หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

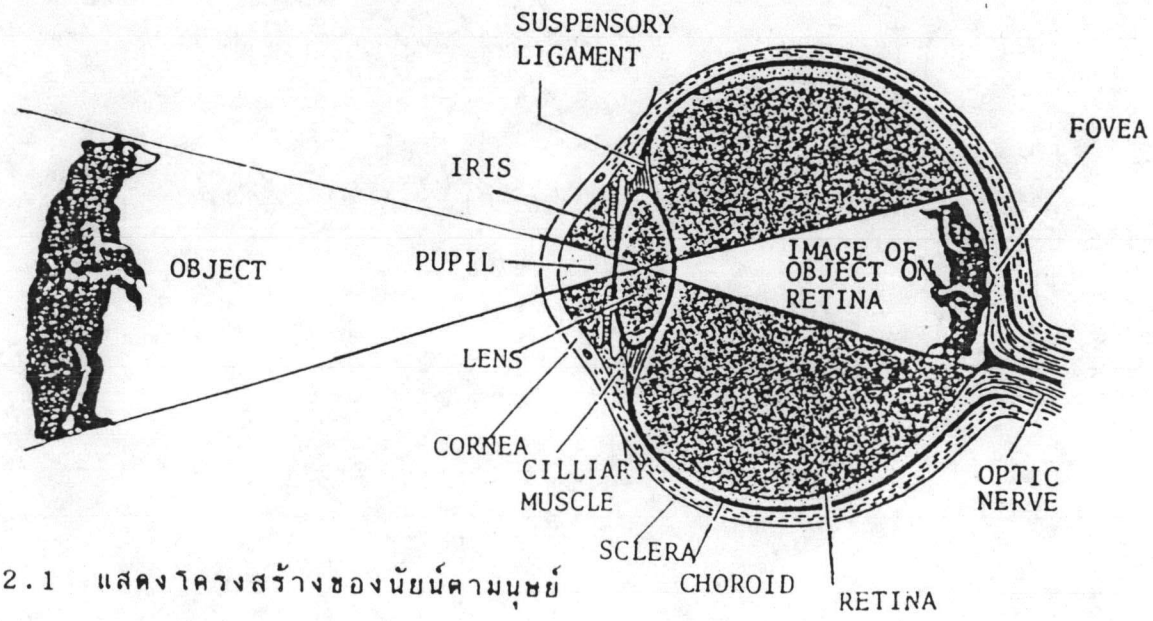
บทที่ 2
แนวความคิดเกี่ยวกับระบบการตรวจรู้อักขระ

2.1 การรับภาพของมนุษย์ (1)

การตรวจรู้อักขระด้วยเครื่องคอมพิวเตอร์เป็นแนวความคิดที่จะพัฒนาให้เครื่องคอมพิวเตอร์สามารถรับข้อมูลที่เป็นรูปภาพเลียนแบบความสามารถของมนุษย์ เพื่อความเข้าใจระบบการรับภาพของเครื่องคอมพิวเตอร์ จึงจำเป็นที่จะต้องเข้าใจระบบการรับภาพของมนุษย์ ซึ่งมีรายละเอียดดังนี้

อวัยวะที่ทำหน้าที่รับภาพของมนุษย์คือนัยน์ตา โครงสร้างของนัยน์ตามนุษย์จะมีรูปร่างกลมมีขนาดเส้นผ่าศูนย์กลางประมาณ 2.5 เซนติเมตร นัยน์ตาประกอบด้วยผนังเนื้อเยื่อเหนียวที่เรียกว่า สเคลอรา (Sclera) คอรอยด์ (Choroid) และ เรตินา (Retina) ห่อหุ้มเป็นชั้นๆ ตามลำดับ (รูปที่ 2.1) โดยผนังแต่ละชั้นมีรายละเอียดดังนี้

ชั้น สเคลอรา เป็นชั้นเนื้อเยื่อเหนียวแต่ไม่ยืดหยุ่นตอนหน้าสุดของเนื้อเยื่อส่วนนี้จะโปร่งใส และนูนออกมาเรียก กระจกตา (Cornea) นอกจากนี้ส่วนที่เป็นนัยน์ตาขาวก็เป็นส่วนหนึ่งของเนื้อเยื่อชั้นสเคลอราเหมือนกัน



