนที่ปรับให้อยู่ในมาตรฐานเดียวกันไปคำนวณค่าความแปรปรวนตามสมการที่ 2.9 ในขั้นตอนสุดท้ายของการสกัดหาตัวอธิบายคุณลักษณะจากเมตริกซ์ความแปรปรวน โดยอาศัยการแยกองค์ประกอบค่าเดียว ซึ่งผลลัพธ์ที่ได้จะเป็นไปตามสมการที่ 2.2 ซึ่ง U , V และ S คือ ไอเกนเวกเตอร์ทางซ้าย ไอเกนเวกเตอร์ทางขวา และค่าเอกฐานตามลำดับดังรูปที่ 2.5 โดยที่ตัวอธิบายคุณลักษณะจะถูกสร้างจากการเลือกไอเกนเวกเตอร์ใน “ลำดับปลาย” คือ ลำดับไอเกนเวกเตอร์ที่ 512 ถึง 704

3.2.1.3 ขั้นตอนการรู้จำภาพตัวอักษรด้วยการฉายภาพเมตริกซ์

การฉายภาพเมตริกซ์ (Projection Matrix) เป็นเทคนิคที่ถูกนำมาประยุกต์ใช้งานในขั้นตอนการรู้จำ โดยการฉายภาพเมตริกซ์นั้น มีลักษณะเด่นทางด้านความคล้ายเชิงมุม การนำการฉายภาพเมตริกซ์มาใช้ในการวัดความคล้ายเชิงมุมของตัวอธิบายคุณลักษณะกับชุดข้อมูลทดสอบ ซึ่งโดยทั่วไปการรู้จำที่ผิดพลาดจะเกิดขึ้นในกรณีที่ตัวอักษรมีความคล้ายคลึงกัน จากรูปที่ 3.5 ได้แสดงให้เห็นถึงความคล้ายคลึงกันของคุณลักษณะของ “ข ข ช ซ” ถ้าพิจารณาเลือกไอเกนเวกเตอร์ใน “ลำดับต้น” (SD1) ของ 4 ตัวอักษรข้างต้นซึ่งถูกแทนด้วย $P(x)$ $P(y)$ $P(z)$ และ $P(w)$ แต่ในทางกลับกัน ถ้าพิจารณาเลือกไอเกนเวกเตอร์ใน “ลำดับปลาย” (SD2) ของ 4 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอักษรจะนำเสนอในรูปแบบของ $H(\psi)$ $H(\varphi)$ $H(\chi)$ และ $H(\zeta)$ จะเห็นได้อย่างชัดเจนว่าไอเกนเวกเตอร์ใน “ลำดับต้น” ของทั้ง 4 ตัวอักษรนั้น มีทิศทางไปในทางเดียวกัน แต่ขณะที่ไอเกนเวกเตอร์ใน “ลำดับปลาย” มีทิศทางที่แตกต่างกันอย่างชัดเจน และในตัวอักษรอื่น ๆ ก็เช่นกันที่ไอเกนเวกเตอร์ใน “ลำดับปลาย” สามารถที่จะแบ่งแยกได้อย่างชัดเจน ในที่นี้ ถ้าเวกเตอร์ทดสอบถูกเขียนแทนด้วย $y(\psi)$ ดังรูปที่ 3.5 และการวัดความคล้ายเชิงมุมระหว่าง Y กับ P ในกรณีนี้การวัดความคล้ายเชิงมุมจะให้ค่าที่สูงสุด หรือให้มุมที่เล็กที่สุดซึ่งจากรูปที่ 3.5 อาจจะทำให้เกิดการรู้จำผิดพลาดได้ แต่ในทางกลับกัน เมื่อนำการวัดความคล้ายเชิงมุมระหว่าง Y กับ H จะให้ค่าความคล้ายเชิงมุมน้อย หรือมุมที่ได้รับจะกว้าง ซึ่งสามารถนำมาใช้แบ่งแยกตัวอักษรภาษาไทยได้และส่งผลให้ค่าความถูกต้องสูงขึ้นตาม



รูปที่ 3.5 การเปรียบเทียบไอเกนเวกเตอร์ “ลำดับต้น” กับ “ลำดับปลาย”

ดังนั้น จึงสรุปได้ว่าไอเกนเวกเตอร์ใน “ลำดับปลาย” สามารถใช้เป็นคุณลักษณะที่ดีเพื่อแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันได้ดี เพราะฉะนั้น ในงานวิจัยนี้จะนำเสนอการสร้างตัวเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อธิบายคุณลักษณะบนพื้นฐานของไอเคนเวกเตอร์ใน “ลำดับปลาย” และใช้การฉายภาพเมตริกซ์ในขั้นตอนการแบ่งแยกภาพตัวอักษร การเลือกใช้ไอเคนเวกเตอร์ใน “ลำดับปลาย” สามารถกล่าวได้ว่าเป็นการสกัดคุณลักษณะแบบเลือกความถี่สูง (แต่ไม่ใช่ความถี่สูงสุด) หรือบริเวณขอบของภาพตัวอักษรนั่นเอง

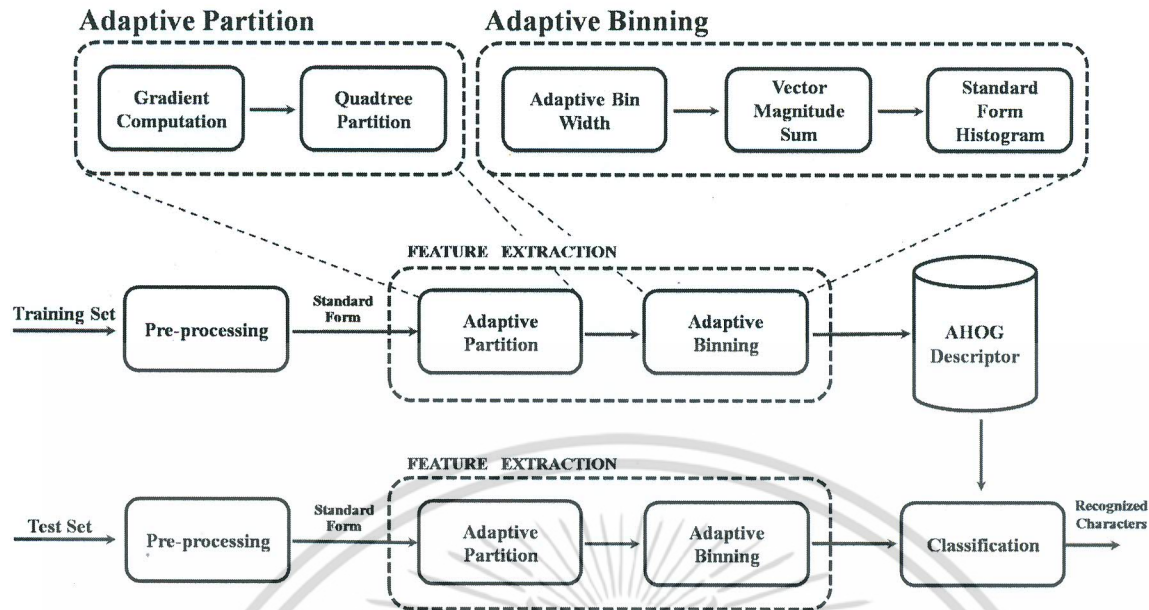
จะเห็นได้ว่าการแยกองค์ประกอบค่าเดียวสำหรับการสกัดคุณลักษณะและใช้การวัดเชิงมุมด้วยการฉายภาพเมตริกซ์นั้น มีประสิทธิภาพสามารถแบ่งแยกกลุ่มของตัวอักษรได้ชัดเจน ดังที่แสดงในบทที่ 4 แต่มีขนาดของคุณลักษณะเท่ากับ 1024×193 หรือ 197,632 จากปัญหานี้ จึงจำเป็นต้องลดขนาดของคุณลักษณะลง เพื่อลดพื้นที่ในการจัดเก็บ อีกทั้งถ้าคุณลักษณะมีขนาดใหญ่จะส่งผลต่อเวลาในการประมวลผลการจำแนกประเภทข้อมูลอีกด้วย

ในหัวข้อถัดไปจะกล่าวถึงการสกัดคุณลักษณะโดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ ซึ่งถูกพัฒนามาจากฮิสโตแกรมของการไล่ระดับเชิงมุม (HOG) และใช้การวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษรเพื่อให้ได้ขนาดของคุณลักษณะที่น้อยและใช้เวลาในการจำแนกประเภทข้อมูลลดลง แต่ยังคงความถูกต้องสูง

3.2.2 การสกัดคุณลักษณะโดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและการวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษร

การรู้จำตัวอักษรจัดเป็นการรู้จำวัตถุรูปแบบหนึ่ง ซึ่งคุณลักษณะเชิงรูปร่างมีความสำคัญมาก สำหรับตัวอักษรภาษาไทยจากที่กล่าวไว้ในงานวิจัยในหัวข้อ 2.4 อีกทั้งภาพตัวอักษรก็มีลักษณะรูปร่างที่แน่นอน และไม่ยืดหยุ่นเท่ากับวัตถุในรูปแบบอื่น เช่น รถยนต์ หรือคน แต่ภาพตัวอักษรมีความคล้ายคลึงกันสูง จึงจำเป็นต้องพิจารณาความต่างอย่างละเอียดหรือเฉพาะที่ โดยทั่วไปรูปทรงของวัตถุจะถูกกำหนดด้วยชั้นของขอบภาพ ดังนั้น เมื่อพิจารณาที่ขอบภาพก็มีความหมายว่า เรากำลังพิจารณาที่ตัวรูปทรงของวัตถุนั้นเอง ซึ่งหลักการทำงานของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะจะเริ่มจากการหาขอบภาพ หลังจากนั้น จะคำนวณหาขนาดและทิศทางในแต่ละพิกเซล และจัดเก็บพลังงานในรูปของดังตามส่วนย่อยของภาพและขั้นตอนสุดท้าย คือ การนำค่าทุกถ่วงมาเรียงต่อกัน ดังรูปที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.6 กระบวนการรู้จำด้วยการสกัดคุณลักษณะ โดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและการวัดระยะทางแบบยุคลิดสำหรับการแบ่งแยกกลุ่มของตัวอักษร

3.2.2.1 ขั้นตอนการเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูล (Preprocessing) ในวิธีการที่สองนี้ ไม่มีความแตกต่างจากการเตรียมข้อมูลในวิธีการแรก ซึ่งประกอบด้วยกระบวนการตัดแบ่งส่วนภาพ (Image Segmentation) การกลับสีภาพ (Image Complementation) การเติมศูนย์เข้าไปในภาพอักษร (Zero Padding) และปรับขนาดของภาพอักษรให้มีขนาดเป็นมาตรฐานเดียวกัน (Image Resizing)

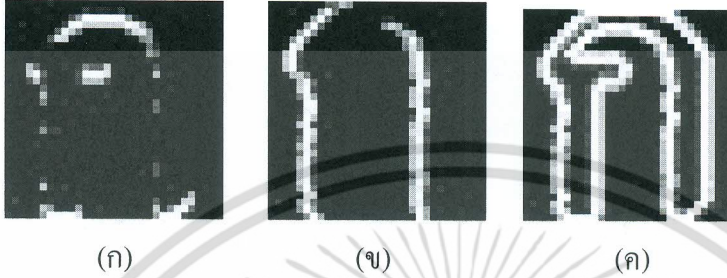
3.2.2.2 การสกัดคุณลักษณะ

ในขั้นตอนนี้ผู้วิจัยได้พัฒนาฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะมาจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบดั้งเดิม โดยในส่วนถัดไปจะอธิบายถึงขั้นตอนการทำงานพร้อมทั้งอธิบายหลักการและเหตุผลในส่วนที่ได้ทำการปรับปรุงแก้ไขเพื่อเพิ่มประสิทธิภาพให้กับฮิสโตแกรมของการไล่ระดับเชิงมุม

(1) การคำนวณขนาดและทิศทาง (Gradient Computation)

เทคนิคการหาขอบภาพมีอยู่ด้วยกันหลายวิธี ซึ่งวิธีการหาขอบภาพด้วยอนุพันธ์อันดับที่ 1 เป็นวิธีที่ง่าย รวดเร็ว และมีประสิทธิภาพ โดยเริ่มจากการหาขอบภาพตามแนวแกน x และแกน y อาศัยหลักการของตัวกรอง $[-1 \ 0 \ 1]$ และ $[-1 \ 0 \ 1]^T$ เพื่อให้ได้เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลลัพธ์ คือ ภาพที่มีแต่ขอบภาพตามแนวแกน x และแกน y ซึ่งจะถูกเก็บใน I_x และ I_y ตามลำดับ ดังรูปที่ 3.7 (ก) และรูปที่ 3.7(ข) ตามลำดับ จากนั้นคำนวณหาขนาด (M) ในแต่ละจุดภาพตามสมการที่ 2.10 และคำนวณทิศทาง (θ) ตามสมการที่ 2.11 หลังจากนั้นให้เปลี่ยน θ จากหน่วยเรเดียนให้เป็นหน่วยองศา เพื่อสะดวกต่อการอ้างอิงในการนำไปคำนวณ



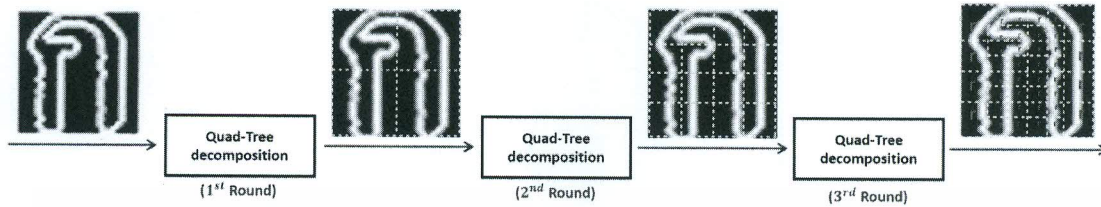
รูปที่ 3.7 การหาขอบภาพ (ก) ผลลัพธ์ของ I_x (ข) ผลลัพธ์ของ I_y (ค) ผลลัพธ์ของ M

(2) การแบ่งภาพออกเป็นส่วนย่อย (Partition)

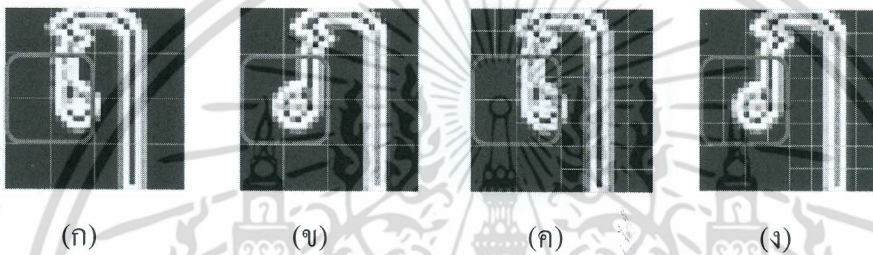
นำผลลัพธ์จากขั้นตอนที่ (1) มาแบ่งภาพออกเป็นภาพย่อย หรือ เซลล์ (cell) ซึ่งจะมีขนาดเป็นจัตุรัส ในส่วนของการกำหนดขนาดของเซลล์มี 2 รูปแบบ คือ แบบคงตัว (Fixed Partition) และ แบบปรับเหมาะ (Adaptive Partition) ซึ่งวิธีดั้งเดิมได้กำหนดขนาดของเซลล์เป็นแบบคงตัวตามขนาดที่กำหนด (ทุกเซลล์มีขนาดเดียวกันหมด) แต่ในงานวิจัยนี้ได้มีการกำหนดขนาดของเซลล์ในแต่ละเซลล์ไม่เท่ากัน โดยขึ้นกับค่าพิสัยเมื่อเทียบกับค่าเกณฑ์ ซึ่งการแบ่งแบบปรับเหมาะ เป็นการเลือกขนาดของเซลล์ให้เหมาะสมกับข้อมูล ดังรูปที่ 3.9 ถ้ากำหนดขนาดของเซลล์ละเอียดในบริเวณที่ข้อมูลไม่มีการเปลี่ยนแปลงอาจส่งผลให้เกิดข้อมูลที่มีความซ้ำซ้อนกัน แต่ถ้ากำหนดขนาดของเซลล์หยาบเกินไปในบริเวณที่ข้อมูลที่มีความแตกต่างกันสูง อาจทำให้สูญเสียรายละเอียดของข้อมูลบางส่วนไป ดังนั้น ในงานวิจัยนี้ ได้นำเสนอการแบ่งขนาดเซลล์แบบปรับเหมาะ โดยใช้การแยกองค์ประกอบแบบต้นไม้จตุนิยมเพื่อแบ่งภาพออกเป็นเซลล์ หรือ r_1, r_2, \dots, r_n โดยค่าเกณฑ์ (Threshold) มีอิทธิพลต่อการแบ่งภาพออกเป็นเซลล์ ซึ่งถ้าค่าเกณฑ์น้อยก็จะแบ่งเซลล์ได้ละเอียด และถ้าค่าเกณฑ์มากจะส่งผลให้แบ่งเซลล์ได้หยาบ ผลจากการทดลองพบว่า ค่าเกณฑ์ 0.25 เป็นค่าที่มีความเหมาะสมสำหรับข้อมูลชุดนี้ และขนาดของแต่ละเซลล์ คือ 4 และ 8 พิกเซล เนื่องจากภาพตัวอักษรภาษาไทยนั้นมีขนาด 32×32 จุดภาพ และจำนวนดังตามขั้นตอนข้างต้น ถ้าขนาดต่อเซลล์เล็กกว่า 4 พิกเซล จะทำให้ค่าในแต่ละดังมีโอกาสดังไม่มีข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และความกว้างของเซลล์ไม่ควรเกิน 8 เพราะไม่มีกรณีใดที่ทำให้เซลล์มีขนาดของเซลล์มากกว่า 8
 ดังรูปที่ 3.8



รูปที่ 3.8 การแบ่งภาพออกเป็นส่วนย่อยโดยอาศัยการแยกองค์ประกอบแบบต้นไม้จตุรมุข

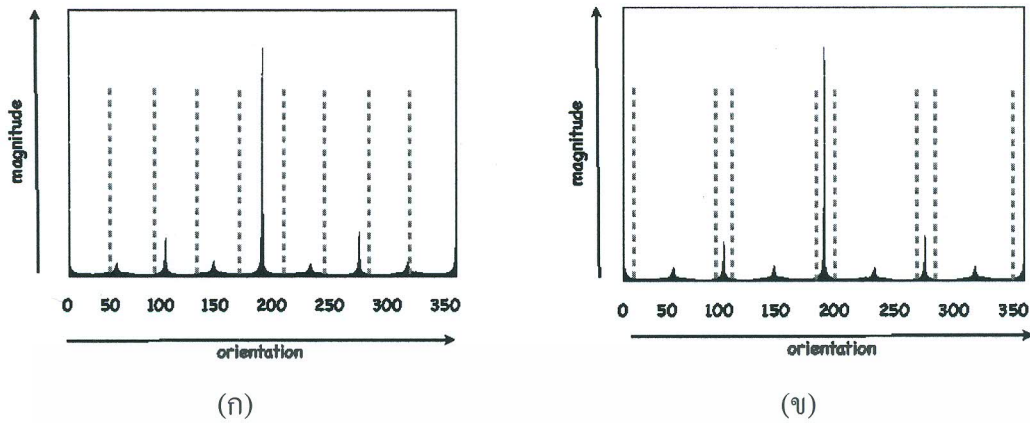


รูปที่ 3.9 ความแตกต่างการแบ่งภาพออกเป็นส่วนย่อย (ก) การแบ่ง “ก” แบบคงตัว (ข) การแบ่ง “ก” แบบคงตัว (ค) การแบ่ง “ก” แบบปรับเหมาะ (ง) การแบ่ง “ก” แบบปรับเหมาะ

(3) การกำหนดความกว้างของถังแบบปรับเหมาะ (Adaptive Binning)

