

## บทที่ 2

### แนวคิดและงานวิจัยที่เกี่ยวข้อง

#### 2.1 แนวคิดเกี่ยวกับข้อมูลภาพและการประมวลผล

โดยทั่วไปแล้วภาพที่นำมาใช้ในการประมวลผลภาพดิจิทัลนั้นแบ่งออกเป็นสองประเภทหลัก ๆ (دنุพล จันทรทรวง, 2552) คือ

1) ภาพเคลื่อนไหว ก็คือภาพนิ่งที่แสดงต่อเนื่องกันโดยจะต้องใช้รูปภาพอย่างน้อย 24 รูปต่อหนึ่งวินาที เนื่องจากสายตาของคนเราเมื่อนำภาพนิ่งมาฉายติดต่อกันมากกว่า 24 รูปต่อหนึ่งวินาที แล้วก็จะมองว่าภาพนั้นเป็นภาพเคลื่อนไหว โดยแยกไม่ออกเนื่องจากมีความเร็วมากเกินไป แต่หากนำภาพนิ่งมาฉายติดต่อกันน้อยกว่า 24 รูปต่อหนึ่งวินาทีแล้ว จะมองเห็นว่าภาพนั้นไม่ต่อเนื่อง

2) ภาพนิ่ง ภาพที่นำมาใช้ในการประมวลผลในคอมพิวเตอร์นั้น ถ้าในระบบ RGB จะใช้ความเข้มแสงสีแดง เขียว และน้ำเงิน ความหมายของภาพนิ่งก็คือมีอยู่เพียงภาพเดียว ภาพนิ่งที่นำมาใช้ก็มีอยู่หลาย Format เช่น .bmp หรือ .jpg เป็นต้น ส่วนใหญ่จะเป็น .bmp เพราะไม่ต้องมาถอดรหัสก่อน เนื่องจากภาพที่เป็น .jpg นั้นมีการบีบอัดภาพให้มีขนาดเล็ก ดังนั้นหากจะนำมาใช้ก็ต้องคลายข้อมูลออกมาก่อนที่จะนำภาพไปประมวลผลต่อไป

เนื่องจากข้อมูลที่ใช้เป็นไฟล์วิดีโอที่ได้จากระบบ APS ซึ่งเป็นภาพเคลื่อนไหวจึงต้องทำการแปลงภาพเคลื่อนไหวให้เป็นภาพนิ่งแล้วจึงสามารถนำภาพไปประมวลผลเบื้องต้นได้ภาพที่ได้คือ Digital Image เป็นภาพที่ประกอบขึ้นจาก Dots หรือ Pixel เรียกว่า Bit-map คือภาพที่เก็บอยู่ในรูปแบบของดิจิทัล ภาพที่มองเห็นด้วยสายตาทั่วไปนั้นเป็นภาพในลักษณะสามมิติ ส่วนภาพถ่ายหรือภาพที่แสดงบนเครื่องคอมพิวเตอร์เป็นการแปลงภาพจากสามมิติ มาเป็นภาพสองมิติ โดยการแปลงสัญญาณไฟฟ้าในรูปแบบอนาล็อก (John C. R., 1994) อย่างเช่น ในกล้องวิดีโอจะมีเซนเซอร์ภายในกล้องจะทำการสแกนหรือวัดผลรวมความเข้มแสงที่จุดต่าง ๆ ไปตามแนวสแกนที่เรียกว่า Raster Scan โดยลักษณะการสแกนจะมีทิศทางโดยเริ่มจากบนลงล่างและจากซ้ายไปขวา ภาพที่ได้จากการสแกนจะเป็นภาพแบบต่อเนื่อง (Continuous) ด้วยความเร็วทั่วไปที่ 24 ภาพต่อวินาที เช่นเดียวกันนี้ในเครื่องรับภาพวิดีโอก็จะรับภาพที่ได้มาจากเครื่องถ่ายวิดีโอ และแสดงผลโดยเริ่มจากบนลงล่างและจากซ้ายไปขวาเช่นเดียวกัน แต่ภาพที่ได้มาจากระบบอนาล็อกนั้นยังเป็นภาพแบบต่อเนื่อง ที่ยังไม่สามารถนำมาใช้ในการประมวลผลได้ ต้องมาทำการแปลงให้เป็นภาพเชิงตัวเลขเสียก่อนโดยแปลงจากภาพเคลื่อนไหวให้เป็นภาพนิ่ง