รูปที่ 2.1 แสดงโครงสร้างของนัยน์ตามนุษย์

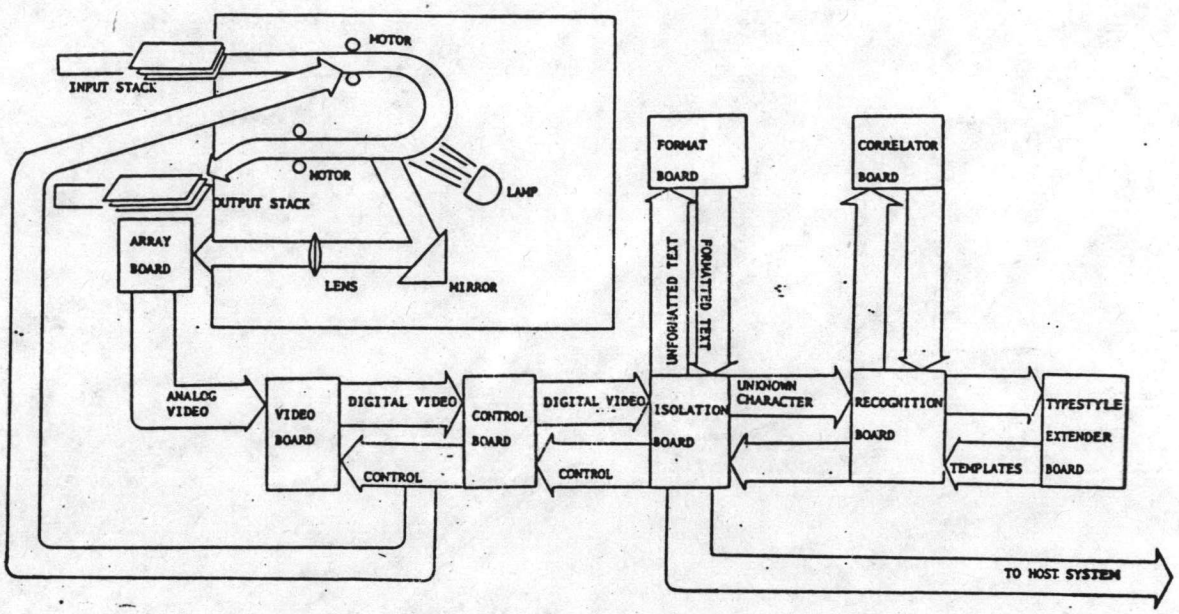
ชั้น คอรอยด์ เป็นชั้นที่มีเส้นเลือดและรงควัตถุแผ่กระจายอยู่เป็นจำนวนมาก เพื่อป้องกันมิให้แสงสว่างส่องทะลุผ่านเนื้อเยื่อชั้นเรตินาไปสู่เนื้อเยื่อชั้นสเคลอราได้ ในชั้นคอรอยด์นี้จะประกอบด้วย ม่านตา (Iris) และเลนส์ (Len) รับภาพวัตถุ ม่านตาจะเป็นเนื้อเยื่อยื่นลงมาจากด้านบนและด้านล่างของผนังคอรอยด์คล้ายกับเป็นผนังกั้นบางส่วนของเลนส์ ส่วนช่องกลางที่เหลือให้แสงส่องผ่านเข้านั้นมีลักษณะกลมเรียก พิวปิล (Pupil) ขนาดของพิวปิลจะแคบหรือกว้างขึ้นอยู่กับม่านตาจะบีบเลนส์มากหรือบีบเลนส์น้อยตามลำดับ

ส่วนเลนส์รับภาพจะมีลักษณะใสกั้นนัยน์ตาออกเป็นสองช่องคือ ช่องหน้าเลนส์ และช่องหลังเลนส์ ภายในช่องทั้งสองจะมีของเหลวใสเรียกว่า น้ำเลี้ยงลูกตาบรรจุอยู่ เลนส์รับภาพนี้มีหน้าที่ปรับขนาดของภาพที่มองเห็นให้ไปตกลงบนเรตินาเพื่อให้นัยน์ตาสามารถรับภาพที่ชัดเจนที่สุด

