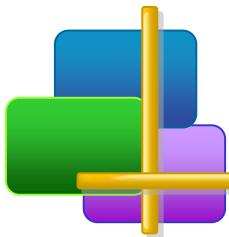


(I'm Fun) Digital Image Fundamentals

Week 7: Morphological Image Processing



Thuong Nguyen Canh

Institute for Datability Science, Osaka University

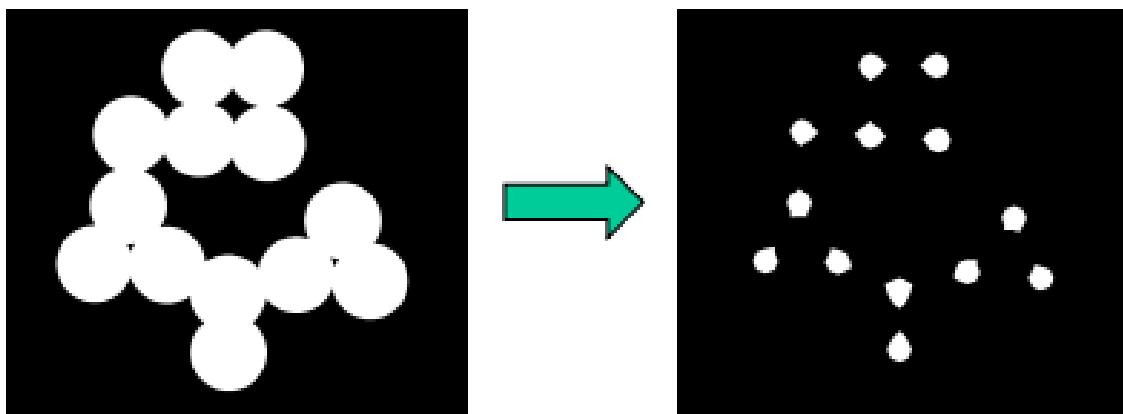
November 2019

Morphological Image Processing

- Non-linear operations related to the shape or morphology of image features
 - Binary dilation and erosion
 - Set-theoretic interpretation
 - Hit-miss filter
 - Binary dilation and erosion
 - Morphological filters for gray-level images
 - Cascading dilations and erosions
 - Rank filters, median filters, majority filters
- Reduce the imperfection of binary images by considering the form & structure

INTEREST-POINT DETECTION

Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].



Binary Image Processing

- ➊ Binary images are common
 - ❼ Intermediate abstraction in a gray-scale/color image analysis system
 - ➌ Thresholding/segmentation
 - ➌ Presence/absence of some image property
 - ❼ Text and line graphics, document image processing
- ➋ Representation of individual pixels as 0 or 1, convention:
 - ❼ foreground, object = 1 (white)
 - ❼ background = 0 (black)
- ➌ Processing by logical functions is fast and simple
- ➍ Shift-invariant logical operations on binary images:
 - ❼ “morphological” image processing

Distorted by noise and texture

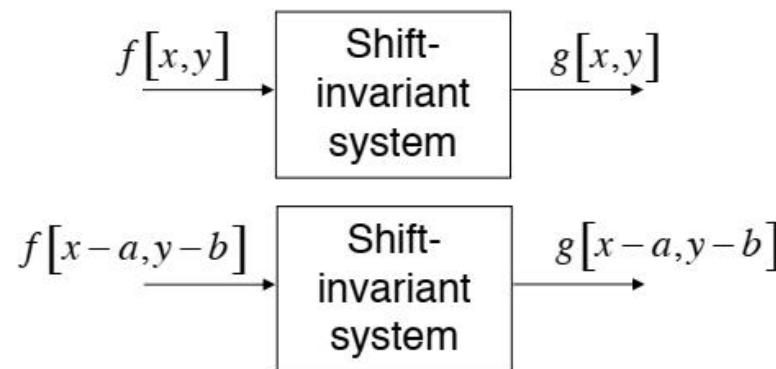
Morphological image processing has been generalized to gray-level images via level sets

Shift-Invariance

- Assume that digital images $f[x,y]$ and $g[x,y]$ have infinite support

$$[x,y] \in \{\dots, -2, -1, 0, 1, 2, \dots\} \times \{\dots, -2, -1, 0, 1, 2, \dots\}$$

. . . then, for all integers a and b



- Shift-invariance does **not** imply linearity (or vice versa).

Structuring Element

- Neighborhood “window” operator

$$W\{f[x, y]\} = \left\{ f[x - x', y - y'] : [x', y'] \in \Pi_{xy} \right\}$$

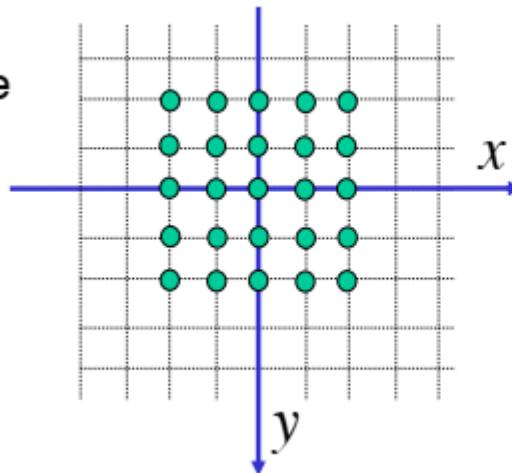
180 degree rotation “structuring element”

$$\hat{W}\{f[x, y]\} = \left\{ f[x + x', y + y'] : [x', y'] \in \Pi_{xy} \right\}$$

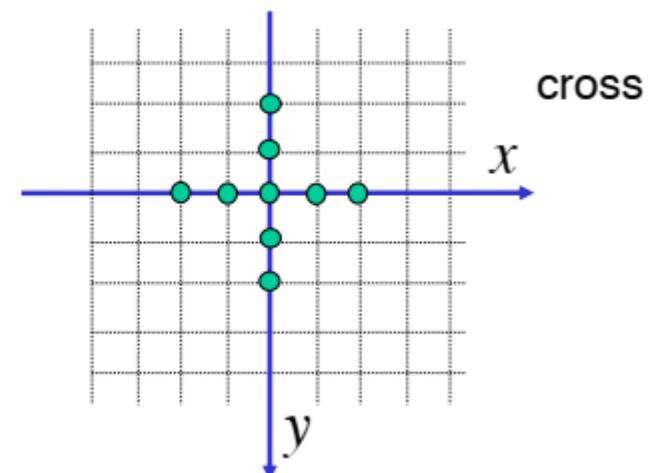
The “hat” notation indicates upright structuring element (i.e., not rotated!)

- Example structuring elements Π_{xy} :

5x5 square



cross



Binary dilation

- Gradually enlarge the boundaries of regions of foreground pixels
 - Areas of foreground pixels grow in size
 - Holes within those regions become smaller.



dilation

/dɪ'leɪʃ(ə)n, ɪv'leɪʃ(ə)n/

noun

1. PHYSIOLOGY

the action or condition of becoming or being made wider, larger, or more open.
"nitric oxide causes dilation of the blood vessels"

2. the action of speaking or writing at length on (a subject).