## 2.2 แนวคิดเกี่ยวกับ Optical Character Recognition (OCR)

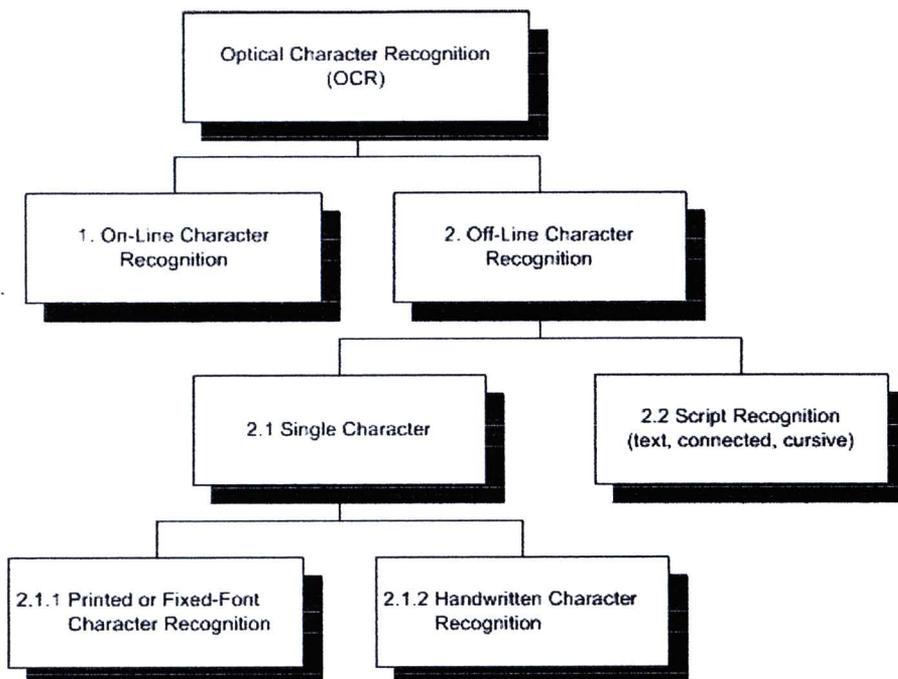
ไอซีอาร์เป็นการรู้จำรูปแบบตัวอักษร ซึ่งเป็นงานวิจัยในสาขาการรู้จำรูปแบบ (Pattern Recognition) เป็นเทคโนโลยีที่ช่วยให้ระบบคอมพิวเตอร์สามารถระบุรูปแบบได้อย่างถูกต้อง การจับคู่แผ่นแบบ (Template Matching) พื้นฐานของเทคนิคคือ ภาพลักษณะตัวอักษรจะถูกกราดตรวจ (Scan) ด้วยตัวตรวจจับลำแสง (Photo detector) ผ่านแผ่นแบบ (Template) ซึ่งมีลำแสงยิงผ่าน และเมื่อผ่านแผ่นแบบที่จับคู่ได้ดีกับภาพลักษณะตัวอักษร ตัวตรวจจับลำแสงก็จะรู้จำภาพลักษณะตัวอักษรได้ (Jesse H., 2005)

### 2.2.1 ประเภทของไอซีอาร์

ไอซีอาร์สามารถแบ่งได้เป็น 2 ประเภท (ธวัชชัย หยุแก้ว, 2551) ได้ดังนี้

1) การรู้จำตัวอักษรแบบออนไลน์ (On-line Character Recognition) วิธีการนำเข้าสู่ข้อมูลของกลุ่มนี้ได้มาจากดิจิทัลไทเซอร์ หรือปากกาอิเล็กทรอนิกส์ ที่ใช้กับเครื่องคอมพิวเตอร์มือถือ ส่วนการวิเคราะห์ตัวอักษร จะทำในขณะที่มีการลากเส้น เพื่อเขียนตัวอักษร (ออนไลน์) ถ้าจะเทียบความยากง่าย กับการรู้จำลายมือเขียนแบบออฟไลน์ กลุ่มนี้จะง่ายกว่า เพราะจะได้ข้อมูลเพิ่มเติมเกี่ยวกับทิศทาง และลำดับการลากเส้นมาช่วยด้วย ไอซีอาร์กลุ่มนี้มักมาพร้อมกับอุปกรณ์การเขียน ที่มีการกำหนดพื้นที่ให้นำเข้าสู่ข้อมูล โดยส่วนใหญ่มักต้องเขียนทีละตัวอักษร โดยมีรหัสพิเศษเพื่อใช้ในการเขียนตัวอักษรแต่ละตัว ความก้าวหน้าของเทคโนโลยีไอซีอาร์ มีผลเป็นอย่างมากต่อการเจริญเติบโตของธุรกิจในวงการเครื่องคอมพิวเตอร์มือถือ ที่เป็นคอมพิวเตอร์ขนาดเล็ก ซึ่งต้องอาศัยการใส่หน้าเข้าจากปากกาอิเล็กทรอนิกส์แทนคีย์บอร์ด

2) การรู้จำตัวอักษรแบบออฟไลน์ (Off-line Character Recognition) ซึ่งสามารถแยกย่อยเป็นตัวอักษรโดด (Single Character) เช่น การรู้จำตัวพิมพ์แบบฟอนต์เฉพาะ (Printed Fixed-Font Character Recognition) การรู้จำลายมือเขียนแบบตัวโดด (Isolated Handprint Character Recognition (ICR) เป็นต้น และการรู้จำลายมือแบบเขียนต่อเนื่อง (Script Recognition) โดยโครงสร้างประเภทของไอซีอาร์แสดงดังแผนผังรูปที่ 2.1



รูปที่ 2.1 แผนผังประเภทของโอซีอาร์ (ธวัชชัย หยูแก้ว, 2551)

ข้อมูลระบบ APS ที่จะใช้ในการทดลองจัดอยู่ในกลุ่มประเภทการรู้จำตัวพิมพ์แบบฟอนต์เฉพาะ (Printed Fixed-Font Character Recognition) เป็นกลุ่มของโปรแกรมที่ใช้ได้กับตัวอักษรประเภทตัวพิมพ์ที่มีการกำหนดกลุ่มของ ฟอนต์ ซึ่งถือว่าอยู่ในระดับง่ายที่สุดในด้านการพัฒนาเทคนิคในการทำงานของโปรแกรม (ABDELWAHAB Z., 2005)

#### 2.2.2 โครงสร้างระบบโอซีอาร์

โครงสร้างของระบบโอซีอาร์โดยทั่วไปประกอบด้วยการทำงานหลัก 3 ขั้นตอนในแต่ละขั้นตอนประกอบด้วยรายละเอียดดังรูปที่ 2.2



รูปที่ 2.2 โครงสร้างระบบโอซีอาร์ (ธวัชชัย หยูแก้ว, 2551)

