บทที่ 2
วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

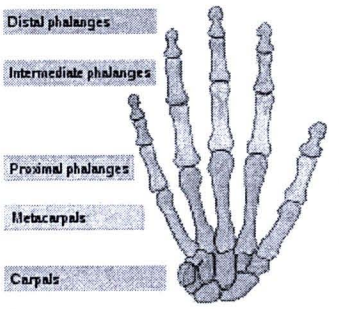
งานวิทยานิพนธ์นี้เป็นการพัฒนาและศึกษาการใช้งานของถุงมือข้อมูลเชิงแสง วรรณกรรมที่เกี่ยวข้องกับงานวิจัยนี้จำแนกออกเป็น 8 หัวข้อคือ

(1) โครงสร้างทางกายภาพของมือ
(2) การจำแนกท่าทางการเคลื่อนไหวของนิ้วมือ
(3) ภาษามือ
(4) ตัวตรวจจับ
(5) ถุงมือข้อมูล (Data glove)
(6) การแยกองค์ประกอบของภาพ
(7) โครงข่ายประสาทเทียม
(8) ขั้นตอนวิธีการส่งค่าย้อนกลับ

1. วรรณกรรมที่เกี่ยวข้อง

1.1 โครงสร้างทางกายภายของมือ

นิ้วมือเป็นอวัยวะส่วนหนึ่งที่สำคัญของมนุษย์ ในมือข้างหนึ่งประกอบด้วยนิ้วทั้งหมด 5 นิ้ว มีชื่อเรียกคือ นิ้วหัวแม่มือ นิ้วชี้ นิ้วกลาง นิ้วนาง และ นิ้วก้อย โดยแต่ละนิ้วจะมีกระดูกนิ้วมือ 3 ท่อน คือ กระดูกนิ้วมือท่อนต้น กระดูกนิ้วมือท่อนกลาง และ กระดูกนิ้วมือท่อนปลาย ทั้งนี้ ยกเว้นนิ้วหัวแม่มือ เท่านั้นที่จะมีเพียงกระดูกนิ้วมือท่อนต้นและท่อนปลาย ดังนั้นกระดูกนิ้วมือในมนุษย์จึงมีทั้งหมดข้างละ 14 ชิ้น



ภาพที่ 2.1 กระดูกมือซ้ายจากมุมมองด้านขวาบน [32]

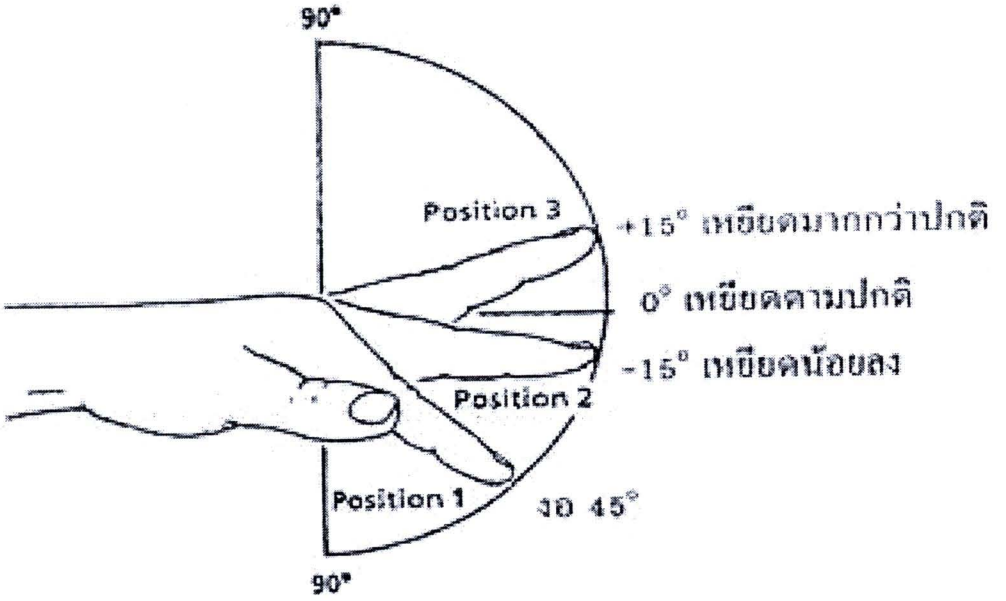
- กระดูกนิ้วมือท่อนต้น (Proximal phalanges) จากภาพที่ 2.1 คือกระดูกนิ้วมือท่อนต้นจะมีลักษณะเว้าเป็นรูปรี เพื่อรับกับปลายของกระดูกฝ่ามือ ขณะที่โคนของกระดูกนิ้วมือท่อนกลางและท่อนปลายจะมีรอยบุ๋มลงไปสองจุดที่คั่นด้วยแนวสันกลางเล็ก ๆ

- กระดูกท่อนกลาง (Intermediate phalanges) จากภาพที่ 2.1 คือกระดูกข้อกลางถัดไปจากกระดูกท่อนต้น มีลักษณะโค้งงอไปทางหลังมือเล็กน้อย และทางด้านข้างจะมีรอยที่เป็นที่เกาะของเอ็นและปลอกหุ้มเอ็นจากกล้ามเนื้อมัดต่าง ๆที่มาควบคุมการเคลื่อนไหวของนิ้วมือ

- กระดูกท่อนปลาย (Distal extremities phalanges) จากภาพที่ 2.1 คือกระดูกส่วนปลายนิ้วมีขนาดเล็กกว่าทางด้านโคนกระดูก และมีลักษณะเป็นปุ่มเล็ก ๆสองปุ่มที่คั่นด้วยร่องแคบ ๆ ในแนวกลาง พื้นผิวด้านฝ่ามือของส่วนปลายสุดของกระดูกนิ้วมือท่อนปลายจะมีลักษณะแบนออกเล็กน้อยเพื่อรองรับผิวหนังที่ใช้ในการรับสัมผัสจากปลายนิ้ว

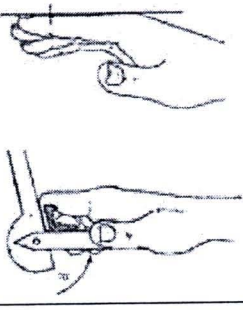
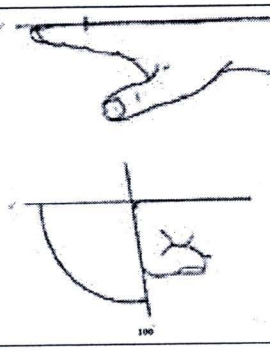
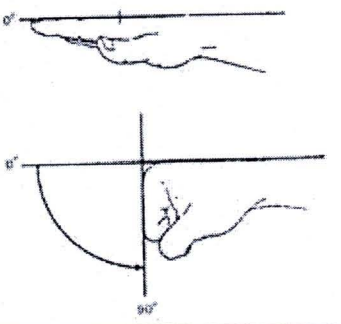
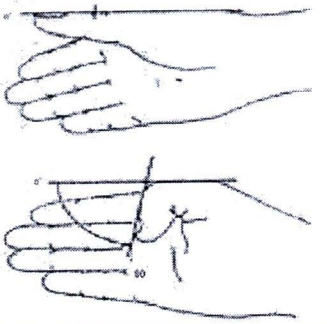
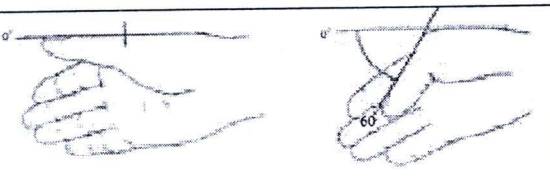
1.2 การจำแนกท่าทางการเคลื่อนไหวของนิ้วมือ

การจำแนกท่าทางการเคลื่อนไหวของนิ้วมือมีความสำคัญต่อการออกแบบถุงมือข้อมูล เนื่องจากถุงมือข้อมูลที่จะพัฒนาขึ้นจะต้องสามารถถ่ายทอดรูปแบบท่าทางการเคลื่อนไหวของนิ้วมือไปยังเครื่องคอมพิวเตอร์ได้ การเคลื่อนไหวของข้อนิ้วมือวัดได้จากจำนวนองศาที่ข้อเคลื่อนที่ได้ระหว่างการงอ และการเหยียดเต็มที่ โดยการเคลื่อนไหวของข้อนิ้วมือเริ่มต้นจาก 0 องศา ซึ่งถือเป็นตำแหน่งกลางของข้อดังภาพที่ 2.2 ข้อนิ้วแต่ละข้อจะมีองศาที่สามารถเคลื่อนไหวได้ไม่เท่ากันตามข้อจำกัดทางกายภาพของแต่ละข้อนิ้ว และสามารถสรุปได้ดังตารางที่ 2.1 ซึ่งแสดงขอบเขตองศาของการงอ และการเหยียดของข้อนิ้วมือแต่ละข้อ และการเคลื่อนไหวของนิ้วมือในลักษณะต่าง ๆ

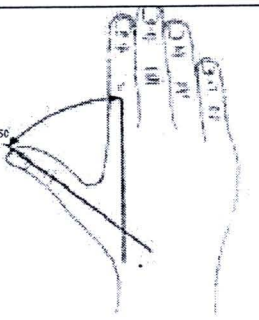
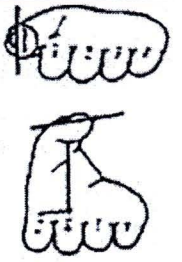


ภาพที่ 2.2 ลักษณะการงอและเหยียดนิ้วมือ [5]

ตารางที่ 2.1 องศาของการงอและการเหยียดของข้อนิ้วมือแต่ละข้อและการเคลื่อนไหวของนิ้วมือในลักษณะต่าง ๆ [5]

หัวข้อ	ภาพ
1) ข้อปลายนิ้ว สามารถเคลื่อนไหวได้จาก 0 องศา (การเหยียดตรง) จนถึง 70 องศา (งอเต็มที่)	
2) ข้อกลางนิ้ว สามารถเคลื่อนไหวได้จาก 0 องศา (การเหยียดตรง) จนถึง 100 องศา (งอเต็มที่)	
3) ข้อโคนนิ้ว สามารถเคลื่อนไหวได้จาก 0 องศา (การเหยียดตรง) จนถึง 90 องศา (งอเต็มที่)	
4) ข้อปลายนิ้วของนิ้วหัวแม่มือ สามารถเคลื่อนไหวได้จาก 0 องศา (การเหยียดตรง) จนถึง 80 องศา (งอเต็มที่)	
5) ข้อโคนนิ้วหัวแม่มือ สามารถเคลื่อนไหวได้จาก 0 องศา (การเหยียดตรง) จนถึง 60 องศา (งอเต็มที่)	

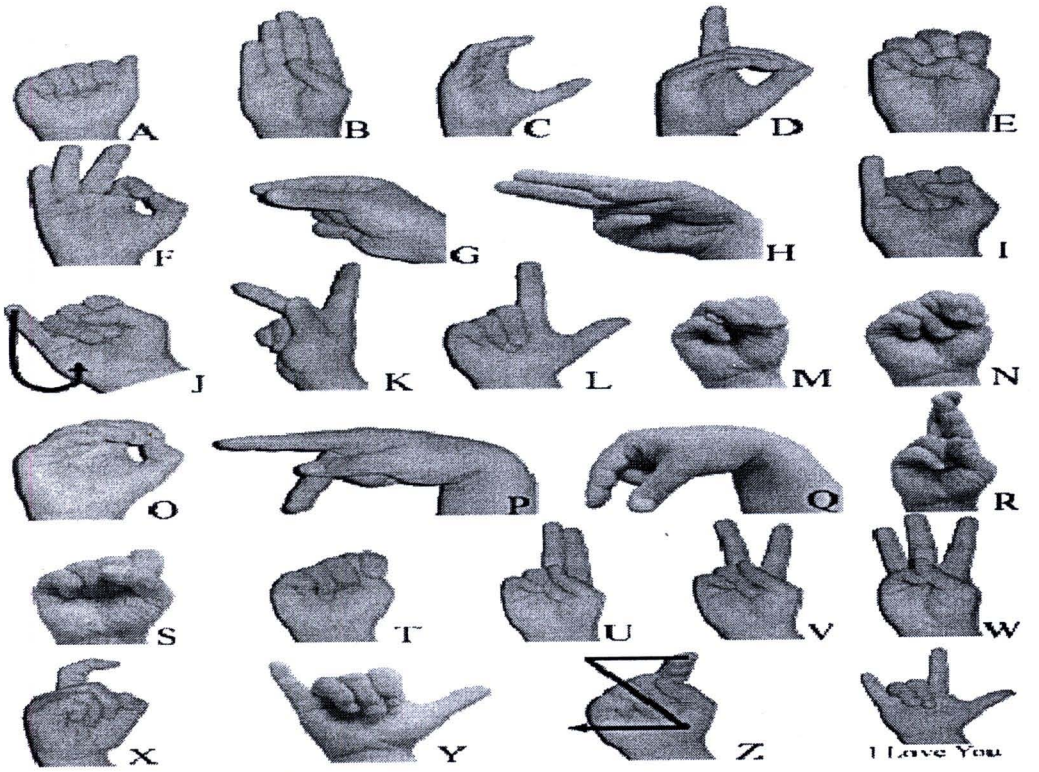
ตารางที่ 2.1 องศาของการงอและการเหยียดของข้อนิ้วมือแต่ละข้อและการเคลื่อนไหวของนิ้วมือในลักษณะต่าง ๆ (ต่อ)

