

บทที่ 2

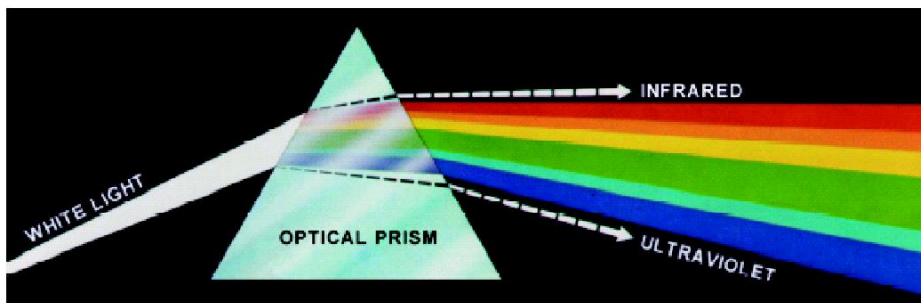
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ทฤษฎีสี (Color Model)

สิ่งต่างๆที่มนุษย์มองเห็นได้เป็นสีต่างๆเกิดจากการที่แสงจากแหล่งกำเนิดวิ่งไปตกกระทบแล้วสะท้อนกลับเข้าตาผู้สังเกตจึงสามารถมองเห็นวัตถุต่างๆมีสีได้ เมื่อมีการทำงานที่เกี่ยวกับสี ความซับซ้อนจึงเกิดขึ้น เพื่อความเข้าใจเกี่ยวกับสีจึงจำเป็นต้องมีการกำหนดระบบ และรูปแบบให้เป็นไปในทิศทางเดียวกัน เพื่อการสื่อสารระหว่างกันได้ โดยรูปแบบที่กำหนดขึ้นมานี้เรียกว่า รูปแบบของสี (Color Model) ซึ่งมีการใช้และยอมรับกันแพร่หลายอยู่ 4 แบบด้วยกัน ได้แก่

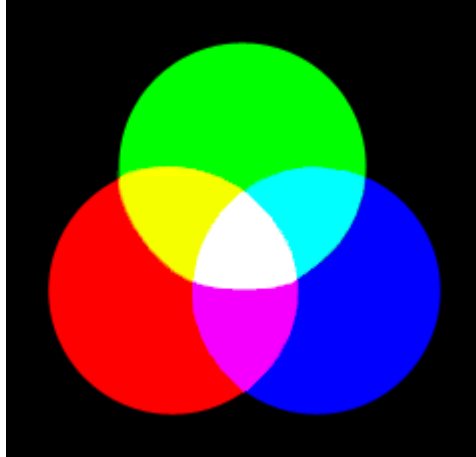
- ทฤษฎีแบบ HSI
- ทฤษฎีแบบ RGB
- ทฤษฎีแบบ CMYK
- ทฤษฎีแบบ LAB
- แต่ที่เกี่ยวข้องกับงานวิจัยนี้มีแค่ 1 รูปแบบ คือ ทฤษฎีแบบ RGB



ภาพที่ 2.1 แสดงภาพของการที่แสงขาวส่องผ่านแท่งปริซึม

ที่มา : Digital Image Processing (2nd Edition), Rafael C. Gonzalez and Richard E. Wood, 2002

รูปแบบสี RGB เป็นระบบสีของแสง ที่เกิดจากการหักเหของแสงสีขาวผ่านแท่งแก้วปริซึม ดังที่แสดงในภาพที่ 1 ซึ่งทำให้เกิดแถบสีรุ้ง (Spectrum) ซึ่งมี 7 สี ได้แก่ ม่วง คราม น้ำเงิน เขียว เหลือง แสด แดง โดยสามารถจัดเป็นพลังงานรูปหนึ่งซึ่งอยู่ในช่วงที่ตาคนเราสามารถมองเห็นได้ และเมื่อศึกษาลึกลงไปจะพบว่าสีทั้งหมดเกิดจากการผสมของสี 3 สี คือ สีแดง สีนํ้าเงิน และสีเขียว จึงถือได้ว่าทั้งสามสีนี้เป็นแม่สีของแสง ซึ่งคนเรานำเอาคุณสมบัติเหล่านี้มาใช้ประโยชน์ทั่วไป เช่น ในภาพวิดีโอ โทรทัศน์ จอคอมพิวเตอร์ ฯลฯ



