บทที่ 2
ทฤษฎีที่เกี่ยวข้อง

2.1 ไอเกนเฟซ

ไอเกนเฟซคือวิธีการสกัดคุณลักษณะเด่นวิธีการหนึ่งซึ่งเป็นที่นิยมในการนำไปประยุกต์ใช้ในด้านคอมพิวเตอร์วิทัศน์โดยผู้คิดค้นวิธีการสกัดคุณลักษณะเด่นนี้คือ L.Sirovich และ M.Kirby [4] ในปี 1987 และต่อมาในปี 1991 M.Turk และ A.Pentland [6] ได้นำวิธีการสกัดคุณลักษณะเด่นนี้ไปใช้ในงานการรู้จำใบหน้าคน (Face Recognition) ซึ่งได้ทำการทดสอบด้วยชุดข้อมูลใบหน้าที่มีแสงแตกต่างกันและผลของการรู้จำใบหน้ามีความถูกต้องถึง 96% วิธีไอเกนเฟซจะมีการใช้หลักการ PCA มาใช้โดยจะเลือกคุณลักษณะเด่นมาใช้เพื่อลดจำนวนมิติของข้อมูลเพื่อให้ประมวลได้เร็ว วิธีการนำไปประยุกต์ใช้ดังนี้

- เตรียมชุดข้อมูลเพื่อนำไปฝึกสอน โดยที่เริ่มต้นรูปฝึกสอนมีขนาด $P \times Q$ กำหนดให้ $S = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_M\}$ โดยที่ S คือเซตของรูปฝึกสอนที่ระดับสีเทา (Grayscale Image) ซึ่งมีจำนวนรูป M รูป โดยที่ Γ_k เป็นคอลัมน์เวกเตอร์ของรูปฝึกสอนขนาด $N \times 1$ ซึ่ง $N = P \times Q$
- หาค่าเฉลี่ยข้อมูลฝึกสอน Ψ ดังสมการที่ 2.1

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i \tag{2.1}$$

- หาค่าความแตกต่าง Φ ระหว่างข้อมูลภาพและค่าเฉลี่ยข้อมูลฝึกสอนดังสมการที่ 2.2

$$\Phi_i = \Gamma_i - \Psi \tag{2.2}$$

- คำนวณหาเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix; C) ดังสมการที่ 2.3

$$C = \frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi^T = AA^T \tag{2.3}$$

โดยที่

$$A = \left[\Phi_1, \Phi_2, \ldots \Phi_M\right]_{N \times M} \tag{2.4}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

- คำนวณหาไอเกนเวกเตอร์ (Eigenvector) และค่าไอเกน (Eigenvalue) ของเมตริกซ์ความแปรปรวนร่วม

$$AA^T v_i = \lambda_i v_i \tag{2.5}$$

โดยที่ v_i เป็นไอเกนเวกเตอร์ของเมตริกซ์ความแปรปรวนร่วม AA^T และ λ_i คือค่าไอเกนสำหรับไอเกนเวกเตอร์ v_i นั้น ไอเกนเวกเตอร์ของเมตริกซ์ความแปรปรวนร่วมหรือใบหน้าที่มีลักษณะพิเศษของรูปฝึกสอนจะถูกเรียกว่าไอเกนเฟซ จากภาพที่ 2.1 ได้แสดงตัวอย่างรูปไอเกนเฟซจากการวิจัยของ M.Turk และ A.Pentland แต่เนื่องจากการหาไอเกนเวกเตอร์ของเมตริกซ์ความแปรปรวนร่วม AA^T โดยตรงนั้นมีความซับซ้อนเชิงการคำนวณมาก เพราะเมตริกซ์ความแปรปรวนร่วมที่มีขนาด $N \times N$ นั้นเป็นเมตริกซ์ที่มีขนาดใหญ่ จึงได้มีการใช้เทคนิคทางคณิตศาสตร์เพื่อลดเวลาในการคำนวณหาไอเกนเวกเตอร์โดยหาไอเกนเวกเตอร์ของเมตริกซ์ $A^T A$ แทน ซึ่งมีขนาดเพียงแค่ $M \times M$ ดังแสดงในสมการที่ 2.6 ซึ่งค่าไอเกนและไอเกนเวกเตอร์ของเมตริกซ์ $A^T A$ ที่ได้จากถูกแทนด้วย λ_i และ u_i เมื่อพิจารณาสมการ 2.6 หากนำเมตริกซ์ A คูณเข้าด้านหน้าของทั้งสองข้างของสมการจะได้ผลลัพธ์ดังสมการ 2.7 และเมื่อเปรียบเทียบสมการ 2.7 กับ 2.5 จะพบว่าเวกเตอร์ Au_i และค่า λ_i ก็จะเป็นไอเกนเวกเตอร์และค่าไอเกนของเมตริกซ์ความแปรปรวนร่วม AA^T ที่ต้องการนั่นเอง

$$AA^T u_i = \lambda_i u_i \tag{2.6}$$

$$AA^T Au_i = \lambda_i Au_i \tag{2.7}$$

- เลือกไอเกนเวกเตอร์ที่สำคัญซึ่งมีการกระจายตัวมากโดยดูจากค่าไอเกนที่มีค่ามาก หลังจากนั้นจะได้ไอเกนเวกเตอร์ R เวกเตอร์ซึ่งจำนวน $R \le M$ จะได้เซตของไอเกนเวกเตอร์ U ดังสมการ 2.8

$$U = [Au_1, Au_2, ..., Au_R]_{N \times R} \tag{2.8}$$

- หลังจากนั้นทำการสกัดคุณลักษณะหรือน้ำหนักของแต่ละไอเกนเฟซโดยการฉายข้อมูลฝึกสอนหรือข้อมูลทดสอบลงบนไอเกนเวกเตอร์หรือไอเกนเฟซที่หามาได้จากก่อนหน้านี้ ดังสมการที่ 2.9 โดยที่ Ω คือเซตของน้ำหนักของแต่ละไอเกนเฟซหรือคุณลักษณะเด่น U คือไอเกนเฟซ Γ คือรูปฝึกสอนหรือทดสอบ w_i คือน้ำหนักของแต่ละไอเกนเฟซที่ i

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$\Omega = U^T(\Gamma - \Psi) = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_R \end{bmatrix}_{Rx1} \tag{2.9}$$

ภาพที่ 2.1 ตัวอย่างไอเกนเฟซ [6]

2.2 พีระมิดฮิสโทแกรมของทิศทางของเกรเดียนต์

HOG (Histogram of Oriented Gradients) คือวิธีการสกัดคุณลักษณะเด่นวิธีหนึ่งซึ่งถูกนิยมนำไปใช้งานทางด้านการประมวลผลภาพและทางด้านคอมพิวเตอร์วิทัศน์ จุดประสงค์ก็เพื่อค้นหาวัตถุในภาพ แนวความคิดที่สำคัญของ HOG คือการปรากฏวัตถุและรูปร่างภายในภาพโดยสามารถอธิบายด้วยการกระจายตัวของค่าความเข้มของเกรเดียนต์และทิศของขอบ (Edge Direction)

ขั้นตอนในการสกัดคุณลักษณะเด่นของ HOG

- คำนวณหาค่าเกรเดียนต์ของภาพจากสมการ 2.10 - 2.12

$$\nabla f \approx \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right| \tag{2.10}$$

โดยที่

$$\frac{\partial f}{\partial x} = \begin{cases} f(x, y+1) - f(x, y) & , x = \{1, M\} \\ \dfrac{f(x, y+1) - f(x, y-1)}{2} & , 1 < x < M \end{cases} \tag{2.11}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$\frac{\partial f}{\partial y} = \begin{cases} f(x+1,y) - f(x,y) & , x = \{1, N\} \\ \dfrac{f(x+1,y) - f(x-1,y)}{2} & 1 < x < N \end{cases} \quad (2.12)$$

(ซึ่ง ∇f คือ เกรเดียนต์ของภาพ $\frac{\partial f}{\partial x}$ คือเกรเดียนต์แกนนอน $\frac{\partial f}{\partial y}$ คือเกรเดียนต์แกนตั้ง M คือจำนวนคอลัมน์ของภาพ และ N คือจำนวนแถวของภาพ

- คำนวณหามุมเกรเดียนต์ซึ่งโดยปกติแล้วมุมเกรเดียนต์จะอยู่ในช่วง $[0, \pi]$ หรือ $[0, 2\pi]$ขึ้นอยู่กับการนำไปใช้งาน ซึ่งสามารถคำนวณได้จากสมการ 2.13