### 1) ขบวนการประมวลผลขั้นต้น (Pre-Processing)

ในการทำงานของโปรแกรมโอซีอาร์นั้น ก่อนที่โปรแกรมจะสามารถบอกได้ว่ารูปภาพที่ส่งเข้าไปประกอบด้วยตัวอักษรอะไรบ้าง จำเป็นจะต้องผ่านขั้นตอนที่สำคัญหลายขั้นตอนดังกล่าวนี้นี้มักถูกเรียกรวมกันว่า กระบวนการประมวลผลขั้นต้น ซึ่งเป็นขั้นตอนในการปรับแต่งและจัดเตรียมข้อมูลให้เหมาะสมกับขั้นตอนการรู้จำต่อไป ขั้นตอนเหล่านี้มีความสำคัญต่อประสิทธิภาพโดยรวมของระบบ เพราะหากมีความผิดพลาดเกิดขึ้นในส่วนนี้ ก็จะส่งผลกระทบต่อประสิทธิภาพโดยรวมของระบบด้วย ขั้นตอนการประมวลผลเบื้องต้นในโปรแกรมโอซีอาร์ที่สำคัญ ได้แก่

- การกรองข้อมูลแทรกซ้อน (Noise Filtering) มีจุดประสงค์เพื่อลดทอนส่วนของรูปภาพที่เป็นสิ่งแปลกปลอมอันไม่พึงประสงค์ออกไป โดยข้อมูลแทรกซ้อนที่เกิดขึ้นส่วนใหญ่มักจะมาจากคุณภาพของเอกสารต้นฉบับที่นำมาทำการอ่าน ซึ่งเป็นต้นเหตุสำคัญที่ทำให้ความถูกต้องของโปรแกรมลดลง จึงจำเป็นที่จะต้องจัดการกับส่วนเกินเหล่านี้ออกไปให้ได้มากที่สุดเท่าที่จะเป็นไปได้ แต่ยังไม่มียุติการใดที่รับรองได้ว่าสามารถจัดการกับข้อมูลแทรกซ้อนได้ โดยสมบูรณ์ ดังนั้นส่วนการรู้จำของโอซีอาร์ก็จะต้องมีความทนทานต่อการแทรกซ้อนเหล่านี้ได้พอสมควร

- การปรับแต่งข้อมูล (Normalization) เป็นการปรับภาพตัวอักษรให้อยู่ในรูปแบบที่ระบบต้องการเพื่อนำไปใช้ในขั้นต่อไป ตัวอย่างการปรับแต่งข้อมูลในโปรแกรมโอซีอาร์ต่างๆ ไป อาทิเช่น การปรับขนาดรูปตัวอักษร, การปรับตัวอักษรที่เอียงให้ตรง, เมื่อได้ภาพ RGB มาแล้วต้องทำการแปลงภาพให้เป็น Gray Scale เพื่อลดปริมาณข้อมูลภายในภาพเพื่อให้การประมวลผลง่ายขึ้นโดยระดับเทา (Gray Level) เป็นค่าซึ่งระบุความสว่างหรือความเข้ม ซึ่งมีค่าตั้งแต่ 0-255 (0 คือระดับมืด 255 คือระดับสว่าง) รวมทั้งพิกัดแนวนอนและแนวตั้ง ซึ่งใช้ระบุตำแหน่งในแถวลำดับภาพ (Image Array) เป็นต้น

- การตัดแบ่งพื้นที่ใช้งาน (Cropping) เป็นการตัดแยกเอาเฉพาะรูปตัวอักษรออกมาจากภาพ เพื่อส่งให้ขั้นตอนการรู้จำในการระบุว่ารูปตัวอักษรนั้นเป็นรหัสอักษรอะไร หลักการพอสั่งเขปที่ใช้สำหรับการตัดรูปตัวอักษรโดยทั่วไปจะใช้พื้นที่สีขาว (สีพื้น) รอบรูปเป็นตัวกำหนดขอบเขตในการตัด ในขั้นตอนนี้มักจะประสบปัญหาที่ส่งผลกระทบต่ออัตราความถูกต้องของระบบโดยรวมอยู่สองปัญหา ปัญหาแรกคือปัญหาตัวติด เกิดจากรูปของตัวอักษรตั้งแต่สองตัวขึ้นไปมีส่วนที่เชื่อมติดกัน ทำให้ไม่สามารถแยกตัวอักษรออกจากกันโดยใช้พื้นที่สีขาวรอบๆ ได้ จำเป็นต้องหาอัลกอริทึมพิเศษมาช่วยในการแยกตัวอักษรออกจากกัน ส่วนปัญหาที่สองในทางตรงกันข้าม เป็นปัญหาตัวขาดที่รูปตัวอักษรหนึ่งๆ ถูกแยกออกเป็นส่วนๆ ทำให้เวลาตัดตัวอักษรจากตัวเดียวจะได้เป็นสองตัว ซึ่งก็ต้องหาวิธีการเฉพาะสำหรับมาจัดการอีกเช่นกัน

