

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีที่เกี่ยวข้อง

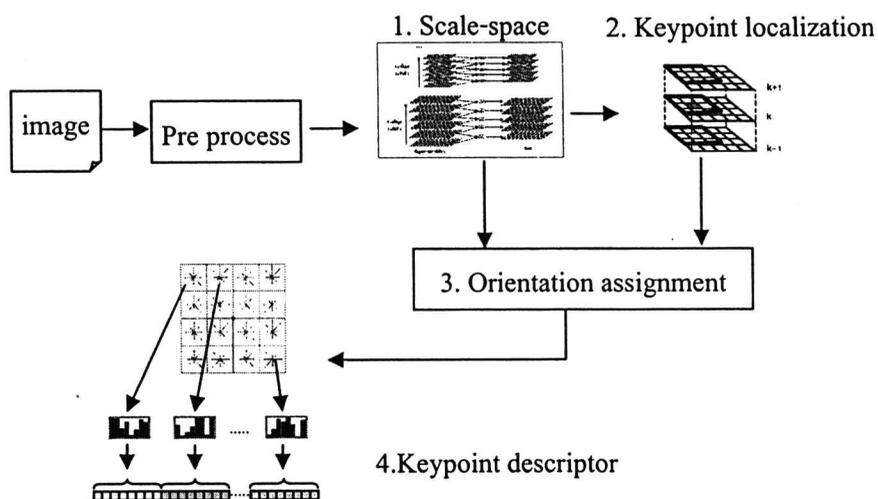
ในส่วนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องในการลดความซับซ้อนของการหาลักษณะเด่นของภาพ ด้วยวิธีการ L-SIFT โดยนำเสนอตามลำดับดังนี้

- 1.1 การหาลักษณะเด่นของภาพแบบ SIFT
- 1.2 การประมาณค่าด้วยตัวก่อกำเนิดพหุนามเป็นช่วง ชนิด p-RPP
- 1.3 การประมาณการเชิงเหตุผล
- 1.4 การหาลักษณะเด่นของภาพแบบ L-SIFT
- 1.5 การสร้างภาพมุมกว้าง (panorama)

1.1 การหาลักษณะเด่นของภาพแบบ SIFT

วิธีการ Scale Invariant Feature Transform (SIFT) เป็นวิธีการหาลักษณะเด่นแบบท้องถิ่นที่ไม่ขึ้นกับการเปลี่ยนขนาดและมุมมองของภาพ [12, 13] โดยอาศัยหลักการที่ว่า ข้อมูลจุดภาพของวัตถุเดียวกัน จะมีทิศทางของจุดภาพชี้ไปในแนวเดียวกัน ดังนั้น เมื่อหมุนข้อมูลภาพให้มีทิศทางเดียวกัน จะสามารถนำลักษณะเด่นมาเปรียบเทียบกันได้อย่างถูกต้อง ขั้นตอนวิธีการ SIFT ประกอบด้วย 4 ขั้นตอนสำคัญ ดังแสดงในภาพที่ 1 ดังนี้

- 1.1.1 การสร้างปริภูมิค่าในมิติขนาดและระยะทาง (scale-space)
- 1.1.2 การกำหนดตำแหน่งจุดสนใจ (keypoint localization)
- 1.1.3 การกำหนดทิศทางของจุดสนใจ (orientation assignment)
- 1.1.4 การสร้างคำอธิบายลักษณะเด่นของภาพ (keypoint descriptor)



ภาพที่ 1 ขั้นตอนการหาลักษณะเด่นของภาพ ด้วย SIFT

1.1.1 การสร้างปริภูมิค่าในมิติขนาดและระยะทาง (scale-space)

ขั้นตอนของการค้นหาจุดสนใจของภาพ จะต้องสร้างระดับชั้นของภาพขึ้นมา โดยข้อมูลภาพในแต่ละชั้นได้จากการสังวัตนาการ (convolution) ระหว่างภาพนำเข้าและตัวกรองแบบเกาส์เซียน ในแต่ละชั้นจะใช้ค่าตัวกรองที่แตกต่างกัน รายละเอียดของภาพจะลดลงเมื่อใช้ตัวกรองที่มีขนาดใหญ่ขึ้น หลังจากนั้นจะนำภาพแต่ละชั้นมาหาค่าความแตกต่างด้วยวิธีการ Difference of Gaussian (DoG) เพื่อนำมาพิจารณาเลือกจุดสนใจ นิยามได้ด้วยฟังก์ชัน $L(x, y, \sigma)$ ซึ่งเป็นชั้นของภาพที่เกิดจากการสังวัตนาการข้อมูลภาพ $I(x, y)$ ด้วยตัวกรองแบบเกาส์เซียนหลายระดับ $G(x, y, \sigma)$ ตามสมการที่ 1 และหาค่าตัวกรองแบบเกาส์เซียนได้จากสมการที่ 2 ดังนี้

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

สำหรับตัวกรองแบบเกาส์เซียนสองมิติจะได้จากผลคูณของส่วนเบี่ยงเบนมาตรฐานในแนวตั้งและแนวนอน ตามสมการที่ 3

$$g(; \sigma_x) * g(; \sigma_y) = g\left(; \sqrt{\sigma_x^2 + \sigma_y^2}\right) \quad (3)$$



การกำหนดค่า σ สำหรับตัวกรองแบบเกาส์เซียนในแต่ละชั้นของภาพจะกำหนดให้มีค่าในช่วง ตั้งแต่ σ ถึง 2σ ซึ่งการกำหนดค่า σ สำหรับชั้นภาพที่ 0 ถึงจำนวนชั้นของภาพ (s) จะอาศัยค่า k ซึ่งเป็นค่าคงที่ กำหนด $k = 2^{\frac{1}{s}}$ จะได้ค่า σ ในช่วง $k^0\sigma$ ถึง $k^s\sigma$ จะคำนวณได้จากสมการที่ 4 ซึ่งแสดงตัวอย่างการคำนวณได้ตามตารางที่ 1 เมื่อกำหนด $\sigma = 1.5$ จำนวนชั้นของแผ่นภาพ $s = 3$

$$g(\cdot, k^{p+1}\sigma) = g(\cdot; \sigma_p) * g(\cdot; k^p\sigma) \tag{4}$$

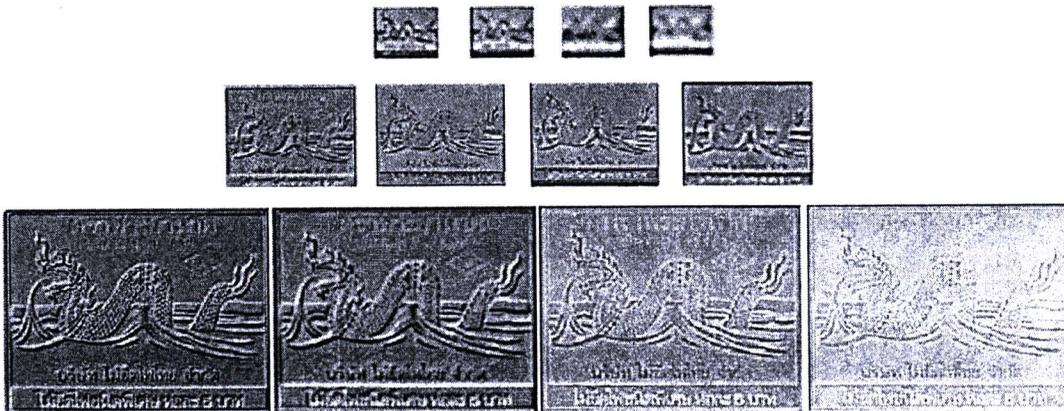
คำนวณหาค่า σ_p เมื่อ p คือลำดับชั้นของภาพ ตามสมการที่ 5

$$\begin{aligned} k^{p+1}\sigma &= \sqrt{\sigma_p^2 + (k^p\sigma)^2} \\ k^{2p+2}\sigma^2 &= \sigma_p^2 + k^{2p}\sigma^2 \\ \sigma_p^2 &= (k^2 - 1)k^{2p}\sigma^2 \\ \sigma_p &= k^p\sqrt{k^2 - 1}\sigma \end{aligned} \tag{5}$$