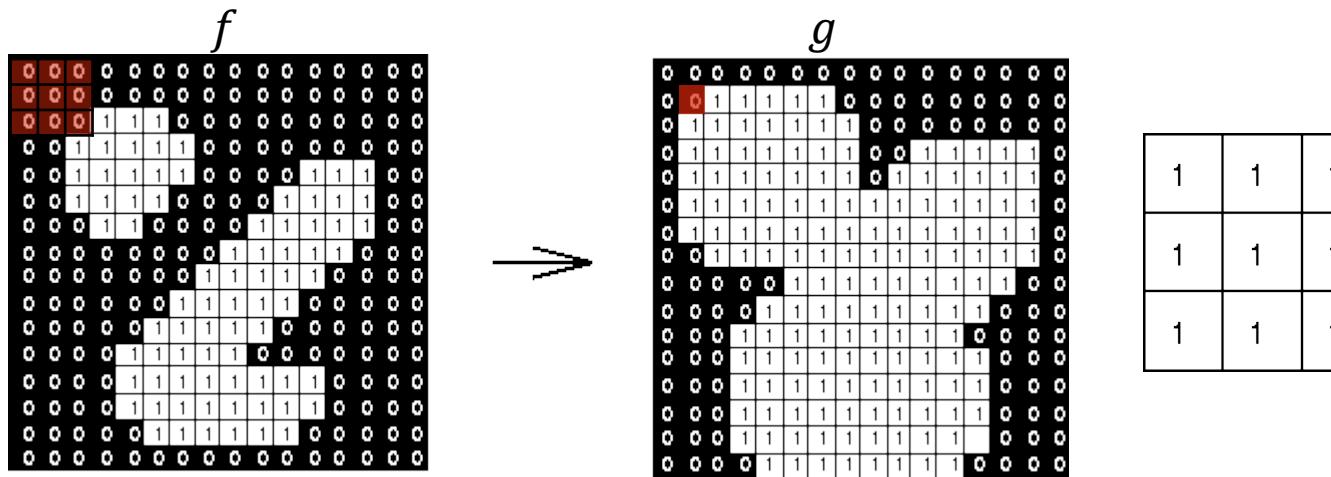
"the main editorial involved no dilation on the privileges or responsibilities of citizenship"

Binary Dilation

Definition of Dilation

$$g[x, y] = OR[W\{f[x, y]\}] \equiv dilate(f, W)$$

- f is the image
 - W is the set of coordination points – structuring element or kernel
- $W\{f[x, y]\}$: translation of W to origin at point $f[x, y]$
- $OR[\quad]$: find intersection points between $W\{f[x, y]\}$ and f



If any pixels is foreground in both f and W , then it is foreground

Binary dilation with square structuring element

$$g[x,y] = OR[W\{f[x,y]\}] := \text{dilate}(f,W)$$



Original (701x781)



dilation with
3x3 structuring element



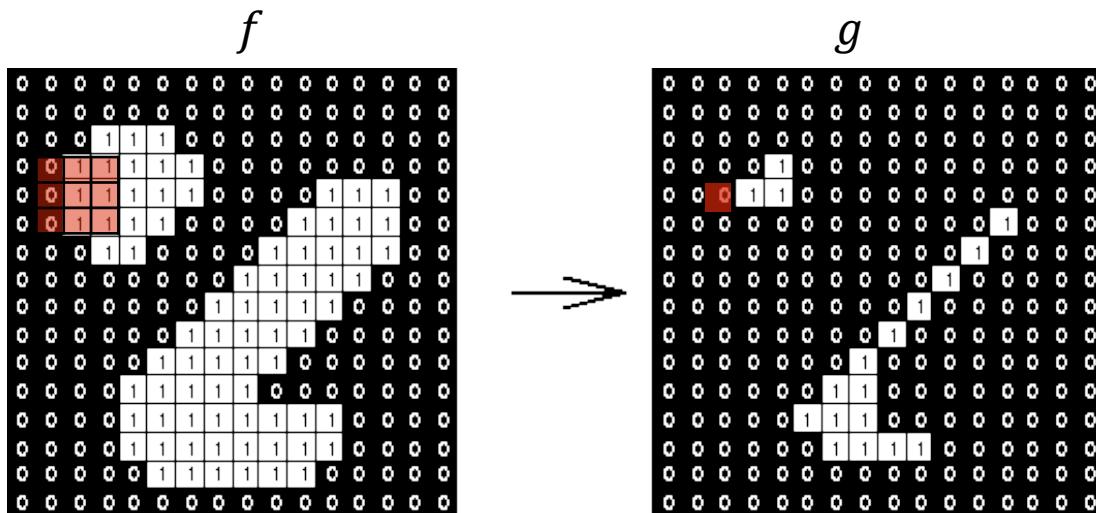
dilation with
7x7 structuring element

- Expands the size of 1-valued objects
- Smoothes object boundaries
- Closes holes and gaps

Binary erosion with square structuring element

- Gradually thinning the boundaries of regions of foreground pixels
 - Areas of foreground pixels reduce in size
 - Holes within those regions become larger.

$$g[x, y] = \text{AND}[\hat{W}\{f[x, y]\}] := \text{erode}(f, W)$$



W

Set of coordinate points =

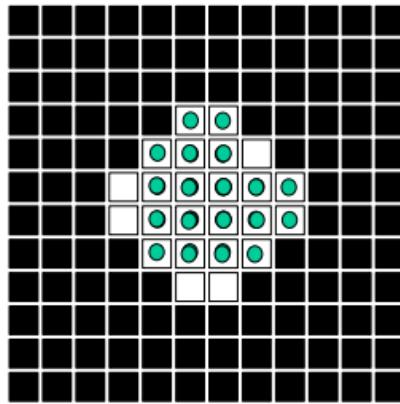
1	1	1
1	1	1
1	1	1

{ (-1, -1), (0, -1), (1, -1),
(-1, 0), (0, 0), (1, 0),
(-1, 1), (0, 1), (1, 1) }

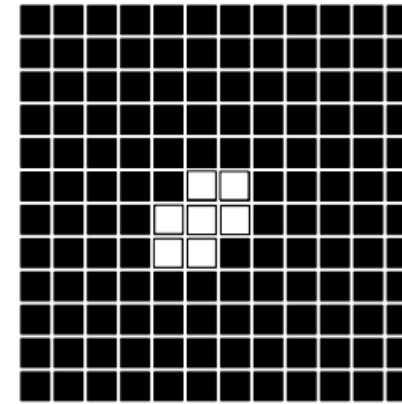
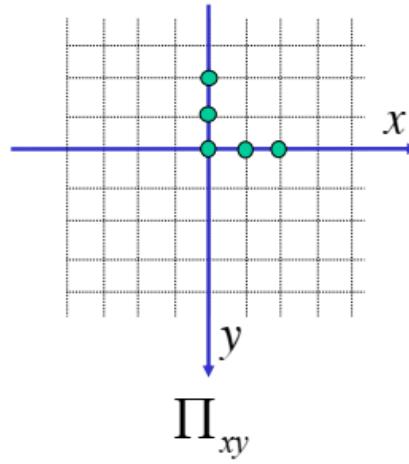
If all pixels is foreground in both f and W , then it is foreground

Binary erosion with square structuring element

$$g[x, y] = \text{AND}[\hat{W}\{f[x, y]\}] := \text{erode}(f, W)$$



$f[x, y]$



$g[x, y]$

Caveat: There is another definition of erosion in the literature, which flips the structuring element, as for dilation.

The Lagunita online videos use that alternative definition. Matlab function *imerode* uses the definition on this slide. To the best of our knowledge, there is no such discrepancy defining dilation.