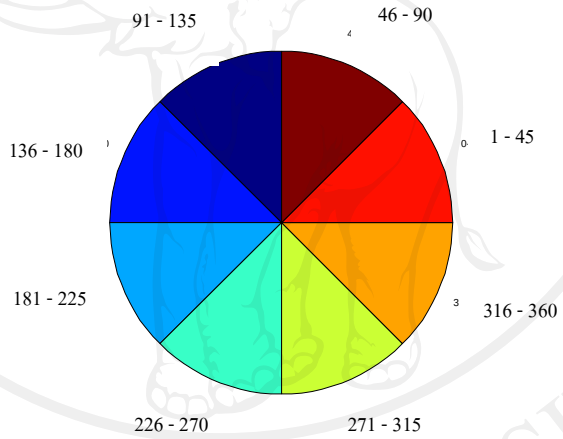
$$A = \arctan\left(\frac{\partial f / \partial y}{\partial f / \partial x}\right) \quad (2.13)$$

โดยที่ A คือ มุมของเกรเดียนต์

- คำนวณหาภาพขอบ (Canny Edge Detection) [18] และติดป้ายหมายเลขตรงที่เป็นขอบ
- กำหนดทิศ (bin) และแบ่งช่วงของมุมเกรเดียนต์ ยกตัวอย่างเช่น ถ้ากำหนดทิศคือ 8 ดังนั้นมุมของเกรเดียนต์จะถูกออกเป็นช่วงทิศดังนี้ช่วงทิศที่ 1 คือมุม 1-45 องศา ช่วงทิศที่ 2 คือ 46-90 องศา ช่วงทิศที่ 3 คือ 91-135 องศา ช่วงที่ 4 คือ 136-180 องศา ช่วงที่ 5 คือ 181-225 องศา ช่วงที่ 6 คือ 226-270 องศา ช่วงที่ 7 คือ 271-315 องศาและช่วงที่ 8 คือ 316-360 องศา ซึ่งแสดงตัวอย่างการแบ่งช่วงของมุมเกรเดียนต์ในภาพที่ 2.3
- หาทิศของมุมเกรเดียนต์ที่เป็นขอบของภาพ
- สร้างฮิสโทแกรมโดยค่าของฮิสโทแกรมคือผลรวมแมกนิจูดเกรเดียนต์ที่มีความสัมพันธ์กับขอบภาพและทิศของมุมเกรเดียนต์ดังนี้ ยกตัวอย่างเช่น จากภาพที่ 2.2 คือภาพตัวอย่างอินพุตในการหาฮิสโทแกรมและภาพที่ 2.4 คือค่าระดับสีเทาของภาพอินพุต จากนั้นนำภาพอินพุตไปคำนวณหาค่าและมุมของเกรเดียนต์ ผลของค่าและมุมของเกรเดียนต์ถูกแสดงในภาพที่ 2.6 และ 2.7 ขั้นตอนต่อไปคือการนำภาพไปหาขอบของภาพและติดป้ายหมายเลขตรงที่เป็นขอบภาพดังในภาพที่ 2.5 จากนั้นนำทิศและค่าของเกรเดียนต์ของขอบภาพดังภาพที่ 2.8 - 2.9 ไปสร้างฮิสโทแกรมและทำการนอร์มัลไลซ์ซึ่งผลลัพธ์ถูกแสดงอยู่ในภาพที่ 2.10 - 2.12

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ภาพที่ 2.2 รูปภาพอินพุต



ภาพที่ 2.3 ช่วงมุมเกรเดียนต์ ซึ่งกำหนดทิศคือ 8

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	255	255	255	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	0	0	0	0	0	0	0	0	0	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255

ภาพที่ 2.4 ค่าระดับสีเทาของรูปภาพอินพุต

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	2	2	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	0	0	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	0	0	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	0	0	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	0	0	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	0	0	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	2	2	2	2	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ภาพที่ 2.5 ผลลัพธ์ของการหาขอบและติดป้ายหมายเลข

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	128	128	128	128	128	128	128	128	128	0	0	0	0	0	0	0
0	0	0	0	128	180	128	128	128	128	128	128	128	180	128	0	0	0	0	0	0
0	0	0	0	128	128	0	0	0	0	0	0	128	128	0	0	0	0	0	0	0
0	0	0	0	128	128	0	0	128	128	128	0	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	180	128	180	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	128	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	0	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	0	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	0	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	0	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	128	128	128	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	128	180	128	180	128	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	0	128	128	128	0	0	128	128	0	0	0	0	0	0
0	0	0	0	128	128	0	0	0	0	0	0	0	128	128	0	0	0	0	0	0
0	0	0	0	128	180	128	128	128	128	128	128	128	180	128	0	0	0	0	0	0
0	0	0	0	0	128	128	128	128	128	128	128	128	128	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ภาพที่ 2.6 ค่าเกรเดียนต์

180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	180	90	90	90	90	90	90	90	90	90	180	180	180	180	180	180	180
180	180	180	180	0	45	90	90	90	90	90	90	90	135	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	270	270	270	180	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	225	270	315	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	0	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	135	90	45	0	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	90	90	90	180	180	180	180	180	180	180	180	180	180
180	180	180	180	0	0	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	0	315	270	270	270	270	270	270	270	225	180	180	180	180	180	180	180
180	180	180	180	180	270	270	270	270	270	270	270	270	270	180	180	180	180	180	180	180
180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180

ภาพที่ 2.7 มุมหรือทิศของเกรเดียนต์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	2	2	2	2	2	2	2	2	0	0	0	0	0	0	0
0	0	0	0	8	1	0	0	0	0	0	0	2	3	0	0	0	0	0	0
0	0	0	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	5	6	7	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	0	0	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	0	0	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	0	0	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	0	0	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	0	0	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	4	4	4	8	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0
0	0	0	0	8	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0
0	0	0	0	8	8	0	0	0	0	0	0	4	4	0	0	0	0	0	0
0	0	0	0	0	7	6	6	6	6	6	6	6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ภาพที่ 2.8 ช่วงทิศที่สัมพันธ์กับขอบภาพและมุมเกรเดียนต์

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	128	128	128	128	128	128	128	128	0	0	0	0	0	0	0
0	0	0	0	128	180	0	0	0	0	0	0	128	180	0	0	0	0	0	0
0	0	0	0	128	0	0	0	0	0	0	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	0	0	0	0	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	180	128	180	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	0	0	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	0	0	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	0	0	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	0	0	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	0	0	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	128	128	128	128	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	0	0	0	0	0	0	128	0	0	0	0	0	0
0	0	0	0	128	0	0	0	0	0	0	0	0	128	0	0	0	0	0	0
0	0	0	0	128	128	0	0	0	0	0	0	0	128	0	0	0	0	0	0
0	0	0	0	0	180	128	128	128	128	128	128	128	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ภาพที่ 2.9 ค่าแมกนิจูดเกรเดียนต์ที่สัมพันธ์กับขอบภาพ

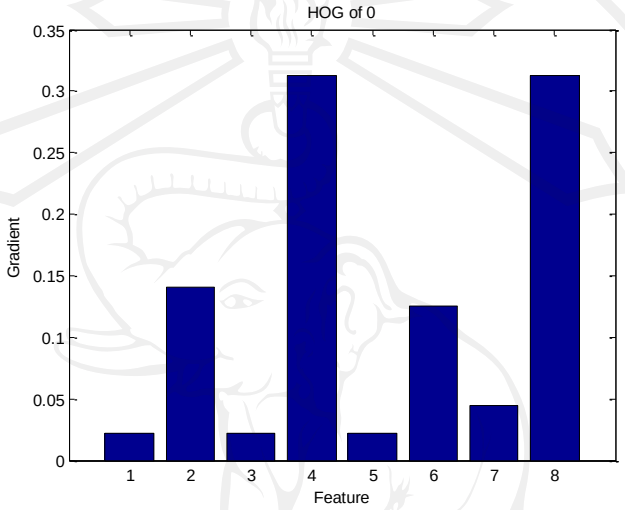
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

180.31	1147.50	180.31	2678.00	180.31	1020.00	360.62	2550.00

ภาพที่ 2.10 ผลรวมแมกนิจูดเกรเดียนต์ของแต่ละทิศเกรเดียนต์

0.02	0.14	0.02	0.32	0.02	0.12	0.04	0.31

ภาพที่ 2.11 ผลรวมแมกนิจูดเกรเดียนต์ของแต่ละทิศเกรเดียนต์ซึ่งถูกนอร์มัลไลซ์



ภาพที่ 2.12 ฮิสโทแกรมของภาพอินพุต

PHOG คือพีระมิดฮิสโทแกรมของทิศทางของเกรเดียนต์ซึ่งเป็นวิธีการสกัดคุณลักษณะเด่นที่นิยมนำไปประยุกต์ใช้งานในการจดจำวัตถุ [5] PHOG จะคล้ายกับการสกัดคุณลักษณะเด่นแบบ HOG แต่จะแตกต่างจาก HOG คือคุณลักษณะที่ได้จาก PHOG จะมีจำนวนมากกว่าหรือเท่ากับ เนื่องจาก PHOG จะมีการแบ่งระดับชั้น L ซึ่งจะทำการแบ่งพื้นที่ย่อยของภาพอินพุตลงไปเรื่อยๆ ยิ่งค่าระดับชั้นยิ่งเยอะ พื้นที่ย่อยก็มีจำนวนมาก ดังนั้นคุณลักษณะที่ได้ก็ยิ่งมีจำนวนมากขึ้น ในการแบ่งพื้นที่ย่อยจะสัมพันธ์กับระดับชั้น โดยที่พื้นที่ย่อยที่ได้จะมีจำนวนเป็น 4 ยกกำลังของระดับชั้นนั้นยกตัวอย่างเช่น ถ้าระดับชั้น $L=2$ จะแบ่งพื้นที่ย่อยได้ 16 ส่วนเท่าๆ กัน ดังภาพที่ 2.13

