

บทที่ 2

ทบทวนวรรณกรรมที่เกี่ยวข้องกับการจัดภาพสามมิติให้ตรงกัน และการซ้อนทับภาพ

2.1 บทนำ

วิธีการจัดภาพสามมิติให้ตรงกันและการซ้อนทับภาพ (Image Registration and Alignment) หมายถึงกระบวนการค้นหาพื้นที่หรือบริเวณของภาพสองภาพที่เหมือนกัน อาจจะเป็นรูปร่าง พื้นผิว ความเข้ม หรือจุดเด่นเอกลักษณ์อื่นใดที่สนใจ อย่างน้อยสองภาพแล้วนำมาซ้อนทับให้ตรงกันพอดี โดยการจับคู่ข้อมูลที่สอดคล้องกันของภาพทั้งสอง โดยที่ข้อมูลนั้นอาจเป็น จุด รูปร่าง ส่วนโค้ง ส่วนเว้า รอยหยัก สัน ปลาย หรือ ความเข้มของภาพ เป็นต้น

การซ้อนทับภาพจำแนกได้สองวิธีหลักคือแบบใช้ความเข้มเป็นหลัก (Intensity-based) และแบบใช้คุณลักษณะเฉพาะเป็นหลัก (Feature-based) การใช้ความเข้มเป็นหลักเป็นการอาศัยความเข้มของแต่ละจุดภาพ (Pixel) ของภาพทั้งสองที่สัมพันธ์กัน ส่วนในภาพสามมิติที่ได้จากเครื่องสร้างภาพทางการแพทย์ เช่น เครื่องเอกซเรย์คอมพิวเตอร์และเครื่องสร้างภาพจากการกำรสแกนแม่เหล็กนั้น ก็สามารถใช้ความเข้มจากวอกซ์เซล (Voxel) ในการศึกษาเป็นการทดลองเพื่อซ้อนทับภาพระยะทางสามมิติ (Range Image) และจัดภาพให้ตรงกัน โดยที่ภาพเหล่านี้ได้จากการถ่ายภาพด้วยเทคนิคการถ่ายภาพสเตอริโอและการสแกนด้วยเลเซอร์

การซ้อนทับภาพและการจัดภาพให้ตรงกันนั้นเป็นขั้นตอนที่สำคัญอย่างยิ่งในการนำภาพมาประยุกต์ใช้งาน เริ่มตั้งแต่การถ่ายภาพ (Image acquired) การผสมภาพ (Image fusion) การทำให้คอมพิวเตอร์เข้าใจทัศนียภาพ (Computer vision) การรู้จักใบหน้ารูปทรง (Recognition) เพื่อการระบุตัวตน (Identification) เป็นต้น ในการทบทวนวรรณกรรมนี้จะนำเสนอวรรณกรรมที่เกี่ยวข้องกับการซ้อนทับภาพในหัวข้อ 2.2 และวรรณกรรมที่เกี่ยวข้องกับการซ้อนทับภาพและการจัดให้ตรงกันของภาพพื้นผิวสามมิติหรือภาพระยะสามมิติในหัวข้อ 2.3 ซึ่งเป็นวรรณกรรมที่เกี่ยวข้องกับงานวิจัยและจัดเป็นแบบใช้คุณลักษณะเฉพาะเป็นหลัก (Feature-based)

2.2 วรรณกรรมของการซ้อนทับภาพ

วิธีการซ้อนทับภาพและจัดภาพให้ตรงกันสามารถแบ่งออกเป็นสองวิธีคือ แบบใช้ความเข้มเป็นหลัก (Intensity-Based) และแบบใช้คุณลักษณะเฉพาะเป็นหลัก (Feature-Based) แบบใช้ความเข้มเป็นหลักอาศัยความเข้มจากพิกเซล (pixel) เมื่อเป็นภาพสองมิติหรืออาศัยข้อมูลความเข้มจากวอกซ์เซล (voxel) ในกรณีที่ภาพนั้นมาจากข้อมูลปริมาตรสามมิติ การทำงานโดยการใช้ความเข้มเป็นหลักนี้จำเป็นต้องคำนวณซ้ำหลายครั้งจนได้ผลดีที่สุดและเพื่อหลีกเลี่ยงการทำงานที่ผลการจัดภาพให้ตรงกันจะผิดพลาด อันเนื่องจากจุดสนใจมีหลายจุด (local minima and local maxima) จึงจำเป็นต้องกำหนดจุดเริ่มต้นในการคำนวณคร่าวๆ เสียก่อน (initial alignment) ก่อนที่จะทำการซ้อนทับภาพ ดังเช่นในงานของ Wood และคณะ [1, 2] โดยการใช้ค่าเบี่ยงเบนมาตรฐานของจุดภาพ

MRI ที่สอดคล้องกันของสองภาพมาใช้ในการจัดภาพให้ตรงกัน Viola and Wells [3] ได้ศึกษาเกี่ยวกับการจัดภาพ MRI ให้ตรงกันโดยใช้ความเข้มสูงสุดที่เหมือนกัน วิธีนี้มีประสิทธิภาพมากในการซ้อนทับภาพระหว่างภาพสามมิติกับภาพสามมิติที่มาจากเครื่องสร้างภาพต่างชนิดกัน เช่น ระหว่างเครื่อง MRI กับ PET หรือ ระหว่างเครื่อง MRI กับ CT ดังปรากฏใน [4-6] แต่ว่าประสิทธิภาพไม่ดีนักในกรณีที่ต้องการจัดภาพให้ตรงกันระหว่างภาพสองมิติกับสามมิติ ส่วนงานอื่นที่พบว่าใช้หลักการของความเข้มเช่นการใช้ normalized cross correlation [7] entropy of difference image [8] gradient correlation [9] และ pattern intensity [10]

แบบใช้คุณลักษณะเฉพาะเป็นหลัก (feature-based) เป็นการหาคุณลักษณะเฉพาะของภาพที่สอดคล้องกันหรือเหมือนกันจากภาพแล้วคำนวณหาตัวแปรการแปลง (transformation parameters) ที่สอดคล้องกันของคุณลักษณะเหล่านั้น วิธีการใช้คุณลักษณะเฉพาะเป็นหลักนี้ยังสามารถแบ่งย่อยได้เป็นแบบการตรวจจับความสมมาตรของพื้นผิว (surface symmetry detection) ตัวแทนโพลิโนเมียลอันดับสูง (high order polynomial representation) ลักษณะที่คงตัว (invariance-based approaches) แผนที่ระยะทาง (distance map technique) หลักการของพลังงาน (energy-based method) แผนที่ของจุดสนใจ (landmark-mapping techniques) และแผนที่ของขอบหรือลายเส้น (boundary/line mapping methods)

สำหรับแบบการตรวจจับความสมมาตรของพื้นผิว นั้นใช้ได้กับสิ่งของที่มีลักษณะสมมาตรเท่านั้น โดยประสิทธิภาพจะลดลงตามความไม่สมมาตรของวัตถุนั้น [11] ส่วนวิธีการที่เป็นที่นิยมมากอันหนึ่งคือ อาศัยลักษณะทางกายภาพที่คงตัว (geometric invariance-based) [12] แต่ต้องทำงานซับซ้อนแม้ว่าจะเป็นพื้นผิวทั่วไป [13] เทคนิคแผนที่ระยะทาง ใช้วิธีการคำนวณระยะทางระหว่างจุดจำเป็นต้องคำนวณซ้ำหลายรอบเพื่อหาจุดที่ใกล้กันมากที่สุด วิธีการนี้เหมาะสมมากสำหรับงานที่มีข้อมูลภาพไม่สมบูรณ์แต่ทำงานได้ค่อนข้างช้าเมื่อจำเป็นต้องมีการแปลงมากกว่าแบบตายตัว เช่น similarity, affine และ weak perspective โดยที่นักวิจัยของ Pelizzari [14] ได้พัฒนาวิธีการคำนวณแบบอัตโนมัติสำหรับการจัดภาพสามมิติที่ได้จากเครื่องสร้างภาพคนละเครื่อง โดยคำนวณค่าการแปลงจากค่าผิดพลาดที่น้อยที่สุดที่คำนวณได้จากค่ากำลังสองเฉลี่ยระหว่างภาพ MRI และ CT และจำเป็นต้องกำหนดตำแหน่งซ้อนทับภาพแบบคร่าว ๆ ไว้ก่อนด้วยเพื่อป้องกันความผิดพลาดที่อาจเกิดขึ้นจากรูปร่างที่คล้ายกัน ส่วนที่มของ Grimson [15] ได้พัฒนาการทำงานแบบอัตโนมัติที่ไม่ต้องกำหนดจุดเริ่มต้นของการคำนวณให้ โดยเขาได้ใช้วิธีการสุ่มจุดเริ่มต้นขึ้นมาแทน

Jian และ Vemuri [16] ได้ใช้ Gaussian mixture model (GMM) ซึ่งมองว่าจุดในพิกัดสามมิติของแต่ละภาพนั้น จะถูกแทนที่ด้วยส่วนที่ผสมกันแบบเกาส์เซียนและคำนวณหาการแปลงที่เหมาะสมจากส่วนผสมของตัวแทนเกาส์เซียนของภาพทั้งสอง วิธีการนี้ Myronenko and Song [17] ได้นำมาใช้ โดยการหา Maximum likelihood จาก centroids ของ GMM แล้วบังคับให้ centroid เคลื่อนที่พร้อม ๆ กัน (coherent) เป็นกลุ่ม และเรียกวิธีนี้ว่า Coherent Point Drift (CPD) ซึ่งเป็นหนึ่งในวิธีที่ผู้วิจัยนำมาเปรียบเทียบ

การซ้อนทับภาพแบบอาศัยหลักการทางพลังงาน (Energy-based registration) เป็นการ
ใช้หลักทางกล โดยใช้ค่าความเค้น ความเครียด และระยะทางที่เปลี่ยนแปลงไป เช่น Terzopoulos
และคณะได้นำเสนอวิธีที่เรียกว่า Active contour หรือ snakes ไว้ใน [18]

Land mark-mapping alignment เป็นการใช้พิกัดที่เด่น มาทำเป็นแผนผังเพื่อหา
ความสัมพันธ์ระหว่างพิกัดเด่นบนภาพทั้งสอง แล้วคำนวณหาค่าการแปลงระหว่างพิกัดเด่นนั้น เรา
มักใช้ลักษณะทางกายภาพมาหาพิกัดเด่นเช่น ร่องข้างจมูก สันจมูก โหนกแก้ม หางตา เป็นต้น
ดังเช่น [19-22] หรืออาจใช้จุดตัดของเส้น [23] บริเวณที่ค่าความโค้งสูงสุด (local curvature
maxima) [24-26] จุดกึ่งกลางของพื้นผิวรูปปิด (centroid of closed-boundary region) [27] จุดที่
เป็นปม [28] (knot point) และจุดหักเห (inflection point) [29] ผู้วิจัยบางคณะก็ใช้วิธีผสมกันของ
จุดเด่นหลายอย่าง ดังเช่น [30] และ [31]

มีการใช้แนวเส้นสัน (ridge line) ที่ได้จากข้อมูลภาพสามมิติเพื่อสร้าง เส้นสันสูงสุด (crest
line) [32, 33] โดยหาแนวเส้นสันจากค่าความโค้งสูงสุด Govindo และ Shecker [34] ได้ใช้ ค่าความ
โค้งของลักษณะทางกายภาพเพื่อหาความเหมือนกันระหว่างภาพสองภาพ

Jian และ Vemuri [16] ได้ใช้ Gaussian mixture model(GMM) ซึ่งมองว่าจุดในพิกัดสาม
มิติของแต่ละภาพนั้น จะถูกแทนที่ด้วยส่วนที่ผสมกันแบบเกาส์เซียนและคำนวณหาการแปลงที่
เหมาะสมจากส่วนผสมของตัวแทนเกาส์เซียนของภาพทั้งสอง วิธีการนี้ Myronenko and Song [17]
ได้นำมาใช้ โดยการหา Maximum likelihood จาก centroids ของ GMM แล้วบังคับให้ centroid
เคลื่อนที่พร้อม ๆ กัน (Coherent) เป็นกลุ่ม และเรียกวิธีนี้ว่า Coherent Point Drift (CPD) ซึ่งเป็นหนึ่ง
ในวิธีที่ผู้วิจัยนำมาเปรียบเทียบการทำงาน

ในการศึกษาครั้งนี้ได้นำเสนอวิธีการที่ผสมผสานการทำงานระหว่าง การใช้จุดเด่นของ
ลักษณะทางกายภาพ (landmark) และเป็นลักษณะทางกายภาพที่คงตัว (geometric invariance)
โดยการคำนวณหาจุดที่เป็นลักษณะทางกายภาพที่คงตัวนี้ด้วยอนุพันธ์อันดับสูงของเส้นโค้งและ
พื้นผิวซึ่งมีความไวสูงในการบอกความแตกต่าง เราได้เสนอวิธีการใหม่ขึ้นมาสองประการด้วยกัน
ประการแรกคือวิธีการหาจุดเด่นใหม่บนพื้นผิว โดยการหาจากผลลัพธ์ของสามเวกเตอร์

$$v_1 = \langle r^{(0)}(t) \times r^{(1)}(t), r^{(2)}(t) \rangle$$

และ

$$v_2 = \langle r^{(1)}(t) \times r^{(2)}(t), r^{(3)}(t) \rangle$$

ซึ่งอยู่บนเส้นโค้งเชิงพาราโบลาที่ซึ่งปริมาตรที่คำนวณได้จากเวกเตอร์เหล่านี้เป็นศูนย์ (Zero-
volume, $v_1 = 0$) และยังคงสภาพอยู่ได้ภายใต้การแปลงแบบแอฟไฟน์ (การหมุน การเปลี่ยน
ขนาด การเลื่อนตำแหน่ง และการเฉือน) และที่ตำแหน่ง (Zero-torsion, $v_2 = 0$) ซึ่งสัมพันธ์กับ
แรงบิดที่วิ่งอยู่บนเส้นโค้งเชิงพาราโบลา ณ ตำแหน่งที่การบิดเปลี่ยนกลับจากลบเป็นบวก หรือ
เปลี่ยนจากด้านบวกเป็นลบจะหมายถึงตำแหน่งที่มีการหักงอของพื้นผิวซึ่งจะคงสภาพเดิมอยู่ได้
ภายใต้การแปลงรูปร่าง ประการที่สองคือ การเสนอแนวคิดใหม่ในการหาจุดที่สอดคล้องกันของภาพ

ทั้งสองโดยการใช้ค่าอัตราส่วนของปริมาตรที่ได้จากผลลัพธ์สามเวกเตอร์ของจุดแรงบิดเป็นศูนย์ที่วิ่งอยู่บนเส้นโค้งเชิงพาราโบลา