หัวข้อ	ภาพ
6) การหุบนิ้วหัวแม่มือ เป็นการเคลื่อนไหวที่นิ้วหัวแม่มือเคลื่อนเข้าหามือในแนวขนานกับฝ่ามือจนกระทั่ง Flexor crease ของข้อปลายนิ้วหัวแม่มือชิดกับ Distal palmar crease เหนือข้อโคนนิ้วก้อย	
7) การกางนิ้วหัวแม่มือ เป็นการเคลื่อนไหวที่นิ้วหัวแม่มือกางออก แยกจากนิ้วชี้ในแนวขนานกับฝ่ามือ การเคลื่อนไหวของการกางนิ้วหัวแม่มือเท่ากับ 0-50 องศา	
8) การหมุนนิ้วหัวแม่มือ เป็นการเคลื่อนไหวที่นิ้วหัวแม่มือกระดกขึ้นจากแนวราบในทิศทางตั้งฉากกับฝ่ามือ และเคลื่อนเข้าหาใจกลางมือ	

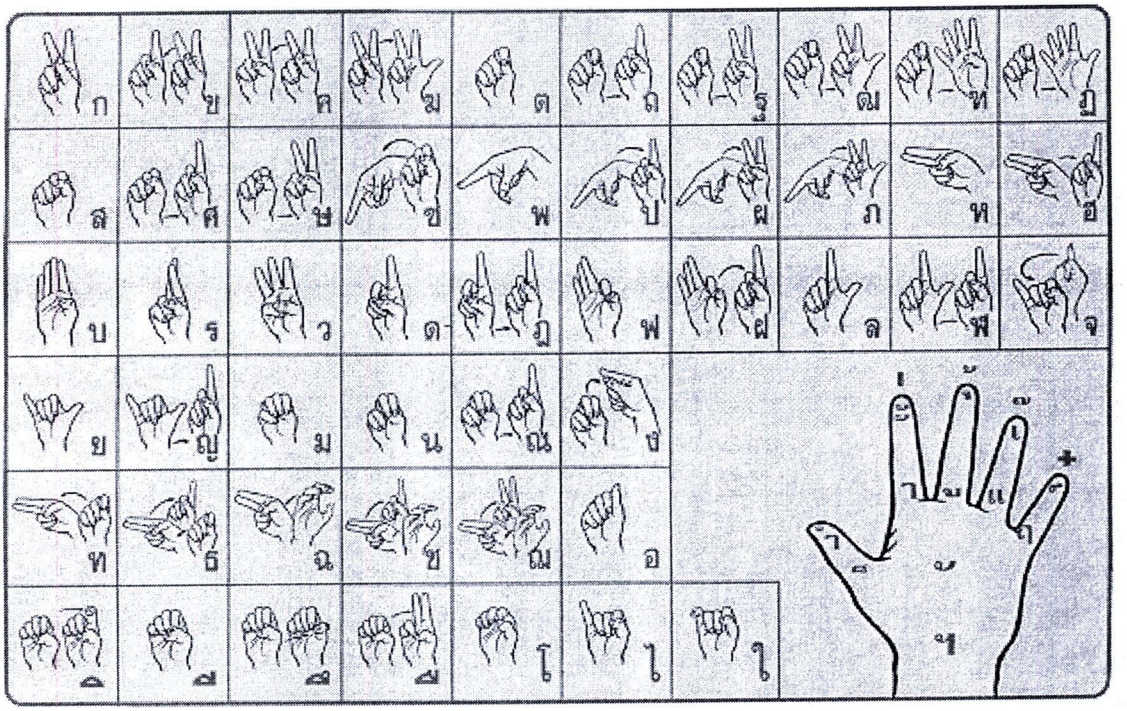
1.3 ภาษามือ

ภาษามือจัดเป็นส่วนหนึ่งของอวัจนภาษา การสื่อสารใช้ลักษณะของมือที่ทำเป็นสัญลักษณ์ การเคลื่อนไหวมือ แขนและร่างกาย และการแสดงความรู้สึกทางใบหน้าเพื่อช่วยในการสื่อสารความคิดของผู้สื่อ ภาษาสัญลักษณ์ส่วนใหญ่มักใช้ในกลุ่มที่เกี่ยวข้องกับผู้พิการทางหู ซึ่งรวมทั้งผู้พิการทางหูเอง ผู้ตีความหมาย (Interpreter) ผู้ร่วมงาน เพื่อน และครอบครัวของผู้พิการทางหูซึ่งอาจจะพอได้ยินบ้างหรือไม่ได้ยินเลย โดยมีการแทนค่าตัวอักษรทั้งในภาษาไทย (ภาพที่ 2.4) และภาษาอังกฤษ รวมถึงการสื่อออกมาแทนคำพูดหรือประโยคที่นิยมใช้ เช่น I love you (ภาพที่ 2.3)

ภาษามือไม่ได้มีแค่สำหรับคนพิการเท่านั้น มนุษย์ก็ได้ใช้ภาษามือในการสื่อสารมาตั้งแต่แรกเกิดแล้วโดยการทำกิริยาต่าง ๆทางมือเพื่อแสดงถึงความรู้สึก เช่น รัก ลาก่อน หรือ แสดงถึงคำมั่นสัญญา เป็นต้น



ภาพที่ 2.3 ท่าภาษามือในตัวอักษรภาษาอังกฤษ [15]



ภาพที่ 2.4 ท่าภาษามือในตัวอักษรภาษาไทย [12]

1.4 ตัวตรวจจับ (Sensor)

ตัวตรวจจับ คืออุปกรณ์ที่ใช้เปลี่ยนรูปแบบของพลังงานจากพลังงานที่เราสนใจไปเป็นพลังงานในรูปแบบที่ตัวตรวจจับสามารถส่งออกมาได้ ตัวตรวจจับมีหลายประเภท เช่น ตัวตรวจจับที่เกี่ยวกับความร้อน ได้แก่ เทอร์โมมิเตอร์ เทอร์โมคัปเปิล เป็นต้น ตัวตรวจจับที่เกี่ยวกับไฟฟ้า เช่น กัลวาโนมิเตอร์ แอมมิเตอร์ โวลต์มิเตอร์ เป็นต้น ตัวตรวจจับที่เกี่ยวกับแก๊สและของเหลว เช่น เครื่องวัดปริมาณการไหล เครื่องวัดปริมาณแก๊ส เป็นต้น ตัวตรวจจับที่เกี่ยวกับคุณสมบัติทางเคมี เช่น ตัวตรวจจับปริมาณของแก๊สชนิดต่างๆ เป็นต้น

ตัวตรวจจับที่เกี่ยวข้องกับงานวิจัยนี้เป็นตัวตรวจจับตำแหน่ง (Position sensor) ซึ่งมีความจำเป็นสำหรับการตรวจจับตำแหน่งของนิ้วหรือการงอของข้อนิ้วข้อต่างๆ ตัวตรวจจับตำแหน่งเมื่อแบ่งตามลักษณะการเคลื่อนที่สามารถแบ่งได้เป็นสองลักษณะใหญ่ได้แก่

1.4.1 ตัวตรวจจับตำแหน่งเชิงเส้น (Linear position sensor) ตัวตรวจจับตำแหน่งประเภทนี้จะให้ข้อมูลตำแหน่งในรูประยะทางตามแนวเส้นตรงหรือเส้นสัมผัสซึ่งมีทั้งแบบสัมพัทธ์และแบบสัมบูรณ์ ตัวตรวจจับประเภทนี้เหมาะกับงานประเภทตรวจจับตำแหน่งการเคลื่อนที่ในแนวราบ

1.4.2 ตัวตรวจจับตำแหน่งเชิงมุม (Angular position sensor) ตัวตรวจจับตำแหน่งประเภทนี้จะให้ข้อมูลตำแหน่งในรูปองศาที่หมุนไปซึ่งมีทั้งแบบสัมพัทธ์และแบบสัมบูรณ์ ตัวตรวจจับประเภทนี้เหมาะกับงานประเภทตรวจจับการหมุน

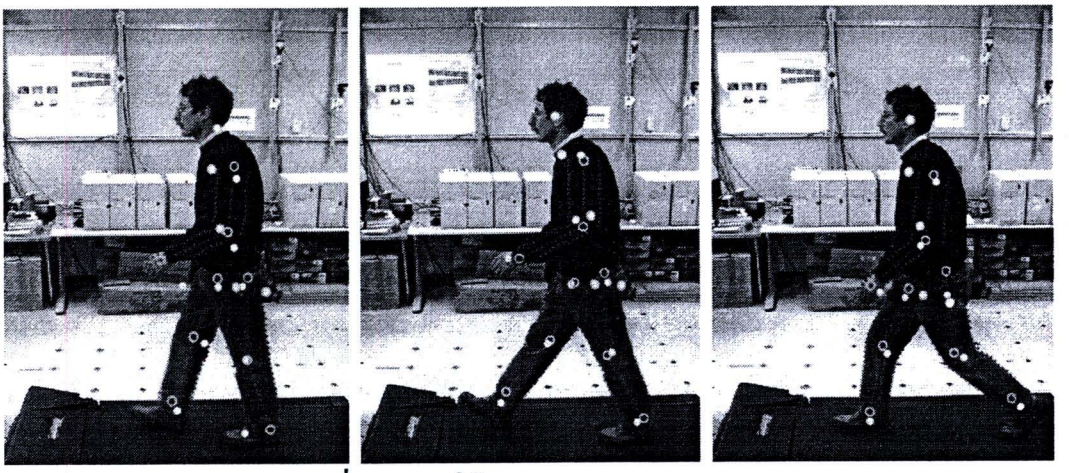
เมื่อแบ่งประเภทตามหลักการทางฟิสิกส์ที่ตัวตรวจจับใช้ในการตรวจจับตำแหน่ง ตัวตรวจจับตำแหน่งสามารถแบ่งได้เป็น 4 ชนิดใหญ่ได้แก่

(1) ตัวตรวจจับตำแหน่งประเภทใช้แสง

ตัวตรวจจับประเภทนี้แบ่งเป็นสองประเภทย่อยได้แก่ ตัวตรวจจับที่อาศัยตัวตรวจจับแสงและแผ่นสลิตสำหรับปิดบังแสงทำงานร่วมกันและตัวตรวจจับตำแหน่งที่ใช้กล้องบันทึกภาพ ในประเภทแรก ตัวตรวจจับที่อาศัยตัวตรวจจับแสงและแผ่นสลิต มีหลักการทำงานคือ แผ่นสลิตจะถูกยึดติดกับชิ้นส่วนที่เคลื่อนไหวของวัตถุที่ต้องการตรวจจับตำแหน่ง เมื่อมีการเคลื่อนที่เกิดขึ้น แผ่นสลิตจะเคลื่อนไปบังแสงและเปิดให้ตัวตรวจจับแสงรับแสงตามระยะทางที่แผ่นสลิตเคลื่อนที่ไปได้ สัญญาณการเปิดปิดของตัวตรวจจับแสงจะถูกนำไปเข้ารหัสเป็นข้อมูลระยะทางการเคลื่อนที่ ซึ่งตัวตรวจจับประเภทนี้สามารถใช้กับการตรวจจับตำแหน่งเชิงเส้นและตำแหน่งเชิงมุมได้ ข้อดีของตัวตรวจจับตำแหน่งประเภทนี้คือ มีความเที่ยงตรงสูงและมีความละเอียดมาก และมีความน่าเชื่อถือ จึงมักเป็นที่นิยมใช้ในงานอุตสาหกรรมทั่วไปรวมไปถึงใช้ในเครื่องใช้ไฟฟ้าต่างๆ

ในประเภทที่สอง ตัวตรวจจับตำแหน่งที่ใช้กล้องบันทึกภาพ เป็นการนำกล้องวิดีโอมาใช้ในการตรวจจับการเคลื่อนไหวของวัตถุในภาพ ซึ่งสามารถใช้ตรวจจับการเคลื่อนไหวของวัตถุทั้งในระยะใกล้และระยะไกลได้ โดยการเคลื่อนที่ของวัตถุจะถูกคำนวณโดยใช้การประมวลผลภาพเชิงดิจิทัล ประสิทธิภาพของระบบตรวจจับตำแหน่งประเภทนี้ขึ้นอยู่กับคุณสมบัติของระบบ

กล้องวิดีโอและคอมพิวเตอร์ที่ใช้ประมวลผล ข้อดีของระบบตรวจจับตำแหน่งประเภทนี้คือสามารถใช้ตรวจจับการเคลื่อนไหวของวัตถุโดยไม่มีการสัมผัส และสามารถตรวจจับวัตถุได้ในระยะไกล ส่วนข้อจำกัดของระบบนี้คือ อัตราการประมวลผลของระบบจะถูกจำกัดโดย Frame rate ของกล้องและความสามารถในการประมวลผลของคอมพิวเตอร์ นอกจากนี้ความโค้งของภาพอันเนื่องมาจากเลนส์กล้องก็สามารถก่อให้เกิดความผิดพลาดในการตรวจจับตำแหน่งได้