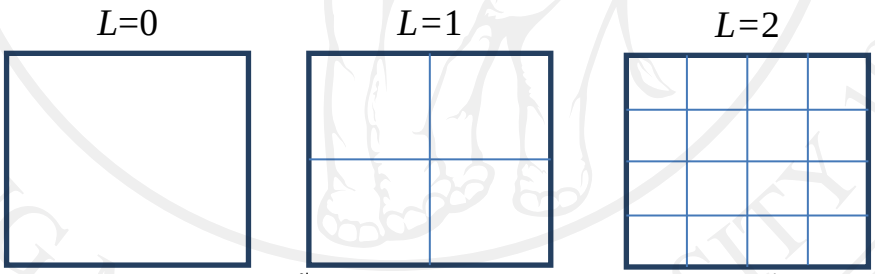
ส่วนในกรณีการหาจำนวนพื้นที่ย่อยทั้งหมดที่จะต้องสกัดคุณลักษณะเด่นของระดับชั้นนั้น คือพื้นที่ย่อยทั้งหมดที่ได้จากการแบ่งของระดับชั้นนั้นและระดับชั้นก่อนหน้า ซึ่งจะมีจำนวนเท่ากับ $\sum_{i=0}^{L} 4^i$ ยกตัวอย่างเช่น ถ้าระดับชั้น $L=2$ เพราะฉะนั้น จำนวนพื้นที่ย่อยทั้งหมดที่จะต้องสกัดคุณลักษณะเด่นของระดับชั้น คือ 21 พื้นที่ย่อยโดยพื้นที่ย่อยนี้ได้จากชั้น $L=0$ จำนวน 1 พื้นที่ซึ่งก็

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

คือไม่มีการแบ่งพื้นที่ย่อย $L=1$ จำนวน 4 พื้นที่ย่อยของภาพอินพุตและ $L=2$ จำนวน 16 พื้นที่ย่อยของภาพอินพุต ดังนั้นรวมทั้งหมด 2 1 พื้นที่ที่จะต้องนำไปสกัดหาคุณลักษณะเด่น

ส่วนจำนวนคุณลักษณะเด่นที่ได้ในแต่ละพื้นที่นั้นขึ้นอยู่กับทิศที่กำหนด ถ้ากำหนดทิศคือ 8 เพราะฉะนั้นพื้นที่ส่วนนั้นจะมีคุณลักษณะ 8 คุณลักษณะ ดังตัวอย่างภาพที่ 2.14 ซึ่งได้แสดงการแบ่งพื้นที่ย่อยในแต่ละระดับชั้นและจำนวนคุณลักษณะในแต่ละส่วนพื้นที่นั้น

จำนวนคุณลักษณะทั้งหมดที่ได้จาก PHOG จะขึ้นอยู่กับระดับชั้นและค่าทิศที่กำหนด ตัวอย่างเช่น ถ้าการสกัดคุณลักษณะคือ PHOG ระดับชั้น $L=2$ และกำหนดให้มีการแบ่งช่วงทิศทางออกเป็น 8 ช่วง จะได้จำนวนคุณลักษณะทั้งหมดจะมีจำนวน 168 คุณลักษณะ (8+32+128) ภาพที่ 2.15ได้แสดงรูปตัวอย่างภาพอินพุตซึ่งคือตราสัญลักษณ์ของผู้ผลิตรถยนต์และฮิสโทแกรมของคุณลักษณะที่ถูกสกัดโดย PHOG ระดับชั้น $L=2$ โดยที่แกน x คือคุณลักษณะที่ถูกสกัดออกมาได้ มีจำนวน 168 คุณลักษณะและ แกน y คือค่าแมกนิจูดเกรเดียนต์ของแต่ละคุณลักษณะ นอกจากนี้จะเห็นได้ว่าฮิสโทแกรมทั้งสองจะมีค่าแมกนิจูดเกรเดียนต์ของคุณลักษณะที่ได้แตกต่างกันซึ่งสามารถนำไปประยุกต์ใช้ในการจดจำวัตถุได้

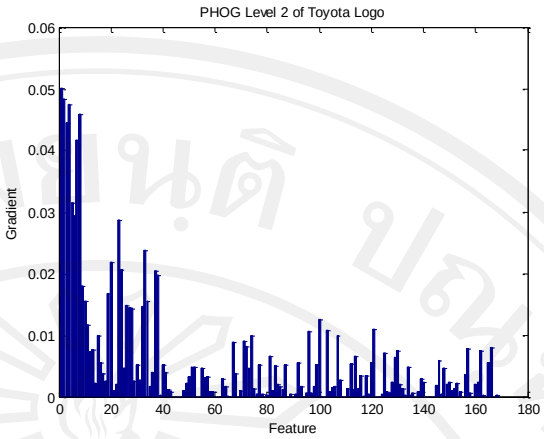


ภาพที่ 2.13 จำนวนพื้นที่ย่อยที่ได้จากการแบ่งของแต่ละระดับชั้น

ภาพที่ 2.14 คุณลักษณะที่ได้ของพื้นที่ย่อยในแต่ละระดับชั้น ถ้ากำหนดทิศทางคือ 8

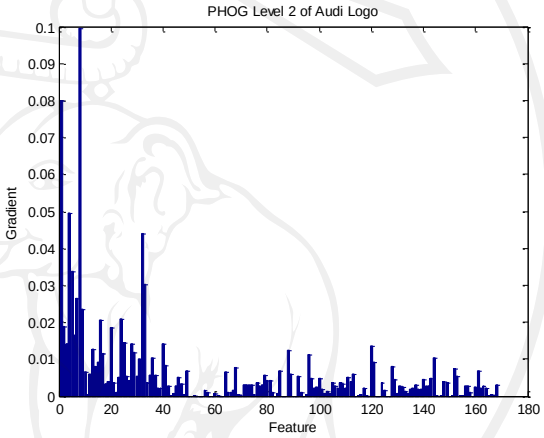
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

(ก)



(ข)

(ค)



(ง)

ภาพที่ 2.15 ตัวอย่างการหาคุณลักษณะ PHOG โดยแบ่งระดับชั้น $L = 2$ (ก) รูปตราสัญลักษณ์ Toyota (ข) ฮิสโทแกรมของรูปตราสัญลักษณ์ Toyota (ค) รูปตราสัญลักษณ์ Audi (ง) ฮิสโทแกรมของ
รูปตราสัญลักษณ์ Audi

2.3 Maximally Stable Extremal Regions (MSER)

MSER คือ พื้นที่ที่มีความเป็นเนื้อเดียวกันโดยที่พื้นที่รอบข้างมีความเข้มสีแตกต่างกัน MSER จะได้จากการแบ่งค่าขีดแบ่งความเข้มสีในรูปภาพระดับสีเทาโดยพื้นที่นั้นมีขนาดไม่เปลี่ยนแปลงเมื่อปรับเปลี่ยนค่าขีดแบ่งความเข้มสี แนวคิดของวิธีการนี้คือค่าความเข้มสีของพิกเซลในระดับภาพสีเทาซึ่งมีค่าต่ำกว่าค่าขีดแบ่งจะให้เป็นค่าสีดำ ส่วนค่าความเข้มสีของพิกเซลที่มีค่ามากกว่าหรือเท่ากับค่าขีดแบ่งจะให้เป็นสีขาวอย่าง เช่น ถ้าเรามองภาพต่อเนื่องเหมือนเราฉายหนังภาพๆหนึ่งตอนแรกกำหนดให้ค่าขีดแบ่งของความเข้มสีเป็น 0 ภาพแรกที่เราเห็นจะเป็นสีขาวทั้ง

