บทที่ 3

การประมวลรูปภาพ

3.1 นำเรื่อง

ในบทนี้จะกล่าวถึงวิธีการที่ใช้ในการนำ กล้องดิจิตัล (CCD camera) มาใช้ในการป้อนกลับเพื่อควบคุมแขนกล ระบบกล้องดิจิตัล (Vision System) มีจุดเด่นตรงที่เป็นอุปกรณ์ตรวจรู้แบบไม่ต้องสัมผัสกับชิ้นงาน มีความยืดหยุ่นสูงมาก และเหมาะสมกับงานที่สภาพแวดล้อมหรือตำแหน่งของวัตถุที่มีความไม่แน่นอน อย่างไรก็ตาม ระบบนี้ก็มีจุดอ่อนเมื่อเทียบกับอุปกรณ์ตรวจรู้แบบอื่นตรงที่มีความละเอียดต่ำ ทำการปรับเทียบได้ยาก และมีความถี่ในการวัดไม่สูงมากนัก

ระบบที่ใช้ระบบกล้องดิจิตัลในการป้อนกลับเพื่อควบคุมจะมีความสลับซับซ้อนมากกว่าระบบโดยทั่วไปเนื่องจากระบบกล้องดิจิตัลมีความถี่ในการวัดต่ำและข้อมูลที่ได้มีความล่าช้ากว่าเหตุการณ์ที่เกิดขึ้น ปัญหาดังกล่าวนี้ทำให้จำกัดระบบควบคุมที่ใช้ระบบกล้องดิจิตัลเฉพาะกับงานที่มีการเคลื่อนที่ไม่เร็วมากนัก กล้องดิจิตัลในทางการค้าโดยทั่วไปจะมีอัตราในการบันทึกภาพ 60 ภาพต่อวินาที และมีความละเอียด 480 × 512 จุดสี ระบบกล้องดิจิตัลที่มีความเร็วและความละเอียดสูงกว่านี้จะมีราคาสูงขึ้นมาก

ต่อจากนี้จะกล่าวถึงกล้องดิจิตัลโดยทั่วไปอย่างสั้นๆ ในตอนต้น เพื่อให้เห็นถึงจุดเด่นและข้อจำกัดของระบบนี้ และอธิบายถึงวิธีการพื้นฐานในการประมวลรูปภาพเพื่อให้ได้ข้อมูลของวัตถุเป้าหมายตามที่ต้องการ และได้สรุปถึงสมรรถนะของระบบกล้องดิจิตัลที่ใช้ในการศึกษานี้ในตอนท้าย อุปกรณ์ที่ใช้ในการศึกษานี้ประกอบด้วย:

- คอมพิวเตอร์ส่วนบุคคล PC Pentium II 266 MHz, 128 MB Ram
- การ์ด DT3155 จากบริษัท Data Translation Inc.
- กล้อง CCD จากบริษัท Camstar รุ่น CIC 741
- โปรแกรม HLImage++ จากบริษัท Western Vision Software
- โปรแกรม Visual Studio 5.0 จากบริษัท Microsoft Corporation

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.2 การบันทึกภาพของกล้องดิจิตัล

กล้องดิจิตัลจะบันทึกรูปภาพในรูปแบบดิจิตัลและจะมีโปรเซสเซอร์ปรับสัญญาณอยู่ภายใน กล้องความละเอียดสูงที่มีขนาดรูปภาพ 1024 X 1024 จุดสี (pixel) และเก็บข้อมูลแบบ 8 บิต โดยมีอัตราในการบันทึกภาพ 60 ภาพต่อวินาที จะต้องเก็บและประมวลข้อมูลรูปภาพ 60 เมกะไบต์ต่อวินาที ระบบกล้องดิจิตัลจะต้องสามารถเก็บและประมวลข้อมูลที่มีขนาดใหญ่ กล้องดิจิตัลรุ่นใหม่ๆ อาจสามารถลดขนาดของข้อมูลที่จะทำการบันทึกหรือประมวลผลโดยยอมให้เครื่องคอมพิวเตอร์ที่ต่ออยู่กับระบบกล้องดิจิตัลสามารถเข้าถึงค่าจุดสีได้อย่างอิสระ ทำให้ไม่จำเป็นต้องทำการบันทึกข้อมูลทั้งรูปภาพ แต่จะสามารถเลือกส่วนหนึ่งส่วนใดของรูปภาพเพื่อทำการบันทึกและประมวลผลได้ เป็นผลทำให้ระบบกล้องดิจิตัลมีอัตราการทำงานที่เร็วขึ้น และบางระบบจะมีโปรเซสเซอร์สำหรับประมวลผลรูปภาพเบื้องต้นเพื่อลดความล่าช้าของข้อมูล