- การสกัดลักษณะสำคัญ (Feature Extraction) เป็นอีกขบวนการหนึ่งที่สำคัญมาก บางตำราใหญ่จะแยกส่วนนี้ออกจากการประมวลผลเบื้องต้น คือจะอยู่ระหว่างขั้นตอนการประมวลผลเบื้องต้นกับขั้นตอนการรู้จำ การสกัดลักษณะสำคัญเป็นการดึงเอาโครงสร้างพื้นฐานที่สำคัญของตัวอักษรนั้นออกมา โดยโครงสร้างพื้นฐานที่ว่าจะต้องมีการกำหนดไว้ก่อนว่าจะมีอะไรบ้าง มีการนิยามอย่างไร ตัวอย่างเช่น สำหรับภาษาไทยอาจกำหนดว่าตัวอักษรภาษาไทยทั้งหมดประกอบด้วยโครงสร้างพื้นฐานคือ เส้นตรง (แนวตั้ง/นอน) เส้นเอียง หัว (วงกลม) ส่วนโค้ง ส่วนเว้า จุดแตกกิ่ง จุดตัด เป็นต้น เมื่อสามารถแยกเอาองค์ประกอบของตัวอักษรแต่ละตัวออกมาได้แล้ว จากนั้นจึงนำเสนอรูปภาพของตัวอักษรนั้นในรูปแบบของรายการขององค์ประกอบพื้นฐานต่างๆ แทน ซึ่งจะถูกส่งต่อสำหรับขั้นตอนการรู้จำต่อไป

## 2) การรู้จำ (Recognition)

ขั้นตอนนี้อถือเป็นหัวใจของระบบ เพราะเป็นส่วนที่จะตัดสินว่ารูปตัวอักษรที่ส่งเข้าไปเป็นรหัสตัวอักษรอะไร เช่นเดียวกับส่วนอื่นๆ ที่มีวิธีการหลากหลายซึ่งนำมาใช้เพื่อให้ได้ผลการทำงานที่ดีที่สุด เทคนิคใหม่ๆ ได้ถูกพัฒนาขึ้นมาอย่างต่อเนื่องเพื่อรองรับกับปัญหาที่เกิดจากเอกสารที่มีความสลับซับซ้อนมากขึ้น เทคนิคเหล่านี้้ออกสามารถแบ่งเป็นกลุ่มตามแนวทางหลักที่ใช้ในการแก้ปัญหาตามทฤษฎี เป็น 4 กลุ่ม คือ วิธีการเข้าคู่รูปแบบ (Template Matching) วิธีการสถิติ (Statistical Approach) วิธีการวิเคราะห์ทางโครงสร้าง (Structural Analysis) วิธีการโครงข่ายประสาทเทียม (Neural Network) (นราพงษ์ อภิรัตน์วรากล และ นเรศ เกื้อปัญญากุล, 2552)

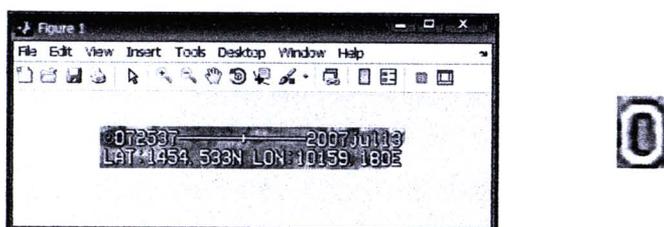
## 3) ขบวนการประมวลผลขั้นปลาย (Post-Processing)

หลังจากที่ผ่านขั้นตอนการรู้จำแล้ว รูปตัวอักษรที่ถูกส่งเข้าไปจะได้ผลลัพธ์ออกมาเป็นรหัสตัวอักษร ซึ่งก็ไม่ได้หมายความว่าผลลัพธ์ที่ได้มาจะถูกต้องทั้งหมดปัจจุบัน ไม่มีผลิตภัณฑ์ไอซีอาร์ตัวใด ไม่ว่าจะเป็ภาษาใดก็ตามที่รับรองความถูกต้อง 100 % ดังนั้นเพื่อเพิ่มความถูกต้องให้แก่โปรแกรมจึงได้มีการเสริมส่วนการตรวจสอบและแก้ไขข้อความเข้ามา โปรแกรมส่วนนี้มักจะทำงานเกี่ยวกับการตรวจสอบความถูกต้องของการสะกดคำและไวยากรณ์ภาษา โดยมักจะใช้พจนานุกรมมาช่วยในการตรวจสอบคำผิด ซึ่งอาจแก้ไขให้โดยอัตโนมัติหรือแสดงเครื่องหมายบางอย่างเพื่อบอกให้ผู้ใช้ทราบว่าคำดังกล่าวอาจไม่ถูกต้อง ซึ่งผู้ใช้อาจแก้หรือไม่แก้ก็ขึ้นกับการตัดสินใจของผู้ใช้เอง นอกเหนือไปจากการตรวจสอบความถูกต้องระดับคำแล้ว โปรแกรมส่วนนี้สามารถเพิ่มอัตราความถูกต้องให้แก่โปรแกรมได้ โดยเฉพาะในส่วนที่นอกเหนือจากความสามารถของส่วนการรู้จำ

### 2.3 การจับคู่แผ่นแบบ (Template matching)

การจับคู่แผ่นแบบเป็นวิธีการแรกๆ ที่มาใช้ในการรู้จำตัวอักษร หลักการโดยทั่วไปคือจะต้องมีรูปแผ่นแบบ (Template) ที่สร้างขึ้นมาสำหรับอ่านตัวอักษร โดยมีการกำหนดตำแหน่งสำคัญที่สามารถใช้แยกแยะความแตกต่างระหว่างตัวอักษรแต่ละตัว โดยการให้นำรูปภาพที่ต้องการอ่านไปทาบบนแบบเพื่อวัดความคล้ายคลึงกันของภาพกับตัวแบบ จากนั้นก็ระบุว่าเป็นรหัสตัวอักษรอะไร โดยใช้ค่าผ่านระดับความเข้มแสงหรือวิธีการบางอย่างในการตัดสินใจ ข้อเสียของวิธีการนี้คือจะค่อนข้างอ่อนไหวต่อข้อมูลแทรกซ้อน ขนาด และการเอียงของตัวอักษร จึงจำเป็นต้องมีขั้นตอนการปรับแต่งข้อมูลที่ดี

