

โครงร่างวิทยานิพนธ์ (THESIS PROPOSAL)

ชื่อเรื่อง (ภาษาไทย)	การวางแผนการจัดวางปลายนิ้วสำหรับการจับวัตถุที่ไม่รู้จัก
ชื่อเรื่อง (ภาษาอังกฤษ)	Fingertip Placement Planning for Unknown Object Grasping
เสนอโดย	นายณชนนท์ วงษ์วิไล
รหัสนิสิต	5470185721
หลักสูตร	วิศวกรรมศาสตรมหาบัณฑิต
ภาควิชา	วิศวกรรมคอมพิวเตอร์
คณะ	วิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
สถานที่ติดต่อ	ภาควิชาคอมพิวเตอร์ วิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
โทรศัพท์	08-1363-1249
อีเมล	Natchanon.W@Student.chula.ac.th
อาจารย์ที่ปรึกษา	อ.ดร.นัทธี นิภานันท์
คำสำคัญ (ภาษาไทย)	การจัดวางปลายนิ้ว, การจับ, วัตถุที่ไม่รู้จัก
คำสำคัญ (ภาษาอังกฤษ)	Fingertip Placement, Grasping, Unknown Object

การวางแผนการจับวางปลายนิ้วสำหรับการจับวัตถุที่ไม่รู้จัก

(Fingertip Placement Planning for Unknown Object Grasping)

นายณัชนนท์ วงษ์วิไล

19 กันยายน 2555

1 ที่มาและความสำคัญของปัญหา

การใช้งานหุ่นยนต์ทั้งในด้านอุตสาหกรรม งานบริการ และการใช้งานในบ้าน จำเป็นที่ต้องมีความสามารถ หลายๆอย่างที่จะช่วยให้สามารถทำงานได้บรรลุวัตถุประสงค์ หนึ่งในความสามารถที่สำคัญคือ ความสามารถในการหยิบจับสิ่งของ เมื่อศึกษารายละเอียดเกี่ยวกับการจับ ในแง่ของการใช้งานจริง สภาพแวดล้อมของการทำงาน มักจะเป็นสภาพแวดล้อมที่ไม่รู้จัก กล่าวคือ หุ่นยนต์ไม่รู้จักตำแหน่ง รูปร่าง และคุณสมบัติของวัตถุที่จะจับ ถือเป็นปัญหาที่สำคัญที่รู้จักกันในชื่อ “ปัญหาการจับวัตถุที่ไม่รู้จัก” ซึ่งการพัฒนาให้หุ่นยนต์สามารถกระทำการจับวัตถุที่ไม่รู้จักได้ ยังคงเป็นปัญหาที่ท้าทาย และมีหลากหลายงานวิจัยที่พยายามพัฒนางานในส่วนนี้ให้ดียิ่งขึ้น

ในปัญหาการจับวัตถุของหุ่นยนต์ ได้มีการศึกษาและพัฒนาในแง่มุมที่หลากหลาย เริ่มตั้งแต่การออกแบบแขนและมือหุ่นยนต์ที่เหมาะสมกับการจับ การวางแผนการเคลื่อนที่ของมือหุ่นยนต์เพื่อเข้าไปจับวัตถุโดยสามารถหลบหลีกสิ่งกีดขวางได้ การคำนวณท่าทางการจับที่สามารถจับวัตถุได้อย่างมีประสิทธิภาพ และองค์ความรู้ในอื่นๆ ที่เสริมความสามารถของการจับให้ดีและใช้งานได้ง่ายขึ้น ซึ่งการคำนวณหาท่าทางการจับวัตถุที่เหมาะสม เป็นหนึ่งในแนวทางที่สำคัญ ที่จำเป็นต้องพิจารณาในการจับวัตถุของหุ่นยนต์ โดยจะอาศัยข้อมูลจากอุปกรณ์รับรู้ที่หลากหลาย อาทิเช่น กล้อง (Camera), อุปกรณ์วัดระยะทางด้วยเลเซอร์ (Laser range finder) และอุปกรณ์วัดแรง (Force and contact sensor) เป็นต้น เพื่อนำข้อมูลเหล่านั้นมาใช้ระบุตำแหน่ง สร้างแบบจำลอง หรือหาคุณสมบัติต่างๆ ของวัตถุ แล้วจึงวิเคราะห์ท่าทางและจุดจับบนพื้นผิวของวัตถุที่เหมาะสม โดยจะคำนึงถึงความสมดุลของแรงที่กระทำต่อวัตถุ เพื่อให้สามารถจับวัตถุได้โดยวัตถุไม่หลุดออกจากมือหุ่นยนต์ ซึ่งมีหลากหลายงานวิจัยที่ได้พัฒนางานในส่วนนี้ และสามารถพิสูจน์ได้ว่าท่าทางการจับที่หามาได้นั้นจะสามารถจับให้วัตถุอยู่ในมือได้อย่างสมดุล แต่ผลการจับวัตถุที่แสดงให้เห็นจริง ยังคงมีผลการจับวัตถุที่ล้มเหลวเป็นส่วนมาก

สาเหตุที่ส่งผลให้การจับวัตถุในงานวิจัยต่างๆ ล้มเหลว เกิดจากหุ่นยนต์ไม่สามารถนำปลายนิ้วไปวางลงบนจุดจับที่ต้องการได้ เนื่องจากสภาพแวดล้อมในโลกจริงนั้น จะมีความไม่แน่นอน (Uncertainty) ที่ส่งผลให้เกิดความผิดพลาดทั้ง การวัดค่าข้อมูลจากอุปกรณ์รับรู้ การควบคุมหุ่นยนต์ รวมไปถึงการประมวลผลต่างๆ ตัวอย่างเช่น ภาพที่ได้จากกล้องมีจุดสีรบกวนเกิดขึ้นแบบสุ่มทำให้การคำนวณแบบจำลองวัตถุผิดไปจากความจริง และการเคลื่อนที่ของหุ่นยนต์มีระยะเกินจากคำสั่งที่สั่งจริงทำให้การระบุตำแหน่ง

ของปลายนิ้วผิดพลาด เป็นต้น ซึ่งผลจากความผิดพลาดเหล่านี้ ส่งผลให้การคำนวณหาจุดจับบนวัตถุ ไม่สามารถทำได้ถูกต้อง และการนำนิ้วไปวางบนจุดจับเหล่านั้น ไม่สามารถนำไปวางให้ตรงจุดได้ โดยอาจกล่าวได้ว่า “แม้สามารถคำนวณหาท่าทางและจุดจับของวัตถุได้ดีเพียงใด หากไม่สามารถวางนิ้วลงบนตำแหน่งที่หามาได้นั้น ผลการจับวัตถุก็ยังคงมีโอกาที่จะล้มเหลวสูง” ซึ่งหลายงานวิจัยจะไม่พิจารณาความผิดพลาดที่เกิดขึ้น เนื่องจากการจัดการกับความผิดพลาดเหล่านี้ ต้องอาศัยการคำนวณที่ยากและซับซ้อน อีกทั้งความผิดพลาดเหล่านี้ เป็นสิ่งที่ไม่สามารถคาดเดาได้ว่าจะเกิดขึ้นแบบใด ณ ระยะเวลาใด จึงเป็นปัญหาที่ยากต่อการแก้ไขอย่างยิ่ง

ด้วยเหตุผลดังกล่าว งานวิจัยนี้จึงมุ่งเน้นเพื่อพัฒนาและวางแผนวิธีการนำปลายนิ้ว ไปวางยังจุดจับที่กำหนดบนวัตถุอย่างแม่นยำ รวมไปถึงการจัดการกับความไม่แน่นอนที่เกิดขึ้น ทั้งความไม่แน่นอนในการสร้างแบบจำลองของวัตถุ ซึ่งส่งผลโดยตรงต่อการหาจุดจับที่เหมาะสม และความไม่แน่นอนในการระบุตำแหน่งที่จะส่งผลต่อความแม่นยำในการจับวัตถุ เพื่อเพิ่มประสิทธิภาพในการจับวัตถุที่ไม่รู้จักให้ดียิ่งขึ้น