Relationship between dilation and erosion

- Duality: erosion is dilation of the background

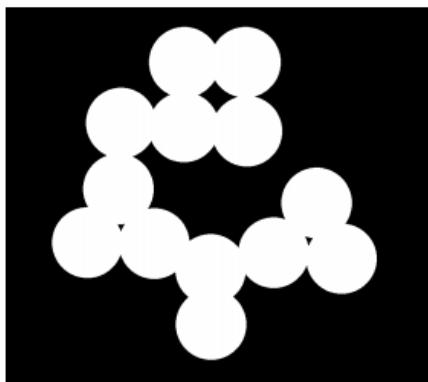
$$\text{dilate}(f, W) = \text{NOT}[\text{erode}(\text{NOT}[f], \hat{W})]$$

$$\text{erode}(f, W) = \text{NOT}[\text{dilate}(\text{NOT}[f], \hat{W})]$$

- But: erosion is not the inverse of dilation

$$\begin{aligned} f[x, y] &\neq \text{erode}(\text{dilate}(f, W), W) \\ &\neq \text{dilate}(\text{erode}(f, W), W) \end{aligned}$$

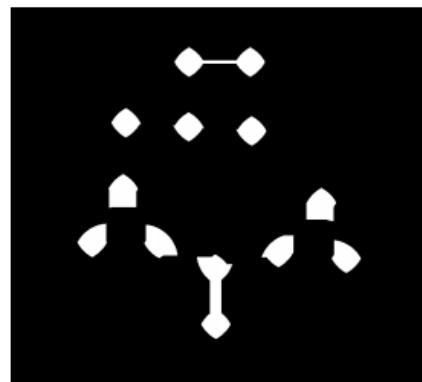
Example: Blob separation/detection by erosion



Original binary image
Circles (792x892)



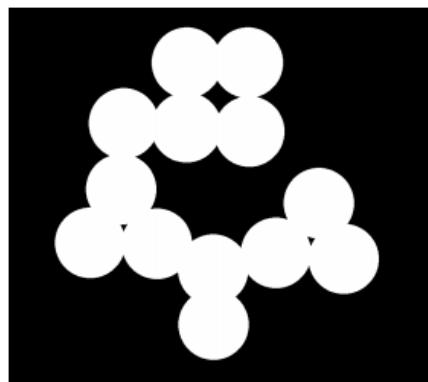
Erosion by 30x30
structuring element



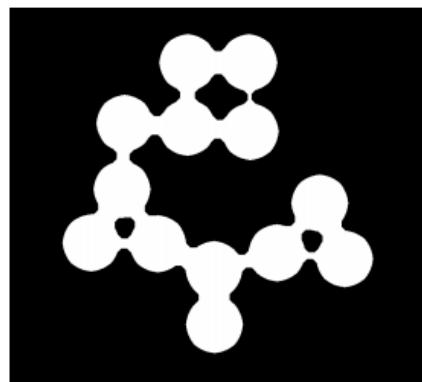
Erosion by 70x70
structuring element



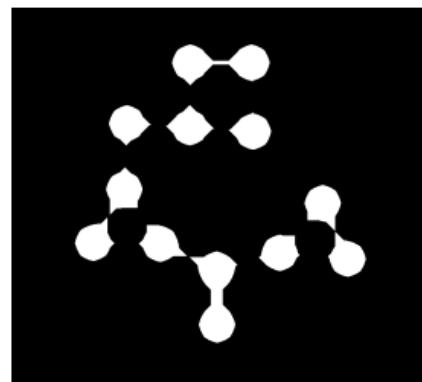
Erosion by 96x96
structuring element



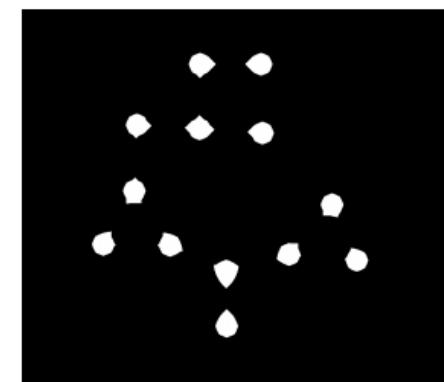
Original binary image
Circles (792x892)



Erosion by disk-shaped
structuring element
Diameter=15

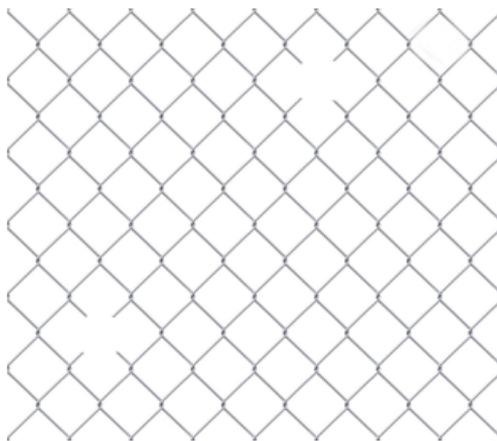


Erosion by disk-shaped
structuring element
Diameter=35

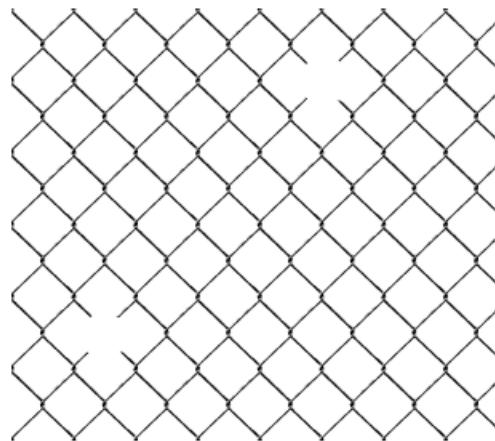


Erosion by disk-shaped
structuring element
Diameter=48

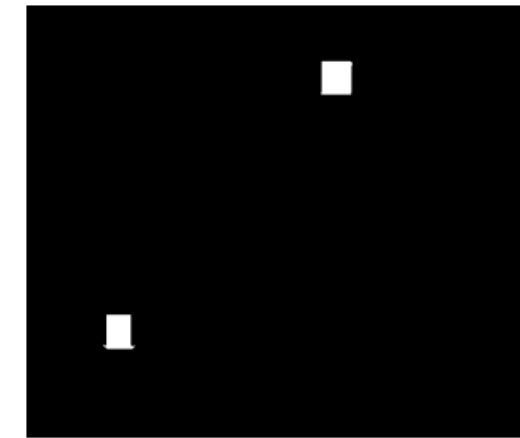
Example: chain link fence hole detection



Original grayscale image
Fence (1023 x 1173)



Fence thresholded
using Otsu's method



Erosion with 151x151
“cross” structuring element

Set-theoretic interpretation

- Set of object pixels

$$F \equiv \{(x, y) : f(x, y) = 1\}$$

*Continuous (x,y).
Works for discrete [x,y]
in the same way.*

- Background: complement of foreground set

$$F^c \equiv \{(x, y) : f(x, y) = 0\}$$

- Dilation is Minkowski set addition

$$G = F \oplus \Pi_{xy}$$

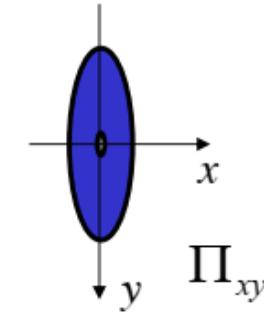
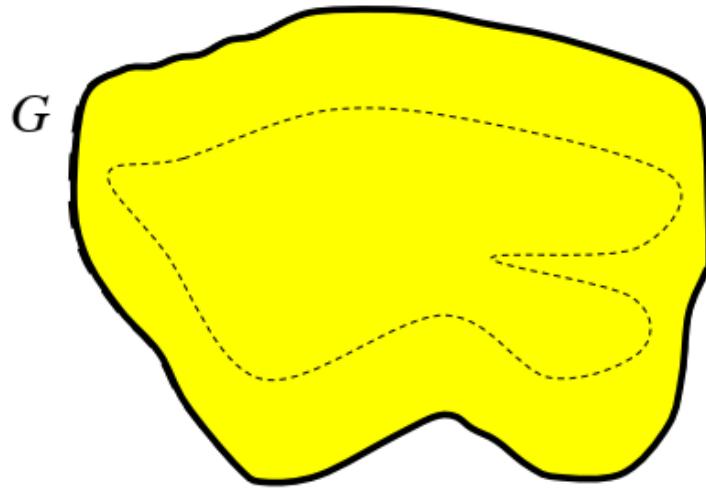
Commutative and associative!