ภาพที่ 2.2 แสดงภาพของแม่สีที่มีการทับซ้อนกัน

ที่มา : Color Theory for Digital Displays: A Quick Reference: Part II, Pabini Gabriel-Petit, 2006

จากภาพที่ 2 จะเห็นได้ว่าเมื่อแม่สีมีการทับซ้อนกัน (Color Overlap) จะเกิดสีใหม่ขึ้น เรียกสีแบบนี้ว่า แม่สีแบบบวก (Additive Color) การแปลงรูปภาพสีแบบ RGB ให้เป็นรูปภาพสีแบบอื่น (RGB Conversion) การแปลงรูปภาพสีแบบ RGB ให้เป็นรูปภาพระดับสีเทา การแปลงรูปภาพสีแบบ RGB ให้เป็นรูปภาพระดับเทานั้น สามารถทำได้โดยใช้ สมการที่ 1 ดังนี้

$$I = \frac{1}{3}(R + G + B) \quad (1)$$

โดยกำหนดให้ ตัวแปรแต่ละตัวมีความหมาย ดังนี้

I คือ จุดภาพหนึ่งๆ ของรูปภาพระดับเทา

R คือ จุดภาพหนึ่งๆ ในเมตริกซ์แผ่นสีแดง ของรูปภาพสีแบบ RGB

G คือ จุดภาพหนึ่งๆ ในเมตริกซ์แผ่นสีเขียว ของรูปภาพสีแบบ RGB

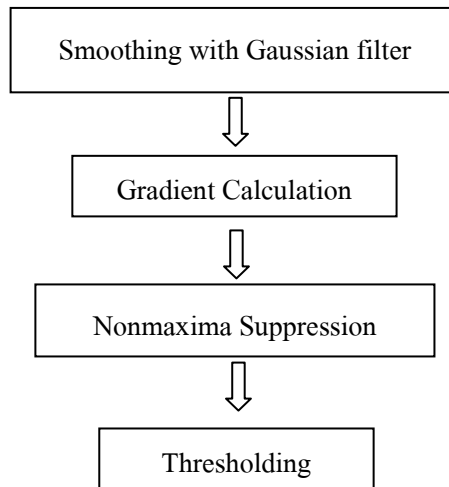
B คือ จุดภาพหนึ่งๆ ในเมตริกซ์แผ่นสีน้ำเงิน ของรูปภาพสีแบบ RGB

2.1.2 การหาขอบ (Edge Detection)

การหาขอบภาพ คือการหาเส้นรอบของวัตถุที่อยู่ในภาพ ซึ่งเมื่อทราบเส้นรอบวัตถุ เราสามารถนำไปประยุกต์ใช้งานได้อย่างมากมายเช่น การหาของวัตถุ หรือรู้จำชนิดของวัตถุ อย่างไรก็ตามการหาขอบของภาพที่สมบูรณ์เป็นเรื่องที่จำเป็น โดยเฉพาะการหาขอบของภาพที่มีคุณภาพต่ำ ภาพที่มีความแตกต่างระหว่างพื้นหน้ากับพื้นหลังน้อย หรือภาพที่มีความสม่ำเสมอทั้งภาพ

ขอบภาพเกิดจากความแตกต่างของแสงจากจุดหนึ่งไปยังอีกจุดหนึ่ง จากค่าความต่างนี้มีค่ามาก ขอบภาพก็จะเห็นได้ชัด แต่ถ้ามีความแตกต่างน้อยขอบภาพก็จะไม่ชัดเจน

ขั้นตอนการหาขอบโดยวิธีของ Canny ประกอบด้วย 4 ขั้นตอน ดังภาพที่ 3



ภาพที่ 2.3 แสดงขั้นตอนของ Canny edge detection

ที่มา : การจำแนกชนิดยาเม็ดจากภาพถ่าย โดยใช้เทคนิคเครือข่ายใยประสาท, จารวี ฉันทสิทธิ์พร, 2548

การทำงานของ Canny edge detection นั้นเริ่มต้นจากการปรับภาพให้เรียบ (Smoothing) ด้วยตัวกรองเกาส์เซียน (Gaussian filter) เพื่อกำจัดสัญญาณรบกวน หลังจากนั้น คำนวณค่าขนาด (magnitude) และทิศทาง (orientation) ของ gradient โดยใช้การหาอนุพันธ์อันดับหนึ่ง ในถัดมา จึงใช้ nonmaxima suppression กับ gradient magnitude เพื่อให้ได้ขอบที่บางลง และในขั้นตอนสุดท้ายใช้ double thresholding algorithm เพื่อระบุพิกเซลที่เป็นขอบและช่วยเชื่อมต่อขอบ (Green, Bill 2002 ; ION528 – Image processing algorithms 2005 ; Canny Operator Links 2005 ; Rubino, Matthew 2005) โดยในแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

2.1.3 Smoothing

ในขั้นตอนแรกของการหาขอบโดยอัลกอริทึมนี้จะต้องกำจัดสัญญาณรบกวนออกก่อน โดยใช้ Gaussian filter ซึ่งสามารถคำนวณได้จากการใช้กรอบ (mask) ขนาดเล็ก ขนาดของ Gaussian mask นี้หากมีขนาดกว้างจะมีผลทำให้ลดสัญญาณรบกวนได้มาก แต่ถ้ากว้างมากเกินไปจะมีผลทำให้ขอบย่อยๆ ที่เป็นส่วนรายละเอียดนั้นหายไป สำหรับการคำนวณหาภาพที่ได้จากการใช้ Gaussian filter เป็นดังสมการที่ 2

$$S[i, j] = G[i, j, \sigma] * I[i, j] \quad (2)$$

กำหนดให้

$I[i, j]$ เป็นภาพที่ต้องการหาขอบ

$G[i, j, \sigma]$ เป็น Gaussian smoothing filter

σ เป็น spread of the Gaussian (ควบคุมระดับของการ smoothing)

2.1.3.1 Gradient Calculation

ในขั้นแรกนำ smoothing image $S[i, j]$ มาสร้าง x, y partial derivatives $P[i, j]$ และ $Q[i, j]$ ตามลำดับ ดังสมการที่ 3 และ 4

$$P[i, j] \approx (S[i, j+1] - S[i, j] + S[i+1, j+1] - S[i+1, j]) / 2 \quad (3)$$

$$Q[i, j] \approx (S[i, j] - S[i+1, j] + S[i, j+1] - S[i+1, j+1]) / 2 \quad (4)$$

หลังจากนั้นนำค่า x, y partial derivatives มาคำนวณด้วยสูตรมาตรฐาน สำหรับการแปลงรูปแบบจาก rectangular ไปเป็น polar (rectangular-to-polar conversion) เพื่อหาขนาดและทิศทางของ gradient ตามสมการที่ 5

$$M[i, j] = \sqrt{P[i, j]^2 + Q[i, j]^2}$$

$$\theta[i, j] = \arctan(Q[i, j], P[i, j]) \quad (5)$$

จากสมการข้างต้นจะสามารถหาค่ามุม θ ออกมาได้เมื่อแทนค่าตัวแปรในฟังก์ชัน $\arctan(x, y)$

2.1.3.2 Nonmaxima Suppression

สำหรับการหาขอบโดย Canny method จุดที่ถือเป็นเส้นขอบได้นั้นต้องเป็นจุดที่ให้ค่าสูงสุดเฉพาะที่และเป็นทิศทางเดียวกับ gradient ด้วย ซึ่งด้วยวิธีดังกล่าวนี้ทำให้ได้ขอบที่บางเพียง 1 พิกเซล ภาพที่ได้หลังการทำ Nonmaxima Suppression จะให้ค่าเป็นศูนย์ในทุกจุดยกเว้นจุดที่เป็น local maxima points ซึ่งจะยังคงค่าเดิมไว้

2.1.3.3 Thresholding

แม้ว่าภาพจะผ่านการ smoothing ในขั้นตอนแรกแล้วก็ตาม ภาพที่ได้อาจยังมีเส้นขอบที่ไม่ใช่ขอบที่แท้จริงปรากฏอยู่เนื่องจากสัญญาณรบกวนหรือลักษณะของวัตถุในภาพเป็นพื้นผิวที่มีลวดลายหรือมีรายละเอียดภายในมาก ดังนั้นเพื่อลดปัญหาดังกล่าวจึงได้มีการกำหนดค่า threshold ขึ้นมา 2 ค่า คือ high threshold ($T1$) และ low threshold ($T2$) โดยพิกเซลที่มีค่ามากกว่า $T1$ จะถูกปรับเป็น 1 (เป็นพิกเซลที่เป็นขอบ) แต่ถ้าน้อยกว่า $T2$ จะถูกปรับเป็น 0 ส่วนค่าที่อยู่ระหว่างค่า threshold ทั้งสอง การปรับเป็นค่า 0 หรือ 1 นั้นขึ้นอยู่กับพิกเซลที่อยู่รอบข้าง หากพบว่าพิกเซลที่อยู่รอบข้างของพิกเซลที่เป็นขอบ (ค่า $> T1$) มีค่ามากกว่า $T2$ แล้ว จะปรับค่าพิกเซลดังกล่าวให้มีค่าเป็น 1 และถือเป็นสมาชิกหนึ่งในภาพขอบด้วยเช่นกัน



ภาพที่ 2.4 ตัวอย่างการหาขอบภาพแบบ Canny

ที่มา : Learning MATLAB Version 6, The Math Works Inc., 2001

2.1.4 การประมวลผลภาพเชิงสัณฐาน (Morphological Image Processing)

การประมวลผลภาพเชิงสัณฐานเป็นวิธีการประมวลผลภาพดิจิทัลพื้นฐานที่สำคัญ ซึ่งช่วยในการแสดงและอธิบายส่วนประกอบต่างๆของวัตถุในภาพเช่น พื้นที่ โครงร่าง(Skelton) ส่วนเว้าและส่วนโค้ง ในการประมวลผลภาพเชิงสัณฐานจะใช้ทฤษฎีของเซต (Set theory) ซึ่งเป็นพื้นฐานทางคณิตศาสตร์ในการประมวลผล ในการวิจัยนี้จะใช้การประมวลผลภาพเชิงสัณฐานกับภาพลักษณะฐานสองเท่านั้น ตัวดำเนินการ (Operator) ที่ใช้ในการประมวลผลภาพเชิงสัณฐานมีดังนี้

1) การขยายขนาด (Dilation)

การขยายขนาดเป็นตัวดำเนินการพื้นฐานของการประมวลผลภาพเชิงสัณฐาน ซึ่งใช้ในการเชื่อมวัตถุสองวัตถุที่อยู่ใกล้กัน หรือขาดออกจากกันและใช้ในการปิดรูหรือกำจัดสัญญาณรบกวน ที่เป็นจุดเล็กน้อยในภาพซึ่งเกิดจากการหาค่าขีดแบ่งในการแปลงภาพให้เป็นภาพลักษณะฐานสอง การขยายขนาดจะแสดงดังสมการที่ 6