2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยการจับวัตถุ นั้น จะสามารถจับวัตถุได้นั้นจะต้องคำนวณหา 2 สิ่งคือ จุดที่จะนำนิ้วไปจับบนวัตถุ และทิศทางการจับวัตถุ ซึ่งเมื่อพิจารณาปัญหาการจับวัตถุที่ไม่รู้จักแล้ว วัตถุที่ไม่รูรูปร่าง จะสามารถหาจุดจับและทิศทางการจับได้นั้น จำเป็นต้องสร้างแบบจำลองของวัตถุขึ้นมา หรือวิเคราะห์หาข้อมูลที่สามารถอธิบายลักษณะของวัตถุได้ ซึ่งมีหลายงานวิจัยได้นำเสนอวิธีการในการนำเสนอแบบจำลองของวัตถุ ที่แตกต่างกันออกไป อาทิเช่น ในสองมิติ [1, 2, 3, 4], ในสามมิติ [5, 6, 7] และในสองมิติครึ่ง [8, 9] เป็นต้น และยังมีงานวิจัย [10, 11] ที่เสนอวิธีการหาจุดจับโดยใช้เพียง features จากภาพมาวิเคราะห์โดยไม่ต้องสร้างแบบจำลองของวัตถุ

อุปกรณ์รับรู้ เป็นสิ่งจำเป็นสำหรับ งานวิจัยในเรื่องการจับวัตถุที่ไม่รู้จักอย่างมาก โดยที่ได้รับความนิยมมากที่สุดคือ อุปกรณ์รับรู้ภาพหรือกล้อง เนื่องจากเป็นอุปกรณ์ที่ใช้งานง่าย ราคาถูก ให้ข้อมูลปริมาณมาก และความเร็วในการส่งข้อมูลสูง แต่กล้องจะให้ข้อมูลภาพที่เป็นสองมิติเท่านั้น ทำให้ในหลายงานวิจัยจะมีลักษณะของการใช้งานกล้องที่แตกต่างกันออกไปเพื่อให้ได้ข้อมูลจากภาพที่มากขึ้น เช่น ในงานวิจัย [12] ได้นำกล้องสเตอริโอ (stereo camera) มาใช้งานเพื่อหาความลึก ขอบวัตถุ และพื้นผิวของวัตถุในสิ่งแวดล้อมที่ซับซ้อน เพื่อการวิเคราะห์หาจุดจับที่ดี และงานในปี ค.ศ. 2012 [13] ได้ใช้ Active vision ในการเปลี่ยนมุมมองเพื่อค้นหาวัตถุบนโต๊ะ และใช้ภาพจากการเปลี่ยนมุมมองหลายมุมมองของกล้องนี้มาใช้หาความลึกและแยกภาพส่วนของวัตถุออกจากฉากหลัง เป็นต้น

การติดกล้องที่มีมือของหุ่นยนต์ (Eye-in-hand camera) เป็นอีกหนึ่งลักษณะที่มีการนำมาใช้ในงานวิจัยการจับวัตถุที่ไม่รู้จัก โดยจะนำประโยชน์ของ Eye-in-hand camera คือสามารถปรับเปลี่ยนมุมมองของกล้องไปสังเกตรอบๆ วัตถุได้ มาใช้ในการพัฒนางานวิจัย เช่นในงาน [14, 15] เป็นงานวิจัยส่วนหนึ่งที่ช่วย

แก้ปัญหาการจับวัตถุที่ไม่รู้จัก โดยการพยายามสร้างแบบจำลองของวัตถุ ด้วยการเคลื่อนที่กล้องไปมารอบวัตถุ ซึ่งผลความถูกต้องของแบบจำลองวัตถุ ก็จะขึ้นอยู่กับเวลาที่ใช้ในการประมวลผล ทำให้ถ้านำไปใช้ร่วมกับการจับวัตถุจริง จะทำให้เสียเวลามากในการจับครั้งหนึ่งๆ และงานในปี ค.ศ. 2011 [2] ได้นำเสนอวิธีการจับวัตถุที่ไม่รู้จัก มีข้อดีคือ ง่าย เร็ว และทนทานต่อความผิดพลาด โดยในงานนี้จะใช้ข้อมูลภาพของวัตถุที่เป็นสองมิติจาก Eye-in-hand camera นำมาหาเส้นโค้งที่สามารถล้อมรอบภาพวัตถุที่ได้ และเสนอว่าจุดที่เป็น concave บนเส้นโค้งนั้นน่าจะเป็นจุดจับที่ดี จากนั้นก็ทำการเคลื่อนกล้องไปมุมต่างๆ เพื่อหาภาพที่ให้ค่าความโค้งมากที่สุด นั่นคือมุมมองที่จะสามารถหาจุดจับที่ดีได้นั่นเอง แต่ในการจับวัตถุจริงก็ยังไม่สามารถจับวัตถุได้ตรงตามที่หาไว้ได้ และไม่รับรองผลของการจับที่เกิดขึ้นในการจับจริงอีกด้วย

โดยงานวิจัยที่ได้กล่าวมานั้นจะถือว่าข้อมูลที่ได้จากอุปกรณ์รับรู้เป็นข้อมูลที่สมบูรณ์ ไม่มีความผิดพลาด และไม่มีการจัดการกับความไม่แน่นอนที่เกิดขึ้น ซึ่งจะส่งผลให้มีโอกาสเกิดความล้มเหลวในการจับวัตถุสูง ในหลายปีที่ผ่านมาจึงได้มีแนวคิด การจัดการภายใต้ความไม่แน่นอน (Manipulation Under Uncertainty) [16] ที่เป็นแนวทางการวิจัยของงานทางด้าน Manipulation ที่อยู่บนพื้นฐานของความไม่แน่นอนที่เกิดขึ้นในสภาพแวดล้อม และพยายามจัดการหรือลดปริมาณความผิดพลาดที่เกิดขึ้น เพื่อเพิ่มความถูกต้องและแม่นยำในการทำงานของหุ่นยนต์ ซึ่งเป็นแนวทางที่มีความสำคัญอย่างยิ่ง ในการพัฒนางานวิจัยการจับวัตถุที่ไม่รู้จัก ภายใต้สภาพแวดล้อมไม่อาจคาดเดาได้ โดยในปัจจุบันมีการพัฒนาภายใต้สภาพแวดล้อมที่หลากหลาย และมีลักษณะการใช้งานที่แตกต่างกันออกไป เช่น การประกอบอุปกรณ์ในโรงงาน [17], การหยิบจับอุปกรณ์ในครัว [18], การทำความสะอาดพื้นโต๊ะ [19] และการยกและวางวัตถุ [20, 21] เป็นต้น

มีหลายแนวทางที่จะจัดการกับสิ่งแวดล้อมที่ไม่แน่นอน ซึ่งบางวิธีจะขึ้นอยู่กับข้อจำกัดทางแบบจำลองของวัตถุ เช่น บริเวณสัมผัสอิสระ (Independent contact region) [22] ที่จะคำนวณพื้นที่ที่จะวางนิ้วมือ แทนการใช้จุด ซึ่งจะช่วยให้สามารถมีความผิดพลาดในการวางนิ้วได้เล็กน้อย แต่อย่างไรก็ตามวิธีนี้จำเป็นต้องรู้แบบจำลองของวัตถุมาก่อน เป็นต้น ในบางวิธีจะใช้การควบคุมหุ่นยนต์เพื่อจัดการกับความไม่แน่นอนที่เกิดขึ้น อาทิเช่น Visual Servoing [2, 13] จะใช้ error ที่หาได้จากภาพ มาควบคุมการเคลื่อนที่ของหุ่นยนต์เพื่อลด error ให้เข้าสู่ค่าศูนย์ แต่วิธีการนี้จะไม่รับรองผลว่าจะสามารถทำให้เข้าสู่ได้