$$= \{(x + p_x, y + p_y) : (x, y) \in F, (p_x, p_y) \in \Pi_{xy}\}$$

$$= \bigcup_{(p_x, p_y) \in \Pi_{xy}} F_{+^{(p_x, p_y)}}$$

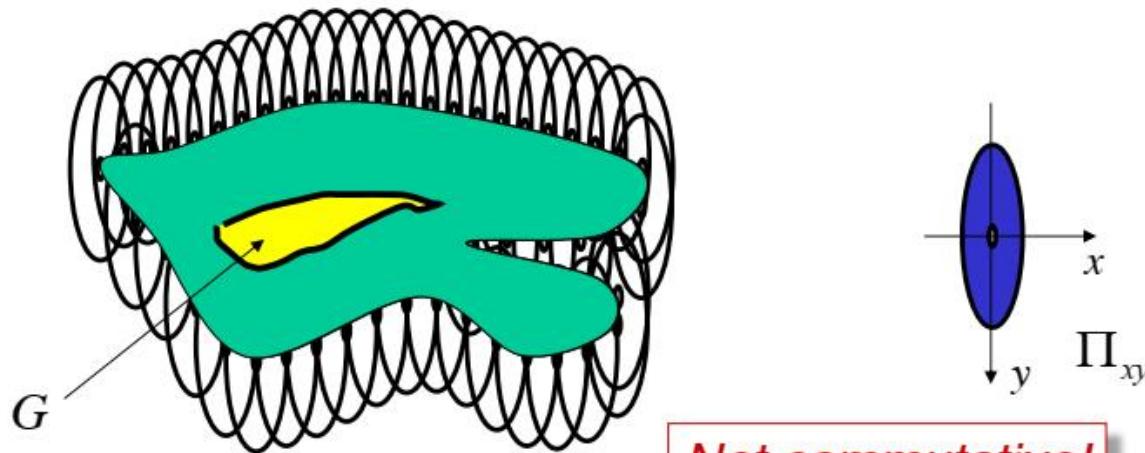
translation of F by vector (p_x, p_y)

Set-theoretic interpretation: dilation



$$\begin{aligned}G &= F \oplus \Pi_{xy} \\&= \left\{ (x + p_x, y + p_y) : (x, y) \in F, (p_x, p_y) \in \Pi_{xy} \right\} \\&= \bigcup_{(p_x, p_y) \in \Pi_{xy}} F_{+(p_x, p_y)}\end{aligned}$$

Set-theoretic interpretation: dilation



*Not commutative!
Not associative!*

Minkowski set subtraction

$$G = \bigcap_{(-p_x, -p_y) \in \Pi_{xy}} F_{+(p_x, p_y)} = F_\Theta \Pi_{xy}$$

Opening and closing

- Goal: smoothing without size change

- Open filter

$$\text{open}(f, W) = \text{dilate}(\text{erode}(f, W), W)$$

- Close filter

$$\text{close}(f, W) = \text{erode}(\text{dilate}(f, W), W)$$

- Open filter and close filter are biased

- Open filter removes small 1-regions
 - Close filter removes small 0-regions
 - Bias is often desired for enhancement or detection!

- Unbiased size-preserving smoothers

$$\text{close-open}(f, W) = \text{close}(\text{open}(f, W), W)$$

$$\text{open-close}(f, W) = \text{open}(\text{close}(f, W), W)$$

- *close-open* and *open-close* are duals, but not inverses of each other.