ตารางที่ 1 ตัวอย่างการคำนวณค่าส่วนเบี่ยงเบนมาตรฐาน ของตัวกรองเกาส์เซียนหลายระดับ

| ชั้นที่ (p) | ค่าตัวกรองแบบเกาส์เซียน | σ_p |
|-------------|--|---|
| 0 | $g(\cdot; 1.5)$ | $\sigma_0 = k^0\sqrt{k^2 - 1}\sigma = 1.1496$ |
| 1 | $g(\cdot, \sqrt{1.1496^2 + 1.5^2}) = g(\cdot; 1.8899)$ | $\sigma_1 = k^1\sqrt{k^2 - 1}\sigma = 1.4484$ |
| 2 | $g(\cdot, \sqrt{1.4484^2 + 1.8899^2}) = g(\cdot; 2.3811)$ | $\sigma_2 = k^2\sqrt{k^2 - 1}\sigma = 1.8249$ |
| 3 | $g(\cdot, \sqrt{1.8249^2 + 2.3811^2}) = g(\cdot; 3.0) = g(\cdot; 2\sigma)$ | |

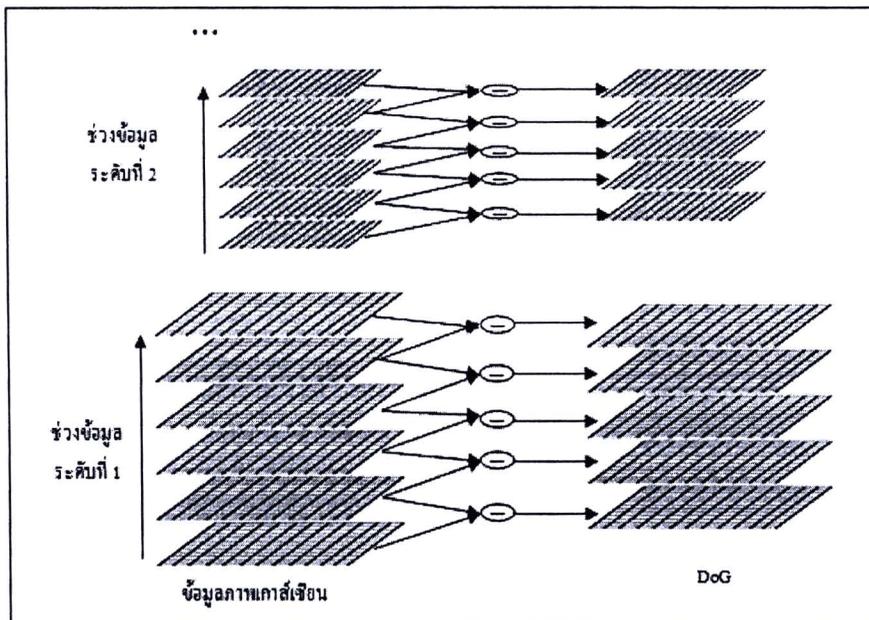
เมื่อสร้างภาพหลายระดับด้วยภาพนำเข้าครบทุกชั้นของภาพแล้ว หรือเรียกว่า 1 ช่วงข้อมูล หลังจากนั้นจะใช้ภาพสุดท้ายของช่วงข้อมูลปัจจุบัน เป็นภาพนำเข้าของช่วงข้อมูลถัดไป โดยจะลดขนาดภาพลงครึ่งหนึ่งในแต่ละมิติของขนาดเดิม ด้วยการสุ่มค่าทั้งสองจุดภาพ ทำซ้ำไปจนครบจำนวนช่วงข้อมูลที่ต้องการ การกำหนดจำนวนช่วงข้อมูลจะขึ้นอยู่กับขนาดของภาพ เพราะเมื่อเพิ่มช่วงข้อมูลขึ้น 1 ช่วง จะต้องลดขนาดภาพหนึ่งในสี่เท่าของขนาดเดิมเสมอ ตัวอย่างระดับชั้นของภาพแสดงตามภาพที่ 2



ภาพที่ 2 ตัวอย่างระดับชั้นของภาพที่สังวัตนาการด้วยตัวกรองแบบเกาส์เซียนหลายระดับ

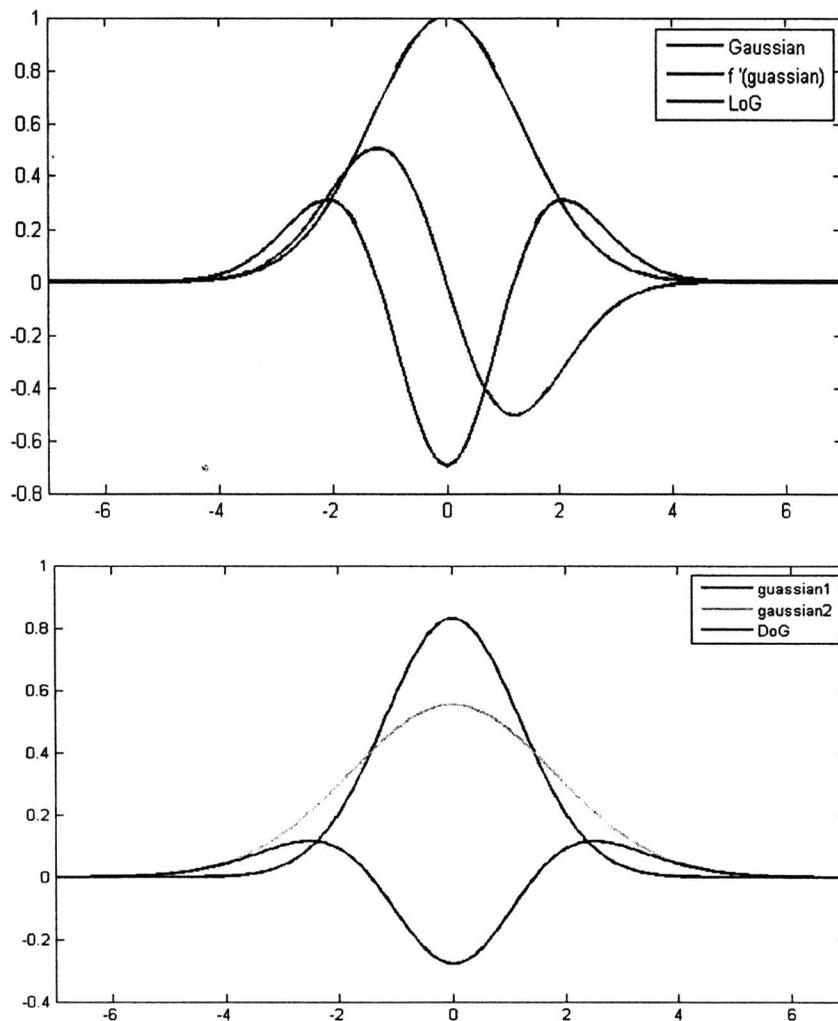
ขั้นตอนถัดไป จะทำการหาค่าความแตกต่างระหว่างชั้นของข้อมูลที่อยู่ติดกัน ด้วยวิธีการที่เรียกว่า Difference of Gaussian (DoG) เพื่อให้ข้อมูลของภาพ ในส่วนของรายละเอียดภาพมีความเด่นชัดมากขึ้น สามารถคำนวณได้จากสมการ 6 รายละเอียดแสดงในภาพที่ 3

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (6)$$



ภาพที่ 3 ขั้นตอนการหาค่า Difference of Gaussian

การคำนวณหาค่าความแตกต่างระหว่างชั้นของภาพ ด้วยวิธีการ Difference of Gaussian (DoG) วิธีการนี้จะเป็นการประมาณค่าของวิธีการ Laplacian of Gaussian (LoG) จะต้องใช้การคำนวณอนุพันธ์อันดับที่สองของฟังก์ชันเกาส์เซียน ซึ่งจะใช้ในการเก็ลี่ยภาพ เพื่อลดความละเอียดของภาพ หรือใช้ในการหาขอบภาพ



ภาพที่ 4 ผลการคำนวณความแตกต่างด้วยวิธีการ LoG (บน) และ DoG (ล่าง)

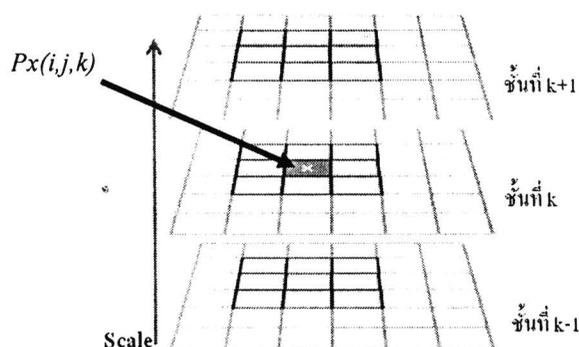
ทั้งวิธีการ LoG และ DoG เป็นตัวกรองแบบแถบความถี่ผ่าน (Band pass Filter) จะยอมให้สัญญาณในช่วงความถี่หนึ่งผ่านไปได้ หากสัญญาณมีความถี่อยู่นอกช่วงที่กำหนดจะถูกลดทอนหรือไม่ยอมให้ผ่านไป ทำให้คงเหลือลักษณะสำคัญของภาพต้นฉบับเอาไว้และกำจัดส่วนที่มีรายละเอียดต่ำออก โดยการเปลี่ยนแปลงสัญญาณจะอยู่ที่จุดสูงสุดหรือจุดต่ำสุด (จุดที่ความชันมีค่าเข้าใกล้ศูนย์)

