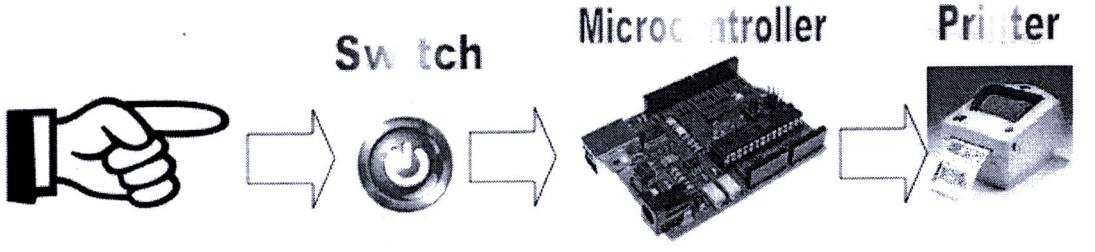
บทที่ 2
งานวิจัยและทฤษฎีที่เกี่ยวข้อง

ในเนื้อหาของบทนี้ ได้มีการศึกษางานวิจัยที่มีความเกี่ยวข้อง และหลักการต่างๆ ที่ได้นำมาใช้ในโครงงานวิจัยนี้ ซึ่งประกอบไปด้วยเนื้อหาส่วนพื้นฐานของการประมวลผลภาพ กล้องC328R และโครงข่ายประสาทเทียม

2.1 กล่าวนำ

ในอดีตการสร้างเครื่องออกบัตรคิวมักสร้างขึ้นอย่างง่ายๆ เพื่อใช้ในการจัดลำดับของผู้ใช้บริการ เช่น ธนาคาร เป็นต้น โดยโครงสร้างการทำงานก็คือเมื่อผู้ใช้บริการกดขอบัตรคิว เครื่องก็จะพิมพ์หมายเลขคิวออกมาให้ ซึ่งองค์ประกอบของเครื่องประกอบไปด้วย สวิตช์กดขอคิว ส่วนควบคุม (ไมโครคอนโทรลเลอร์) และส่วนของเครื่องพิมพ์ โดยปกติมักจะใช้แบบเครื่องพิมพ์ความร้อน ดังแสดงในรูปที่ 2.1



รูปที่ 2.1 แผนผังการทำงานของเครื่องออกบัตรคิวแบบทั่วไป

สำหรับงานบริการบางรูปแบบที่ต้องการระบุตัวตนของผู้กด เช่น การบริการทางการแพทย์ เครื่องออกบัตรคิวอย่างง่ายดังกล่าวแล้วข้างต้นไม่สามารถทำได้ ดังนั้นในงานวิจัยนี้จึงได้นำเสนอ การใช้ระบบประมวลผลผลสัญญาณภาพมาทำงานร่วมกับการออกบัตรคิวแบบดั้งเดิม เพื่อให้สามารถระบุตัวตนของผู้กดได้ โดยเครื่องออกบัตรคิวในโครงการวิจัยนี้จะเป็นประกอบไปด้วย ส่วนของสวิตช์กดขอคิว กล้องถ่ายภาพบัตรประจำตัวผู้ใช้บริการและเครื่องพิมพ์ การทำงานของเครื่องคือ เมื่อผู้ใช้บริการนำบัตรไปทาบ ณ ตำแหน่งที่กำหนด กล้องจะถ่ายภาพบัตร และส่งเข้าประมวลผลเพื่ออ่านเลขประจำตัวบนบัตร และเมื่อได้เลขประจำตัวผู้ใช้บริการแล้วก็จะทำตรวจสอบกับฐานข้อมูลของโรงพยาบาลว่าเป็นผู้มาใช้บริการจริงหรือไม่ หากจริงก็จะทำการพิมพ์ชื่อและหมายเลขประจำตัวของผู้มาใช้บริการลงบนบัตรคิวพร้อมทั้งหมายเลขลำดับคิว ซึ่งก็จะทำให้สามารถระบุตัวตนผู้ใช้บริการได้ในระดับหนึ่ง

2.2 การประมวลผลภาพเชิงตัวเลข (Digital image processing)

การประมวลผลภาพเชิงตัวเลข หรือ DIP (Digital image processing) หมายถึงการนำภาพมาประมวลผลด้วยคอมพิวเตอร์โดยภาพที่นำมาประมวลผลนี้จะถูกแทนที่ด้วยตัวเลขให้อยู่ในรูปแบบของเมตริกซ์ แต่ภาพที่ได้โดยส่วนมากแล้วจะเป็นภาพที่ได้จากตัวรับสัญญาณซึ่งอยู่ในรูปของฟังก์ชัน f(x,y) ที่ต่อเนื่องในระนาบสองมิติ (คือแกน x และแกน y) โดยจะเป็นสัดส่วนกับความสว่างหรือความเข้มของภาพที่ตำแหน่ง x,y ซึ่งเรียกว่าระดับสีเทา (Grey level)