Small Hole Removal by Closing



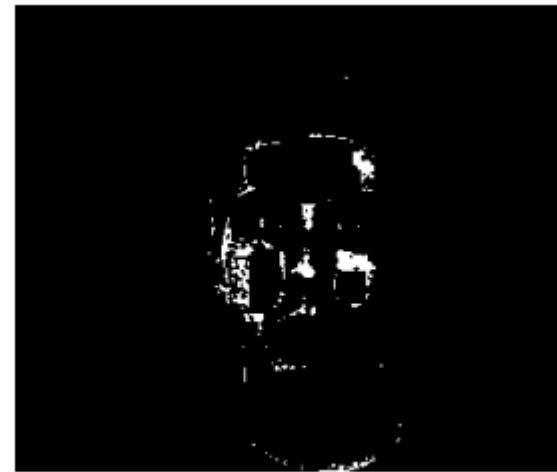
Original binary mask



Dilation
10x10



Closing 10x10

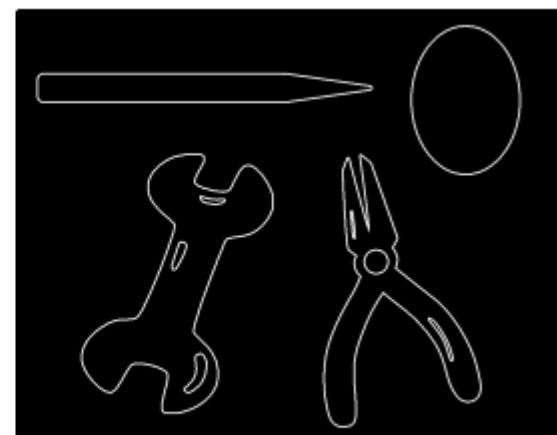


Difference to original mask

Morphological edge detectors

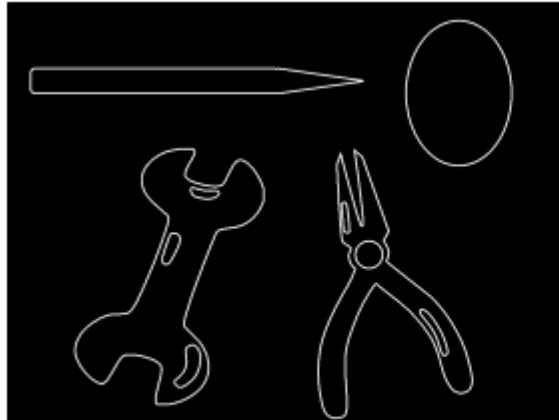


$$f[x, y]$$



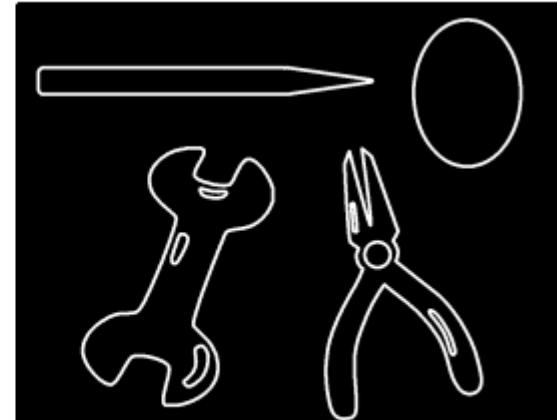
$$\text{dilate}(f, W) \neq f$$

internal gradient



$$\text{erode}(f, W) \neq f$$

Morphological gradient



$$\text{dilate}(f, W) \neq \text{erode}(f, W)$$

Recognition by erosion

Binary image f

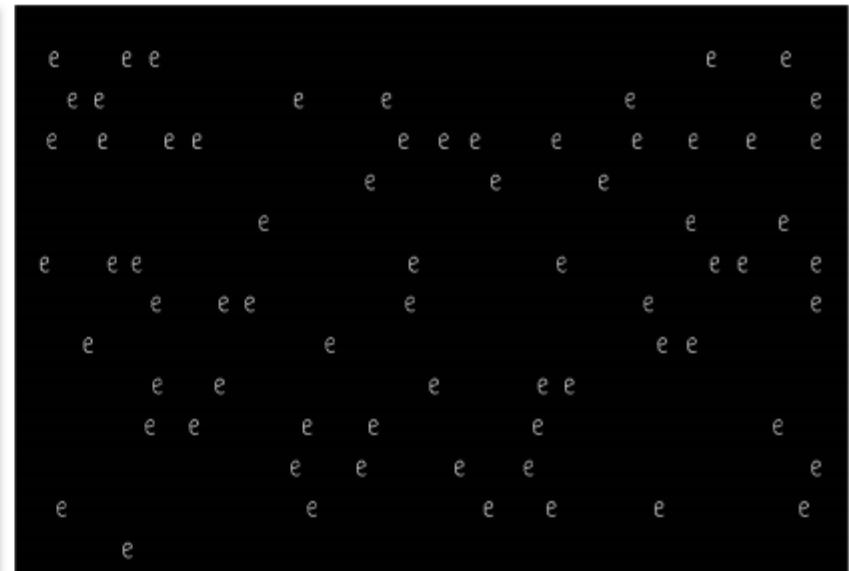
1400

INTEREST-POINT DETECTION

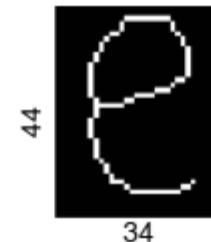
Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

2000

$$\text{dilate}(\text{erode}(\text{NOT}[f], W), W)$$



Structuring
element W



44

34

Recognition by erosion

INTEREST-POINT DETECTION

Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

Structuring
element W

44



34

Recognition by erosion

Binary image f

INTEREST-POINT DETECTION

Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

1400

2000

$dilate(erode(NOT[f],W),W)$



Structuring
element W



18

22

Recognition by erosion

INTEREST-POINT DETECTION

Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

Structuring
element W



Hit-miss Filter

- Two structure V and W are mutually exclusive $V \cap W = \emptyset$

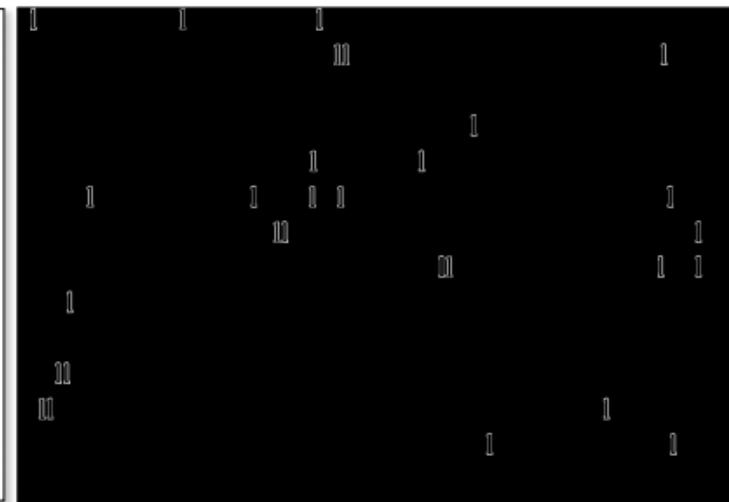
Binary image f

INTEREST-POINT DETECTION
Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

1400

2000

$dilate(erode(NOT[f],V) \& erode(f,W),W)$



Structuring element V



Structuring element W



A point x belongs to the hit-or-miss transform output if V translated to x fits in F , and W translated to x misses F (fits the background of F).

Hit-miss filter

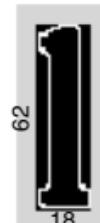
INTEREST-POINT DETECTION

Feature extraction typically starts by finding the salient interest points in the image. For robust image matching, we desire interest points to be repeatable under perspective transformations (or, at least, scale changes, rotation, and translation) and real-world lighting variations. An example of feature extraction is illustrated in Figure 3. To achieve scale invariance, interest points are typically computed at multiple scales using an image pyramid [15]. To achieve rotation invariance, the patch around each interest point is canonically oriented in the direction of the dominant gradient. Illumination changes are compensated by normalizing the mean and standard deviation of the pixels of the gray values within each patch [16].

Structuring
element V



Structuring
element W



Morphological filters for gray-level images

- Threshold sets of a gray-level image $f[x,y]$ (aka level sets)

$$T_\theta(f[x,y]) = \{[x,y] : f[x,y] \geq \theta\}, \quad -\infty < \theta < +\infty$$

- Reconstruction of original image from threshold sets

$$f[x,y] = \sup\{\theta : [x,y] \in T_\theta(f[x,y])\}$$

- Idea of morphological operators for multi-level (or continuous-amplitude) signals

- Decompose into threshold sets
 - Apply binary morphological operator to each threshold set
 - Reconstruct via supremum operation
 - Gray-level operators thus obtained: **flat operators**
- ➔ *Flat morphological operators and thresholding are commutative*

Dilation/Erosion for gray-level images

- Explicit decomposition into threshold sets not required in practice
- Flat dilation operator: local maximum over window W