พื้นฐานทั่วไปของกล้องที่มี Interface แบบ RS-170

ระบบกล้องดิจิตัลโดยทั่วไปจะใช้กล้องที่มี Interface แบบ RS-170 ซึ่งจะส่งสัญญาณภาพแบบแอนะลอกด้วยความเร็ว 60 ภาพต่อวินาที ความละเอียดของภาพโดยทั่วไปจะเท่ากับ 480 X 512 จุดสี วงรอบการทำงานของระบบกล้องดิจิตัลจะประกอบด้วย 4 ขั้นตอน คือ

1. การบันทึกภาพของกล้องดิจิตัล
2. การถ่ายโอนข้อมูลเข้าสู่เครื่องคอมพิวเตอร์ส่วนบุคคล
3. การประมวลรูปภาพ
4. การส่งข้อมูลไปใช้งาน

ถ้าภาพแต่ละภาพได้รับการประมวลเพื่อใช้ในงานควบคุม สัญญาณที่ได้โดยทั่วไปจะมีความล่าช้าจากเวลาที่เริ่มบันทึกภาพอย่างน้อย 1/20 วินาที หรือสามเท่าของอัตราการในการบันทึกภาพ

การบันทึกภาพของกล้องดิจิตัล

กล้องดิจิตัลจะมีการเปิดหน้ากล้องแบบอิเล็กทรอนิกส์ซึ่งจะทำให้ แอเรย์ดิจิตัล (CCD array) เกิดการชาร์จตามช่วงเวลาที่กำหนด (โดยทั่วไปแล้วจะประมาณ 1/60 วินาที) หลังจากหน้ากล้องปิด ประจุที่ชาร์จจะถ่ายโอนไปที่ Shift Register และแอเรย์ดิจิตัลก็จะว่าง พร้อมที่จะรับการบันทึกครั้งต่อไป เวลาที่ใช้ในการถ่ายโอนประจุจากแอเรย์ดิจิตัลไปที่ Shift Register นั้น

น้อยมาก กล้องดิจิตัลบางตัวมีความเร็วหน้ากล้องถึง 1/1000 วินาที หน้ากล้องที่เร็วจะทำให้สามารถถ่ายภาพเคลื่อนไหวได้คมและลดเวลาที่ใช้ในการบันทึกภาพ

การถ่ายโอนข้อมูลเข้าสู่เครื่องคอมพิวเตอร์ส่วนบุคคล

กล้องดิจิตัลจะส่งข้อมูลรูปภาพจาก Shift Register ไปที่บอร์ดบันทึกภาพ (Frame Grabber Board) ที่ติดตั้งบนเครื่องคอมพิวเตอร์ การส่งเป็นแบบอนุกรม โดยทั่วไปมีอัตราการส่ง 1/60 วินาที กล้องจะส่งข้อมูลค่าของจุดสีแบบแอนะลอกจากตำแหน่งแถวปัจจุบัน และบอร์ดบันทึกภาพจะทำการเปลี่ยนเป็นค่าดิจิตัลและบันทึกค่าลงบนหน่วยความจำบนบอร์ด โดยทั่วไปแล้วบอร์ดบันทึกภาพจะเก็บค่าของแถวที่ส่งมาจากกล้องโดยแบ่งออกเป็น 512 หลัก และจะทำการบันทึกทั้งหมด 240 แถวต่อเฟรม แต่ละเฟรมจะประกอบด้วยแถวคู่ทั้งหมดหรือแถวคี่ทั้งหมดของรูปภาพ กล้องจะทำการส่งสลับระหว่างเฟรมคู่กับเฟรมคี่ ข้อมูลจากสองเฟรมรวมกันจะให้ข้อมูลของรูปภาพที่สมบูรณ์ ซึ่งจะประกอบด้วย 480 แถว และ 512 หลัก (480 X 512 จุดสี)

การประมวลรูปภาพ

หลังจากที่ได้ของมูลของรูปภาพที่สมบูรณ์บนหน่วยความจำของบอร์ดบันทึกภาพเครื่องคอมพิวเตอร์ก็สามารถที่จะเริ่มขบวนการประมวลรูปภาพเพื่อให้ได้ข้อมูลตามที่ต้องการ การประมวลรูปภาพควรที่จะใช้เวลาน้อยกว่า 1/60 วินาที เพื่อให้กล้องสามารถทำการบันทึกภาพได้อย่างต่อเนื่อง เวลาที่ใช้ในการประมวลรูปภาพจะขึ้นอยู่กับอุปกรณ์พิเศษสำหรับประมวลภาพและประสิทธิภาพของวิธีการที่ใช้ ในบทนี้ได้อธิบายถึงกรรมวิธีในการประมวลรูปภาพที่ใช้ในการศึกษานี้ซึ่งจะกล่าวถึงต่อไป