การทำ Quantization คือ กระบวนการกำหนดขนาดแต่ละช่วงของข้อมูล ซึ่งจะถูกแทนด้วยจำนวนบิตลงไปในแต่ละช่วง ขนาดของช่วงส่งผลต่อความละเอียดของข้อมูล และความผิดพลาดอันเนื่องมาจากการกำหนดขนาดของช่วง (Quantization Error) เช่น ถ้าขนาดของช่วงกว้างจะส่งผลทำให้ความละเอียดของสัญญาณน้อยลง และส่งผลให้มีความผิดพลาดมากขึ้น อันเนื่องมาจากขนาดของช่วงกว้าง แต่ถ้าขนาดของช่วงแคบจะทำให้ความละเอียดของสัญญาณสูงขึ้น และส่งผลให้ความผิดพลาดน้อยลง อันเนื่องมาจากการกำหนดขนาดของช่วงแคบ การทำ Quantization มี 2 แบบ คือ Uniform Quantization และ Non-uniform Quantization ดังรูปที่ 3.10 ซึ่ง Uniform Quantization เป็นการกำหนดขนาดของแต่ละช่วงของข้อมูลให้มีขนาดเท่ากันหมดในทุกช่วง

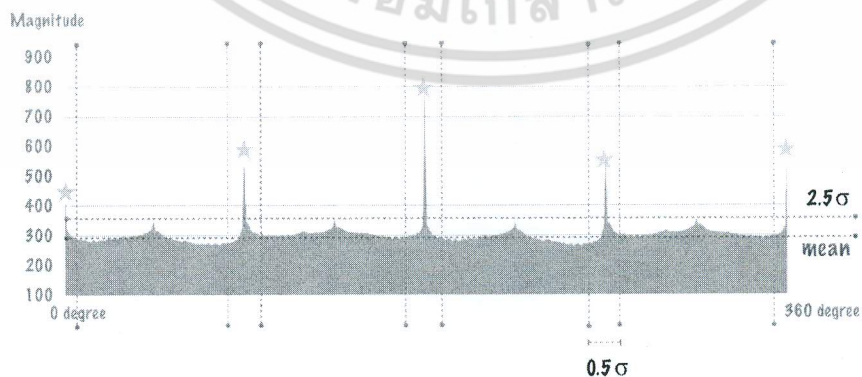
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.10 การกำหนดความกว้างของถัง (ก) กำหนดแบบ Uniform (ข) กำหนดแบบ Non-uniform

ในส่วนของ Non-uniform Quantization จะแบ่งขนาดของแต่ละช่วงของข้อมูลออกเป็นไม่เท่ากัน ขึ้นอยู่กับความน่าจะเป็นของการเกิดข้อมูล ถ้าข้อมูลมีโอกาสเกิดขึ้นมาก ก็จะกำหนดให้ขนาดของช่วงนั้นแคบ แต่ถ้ามีโอกาสเกิดน้อยก็จะกำหนดขนาดของช่วงข้อมูลกว้าง ซึ่งหากเก็บข้อมูลที่มีโอกาสเกิดมากแบบละเอียด (ขนาดแคบ) จะทำให้เกิดข้อผิดพลาดขึ้นเนื่องมาจากการกำหนดขนาดของช่วงน้อยลง ดังนั้น ในงานวิจัยนี้จึงเลือกปรับปรุงการแบ่งถังแบบเดิมที่ทุกถังมีความกว้างเท่ากันมาเป็นแบบไม่เท่ากัน ซึ่งขึ้นอยู่กับโอกาสการเกิดของข้อมูลนั้น เพื่อลดการเกิดความผิดพลาดขึ้นเนื่องมาจากการกำหนดขนาดของช่วงที่ไม่เหมาะสม

จากการคำนวณพลังงานตามสมการที่ 2.10 ของทุกองศาในทุกภาพของตัวอักษรของข้อมูลชุดทดสอบ เพื่อคำนวณหาค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของพลังงาน สุดท้ายเลือกองศาที่มีพลังงานมากกว่าค่าเฉลี่ยรวมกับส่วนเบี่ยงเบน (2.5 σ) และกำหนดความกว้างของถังให้มีค่า 0.5 σ ดังรูปที่ 3.11 โดยที่แนวแกนตั้ง คือ พลังงาน และแนวแกนนอน คือ องศา



รูปที่ 3.11 การกำหนดจำนวนถัง และความกว้างแต่ละถัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

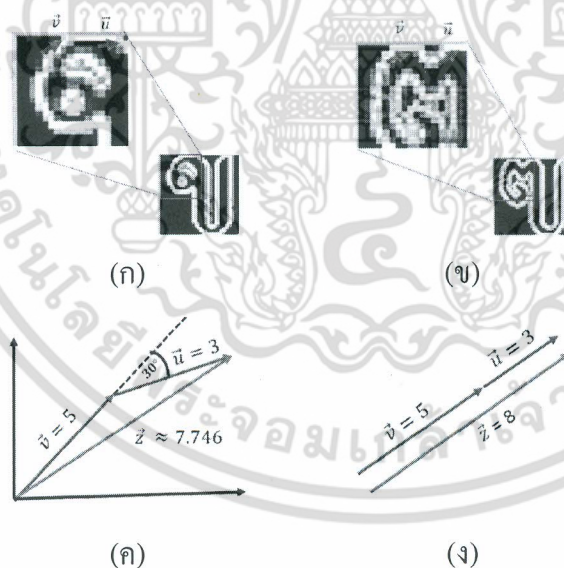
(4) การคำนวณขนาดตามทิศทางจากการแบ่งภาพออกเป็นส่วนย่อย

ในขั้นตอนนี้เป็นการรวมค่าฮิสโตแกรมโดยแบ่งตามทิศทาง ซึ่งการรวมฮิสโตแกรมของการไล่ระดับเชิงมุมแบบดั้งเดิมจะเป็นการรวมแบบสเกลาร์ ข้อดีคือง่ายและรวดเร็ว แต่ขาดความแม่นยำ ซึ่งตรงข้ามกับการรวมค่าแบบเวกเตอร์ตามสมการที่ 3.1 จะอาศัยทั้งขนาดและทิศทางในการรวมค่าฮิสโตแกรม ส่งผลให้มีความแม่นยำสูงขึ้น

$$\vec{z} = \sqrt{\|\vec{n}\|^2 + \|\vec{c}\|^2 + 2\|\vec{n}\|\|\vec{c}\|\cos\theta} \quad (3.1)$$

$$\vec{z} = \|\vec{n}\| + \|\vec{c}\| \quad (3.2)$$

จากตัวอย่างในรูปที่ 3.12(ก) และ 3.12(ข) ตัวอักษร “ข-ช” เป็นตัวอักษรที่มีความคล้ายคลึงกัน ถ้าพิจารณาในส่วนที่มีความแตกต่างกัน คือ บริเวณส่วนหัวของตัวอักษร เมื่อรวมค่าฮิสโตแกรมแบบสเกลาร์ตามสมการที่ 3.2 ดังรูปที่ 3.12(ง) ส่งผลให้ค่าฮิสโตแกรมของ “ข-ช” ไม่มีความแตกต่างกัน แต่ถ้าวรวมค่าฮิสโตแกรมแบบเวกเตอร์ดังรูปที่ 3.12(ค) ซึ่งมุม θ ที่ใช้จะต่างกันส่งผลให้ค่าฮิสโตแกรมที่ได้มีความแตกต่างกันด้วย



รูปที่ 3.12 เปรียบเทียบความแตกต่างของการรวมขนาดแบบสเกลาร์และเวกเตอร์ (ก) แสดง \vec{n} และ \vec{c} ของภาพอักษร “ข” (ข) แสดง \vec{n} และ \vec{c} ของภาพอักษร “ช” (ค) แสดงการรวม \vec{n} และ \vec{c} แบบเวกเตอร์ (ง) แสดงการรวม \vec{n} และ \vec{c} แบบสเกลาร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น การรวมค่าฮิสโตแกรมโดยอาศัยทั้งขนาดและทิศทางแบบเวกเตอร์จะส่งผลให้เกิดความแม่นยำ และถูกต้องกับภาพประเภทตัวอักษร ในงานวิจัยนี้ได้นำเสนอการคำนวณค่าฮิสโตแกรมตามทิศทางในแต่ละส่วนย่อยแบบเวกเตอร์

(5) การจัดรูปแบบฮิสโตแกรมมาตรฐาน

ในขั้นตอนสุดท้ายเป็นการจัดรูปแบบฮิสโตแกรมมาตรฐาน (Standard Form Histogram) จะพิจารณาผลลัพธ์ที่ได้จากกระบวนการข้างต้น ในส่วนของเซลล์ที่มีขนาด 8×8 จุดภาพ จะพิจารณาแยกเซลล์ให้มีขนาด 4×4 จุดภาพ จำนวน 4 เซลล์ และใช้ค่าฮิสโตแกรมเดียวกันทั้ง 4 เซลล์ (H_i) ดังรูปที่ 3.13 เนื่องจากค่าความแปรปรวนรวมภายในเซลล์มีค่อนข้างน้อย อีกทั้งช่วยลดภาระในการคำนวณเซลล์ขนาด 4×4 จุดภาพ จากจำนวน 4 รอบให้เหลือการคำนวณเพียงรอบเดียว หลังจากนั้น จึงนำค่าฮิสโตแกรมในแต่ละเซลล์มาเรียงต่อกันเป็นเวกเตอร์คุณลักษณะ ($F = \{H_1, H_1, H_2, H_2, \dots, H_{40}\}$) แบบหลักต่อหลัก

ตัวอธิบายคุณลักษณะจะถูกสร้างจากกระบวนการฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะจากรูป 3.12(ข) แล้วทำการแปลงให้เป็นเวกเตอร์ตามหัวข้อ 2.2.3 หรือเรียกว่า SD-AHOG

H_1	H_5	H_{13}	H_{21}	H_{29}	H_{37}
	H_6	H_{14}	H_{22}	H_{30}	
H_2	H_7	H_{15}	H_{23}	H_{31}	H_{38}
	H_8	H_{16}	H_{24}	H_{32}	
H_3	H_9	H_{17}	H_{25}	H_{33}	H_{39}
	H_{10}	H_{18}	H_{26}	H_{34}	
H_4	H_{11}	H_{19}	H_{27}	H_{35}	H_{40}
	H_{12}	H_{20}	H_{28}	H_{36}	

(ก)

H_1	H_1	H_5	H_{13}	H_{21}	H_{29}	H_{37}	H_{37}
H_1	H_1	H_6	H_{14}	H_{22}	H_{30}	H_{37}	H_{37}
H_2	H_2	H_7	H_{15}	H_{23}	H_{31}	H_{38}	H_{38}
H_2	H_2	H_8	H_{16}	H_{24}	H_{32}	H_{38}	H_{38}
H_3	H_3	H_9	H_{17}	H_{25}	H_{33}	H_{39}	H_{39}
H_3	H_3	H_{10}	H_{18}	H_{26}	H_{34}	H_{39}	H_{39}
H_4	H_4	H_{11}	H_{19}	H_{27}	H_{35}	H_{40}	H_{40}
H_4	H_4	H_{12}	H_{20}	H_{28}	H_{36}	H_{40}	H_{40}

(ข)

รูปที่ 3.13 เปรียบเทียบผลลัพธ์ก่อนและหลังผ่านกระบวนการการจัดรูปแบบฮิสโตแกรมมาตรฐาน (ก) ผลลัพธ์ก่อนจัดให้อยู่ในรูปแบบฮิสโตแกรมมาตรฐาน (ข) ผลลัพธ์หลังจัดให้อยู่ในรูปแบบฮิสโตแกรมมาตรฐาน

3.2.2.3 การรู้จำตัวอักษรด้วยการหาระยะทางยุคลิด

การวัดความคล้ายด้วยระยะทางยุคลิดนั้น จะทำการเทียบเวกเตอร์คุณลักษณะของข้อมูลชุดฝึกสอนกับข้อมูลของชุดทดสอบ โดยที่ระยะทางระหว่างเวกเตอร์คุณลักษณะคู่ใดน้อยที่สุด แสดงว่าสองเวกเตอร์คุณลักษณะนั้นมีความเหมือนกันมากที่สุด เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น ไม่นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากความหลากหลายของชุดแบบอักษร ชนิด และขนาดของอักษร อีกทั้งความคล้ายคลึงกันของอักษรภาษาไทยเป็นปัจจัยที่ส่งผลต่อประสิทธิภาพการรู้จำตัวอักษรภาษาไทย ซึ่งในบทนี้ได้ ออกแบบกระบวนการรู้จำตัวอักษรภาษาไทยมา 2 แนวทาง โดยแนวทางแรกทำการสกัดคุณลักษณะด้วยวิธีการแยกองค์ประกอบค่าเดียวเพื่อให้ได้ตัวอธิบายคุณลักษณะที่สามารถแบ่งแยกภาพตัวอักษรที่มีความคล้ายคลึงกันได้ อีกทั้งยังมีความทนทานต่อการเปลี่ยนแปลงของชุดแบบอักษร ชนิด และขนาดของภาพตัวอักษรดังผลการทดลองที่ 1 ในบทที่ 4 แต่ขนาดของคุณลักษณะที่ได้ยังมีขนาดใหญ่ ดังนั้น การรู้จำตัวอักษรภาษาไทยแบบที่สอง จึงได้สกัดคุณลักษณะจากวิธีฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ ซึ่งเป็นคุณลักษณะที่สามารถแบ่งแยกภาพตัวอักษรที่มีความคล้ายคลึงกันได้ แล้วยังมีความทนทานต่อการเปลี่ยนแปลงของชุดแบบอักษร ชนิดและขนาดของภาพตัวอักษร อีกทั้งคุณลักษณะที่ได้ยังมีขนาดเล็กอีกด้วยดังผลการทดลองที่ 2 ในบทที่ 4 นอกจากนี้ ในบทที่ 4 ยังมีผลการทดลองเปรียบเทียบประสิทธิภาพระหว่างวิธีการสกัดคุณลักษณะด้วยการแยกองค์ประกอบค่าเดียวในลำดับที่ 512-704 (SD2) กับการสกัดคุณลักษณะโดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (SD-AHOG)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการวิจัย

ในบทนี้จะกล่าวถึงผลการทดลองโดยชุดข้อมูลทดลองประกอบด้วยชุดแบบอักษรทั้งหมด 12 ชุดแบบอักษร ในแต่ละชุดแบบอักษรมีขนาด 8 , 10 , 12 , 14 , 16 , 18 , 20 และ 22 พอยต์ รูปแบบปกติ หนา เอียง และหนาปนเอียง โดยประกอบด้วย พยัญชนะ สระ วรรณยุกต์ และเลขไทย เนื่องจากภาษาไทยมี 3 ระดับ คือ บน กลาง และล่าง ซึ่งมีจำนวนสมาชิกคือ 12, 65, และ 2 ตามลำดับ ในทุกระดับมีความสำคัญ แต่ในตัวอักษรภาษาไทยในระดับกลางจะมีความสำคัญสูงสุด เพราะมีจำนวนสมาชิกมาก ความละเอียดของภาพที่ใช้ในงานวิจัยนี้คือ 400 dpi โดยในแต่ละภาพจะเป็นตัวอักษรเดี่ยวตัวอักษรสีดำ พื้นหลังสีขาว ไม่มีสัญญาณรบกวน ตัวอักษรมีลักษณะสมบูรณ์ไม่ขาดหาย และเป็นตัวพิมพ์เท่านั้น

ในบทนี้แสดงผลการทดลองทั้งหมด 4 การทดลอง ในการทดลองแรกเป็นการทดลองที่เกี่ยวข้องกับงานวิจัยของ Woraratpanya and Titijaronroj (2013: 165-173) โดยมีวัตถุประสงค์เพื่อทดสอบหาจำนวนและลำดับไอเจนเวกเตอร์ที่เหมาะสมสำหรับใช้เป็นตัวอธิบายคุณลักษณะ การทดสอบค่าความถูกต้องของการรู้จำ และการทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อการเปลี่ยนแปลงขนาด และชุดแบบของตัวอักษร โดยใช้การเลือกสุ่มข้อมูลแบบเที่ยงตรง K กลุ่ม (K-Fold Cross Validation) แบบมีเงื่อนไขสำหรับการแบ่งข้อมูล ในการทดลองที่สองเป็นการทดลองที่เกี่ยวข้องกับงานวิจัยของ Woraratpanya and Titijaronroj (2014: 83-92) เพื่อเปรียบเทียบประสิทธิภาพการรู้จำของงานวิจัยนี้กับกระบวนการในกลุ่มของฮิสโตแกรมของการไล่ระดับเชิงมุม และการทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อการเปลี่ยนแปลงขนาด และชุดแบบของตัวอักษร โดยการเลือกสุ่มข้อมูลแบบเที่ยงตรง K กลุ่มแบบมีเงื่อนไขสำหรับการแบ่งข้อมูล เพื่อการวัดประสิทธิภาพของการทดลองดังกล่าว ในการทดลองที่สามคือการเปรียบเทียบงานวิจัยของ Woraratpanya and Titijaronroj (2013: 165-173) กับงานวิจัยของ Woraratpanya and Titijaronroj (2014: 83-92) ในส่วนของประสิทธิภาพของตัวอธิบายคุณลักษณะ โดยประเมินจากความถูกต้องในการรู้จำ จำนวนคุณลักษณะ เวลาที่ใช้ในการสกัดคุณลักษณะ และความทนทานต่อขนาด และชุดแบบของตัวอักษร โดยทุกการทดลองได้นำข้อมูลจากการทดลองที่หนึ่งและการทดลองที่สองมาใช้ในการเปรียบเทียบและวิเคราะห์ผล และในส่วนสุดท้ายคือการทดลองเปรียบเทียบประสิทธิภาพการรู้จำโดยจำแนกตามกรณีต่าง ๆ ได้แก่ การรู้จำถูกต้อง (TP) การรู้จำผิดพลาดในกลุ่มที่มีความคล้ายคลึง เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กัน (FP) และการรู้จำผิดพลาดในกลุ่มที่ไม่มีความคล้ายคลึงกัน (FN) ของ SD1, SD2, SD-AHOG, HOG, SSHOG, PHOG, MHOG และ HOGC

4.1 ผลการทดลองที่ 1: การวัดประสิทธิภาพคุณลักษณะที่ได้จากการแยกองค์ประกอบค่าเดียว

ในการทดลองนี้มีวัตถุประสงค์เพื่อวัดประสิทธิภาพของ SD1 และ SD2 โดยทดสอบว่าตัวอธิบายคุณลักษณะใดสามารถแบ่งแยกภาพตัวอักษรที่มีความคล้ายคลึงกันได้ดี อีกทั้งยังมีความทนทานต่อการเปลี่ยนแปลงของชุดแบบอักษร ชนิด และขนาดของภาพตัวอักษร ซึ่งการเลือกไอเทมเวกเตอร์ทั้ง 2 แบบ กล่าวคือ SD1 มีขนาดของคุณลักษณะต่อหนึ่งภาพอักษรเท่ากับ 7,168 ขณะที่ SD2 มีขนาดของคุณลักษณะเท่ากับ 197,632 และใช้เวลาในการสกัดคุณลักษณะต่อหนึ่งภาพตัวอักษร 0.0641 วินาทีเท่ากัน ผลการทดลองในหัวข้อนี้ถูกแบ่งออกเป็น 3 ส่วน ได้แก่ การเลือกองค์ประกอบเมตริกซ์คุณลักษณะ การทดสอบค่าความถูกต้องของตัวอธิบายคุณลักษณะ และการทดสอบความทนทานของตัวอธิบายคุณลักษณะ

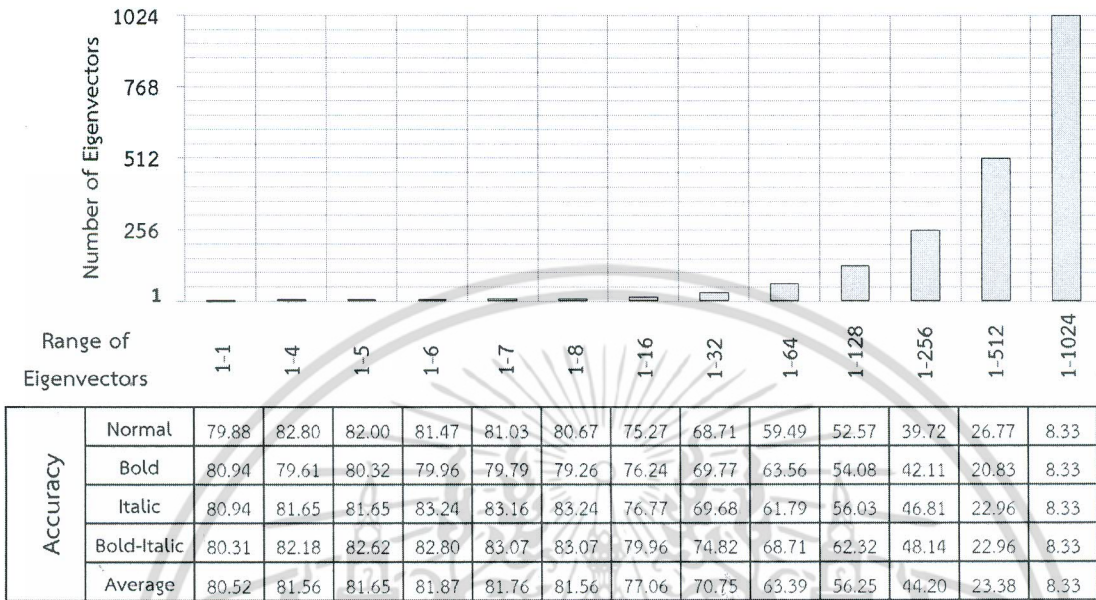
4.1.1 การเลือกองค์ประกอบเมตริกซ์คุณลักษณะที่เหมาะสม