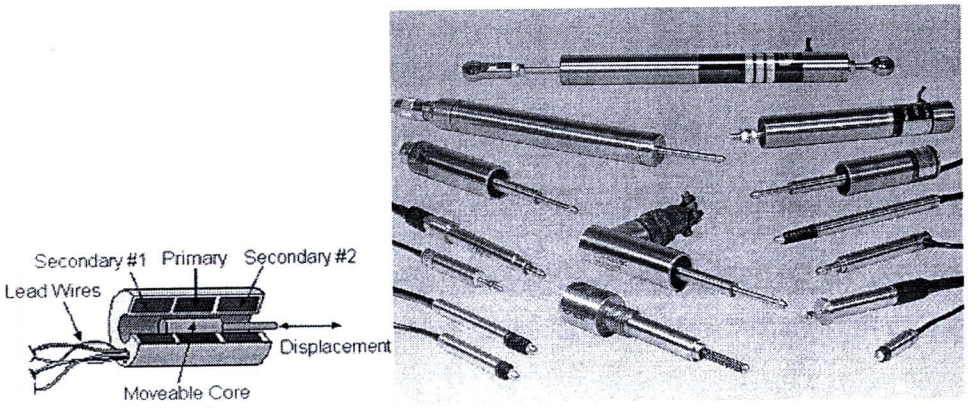


ภาพที่ 2.5 การใช้งาน Motion Tracking [26]

ภาพที่ 2.5 แสดงตัวอย่างการตรวจจับการเคลื่อนไหว โดยการนำภาพตำแหน่งของร่างกายที่ได้จากกล้องมาทำการเปรียบเทียบกับจุดอ้างอิง ซึ่งวงกลมสีดำเป็นจุดอ้างอิง ส่วนจุดสีขาวเป็นจุดที่ได้จากการคำนวณค่าข้อมูลที่รับเข้ามา ตัวอย่างนี้เป็นการตรวจจับการเคลื่อนไหวใน 2 มิติ

(2) ตัวตรวจจับตำแหน่งประเภทใช้สนามแม่เหล็ก

อุปกรณ์สำหรับตรวจจับตำแหน่งประเภทที่ใช้สนามแม่เหล็กที่มีใช้โดยทั่วไปได้แก่ Linear Differential Variable Transformer (LDVT) ซึ่งเป็นหม้อแปลงไฟฟ้าที่ประกอบด้วยขดลวดหลายขดพันรอบแกนเหล็กในทิศทางตรงข้ามกัน แกนเหล็กถูกออกแบบมาให้เคลื่อนที่ได้และถูกยึดติดกับชิ้นส่วนเคลื่อนไหวที่ต้องการระบุตำแหน่ง ขนาดของแรงดันไฟฟ้ากระแสสลับที่ออกจากขดลวดจะแปรผันตรงกับระยะทางการเคลื่อนที่ของขดลวด อุปกรณ์ชนิดนี้สามารถใช้ตรวจจับตำแหน่งของวัตถุได้ในระยะสั้น ๆตามความยาวของแกนเหล็ก นอกจาก LDVT แล้วยังมีตัวตรวจจับตำแหน่งประเภท Hall device ซึ่งเป็นอุปกรณ์อิเล็กทรอนิกส์ขนาดเล็กที่ให้ค่าแรงดันไฟฟ้าแปรเปลี่ยนไปตามความเข้มของสนามแม่เหล็ก อุปกรณ์ประเภทนี้มักใช้ในมอเตอร์ไฟฟ้าเพื่อทำหน้าที่วัดตำแหน่งของโรเตอร์

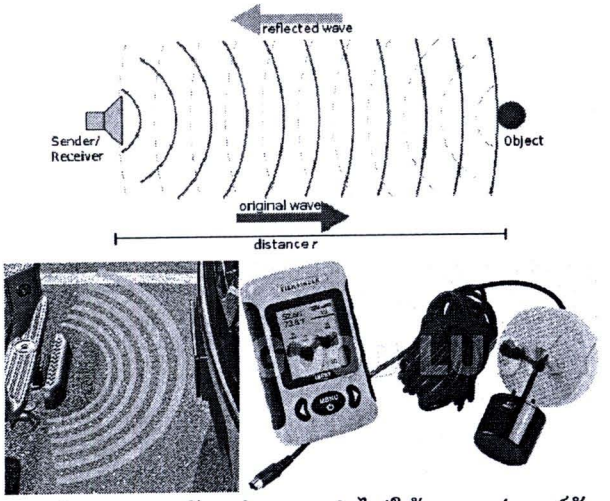


ภาพที่ 2.6 ส่วนประกอบ และตัวอย่างของ LDVT [24], [25]

ภาพที่ 2.6 แสดงส่วนประกอบ และตัวอย่างอุปกรณ์ LDVT ชนิดต่าง ๆ ที่มีจำหน่ายในท้องตลาด ซึ่งส่วนใหญ่จะเป็นอุปกรณ์ที่ใช้ในการตรวจจับการเคลื่อนไหวเพียงแกนเดียว

(3) ตัวตรวจจับตำแหน่งประเภทใช้คลื่นเสียงอัลตราโซนิกส์

อุปกรณ์ประเภทนี้ใช้หลักการของการเดินทางของคลื่นเสียงในตัวกลาง โดยตัวตรวจจับประกอบด้วยหัวส่งและหัวรับคลื่นอัลตราโซนิกส์ การวัดระยะทางทำได้โดยการกระตุ้นให้หัวส่งส่งคลื่นอัลตราโซนิกส์ไปกระทบวัตถุที่ต้องการวัดระยะทางแล้วสะท้อนกลับมายังหัวรับคลื่นอัลตราโซนิกส์ ระยะทางสามารถคำนวณจากเวลาที่คลื่นอัลตราโซนิกส์เดินทางสะท้อนกลับจากวัตถุมายังหัวรับคลื่นคูณด้วยความเร็วของคลื่นอัลตราโซนิกส์ในตัวกลางนั้น ๆ ข้อดีของตัวตรวจจับประเภทนี้คือสามารถใช้วัดระยะทางได้ในระยะไกลโดยไม่มีการสัมผัสกับวัตถุ ดังนั้นตัวตรวจจับระยะทางประเภทนี้จึงนิยมใช้ในการวัดระดับของของเหลวในถัง หรือใช้วัดระดับความสูงของวัตถุที่บรรจุภายในไซโล



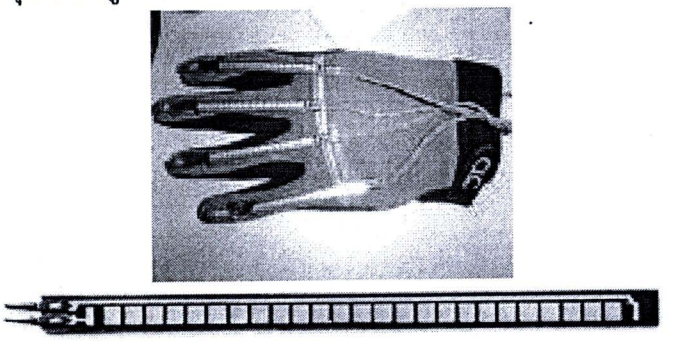
ภาพที่ 2.7 หลักการทำงานและตัวอย่างการนำไปใช้ของอุปกรณ์วัดระยะทางโดยใช้คลื่นอัลตราโซนิกส์ [21], [27]

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

ภาพที่ 2.7 แสดงหลักการทำงานเบื้องต้น และตัวอย่างอุปกรณ์วัดระยะทางโดยใช้คลื่นอัลตราโซนิกส์ที่มีการนำมาประยุกต์ใช้งาน

(4) ตัวตรวจจับตำแหน่งประเภทใช้ความต้านทาน

อุปกรณ์ตรวจจับตำแหน่งประเภทนี้ใช้ตัวต้านทานไฟฟ้าชนิดพิเศษที่สามารถเปลี่ยนค่าความต้านทานไฟฟ้าตามการยืดหรือโค้งงอของตัวต้านทานได้ ตัวอย่างตัวต้านทานประเภทนี้ได้แก่ Strain gauge และ Flex sensor โดย Strain gauge เป็นตัวต้านทานที่สามารถเปลี่ยนค่าความต้านทานไฟฟ้าได้เมื่อมีการยืดหรือหด ดัง Strain gauge จึงนิยมใช้ใน Load cell สำหรับชั่งน้ำหนักหรือสำหรับวัดระยะทางการเคลื่อนที่ในระดับเล็ก ๆของโครงสร้างขนาดใหญ่ สำหรับ Flex sensor เป็นตัวต้านทานไฟฟ้าที่มีลักษณะเป็นแถบยาวและสามารถเปลี่ยนค่าความต้านทานไฟฟ้าตามการโค้งงอของตัวต้านทานได้ ตัวตรวจจับประเภทนี้นิยมใช้ในการวัดการงอของนิ้วมือในถุงมือข้อมูลที่มีจำหน่ายทั่วไปในท้องตลาด เนื่องจากตัวตรวจจับประเภทนี้มีขนาดเล็ก อย่างไรก็ตาม Flex sensor ยังมีราคาแพง และในถุงมือข้อมูลแต่ละข้างจะต้องใช้ Flex sensor เป็นจำนวนมาก จึงทำให้ถุงมือข้อมูลมีราคาแพง



ภาพที่ 2.8 Flexible bend sensor และตัวอย่างการนำไปใช้งาน [22], [23]

ภาพที่ 2.8 แสดงตัวอย่างของ Flexible bend sensor และตัวอย่างการนำไปใช้งานบนถุงมือข้อมูล

ตัวตรวจจับตำแหน่งที่กล่าวมาทั้งหมดนี้ ส่วนใหญ่ไม่เหมาะสมที่จะนำมาติดตั้งบนถุงมือข้อมูลเชิงแสงที่พัฒนาขึ้นเนื่องมาจาก ตัวตรวจจับการงอของนิ้วจะต้องมีขนาดเล็กพอที่จะติดตั้งบนถุงมือได้ ซึ่งจะมีเพียง Flex sensor หรือ Hall device ที่สามารถนำมาใช้ได้ แต่ตัวตรวจจับเหล่านี้มีราคาแพงและต้องการวงจรขับสัญญาณที่ซับซ้อนมาควบคุม ดังนั้นผู้วิจัยจึงได้เสนอแนวทางใหม่ในการพัฒนาตัวตรวจจับตำแหน่งโดยใช้เส้นเอ็นและกล้องเว็บทำงานร่วมกันซึ่งจะได้กล่าวถึงในรายละเอียดต่อไป

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย
วันที่ 22 มิ.ย. 2555
เลขทะเบียน 190701
เลขเรียกหนังสือ

1.5 ถุงมือข้อมูล

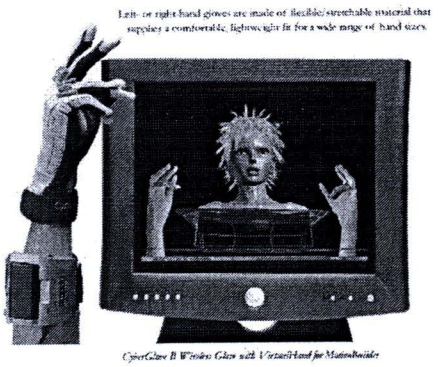
ถุงมือข้อมูล (Data glove) เป็นอุปกรณ์ถ่ายทอดข้อมูลท่าทางของมือเข้าสู่คอมพิวเตอร์ โดยใช้งานควบคู่กันระหว่างถุงมือกับตัวตรวจจับ เพื่อแปลงข้อมูลทางกายภาพของมือ เช่น การงอของนิ้ว ให้เป็นข้อมูลเชิงดิจิทัล เพื่อให้คอมพิวเตอร์นำไปประมวลผลต่อไป

ในปัจจุบันได้มีบริษัทหลายแห่งผลิตถุงมือข้อมูลออกมาเพื่อใช้งาน เช่น บริษัท 5DT นินเท็นโด (Nintendo) เป็นต้น แต่ภายในประเทศไทยยังไม่มีการใช้งานถุงมือข้อมูลเพราะยังไม่มีเทคโนโลยีที่รองรับการใช้งานภายในประเทศถุงมือข้อมูลที่มีขายตามท้องตลาดนั้นยังมีราคาที่แพงมากเนื่องจากเป็นการใช้ในงานเฉพาะอย่าง จึงทำให้ไม่มีการใช้งานอย่างแพร่หลาย ตัวอย่างถุงมือข้อมูลที่มีจำหน่ายในปัจจุบัน ได้แก่

CyberGlove เป็นถุงมือข้อมูลแบบ Wireless มี Sensor ที่ประกอบด้วย Bend sensors, Abduction sensors และ Plus sensors ที่ตรวจจับการเคลื่อนไหวของนิ้วโดยอาศัยการงอตัวและการกางของนิ้วมือไปเป็นข้อมูลเชิงดิจิทัลแบบเวลาจริงที่ให้ค่าเป็น 3 แกนได้ ให้ความละเอียดของข้อมูลที่ 12 bits และมีจำนวนตัวตรวจจับ 22 ตัว (ภาพที่ 2.9 ซ้าย) ถุงมือนี้สามารถนำไปใช้สร้างงาน 3D ได้ ซึ่งเหมาะสำหรับการสร้างภาพยนตร์ Animation การสร้างเกม การสร้างการ์ตูน และการสร้างงานศิลปะต่างๆ ถุงมือใช้เทคโนโลยีบลูทูท (Bluetooth) สามารถรับส่งสัญญาณได้ไกล 30 ft การใช้งาน CyberGlove ยังมีอุปกรณ์เสริม ที่สามารถนำไปใช้กับงานเฉพาะอย่างได้เช่น

