Basic Segmentation and Binary image processing

Dr. Praisan Padungweang

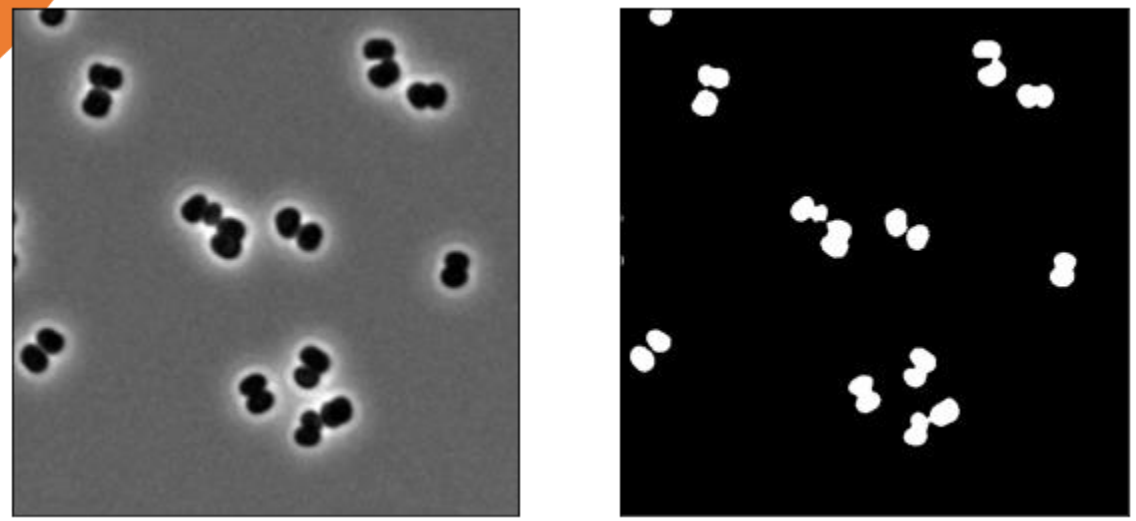
Credit: https://vincmazet.github.io/bip/segmentation/intro.html

BASIC SEGMENTATION

การแบ่งส่วนภาพคืออะไร?

https://vincmazet.github.io/bip/segmentation/intro.html

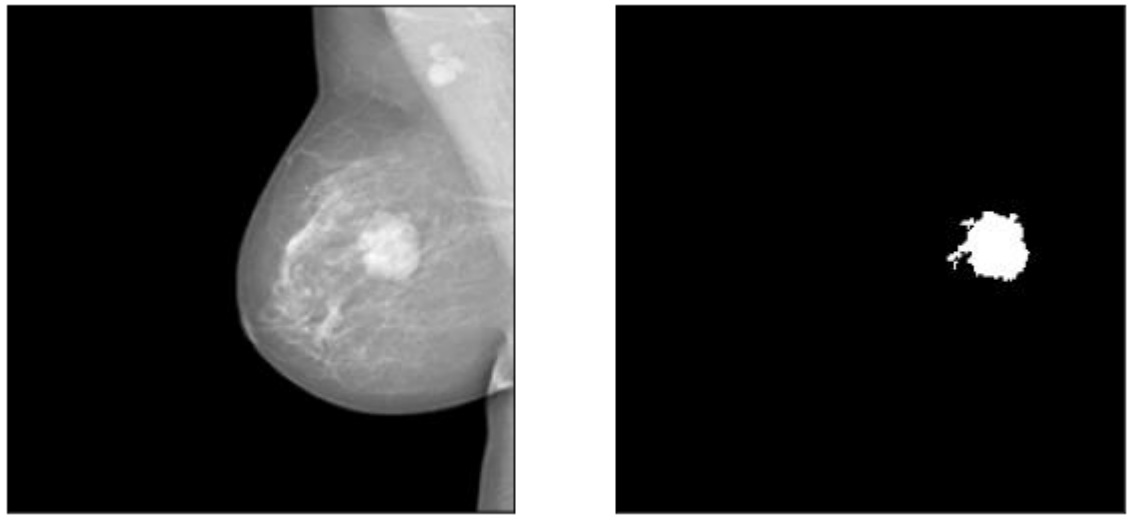
- การแบ่งส่วนภาพคือกระบวนการแบ่งภาพ f ตามเกณฑ์ที่กำหนด
- ภาพจะถูกแบ่งออกเป็นบริเวณ (Ri) ซึ่ง:
  - ไม่มีส่วนที่ซ้อนทับกัน
  - ครอบคลุมพื้นที่ทั้งหมดของภาพ
- พิกเซลในบริเวณเดียวกันต้องเป็นไปตามเกณฑ์ที่กำหนด
- พิกเซลในบริเวณที่ติดกันจะไม่เป็นไปตามเกณฑ์เดียวกัน



Example of segmentation on gray levels

Example of segmentation on color with a constraint of the region size

Example of segmentation on color



Example of segmentation on color

ข้อควรพิจารณา

- ผลลัพธ์การแบ่งส่วนไม่มีรูปแบบที่ตายตัว
- ผลลัพธ์แตกต่างกันขึ้นอยู่กับ:
  - เกณฑ์ที่เลือกใช้
  - วิธีการแบ่งส่วน
  - การกำหนดค่าเริ่มต้น
  - พารามิเตอร์อื่นๆ

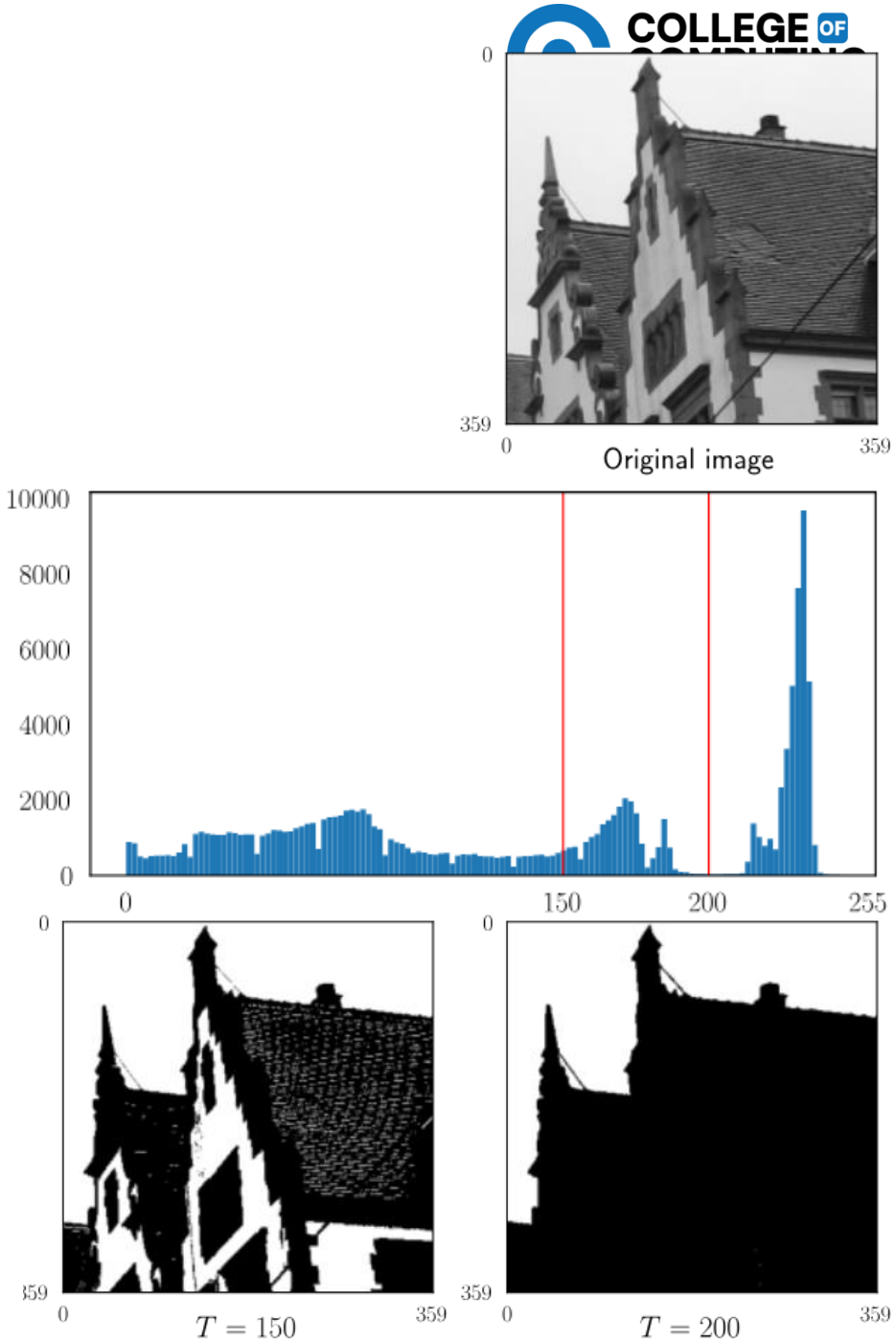
การแบ่งส่วนภาพด้วยฮิสโตแกรม

1. การกำหนดค่าขีดแบ่งแบบไบนารี

- เป็นวิธีการแบ่งส่วนอย่างง่ายที่กำหนดค่าไบนารีให้กับพิกเซลตาม:
  - ความเข้มของพิกเซล
  - ค่าขีดแบ่ง (T)

