Бинарные изображения Морфология

Лекция 3

Лектор: Демяненко Я.М., ФИИТ, 2026

Где мы находимся

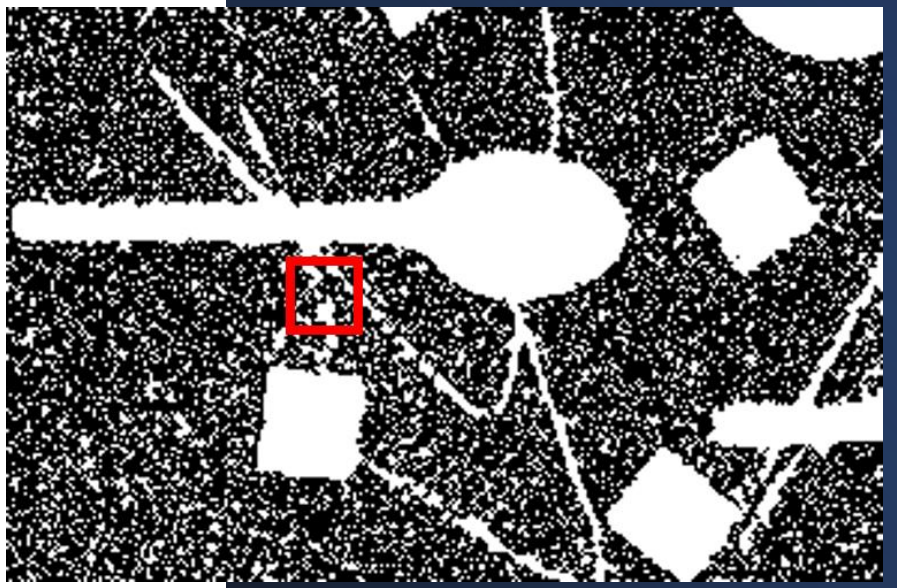
- **Напоминание**:
  - лекция 2 — яркостные преобразования
  - лекция 3 — бинаризация и Оцу
  - лекция 4 — морфология бинарных изображений

- **Схема pipeline**: изображение → grayscale → бинаризация → морфология → анализ

- **Тезис**: бинаризация даёт маску, но маска грязная

Проблема сырой маски

Пример: маска после Оцу с шумом, разрывами, слипшимися объектами



Три вопроса:

- как убрать шум?
- как разделить слипшееся?
- как измерить объект?

Тезис лекции: морфология — инструмент, превращающий сырую маску в пригодную для анализа

План лекции

- **Бинарные изображения и связность**
- Математическая морфология: базис
- Составные операции
- Пять прикладных задач
- Мостик к CNN

Бинарное изображение

Определение: 1 бит на пиксель, два значения (0/1, фон/объект)

Где возникает: пороговая бинаризация (лекция 3), маски CNN, ROI







Связность: 4 и 8

	1	
2	*	3
	4	

4-связность

1	2	3
4	*	5
6	7	8

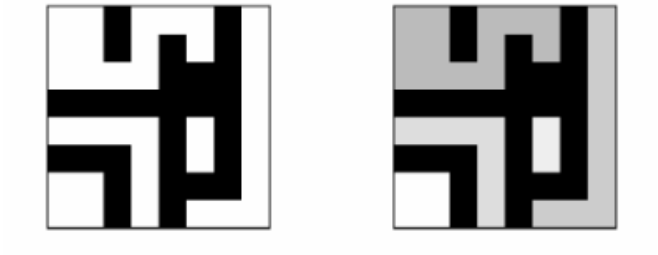
8-связность



|x1-x2| + |y1-y2|<=1

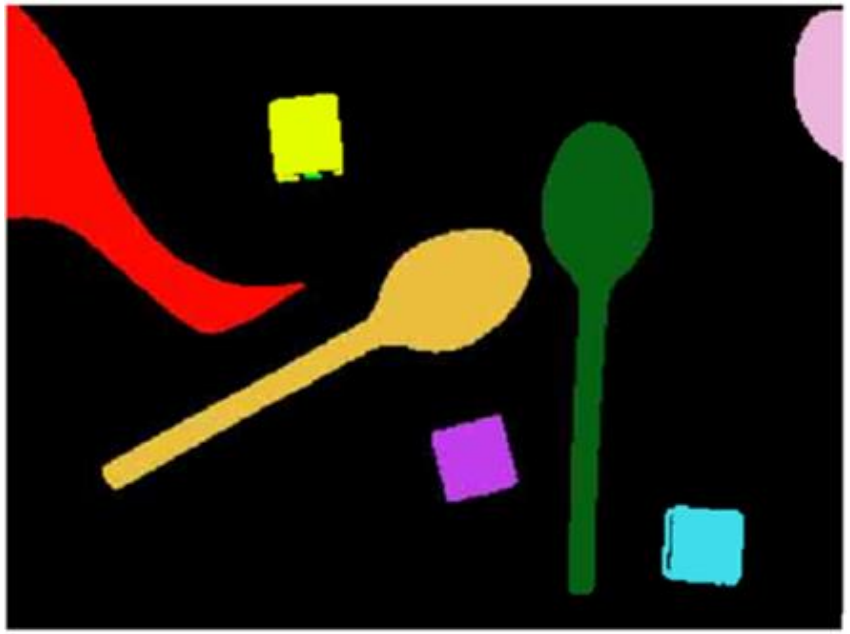
|x1-x2|<=1 и |y2-y1|<=1

Связные компоненты



1	1	0	1	1	1	0	1
1	1	0	1	0	1	0	1
1	1	1	1	0	0	0	1
0	0	0	0	0	0	0	1
1	1	1	1	0	1	0	1
0	0	0	1	0	1	0	1
1	1	0	1	0	0	0	1
1	1	0	1	0	1	1	1

1	1	0	1	1	1	0	2
1	1	0	1	0	1	0	2
1	1	1	1	0	0	0	2
0	0	0	0	0	0	0	2
3	3	3	3	0	4	0	2
0	0	0	3	0	4	0	2
5	5	0	3	0	0	0	2
5	5	0	3	0	2	2	2



Определение: максимальное множество связанных пикселей

Тезис: компоненты — основа подсчёта и фильтрации объектов

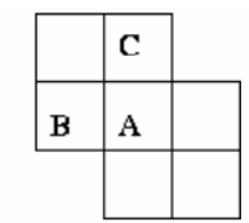
Визуализация: объекты разными цветами

Маркировка компонент: идея

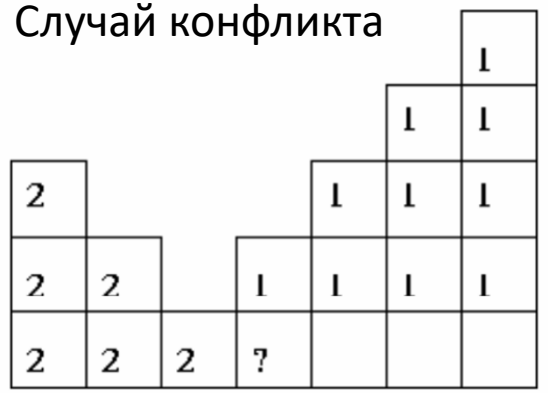
Два подхода

- Рекурсивный алгоритм маркировки (в памяти целиком) — алгоритм заливки
- Алгоритм построчной маркировки (большие изображения — обрабатываются порциями по 2 строки)

```
if A = 0  do nothing
else if (not B labeled) and (not C labeled)
        increment label numbering and label A
else if B xor C labeled
        copy label to A
else if B and C labeled
        if B label = C label
                copy label to A
else
        copy either B label or C label to A
```



Случай конфликта



Маркировка компонент: идея

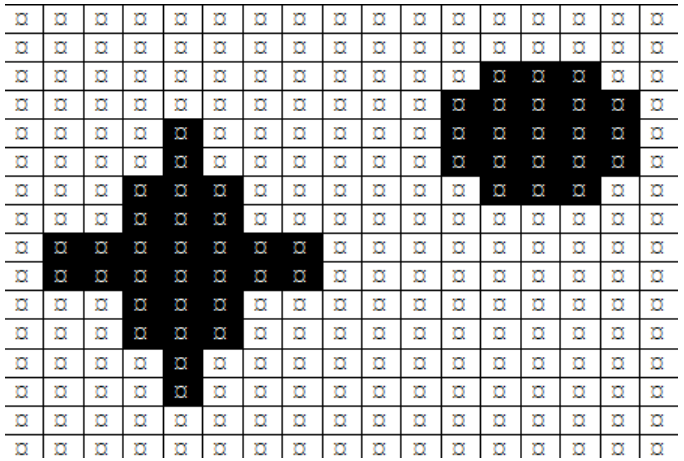
1	1	0	2	2	2	0	3
1	1	0	2	0	2	0	3
1	1	1	1	0	0	0	3
0	0	0	0	0	0	0	3
4	4	4	4	0	5	0	3
0	0	0	4	0	5	0	3
6	6	0	4	0	0	0	3
6	6	0	4	0	7	7	3

1	1	0	1	1	1	0	3
1	1	0	1	0	1	0	3
1	1	1	1	0	0	0	3
0	0	0	0	0	0	0	3
4	4	4	4	0	5	0	3
0	0	0	4	0	5	0	3
6	6	0	4	0	0	0	3
6	6	0	4	0	3	3	3

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

Резюме: O(N) по времени, память зависит от подхода

Метрики: геометрия



Площадь

$$A = \sum_{(r,c)\in R} 1$$

Центр тяжести

$$\bar{r} = \frac{1}{A}\sum (r,c) \in R^r$$

$$\bar{c} = \frac{1}{A}\sum (r,c) \in R^c$$

Периметр

$$|P| = \left|\left\{k \middle| (r_{r+1}, c_{r+1}) \in N_4(r_r, c_r)\right\}\right|$$

$$+\sqrt{2}\left|\left\{k \middle| (r_{r+1}, c_{r+1}) \in N_8(r_r, c_r) - N_4(r_r, c_r)\right\}\right|$$

Метрики: форма

Округлость

$$C_1 = \frac{|P|^2}{A}$$

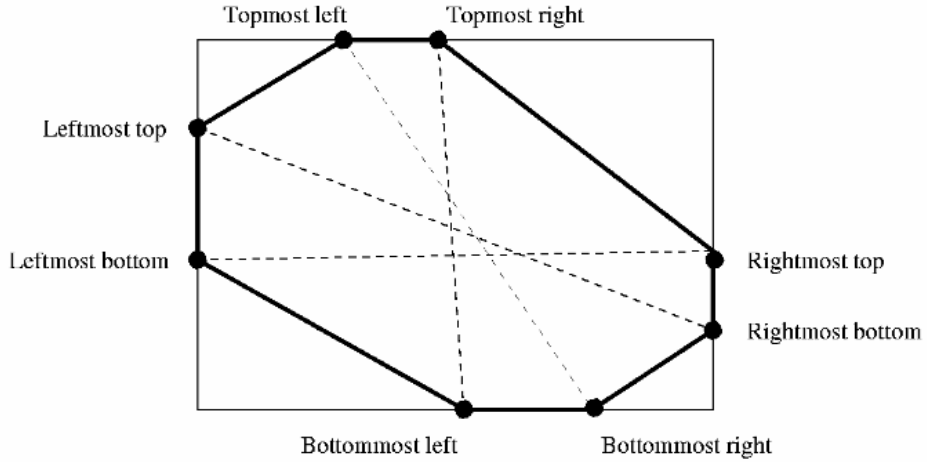
Округлость Харалика 1974

$$C_2 = \frac{\mu_R}{\sigma_R}$$

$$\mu_R = \frac{1}{K}\sum_{k=0}^{K-1}\left\|(r_k, c_k) - (\bar{r}, \bar{c})\right\|$$

$$\sigma_R = \left(\frac{1}{K}\sum_{k=0}^{K-1}\left[\left\|(r_k, c_k) - (\bar{r}, \bar{c})\right\| - \mu_R\right]^2\right)^{\frac{1}{2}}$$

Описывающий прямоугольник и экстремальные точки



$$D = \sqrt{(r_2 - r_1)^2 + (c_2 - c_1)^2} + Q(\Theta)$$

$$Q(\Theta) = \begin{cases} \dfrac{1}{|\cos\Theta|} & : \ |\Theta| < 45° \\ \dfrac{1}{|\sin\Theta|} & : \ |\Theta| > 45° \end{cases}$$