การจับคู่แผ่นแบบ เป็นวิธีการค้นหาตำแหน่งของวัตถุที่สนใจในภาพ (Source) หรือเรียกว่า Region of Interest (ROI) เป็นการกำหนดพื้นที่บริเวณที่สนใจในที่นี่ ROI คือ บริเวณตัวเลขเวลา วัน เดือน ปี ค่า ละติจูดและลองจิจูดเพื่อลดเวลาของการประมวลผลลง โดยการเทียบรูปของแผ่นแบบ (Template) กับทุกๆ บริเวณในภาพที่ต้องการค้นหาตำแหน่งแสดงดังรูปที่ 2.3



รูปที่ 2.3 ภาพที่ต้องการค้นหา (Source) และภาพแผ่นแบบ (Template)

การจับคู่แผ่นแบบมีขั้นตอนดังต่อไปนี้

- แผ่นแบบจะถูกเลื่อนไปบนทุกพิกเซลของภาพนำเข้าที่ต้องการ ยกเว้นบริเวณขอบโดยรอบของภาพ ซึ่งบริเวณขอบที่ยกเว้นนี้จะขึ้นอยู่กับขนาดของแผ่นแบบ
- ในขณะที่แผ่นแบบวางตัวอยู่บนภาพนำเข้า โดยจะทำการคำนวณหาค่าความแตกต่างของค่า Intensity ของทุกๆ พิกเซลที่อยู่บนแผ่นแบบกับทุกๆ พิกเซลของภาพนำเข้าที่อยู่ตำแหน่งเดียวกัน
- เมื่อแผ่นแบบเคลื่อนที่ไปจนครบทั่วทั้งภาพนำเข้าแล้ว และจะทำการหาค่าต่ำที่สุดของค่าความแตกต่างที่คำนวณได้
- ตำแหน่งของค่าต่ำที่สุดที่หาได้ คือ ตำแหน่งหรือบริเวณในภาพนำเข้าที่มีวัตถุเหมือนกับภาพแผ่นแบบที่ใช้ในการค้นหานั้นเอง

โดยหลักการของการจับคู่แผ่นแบบมีหลายชนิดซึ่งแต่ละชนิดขึ้นอยู่กับความเหมาะสมกับข้อมูลที่ต่างกันซึ่งมีผลกับความถูกต้องของงานวิจัยโดยวิธีที่ง่ายและใช้โดยทั่วไปมี 3 วิธีคือ Sum

of Absolute Differences (SAD), Sum of Squared Differences (SSD) และ Maximum Absolute Difference (MaxAD) (Matlab and Simulink, 2010)

2.3.1 Sum of Absolute Differences (SAD) ผลรวมของความแตกต่างสัมบูรณ์ เป็นผลรวมค่าสัมบูรณ์ของความแตกต่างระหว่างพิกเซลในภาพต้นฉบับและพิกเซลที่สอดคล้องกันในภาพแม่แบบ เมทริกซ์นี้จะเป็นบรรทัดฐานของภาพที่แตกต่างกัน SAD โดยใช้ค่าที่น้อยที่สุดในการประมาณตำแหน่งที่ดีที่สุดของแม่แบบในการค้นหารูปภาพ ระยะผลรวมของความแตกต่างสัมบูรณ์จะกลายเป็นตัวชี้วัดทั่วไป

$$d_1(I_j, T) = \sum_{i=1}^n |I_{i,j} - T_{i,j}| \dots\dots\dots (2.1)$$

2.3.2 Sum of Squared Differences (SSD) ผลรวมของความแตกต่าง Squared เป็นผลรวมกำลังสองของความแตกต่างกันแน่นอนระหว่างพิกเซลในภาพต้นฉบับและพิกเซลที่สอดคล้องกันในภาพแม่แบบ เมทริกซ์นี้จะเป็นบรรทัดฐานที่สองของความแตกต่างของภาพ SSD ตัวชี้วัดโดยทั่วไปจะกลายเป็นระยะทาง:

$$d_2(I_j, T) = \sum_{i=1}^n |I_{i,j} - T_i|^2 \dots\dots\dots (2.2)$$

2.3.3 Maximum Absolute Difference (MaxAD) ความแตกต่างสูงสุดของผลลัพธ์ ตัวชี้วัดนี้เป็นที่รู้จักกันเป็นระยะทางเมทริกซ์ Uniform โดยใช้จำนวนค่ามากที่สุดของค่าสัมบูรณ์ของความต่างระหว่างพิกเซลในภาพต้นฉบับและพิกเซลที่สอดคล้องกันในภาพแม่แบบนี้ตัวชี้วัดระยะทางให้กฎเกณฑ์ของภาพแตกต่าง ระยะทางจะกลายเป็นตัวชี้วัดทั่วไป MaxAD :

$$d_\infty(I_j, T) = \lim_{x \rightarrow \infty} \sum_{i=1}^n |I_{i,j} - T_i|^x \dots\dots\dots (2.3a)$$