การเลือกองค์ประกอบเมตริกซ์คุณลักษณะที่เหมาะสมสำหรับตัวอธิบายคุณลักษณะเป็นการทดลองเลือกใช้องค์ประกอบเมตริกซ์คุณลักษณะที่ได้จากการสกัดโดยใช้การแยกองค์ประกอบค่าเดียวในตำแหน่งที่แตกต่างกันเพื่อใช้เป็นตัวอธิบายคุณลักษณะ ซึ่งผลการทดลองในการเลือกองค์ประกอบของเมตริกซ์คุณลักษณะตำแหน่งต่าง ๆ ของแบบ SD1 และ SD2 ได้แสดงในรูปที่ 4.1 ถึง 4.6 โดยใช้ชุดฝึกฝน (Training Set) และ ชุดทดสอบ (Test Set) จำนวน 2 ชุดที่มีภาพอักษรทุกประเภท ทุกชนิด และทุกขนาด

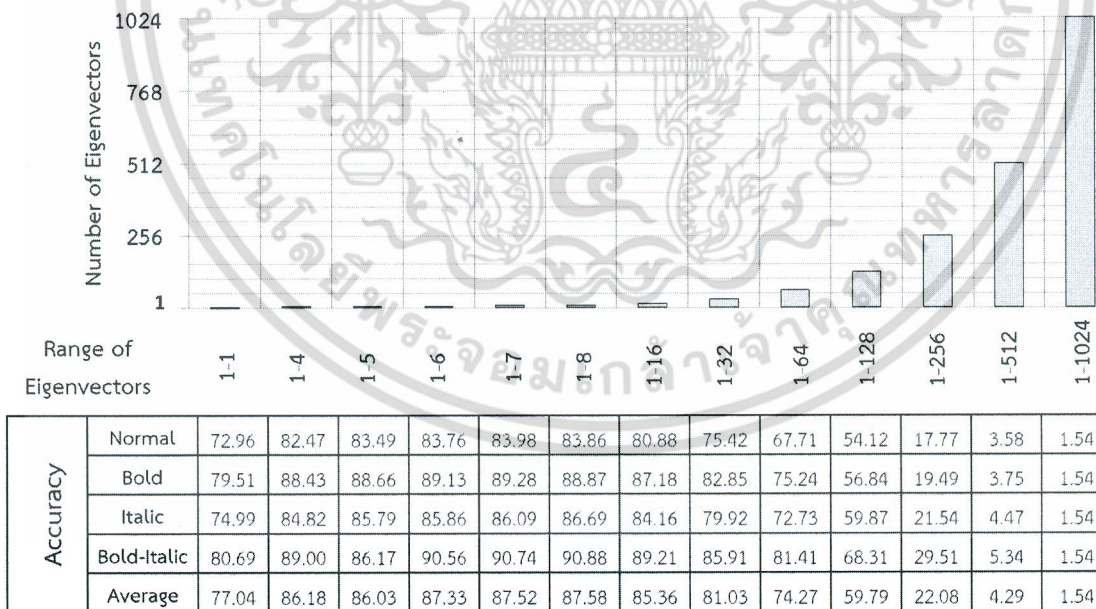
จากผลการทดลองพบว่า ช่วงที่เหมาะสมของไอเทมเวกเตอร์แบบ SD1 ให้ความถูกต้องเฉลี่ยของการรู้จำสูงสำหรับตัวอักษรภาษาไทยในระดับบน กลาง และล่าง คือ ลำดับที่ 1-6, 1-7 และ 1 ตามลำดับ และส่วน SD2 ของตัวอักษรภาษาไทยในระดับบน กลาง และล่าง คือ ลำดับที่ 512-704, 512-1024 และ 512-1024 ตามลำดับ เนื่องด้วยสมาชิกในแต่ละลำดับมีจำนวนที่แตกต่างกัน ตัวอักษรภาษาไทยในระดับกลางจึงมีความสำคัญสูงสุดเพราะมีจำนวนสมาชิกมาก ดังนั้น ในกรณีของ SD1 จึงเลือกใช้คุณลักษณะที่สร้างมาจากไอเทมเวกเตอร์ในลำดับที่ 1-7 โดยยึดผลการทดลองของอักษรไทยในระดับกลาง และในกรณีของ SD2 จะเลือกใช้คุณลักษณะที่สร้างมาจากไอเทมเวกเตอร์ในลำดับที่ 512-704 ถึงแม้ว่าไอเทมเวกเตอร์ในลำดับที่ 512-1024 จะพบว่าได้ค่าความถูกต้องสูงสุด แต่จำนวนคุณลักษณะที่ใช้จะสูงตามไปด้วย แต่สำหรับช่วง 512-704 เป็นช่วงที่ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะน้อยกว่าในช่วง 512-1024 อยู่ 2.67 เท่า แต่มีค่าความถูกต้องลดลงเฉลี่ย 0.25% ซึ่งสามารถยอมรับได้

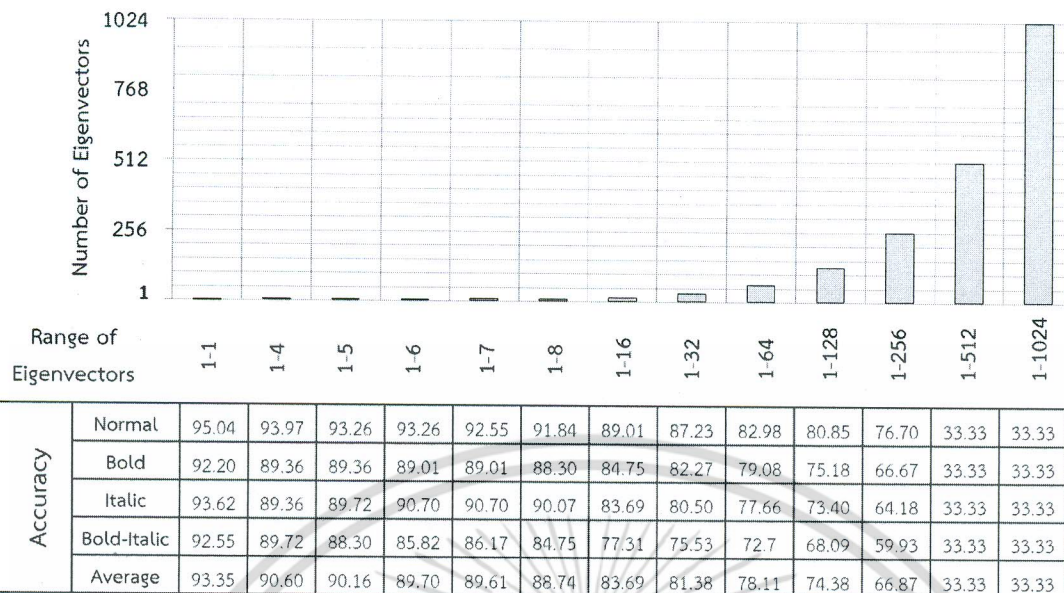


รูปที่ 4.1 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดสอบ ภาพอักษรภาษาไทยระดับบน

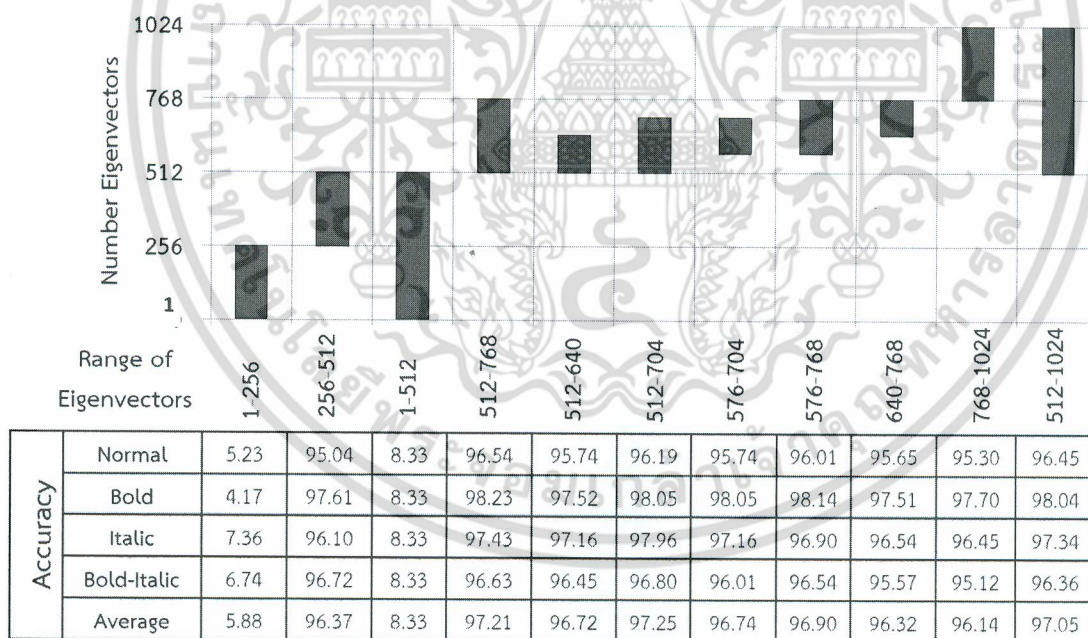


รูปที่ 4.2 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดสอบ ภาพอักษรภาษาไทยระดับกลาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

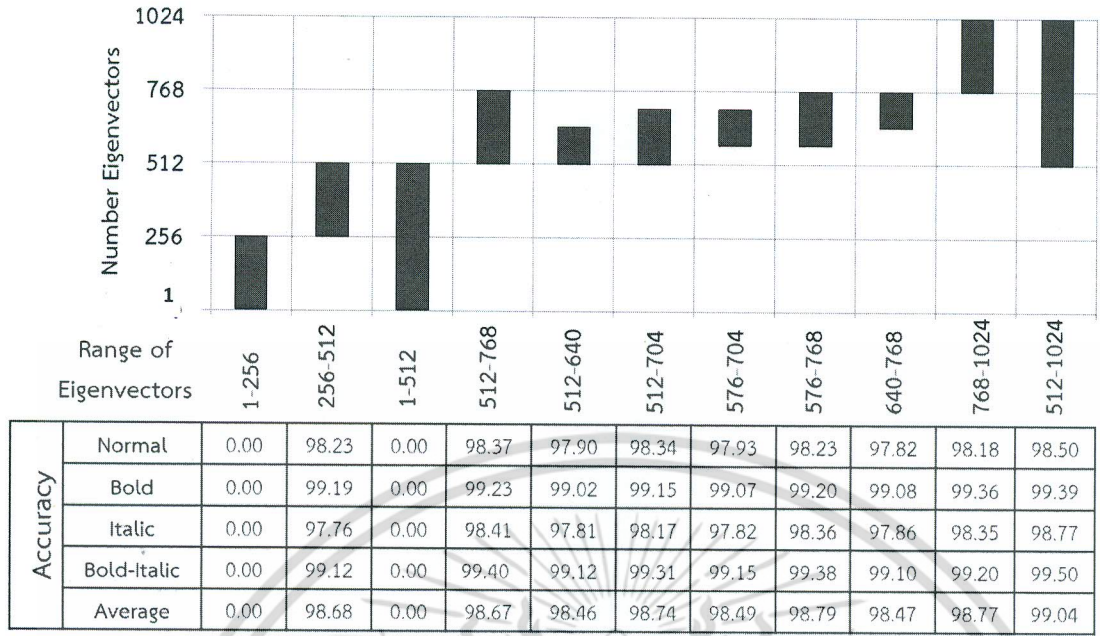


รูปที่ 4.3 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD1 สำหรับชุดทดลอง ภาพอักษรภาษาไทยระดับล่าง

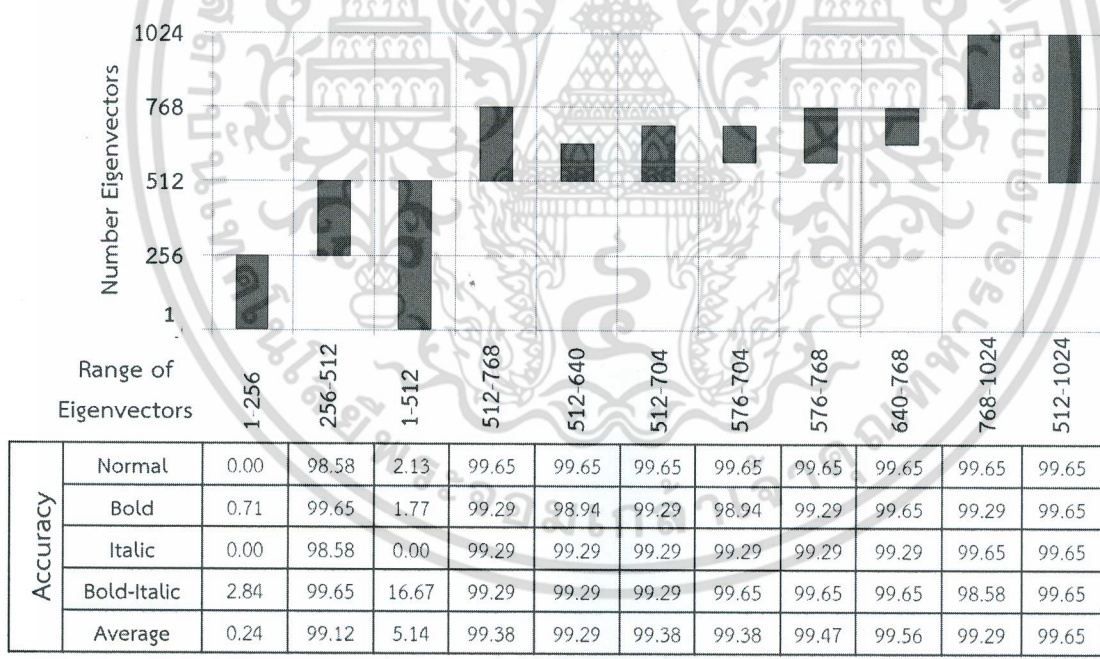


รูปที่ 4.4 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลอง ภาพอักษรภาษาไทยระดับบน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับกลาง



รูปที่ 4.6 ความถูกต้องของการรู้จำเมื่อเลือกองค์ประกอบที่แตกต่างกันของ SD2 สำหรับชุดทดลองภาพอักษรภาษาไทยระดับล่าง

4.1.2 การทดสอบความถูกต้องตัวอธิบายคุณลักษณะ SD1 และ SD2

ในการทดสอบโดยใช้การเลือกข้อมูลแบบเที่ยงตรง K แบบโดยที่ K = 30 ซึ่งได้ผลการทดลองด้วยวิธี SD1 และ SD2 จากตารางที่ 4.1 จะแสดงค่าเฉลี่ยและส่วนเบี่ยงเบนเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำมาเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มาตรฐานที่ได้จากค่าความถูกต้องโดยคำนวณจากสมการที่ 2.4 โดยใช้การเลือกข้อมูลแบบเที่ยงตรง K แบบ โดยที่ $K = 30$ จากตารางที่ 4.3

ตารางที่ 4.1 ผลการเปรียบเทียบความถูกต้องของการรู้จำของ SD1 และ SD2

	SD1 (%)			SD2 (%)		
	Upper Level	Middle Level	Lower Level	Upper Level	Middle Level	Lower Level
Normal	80.27±1.64	83.48±0.52	99.57±0.40	94.76±0.97	97.62±0.19	99.52±0.43
Bold	78.54±1.64	87.87±0.52	98.83±0.73	96.08±0.67	98.52±0.16	98.94±0.73
Italic	82.68±1.61	85.19±0.51	100.00±0.00	95.37±0.71	97.14±0.27	99.95±0.16
Bold-Italic	83.29±2.06	89.36±0.38	98.65±0.87	94.88±0.79	98.41±0.15	99.13±0.49
AVG	81.20	86.48	99.26	<u>95.27</u>	<u>97.92</u>	<u>99.39</u>
		86.00			<u>97.55</u>	

จากผลการทดลองพบว่า ค่าความถูกต้องของคุณลักษณะแบบ SD2 จะได้ค่าความถูกต้องสูงกว่า SD1 ในทุกระดับและทุกชนิดของตัวอักษร โดยมีค่าความถูกต้องเฉลี่ยสูงกว่าโดยเฉลี่ย 11.55% เนื่องจาก SD2 สามารถแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันดีกว่า SD1 จึงส่งผลให้ SD2 เป็นตัวอธิบายคุณลักษณะที่ดีในงานวิจัยของ Woraratpanya and Titijaronroj (2013: 165-173) เนื่องจากการเลือกใช้ไอเทมเวกเตอร์ใน “ลำดับปลาย” หรือ SD2 สามารถกล่าวได้ว่าเป็นการสกัดคุณลักษณะแบบเลือกความถี่สูง หรือบริเวณขอบของภาพตัวอักษร ซึ่งจะพบว่าชุดข้อมูลตัวอักษรภาษาไทยมีลักษณะเด่นในเชิงรูปร่างจึงเป็นสาเหตุที่ SD2 มีความถูกต้องในการรู้จำที่ดีกว่า SD1

4.1.3 การทดสอบความทนทานของตัวอธิบายคุณลักษณะ

การทดสอบความทนทาน (Robustness) ของตัวอธิบายคุณลักษณะในการการรู้จำตัวอักษร แบ่งเป็น 2 ส่วน คือ ความทนทานต่อความแตกต่างของชุดแบบอักษร และความทนทานต่อความแตกต่างของขนาดตัวอักษร ด้วยวิธีการของ SD1 และ SD2 ซึ่งให้ผลการทดสอบดังแสดงในตารางที่ 4.2 โดยใช้การเลือกข้อมูลแบบเที่ยงตรง K แบบ มีเงื่อนไขการแบ่งข้อมูลตามขนาดของตัวอักษร จะสามารถแบ่งได้ 8 กลุ่ม เพื่อนำไปใช้ทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อขนาด และแบ่งข้อมูลตามชุดแบบของตัวอักษร จะสามารถแบ่งได้ 12 กลุ่ม เพื่อนำไปใช้ทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อชุดแบบอักษร

ตารางที่ 4.2 ผลการทดสอบความทนทานในการรู้จำที่มีความแตกต่างของชุดแบบของตัวอักษรกับขนาดของตัวอักษรด้วยวิธี SD1 และ SD2

Method	Robustness Font			Robustness Size		
	Upper Level	Middle Level	Lower Level	Upper Level	Middle Level	Lower Level
SD1	73.78	73.60	<u>97.92</u>	74.71	72.12	96.88
SD2	<u>85.65</u>	<u>86.49</u>	95.67	<u>90.75</u>	<u>96.07</u>	<u>97.48</u>

จากผลการทดลองจะพบว่า ตัวอธิบายคุณลักษณะแบบ SD1 และ SD2 มีความทนทานเฉลี่ยต่อขนาดของตัวอักษรถึง 81.24% และ 94.77% ตามลำดับ จะเห็นได้ว่าค่าความถูกต้องลดลงจากการทดลองในตารางที่ 4.1 อยู่ 4.76% และ 2.78% ในส่วนของ SD1 และ SD2 ตามลำดับ ซึ่งเมื่อพิจารณาจะพบว่า SD2 มีค่าความถูกต้องลดลงน้อยกว่า SD1 อีกทั้งค่าความถูกต้องของ SD2 ลดลงเล็กน้อย จึงสามารถสรุปได้ว่า SD2 มีความทนทานต่อขนาดของตัวอักษร และทนทานมากกว่า SD1