ภาพข้อมูลแบบดิจิตอล (Digital image) เป็นภาพที่ถูกแปลงมาจากภาพแอนาลอกให้อยู่ในรูปของตัวเลขโดยทำการนำภาพแอนาลอกมาแบ่งเป็นพื้นที่สี่เหลี่ยมเล็กๆที่เรียกว่าพิกเซล (pixel) ในแต่ละพิกเซลจะถูกระบุตำแหน่งโดย x,y และค่าระดับสีเทาของพิกเซลโดยเราสามารถแปลงภาพเป็นข้อมูลแบบดิจิตอลได้โดยมีขั้นตอนและวิธีการดังนี้ เมื่อเรานำสัญญาณแอนาลอกที่ต้องการการประมวลผลมาผ่านส่วนของ ดิจิไท

เซอร์ (Digitizer) ซึ่งจะมีหน้าที่ในการเปลี่ยนสัญญาณแอนาลอกให้เป็นสัญญาณดิจิตอล จากนั้นจึงเข้าสู่กระบวนการควอนไทซ์ (Quantizing) เพื่อที่จะประมวลผลสัญญาณด้วยระบบคอมพิวเตอร์ ฟังก์ชันของภาพ f(x,y) จะถูกทำให้เป็นสัญญาณแบบไม่ต่อเนื่องทั้งระนาบของภาพซึ่งเราเรียกว่าการสุ่มภาพ (Image sampling) ของฟังก์ชันที่ได้เรียกว่าการควอนไทเซชั่นระดับสีเทา (Grey level quantization) ก็จะได้ข้อมูลที่เป็นดิจิตอล

การวิเคราะห์ภาพดิจิตอล คือการอธิบายถึงลักษณะที่มีความสำคัญ รวมถึงการจดจำข้อมูลภาพดิจิตอล ซึ่งอินพุตของระบบเป็นข้อมูลภาพดิจิตอลและเอาต์พุตจะเป็นเครื่องหมาย ที่ใช้แทนข้อมูลภาพดิจิตอล ในการวิเคราะห์ภาพนั้นมีอยู่ด้วยกันหลายวิธี เช่นวิเคราะห์จากการทำงานของตามนุษย์ (Human vision) ซึ่งการมองเห็นของตามนุษย์นับว่าเป็นกระบวนการที่ซับซ้อน สำหรับงานที่มีความซับซ้อนซึ่งได้แก่ การพัฒนางานทางความคิด กระบวนการจดจำและการวิเคราะห์จำเป็นต้องใช้กระบวนการทางคอมพิวเตอร์เข้ามาช่วยในการวิเคราะห์

2.3 การแยกข้อมูลภาพออกเป็นส่วนๆ (Image segmentation)

การทำการแยกข้อมูล (Segmentation) จะทำให้สามารถแยกข้อมูลภาพของส่วนที่ต้องการออกมาได้ (ข้อมูลที่มีลักษณะเหมือนกับข้อมูลตัวอย่าง) วิธีการพื้นฐานสำหรับการแยกข้อมูลคือการพิจารณาแอมปลิจูดภาพ (Image amplitude) ได้แก่การพิจารณาความสว่างของภาพสำหรับภาพแบบสเกลสีเทาและความแตกต่างของสีสำหรับภาพสี นอกจากนี้ขอบของภาพและลักษณะของโครงร่าง (Texture) ก็เป็นองค์ประกอบหนึ่งที่จะทำให้สามารถทำการแยกข้อมูลได้สะดวกยิ่งขึ้น

2.3.1 วิธีแยกข้อมูลโดยแอมปลิจูด (Amplitude segmentation methods)