การใช้อุปกรณ์รับรู้เป็นหลัก (Sensor based) เป็นอีกหนึ่งแนวทาง ที่จะนำข้อมูลสภาพแวดล้อมที่วัดได้จากอุปกรณ์รับรู้ มาวางแผนการทำงานของหุ่นยนต์ ให้เหมาะสมกับสถานการณ์ ณ ขณะนั้นๆ ตัวอย่างเช่น Teichmann และ Mishra [23] ได้นำเสนอวิธีการจัดวาง gripper ให้ตรงตำแหน่งโดยใช้อุปกรณ์วัดลำแสง (Light beam sensor) ซึ่งมีงานวิจัยอื่นๆ ได้ใช้แนวคิดในลักษณะเดียวกัน แต่มีการเปลี่ยนอุปกรณ์รับรู้เป็น Infra-red (IR) [24], อุปกรณ์วัดการสัมผัส (Tactile sensor) [25, 26] และอุปกรณ์วัดแรงและการสัมผัส (Force and tactile feedback) [27] ซึ่งวิธีเหล่านี้จำเป็นต้องนำอุปกรณ์รับรู้ เข้าไปใกล้กับวัตถุหรือสัมผัสวัตถุก่อนจึงจะสามารถวัดค่าข้อมูลได้ ซึ่งอาจทำให้มือหุ่นยนต์ชนกับวัตถุเกิดความเสียหายได้ อีกทั้งไม่สามารถระบุได้ว่า บริเวณที่นิ้วเข้าไปสัมผัสเป็นส่วนใดของวัตถุ จึงไม่สามารถเลือกกำหนดจุดจับที่ต้องการจาก

แบบจำลองวัตถุได้ ซึ่งแตกต่างจากอุปกรณ์รับรู้ภาพหรือกล้อง ถือว่าเป็นอุปกรณ์รับรู้หนึ่งที่มีการใช้งานในงานวิจัยทางการจับวัตถุเป็นจำนวนมาก [1, 2, 5, 10] ที่ให้ข้อมูลเป็นจำนวนมาก และสามารถบ่งบอกคุณลักษณะของวัตถุได้ดีกว่าอุปกรณ์รับรู้ชนิดอื่นๆ แต่ปัญหาสำหรับกล้องตัวเดียวคือ ข้อมูลที่ได้จากกล้องนั้นเป็นภาพในสองมิติ ซึ่งเกิดจากการ projection สิ่งแวดล้อมในสามมิติลงมาบนภาพ จึงทำให้ข้อมูลระยะทางจากกล้องถึงวัตถุสูญหายไป อีกทั้งอุปกรณ์รับรู้ชนิดนี้มีความไวต่อสภาพแสง เกิดสิ่งรบกวนได้ง่าย จึงยากแก่การจัดการกับความไม่แน่นอนหรือสิ่งรบกวนต่างๆ เหล่านี้

จากงานวิจัยที่ได้กล่าวมาข้างต้น จะเห็นได้ว่าอุปกรณ์รับรู้ภาพหรือกล้อง มีการนำมาใช้ในงานวิจัยด้านการจับวัตถุที่ไม่รู้จักกันอย่างแพร่หลาย อีกทั้งในปัจจุบันอุปกรณ์ชนิดนี้ มีขนาดเล็ก น้ำหนักเบา และให้ความละเอียดของภาพที่สูง สามารถนำไปติดตั้งยังส่วนต่างๆ ของหุ่นยนต์ได้ง่าย แต่การใช้กล้องในงานวิจัยด้านนี้ยังคงมีปัญหามารจัดการกับความไม่แน่นอนหรือสิ่งรบกวนที่เกิดขึ้น ซึ่งเป็นปัญหาที่มีความน่าสนใจและท้าทายในการทำงานวิจัยในด้านนี้อย่างยิ่ง วิทยานิพนธ์นี้จะอาศัยข้อมูลจากอุปกรณ์รับรู้เป็นกล้องเพียงหนึ่งตัว ในการพัฒนาระเบียบวิธี การวางแผนการจ้วงวางปลายนิ้ว สำหรับการจับวัตถุที่ไม่รู้จักนั้น ซึ่งมีเป้าหมายเพื่อเพิ่มความถูกต้องแม่นยำในการนำปลายนิ้วของมือหุ่นยนต์ ไปวางลงบนพื้นผิวของวัตถุที่ไม่รู้จัก ในจุดที่กำหนดไว้ รวมไปถึงการจัดการกับความไม่แน่นอนที่เกิดขึ้นภายในระบบด้วย

3 วัตถุประสงค์ของงานวิจัย

เพื่อพัฒนาระเบียบวิธี ในการวางแผนการจ้วงวางปลายนิ้วลงบนวัตถุอย่างแม่นยำ สำหรับการจับวัตถุที่ไม่รู้จัก ภายใต้สภาพแวดล้อมที่ไม่แน่นอน (Uncertainty) โดยใช้อุปกรณ์รับรู้เป็นกล้องหนึ่งตัวเท่านั้น

4 ทฤษฎีและอุปกรณ์ที่เกี่ยวข้อง

4.1 กล้องและการปรับแก้พารามิเตอร์ภายในของกล้อง (Camera and Intrinsic Parameter Calibration)

ลักษณะการทำงานของกล้องโดยทั่วไปนั้น จะเป็นการทำงานแบบกล้องรูเข็ม [28] โดยกล้องจะทำการฉายภาพวัตถุในพิกัดสามมิติผ่านเลนส์ที่เป็นรูขนาดเล็ก ให้มาอยู่ในพิกัดสองมิติบนภาพ ซึ่งสามารถอธิบายได้โดยสมการดังต่อไปนี้

$$s \cdot p = A[R|t]P$$

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

โดยกำหนดให้

s	เป็นค่าคงที่
$[R t]$	เป็นเมทริกซ์อธิบายการเลื่อนตำแหน่ง และการหมุน
(X, Y, Z)	เป็นพิกัดสามมิติของวัตถุในกรอบอ้างอิงโลก
(u, v)	เป็นพิกัดของจุดที่ฉายลงบนระนาบสองมิติในกรอบอ้างอิงของภาพ
A	เป็นเมทริกซ์อธิบายพารามิเตอร์ภายในของกล้อง
(c_x, c_y)	เป็นจุดกึ่งกลางภาพ
f_x, f_y	เป็นความยาวโฟกัส มีหน่วยเป็นพิกเซล

ค่าคงที่ s จะมีการเปลี่ยนแปลงค่า ในกรณีที่ภาพที่ได้จากกล้องถูกขยายขึ้นหรือย่อลงด้วยค่าคงที่หนึ่ง ซึ่งจะส่งผลให้พารามิเตอร์ในเมทริกซ์ A จะต้องถูกคูณด้วยค่าคงที่นั้น จึงต้องมีการเปลี่ยนแปลงค่าคงที่ s เพื่อรักษาให้เมทริกซ์ A มีค่าเท่าเดิม จึงเรียกเมทริกซ์ A ว่าพารามิเตอร์ภายในของกล้อง (Intrinsic Parameters) ซึ่งค่าพารามิเตอร์นี้จะขึ้นอยู่กับตัวกล้อง ไม่ขึ้นกับมุมมองของกล้อง และเมื่อถูกวัดออกมาแล้วจะสามารถนำไปใช้ซ้ำกับกล้องตัวเดิมได้เสมอ โดยที่ความยาวโฟกัสจะต้องไม่ถูกเปลี่ยนแปลง หรือไม่มีการซูมนั่นเอง ในส่วนของเมทริกซ์การเลื่อนตำแหน่งและการหมุน $[R|t]$ จะถูกเรียกรวมกันว่า พารามิเตอร์ภายนอกของกล้อง (Extrinsic Parameters) ซึ่งเป็นเมทริกซ์อธิบายการวางตัวของกล้องเมื่อเทียบกับกรอบอ้างอิงของโลก ที่เปลี่ยนตำแหน่ง (X, Y, Z) ไปเป็นตำแหน่ง (x, y, z) ในกรอบอ้างอิงที่คงที่เมื่อเทียบกับตำแหน่งของกล้อง ความสัมพันธ์ข้างต้นสามารถจัดรูปแบบใหม่ได้เป็น

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + t$$
$$x' = x/z$$
$$y' = y/z$$
$$u = f_x \cdot x' + c_x$$
$$v = f_y \cdot y' + c_y$$

แต่เนื่องจากเลนส์โดยทั่วไปที่มีใช้งานอยู่ในปัจจุบันนั้น จะมีปัญหาในเชิงเทคนิคที่ทำให้ภาพที่เกิดขึ้นเกิดการบิดงอ (Distortion) ซึ่งโดยส่วนใหญ่จะเป็นการบิดงอในแนวแกนรัศมีจากศูนย์กลางของเลนส์ ดังนั้นเราจึงต้องปรับแก้สมการข้างต้นใหม่ให้เป็น

$$x'' = x'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 x' y' + p_2 (r^2 + 2x'^2)$$
$$y'' = y'(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1 (r^2 + 2y'^2) + 2p_2 x' y'$$