ในส่วนของความทนทานต่อชุดแบบของตัวอักษรจะพบปัญหาในกรณีของชุดแบบอักษร JasmineUPC เพราะชุดแบบอักษรดังกล่าวจะมีลักษณะรูปร่างที่แตกต่างจากชุดแบบอักษรอื่น ดังรูปที่ 4.7 บริเวณที่เป็นแถบสีแดงจะพบว่าชุดแบบของตัวอักษร JasmineUPC จะมีหัวทั้ง 2 ข้างเป็นจุดที่ทำให้ไปคล้ายกับตัวอักษรอื่น อีกทั้งความคล้ายคลึงของ “ก-ภ-ถ” ในชุดแบบตัวอักษร JasmineUPC มีค่อนข้างสูง ดังนั้น จึงทำให้ค่าความทนทานเฉลี่ยต่อชุดแบบอักษรลดลงจากการทดลองในตารางที่ 4.1 เหลือเพียง 81.77% และ 89.27% ในส่วนของ SD1 และ SD2 ตามลำดับ ถ้าพิจารณาค่าความถูกต้องที่ลดลงจากตารางที่ 4.1 ของ SD1 และ SD2 คือ 4.23% และ 8.28% ตามลำดับ จะพบว่า SD1 มีความทนทานต่อชุดแบบอักษรมากกว่า SD2 แต่ค่าความถูกต้องเฉลี่ยในส่วนของค่าความทนทานต่อชุดแบบอักษรของ SD2 มากกว่า SD1

	ก	ภ	ถ
BrowalliaUPC			
CordiaUPC			
JasmineUPC			

รูปที่ 4.7 ตัวอย่างภาพอินพุทของ “ก-ภ-ถ” ที่มีความคล้ายคลึงกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการทดลองที่ 2: การวัดประสิทธิภาพคุณลักษณะที่สร้างจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

ในการทดลองนี้ มีวัตถุประสงค์เพื่อวัดประสิทธิภาพจากการสกัดคุณลักษณะที่สร้างจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและแบ่งแยกประเภทกลุ่มข้อมูล ด้วยวิธีการวัดระยะทางแบบยุคลิด เพื่อหาตัวอธิบายคุณลักษณะที่สามารถแบ่งแยกภาพตัวอักษรที่มีความคล้ายคลึงกันได้ อีกทั้งยังมีความทนทานต่อการเปลี่ยนแปลงของชุดแบบของตัวอักษร ชนิด และขนาดของภาพตัวอักษร ซึ่งการทดลองจะถูกแบ่งออกเป็น 3 ส่วนได้แก่ 1) การทดสอบค่าความถูกต้องของตัวอธิบายคุณลักษณะที่ได้จากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ 2) การทดสอบขนาดของตัวอธิบายคุณลักษณะ และ 3) เวลาที่ใช้ในการสกัดคุณลักษณะที่ได้จากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ และการทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อความหลากหลายของขนาด และชุดแบบของตัวอักษร

4.2.1 การทดสอบความถูกต้องของตัวอธิบายคุณลักษณะ SD-AHOG

การทดลองนี้ได้แสดงค่าความถูกต้องของการรู้จำตัวอักษร โดยอาศัยฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะสำหรับการสกัดคุณลักษณะ โดยใช้การเลือกกลุ่มข้อมูลแบบเที่ยงตรง K แบบโดยที่ $K = 30$ จากตารางที่ 4.3 แสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานที่ได้จากค่าความถูกต้องที่คำนวณจากสมการที่ 2.4

ตารางที่ 4.3 ผลการเปรียบเทียบความถูกต้องของการรู้จำของ HOG, SSHOG, PHOG, MHOG, HOGC และ SD-AHOG

Method	Upper Level (%)				Middle Level (%)			
	Regular	Bold	Italic	Bold- Italic	Regular	Bold	Italic	Bold- Italic
HOG	94.02 ± 0.63	92.08 ± 0.63	94.85 ± 0.54	92.54 ± 0.61	96.28 ± 0.22	96.55 ± 0.19	96.36 ± 0.27	97.38 ± 0.20
SSHOG	94.62 ± 0.61	93.31 ± 0.60	95.25 ± 0.53	93.09 ± 0.71	97.24 ± 0.19	97.41 ± 0.17	97.32 ± 0.22	98.05 ± 0.16
PHOG	90.74 ± 0.70	91.34 ± 0.62	91.30 ± 0.77	90.60 ± 0.81	92.90 ± 0.27	93.89 ± 0.25	92.22 ± 0.32	93.67 ± 0.25
HOGC	86.96 ± 0.84	86.20 ± 0.94	86.87 ± 0.84	85.65 ± 0.82	90.16 ± 0.31	90.93 ± 0.35	89.83 ± 0.30	91.10 ± 0.28
MHOG	93.75 ± 0.67	93.04 ± 0.67	93.63 ± 0.74	91.84 ± 0.67	94.12 ± 0.29	94.71 ± 0.30	93.09 ± 0.31	94.29 ± 0.22
SD-AHOG	94.41 ± 0.78	94.40 ± 0.61	94.93 ± 0.55	94.54 ± 0.64	97.91 ± 0.18	98.22 ± 0.18	97.91 ± 0.22	98.24 ± 0.18

จากตารางที่ 4.3 ได้แสดงผลการทดลองของตัวอักษรภาษาไทยในระดับบนและระดับกลางเท่านั้น จะเห็นว่าค่าความถูกต้องโดยเฉลี่ยของ SD-AHOG จะมีสูงกว่าวิธีการอื่น ยกเว้นกรณีตัวเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปกติ และตัวเอียงในตัวอักษรภาษาไทยในระดับบนที่จะน้อยกว่า SSHOG โดยเฉลี่ย 0.21% และ 0.32% ตามลำดับ ในส่วนของตัวอักษรในระดับล่างนั้นมีค่าความถูกต้อง 100% ทุกกรณี เนื่องจาก AHOG ได้มีการปรับปรุงการแบ่งภาพออกเป็นส่วนย่อยแบบปรับเหมาะเพื่อให้คุณลักษณะที่ได้มีความละเอียดมากขึ้น และการกำหนดความกว้างของถังแบบปรับเหมาะยังช่วยลดการเกิดข้อผิดพลาดของการทำการควอนไทซ์ อีกทั้งยังเพิ่มความแม่นยำให้ SD-AHOG โดยการรวมพลังงานแบบเวกเตอร์ซึ่งอาศัยทั้งขนาดและทิศทาง โดยที่ค่าความถูกต้องเฉลี่ยของ AHOG, HOGC, SSHOG, PHOG, MHOG และ HOG คือ 97.59%, 90.13%, 97.04%, 93.01%, 94.05% และ 96.23% ตามลำดับ ดังนั้น SD-AHOG สามารถแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันได้ดีกว่าวิธีอื่น จึงส่งผลให้ SD-AHOG มีค่าความถูกต้องสูงกว่าวิธีอื่น

4.2.2 การเปรียบเทียบขนาดของตัวอธิบายคุณลักษณะและเวลาที่ใช้ในการสกัดตัวอธิบายคุณลักษณะ

ผลการทดลองแสดงให้เห็นถึงขนาดของคุณลักษณะของฮิสโตแกรมของการได้ระดับเชิงมุมแบบปรับเหมาะและเวลาสำหรับการสกัดคุณลักษณะดังตารางที่ 4.4 เมื่อเปรียบเทียบกับ HOG, SSHOG, PHOG, MHOG และ HOGC

ตารางที่ 4.4 ผลการเปรียบเทียบเวลาในการสกัดคุณลักษณะและจำนวนคุณลักษณะของ HOG, SSHOG, PHOG, MHOG, HOGC และ SD-AHOG

Method	HOG	SSHOG	PHOG	MHOG	HOGC	SD-AHOG
Feature Size	1,764	7,056	765	16,992	720	<u>576</u>
Computing Time	<u>0.008 s</u>	0.031 s	0.012 s	0.111 s	0.018 s	0.034 s

จากการเปรียบเทียบดังตารางที่ 4.4 พบว่า ขนาดของคุณลักษณะของฮิสโตแกรมของการได้ระดับเชิงมุมแบบปรับเหมาะมีขนาดเล็กที่สุด เพราะ AHOG ไม่มีการทำกระบวนการคาบเกี่ยวระหว่างเซลล์ (Overlap) เหมือนใน HOG แบบดั้งเดิม จึงทำให้ขนาดของคุณลักษณะลดลง และใช้เวลาในการสกัดคุณลักษณะน้อยกว่า MHOG แต่ใช้เวลามากกว่า HOG, SSHOG, PHOG และ HOGC เนื่องจากขั้นตอนการรวมพลังงานในแต่ละถังจะเป็นการรวมแบบเวกเตอร์ อีกทั้งในกระบวนการแบ่งภาพออกเป็นส่วนย่อย ๆ แบบปรับเหมาะของ AHOG จะต้องพิจารณาข้อมูลในแต่ละ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ละส่วนของภาพก่อนการกำหนดขนาด ซึ่งจะต้องใช้เวลามากกว่าการแบ่งภาพออกเป็นส่วนย่อยแบบคงที่

4.2.3 การทดสอบความทนทานของตัวอธิบายคุณลักษณะ

การทดสอบความทนทานของตัวอธิบายคุณลักษณะในการรู้จำตัวอักษร แบ่งเป็น 2 ส่วน คือ ความทนทานต่อความแตกต่างของชุดแบบของตัวอักษร และความทนทานต่อความแตกต่างของขนาดตัวอักษรด้วยวิธีการ AHOG, HOG, SSHOG, PHOG, MHOG และ HOGC ให้ผลการทดสอบดังแสดงในตารางที่ 4.5 โดยใช้การเลือกสุ่มข้อมูลแบบเที่ยงตรง K แบบ โดยมีเงื่อนไขสำหรับการแบ่งข้อมูล สำหรับการแบ่งข้อมูลตามขนาดของตัวอักษรจะสามารถแบ่งได้ 8 กลุ่ม เพื่อนำไปใช้ทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อขนาด และแบ่งข้อมูลตามชุดแบบของตัวอักษรจะสามารถแบ่งได้ 12 กลุ่ม เพื่อนำไปใช้ทดสอบความทนทานของตัวอธิบายคุณลักษณะต่อชุดแบบของตัวอักษร

ตารางที่ 4.5 ผลการทดสอบความทนทานในการรู้จำที่มีความแตกต่างของชุดแบบของตัวอักษรกับขนาดของตัวอักษรด้วย SD-AHOG, HOG, SSHOG, PHOG, MHOG และ HOGC

Method	Robustness Font			Robustness Size		
	Upper	Middle	Lower	Upper	Middle	Lower
HOG	87.14	86.86	100.00	93.48	96.97	100.00
SSHOG	87.70	88.01	100.00	94.50	97.90	100.00
PHOG	87.30	81.68	99.22	91.44	92.56	99.71
MHOG	86.88	83.52	100.00	92.65	93.83	100.00
HOGC	78.94	77.70	100.00	85.90	89.20	100.00
SD-AHOG	88.61	86.84	100.00	95.22	98.37	100.00

จากผลการทดลองจะพบว่า ตัวอธิบายคุณลักษณะแบบ SD-AHOG มีความทนทานต่อขนาดของตัวอักษรเฉลี่ยถึง 97.93% เมื่อเปรียบเทียบกับ HOG, SSHOG, PHOG, MHOG และ HOGC ซึ่งได้ความทนทานต่อขนาดของตัวอักษรเฉลี่ย 96.52%, 97.43%, 92.57%, 93.80% และ 88.98% ตามลำดับ แต่จะพบว่าความทนทานต่อชุดแบบของตัวอักษร จะพบปัญหาในกรณีเดียวกับคุณลักษณะที่ได้จากการแยกองค์ประกอบค่าเดียว คือ ชุดแบบของตัวอักษร JasmineUPC ทำให้ค่าความทนทานเฉลี่ยต่อชุดแบบอักษรของ SD-AHOG, SSHOG, PHOG, MHOG, HOG และ HOGC ลดลงเหลือ 87.45%, 88.27%, 82.98%, 84.44%, 87.23% และ 78.45% ตามลำดับ

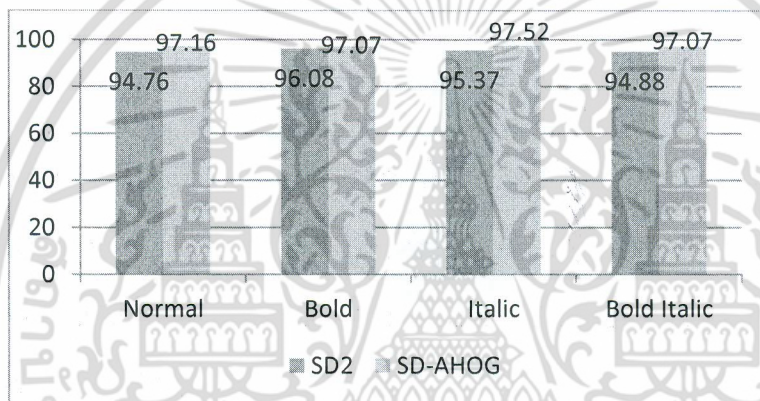
เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 ผลการทดลองที่ 3: การเปรียบเทียบประสิทธิภาพการรู้จำของ SD2 และ SD-AHOG

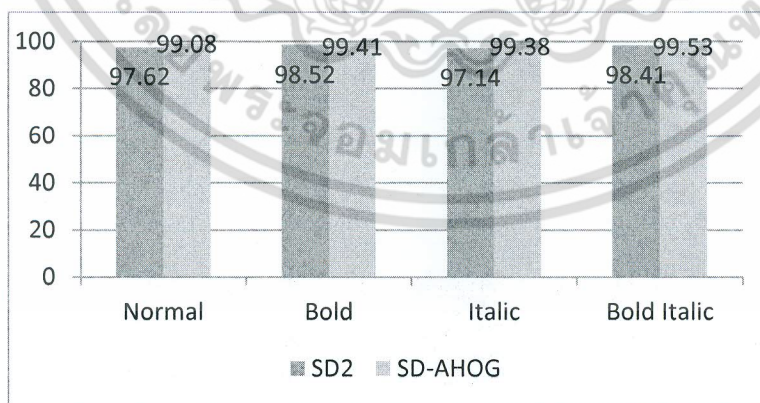
ผลการทดลองเปรียบเทียบประสิทธิภาพการรู้จำและขนาดของคุณลักษณะที่ได้จากการใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (SD-AHOG) กับการสกัดคุณลักษณะโดยการแยกองค์ประกอบค่าเดี่ยว (SD2)

4.3.1 การเปรียบเทียบความถูกต้องของตัวอธิบายคุณลักษณะ

ผลการทดลองความถูกต้องของการรู้จำตัวอักษรภาษาไทยโดยการใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะกับการสกัดคุณลักษณะโดยการแยกองค์ประกอบค่าเดี่ยวจะมี 3 ส่วนโดยแบ่งตามระดับของตัวอักษรดังรูปที่ 4.8-4.10

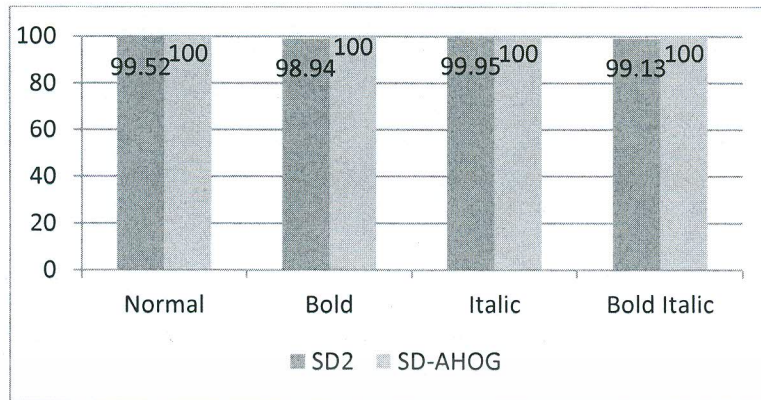


รูปที่ 4.8 การเปรียบเทียบความถูกต้องของการรู้จำระหว่าง SD2 กับ SD-AHOG ของตัวอักษรภาษาไทยในระดับบน



รูปที่ 4.9 การเปรียบเทียบความถูกต้องของการรู้จำระหว่าง SD2 กับ SD-AHOG ของตัวอักษรภาษาไทยในระดับกลาง

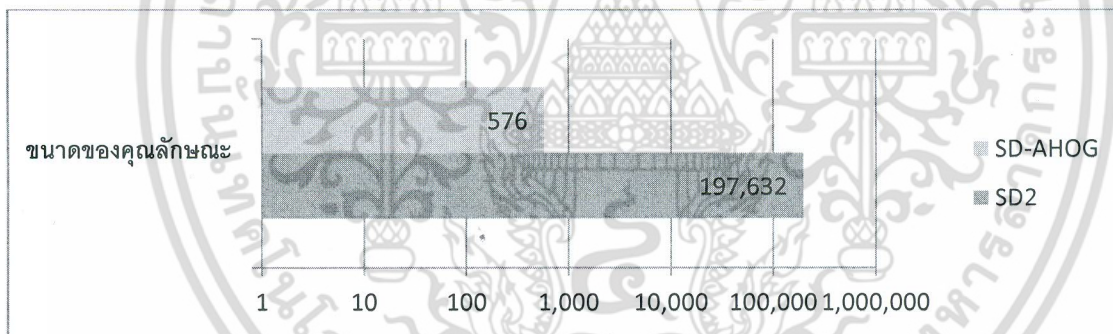
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



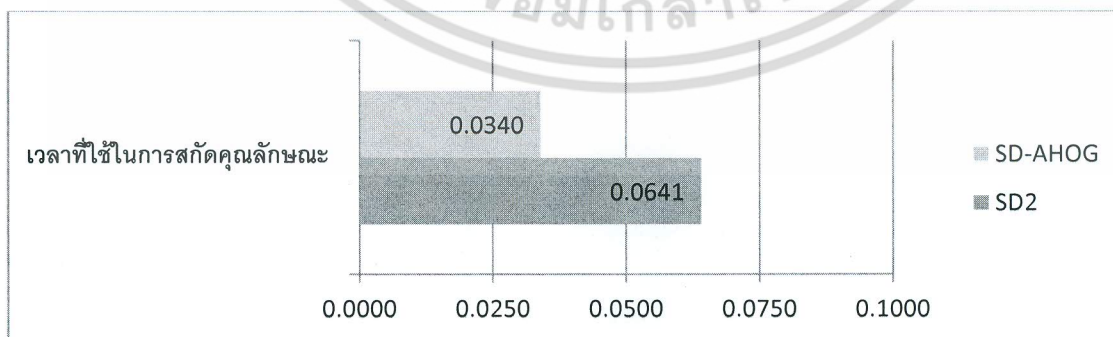
รูปที่ 4.10 การเปรียบเทียบความถูกต้องของการรู้จำระหว่าง SD2 กับ SD-AHOG ของตัวอักษรภาษาไทยในระดับบนล่าง

4.3.2 การเปรียบเทียบขนาดคุณลักษณะและเวลาที่ใช้ในการสกัดคุณลักษณะ

ผลการทดลองแสดงขนาดของคุณลักษณะและเวลาที่ใช้ในการสกัดคุณลักษณะระหว่างคุณลักษณะของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและคุณลักษณะของการแยกองค์ประกอบค่าเดี่ยวดังรูปที่ 4.11 และ 4.12



รูปที่ 4.11 การเปรียบเทียบขนาดของคุณลักษณะระหว่าง SD2 กับ SD-AHOG



รูปที่ 4.12 การเปรียบเทียบเวลาที่ใช้ในการสกัดคุณลักษณะระหว่าง SD2 กับ AHOG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.3 การเปรียบเทียบความทนทานของตัวอธิบายคุณลักษณะ

ผลการทดลองเพื่อเปรียบเทียบความทนทานของตัวอธิบายคุณลักษณะที่ได้จากการสกัดคุณลักษณะด้วยวิธีฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะและการแยกองค์ประกอบค่าเดียวในด้านความทนทานต่อการเปลี่ยนแปลงขนาด และการเปลี่ยนแปลงชุดแบบของตัวอักษรแสดงดังตารางที่ 4.6

ตารางที่ 4.6 การเปรียบเทียบความทนทานต่อขนาดและชุดแบบอักษรระหว่าง SD2 กับ SD-AHOG

	Robustness Font			Robustness Size		
	Upper Level	Middle Level	Lower Level	Upper Level	Middle Level	Lower Level
SD2	85.65	86.49	95.67	90.75	96.07	97.48
SD-AHOG	88.61	86.84	100.00	95.22	98.37	100.00

จากผลการทดลองในหัวข้อ 4.3 พบว่า การรู้จำตัวอักษร โดยอาศัยคุณลักษณะที่ได้จากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ ให้ผลความถูกต้องมากกว่าวิธีการแยกองค์ประกอบค่าเดียว อีกทั้งขนาดของตัวอธิบายคุณลักษณะยังมีขนาดเล็กอีกด้วย และเมื่อเปรียบเทียบ SD-AHOG กับ SD2 แล้วพบว่า SD-AHOG ยังมีความทนทานต่อขนาดและชุดแบบของตัวอักษรมากกว่า SD2 นอกจากนี้ วิธีฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะยังใช้เวลาสำหรับการสกัดคุณลักษณะที่น้อยกว่าการแยกองค์ประกอบค่าเดียว ดังนั้น จึงสามารถสรุปได้ว่า SD-AHOG มีประสิทธิภาพดีกว่าในด้านของความถูกต้องและขนาดของคุณลักษณะที่ใช้งาน

4.4 ผลการทดลองที่ 4: การเปรียบเทียบประสิทธิภาพการรู้จำโดยจำแนกตามกรณี

ผลการทดลองเปรียบเทียบประสิทธิภาพการรู้จำโดยจำแนกตามกรณีต่าง ๆ ได้แก่ การรู้จำถูกต้อง (TP), การรู้จำผิดพลาดในกลุ่มที่มีความคล้ายคลึงกัน (FP) และการรู้จำผิดพลาดในกลุ่มที่ไม่มีความคล้ายคลึงกัน (FN) ของ SD1, SD2, SD-AHOG, SSHOG, PHOG, MHOG, HOG และ HOGC โดยจัดกลุ่มของตัวอักษรที่มีความคล้ายคลึงกันตามการแข่งขัน NSC2013 และปรับเปลี่ยนกลุ่มของตัวอักษรที่มีความคล้ายคลึงกันตามผลการทดลองของ SD1 รวมถึงลักษณะทางกายภาพที่มีความคล้ายคลึงกันรายละเอียดดังตารางที่ 4.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และ MHOG มีค่า FN น้อยกว่า SD2 และ SD-AHOG โดยรูปที่ 4.13 จะแสดงให้เห็นถึงกรณี
SD-AHOG รู้จำผิดพลาดในกรณีของ FN แต่ SSHOG สามารถรู้จำตัวอักษรได้ถูกต้อง



รูปที่ 4.12 ผลลัพธ์ของ AHOG ในกรณีการรู้จำผิดพลาดในกลุ่มที่ไม่มีความคล้ายคลึงกัน
(ก) ภาพอินพุตคือ “อ” ผลลัพธ์คือ “อ”, (ข) ภาพอินพุตคือ “อ” ผลลัพธ์คือ “อ”, (ค) ภาพอินพุตคือ
“อ” ผลลัพธ์คือ “อ” และ (ง) ภาพอินพุตคือ “อ” ผลลัพธ์คือ “อ”

ในส่วน of ตัวอักษรภาษาไทยในระดับกลางจะพบว่า SD-AHOG และ SD2 มีความถูกต้อง
มากที่สุด และสามารถแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันได้ดี เมื่อพิจารณาจากค่า TP และ FP
ตามลำดับ แต่เมื่อพิจารณาค่า FN จะพบว่า SSHOG สามารถแบ่งแยกตัวอักษรที่ไม่มีความคล้ายกัน
ได้ดีกว่า SD-AHOG โดยเฉลี่ย 0.03% ดังรูปที่ 4.14 ซึ่งแสดงกรณีที่ SD-AHOG รู้จำผิดพลาด ใน
กรณีของ FN แต่ SSHOG สามารถรู้จำตัวอักษรได้ถูกต้อง ในตัวอักษรภาษาไทยในระดับล่างนั้น
SD-AHOG, HOG, SSHOG, MHOG และ HOGC มีค่า TP และ FN เท่ากันคือ 100%, 0% ตามลำดับ
ซึ่งดีกว่า SD1 และ SD2



รูปที่ 4.13 ผลลัพธ์ของ AHOG ในกรณีการรู้จำผิดพลาดในกลุ่มที่ไม่มีความคล้ายคลึงกัน
(ก) ภาพอินพุตคือ “จ” ผลลัพธ์คือ “จ”, (ข) ภาพอินพุตคือ “อ” ผลลัพธ์คือ “อ”, (ค) ภาพอินพุตคือ
“ช” ผลลัพธ์คือ “ช”, (ง) ภาพอินพุตคือ “ฮ” ผลลัพธ์คือ “ฮ”, (จ) ภาพอินพุตคือ “ท”
ผลลัพธ์คือ “ท”, (ฉ) ภาพอินพุตคือ “ค” ผลลัพธ์คือ “ค”, (ช) ภาพอินพุตคือ “ว” ผลลัพธ์คือ “ว”, (ซ) ภาพอินพุตคือ
“ล” ผลลัพธ์คือ “ล”, (ญ) ภาพอินพุตคือ “ท” ผลลัพธ์คือ “ท” และ (ฒ) ภาพอินพุตคือ “จ” ผลลัพธ์
คือ “จ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้มุ่งเน้นเพื่อเพิ่มประสิทธิภาพของการรู้จำตัวพิมพ์อักษรภาษาไทยโดยการแก้ปัญหาจาก 2 ปัจจัยหลัก คือ ปัจจัยความคล้ายกันของตัวอักษรภาษาไทย และปัจจัยความหลากหลายของชุดแบบตัวอักษร ชนิดของตัวอักษร และขนาดตัวอักษร โดยการสร้างตัวอธิบายคุณลักษณะจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ ที่สามารถแบ่งแยกข้อมูลที่มีความคล้ายคลึงกันและทนทานต่อความหลากหลายของชุดแบบตัวอักษร ชนิดของตัวอักษร และขนาดตัวอักษร

งานวิจัยนี้ได้พัฒนาขั้นตอนการเตรียมข้อมูล การสกัดคุณลักษณะ และการรู้จำตัวอักษรภาษาไทยด้วยโปรแกรม Matlab ซึ่งในส่วนของ การรู้จำตัวอักษรภาษาไทยนั้นจะมี 2 วิธีการ คือ การสร้างตัวอธิบายคุณลักษณะโดยอาศัยการแยกองค์ประกอบค่าเดี่ยวซึ่งตัวอธิบายคุณลักษณะ SD1 และ SD2 โดยสร้างมาจากการเลือกไอเกนเวกเตอร์ในลำดับที่แตกต่างกันโดยที่ SD1 สร้างจากไอเกนเวกเตอร์ใน “ลำดับต้น” และ SD2 สร้างมาจากไอเกนเวกเตอร์ใน “ลำดับปลาย” และการสร้างตัวอธิบายคุณลักษณะจากฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (SD-AHOG)

5.1.1 ผลงานวิจัยของการสกัดคุณลักษณะโดยการแยกองค์ประกอบค่าเดี่ยว

ในการทดลองที่ 1 ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของการรู้จำตัวอักษรระหว่าง SD1 กับ SD2 ซึ่งจากผลการทดลองพบว่าการรู้จำด้วยวิธีการรู้จำด้วย SD1 ได้ค่าความถูกต้องเฉลี่ย 86% แต่ในวิธีการรู้จำด้วยวิธี SD2 ได้ค่าความถูกต้องเฉลี่ย 97.55% ดังนั้น จึงสรุปได้ว่า การรู้จำด้วยวิธี SD2 มีประสิทธิภาพที่ดีกว่าการรู้จำด้วยวิธี SD1 เมื่อภาพตัวอักษรไม่มีสัญญาณรบกวน ซึ่งช่วงที่เหมาะสมของ SD2 คือ ไอเกนเวกเตอร์ในช่วง 512 ถึง 704

ผู้วิจัยได้ทำการทดสอบความทนทาน (Robustness) ของตัวอธิบายคุณลักษณะในการการรู้จำตัวอักษรด้วยวิธีการของ SD2 ซึ่งการทดลองถูกแบ่งออกเป็น 2 ส่วน คือ ความทนทานเมื่อความแตกต่างของขนาด (Size) ตัวอักษร ซึ่งมีค่าความถูกต้องเฉลี่ย 94.77% และมากกว่า SD1 อยู่ 13.53% ในส่วนของความทนทานเมื่อความแตกต่างของชุดแบบตัวอักษร (Font) ซึ่งได้ค่าความถูกต้องเฉลี่ย 89.27% โดยมากกว่า 7.50% ของ SD1 เนื่องจากปัญหาในชุดแบบของตัวอักษร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

JasmineUPC ซึ่งมีลักษณะรูปร่างที่แตกต่างจากชุดแบบของตัวอักษรอื่น และคล้ายคลึงกับพยัญชนะที่มีความคล้ายกันในชุดแบบของตัวอักษรอื่น

จากผลการทดลองสามารถสรุปได้ว่าการรู้จำโดยใช้ตัวอธิบายคุณลักษณะจากการแยกองค์ประกอบค่าเดียวมีความทนทานต่อขนาดที่มีความหลากหลาย และสามารถทนทานได้ดีกับชุดแบบของตัวอักษรที่มีความหลากหลาย แต่จะเห็นว่าขนาดของคุณลักษณะที่ได้จากการแยกองค์ประกอบค่าเดียวนั้น สูงถึง 197,632 ตัว ซึ่งส่งผลต่อพื้นที่ในการจัดเก็บตัวอธิบายคุณลักษณะ อีกทั้งยังส่งผลต่อเวลาในการแบ่งแยกประเภทภาพอักษรภาษาไทย

5.1.2 ผลงานวิจัยของการสกัดคุณลักษณะโดยใช้ฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ

ในการทดลองที่ 2 ผู้วิจัยได้ทำการทดลองเปรียบเทียบประสิทธิภาพของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ (SD-AHOG) ที่ผู้วิจัยเป็นผู้พัฒนาขึ้น เมื่อเทียบกับวิธีการในกลุ่มของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบอื่นๆ เช่น HOG, PHOG, HOGC, MHOG และ SSHOG ซึ่งจากผลการทดลองพบว่าตัวอธิบายคุณลักษณะของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะมีค่าการรู้จำถูกต้องโดยเฉลี่ย 97.59% ซึ่งถือว่ามีประสิทธิภาพมากกว่าตัวอธิบายคุณลักษณะของ HOG, SSHOG, PHOG, MHOG และ HOGC ที่มีค่าความถูกต้องอยู่ที่ 96.23%, 97.04, 93.01%, 94.05% และ 90.13% ตามลำดับ อีกทั้งเมื่อพิจารณาเปรียบเทียบขนาดคุณลักษณะของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะ กับขนาดของคุณลักษณะของ HOG, PHOG, HOGC, MHOG และ SSHOG พบว่าขนาดของคุณลักษณะของฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะเท่ากับ 576 ตัว ซึ่งมีขนาดเล็กที่สุดเมื่อเปรียบเทียบกับขนาดของคุณลักษณะของ HOG, PHOG, HOGC, MHOG และ SSHOG

ผู้วิจัยได้ทำการทดสอบความทนทานของตัวอธิบายคุณลักษณะในการการรู้จำตัวอักษรด้วยวิธีการฮิสโตแกรมของการไล่ระดับเชิงมุมแบบปรับเหมาะซึ่งการทดลองถูกแบ่งออกเป็น 2 ส่วน คือ ความทนทานเมื่อความแตกต่างของขนาดตัวอักษรได้ค่าความถูกต้องเฉลี่ย 97.93% และความทนทานเมื่อความแตกต่างของชุดแบบตัวอักษรซึ่งได้ค่าความถูกต้องเฉลี่ย 87.45% การที่ค่าความถูกต้องลดลงในกรณีนี้ มีเหตุผลสนับสนุนเดียวกับ 5.1.1 ในส่วนของการทดลองความทนทาน

5.1.3 ผลงานวิจัยของการเปรียบเทียบการสกัดคุณลักษณะ

เมื่อนำตัวอธิบายคุณลักษณะของวิธีการแรก (SD2) กับตัวอธิบายคุณลักษณะของวิธีการสอง (SD-AHOG) มาเปรียบเทียบประสิทธิภาพในด้านค่าความถูกต้องของการรู้จำกับขนาดของคุณลักษณะที่ใช้จะพบว่า SD-AHOG สามารถรู้จำได้สูงกว่า SD2 โดยเฉลี่ย 0.05% อีกทั้ง SD-เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

AHOG ยังมีขนาดของคุณลักษณะน้อยกว่า SD2 ถึง 343 เท่า เวลาที่ใช้สำหรับการสกัดคุณลักษณะของฮิสโตแกรมของการไต่ระดับเชิงมุมแบบปรับเหมาะน้อยกว่า 2 เท่า ในส่วนของความทนทานต่อการเปลี่ยนแปลงต่อขนาดและชุดแบบของตัวอักษรจะพบว่า SD-AHOG สามารถทนทานได้ดีกว่า SD2 อยู่ที่ 2.64% และ 0.85% ตามลำดับ ดังนั้น จึงสามารถสรุปได้ว่าตัวอธิบายคุณลักษณะของงานวิจัยนี้ คือ SD-AHOG เป็นตัวอธิบายคุณลักษณะที่มีประสิทธิภาพสำหรับภาพตัวอักษรภาษาไทยและยังมีขนาดของคุณลักษณะที่เหมาะสม เมื่อภาพตัวอักษรไม่มีสัญญาณรบกวน

5.1.4 ผลงานวิจัยของการเปรียบเทียบประสิทธิภาพการรู้จำโดยจำแนกตามกรณีต่าง ๆ

กรณีตัวอักษรภาษาไทยในระดับบน SD2 มีประสิทธิภาพที่ดีกว่า SD-AHOG ในเรื่องของความถูกต้องของการรู้จำ และความสามารถในการแยกตัวอักษรที่มีความคล้ายคลึงกันอยู่ 0.7% และ 1.09% ตามลำดับ ในส่วนของตัวอักษรภาษาไทยในระดับกลางพบว่า SD-AHOG มีความถูกต้องมากที่สุด และสามารถแบ่งแยกตัวอักษรที่คล้ายคลึงกันได้ดี ในส่วนของตัวอักษรภาษาไทยในระดับล่างพบว่า SD-AHOG, HOG, SSHOG, MHOG และ HOGC มีประสิทธิภาพที่ดีกว่า PHOG, SD1 และ SD2 ในรูปที่ 4.13 และ 4.14 ได้แสดงกรณี SD-AHOG รู้จำผิดพลาด แต่ SSHOG สามารถรู้จำถูกต้อง ซึ่งสามารถกล่าวได้ว่า SD2 และ SD-AHOG มีความสามารถในการแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันได้มากกว่าวิธีการอื่น โดยพิจารณาจากค่า FP แต่ถ้าพิจารณาค่า FN ของ SD2 และ SD-AHOG จะพบว่า SD-AHOG มีค่าที่มากกว่า SSHOG แต่เมื่อเทียบกับประสิทธิภาพการรู้จำตัวอักษรโดยภาพรวม ขนาดของคุณลักษณะและความสามารถในการแบ่งแยกตัวอักษรที่มีความคล้ายคลึงกันจะพบว่า SD2 และ SD-AHOG มีประสิทธิภาพที่ดีกว่าวิธีการอื่น

5.2 ข้อเสนอแนะ

ในงานวิจัยนี้ได้ทำการสร้างตัวอธิบายคุณลักษณะที่ได้จากชุดข้อมูลฝึกฝน โดยวิธีการแยกองค์ประกอบค่าเดียวกับฮิสโตแกรมของการไต่ระดับเชิงมุมแบบปรับเหมาะและปัญหาที่ได้อพบในการทำงานวิจัยครั้งนี้ รวมถึงแนวคิดที่จะใช้ปรับปรุงในอนาคต ผู้วิจัยมีข้อเสนอแนะดังนี้

- ในงานวิจัยส่วนของ SD1 และ SD2 จะพบว่า SD2 มีประสิทธิภาพมากกว่า SD1 แต่เนื่องจาก SD2 ถูกสร้างโดยไอเกนเวกเตอร์ในลำดับปลาย ดังนั้น SD2 อาจจะมีไหวต่อสัญญาณรบกวนมากกว่า SD1 ที่ถูกสร้างจากไอเกนเวกเตอร์ในลำดับต้น และอาจจะทำให้ความถูกต้องลดลงเมื่อภาพอักษรมีสัญญาณรบกวน
- ในงานวิจัยส่วนของขั้นตอนการรู้จำตัวอักษรโดยอาศัยหลักการการฉายภาพเมตริกซ์และการวัดระยะทางแบบยุคลิด ซึ่งใช้เวลาค่อนข้างมากในการรู้จำ ดังนั้น ถ้ามีการประยุกต์ใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลักการในการรู้จำด้วยวิธีอื่น เช่น Support Vector Machine (SVM) จะช่วยลดเวลาที่ใช้ในขั้นตอนการแบ่งแยกภาพอักษรได้

- ในงานวิจัยนี้ถ้าพิจารณาผลการทดลองในส่วนความทนทานต่อขนาดและชุดแบบของตัวอักษรจะพบว่าตัวอธิบายคุณลักษณะที่ได้จากระบบการกับฮิสโตแกรมของการได้ระดับเชิงมุมแบบปรับเหมาะมีความทนทานต่อขนาดที่ดีมาก แต่ในทางกลับกัน ความทนทานต่อชุดแบบของตัวอักษรนั้นจะน้อยกว่าความทนทานต่อขนาดอยู่เล็กน้อย



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- Andrew, J.N., Lewis, D.G.: Multiscale Histogram of Oriented Gradient Descriptors for Robust Character Recognition. **In: 2011 International Conference on Document Analysis and Recognition**, pp. 1085–1089. IEEE Press, New York (2011)
- Anna, B., Andrew, Z., Xavier, M.: Representing Shape with a Spatial Pyramid Kernel. **In: the 6th ACM International Conference on Image and Video Retrieval**, (2007)
- Bovik, A. 2000. **Handbook of Image & Video Processing**. USA : Academic Press.
- Duangphasuk, S. “**Thai Printed Character Recognition Using Feature Matching Method and ART1**” M.S. Thesis of Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok. 2001
- Farag, A.A. “**A Tutorial on Principal Component Analysis.**” [Online]. Available : <http://dai.fmph.uniba.sk/courses/ml/sl/PCA.pdf>. 2009.
- Gerbrands, J.J.: ON THE RELATIONSHIPS BETWEEN SVD, KLT AND PCA. **In: 1981 Pattern Recognition**, pp. 375–381. Pergamon Press, Great Britain (1981)
- Jitsopat, O. 2009. **Digital Image Processing**. Thailand : Skprintingmedia.
- Kimpan, C., Itoh, A., Kawanishi, K. “Fine Classification of Printed Thai Character Recognition Using the Karhunen-Loeve Expansion” **In: IEE Proceedings**, vol.134. 1987. pp. 257- 264.
- Kruatrachue, B., Pantrakarn, N., Siriboon, K. “Automatic State Machine Induction for String Recognition” **In: Proceedings of World Congress on Engineering**, 2007. pp. 153-158.
- Lauffenburger, D., Matsudaira, P., Khademhosseini, A. “**Singular Value Decomposition (SVD) tutorial.**” [Online]. Available : http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm. 2002.

บรรณานุกรม (ต่อ)

- Navneet, D., Bill, T.: Histograms of Oriented Gradient for Human Detection. **In: IEEE International Conference on Computer Vision and Pattern Recognition**, pp. 1–8. IEEE Press, New York (2005)
- Ning, H., Jiaheng, C., Lin, S.: Scale Space Histogram of Oriented Gradient for Human Detection. **In: 2008 International Symposium on Information Science and Engineering**, pp. 167–170. IEEE Press, New York (2008)
- Subhransu, M., Alexander, C.B., Jitendra, M.: Classification Using Intersection Kernel Support Vector Machines is Efficient. **In: IEEE International Conference on Computer Vision and Pattern Recognition**, IEEE Press, New York (2008)
- Suttapakti, U., Woraratpanya, K., Pasupa, K.: Font Descriptor Construction for Printed Thai Character Recognition. **In: 2013 International Conference on Machine Vision Applications**, pp. 45–48. MVA Press, Kyoto, Japan (2013)
- Suparatei, S. 2011. คู่มือการใช้ MATLAB ฉบับสมบูรณ์. Thailand : IDC Premier.
- Tanprasert, C., Koanantakool, T. “Thai OCR : A Neural Network Application” **In: Proceedings of the 1996 IEEE TENCON – Digital Signal Processing Applications**, 1996. pp. 90-95
- Tanprasert, C., Sae-tang, S.: Thai type style recognition. **In: 1999 IEEE International Symposium on Circuits and System**, pp. 336–339. IEEE Press, Orlando, FL (1999)
- Tangsurakit, N. “Thai Printed Consonant Recognition Based on Character Density and Strip Feature” M.S. Thesis of Faculty of Telecommunication Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok. 2005
- Udomhunsakun, S. 2011. **Fundamentals of Digital Image Processing**. Thailand : Top Publishing.
- Woraratpanya, K., Titijaroonroj, T.: Printed Thai Character Recognition Using Standard Descriptor. **In: 9th International Conference on Computing and Information Technology**, pp. 165–173, Springer-Verlag Berlin Heidelberg (2013)

บรรณานุกรม (ต่อ)

Woraratpanya, K., Titijaronroj, T.: Adaptive Histogram of Oriented Gradient for Printed Thai Character Recognition. In: **10th International Conference on Computing and Information Technology**, pp. 336–339. Springer-Verlag Berlin Heidelberg, Phuket, Thailand (2014)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก.