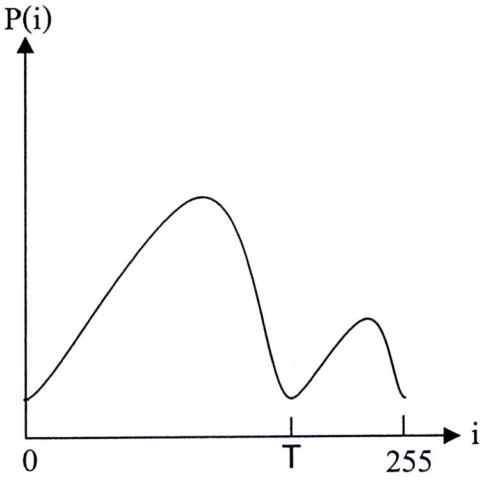
เป็นการพิจารณาความเข้มของจุดต่างๆภายในพิกเซล ซึ่งผลของการแยกข้อมูลจะขึ้นอยู่กับวิธีการเทรสโฮลส์ (Threshold) ของส่วนประกอบที่เป็นความเข้มหรือสีของภาพ ซึ่งคือการทำเทรสโฮลส์ดิงค์ความเข้มแสงสองระดับ (Bilevel luminance thresholding) สำหรับภาพบางชนิดจะมีลักษณะวัตถุที่เราสนใจซึ่งมีความเข้มคงที่เมื่อเทียบกับพื้นหลังตัวอย่างได้แก่ ภาพของตัวอักษร (Text) เป็นต้น ซึ่งภาพเหล่านี้จะมีความเข้มของวัตถุที่เราสามารถแยกออกพื้นหลังได้อย่างชัดเจน (มีความเข้มข้นสองระดับได้แก่ความเข้มของวัตถุและความเข้มของพื้นหลัง)

การทำการแยกข้อมูลสามารถทำได้โดยกำหนดค่าเทรสโฮลส์ซึ่งเป็นค่าความเข้ม เพื่อให้สามารถแยกความแตกต่างของวัตถุและพื้นหลังได้ตัวอย่างเช่น ภาพของตัวอักษรที่มีความเข้มของตัวอักษรเป็น 0 (สีดำ) และมีความเข้มของพื้นหลังเป็น 255 (สีขาว) ดังนั้นค่าเทรสโฮลส์จึงควรจะมีค่าเท่ากับ 128 เพื่อที่จะให้สามารถแยกวัตถุออกจากพื้นหลังได้ ปกติแล้วการเลือกค่าเทรสโฮลส์จะขึ้นอยู่กับฮีสโตแกรม (Histogram) ของภาพตามรูปที่ 2.2 แสดงการหาค่าเทรสโฮลส์โดยค่าเทรสโฮลส์ควรจะเลือกค่า ฮีสโตแกรม ที่อยู่ที่จุดต่ำสุดซึ่งอยู่ระหว่างจุดสูงสุด (Peaks) โดยที่ถ้าค่าข้อมูลที่ตำแหน่งใดๆมากกว่าค่า เทรสโฮลส์จะมีค่าเป็น 1 (สีขาว) และที่ค่าอื่นๆจะเป็น 0 (สีดำ) ดังสมการ 2.1

$$g(x,y) = \begin{cases} 1 & \text{if } g(x,y) > T \\ 0 & \text{otherwise} \end{cases} \qquad (2.1)$$

เมื่อ $g(x,y)$ เป็นข้อมูลภาพ ณ ตำแหน่งที่ x,y

T เป็นค่าเทรสโฮลส์



รูปที่ 2.2 การทำเทรสโฮลส์ดิงค์ความเข้มแสงสองระดับ

ในการทำภาพไบนารี่โดยการทำเทรสโฮลส์ดิงค์ให้ได้ภาพที่ดีและคมชัด ต้องเกิดจากการเลือกค่าเทรสโฮลส์ที่เหมาะสม ถ้าเลือกค่าเทรสโฮลส์ไม่เหมาะสมเช่น ค่าเทรสโฮลส์ที่มากหรือน้อยจนเกินไปภาพที่ได้จะขาดความคมชัดหรืออาจทำให้รายละเอียดของภาพขาดหายไป หรือภาพที่ได้อาจจะมืดเกินไปหรือสว่างเกินไป หรือเกิดภาพที่มีสิ่งรบกวน (Noise) เกิดขึ้นทำให้ภาพผลลัพธ์ที่ได้ไม่ชัดเจน

ค่าเทรสโฮลส์นั้นสามารถหาได้จากฮีสโตแกรมของภาพแต่ในหลายๆกรณีที่การเปลี่ยนแปลงของฮีสโตแกรมไม่สามารถบอกการเปลี่ยนแปลงระหว่างวัตถุได้อย่างชัดเจน วิธีการที่ง่ายที่สุดที่สามารถทำให้ฮีสโตแกรมหาค่าเทรสโฮลส์ได้ง่ายขึ้นก็คือการใช้วิธีการหาขอบ (Edge detection) เพื่อพิจารณาพิกเซลต่างๆของภาพว่าเป็นขอบของวัตถุ

2.3.2 หลักการหาขอบภาพ (Edge detection)