  $g(m,n) = 1$ ถ้า $f(m,n) \geq T$

  $g(m,n) = 0$ ถ้า $f(m,n) < T$
- ผลลัพธ์เป็นการแบ่งภาพระดับสีเทาเป็นสองกลุ่ม
- การเลือกค่าขีดแบ่งมีผลต่อคุณภาพการแบ่งส่วน
- การวิเคราะห์ฮิสโตแกรมช่วยในการเลือกค่าขีดแบ่ง



การแบ่งส่วนภาพด้วยฮิสโตแกรม

2. วิธีของ Otsu

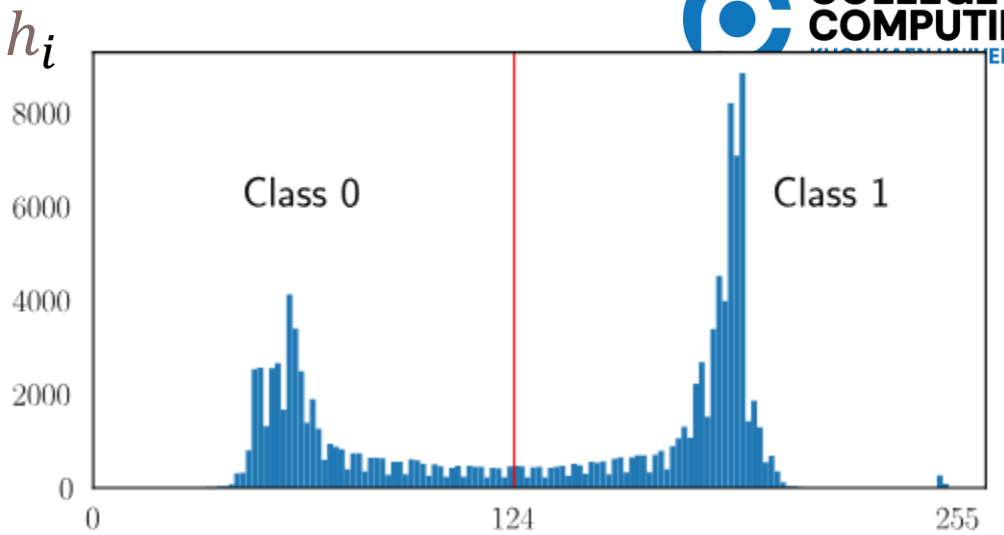
ภาพรวม

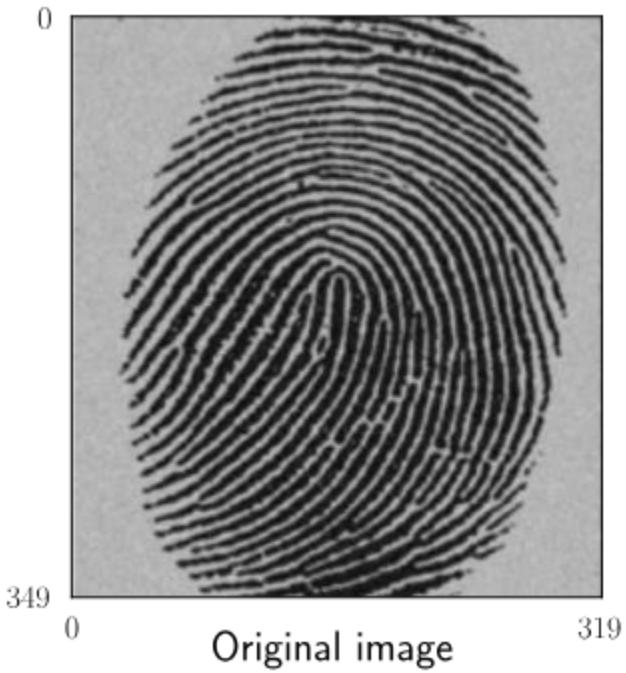
- เทคนิคการกำหนดค่าขีดแบ่งอัตโนมัติ (T)
- แบ่งฮิสโตแกรมเป็นสองกลุ่ม (0 และ 1) โดย

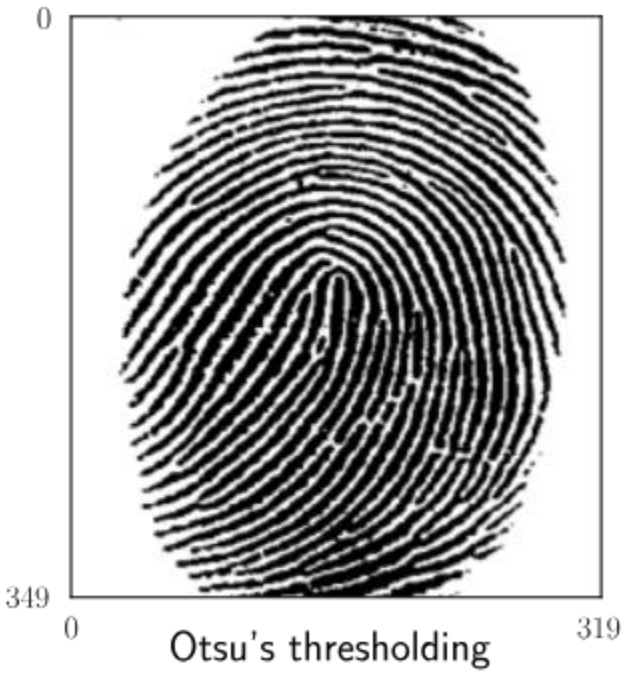
Minimize $\sigma_w^2(T)$

$$\sigma_w^2(T) = w_0(T)\sigma_0^2(T) + w_1(T)\sigma_1^2(T).$$

ทดลองทุดความเป็นไปได้ $T = \{0, \ldots, L-1\}$ แล้ว เลือกค่า T ที่ทำให้สมการมีค่าน้อยสุด