การส่งข้อมูลไปใช้งาน

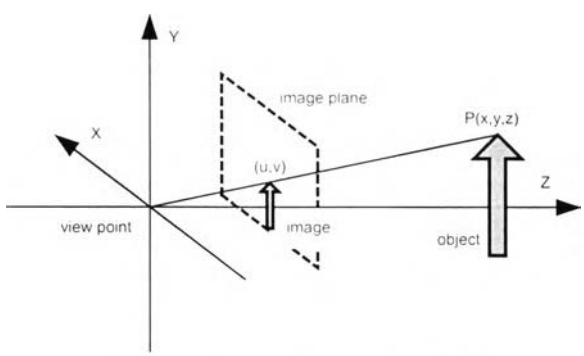
หลังจากเสร็จสิ้นการประมวลรูปภาพ ข้อมูลที่ถอดออกมาได้จากรูปภาพจะถูกส่งไปเพื่อใช้ในงานควบคุม เวลาที่ใช้ในการส่งจะขึ้นอยู่กับขนาดข้อมูลที่ได้ โดยทั่วไปแล้ว ข้อมูลที่พูดถึงนี้ก็คือตำแหน่งของวัตถุนั่นเอง

3.3 แบบจำลองทางคณิตศาสตร์ของกล้อง

ส่วนนี้ จะอธิบายถึงแบบจำลองทางคณิตศาสตร์สำหรับกล้อง ซึ่งจะเป็นความสัมพันธ์ระหว่างตำแหน่งของวัตถุบนแกนอ้างอิงกับตำแหน่งของวัตถุที่ปรากฏบนฉากรับภาพหรือบนภาพที่บันทึกได้จากกล้อง กล้องจะมีเลนส์ที่จะโปรเจ็ควัตถุจริง 3 มิติมาปรากฏเป็นรูป 2 มิติบน

ฉากรับภาพ ลักษณะการโปรเจ็กภาพแบบนี้จะทำให้ข้อมูลความลึกของวัตถุหายไป และตำแหน่งบนภาพก็จะสัมพันธ์กับเส้นทางเดินแสงใน 3 มิติ ดังนั้น จำเป็นต้องมีข้อมูลอื่นประกอบเพื่อระบุตำแหน่งใน 3 มิติของจุดหนึ่งๆ บนภาพ ข้อมูลนี้อาจมาจากการใช้กล้องหลายตัว หรือใช้กล้องเดียวแต่ถ่ายหลายๆ มุมมอง หรืออาศัยลักษณะรูปร่างของวัตถุที่ทราบมาก่อนล่างหน้ามาพิจารณาประกอบ แบบจำลองทางคณิตศาสตร์สำหรับแสดงความสำพันธ์ระหว่างตำแหน่งของวัตุจริงกับตำแหน่งของวัตถุบนภาพที่บันทึกได้ที่ใช้กันทั่วไปมี 3 แบบจำลอง คือแบบจำลองเพอร์สเพคทีพ (Perspective projection) แบบจำลองออร์โทกราฟิคแบบลดขนาด (Scaled Orthographic Projection) และแบบจำลองแอฟฟีน (Affine Projection) และต่อไปจะได้กล่าวถึงแบบจำลองทั้งสามแบบสั้นๆ

สำหรับแบบจำลองทั้งสามที่จะกล่าวถึง กำหนดให้ฉากรับภาพอยู่บนระนาบที่ขนานกับระนาบ xy และอยู่ห่างจากจุด origin เป็นระยะทาง λ (ระยะโฟกัส) ในขณะที่แกน z จะตั้งฉากกับระนาบนี้ ตำแหน่งของวัตถุบนฉากรับภาพจะระบุด้วยแกน u และ v ดังแสดงในรูปที่ 1



รูปที่ 3.1 แกนอ้างอิงสำหรับระบบกล้อง

1. **แบบจำลองเพอร์สเพคทีพ**: แบบจำลองนี้จะมีลักษณะเหมือนของจริงที่สุดโดยใช้หลักการพิจารณาทางเดินแสงเพื่อหาตำแหน่งที่ทางเดินนี้ตัดกับฉากรับภาพ จุด $^{c}P=[x,y,z]^{T}$ ซึ่งมีตำแหน่งสัมพัทธ์กับแกนอ้างอิงของกล้อง, c, จะโปรเจ็คลงบนฉากรับภาพที่ตำแหน่ง $p=[u,v]^{T}$ ตามสมการ

$$\pi(x,y,z)=\begin{bmatrix}u\\v\end{bmatrix}=\frac{\lambda}{z}\begin{bmatrix}x\\y\end{bmatrix}$$