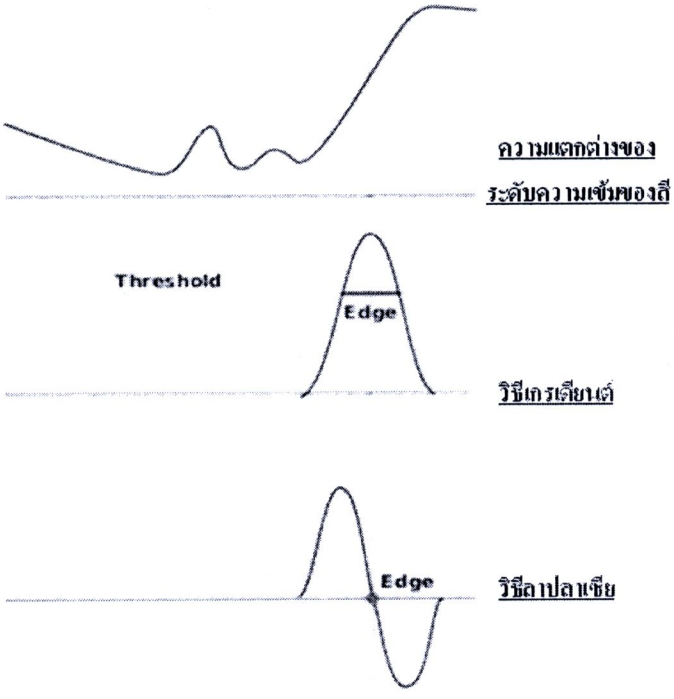
การหาขอบภาพคือการตรวจสอบว่าเส้นขอบลากผ่านหรือใกล้เคียงกับจุดใด โดยวัดจากการเปลี่ยนแปลงของความเข้มในตำแหน่งที่ใกล้เคียงกับจุดดังกล่าว ซึ่งวิธีการหาขอบนั้นมีด้วยกันหลายวิธี แต่อย่างไรก็ตามสามารถแบ่งได้เป็น 2 กลุ่มหลัก คือ

1. วิธีเกรเดียนต์ (Gradient method)

วิธีนี้จะหาขอบโดยการหาจุดต่ำสุดและจุดสูงสุดในรูปของอนุพันธ์อันดับหนึ่งของภาพ โดยจุดที่เป็นขอบจะอยู่ในส่วนที่เหนือค่าเทรสโฮลส์ดังรูปที่ 2.3 จึงอาจทำให้เส้นขอบที่ได้มีลักษณะหนา ตัวอย่างวิธีการหาขอบของกลุ่มนี้ เช่น Roberts, Prewitt, Sobel และ Canny เป็นต้น

2. วิธีลาปลาเซีย (Laplacian method)

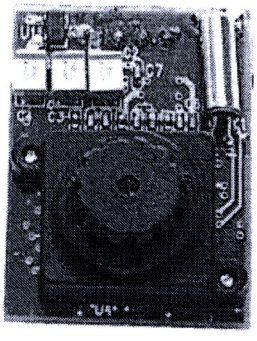
จะหาขอบโดยใช้อนุพันธ์อันดับ 2 โดยใช้จุดที่ค่าในแกนแนวตั้ง เป็น 0 (Zerocrossing) ซึ่งวิธีนี้จะใช้เวลาในการคำนวณมากกว่า วิธีเกรเดียนต์ ตัวอย่างวิธีการหาขอบของกลุ่มนี้ เช่น ลาปาเซียออฟเกาส์เซียน (Laplacian of Gaussian) และ มาร์-เฮลเดิร์ส (Marrs-Hildreth) เป็นต้น



รูปที่ 2.3 การหาขอบด้วยวิธีเกรเดียนต์และวิธีลาปลาเซียน

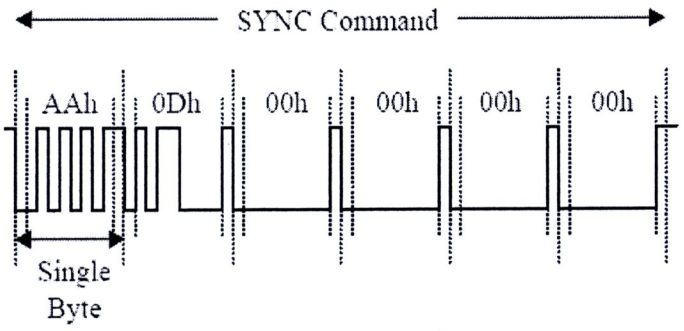
2.4 กล้อง C328R

เป็นกล้องแบบวีจีเอ (VGA) ดังรูปที่ 2.4 โดยสามารถส่งข้อมูลผ่านระบบเครือข่ายท้องถิ่นแบบไร้สายได้และผู้ใช้สามารถสั่งการกล้องให้เก็บภาพได้เมื่อต้องการโดยส่งคำสั่งเข้าตัวกล้องได้โดยตรง ซึ่งข้อมูลภาพจะถูกส่งกลับมาในรูปแบบของไฟล์เจเปค (JPEG) ผ่านทางพอร์ตสื่อสารแบบอนุกรม