1.1.2 การกำหนดตำแหน่งจุดสนใจ (keypoint localization)

ในขั้นตอนถัดมา คือ การหาค่าสูงสุดและต่ำสุดของข้อมูลที่ได้จากการหาค่าความแตกต่างระหว่างชั้นของข้อมูลแล้ว นำมาพิจารณาเลือกจุดสนใจ โดยเลือกจุดที่มีค่าสูงสุดหรือต่ำสุดเมื่อเทียบกับจุดภาพรอบข้างในชั้นเดียวกันและชั้นข้างเคียงจะถูกเลือกเป็นจุดสนใจ ตามสมการ 7

$$\text{if } Px(i,j,k) = \text{Max}_{m=i-1}^{i+1} \text{Max}_{n=j-1}^{j+1} \text{Max}_{s=k-1}^{k+1} (Px(m,n,s)) \quad (7)$$

เมื่อ $Px(i,j,k)$ คือจุดภาพที่พิจารณาในชั้นที่ k แสดงได้ในภาพที่ 5



ภาพที่ 5 การเปรียบเทียบข้อมูลจุดภาพกับจุดภาพข้างเคียง

เนื่องจากการใช้ค่า DoG จะเป็นค่ามีตอบสนองต่อเส้นขอบของภาพได้ดี ทำให้จุดสนใจที่เลือกมา อาจจะมาจกจุดภาพที่เป็นสัญญาณรบกวน ดังนั้นต้องพิจารณาความโค้งหลักของจุดภาพ (principle curvatures) โดยใช้ 2×2 Hessian matrix ตามสมการที่ 8

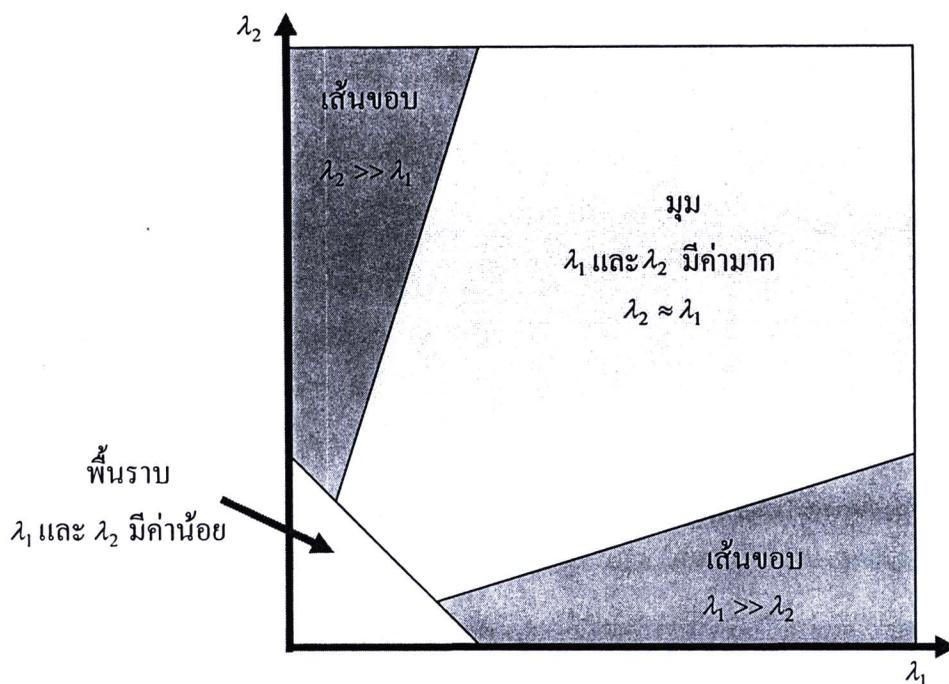
$$M = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (8)$$

การพิจารณาการตอบสนองต่อเส้นขอบ จะพิจารณาจากสมการดังนี้

$$R = \det M - k(\text{trace } M)^2 \quad (9)$$

เมื่อ $\det M = \lambda_1 \lambda_2$, $\text{trace } M = \lambda_1 + \lambda_2$ โดย λ คือค่าลักษณะเฉพาะ (eigenvalues) และ k เป็นค่าคงที่ อยู่ในช่วง 0.04 - 0.06 การตอบสนองต่อเส้นขอบจะพิจารณาจากอัตราส่วนของค่าลักษณะเฉพาะ (eigenvalues) ที่มากที่สุดต่อค่าลักษณะเฉพาะที่น้อยที่สุด จุดสนใจที่ดีต้องมีค่า

ลักษณะเฉพาะที่มีค่าใกล้เคียงกัน ถ้าค่าความโค้งในทิศทางหนึ่งมีค่ามาก แต่อีกทิศทางมีค่าน้อย อาจจะเป็นจุดสนใจที่วางแนวไปตามขอบภาพ หรือได้มาจากสัญญาณรบกวน ดังรายละเอียดใน ภาพที่ 6



ภาพที่ 6 การพิจารณาการตอบสนองต่อเส้นขอบ

ในการพิจารณาการตอบสนองต่อเส้นขอบของวิธีการ SIFT จะใช้การคำนวณตามสมการที่ 10 โดยไม่ต้องคำนวณหาค่าลักษณะเฉพาะ กำหนดให้ r คือ ค่าคงที่ของค่าขีดแบ่งอัตราส่วนของค่าลักษณะเฉพาะ ตามวิธีการมาตรฐานของ SIFT กำหนดให้ $r=10$ ดังนั้นจุดสนใจที่มีค่าอัตราส่วนของค่าลักษณะเฉพาะมากกว่า 12.1 จะถูกตัดออกจากการพิจารณาตามสมการ

$$\frac{D_{xx} + D_{yy}}{D_{xx}D_{yy} - (D_{xy})^2} < \frac{(r+1)^2}{r} \quad (10)$$

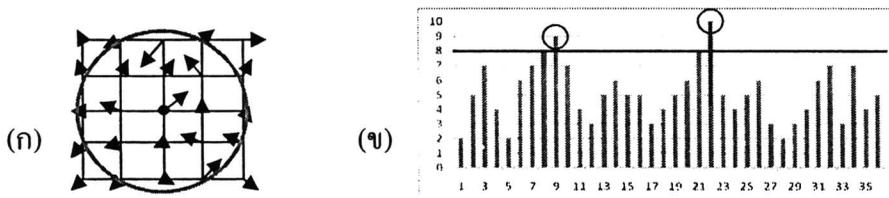
1.1.3 การกำหนดทิศทางของจุดสนใจ (orientation assignment)

ในขั้นตอนนี้จะทำการหามุมของจุดสนใจ โดยหาค่าทิศทางและขนาดของความชันของจุดภาพในแต่ละระดับชั้น $L(x, y)$ สามารถคำนวณหาขนาดของความชัน $m(x, y)$ และค่าทิศทางของความชัน $\theta(x, y)$ ตามสมการที่ 11 และ 12

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (11)$$

$$\theta(x, y) = \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (12)$$

เมื่อได้ทิศทางและขนาดของความชันของจุดภาพในแต่ละระดับชั้นแล้ว จะหาแถบความถี่ของขนาดความชันของจุดภาพตามทิศทางรอบจุดสนใจ ในพื้นที่วงกลมรัศมีเท่ากับ 1.5 เท่าของส่วนเบี่ยงเบนมาตรฐานของแผ่นภาพที่มีจุดสนใจปรากฏอยู่ เพื่อกำหนดเป็นทิศทางหลักของจุดสนใจ ตามตัวอย่างในภาพที่ 7 แถบความถี่ของขนาดความชัน จะแบ่งเป็น 36 กลุ่ม จะพิจารณาค่าแถบความถี่ที่มีค่ามากกว่าร้อยละ 80 จะกำหนดทิศทางจากแถบความถี่นั้นเป็นทิศทางหลักของจุดสนใจ ซึ่งจะเกิดจุดสนใจทับซ้อนในตำแหน่งเดียวกัน แต่จะมีค่าทิศทางของจุดสนใจที่ต่างกัน เมื่อนำไปใช้ในการเปรียบเทียบจุดสนใจจะทำให้มีความถูกต้องมากกว่าการเลือกทิศทางหลักที่มากที่สุดเพียงค่าเดียว