$$g[x, y] = \max\{W\{f[x, y]\}\} := \text{dilate}(f, W)$$

- Flat erosion operator: local minimum over window W

$$g[x, y] = \min\{\hat{W}\{f[x, y]\}\} := \text{erode}(f, W)$$

- Binary dilation/erosion operators contained as special case

1D Illustration of erosion and dilation

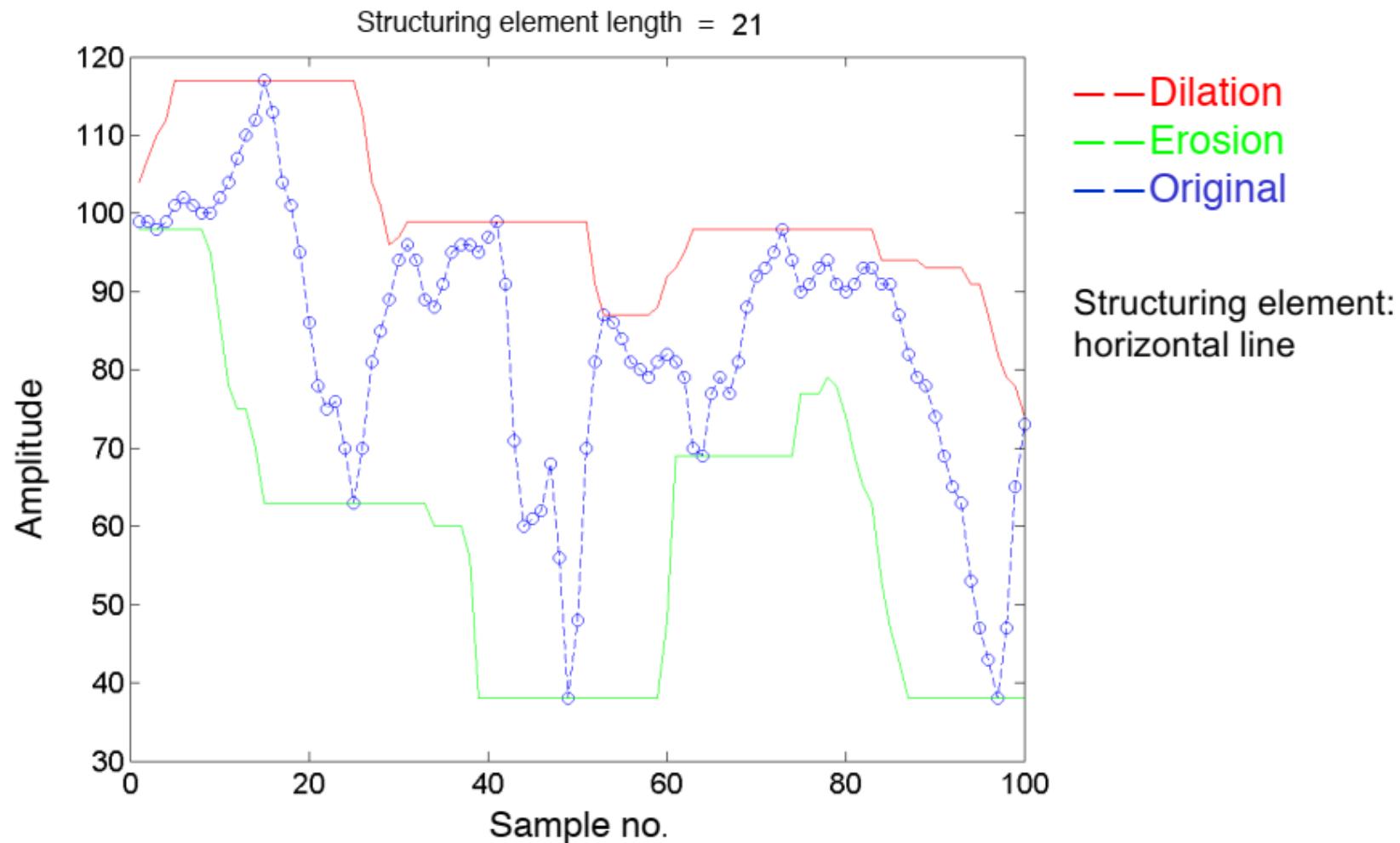


Image Example



Original 394 x 305



Dilation 10x10 square



Erosion 10x10 square

Flat Dilation with Different Structuring Elements



Original



Diamond



Disk



20 degree line



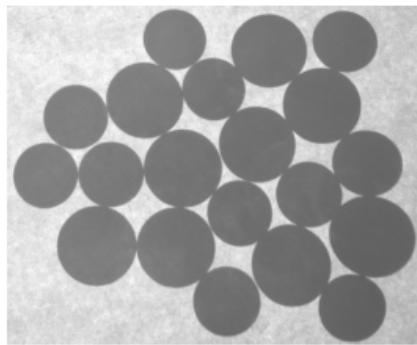
9 points



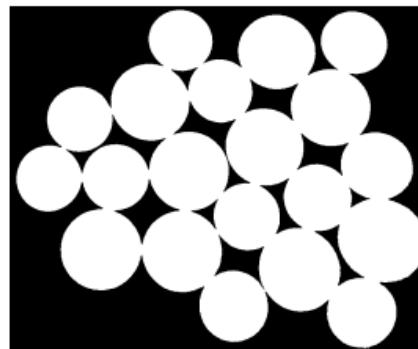
2 horizontal lines



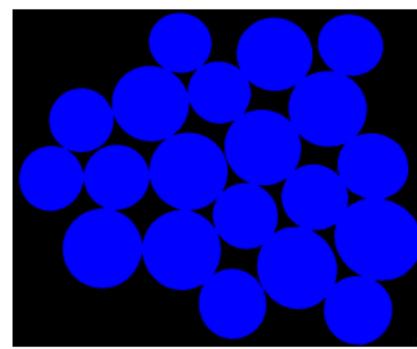
Example: Counting coins



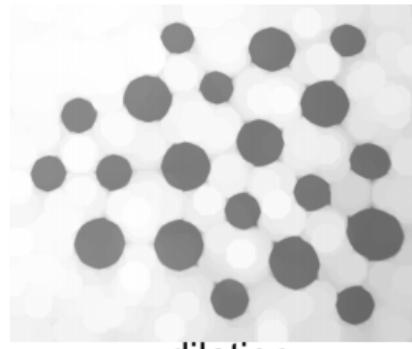
Original



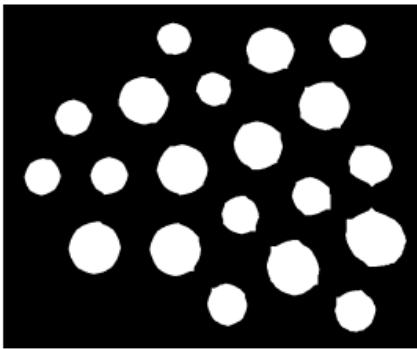
thresholded



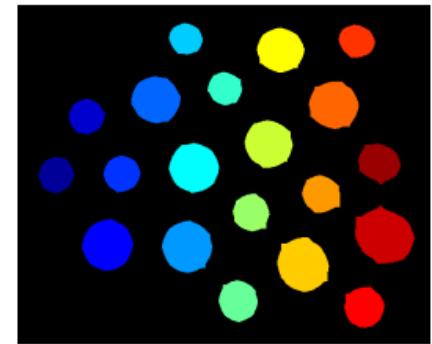
1 connected component



dilation



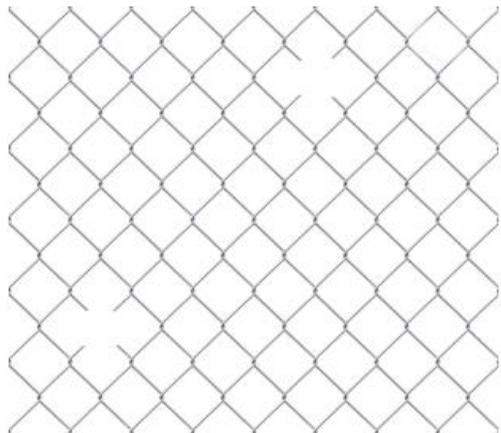
thresholded after dilation



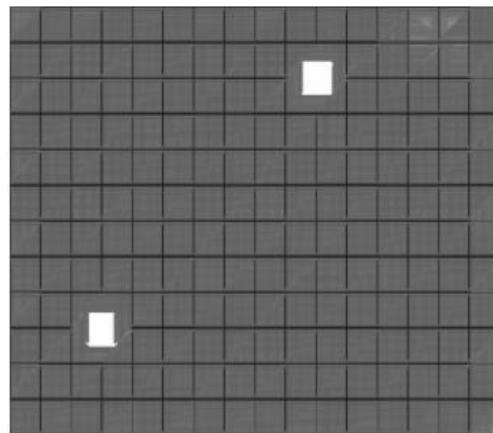
20 connected
components



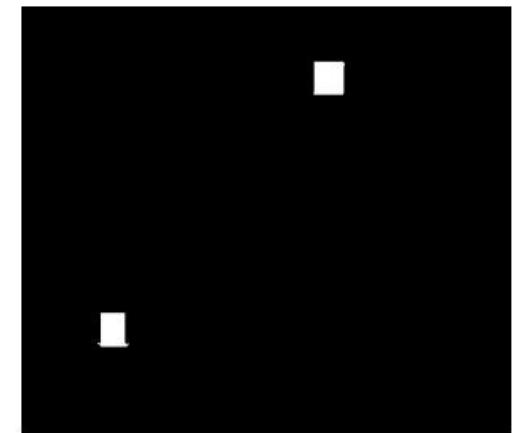
Example: Chain link fence hole detection



Original grayscale image
Fence (1023 x 1173)