โดยที่

$$r^2 = x'^2 + y'^2$$

และ

$$u = f_x \cdot x'' + c_x$$
$$v = f_y \cdot y'' + c_y$$

โดยกำหนดให้

- k_1, k_2, k_3 เป็นค่าคงที่ความบิดงอในแนวรัศมี
 p_1, p_2 เป็นค่าคงที่ความบิดงอในแนวตั้งฉาก
 (u, v) เป็นพิกัดของจุดที่ฉายลงบนระนาบสองมิติในกรอบอ้างอิงของภาพ
 A เป็นเมทริกซ์ออร์โธโกนอลของพารามิเตอร์ภายในของกล้อง
 (c_x, c_y) เป็นจุดกึ่งกลางภาพ
 f_x, f_y เป็นความยาวโฟกัส มีหน่วยเป็นพิกเซล

ค่าคงที่ในข้างต้นถือว่าเป็นส่วนหนึ่งของค่าพารามิเตอร์ภายในของกล้อง เนื่องจากค่าคงที่ดังกล่าวจะขึ้นกับตัวกล้อง ไม่ขึ้นกับตำแหน่งและมุมมองของกล้อง และค่าคงที่เหล่านี้จะไม่ขึ้นกับขนาดของภาพด้วย โดยการปรับแก้พารามิเตอร์เหล่านี้ จะช่วยเพิ่มความถูกต้องในการอ้างอิงตำแหน่งของวัตถุในสามมิติ กับตำแหน่งวัตถุภายในภาพให้มีความแม่นยำขึ้น ซึ่งวิธีการปรับแก้พารามิเตอร์เหล่านี้มีผู้นำเสนอไว้หลากหลายวิธี และมีการเผยแพร่โปรแกรมสำหรับการหาค่าพารามิเตอร์ภายในให้ใช้ แบบไม่คิดมูลค่าอีกด้วย

โดยในงานวิจัยนี้จะใช้เครื่องมือการหาค่าพารามิเตอร์ภายในจาก OpenCV [29] โดยใช้กระดานตารางสีขาวและดำ (รูปที่ 1) ที่รู้จำนวน ขนาด และตำแหน่งของแต่ละช่องอยู่ก่อนแล้ว จึงทำให้ทราบความสัมพันธ์ระหว่างจุดสองมิติในภาพ (p) กับตำแหน่งในโลกสามมิติ (P) จากนั้นจึงนำข้อมูลความสัมพันธ์ดังกล่าวมาแก้สมการหาค่าของพารามิเตอร์ภายในของกล้อง



รูปที่ 1 อุปกรณ์สำหรับการปรับแก้พารามิเตอร์ภายในของกล้อง



รูปที่ 2 กล้องรุ่น Logitech c920

สำหรับกล้องที่ใช้ในงานวิจัยนี้ จะเป็นกล้องยี่ห้อ Logitech รุ่น Logitech c920 [30] (รูปที่ 2) โดยมีคุณสมบัติต่างๆ ที่น่าสนใจดังนี้

ชื่อรุ่น	Logitech c920
Resolution	1920 x 1080
FOV	78°
Frame Rate (Max)	1080p @ 30 fps
Weight	162 g

ตารางที่ 1 คุณสมบัติของกล้องรุ่น Logitech c920

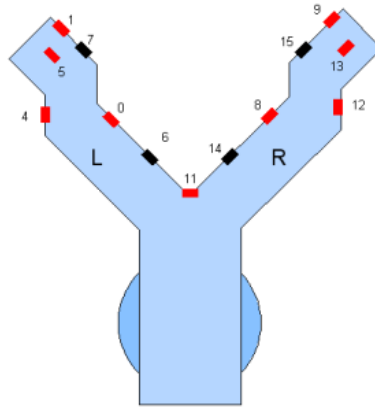
4.2 Katana Robotic Arm

Katana เป็นแขนหุ่นยนต์ที่ผลิตโดยบริษัท Neuronics จากประเทศสวีตเซอร์แลนด์ [31] ซึ่งได้มีการพัฒนาและผลิตออกมาหลากหลายรุ่น โดยในงานวิจัยนี้จะใช้แขนหุ่นยนต์รุ่น Katana 6M 180 (รูปที่ 3) ซึ่งเป็นแขนหุ่นยนต์ที่มีน้ำหนักเบา สามารถเคลื่อนที่ได้ 6 องศาอิสระ (degrees of freedom) โดยมาพร้อมด้วย gripper หนึ่งอัน ซึ่ง Katana 6M ขับเคลื่อนโดยมอเตอร์กระแสตรงจำนวน 6 ตัว ที่ควบคุมโดยการส่งคำสั่งจาก PC ผ่านสาย serial ไปยัง microcontroller ที่อยู่ภายในกล่องควบคุมของ Katana โดยตำแหน่งของมอเตอร์สามารถวัดได้โดยการอ่านค่าจาก encoders ของมอเตอร์แต่ละตัว



รูปที่ 3 (ซ้าย) แขนหุ่นยนต์ Katana 6M 180, (ขวา) ขอบเขตที่สามารถเคลื่อนที่เข้าไปถึงได้ของแขน Katana

สำหรับ gripper นั้นสามารถติดตั้งได้ 2 รูปแบบคือ ติดตั้งในทิศทาง 180 องศาเทียบกับข้อมือ (Katana 6M 180) และติดตั้งในทิศทาง 90 องศา (Katana 6M 90) ซึ่งอุปกรณ์รับรู้ที่ติดตั้งอยู่บน gripper มีทั้งหมด 15 ชิ้น ประกอบไปด้วยอุปกรณ์วัดระยะ (infra-red : IR) และอุปกรณ์วัดแรง (force and contact sensors) แสดงให้เห็นดังรูปที่ 4



รูปที่ 4 อุปกรณ์รับรู้อุปกรณ์ gripper (สีแดง) infra-red, (สีดำ) force and contact sensors

ในส่วนของอุปกรณ์ควบคุมของ Katana นั้นทำงานอยู่บนระบบปฏิบัติการ Linux ซึ่งเป็น open source จึงอนุญาตให้สามารถเข้าถึงส่วนควบคุมระดับล่าง และสามารถแก้ไขและพัฒนาให้ตรงกับความต้องการในการใช้งานที่แตกต่างกันได้ โดยในส่วนควบคุมนี้มีอุปกรณ์และคุณสมบัติที่น่าสนใจต่างๆ ดังนี้

Processor	Freescale MPC5200
Embedded controllers	6 x TI TMS320 32bit motor controllers
Memory	64MB RAM
Flash	32MB flash
CAN bus	1 x control bus; 1 x sensor bus; external CAN connector
Networking	1 x Ethernet hub
USB	1 x host; 1 x device
Other I/O	integrated digital I/O extension board
Operating system	Linux 2.4.25 (Linux kernel 2.6 optional) with Xenomai hard real-time extensions; driver patches for control board

ตารางที่ 2 อุปกรณ์ภายในส่วนควบคุมของ Katana

4.3 การระบุตำแหน่งพร้อมกับการสร้างแผนที่ (Simultaneous Localization and Mapping : SLAM)

วิธีการระบุตำแหน่งพร้อมกับการสร้างแผนที่ (SLAM) [32] เป็นกระบวนการที่หุ่นยนต์ จะสร้างแผนที่ของสภาพแวดล้อมในขณะที่กำลังเคลื่อนที่ และระบุตำแหน่งของตัวเองไปพร้อมๆ กัน โดยที่หุ่นยนต์จะไม่รู้ข้อมูลของสิ่งแวดล้อมนั้นมาก่อน ซึ่งปัญหาการระบุตำแหน่ง และการสร้างแผนที่ เป็นปัญหาที่มีความสอดคล้องและเกี่ยวพันกันเป็นอย่างมาก ในขณะที่เริ่มต้น หุ่นยนต์จะไม่รู้จักตำแหน่งของตัวเองมาก่อน และจะค่อยๆ สร้างแผนที่ของสิ่งแวดล้อมรอบๆ ตัวมันขึ้นมาเรื่อยๆ ในขณะที่เดียวกันก็จะใช้แผนที่นั้น ในการระบุ