รูปที่ 2.4 กล้อง C328R

ในการเริ่มการติดต่อสื่อสารกับกล้อง C328R นั้นจะต้องเริ่มโดยการส่งสัญญาณคำสั่งซิงค์ (SYNC command) ดังรูปที่ 2.5 จำนวน 25-60 ครั้ง และกล้องจะทำการส่งสัญญาณตอบรับกับสัญญาณซิงค์กลับมาพร้อมกับสัญญาณซิงค์ เมื่อผู้ใช้ส่งสัญญาณตอบรับกับสัญญาณซิงค์ของกล้องกลับไป ถือเป็นอันเสร็จสิ้นการเริ่มการติดต่อสื่อสารกับกล้อง และพร้อมสำหรับการส่งสัญญาณคำสั่งอื่นๆต่อไป โดยขั้นตอนนี้ต้องกระทำทันทีเมื่อจ่ายไฟให้กับกล้อง



รูปที่ 2.5 สัญญาณซิงค์

ในส่วนของคำสั่งของกล้อง C328R นั้นจะมีทั้งหมด 11 คำสั่ง ดังตารางที่ 2.1 โดยที่ทุกครั้งที่ส่งคำสั่งไปยังกล้อง กล้องจะต้องส่งสัญญาณตอบรับกับคำสั่งนั้นๆกลับมาเสมอทุกครั้ง หากไม่ส่งกลับมาจะไม่สามารถสั่งคำสั่งอื่นต่อไปได้ และต้องเริ่มการเชื่อมต่อกับกล้องใหม่อีกครั้ง

ตารางที่ 2.1 คำสั่งของกล้อง C328R

Command	ID Number	Parameter1	Parameter2	Parameter3	Parameter4
Initial	AA01h	00h	Color Type	RAW Resolution (Still image only)	JPEG Resolution
Get Picture	AA04h	Picture Type	00h	00h	00h
Snapshot	AA05h	Snapshot Type	Skip Frame Low Byte	Skip Frame High Byte	00h
Set Package Size	AA06h	08h	Package Size Low Byte	Package Size High Byte	00h
Set Baudrate	AA07h	1st Divider	2nd Divider	00h	00h
Reset	AA08h	Reset Type	00h	00h	xxh*
Power Off	AA09h	00h	00h	00h	00h
Data	AA0Ah	Data Type	Length Byte 0	Length Byte 1	Length Byte 2
SYNC	AA0Dh	00h	00h	00h	00h
ACK	AA0Eh	Command ID	ACK counter	00h / Package ID Byte 0	00h / Package ID Byte 1
NAK	AA0Fh	00h	NAK counter	Error Number	00h
Light Frequency	AA13h	Frequency Type	00h	00h	00h

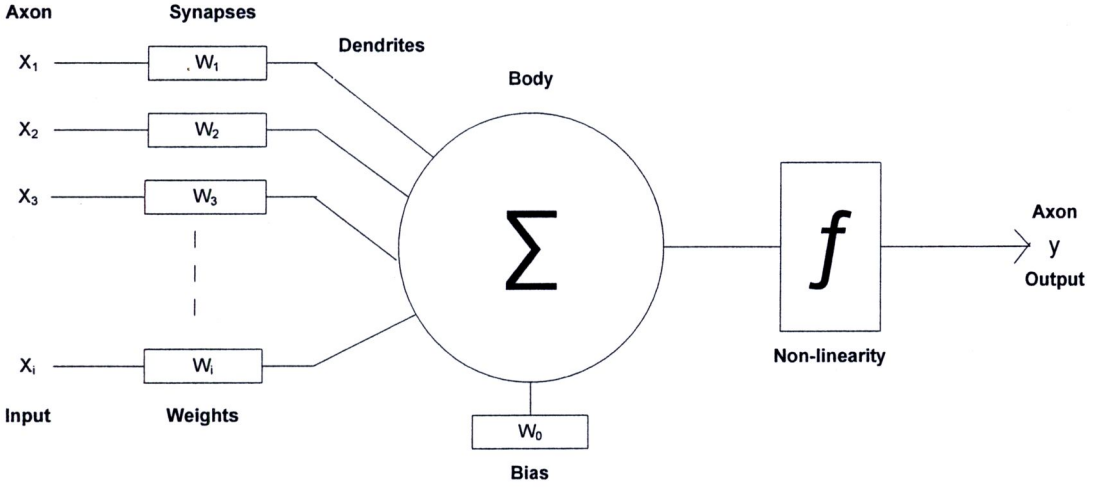
2.5 โครงข่ายประสาทเทียม (Artificial neural network)