ผลงานวิจัยที่ได้รับการตีพิมพ์

- [1] Woraratpanya, K., Titijaroonroj, T.: **Printed Thai Character Recognition Using Standard Descriptor.** In: 9th International Conference on Computing and Information Technology, pp. 165–173, Springer-Verlag Berlin Heidelberg (2013)
- [2] Woraratpanya, K., Titijaroonroj, T.: **Adaptive Histogram of Oriented Gradient for Printed Thai Character Recognition.** In: 10th International Conference on Computing and Information Technology, pp. 83-91, Springer-Verlag Berlin Heidelberg (2014)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING 209

Phayung Meesad
Herwig Unger
Sirapat Boonkrong (Eds.)

**The 9th International
Conference on Computing
and Information
Technology (IC²IT2013)**

May 9th–10th, 2013

King Mongkut's University of Technology North Bangkok,
Bangkok, Thailand

 Springer

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Printed Thai Character Recognition Using Standard Descriptor

Kuntpong Woraratpanya and Taravichet Titijaroonroj

Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang, Thailand
(kuntpong, bank2533)@gmail.com

Abstract. The various font-types, font-sizes, and font-styles have a great impact on recognition performance of optical character recognition (OCR) systems. This becomes a grand challenge for recognition improvement. In order to enhance the performance, this paper proposes the printed Thai character recognition using a standard descriptor. The descriptor construction consists of two principal phases—preprocessing and feature extraction. In the former phase, the preprocessing provides a standard form for each character image. In the latter phase, the singular value decomposition (SVD) is applied to all font-type, font-size, and font-style character images to extract features. Then the standard descriptor is constructed from the suitable order selection of the SVD feature decomposition. Finally, the projection matrix technique is applied to the recognition phase in order to measure the cosine similarity between the standard descriptor and test set. The experimental results show that the proposed method achieves a high recognition rate and is invariant to font-types, font-sizes, and font-styles.

Keywords: optical character recognition (OCR), standard descriptor, singular value decomposition (SVD), feature extraction.

1 Introduction

An evolution of a variety of printed Thai fonts leads to the difficulty of retaining a high recognition rate of optical character recognition (OCR) systems. This becomes a grand challenge for recognition improvement. Over the last 20 years, printed Thai-character recognitions have been continually researched. Most of the approaches focus on an improvement of recognition rate. For instances, Kimpan et al. [1] introduced to fine classification of printed Thai character recognition using the Karhunen-Loeve expansion. This work uses a two-step approach, i.e., the lowest order of eigenvectors for rough classification and the higher order of eigenvectors for fine classification. It achieves the high recognition rate in testing with a standard font. Duang-phasuk et al. [2] presented printed Thai character recognition using feature matching and adaptive resonance theory I (ART I). This approach also requires two steps for feature extraction and recognition, i.e., global features for rough classification and local features for fine classification. It provides the high recognition rate

P. Meccad et al. (Eds.): *IC²IT2013*, AISC 209, pp. 165–173.

DOI: 10.1007/978-3-642-37371-8_20 © Springer-Verlag Berlin Heidelberg 2013

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

when experimented on various fonts and a few font-sizes, but takes more time-consuming. Tangsurakit et al. [3] proposed printed Thai consonant recognition based on character density and strip features. This approach is evaluated with only five basic fonts of consonance characters. In addition, Kruatrachue et al. [4] presented automatic state machine induction for string recognition that focuses on an enhancement of character classifications. The result of recognition is rather high efficiency, but in case of very similar characters the recognition rate is low.

However, as discussed previously, none of these papers point out the problems of a variety of printed Thai font-types, font-sizes, and font-styles. A few papers studied on font styles and types. For examples, Tanprasert et al [5] proposed Thai type style recognition, and Thammano et al. [6] presented hierarchical cross-correlation ARTMAP neural network for recognizing printed Thai characters of no-head fonts.

Therefore, this paper proposes a standard descriptor, which is a modularity method, to improve the recognition performance of OCR systems. The standard descriptor is made from all font-type, font-size, and font-style character features extracted by using SVD. It is believed that the proposed descriptor helps improve the recognition performance of OCR systems without reengineering software.

2 Standard Descriptor Construction

A standard descriptor plays an important role in representative of various font-types, font-sizes and font-styles, thus leading to the performance improvement of OCR. This section describes the effective standard descriptor construction consisting of two principal procedures, preprocessing and feature extraction. In addition, projection matrix technique applied to a recognition procedure is presented in the last subsection.

2.1 Preprocessing Procedure

Each character segmented from a document image is usually in different sizes. In order to provide a standard form for character segmentation images to construct the standard descriptor, image complementation, zero padding, and resizing procedures are applied to each character image as shown in Fig. 1(a). As a result, the character segmentation image is in a standard form of 32×32 pixels and zero backgrounds. Fig. 1(b) and (c) show a matrix of different font-types and font-sizes and a matrix of a standard form, respectively.

2.2 Feature Extraction

All matrices of a standard form in Fig. 1(c) are transformed into image vectors as depicted in Fig. 2. Then a training matrix is formed from such image vectors, which are contained features of all font-types, font-sizes, and font-styles. The procedure to construct a standard descriptor can be demonstrated in Fig. 3.

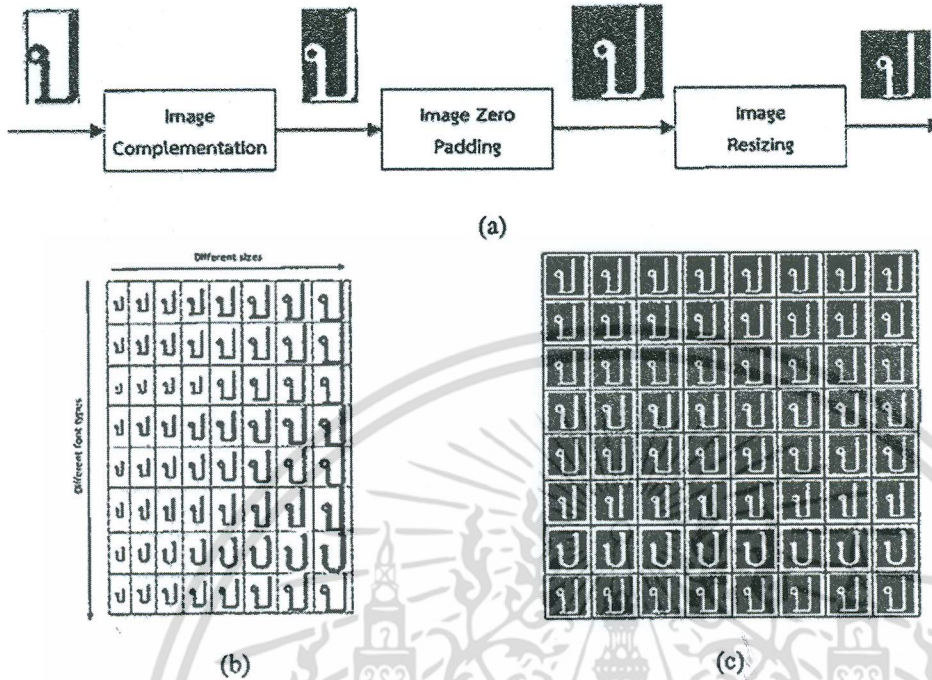


Fig. 1. Preprocessing procedures: (a) image complementation, zero padding, and resizing procedures, (b) matrices of different font-types and font-sizes, and (c) matrices of standard forms

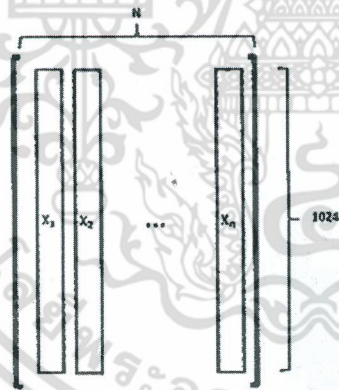


Fig. 2. A training matrix formed from image vectors

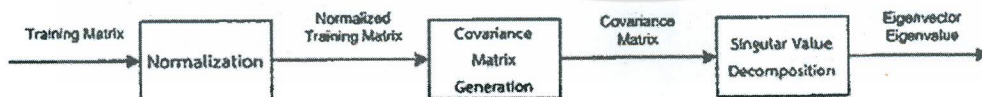


Fig. 3. Standard descriptor construction procedures

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Initially, the training matrix is normalized by using Eq. (1). The result is a normalized training matrix. Then a covariance matrix is generated from the normalized training matrix by subtracting mean and determining matrix elements by using Eq. (2).

$$\vec{u} = \frac{u}{\|u\|}; \quad \|u\| = \sqrt{u_1^2 + u_2^2 + \dots + u_N^2} \quad (1)$$

$$C_x = \frac{1}{N} \sum_{i=1}^N (X_i - M)(X_i - M)^T, \quad (2)$$

where X_i , N , and M are an image vector, a number of image vectors, and a mean value, respectively. Finally, the features are extracted from the covariance matrix by using SVD defined as Eq (3).

$$A = U \times S \times V^T, \quad (3)$$

where U , V , and S denote left eigenvector, right eigenvectors, and eigenvalue, respectively. The left and right eigenvectors are represented as $U=AA^T$ and $V=A^T A$, respectively. Finally, the standard descriptor is constructed by means of the suitable range of the higher order eigenvectors.

2.3 Recognition Procedure

The projection matrix technique is applied to the recognition phase in order to measure the cosine similarity between the standard descriptor and test set. In general, the misclassification occurs when the shape of characters is similar. It can be demonstrated with Fig. 4. This example shows features of four similar characters—ช, ฅ, ซ, ซ—in a vector space. As mentioned in the previous section, the lower order eigenvectors of such characters are represented by $P(\text{ช})$, $P(\text{ฅ})$, $P(\text{ซ})$, and $P(\text{ซ})$. On the other hand, the higher order eigenvectors are represented by $H(\text{ช})$, $H(\text{ฅ})$, $H(\text{ซ})$, and $H(\text{ซ})$. It is evident that the lower order eigenvectors of four similar characters have the same directions, while the higher order eigenvectors have different directions. In other words, the higher order eigenvectors provide an evident classification. Suppose that there is a test vector $y(\text{ช})$ in a vector space as depicted in Fig. 4 and the cosine similarity is applied to measure the angle between P and y . In this case the cosine function gives a maximum value, since the angle is very small. This leads to the misclassification; that is, ช can be recognized as ฅ, ซ, or ซ. Conversely, when the cosine similarity is applied to measure the angle between H and y , and the classification criteria is minimal; it is a largest angle. In this case the test vector can be recognized properly. It is summarized that the higher order eigenvectors provide the good features to classify

the similar character shapes. Therefore, this paper proposes the construction of the efficient standard descriptor based on the higher order eigenvectors and the use of the projection matrix technique with minimal cosine similarity measure for recognition.

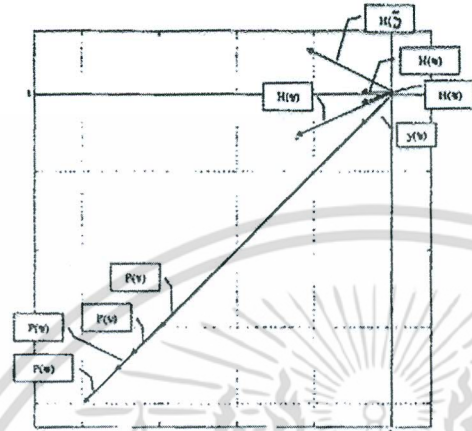


Fig. 4. A plot of the lower order eigenvectors versus the higher order eigenvectors

The overall recognition procedure is illustrated in Fig. 5. The manipulation begins with a test vector, y , is normalized with standard descriptor norms— $\|sd_1\|$, $\|sd_2\|$, ..., $\|sd_n\|$. This step generates n test unit vectors— y_1, y_2, \dots, y_n . Then the projection matrix technique is applied to obtain cosine similarity values— s_1, s_2, \dots, s_n —calculated from inner product of test unit vectors and standard descriptors. Finally, the minimum value of a set s_n is selected for recognition.

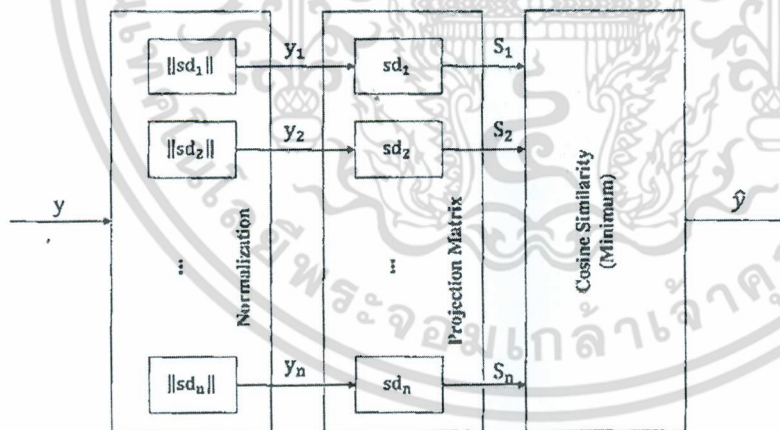


Fig. 5. A recognition procedure by using projection matrix technique

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3 Experimental Results

As mentioned in the subsection 2.2, the standard descriptor constructed by SVD method depends on the suitable feature selection. Hence, the feature selection based on experiments is analyzed in order to generate the efficient standard descriptor. Furthermore, the evaluation of recognition performance is also illustrated.

3.1 Data Preparation

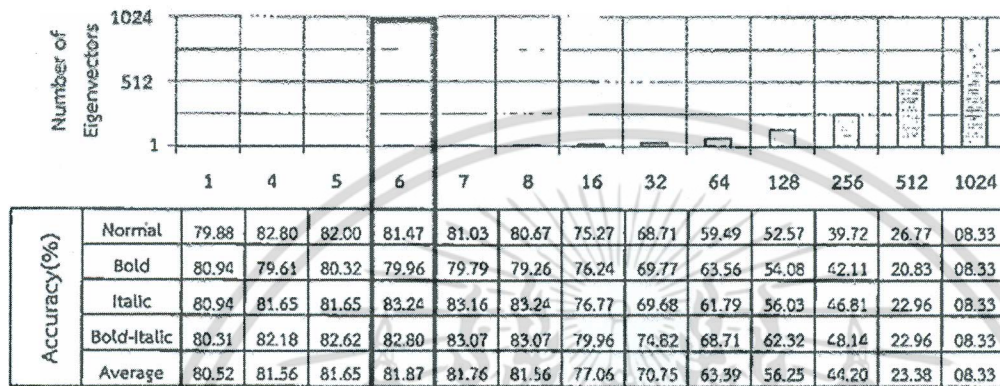
The test sets used in all experiments are Thai character image corpus consisting of consonants, vowels, and tones. A resolution of such images is a 400 dpi. Font types are composed of AngsanaUPC, BrowalliaUPC, CordiaUPC, DilleniaUPC, EucrosiaUPC, FreesiaUPC, IrisUPC, JasmineUPC, and unknown fonts, and font sizes are composed of 8, 10, 12, 14, 16, 18, 20, and 22. Font styles are regular, bold, italic, and bold-italic. There are totally 58,656 samples which are equally divided into training set for constructing the standard descriptor and test set.

3.2 Suitable Feature Selection and Recognition Accuracy

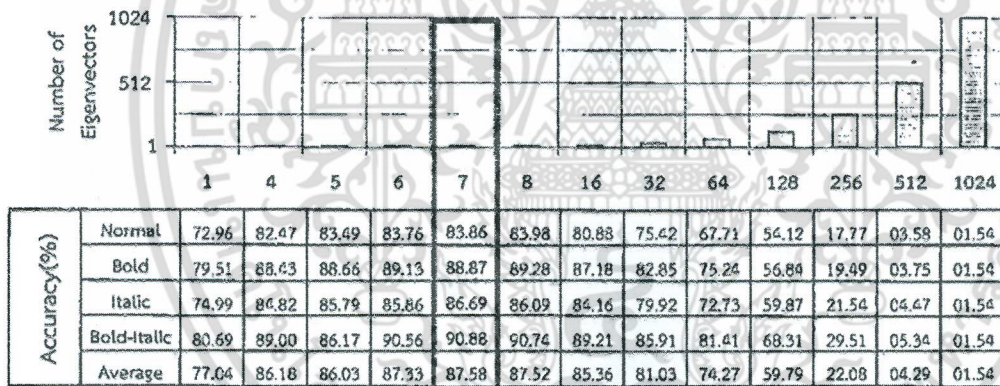
In this subsection, the appropriate feature selection is tested with two experiments. Then the results are analyzed. Hence, two standard descriptors, namely SD1 and SD2, are constructed by selecting the lower-order eigenvector and the higher-order eigenvector of SVD, respectively. Both standard descriptors are tested with Thai character images divided into three levels, upper, middle, and lower levels. In the first experiment, SD1 provides the lower recognition rate as illustrated in Fig. 6. It yields the good recognition accuracy, on average 87.58%, for all font styles, when seven components in lower-order eigenvectors are selected as depicted in Fig. 6(b). In addition, this experiment also reveals that the increase of a selected number of eigenvectors does not help improve the accuracy. In the same way, Fig. 6(a) and (c) exhibits the low recognition accuracy of upper and lower levels, 81.87% and 93.35%, respectively. On the other hand, SD2 is constructed by selecting different ranges of higher-order eigenvectors as shown in Fig. 7. The purpose of this experiment is to obtain a suitable eigenvector order to provide the high recognition accuracy with the minimum number of eigenvectors. The experimental results prove that the suitable range and order of eigenvectors are from 512 to 704, 18% of total eigenvectors, and point out by a block as shown in Fig. 7. The recognition accuracy of SD2 is up to 98.74% for middle level characters. The recognition accuracy in comparison of SD1 and SD2 is summarized in Table 1. It is obvious that SD2 outperforms SD1 in terms of recognition accuracy; however, the high accuracy of recognition rate trades off the increase of a number of components.

Furthermore, no-head fonts of printed Thai characters cause a serious problem in decreasing the recognition rate [6]. However, based on the experiment, SD2 can also provide the higher recognition accuracy, 94.05%, when compared to the method in [6], 83.77%.

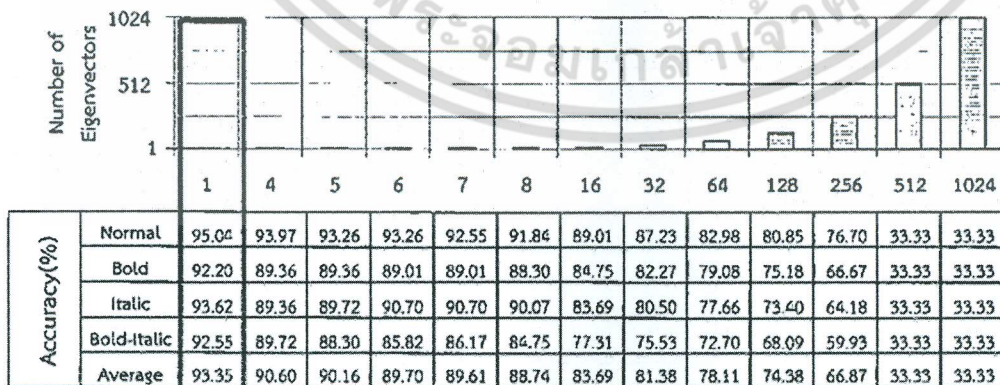
As reviewed in Introduction section, the methods proposed in [1-2] need two steps to extract features, whereas SD2 uses only single-step to extract features. Fewer steps lead to the reduction of computing time. Moreover, SD2 outperforms the method in [4] in terms of misclassification, when the shape of characters is similar.



(a)



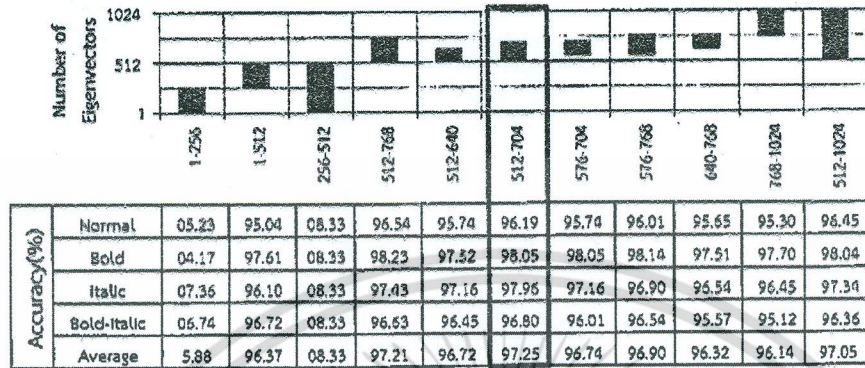
(b)



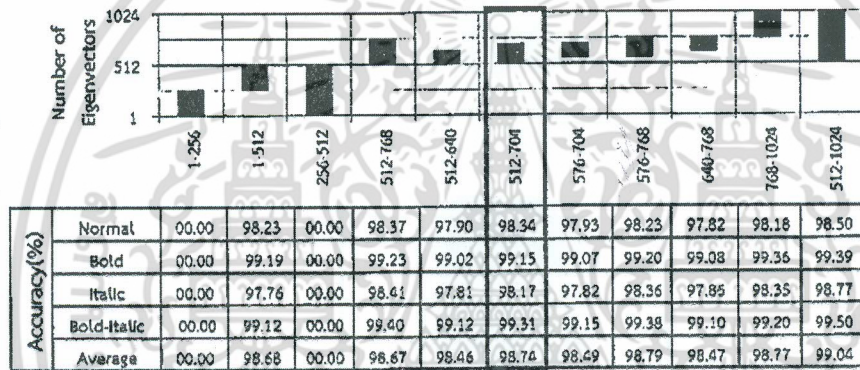
(c)

Fig. 6. Accuracy of the standard descriptor using the lower-order eigenvector of SVD in (a) upper level, (b) middle level, and (c) lower level

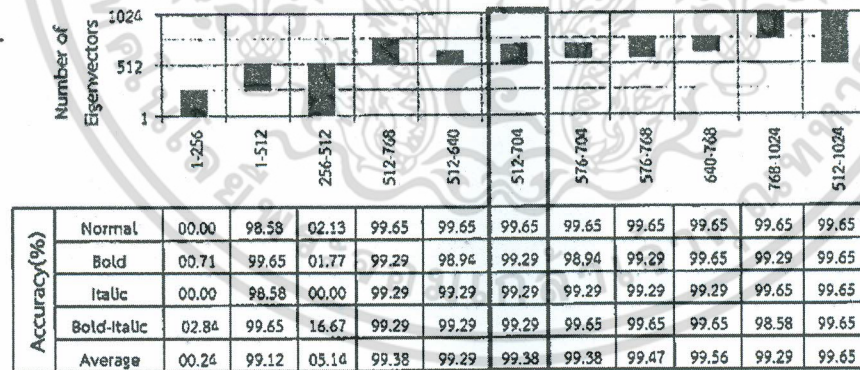
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



(a)



(b)



(c)

Fig. 7. Accuracy of the standard descriptor using the higher-order eigenvector of SVD in (a) upper level, (b) middle level, and (c) lower level

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Table 1. A comparison of accuracy of SD1 and SD2

Level	Accuracy (%)							
	SD1				SD2			
	Normal	Bold	Italic	Bold-Italic	Normal	Bold	Italic	Bold-Italic
Upper	81.03	79.79	83.16	83.07	96.19	98.05	97.96	96.80
Middle	83.98	89.28	86.09	90.74	98.34	99.15	98.17	99.31
Lower	92.55	89.01	90.70	86.17	99.65	99.29	99.29	99.29
Average	85.85	86.03	86.65	86.66	98.06	98.83	98.47	98.47

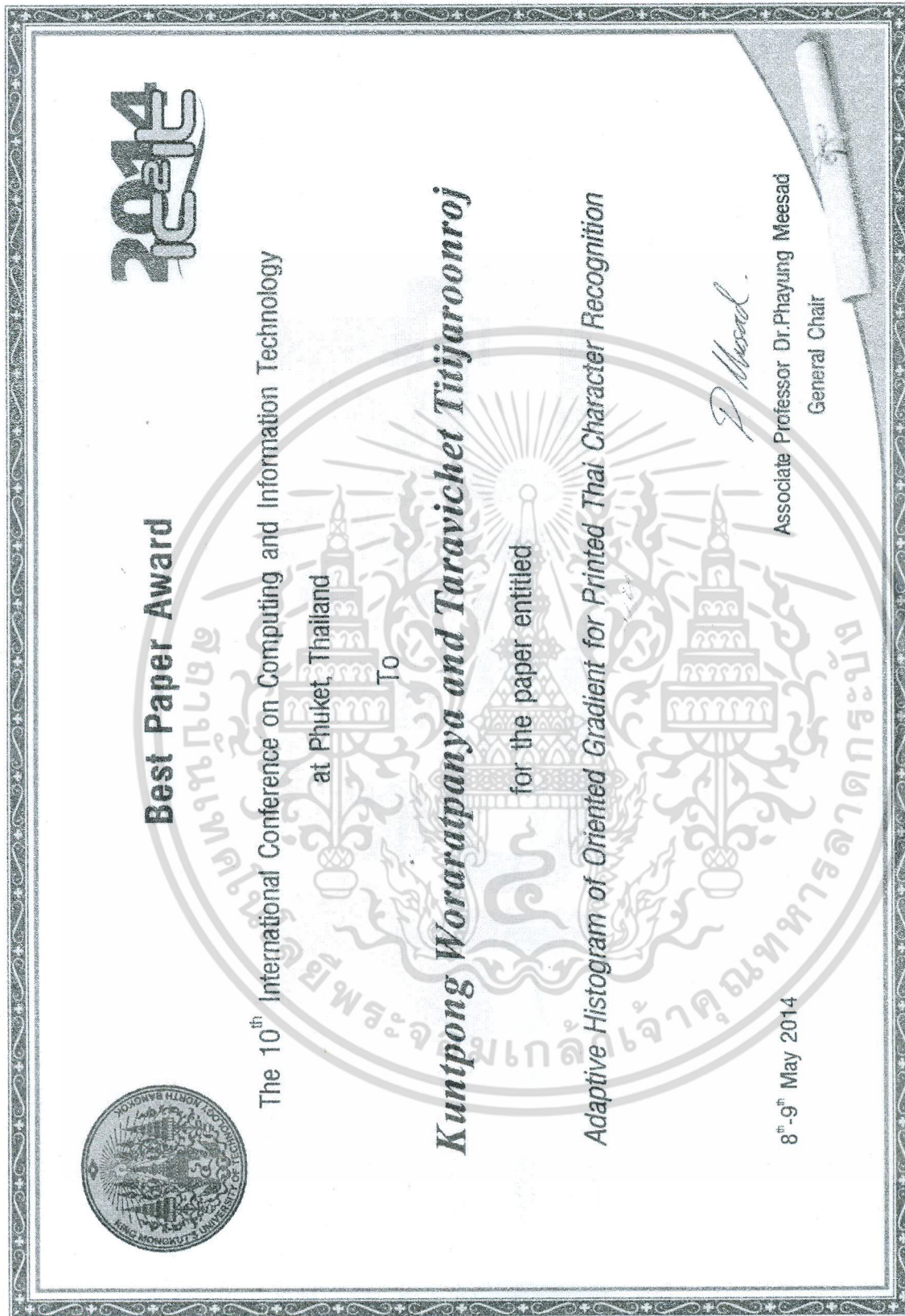
4 Conclusions

In this paper, an efficient standard descriptor for character recognition is proposed. The main contribution of this paper is constructing the standard descriptor based on the higher order eigenvectors and recognizing the character with the projection matrix technique. The experiment results show that the proposed method, SD2, evidently outperforms the traditional method, SD1, in terms of a recognition rate. It is concluded that the proposed standard descriptor helps improve the performance of OCR systems without reengineering software.

References

1. Kimpan, C., Itoh, A., Kawanishi, K.: Fine Classification of Printed Thai Character Recognition Using the Karhunen-Loeve Expansion. In: IEEE Proceedings, vol. 134, pp. 257-264 (1987)
2. Duangphasuk, S.: Thai Printed Character Recognition Using Feature Matching Method and ART1. M.S. Thesis, Dept. of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok (2002) (in Thai)
3. Tangsurakit, N.: Thai Printed Consonant Recognition Based on Character Density and Strip Features. M.S. Thesis, Dept. of Telecommunication Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok (2005) (in Thai)
4. Kruatrachue, B., Pantrakarn, N., Siriboon, K.: Automatic State Machine Induction for String Recognition. In: Proceedings of World Congress on Engineering, pp. 153-158 (2007)
5. Tanprasert, C., Sae-Tang, S.: Thai type style recognition. In: Proceedings of the 1999 IEEE International Conference on Circuits and Systems, pp. 336-339 (1999)
6. Thammano, A., Duangphasuk, P.: Printed Thai Character Recognition Using the Hierarchical Cross-correlation ARTMAP. In: Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence, pp. 695-698 (2005)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Best Paper Award

The 10th International Conference on Computing and Information Technology

at Phuket, Thailand

To

Kuntpong Woratpanya and Taravichet Titijaronroj

for the paper entitled

Adaptive Histogram of Oriented Gradient for Printed Thai Character Recognition

P Meesad

Associate Professor Dr. Phayung Meesad
General Chair

8th-9th May 2014


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Advances in Intelligent Systems and Computing 265

Sirapat Boonkrong
Herwig Unger
Phayung Meesad *Editors*

Recent Advances in Information and Communication Technology

Proceedings of the 10th International Conference
on Computing and Information Technology (IC²IT2014)

 Springer

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Adaptive Histogram of Oriented Gradient for Printed Thai Character Recognition

Kuntpong Woraratpanya and Taravichet Titijaronroj

Faculty of Information Technology,
King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
{Kuntpong, bank2533}@gmail.com

Abstract. A similarity of printed Thai characters is a grand challenge of optical character recognition (OCR), especially in case of a variety of font types, sizes, and styles. This paper proposes an effective feature extraction, adaptive histogram of oriented gradient (AHOG), for overcoming the character similarity. The proposed method improves the conventional histogram of oriented gradient (HOG) in two principal phases, which are (i) adaptive partition for gradient images and (ii) adaptive binning for oriented histograms. The former is implemented with quadtree partition based on gradient image variance so as to provide for an effective local feature extraction. The later is implemented with non-uniform mapping technique, so that the AHOG descriptor can be constructed with minimal errors. Based on 59,408 single character images equally divided into training and testing samples, the experimental results show that the AHOG method outperforms the conventional HOG and state-of-the-art methods, including scale space histogram of oriented gradient (SSHOG), pyramid histogram of oriented gradient (PHOG), multilevel histogram of oriented gradient (MHOG), and HOG column encoding algorithm (HOG-Column).

Keywords: Printed Thai Character Recognition, Pattern Recognition, Histogram of Oriented Gradient (HOG), Adaptive Histogram of Oriented Gradient (AHOG), Feature Extraction.

1 Introduction

An evolution of printed Thai fonts has continually developed and created for a variety of print medias, such as magazines, books, brochures, newspapers, and so on. These are appealing to readers, but are useless to visually impaired persons. In addition, a variety of font types, sizes, and styles is a key factor to degrade recognition rate of optical character recognition (OCR), which is a process of converting document images to editable text and Braille books. Therefore, the improvement of OCR based on a variety of print medias in Thai language is essential for visually impaired persons.

Although the printed Thai character recognition has been continually researched over the past two decades [1], it still requires the performance improvement for applying to real applications [2–4]. One of the main problems is the diversity of new

Thai character fonts, i.e., the more new fonts are created, the more recognition errors are increased. Furthermore, a similarity of printed Thai characters becomes a grand challenge of the Thai OCR. In order to achieve the higher performance of character recognition, the effective feature extraction method is required. Based on observation regarding a Thai character structure, its shape is a significant feature. For this reason, this paper investigates on the effective shape feature extraction. One of the successful methods for shape recognition is the histogram of oriented gradient (HOG) introduced by Navneet D. et al. [5]. Since then, many research papers have proposed modified HOGs, for example, scale space histogram of oriented gradient [6], pyramid histogram of oriented gradient [7], and multilevel histogram of oriented gradient [8]. Most of these have been successful in high efficiency for object recognition, but few papers [9] studied on HOG for character recognition.

Even though the conventional HOG and modified HOG methods have been successful for object recognition, these fail to printed Thai character recognition. The main obstacle is the similarity of Thai characters. Therefore, this paper proposes an AHOG method to overcome this problem. This method improves the conventional HOG algorithm in two principal phases, (i) adaptive partition of gradient images and (ii) adaptive binning of oriented histograms. The former phase is implemented with quadtree decomposition based on gradient image variance so as to provide for effectively extracting local shape features. The later phase is implemented with non-uniform quantization, so that the AHOG descriptor can be constructed with minimal errors. In this way, the recognition rate of Thai characters is significantly improved.

The rest of this paper is organized as follows: Section 2 reviews backgrounds of the conventional HOG and modified HOG algorithms, and points out their drawbacks. In Section 3, an AHOG method is proposed. Section 4 shows experimental results and discussions. Finally, conclusions are presented in Section 5.

2 Related Works

In this section, the important issues of conventional HOG and state-of-the-art methods, including SSHOG, PHOG, MHO, and HOG-Column, are reviewed and pointed out their drawbacks when applied to Thai character recognition.

A similarity of printed Thai characters becomes a critical issue in recognition as shown in Fig. 1. In this case, one way to classify similar characters effectively is using local shape features. This paper investigates on an improvement of feature extractions based on shape recognition. One of the effective methods for shape recognition is HOG, which was first proposed by Navneet D. et al. [5] and was successfully applied to human detection. The HOG algorithm is composed of five procedures, (i) gamma and color normalization, (ii) gradient computation, (iii) spatial and oriented binning, (iv) normalization and descriptor blocks, and (v) detector window and context. Then the features extracted from the HOG algorithm are classified by a linear support vector machine (SVM) to identify persons or non-persons. Although the HOG method achieves in human detection with high performance, it is not suitable for Thai characters with high similarity. The main reason is that the HOG approach was designed for coarse scale of object recognition, thus making it difficult to classify similar shape objects.



Fig. 1. An example of three groups of printed Thai characters with high similarity

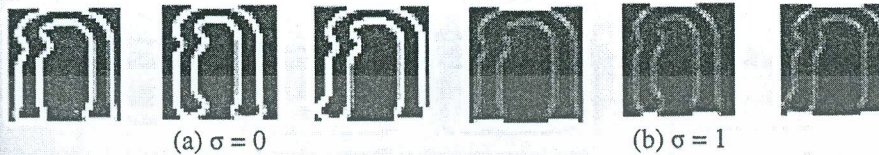


Fig. 2. An effect of applying two scale spaces, $\sigma = 0$ and $\sigma = 1$, to gradient images

SSHOG was introduced by Ning H. et al. [6] to solve complex objects and human detections which can be better to perceive at different scales. The main idea of this method is applying Gaussian kernel with multiple scales as demonstrated in Fig. 2. The conventional HOG differs from SSHOG in that it has only uni-scale ($\sigma = 0$). The experimental results show that the SSHOG method outperforms the conventional HOG method. Nonetheless, the higher scale of the SSHOG method filters out fine shape information. This leads to the misclassification when the SSHOG method applies to Thai characters. Furthermore, the more increased scales construct the larger feature size, thus consuming the computational time.

Representing shape with a spatial pyramid kernel was introduced by Anna B. et al. [7]. Its main idea is to extend a resolution of the conventional HOG method to multi-resolutions. In this method, the resolution implies a number of sub-region images divided from the original image. The descriptor made from the multi-resolutions is called PHOG. Its efficiency is better than the traditional HOG method. However, each resolution of the PHOG method is allocated with a fixed size. This technique works well with coarse shape objects, but it leads to the less recognition rate when applied to Thai characters with high similarity.

MHOG was proposed by Subhransu M. et al. [8]. Due to this method derived from the PHOG algorithm, it cannot extract the fine shape information to generate effective features. Furthermore, feature vectors are increasingly larger size.

Andrew J.N. et al. [9] introduced HOG column encoding algorithm (HOG-Column) and showed that the descriptor constructed from this algorithm is robust to character recognition. This approach is an extension of the conventional HOG descriptor by including features extracted from multiple scales, a base scale, σ_{base} , and a coarser scale, $r\sigma_{\text{base}}$, where r is a scale ratio. It is successful to recognize imperfect characters acquired from natural scenes. However, in case of a higher scale, there is no difference when compared to the SSHOG method.

Based on an analysis of Thai character patterns and conventional HOG algorithms, it can be summarized into two issues. (i) The Thai characters require fine shape features for classification, whereas (ii) the conventional HOG method needs to be improved for fine shape recognition. In order to meet the requirements, the critical issue of the HOG algorithm is regarded as three factors, (i) a fixed partitioning, (ii) a fixed binning, and (iii) a scalar magnitude sum. Fortunately, such an issue can be solved by the proposed method which is described in details in the next section.

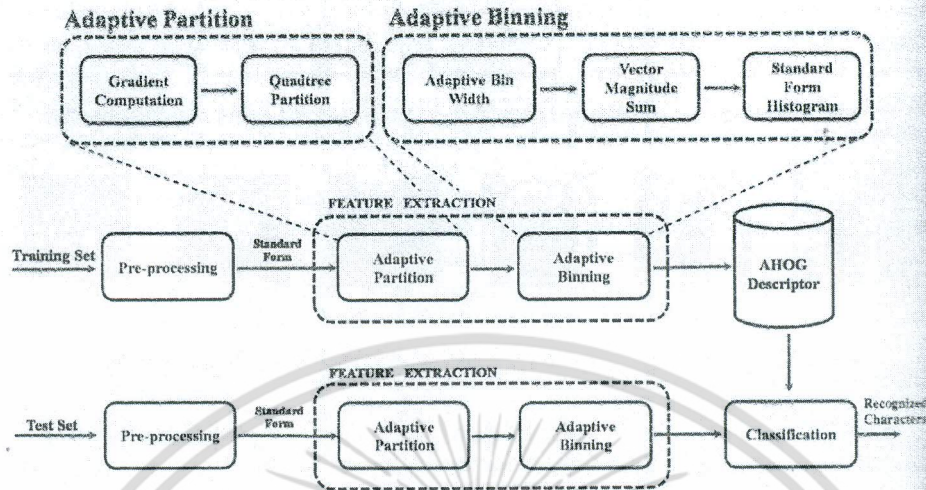


Fig. 3. A framework of AHOG for printed Thai character recognition

3 Proposed Method

In this paper, AHOG, which is an effective feature extraction, is proposed for overwhelming Thai character recognition with high similarity. A framework of the proposed method is schematically depicted in Fig. 3 and the manipulation details of each phase are described as follows.

3.1 Adaptive Partition

In general, printed Thai character recognition can be viewed as object recognition. Features extracted from printed Thai characters are represented in shape, which has a certain pose with a single view point. The best way to achieve in extracting the effective descriptor is using adaptive partition. In this procedure, there are two important phases, gradient computation and quadtree partition, to extract effective features.

Gradient Computation. Gradient images can be calculated by several methods. The mostly used technique is the first order derivative of Gaussian filters, such that $S_x = [-1 \ 0 \ 1]^T$ and $S_y = [-1 \ 0 \ 1]^T$, where superscript T is a matrix transpose operator. After that, two features, magnitude, M , and orientation, θ , of the gradient images are calculated by using equations (1) and (2) for each pixel, respectively, and the results are illustrated in Fig. 4.

$$M = \sqrt{I_x^2 + I_y^2} \quad (1)$$

$$\theta = \tan^{-1} \left(\frac{I_y}{I_x} \right); \quad -\pi \leq \theta \leq \pi \quad (2)$$

where I_x and I_y are gradient images.