ภาพที่ 7 การหาค่าทิศทางหลักของจุดสนใจ (ก) ทิศทางของจุดภาพรอบจุดสนใจ
(ข) แถบความถี่ของทิศทางรอบจุดสนใจ

1.1.4 การสร้างคำอธิบายลักษณะเด่นของภาพ (keypoint descriptor)

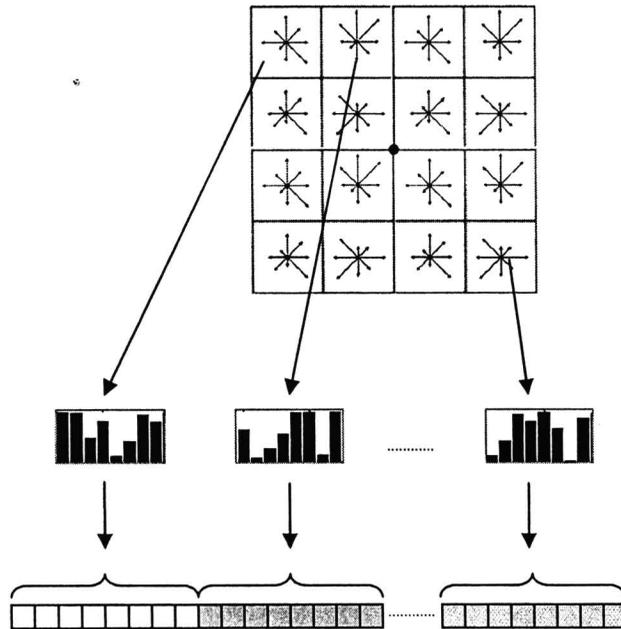
เมื่อได้ทิศทางหลักของจุดสนใจแล้ว หลังจากนั้นจะสร้างลักษณะเด่นของจุดสนใจจากจะสร้างพื้นที่สนใจที่มีจุดสนใจเป็นจุดศูนย์กลาง แล้วแบ่งพื้นที่สนใจเป็นพื้นที่ย่อยขนาด 4×4 ดังรูปที่ 8 คำนวณภาพแห่งความถี่ของทิศทางเกรเดียนต์ ตามขนาดของเกรเดียนต์ ในแต่ละพื้นที่ย่อย ทิศทางของจุด จะถูกแบ่งออกเป็น 8 กลุ่ม คำนวณค่าแห่งความถี่ของแต่ละพื้นที่ย่อย ได้ดังนี้

$$h_{r(l,m)}(k) = \sum_{x,y \in r(l,m)} M(x,y) (1 - |\Theta(x,y) - c_k| / \Delta_k) \quad (13)$$

เมื่อ $\Theta(x,y)$ เป็นสมาชิกของกลุ่ม k , c_k คือ ทิศทางของศูนย์กลาง, Δ_k คือช่วงข้อมูลทิศทางของกลุ่ม และ (x,y) คือจุดร่วมของภาพในพื้นที่ย่อย เมื่อคำนวณครบทุกพื้นที่ย่อย แล้วนำค่าแห่งความถี่ของทิศทางเกรเดียนต์มาเรียงต่อกันจะได้ลักษณะเด่นของภาพที่หาจากวิธีการ SIFT ขนาด 128 ข้อมูล

$$u = (h_{r(1,1)}, \dots, h_{r(l,m)}, \dots, h_{r(4,4)}) \quad (14)$$

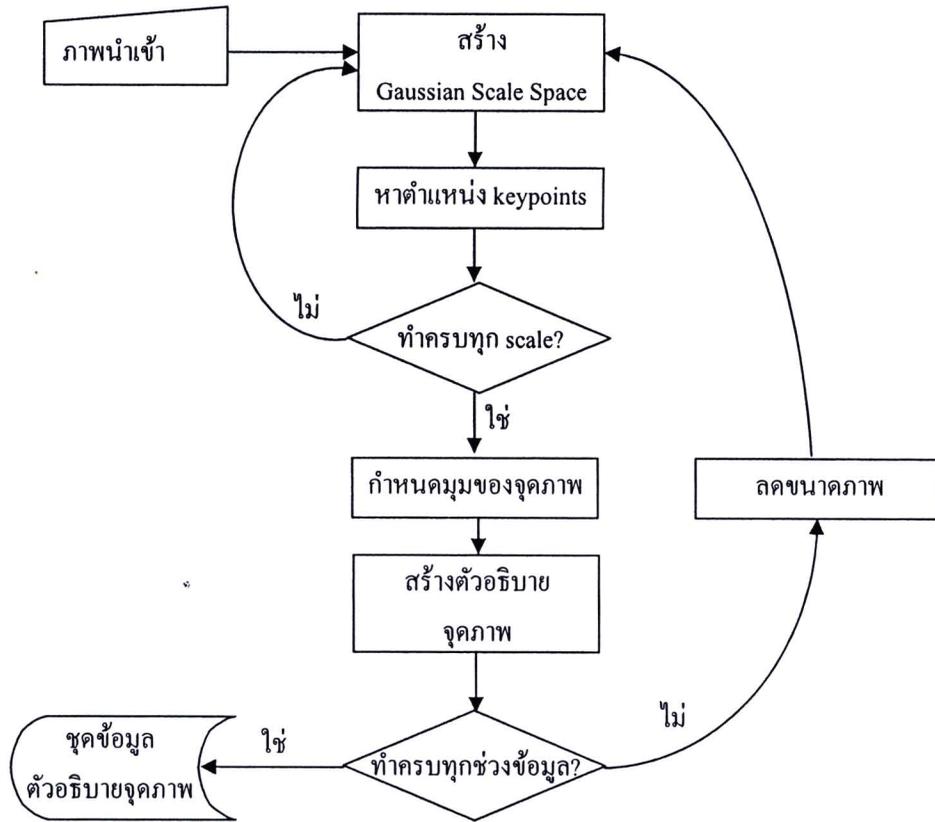
ทำการคำนวณหาลักษณะเด่นจนครบทุกจุดสนใจ จะได้ลักษณะเด่นของภาพทั้งหมด $n \times 128$ มิติ ข้อมูล เมื่อ n คือจำนวนจุดสนใจของภาพ



ภาพที่ 8 การสร้างคำอธิบายลักษณะเด่นของภาพด้วยวิธีการ SIFT



โดยสามารถสรุปขั้นตอนการคำนวณหาลักษณะเด่นของภาพด้วยวิธีการ SIFT ได้ตามภาพที่ 9



ภาพที่ 9 ฟังก์ชันการทำงานของวิธีการ SIFT

1.1.5 ข้อดีและข้อเสียของวิธีการ SIFT

1.1.5.1 ข้อดีของวิธีการ SIFT

สามารถสร้างการอธิบายลักษณะเด่นของภาพถ่ายวัตถุเดียวกัน แต่ภาพมีขนาดและทิศทาง หรือมุมมองที่แตกต่างกันได้

1.1.5.2 ข้อเสียของวิธีการ SIFT

ในการสร้างตัวอธิบายลักษณะเด่น จะใช้มิติข้อมูลจำนวนมาก และการทำงานจะจำกัดเฉพาะภาพที่เป็นค่าสีระดับเทา การคำนวณมีความซับซ้อนในสองส่วนที่สำคัญ คือ ส่วนของขั้นตอนการหาจุดสนใจและส่วนของการกำหนดทิศทางของจุดสนใจ ซึ่งในขั้นตอนดังกล่าวจะใช้ฟังก์ชันทางคณิตศาสตร์ที่ซับซ้อนเช่น $\exp()$, $\arctan()$ เป็นต้น

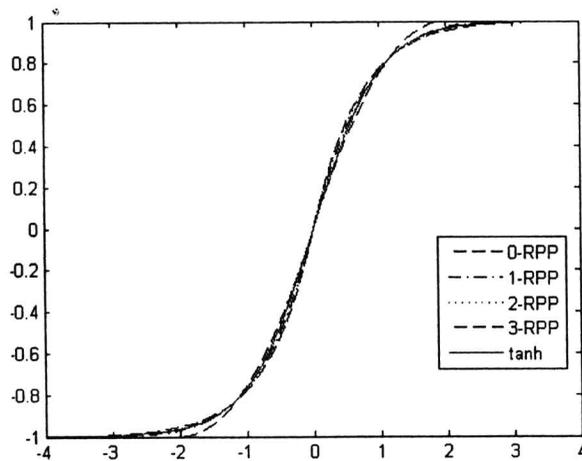


1.2 การประมาณค่าด้วยตัวก่อกำเนิดพหุนามเป็นช่วง ชนิด p-RPP

ตัวก่อกำเนิดพหุนามเป็นช่วง ชนิด p-RPP เป็นวิธีการที่พัฒนาขึ้นเพื่อใช้ประมาณค่าฟังก์ชัน \tanh [1, 18] การประมาณค่าด้วยวิธีการนี้ จะมีข้อดีคือ สามารถขยายช่วงข้อมูลนำเข้าอยู่ในช่วง $[-2^{p+1}, 2^{p+1}]$ เมื่อ $p = 0, 1, 2, 3$ ตามสมการที่ 15

$$G(\hat{x}, p) = \begin{cases} -1 & , \hat{x} \leq -2^{p+1} \\ -1 + (1 + 2^{-(p+1)} \times \hat{x})^{3 \times 2^p - 1} & , -2^{p+1} < \hat{x} < 0 \\ +1 - (1 - 2^{-(p+1)} \times \hat{x})^{3 \times 2^p - 1} & , 0 < \hat{x} < 2^{p+1} \\ +1 & , \hat{x} \geq 2^{p+1} \end{cases} \quad (15)$$

เปรียบเทียบผลการคำนวณกับฟังก์ชัน \tanh ผลลัพธ์ที่ได้จะมีค่าใกล้เคียงกัน ดังแสดงตามภาพที่ 10 เมื่อเพิ่มค่า p ผลลัพธ์ที่ได้จากการคำนวณจะมีค่าใกล้เคียงกับฟังก์ชัน \tanh มากยิ่งขึ้น

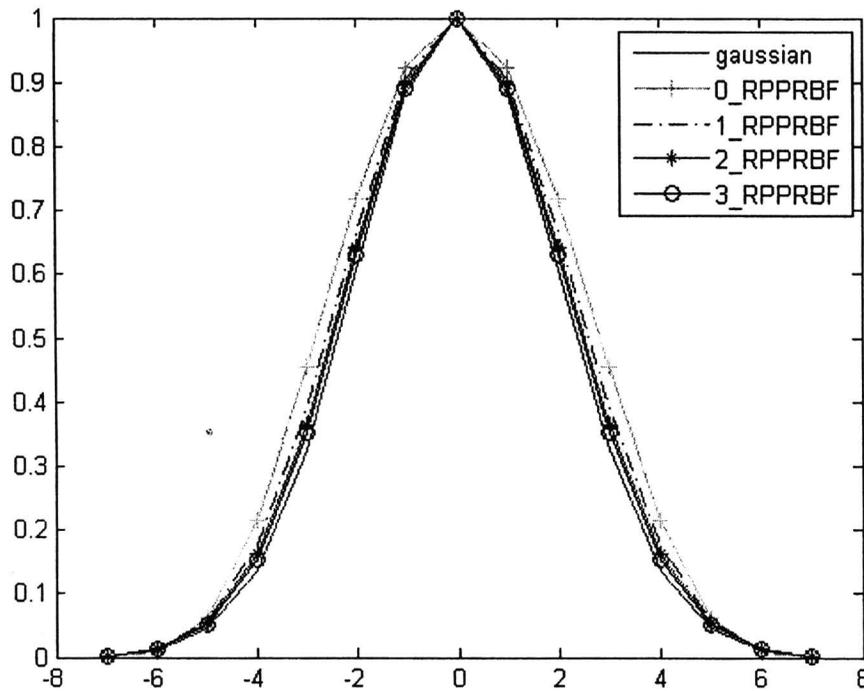


ภาพที่ 10 การประมาณค่า tanh-like เทียบกับ \tanh

เมื่อกำหนดหาค่าอนุพันธ์ของฟังก์ชัน p-RPP เพื่อใช้ในการประมาณค่าฟังก์ชันเกาส์เซียน จะได้ตามสมการ 16

$$G'(\hat{x}, p) = \begin{cases} 0, & \hat{x} \leq -2^{p+1} \\ \frac{3 \times 2^p - 1}{2^{p+1}} (1 + 2^{-(p+1)} \times \hat{x})^{3 \times 2^p - 2}, & -2^{p+1} < \hat{x} < 0 \\ \frac{3 \times 2^p - 1}{2^{p+1}} (1 - 2^{-(p+1)} \times \hat{x})^{3 \times 2^p - 2}, & 0 < \hat{x} < 2^{p+1} \\ 0, & \hat{x} \geq 2^{p+1} \end{cases} \quad (16)$$

จากสมการที่ 16 จะนำไปใช้ในการคำนวณหาค่าตัวกรองแบบเกาส์เซียน เรียกว่า p-recursive piecewise polynomial radial basis function (p-RPPRBF) เมื่อกำหนดค่าฟังก์ชัน p_RPPRBF ที่เปลี่ยนค่า $p = 0, 1, 2, 3$ เทียบกับค่าที่คำนวณด้วยฟังก์ชันเกาส์เซียน เมื่อค่า p เพิ่มมากขึ้น ค่าที่คำนวณได้จะมีค่าใกล้เคียงกับฟังก์ชันเกาส์เซียนมากยิ่งขึ้น แสดงได้ตามภาพที่ 11



ภาพที่ 11 การประมาณค่าด้วย p-RPPRBF เทียบกับค่าจากการคำนวณด้วยฟังก์ชันเกาส์เซียน

1.3 การประมาณเชิงเหตุผล

การประมาณเชิงเหตุผล (Rational Approximation) เป็นการแก้ปัญหาพีชคณิตที่อยู่ในรูป $r_{n,m}(t) = p_n(t)/q_m(t)$ ที่ให้ค่าผลลัพธ์ใกล้เคียงกับฟังก์ชัน $f(t)$ โดย $p_n(t)$ และ $q_m(t)$ เป็นสมการพหุนาม [4] อยู่ในรูปสมการ คือ

$$\begin{aligned}
 p_n(t) &= \sum_{i=0}^n p_i t^i = p_0 + p_1 t + p_2 t^2 \dots + p_n t^n \\
 q_m(t) &= 1 + \sum_{i=1}^m q_i t^i = 1 + q_1 t + q_2 t^2 \dots + q_m t^m
 \end{aligned}
 \tag{17}$$



ทำการแก้สมการหลายตัวแปรเพื่อหาค่า $p(t)$ และ $q(t)$ ที่ทำให้ $r_{n,m}(t)$ มีค่าใกล้เคียง $\arctan(t)$ มากที่สุด สำหรับงานวิจัยนี้ จะใช้การประมาณค่าพหุนามอันดับสอง กำหนดให้ $x \in [-1, 1]$ จะประมาณค่า $\arctan(x)$ ได้ดังนี้

กำหนดให้

$$f(x) = \text{Arc tan}(x) \quad (18)$$

$$p_n(x) = p_0 + p_1x + p_2x^2 \quad (19)$$

$$q_m(x) = 1 + q_1x + q_2x^2 \quad (20)$$

$$R(x) = \frac{p_0 + p_1x + p_2x^2}{1 + q_1x + q_2x^2} \quad (21)$$

คำนวณหาค่า $f(x) = \text{Arc tan}(x)$ ในช่วง $x \in [-1, 1]$ เพื่อใช้เป็นคำตอบของการประมาณค่า และจัดสมการให้อยู่ในรูปดังนี้

$$R(x) = \frac{p_0 + p_1x + p_2x^2}{1 + q_1x + q_2x^2} = f(x) = \text{Arc tan}(x) \quad (22)$$

จะได้

$$p_0 + p_1x + p_2x^2 = f(x)(1 + q_1x + q_2x^2) \quad (23)$$

แทนค่า x ทุกตัวที่คำนวณได้จาก $f(x) = \text{Arc tan}(x)$ ในสมการที่ 23 จะได้สมการดังนี้

$$\begin{aligned} p_0 + p_1x_1 + p_2x_1^2 &= f(x_1)(1 + q_1x_1 + q_2x_1^2) \\ p_0 + p_1x_2 + p_2x_2^2 &= f(x_2)(1 + q_1x_2 + q_2x_2^2) \\ p_0 + p_1x_3 + p_2x_3^2 &= f(x_3)(1 + q_1x_3 + q_2x_3^2) \\ &\vdots \\ p_0 + p_1x_k + p_2x_k^2 &= f(x_k)(1 + q_1x_k + q_2x_k^2) \end{aligned} \quad (24)$$

เมื่อ k คือจำนวนข้อมูลที่เป็นคำตอบของ $f(x)$ คำนวณแก้สมการหลายตัวแปร เพื่อหาค่าของ p และ q แล้วแทนค่าลงในสมการที่ 22 โดยค่า p และ q ที่คำนวณได้จะมีค่าแตกต่างกันในแต่ละครั้ง ที่คำนวณ ขึ้นอยู่กับช่วงค่าที่คำนวณในฟังก์ชัน $f(x)$ แต่สามารถจัดอยู่ในรูปทั่วไปได้ดังนี้

$$R(x) = \frac{\alpha x}{1 + \beta x^2} \quad (25)$$

โดยค่าที่ทำให้ $R(x)$ ใกล้เคียงกับ $f(x)$ มากที่สุดคือ $\beta = 0.280872$ และ α มีค่าเข้าใกล้ 1 ซึ่งในงานวิจัยนี้ต้องการระดับการคำนวณที่ทศนิยมตำแหน่งที่สอง จึงได้สมการในการประมาณค่า ดังนี้

$$R(x) = \frac{x}{1 + 0.28x^2} \quad (26)$$

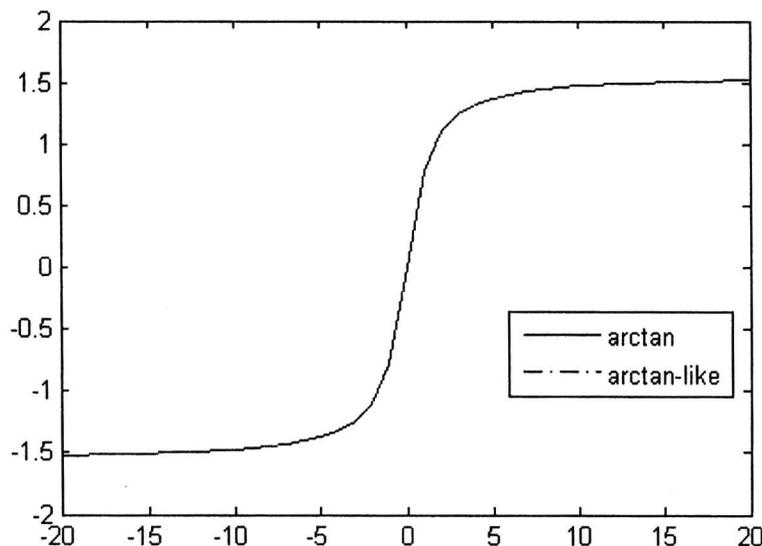
เนื่องจากการใช้ rational approximation function เพียง 1 ฟังก์ชัน จะครอบคลุมช่วง $-1 \leq x \leq 1$ เท่านั้น เพื่อขยายให้เต็มช่วงที่สนใจ จึงต้องใช้ฟังก์ชันลักษณะเดียวกันมากำหนดค่าข้อมูลที่อยู่นอกขอบเขตดังกล่าว เราเรียกว่า piecewise rational approximation function ในงานวิจัยนี้แบ่งโดเมนเป็น 3 ช่วง คือ $x < -1$, $-1 \leq x \leq 1$ และ $x > 1$

ดังนี้

$$\text{atan_like}(x) = \begin{cases} \frac{\pi}{2} - \left(\frac{x}{x^2 + 0.28}\right), & x > 1 \\ \frac{x}{1 + 0.28x^2}, & -1 \leq x \leq 1 \\ -\frac{\pi}{2} - \left(\frac{x}{x^2 + 0.28}\right), & x < -1 \end{cases} \quad (27)$$

และผลการคำนวณเปรียบเทียบค่าจากการประมาณและค่าที่ได้จากการคำนวณด้วยฟังก์ชัน $\arctan()$ ตามภาพที่ 12 ฟังก์ชัน $\arctan()$ [14] จะใช้การคำนวณตามสมการที่ 28 ดังนี้

$$\arctan(x) = \frac{i}{2} \log\left(\frac{i+x}{i-x}\right), \text{ เมื่อ } i \text{ คือหน่วยจินตภาพ } (\sqrt{-1}) \quad (28)$$



ภาพที่ 12 การประมาณค่า $\arctan\text{-like}()$ เทียบกับ $\arctan()$

คำนวณหาค่า MSE ระหว่างค่าที่ได้จากการประมาณด้วย arctan-like() และค่าที่คำนวณได้ ฟังก์ชัน arctan() คำนวณซ้ำ 100 รอบ มีค่า MSE เท่ากับ 0.0000016 จะเห็นว่าความผิดพลาดจากการประมาณจะมีความแตกต่างจากค่าที่คำนวณได้จากสมการหลักน้อยมาก แต่การดำเนินการทางคณิตศาสตร์ด้วยวิธีการ arctan-like() จะใช้การบวก 1 ครั้ง การลบ 1 ครั้ง การคูณ 2 ครั้งและการหาร 1 ครั้ง โดยเฉลี่ย จึงมีความซับซ้อนน้อยกว่าการคำนวณฟังก์ชัน arctan() ที่ใช้การคูณและการหารในขั้นตอนการคำนวณมากกว่า

การประมาณค่า arctan อีกวิธีหนึ่งคือการใช้ look up table (LUT) โดยจะเก็บค่าข้อมูลการคำนวณที่เป็นไปได้ไว้ในตารางหลายมิติ โดยแต่ละข้อมูลจะกำหนดดัชนีในการเข้าถึงข้อมูลเอาไว้ การประมาณค่าด้วย LUT เป็นวิธีการประมาณค่าที่ง่ายในการใช้งาน เพราะไม่ต้องคำนวณสมการที่ซับซ้อน สามารถนำค่าที่คำนวณแล้วไปใช้งานได้ แต่ต้องใช้เวลาในการคำนวณตำแหน่งของข้อมูล que เก็บไว้ในตาราง เช่น hashing และข้อมูลถูกเก็บในหน่วยความจำ ทำให้ต้องใช้เวลาส่วนหนึ่งในการเข้าถึงหน่วยความจำด้วย จึงใช้เวลามากกว่าวิธีการที่คำนวณด้วยหน่วยประมวลผลกลางโดยตรง

1.4 การหาลักษณะเด่นของภาพแบบ L-SIFT

วิธีการหาลักษณะเด่นของภาพแบบ Low complexity SIFT (L-SIFT) เป็นวิธีการหาลักษณะเด่นที่ได้ลดความซับซ้อนในการคำนวณของขั้นตอนวิธีการ SIFT จะแก้ไขขั้นตอนวิธีของ SIFT 2 ส่วน คือ ส่วนที่ 1 การสร้างปริภูมิค่าในมิติขนาดและระยะทางเพื่อหาจุดสนใจของภาพ โดยจะเปลี่ยนวิธีการสร้างภาพหลายระดับชั้น จากการใช้ตัวกรองแบบเกาส์เซียนสำหรับ DoG เป็นวิธีการประมาณค่าด้วย p-RPPRBF สำหรับ Difference of p-recursive piecewise polynomial RBF (Dop-RPPRBF) และส่วนที่ 2 การกำหนดทิศทางการหลักของจุดสนใจ จากการใช้ฟังก์ชัน arctan() เป็นการประมาณค่าด้วย piecewise rational approximation function สำหรับ arctan-like()

การสร้างปริภูมิค่าในมิติขนาดและระยะทาง จะหาความแตกต่างระหว่างภาพที่สังวัตนาการด้วยตัวกรองที่ประมาณค่าฟังก์ชันแบบ p-RPPRBF เมื่อนำข้อมูลภาพนำเข้า $I(x, y)$ มาสังวัตนาการกับตัวกรอง $\phi(x, y, \sigma, d)$ ที่ได้จาก p-RPPRBF ตามสมการ

$$L(x, y, \sigma) = \phi(x, y, \sigma, d) * I(x, y) \quad (29)$$

การคำนวณหาค่า d จะใช้ค่าส่วนเบี่ยงเบนมาตรฐานมาพิจารณาร่วมด้วย โดย กำหนดค่า d จากสมการ

$$d = (x^2 + y^2) / 2\sigma^2 \quad (30)$$

เมื่อ (x, y) คือ ค่าจตุรรวมของตัวกรองในแนวตั้งและแนวนอน และ σ คือ ค่าส่วนเบี่ยงเบนมาตรฐาน โดยจะใช้ค่าเดียวกันกับวิธีการของ SIFT เดิม เพื่อให้ได้ค่าตัวกรองที่มีลักษณะเดียวกับฟังก์ชันเกาส์เซียน จะเขียนในรูปรหัสคำสั่งได้ตามภาพที่ 13 เพื่อใช้หาค่าชั้นของข้อมูลภาพตามวิธีการของ SIFT ซึ่งในการคำนวณค่า p-RPPRBF ใช้เพียงตัวดำเนินการทางคณิตศาสตร์พื้นฐานเท่านั้น

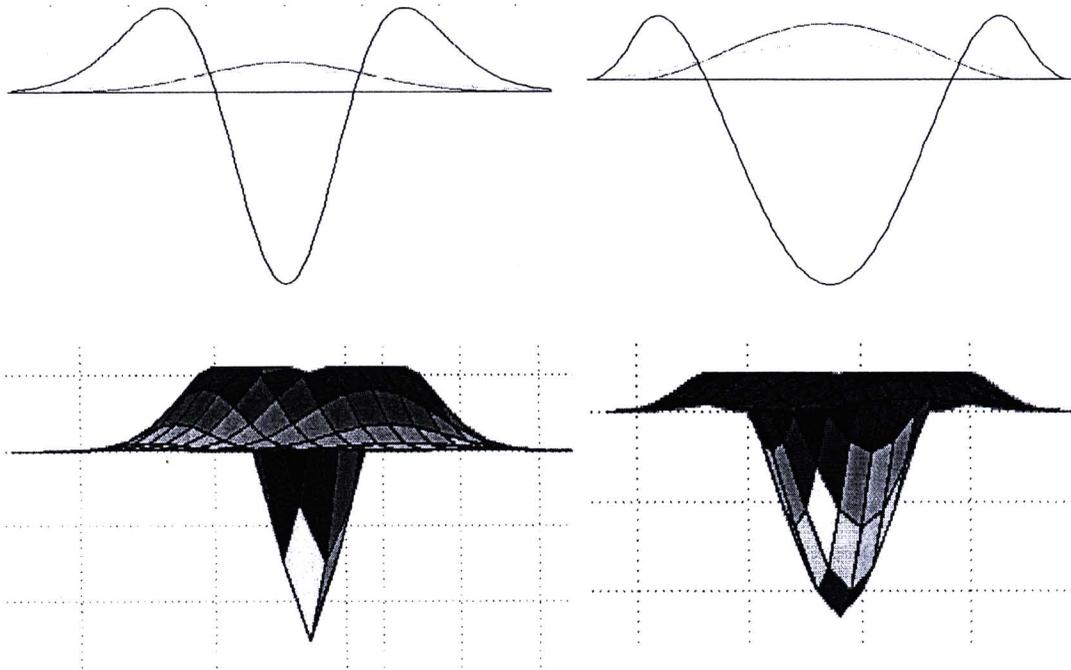
```
function output= pRPPRBF(d,p)
    delta = 1-(d./(2.^(p+1)));
    delta(delta<0) = 0;
    output = delta.*delta;
    for n=1:p
        output = delta.*output.*output;
    end
end
```

ภาพที่ 13 รหัสคำสั่งภาษา matlab สำหรับ p-RPPRBF

การหาความแตกต่างระหว่างชั้นของข้อมูลที่อยู่ติดกันจะคำนวณได้จากสมการ

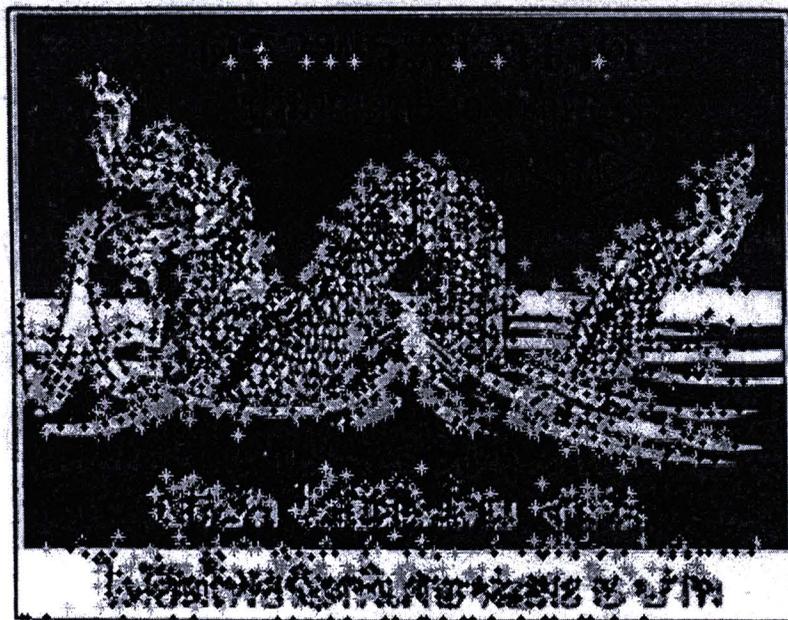
$$\begin{aligned} L(x, y, \sigma_1) &= (\phi(x, y, \sigma_1, d_1) * I(x, y)) \\ L(x, y, \sigma_2) &= (\phi(x, y, \sigma_2, d_2) * I(x, y)) \\ Dop - RPPRBF &= L(x, y, \sigma_2) - L(x, y, \sigma_1) \end{aligned} \quad (31)$$

เมื่อคำนวณหาความแตกต่างของข้อมูลภาพที่ผ่านการสังวัตนาการจากวิธีการ DoG และ วิธีการ Dop-RPPRBF จะมีลักษณะคล้ายคลึงกัน ดังภาพที่ 14 ที่แสดงลักษณะของข้อมูลที่ผ่านวิธีการ DoG และวิธีการ Dop-RPPRBF ทั้งแบบสองมิติและ สามมิติ ผลลัพธ์ที่ได้จะมีลักษณะคล้ายกัน



รูปที่ 14 ลักษณะของผลลัพธ์ที่ได้จากวิธีการ (ซ้าย) DoG (ขวา) Dop-RPPRBF

เมื่อเปรียบเทียบผลลัพธ์ที่ได้ของวิธี DoG และ Dop-RPPRBF จากรูปที่ 15 ตำแหน่งของจุดสนใจที่ได้จากวิธี DoG (จุด * สีเขียว) และ Dop-RPPRBF (จุด o สีน้ำเงิน) จุดสนใจที่ได้จาก 2 วิธีจะอยู่ในบริเวณที่ใกล้เคียงกัน โดยจุดสนใจจะปรากฏตามขอบภาพ



รูปที่ 15 จุดสนใจที่ได้จากวิธี DoG และ Dop-RPPRBF

การกำหนดทิศทางหลักของจุดสนใจ จะใช้การประมาณค่าด้วย piecewise rational approximation function สำหรับ arctan-like() จะใช้รหัสคำสั่งสำหรับการคำนวณตามภาพที่ 16

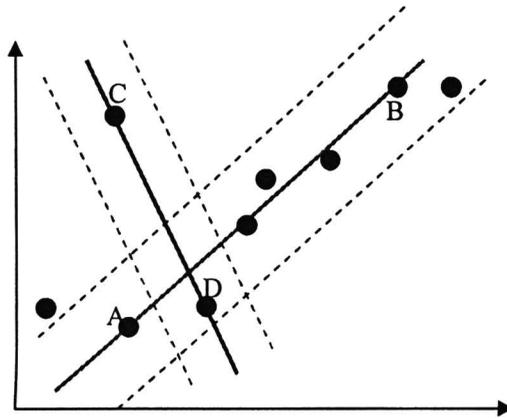
```
function output = atan_like(x)
    output = zeros(size(x));
    id = find(x(:)>1);
    xid = x(id);
    output(id) = pi./2-(xid./((xid.^2)+0.28));
    id = find(x(:)>=-1 & x(:)<=1);
    xid = x(id);
    output(id) = xid./(1+0.28.*xid.^2);
    id = find(x(:)<-1);
    xid = x(id);
    output(id) = -pi./2-(xid./((xid.^2)+0.28));
end
```

ภาพที่ 16 รหัสคำสั่งภาษา matlab สำหรับ arctan-like

1.5 การสร้างภาพมุมกว้าง (panorama)

การสร้างภาพมุมกว้าง คือการนำภาพถ่ายหลาย ๆ ภาพมาต่อกันให้เป็นภาพเดียว หรือการสร้างภาพวิวที่เห็นสมบูรณ์ ซึ่งไม่สามารถถ่ายภาพได้ในครั้งเดียว การเชื่อมต่อภาพจะต้องหาจุดสนใจที่มีประสิทธิภาพ เพื่อให้เกิดการเชื่อมต่อที่สมบูรณ์ สำหรับงานวิจัยนี้จะนำมาใช้ในการวัดประสิทธิภาพของจุดสนใจที่หาได้จากวิธีการ SIFT และ L-SIFT การคำนวณการเชื่อมต่อภาพจะใช้วิธีการของ Automatic Panoramic [3] โดยอาศัยหลักการของ RANSAC (Random Sample Consensus) โดยมีหลักการ ดังนี้

1. สุ่มเลือกจุดสนใจจากกลุ่มของจุดสนใจที่มีความคล้ายคลึงกันมา 2 จุด
2. สร้างสมการเส้นตรงจากจุดสนใจที่เลือกมาและหาระยะทางจากจุดทุกจุดมายังเส้นตรง
3. นับจำนวนจุดสนใจที่มีระยะทางจากเส้นตรงน้อยกว่าค่าขีดแบ่งระยะทาง เรียกว่า support
4. ทำซ้ำ 1 - 3 ตามจำนวน N รอบที่กำหนด เพื่อหาเส้นที่มีจำนวน support มากที่สุด
5. เลือกเส้นที่มี support มากที่สุดเป็นตัวแทนโครงสร้าง และจุดที่อยู่ภายใต้เส้นเขตแบ่งของระยะทาง เรียกว่า inliers



ภาพที่ 17 การสุ่มจุดสนใจและการพิจารณา support ของวิธีการ RANSAC

จากภาพที่ 17 เส้นตรง AB จะมี support ทั้งหมด 7 จุด และเส้นตรง CD จะมี support ทั้งหมด 2 จุด ซึ่งการพิจารณาค่าขีดแบ่งระยะทาง (distance threshold) อาจจะใช้การกำหนดเป็นค่าคงที่ หรือจะใช้การควรถ้านวน เช่น ฟังก์ชันเกาส์เซียนที่ทราบค่าส่วนเบี่ยงเบนมาตรฐาน เป็นต้น จำนวนรอบของการคำนวณซ้ำ (N) จะต้องกำหนดให้มีจำนวนมากพอที่จะคำนวณให้เกิดประสิทธิภาพสูงสุด สามารถคำนวณได้ดังนี้

กำหนดให้

ϵ คือ สัดส่วนของข้อมูลที่ผิดพลาด (ไม่อยู่ใน inliers)

$(1 - \epsilon)$ คือ ความน่าจะเป็นที่จะได้ข้อมูลที่ถูกต้อง 1 ตัว

$(1 - \epsilon)^s$ คือ ความน่าจะเป็นที่จะได้ข้อมูลที่ถูกต้อง s ตัว

$(1 - (1 - \epsilon)^s)^N$ คือ ความน่าจะเป็นที่จะเกิดความผิดพลาดทั้งหมด

p คือ ความน่าจะเป็นที่จะได้ข้อมูลที่ถูกต้อง

จะได้

$$1 - p = (1 - (1 - \epsilon)^s)^N \quad (32)$$

ดังนั้น จำนวนรอบสูงสุด หาได้จาก

$$N = P(\epsilon, p) = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^s)} \quad (33)$$

รหัสคำสั่งเทียมที่ใช้ในการคำนวณค่า RANSAC แสดงได้ดังต่อไปนี้

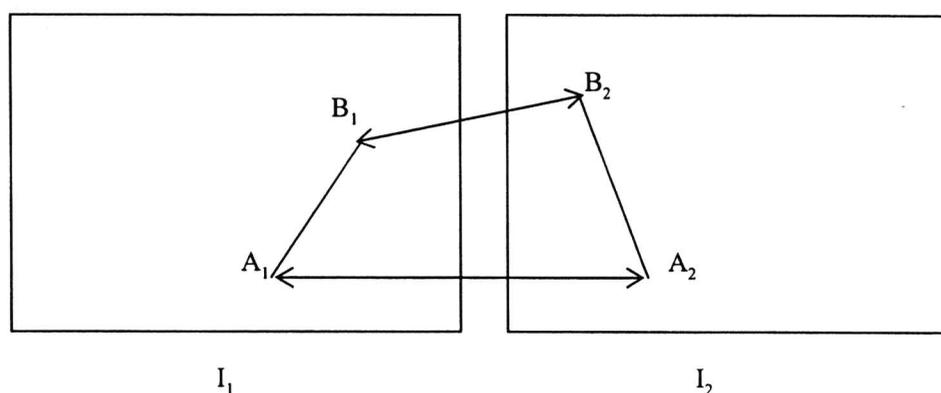
กำหนด $N \leftarrow \infty$ และ $sample_count \leftarrow 0$

ทำซ้ำเมื่อ $N > sample_count$

1. สุ่มจุดข้อมูล s จุดและกำหนดให้เป็นค่าเริ่มต้นของ model
2. กำหนดกลุ่มของ inliers เมื่อจุดข้อมูลอยู่ในค่าขีดแบ่งระยะทาง (t)
3. หาคำนวณน่าจะเป็นของ ϵ จากกลุ่มข้อมูลที่ผิดพลาด
4. กำหนดค่า $N \leftarrow \min(N, P(\epsilon, p))$
5. เพิ่มค่า $sample_count = sample_count + 1$

เลือกกลุ่มของ inliers ที่มีขนาดใหญ่ที่สุด และสร้าง model ขึ้นใหม่อีกครั้ง

เมื่อได้ model แล้วจะนำไปใช้ในการสร้างจุดร่วมที่สัมพันธ์กันระหว่างภาพสองภาพ ซึ่งจะต้องมีอัตราส่วนที่เท่ากัน การเชื่อมต่อภาพแบบ 2 มิติ จะอาศัยจุดร่วมจากภาพที่ 2 ย้ายตำแหน่งไปยังจุดร่วมในภาพที่ 1 ลักษณะการสร้างจุดร่วมเป็นไปดังภาพที่ 18



ภาพที่ 18 การใช้เส้นสองเส้นเพื่อปรับภาพ I_1 และ I_2 ให้เชื่อมต่อกัน

จากภาพที่ 18 จะมีเส้นตรงที่ลากผ่านจุด A_1 ไปยัง B_1 ในภาพ I_1 และ A_2 ไปยัง B_2 ในภาพ I_2 ซึ่งจะใช้เป็นจุดยึดในการหมุนภาพให้เส้นทั้งสองทาบกันได้พอดี คู่ของจุดภาพ A_1, A_2 และคู่ของจุดภาพ B_1, B_2 คือจุดภาพที่มีความคล้ายคลึงกัน ขั้นตอนในการแปลงภาพ สามารถทำได้ ดังนี้

และ [14] กล่าวว่า วิธีการ SIFT เป็นวิธีการหาลักษณะเด่นของภาพแบบท้องถิ่นที่มีประสิทธิภาพดี วิธีการเลือกจุดสนใจของภาพโดยการหาค่าสูงสุดหรือต่ำสุดจากการคำนวณด้วยวิธี DoG จุดสนใจที่พบจะปรากฏอยู่ตามขอบของภาพ หรือตามบริเวณเส้นชั้นความสูง จุดสนใจที่หาได้จากวิธีการนี้จะมีควมไวต่อการเปลี่ยนแปลงข้อมูลและสัญญาณรบกวน จึงเสนอวิธีในการเลือกจุดสนใจโดยใช้ affine Invariant Detectors เพื่อแก้ไขปัญหาดังกล่าว และได้ศึกษาเปรียบเทียบประสิทธิภาพของการหาพื้นที่สนใจด้วยวิธีการต่าง ๆ 6 วิธี คือ Harris-Affine, Hessian-Affine, MSER, IBR, EBR และ Salient Regions พบว่า ประสิทธิภาพการทำงานที่ดี คือ MSER, Hessian-Affine และ Harris-Affine ตามลำดับ วิธีการที่ทำงานได้ประสิทธิภาพต่ำสุดคือ Salient Regions

มีการพัฒนาและปรับปรุงวิธีการ SIFT อีกหลายงานวิจัย เช่น [10] ได้ปรับปรุงวิธีการ SIFT โดยใช้การสร้างลักษณะเด่นในพื้นที่ของวงกลม ทิศทางของจุดสนใจ จะมีทิศทางพุ่งออกจากจุดศูนย์กลาง เรียกว่า rotation-invariant generalization of SIFT (RIFT) ต่อมา [2] ได้พัฒนาวิธีการ Speeded-Up Robust Features (SURF) เพื่อเพิ่มประสิทธิภาพของ SIFT ให้มีความสามารถในการตอบสนองต่อขนาดและทิศทางของวัตถุที่เปลี่ยนแปลง โดยใช้วิธีการ Haar wavelet ในการค้นหาจุดสนใจ และ [9] ได้ปรับปรุงวิธีการ SIFT โดยลดขนาดของลักษณะเด่น จาก 128 ข้อมูล ให้เหลือ 36 ข้อมูล โดยใช้หลักการของ PCA มาช่วยในการลดมิติ เรียกว่า PCA-SIFT รวมทั้งงานวิจัย [20] ได้ปรับปรุงวิธีการหาจุดสนใจของ SIFT จากวิธีการ DoG เป็น FAST corner detector กำหนดขนาดของพื้นที่ย่อยของภาพแบบคงที่ และใช้ขนาดของลักษณะเด่น 36 มิติ