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (6)$$

โดยที่ A เป็นภาพต้นฉบับ

B เป็นสมาชิกโครงสร้างที่ใช้ดำเนินการกับภาพ

Z เป็นตำแหน่งที่ใช้ในการเลื่อนสมาชิกโครงสร้าง

B' เป็นสมาชิกโครงสร้างที่ถูกพลิกกลับหัว

จากสมการที่ 8 พบว่าในการขยายขนาดจะมีสมาชิกโครงสร้าง (Structure element) มาดำเนินการกับภาพ ซึ่งจะทำให้การเลื่อน (Translation) สมาชิกโครงสร้างไปบนภาพโดยใช้จุดศูนย์กลางของสมาชิกโครงสร้างเป็นจุดหลักแล้วทำการขยายขนาดตามสมการที่ 6 กระบวนการขยายขนาดมีขั้นตอนดังนี้

- 1) เมื่อจุดศูนย์กลางของสมาชิกโครงสร้างอยู่ส่วนใดๆในภาพ ซึ่งไม่อยู่ตรงบริเวณของวัตถุจะไม่ดำเนินการใดๆ และทำการเลื่อนสมาชิกโครงสร้างไปยังจุดภาพถัดไป
- 2) เมื่อจุดศูนย์กลางของสมาชิกโครงสร้างตรงกับบริเวณของวัตถุจะดำเนินการด้วยตัวดำเนินการทางตรรกะ OR ระหว่างวัตถุกับสมาชิกโครงสร้าง ณ ตำแหน่งที่ตรงกับบริเวณของวัตถุ



ภาพที่ 2.5 ตัวอย่างภาพในการขยาย

ที่มา : Learning MATLAB Version 6, The Math Works Inc., 2001

ตัวอย่างของการขยายขนาดแสดงดังภาพที่ 5 โดยภาพที่ 5(ก) เป็นภาพต้นฉบับ ภาพที่ 5(ข) เป็นภาพที่ทำการขยายขนาดแล้ว

2) การกร่อน (Erosion)

การกร่อนเป็นตัวดำเนินการพื้นฐานของการประมวลผลภาพเชิงสัญญาณ เช่นเดียวกับการขยายขนาด ซึ่งใช้ในการลดขนาดวัตถุให้เล็กลง หรือใช้ในการกำจัดสัญญาณรบกวนที่เป็นจุดเล็กน้อยในภาพ การกร่อนจะแสดงดังสมการที่ 7

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (7)$$

โดยที่ A เป็นภาพต้นฉบับ

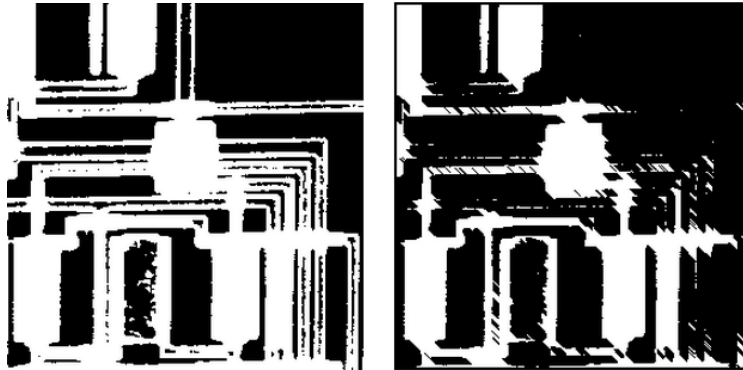
B เป็นสมาชิกโครงสร้างที่ใช้ดำเนินการกับภาพ