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

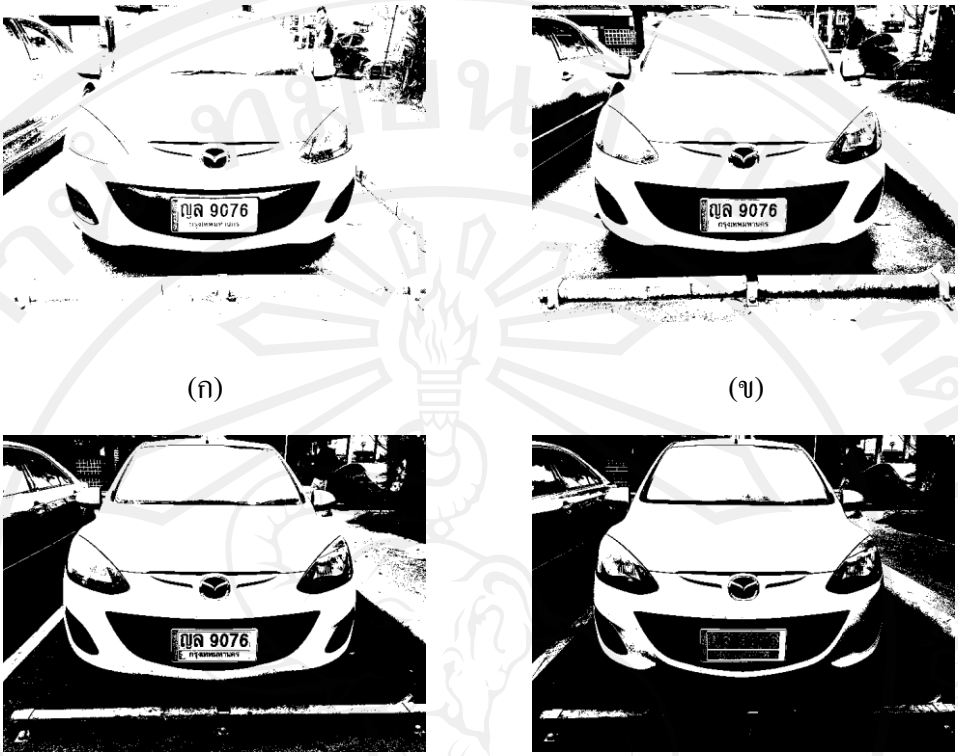
ภาพ ต่อมาเมื่อเพิ่มค่าขีดแบ่งมากขึ้น ค่าความเข้มสีของพิกเซลที่มีค่าน้อยกว่าค่าขีดแบ่งจะถูกกำหนดให้เป็นสีดำ ส่วนค่าความเข้มสีของพิกเซลที่มีค่ามากกว่าค่าขีดแบ่งจะมีค่าเป็นสีขาวเหมือนเดิม ดังนั้นเมื่อเพิ่มค่าขีดแบ่ง พิกเซลจะเป็นสีดำเพิ่มมากขึ้นจนในที่สุดภาพก็จะมีสีดำทั้งภาพ ยกตัวอย่างเช่น ดังภาพที่ 2.16 (ก) คือภาพอินพุตและภาพที่ 2.16 (ข) คือภาพระดับสีเทา จากนั้นเมื่อทำการปรับเปลี่ยนค่าขีดแบ่งเพิ่มขึ้นพิกเซลที่มีค่าต่ำกว่าค่าขีดแบ่ง พิกเซลนั้นจะเป็นสีดำ นอกนั้นให้เป็นสีขาว ลำดับภาพที่ได้จากการปรับเปลี่ยนค่าขีดแบ่งเพิ่มขึ้นถูกแสดงในภาพที่ 2.17 โดยที่ภาพที่ 2.17 (ก) ค่าขีดแบ่งคือ 40 ภาพที่ 2.17 (ข) ค่าขีดแบ่งคือ 63 ภาพที่ 2.17 (ค) ค่าขีดแบ่งคือ 100 และภาพที่ 2.17 (ง) ค่าขีดแบ่งคือ 126 ซึ่งจะเห็นได้ว่าตัวอักษรป้ายทะเบียนในภาพที่ 2.17 (ก) – 2.17 (ค) ยังไม่มีการเปลี่ยนแปลงแม้จะมีการปรับเปลี่ยนค่าขีดแบ่งแล้วก็ตาม

พื้นที่ (Regions) คือพื้นที่ที่มีค่าเข้มสีเหมือนกัน พื้นที่ขีดสุด (Extremal Regions) คือพื้นที่ที่มีความเข้มสีสูงหรือต่ำกว่าเมื่อเทียบกับบริเวณรอบๆ ส่วน MSER คือพื้นที่ขีดสุดที่ปรับเปลี่ยนค่าขีดแบ่งแล้วขนาดของพื้นที่นั้นไม่เปลี่ยนแปลงจากเดิมหรือพื้นที่ที่มีอัตราการเปลี่ยนแปลงของฟังก์ชันพื้นที่น้อยที่สุด

(ก) (ข)

ภาพที่ 2.16 รูปในการหาความเปลี่ยนแปลงค่าขีดแบ่งความเข้มสี
(ก) รูปภาพอินพุต (ข) รูปภาพระดับสีเทา

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



(ก) (ข)

(ค) (ง)

ภาพที่ 2.17 รูปที่ได้หลังจากเปลี่ยนแปลงค่าขีดแบ่งของความเข้มสี (ก) ค่าขีดแบ่งคือ 40 (ข) ค่าขีดแบ่งคือ 63 (ค) ค่าขีดแบ่งคือ 100 (ง) ค่าขีดแบ่งคือ 126

Image I is a mapping $I : \mathcal{D} \subset \mathbb{Z}^2 \rightarrow \mathcal{S}$. Extremal regions are well defined on images if:

1. $\mathcal{S}$ is totally ordered, i.e. reflexive, antisymmetric and transitive binary relation $\leq$ exists. In this paper only $S = \{0, 1, \ldots, 255\}$ is considered, but extremal regions can be defined on e.g. real-valued images ($\mathcal{S} = R$).
2. An adjacency (neighbourhood) relation $A \subset \mathcal{D} \times \mathcal{D}$ is defined. In this paper 4-neighbourhoods are used, i.e. $p, q \in \mathcal{D}$ are adjacent (pAq) iff $\sum_{i=1}^{d} |p_i - q_i| \leq 1$.

Region $\mathcal{Q}$ is a contiguous subset of $\mathcal{D}$, i.e. for each $p, q \in \mathcal{Q}$ there is a sequence $p, a_1, a_2, \ldots, a_n, q$ and $pAa_1, a_iAa_{i+1}, a_nAq$.

(Outer) Region Boundary $\partial\mathcal{Q} = \{q \in \mathcal{D} \setminus \mathcal{Q} : \exists p \in \mathcal{Q} : qAp\}$, i.e. the boundary $\partial\mathcal{Q}$ of $\mathcal{Q}$ is the set of pixels being adjacent to at least one pixel of $\mathcal{Q}$ but not belonging to $\mathcal{Q}$.

Extremal Region $\mathcal{Q} \subset D$ is a region such that for all $p \in \mathcal{Q}, q \in \partial\mathcal{Q} : I(p) > I(q)$ (maximum intensity region) or $I(p) < I(q)$ (minimum intensity region).

Maximally Stable Extremal Region (MSER). Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_{i-1}, \mathcal{Q}_i, \ldots$ be a sequence of nested extremal regions, i.e. $\mathcal{Q}_i \subset \mathcal{Q}_{i+1}$. Extremal region $\mathcal{Q}_{i^*}$ is maximally stable iff $q(i) = |\mathcal{Q}_{i+\Delta} \setminus \mathcal{Q}_{i-\Delta}|/|\mathcal{Q}_i|$ has a local minimum at i^* ($|.|$ denotes cardinality). $\Delta \in \mathcal{S}$ is a parameter of the method.

ภาพที่ 2.18 นิยามและคำจำกัดความของ MSER [19]

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

นิยาม หรือ คำจำกัดความของ MSER ถูกแสดงดังภาพที่ 2.18 และอธิบายดังนี้

- **รูปภาพ** I เป็นการแสดงความสัมพันธ์ระหว่างตำแหน่งพิกเซลในเซต $D \subset Z^2$ กับค่าความเข้มสี S โดยที่ S เป็นเซตของความเข้มสีที่เรียงลำดับ $S = \{1,2,3,...,255\}$
- **พื้นที่** Q คือพื้นที่โดยที่พิกเซล $p \in Q$
- **ขอบนอก (Outer Region Boundary)** คือเซตของพิกเซลที่ติดกับพิกเซลใดๆในเซต Q อย่างน้อย 1 ตัวแต่พิกเซลเหล่านั้นไม่ได้เป็นสมาชิกของเซต Q โดยที่พิกเซล $q \in \partial Q$
- **พื้นที่ขีดสุด** คือบริเวณ Q ที่เป็นสับเซตของ D ที่ทุกๆ $p \in Q, q \in \partial Q$ โดยที่ $I(p) > I(q)$ (บริเวณนั้นมีค่าความเข้มสูงเมื่อเทียบกับบริเวณรอบๆ) ซึ่งเรียกว่า Maximal Regions หรือ $I(p) < I(q)$ (บริเวณนั้นมีค่าความเข้มต่ำเมื่อเทียบกับบริเวณรอบๆ) ซึ่งเรียกว่า Minimal Regions
- **MSER** $Q_1,...Q_{i-1}, Q_i,...$ เป็นลำดับของในพื้นที่ขีดสุดที่ซ้อนกันไปเรื่อยๆโดยการปรับค่าขีดแบ่งความเข้มสี เช่น $Q_i \subset Q_{i+1}$ ซึ่งพื้นที่ที่เป็น MSER คือเซตพื้นที่ใน Q_{i^*} ที่ i^* เป็นจุดต่ำสุดสัมพัทธ์ของฟังก์ชัน $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}|/|Q_i|$ เมื่อ $|.|$ คือจำนวนสมาชิกของเซตนั้นๆ และ $\Delta \in S$ คือพารามิเตอร์ที่ใช้ในการปรับค่าขีดแบ่งความเข้มสี ซึ่งก็คือพื้นที่ขีดสุดที่ซึ่งมีอัตราการเปลี่ยนแปลงน้อยที่สุด