หรือ

$$d_\infty(I_j, T) = \max_i^n |I_{i,j} - T_i|^p \dots\dots\dots (2.3b)$$

งานวิจัยนี้ได้เลือกใช้วิธี SAD ในการจับคู่แผนแบบ (โสภณ ผู้มีจรรยา, 2553) (ตัวอย่างการจับคู่แผนแบบแสดงการคำนวณอยู่ในภาคผนวก ก.) เนื่องจากข้อมูลอยู่ในที่ใช้ทดสอบ OCR ประเภท Fixed-Front Character Recognition ซึ่งไม่มีการเปลี่ยนแปลงการบิดเบือนของลักษณะอักษรและมุกกล่องมาเกี่ยวข้องซึ่งเป็นวิธีที่ไม่ยุ่งยากและประมวลผลไม่ซับซ้อนซึ่งพอเพียงสำหรับงานวิจัยนี้

## 2.4 Window Mask

เป็นลักษณะของการเคลื่อนที่ของแผ่นแบบในการวิเคราะห์บริเวณจุดที่สนใจโดยการกำจัดสิ่งรบกวนของภาพโดยวิธีการ Window Mask ช่วยในการประมวลผล ซึ่งจะสามารถปรับเปลี่ยนขนาดได้ การปรับเปลี่ยนขนาดของ Window Mask จะทำให้การกำจัดสิ่งรบกวนมีประสิทธิภาพมากขึ้น แต่ก็มีผลทำให้สูญเสียความคมชัดของภาพโดยนำมาใช้ในการ Average Filter โดยที่ Average Filter มีวิธีการหาค่าเฉลี่ยที่ได้จาก Window Mask สามารถหาค่าได้จากสมการที่ 3 ซึ่งมีลักษณะดังนี้ (อวิชัย หยูแก้ว, 2551)

$f(i-1, j-1)$	$f(i, j-1)$	$f(i+1, j-1)$
$f(i-1, j)$	$f(i, j)$	$f(i+1, j)$
$f(i-1, j+1)$	$f(i, j+1)$	$f(i+1, j+1)$

ตัวอย่างลักษณะของ Window Mask ขนาด 3x3

$$g(i, j) = \frac{1}{n} \sum_{i=1, j=1}^{m, n} f(i, j) \dots\dots\dots (2.4)$$

$g(i, j)$  คือ ค่าเฉลี่ยของตำแหน่ง  $(i, j)$  ที่นำ Window Mask ไปครอบ

$f(i, j)$  คือ ตำแหน่งที่นำ Window Mask ไปครอบ

$M, n$  คือ ขนาดของ Window Mask ซึ่งมีขนาดเป็น Square Matrix

## 2.5 ฐานข้อมูล PostgreSQL

PostgreSQL คือ ระบบจัดการฐานข้อมูลเชิงวัตถุ-สัมพันธ์ (Object-Relational DataBase Management System หรือ ORDBMS) โดยสามารถใช้รูปแบบของภาษา SQL ได้เกือบทั้งหมด และสามารถใช้ Subselects, Transactions, User-defined types และ Functions ระบบจัดการฐานข้อมูล ที่คิดค้นขึ้นมาตั้งแต่ Version แรกๆ และพัฒนาปรับปรุงอย่างต่อเนื่องจนถึงปัจจุบัน

โดย PostgreSQL มีชนิดข้อมูลทางเรขาคณิตแทนสองมิติมิติวัตถุ แสดงดังตารางที่ 2.1 แสดงรูปแบบทางเรขาคณิตที่มีอยู่ใน PostgreSQL ชนิดพื้นฐานที่สุด จุด เป็นพื้นฐานสำหรับชนิดอื่นๆ ทั้งหมด (PostgreSQL, 1996-2011)



## ตารางที่ 2.1 ชนิดการเก็บข้อมูล Geometry

Name	Storage Size	Representation	Description
point	16 bytes	Point on a plane	(x,y)
line	32 bytes	Infinite line (not fully implemented)	((x1,y1),(x2,y2))
lseg	32 bytes	Finite line segment	((x1,y1),(x2,y2))
box	32 bytes	Rectangular box	((x1,y1),(x2,y2))
path	16+16n bytes	Closed path (similar to polygon)	((x1,y1),...)
path	16+16n bytes	Open path	[(x1,y1),...]
polygon	40+16n bytes	Polygon (similar to closed path)	((x1,y1),...)
circle	24 bytes	Circle	<(x,y),r> (center point and radius)

นอกจากการเก็บข้อมูลแบบเรขาคณิตฐานข้อมูล PostgreSQL ยังสนับสนุนการเก็บข้อมูลเวลา Date/Time วันและเวลาที่การป้อนข้อมูลจะยอมรับในเกือบทุกรูปแบบสมเหตุสมผล สำหรับรูปแบบบาง วัน เดือน และปีในวันที่ป้อนข้อมูลการสั่งที่จะคลุมเครือ และไม่มีการสนับสนุนสำหรับการระบุคำสั่งไว้ของเซตข้อมูลเหล่านี้ การตั้งค่าของพารามิเตอร์ DateStyle เป็น MDY เพื่อเลือกแปลงเดือนวันปี DMY เพื่อเลือกแปลงวันเดือนปี หรือ YMD เพื่อเลือกการตีความเนื้อหาของปีเดือนวัน PostgreSQL คือมีความยืดหยุ่นในการจัดการวัน / เวลาว่าต้องใช้ SQL มาตรฐานในการป้อนข้อมูล Date/Time สามารถตั้งค่ารูปแบบผลลัพธ์ชนิดวัน / เวลาหนึ่งตามลักษณะ ISO 8601, SQL (Ingres), POSTGRES ดั้งเดิม (รูปแบบวัน Unix), หรือภาษาเยอรมัน ค่าเริ่มต้นคือ รูปแบบ ISO (SQL มาตรฐานต้องใช้รูปแบบ ISO 8601 ชนิดวันและเวลาตามมาตรฐาน)