- Cyber Grasp สามารถใช้ถุงมือจำลองสภาพในการจับวัสดุอุปกรณ์ต่างๆได้ ในการใช้งานจะมีอุปกรณ์ที่ใช้ปรับแรงต้านที่นิ้วแต่ละนิ้วมือติดที่ติดเพิ่มเข้ามาที่ CyberGlove แล้วปรับแรงต้านตามลักษณะงานที่ใช้ได้ สามารถให้แรงต้านมากสุดที่ 12N ต่อนิ้ว การประยุกต์ใช้งาน Cyber Grasp สามารถนำไปใช้ได้อย่างหลากหลาย เช่น เป็นเครื่องมือทางการแพทย์ การสร้างแบบจำลอง การออกแบบ (CAD) เครื่องมือควบคุมวัสดุที่เป็นอันตราย

Specifications	
Sensor Resolution	< 1 degree
Sensor Repeatability	3 degrees
Sensor Linearity	0.6 %
Sensor Data Rate	90 records/sec
Glove	One size fits most; 3 oz.
Interface Unit	7.62 x 11.56 x 2.64 cm
Cable	7.62 m
Interface	RS-232



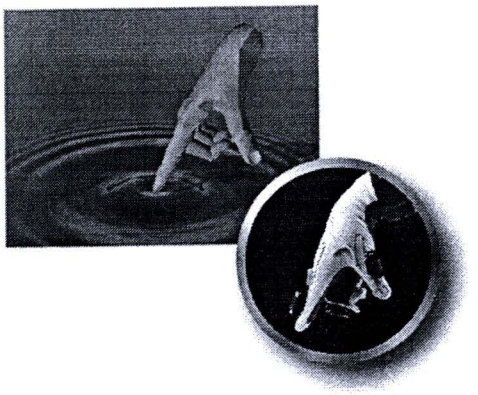
CyberGlove II Wireless Glove with VirtualHand for MotionBuilder

ภาพที่ 2.9 รูปแบบการใช้งาน CyberGlove และคุณลักษณะ [19]

- CyberTouch สามารถใช้ถุงมือจำลองสภาพในการแตะสัมผัสพื้นผิวต่าง ๆหรือสัมผัสส่วนต่าง ๆได้ โดยการใช้งานจะมีอุปกรณ์ขนาดเล็กที่ติดที่นิ้วแต่ละนิ้วและบริเวณอุ้งมือ ผู้ใช้สามารถทดลองตรวจจับการทำงานแบบง่าย ๆ เช่น ตรวจจับชีพจร หรือสัมผัสการสั่นสะเทือนต่าง ๆ สามารถนำไปใช้ในการจำลองสภาพต่าง ๆ เหมาะสำหรับการนำไปใช้ในงานที่ต้องการการตอบสนองกลับมา เช่น การใช้นิ้วมือแตะที่ผิวน้ำแล้วจะเกิดการกระเพื่อมของน้ำกลับมา ดังแสดงในภาพที่ 2.10

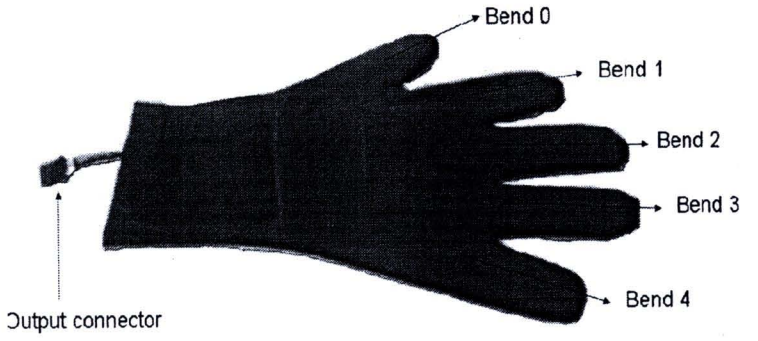
- CyberForce จะใช้ควบคู่กับ CyberGlove และ CyberGrasp โดยจะใช้สวมเข้าไปใน CyberGlove และจับยึดเข้ากับ CyberGrasp เพื่อใช้ในการคำนวณความสัมพันธ์ในการออกแรงของมือและแขนของผู้ใช้ สารารถวัดการออกแรงในลักษณะที่เป็นการหมุน และการย้ายที่ได้ เป็นการเพิ่มความสามารถให้กับแบบจำลองให้มากขึ้นตัวอย่างการใช้งานเช่นการบังคับพวงมาลัยในการเลี้ยวของวงล้อ

X-IST Data glove เป็นถุงมือข้อมูลที่มีประโยชน์ในด้านการควบคุมต่าง ๆ เช่น การควบคุมดนตรี คือสามารถจะใส่ไฟล์ MIDI ลงไปได้ หรือจะใช้ในการควบคุมระบบแสงสีก็ได้ โดยตัวตรวจจับที่ใช้มี 2 ชนิด ได้แก่ Bend sensor และ Pressure fingertip sensor ซึ่งสามารถเลือกทำงานตามที่เหมาะสมได้ โดย Bend sensor ใช้สำหรับตรวจวัดการงอของนิ้วแต่ละนิ้ว และ Pressure fingertip ใช้สำหรับตรวจวัดการสัมผัสและการวัดค่าของแรงกดดันของปลายนิ้ว



ภาพที่ 2.10 การใช้งาน CyberTouch [19]

DG5-VHand Glove เป็นถุงมือข้อมูลอเนกประสงค์ โดยใช้ตัวตรวจจับประเภท Flex sensor (ภาพที่ 2.11) ที่มีความแม่นยำสูงในการวัดถุงมือนี้มีความสามารถมากกว่าถุงมือข้อมูลทั่วไปคือ นอกจากจะวัดการงอของนิ้วแล้วยังสามารถวัดการม้วน การเหวี่ยง และการเอียงการเฉียงของมือได้ ถุงมือนี้เหมาะกับงานหลาย ๆด้าน เช่น การควบคุมกลไก หรือ การใช้งานด้าน Virtual reality ต่าง ๆ DG5-VHand Glove นับเป็นอุปกรณ์ที่มีความคุ้มค่า เพราะนอกจากสามารถนำไปใช้ในด้านต่าง ๆ ยังขนาดที่เล็กบางอีกด้วย อย่างไรก็ตามถุงมือชนิดนี้มีข้อจำกัดด้วยจำนวนตัวตรวจจับ 1 ตัวต่อ 1 นิ้ว ทำให้ไม่สามารถแยกแยะการงอของข้อนิ้วแต่ละข้อ



Technical Characteristics	
Power Supply	From 3.3 V to 5.0 V
Load Current	20 mA
Operating Temperature	From 0 to 50°C
Storage Temperature	From 0 to 70°C
Finger Sensing Resolution	10 bit (1024 step)
Number of Finger Sensor	5 (one per finger)
Hand Orientation Resolution	0.5 degree
Measured Hand Acceleration	From -2g to 2g
Sampling Rate	25 Hz
Glove	One size fits many sizes.
Output Connector	Standard AMP connector, 4 ways (GND, Vcc, TX and RX signals)
Output High (1)	From 2.8 V to 3.3 V
Output Low (0)	From 0.0 V to 3.3 V

ภาพที่ 2.11 รูปแบบของถุงมือ DG5-VHand Glove และคุณลักษณะ [19]

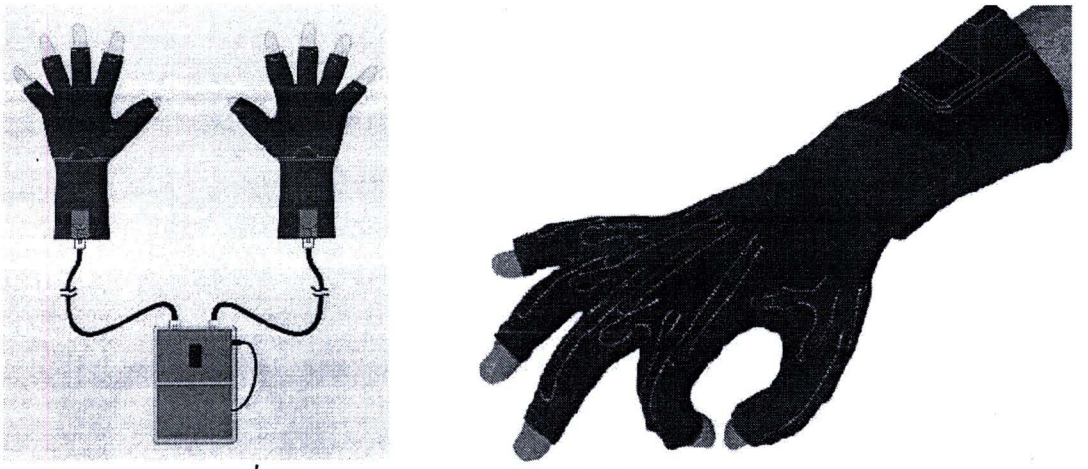
Pinch Glove เป็นถุงมือที่ออกแบบมาเพื่องานสร้าง การจำลองปฏิกิริยา ในงานทางด้าน 3 มิติ เช่นการหยิก การบีบ หรือปฏิกิริยาที่ใช้มือไปกระทำ เพื่อความเสมือนจริง ถุงมือนี้จะอ่านลักษณะได้อย่างหลากหลาย โดยที่ถุงมือจะมีลายแถบสำหรับวัดการยืด หดของถุงมือ ภายในถุงมือจะมีตัวตรวจจับใส่ไว้ที่ปลายนิ้วแต่ละนิ้ว สามารถใช้กับการจำลองขับรถ การบังคับเครื่องบิน การสร้างเกม หรืองานอื่น ๆที่ต้องการการแสดงลักษณะท่าทางที่กว้าง ๆได้

MidiGlove เป็นนวัตกรรมใหม่สำหรับนักดนตรี ใช้ควบคุมเพลงทั้งเพลงโดยใช้แค่เพียงมือเดียวได้ โดยผู้ใช้ต้องสวม MidiGlove แล้วต่อสายใส่เข้าไปใน GloveBox โดย GloveBox นี้เป็น

ตัวจัดการเกี่ยวกับไฟล์ Midi ที่ได้ใส่เข้าไป อุปกรณ์นี้สามารถควบคุมไฟล์ Midi ได้มากที่สุด 12 เสียง อุปกรณ์นี้ใช้ตัวตรวจจับชนิด Bend sensor

P5 Glove อุปกรณ์ชิ้นนี้ออกแบบมาเพื่อเพิ่มศักยภาพในการเล่นเกมโดยเฉพาะ สามารถสร้างการเคลื่อนไหวได้ 6 ระดับคือ X Y Z เอียง เหวี่ยง และ หมุน ตัวตรวจจับที่ใช้เป็น Bend sensor และ Optical tracking technology

5DT Data glove 5 Ultra อำนวยความสะดวกในการสร้างสรรค์ผลงานที่แสดงการเคลื่อนไหวต่าง ๆ ถุงมือมีขนาดที่เล็กกะทัดรัด ตัวตรวจจับที่ใช้เป็น Bend sensor (ภาพที่ 2.12) อุปกรณ์นี้สามารถส่งสัญญาณได้ไกลถึง 20 เมตรและใช้งานต่อเนื่องได้นานถึง 8 ชั่วโมง



ภาพที่ 2.12 รูปแบบของถุงมือ 5DT Data glove 5 Ultra [19]

1.6 การแยกองค์ประกอบของภาพ (Image segmentation)

การแยกองค์ประกอบของภาพเป็นการแยกรายละเอียดต่าง ๆ ในภาพ เช่น วัตถุและฉากหลังออกจากกันเพื่อให้รายละเอียดเหล่านี้อยู่ในรูปโครงสร้างข้อมูลที่เหมาะสมซึ่งจะช่วยลดจำนวนข้อมูลที่ไม่จำเป็นในการวิเคราะห์ภาพลง การแยกองค์ประกอบของภาพอาศัยข้อมูลความสัมพันธ์ระหว่างส่วนประกอบในภาพ เช่น ความเหมือนกัน (Similarity) ของคุณสมบัติของพิกเซลในพื้นที่เดียวกัน หรือความไม่ต่อเนื่องกัน (Discontinuity) ของคุณสมบัติของพิกเซลบริเวณรอยต่อระหว่างวัตถุกับฉากหลัง เพื่อใช้ในการแยกแยะพื้นที่

การแยกองค์ประกอบของภาพมีความจำเป็นสำหรับการพัฒนาถุงมือข้อมูลเชิงแสงในงานวิจัยนี้โดยเป็นส่วนหนึ่งของขั้นตอนวิธีการประมวลผลภาพเชิงดิจิตอลเพื่อคำนวณหาระยะทางการเคลื่อนที่ของเส้นเอ็นที่ใช้ตรวจจับการงอของนิ้ว เทคนิคการแยกองค์ประกอบของภาพแบ่งได้ 3 ประเภทหลัก ๆ คือ

1.6.1 Pixel oriented image segmentation เป็นวิธีการแยกองค์ประกอบของภาพโดยพิจารณาตามความเหมือนกันของคุณสมบัติของพิกเซลเพียงอย่างเดียว วิธีการที่นิยมมากที่สุดได้แก่ วิธีการ Thresholding และวิธีการ Clustering โดยวิธีการ Thresholding เป็นวิธีการแยกวัตถุออกจาก