คุณสมบัติต่างๆของ MSER มีดังนี้

- ทนทานต่อการหมุน
- สามารถหาวัตถุได้ทุกขนาด
- เซตพื้นที่ขีดสุดทั้งหมดสามารถหาได้ด้วยความซับซ้อน $O(n \log(\log\ n))$ โดยที่ n คือจำนวนพิกเซลในภาพ

2.4 คอนโวลูชันและสหสัมพันธ์ (Convolution and Correlation)

คอนโวลูชันคือกระบวนการทางคณิตศาสตร์ระหว่าง 2 ฟังก์ชัน f กับ g ซึ่งจะได้ฟังก์ชันผลลัพธ์หนึ่งอันออกมา การทำคอนโวลูชันคือการอินทิเกรตของผลคูณของฟังก์ชันอินพุต $f(n)$ กับฟังก์ชันที่มีเอกลักษณ์เฉพาะตัว $g(n)$ หรือเรียกอีกอย่างว่า เคอร์เนลฟังก์ชัน (Kernel function) ที่สะท้อนและเลื่อนขนานและผลลัพธ์ที่ได้คือฟังก์ชันเอาต์พุต $y(n)$ นิยามของการคอนโวลูชัน 1 มิติสำหรับฟังก์ชันไม่ต่อเนื่องถูกแสดงในสมการ 2.14

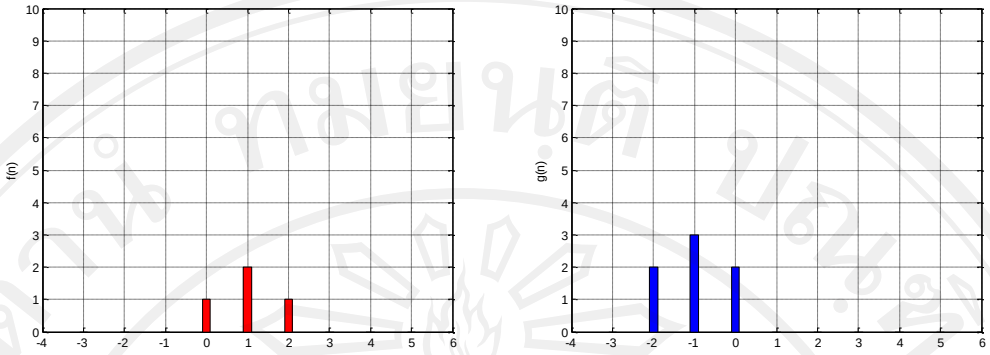
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$y(n) = f(n) * g(n) = \sum_j f(j)g(n-j) \qquad (2.14)$$

สหสัมพันธ์คือสถิติที่ใช้ในการหาความสัมพันธ์ระหว่างฟังก์ชัน 2 ฟังก์ชันหรือสัญญาณ 2 สัญญาณ เช่น การหาสหสัมพันธ์ของเงินเดือนพนักงานกับประสิทธิภาพในการทำงานของพนักงาน คุณสมบัติอย่างหนึ่งที่สำคัญที่ได้จากการหาค่าสหสัมพันธ์คือ ถ้าฟังก์ชันหนึ่งกับอีกฟังก์ชันหนึ่งที่นำมาหาความสัมพันธ์มีความสัมพันธ์หรือคล้ายกันมาก ค่าสหสัมพันธ์ก็จะมีค่ามากที่สุด ดังตัวอย่างในภาพที่ 2.19 (ก) คือฟังก์ชันที่หนึ่ง ภาพที่ 2.19 (ข) คือฟังก์ชันที่สอง การหาค่าสหสัมพันธ์ทำได้โดยการเลื่อนฟังก์ชันใดฟังก์ชันหนึ่งไปยังตำแหน่งที่เลื่อน n และทำการคำนวณตามสมการที่ 2.15 เพื่อหาค่าสหสัมพันธ์ในตำแหน่งที่เลื่อน n ดังตัวอย่างในภาพที่ 2.20 ภาพที่ 2.20 (ก) คือการหาสหสัมพันธ์โดยการเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=-1$ เมื่อทำการคำนวณพบว่าค่าสหสัมพันธ์เท่ากับ 0 และจากการสังเกตพบว่าการเลื่อนฟังก์ชันที่สองไปตำแหน่งที่ $n=-1$ ฟังก์ชันทั้งสองไม่ซ้อนทับกันทำให้ค่าสหสัมพันธ์ที่คำนวณได้มีค่าเท่ากับ 0 ดังนั้นเมื่อทำการเลื่อนฟังก์ชันที่สองไปยังตำแหน่งที่น้อยกว่า -1 ฟังก์ชันทั้งสองจะไม่ซ้อนทับกันดังนั้นค่าสหสัมพันธ์ที่ได้เมื่อเลื่อนฟังก์ชันที่สองไปยังในตำแหน่งที่น้อยกว่า -1 จะมีค่าสหสัมพันธ์เท่ากับ 0 ภาพที่ 2.20 (ข) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=0$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 2 ภาพที่ 2.20 (ค) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=1$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 7 ภาพที่ 2.20 (ง) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=2$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 10 ภาพที่ 2.20 (จ) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=3$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 7 ภาพที่ 2.20 (ฉ) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=4$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 2 ภาพที่ 2.20 (ช) คือการหาสหสัมพันธ์โดยเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n=5$ และค่าสหสัมพันธ์ที่ได้ในตำแหน่งนี้คือ 0 จากการสังเกตพบว่าเมื่อฟังก์ชันที่สองได้เลื่อนไปยังตำแหน่งนี้ฟังก์ชันทั้งสองไม่ซ้อนทับกันทำให้ค่าสหสัมพันธ์ที่ได้มีค่าเท่ากับ 0 ดังนั้นเมื่อทำการเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง $n>5$ ฟังก์ชันทั้งสองจะไม่ซ้อนทับกัน ค่าสหสัมพันธ์ที่ได้จะมีค่าเท่ากับ 0 ค่าสหสัมพันธ์ของแต่ละตำแหน่งถูกแสดงในภาพที่ 2.21 ได้แสดงค่าสหสัมพันธ์ของแต่ละตำแหน่งที่ฟังก์ชันที่สองเลื่อนและจากการสังเกตพบว่าในตำแหน่งที่ $n=2$ คือตำแหน่งที่ฟังก์ชันทั้งสองซ้อนทับกันพอดี ดังนั้นจึงมีค่าสหสัมพันธ์มากสุดตรงตามคุณสมบัติที่บอกไว้ในตอนแรก การหาสหสัมพันธ์ของสองฟังก์ชันจะมีนิยามที่คล้ายคลึงกับการหาคอนโวลูชันแต่แตกต่างกันตรงที่จะมีเพียงการเลื่อนแต่ไม่มีการสะท้อนดังแสดงในสมการ 2.15

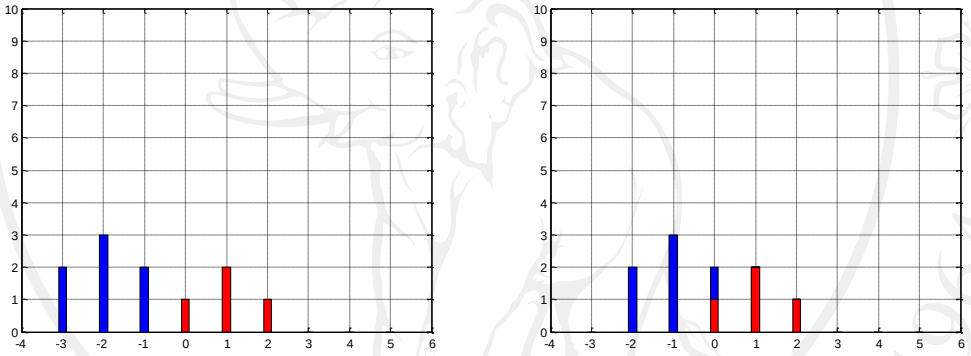
$$f(n) \otimes g(n) = f(n) * g(-n) = \sum_j f(j)g(n+j) \qquad (2.15)$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