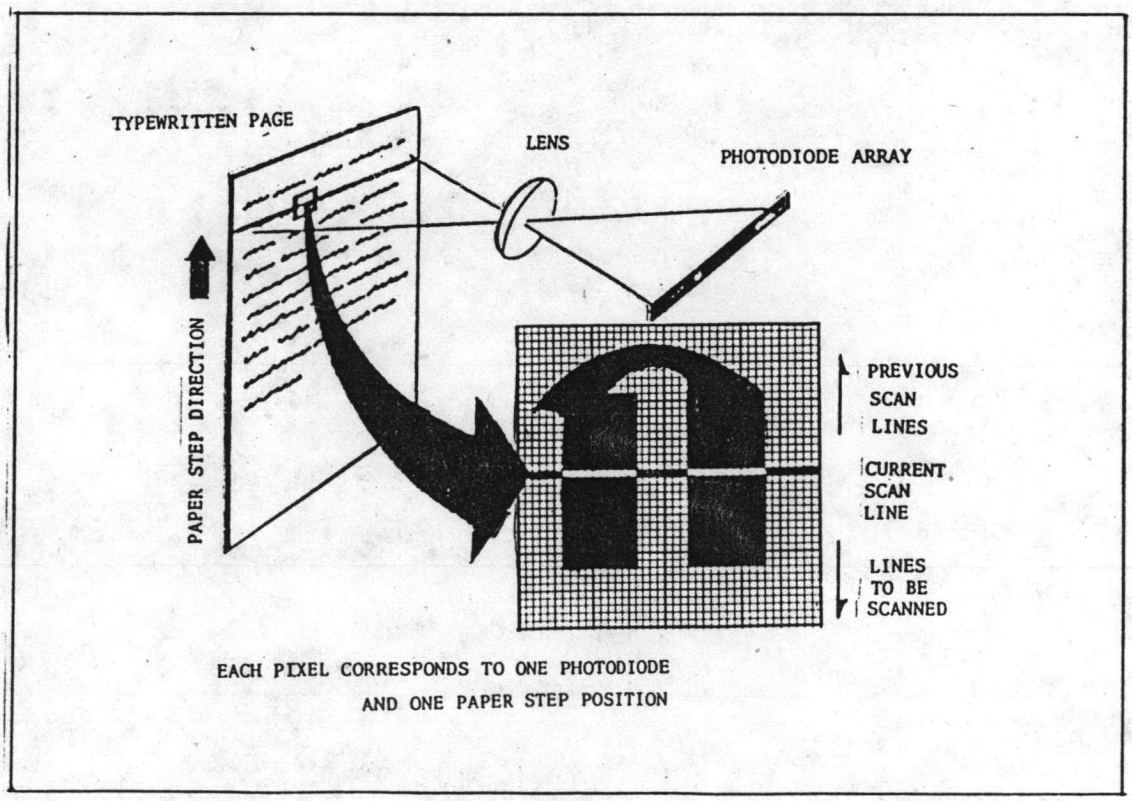
ชั้น เรตินา เป็นผนังชั้นที่มีเซลรับแสงอยู่มากมายเซลรับแสงนี้จะมีสารสีม่วงแดงชื่อ โรดอปซิล (Rhodopsin) สารนี้เมื่อถูกแสงสว่างจะเปลี่ยนโครงสร้างไปเป็นเรตินีน (Retinene) และออปซิน (Opsin) ซึ่งจะทำให้สีของโรดอปซิลซีดลง ขณะที่มีการเปลี่ยนสีของสารดังกล่าวก็จะมีกระแสประสาทเกิดขึ้น กระแสประสาทนี้จะถูกส่งไปตามเส้นประสาทนัยน์ตา เพื่อไปประมวลผลภาพยังสมองต่อไป เมื่อไม่มีแสงสว่างสารเรตินีน และออปซินก็จะเปลี่ยนกลับเป็นสารโรดอปซิลดังเดิม

2.2 ระบบการตรวจรู้อักขระ (2)

OCR (Optical Character Recognition)เป็นระบบการตรวจรู้อักขระที่มนุษย์สร้างเลียนแบบนัยน์ตาของมนุษย์ขึ้น เพื่อช่วยให้เครื่องคอมพิวเตอร์สามารถตรวจรู้อักขระได้เหมือนอย่างมนุษย์ ส่วนประกอบและหน้าที่ของเครื่อง OCR (รูปที่ 2.2) มีลักษณะทั่วๆ ไปดังนี้