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

ฉากหลังโดยการเปรียบเทียบค่าความเข้มของแสงของพิกเซลกับค่า Threshold ผลลัพธ์ที่ได้จากวิธีการ Thresholding จะได้ภาพไบนารี (Binary) ที่แสดงถึงพิกเซลที่อยู่ในช่วงความเข้มของแสงช่วงที่กำหนด วิธีการนี้ใช้ได้ดีกับข้อมูลภาพที่ชัดเจนไม่มีสัญญาณรบกวนและมีความแตกต่างของสี (Contrast) มาก ข้อดีของวิธีการนี้คือใช้เวลาในการคำนวณน้อย แต่อย่างไรก็ตามวิธีการ Thresholding ไม่เหมาะกับข้อมูลภาพที่มีสัญญาณรบกวนหรือมีความสว่างไม่คงที่เนื่องจากมักได้ผลลัพธ์ที่มีความผิดพลาดสูง

วิธีการแยกองค์ประกอบของภาพนั้น สามารถหาได้จากสมการ

$$P_{out} = i \qquad where \qquad D_i = \min\{D_R, D_G, D_B\}$$

$$D_R = \sqrt{(r_{in} - r_R)^2 + (g_{in} - g_R)^2 + (b_{in} - b_R)^2}$$

$$D_G = \sqrt{(r_{in} - r_G)^2 + (g_{in} - g_G)^2 + (b_{in} - b_G)^2}$$

$$D_B = \sqrt{(r_{in} - r_B)^2 + (g_{in} - g_B)^2 + (b_{in} - b_B)^2}$$

กำหนดให้ i คือ $\{1(RED),2(Green),3(Blue)\}$

r, g, b คือ ค่าต้นแบบของแต่ละจุดสี

r_{in} คือ ค่าของจุดสีส่วนที่เป็นสีแดงที่ได้รับเข้ามา

g_{in} คือ ค่าของจุดสีส่วนที่เป็นสีเขียวที่ได้รับเข้ามา

b_{in} คือ ค่าของจุดสีส่วนที่เป็นสีน้ำเงินที่ได้รับเข้ามา

โดยหาก D_R, D_G, D_B ตัวใดมีค่าน้อยที่สุดให้ถือว่าจุดสีที่ได้รับค่าเข้ามาเป็นสีนั้น

ส่วนวิธีการ Clustering เป็นวิธีการแบ่งกลุ่มพิกเซลตามค่าสี หรือความเข้มของแสงของพิกเซลที่มีค่าใกล้เคียงกัน โดยการวัดระยะทางใน Feature space การแบ่งกลุ่มด้วยวิธีนี้ได้นำเอาหลักการของปัญญาประดิษฐ์ (Artificial intelligence) มาใช้เพื่อให้สามารถเรียนรู้และปรับตัวเองได้ เช่น การแบ่งกลุ่มโดยใช้โครงข่ายประสาทเทียมด้วยวิธี Self Organizing Map (SOM) และ Fuzzy C-Mean (FCM) วิธีการเหล่านี้มีการกำหนดจำนวนกลุ่มของข้อมูลที่ต้องการจะแบ่งและค่ากลางของแต่ละกลุ่ม จากนั้นจะมีการเปรียบเทียบระยะทางระหว่างค่าสีของพิกเซลที่กำลังพิจารณากับค่ากลางของกลุ่ม แล้วเลือกค่ากลางของกลุ่มที่ใกล้ที่สุดเพื่อแบ่งกลุ่มพิกเซล วิธีการเหล่านี้มีกระบวนการปรับตัวและเรียนรู้ด้วยตัวเองโดยไม่ต้องมีผู้สอน ทุกครั้งที่มีการเปรียบเทียบจะมีการปรับปรุงค่ากลางของแต่ละกลุ่มใหม่จนกระทั่งได้ผลการแบ่งกลุ่มเป็นที่น่าพอใจ

1.6.2 Region oriented image segmentation เป็นเทคนิคการแยกองค์ประกอบของภาพที่ ดีกว่าวิธีการ Pixel oriented image segmentation เพราะวิธีการนี้จะแยกองค์ประกอบของภาพโดยพิจารณาจากความสัมพันธ์ระหว่างพิกเซลภายในภาพและคุณสมบัติของพิกเซลที่เหมือนกัน โดยพิกเซลที่อยู่ติดกันและมีคุณสมบัติเหมือนกันจะถูกจัดให้เข้าในกลุ่มเดียวกัน คุณสมบัติของพิกเซลที่สามารถนำมาพิจารณาได้แก่ ค่าความเข้มของแสง ค่าสี ค่าความแปรปรวน หรือค่าเกรเดียนต์ เป็นต้น ผลลัพธ์ที่ได้จากวิธีการนี้จะเป็นภาพกลุ่มของพิกเซลที่มีพื้นที่ต่อเนื่องกัน

อย่างไรก็ตามวิธีการนี้ใช้เวลานานในการประมวลผลภาพ ตัวอย่างของวิธีการนี้ได้แก่ Region growing และ Region splitting and merging มีรายละเอียดดังนี้

วิธี Region growing เริ่มต้นจากการกำหนดพิกเซลของจุดเริ่มต้น (Seed pixel) จากนั้นจึงทำการรวมกลุ่มพิกเซลข้างเคียงที่มีคุณสมบัติเหมือนกันเข้ากลุ่มเดียวกันทำให้พื้นที่โตขึ้นเรื่อย ๆ ระหว่างการทำงานพื้นที่สามารถขยายไปได้ทุกทิศทางจนบรรจบกับขอบเขตของพื้นที่อื่น กระบวนการทำงานจะวนรอบไปเรื่อย ๆ จนกระทั่งทุกพิกเซลได้ถูกกำหนดให้อยู่ในพื้นที่ใดพื้นที่หนึ่ง วิธีการนี้รับประกันได้ว่าผลลัพธ์ที่ได้จะเป็นพื้นที่ต่อเนื่องและไม่เกิดช่องว่างภายในพื้นที่ อย่างไรก็ตามความยุ่งยากของวิธีการนี้คือ การกำหนดวิธีการเลือกตำแหน่งของพิกเซลเริ่มต้น และการกำหนดเงื่อนไขในการรวมพิกเซลเข้าด้วยกัน และเงื่อนไขในการหยุดขยายตัว นอกจากนั้นวิธีการนี้ยังต้องพิจารณาพิกเซลทีละพิกเซลจนครบทั้งภาพทำให้เสียเวลาในการคำนวณมาก

ส่วนวิธี Region splitting and merging เป็นวิธีการแยกองค์ประกอบของภาพโดยการแบ่งภาพตามด้วยการรวมภาพ โดยเริ่มต้นจากการแบ่งภาพทั้งหมดออกเป็นเซกเมนต์ (Segment) ย่อย ๆ โดยที่เซกเมนต์ใดประกอบด้วยพิกเซลที่มีคุณสมบัติแตกต่างกันจะถูกแบ่งเป็นเซกเมนต์ย่อยลงไปอีกจนกว่าจะได้เซกเมนต์ที่มีพิกเซลที่มีคุณสมบัติเหมือนกัน หลังจากนั้นจึงเป็นการรวมเซกเมนต์ที่มีคุณสมบัติเหมือนกันและอยู่ติดกันเข้าเป็นพื้นที่เดียวกัน วิธีการนี้มีข้อดีคือจะให้ผลลัพธ์เป็นภาพของการแยกองค์ประกอบของภาพทั้งภาพแต่จะมีข้อเสียคือต้องใช้โครงสร้างข้อมูลที่ซับซ้อน และกำหนดเงื่อนไขในการแบ่งและการรวมเซกเมนต์ที่เหมาะสมทำได้ลำบาก

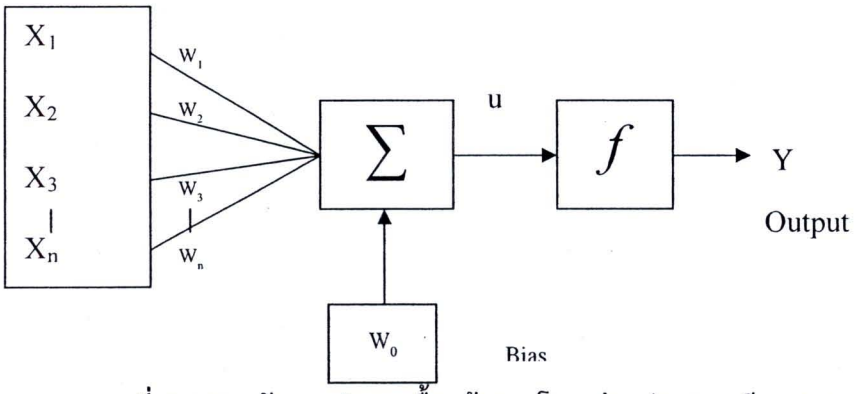
1.6.3 Edge oriented image segmentation เป็นวิธีการแยกองค์ประกอบของภาพโดยอาศัยความไม่ต่อเนื่องกันของคุณสมบัติของพิกเซลบริเวณรอยต่อระหว่างองค์ประกอบในภาพ วิธีการนี้มุ่งเน้นที่ขอบของวัตถุเป็นหลัก ผลลัพธ์ที่ได้จะอยู่ในรูปของขอบเขต (Boundary) ที่แบ่งระหว่างพื้นที่ต่าง ๆ การหาขอบของวัตถุทำได้หลายวิธี เช่น ใช้การเปรียบเทียบระหว่าง Edge model กับภาพจริง หรือเกรเดียนต์ของภาพดังในวิธีการ Spine line หรือใช้วิธีการตรวจหาขอบวัตถุด้วยวิธีการอื่น ๆ เช่น การวัดค่า Threshold ของขนาดภาพเกรเดียนต์ การหาจุดผ่านศูนย์ของภาพลาปลาเซียน โดยที่ตำแหน่งขอบของวัตถุจะสอดคล้องกับตำแหน่งที่เป็นจุดสูงสุดหรือจุดต่ำสุดของภาพเกรเดียนต์และตำแหน่งที่เป็นจุดผ่านศูนย์ของภาพลาปลาเซียน

1.7 โครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

โครงข่ายประสาทเทียมมีประวัติความเป็นมาย้อนหลังไปประมาณ 60 กว่าปีก่อน ในปี ค.ศ. 1943 Mcculoch และ Pitts แห่งมหาวิทยาลัยชิคาโก ประเทศสหรัฐอเมริกา ได้นำเสนอบทความวิชาการ “Boolean brain” ซึ่งได้กลายเป็นจุดกำเนิดของการจัดรูปแบบคณิตศาสตร์ของประสาทเทียม ต่อมาได้มีนักวิจัยได้คิดค้นรูปแบบโครงข่ายประสาทเทียมประกอบกับวิธีการสอนโครงข่ายแบบต่าง ๆ มากมาย ซึ่งวิธีการต่าง ๆ จะมีความซับซ้อนแตกต่างกันไป [14]

โครงข่ายประสาทเทียม เป็นการจำลองการทำงานโครงข่ายประสาทของมนุษย์ (Biological neurons) ซึ่งประกอบด้วยส่วนของการประมวลผลที่เรียกว่านิวรอน (Neuron) ทุกๆ นิวรอนสามารถมีอินพุตได้หลายอินพุตแต่มีเอาต์พุตเพียงเอาต์พุตเดียว และทุกๆ เอาต์พุตจะแยกไปยังอินพุตของนิวรอนอื่นๆ ภายในโครงข่ายการติดต่อกันภายในระหว่างนิวรอนไม่ใช่ลักษณะการต่อแบบธรรมดาทุกๆ อินพุตจะมีน้ำหนักเป็นตัวกำหนดกำลังของการติดต่อภายในและช่วยในการตัดสินใจ การทำงานของนิวรอนในบางโครงข่ายจะถูกกำหนดไว้ตายตัว แต่บางโครงข่ายสามารถที่จะปรับแต่งได้ซึ่งอาจจะเป็นการปรับแต่งจากภายนอกโครงข่ายหรือนิวรอนสามารถปรับได้ด้วยตัวของมันเอง จุดนี้แสดงถึงความสามารถในการเรียนรู้และจดจำของโครงข่ายประสาทเทียม

หลักการทำงานเบื้องต้นของโครงข่ายประสาทเทียมแสดงไว้ในภาพที่ 2.14 โดย Input แต่ละค่าถ่วงน้ำหนัก (Weight) เป็นตัวกำหนดน้ำหนักของ Input โดยนิวรอนแต่ละหน่วยจะมีค่า Threshold เป็นตัวกำหนดว่าน้ำหนักรวมของ Input ต้องมากขนาดไหนจึงจะสามารถส่ง Output ไปยังนิวรอนตัวอื่นได้ เมื่อนำนิวรอนแต่ละหน่วยมาต่อกันให้ทำงานร่วมกันการทำงานนี้ในทางตรรกหลักการนี้หมือนกับปฏิกิริยาเคมีที่เกิดในสมอง เพียงแต่ในคอมพิวเตอร์ทุกอย่างเป็นตัวเลขเท่านั้นเอง ซึ่งสามารถอธิบายได้ว่า ในการทำงานของโครงข่ายอาจจะมี Input หลายตัวที่เป็นตัวกระตุ้นให้เกิด output โดยที่แต่ละ Input จะมีอิทธิพลต่อ Output ต่างกัน ซึ่งวัดได้จากค่าถ่วงน้ำหนัก