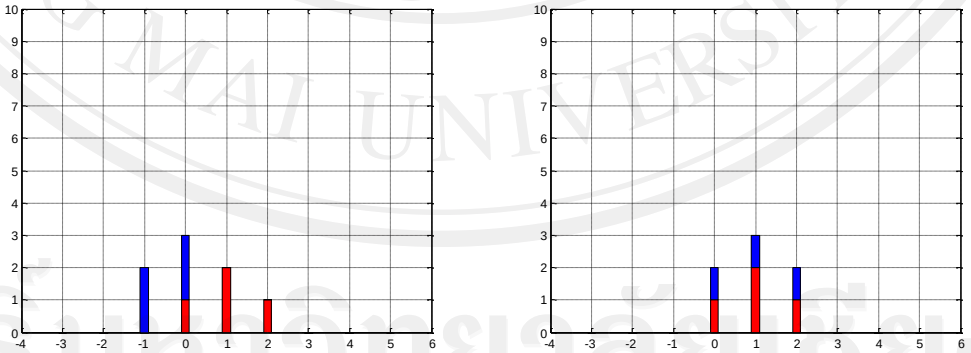


(ก) (ข)

ภาพที่ 2.19 ตัวอย่างฟังก์ชัน (ก) ฟังก์ชันที่หนึ่ง (ข) ฟังก์ชันที่สอง



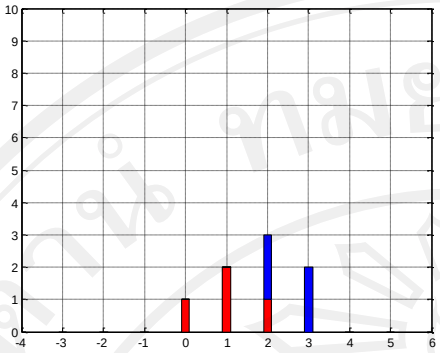
(ก) (ข)



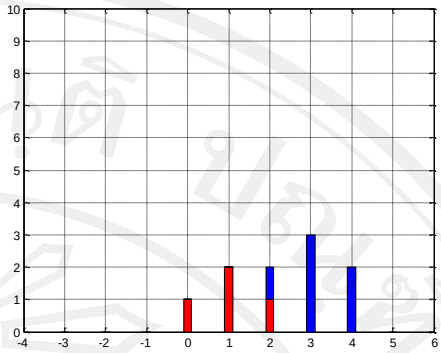
(ค) (ง)

ภาพที่ 2.20 การหาค่าสหสัมพันธ์โดยการเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง n (ก) $n \le -1$ (ข) $n = 0$ (ค) $n = 1$ (ง) $n = 2$ (จ) $n = 3$ (ฉ) $n = 4$ (ช) $n \ge 5$

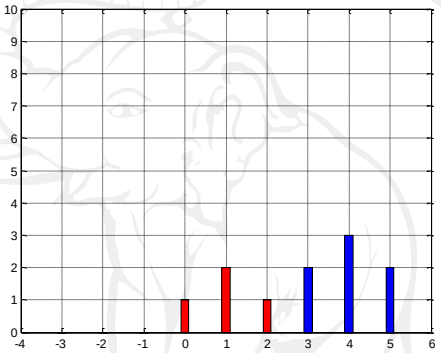
ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved



(จ)

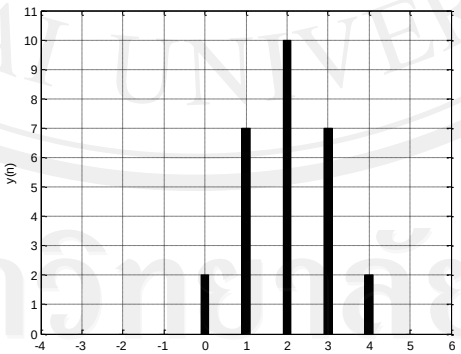


(ฉ)



(ช)

ภาพที่ 2.20 (ต่อ) การหาค่าสหสัมพันธ์โดยการเลื่อนฟังก์ชันที่สองไปยังตำแหน่ง n (ก) $n \le -1$ (ข) $n = 0$ (ค) $n = 1$ (ง) $n = 2$ (จ) $n = 3$ (ฉ) $n = 4$ (ช) $n \ge 5$



ภาพที่ 2.21 ค่าสหสัมพันธ์ในแต่ละตำแหน่งการเลื่อนของฟังก์ชันที่สอง

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.5 ความเด่นเชิงการมองโดยมีพื้นฐานมาจากกราฟ (Graph-Base Visual Saliency; GBVS)

ความเด่นเชิงการมองคือการหาความเด่นของภาพโดยใช้หลักการของการมองเห็น ซึ่งการหาความเด่นหลักการนี้มีอยู่หลากหลายวิธีและหลายทฤษฎีที่เกี่ยวข้อง เช่นวิธีของ Itti, Koch และ Niebur [20] โดยวิธีนี้จะทำการหาแผนที่คุณลักษณะ 3 ชนิด (Channel) คือ สี (Color) ความเข้ม (Intensity) และทิศทาง (Orientation) ของภาพอินพุต จากนั้นแผนที่คุณลักษณะสีและความเข้มจะถูกนำไปสร้างพีระมิดเกาส์เซียน (Gaussian Pyramid) ส่วนแผนที่ทิศทางจะถูกนำไปสร้างพีระมิดเกเบอร์ (Gabor Pyramid) ซึ่งจะทำให้ได้ภาพหรือแผนที่ที่มีขนาดแตกต่างกัน จากนั้นนำแผนที่ที่ได้จากการสร้างพีระมิดบางขนาดมาจับคู่และปรับขนาดให้เท่ากันแล้วทำการลบกันในแต่ละชนิดแผนที่ เมื่อนำแผนที่บางขนาดของแต่ละชนิดแผนที่มาลบกันแล้ว ให้แผนที่ของแต่ละชนิดมารวมกันให้เหลือเป็นแผนที่แผ่นเดียวในแต่ละชนิดซึ่งเรียกแผนที่หลังจากรวมกันแล้วว่าแผนที่ชัดแจ้ง (Conspicuity map) ต่อจากนั้นทำการหาแผนที่ความเด่น (Saliency map) โดยทำการหาค่าเฉลี่ยของแผนที่ชัดแจ้งและทำการปรับขนาดให้เท่ากับภาพอินพุต

GBVS คือการหาลักษณะความเด่นของภาพจากการมองเห็นวิธีหนึ่งโดยมีพื้นฐานมาจากกราฟซึ่งสร้างจากการมุ่งเน้นความสนใจไปที่แต่ละพิกเซลของภาพ ลักษณะความเด่นของแต่ละพิกเซลจะตอบสนองจากความแตกต่างของพิกเซลนั้นกับพิกเซลรอบๆ (Neighbor) รวมไปถึงระยะห่างจากพิกเซลนั้นกับพิกเซลรอบๆ ถ้าพิกเซลนั้นมีความแตกต่างที่มากและระยะห่างกับพิกเซลรอบๆ น้อย พิกเซลนั้นจะมีค่าความเด่นสูงมาก ในทางกลับกัน ถ้าพิกเซลนั้นไม่ค่อยมีความแตกต่างกับพิกเซลรอบๆ และมีระยะห่างกับพิกเซลรอบๆ มาก ดังนั้นพิกเซลนั้นจะไม่เป็นที่สนใจในการมองหรือพิกเซลนั้นมีความเด่นน้อย ผลลัพธ์ที่ได้จากการหาความเด่นของภาพจากการมองเห็นโดยมีพื้นฐานมาจากกราฟคือแผนที่ความเด่นซึ่งมีขนาดเท่ากับรูปภาพอินพุต

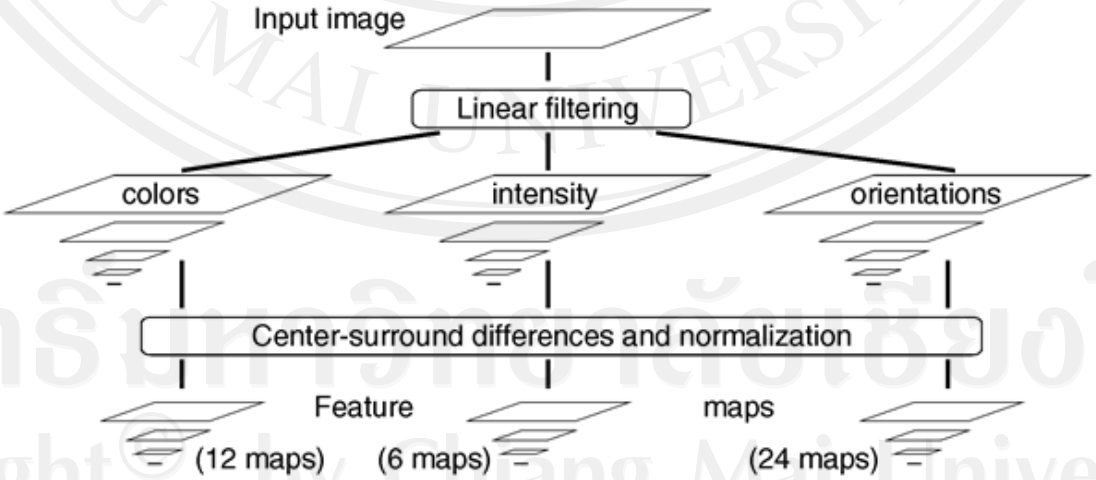
วิธีการหาความเด่นของภาพจากการมองเห็นโดยมีพื้นฐานจากกราฟแบ่งออกได้ 3 ขั้นตอนดังต่อไปนี้คือ