สำนักงานคณะกรรมการวิจัยแห่งชาติ  
ห้องสมุดงานวิจัย  
วันที่ 29 มี.ย. 2555  
เลขทะเบียน 247064  
เลขเรียกหนังสือ

## 2.6 Matlab

งานวิจัยนี้ได้เลือกใช้ Matlab เนื่องจากมีเครื่องมือที่สนับสนุนการทำงานด้วย Multimedia และ Image processing และมีเครื่องมือเพื่อช่วยในการเชื่อมต่อโปรแกรม Matlab กับโปรแกรมภาษาอื่น(มนัส สังวรศิลป์ และ วรรัตน์ ภัทรอมรกุล, 2543) พัฒนาเครื่องมือทางด้านการประมวลผลข้อมูลด้วยวีดีโอซึ่งแบ่งแยกเป็นสัดส่วนและสามารถเรียนรู้ได้ด้วยตนเอง มีการสนับสนุนเครื่องมือในการทำงานด้าน Multimedia อย่างเช่น การนำเข้าภาพถ่ายและวีดีโอ การดูข้อมูลเบื้องต้นวีดีโอ การวิเคราะห์และการเพิ่มคุณภาพ การแปลงข้อมูล การแปลงทางเลขาคณิต การดำเนินงานทางสัตวศาสตร์ ตัวอย่างการประยุกต์ใช้ วัตถุและระบบการใช้งาน ระบบการประมวลผลเพื่อระบุวัตถุในวีดีโอ การสนับสนุนวีดีโอและการประมวลผลด้วยสมองกล (Matlab Matlab and Simulink, 2010)

Matlab คือ โปรแกรมคำนวณเชิงตัวเลขที่มีสิ่งแวดล้อมในการคำนวณของตัวเอง (Numerical Computing Environment) และมีภาษาเฉพาะตัวในการเขียนโปรแกรมได้ โดย Matlab มาจากคำ 2 คำรวมกัน คือ Matrix และ Laboratory ซึ่งหมายถึง ห้องปฏิบัติการเมทริกซ์

การพัฒนาโปรแกรมด้วย Matlab มีความง่ายและรวดเร็ว เพราะมีไลบรารีจำนวนมาก รองรับการทำงานหลายรูปแบบ และด้วยลักษณะการทำงานเชิงเมทริกซ์ทำให้สามารถจัดการกับอาร์เรย์ได้อย่างสะดวกอีกทั้ง รองรับการทำงานกราฟิกรวมถึง GUI ทำให้สะดวกในการป้อนค่าและแสดงผล นอกจากนั้น ยังสามารถติดต่อกับฮาร์ดแวร์และโปรแกรมภาษาอื่นๆ ได้โดยสามารถแบ่งประโยชน์ของ Matlab ออกเป็น 3 ประเภท (ปริญญา สงวนสัตย์, 2553) ได้แก่

1) เป็นโปรแกรมคำนวณ ที่รองรับทั้ง เชิงตัวเลข (Numeric) เป็นเครื่องมือในการคำนวณธรรมดา หรือใช้งานฟังก์ชันทางคณิตศาสตร์ขั้นสูงได้และ เชิงสัญลักษณ์ (Symbolic) สามารถคำนวณในเชิงตัวแปรได้ เช่น การอินทิเกรต หรือ การแก้สมการต่างๆ แบบติดตัวแปร

2) สามารถเขียนโปรแกรมได้สามารถเขียนแบบ Script ซึ่งทำงานในลักษณะชุดคำสั่งต่อเนื่องหรือเขียนเป็น ฟังก์ชัน เพื่อใช้งานก็ได้ ใช้งานได้ทั้งแบบ Interpret หรือ Compile โดยสามารถ Compile โปรแกรม Matlab ออกมาได้หลายชนิดทั้งแบบ Standalone หรือ Library เช่น .exe หรือ .dll ทั้งยังมี GUI รองรับ โดยสามารถเขียนได้ทั้งแบบใช้ GUIDE (คล้าย Visual Basic) หรือแบบไม่ใช้ก็ได้ รองรับการทำงานโปรแกรมเชิงวัตถุทั้งคลาสของ Matlab เองหรือคลาสของภาษาอื่น Java หรือ .NET สามารถ Debug โปรแกรมได้ และในส่วนของติดต่อกับภาษาอื่นๆ สามารถ Compile ไปเพื่อทำการ Debug ในโปรแกรมอื่นได้ด้วย

3) สามารถติดต่อ หรือใช้งานร่วมกับโปรแกรม ภาษา ฮาร์ดแวร์ หรือแฟ้มข้อมูลรูปแบบต่างๆ โดยสามารถเชื่อมต่อกับภาษา หรือโปรแกรมอื่นๆ ได้ เช่น Java, C/C++, .NET โดยอาจให้

โปรแกรมหลักเขียนโดย Matlab แล้วเชื่อมต่อเพื่อเรียกใช้งานภาษาอื่น หรือโปรแกรมอื่นเป็นหลัก แล้วเรียกใช้งาน Matlab ก็ได้ สามารถอ่านหรือเขียนแฟ้มข้อมูลสื่อสารมาตรฐานได้ เช่น ข้อความ วีดีโอ รูปภาพ เสียง วีดีโอ เป็นต้น