Центральные моменты второго порядка

$$\mu_{rr} = \frac{1}{A}\sum_{(r,c)\in R}(r - \bar{r})^2$$

$$\mu_{cc} = \frac{1}{A}\sum_{(r,c)\in R}(c - \bar{c})^2$$

$$\mu_{rc} = \frac{1}{A}\sum_{(r,c)\in R}(r - \bar{r})(c - \bar{c})$$

Сравните значения метрик округлости

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	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	0	0
2	2	2	2	0	0	0	0	0	1	1	1	1	1	1	0
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	0	1	1	1	1	1	1	0
2	2	2	2	0	0	0	0	0	0	1	1	1	1	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0

labeled connected-components image

region num.	region area	row of center	col of center	perim. length	circu-larity$_1$	circu-larity$_2$	radius mean	radius var.
1	44	6	11.5	21.2	10.2	15.4	3.33	.05
2	48	9	1.5	28	16.3	2.5	3.80	2.28
3	9	13	7	8	7.1	5.8	1.2	0.04

properties of the three regions

Мостик

Компоненты — подсчёт объектов, фильтрация шума, признаки для ML

Но: шум и слипания мешают

Нужна морфология

План лекции

- Бинарные изображения и связность
- **Математическая морфология: базис**
- Составные операции
- Пять прикладных задач
- Мостик к CNN

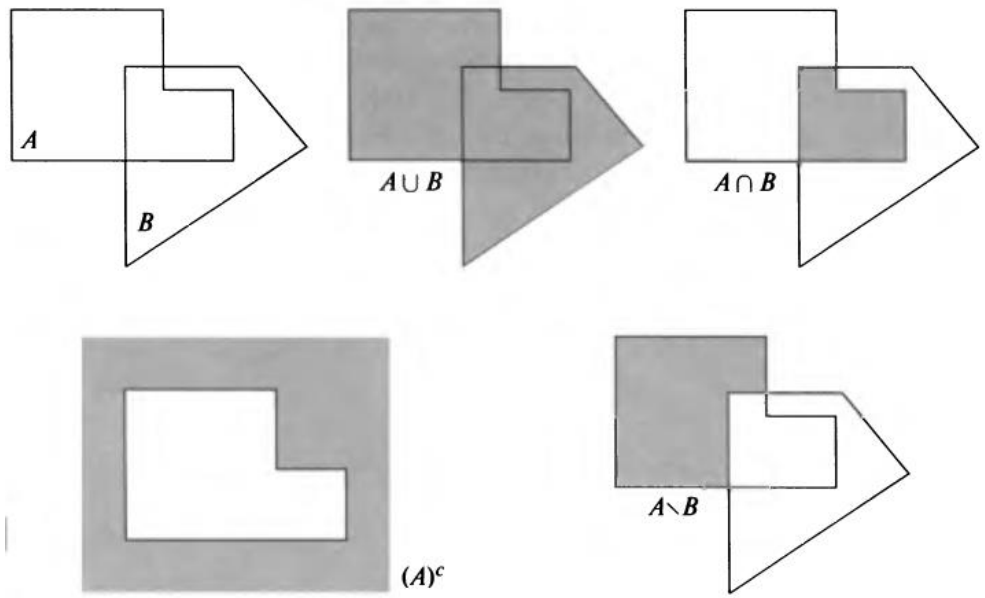
Что такое математическая морфология

Определение: теория анализа форм на основе теории множеств и решёток

История: Serra, Matheron, 1960-е

Тезис: морфология — не набор трюков, а алгебра форм

Теория множеств: базовые операции



- Два множества A и B
- Объединение множеств A и B
- Пересечение множеств A и B
- Дополнение множеств A и B
- Разность множеств A и B

Центральное отражение

$$\hat{B} = \{w \mid w = -b,\ b \in B\}$$

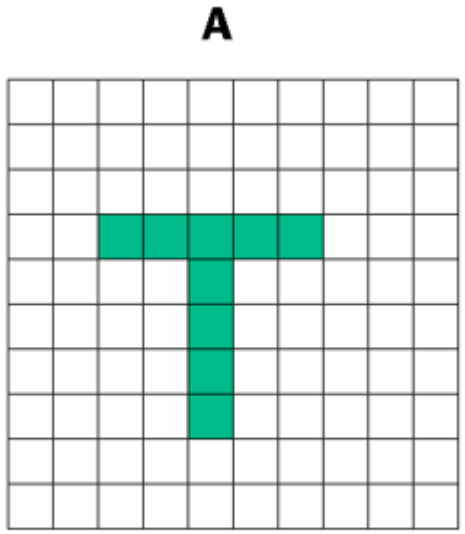
Параллельный перенос

$$(A)_z = \{c \mid c = a + z,\ a \in A\}$$

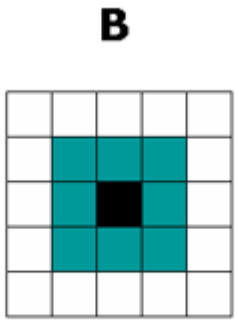
Базовые операции морфологии

- Расширение или наращивание (Dilation)
- Сужение или эрозия (Erosion)
- Замыкание (closing)
- Размыкание (opening)

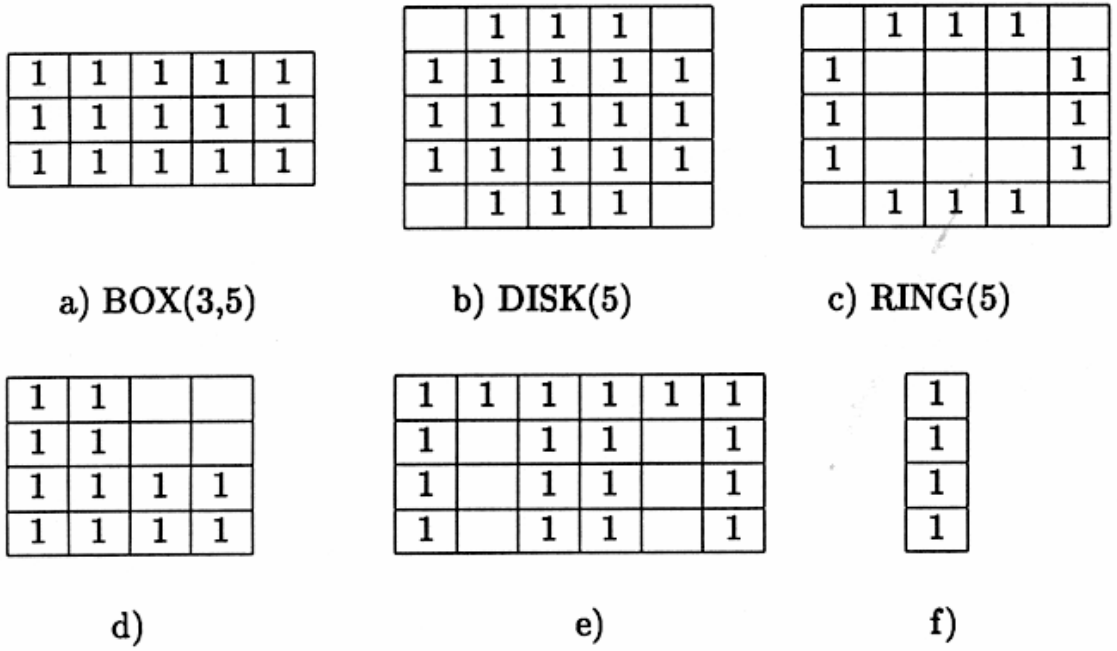
Структурный элемент



Объект обработки



Структурный элемент – инструмент



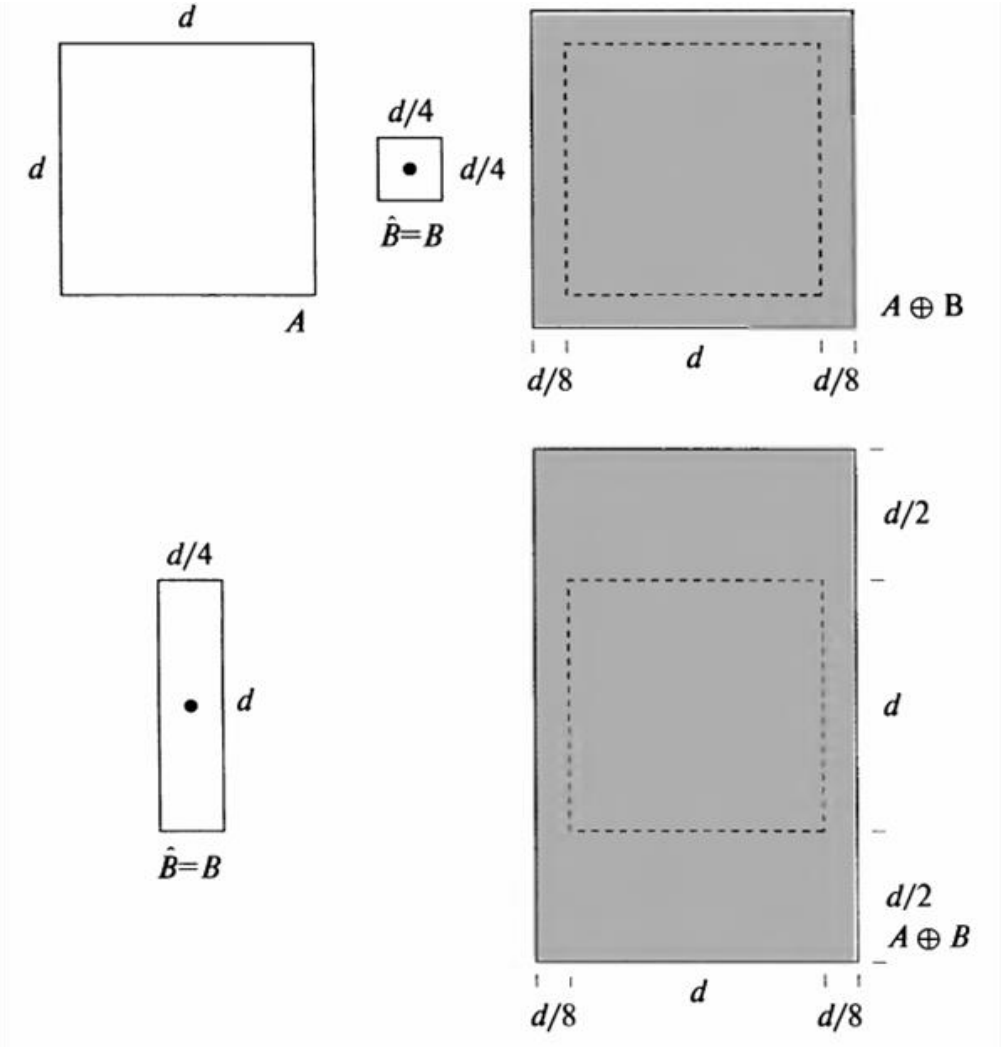
Примеры структурирующих элементов

Дилатация множества A по множеству B. Определение

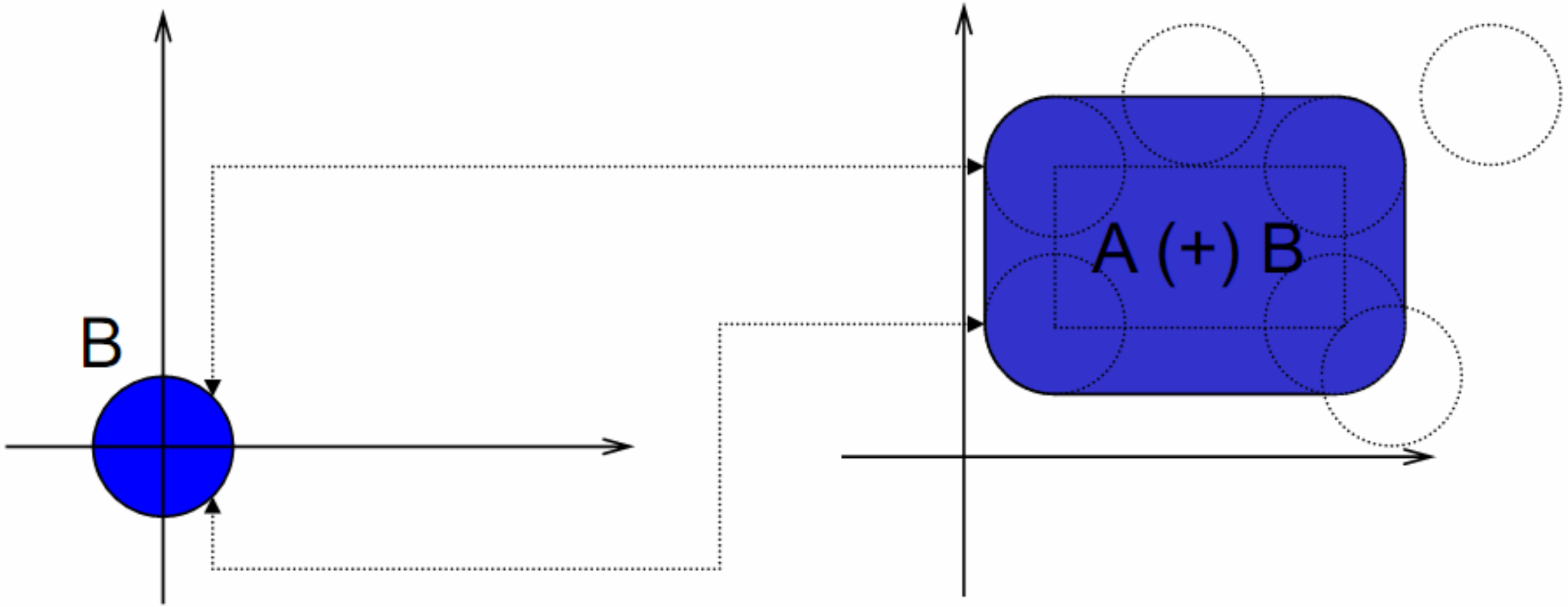
Эквивалентные определения

$$A \oplus B = \left\{ z \mid (\hat{B})_z \cap A \neq \varnothing \right\}$$

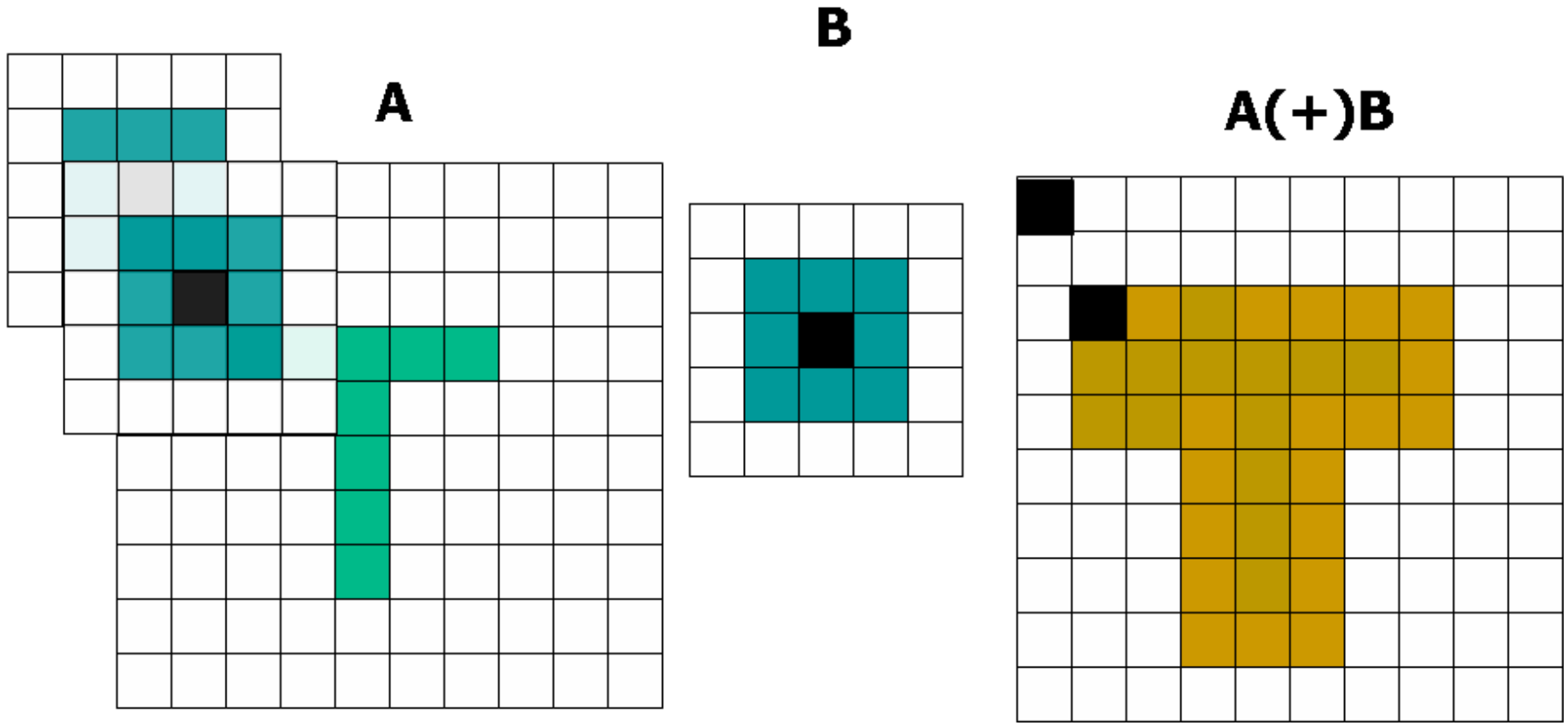
$$A \oplus B = \bigcup_{a \in A} B_a$$



Расширение или наращивание (Dilation): геометрическая интерпретация



Дилатация: пример



Эффект: объект растёт, дырки заполняются, разрывы соединяются

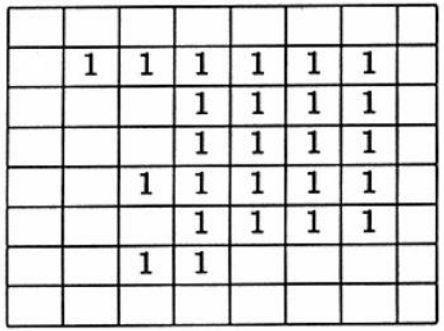
Дилатация: связь с суммой Минковского

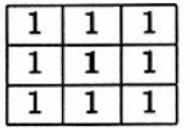
$A \oplus B = A + \hat{B}$ (Minkowski sum)

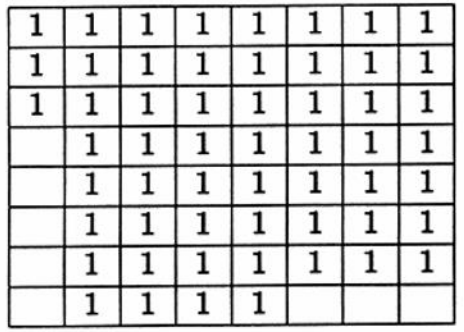
Тезис: это алгебраическая операция

Алгоритм реализации

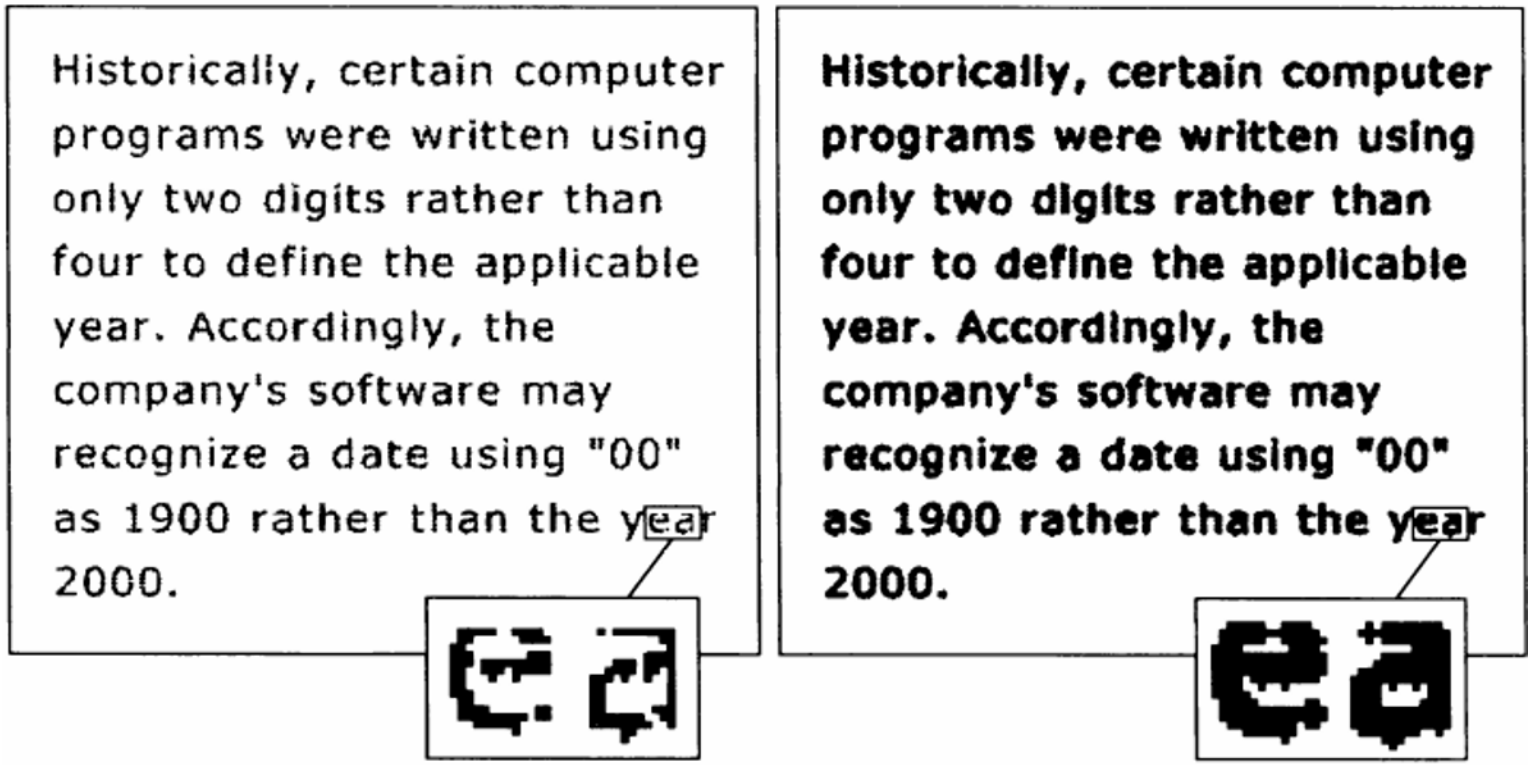
```
void Dilation(BIT*src[], bool*mask[], BIT*dst[]) {
//W,H – размеры исходного и результирующего изображений
//MW,MH – размеры структурного множества
for(y = MH/2; y <  H – MH/2;y++)
  for(x = MW/2; x < W – MW/2;x++) {
    BIT max=0;
    for(j = - MH/2; j <= MH/2; j++)
      for( i= - MW/2; i <= MW/2; i++)
        if (mask[i][j] && src[x+i][y+j]>max)
          max=src[x+i][y+j];
    dst[x][y] = max;
  }
}
```







Применение: соединение разрывов в тексте, OCR



Max длина разрыва = 2 пиксела

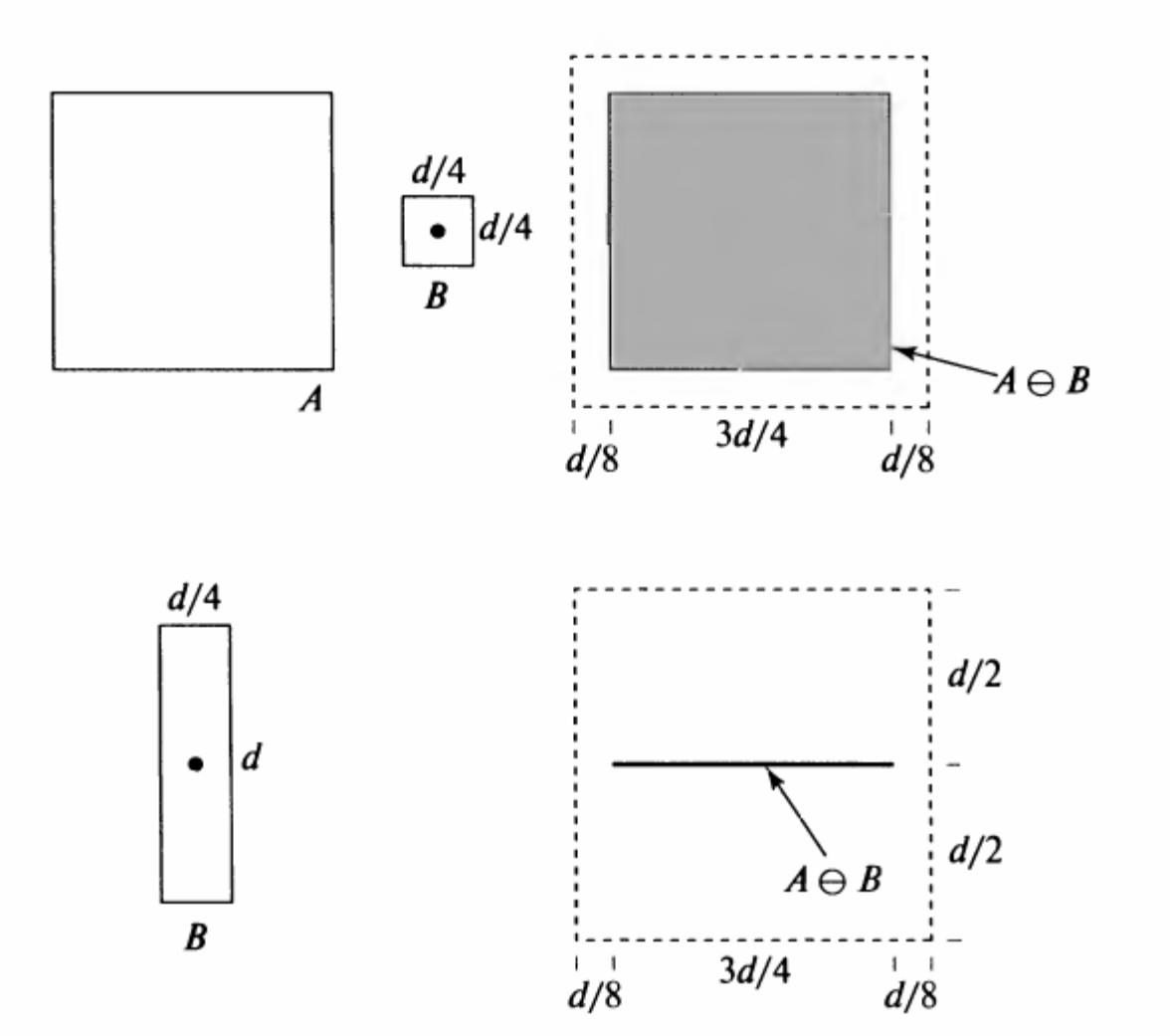
0	1	0
1	1	1
0	1	0

Эрозия множества A по множеству B . Определение

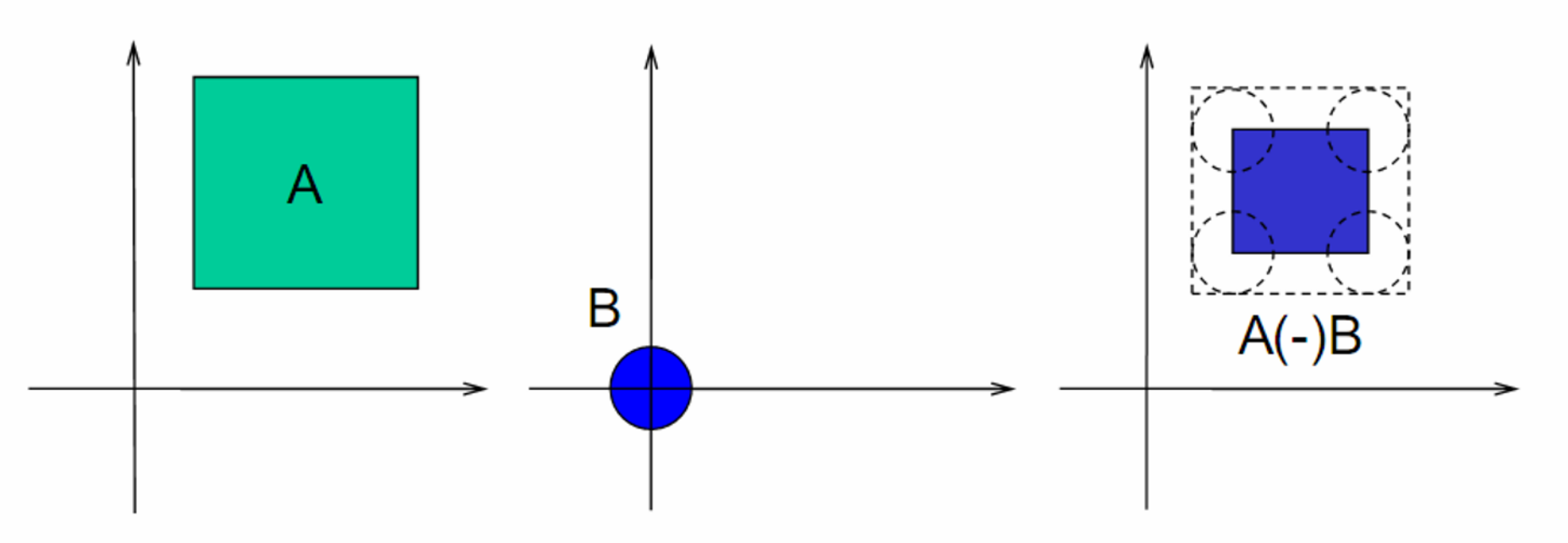
$$A \ominus B = \{z \mid (B)_z \subseteq A\}$$

$$A(-)B = \{a \mid a + b \in A \forall b \in B\}$$

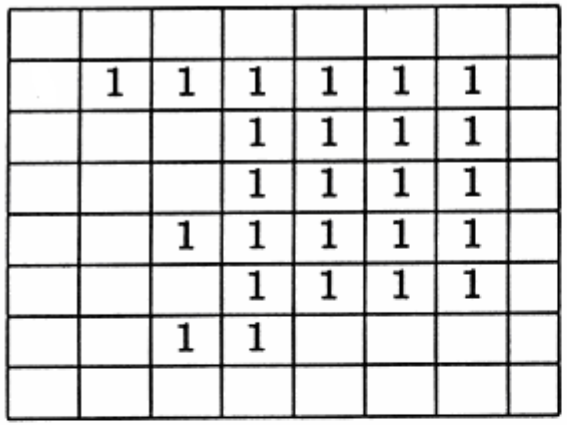
A (-) B = (A^C (+) B)^C,
где A^C – дополнение A

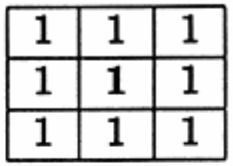


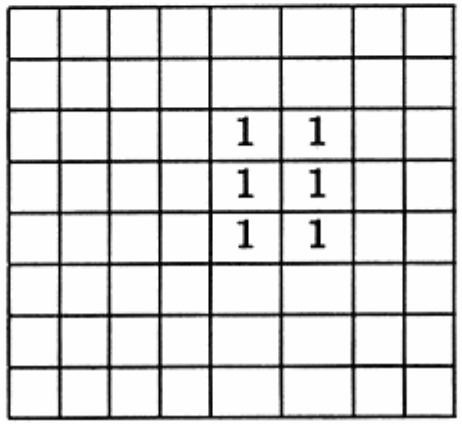
Сужение или эрозия (Erosion): геометрическая интерпретация



Эрозия: пример







Эффект:

- объект сжимается
- мелкий шум исчезает
- тонкие связи рвутся

Эрозия: связь с разностью Минковского

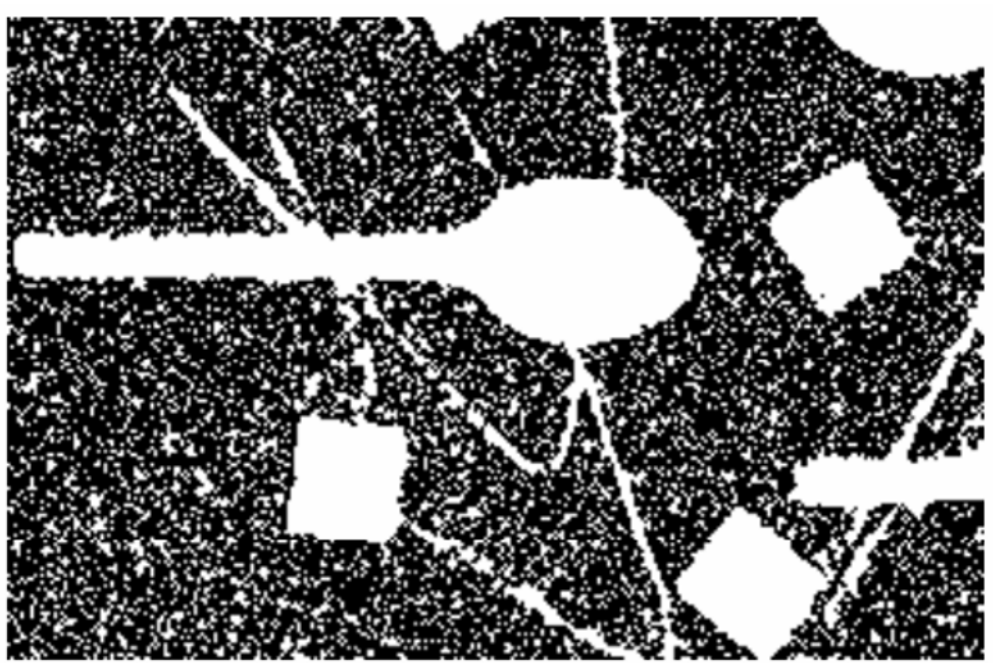
$A \ominus B = A \setminus \check{B}$

Алгоритм морфологического сужения

```
void Erosion(BIT*src[],bool*mask[],BIT*dst[]) {
//W,H – размеры исходного и результирующего изображений
//MW,MH – размеры структурного множества
for(y = MH/2; y < H – MH/2 ; y++)
  for(x = MW/2; x < W – MW/2; x++) {
    BIT min = MAXBIT;
    for(j = -  MH/2; j <= MH/2; j++)
       for( i = - MW/2; i <= MW/2; i++)
          if (mask[i][j] && src[x+i][y+j]<min)
             min=src[x+i][y+j];
     dst[x][y] = min;
    }
}
```

Применение эрозии

- удаление шума
- разделение тонких связей
- выделение ядра объекта





Результат применения эрозии с разными ядрами



$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & [1] & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & [1] & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & [1] & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Двойственность

$$(A \ominus B)^c = A^c \oplus \hat{B} \qquad (A \oplus B)^c = A^c \ominus \hat{B}$$

Дилатация и **эрозия** являются **двойственными** операциями по отношению к теоретико-множественным операциям дополнения и центрального отражения

A (-) B = (A^C (+) B)^C,
где A^C – дополнение A

Тезис: erosion и dilation — зеркала друг друга

Формальное доказательство

$$(A \ominus B)^c = A^c \oplus \hat{B}$$

$$(A \ominus B)^c = \{z \mid (B)_z \subseteq A\}^c$$

$$(A \ominus B)^c = \{z \mid (B)_z \cap A^c = \varnothing\}^c$$

$$(A \ominus B)^c = \{z \mid (B)_z \cap A^c \neq \varnothing\} = A^c \oplus \hat{B}$$

Свойства: коммутативность и ассоциативность

Дилатация

$A \oplus B = B \oplus A$; $A \oplus (B \oplus C) = (A \oplus B) \oplus C$

Эрозия

— не коммутативна

Коммутативный закон

- A (+) B = B (+) A
- A (-) B < > B (-) A

Ассоциативный закон

- A (+) (B (+) C) = (A (+) B) (+) C
- A (-) (B (-) C) = (A (-) B) (-) C

Свойства: монотонность и экстенсивность

Монотонность: $A \subseteq B \Longrightarrow A \oplus C \subseteq B \oplus C$

Экстенсивность дилатации: $A \subseteq A \oplus B$

Анти-экстенсивность эрозии: $A \ominus B \subseteq A$

Тезис: эти свойства — основа для доказательства свойств opening/closing

Результат применения эрозии и дилатации



(а) Изображение, содержащее квадраты со сторонами 1, 3, 5, 7, 9 и 15 пикселей. (б) Результат эрозии изображения (а) по квадратному примитиву размерами 13×13 пикселей, заполненному единицами. (в) Результат дилатации изображения (б) по тому же примитиву.

Greyscale-морфология

Тезис: морфология работает не только на бинарных, но и на полутоновых изображениях

Морфология Серра или **функциональная** (или **плоская**) морфология
Формулы Серра: $(f \oplus b)(x) = \sup_y \{f(x-y) + b(y)\}$; $(f \ominus b)(x) = \inf_y \{f(x+y) - b(y)\}$
здесь:
f — исходная функция (изображение),
b — структурирующий элемент — маленькая функция-шаблон,
x, y — точки пространства (обычно координаты пикселя),
sup — верхняя грань (максимум), inf — нижняя грань (минимум).

Смысл: дилатация раздувает яркие области. Если b — плоский элемент (например, b(y) = 0 в некоторой окрестности и $-\infty$ вне неё), то формула сводится к локальному максимуму:
$(f \oplus b)(x) = \max_{y \in B} f(x - y)$
то есть в каждой точке берётся максимум яркости в окрестности. Тёмные детали уменьшаются, светлые расширяются.
Если b не плоский, он добавляет «рельеф»: пики b усиливают соответствующие смещения

Пример: top-hat для выделения мелких светлых объектов

Резюме блока

Дилатация и эрозия — **базис**

Свойства: двойственность, монотонность, экстенсивность

Следующий шаг: составные операции (размыкание и замыкание)

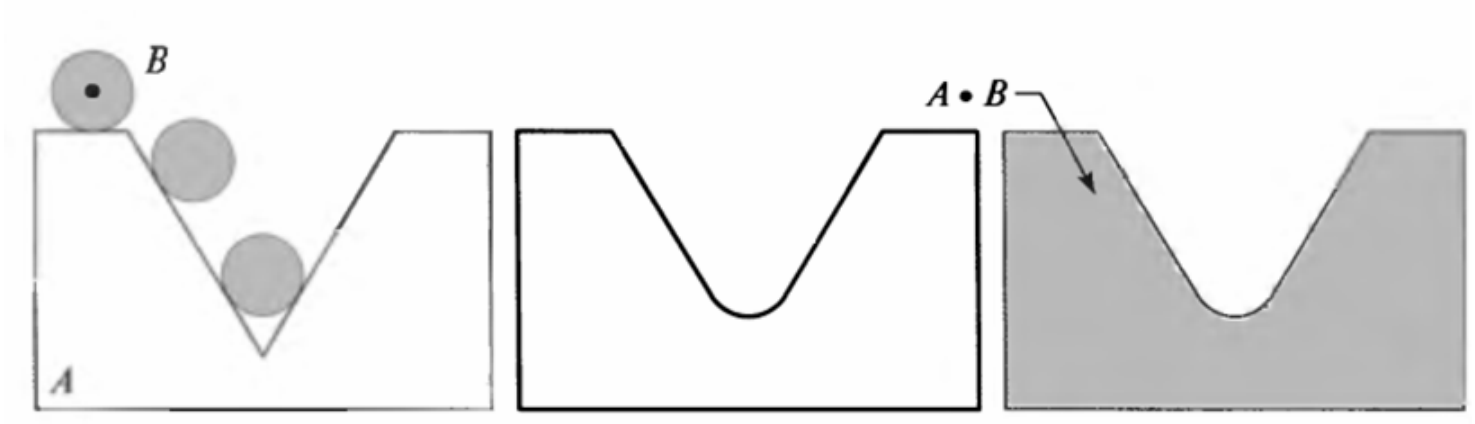
План лекции

- Бинарные изображения и связность
- Математическая морфология: базис
- **Составные операции**
- Пять прикладных задач
- Мостик к CNN

Замыкание, закрытие (closing)

Def: $A \bullet B = (A \oplus B) - B$

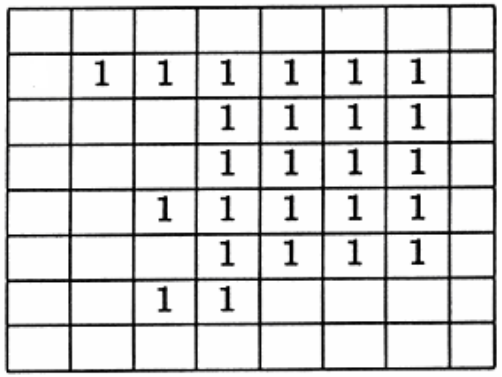
close(A, B) = (A (+) B) (-) B



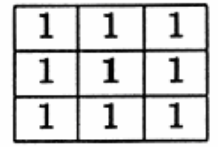
Геометрическая интерпретация: прокатывание B снаружи A
Эффект: заполняет впадины и дырки

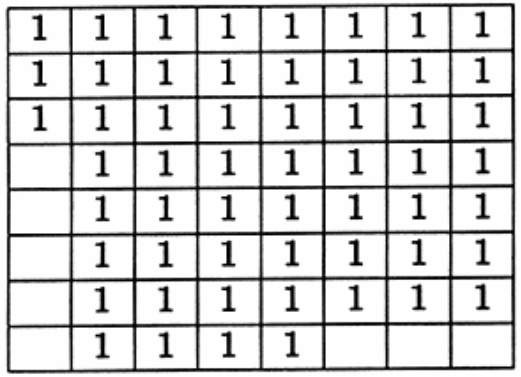
Пример замыкания

$$B \bullet S = (B \oplus S) - S$$

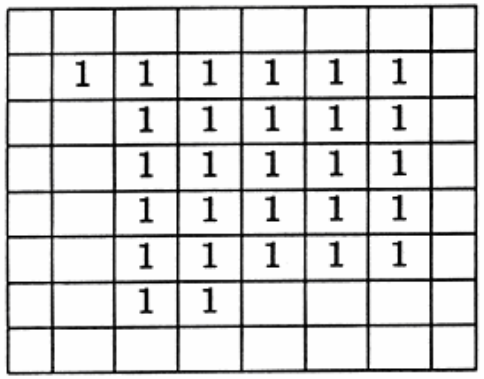


a) Binary image B





c) Dilation $B \oplus S$

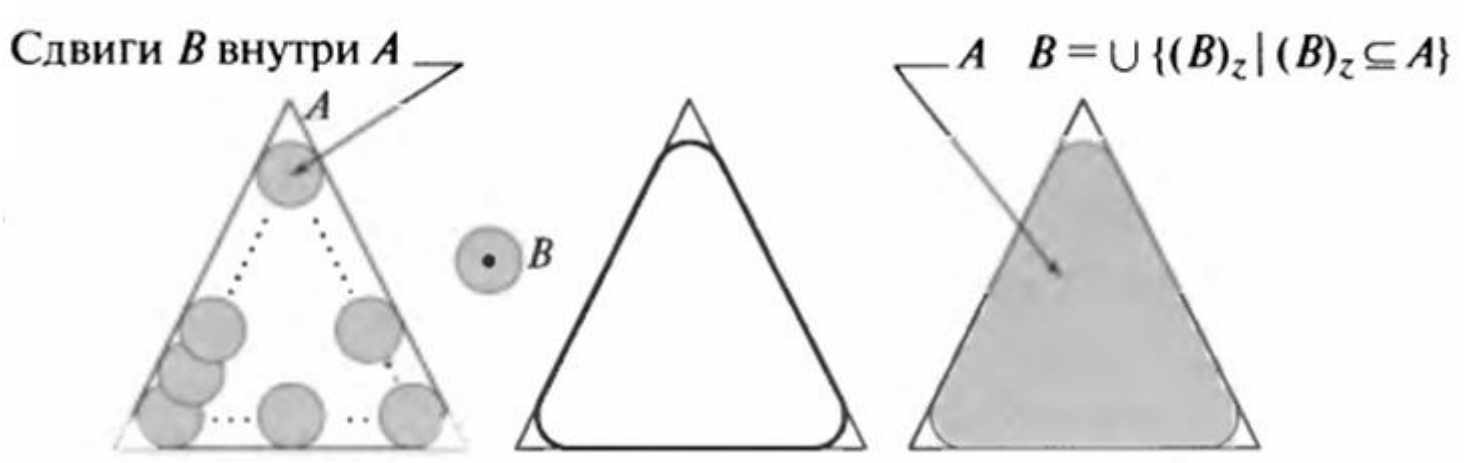


e) Closing $B \bullet S$

Интуиция: closing сглаживает контур снаружи

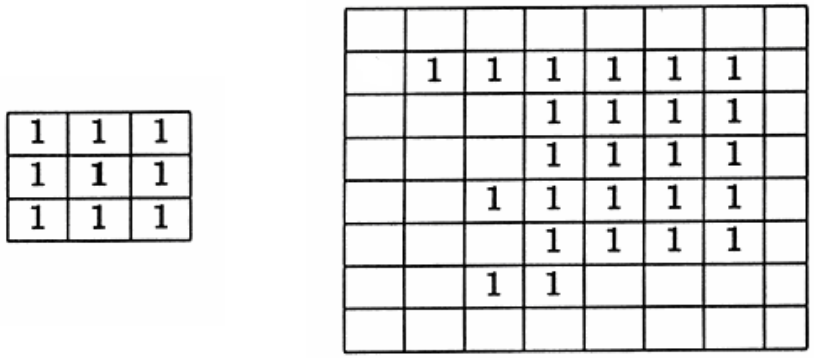
Размыкание, раскрытие (opening)

Def: $A \circ B = (A(-)B) \oplus B$



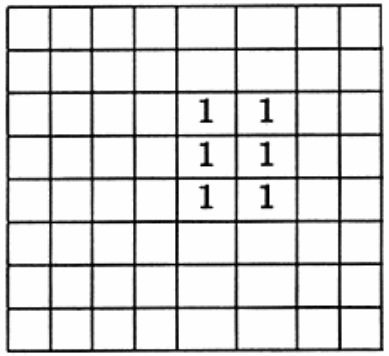
Геометрическая интерпретация: прокатывание B внутри A, объединение всех положений
Эффект: удаляет выступы

Пример размыкания

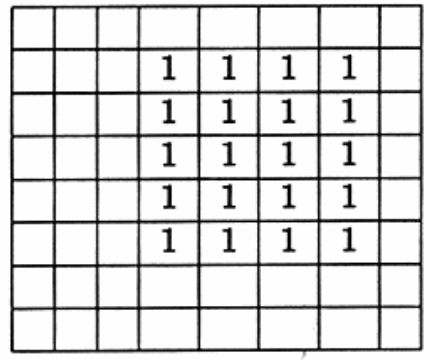


a) Binary image B

$$B \circ S = (B - S) \oplus S$$



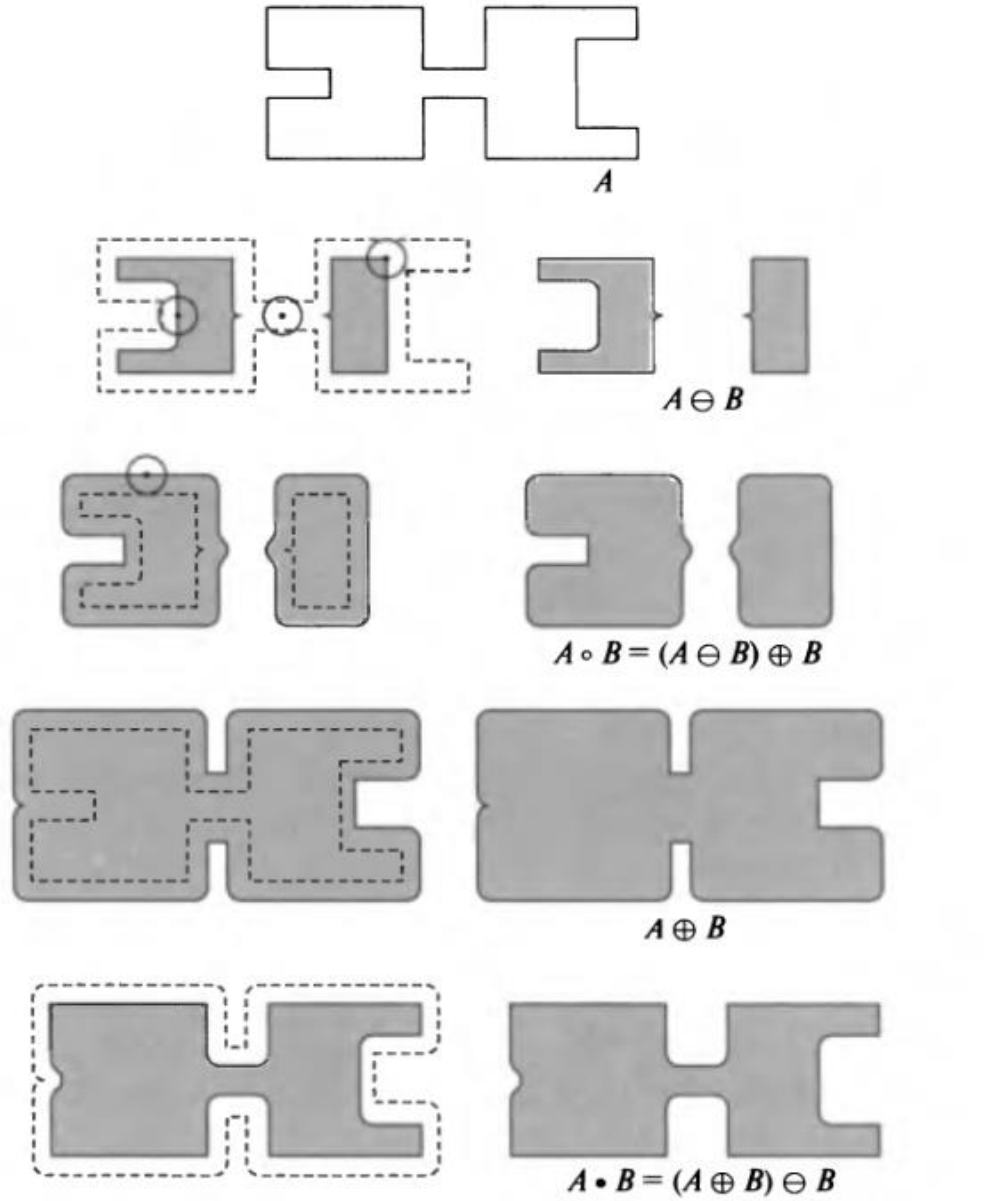
d) Erosion $B \ominus S$



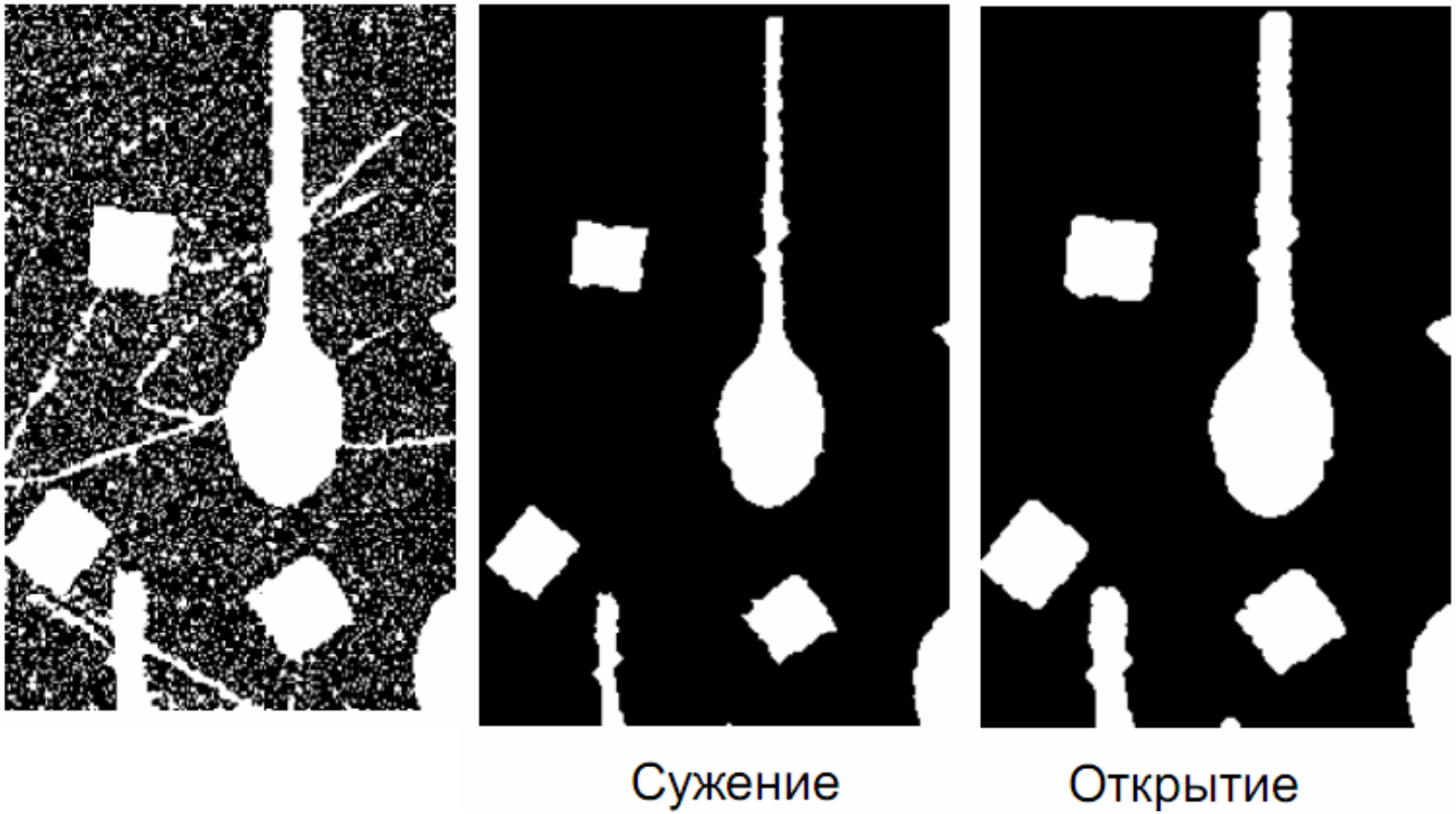
f) Opening $B \circ S$

Интуиция: opening «глаживает контур изнутри

Размыкание и замыкание



Сужение vs Размыкание



Свойства opening и closing

Экстенсивность closing:
$A \subseteq A \bullet B$

Анти-экстенсивность opening: $A \circ B \subseteq A$

Идемпотентность:

$(A \circ B) \circ B = A \circ B;$

$(A \bullet B) \bullet B = A \bullet B$

Монотонность

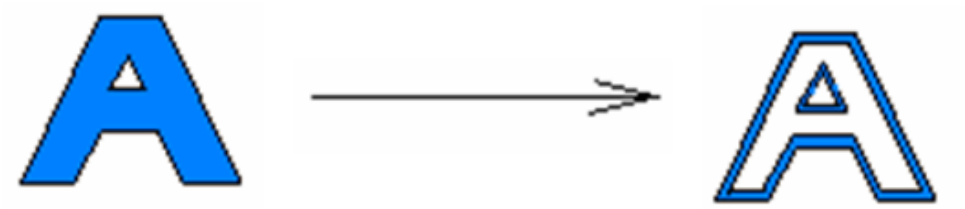
Тезис:
идемпотентность — почему повторное применение не меняет результат

Двойственность opening/closing

$$(A \bullet B)^c = A^c \circ \hat{B}$$

Тезис: opening и closing — зеркала, как erosion и dilation

Морфологический градиент



$\beta(A) = A \oplus B - A$ для внешней границы

или

$\beta(A) = A - A \ominus B$ для внутренней границы

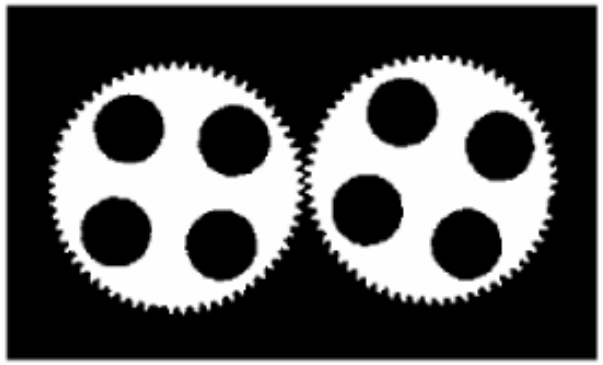
Применение: выделение контуров, альтернатива Canny для бинарных масок

Морфологическая фильтрация



Исходное изображение предоставлено Национальным институтом стандартов и технологии США

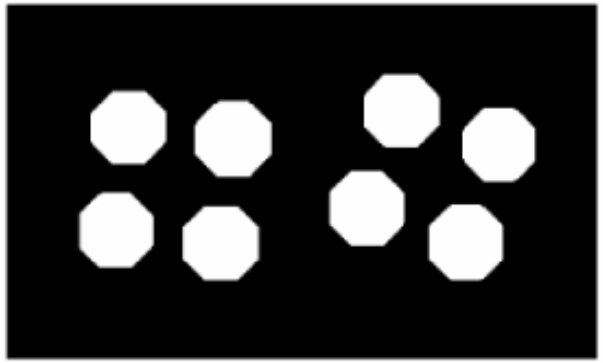
Контроль шестерёнок



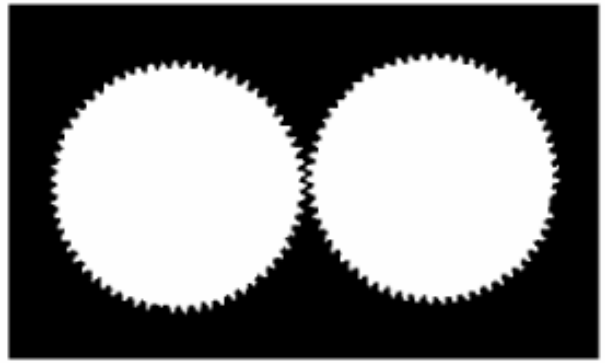
a) original image B



b) B1 = B $\ominus$ hole_ring

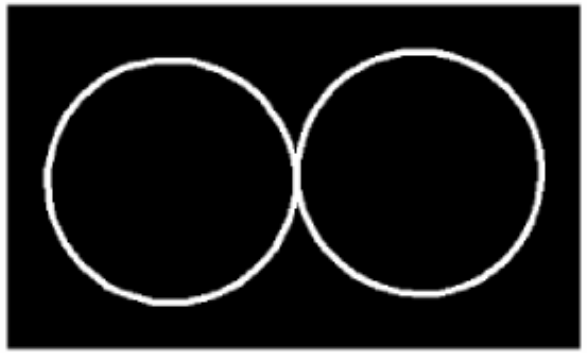


c) B2 = B1 $\oplus$ hole_mask

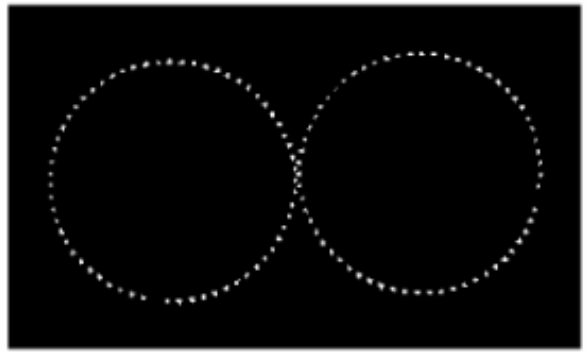


d) B3 = B OR B2

Контроль шестерёнок 2



e) B7 (see text)



f) B8 = B AND B7



g) B9 = B8 $\oplus$ tip_spacing



h) RESULT = ((B7 - B9) $\oplus$ defect_cue) OR B9

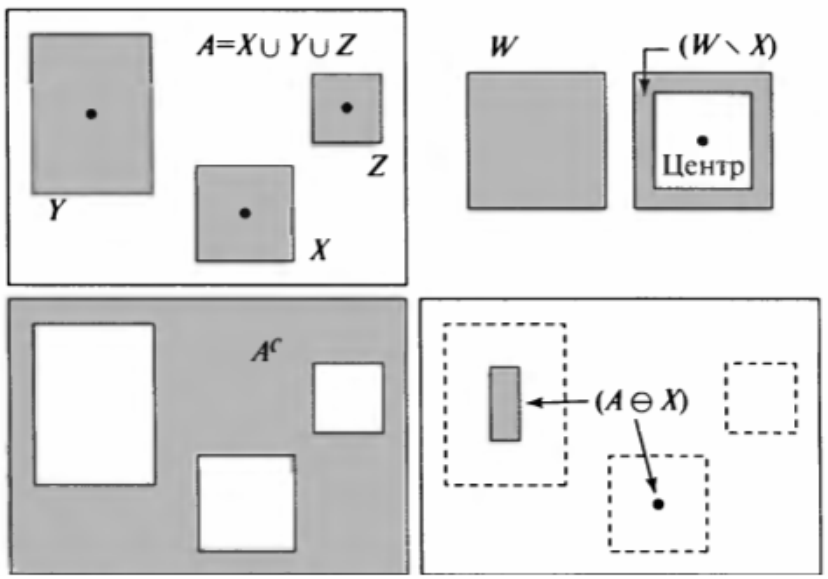
Top-hat и black-hat

Top-hat: $A - (A \circ B)$ — выделяет мелкие светлые объекты
Black-hat: $(A \bullet B) - A$ — выделяет мелкие тёмные объекты
Применение:

- выделение мелких объектов на сложном фоне (микроорганизмы, дефекты, трещины, сосуды, царапины)
- текста на неравномерном фоне
- локальное контрастирование (прибавив top-hat к изображению, усиливают мелкие детали, не трогая крупные)
- обнаружение дефектов
- выравнивает фон

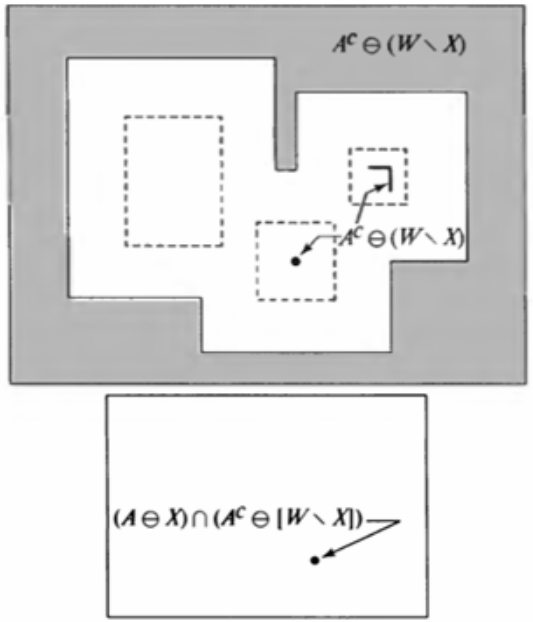
Преобразование попадание пропуск (hit or miss transform)

Назначение: поиск конкретных конфигураций (углы, тонкие линии, окончания)



$$A \circledast B = (A \ominus X) \cap \left[A^c \ominus (W \setminus X) \right]$$

$$A \circledast B = (A \ominus B_1) \cap (A^c \ominus B_2)$$



Skeleton (скелетизация)

Определение: множество центров максимальных вписанных шаров
Проще: для каждой точки $c \in X$ найдём **наибольший** шар с центром в c, целиком лежащий в X. Если этот шар **не содержится** ни в каком другом шаре внутри X (то есть он максимален по включению), то c принадлежит скелету.

Алгоритм: итеративное thinning (hit-or-miss + вычитание)

Свойства: связность, гомотопия, минимальная ширина

Применение: анализ сосудов, OCR, анализ дорог

Distance transform — преобразование расстояний

Определение: DT(x) = min расстояние до фона
Связь с эрозией: erosion с DISK(r) убирает пиксели с DT < r
Применение: watershed, boundary loss в CNN, скелет

Boundary loss в CNN (медицинская сегментация)
Проблема: обычные функции потерь (Dice, cross-entropy) плохо штрафуют ошибки на границах, особенно при сильном дисбалансе классов (тонкие структуры — сосуды, опухоли)
Идея: если предсказание промахивается мимо границы, разница ϕ_G даёт градиент, который тянет предсказание к правильной границе. DT здесь — **геодезическая карта расстояний до границы**, которая задаёт направление и величину ошибки

Резюме блока

Opening/closing — фильтры шума

Gradient, top-hat, hit-or-miss, skeleton — инструменты анализа

.

Следующий шаг: 5 прикладных задач

План лекции

- Бинарные изображения и связность
- Математическая морфология: базис
- Составные операции
- **Пять прикладных задач**
- Мостик к CNN

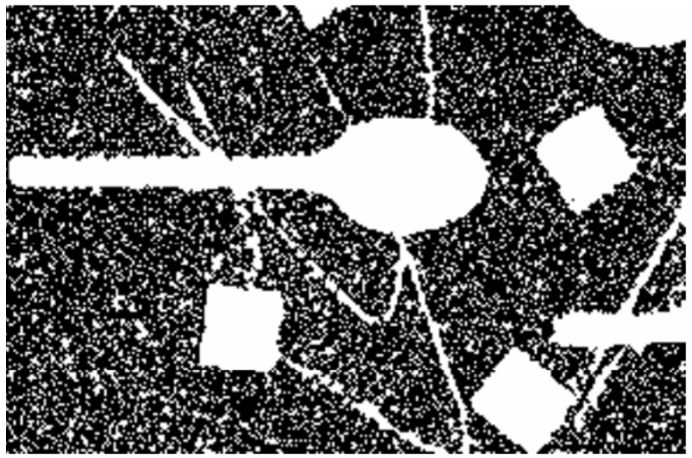
Обзор пяти прикладных задач

1. Шумоподавление
2. Заполнение дырок
3. Выделение границ
4. Скелетизация
5. Watershed

Формат: постановка → морфологическое решение → результат

Задача 1: шумоподавление

Постановка: маска с солью-перцем, мелкими точками
Решение: opening (убирает светлый шум) → closing (убирает тёмный шум)







$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & [1] & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



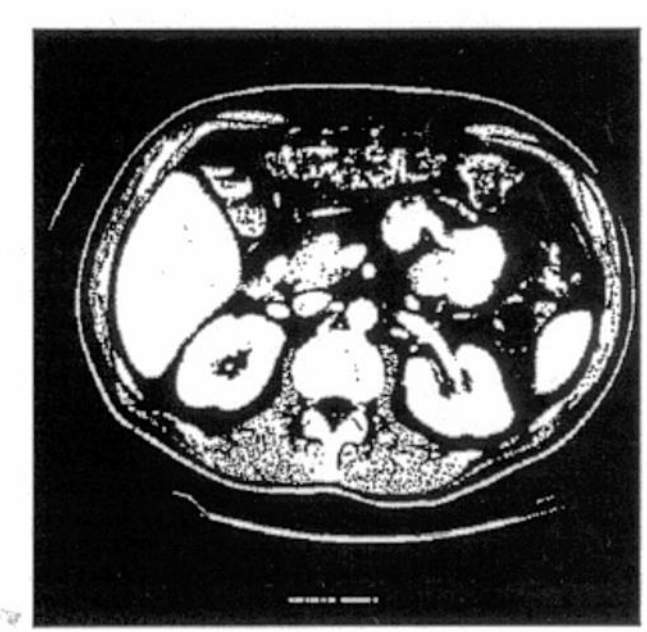
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & [1] & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

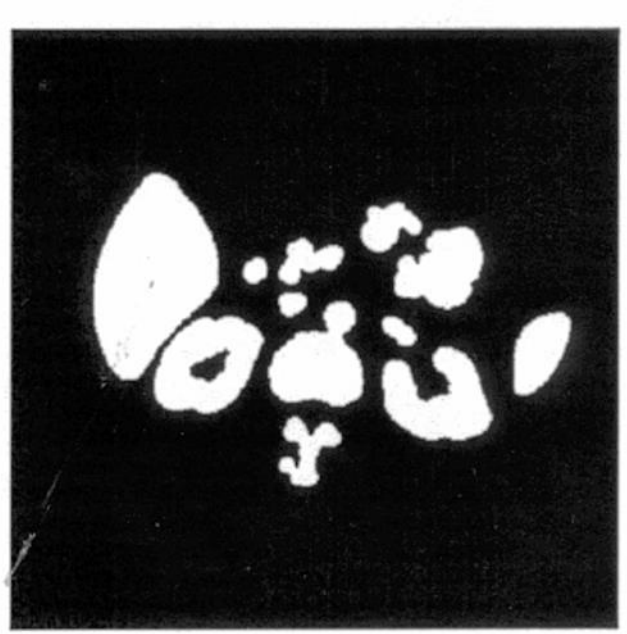


$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & [1] & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Задача 2: заполнение дырок

Постановка: объект с внутренними дырками (например, маска опухоли)
Решение: closing + flood fill от границы



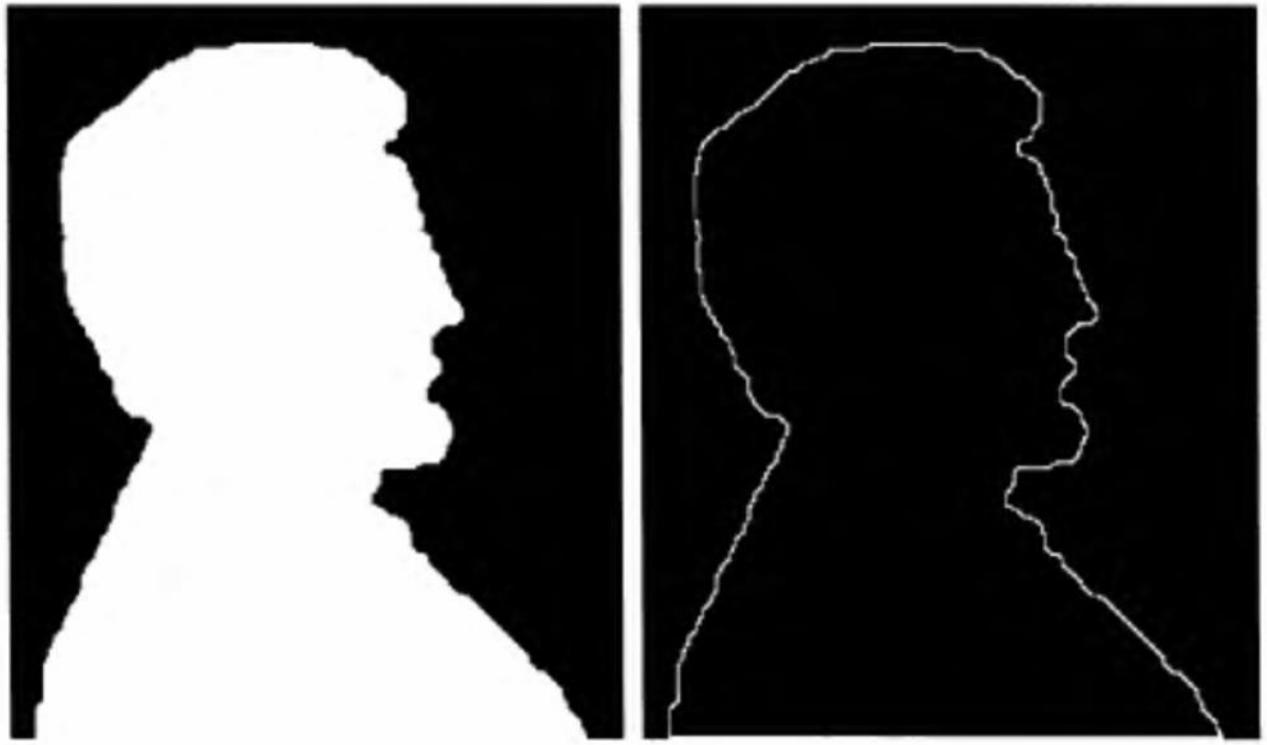


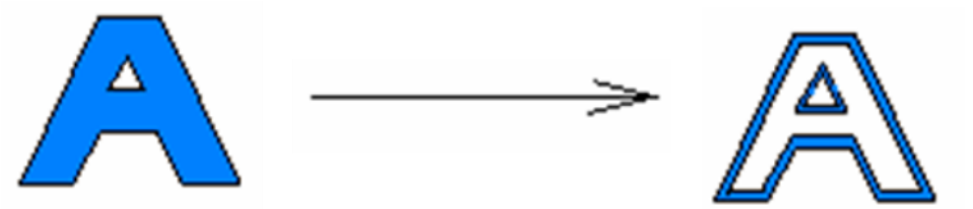
Обсуждение: closing не всегда достаточен, flood fill надёжнее

Задача 3: выделение границ

Постановка: нужен контур объекта

Решение: морфологический градиент $\beta(A) = A \oplus B - A \ominus B$



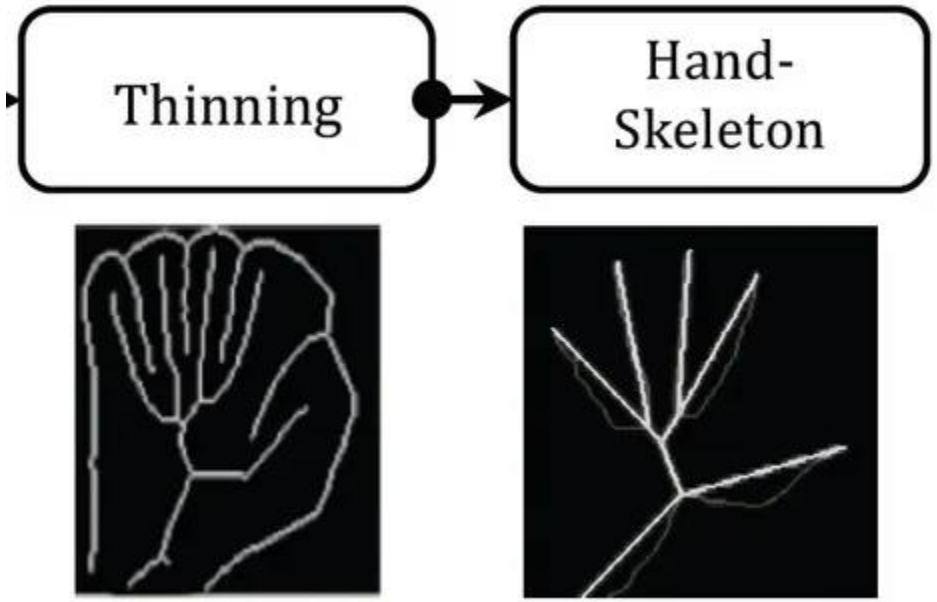


Морфологический градиент vs Canny vs Sobel

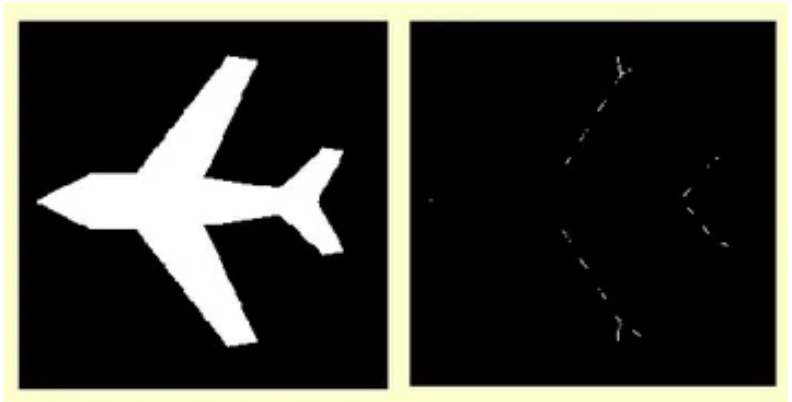
Тезис: для бинарных масок морфологический градиент проще и надёжнее

Задача 4: скелетизация

Постановка: нужна центральная линия объекта
Решение: thinning / skeleton
Пример: сосуды, дороги, рукописный текст
Обсуждение: чувствительность к шуму, предварительное

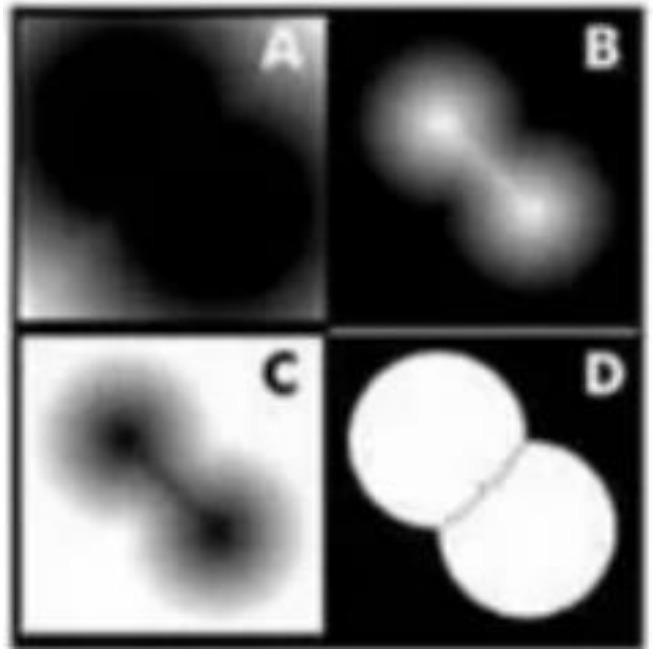


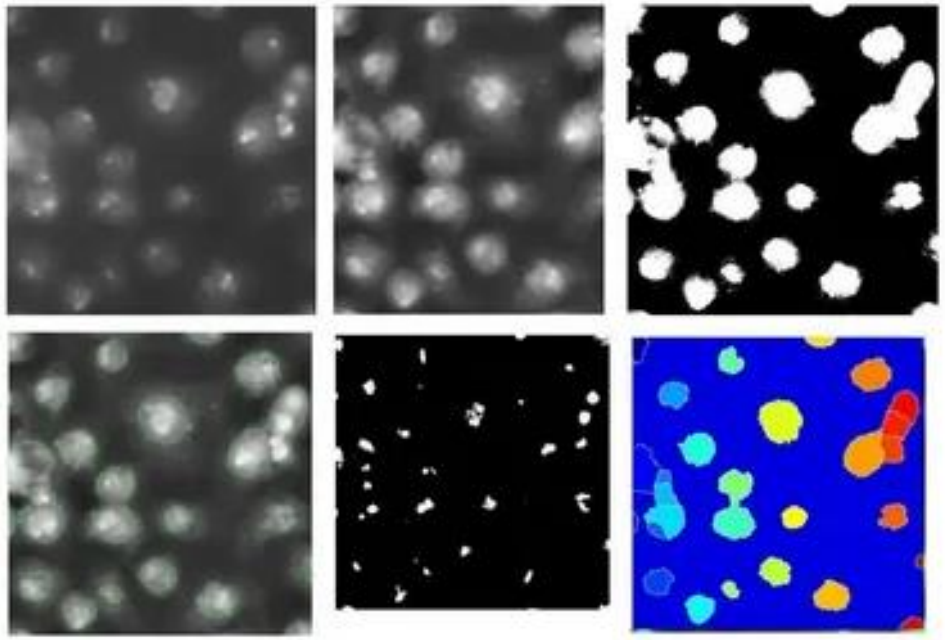




Задача 5: watershed (водораздела)

Постановка: разделить слипшиеся объекты (клетки, монеты)
Решение: distance transform → markers → watershed
Обсуждение: пере-сегментация, выбор маркеров





Итог: сырая маска → пригодная для измерения и распознавания

План лекции

- Бинарные изображения и связность
- Математическая морфология: базис
- Составные операции
- Пять прикладных задач
- **Мостик к CNN**

Морфология и CNN. Pooling как частный случай

Pooling как частный случай: max-pooling ≈ dilation + subsampling; min-pooling ≈ erosion

Морфологические операторы — это частные случаи или предельные формы того, что делают CNN, а pooling — прямое воплощение дилатации и эрозии

Max-pooling — это дилатация, за которой идёт прореживание. Морфологически это расширение ярких областей + сжатие разрешения

Min-pooling — это эрозия + субдискретизация. Сужение ярких областей + сжатие разрешения

Среднее — **линейный** оператор, а морфология — **нелинейная** (решёточная). Average-pooling ближе к свёртке/размытию, а не к морфологии

Морфология и CNN. Receptive field и структурный элемент

Receptive field и структурный элемент — **концептуальная связь**

Receptive field (RF) нейрона — область входа, влияющая на его отклик
Структурирующий элемент (SE) — область, по которой берётся max/min/свёртка в морфологии

Это одно и то же понятие:
RF описывает **геометрию локальной окрестности**, в которой собирается информация
SE описывает **форму этой окрестности** (диск, квадрат, линия, крест)

Различие:
CNN: RF растёт с глубиной (композиция свёрток и pooling)
Морфология: SE задаётся вручную, но композиция операторов даёт эффективный SE

Ключевая идея: морфология даёт геометрическую интерпретацию RF: каждый нейрон видит окрестность определённой формы, и эта форма — эффективный структурирующий элемент композиции слоёв

Историческая перспектива

1960-е: Matheron, Serra — математическая морфология
1980-е: морфологические фильтры в обработке изображений
1990-е: морфологические нейронные сети (Ritter, Davidson)
2010-е: CNN доминируют, но внутри — те же max/min-операции
2020-е: возрождение интереса — MorphNet, max-plus networks, boundary loss, SDF, differentiable morphology

Морфология была переоткрыта под новыми именами

Morphological layers

Работы: Shih et al. (2020), morphological pooling

Идея: заменить max-pooling на learnable morphological операцию

Применение: robustness к шуму, интерпретируемость

Морфология как препроцессинг

Очистка масок перед U-Net / Mask R-CNN

Пример: Оцу + opening + closing → CNN (малярия, 97.96% — из лекции 3)

Тезис: классика + deep learning = лучший результат

Watershed и deep learning

Marker-controlled watershed в instance segmentation
(Управляемый маркерами водораздел в сегментации экземпляров)

Современные аналоги: Mask2Former, SAM

Mask2Former отвечает на вопрос «что это за объект и где его границы», а SAM — на вопрос «где границы объекта, который я тебе указал»

Distance transform в loss-функциях (boundary loss, Hausdorff)

Distance Transform (DT) в функциях потерь используется для того, чтобы заставить нейросеть точнее очерчивать границы объектов

Обычные лоссы (Dice, Cross-Entropy) оценивают совпадение областей, но не штрафуют за смещение контура: предсказание может съехать на несколько пикселей и сохранить высокий Dice, хотя граница будет визуально неверной

Суть: DT здесь — это карта штрафов, которая превращает геометрическое отклонение контура в дифференцируемый сигнал для обучения. Без DT лоссы не видят границ, с DT — начинают их оптимизировать

Скелет в современных задачах

Медицина: анализ сосудов, нейронов
OCR: распознавание рукописей
Робототехника: планирование пути

Сравнительная таблица

Критерий	Классическая морфология	Deep learning
Скорость	O(N), реальное время	Требует GPU
Данные	Не нужны	Тысячи примеров
Интерпретируемость	Полная	Чёрный ящик
Когда применять	Простые маски, реальное время	Сложные сцены

Резюме

- Бинарные изображения: связность, компоненты, метрики
- Морфология: dilation, erosion, opening, closing, gradient, top-hat, skeleton, watershed
- Пять прикладных задач
- Мостик к CNN: pooling, препроцессинг, гибриды