2. **แบบจำลองออร์โทกราฟิคแบบลดขนาด**: แบบจำลองเพอร์สเพคทีพจะมีลักษณะเป็นแบบไม่เชิงเส้นซึ่งแม้ว่าจะมีความสมบูรณ์แต่ก็มีความสลับซับซ้อน ในบางกรณี อาจประมาณความสัมพันธ์ให้เป็นแบบเชิงเส้นเพื่อให้แบบจำลองมีโครงสร้างที่ง่ายขึ้น ในแบบจำลองแบบนี้ (ออร์โทกราฟิคแบบลดขนาด) จุด $^{c}P=[x,y,z]^{T}$ จะระบุโดย

$$\pi(x, y, z) = \begin{bmatrix} u \\ v \end{bmatrix} = s \begin{bmatrix} x \\ y \end{bmatrix}$$

โดยที่ s ก็คือตัวปรับลดที่มีค่าคงที่

แบบจำลองออร์โทกราฟิคแบบลดขนาดนี้จะมีความแม่นยำในกรณีที่ความลึกสัมพัทธ์ระหว่างจุดต่างๆ มีค่าน้อย เมื่อเทียบกับระยะทางระหว่างกล้องกับฉากรับภาพ

3. **แบบจำลองแอฟฟิน**: เป็นแบบจำลองเชิงเส้นที่ใช้ในการประมาณเพอร์สเพคทีพโปรเจ็คชั่นในอีกลักษณะหนึ่ง หรืออาจกล่าวได้ว่าแบบจำลองออร์โทกราฟิคแบบลดขนาดก็เป็นแบบจำลองแอฟฟินแบบหนึ่ง แบบจำลองแอฟฟินเป็นเพียงการประมาณเชิงเส้นโดยไม่คำนึงถึงลักษณะในการบันทึกภาพ จุด $^{c}P = [x, y, z]^{t}$ จะระบุโดย

$$\pi(x, y, z) = \begin{bmatrix} u \\ v \end{bmatrix} = A\,{}^{c}P + c$$

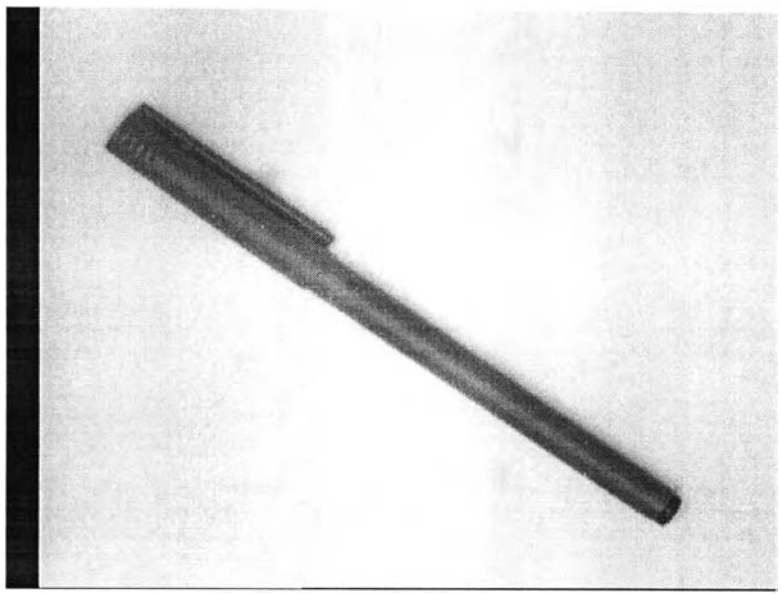
โดยที่ A คือเมตริกซ์ขนาด 2 × 3 และ c เป็นเวคเตอร์ 1 × 2

แบบจำลองแอฟฟินจะรวมทั้งโครงสร้างภายใน (เช่นเลนส์) และภายนอก (เช่นตำแหน่งของกล้อง) ของกล้อง และเนื่องจากว่าแบบจำลองเป็นแบบเชิงเส้น ทำให้สามารถคำนวณหา A และ c ได้อย่างไม่ยากนักโดยใช้วิธีการถดถอยเชิงเส้น (Linear Regression Technique) และจะทำให้สามารถปรับเทียบกล้องได้อย่างง่าย