โครงข่ายประสาทเทียมเป็นโครงสร้างการคำนวณที่นำเอาหลักการการทำงานของข่ายประสาทในสิ่งมีชีวิตมาใช้ในส่วนของการเชื่อมโยงต่างๆ ในโครงข่ายประสาทเทียมนี้การรับข้อมูลเข้าเทียบได้กับการรับข้อมูลเข้าของเซลล์ประสาททางชีววิทยาที่รับจากเส้นใยประสาทอื่นๆ และในทำนองเดียวกันการส่งข้อมูลออกของประสาทเทียมนี้เทียบได้กับสัญญาณที่ถูกส่งออกจากเซลล์ประสาททางชีววิทยาเช่นกัน สัญญาณเทียมเหล่านี้จึงอาจเปรียบได้กับสัญญาณกระแสไฟฟ้าอ่อนๆในสมองคนได้ หรืออีกนัยหนึ่งสามารถพูดได้ว่า

เซลล์ในโครงข่ายใยประสาทเทียมรับข้อมูลจากใยประสาทอื่น หรือจากแหล่งข้อมูลภายนอกแล้วแปรรูปเป็นสารสนเทศและผ่านต่อไปยังใยประสาทอื่นหรือข้อมูลออกสู่ภายนอกต่อไป

2.5.1 เซลล์ประสาทเทียม (Artificial neurons)

คือหน่วยรากฐานของโครงข่ายใยประสาทเทียมซึ่งเซลล์ประสาทเทียมนั้นไม่สามารถใช้เป็นแบบในการอธิบายการทำงานของเซลล์ประสาทของสมองมนุษย์ได้ถูกต้อง แต่เป็นการนำแนวคิดที่ได้จากการทำงานของเซลล์ประสาทของสมองมนุษย์มาประยุกต์ใช้ แบบจำลองพื้นฐานของเซลล์ประสาทเทียมแสดงดังรูปที่ 2.6 โดยอินพุตที่เข้ามายังตัวเซลล์ประสาทเทียมเปรียบได้กับสัญญาณกระตุ้น หากผลรวมของสัญญาณเข้ามีค่าสูงเกินค่าระดับเทรสโฮลส์ก็เปรียบได้กับการส่งสัญญาณไฟฟ้าออกทางแอคซอน (Axons) จากเซลล์ประสาทของสิ่งมีชีวิตนั่นเอง



รูปที่ 2.6 แบบจำลองเซลล์ประสาทเทียม

จากรูปที่ 2.6 แสดงแบบจำลองเซลล์ประสาทเทียมที่สามารถสอนให้โครงข่ายตัดสินใจได้โดยมี X_i เป็นสัญญาณอินพุตและ W_i เป็นค่าถ่วงน้ำหนัก (Weights) ที่ได้จากการฝึกสอนโครงข่าย เทียบได้กับไซแนปส์ (Synapses) ที่ใช้ในการส่งผ่านสัญญาณโดยการกระตุ้นหรือยับยั้งสัญญาณจะขึ้นอยู่กับค่าถ่วงน้ำหนัก W_i และแต่ละโหนดของโครงข่ายจะใช้แทนเซลล์ประสาทแต่ละเซลล์ส่วน y เป็นเอาต์พุตการคำนวณสัญญาณเอาต์พุตสามารถทำได้ดังสมการที่ (2.2)

$$y = f(\sum_{i=1}^{N} W_i X_i + W_0) \qquad (2.2)$$

โดยที่ W เป็นเวกเตอร์ (Vector) ของค่าถ่วงน้ำหนักซึ่งสามารถกำหนดได้ดังนี้

$$X = (W_1 + W_2 + W_3 + ... + W_n)^t \qquad (2.3)$$

โดยที่ X เป็นเวกเตอร์อินพุตและ t เป็นตัวดำเนินการทรานสโพสต์ของเมตริกซ์ฟังก์ชันกำหนดสัญญาณเอาต์พุตในสมการที่ (2.2) ถูกเรียกว่าฟังก์ชันกระตุ้น (Activation function)

2.5.2 ฟังก์ชันกระตุ้น (Activation function)

ฟังก์ชันกระตุ้นสองชนิดคือชนิดที่เป็นเชิงเส้นและชนิดที่ไม่เป็นเชิงเส้นแสดงดังรูปที่ 2.7 ฟ งก์ชันกระตุ้นที่พบบ่อยมีดังนี้