Z เป็นตำแหน่งที่ใช้เลื่อนสมาชิกโครงสร้าง

ขั้นตอนการกร่อนมีลักษณะเช่นเดียวกับการขยายขนาด ซึ่งมีขั้นตอนดังนี้

- 1) เมื่อจุดศูนย์กลางของสมาชิกโครงสร้างอยู่จุดใดๆ ในภาพซึ่งไม่อยู่ตรงบริเวณของวัตถุจะไม่ดำเนินการใดๆ และทำการเลื่อนสมาชิกโครงสร้างไปยังจุดภาพถัดไป
- 2) เมื่อจุดศูนย์กลางของสมาชิกโครงสร้างตรงกับบริเวณของวัตถุ จะทำการพิจารณาว่าจุดภาพของวัตถุเป็นสับเซต (Subset) ของวัตถุหรือไม่ ถ้าจุดภาพตรงกันทุกจุดจะ

ให้จุดศูนย์กลางของสมาชิกโครงสร้างเป็นบริเวณของวัตถุ แต่ถ้าจุดภาพไม่ตรงกันทุกจุด จะให้เป็นบริเวณพื้นหลังของวัตถุ



ภาพที่ 2.6 ตัวอย่างภาพในการกร่อน

ที่มา : Learning MATLAB Version 6, The Math Works Inc., 2001

ตัวอย่างการกร่อนแสดงดังรูปที่ 6 โดยรูปที่ 6(ก) เป็นภาพต้นฉบับรูปที่ 6(ข) เป็นภาพที่ทำการกร่อนแล้ว

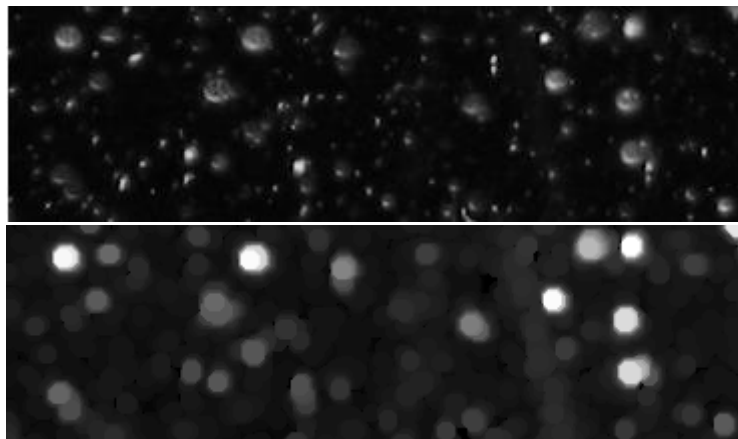
3) การเปิด (Opening)

การเปิดเป็นการนำตัวดำเนินการพื้นฐานทางการประมวลผลภาพเชิงสัญญาณมารวมกัน ซึ่งใช้ในการปรับเรียบขอบที่มีลักษณะเป็นหลุมการเปิด แสดงดังสมการที่ 8

$$A \circ B = (A \oplus B) \ominus B \quad (8)$$

โดยที่ A เป็นภาพต้นฉบับ

B เป็นสมาชิกโครงสร้างที่ใช้ดำเนินการกับภาพ



ภาพที่ 2.7 ตัวอย่างภาพที่ทำการเปิด (Opening)

ที่มา : Learning MATLAB Version 6, The Math Works Inc., 2001

4) การปิด (Closing)

การปิดเป็นการนำตัวดำเนินการพื้นฐานทางการประมวลผลภาพเชิงสัญญาณมา รวมกันเช่นเดียวกับการเปิด ซึ่งใช้ในการปรับเรียบขอบที่มีลักษณะเป็นเนิน การปิดแสดงดังสมการที่ 9

$$A \circ B = (A \oplus B) \ominus B \quad (9)$$

โดยที่ A เป็นภาพต้นฉบับ

B เป็นสมาชิกโครงสร้างที่ใช้ดำเนินการกับภาพ

2.1.5 การดึงลักษณะเด่น (Feature Extraction)