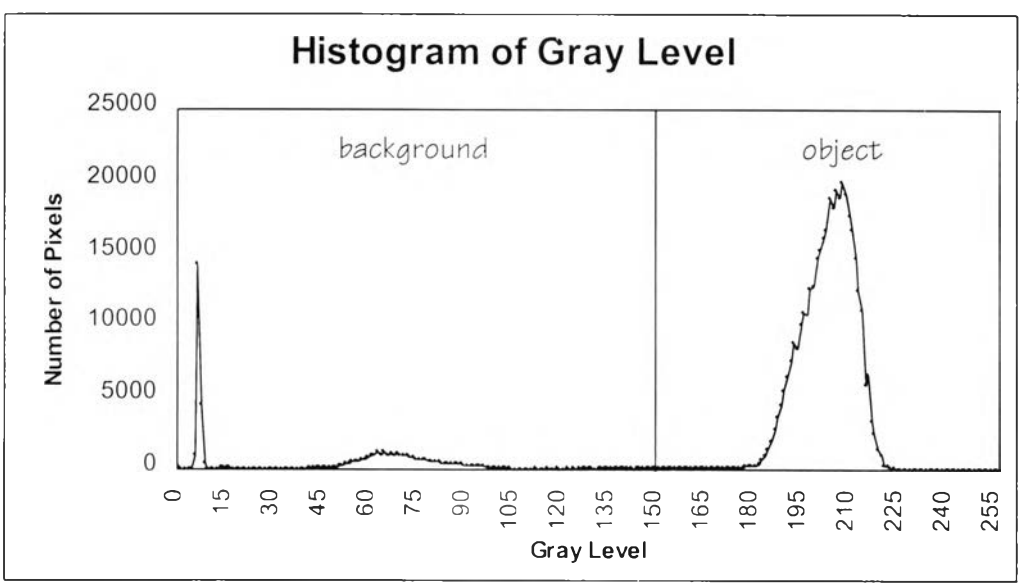
3.4 การประมวลรูปภาพเพื่อหาตำแหน่งของวัตถุ

ในส่วนนี้จะอธิบายขั้นตอนในการหาตำแหน่งของวัตถุจากรูปภาพอย่างละเอียด โดยจะเน้นถึงเฉพาะกรรมวิธีที่ใช้ในงานวิจัยนี้ และอธิบายถึงวิธีอื่นที่มีใช้อย่างสั้นๆ

ขั้นตอนการประมวลรูปภาพที่จะกล่าวถึงนี้ จะเป็นการพิจารณารูปภาพแบบบิตแมปเพื่อหาตำแหน่งจุดศูนย์กลางของวัตถุเทียบกับแกนของรูปภาพ (Image Coordinate Frame) โดยจะมี 3 ขั้นตอนคือ การแยกวัตถุออกจากพื้นโดยใช้วิธีการตัดสี (Thresholding) การหาขอบของวัตถุ (Edge Detection) และการหาจุดศูนย์กลางของวัตถุ (Centroid Determination)



รูปที่ 3.2 รูปขาวดำ 8 บิต ของปากกา



รูปที่ 3.3 การกระจายของระดับความเข้มของจุดสี

เริ่มต้นให้พิจารณาภาพปากกาซึ่งเป็นภาพขาวดำ 8 บิต ขนาด 768 × 576 จุด ดังแสดงในรูปที่ 3.2 จากภาพจะพบว่าวัตถุจะมีระดับสีเข้มกว่าพื้นหรือสิ่งแวดล้อม การแยกวัตถุออกมาจากสิ่งแวดล้อมก็จะอาศัยการพิจารณาความแตกต่างนี้ เมื่อพิจารณา Histogram ของระดับความเข้มของจุดสีต่างๆ ในภาพดังแสดงในรูปที่ 3.3 พบว่าความเข้มสีของวัตถุ (ยอดสูงด้านขวามือ) แตกต่างจากความเข้มสีของพื้นอย่างชัดเจน ถ้าใช้ระดับความเข้ม 125 เป็นจุดแบ่งระหว่างวัตถุกับพื้น (Thresholding) ก็จะทำให้สามารถแยกวัตถุออกมาจากพื้นได้ รูปที่ 3.4 แสดงภาพสองสีที่ประมวลได้จากการใช้วิธีแยกสี สีขาวในรูปจะหมายถึงพื้นและสีดำในรูปจะหมายถึงวัตถุ

รูปที่ 3.4 ภาพสองสีแสดงวัตถุที่ประมวลได้โดยใช้การดัดสี

ค่าดัดสีที่ใช้ในการแยกระหว่างวัตถุกับพื้นโดยทั่วไปจะเขียนอยู่ในรูป:

$$T = T[x, y, p(x, y), f(x, y)]$$

โดยที่ p(x,y) คือคุณสมบัติที่พิจารณา ณ บริเวณรอบตำแหน่งของจุดสี
f(x,y) คือระดับความเข้มของจุดสีที่ตำแหน่ง (x,y)