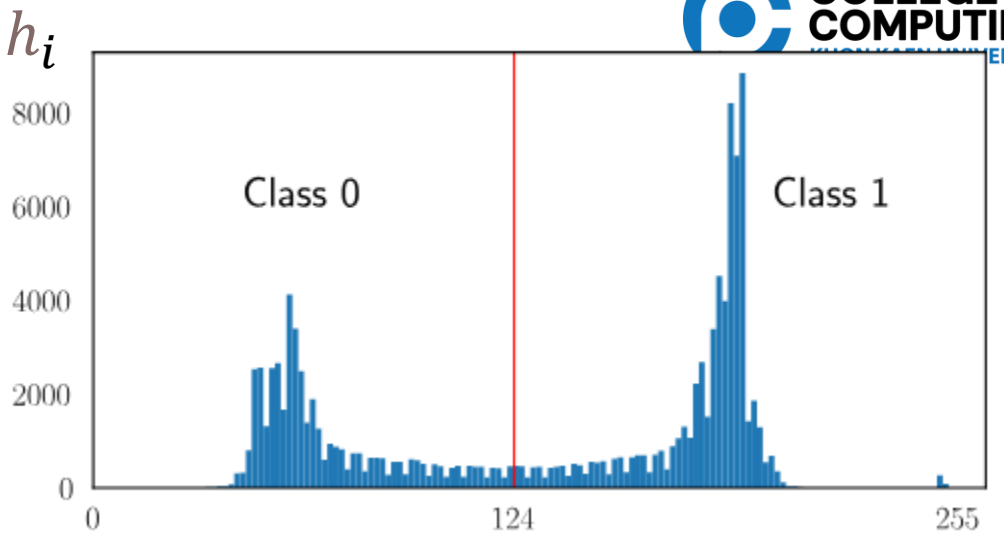




การแบ่งส่วนภาพด้วยฮิสโตแกรม

- 2. วิธีของ Otsu

$$\sigma_w^2(T) = w_0(T)\sigma_0^2(T) + w_1(T)\sigma_1^2(T).$$



Class 1

$$w_0(T) = \sum_{i=0}^{T} h(i)$$

$$\mu_0(T) = \frac{1}{w_0(T)} \sum_{i=0}^{T} ih(i)$$

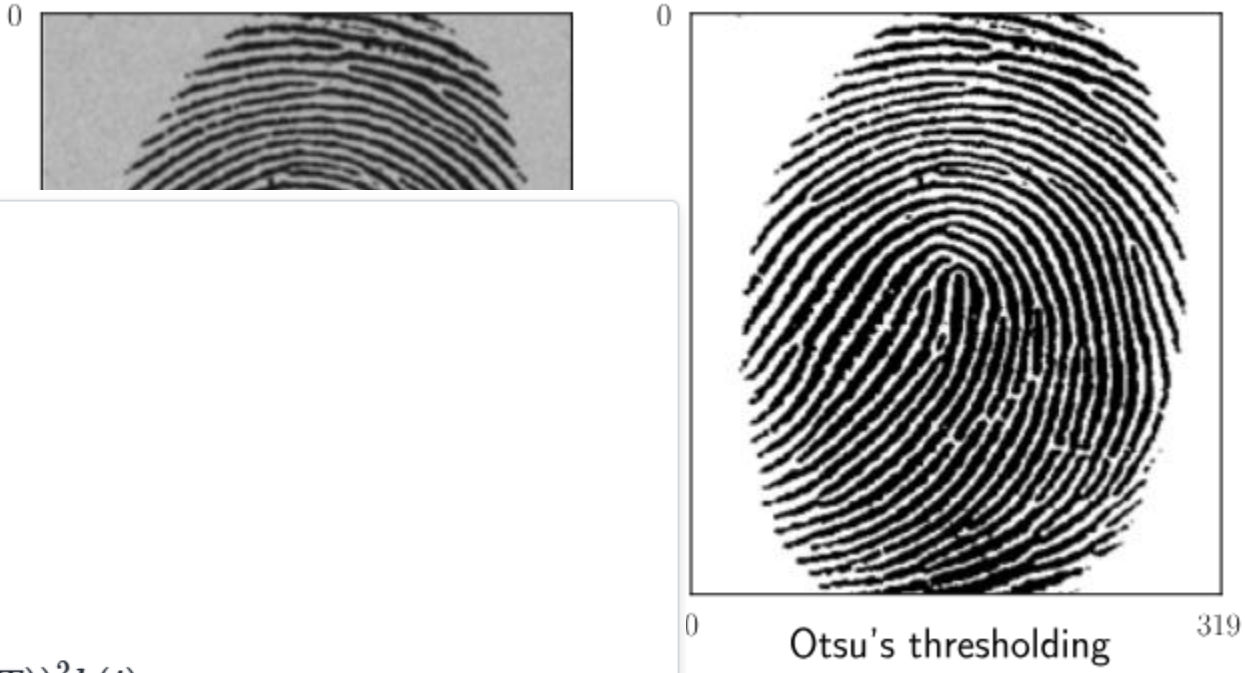
$$\sigma_0^2(T) = \frac{1}{w_0(T)} \sum_{i=0}^{T} (i - \mu_0(T))^2 h(i)$$

Class 2

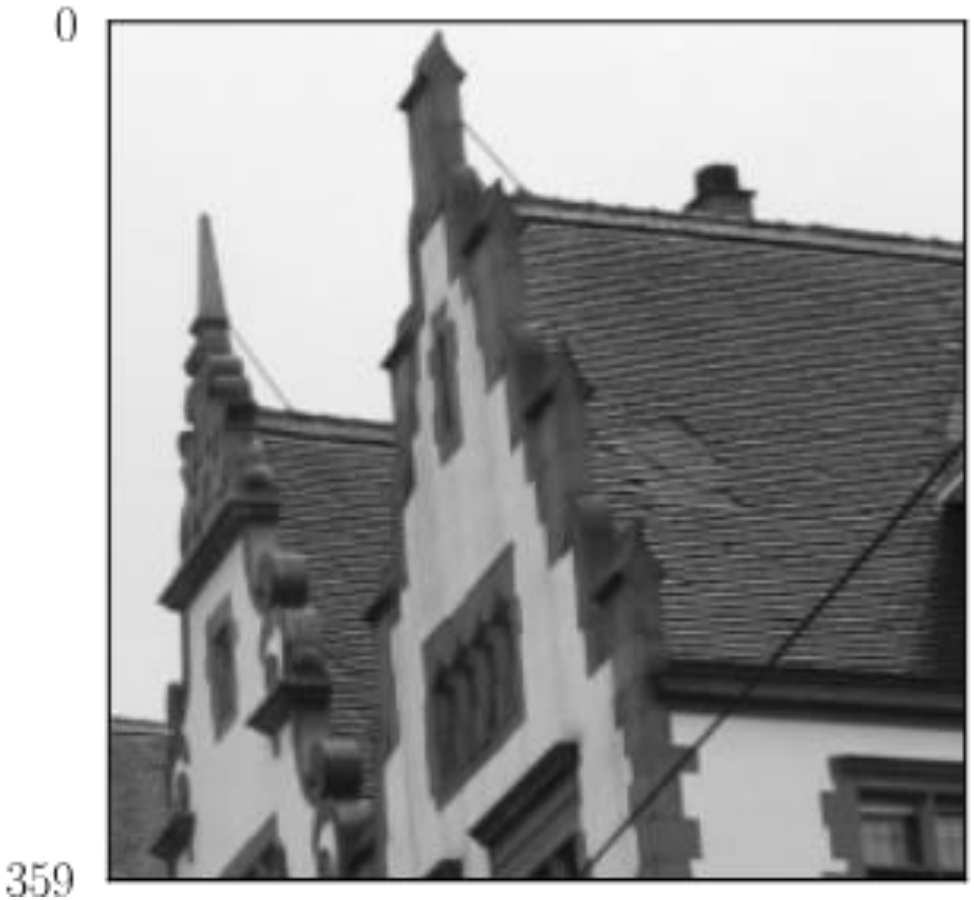
$$w_1(T) = \sum_{i=T+1}^{L-1} h(i)$$

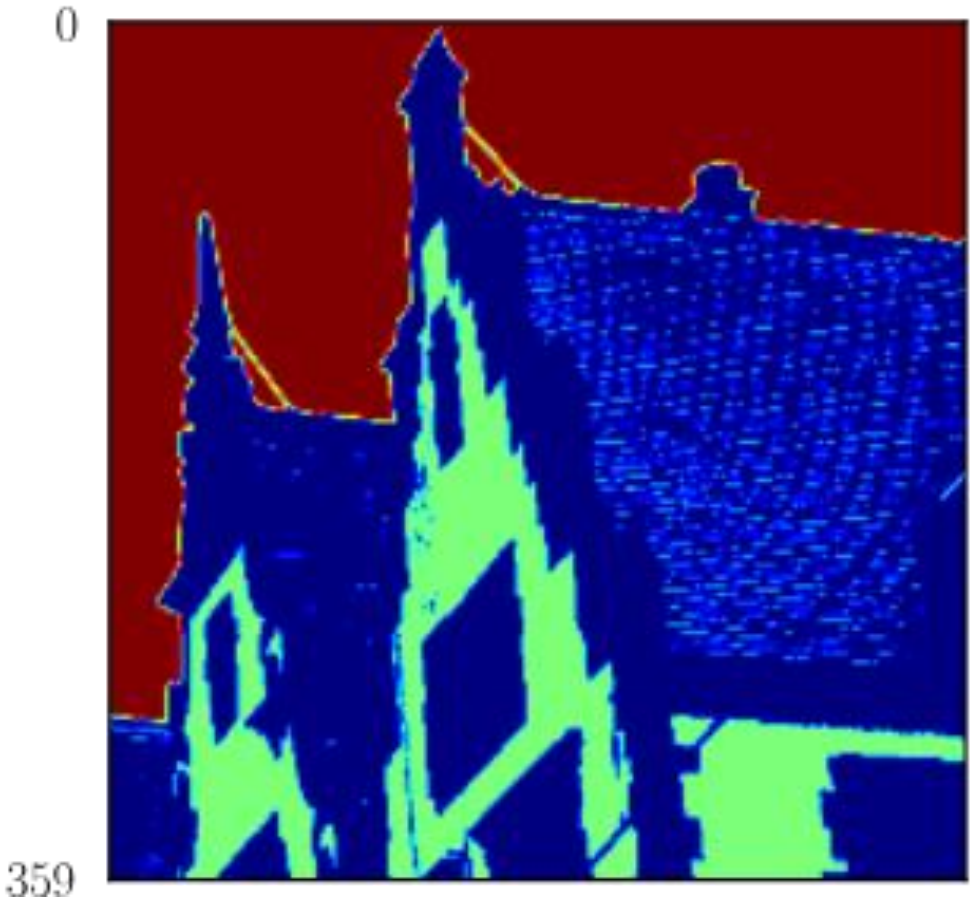
$$\mu_1(T) = \frac{1}{w_1(T)} \sum_{i=T+1}^{L-1} ih(i)$$

$$\sigma_1^2(T) = \frac{1}{w_1(T)} \sum_{i=T+1}^{L-1} (i - \mu_1(T))^2 h(i)$$



ภาพสำหรับฝึก
https://en.wikipedia.org/wiki/Otsu%27s_method#/media/File:Image_processing_pre_otsus_algorithm.jpg



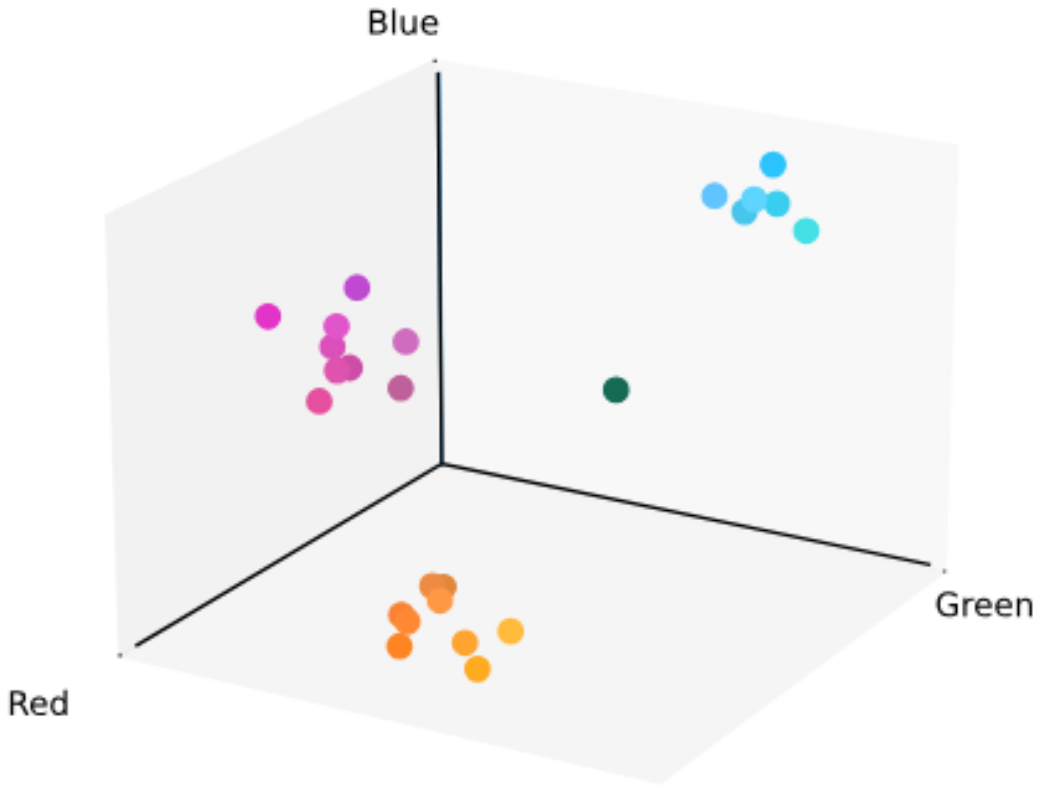


การใช้ค่าขีดแบ่งหลายค่า

- สามารถแบ่งส่วนเป็นหลายกลุ่ม
- สามารถขยายวิธีของ Otsu ได้
- ความซับซ้อนในการคำนวณเพิ่มขึ้นตามจำนวนกลุ่ม
- เหมาะสำหรับการแบ่งส่วนภาพที่ซับซ้อน

การแบ่งส่วนภาพด้วยการจัดกลุ่ม

- ใช้ได้กับภาพสีและภาพหลายแถบความถี่
- แทนพิกเซลด้วยจุดในปริภูมิ B มิติ
- สีที่คล้ายกันจะรวมกลุ่มในปริภูมินี้



การแบ่งส่วนภาพด้วยการจัดกลุ่ม

อัลกอริทึม K-means

Algorithm: K-means

1. Initialize (randomly) the K centroids
2. While the centroids vary:
    1. STEP A For each point:
        1. Calculate the distances from the point to all centroids
        2. Assign the point to the nearest group
    2. STEP B Calculate the centroid of each group

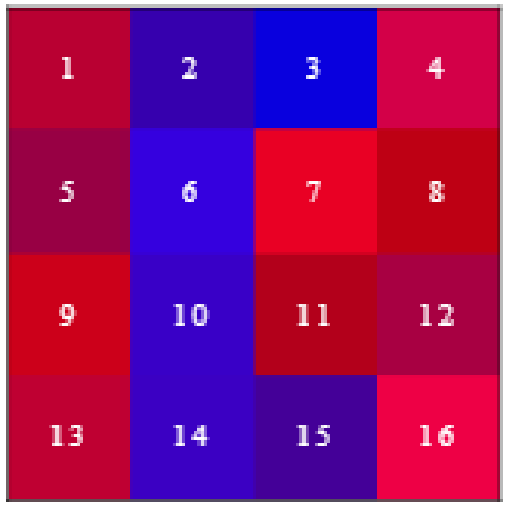
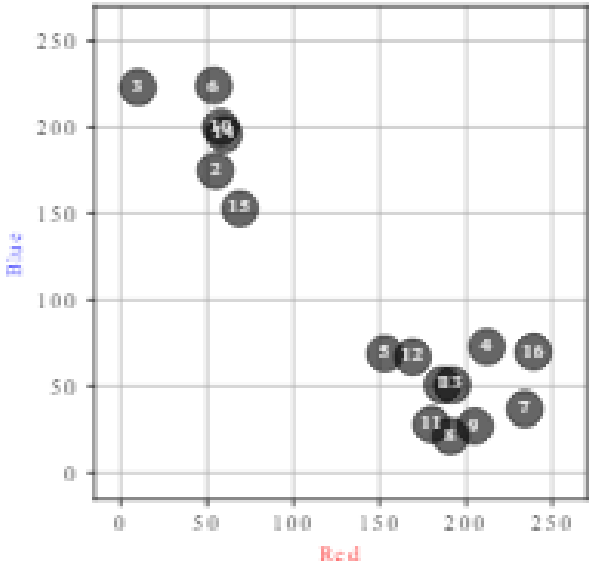


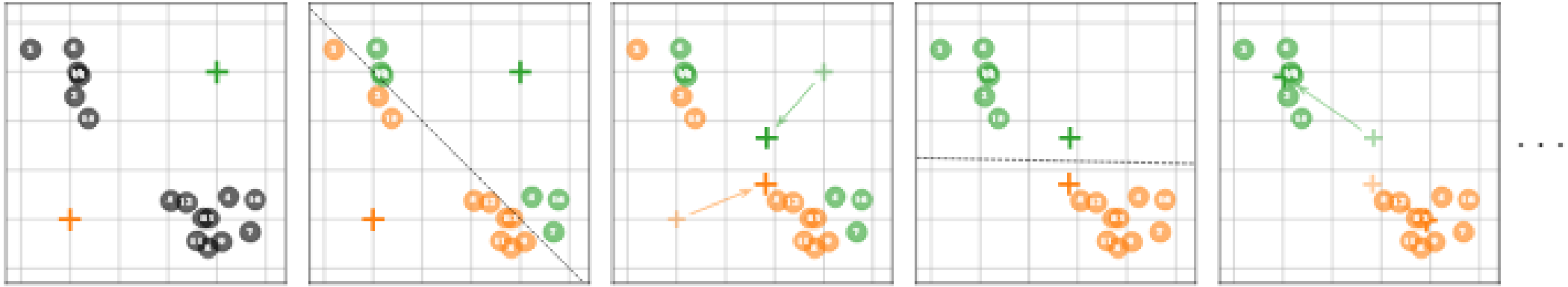
Image with bands red and blue.
Numbers are the pixel indices.

186 51	55 175	10 223	212 73
153 69	54 224	234 37	191 21
205 27	58 200	180 28	169 67
192 51	60 196	69 153	239 70

Intensities (red and blue)
of the pixels.



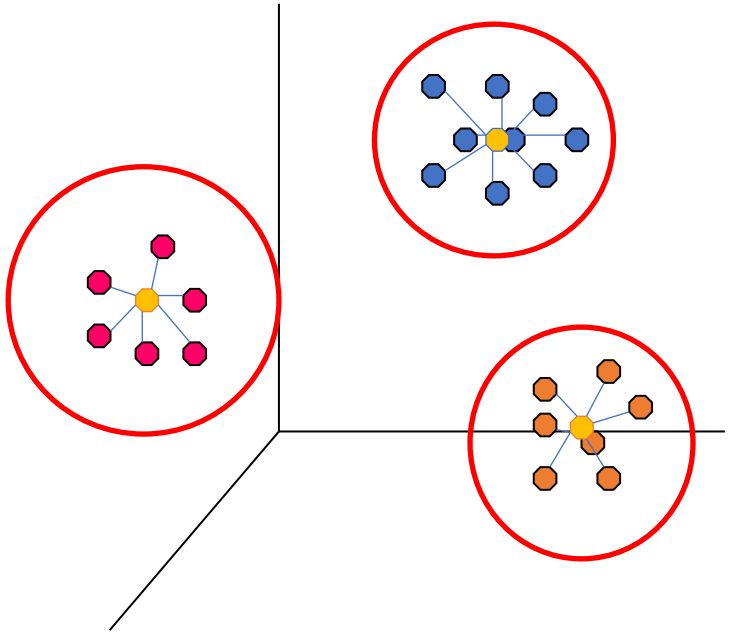
Pixels plotted in
the (red,blue) space.



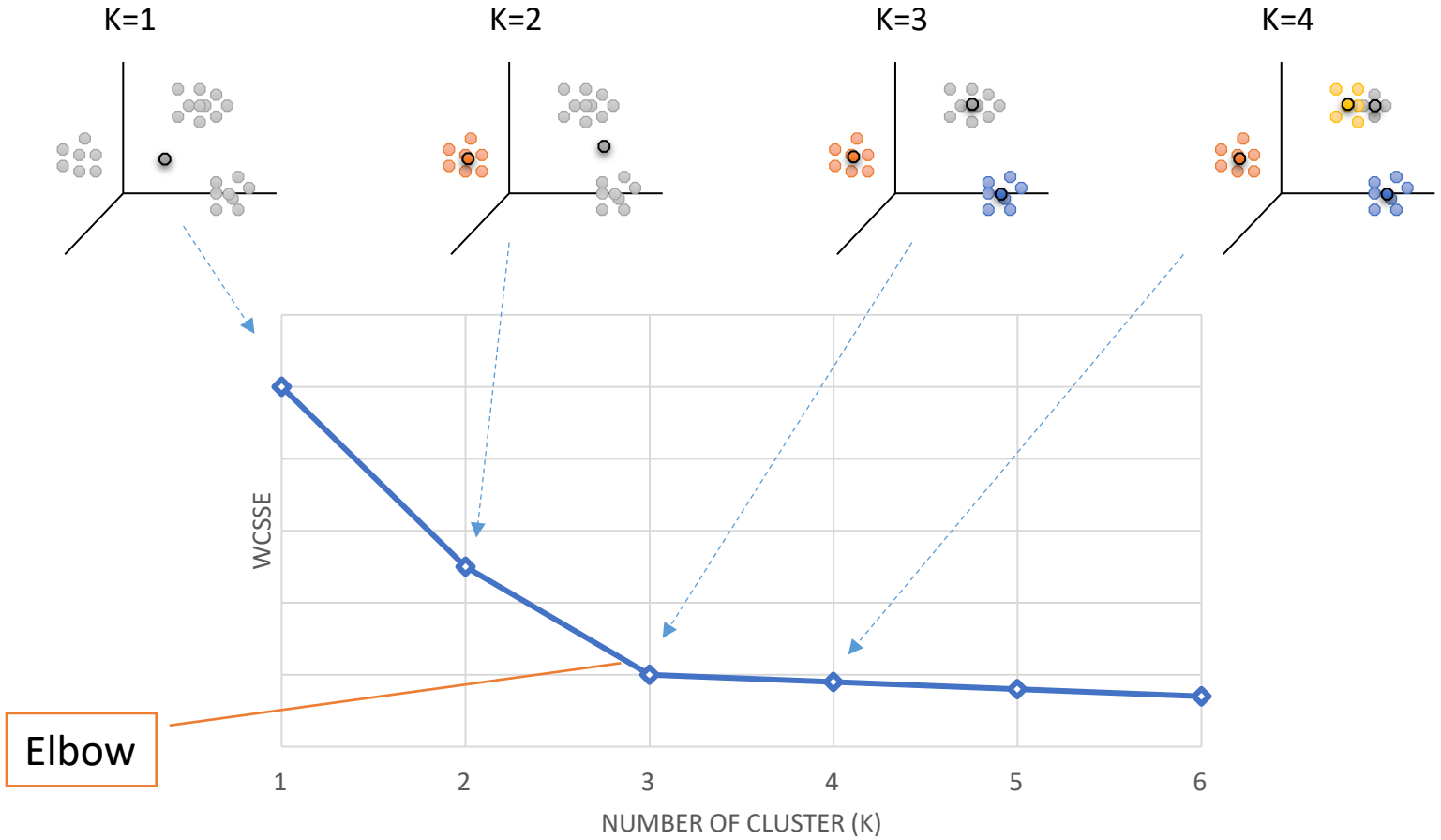
Initialization | Iteration 1 Step A | Iteration 1 Step B | Iteration 2 Step A | Iteration 2 Step B

Progression of the K-means algorithm (only the initialization and the first two iterations are illustrated).
Centroids are represented by +, the dotted line indicates the boundary between the two classes.

การเลือกจำนวนกลุ่มอัตโนมัติโดยใช้ WCSSE และจุดหักศอก



ทดลองจัดกลุ่มโดยเพิ่มจำนวนกลุ่มไปเรื่อย ๆ แล้วพิจารณาจุดหักศอกของ WCSSE



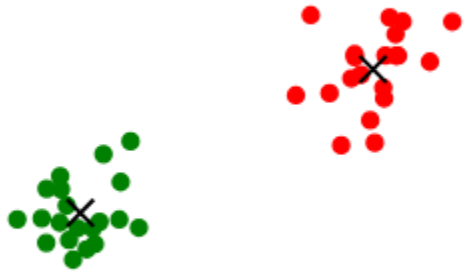
Cluster Validity Index (CVI)

WCSSE is the **S**um of the **S**quared differences (**E**rror) between each observation and its group's mean (**W**ithin **C**luster)

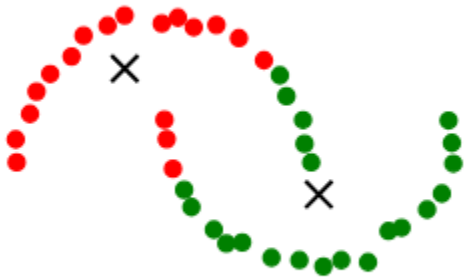
การแบ่งส่วนภาพด้วยการจัดกลุ่ม

K=2

K=4



Spherical clusters



Non-spherical clusters:
the k-means algorithm fails.

Presence of outliers:
the k-means algorithm fails
in this example.

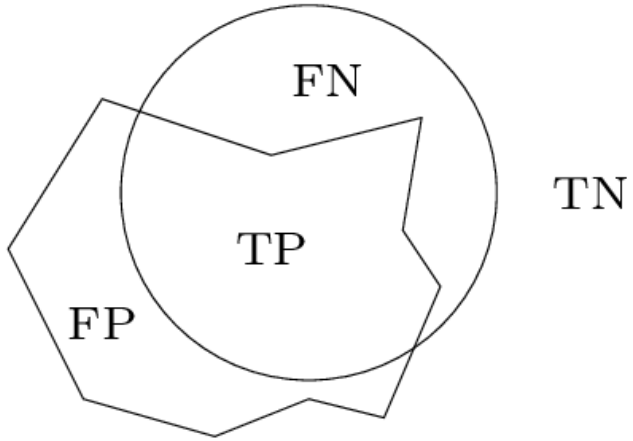
Other methods: https://scikit-learn.org/1.5/modules/clustering.html

การประเมินผลการแบ่งส่วนภาพ

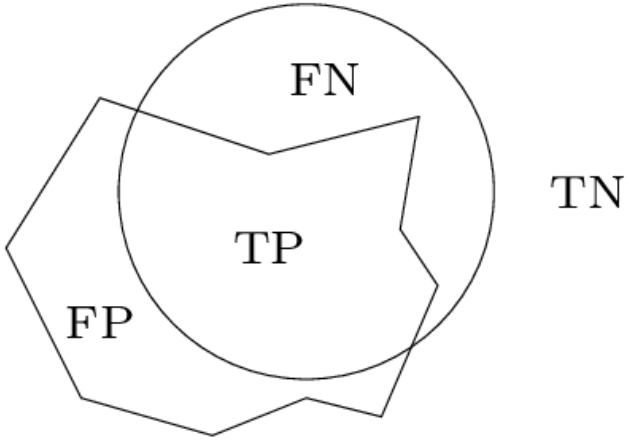
เมื่อเปรียบเทียบผลการแบ่งส่วนกับ ground truth มีพื้นที่สำคัญ 4 ประเภท

1. True Positives (TP)
    1. พิกเซลที่ระบุว่าเป็นวัตถุได้ถูกต้อง
    2. ปรากฏทั้งในผลการแบ่งส่วนและ ground truth
2. True Negatives (TN)
    1. พิกเซลที่ระบุว่าเป็นพื้นหลังได้ถูกต้อง
    2. ไม่ปรากฏทั้งในผลการแบ่งส่วนและ ground truth
3. False Positives (FP)
    1. พิกเซลที่ระบุว่าเป็นวัตถุอย่างไม่ถูกต้อง
    2. ปรากฏในผลการแบ่งส่วนแต่ไม่มีใน ground truth
4. False Negatives (FN)
    1. พิกเซลที่ระบุว่าเป็นพื้นหลังอย่างไม่ถูกต้อง
    2. ปรากฏใน ground truth แต่ไม่พบในผลการแบ่งส่วน

Ground truth segmentation



การประเมินผลการแบ่งส่วนภาพ



Sensibility (*sensibilité*)

$$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$$

Specificity (*spécificité*)

$$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$$

Dice coefficient (*coefficient de Dice*)

$$\frac{2\,\mathrm{TP}}{2\,\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} = \frac{2\,|f \cap f^*|}{|f|+|f^*|}$$

Jaccard coefficient (*coefficient de Jaccard*)

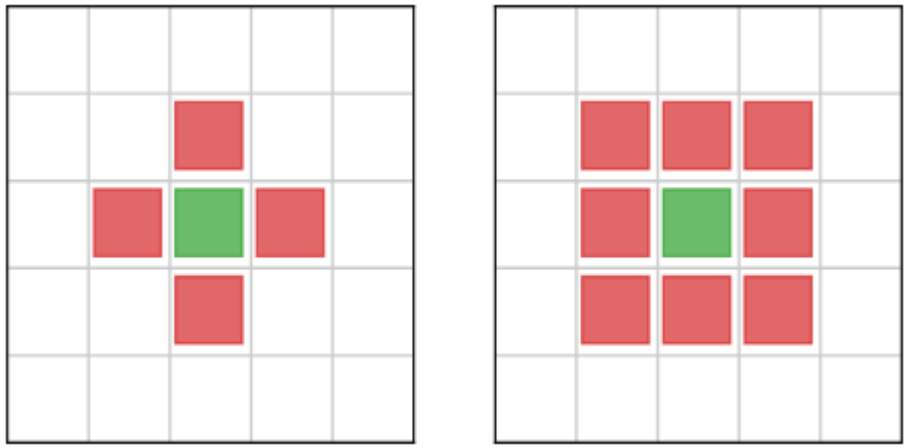
$$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} = \frac{|f \cap f^*|}{|f \cup f^*|}$$

BINARY IMAGE PROCESSING

Binary image processing

ความเป็นเพื่อนบ้านและการเชื่อมต่อของพิกเซล (Neighborhood and connectivity)

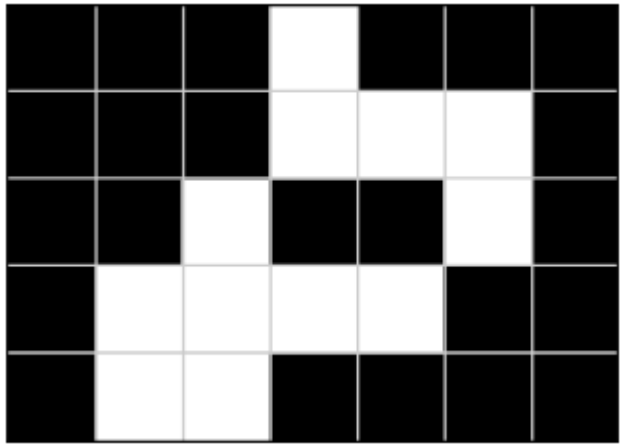
- การเชื่อมต่อแบบ 4 ทิศทาง (4-connectivity)
- การเชื่อมต่อแบบ 8 ทิศทาง (8-connectivity)



4-connectivity 8-connectivity

ส่วนประกอบที่เชื่อมต่อกัน (Connected component)

- เซต S ของพิกเซลในภาพ
- พิกเซลที่เชื่อมต่อ: มีเส้นทางระหว่างพิกเซลภายในเซต S



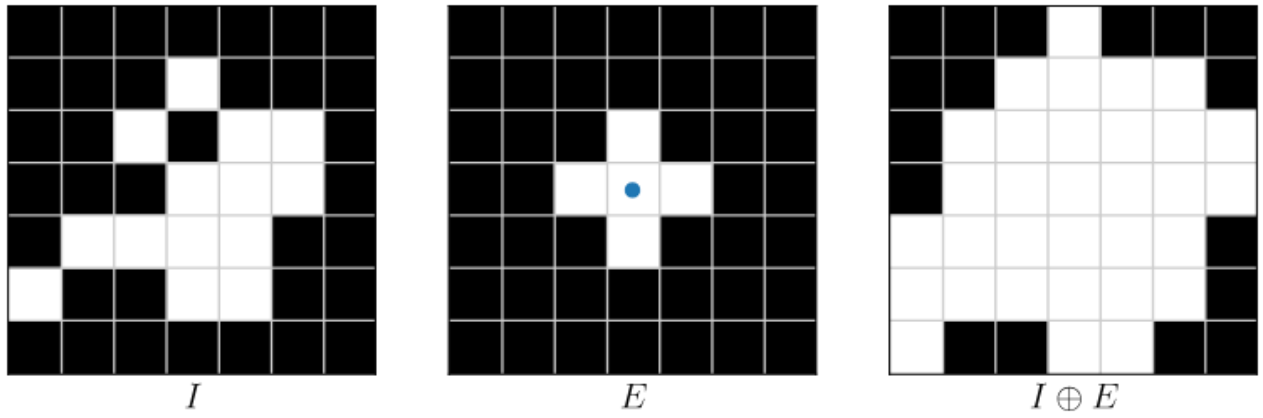
มี 2 ส่วนประกอบ หากพิจารณาตาม 4-connectivity

มี 1 ส่วนประกอบ หากพิจารณาตาม 8-connectivity

Morphology

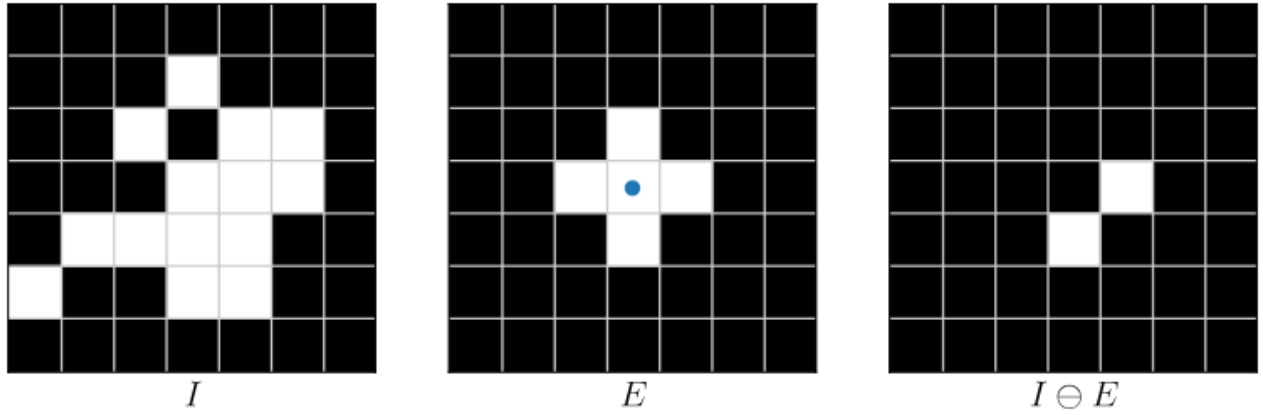
Dilation (การขยายวัตถุ) $I \oplus E_c = \{E_c \mid c \in I\}$

- เลื่อน kernel ไปยังแต่ละจุด
- เก็บพิกเซลทั้งหมดที่มีส่วนของ kernel คลุมถึง
- ผลลัพธ์: วัตถุขยายใหญ่ขึ้น รูเล็กลง
- (การขยายจุดสว่างในภาพสี $\max(K \cup I)$)



Erosion (การหดวัตถุ) $I \ominus E_c = \{c \mid E_c \subseteq I\}$

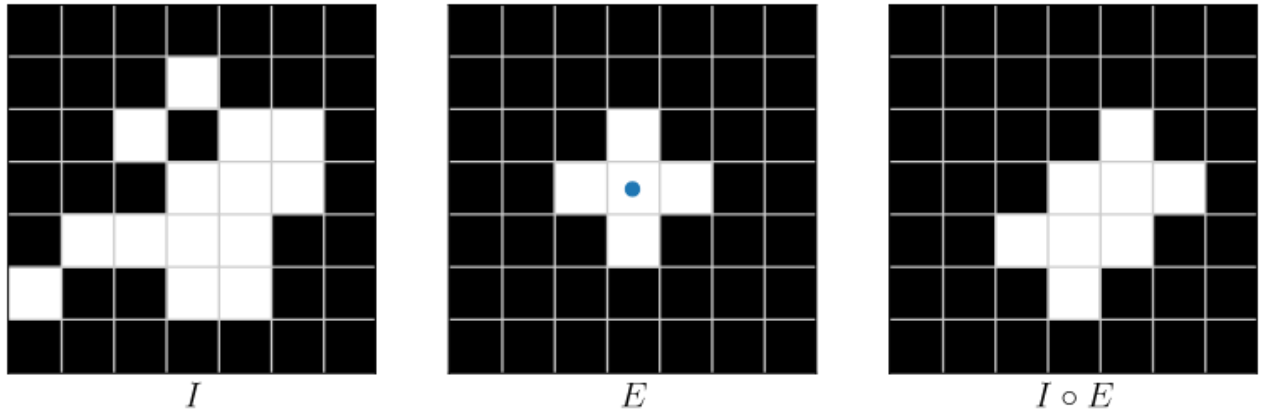
- เลื่อน kernel ไปยังแต่ละจุด
- เก็บเฉพาะพิกเซลคลุม kernel ทั้งหมด
- ผลลัพธ์: วัตถุเล็กลง รูใหญ่ขึ้น
- (การขยายจุดมืดในภาพสี $\min(K \cup I)$)



Morphology

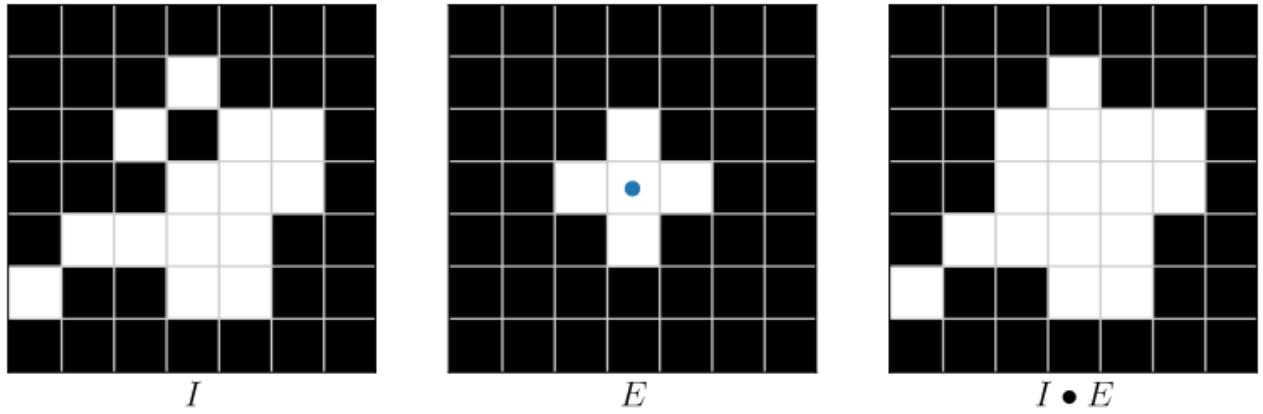
Opening

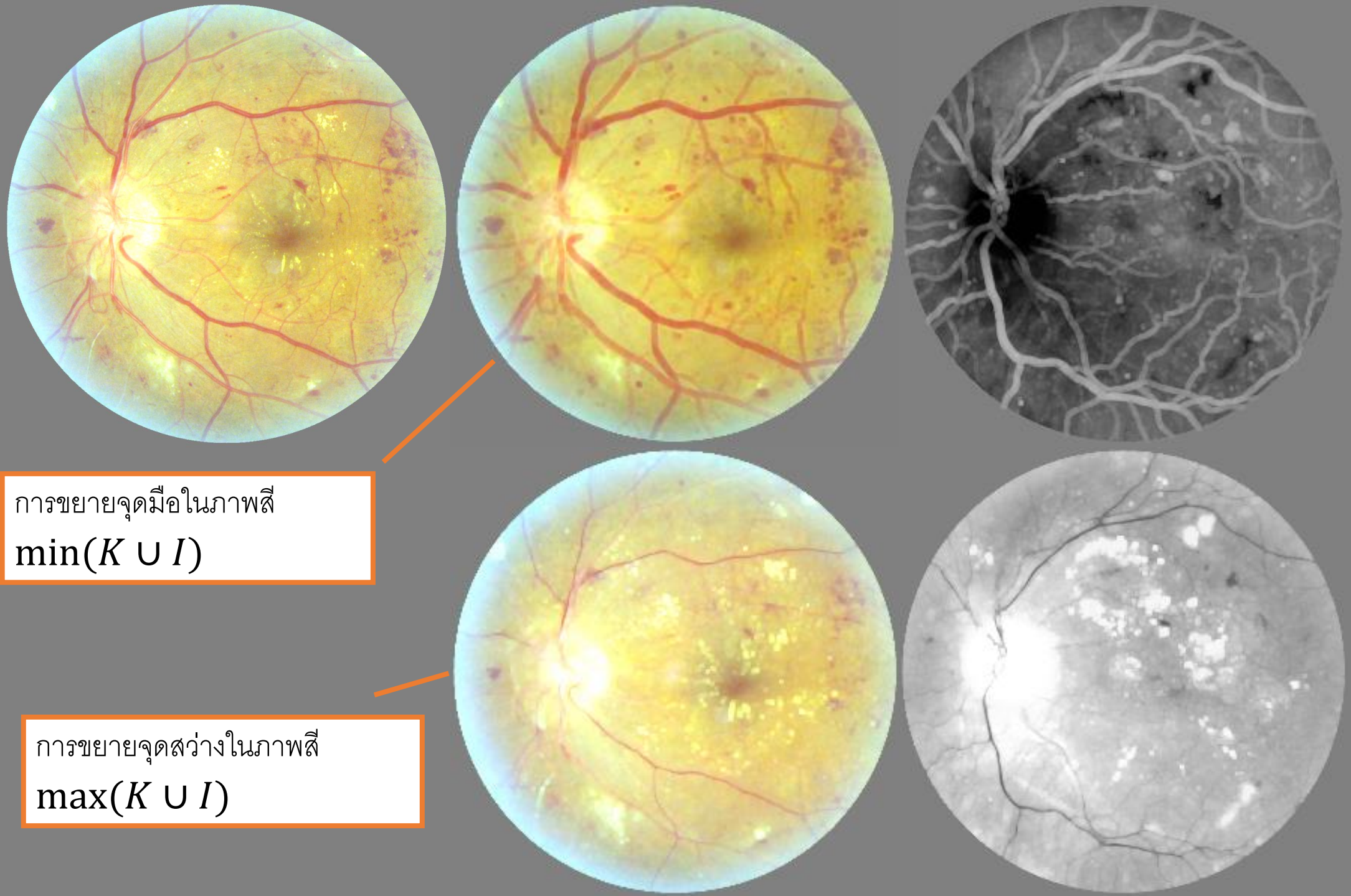
- erosion followed by a dilation (การหดวัตถุแล้วการขยายวัตถุ)
- $I \circ E = (I \ominus E_c) \oplus E_c$



Closing

- dilation followed by a erosion (การขยายวัตถุแล้วการหดวัตถุ)
- $I \cdot E = (I \oplus E) \ominus E$





การขยายจุดมืดในภาพสี

$\min(K \cup I)$

การขยายจุดสว่างในภาพสี

$\max(K \cup I)$

การประยุกต์ใช้งาน

Morphological operators (การหด การขยาย การเปิด และการปิด) สามารถนำมารวมกันเพื่อทำงานที่ซับซ้อนมากขึ้น

- **การแปลง Top-Hat**

  ใช้ในการแยกวัตถุในภาพที่มีขนาดเล็กกว่าโครงสร้าง (structuring element) มีการแปลงท็อปแฮตสองแบบคือ:

  - การแปลงท็อปแฮตสีขาว: เน้นรายละเอียดที่สว่างบนพื้นหลังสีเข้ม คำนวณได้จากความแตกต่างระหว่างภาพต้นฉบับและการเปิด (opening)

$$T_w = I \setminus I \circ E$$

  - การแปลงท็อปแฮตสีดำ:เน้นรายละเอียดสีเข้มบนพื้นหลังสีสว่าง คำนวณได้จากความแตกต่างระหว่างการปิดภาพ (closing) และภาพต้นฉบับ

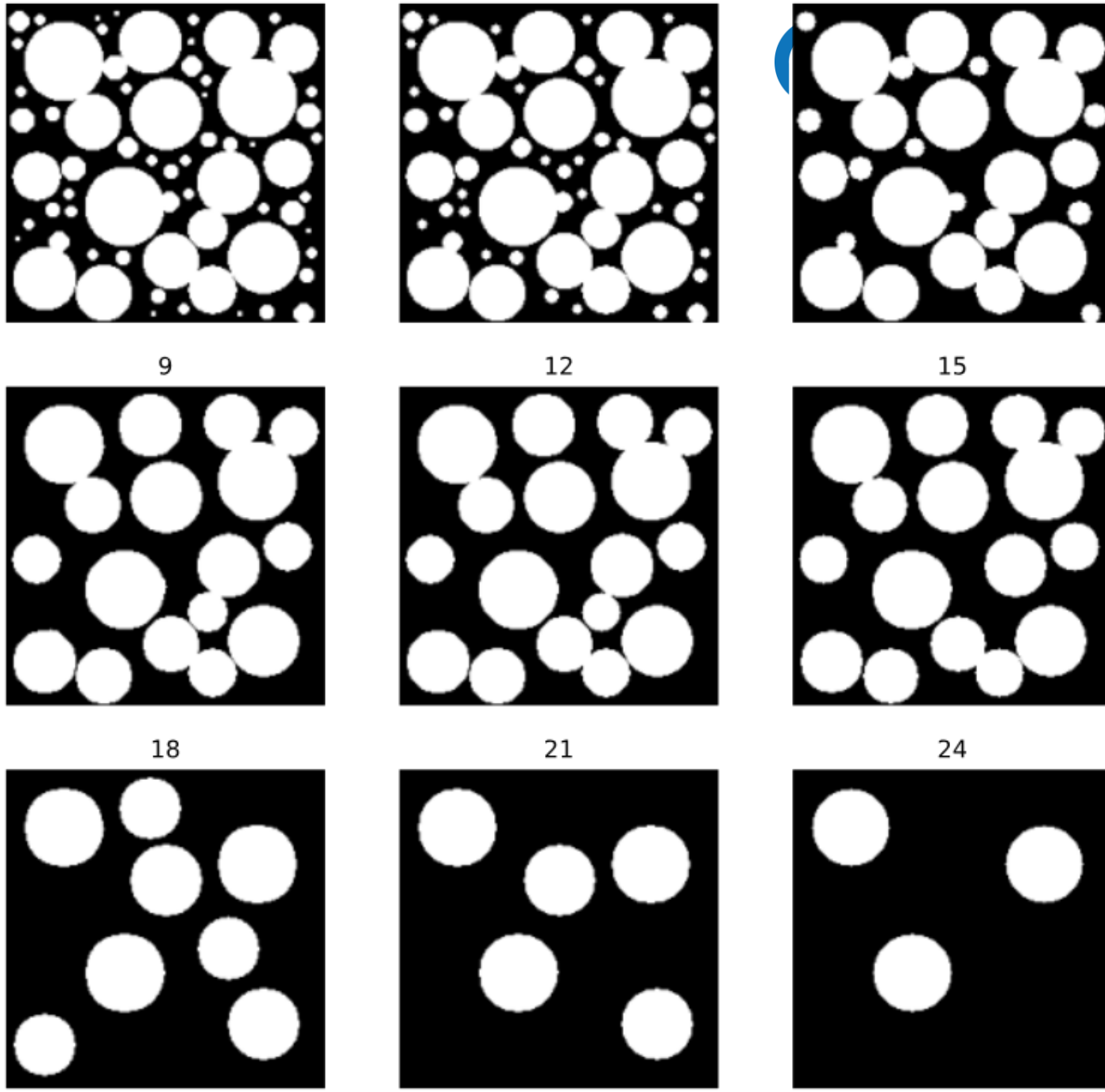
$$T_b = I \cdot E \setminus I$$

white top-hat transform

black top-hat transform

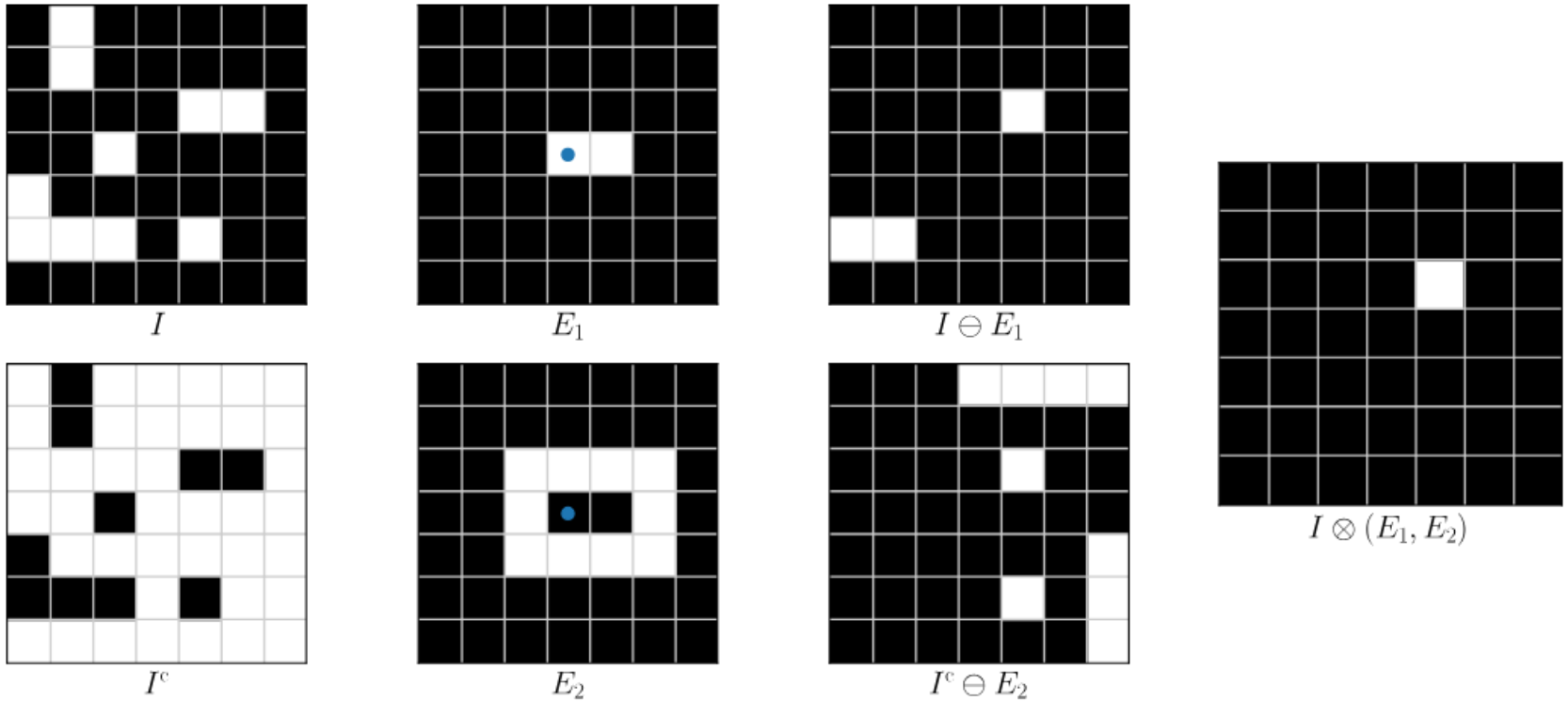
การประยุกต์ใช้งาน

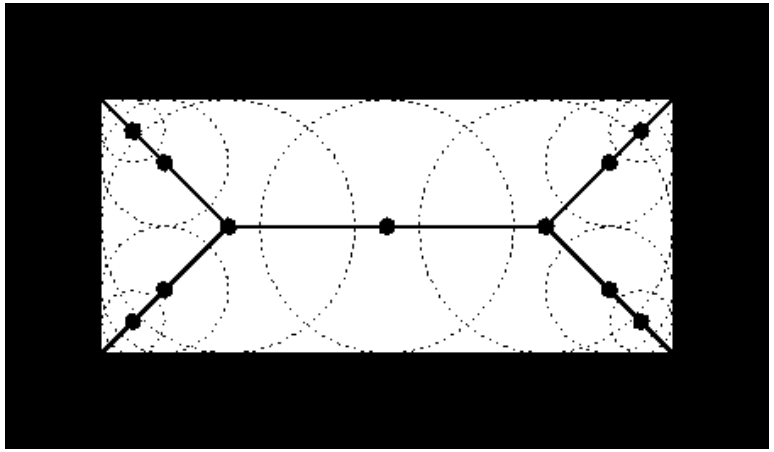
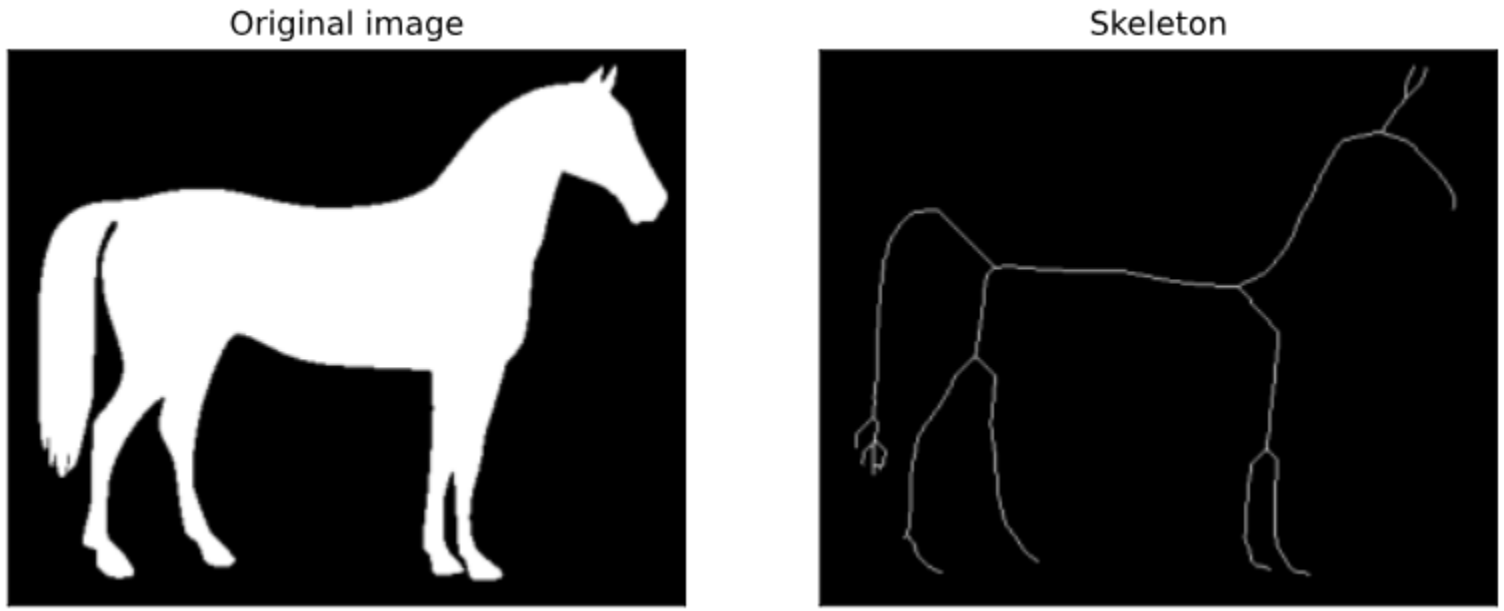
- **กรานูโลเมทรี (Granulometry)** เป็นเทคนิคสำหรับเลือกวัตถุที่มีขนาดเพิ่มขึ้น โดยการเปิดภาพ (opening) ด้วยโครงสร้างที่มีขนาดเพิ่มขึ้นเรื่อย ๆ (เช่น แผ่นดิสก์ที่มีรัศมี 1, 2, 3 เป็นต้น)



การประยุกต์ใช้งาน

- **การแปลงฮิตออร์มิส (Hit-or-miss transform)** ใช้ในการตรวจจับวัตถุที่มีรูปทรงเฉพาะ
  - the erosion of the image by $I \ominus E_1$
  - the erosion of the image background by $I^c \ominus E_2$
  - $I \otimes (E_1, E_2) = (I \ominus E_1) \cap (I^c \ominus E_2) = (I \ominus E_1) \cap (I \oplus E_2)^c$



การประยุกต์ใช้งาน

- การสร้างโครงกระดูก (Skeletonization) ทำให้วัตถุบางลงจนถึงโครงสร้างที่จำเป็น (เส้นกว้างหนึ่งพิกเซล) โดยรักษาการเชื่อมต่อและโครงสร้างไว้