1. ฟังก์ชันเชิงเส้น (Linear function)

$$f_{AN} = (net - \theta) = \beta(net - \theta) \tag{2.4}$$

ฟังก์ชันเชิงเส้นจะสร้างเอาต์พุตที่ถูกมอดูเลตแบบเชิงเส้นโดยที่ β เป็นค่าคงที่

2. ฟังก์ชันขั้นบันได (Step function)

$$f_{AN}(net - \theta) = \begin{cases} \beta_1 \text{ if } net \geq \theta \\ \beta_2 \text{ if } net < \theta \end{cases} \tag{2.5}$$

ฟังก์ชันขั้นบันไดจะสร้างค่าเอาต์พุตที่เป็นสเกลาร์หนึ่งในสองค่าขึ้นอยู่กับค่าเทรสโฮลส์ θ

3. ฟังก์ชันลาดเอียง (Ramp function)

$$f_{AN}(net - \theta) = \begin{cases} \beta & \text{if } net - \theta \geq \beta \\ net - \theta & \text{if } |net - \theta| < \beta \\ -\beta & \text{if } net - \theta \leq \beta \end{cases} \tag{2.6}$$

ฟังก์ชันลาดเอียงเป็นการรวมฟังก์ชันเชิงเส้นและขั้นบันไดเข้าด้วยกัน

4. ฟังก์ชันซิกมอยด์ (Sigmoid function)

$$f_{AN}(net - \theta) = \frac{1}{1 + e^{-y(net - \theta)}} \tag{2.7}$$

เป็นฟังก์ชันที่นิยมใช้กันมากที่สุด

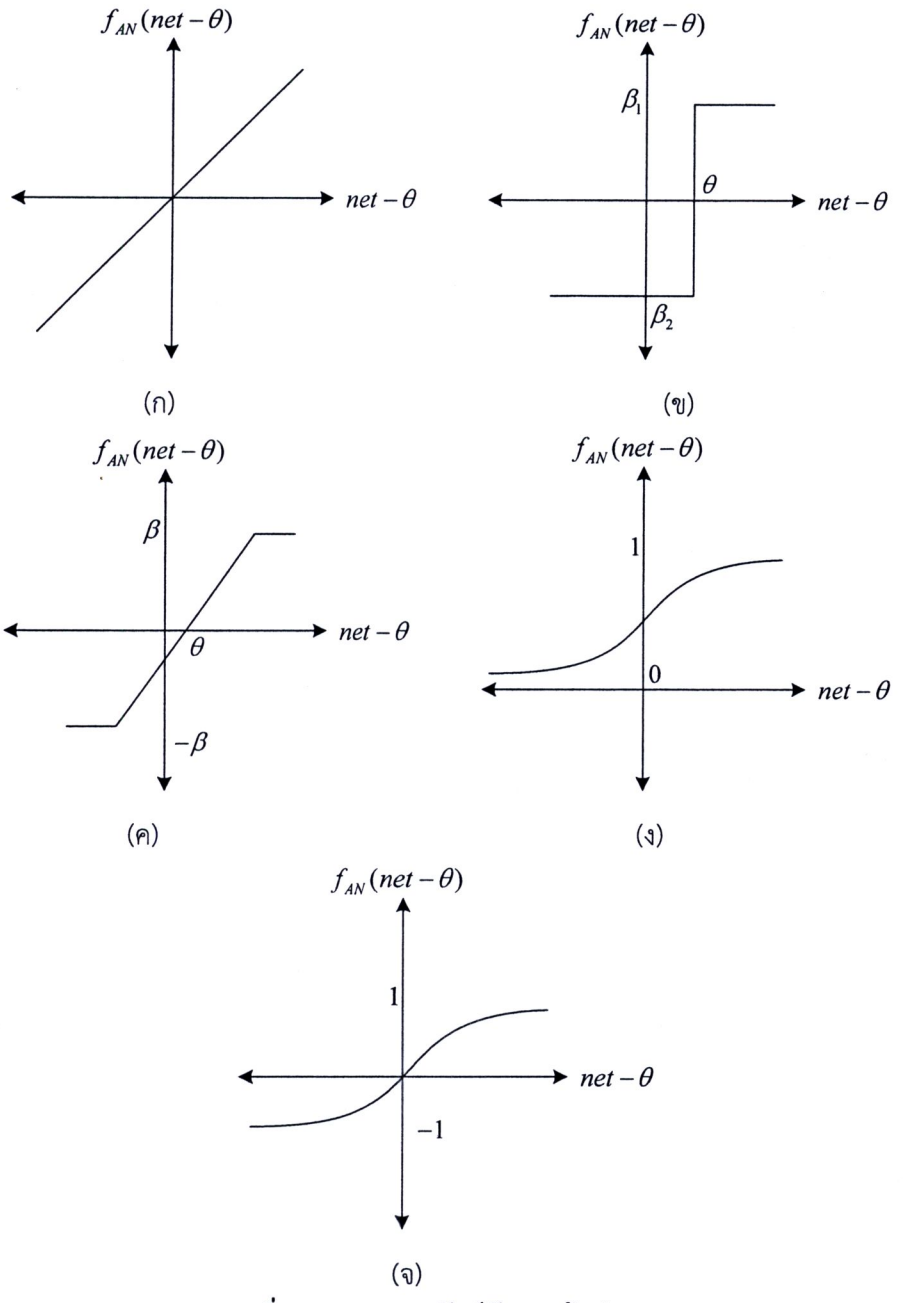
5. ฟังก์ชันไฮเปอร์แทนเจนต์ (Hyperbolic tangent function)

$$f_{AN}(net - \theta) = \frac{e^{\lambda(net - \theta)} - e^{-\lambda(net - \theta)}}{e^{\lambda(net - \theta)} + e^{-\lambda(net - \theta)}} \tag{2.8}$$

หรืออาจนิยามเป็น

$$f_{AN}(net - \theta) = \frac{2}{1 - e^{-\lambda(net - \theta)}} - 1 \tag{2.9}$$

เอาต์พุตของไฮเปอร์โบลิกแทนเจนต์อยู่ในช่วง (-1,1)



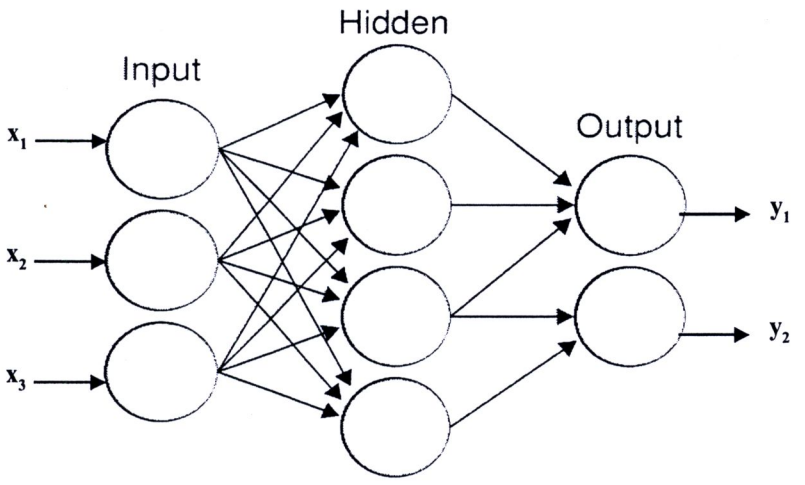
รูปที่ 2.7 กราฟของฟังก์ชันกระตุ้นต่างๆ

(ก) ฟังก์ชันเชิงเส้น (ข) ฟังก์ชันขั้นบันได (ค) ฟังก์ชันลาดเอียง

(ง) ฟังก์ชันซิกมอยด์ (จ) ฟังก์ชันไฮเปอร์แทนเจนต์

2.5.3 โครงสร้างของโครงข่ายใยประสาทเทียม

โดยปกติโครงสร้างของโครงข่ายใยประสาทเทียมจะจัดอยู่ในรูปของชั้น (Layer) กำหนดให้ชั้นแรกคือ ชั้นอินพุต (Input layer) และชั้นสุดท้ายคือชั้นเอาต์พุต (Output layer) ส่วนชั้นที่อยู่ตรงกลางเรียกว่าชั้นซ่อน (Hidden layer) ซึ่งอาจมีหนึ่งหรือมากกว่าหนึ่งชั้นก็ได้ ดังรูปที่ 2.8 จำนวนเซลล์ประสาทเทียม (Neural) ในชั้นอินพุตจะเท่ากับจำนวนอินพุต (X) และจำนวนเซลล์ประสาทเทียมในชั้นเอาต์พุตจะเท่ากับจำนวนเอาต์พุต (Y) ส่วนเซลล์ประสาทเทียมในชั้นซ่อนคือเซลล์ประสาทเทียมที่ทำหน้าที่ในการประมวลผลข้อมูลเพื่อแปลงอินพุตเป็นเอาต์พุต



รูปที่ 2.8 รูปแบบการเชื่อมโยงของเซลล์ประสาทเทียมในโครงข่ายประสาทเทียม