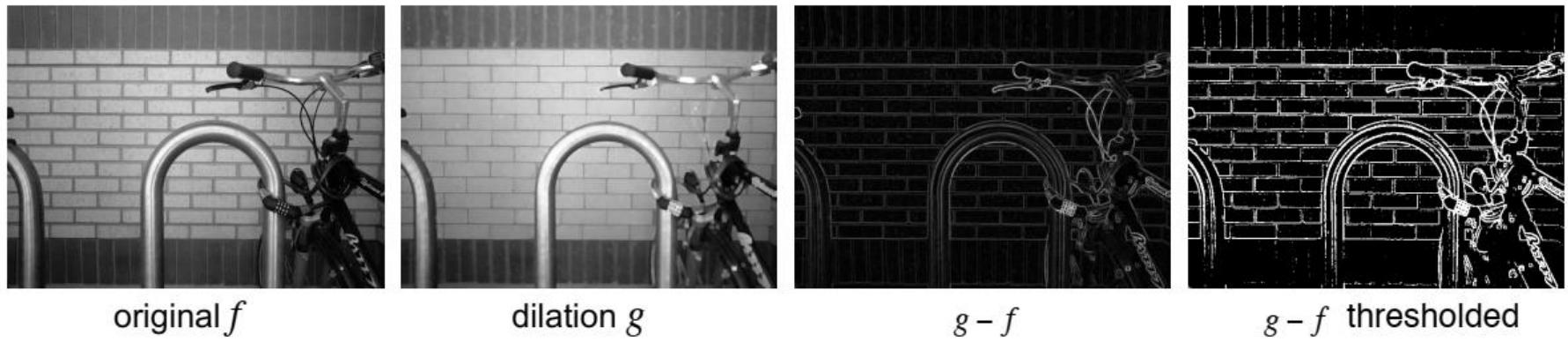


Flat erosion with 151x151
“cross” structuring element



Binarized by Thresholding

Morphological Edge Detector



Beyond Flat Morphological Operators

- General dilation operator

$$g[x, y] = \sup_{\alpha, \beta} \{ f[x - \alpha, y - \beta] + w[\alpha, \beta] \} = \sup_{\alpha, \beta} \{ w[x - \alpha, y - \beta] + f[\alpha, \beta] \}$$

- Like linear convolution, with sup replacing summation, addition replacing multiplication
- Dilation with “unit impulse”

$$d[\alpha, \beta] = \begin{cases} 0 & \alpha = \beta = 0 \\ -\infty & \text{else} \end{cases}$$

does not change input signal:

$$f[x, y] = \sup_{\alpha, \beta} \{ f[x - \alpha, y - \beta] + d[\alpha, \beta] \}$$

Flat Dilation as Special Case

- Find $w[\alpha, \beta]$ such that

$$f[x, y] = \sup_{\alpha, \beta} \{ f[x - \alpha, y - \beta] + w[\alpha, \beta] \} = \text{dilate}(f, W)$$

- Answer:

$$w[\alpha, \beta] = \begin{cases} 0 & [\alpha, \beta] \in \Pi_{xy} \\ -\infty & \text{else} \end{cases}$$

- Hence, write in general

$$\begin{aligned} g[x, y] &= \sup_{\alpha, \beta} \{ f[x - \alpha, y - \beta] + w[\alpha, \beta] \} \\ &= \text{dilate}(f, w) = \text{dilate}(w, f) \end{aligned}$$

General Erosion for Gray-Level Images

- General erosion operator

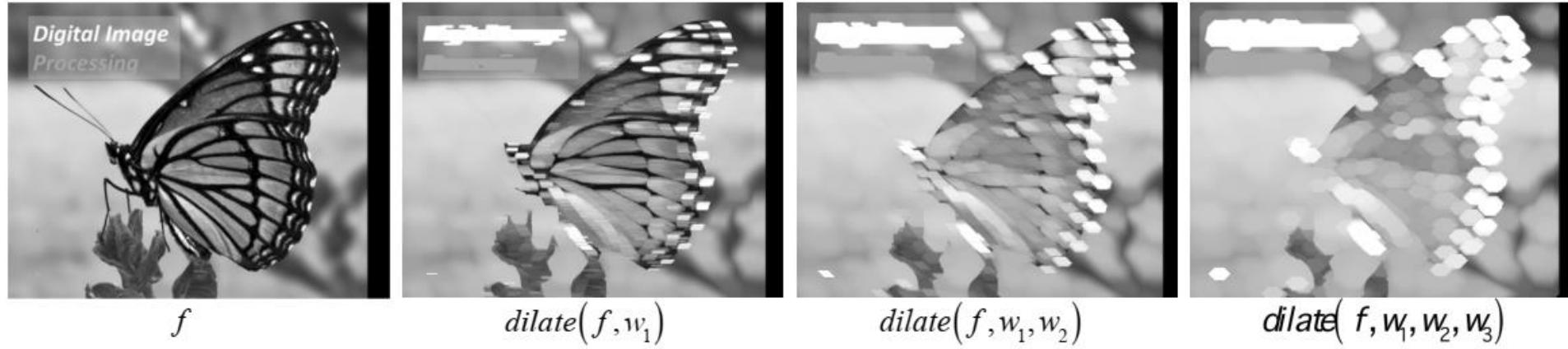
$$g[x, y] = \inf_{\alpha, \beta} \{ f[x + \alpha, y + \beta] - w[\alpha, \beta] \} = \text{erode}(f, w)$$

- Dual of dilation

$$\begin{aligned} g[x, y] &= \inf_{\alpha, \beta} \{ f[x + \alpha, y + \beta] - w[\alpha, \beta] \} \\ &= -\sup_{\alpha, \beta} \{ -f[x + \alpha, y + \beta] + w[\alpha, \beta] \} = -\text{dilate}(-f, \hat{w}) \end{aligned}$$

- Flat erosion contained as a special case

Cascaded dilations



$$\text{dilate} \left[\text{dilate} \left(f, w_1 \right), w_2 \right] = \text{dilate} \left(f, w \right)$$

where $w = \text{dilate} \left(w_1, w_2 \right)$

Cascaded Erosions

- Cascaded erosions can be lumped into single erosion

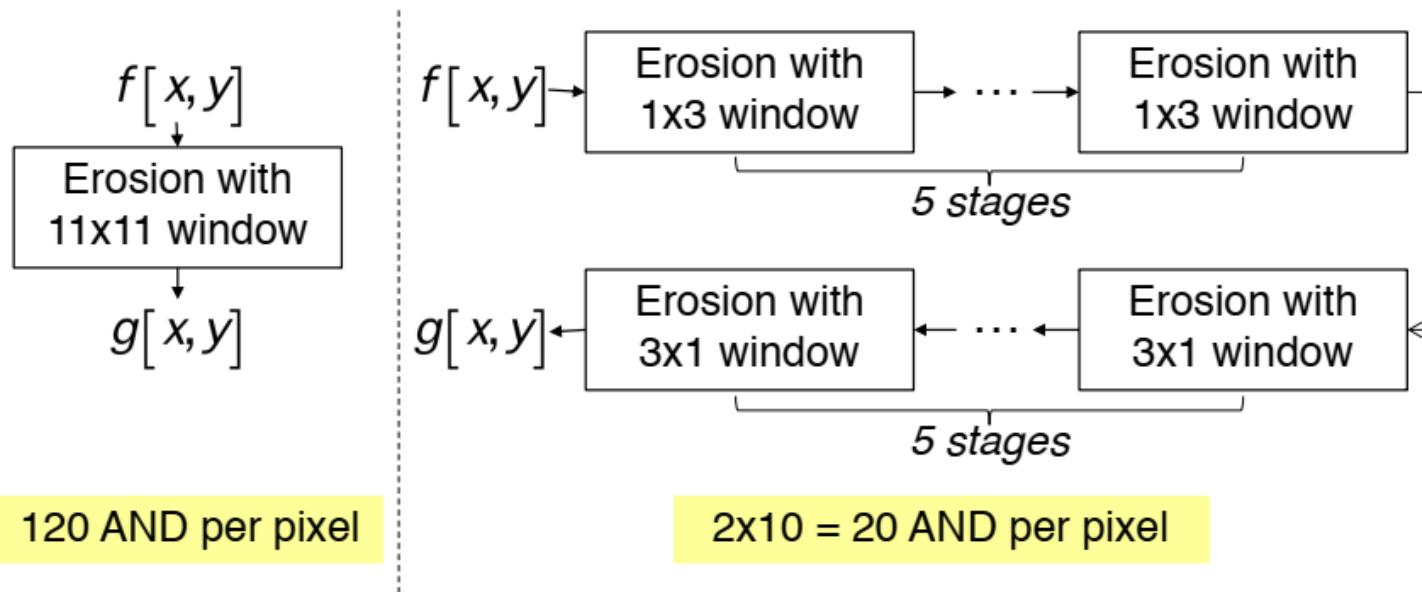
$$\begin{aligned} \text{erode}[\text{erode}(f, w_1), w_2] &= \text{erode}[-\text{dilate}(-f, \hat{w}_1), w_2] \\ &= -\text{dilate}[\text{dilate}(-f, \hat{w}_1), \hat{w}_2] \\ &= -\text{dilate}(-f, \hat{w}) \\ &= \text{erode}(f, w) \end{aligned}$$