คือการวิเคราะห์เพื่อดึงข้อมูลภาพเฉพาะส่วนที่สนใจออกมาเป็นเวกเตอร์เพื่อนำไปเป็น ภาพอินพุตให้กับตัวคัดแยก (Classifier) ซึ่งในงานวิจัยนี้เราจะใช้ตัวอธิบายรูปร่างด้วยฟูรีเยร์ (Fourier Descriptor) เนื่องจากงานวิจัยนี้ ภาพลำดับด้านในของกึ่งมีโครงสร้างเป็นเส้นโค้ง แต่ลักษณะความโค้งนั้นมี ลักษณะที่แตกต่างกันเล็กน้อยในโดเมนพื้นที่ (Spatial Domain) พอทำให้แยกกลุ่มภาพออกมาได้จึงทำให้ สนใจที่จะนำข้อมูลขอบภาพ (Contour) มาวิเคราะห์ในโดเมนความถี่ (Spectral Domain) โดยใช้ทฤษฎี การแปลงฟูรีเยร์ (Fourier Transformation) ในการสร้างค่าการอธิบายรูปร่างด้วยฟูรีเยร์

การหาค่าการอธิบายรูปร่างด้วยฟูรีเยร์ เริ่มจากการหาขอบภาพของวัตถุก่อน แล้วจึง ทำการเรียงลำดับจุดภาพบนขอบภาพ ซึ่งหาได้ในลักษณะ $x(n) = (x(n)_1, x(n)_2)$ ให้อยู่ในรูปแบบ ตัวเลขเชิงซ้อน (Complex Number) หาได้จากสมการที่ 8

$$u(n) = x_1(n) + jx_2(n) \quad (10)$$

เนื่องจากขอบภาพที่หามาได้เป็นข้อมูลแบบดิสครีต โดยกำหนดให้จุดภาพที่ต้องการ ซักตัวอย่าง (Sampling) เป็น N แล้ว สามารถคำนวณหาค่าสัมประสิทธิ์ในโดเมนความถี่ $a(k)$ ของ $u(n)$ ทั้ง N ค่า ด้วยทฤษฎีการแปลงฟูรีเยร์แบบดิสครีต (Discrete Fourier Transform, DFT) หาได้จากสมการที่ 9

$$a(k) = \sum_{n=0}^{N-1} u(n)e^{-2\pi jkn/N} \quad (11)$$

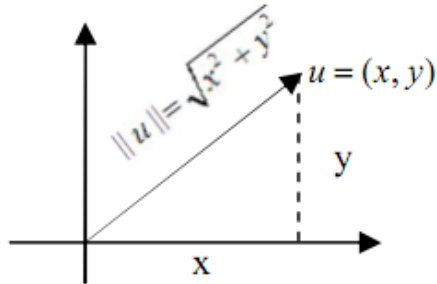
โดย $k = 0, 1, \dots, N-1$ เรียกค่าสัมประสิทธิ์ $a(k)$ ว่าเป็นการหาตัวอธิบายรูปร่าง ด้วยฟูรีเยร์ (Fourier Descriptor) การหาขนาดของ $a(k)$ หรือการหาขนาดของจำนวนเชิงซ้อน (Complex norm) ซึ่งเรียกได้ว่าเป็นค่า Magnitude หาได้จากสมการที่ 10

$$|z| = |x + jy| = \sqrt{x^2 + y^2} \quad (12)$$

2.1.6 ยูคลิเดียนเวกเตอร์นอร์ม (Euclidean Vector Norm)

เวกเตอร์จะถูกใช้เพื่อแสดงขนาด หรือทิศทางของสิ่งต่างๆ โดยจะเขียนเป็นสัญลักษณ์

ดังภาพที่ 2.8



ภาพที่ 2.8 แสดงการหาขนาดของเวกเตอร์ 2 มิติ

และหาค่าขนาดโดยใช้สมการที่ 13

$$\|u\| = \sqrt{x^2 + y^2} \quad (13)$$