ตำแหน่งของหุ่นยนต์ จึงกล่าวได้ว่า ความไม่แน่นอนของตำแหน่งของหุ่นยนต์ จะขึ้นอยู่กับความไม่แน่นอนของแผนที่ และในทางกลับกัน ความไม่แน่นอนของแผนที่ ก็จะขึ้นอยู่กับความไม่แน่นอนของตำแหน่งหุ่นยนต์ด้วย ซึ่งวิธีการระบุตำแหน่งพร้อมกับสร้างแผนที่นี้ จะใช้วิธีการทางความน่าจะเป็นในการจัดการกับความไม่แน่นอนเหล่านี้

เพื่อความสะดวกในการอธิบาย จะขอกำหนดตัวแปรดังต่อไปนี้ เมื่อหุ่นยนต์เคลื่อนที่ผ่านสิ่งแวดล้อมแล้วพบจุดสังเกต (Landmarks) ณ เวลา k กำหนดให้

- x_k เป็นสถานะของหุ่นยนต์ ประกอบด้วย ข้อมูลตำแหน่งและทิศทางของหุ่นยนต์ เมื่อเวลา k
- u_k เป็นคำสั่งของหุ่นยนต์ที่สั่งให้หุ่นเคลื่อนที่จากเวลา $k-1$ ไปสถานะ x_k ที่เวลา k
- m_i เป็นเวกเตอร์บอกถึงตำแหน่งของจุดสังเกต (Landmarks) จุดที่ i โดยสมมติว่าตำแหน่งที่แท้จริงไม่เปลี่ยนแปลงตามเวลา
- $z_{i,k}$ เป็นตำแหน่งของจุดสังเกตที่ i ที่ได้วัดจากหุ่นยนต์ที่เวลา k ซึ่งสามารถเขียนรวมทุกจุดสังเกตได้เป็น

z_k
นอกจากนี้ยังขอกำหนดให้

- $X_{0:k} = \{x_0, x_1, \dots, x_k\} = \{X_{0:k-1}, x_k\}$ เป็นบันทึก (History) ของสถานะหุ่นยนต์
- $U_{0:k} = \{u_0, u_1, \dots, u_k\} = \{U_{0:k-1}, u_k\}$ เป็นบันทึก (History) ของคำสั่งของหุ่นยนต์
- $m = m_0, m_1, \dots, m_n$ เป็นเซตของจุดสังเกตทั้งหมด
- $Z_{0:k} = \{z_0, z_1, \dots, z_k\} = \{Z_{0:k-1}, z_k\}$ เป็นเซตของข้อมูลการวัดจุดสังเกตทั้งหมดจากหุ่นยนต์

4.3.1 การระบุตำแหน่งพร้อมกับการสร้างแผนที่ในเชิงความน่าจะเป็น

รูปแบบของความน่าจะเป็นของปัญหาการระบุตำแหน่งพร้อมกับการสร้างแผนที่ (SLAM) สามารถเขียนในรูปแบบ การกระจายของความน่าจะเป็น (probability distribution) ได้ดังนี้

$$P(x_k, m | Z_{0:k}, U_{0:k}, x_0) \tag{1}$$

จากสมการแสดง การกระจายความน่าจะเป็น ของความหนาแน่นร่วมในภายหลัง (joint posterior density) ของสถานะหุ่นยนต์ (x_k) ซึ่งในที่นี้คือตำแหน่งของหุ่นยนต์ และตำแหน่งของจุดสังเกต (m) ที่เวลา k เมื่อกำหนดเซตของจุดสังเกตทั้งหมดที่วัดได้จากหุ่นยนต์ ($Z_{0:k}$), คำสั่งทั้งหมดของหุ่นยนต์ ($U_{0:k}$) ตั้งแต่เวลา 0 ถึง k และสถานะเริ่มต้นของหุ่นยนต์ (x_0)

ซึ่งโดยทั่วไปแล้วผลลัพธ์แบบเวียนบังเกิด (recursive solution) ของปัญหา SLAM จะเริ่มจากการประมาณการกระจายของ $P(x_{k-1}, m | Z_{0:k-1}, U_{0:k-1})$ ที่เวลา $k-1$ หลังจากนั้นจึงคำนวณ joint posterior จากคำสั่งของหุ่นยนต์ u_k และจุดสังเกตทั้งหมดที่วัดได้จากหุ่นยนต์ z_k ที่เวลา k โดยอาศัย Bayes theorem ซึ่งการคำนวณนั้นจำเป็นต้องอาศัยโมเดลการเปลี่ยนแปลงของสถานะหุ่นยนต์ (state transition model) และโมเดลการวัดจุดสังเกต (observation model) ที่แปรผันตามคำสั่งของหุ่นยนต์ และจุดสังเกต ตามลำดับ

โมเดลการวัดจุดสังเกต (observation model) จะบรรยายความน่าจะเป็นของจุดสังเกตทั้งหมดที่วัดได้จากหุ่นยนต์ (z_k) เมื่อรู้สถานะของหุ่นยนต์ (x_k) และตำแหน่งของจุดสังเกต m ที่แน่นอน ซึ่งเขียนได้เป็น

$$P(z_k|x_k, m) \quad (2)$$

ในที่นี้จะสมมติว่า จุดสังเกตที่วัดได้ (z_k) จะขึ้นอยู่กับสถานะหุ่นยนต์ (x_k) และตำแหน่งของจุดสังเกต m ปัจจุบันเท่านั้น

โมเดลการเปลี่ยนแปลงสถานะหุ่นยนต์ (state transition model) จะบรรยายการกระจายความน่าจะเป็นได้ดังนี้

$$P(x_k|x_{k-1}, u_k) \quad (3)$$

State transition ทั้งหมดถูกสมมติให้เป็น Markov process ได้ซึ่งสถานะถัดไป (x_k) จะขึ้นอยู่กับสถานะก่อนหน้า (x_{k-1}) และคำสั่งของหุ่นยนต์ (u_k) เท่านั้น โดยขึ้นกับจุดสังเกตที่วัดได้ (z_k) และแผนที่

SLAM อัลกอริทึม จะสามารถอธิบายได้ด้วย อัลกอริทึมแบบเวียนบังเกิดสองขั้นตอนคือ การประมาณค่า (prediction, time-update) และการปรับแก้ค่า (correction, measurement-update) ซึ่งอยู่ในรูป Prediction

$$P(x_k, m|Z_{0:k-1}, U_{0:k}, x_0) = \int P(x_k|x_{k-1}, u_k) \cdot P(x_{k-1}, m|Z_{0:k-1}, U_{0:k-1}, x_0) \cdot dx_{k-1} \quad (4)$$

Correction

$$P(x_k, m|Z_{0:k}, U_{0:k}, x_0) = \frac{P(z_k|x_k, m) \cdot P(x_k, m|Z_{0:k-1}, U_{0:k}, x_0)}{P(z_k|Z_{0:k-1}, U_{0:k})} \quad (5)$$

ในสมการที่ 4 และ 5 เป็นการคำนวณแบบเวียนบังเกิด สำหรับคำนวณหา ความน่าจะเป็นร่วม ภายหลัง (joint posterior) $P(x_k, m|Z_{0:k}, U_{0:k}, x_0)$ สำหรับสถานะหุ่นยนต์ (x_k) และเซตของจุดสังเกตทั้งหมดในแผนที่ m ณ เวลา k โดยขึ้นกับ เซตของจุดสังเกตทั้งหมดที่วัดได้ ($Z_{0:k}$) และคำสั่งทั้งหมดของหุ่นยนต์ ($U_{0:k}$) การคำนวณแบบเวียนบังเกิดนี้ จะอยู่ในรูปโมเดลการเปลี่ยนแปลงของสถานะหุ่นยนต์ (state transition model) $P(x_k|x_{k-1}, u_k)$ และโมเดลการวัดจุดสังเกต (observation model) $P(z_k|x_k, m)$

สำหรับปัญหาการสร้างแผนที่เพียงอย่างเดียวสามารถเขียน การกระจายความน่าจะเป็น ได้เป็น $P(m|X_{0:k}, Z_{0:k}, U_{0:k})$ โดยกำหนดว่ารู้สถานะของหุ่นยนต์ (x_k) ที่แน่นอนตลอดเวลา ดังนั้นแผนที่จะเป็นการรวมกันระหว่างจุดสังเกตที่ตำแหน่งของหุ่นยนต์ต่างๆ กัน ในทางกลับกัน สำหรับปัญหาที่ต้องการการระบุตำแหน่งเพียงอย่างเดียว สามารถเขียนการกระจายความน่าจะเป็นได้เป็น $P(x_k|Z_{0:k}, U_{0:k}, m)$ โดยสมมติว่ารู้ตำแหน่งของจุดสังเกตที่แน่นอนตลอดเวลา ดังนั้นจุดมุ่งหมายจะกลายเป็นการประมาณสถานะหุ่นยนต์เทียบกับจุดสังเกตทั้งหลาย

4.3.2 การแก้ปัญหา SLAM

การแก้ปัญหา SLAM เป็นการหารูปแบบการนำเสนอของ โมเดลการเปลี่ยนแปลงของสถานะหุ่นยนต์ (state transition model) (3) และโมเดลการวัดจุดสังเกต (observation model) (2) ที่มีประสิทธิภาพ

เพียงพอที่จะคำนวณ การประมาณค่า (4) และการปรับค่า (5) ซึ่งโดยทั่วไปแล้วจะถูกนำเสนออยู่ในรูป โมเดลของ state-space ที่ใช้ Gaussian noise เข้าร่วมด้วย ซึ่งวิธีการแก้ปัญหาของ SLAM มีผู้เสนอไว้มากมาย ยกตัวอย่างเช่น Extended Kalman Filter (EKF) [33, 34] เป็นต้น

5 แนวคิดของการวิจัยและวิธีการดำเนินงาน

การพัฒนาระเบียบวิธี ในการวางแผนการจ้วงปลายนิ้ว สำหรับการจับวัตถุที่ไม่รู้จักนั้น มีเป้าหมาย เพื่อเพิ่มความถูกต้องแม่นยำในการนำปลายนิ้วของมือหุ่นยนต์ ไปวางลงบนพื้นผิวของวัตถุที่ไม่รู้จัก ในจุดที่กำหนดไว้ ภายใต้สภาพแวดล้อมที่มีความไม่แน่นอนเกิดขึ้น โดยอาศัยข้อมูลจากอุปกรณ์รับรู้เป็นกล้อง ความละเอียดสูงเพียงหนึ่งตัวเท่านั้น

ปัญหาหนึ่งที่ต้องพิจารณาในงานวิจัยนี้คือ ปัญหาการระบุตำแหน่งของปลายนิ้ว เทียบกับจุดจับบน พื้นผิวของวัตถุที่ต้องการ โดยในงานวิจัยนี้จะใช้กล้องความละเอียดสูงหนึ่งตัว ติดตั้งบนมือหุ่นยนต์ (Eye-in-hand camera) ซึ่งมีข้อดีคือ สามารถปรับเปลี่ยนมุมมองของกล้องได้หลายหลายมุมมอง ช่วยเพิ่มข้อมูลที่จะ นำมาใช้ในการระบุตำแหน่ง อีกทั้งในการระบุตำแหน่งของกล้อง ก็จะทำให้สามารถระบุตำแหน่งของมือ หุ่นยนต์และปลายนิ้วได้โดยง่าย เมื่อพิจารณาการจับวัตถุที่ไม่รู้จัก ที่ไม่รู้ทั้งรูปร่าง ขนาด และคุณสมบัติ ต่างๆ ของวัตถุที่จะจับ ทำให้ไม่สามารถระบุตำแหน่งของปลายนิ้วเทียบกับจุดจับได้โดยตรง จึงมีความ จำเป็นต้องสร้างแบบจำลอง หรือหาข้อมูลที่บ่งบอกถึง รูปร่างและคุณลักษณะสำคัญ ที่สามารถนำมาใช้ วิเคราะห์หาจุดจับที่ติดบนวัตถุนั้นๆ ซึ่งการใช้กล้องเพียงหนึ่งตัว จะทำให้ปัญหานี้มีความยากและท้าทายเป็น อย่างยิ่ง ทำให้มีหลากหลายงานวิจัยได้พยายามพัฒนางานวิจัย เพื่อแก้ไขปัญหานี้ [10, 11, 12] โดย ภาพรวมแล้วมีหลายงานวิจัย ที่ใช้ข้อมูลภาพจากกล้องเพียงมุมมองเดียวในการวิเคราะห์หาจุดจับ ซึ่งทำให้ ไม่ได้ข้อมูลที่สำคัญของวัตถุที่อาจอยู่ในมุมมองอื่นๆ และมีงานวิจัยอีกกลุ่มหนึ่งที่พยายามสร้างแบบจำลอง ของวัตถุใน 3 มิติ (3D Reconstruction) [14] ก่อนที่จะนำมาใช้วิเคราะห์หาจุดจับ ซึ่งจะใช้เวลาานเพื่อให้ แบบจำลองที่ได้ออกมามีความแม่นยำ

ในงานวิจัยนี้จึงได้นำแนวคิดจากปัญหา การระบุตำแหน่งพร้อมกับสร้างแผนที่ (Simultaneous Localization and Mapping : SLAM) [32] เป็นกระบวนการที่หุ่นยนต์สร้างแผนที่ของสภาพแวดล้อมใน ขณะที่กำลังเคลื่อนที่ และระบุตำแหน่งของตัวเองไปพร้อมๆ กัน โดยในงานวิจัยนี้จะใช้การเคลื่อนที่ของ กล้องไปรอบๆ วัตถุ เพื่อสร้างแผนที่หรือก็คือการสร้างแบบจำลองของวัตถุ พร้อมกับระบุตำแหน่งของ กล้องที่ติดตั้งบนมือหุ่นยนต์ไปพร้อมกัน อีกทั้งในการแก้ไขปัญหานี้ SLAM จะมีการพิจารณาและจัดการกับ ความไม่แน่นอนที่เกิดขึ้น ซึ่งจะช่วยให้การระบุตำแหน่งของปลายนิ้วมีความแม่นยำมากยิ่งขึ้น โดยปัญหา ของงานวิจัยนี้ จะแบ่งออกเป็น 2 ประการหลักดังนี้

- 1) การระบุตำแหน่งพร้อมกับสร้างแบบจำลองวัตถุ
- 2) การวางแผนการจ้วงปลายนิ้ว

5.1 การระบุตำแหน่งพร้อมทั้งสร้างแบบจำลองวัตถุ

ปัญหาในส่วนการระบุตำแหน่งพร้อมทั้งสร้างแบบจำลองวัตถุนี้ สิ่งที่ต้องการคือ ตำแหน่งของกล้องที่ติดตั้งบนมือหุ่นยนต์ และแบบจำลองของวัตถุ ซึ่งในวิทยานิพนธ์นี้ จะมุ่งเน้นความแม่นยำเป็นหลัก ดังนั้นในการพัฒนางานส่วนนี้จึงต้องคำนึงถึงความผิดพลาดที่จะเกิดขึ้น ในการประมาณตำแหน่งของกล้อง และการสร้างแบบจำลองของวัตถุ โดยการนำแนวคิดของ SLAM มาช่วยแก้ไขปัญหาก็จะมีประเด็นหลักที่จะต้องพิจารณาอยู่ 2 ประเด็น นั่นคือ Observation model และ Map representation

5.1.1 Observation model

โมเดลการวัดจุดสังเกต (observation model) จะบรรยายความน่าจะเป็นของจุดสังเกตทั้งหมดที่วัดได้ ในการแก้ไขปัญห SLAM จะเป็นการหารูปแบบการนำเสนอของโมเดลการเปลี่ยนแปลงสถานะของหุ่นยนต์ (state transition model) และ โมเดลการวัดจุดสังเกต (observation model) ที่มีประสิทธิภาพ ซึ่งได้อธิบายไว้ในทฤษฎีที่เกี่ยวข้อง โดยในการวัดค่าจุดสังเกตที่สำคัญ จะต้องพิจารณาเลือก feature ภายในภาพ และหาความสัมพันธ์ของ feature ที่ตรวจพบกับ landmark ซึ่งมีแนวทางดังนี้

1) Features detection

ภาพที่ได้จากกล้องนั้น เป็นข้อมูลของจุดสีเป็นปริมาณมาก ซึ่งเป็นข้อมูลที่คอมพิวเตอร์นั้นไม่สามารถเข้าใจได้ ในขั้นตอนนี้จึงต้องพิจารณาหาส่วนของภาพที่มีลักษณะเด่นในสิ่งแวดล้อม โดยต้องคำนึงถึงความคงทนต่อการเปลี่ยนแปลงต่อสภาพแสง เพื่อให้สามารถตรวจพบได้อย่างแม่นยำ และสามารถแยกแยะออกจากภาพส่วนอื่นๆ ได้ เช่น เสา, มุมห้อง และขอบประตู เป็นต้น โดยลักษณะเด่นที่จะปรากฏอยู่ในภาพส่วนมากจะมีลักษณะ เป็นเส้นขอบ และเหลี่ยมมุม ซึ่งสามารถดึงเอาส่วนสำคัญของภาพเหล่านี้ออกมาได้ด้วยวิธีการทาง Image processing

2) Feature associations

นอกจากการตรวจหา feature ในภาพแล้ว ขั้นตอนสำคัญอีกขั้นตอนหนึ่งคือ การหาความสัมพันธ์ระหว่าง feature ที่ตรวจพบภายในภาพกับจุดสังเกต (landmark) ที่กล้องเคยตรวจพบแล้วในอดีต วิธีการหาความสัมพันธ์ระหว่าง feature กับ landmark สามารถแบ่งออกได้เป็น 2 แบบคือ

2.1) การหาความสัมพันธ์จากภาพ เป็นการหาความสัมพันธ์ของ feature ปัจจุบันเทียบกับ feature ในภาพก่อนโดยพิจารณาความสัมพันธ์จากภาพ ตัวอย่างอัลกอริทึมที่หาความสัมพันธ์ของ feature จากภาพเช่น optical flow จะเป็นอัลกอริทึมซึ่งได้จากการ track feature เพื่อติดตามตำแหน่งของ feature ทำให้ได้ข้อมูลการเคลื่อนของ feature จากอดีตสู่ปัจจุบัน แล้วจึงสามารถกำหนดความสัมพันธ์ของ feature ปัจจุบันกับ landmark ได้

2.2) การหาความสัมพันธ์จากข้อมูลการวัด การหาความสัมพันธ์จากข้อมูลการวัด จะใช้ก็ต่อเมื่อไม่สามารถหาความสัมพันธ์ของ feature และ landmark จากภาพได้ สาเหตุเนื่องจาก landmark ที่เคยตรวจพบนั้นหายไปจากภาพ แต่ได้กลับเข้ามาในภาพอีกครั้ง ด้วยสาเหตุนี้จึงไม่สามารถ

track feature ได้ ส่วนวิธีการหาความสัมพันธ์นั้น จะเปรียบเทียบความแตกต่างระหว่างค่าจุดสังเกตจริง กับค่าการวัดที่ประมาณได้จากสถานะจุดสังเกตของ SLAM สำหรับขั้นตอนการหาความสัมพันธ์ จะต้องทำหลังจากที่ SLAM ได้ prediction state เป็นที่เรียบร้อยแล้วจึงจะเกิดความแม่นยำ

5.1.2 Map representation

การกำหนดรูปแบบการนำเสนอของแผนที่ (map representation) เป็นการนำเสนอรูปแบบของแผนที่ของสภาพแวดล้อมที่จะใช้ใน SLAM ซึ่งมีหลากหลายรูปแบบอาทิเช่น point cloud และ mesh เป็นต้น โดยลักษณะของแผนที่ที่แตกต่างกันก็จะให้ข้อมูลที่แตกต่างกัน และมีประโยชน์ในด้านการใช้งานที่ต่างกัน ซึ่งในงานวิจัยนี้แผนที่ของสภาพแวดล้อม จะเปรียบได้กับแบบจำลองของวัตถุ การเลือกรูปแบบของแบบจำลองที่เหมาะสม จึงจะส่งผลต่อความแม่นยำในการสร้างแบบจำลองของวัตถุโดยตรง

5.2 การวางแผนการจัดวางปลายนิ้ว

เนื่องด้วยการจับวัตถุในแต่ละครั้ง แผนที่หรือแบบจำลองวัตถุจะถูกสร้างขึ้นใหม่ ทำให้ในตอนเริ่มการทำงาน จะไม่มีทั้งตำแหน่ง และแบบจำลองของวัตถุ จึงจำเป็นที่จะต้อง วางแผนการเคลื่อนที่เพื่อสำรวจวัตถุ ก่อนการจับวัตถุจริง

โดยในขั้นตอนนี้จะต้อง วางแผนการเคลื่อนที่ของกล้องรอบวัตถุ เพื่อสำรวจและสร้างแผนที่หรือแบบจำลองของวัตถุ ซึ่งต้องมีการตรวจสอบด้วยว่าแบบจำลองที่ได้เพียงพอสำหรับการหาจุดจับหรือไม่ อีกทั้งในการเคลื่อนกล้องเพื่อสำรวจนั้นอาจทำให้พบมุมมองที่เห็นชิ้นส่วนใหม่ของวัตถุ เช่น หูแก้วน้ำ จึงต้องมีการวางแผนและคิดอัลกอริทึมที่ใช้ในการตัดสินใจ เคลื่อนที่ไปในทิศทางที่คาดว่าจะพบมุมมองที่ปรากฏจุดจับที่ได้อยู่ด้วย โดยในวิทยานิพนธ์นี้จะมุ่งเน้นเฉพาะการวางแผนการนำปลายนิ้ว ไปวางบนจุดจับที่เหมาะสมบนวัตถุอย่างแม่นยำ โดยจะไม่พิจารณาวิธีการหาจุดจับบนวัตถุ ซึ่งจะนำวิธีการหาจุดจับที่เหมาะสมนี้มาจากงานวิจัยอื่นๆ

และในระหว่างการเคลื่อนที่ของของกล้อง อาจมีบางสถานการณ์ที่ส่งผลกระทบต่อการทำงานของระบบผิดพลาด เช่น กล้องเคลื่อนที่เข้าใกล้วัตถุมากเกินไป ไม่สามารถมองเห็น feature บนวัตถุได้ ทำให้การระบุตำแหน่งผิดพลาดได้ และเนื่องด้วยข้อจำกัดทางด้าน kinematic ของแขนหุ่นยนต์ ทำให้การเคลื่อนที่ของกล้องไม่เป็นไปตามที่วางแผน อาจส่งผลให้วัตถุหลุดออกจากมุมมองของกล้องได้ เป็นต้น จึงจำเป็นต้องมีกลยุทธ์ที่ใช้แก้ไขสถานการณ์ต่างๆ เหล่านี้ เพื่อให้การทำงานของระบบสามารถทำต่อไปได้อย่างถูกต้อง

5.3 การวัดผลการทดลอง

ในวิทยานิพนธ์นี้จะแบ่งการทดลองออกเป็น 2 การทดลองดังนี้

- 1) การทดลองเพื่อวัดผลความแม่นยำในการวางปลายนิ้ว ซึ่งจะวัดตำแหน่งที่ปลายนิ้ววางลงบนพื้นผิววัตถุ โดยจะเปรียบเทียบกับข้อมูลจุดจับที่เป็น ground truth
- 2) การทดลองเพื่อวัดความถูกต้องในการสร้างแบบจำลองวัตถุ โดยจะนำแบบจำลองที่สร้างได้จากวิธีที่นำเสนอ มาเปรียบเทียบกับแบบจำลองวัตถุที่มาจากข้อมูลที่เชื่อถือได้ เช่น จากฐานข้อมูลมาตรฐาน และจากข้อมูล Kinect เป็นต้น