ภาพที่ 2.13 หลักการทำงานเบื้องต้นของโครงข่ายประสาทเทียม [4]

จากรูปข้างต้นเป็นตัวอย่างแบบจำลองเซลล์ประสาทที่ McCulloch-Pitts ได้เสนอไว้ โดย $X = [X_1, X_2, ..., X_n]^T$ คือ ข้อมูลนำเข้า Input

$W = [W_1, W_2, ..., W_n]^T$ คือ น้ำหนักความสำคัญที่ให้กับข้อมูล Input แต่ละตัว

ฟังก์ชันของข่ายงาน (Network functions) คือผลรวมทั้งหมดของผลคูณของข้อมูลนำเข้า input กับน้ำหนักความสำคัญที่ให้กับข้อมูล input แต่ละตัวบวกกับ Biased term โดยฟังก์ชัน

ของข่าย (Net functions) อาจจะเป็นความสัมพันธ์จาก Input ไปเป็น Output แบบเป็นเส้นตรงหรือไม่ใช่เส้นตรงก็ได้ ในที่นี้เราเรียกว่า Hyperplane

$$u = \sum_{j=1}^{N} w_{ij} x_j + \theta_i = w_i^T x + \theta_i = \begin{bmatrix} \theta_i & w_i^T \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}$$

โดยทั่วไปโครงสร้างของโครงข่ายประสาทเทียมนั้นประกอบไปด้วย 3 ส่วนดังนี้

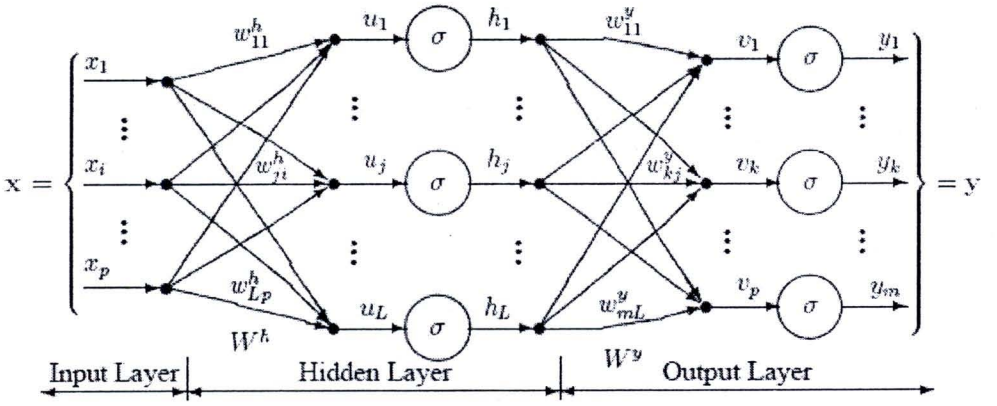
(1) สถาปัตยกรรมของโครงข่าย (Network architecture)

(2) การปรับค่าน้ำหนัก (Adjusting weight)

(3) ฟังก์ชันกระตุ้น (Activation function)

1.7.1 สถาปัตยกรรมของโครงข่าย (Network architecture)

โดยทั่วไปโครงข่ายประสาทเทียมประกอบไปด้วยชั้น (Layer) ของโครงข่าย 3 ชั้นหลัก คือ ชั้นรับข้อมูลเข้า (Inpput layer) ชั้นเซลล์ประสาทซ่อน (Hidden layer) และชั้นข้อมูลออก (Output layer) ดังแสดงในภาพที่ 2.15 ซึ่งสถาปัตยกรรมของโครงข่ายประสาทเทียมมีหลายประเภทที่น่าสนใจ เช่น แบบป้อนไปข้างหน้า (Feed-forward) แบบย้อนกลับ (Feedback) แบบเชื่อมกันอย่างสมบูรณ์ (Fully interconnected net) และแบบแข่งขัน (Competitive net)



ภาพที่ 2.14 การเชื่อมโยงกันระหว่างโหนดของระบบโครงข่ายประสาทเทียม [4]

1.7.2 การปรับค่าน้ำหนัก (Adjusting weight)

การประมวลผลของโครงข่ายประสาทเทียมนั้น ยินยอมให้มีกระบวนการปรับค่าน้ำหนักในระหว่างการเรียนรู้ของระบบ เพื่อให้โครงข่ายสามารถเรียนรู้พฤติกรรมของข้อมูลใช้ฝึกสอน (Training data) จนกระทั่งบรรลุวัตถุประสงค์ของการเรียนรู้ ซึ่งส่วนใหญ่การเรียนรู้สำหรับตัว Neural network จะแบ่งออกเป็น 2 แบบด้วยกัน คือ

(1) การเรียนรู้แบบมีการสอน (Supervised lerning) เป็นการเรียนแบบที่มีการตรวจคำตอบเพื่อให้วงจรข่ายปรับตัว ชุดข้อมูลที่ใช้สอนวงจรข่ายจะมีคำตอบไว้คอยตรวจดูว่า

วงจรข่ายให้คำตอบที่ถูกหรือไม่ ถ้าตอบไม่ถูก วงจรข่ายก็จะปรับตัวเองเพื่อให้ได้คำตอบที่ดีขึ้น (เปรียบเทียบกับคน เหมือนกับการสอนนักเรียนโดยมีครูผู้สอนคอยแนะนำ) ตัวอย่างชนิดของโครงข่ายประสาทเทียมที่มีการเรียนรู้แบบมีผู้สอน ได้แก่ Perceptron, Learning vector quantization, Back-propagation, Boltzmann machine

(2) การเรียนรู้แบบไม่มีการสอน (Unsupervised lerning) เป็นการเรียนแบบไม่มีผู้แนะนำ ไม่มีการตรวจคำตอบว่าถูกหรือผิด วงจรข่ายจะจัดเรียงโครงสร้างด้วยตัวเองตามลักษณะของข้อมูล ผลลัพธ์ที่ได้ วงจรข่ายจะสามารถจัดหมวดหมู่ของข้อมูลได้ (เปรียบเทียบกับคน เช่น การที่เราสามารถแยกแยะพันธุ์พืช พันธุ์สัตว์ตามลักษณะรูปร่างของมันได้เองโดยไม่มีใครสอน) ตัวอย่างชนิดของโครงข่ายประสาทเทียมที่มีการเรียนรู้แบบไม่มีผู้สอน เช่น Self-organization map, Competitive learning, Hopefield, Counter propagation

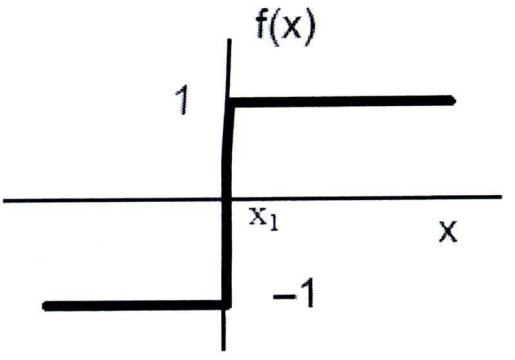
1.7.3 ฟังก์ชันกระตุ้น (Activation function)

ฟังก์ชันกระตุ้นเป็นฟังก์ชันทางคณิตศาสตร์ที่เป็นเส้นที่รูปร่างที่เป็นเส้นเดี่ยวและสามารถครอบคลุมการจำแนกของโครงข่ายประสาทเทียมได้ทั้งหมด ซึ่งการเลือกใช้ฟังก์ชันกระตุ้นนั้นขึ้นอยู่กับประเภทของข้อมูลที่ต้องการจะจำแนก ซึ่งข้อมูลในเชิงจำนวนนั้นจะแบ่งได้เป็น 2 ประเภท

1. ข้อมูลแบบไม่ต่อเนื่อง (Discrete data) เป็นลักษณะของข้อมูลที่เป็นข้อมูลเชิงคุณภาพ (Quality data) ซึ่งไม่สามารถนำมาระบุเป็นค่าตัวเลขได้เช่น สีผม ความชอบ คุณลักษณะ ซึ่งข้อมูลที่จำแนกให้เป็นข้อมูลที่ไม่ต่อเนื่องนั้นมักจะใช้ฟังก์ชันทำนบ หรือที่เรียกว่า Threshold function

$$f(x)\begin{cases} 1 \;; x \geq a \\ -1; x < a \end{cases} \qquad a \text{ คือค่าคงที่ใดๆ}$$

ลักษณะที่จะจำแนกได้จะแสดงได้โดยกราฟนี้



ภาพที่ 2.15 กราฟของฟังก์ชัน Threshold

2. ข้อมูลแบบต่อเนื่อง (Continuous data) เป็นลักษณะของข้อมูลที่เป็นข้อมูลเชิงปริมาณ (Quantity data) ซึ่งสามารถนำมาระบุเป็นค่าตัวเลขได้เช่น ความสูง น้ำหนัก อุณหภูมิ อายุ ซึ่งข้อมูลที่จำแนกให้เป็นข้อมูลที่ต่อเนื่องนั้นมักจะใช้ฟังก์ชันหลากหลาย ซึ่งจะครอบคลุมการจำแนกได้ ซึ่งฟังก์ชันที่เป็นที่นิยมคือ

1) ฟังก์ชันเชิงเส้น (Linear function) เป็นฟังก์ชันที่มีลักษณะแบ่งข้อมูลออกเป็นสองส่วนแบบแนวระนาบ ข้อมูลที่ถูกแบ่งนั้นจะได้เป็นส่วนพื้นที่ที่เท่ากัน

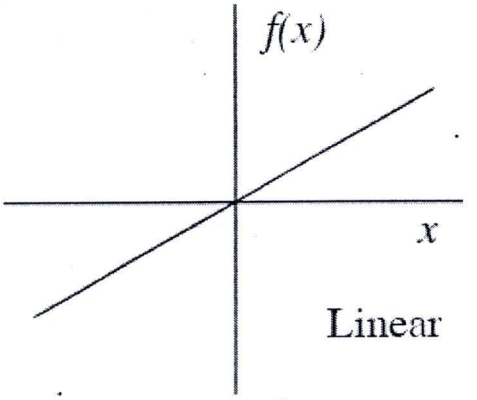
$$f(x) = mx + c$$

เมื่อ m คือความชัน

c คือค่าคงที่ใด ๆ

อนุพันธ์ของ $f(x)$ คำนวณได้จาก $\frac{d}{dx} f(x) = m$

ลักษณะที่จะจำแนกได้จะแสดงได้โดยกราฟนี้



ภาพที่ 2.16 กราฟของฟังก์ชันเส้นตรง (Linear function)

2) ฟังก์ชัน Gaussian เป็นฟังก์ชันที่มีลักษณะแบ่งข้อมูลออกเป็นสองส่วนที่มีการกระจายตัวเป็นการแจกแจงแบบปกติ

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

เมื่อ μ คือค่าเฉลี่ยของประชากร

σ คือส่วนเบี่ยงเบนมาตรฐานของประชากร

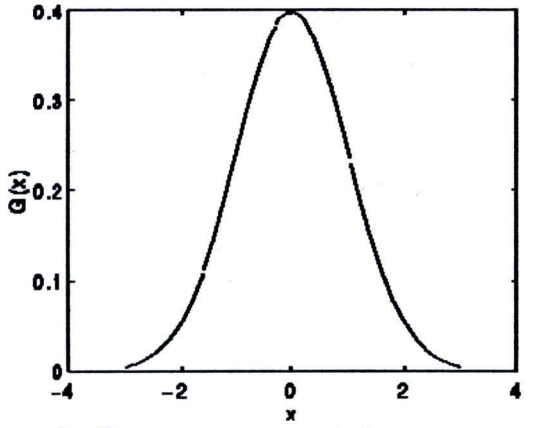
e คือค่าคงที่ของเนเปียร์ ประมาณ 2.7183

π คือค่าคงที่ประมาณ 3.14159

อนุพันธ์ของ $f(x)$ คำนวณได้จาก

$$\frac{d}{dx}f(x) = -\frac{x-\mu}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = -\left(\frac{x-\mu}{\sigma}\right)f(x)$$

ลักษณะที่จะจำแนกได้จะแสดงได้โดยกราฟนี้



ภาพที่ 2.17 กราฟของฟังก์ชันการแจกแจงแบบปกติ (Normal distribution function)

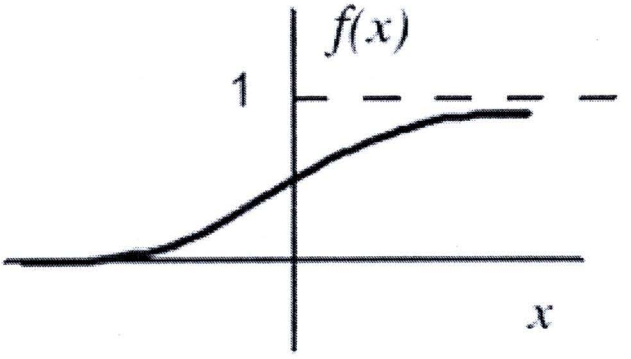
3) ฟังก์ชันซิกมอย (Sigmoid function) เป็นฟังก์ชันที่มีลักษณะการกระจายตัวเป็นเส้นโค้งตามอัตราการเจริญเติบโต (λ) ซึ่งค่าของข้อมูลนั้นจะมีค่าที่จะได้รับจากฟังก์ชันนี้นั้นอยู่ระหว่าง $(0,1)$ เท่านั้น

$$f(x) = \frac{1}{1+e^{-\lambda x}}$$

เมื่อ λ คือค่าสัมประสิทธิการเติบโตของฟังก์ชัน

อนุพันธ์ของ $f(x)$ คำนวณได้จาก $\frac{d}{dx}f(x) = -\lambda(1-f(x))f(x)$

ลักษณะที่จะจำแนกได้จะแสดงได้โดยกราฟนี้



ภาพที่ 2.18 กราฟของฟังก์ชันซิกมอย (Sigmoid function)