รูปที่ 2.2 แสดงส่วนประกอบของเครื่อง OCR ทั่วๆ ไป



รูปที่ 2.3 แสดงส่วนรับภาพของเครื่อง OCR

2.2.1 ส่วนเคลื่อนย้ายเอกสาร (Paper Transport)

ส่วนนี้ทำหน้าที่เคลื่อนย้ายเอกสารเข้าสู่ระบบตรวจรู้อักขระในขนาดความเร็วที่กำหนด ปัญหาของส่วนเคลื่อนย้ายเอกสารนี้มีอยู่มากมายเช่น ลักษณะ ขนาด ความหนาบางของเอกสารที่ต้องการจะเคลื่อนย้ายเข้าสู่ระบบการตรวจรู้อักขระ ความเร็ว (Speed) ของการเคลื่อนย้ายเอกสารก็เป็นอีกปัญหาหนึ่งของส่วนเคลื่อนย้ายฯ นี้

2.2.2 ส่วนรับภาพ (Optical System)

เครื่องคอมพิวเตอร์ไม่สามารถนำคลื่นแสงที่แทนภาพอักขระหรือสัญญลักษณ์ต่างๆ ในเอกสารมาทำการประมวลผลโดยตรงได้ เพราะอักขระหรือสัญญลักษณ์ต่างๆ นั้นเป็นภาพสองมิติแต่คลื่นแสงซึ่งเป็นสัญญาณทางคลื่นแม่เหล็กไฟฟ้า (Electromagnetic Wave) นั้นมีมิติเดียวคือ เป็นคลื่นที่แทนค่าความเข้มของแสงมากหรือน้อยเท่านั้น จึงไม่สามารถนำคลื่นแสงมาแทนสัญญลักษณ์ที่มีสองมิติได้ ฉะนั้นส่วนรับภาพนี้จึงมีหน้าที่แปลงข้อมูลที่เป็นอักขระหรือสัญญลักษณ์ต่างต่างที่เป็นคลื่นแสงให้เป็นสัญญาณดิจิตอล (Digital) เพื่อให้เครื่องคอมพิวเตอร์สามารถประมวลผลอักขระหรือสัญญลักษณ์เหล่านั้นได้ การปฏิบัติงานของส่วนแปลงสัญญาณนี้ (รูปที่ 2.3) เริ่มจากการให้ลำแสงที่ตกกระทบอักขระบนเอกสารผ่านเลนส์ โดยที่เลนส์จะมีหน้าที่ปรับขนาดของอักขระให้มีขนาดตามที่ระบบการตรวจรู้สามารถจะนำไปประมวลผลได้ ภาพของอักขระที่ส่องผ่านเลนส์นี้จะอยู่ในรูปของคลื่นแสง ซึ่งสามารถแปลงให้อยู่ในรูปของสัญญาณไฟฟ้าโดยใช้แผงโฟโตไดโอด (Photodiode) ช่วยในการแปลง (โฟโตไดโอดเป็นอุปกรณ์อิเลคโทรนิคที่ใช้แสงเป็นตัวกระตุ้นการทำงาน) เมื่อลำแสงวิ่งผ่านบริเวณที่เป็นสีดำหรือบริเวณที่เป็นตัวอักขระบนเอกสารผ่านเลนส์ไปตกกระทบบนแผงโฟโตไดโอดบริเวณใดบริเวณที่ไม่ได้รับแสงนั้นก็จะมีสัญญาณไฟฟ้าเกิดขึ้น จากนั้นจะใช้ Analog-to-Digital Converter (A/D Converter) เป็นตัวแปลงสัญญาณไฟฟ้าบนแผงโฟโตไดโอดให้เป็นดิจิตอล

2.2.3 ส่วนตรวจรู้และตัดสินผลการตรวจรู้อักขระ (Recognition & Decision Logic)

เครื่องคอมพิวเตอร์ไม่สามารถจะแยกค่าดิจิตอลที่แปลงมาจากสัญญาณภาพของอักขระแต่ละรูปได้ว่าค่าดิจิตอลนั้นแทนอักขระรูปใด ส่วนตรวจรู้ และตัดสินผลการตรวจรู้อักขระจึงมีหน้าที่ตรวจรู้ให้ได้ว่าค่าดิจิตอลที่รับเข้ามาประมวลผลนั้นแทนค่าอักขระรูปใด โดยการใช้แบบการคำนวณเปลี่ยนค่าดิจิตอลให้เป็นรหัสที่แทนอักขระแต่ละรูป และสามารถนำรหัสเหล่านั้นไปประมวลผลด้วยเครื่องคอมพิวเตอร์ได้