1) ขั้นตอนของการสกัดคุณลักษณะของภาพ (Extraction) เพื่อนำไปคำนวณหาค่าความเด่นของภาพ
2) ขั้นตอนการกระตุ้น (Activation) คือขั้นตอนในการนำคุณลักษณะที่ได้จากขั้นตอนก่อนหน้ามาคำนวณหาค่าความเด่นของแต่ละพิกเซลภาพ
3) ขั้นตอนการนอร์มัลไลซ์หรือการรวม (Normalization/Combination) คือขั้นตอนการปรับค่าความเด่นที่ได้ให้อยู่ในช่วงที่เหมาะสมเพื่อนำไปใช้ต่อไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.5.1 ขั้นตอนการสกัดคุณลักษณะ (Extraction)

ในขั้นตอนนี้คือการสกัดคุณลักษณะของภาพอินพุตและสร้างเป็นแผนที่คุณลักษณะในการหาความเด่นของภาพจากการมองเห็นโดยมีพื้นฐานมาจากกราฟจะมีขั้นตอนการสกัดเหมือนกับวิธีการสกัดคุณลักษณะเด่นของ Itti, Koch และ Niebur [20] โดยการสกัดคุณลักษณะจะแบ่งเป็น 3 ชนิดคือ การสกัดคุณลักษณะจากสี เนื่องจากบางคู่สีจะมีความไวในการตอบสนองของเซลล์ประสาทสูง ความเข้ม เนื่องจากเซลล์ประสาทของสัตว์เลี้ยงลูกด้วยนมมีความไวในการตอบสนองต่อพื้นที่ที่มีความแตกต่างความเข้มสี เช่น พื้นที่นั้นมืดแต่พื้นที่รอบๆ สว่างและการสกัดคุณลักษณะทิศทาง หลังจากนั้นคุณลักษณะของแต่ละชนิดซึ่งมีขนาดเท่ากับภาพอินพุตจะถูกนำไปสร้างเป็นพีระมิดเกาส์เซียน หลายระดับชั้น (Levels) $\sigma \in [0,8]$ ซึ่งจะได้รูปภาพที่เรียบขึ้นและมีหลายขนาด ขนาดของภาพจะขึ้นอยู่แต่ละระดับชั้นของพีระมิดเกาส์เซียนโดยขนาดของภาพจะมีอัตราส่วน $1:2^{\sigma}$ ของภาพอินพุต เช่นขนาดของรูปภาพของพีระมิดเกาส์เซียนในระดับชั้นที่ 1 จะมีขนาดเป็นครึ่งหนึ่งของขนาดภาพอินพุต พีระมิดเกาส์เซียนสามารถสร้างได้โดยการนำตัวกรองเกาส์เซียน (Gaussian filter) ไปคอนโวลูชันกับภาพอินพุตจากนั้นทำการย่อขนาดของภาพ (Subsampling) และทำตามขั้นตอนก่อนหน้านี้ซ้ำเพื่อได้รูปภาพหลายขนาด หลังจากที่สร้างพีระมิดเกาส์เซียนแล้ว ภาพบางขนาดในแต่ละชนิดจะถูกจับคู่และทำการปรับขนาดภาพที่มีขนาดเล็กให้มีขนาดเท่ากับขนาดใหญ่ภายในการจับคู่นั้น (Interpolation) และหาค่าความแตกต่างของภาพที่จับคู่นั้น กระบวนการนี้ถูกเรียกว่า “Center-Surround Differences” ซึ่งการปรับขนาดและหาค่าความแตกต่างจะถูกแทนด้วยสัญลักษณ์ Θ ผลลัพธ์ที่ได้คือแผนที่คุณลักษณะของแต่ละชนิดซึ่งแต่ละชนิดจะมีจำนวนแผนที่แตกต่างกัน ขั้นตอนในการสร้างแผนที่คุณลักษณะถูกสร้างเป็นแผนภาพดังภาพที่ 2.22



ภาพที่ 2.22 แผนภาพในการสร้างแผนที่คุณลักษณะ[20]

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.5.1.1 ขั้นตอนการสร้างแผนที่คุณลักษณะชนิดความเข้ม

ขั้นตอนแรกของการสร้างแผนที่ชนิดนี้คือการหาค่าความเข้มของภาพอินพุต โดยคำนวณค่าเฉลี่ยของค่าสีแดง สีเขียว สีน้ำเงินของภาพดังสมการที่ 2.16 โดยค่าของพิกเซลในภาพความเข้มสีจะต้องมากกว่า 1 ใน 10 ของค่าสูงสุดของพิกเซลในภาพความเข้มสีนั้น พิกเซลที่มีค่าน้อยกว่าให้กำหนดพิกเซลนั้นเป็น 0 จากนั้นสร้างพีระมิดเกาส์เซียนของรูปภาพความเข้ม ดังนั้นจะได้ภาพความเข้มสีหลากหลายขนาดซึ่งถูกแทนด้วย $I(\sigma)$ ขั้นตอนต่อไปคือการจับคู่ภาพความเข้มบางขนาดมาทำการปรับขนาดให้เท่ากับขนาดของภาพที่ใหญ่กว่าแล้วทำการหาความแตกต่างพิกเซลของภาพคู่นั้นซึ่งสามารถคำนวณได้จากสมการที่ 2.17 โดยที่ค่า c, s คือระดับชั้นในของพีระมิดเกาส์เซียน $c \in \{2,3,4\}$ และ $s = c + \delta, \delta \in \{3,4\}$ ผลลัพธ์หลังจากหาความแตกต่างพิกเซลคือแผนที่คุณลักษณะ ซึ่งในการหาแผนที่คุณลักษณะชนิดความเข้ม $I(c,s)$ จะได้ทั้งหมด 6 ภาพ

$$I = \frac{r+g+b}{3} \tag{2.16}$$

$$I(c,s) = \left| I(c) \Theta I(s) \right| \tag{2.17}$$

2.5.1.2 ขั้นตอนการสร้างแผนที่คุณลักษณะชนิดสี

การสร้างแผนที่คุณลักษณะชนิดสีนี้จะคล้ายกับการสร้างแผนที่คุณลักษณะชนิดความเข้ม ขั้นตอนแรกคือการสร้างภาพอินพุตขึ้นมาใหม่ 4 ภาพคือ

1) ภาพอินพุตสีแดง R
2) ภาพอินพุตสีเขียว G
3) ภาพอินพุตสีเหลือง Y
4) ภาพอินพุตสีน้ำเงิน B

ซึ่งสามารถคำนวณได้จากสมการ 2.18 - 2.21 โดยที่ r คือค่าสีแดงของภาพอินพุตตั้งต้น g คือค่าสีเขียวของภาพอินพุตตั้งต้น b คือค่าสีน้ำเงินของภาพอินพุตสีน้ำเงิน ค่าพิกเซลของแต่ละภาพอินพุตใหม่ที่มีค่าติดลบจะถูกปรับให้มีค่าเป็น 0 จากนั้นสร้างพีระมิดเกาส์เซียนของแต่ละภาพอินพุตใหม่โดยมีระดับชั้นเท่ากับขั้นตอนการสร้างแผนที่คุณลักษณะความเข้ม ทำการจับคู่ภาพอินพุตใหม่สีแดงกับสีเขียวบางขนาดและจับคู่ภาพอินพุตใหม่สีน้ำเงินกับสีเหลืองบางขนาด นำภาพแต่ละคู่ผ่านกระบวนการปรับขนาดและหาความแตกต่างของพิกเซลภาพคู่นั้น การจับคู่สีเหล่านี้เนื่องจากมีความไวการตอบสนองในการมองเห็นสูง ในการจับคู่สีภาพอินพุตบาง