## 2.7 งานวิจัยที่เกี่ยวข้อง

Burr D., (1981) การเข้าคู่แผนแบบเป็นวิธีที่ง่ายวิธีหนึ่งที่ใช้ในการรู้จำตัวอักษร ในการค้นหาและตัดสินใจระหว่างสิ่งของสองสิ่งที่มีลักษณะคล้ายคลึงกันด้วยวิธีการเข้าคู่แผนแบบ แผนแบบใช้เป็นต้นแบบหรือแบบจำลองในการรู้จำโดยการนำรูปร่างของแผนแบบลักษณะต่างๆที่จัดเก็บไปประกบกับข้อมูลนำเข้า โดยทั่วไปแผนแบบจะถูกจัดเก็บรวมๆกันเป็นชุดข้อมูลเพื่อนำไปใช้ประมวลผล ผลลัพธ์ที่ได้ขึ้นอยู่กับความสามารถในการประมวลผลของแต่ละลักษณะการใช้งาน

Pratt W., (1991) ลักษณะสำคัญของรูปภาพเองสามารถพิจารณาโดยใช้ ลักษณะสำคัญของเวกเตอร์ ของโครงสร้างพื้นฐานของความคล้ายคลึงระหว่างแผนแบบและรูปที่ภาพที่สนใจโดยอ้างอิงจากภาพ โดยที่ D ความต่างของระยะข้อมูลที่ไม่เหมือนกัน Z คือภาพที่สนใจ Tj คือแผนแบบโดยการคำนวณหาค่าเฉลี่ยของระยะผลต่างกำลังสองแสดงดังสมการที่

$$D_j = \sum_{i=1}^M |Z(X_i, Y_i) - T_j(X_i, Y_i)|^2 \quad \dots\dots\dots (2.6)$$

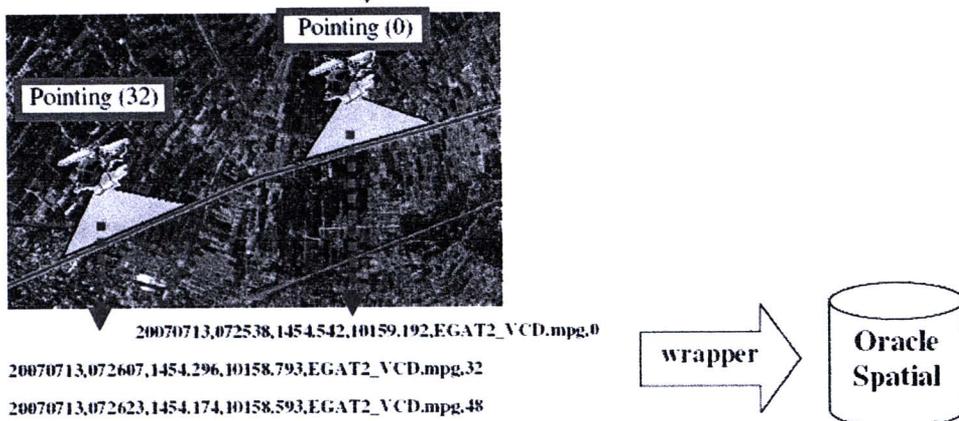
ภัทธาภรณ์ อริยปรีชากุล และ นางลักษณ โควาวิสารัช(2548) ได้นำเสนอการทดลองหาประสิทธิภาพของระบบในการระบุบุคคลหาได้จากค่าอัตราการยอมรับบุคคลผิดและอัตราการปฏิเสธบุคคลผิดโดยนำแผนแบบที่ได้ของภาพทดสอบมาหาค่าระยะทางกับแผนแบบที่อยู่ในฐานข้อมูลทั้งหมดแล้วกำหนดค่าขีดแบ่งขึ้น สำหรับการหาระยะทางแบบ Regional Correlation เมื่อทำการทดลองโดยการสกัดลักษณะด้วยวิธีการแปลงแบบเรดอนแล้วใช้ฟังก์ชันระยะทาง 4 แบบได้ผลดังตารางที่ 2.2 ของการระบุบุคคลของฟังก์ชันระยะทาง 4 ชนิด

ตารางที่ 2.2 ร้อยละความผิดพลาดของการระบุบุคคล

ร้อยละความผิดพลาดของการระบุบุคคล	
ฟังก์ชันระยะทาง	EER (%)
1. Regional correlation	23
2. Euclidean	18
3. Absolute	44
4. S1	56

Soravis S. and Sanphet C., (2009), การจัดการพื้นที่บริเวณเสาไฟฟ้าแรงสูง การไฟฟ้าฝ่ายผลิตแห่งประเทศไทยได้ทำการส่งผ่านพลังงานไฟฟ้าผ่านสายส่งไฟฟ้าบริเวณพื้นที่เสาไฟฟ้าแรงสูง โดยใช้ข้อมูลจากระบบ APS ในการสังเกตการณ์และตรวจสอบแนวสายส่งไฟฟ้าประกอบด้วยภาพถ่ายและวิดีโอ ซึ่งมีข้อมูลภาพถ่ายและวิดีโอจำนวนมากที่ได้ทำการบันทึกและจัดเก็บยากต่อการค้นหา การพัฒนาเว็บไซต์ฐานข้อมูล Multimedia GIS เพื่อให้ผู้ใช้สามารถค้นหาข้อมูลเพื่อให้สามารถวิเคราะห์ การจัดการและการรุกร้าบริเวณแนวเสาส่งไฟฟ้าแรงสูงได้ทันท่วงที โดยพัฒนาเวปฐานข้อมูลอยู่บนมาตรฐานของ WMS และ WFS แสดงดังรูปที่ 2.4

<year-month-date>,<time stamp>,<latitude coordinate>,<longitude coordinate>,<file name>,<time offset>



รูปที่ 2.4 ลักษณะการเก็บข้อมูลของเวปฐานข้อมูล Multimedia GIS