เมื่อตรวจรู้และตัดสินผลการตรวจรู้อักขระได้แล้วส่วนตรวจรู้และตัดสินผลนี้จะนำรหัสของอักขระที่ตรวจรู้ได้เก็บไว้ในหน่วยความจำสำรอง (Temporary Storage) ก่อนจะส่งไปยังหน่วยประมวลผลข้อมูล (Data Processor) หรือจะส่งไปประมวลผลยังหน่วยประมวลผลข้อมูลโดยตรงแบบออนไลน์ (Online) ก็ได้แล้วแต่ลักษณะการออกแบบระบบงานนั้นๆ

ปัญหาของส่วนตรวจรู้และตัดสินผลการตรวจรู้อักขระมีดังนี้

2.2.3.1 ขนาดของอักขระที่จะตรวจรู้จะต้องมีขนาดที่พอเหมาะ เพราะอักขระที่มีขนาดเล็กการตรวจรู้อักขระจะเป็นไปด้วยความยากลำบาก

2.2.3.2 รูปแบบของอักขระที่จะตรวจรู้ถ้าเป็นรูปแบบที่เป็นมาตรฐาน หรือเป็นรูปแบบที่ใช้ในเครื่องพิมพ์ดีดทั่วๆ ไป การตรวจรู้อักขระก็จะทำได้ง่ายกว่าอักขระที่มีรูปแบบที่เขียนด้วยลายมือ

2.2.3.3 ลักษณะการวางอักขระจะต้องไม่เอียงหรือตะแคง เพราะอักขระที่เอียงหรือตะแคงนั้นรูปแบบของอักขระจะเปลี่ยนไป

2.2.3.4 ความเร็วของการตรวจรู้อักขระแต่ละรูปแบบการคำนวณที่ดีจะต้องใช้เวลาในการตรวจรู้อักขระแต่ละรูปไม่มากจนเกินไป

012099

2.2.4 ส่วนควบคุมการทำงานของระบบ (Control Unit)

หลังจากตรวจรู้อักขระแต่ละรูปในเอกสารได้แล้วส่วนนี้จะส่งผลลัพธ์ของการตรวจรู้ไปจัดรูปประโยค เพื่อที่จะให้เครื่องคอมพิวเตอร์นำผลลัพธ์ที่อยู่ในรูปประโยคไปใช้ในการประมวลผลต่อไป ในขณะเดียวกันนั้นส่วนควบคุมการทำงานของระบบก็จะสั่งงานไปยังส่วนเคลื่อนย้ายเอกสารให้เลื่อนเอกสารที่จะตรวจรู้ชุดถัดไปเข้าสู่ระบบ

ปัญหาของส่วนควบคุมการทำงานของระบบคือ ความสัมพันธ์ด้านความเร็วของส่วนเคลื่อนย้ายเอกสาร และส่วนตรวจรู้อักขระ หากส่วนหนึ่งส่วนใดปฏิบัติงานช้ากว่าอีกส่วนหนึ่งก็จะต้องมีการสร้างที่พักข้อมูลชั่วคราว (Buffer) สำหรับใช้เป็นที่เก็บข้อมูลก่อนการประมวลผล

2.3 ประวัติการค้นคว้าเกี่ยวกับระบบการตรวจรู้อักขระ (2)

การค้นคว้าเรื่องการตรวจรู้อักขระนี้ได้เริ่มมีการค้นคว้าตั้งแต่ปี ค.ศ. 1809 คือมีการพัฒนาเครื่องช่วยอ่านหนังสือสำหรับคนพิการตาบอดขึ้น ส่วนเครื่อง OCR ได้มีการพัฒนามากว่า 100 ปีแล้ว ดังมีรายละเอียดบางส่วนดังนี้

ในปี ค.ศ.1870 C.R.Carey แห่งเมืองBoston,Massachusetts ได้พัฒนาเครื่อง Retina Scanner ได้เป็นเครื่องแรก โดยใช้อุปกรณ์ที่เรียกว่า Photocells เป็นตัวพิเคราะห์อักขระแต่ละรูป

ในปี ค.ศ.1890 P.Nipkow แห่ง Poland ได้พัฒนาเครื่อง Scanning Disk ซึ่งเป็นจุดเริ่มต้นของการพัฒนาเครื่อง Television Cameras ในเวลาต่อมา

ต่อมา Emmanuel Goldberg แห่งเมือง Chicago ได้พัฒนาวิธีการพิเคราะห์ข้อมูลที่เป็นข้อความแล้วแทนด้วยรหัสมอส (Morse Code)

ในปี ค.ศ.1950 Mark Sheppard ได้ประดิษฐ์หุ่นยนต์โรบอท (Robot) ชื่อ GISMO ซึ่งเป็นหุ่นยนต์ที่สามารถอ่านและเขียนหนังสือได้ ซึ่งทำให้การค้นคว้าเรื่อง OCR ได้รับความสนใจอย่างกว้างขวางมากขึ้น

ในปี ค.ศ.1954 Jacob Rabinow ได้พัฒนาเครื่องอ่านอักขระตัวพิมพ์ใหญ่ภาษาอังกฤษที่พิมพ์ด้วยเครื่องพิมพ์ดีดได้ด้วยความเร็ว 1 ตัวอักษรต่อนาที

ในปี ค.ศ.1956 American Banker's Association ได้คิดแบบพิมพ์มาตรฐาน (Standard Font) ของอักขระที่ใช้สารแม่เหล็กเป็นส่วนผสม เรียก Magnetic Ink Character Recognition (MICR)สำหรับตรวจสอบความถูกต้องของเอกสารทางธุรกิจเช่น หมายเลขเช็ค เป็นต้น

จากความสำเร็จของการนำเทคนิค MICR มาใช้ทำให้บริษัทต่างๆ ได้เริ่มต้นหันมาใช้เทคนิค MICR ในธุรกิจของตน โดยที่แต่ละบริษัทก็จะค้นคิดแบบพิมพ์ของตนขึ้น จนในปี ค.ศ. 1966 American National Standard Institute ได้กำหนดแบบพิมพ์มาตรฐานที่ใช้ในระบบการตรวจรู้อักขระเรียกว่า USASI-A (OCR-A) และต่อมากลุ่มประเทศทางแถบยุโรปได้มีการพัฒนาแบบพิมพ์มาตรฐานของตนขึ้นใช้ เรียก OCR-B

หลังปี ค.ศ.1970 รูปแบบของอักขระที่ใช้กับเครื่อง OCR นอกจากจะเป็นตัวอักขระแล้วยังมีการพัฒนาให้เครื่อง OCR สามารถตรวจรู้รหัสที่เป็นกลุ่มเส้นขนานขนาดหนาบางไม่เท่ากันวางเรียงกันเป็นแถบที่เรียกว่า Bar Code จำนวนเส้นและขนาดความหนาบางของ Bar Code นี้สามารถใช้แทนความหมายต่างๆ ทางธุรกิจได้มากมายเช่น ชื่อสินค้า ราคาสินค้า จำนวนสินค้าคงคลัง ฯลฯ

ปัจจุบันนี้บริษัทผู้นำในวงการคอมพิวเตอร์หลายแห่งไม่ว่าจะเป็นบริษัท IBM หรือ Bell Laboratory ต่างก็พยายามค้นคว้าวิจัยและ พัฒนาเทคโนโลยีเรื่องการตรวจรู้อักขระนี้กันอย่างกว้างขวาง

2.4 เทคนิคการตรวจรู้อักขระ

เทคนิคการตรวจรู้อักขระนี้แบ่งออกเป็น 3 ลักษณะใหญ่ๆ คือ

2.4.1 การจับคู่เปรียบเทียบกับอักขระต้นแบบ (Template Matching)