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ขนาด ปรับขนาดและหาความแตกต่างของพิกเซลภาพคู่นั้นเพื่อสร้างแผนที่คุณลักษณะชนิดสี สามารถคำนวณได้จากสมการ 2.22 -2.23 แผนที่คุณลักษณะชนิดสี $RG(c,s)$ และ $BY(c,s)$ จะมีจำนวนรวมทั้งสิ้น 12 ภาพ โดยที่ขนาดภาพของระดับชั้นพีระมิด $c \in \{2,3,4\}$ และ $s = c + \delta, \delta \in \{3,4\}$

$$R = r - (g + b)/2 \tag{2.18}$$

$$G = g - (r + b)/2 \tag{2.19}$$

$$B = b - (r + g)/2 \tag{2.20}$$

$$Y = (r + g)/2 - |r - g|/2 - b \tag{2.21}$$

$$RG(c,s) = |(R(c) - G(c)) \Theta (G(s) - R(s))| \tag{2.22}$$

$$BY(c,s) = |(B(c) - Y(c)) \Theta (Y(s) - B(s))| \tag{2.23}$$

2.5.1.3 ขั้นตอนการสร้างแผนที่คุณลักษณะชนิดทิศทาง

เริ่มต้นจากสร้างพีระมิดเกเบอร์ จากภาพความเข้มอินพุตจำนวน 4 แบบ คือ

1) พีระมิดในทิศทางที่ $0°$
2) พีระมิดในทิศทางที่ $45°$
3) พีระมิดในทิศทางที่ $90°$
4) พีระมิดในทิศทางที่ $135°$

โดยพีระมิดแต่ละแบบจะมีระดับชั้นเท่ากับการสร้างแผนที่คุณลักษณะในขั้นตอนก่อนหน้า จากนั้นทำการจับคู่ภาพของแต่ละแบบพีระมิดบางขนาด ทำการปรับขนาดและหาความแตกต่างระหว่างพิกเซลภาพคู่นั้นเพื่อสร้างแผนที่คุณลักษณะชนิดทิศทาง โดยสามารถคำนวณได้จากสมการที่ 2.24 ซึ่ง c, s มีสมาชิกเหมือนกับขั้นตอนการสร้างแผนที่คุณลักษณะก่อนหน้า ดังนั้นแผนที่คุณลักษณะชนิดทิศทาง $O(c,s,\theta)$ จึงมีจำนวนทั้งหมด 24 แผนที่

$$O(c,s,\theta) = O(c,\theta) \Theta O(s,\theta) \tag{2.24}$$

2.5.2 ขั้นตอนการกระตุ้น (Activation)

ขั้นตอนนี้คือการสร้างแผนที่กระตุ้น (Activation Map) ที่ได้จากการคำนวณแผนที่คุณลักษณะแต่ละชนิด แผนที่กระตุ้นคือแผนที่ที่แสดงคุณลักษณะความเด่นในภาพว่าพิกเซลไหนมีความเด่น พิกเซลไหนไม่มีความเด่น ซึ่งจะสัมพันธ์กับความแตกต่างและระยะห่างพิกเซลนั้นกับ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

พิกเซลที่อยู่รอบๆ โดยที่ความแตกต่างของพิกเซลนั้น (i, j) กับพิกเซลรอบๆ (p,q) สามารถคำนวณได้จากสมการที่ 2.25 ซึ่งถูกเรียกว่า “Dissimilarity Function” หรือ $d((i, j) \| (p,q))$

$$d((i, j) \| (p,q)) = \left| \log \frac{M(i, j)}{M(p,q)} \right| \tag{2.25}$$

เมื่อ $M(i, j)$ เป็นค่าในพิกเซลของแผนที่คุณลักษณะที่จะหาความเด่นในพิกเซลนั้น และ $M(p,q)$ ค่าในพิกเซลอื่นๆ และระยะห่างของพิกเซลนั้นกับพิกเซลรอบๆ ซึ่งถูกแทนด้วย $F(a,b)$ สามารถคำนวณได้จากสมการที่ 2.26

$$F(a,b) = \exp\left(-\frac{a^2 + b^2}{2\sigma^2} \right) \tag{2.26}$$

โดยที่ a,b ระยะห่างในแนวแกนนอนและแกนตั้งตามลำดับ σ คือพารามิเตอร์ที่ปรับค่าได้อย่างอิสระ ก่อนที่จะทำการคำนวณหาความแตกต่างและระยะห่างระหว่างพิกเซลนั้นจะต้องทำการปรับขนาดแผนที่คุณลักษณะให้มีขนาดเท่ากันก่อน ในการสร้างแผนที่กระตุ้นนั้นจะใช้หลักการห่วงโซ่มาร์คอฟ (Markov Chain) โดยมองว่าพิกเซลคือโหนด (Node) หรือสถานะ (State) และมองความน่าจะเป็นในการเปลี่ยนแปลงจากโหนดหนึ่งไปอีกโหนดหนึ่งคือค่าน้ำหนัก w_1 ที่ได้จากความสัมพันธ์ของความแตกต่างและระยะห่างของพิกเซลนั้นกับพิกเซลอื่นซึ่งสามารถคำนวณได้จากสมการที่ 2.27

$$w_1((i, j),(p,q)) = d((i, j)) \| (p,q)) * F(i-p, j-q) \tag{2.27}$$

โดยจะมีการทำการนอร์มัลไลซ์ให้ค่าอยู่ในช่วง 0 ถึง 1 เพื่อใช้เป็นความน่าจะเป็น เมื่อได้ค่าความน่าจะเป็นของการเปลี่ยนแปลงจากโหนดหนึ่งไปโหนดหนึ่งทั้งหมดแล้ว ให้ทำการสร้างเป็นกราฟที่เชื่อมต่อกันทั้งหมดของพิกเซลนั้นกับพิกเซลอื่นๆ (Fully-Connected Graph) แทนสัญลักษณ์ด้วย G_A ต่อจากนั้นทำการหาความน่าจะเป็นของโหนดนั้นในสถานะคงที่ (Steady State) ซึ่งค่าที่ได้นั้นจะไม่เปลี่ยนแปลงจากสถานะก่อนหน้าหรือหลังจากนั้น ความน่าจะเป็นของโหนดนั้นในสถานะคงที่ก็จะถูกนำมาใช้เป็นค่าความเด่นของพิกเซลนั้นๆ นั่นเอง เมื่อได้ความเด่นในแต่ละพิกเซลทั้งหมดในแต่ละแผนที่คุณลักษณะจะถูกเรียกว่าแผนที่กระตุ้นที่แทนสัญลักษณ์ด้วย A

2.5.3 ขั้นตอนการนอร์มัลไลเซชันหรือการรวม (Normalization/Combination)

ในการนอร์มัลไลซ์คือการนำความน่าจะเป็นของแต่ละโหนดในสถานะคงที่ที่ได้จากการคำนวณก่อนหน้านี้ไปคำนวณโดยใช้ทฤษฎีของห่วงโซ่มาร์คอฟอีกหนึ่งครั้ง แต่ค่าน้ำหนักจากโหนดหนึ่งไปอีกโหนดหนึ่งจะถูกกำหนดดังสมการที่ 2.28

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$w_2((i,j),(p,q)) = A(p,q) * F(i-p, j-q) \quad (2.28)$$

โดยที่ $A(p,q)$ ค่าความน่าจะเป็นของโหนดอื่นของแผนที่กระตุ้น จากนั้นหาความน่าจะเป็นของแต่ละโหนดในสถานะคงที่อีกครั้งหนึ่ง เมื่อคำนวณความน่าจะเป็นของแต่ละโหนดในสถานะคงที่ของแต่ละแผนที่คุณลักษณะแต่ละชนิดแล้วซึ่งยังคงจำกันได้คือ แผนที่คุณลักษณะชนิดความเข้มมี 6 ภาพ แผนที่คุณลักษณะชนิดสีมี 12 ภาพและแผนที่คุณลักษณะชนิดทิศทางมี 24 ภาพ ให้ทำการเฉลี่ยแต่ละพิกเซลในแต่ละชนิดของแผนที่คุณลักษณะ เช่น เฉลี่ยแผนที่คุณลักษณะความเข้มโดยนำค่าของพิกเซลที่มีตำแหน่งตรงกันทั้ง 6 ภาพนำมารวมกันแล้วหาร 6 ก็จะได้ค่าเฉลี่ยของพิกเซลตำแหน่งนั้น ต่อจากนั้นเมื่อทำการเฉลี่ยพิกเซลครบทั้งหมดแล้วผลลัพธ์ที่ได้จะได้แผนที่คุณลักษณะชนิดความเข้มเฉลี่ย แผนที่คุณลักษณะชนิดสีเฉลี่ย และแผนที่คุณลักษณะชนิดทิศทางเฉลี่ยอย่างละ 1 ภาพ ในขั้นตอนรองสุดท้ายให้ทำการนำค่าพิกเซลที่ตรงกันในแต่ละชนิดแผนที่คุณลักษณะมาทำการรวมกันให้เหลือเพียงแผนที่เดียวคือแผนที่ความเด่น ในขั้นตอนสุดท้ายให้ทำการปรับขนาดให้เท่ากับขนาดภาพอินพุตเพื่อนำไปประยุกต์ใช้งานต่อไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved