บทที่ 2
งานวิจัยและทฤษฎีที่เกี่ยวข้อง

1. งานวิจัยที่เกี่ยวข้อง

Fukunaga และ Hostetler กล่าวว่าการประมาณความหนาแน่น (gradient estimates) อยู่บนพื้นฐานของฟังก์ชันความหนาแน่นที่มีความต่อเนื่อง สามารถประมาณได้โดยสำรวจข้อมูลตัวอย่างภายในบริเวณเล็ก ๆ รอบจุดนั้น ปัญหาโดยส่วนใหญ่ของการรู้จำรูปแบบ (pattern recognition) คือ การหาความน่าจะเป็นของความหนาแน่นข้อมูล ที่จำเป็นต้องหาวิธีการหรือเทคนิคเพื่อให้ทราบถึงรูปแบบความหนาแน่นข้อมูล และตัวแปร (parameter) ที่จำเป็นต้องใช้ในวิธีการหรือเทคนิคนั้น ๆ โดยควรกำหนดโดยพิจารณาจากพื้นฐานของข้อมูลที่ใช้ ดังนั้นเทคนิคการประมาณแบบไม่กำหนดตัวแปร (nonparametric estimation) จึงมีความเหมาะสมที่จะนำมาใช้แก้ปัญหาดังกล่าว และเป็นที่มาของการคิดค้นและนำเสนอ “mean-shift algorithm” [1]

Cheng ได้ศึกษาขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเพิ่มเติมและพิสูจน์ให้เห็นถึงความเป็นไปได้ที่จะนำมาใช้ในการจัดกลุ่มข้อมูล (clustering) และได้พิสูจน์ให้เห็นว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเป็นลักษณะที่เรียกว่า “mode-seeking process” [2] โดยกระบวนการนี้เป็นที่นิยมใช้กันอย่างกว้างขวางในงานที่เกี่ยวข้องกับภาพ ไม่ว่าจะเป็น การแบ่งส่วนภาพ (Image segmentation) การทำให้ภาพเรียบที่ยังคงความไม่ต่อเนื่องของภาพไว้ (discontinuity - preservingsmoothing) รวมทั้งการติดตามภาพ (tracking)

Comaniciu และ Meer ได้พิสูจน์ให้เห็นว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยสามารถนำมาใช้ได้อย่างมีประสิทธิภาพกับการแบ่งส่วนภาพ (Image segmentation) การทำให้ภาพเรียบ (smoothing) พร้อมทั้งได้นำเสนอการติดตามภาพเคลื่อนไหว (tracking) โดยใช้เทคนิคการนำ histogram ของภาพสีของขอบเขตที่พิจารณา (search window) มาเป็นข้อมูลในการพิจารณา [3,4] ทำให้การติดตามภาพเคลื่อนไหวมีความรวดเร็วและถูกต้องสูง หลังจากนั้นก็มีการนำเทคนิคนี้ไปประยุกต์ใช้อย่างกว้างขวาง

Carreira-Perpiñán มีการนำเสนอการเพิ่มความเร็วแบบพิจารณาตามข้อมูลที่ย้ายไปแล้วในรอบก่อนหน้า (burring mean shift algorithm) โดยมีการจัดกลุ่มของข้อมูลในแต่ละรอบและทำการตัดข้อมูลกลุ่มที่มีสมาชิกบางตัวย้ายไปตำแหน่งสูงสุดแล้วทำให้ขั้นตอนต่อไปมีจำนวนข้อมูลน้อยลง [6] การคำนวณจึงเร็วขึ้นตามไปด้วย

Edan Lerner ได้นำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยมาช่วยทำให้ภาพเรียบ (smoothing) สำหรับภาพที่มีสัญญาณรบกวนมาก (noise) ก่อนที่จะหาขอบของภาพ โดยใช้ข้อมูลตำแหน่งบนภาพและความเข้มของภาพเป็นข้อมูลในการทำให้ภาพเรียบ [7] พบว่าได้ผลเป็นอย่างดี ปัญหาที่พบในการดำเนินการเกี่ยวกับภาพคือต้องใช้จำนวนรอบในการคำนวณที่มากส่งผลให้เวลาที่ใช้ในการคำนวณมากตามไปด้วย

Mark Fashing และ Carlo Tomasi ได้ทำการศึกษาขั้นตอนวิธีการย้ายตามค่าเฉลี่ยพบว่าเป็นขั้นตอนที่มีทิศทางการย้ายจุดศูนย์กลางถูกต้อง (Bound Optimization) [8]

Praisan Padungweang, Sirapat Chiewchanwattana และ Khamron Sunat มีการนำเสนอการเพิ่มความเร็วของการย้ายตามค่าเฉลี่ยโดยใช้ขั้นตอนวิธี Resilient backpropagation (Rprop) ซึ่งเป็นขั้นตอนวิธีในการเรียนรู้ของโครงข่ายประสาทเทียม (neural network) ทำให้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยทำงานได้เร็วขึ้น [9] เนื่องจากใช้จำนวนรอบน้อยลงและยังมีความถูกต้องเมื่อเทียบกับวิธีการเดิม

Zhang, Kwok และ Tang มีการนำเสนอการลดจำนวนข้อมูลด้วยการจัดกลุ่มของข้อมูลเพื่อหาตัวแทนก่อนทำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแบบพิจารณาตามข้อมูลที่ย้ายไปแล้วในรอบก่อนหน้า [10] วิธีนี้ให้ผลเร็วแต่มีความคลาดเคลื่อนสูงของทั้งจำนวนกลุ่มข้อมูลและสมาชิกในแต่ละกลุ่ม

Carreira-Perpiñán มีการนำเสนอการเพิ่มความเร็วสำหรับการแยกส่วนประกอบของภาพสี่วิธีคือ วิธีที่หนึ่งรวมตำแหน่งของภาพที่อยู่ใกล้กันเป็นเซลเดียวกัน จากนั้นใช้วิธีการย้ายตามค่าเฉลี่ย และเมื่อเซลใดย้ายไปยังเซลที่มีการย้ายตำแหน่งไปก่อนหน้าแล้ว ก็จะหยุดคำนวณ และถือว่าเซลนั้นเป็นกลุ่มเดียวกัน วิธีที่สองหาตัวแทนของจุดที่อยู่ใกล้กันในโดเมนของภาพกำหนดโดยระยะห่างคงที่ค่าหนึ่ง แล้วดำเนินการคล้ายกับวิธีแรก ทั้งสองวิธีได้ผลเร็วมาก แต่ความคลาดเคลื่อนจะเกิดขึ้นตั้งแต่ตอนแรกที่มีการรวมกลุ่มของจุดบนภาพ และอีกสองวิธีเป็นการลดการคำนวณของฟังก์ชันหาความหนาแน่นและใช้วิธีลู่เข้าแบบกำลังสี่ทำให้จำนวนรอบในการย้ายน้อยกว่า [11] แต่ใช้การคำนวณที่สูงมากจึงทำให้ช่วยเพิ่มความเร็วไม่มากเท่าที่ควร

Luyieng Zheng, Jing taozhang และ Qianyu wang ได้มีการนำขั้นตอนวิธีการย้ายตามค่าเฉลี่ยมาใช้ในการแยกส่วนประกอบของภาพพืชผักสีเขียวจากพื้นดินและเงาของพืช แนวความคิดประกอบด้วยสองส่วนที่สำคัญ คือ การสกัดลักษณะเด่น และ การแยกส่วนประกอบภาพ โดยที่ใช้การสกัดลักษณะเด่นออกมาจากภาพเป็น 7 มิติ คือ ตำแหน่งในแนวตั้ง ตำแหน่งในแนวนอน, ความแตกต่างระหว่างความเข้มสีเขียวและสีแดง, ความแตกต่างระหว่างความเข้มสีเขียวและสีน้ำเงิน , และอีกสามมิติมาจากค่าสีในระบบ HSV ส่วนการแยกส่วนประกอบของภาพนั้นใช้ขั้นตอนวิธีการย้ายค่าเฉลี่ย และวิธีทางนิวรอลเน็ตเวิก ในการแยกส่วนของภาพออกเป็นสองส่วนคือส่วนที่เป็นสีเขียวและส่วนที่ไม่ใช่สีเขียว [12] โดยได้ศึกษาเทียบกับวิธีการอื่น ซึ่งพบว่าวิธีที่นำเสนอขึ้นนี้ได้ผลดีที่สุด

จากวรรณกรรมและงานวิจัยที่เกี่ยวข้องจะเห็นว่าขั้นตอนวิธีการย้ายตามค่าเฉลี่ยสามารถประยุกต์ใช้งานได้อย่างหลากหลาย แต่ยังมีปัญหาเกี่ยวกับการคำนวณที่ต้องใช้เวลามาก ผู้วิจัยจึงสนใจศึกษาการลดการย้ายตำแหน่งที่ซ้ำซ้อนของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเดิม ในการดำเนินการกับข้อมูลจำนวนมากที่ยังคงความถูกต้องของวิธีการเดิม ซึ่งโดยทั่วไปมีข้อมูลที่ต้องย้ายตำแหน่งไปในเส้นทางเดียวกันหรือใกล้เคียงเป็นจำนวนมาก สุดท้ายตำแหน่งเหล่านี้ก็จะย้ายไปยังตำแหน่งที่มีความหนาแน่นสูงสุดตำแหน่งเดียวกัน ดังนั้นการลดจำนวนจุดเหล่านี้จะช่วยเพิ่มความเร็วและลดความซ้ำซ้อนให้กับขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแบบเดิม

2. การประมวลผลภาพ และแบบจำลองสี

ภาพดิจิตอลประกอบด้วยตารางสี่เหลี่ยมเล็กๆ เรียกว่า พิกเซล (pixel) แต่ละพิกเซลจะมีค่าความเข้มของสีต่าง ๆ ระบุโดยแบบจำลองของสีในแต่ละแบบ เช่น แบบจำลองสีแบบ RGB, แบบจำลองสีแบบ HSV เป็นต้น

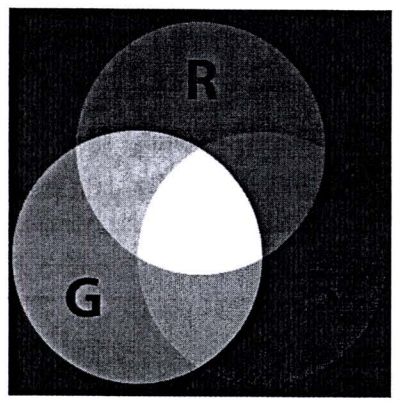
2.1 การประมวลผลภาพ (Image Processing) [13]

การประมวลผลภาพ เรียกอีกอย่างหนึ่งว่า Image Processing หมายถึง การใช้ขั้นตอน หรือกรรมวิธีใดๆ มากระทำกับภาพ โดยมีวัตถุประสงค์เพื่อปรับปรุงคุณภาพของภาพให้ได้ภาพใหม่ที่มีคุณสมบัติตามต้องการ เช่น ความคมชัด หรือการแยกบริเวณที่สนใจออกจากบริเวณอื่น ซึ่งสามารถนำมาใช้กับงานรู้จำลักษณะ หรืองานตามจุดประสงค์ที่ต้องการได้ดีขึ้น แต่ละภาพจะประกอบไปด้วยองค์ประกอบต่าง ๆ รวมถึงสิ่งที่ไม่ต้องการ และไม่จำเป็นในการประมวลผลภาพในงานที่เราจะนำไปใช้ ซึ่งรวมไปถึงสัญญาณรบกวน (Noise) ทำให้ผลลัพธ์ที่ได้ไม่เป็นไปตามความต้องการเท่าที่ควร ดังนั้นการประมวลผลภาพสามารถกระทำกับภาพได้ดังนี้ คือ กำจัดสัญญาณรบกวน ทำภาพให้ชัด (High Pass Filter) หรือ เรียบขึ้น (Low pass Filter) หาขอบภาพ (Edge Detection) แปลงภาพเป็น Binary Image และแยกส่วนประกอบภาพ (Image Segmentation) เป็นต้น

หลังจากนั้นเราสามารถนำภาพไปใช้ตามจุดประสงค์ที่เราต้องการ เช่น การรู้จำใบหน้า รู้จำวัตถุ แยกประเภท หรือรู้จำตัวอักษร เป็นต้น

2.2 แบบจำลองสี RGB

เป็นระบบที่เกิดขึ้นจากการผสมกันระหว่างสีของแสงที่เป็นแม่สีหลัก (primary color) เข้าด้วยกัน คือ แสงสีแดง แสงสีเขียวและแสงสีน้ำเงิน (Red Green Blue, RGB) [14, 15] ผลจากการผสมทำให้เกิดเป็นสีใหม่ ได้แก่ สีฟ้า สีม่วงแดง และสีเหลือง (cyan, magenta, yellow) และหากสีหลักทั้งสามที่มีความเข้มสูงสุดผสมเข้าด้วยกันจะได้เป็นแสงขาว การแสดงผลสีในเครื่องคอมพิวเตอร์ส่วนใหญ่จะใช้การนำความเข้มของแม่สีหลักมาผสมเพื่อให้เกิดเป็นสีอื่น ๆ ซึ่งเป็นหลักที่ว่าด้วยการรับรู้สีสันจากแสงสว่างที่เรามองเห็นด้วยตาเปล่า ทั้งนี้เป็นผลจากการผสมกันของแม่สี โดยมีอัตราส่วนความเข้มของแม่สีแต่ละตัวเป็นตัวแปรสำคัญทำให้ปรากฏสีสันต่าง ๆ ที่หลากหลายในธรรมชาติ ซึ่งแต่ละแม่สีอาจมีค่าความอ่อนแก่เป็นตัวเลขในช่วงระหว่าง 0 ถึง 1 หรืออาจนำมาคำนวณแปลงเป็น ช่วงตัวเลขที่มีค่าตั้งแต่ 0 จนถึง 255



ภาพที่ 1 การผสมสีของแบบจำลองสี RGB

เนื่องจากแบบจำลองสี RGB เป็นการรวมกันของความสว่างและสี เพื่อตัดองค์ประกอบของความสว่างทิ้ง คือ เมื่อภาพในระบบสี RGB เป็นสีเดียวกันแต่ความสว่างต่างกันจะถูกกำหนดให้เป็นสีเดียวกันเกิดเป็นระบบนอร์มอลไลซ์ RGB (normalized RGB) ซึ่งหาค่าได้ดังนี้

$$r = \frac{R}{(R+G+B)}, \quad g = \frac{G}{(R+G+B)}, \quad b = \frac{B}{(R+G+B)}$$

เช่น ค่าในระบบสี RGB (220, 0, 0) สามารถแปลงไปสู่ rgb ได้ (1, 0, 0) และค่าในระบบสี RGB (28, 0, 0) แปลงไปสู่ rgb ได้ (1, 0, 0) เช่นกัน ค่าทั้งสองจะมีเฉพาะองค์ประกอบสีแดงที่มีค่าความเข้มแสงต่างกัน ส่วนในระบบ rgb นั้นไม่สนใจองค์ประกอบการพิจารณาความสว่างแสงโดยถือว่า 2 ค่านี้เป็นค่าเดียวกัน

2.3 แบบจำลองสี HSV (Hue Saturation Value)

เป็นการพิจารณาสีโดยใช้ HSV ซึ่ง H (Hue) คือ ค่าสี, S (Saturation) คือความอิ่มตัวสี และ V (Value) คือความสว่างของสี [16] สามารถนำมาสร้างเป็นแบบจำลองรูปทรงกรวยซึ่งบรรจุค่าสี ความอิ่มตัวสี และความสว่างของสี โดยส่วนบนจะสว่างที่สุดและส่วนล่างจะมืดที่สุด จะพบว่าเมื่อองค์ประกอบทั้งสามถูกนำมาใช้ร่วมกันก็จะทำให้เกิดสีสันอื่น ๆ ขึ้นมา ตัวอย่างเช่น หากพิจารณาเฉพาะแม่สีแดงจะพบว่าเมื่อทำสีแดงให้สว่างที่สุด สีแดงจะกลายเป็นสีขาว และเมื่อทำให้มืดที่สุดจนถึงสีดำ เงื่อนไขเหล่านี้ถูกนำมาใช้กับสีทุกสีที่เกิดขึ้นจากการนำแม่สีแสงมาผสมในอัตราส่วนต่าง ๆ กันเกิดเป็นสีสันต่าง ๆ สามารถคำนวณได้จากระบบสี RGB ได้ดังนี้

$$H = \frac{240B_h + 120G_h}{B_h + G_h}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}$$

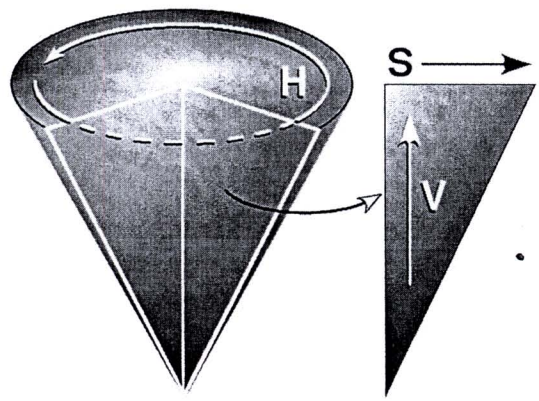
$$V = \max(R,G,B)$$

โดยที่

$R_h = R - \min(R, G, B)$

$G_h = G - \min(R, G, B)$

$B_h = B - \min(R, G, B)$



ภาพที่ 2 แบบจำลองสี HSV

3 การแยกส่วนประกอบภาพ (Image segmentation)

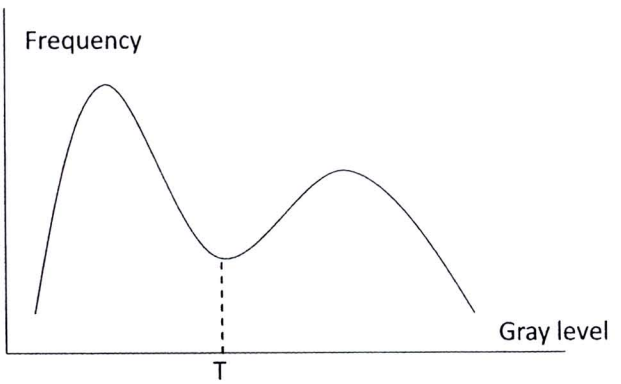
การแยกส่วนประกอบภาพมีจุดประสงค์เพื่อแยกองค์ประกอบต่าง ๆของรูปภาพออกจากกันตามลักษณะสำคัญที่เราพิจารณา ซึ่งขั้นตอนนี้เป็นขั้นตอนสำคัญของการวิเคราะห์ภาพ [17] ประโยชน์ของการแยกส่วนประกอบภาพ คือ ลดจำนวนข้อมูลในรูปภาพที่ไม่จำเป็นในการวิเคราะห์ พร้อมกับจัดระเบียบข้อมูลในรูปภาพให้เป็นกลุ่มได้ดีขึ้นและแสดงข้อมูลในรูปที่เข้าใจง่ายขึ้น โดยในการแยกองค์ประกอบของภาพสามารถ

แยกออกเป็นสองแบบ คือ การแยกองค์ประกอบโดยดูจากความไม่ต่อเนื่อง (Discontinuity) ของคุณสมบัติของพิกเซลบริเวณรอยต่อระหว่างวัตถุในภาพกับฉากหลัง และการแยกองค์ประกอบตามความเหมือนกัน (Similarity) ของคุณสมบัติของพิกเซลในรูปภาพ

3.1 การแยกองค์ประกอบโดยดูจากความไม่ต่อเนื่อง (Discontinuity)

เป็นวิธีพื้นฐานสำหรับการแยกส่วนประกอบภาพ โดยการพิจารณาความสว่างของภาพสำหรับภาพระดับสีเทา ความแตกต่างของสีสำหรับภาพสี ลักษณะพื้นผิวและขอบวัตถุก็เป็นส่วนที่ทำให้การแยกส่วนประกอบภาพได้สะดวกยิ่งขึ้น วิธีนี้จะพิจารณาความเข้มของจุดต่าง ๆ ภายในภาพ ซึ่งสามารถใช้การกำหนดค่าขีดแบ่ง (Threshold) ของส่วนประกอบที่เป็นความเข้มหรือสีของภาพ การพิจารณาขีดแบ่งที่สามารถแยกความต่างของวัตถุ และพื้นหลังมี 2 แบบ คือ การหาค่าขีดแบ่งเชิงเดี่ยว (Single threshold) และ การหาค่าขีดแบ่งแบบหลายเชิง (Multiple threshold) [16,18]

การหาค่าขีดแบ่งเชิงเดี่ยว (Single threshold) เหมาะสำหรับภาพที่มีความเข้มคงที่เมื่อเทียบกับพื้นหลัง ซึ่งภาพเหล่านี้จะมีความเข้มของวัตถุที่สามารถแยกออกจากพื้นหลังได้อย่างชัดเจน ซึ่งมีความเข้ม 2 ระดับ คือ ความเข้มของวัตถุ และพื้นหลัง การแยกส่วนประกอบของภาพสามารถทำได้โดยการกำหนดค่า Threshold ส่วนใหญ่แล้วการเลือกค่า Threshold จะขึ้นอยู่กับ Histogram ของภาพดังภาพที่ 3 การแบ่งแยกวัตถุออกจากพื้นหลังทำโดยหาค่าขีดแบ่ง T ที่สามารถแบ่ง 2 กลุ่มดังกล่าวออกจากกัน โดยการพิจารณาค่าขีดแบ่งจากฮีสโทแกรมที่อยู่ที่จุดต่ำสุดที่อยู่ระหว่างจุดสูงสุดของวัตถุ และพื้นหลัง



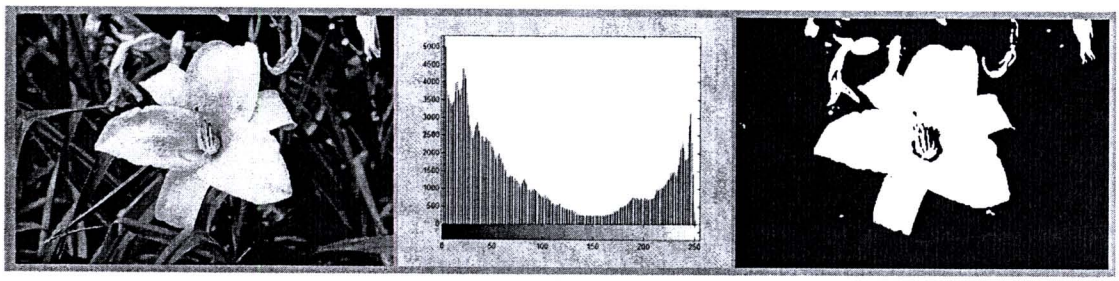
ภาพที่ 3 ค่าขีดแบ่งเชิงเดี่ยวจากฮีสโทรแกรมของความเข้มระดับสีเทา

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

เมื่อ g(x,y) เป็นข้อมูลภาพ ณ ตำแหน่ง x,y

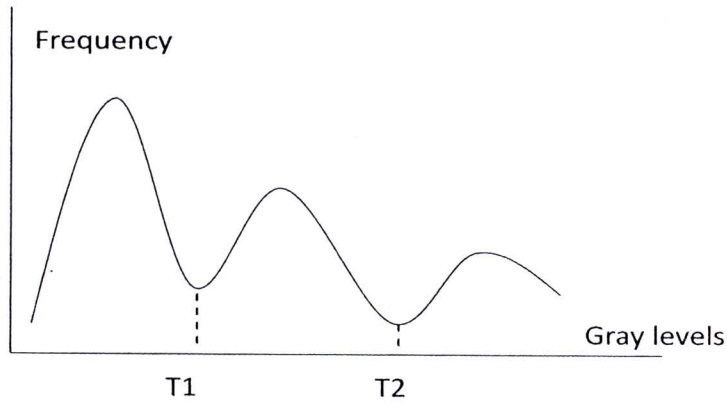
T เป็นค่า Threshold

จากนั้นกำหนดจุดภาพ f(x,y) > T เป็นจุดภาพของวัตถุ ส่วนจุดอื่น ๆ ที่เหลือเป็นจุดภาพพื้นหลังดังแสดงตัวอย่างในภาพที่ 4



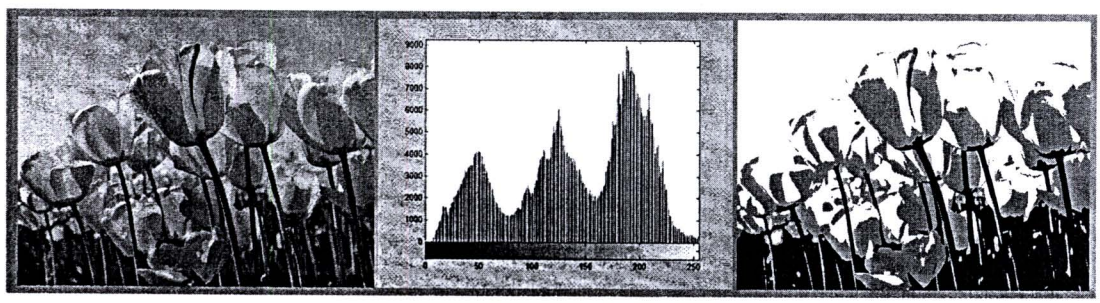
ภาพที่ 4 การแยกส่วนประกอบภาพโดยใช้ค่าขีดแบ่งเชิงเดี่ยว (Single threshold)

การหาค่าขีดแบ่งแบบหลายเชิง (Multiple threshold) สำหรับภาพที่ประกอบด้วยหลาย ๆ วัตถุ สามารถทำการแยกส่วนประกอบภาพได้โดยการกำหนดค่า Threshold หลายค่า ค่าขีดแบ่งเมื่อรูปภาพมีฮีสโทแกรมแบบ 3 ฐานนิยม แสดงดังภาพที่ 5



ภาพที่ 5 ค่าขีดแบ่งเชิงแบบหลายเชิงจากฮีสโทรแกรมของค่าระดับเทา

จำแนกจุดภาพ (x,y) ว่าเป็นจุดภาพของวัตถุแรกถ้า $T_1 < f(x,y) < T_2$ และเป็นจุดภาพของวัตถุที่สองถ้า $f(x,y) > T_2$ และเป็นจุดภาพพื้นหลังถ้า $f(x,y) < T_1$ ดังแสดงตัวอย่างในภาพที่ 6



ภาพที่ 6 การแยกส่วนประกอบภาพโดยใช้ค่าขีดแบ่งแบบหลายเชิง (Multiple threshold)

ภาพที่ 7 ขบวนการรู้จำแบบ (pattern recognition)

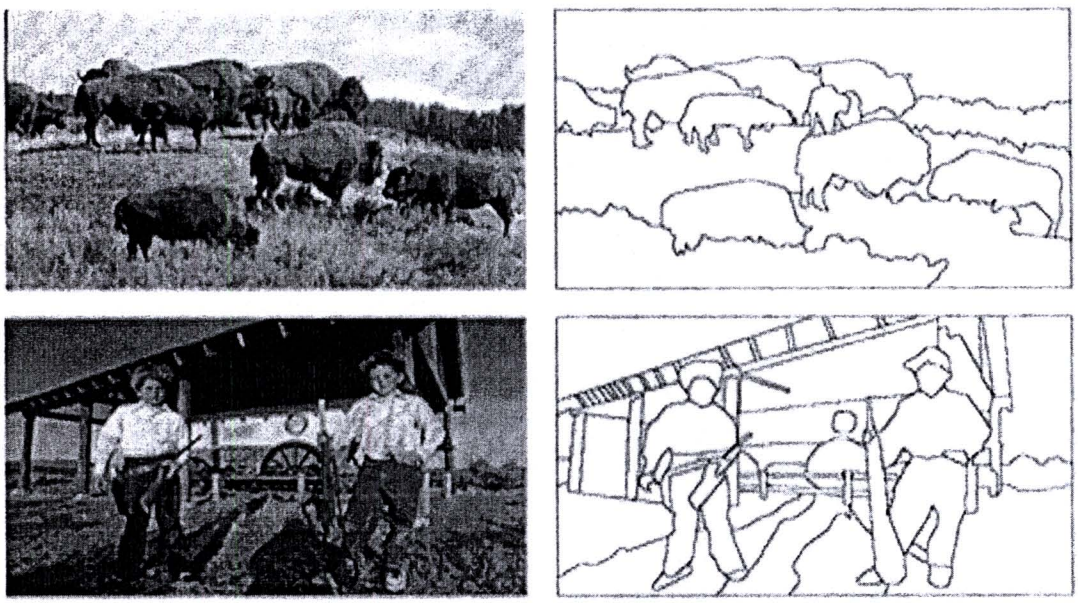
จากภาพที่ 7 เป็นตัวอย่างองค์ประกอบของการวิเคราะห์ภาพ (Element of Image Analysis) ซึ่งขั้นต้นเป็นการได้มาซึ่งภาพและการปรับปรุงภาพให้ดีขึ้น ขั้นกลางเป็นการแยกองค์ประกอบและการดึงเอาลักษณะที่สำคัญออกมา และขบวนการขั้นสูงเป็นการตีความหมายและการรู้จำภาพ

3.2 การแยกองค์ประกอบตามความเหมือนกัน (Similarity)

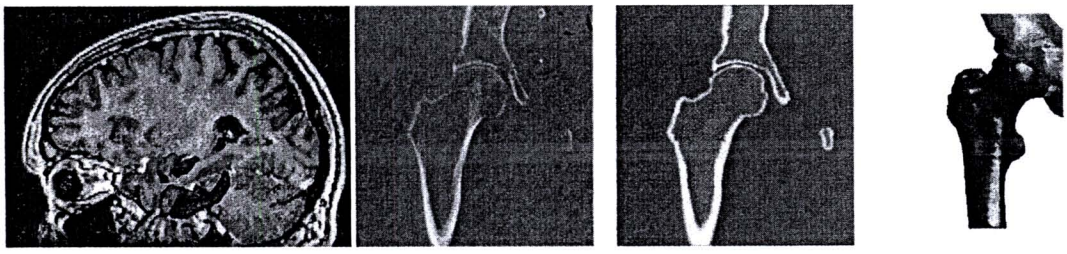
วิธีนี้สามารถใช้ขั้นตอนวิธีในการจัดกลุ่มตำแหน่งต่าง ๆ ของภาพให้อยู่ในกลุ่มเดียวกัน และแทนแต่ละกลุ่มด้วยค่าสีที่เหมือนกัน ซึ่งวิธีนี้กำลังเป็นที่นิยมเพราะมีความยืดหยุ่นและประยุกต์ใช้ได้หลากหลาย สามารถดำเนินการก่อนทำการจัดกลุ่ม เพื่อให้บริเวณที่มีความคล้ายกันแยกตัวไปรวมกันในบริเวณที่มีความหนาแน่นของกลุ่มเดียวกันมาก ก่อนที่จะมีการจัดกลุ่มบริเวณเหล่านั้นเข้าด้วยกัน ซึ่งจะแยกส่วนประกอบของภาพที่มีลักษณะใกล้เคียงกันเป็นกลุ่มเดียวกัน โดยแต่ละกลุ่มจะเป็นข้อมูลที่มีความสัมพันธ์กันมากที่สุดแสดงตัวอย่างในภาพที่ 8

ภาพที่ 8 การแยกส่วนประกอบภาพโดยการจัดกลุ่มบริเวณที่เหมือนกัน

การแยกส่วนประกอบภาพสามารถนำมาประยุกต์ใช้กับงานต่าง ๆ เช่น การหาขอบของภาพดังแสดงในภาพที่ 9 การนำมาใช้ในการแยกส่วนบริเวณที่ผิดปกติของภาพ CT scans เพื่อช่วยให้แพทย์วินิจฉัยโรคได้ง่ายและถูกต้องมากขึ้นดังแสดงในภาพที่ 10



ภาพที่ 9 การประยุกต์การแยกส่วนประกอบภาพนำมาใช้ในการหาขอบของภาพ
(ที่มา: Berkeley segmentation database)



ภาพที่ 10 ตัวอย่างการใช้การแยกส่วนประกอบภาพในการช่วยแพทย์วินิจฉัยโรค และสร้างแบบจำลองกระดูกขาและกระดูกเชิงกรานจากการแยกส่วนประกอบของภาพหลายภาพ