และจะแยกระหว่างวัตถุกับพื้นได้จาก

(x,y) เป็นส่วนหนึ่งของวัตถุ ถ้า f(x,y) > T

สมมุดให้วัตถุมีความเข้มมากกว่าพื้น ถ้า T ขึ้นอยู่กับ f(x,y) เท่านั้นจะเรียกว่า การดัดสีแบบสม่ำเสมอ (Global Threshold) ถ้า T ขึ้นอยู่กับทั้ง f(x,y) และ p(x,y) จะเรียกว่า การดัดสีตามตำแหน่ง (Local Threshold) และถ้า T ขึ้นอยู่กับตำแหน่งของ (x,y) จะเรียกว่า การดัดสีพลวัด (Dynamic Threshold) หรือจะกล่าวอีกนัยหนึ่งก็คือ การดัดสีแบบสม่ำเสมอจะทำการแยกวัตถุโดยใช้ค่าดัดสีค่าเดียวกับทุกจุดในรูปภาพ ในขณะที่การดัดสีพลวัดจะแบ่งภาพใหญ่ออกเป็นภาพย่อยๆ และจะทำการดัดสีในภาพย่อยๆ เหล่านี้

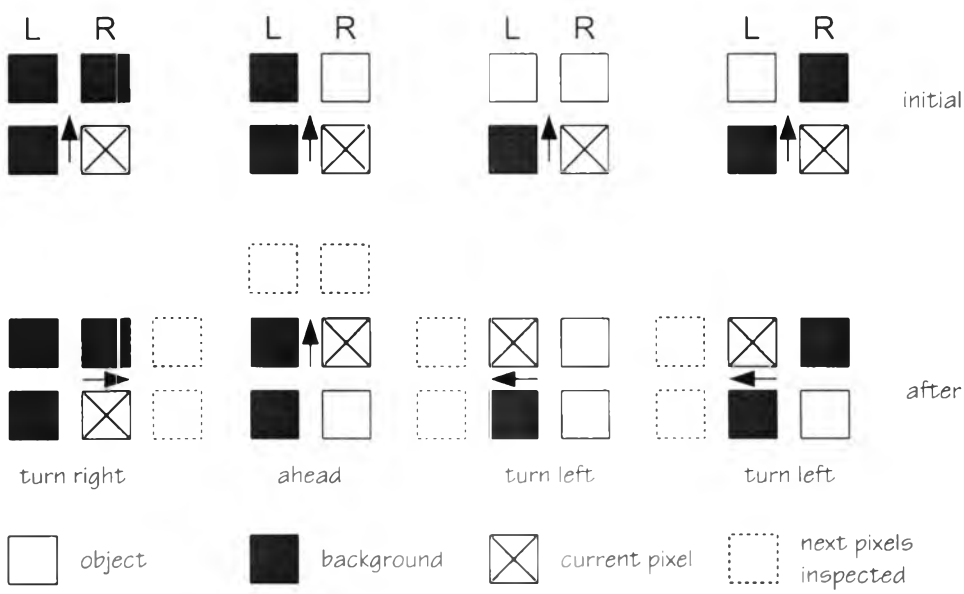
เมื่อได้ภาพสองสีที่แยกระหว่างวัตถุกับพื้นอย่างชัดเจน ก็จะนำภาพสองสีนี้มาหาขอบของวัตถุ วิธีการหาขอบของวัตถุจะใช้หลักการดังนี้ (รูปที่ 3.5) เริ่มต้น จะสแกนรูปภาพเพื่อหา

จุดสีที่เป็นวัตถุ โดยอาจเริ่มจากการพิจารณาจุดกึ่งกลางด้านซ้ายสุดของรูปภาพก่อน แล้วสแกนมาทางขวาเพื่อหาจุดสีที่เป็นวัตถุจุดแรก ถ้าไม่พบก็ทำการสแกนบรรทัดถัดไป เมื่อตรวจพบจุดสีที่เป็นวัตถุ จุดนี้จะเป็นจุดที่อยู่บนขอบของวัตถุเนื่องจากอยู่ติดกับจุดสีที่เป็นพื้น การตรวจหาจุดขอบถัดไปก็ทำได้โดยการพิจารณาสองจุดสีที่อยู่ข้างหน้า (ทั้งด้านซ้ายและด้านขวา) และพยายามรักษาให้วัตถุอยู่ทางด้านขวาโดยอาศัยกฎต่อไปนี้ (X คือพื้น และ O คือวัตถุ)