6 ขอบเขตของงานวิจัย

- 1) งานวิจัยนี้จะพัฒนาระเบียบวิธี ในการวางแผนการจัดวางปลายนิ้วลงบนวัตถุอย่างแม่นยำ สำหรับการจับวัตถุที่ไม่รู้จัก โดยใช้อุปกรณ์รับรู้เป็นกล้องหนึ่งตัวเท่านั้น
- 2) งานวิจัยนี้จะสนใจเฉพาะการเคลื่อนแขนหุ่นยนต์ ไปวางนิ้วให้ตรงกับจุดจับที่กำหนดไว้ โดยไม่รวมการพัฒนาอัลกอริทึมการหาจุดจับ และไม่รวมการหลบหลีกสิ่งกีดขวาง
- 3) สภาพแวดล้อมจะต้องเหมาะสมกับการทำงานของแขนหุ่นยนต์ และอุปกรณ์รับรู้ โดยไม่มีสิ่งกีดขวางบดบังการทำงานของระบบ
- 4) วัตถุที่ใช้ในการจับจะมีรูปร่างไม่ซับซ้อน และมีลวดลายที่เด่นชัด

7 ขั้นตอนการดำเนินงาน

- 1) ศึกษาทฤษฎีพื้นฐานของการจับ และการทำงานของแขนหุ่นยนต์
- 2) ศึกษาทฤษฎีพื้นฐานของการระบุตำแหน่ง และการสร้างแบบจำลองวัตถุ
- 3) ศึกษาและพัฒนาระเบียบวิธีในการตรวจหา features จากกล้อง การแปลงและเลือกใช้ features ที่เหมาะสม
- 4) พัฒนาระเบียบวิธีการประมาณตำแหน่งของกล้อง และการสร้างแบบจำลองวัตถุ รวมถึงพัฒนาการจัดการความไม่แน่นอนที่เกิดขึ้นในระบบ
- 5) พัฒนาระเบียบวิธีในการวางแผนการเคลื่อนที่แขนหุ่นยนต์และกล้อง ในการปรับเปลี่ยนมุมมอง และการเข้าจับวัตถุ
- 6) วัดผลการทดลอง สร้างผลการทดลองพร้อมจัดทำรูปเล่ม และเนื้อหาวิทยานิพนธ์

8 เอกสารอ้างอิง

- [1] E. Chinellato, A. Morales, R. B. Fisher and A. P. del Pobil, “Visual quality measures for characterizing planar robot grasps,” *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 35, no. 1, pp. 30–41, 2005.
- [2] B. Calli, M. Wisse and P. Jonker, “Grasping of unknown objects via curvature maximization using active vision,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 996-1001, 2011.
- [3] C. Borst, M. Fischer, and G. Hirzinger, “Unknown object grasping using statistical pressure models,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 2, 2000, pp. 1054–1059.
- [4] J. Bohg and D. Kragic, “Grasping familiar objects using shape context,” in *Proceedings of the International Conference on Advanced Robotics (ICAR)*, 2009, p. 16.
- [5] K. Hübner, S. Ruthotto, and D. Kragic, “Minimum volume bounding box decomposition for shape approximation in robot grasping,” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '08)*, 2008, pp. 1628–1633.
- [6] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, “Automatic grasp planning using shape primitives,” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '03)*, 2003, pp. 1824–1829.
- [7] C. Goldfeder, P. K. Allen, C. Lackner, and R. Pelossof, “Grasp planning via decomposition trees,” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '07)*, 2007.
- [8] M. Richtsfeld and M. Vincze, “Grasping of unknown objects from a table top,” in *Workshop on Vision in Action: Efficient strategies for cognitive agents in complex environments*, 2008.
- [9] M. Richtsfeld and M. Zillich, “Grasping unknown objects based on 21/2d range data,” in *Proceedings of the IEEE Conference on Automation Science and Engineering*, 2008, pp. 691–696.
- [10] A. Saxena, J. Driemeyer and A. Y. Ng, “Robotic grasping of novel objects using vision,” *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 157–173, 2008.
- [11] A. Saxena, J. Driemeyer, J. Kearns, C. Osondu, and A. Y. Ng, *Experimental Robotics*. Springer Berlin / Heidelberg, 2008, vol. 39, ch. Learning to Grasp Novel Objects Using Vision, pp. 33–42.

- [12] M. Popović, G. Kootstra, J. Alison J., D. Kragic and N. Krüger, “Grasping unknown objects using an early cognitive vision system for general scene understanding,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 987-994, 2011.
- [13] X. Gratal, J. Romero, J. Bohg and D. Kragic, “Visual servoing on unknown objects,” *Mechatronics*, vol. 22, pp. 423-435, 2012.
- [14] G. Walck and M. Drouin, “Automatic Observation for 3D Reconstruction of Unknown Objects using Visual Servoing,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2727-2732, 2010.
- [15] V. Lippiello, F. Ruggiero and B. Siciliano, “Floating Visual Grasp of Unknown Objects Using an Elastic Reconstruction Surface,” *Robotics Research*, STAR 70, pp. 329–344, 2011.
- [16] “ICRA 2011 Workshop on Manipulation Under Uncertainty”. [Online]. Available: <http://www.willowgarage.com/muu11>. [Accessed: 13-Aug-12].
- [17] G. Zhao, Y. Jia and Z. Ou, “Construction of Vision-based Manipulation System for 3D Industrial Objects,” *Proceedings of the 2009 IEEE International Conference on Robotics and Biomimetics*, 2009, pp. 1051-1056.
- [18] L. Chang and N. Pollard, “Posture optimization for pre-grasp interaction planning,” *ICRA 2011 Workshop: Manipulation Under Uncertainty*, 2011.
- [19] J. Hess, J. Sturm and W. Burgard, “Learning to Efficiently Clean Surfaces with Mobile Manipulation Robots,” *ICRA 2011 Workshop: Manipulation Under Uncertainty*, 2011.
- [20] S. Tellex, T. Kollar, S. Dickerson, M. R. Walter, A. G. Banerjee, S. Teller and N. Roy, “Interpreting Robotic Mobile Manipulation Commands Expressed in Natural Language,” *ICRA 2011 Workshop: Manipulation Under Uncertainty*, 2011.
- [21] J. Felip, J. Bernabe and A. Morales, “Emptying the box using blind haptic manipulation primitives,” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011.
- [22] M. Roa and R. Suarez, “Computation of independent contact regions for grasping 3-d objects,” *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 839 –850, 2009.
- [23] M. Teichmann and B. Mishra, “Reactive algorithms for grasping using a modified parallel jaw gripper,” in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, pp. 1931 –1936 vol.3, May 1994.

- [24] K. Hsiao, P. Nangeroni, M. Huber, A. Saxena, and A. Y. Ng, "Reactive grasping using optical proximity sensors," in *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pp. 2098–2105, May 2009.
- [25] K. Hsiao, S. Chitta, M. Ciocarlie, and E. Jones, "Contact-reactive grasping of objects with partial shape information," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 1228–1235, 2010.
- [26] D. Gunji, Y. Mizoguch, S. Teshigawara, A. Ming, A. Namiki, M. Ishikawa, and M. Shimojo, "Grasping force control of multifingered robot hand based on slip detection using tactile sensor," in *SICE Annual Conference, 2008*, pp. 894–899, 2008.
- [27] J. Felip and A. Morales, "Robust sensor-based grasp primitive for a three-finger robot hand," in *Intelligent Robots and Systems 2009. IROS 2009. IEEE/RSJ International Conference on*, pp. 1811–1816, 2009.
- [28] Zhengyou Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol.1, pp.666-673, 1999.
- [29] "OpenCV: Open Source Computer Vision". [Online]. Available: <http://opencv.willowgarage.com/wiki/>. [Accessed: 13-Aug-12].
- [30] "Logitech HD Pro Webcam C920". [Online]. Available: <http://www.logitech.com/en-us/webcam-communications/webcams/hd-pro-webcam-c920>. [Accessed: 13-Aug-12].
- [31] "Neuronics". [Online]. Available: <http://www.neuronics.ch/>. [Accessed: 13-Aug-12].
- [32] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping:part i,ii," *IEEE Robotics & Automation Magazine*, vol. June, 2006.
- [33] P. Moutarlier and R. Chatila, "An experimental system for incremental environment modeling by an autonomous mobile robot," in *In 1st International Symposium on Experimental Robotics*, 1989.
- [34] R. Chatila and P. Moutarlier, "Stochastic multisensory data fusion for mobile robot location and environment modeling," in *In 5th Int. Symposium on Robotics Research*, 1989.