where $w = \text{dilate}(w_1, w_2)$

- New structuring element (SE) is not the erosion of one SE by the other, but dilation.

Fast Dilation and Erosion

- Idea: build larger dilation and erosion operators by cascading simple, small operators
- Example: binary erosion by 11×11 window



Rank Filters

- Generalisation of flat dilation/erosion: in lieu of min or max value in window, use the p-th ranked value
- Increases robustness against noise
- Best-known example: median filter for noise reduction
- Concept useful for both gray-level and binary images
- All rank filters are commutative with thresholding

Median Filter

- Gray-level median filter

$$g[x, y] = \text{median}\left[W\{f[x, y]\}\right] := \text{median}(f, W)$$

- Binary images: majority filter

$$g[x, y] = MAJ\left[W\{f[x, y]\}\right] := \text{majority}(f, W)$$

- Self-duality

$$\text{median}(f, W) = -[\text{median}(-f, W)]$$

$$\text{majority}(f, W) = NOT\left[\text{majority}(NOT[f], W)\right]$$

Examples

Majority filter:



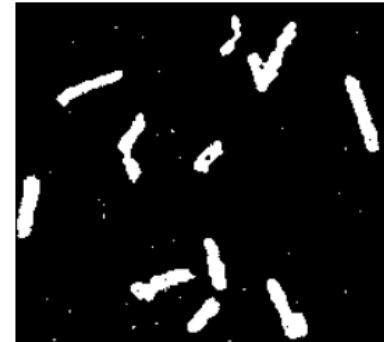
Binary image with
5% 'Salt&Pepper' noise



3x3 majority filter



20% 'Salt&Pepper' noise



3x3 majority filter

Median filter



Original image



5% 'Salt&Pepper' noise



3x3 median filtering



7x7 median filtering

Example: non-uniform lighting compensation



Original image
1632x1216 pixels

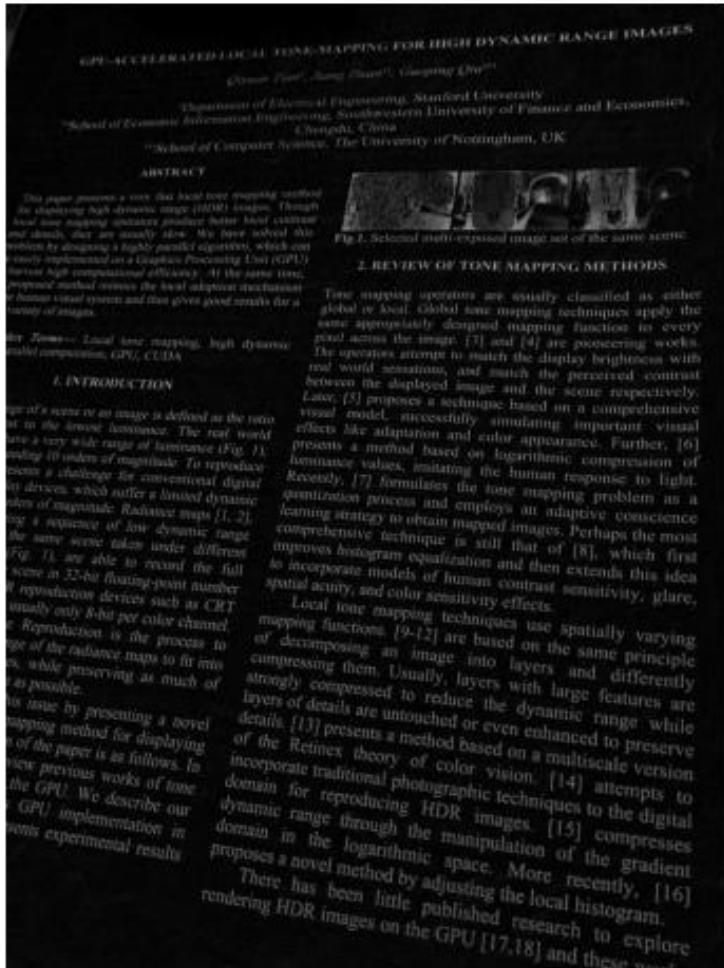


Dilation (local max)
61x61 structuring element

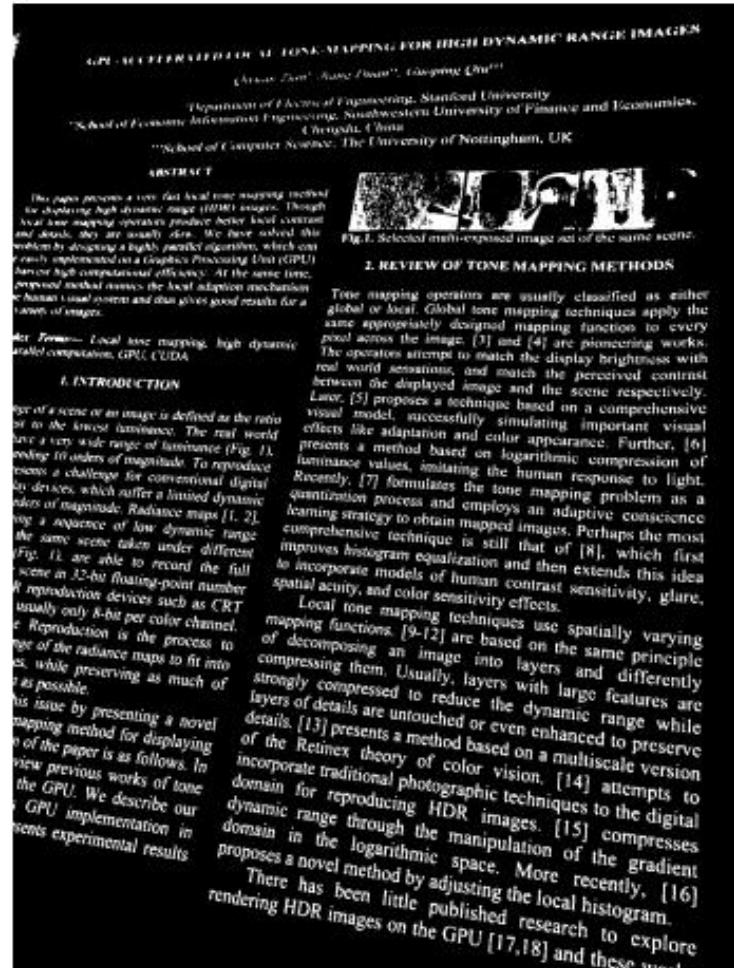


Rank filter
10th brightest pixel
61x61 structuring element

Example: Non-uniform lighting compensation



Background – original image



After global thresholding