ตารางที่ 3.1 ทิศทางการเดินหาขอบ

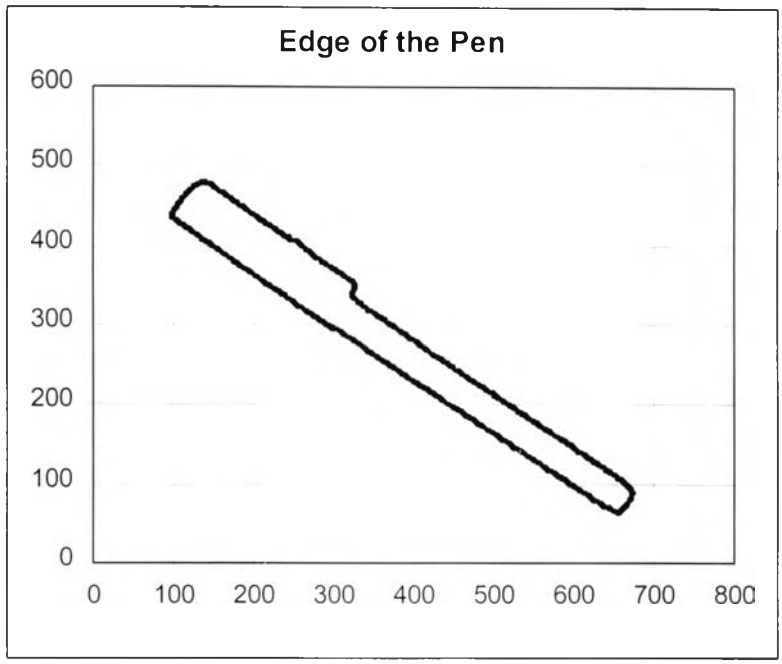
ซ้าย	ขวา	ดำเนินการ
X	X	เลี้ยวขวา
X	O	จุดนี้คือจุดขอบ เดินต่อไป
O	O	จุดนี้คือจุดขอบ เลี้ยวซ้าย
O	X	เหมือนข้างบน

การตรวจหาขอบจะสิ้นสุดเมื่อตรวจพบจุดเริ่มต้นในทิศทางเดียวกัน



รูปที่ 3.5 การตรวจหาขอบ

เมื่อพิจารณาถึงจุดสีจุดหนึ่ง พบว่าจะติดกับจุดรอบด้านทั้งหมด 8 จุด กรรมวิธีตรวจหาขอบวิธีนี้จะทำให้เกิดทางเดินใน 4 ทิศทาง (จากที่เป็นไปได้ทั้งหมด 8 ทิศทาง) และจะไม่มีการพิจารณาจุดสีซ้ำจุดเดิม ทำให้วิธีการนี้ใช้เวลาค่อนข้างน้อย รูปที่ 3.6 แสดงผลลัพธ์ที่ได้จากการตรวจหาขอบของรูปปากกาที่แสดงไว้ในตอนต้น ซึ่งประกอบไปด้วยจุดขอบ 1411 จุด



รูปที่ 3.6 ขอบของวัตถุ

เพื่อป้องกันความผิดพลาดในกรณีที่ตรวจเจอรอยเปื้อนของรูปภาพแล้วเข้าใจผิดว่าเป็นวัตถุ ระบบจะพิจารณาขนาดของวัตถุที่ตรวจพบ ถ้ามีขนาดเล็กกว่าที่กำหนดไว้ก็หมายความว่าตำแหน่งที่ตรวจพบเป็นเพียงรอยเปื้อนบนรูปภาพเท่านั้น

เมื่อได้จุดขอบทั้งหมด ก็สามารถนำจุดเหล่านี้มาคำนวณหาตำแหน่งศูนย์กลาง (Centroid) ของวัตถุได้ตามสมการต่อไปนี้

$$x_c = \frac{\sum_{i=0}^{n-1} x_i}{n} \quad \text{และ} \quad y_c = \frac{\sum_{i=0}^{n-1} y_i}{n}$$

และพื้นที่ของวัตถุจะสามารถคำนวณได้จาก

$$A = \left| \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i) \right|$$

ตำแหน่งศูนย์กลางที่คำนวณได้นี้ จะเป็นตำแหน่งที่เทียบกับแกนอ้างอิงของรูปภาพ และจากตัวอย่างรูปปากกา (รูปที่ 3.2) สามารถหาตำแหน่งศูนย์กลางและพื้นที่ได้ดังนี้

ตำแหน่งศูนย์กลางในแนวแกน X เท่ากับ 362.000 และ
ตำแหน่งศูนย์กลางในแนวแกน Y เท่ากับ 288.000 และ
พื้นที่เท่ากับ 33395 จุดสี^{2}