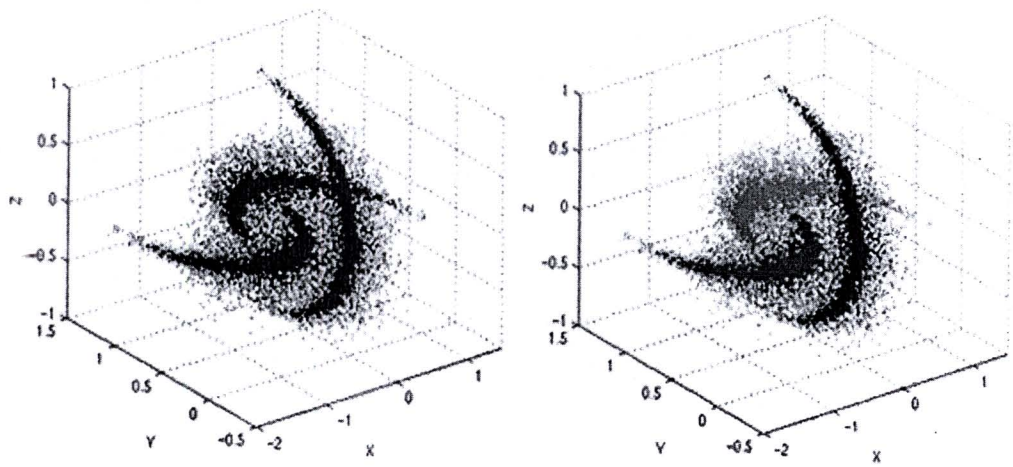
4 การจัดกลุ่มข้อมูล

การจัดกลุ่มข้อมูล (Clustering)* [19] เป็นวิธีการวิเคราะห์ข้อมูล ซึ่งใช้ในการเรียนรู้ของเครื่อง การทำเหมืองข้อมูล โดยจะแบ่งชุดข้อมูลซึ่งโดยทั่วไปเป็นเวกเตอร์ ออกเป็นกลุ่ม (cluster) นำข้อมูลที่มีคุณลักษณะเหมือนกัน หรือคล้ายกันจัดไว้ในกลุ่มเดียวกัน ขั้นตอนวิธีที่ใช้ในการแบ่งกลุ่มจะอาศัยความเหมือน (similarity) โดยคำนวณจากการวัดระยะระหว่างเวกเตอร์ของข้อมูล โดยใช้การวัดระยะแบบต่าง ๆ เช่น การวัดระยะแบบ

ยูคลิด (Euclidean distance) การวัดระยะแบบแมนฮัตตัน (Manhattan distance) การวัดระยะแบบเชบิเชฟ (Chebychev distance) เป็นต้น

การแบ่งกลุ่มข้อมูลจะแตกต่างจากการแบ่งประเภทข้อมูล (Classification) โดยจะแบ่งกลุ่มข้อมูลจากความคล้าย โดยไม่มีการกำหนดประเภทของข้อมูลไว้ก่อน จึงกล่าวได้ว่าการจัดกลุ่มข้อมูลเป็นการเรียนรู้แบบไม่มีผู้สอน ขั้นตอนวิธีการแบ่งกลุ่ม ได้แก่ k-Means clustering, Hierarchical clustering, Self-Organizing Map (SOM), การใช้คุณสมบัติการเชื่อมต่อกันของกราฟ (Graph connected component) และมีการใช้ mean shift algorithmในการแบ่งกลุ่มข้อมูลด้วย

การแบ่งกลุ่มข้อมูลอาจใช้เป็นขั้นตอนเบื้องต้นของการวิเคราะห์ข้อมูล เพื่อช่วยในการลดขนาดข้อมูลก่อนที่จะนำไปวิเคราะห์ด้วยวิธีการอื่นต่อไป อัลกอริทึมในการจัดกลุ่มข้อมูลโดยทั่วไปแบ่งได้เป็น 2 ประเภท คือ การแบ่งแบบตัดเป็นส่วน (partition) เป็นการแบ่งที่ทำเพียงครั้งเดียว และ การแบ่งแบบเป็นลำดับขั้น (hierarchical) จะมีการแบ่งกลุ่มจากกลุ่มย่อยที่ถูกแบ่งไว้ก่อนหน้านั้นซ้ำหลายครั้ง โดยจะมี 2 ลักษณะคือ แบบล่างขึ้นบน (bottom-up) หรือ เป็นการแบ่งแบบรวมกลุ่มจากกลุ่มย่อยให้ใหญ่ขึ้นไปเรื่อย ๆ โดยเริ่มจากกลุ่มเล็กสุดคือในแต่ละกลุ่มมีข้อมูลเพียงตัวเดียว และ แบบบนลงล่าง (top-down) หรือ เป็นการแบ่งแบบกลุ่มจากกลุ่มใหญ่ให้ย่อยไปเรื่อย ๆ โดยเริ่มจากกลุ่มใหญ่ที่สุด คือกลุ่มเดียวมีข้อมูลทุกตัวอยู่ในกลุ่ม

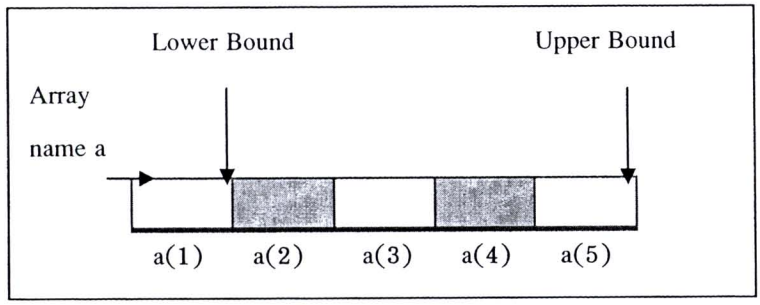


ภาพที่ 11 การจัดกลุ่มข้อมูล 3 กลุ่มใน 3 มิติโดยใช้ mean shift algorithm

5 โครงสร้างข้อมูลแบบอาร์เรย์และกราฟ

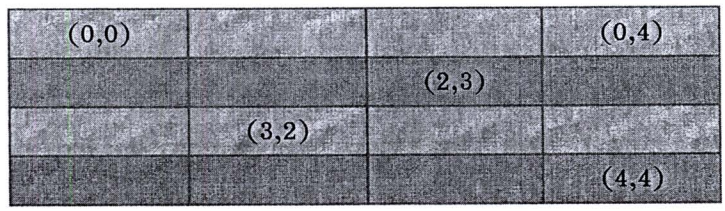
5.1 อาร์เรย์

อาร์เรย์ (Array) หรือแถวลำดับ คือการรวมกลุ่มของตัวแปรที่สามารถใช้ตัวแปรชื่อเดียวแทนข้อมูลสมาชิกได้หลาย ๆ ตัวในคราวเดียวกัน ด้วยการใช้เลขดรรชนี (Index) หรือซับสคริปต์ (Subscript) เป็นตัวอ้างอิงตำแหน่งสมาชิกบนแถวลำดับนั้น ๆ โครงสร้างข้อมูลแบบอาร์เรย์นั้นจัดเป็นโครงสร้างข้อมูลพื้นฐานที่เข้าใจง่ายที่สุดเมื่อเทียบกับบรรดาโครงสร้างข้อมูลชนิดอื่น ๆ และจัดเป็นโครงสร้างข้อมูลพื้นฐานที่สำคัญต่อโครงสร้างอื่น ๆ อีกด้วย [20]



ภาพที่ 12 ตัวอย่างรายละเอียดของอาร์เรย์หนึ่งมิติ

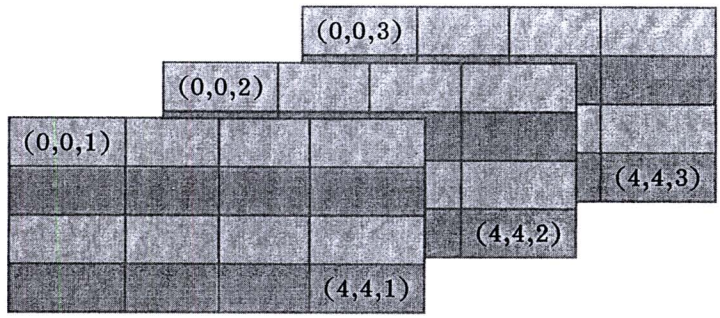
การจัดเก็บอาร์เรย์ในหน่วยความจำ อาร์เรย์ที่จัดเก็บอยู่ในหน่วยความจำคอมพิวเตอร์จะมีลักษณะเป็นลำดับต่อเนื่องกัน และใช้เนื้อที่ในการจัดเก็บข้อมูลสมาชิกของแต่ละตัวในขนาดเท่า ๆ กัน รวมถึงสมาชิกในอาร์เรย์ทุกตัวต้องเป็นข้อมูลชนิดเดียวกัน (Homogenous Elements) เราสามารถเข้าถึงข้อมูลทุกตำแหน่งในอาร์เรย์ได้โดยตรง สำหรับการเข้าถึงตำแหน่งสมาชิกในลำดับแรกกับตำแหน่งสมาชิกในลำดับท้าย ๆ จะใช้เวลาเท่ากันกล่าวคือ เวลาในการเข้าถึงข้อมูลในอาร์เรย์จะเป็นไปตามฟังก์ชั่น O(1) ซึ่งเป็นค่าคงที่ที่ไม่ขึ้นกับขนาดของข้อมูลแต่อย่างใด อย่างไรก็ตาม สำหรับในกรณีอื่น ๆ อย่างเช่น การค้นหาข้อมูลในอาร์เรย์ การหาค่าต่ำสุดหรือการหาค่าสูงสุดในอาร์เรย์ เวลาที่ใช้จะเป็นในลักษณะเชิงเส้นซึ่งเป็นไปตามประสิทธิภาพของฟังก์ชั่น O(n)



ภาพที่ 13 อาร์เรย์สองมิติและการระบุตำแหน่ง

จากภาพที่ 13 เป็นอาร์เรย์สองมิติ (Two Dimension Array) โครงสร้างอาร์เรย์สองมิติจะมีรูปแบบตารางที่ประกอบด้วยแถว (Row) และคอลัมน์ (Column) ในเชิงคณิตศาสตร์ คือ เมทริกซ์ (Matrix) การอ้างอิงข้อมูลในอาร์เรย์สองมิติจึงต้องระบุตำแหน่งแถวและคอลัมน์อาร์เรย์สองมิติมักถูกนำมาใช้งานอย่างแพร่หลาย อันเนื่องมาจากการจัดเก็บข้อมูลโดยทั่วไปมักอยู่ในรูปแบบของตารางสองมิติ

อย่างไรก็ตาม การจัดเก็บอาร์เรย์สองมิติในหน่วยความจำสามารถจัดเก็บได้ 2 วิธีด้วยกันคือ หนึ่งคือจัดเก็บด้วยการเรียงแถวเป็นหลัก (Row Major Order) และสองคือ การจัดเก็บด้วยการเรียงคอลัมน์เป็นหลัก (Column Major Order) ในกรณีการจัดเก็บอาร์เรย์สองมิติในหน่วยความจำด้วยการเรียงแถวเป็นหลัก การจัดเรียงจะเริ่มต้นตั้งแต่แถวแรกและเรียงลำดับต่อไปในแต่ละคอลัมน์จนครบ จากนั้นก็ขึ้นแถวใหม่ไปเรื่อย ๆ จนกระทั่งแถวสุดท้าย



สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

ภาพที่ 14 อาร์เรย์สามมิติและการระบุตำแหน่ง

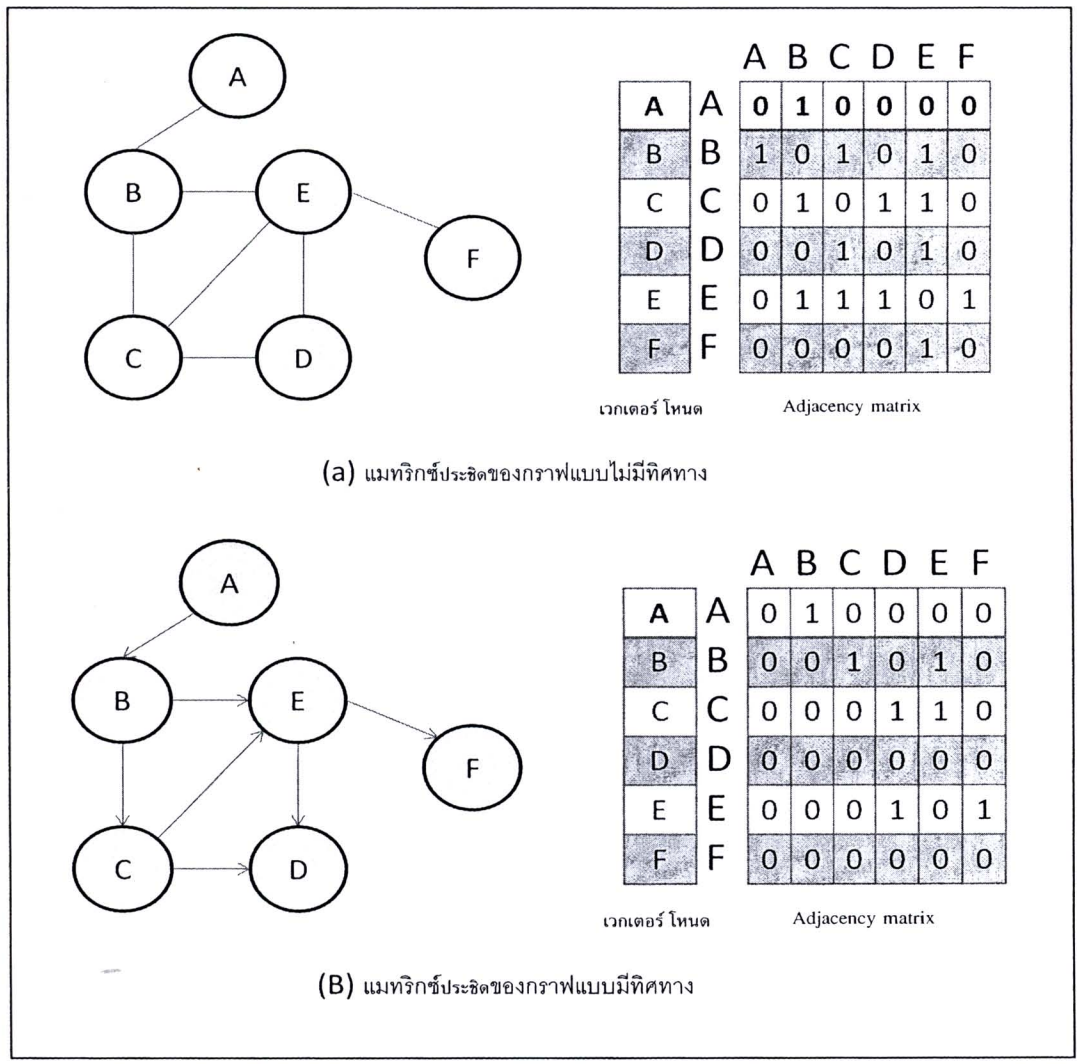
อาร์เรย์สามมิติ (Three Dimension Array) คือการนำอาร์เรย์สองมิติมาเรียงซ้อนกันหลาย ๆ ชั้น (Page) ทำให้เกิดเป็นอาร์เรย์สามมิติ นอกจากจะมีแถวและคอลัมน์แล้ว ก็จะมีความลึกเพิ่มขึ้นอีก ซึ่งความลึกนี้เองที่เกิดขึ้นจากการนำอาร์เรย์สองมิติมาเรียงซ้อนกันดังแสดงในภาพที่ 14

5.2 กราฟ

โครงสร้างการจัดเก็บข้อมูลของกราฟ (Graph Storage Structures) จำเป็นต้องจัดเก็บข้อมูลอยู่ 2 กลุ่ม คือเส้นเชื่อม และ โหนด ซึ่งโครงสร้างทั้งสองจะใช้อาร์เรย์และลิงก์ลิสต์ หรือ เมทริกซ์ประชิดในการจัดเก็บข้อมูล

เมทริกซ์ประชิด (Adjacency Matrix) ใช้เวกเตอร์ (อาร์เรย์หนึ่งมิติ) เพื่อจัดเก็บโหนด และใช้เมทริกซ์ (อาร์เรย์สองมิติ) เพื่อจัดเก็บเส้นเชื่อม หากโหนดไหนมีเส้นเชื่อมโยงระหว่างกัน ค่าของตำแหน่งนั้นในเมทริกซ์ประชิดจะเป็น 1 ถ้าไม่มีเส้นเชื่อมโยงจะมีค่าเป็น 0 ภาพที่ 15 (a) เมทริกซ์ประชิดตำแหน่ง (B, C) และ (C, B) ต่างก็มีค่าเป็น 1 เนื่องจากเป็นกราฟแบบไม่มีทิศทาง ในกรณีที่เป็นกราฟแบบมีทิศทางหรือไดกราฟ เมทริกซ์ประชิดจะมีลูกศรเป็นตัวกำหนดทิศทาง เช่นภาพที่ 15 (b) โหนด B และ C มีเส้นเชื่อมโยงระหว่างกันโดยมีทิศทางชี้จาก B ไป C ดังนั้นเมทริกซ์ประชิดตำแหน่ง (B, C) จะมีค่าเป็น 1 ในขณะที่เมทริกซ์ประชิดตำแหน่ง (C, B) จะมีค่าเป็น 0

สำนักงานคณะกรรมการวิจัยแห่งชาติ
ห้องสมุดงานวิจัย
วันที่ 20 ก.ค. 2555
เลขทะเบียน 247562
เลขเรียกหนังสือ



ภาพที่ 15 การจัดเก็บข้อมูลในกราฟโดยใช้เมทริกซ์ประชิด

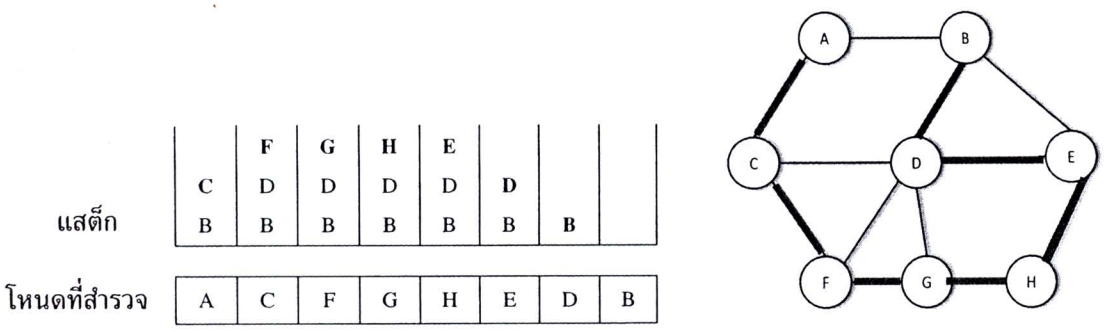
5.3 การเชื่อมต่อกันของกราฟ (Graph Connect Component)

คุณสมบัติการเชื่อมต่อกันของกราฟสามารถนำมาใช้ในการจัดกลุ่มข้อมูลได้ โดยสามารถใช้กราฟแบบไม่มีทิศทางในการหาส่วนของกราฟที่ทุก ๆ โหนดมีเส้นทาง (Path) เชื่อมต่อถึงกัน แต่ละโหนดของส่วนของกราฟนี้สามารถแทนด้วยตำแหน่งของจุด ดังนั้นส่วนของกราฟที่เชื่อมต่อถึงกัน จึงถือได้ว่าเป็นกลุ่มหนึ่งกลุ่ม การหาส่วนของกราฟที่เชื่อมต่อกันนี้สามารถนำมาใช้ในการจัดกลุ่มข้อมูลหลังจากดำเนินการตามขั้นตอนการย้ายค่าเฉลี่ยแล้ว ซึ่งตำแหน่งสุดท้ายของข้อมูลที่อยู่กลุ่มเดียวกันจะอยู่ในบริเวณใกล้กัน ดังนั้น เมื่อให้ข้อมูลเหล่านี้เป็นโหนด ข้อมูลที่อยู่ใกล้กันในระยะที่ยอมรับได้ จะถือว่ามีเส้นเชื่อมต่อถึงกัน จากนั้นก็สามารถใช้ขั้นตอนวิธีการหาส่วนของกราฟที่เชื่อมต่อกันจัดกลุ่มข้อมูลเหล่านี้ได้ กำหนดให้ n เป็นจำนวนโหนด และ E เป็นจำนวนเส้นเชื่อมการหาส่วนของกราฟที่เชื่อมต่อถึงกัน สามารถหาได้โดยใช้ขั้นตอนวิธีแบบค้นหาเชิงลึก ซึ่งวิธีนี้จะใช้เวลาเป็น O(n+E) [21]

การค้นหาเชิงลึก (Depth first search) [22] เป็นการค้นหาที่สามารถหาส่วนที่เชื่อมต่อกันของกราฟได้ด้วย โดยเริ่มจากโหนดเริ่มต้นใด ๆ จากนั้นให้นำโหนดที่อยู่ติดกับโหนดที่กำลังสำรวจอยู่ ที่ยังไม่ได้ทำการ

สำรวจและยังไม่ได้อยู่ในแสต็ก มาเก็บไว้ในสแต็กเมื่อสำรวจโหนดนั้นเสร็จ ก็ดำเนินการพอพ (pop) ตัวบนสุดของโหนดออกมาทำการสำรวจ แล้วพุชโหนดข้างเคียงทั้งหมดที่ยังไม่ได้สำรวจใส่เพิ่มในแสต็ก จากนั้นพอพตัวบนสุดออกมาสำรวจ ทำเช่นนี้เรื่อย ๆ จนกระทั่งพบโหนดที่ต้องการ หรือสำรวจครบทุกโหนด

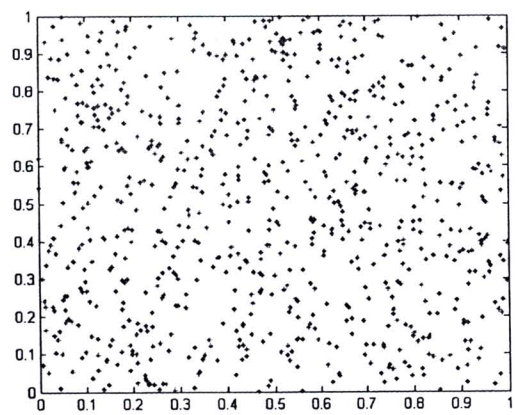
พิจารณาภาพที่ 16 โดยการสำรวจเริ่มต้นที่โหนด A และนำโหนดข้างเคียง B และ C มาเก็บไว้ในแสต็ก เมื่อสำรวจ A เสร็จ ก็พอพข้อมูลจากแสต็กออก ซึ่งเป็นโหนด C ทำการสำรวจโหนด C และนำโหนดข้างเคียงกับ C ที่ยังไม่ได้ทำการสำรวจและยังไม่ได้อยู่ในแสต็กมาใส่แสต็กนั่นคือ D และ F พุช (Push) ใส่แสต็ก ดังนั้นในแสต็กตอนนี้มี B D และ F อยู่ เมื่อสำรวจ C เสร็จ พอพ F ออกมาทำการสำรวจ แล้วนำโหนดข้างเคียงที่ยังไม่ได้สำรวจและยังไม่ได้อยู่ในแสต็กมาใส่แสต็ก ซึ่งก็คือ G ดังนั้นข้อมูลในแสต็กจะเป็น B D G ทำเช่นนี้ไปเรื่อย ๆ จนจบการทำงานก็จะได้ลำดับการสำรวจคือ (A C F G H E D B)



ภาพที่ 16 ลำดับการค้นหาเชิงลึก

6. การวิเคราะห์องค์ประกอบหลัก (PCA)

การวิเคราะห์องค์ประกอบหลัก (PCA) เป็นวิธีเบื้องต้นทางสถิติประยุกต์ในการแปลงมิติของข้อมูลไปยังมิติใหม่ แกนของระบบพิกัดใหม่จะถูกสร้างโดยการย้ายแกนเดิมไปยังพิกัดของค่าเฉลี่ยของข้อมูล (zero means) และทำการหมุนแกนของข้อมูล แกนหลักของการวิเคราะห์องค์ประกอบหลักจะอยู่ในแนวที่มีความแปรปรวนของข้อมูลสูงที่สุด แกนรองจะมีทิศทางตั้งฉากกับแกนหลัก และมีความแปรปรวนเป็นอันดับสองรองจากแกนหลัก



ภาพที่ 17 ข้อมูลที่นำ 2 มิติแรกมาพิจารณา

ตัวอย่างเช่น ข้อมูลมีขนาด 800 ข้อมูล 5 มิติ (X = (5x800)) แบ่งเป็นข้อมูล 2 กลุ่ม โดย 3 มิติแรกนั้นเป็นข้อมูลแบบสุ่มโดยใช้การกระจายตัวแบบปกติ (normal distribution) 2 มิติหลังสุ่มข้อมูลแบบการกระจายตัวปกติเช่นกันแต่ข้อมูลแยกกันชัดเจนเป็น 2 กลุ่ม กราฟของ 2 มิติแรก แสดงภาพที่ 17 จะเห็นว่าไม่เห็นลักษณะการแยกกลุ่มของข้อมูล ซึ่งโดยปกติเราไม่รู้ลักษณะธรรมชาติของข้อมูล การพิจารณาโดยใช้กราฟจึงไม่ใช่เรื่องง่ายหากข้อมูลมีจำนวนมิติมาก ดังนั้นการวิเคราะห์องค์ประกอบหลัก (PCA) จึงมีความสำคัญโดยนำองค์ประกอบหลักซึ่งโดยทั่วไปสามารถเลือกองค์ประกอบที่มิติน้อยกว่ามิติของข้อมูล มาช่วยในการวิเคราะห์ ในงานวิจัยนี้ใช้องค์ประกอบหลักที่สำคัญที่สุด 2 องค์ประกอบแรกมาช่วยในการวิเคราะห์ลักษณะการกระจายตัวของความหนาแน่นของข้อมูล

การวิเคราะห์องค์ประกอบหลัก (PCA) สามารถทำได้โดยอาศัยแนวคิดของการหา eigenvalues และ eigenvectors ซึ่งเป็นแนวความคิดของการแยกเมตริกซ์ที่ว่า

$$AU = \lambda U$$

โดยที่ทราบค่าของ A ส่วน λU ต้องการหาโดยมีเงื่อนไขว่า

A คือ เมตริกซ์จัตุรัสใด ๆ แต่ถ้าเป็น PCA แล้ว A จะเป็นเป็นเมตริกซ์ความแปรปวนร่วม (covariance matrix) ขนาด (mxm)

λ คือ eigenvalue เป็นเมตริกซ์แนวทะแยง (diagonal matrix) ขนาด (mxm)

U คือ eigenvector เป็นเมตริกซ์ของเวกเตอร์หนึ่งหน่วย ขนาด (mxm)

เวลาใช้จะเลือกเวกเตอร์ใน U ซึ่งในที่นี้คือแกนใหม่ของข้อมูล โดยใช้เวกเตอร์ตามลำดับ λ ค่าที่มากสุดโดยถือว่ามีความสำคัญสูงกว่ามาแปลงข้อมูลโดยการคูณกับข้อมูลเดิมเพื่อให้ได้องค์ประกอบหลักโดยจะอธิบายเพิ่มเติมโดยใช้ข้อมูลจากตัวอย่างด้านบนดังนี้

Covariance matrix A คือ ความแปรปรวนร่วมเป็นตัวชี้วัดขอบเขตองค์ประกอบที่ตรงกันจากสองชุดข้อมูลไปในทิศทางเดียวกัน ใช้สูตรการคำนวณความแปรปรวนร่วมดังสมการด้านล่างต่อไปนี้

$$A = Cov(X,Y) = \frac{\sum (X_i - X)(Y_i - Y)}{N} = \frac{\sum x_i y_j}{N}$$

เมื่อ

N คือ จำนวนคะแนนในแต่ละชุดข้อมูล

X คือ คือค่าเฉลี่ยของคะแนน N ในข้อมูลชุดแรก

X_i คือ คะแนนแถวที่ i ในข้อมูลชุดแรก

x_i คือ คะแนนค่าเบี่ยงเบนใน ith ในข้อมูลชุดแรก

Y คือ ค่าเฉลี่ยของคะแนน N ในข้อมูลชุดที่สอง

Y_i คือ คะแนนแถวที่ i ในข้อมูลชุดที่สอง

y_i คือ คะแนนค่าเบี่ยงเบนใน ith ในข้อมูลชุดที่สอง

A = Cov(X, Y) คือ ความแปรปรวนของคะแนนที่สอดคล้องกันในสองชุดข้อมูล

จากตัวอย่างการกำหนดข้อมูลข้างต้น (X = (5x800)) เมื่อคำนวณ A ตามสมการ Covariance matrix A แล้วจะได้เมตริกซ์ A ขนาด 5x5 มิติ จากการคำนวณที่ได้ A จะทำให้ได้ค่า eigenvectors (U) และ eigenvalues (λ) ถ้าพิจารณาที่ค่า eigenvalues จะมีค่าหนึ่งที่สูงกว่าค่าอื่นมาก ๆ จากนั้นทำการเรียงลำดับ eigenvectors ใหม่ตาม eigenvalues จากมากไปน้อย จากค่า eigenvalues สามารถบอกได้ว่าข้อมูลนี้แปลงไปแล้วเหลือ 1 มิติน่าจะเพียงพอเนื่องจากค่าที่สูงจากการคำนวณได้ต่างจากค่าอื่น ๆ มาก

Covariance matrix A

0.084	0.001	-0.005	-0.036	-0.033
0.001	0.086	0.002	0.024	-0.010
-0.005	0.002	0.081	0.036	0.046
-0.036	0.024	0.036	25.824	24.736
-0.033	-0.010	0.046	24.736	25.725

eigenvectors(U)

-0.563	-0.434	0.703	-0.002	-0.001
0.230	-0.899	-0.371	0.025	0.000
-0.794	0.047	-0.607	-0.008	0.001
-0.009	0.015	0.005	0.706	0.708
0.010	-0.016	-0.003	-0.708	0.706

eigenvalues (λ)

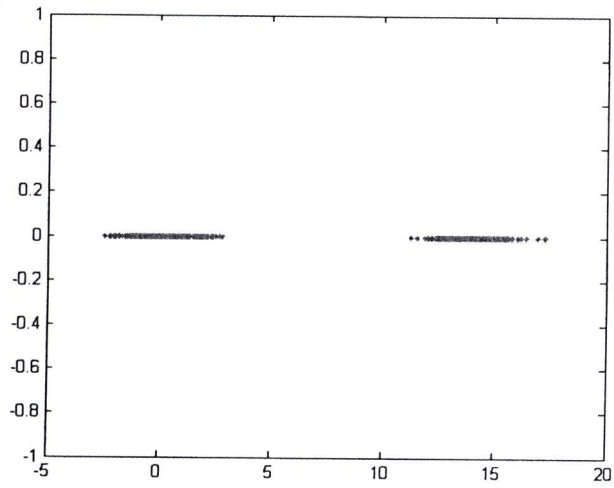
0.077	0	0	0	0
0	0.086	0	0	0
0	0	0.087	0	0
0	0	0	1.039	0
0	0	0	0	50.511

eigenvectors (U) ใหม่

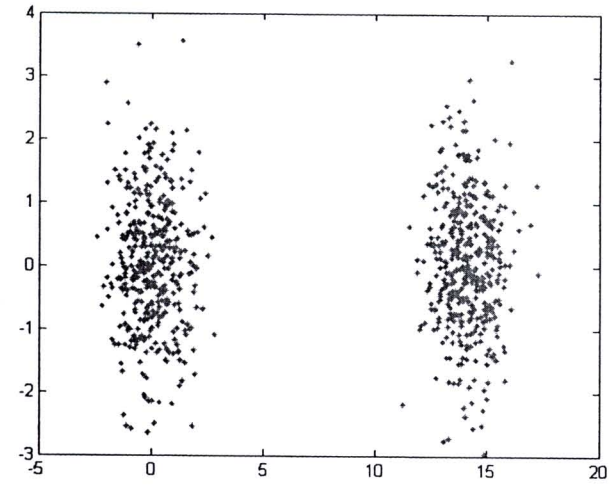
-0.001	-0.002	0.703	-0.434	-0.563
0.000	0.025	-0.371	-0.899	0.230
0.001	-0.008	-0.607	0.047	-0.794
0.708	0.706	0.005	0.015	-0.009
0.706	-0.708	-0.003	-0.016	0.010

ภาพที่ 18 Covariance matrix A, eigenvectors(U), eigenvalues (λ) และ eigenvectors (U) ใหม่

จะเห็นว่า เวกเตอร์แรกของ eigenvectors (U) ใหม่ (column แรก) จะให้ความสำคัญกับมิติที่ 4 กับ 5 สูงกว่า 3 มิติแรกมาก เมื่อนำมิติเดียวคือมิติแรกมาวาดกราฟ โดย Y = X*U(:,1) จะแสดงได้ดังภาพที่ 19 ซึ่งจะเห็นว่ามีการแยกออกเป็น 2 กลุ่มชัดเจน และเมื่อนำ 2 มิติแรกมาวาดกราฟ โดย Y=X*U(:,1:2) จะได้รูปดังด้านล่าง ซึ่งไม่ได้ช่วยให้ความหมายดีขึ้นเนื่องจากเมื่อโปรเจคลงไป ข้อมูลจะไปกองกันทั้งหมด จะเห็นว่าแต่ละมิติของ Y คือส่วนผสมของทุกมิติใน X ที่มีน้ำหนักแตกต่างกัน โดยวิธีที่หาค่า Y เรียก Principal Component Analysis (PCA) และแต่ละมิติของ Y จะเรียกว่า Principal Component



ภาพที่ 19 นำมิติแรกของ eigenvectors (U) มาใช้กับข้อมูลเดิม



ภาพที่ 20 นำมิติหนึ่งและสองของ eigenvectors (U) มาใช้กับข้อมูลเดิม

7. ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย (Mean Shift algorithm)

ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย เป็นขั้นตอนที่แต่ละจุดจะย้ายไปหาจุดที่มีความหนาแน่นสูงสุด โดยการย้ายในแต่ละรอบนั้นจะย้ายไปยังตำแหน่งของค่าเฉลี่ยแบบถ่วงน้ำหนัก แต่ละจุดก็จะดำเนินการตามขั้นตอนวิธีในลักษณะเดียวกัน จากภาพที่ 21 เป็นเส้นทางการเดินทางจากจุดเริ่มต้นของตำแหน่งที่ทำการย้ายจนถึงจุดสิ้นสุดการย้ายโดยค่าที่อยู่ใน search window จะมีค่าเป็น 1 ค่าที่อยู่นอกจะมีค่าเป็น 0 ซึ่งเป็นวิธีการย้ายตามค่าเฉลี่ยแบบหนึ่ง อย่างไรก็ตามในบทความนี้ได้ใช้การกำหนดน้ำหนักแบบเกาส์เซียนซึ่งจะมีการกำหนดให้กับทุก ๆ จุดและจะมีค่าน้อยลงเรื่อย ๆ ถ้าตำแหน่งใดอยู่ไกลจากตำแหน่งที่พิจารณามาก

กำหนดให้ข้อมูลแทนด้วยเวกเตอร์จำนวน n ข้อมูล มีขนาด m มิติ แทนด้วย $\mathbf{X} \in R^{m \times n}$ โดยที่ $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ และ $\mathbf{x}_i = [x_{1,i}, x_{2,i}, ..., x_{m,i}]^T$ การประมาณความหนาแน่นของข้อมูล ณ ตำแหน่ง x ใด ๆ สามารถคำนวณได้ดังสมการ

$$p(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K\left(\left\|(\mathbf{x} - \mathbf{x}_i)/\sigma\right\|^2\right) \tag{1}$$

โดยที่ σ เป็นค่าคงที่ และ $K(t)$ เป็นฟังก์ชันเคอร์เนล ตำแหน่งที่มีความหนาแน่นสูงสุดของบริเวณที่พิจารณานี้คือตำแหน่งที่ $\nabla p(\mathbf{x}) = 0$ ซึ่งสามารถหาได้โดยการคำนวณเพื่อย้ายตำแหน่งหลาย ๆ ครั้งโดยการย้ายตำแหน่งในแต่ละรอบสามารถคำนวณได้ดังสมการ

$$\mathbf{x}^{(\tau+1)} = f(\mathbf{x}^{(\tau)}) \tag{2}$$

โดยที่

$$f(\mathbf{x}) = \frac{\sum_{i=1}^{n} K'\left(\left\|(\mathbf{x} - \mathbf{x}_i)/\sigma\right\|^2\right)\mathbf{x}_i}{\sum_{j=1}^{n} K'\left(\left\|(\mathbf{x} - \mathbf{x}_j)/\sigma\right\|^2\right)} \tag{3}$$

เมื่อ $K'(t) = \frac{\partial K(t)}{\partial t}$ ขั้นตอนวิธีนี้เรียกว่า ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย จาก [23] ถ้าฟังก์ชันเคอร์เนลเป็นฟังก์ชันเกาส์เซียน ($K(t) = e^{-t/2}$) สมการที่ (2), (3) จะเป็นดังสมการ

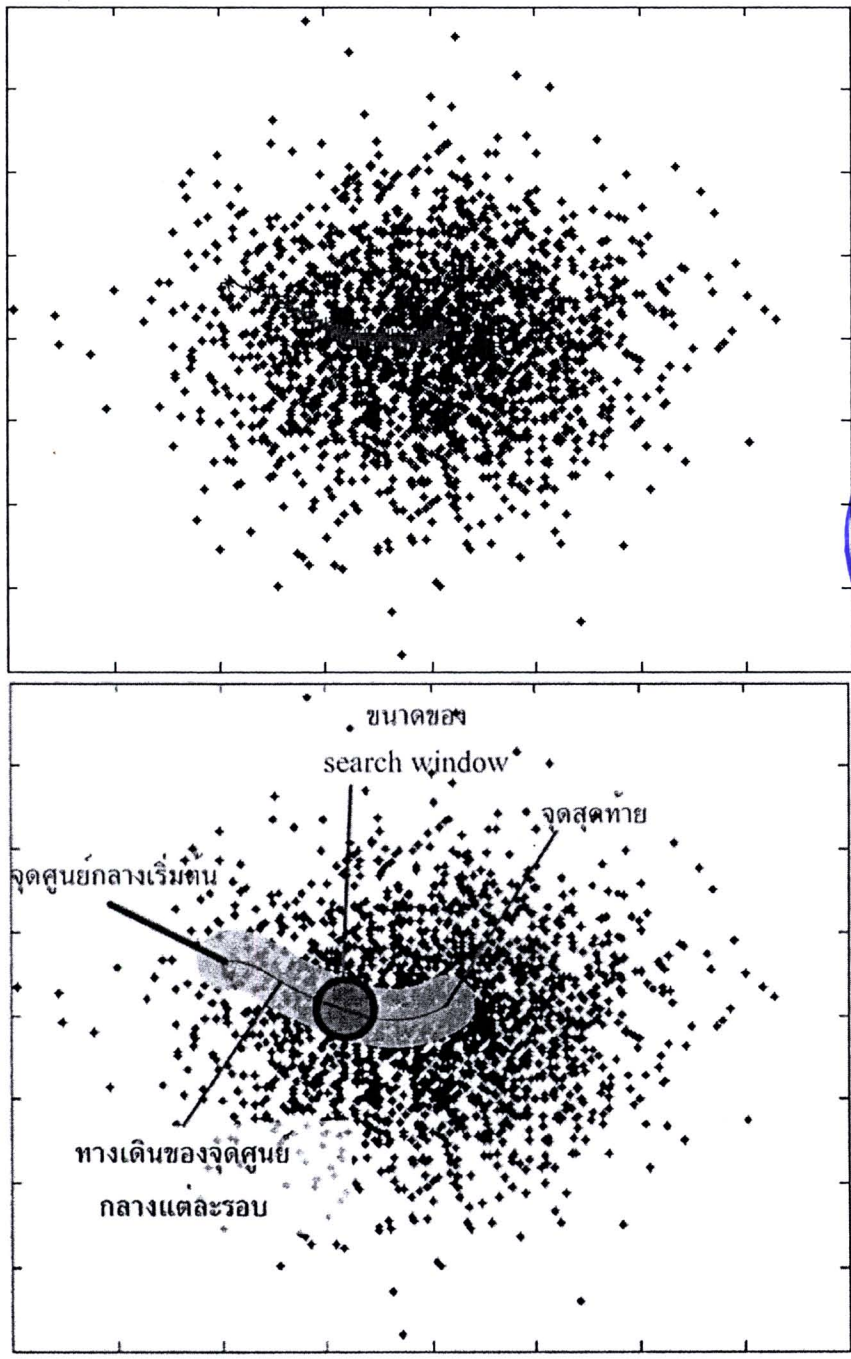
$$\mathbf{x}^{(\tau+1)} = \sum_{i=1}^{n} p(i \mid \mathbf{x}^{(\tau)})\mathbf{x}_i \tag{4}$$

$$p(i \mid \mathbf{x}^{(\tau)}) = \frac{\exp(-\frac{1}{2}\left\|(\mathbf{x}^{(\tau)} - \mathbf{x}_i)/\sigma\right\|^2)}{\sum_{j=1}^{n}\exp(-\frac{1}{2}\left\|(\mathbf{x}^{(\tau)} - \mathbf{x}_j)/\sigma\right\|^2)} \tag{5}$$

การดำเนินการแต่ละตำแหน่งในแต่ละรอบจะดำเนินการตามสมการ (5) จนกว่าไม่มีการย้ายตำแหน่ง หรือตำแหน่งนั้นย้ายน้อยกว่าค่าที่ยอมรับได้ (Threshold) ซึ่งโดยส่วนมากจะมีค่าน้อยมาก ในงานวิจัยนี้กำหนดค่าเป็น 10^{-7} เงื่อนไขที่กำหนดเป็นดังสมการ (6)

$$\| \mathbf{x}^{\tau} - \mathbf{x}^{\tau-1} \|^2 \leq threshold \tag{6}$$

ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยจะสิ้นสุดการดำเนินงานเมื่อได้ดำเนินการกับทุกตำแหน่งแล้ว หลังจากนั้นแต่ละตำแหน่งจะถูกย้ายไปที่ตำแหน่งที่มีความหนาแน่นสูงสุดในบริเวณของตำแหน่งนั้น ซึ่งก็จะสามารถดำเนินการจัดกลุ่มข้อมูลโดยใช้ขั้นตอนวิธีในการจัดกลุ่มต่อได้โดยง่าย



สำนักงานคณะกรรมการวิจัยแห่งชาติ ห้องสมุดงานวิจัย

ภาพที่ 21 แสดงจุดศูนย์กลางของ search window ในแต่ละรอบ (บน) และภาพจำลองทางเดินของจุดศูนย์กลางของ search window (ล่าง)

Mean shift algorithm

1. เลือกจุดศูนย์กลางการคำนวณเริ่มต้น x
2. คำนวณหา Sample mean **m**(x) โดย

$$\mathbf{m}(\mathbf{x}) = \frac{\sum_{\mathbf{x}_i \in \mathbf{X}} K(1/\sigma \| \mathbf{X} - \mathbf{x}_i \|^2)\mathbf{x}_i}{\sum_{\mathbf{x}_i \in \mathbf{X}} K(1/\sigma \| \mathbf{X} - \mathbf{x}_i \|^2)} \qquad (7)$$

3. กำหนดให้ m(x) เป็นจุดศูนย์กลางการคำนวณใหม่ $x^{(t+1)} = m(x^{(t)})$
4. ทำตามขั้นตอนที่ 2 และ 3 จนจุดศูนย์กลางการคำนวณไม่ย้ายหรือมีการย้ายน้อยกว่าค่าที่กำหนด
5. ทำตามขั้นตอนที่ 1- 4 จนครบทุกข้อมูล

ข้อดีของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย คือ ถูกออกแบบมาให้สามารถหาบริเวณที่มีความหนาแน่นของข้อมูลมากอย่างอัตโนมัติ มีการกำหนดค่าพารามิเตอร์น้อย คือกำหนดเพียงค่าแบนด์วิดท์ ซึ่งเป็นค่าที่กำหนดความกว้างของฟังก์ชันประมาณค่าความหนาแน่น และ สามารถประยุกต์ใช้งานได้อย่างหลากหลาย

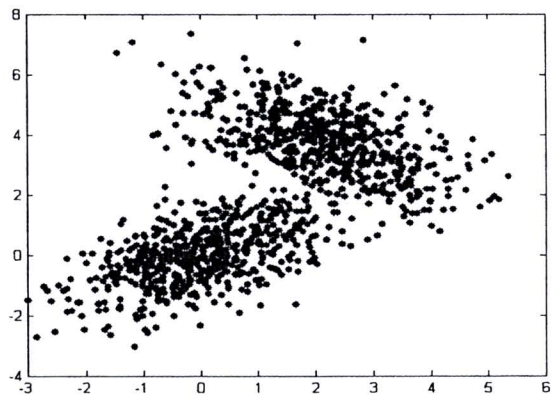
8. การใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยกับการจัดกลุ่มข้อมูล

ขั้นตอนวิธีการย้ายตามค่าเฉลี่ย สามารถนำมาใช้ในการจัดกลุ่มข้อมูลได้ ซึ่งโดยทั่วไปข้อมูลที่ต้องการจัดกลุ่มจะอยู่ในรูปของเวกเตอร์ดังนั้นสามารถใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยดำเนินการกับข้อมูลเหล่านี้ได้ โดยการเลือกแบนด์วิดท์ที่เหมาะสม ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยจะย้ายจุดที่ควรจะอยู่ในกลุ่มเดียวกันให้ชิดกันในบริเวณที่มีความหนาแน่นของกลุ่มนั้นมากที่สุด โดยไม่จำเป็นที่กลุ่มเหล่านั้นจะมีลักษณะเป็นทรงกลมหรือไม่ หลังจากนั้นก็จะสามารถจัดกลุ่มข้อมูลเหล่านี้ให้เป็นกลุ่มเดียวกันได้ง่ายขึ้น เนื่องจากข้อมูลที่อยู่กลุ่มเดียวกันจะอยู่ที่ตำแหน่งเดียวกันหรือใกล้กันมากเมื่อเทียบกับข้อมูลที่อยู่ต่างกลุ่ม ขั้นตอนวิธีที่มีประสิทธิภาพในการจัดกลุ่มข้อมูลหลังจากดำเนินการตามขั้นตอนวิธีการย้ายตามค่าเฉลี่ยแล้วขั้นตอนหนึ่ง คือ การแทนจุดต่าง ๆ ด้วยกราฟ โดยให้มีเส้นเชื่อมของจุดที่อยู่ใกล้กันจากนั้นใช้การหาส่วนของกราฟที่เชื่อมต่อกัน จะสามารถจัดกลุ่มข้อมูลเหล่านี้ได้อย่างง่ายดาย

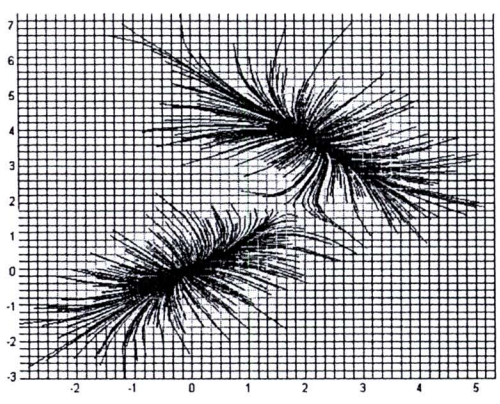
9. การใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยกับการแยกส่วนประกอบภาพ

เนื่องจากขั้นตอนวิธีการย้ายตามค่าเฉลี่ยเป็นหนึ่งในวิธีที่มีประสิทธิภาพในการรวมจุดข้อมูลที่มีความคล้ายคลึงกันไปอยู่ในบริเวณที่มีความหนาแน่นของกลุ่มเดียวกันสูง ทำให้การจัดกลุ่มง่ายและมีประสิทธิภาพ

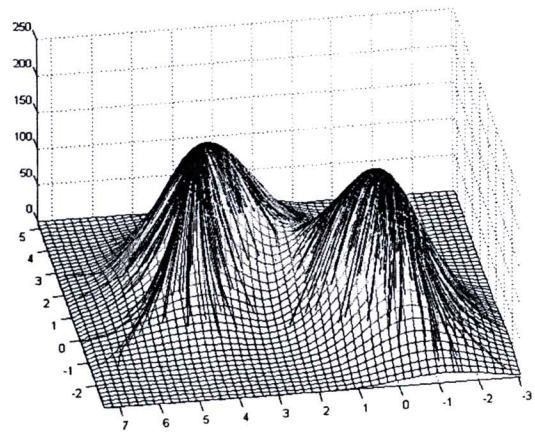
มาก ซึ่งการแยกส่วนประกอบภาพจุดประสงค์ก็เพื่อแยกส่วนที่เหมือนกันให้อยู่ในกลุ่มเดียวกัน ดังนั้นการแยกส่วนประกอบของภาพก็เปรียบเสมือนการจัดกลุ่มของตำแหน่งบนภาพนั้นเอง ดังนั้นขั้นตอนแรกในการแยกส่วนประกอบของภาพโดยใช้วิธีในลักษณะนี้ จำเป็นต้องแทนจุดต่าง ๆ บนภาพด้วยเวกเตอร์เพื่อให้สามารถใช้ขั้นตอนวิธีการย้ายตามค่าเฉลี่ยในการดำเนินการกับตำแหน่งเหล่านี้ได้ ซึ่งหลังจากนั้นสามารถแทนค่าสีของตำแหน่งเหล่านี้ด้วยค่าสีของตำแหน่งสุดท้ายของแต่ละจุดที่ได้จากการดำเนินการตามขั้นตอนวิธีการย้ายตามค่าเฉลี่ยจะได้ภาพที่แบ่งเป็นส่วน ซึ่งแต่ละส่วนคือบริเวณที่ควรจะอยู่ในกลุ่มเดียวกัน ข้อมูลที่นิยมนำมาเป็นเวกเตอร์แทนตำแหน่งบนภาพ คือ ตำแหน่งในแนวนอน ตำแหน่งในแนวตั้ง และค่าสีต่าง ๆ ทำให้ได้ข้อมูลเป็นเวกเตอร์จำนวนเท่ากับจำนวนพิกเซลในภาพ แต่ละเวกเตอร์มีขนาด 5 มิติ (x, y, R, G, B), 3 มิติ (x, y, I), (R, G, B) หรือ มิติเดียว (I) หรืออาจใช้สีในระบบสีแบบอื่น เช่น ระบบสีแบบ HSV เป็นต้น แล้วสามารถนำไปประยุกต์ใช้ในลักษณะเดียวกันกับการจัดกลุ่มข้อมูลดังที่ได้กล่าวมาแล้วข้างต้น



(1)



(2)



(3)

ภาพที่ 22 ธรรมชาติของขั้นตอนการย้ายตามค่าเฉลี่ย (1) ข้อมูล 2 กลุ่ม (2) เส้นทางเดิน ของขั้นตอนการย้าย ตามค่าเฉลี่ย (3) เส้นทางเดินของขั้นตอนการย้ายตามค่าเฉลี่ยในมุมมอง 3 มิติ

10. ปัญหาของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย

การดำเนินการของขั้นตอนวิธีการย้ายตามค่าเฉลี่ย ต้องมีการย้ายตำแหน่งของจุดแต่ละจุดไปยังจุดที่มีความหนาแน่นสูงสุดของตัวเอง โดยแต่ละจุดจะดำเนินหลายรอบขึ้นอยู่กับเงื่อนไขการหยุด สำหรับการกำหนดแบนด์วิดท์ซึ่งยังคงเป็นปัญหาเปิดที่สามารถศึกษาวิจัยเพื่อหาค่าที่เหมาะสมในแต่ละลักษณะของข้อมูลชุดนั้น ๆ โดยทุก ๆ จุดของข้อมูลต้องดำเนินการในลักษณะเดียวกันนี้ถึงแม้ว่าจุดเหล่านั้นจะอยู่ในตำแหน่งเดียวกัน หรือต้องย้ายไปในทิศทางเดียวกันกับจุดที่เคยดำเนินการไปก่อนหน้านี้แล้วก็ตาม ซึ่งจะเห็นชัดเจนจากภาพที่ 22 (2),(3) ที่แสดงเส้นทางเดินของแต่ละจุดไปยังจุดยอดแสดงให้เห็นว่า ธรรมชาติของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยจะมีจุดที่ต้องย้ายไปในเส้นทางเดียวกันหรือใกล้เคียงกันเป็นจำนวนมาก ซึ่งเป็นการทำงานที่ซ้ำซ้อนอย่างไม่จำเป็น ส่งผลให้การดำเนินการกับข้อมูลที่มีจำนวนมากต้องใช้เวลาในการดำเนินการมากตามไปด้วย

ดังนั้นผู้วิจัยจึงสนใจที่จะลดความซ้ำซ้อนในการทำงานของขั้นตอนวิธีการย้ายตามค่าเฉลี่ยนี้โดยอาศัยตัวแทนในการย้ายเพื่อลดการคำนวณการย้ายตำแหน่งข้อมูล ในงานวิจัยนี้ผู้วิจัยได้พัฒนาขั้นตอนวิธีการย้ายตามค่าเฉลี่ยความเร็วสูง เพื่อใช้สำหรับการจัดกลุ่มข้อมูลและการแยกส่วนประกอบของภาพ ซึ่งรายละเอียดจะกล่าวถึงในบทต่อไป