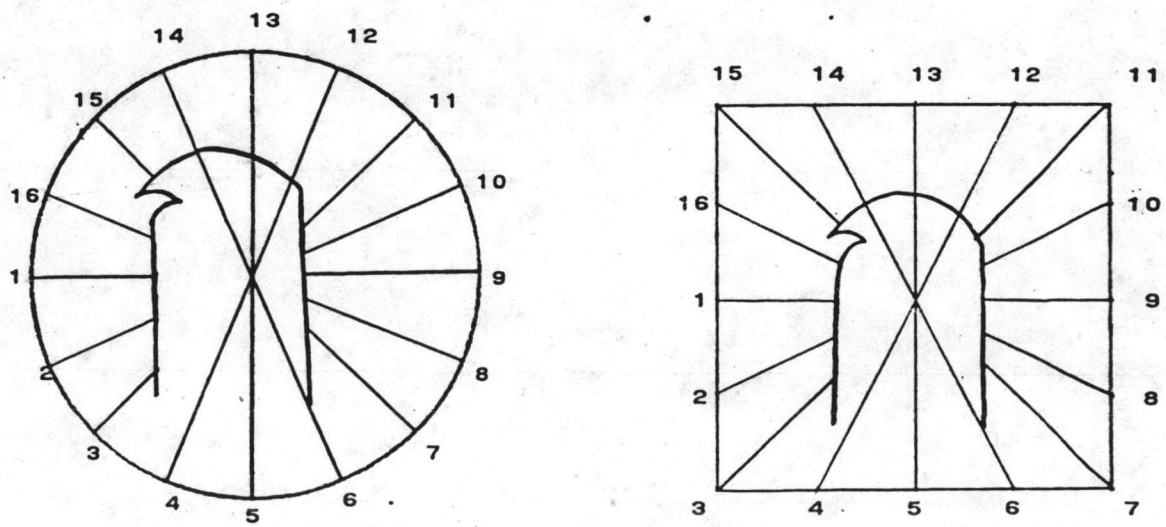
การตรวจรู้อักขระด้วยวิธีนี้มีหลักการปฏิบัติโดยการนำอักขระที่ต้องการตรวจรู้มาจับคู่เปรียบเทียบกับอักขระต้นแบบที่เก็บไว้ในสื่อข้อมูล เมื่ออักขระที่ต้องการตรวจรู้สามารถจับคู่กับอักขระต้นแบบรูปใดก็จะได้ผลลัพธ์ของการตรวจรู้อักขระนั้น หากไม่สามารถจับคู่อักขระที่ต้องการตรวจรู้กับอักขระต้นแบบแสดงว่าไม่สามารถตรวจรู้อักขระนั้นได้ เครื่อง OCR ก็จะหาวิธีการต่างๆ ที่จะเปลี่ยนรูปแบบของอักขระให้ต่างไปจากเดิมเล็กน้อยเช่น การตัดทอนอักขระบางส่วนเพื่อให้รูปแบบของอักขระมีความคมชัดยิ่งขึ้น ฯลฯ แล้วจึงเริ่มทำการจับคู่เปรียบเทียบกับอักขระต้นแบบใหม่ จนกว่าจะสามารถตรวจรู้อักขระเหล่านั้นได้ หากเปลี่ยนรูปแบบของอักขระแล้วยังไม่สามารถตรวจรู้อักขระได้ เครื่อง OCR บางระบบจะให้ผู้ควบคุมเครื่อง OCR นั้นใส่ค่าที่ถูกต้องของอักขระที่ไม่สามารถจะตรวจรู้นั้น

2.4.2 การตรวจรู้จากข้อมูลทางสถิติ (3) (Statistical Methods)

การตรวจรู้อักขระวิธีนี้จะใช้ข้อมูลที่เป็นตัวเลขในการวิเคราะห์หารหัสของอักขระ ตัวเลขดังกล่าวได้จากขั้นตอนการจัดเตรียมภาพก่อนการประมวลผล (Preprocessing) และใช้วิธีการวิเคราะห์และแยกภาพโดยสมการทางคณิตศาสตร์ (Deterministic Classification Techniques) ในการหา หรือในบางวิธีจะใช้ทฤษฎีทางสถิติเรื่องความน่าจะเป็นเข้าช่วยเช่น วิธีของ Bayesian Estimation เป็นต้น