ความเร็วในการประมวลผลรูปภาพ

เมื่อนำการประมวลรูปภาพมาทดสอบบนเครื่อง Pentium II ความเร็ว 266 MHz และมีหน่วยความจำ 128 MB รุ่น Optima MT ของบริษัท ALR โดยติดตั้งการ์ด DT3155 จากบริษัท Data Translation Inc. และใช้กล้อง CCD จากบริษัท Camstar รุ่น CIC 741 ได้ผลลัพธ์ดังนี้

ตารางที่ 3.2 ความเร็วในการบันทึกและประมวลภาพ

ขั้นตอน	เวลาที่ใช้ (มิลลิวินาที)
การบันทึกภาพ	80.13
การดัดสี การตรวจหาขอบ และการหาจุดศูนย์กลาง	288.40
รวม	368.53

หมายเหตุ รูปที่บันทึกเป็นภาพขาวดำ 8 บิต ขนาด 768 × 576 จุด

3.5 อภิปรายและสรุป

จากที่ได้อธิบายในตอนต้นถึงหลักการพื้นฐานในการบันทึกภาพของกล้องดิจิตัลจะพบว่าในการบันทึกภาพจะต้องใช้เวลาในการถ่ายโอนข้อมูลรูปภาพจากตัวกล้องเข้าสู่เครื่องคอมพิวเตอร์ส่วนบุคคล ทำให้ภาพที่ปรากฏมีความล่าช้ากว่าเหตุการณ์ที่เกิดขึ้น ระบบความคุมที่ใช้รูปภาพเป็นข้อมูลป้อนกลับจึงจำเป็นจะต้องสามารถรับมือกับความล่าช้าที่เกิดขึ้นนี้ เป็นผลทำให้ระบบกล้องดิจิตัลจำกัดอยู่เฉพาะกับงานที่มีความเร็วไม่มากนัก นอกจากนั้นแบบจำลองทางคณิตศาสตร์สำหรับการบันทึกภาพยังประกอบด้วยพารามิเตอร์หลายตัว ซึ่งรวมถึงความยาวโฟกัสของเลนส์และตำแหน่งของกล้อง ถ้าพารามิเตอร์เหล่านี้ไม่มีความแม่นยำเพียงพอ ก็จะทำให้ตำแหน่งของวัตถุที่ประมวลได้จากรูปภาพมีความคลาดเคลื่อนจากตำแหน่งจริง และเป็นผลทำให้ระบบควบคุมไม่มีประสิทธิภาพดีตามที่ต้องการ

เพื่อจัดการกับปัญหาที่กล่าวมา ในการวิจัยนี้ จึงออกแบบระบบควบคุมที่นำข้อมูลพารามิเตอร์ในรูปภาพไปใช้โดยตรงโดยที่ไม่จำเป็นต้องทราบพารามิเตอร์ต่างๆ ของแขนกลและของกล้อง ซึ่งจะสอดคล้องกับการทำงานของมนุษย์ซึ่งไม่ทราบพารามิเตอร์ต่างๆ ของร่างกาย เช่น ความยาวแขน ตำแหน่งของหัวไหล่ ระยะโฟกัสของดวงตา และอื่นๆ อีกมากมาย การออกแบบระบบควบคุมที่มีโครงสร้างที่ไม่สลับซับซ้อนและไม่พึ่งความแม่นยำของพารามิเตอร์ต่างๆ จะทำให้การควบคุมหุ่นยนต์โดยใช้กล้องดิจิตัลมีความเป็นไปได้ในเชิงปฏิบัติ ส่วนปัญหาในเรื่องความล่าช้าของภาพที่บันทึกได้กับสิ่งที่เกิดขึ้นจริง ก็แก้ไขได้โดยอาศัยเทคนิคในการประมาณ

ตำแหน่งของวัตถุในเวลาจริงโดยอาศัยข้อมูลจากรูปภาพที่บันทึกได้ในอดีต (โดยทั่วไปจะใช้ Kalman filter) ในงานวิจัยนี้ เน้นที่การพัฒนาระบบควบคุมเพื่อให้ใช้งานได้จริงในเชิงปฏิบัติ โดยที่จะลดความเร็วของการควบคุมลงเพื่อให้มีเวลามากพอในการประมวลผลรูปภาพ ซึ่งก็จะทำให้สามารถรับมือกับปัญหาความล่าช้าของภาพที่บันทึกได้ด้วย อนึ่ง ปัญหานี้จะลดความสำคัญลงในอนาคตเนื่องจากขีดความสามารถในการประมวลผลมีประสิทธิภาพมากขึ้นเรื่อยๆ