4) ฟังก์ชันแทนไฮเปอร์โบลิค (Tangent hyperbolic function) เป็นฟังก์ชันที่มีลักษณะการกระจายตัวเป็นเส้นโค้งตามอัตราการเจริญเติบโต (T) ซึ่งค่าของข้อมูลนั้นจะมีค่าที่จะได้รับจากฟังก์ชันนี้นั้นอยู่ระหว่าง $(-1,1)$ เท่านั้น

$$f(x) = \frac{\sinh(x)}{\cosh(x)}$$

$$\sinh(x) = \frac{e^{Tx} - e^{-TX}}{2}$$

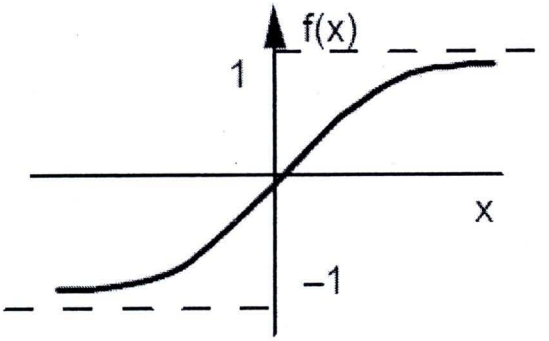
$$\cosh(x) = \frac{e^{Tx} + e^{-TX}}{2}$$

$$f(x) = \frac{e^{Tx} - e^{-TX}}{e^{Tx} + e^{-TX}}$$

เมื่อ T คือค่าสัมประสิทธิการเติบโตของฟังก์ชัน

อนุพันธ์ของ $f(x)$ คำนวณได้จาก $\frac{d}{dx} f(x) = \frac{1}{T}(1 - f^2(x))$

ลักษณะที่จะจำแนกได้จะแสดงได้โดยกราฟนี้



ภาพที่ 2.19 กราฟของฟังก์ชันแทนไฮเปอร์โบลิค (Tangent hyperbolic function)

1.8 ขั้นตอนวิธีการส่งค่าย้อนกลับ (Back-propagation algorithm)

1. ป้อนข้อมูลเข้าไปในโครงข่าย $(x_1, x_2, x_3, ..., x_N)$ พร้อมทั้งทำการกำหนดค่าที่จำแนก $(Tar_1, Tar_2, Tar_3, ..., Tar_N)$ และกำหนดค่าถ่วงน้ำหนักที่เชื่อมต่อในแต่ละเครือข่ายให้เป็นตัวเลขที่ต่ำ ๆ

$$\{w_{1j}, w_{2j}, w_{3j}, ..., w_{Nj}\} \in (0,1]$$

2. คำนวณค่าผ่านทางฟังก์ชันผลรวม ในระดับชั้นนั้น และหาค่าผ่านทางฟังก์ชันกระตุ้น

$$Input_i = \sum_{j=1}^{N} w_j x_j - \mu_j \quad \text{หรือ} \quad Input_i = \prod_{j=1}^{N} w_j x_j - \mu_j$$

$$Output_i = f(Input_i) \qquad Output_i = f(Input_i)$$

3. คำนวณค่าที่ได้จากการผ่านฟังก์ชั้นกระตุ้นผ่านทางฟังก์ชันผลรวม ในระดับชั้นที่ผ่านมากับค่าน้ำหนักในชั้นนั้น และหาค่าผ่านทางฟังก์ชันกระตุ้น

$$\varepsilon = Tar - Net$$

4. คำนวณค่าความผิดพลาด (Error, ε) ในชั้นข้อมูลขาออกที่ได้จากการคำนวณในทุกระดับชั้นแล้ว แล้วทำการปรับปรุงค่าถ่วงน้ำหนักในแต่ละระดับชั้น ซึ่งมีวิธีการที่มีความนิยมดังนี้

1) Gradient descent method

2) Gradient descent with momentum method

3) Newton's method

4) Levenberg-marquardt back-propagation method

5. ทำซ้ำในขั้นตอนที่ 2 – 4 จนกว่าค่าความผิดพลาดที่ได้มาน้อยกว่าค่าที่กำหนดไว้ $|\varepsilon| < Threshold$

วิธีการสำหรับการปรับปรุงค่าน้ำหนักของ Multi layer perceptron จะสามารถทำได้จากสูตรนี้

$$w_i(New) = w_i(Old) + \Delta w_i$$
$$\mu_i(New) = \mu_i(Old) + \Delta \mu_i$$
$$\varepsilon_i = Tar_i - f(Input_i)$$

โดยวิธีการที่เป็นที่นิยมนั้นมีอยู่ 4 แบบดังนี้

1. Gradient descent method

วิธีการนี้ปรับปรุงค่าน้ำหนักตามทิศทางตรงข้ามกับทิศทางของ Gradient ของ Error surface เมื่อเทียบกับค่าน้ำหนัก ค่าสำหรับปรับปรุงค่าน้ำหนัก Δw_i คำนวณได้จากสูตร

$$\Delta w_i = \alpha \frac{\partial}{\partial w_i} \varepsilon^2 \qquad \text{เมื่อ } \alpha \text{ คืออัตราการเรียนรู้ของเครื่องจักรมีค่า } [0,1]$$

$$\Delta w_i = \alpha \frac{\partial}{\partial w_i} \left(Tar_i - f(Input_i)\right)^2$$

$$\Delta w_i = \alpha \frac{\partial}{\partial w_i} \left(Tar_i - f(\sum_{i=1}^{N} w_i x_i) \right)^2$$

$$\Delta w_i = 2\alpha \left(Tar_i - f(\sum_{i=1}^{N} w_i x_i) \right) \frac{\partial}{\partial w_i} \left(Tar_i - f(\sum_{i=1}^{N} w_i x_i) \right)$$

$$\Delta w_i = 2\alpha \left(Tar_i - f(\sum_{i=1}^{N} w_i x_i) \right) \left(\frac{\partial}{\partial w_i} Tar_i - \frac{\partial}{\partial w_i} f(\sum_{i=1}^{N} w_i x_i) \right)$$

$$\Delta w_i = 2\alpha\left(Tar_i - f(\sum_{i=1}^{N} w_i x_i)\right)\left(-f'(\sum_{i=1}^{N} w_i x_i)\frac{\partial}{\partial w_i}\sum_{i=1}^{N} w_i x_i\right)$$

$$\Delta w_i = -2\alpha\left(Tar_i - f(\sum_{i=1}^{N} w_i x_i)\right)\left(f'(\sum_{i=1}^{N} w_i x_i)x_i\right)$$

$$\Delta w_i = -2\alpha\left(Tar_i - f(\sum_{i=1}^{N} w_i x_i)\right)f'(\sum_{i=1}^{N} w_i x_i)x_i$$

$$\Delta w_i = -2\alpha\varepsilon_i f'(Input_i)x_i$$

ค่าน้ำหนักค่าใหม่คำนวณได้จาก $w_i^{new} = w_i^{old} + \Delta w_i$

ในทำนองเดียวกันนี้สามารถนำไปใช้กับการปรับค่าน้ำหนักของ Bias ได้ซึ่งจะได้ดังนี้

$$\Delta\mu_i = \alpha\frac{\partial}{\partial\mu_i}\varepsilon_i^2 = -2\alpha\varepsilon_i f'(Input_i)$$

2. Gradient descent with momentum method

วิธีการนี้เป็นการนำหลักการของ Gradient descent มาต่อเติมโดยการคำนวณค่าความผิดพลาดเพิ่มเติมที่เรียกว่า Momentum ซึ่งเป็นการเพิ่มนิพจน์เข้าไปจากเดิม

$$\Delta w_i = \alpha\frac{\partial}{\partial w_i}\varepsilon_i^2 + \beta w(Old)$$

โดยที่ $0 < \beta < 1$

$$\Delta\mu_i = \alpha\frac{\partial}{\partial\mu_i}\varepsilon_i^2 = -2\alpha\varepsilon_i f'(Input_i)$$

3. Newton's method

การประยุกต์ใช้จากอนุกรม Taylor มาใช้สำหรับการปรับปรุงค่าถ่วงน้ำหนักหากเกิดความผิดพลาดกังนี้

$$f(x) = \sum_{n=0}^{\infty}\frac{f^n(a)(x-a)^n}{n!}$$

จะทำให้ได้สมการของการปรับค่าน้ำหนักดังนี้

$$E(w) = E(w_0) + (w - w_0)\nabla E(w_o) + \frac{1}{2}H(w - w_0)^2 + \ldots.$$

โดยที่ H คือ Hessian matrix;

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1^2} \end{bmatrix}$$

E คือ Error function (ε)

จะได้ว่า

$$\Delta w_i = H_{ij}\varepsilon + K$$

$$H_{ij} = \frac{\partial}{\partial w_i}\frac{\partial}{\partial w_j}E$$

4. Levenberg-Marquardt back-propagation method

$$JJ = J(w) \times J(w)$$

$$JE = J(w) \times E$$

$$\Delta w = -(JJ + \mu I) \cdot JE^{-1}$$

โดยที่ J คือ Jacobian matrix;

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

E คือ Error function (ε)

I คือ Identity matrix;

$$I = \begin{bmatrix} 1 & . & . & 0 & 0 \\ . & 1 & . & . & 0 \\ . & . & . & . & . \\ 0 & . & . & . & . \\ 0 & 0 & . & . & 1 \end{bmatrix}$$

2. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องนี้จะแบ่งออกเป็น 2 กลุ่มตามหัวข้อต่อไปนี้ 1) งานวิจัยที่เกี่ยวข้องกับการถ่ายทอดและวิเคราะห์รูปแบบท่าทางของมือ และ 2) งานวิจัยด้านการแปลความหมายระหว่างรูปแบบท่าทางมือกับสัญลักษณ์ที่เชื่อมโยงกัน

2.1 งานวิจัยที่เกี่ยวข้องกับการถ่ายทอดและวิเคราะห์รูปแบบท่าทางมือ

2.1.1 Hand Gesture Modeling Analysis and Synthesis [29]

งานวิจัยกล่าวว่า การแสดงท่าทางต่าง ๆ ของมือเป็นรูปแบบหนึ่งของการสื่อสารของมนุษย์ ในขณะที่เราพบข้อจำกัดในการติดต่อกับคอมพิวเตอร์ในการป้อนข้อมูลบางประเภท การนำเอาท่าทางต่าง ๆ ของมือมาใช้ในด้านการติดต่อกับคอมพิวเตอร์จึงเป็นเรื่องที่น่าสนใจ อุปกรณ์ในรูปของถุงมือได้ถูกพัฒนาขึ้นเพื่อถ่ายทอดท่าทางการเคลื่อนไหวของนิ้วมือ แล้วนำข้อมูลที่ได้มาวิเคราะห์ประมวลผล และแสดงผลในโลกเสมือนจริง ในอนาคตการนำรูปแบบและการแสดงออก

สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

ต่าง ๆของมือมาใช้ในด้านการติดต่อกับคอมพิวเตอร์จะมีประสิทธิภาพและแพร่หลายมากขึ้น นอกจากนี้ในงานวิจัยนี้ยังได้มีการแบ่งรูปแบบของวิธีการวิเคราะห์การแสดงท่าทางต่าง ๆ ของมือออกเป็น 4 ประเภท คือ

1) วิธีการประเภทที่ใช้ถุงมือ (Glove-Based Analysis)

การนำถุงมือมาใช้ในการวิเคราะห์การเคลื่อนไหวของนิ้วมือนั้นได้เริ่มขึ้นเมื่อประมาณปี 1970 โดยอุปกรณ์ประกอบด้วยถุงมือที่มี Sensor ติดตั้งอยู่ และทำหน้าที่เปลี่ยนการเคลื่อนไหวของนิ้วมือมาอยู่ในรูปสัญญาณทางไฟฟ้า เพื่อนำไปประมวลเป็นท่าทางของมือ

2) วิธีการประเภทที่ใช้การจับภาพของมือ (Vision-Based Analysis)

การใช้การจับภาพมือมาวิเคราะห์การเคลื่อนไหวของนิ้วมือเป็นวิธีที่เป็นธรรมชาติที่สุดที่ใช้งานในด้านการถ่ายทอดข้อมูลให้กับคอมพิวเตอร์ โดยมีการนำกล้องตั้งแต่หนึ่งตัวขึ้นไป มาจับภาพของมือในขณะนั้น จากนั้นภาพที่ ได้จะถูกนำไปแยกแยะภาพมือออกจากฉากหลังและนำไปเปรียบเทียบกับรูปแบบของมือที่เก็บไว้ในฐานข้อมูล ในกรณีที่ใช้กล้องมากกว่า 1 ตัว จะได้ภาพมือในหลายมุมมองทำให้การตีความหมายรูปแบบท่าทางของมือแม่นยำมากขึ้น

3) วิธีการประเภทที่วิเคราะห์รูปแบบท่าทางของมือจากการวาดเขียน (Analysis of Drawing Gestures)