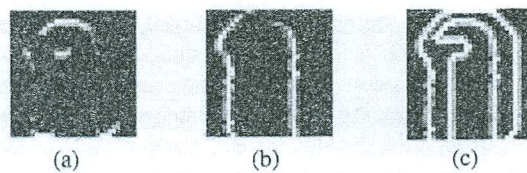


Fig. 4. (a) gradient image convolved by S_x , (b) gradient image convolved by S_y , and (c) a combination of gradient images from (a) and (b)

Quadtree Partition. In this phase, magnitudes of an image gradient are divided into sub-images or cells whose sizes are a square. The conventional HOG algorithm uses a fixed partition technique, i.e., all cells are equal sizes. Nevertheless, this technique has a disadvantage. As illustrated in Fig. 5, two similar shape character images are partitioned by means of an adaptive decomposition based on variance of the gradient images and are partitioned by means of a fixed partition. It is noticed that in the middle-left side of Fig. 5(a) and 5(b), the fixed partition does not provide an effective local shape feature to distinguish between two similar characters. In order to improve this, the AHOG algorithm makes use of adaptive partition based on gradient image variance as depicted in Fig. 5(c) and 5(d). It is evident that the adaptive partition is able to provide a better local shape feature.

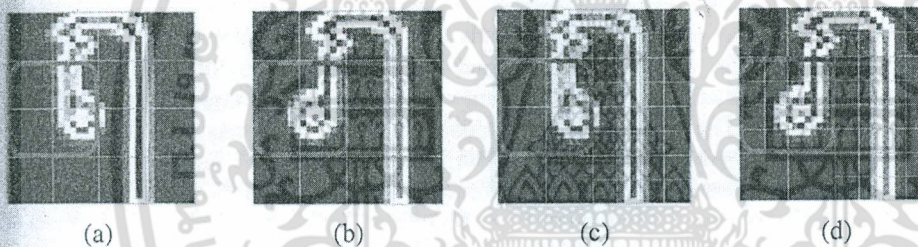


Fig. 5. A comparison of fixed and adaptive partitions

This paper implements the cell size with two levels, 4×4 pixels and 8×8 pixels, since the template of printed Thai characters is designed with size 32×32 pixels. If cell size is smaller than 4×4 pixels, some bins have zero data. This leads to those bins having no vectors in orientation. On the other hand, it is impossible to use cell size greater than 8×8 pixels, because the cell's data is coarser for classification.

3.2 Adaptive Binning

An adaptive binning is an important procedure to minimize error of feature extraction. This procedure is composed of three phases as explained in the following subsections.

Adaptive Bin Width. A bin width is an important factor having an impact on the quality of an AHOG descriptor. In this paper, the bin width assignment makes use of a non-uniform quantization technique. Such a technique is a many-to-few mapping, thus the bin

width becomes a key factor to maintain the significant feature as well as possible. The main idea of an adaptive method is assigning the suitable bin width for the high density of information. Fig. 6 shows a comparison of fixed and adaptive bin-ning of oriented histograms. In case of fixed binning, the bin width is equally defined for the oriented histogram without regarding their density. Hence, it is difficult to maintain the significant features. On the other hand, in case of adaptive binning, the bin width is adaptively defined for those features with regarding the density of the oriented histogram. The higher density of features, the finer bin width is defined to preserve significant features. Here, the suitable bin width can be obtained by using this criterion 2.5σ , where σ is a standard deviation of oriented histograms of a training set.

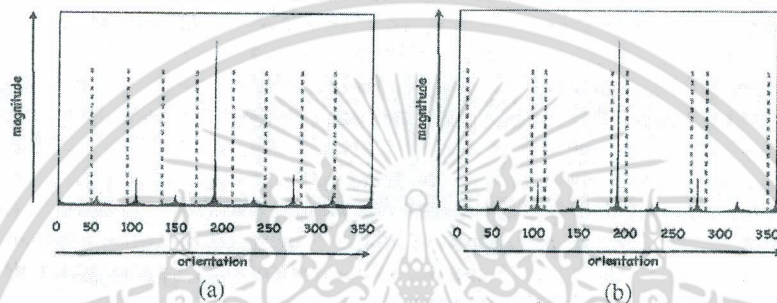


Fig. 6. A comparison of bin width assignments: (a) fixed binning of oriented histogram and (b) adaptive binning of oriented histogram

Vector Magnitude Sum. In the second phase, magnitudes of each partitioned histograms of orientated gradient, which are vector forms, are summed so as to generate a feature vector. The traditional HOG algorithm uses scalar magnitude sum, which cannot provide a more accurate feature vector. Hence, in this paper, the AHOG algorithm makes use of a vector magnitude sum to improve the efficiency of the feature vector. Fig. 7(a) and 7(b) show two similar characters with a fine shape feature located at the head of such characters (left-top sub-images). Fig. 7(c) and 7(d) graphically illustrate results of the fine shape feature by using vector and scalar magnitude sums, respectively. The results from both techniques are different. The vector magnitude sum presents a more accurate result; therefore, it is more appropriate for printed Thai characters. The outcome of this phase is a two-level AHOG descriptor as depicted in Fig. 8(a).

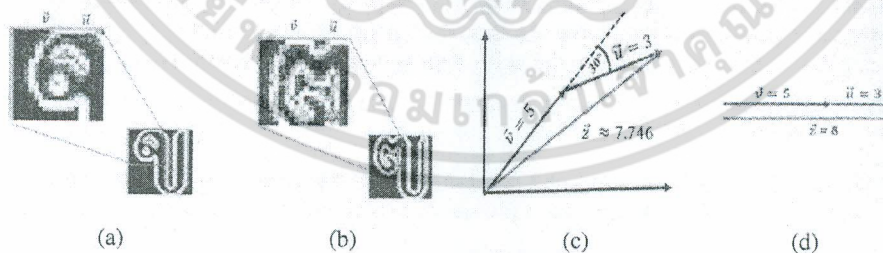


Fig. 7. (a) a sub-image of 'ช', (b) a sub-image of 'ช', (c) a vector addition, and (d) a scalar addition

4.2 Pre-processing

Pre-processing is the first procedure to transform all images in a dataset to a standard form. The procedure has three parts consisting of image complementation, image zero padding, and image resizing. These are applied to each character image as shown in Fig. 9. The standard form of images can be described as follows: the image size is 32×32 pixels, character color is white, and background color is black.

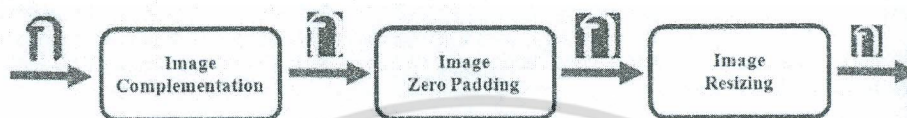


Fig. 9. Pre-processing procedure: image complementation, image zero padding, and image resizing

4.3 Performance Evaluation

This paper focuses on the feature extraction for constructing the effective descriptor. Thus, the descriptor performance is evaluated by means of the simple measurement, Euclidean distance, for classification.

Two experiments are set up in order to evaluate the performance of the AHOG algorithm. The first experiment aims to test the accuracy of the proposed method in terms of recognition rate. There are four state-of-the-art methods—including SSHOG, PHOG, MHOG, and HOG-column—and one conventional HOG, which are implemented as baseline algorithms. For all algorithms, the number of bins of oriented histograms is set to 9 as recommended in [5]. The experimental results are shown in Table 2. It is found that the recognition rate of the AHOG algorithm outperforms all of state-of-the-art algorithms.

In upper level, there are totally 1,128 character images for test. 1.00% of upper level characters are approximately 11 images. In this case, the maximum and minimum improvements of recognition rate in regular style are 7.10% and 0.09%, when compared to the HOG-Column and SSHOG methods. That is, the number of characters correctly recognized in the maximum improvement case is increasing about 80 character images. On the other hand, at least one character is correctly recognized in the minimum improvement case. In bold style, the maximum and minimum improvements of the recognition rate are 5.67% and 1.41%, when compared to the HOG-Column and SSHOG methods, respectively. It means that there are 64 characters greatly improved for recognition when compared to the HOG-Column method whereas there are at least 16 characters improved for recognition when compared to SSHOG. In the same way, in italic style, the maximum and minimum recognition rates of improvements are 6.56% and 0.10% when compared to the HOG-Column and SSHOG methods, respectively. In bold-italic style, the maximum and minimum recognition rates of improvements are 7.62% and 1.06% when compared to the HOG-Column and MHOG methods, respectively.

Middle level has 6,110 character images for test. 1.00% of middle level is approximate 61 character images. The great improvements of recognition rate in

regular, bold, italic, and bold-italic styles are 5.61%, 4.91%, 5.06%, and 4.13%, respectively, when compared to the HOG-Column method. Simultaneously, when compared to the SSHOG method, the AHOG method outperforms with the better recognition rate, 0.50%, 0.64%, 0.59%, and 0.28% for regular, bold, italic, and bold-italic styles, respectively.

In lower level, there are 188 characters for test. The experimental results illustrate that all methods yield the same results, 100%.

Table 2. A comparison of recognition accuracy of HOG, SSHOG, PHOG, MHOG, HOG-Column, and AHOG methods

Method	Upper Level				Middle Level				Lower Level			
	Regular	Bold	Italic	Bold-Italic	Regular	Bold	Italic	Bold-Italic	Regular	Bold	Italic	Bold-Italic
HOG	96.54	94.77	96.90	95.48	97.50	98.67	98.28	98.89	100.00	100.00	100.00	100.00
SSHOG	97.07	95.66	97.42	95.92	98.58	98.77	98.79	99.25	100.00	100.00	100.00	100.00
PHOG	94.24	94.59	94.33	93.44	95.01	95.99	95.52	96.45	100.00	100.00	100.00	100.00
MHOG	96.19	95.66	96.63	96.01	97.07	97.23	96.73	96.87	100.00	100.00	100.00	100.00
HOG Column	90.06	91.40	90.96	89.45	93.47	94.50	94.32	95.40	100.00	100.00	100.00	100.00
AHOG	97.16	97.07	97.52	97.07	99.08	99.41	99.38	99.53	100.00	100.00	100.00	100.00

Table 3. A comparison of feature size and computing time of HOG, SSHOG, PHOG, MHOG, HOG-Column, and AHOG methods

Method	HOG	SSHOG	PHOG	MHOG	HOG Column	AHOG
Feature Size	1,764	7,056	765	16,992	720	576
Computing Time	0.008 s	0.031 s	0.012 s	0.111 s	0.018 s	0.034 s

The second experiment aims to test the proposed method in terms of computing time and feature size. The experimental results are shown in Table 3. The proposed method has the minimal feature, 576 feature vectors. However, the computing time of proposed method is in an average, 0.034 sec, when compared with all methods. Based on the experimental results, it is summarized that the proposed method, AHOG, achieves the higher recognition rate in comparison with the state-of-the-art methods. In addition, the AHOG descriptor is constructed with the smallest size of feature vectors.

5 Conclusions

In this paper, an AHOG method has been proposed to improve the efficiency of extracting local shape features and to increase recognition rate of printed Thai characters with high similarity. In order to achieve these purposes, the proposed

method enhances the conventional HOG algorithm in two principal phases, i.e., (i) using an adaptive partition for gradient images to increase the efficiency of a local shape feature extraction, and (ii) applying an adaptive binning for oriented histograms to reduce the error of the local shape feature extraction. Based on these improvements, the AHOG descriptor is minimal, 576 features, when compared with baseline methods, whereas the computing time of feature extraction is in an average, 0.034 sec. Furthermore, the AHOG algorithm is evaluated in terms of recognition accuracy. Based on 59,408 single character images equally divided into training and testing samples, the experimental results show that the AHOG algorithm outperforms the conventional HOG and state-of-the-art algorithms, including SSHOG, PHOG, MHOG, and HOG-column.

References

1. Kimpan, C., Itoh, A., Kawanishi, K.: Fine Classification of Printed Thai Character Recognition Using the Karhunen-Loeve Expansion. *IEE Proceedings* 134, 257–264 (1987)
2. Tanprasert, C., Sae-Tang, S.: Thai type style recognition. In: 1999 IEEE International Conference on Circuits and Systems, pp. 336–339. IEEE Press, New York (1999)
3. Thammano, A., Duangphasuk, P.: Printed Thai Character Recognition Using the Hierarchical Cross-correlation ARTMAP. In: the 17th IEEE International Conference on Tools with Artificial Intelligence, pp. 695–698. IEEE Press, New York (2005)
4. Woraratpanya, K., Titijaronroj, T.: Printed Thai Character Recognition Using Standard Descriptor. In: Meesad, P., Unger, H., Boonkrong, S. (eds.) *IC²IT2013. AISC*, vol. 209, pp. 165–173. Springer, Heidelberg (2013)
5. Navneet, D., Bill, T.: Histograms of Oriented Gradients for Human Detection. In: IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE Press, New York (2005)
6. Ning, H., Jiaheng, C., Lin, S.: Scale Space Histogram of Oriented Gradients for Human Detection. In: 2008 International Symposium on Information Science and Engineering, pp. 167–170. IEEE Press, New York (2008)
7. Anna, B., Andrew, Z., Xavier, M.: Representing Shape with a Spatial Pyramid Kernel. In: The 6th ACM International Conference on Image and Video Retrieval (2007)
8. Subhransu, M., Alexander, C.B., Jitendra, M.: Classification Using Intersection Kernel Support Vector Machines is Efficient. In: IEEE International Conference on Computer Vision and Pattern Recognition, IEEE Press, New York (2008)
9. Andrew, J.N., Lewis, D.G.: Multiscale Histogram of Oriented Gradient Descriptors for Robust Character Recognition. In: 2011 International Conference on Document Analysis and Recognition, pp. 1085–1089. IEEE Press, New York (2011)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลประวัติผู้วิจัย

ประวัติส่วนตัว

ชื่อ-สกุล นายกันต์พงษ์ วรรณปัญญา

ตำแหน่งปัจจุบัน ผู้ช่วยศาสตราจารย์ สังกัดคณะเทคโนโลยีสารสนเทศ สจล

ประวัติการศึกษา

ชื่อปริญญา	สาขา	สถาบันที่จบ	ปีที่จบ
B. Ind. Tech	Computer Technology	สพระจอมเกล้าลาดกระบัง.	1992
M. Eng	Electrical Engineering	สพระจอมเกล้าลาดกระบัง.	1997
D. Eng	Electrical Engineering	สพระจอมเกล้าลาดกระบัง.	2005

สาขาวิจัยที่มีความชำนาญพิเศษ

- Pattern Recognition and Image Processing
- Multimedia Coding
- Data Compression
- Stereoscopic Acquisition and Compression
- Signal Processing

รางวัลด้านวิชาการ/ด้านวิจัย/งานสร้างสรรค์ ที่ได้รับ (ด้านศิลปะ หรืออื่นๆ)

ปี พ.ศ.	ชื่อรางวัล	สถาบันที่ให้
-	-	-

ทุนการศึกษาและทุนวิจัยที่เคยได้รับ

ปี พ.ศ.	ทุนการศึกษาและทุนวิจัย	สถาบันที่ให้
-	-	-

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลงานวิจัย/งานสร้างสรรค์

ผลงานวิจัย/งานสร้างสรรค์ที่ตีพิมพ์เผยแพร่ (ระดับชาติและนานาชาติ)

ผลงานวิจัยที่ตีพิมพ์เผยแพร่ระดับนานาชาติภายในระยะเวลา 5 ปีที่ผ่านมา

- [1] Thanyarat Nomponkrang and **Kuntpong Woraratpanya**, "Thai-Sentence Classification Using Conceptual Graph," *Proceedings of ICEIT 2010 (2010 International Conference on Educational and Information Technology)*, pp. 479-483, Chongqing, China, Sept 17-19, 2010.
- [2] Surachet Julakananusat, **Kuntpong Woraratpanya**, Supot Nitsuwat, and Ruttikorn Varakulsiripunth, "3D Object Model Retrieval Using Enhanced Elevation Descriptor," *Proceedings of ISTS 2011 (1st International Symposium on Technology for Sustainability)*, pp. 225-228, Bangkok, Thailand, Jan. 26-29, 2012.
- [3] **Kuntpong Woraratpanya**, Donyarut Kakanopas, and Ruttikorn Varakulsiripunth, "Triangle-box Counting Method for Fractal Dimension Estimation," *ASEAN Engineering Journal Part D*, pp. 5-16, Vol.1 No.1 Sept. 2012.
- [4] **Kuntpong Woraratpanya** and Taravichet Titijaronroj, "Printed Thai Character Recognition Using Standard Descriptor," *the 9th International Conference on Computing and Information Technology (IC²IT2013)*, Bangkok, Thailand, 9-10 May 2013 (Published in *Advances in Intelligent Systems and Computing*, Springer Link, Vol. 209, 2013, pp 165-173).
- [5] Ungsumalee Suttapakti, **Kuntpong Woraratpanya**, and Kitsuchart Pasupa, "Font Descriptor Construction for Printed Thai Character Recognition," *Proceedings of IAPR2013 (the 13th IAPR International Conference on Machine Vision Applications)*, pp.45-48, Kyoto, Japan, 20-23 May 2013.
- [6] **Kuntpong Woraratpanya**, Pimlak Boonchukusol, Yoshimitsu Kuroki, and Yasushi Kato "Improved Thai Text Detection from Natural Scenes," *Proceedings of ICITEE2013 (the 5th International Conference on Information Technology and Electrical Engineering)*, pp. 137-142, Yokjakata, Indonesia, 7-8 Oct. 2013.
- [7] Teruaki Fujiyoshi, Taihei Oki, **Kuntpong Woraratpanya**, and Yoshimitsu Kuroki, "On Normalization of Subspace Method in DCT Domain," *Proceedings of ISTS2013 (the 3rd International Symposium on Technology for Sustainability)*, Hong Kong, 20-22 Nov. 2013.
- [8] Ryo Okutani, Makoto Tokunaga, **Kuntpong Woraratpanya**, and Yoshimitsu Kuroki, "Orthogonal Cone-Restricted Subspace Method for Face Recognition," *Proceedings of IWAIT2014 (the 2014 International Workshop on Advanced Image Technology)*, Bangkok, Thailand, 6-8 Jan. 2014.
- [9] **Kuntpong Woraratpanya** and Taravichet Titijaronroj, "Adaptive Histogram of Oriented Gradient for Printed Thai Character Recognition," **Best Paper Award of the 10th International Conference on Computing and Information Technology (IC²IT2014)**, pp. 83-92, Phuket, Thailand, 8-9 May 2014 (Published in *Advances in Intelligent Systems and Computing*, Springer Link, Vol. 265, 2014, pp 83-92).
- [10] **Kuntpong Woraratpanya**, Ungsumalee Suttapakti, Pimlak Boonchukusol, Taravichet Titijaronroj, Ratthapon Hogging, Kitsuchart Pasupa, Yoshimitsu Kuroki, and Yasushi Kato "Text-Background Decomposition for Thai Text Localization and Recognition in Natural Scenes," Accepted for presentation at *the 6th International Conference on Information Technology and Electrical Engineering*, Yokjakata, Indonesia, 7-8 Oct. 2014.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลงานวิจัยที่ตีพิมพ์เผยแพร่ระดับชาติภายในระยะเวลา 5 ปีที่ผ่านมา

- [1] Ungsumalee Suttapakti and **Kuntpong Woraratpanya**, "Document Image Classification Using Texture Features," *Proceedings of the 6th National Conference on Computing and Information Technology*, pp. 647-652, Bangkok, Thailand, 2010.
- [2] Kannika Keawchuy and **Kuntpong Woraratpanya**, "Online computer education journal system," *Proceedings of the 6th National Conference on Computing and Information Technology*, pp. 868-872, Bangkok, Thailand, 2010.
- [3] Sirawit Pokaew, **Kuntpong Woraratpanya**, and Supot Nitsuwat, "Fractal Image Coding Using Adaptive Domain Pool Size," *Proceedings of the 7th National Conference on Computing and Information Technology*, pp.367-372, Bangkok, Thailand, 2011.
- [4] **Kuntpong Woraratpanya** and Pimlak Boonchukusol, "Thai Text Detection from Natural Scenes by Using Boundary Clustering and Connected Component Analysis," *Proceedings of the 9th National Conference on Computing and Information Technology*, pp.665-672, Bangkok, Thailand, 2013.

การเสนอผลงานวิชาการ

.....

.....

.....

.....

ผลงานสิทธิบัตร/สิ่งประดิษฐ์/งานสร้างสรรค์ (ศิลปะ หรือ อื่นๆ)

.....

.....

.....

อื่นๆ

.....

.....

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หนังสือเป็นสมบัติของท่าน
โปรดช่วยกันรักษา

www.lib.kmitl.ac.th

สำนักหอสมุดกลาง โทร. 02-329-8544-5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้