ตัวอย่างต่อไปนี้เป็นลักษณะหนึ่งของการตรวจรู้อักขระโดยวิธีการทางสถิติ กล่าวคือจัดเตรียมอักขระที่ต้องการจะตรวจรู้ไว้ในพื้นที่ของรูปวงกลมหรือสี่เหลี่ยมที่มีขนาดพอเหมาะ โดยพื้นที่ดังกล่าวจะมีเส้นตรงลากแบ่งพื้นที่นั้นออกเป็นส่วนๆ (รูปที่ 2.4) เส้นตรงที่ลากแบ่งพื้นที่ดังกล่าวจะลากแบ่งอักขระที่จะตรวจรู้ออกเป็นส่วนๆ ตามแต่ที่เส้นตรงนั้นจะลากมาพบอักขระนั้นณ ที่ตำแหน่งใด เมื่ออักขระที่ต้องการจะตรวจรู้ถูกแบ่งออกเป็นส่วนๆ แล้ว อักขระแต่ละส่วนก็จะมีค่าเวคเตอร์ (Vector)เฉพาะ หากนำหลักวิชาทางสถิติมาใช้หาค่าต่างๆที่สามารถใช้ตรวจรู้อักขระนั้นๆได้เช่น ค่าผลเฉลี่ยของเวกเตอร์(Mean Vector) ของอักขระแต่ละรูป เป็นต้น เราก็สามารถนำค่าทางสถิติดังกล่าวมาใช้ในการตรวจรู้อักขระได้ เพราะค่าสถิติเหล่านั้นจะมีค่าเฉพาะสำหรับอักขระแต่ละรูป

การตรวจรู้อักขระวิธีนี้สามารถนำไปใช้ในการตรวจรู้อักขระที่มีลักษณะคะแคงหรือหัวกลับได้ เพราะลักษณะการจัดวางอักขระจะไม่มีผลต่อค่าทางสถิติที่ใช้ในการตรวจรู้ อย่างไรก็ตามการตรวจรู้วิธีนี้เหมาะสำหรับการตรวจรู้ภาพที่ไม่ต้องการความถูกต้องแม่นยำสูงเช่น การตรวจรู้ภาพถ่ายจากดาวเทียม เป็นต้น แต่การตรวจรู้วิธีนี้ไม่เหมาะสำหรับการตรวจรู้อักขระ เพราะผลลัพธ์ของการตรวจรู้อักขระวิธีนี้ยังไม่ถูกต้องสมบูรณ์



รูปที่ 2.4 แสดงการแบ่งอักขระ "ก" เพื่อการตรวจรู้จากข้อมูลทางสถิติของอักขระ

2.4.3 การตรวจรู้จากหลักโครงสร้างของอักขระ (Syntactic Methods)

การตรวจรู้อักขระวิธีนี้จะอาศัยแบบการคำนวณที่สามารถเปลี่ยนรูปแบบของอักขระให้เป็นรหัสที่สามารถใช้อ้างอิงถึงอักขระนั้นๆ ได้ โดยรหัสเหล่านั้นจะต้องเป็นรหัสที่ไม่ซ้ำกันสำหรับอักขระที่ต่างรูปกัน และจะมีค่าของรหัสเหมือนกันหรือใกล้เคียงกันสำหรับอักขระรูปเดียวกันแต่ต่างแบบพิมพ์กันเมื่อนำรหัสของอักขระทุกรูปมารวมกันก็จะได้เป็นรหัสต้นแบบที่สามารถใช้ในการเปรียบเทียบกับรหัสของอักขระที่ต้องการจะตรวจรู้ได้ วิธีการที่ใช้ในการค้นหารหัสที่จะตรวจรู้กับรหัสต้นแบบนั้นจะใช้วิธีการสร้างกฎเกณฑ์หรือหลักไวยกรณ์ (Pattern Grammar) ของรหัสของอักขระ เพื่อใช้ในการค้นหารหัสของอักขระที่ต้องการจะตรวจรู้กับรหัสต้นแบบ เพราะอักขระรูปเดียวกันแต่ต่างแบบพิมพ์กันย่อมจะมีค่ารหัสที่คล้ายกันจนสามารถใช้หลักเกณฑ์บางอย่างซึ่งสามารถใช้เป็นหลักไวยกรณ์ของรหัสนั้นมาอ้างอิงถึงได้

สำหรับงานวิจัยนี้ได้ใช้วิธีการตรวจรู้อักขระแบบการตรวจรู้จากรหัสของอักขระแต่ละรูป ด้วยการพิจารณารูปแบบของอักขระแต่ละรูปตามแนวแถว (Row) และแนวสดมภ์ (Column) เพื่อกำหนดเป็นรหัส สำหรับใช้ในการตรวจรู้อักขระต่อไป