การวิเคราะห์รูปแบบท่าทางของมือจากการวาดเขียนเป็นการตรวจจับการเคลื่อนไหวของนิ้วมือ โดยดูจากข้อมูลการวาดเขียนที่ได้จากการลาก (Drag) หรือกด (Click) เมาส์หรือปากกา (Stylus) แล้วนำข้อมูลที่ได้ไปแปลความหมายแบบอัตโนมัติ

4) วิธีการประเภทอื่น ๆ (Other Hand Gesture Analysis Techniques)

นอกเหนือจากวิธีการหลักทั้ง 3 แบบที่กล่าวมาข้างต้นแล้ว ยังมีวิธีการอื่น ๆที่ใช้ถ่ายทอดและวิเคราะห์รูปแบบท่าทางของมือด้วยรูปแบบพิเศษ เช่น การรับข้อมูลรูปแบบท่าทางของมือโดยใช้เครื่อง Electromyogram ซึ่งใช้ทั้งการจับภาพมือและการใช้ถุงมือพิเศษไว้ด้วยกัน

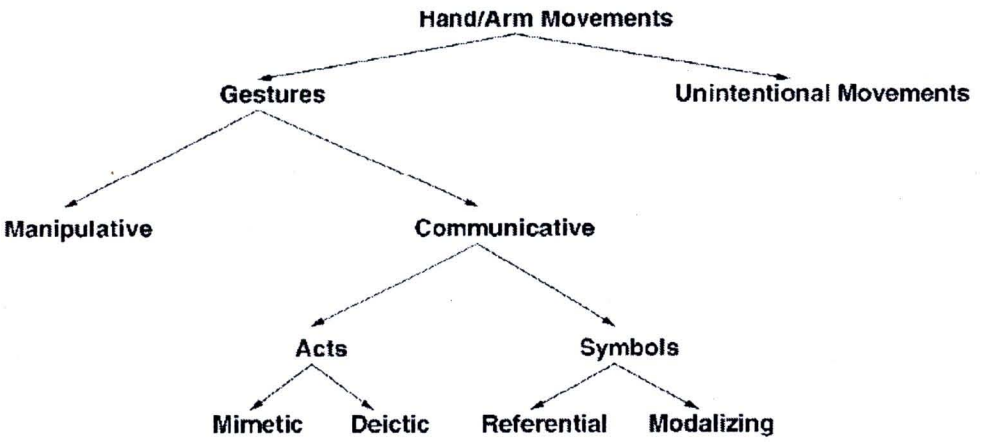
2.1.2 Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review [31]

งานวิจัยกล่าวว่า การใช้ท่าทางของมือในการสื่อสารกับอุปกรณ์รับข้อมูลของคอมพิวเตอร์ เป็นการเพิ่มทางเลือกในการรับข้อมูลสำหรับการทำงานด้าน Human-Computer Interaction (HCI) ในปัจจุบันการจำลองการเคลื่อนไหวท่าทางของมือ ถูกนำไปใช้ด้านนี้อย่างแพร่หลายเนื่องจากความง่ายในการใช้งานและแปลความหมายท่าทางการเคลื่อนไหวของมือ นอกจากนี้ในงานวิจัยนี้ยังได้ศึกษาถึงหลักการการจำลองท่าทางการเคลื่อนไหวของมือ และลักษณะของการนำไปใช้งาน ดังนี้

1) การจำลองรูปแบบการเคลื่อนไหวของมือ

การจำลองรูปแบบการเคลื่อนไหวของมือเป็นการศึกษารูปแบบของท่าทางมือทั้งหมด ดังภาพที่ 2.23 ไม่ว่าจะมีความหมายหรือไม่โดยเริ่มต้นจากการเคลื่อนไหวของมือหรือ

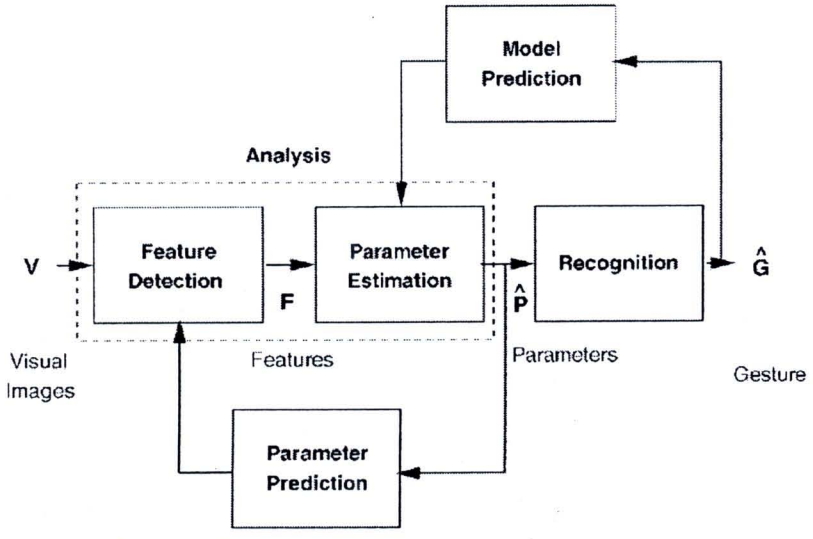
แขน จากนั้นนำไปตีความหมายว่าเป็นการเคลื่อนไหวตามสัญลักษณ์ที่ตกลงกันไว้หรือไม่ ถ้าตรงการเคลื่อนไหวนั้นเป็นรูปแบบใดระหว่างการควบคุมอุปกรณ์หรือการสื่อสาร ถ้าเป็นการสื่อสาร จะเป็นการสื่อสารแบบใดระหว่างการแสดงท่าทางที่แทนสัญลักษณ์หรือแสดงท่าทางที่เป็นภาษากาย ถ้าเป็นภาษากายท่าทางนั้นเป็นการจำลองท่าทางหรือทำท่าทางนั้นจริง ๆ ถ้าเป็นแบบแทนสัญลักษณ์ สัญลักษณ์นั้นแปลความหมายได้เลยหรือไม่หรือต้องนำไปประกอบกับถ้าท่าทางอื่น ๆ ต่อไป



ภาพที่ 2.20 การจำลองรูปแบบการเคลื่อนไหวของมือ [31]

2) การวิเคราะห์การเคลื่อนไหวของมือ

การวิเคราะห์การเคลื่อนไหวของมือเป็นส่วนที่ใช้ทำการจำแนกการเคลื่อนไหวของมือ จากอินพุตที่รับเข้ามาจากนั้นนำไปประมวลผลให้ได้คุณสมบัติของข้อมูลตรงกับข้อมูลที่สามารถนำไปใช้แทนท่าทางการเคลื่อนไหวของมือได้ ดังภาพที่ 2.24 เพื่อที่จะได้นำไปแปลความหมายของท่าทางนั้นต่อไป



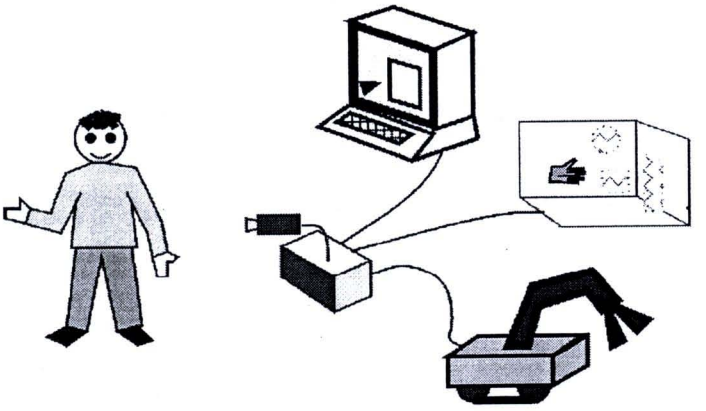
ภาพที่ 2.21 ขั้นตอนการวิเคราะห์การเคลื่อนไหวของมือ [31]

3) การรู้จำท่าทางการเคลื่อนไหวของมือ

การรู้จำท่าทางการเคลื่อนไหวของมือเป็นขั้นตอนที่นำข้อมูลที่ได้จากการวิเคราะห์มาหาความสัมพันธ์กับชุดข้อมูลที่ได้เก็บไว้ว่าตรงกับข้อมูลตัวใด จากนั้นทำการการแสดงผลการทำงานด้วยการนำเสนอข้อมูลที่ทำการเก็บไว้ ส่วนการตรวจสอบความถูกต้องนั้นจำเป็นต้องใช้การตรวจดูข้อมูลที่ได้จากการแปลข้อมูลกลับไปเป็นของข้อมูลเริ่มต้น แล้วทำการเปรียบเทียบกับข้อมูลที่รับเข้ามา

4) การนำท่าทางการเคลื่อนไหวของมือไปใช้งาน

การนำท่าทางการเคลื่อนไหวของมือไปใช้งาน จะแบ่งออกเป็นการนำเสนอตามการจำลองรูปแบบการเคลื่อนไหวของมือในข้อที่ 1 ในภาพที่ 2.25 เป็นการแสดงตัวอย่างการใช้งานด้านการสั่งงานอุปกรณ์ผ่านรูปแบบการเคลื่อนไหวของมือ



ภาพที่ 2.22 การนำไปประยุกต์ใช้งาน [31]

2.2 งานวิจัยที่เกี่ยวข้องกับการเชื่อมโยงระหว่างท่าทางมือกับสัญลักษณ์

2.2.1 Progress in Automated Computer Recognition of Sign Language [16]

ศึกษาถึงโครงสร้างการทำงานของระบบการรู้จำภาษามือ โดยโครงสร้างระบบจะแบ่งออกเป็น 3 ส่วน คือ ส่วนรับข้อมูล ส่วนที่ใช้ตัดสินใจ และส่วนของการเก็บข้อมูล

2.2.2 Linguistic Properties Based on American Sign Language Recognition with Artificial Neural Networks Using a Sensory Glove and Motion Tracker [17]

ศึกษาเรื่องการแปลความหมายรูปแบบท่าทางมือ และการเคลื่อนไหวของแขนเป็นคำในภาษาอังกฤษ โดยใช้ถุงมือข้อมูล และอุปกรณ์ตรวจจับความเคลื่อนไหว (Cyber Glove and Flock of Birds) ที่มีจำหน่ายในท้องตลาดมาประมวลผลร่วมกับการใช้โครงข่ายประสาทเทียม

2.2.3 Recognition of Finger Spelling of American Sign Language with Artificial Neural Network Using Position/Orientation Sensors and Data Glove Recognition of Finger Spelling [18]

ศึกษาเรื่องการแปลท่าทางมือเป็นตัวอักษรในภาษาอังกฤษ โดยใช้ถุงมือรับข้อมูลที่มีจำหน่ายในท้องตลาดมาประมวลผลร่วมกับการใช้โครงข่ายประสาทเทียม ซึ่งผลที่ได้ออกมาในระดับน่าพอใจ เมื่อให้จำนวนตัวอย่างในการฝึกสอนโครงข่ายประสาทเทียมมากกว่า 15 ตัวอย่างต่อ 1 สัญลักษณ์

2.2.4 การติดตามท่าทางของมือจากข้อมูลกลศาสตร์ 3 มิติ ที่ได้จากภาพถ่าย 2 มิติ [1]

เสนอระบบติดตามท่าทางของมือจากข้อมูลกลศาสตร์ 3 มิติ ที่ได้จากภาพถ่าย 2 มิติ การทำงานอาศัยหลักการพื้นฐานจากถุงมือ และการจำลองแบบมาประยุกต์ร่วมกันโดยใช้ภาพถ่ายฝ่ามือข้างขวาซึ่งสวมถุงมือที่ทำสัญลักษณ์ด้วยสีบริเวณปลายนิ้ว ส่วนการออกแบบแบบจำลองมือได้วิเคราะห์จากโครงสร้างมือมนุษย์ โดยแบบจำลองมือมีจำนวนมุมหมุนอิสระ (Degree of Freedom: DOF) จำนวน 21 DOF ซึ่งแบบจำลองมือได้ถูกสร้างภายใต้โปรแกรมจำลองหุ่นยนต์ 3 มิติ โรโบซิม 2 (Robosim 2) ส่วนการควบคุมการขยับปลายนิ้วเข้าสู่ตำแหน่งที่ต้องการได้ใช้ระบบควบคุมแบบฟัซซี่ ร่วมกับเงื่อนไขหรือข้อกำหนดในการขยับมุมหมุนที่ข้อต่อของนิ้วเขียนอัลกอลิทึมในการควบคุมภายใต้โปรแกรม แม็ทแลป (Matlab)

จากการทดสอบได้ใช้สัญญาณภาษามืออเมริกัน (American Sign Language: ASL) ของพยัญชนะต่างๆ จำนวนทั้งสิ้น 24 ท่าทาง สามารถแสดงสัญญาณภาษามืออเมริกันของพยัญชนะต่างๆ ได้จำนวนทั้งสิ้น 18 ท่าทาง ได้แก่พยัญชนะ A B C D E F G H I K L O P Q U W X และ Y โดยจำนวนรอบในการแสดงท่าทางโดยเฉลี่ยประมาณ 17 รอบ และกำหนดให้ค่าผิดพลาดในการลู่เข้าสู่ตำแหน่งในแนวแกน X และ Y ของปลายนิ้ว ไม่เกินกว่า 4 พิกเซล