

Курс “Анализ изображений”

Лекция#1. Выделение краев. Спецэффекты. Бинаризация. Математическая морфология. Текстура.

Большинство слайдов взято из лекций Конушина А.

ФИВТ МФТИ

2017



Выделение краёв



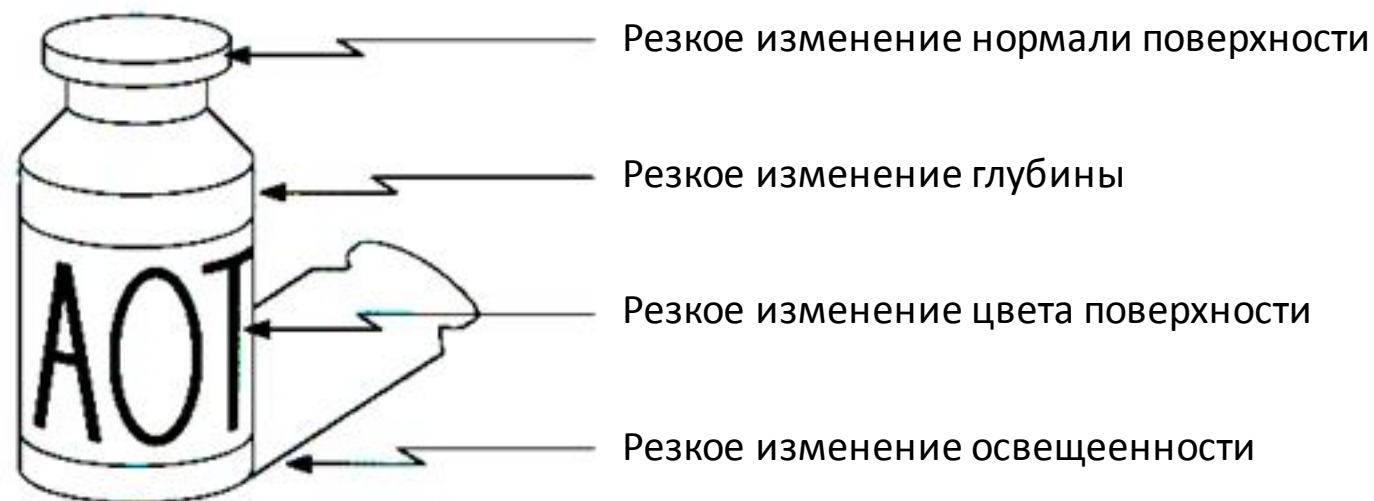
Выделение краев

- **Задача:** Выделить резкие изменения (разрывы) изображения
- Интуитивно понятно, что основная информация в картинке содержится как раз в краях (границах)
 - Компактное представление
 - Соответствует устройству мозга
- **Идеал:** рисунок художника (но артист уже пользуется своими знаниями об объектах)





Откуда берутся края



Резкое изменение = «разрыв»

Существует множество причин формирования краев на изображении



Описание «края»

Край – это точка резкого изменения значений функции интенсивности изображения



Края соответствуют
экстремумам производной



Градиент изображения

- Градиент изображения: $\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$

- $\nabla f = \left[\frac{\partial f}{\partial x}, 0 \right]$ $\nabla f = \left[0, \frac{\partial f}{\partial y} \right]$ $\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$

Градиент направлен в сторону наибольшего изменения интенсивности

Направления градиента задается как: $\theta = \tan^{-1} \left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x} \right)$

- Как направление градиента соответствует направлению края?
- Сила края* задается величиной (нормой) градиента:

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$



Дифференцирование и свёртка

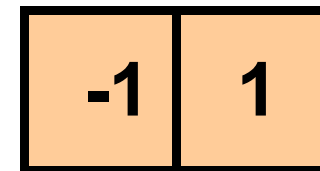
Для функции 2х
переменных, $f(x,y)$:

$$\frac{\partial f}{\partial x} = \lim_{\varepsilon \rightarrow 0} \left(\frac{f(x + \varepsilon, y) - f(x, y)}{\varepsilon} \right)$$

Разностная производная:

$$\frac{\partial f}{\partial x} \approx \frac{f(x_{n+1}, y) - f(x_n, y)}{\Delta x}$$

- Разностная производная - линейная и инвариантная к переносу
- Можно записать как свёртку



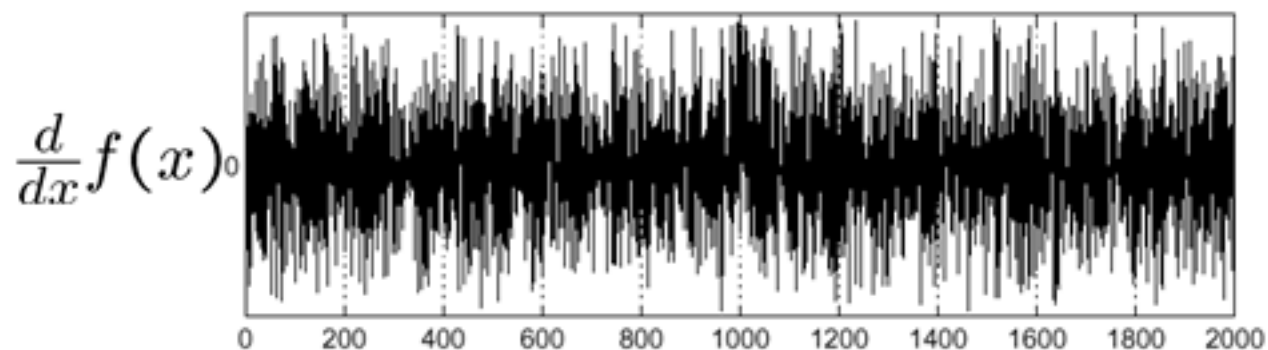
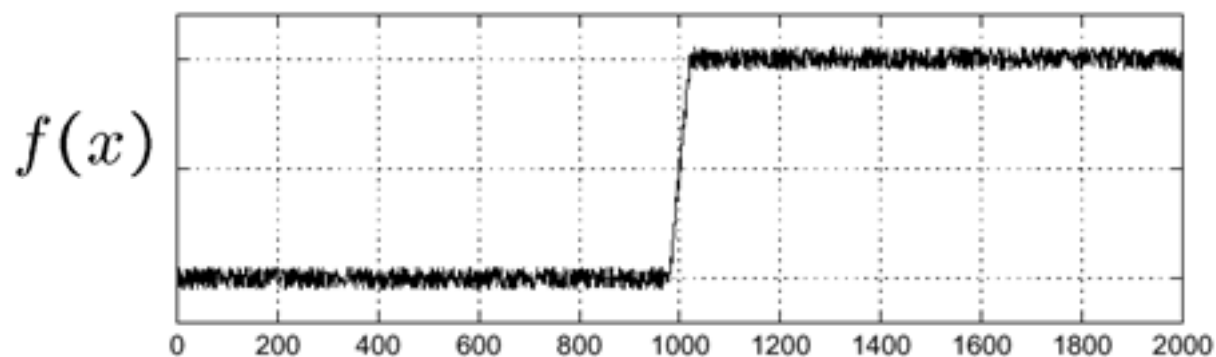
Простейший фильтр



Влияние шума

Рассмотрим строку или столбец изображения

- Интенсивность от положения можно рассматривать как сигнал



Край исчез

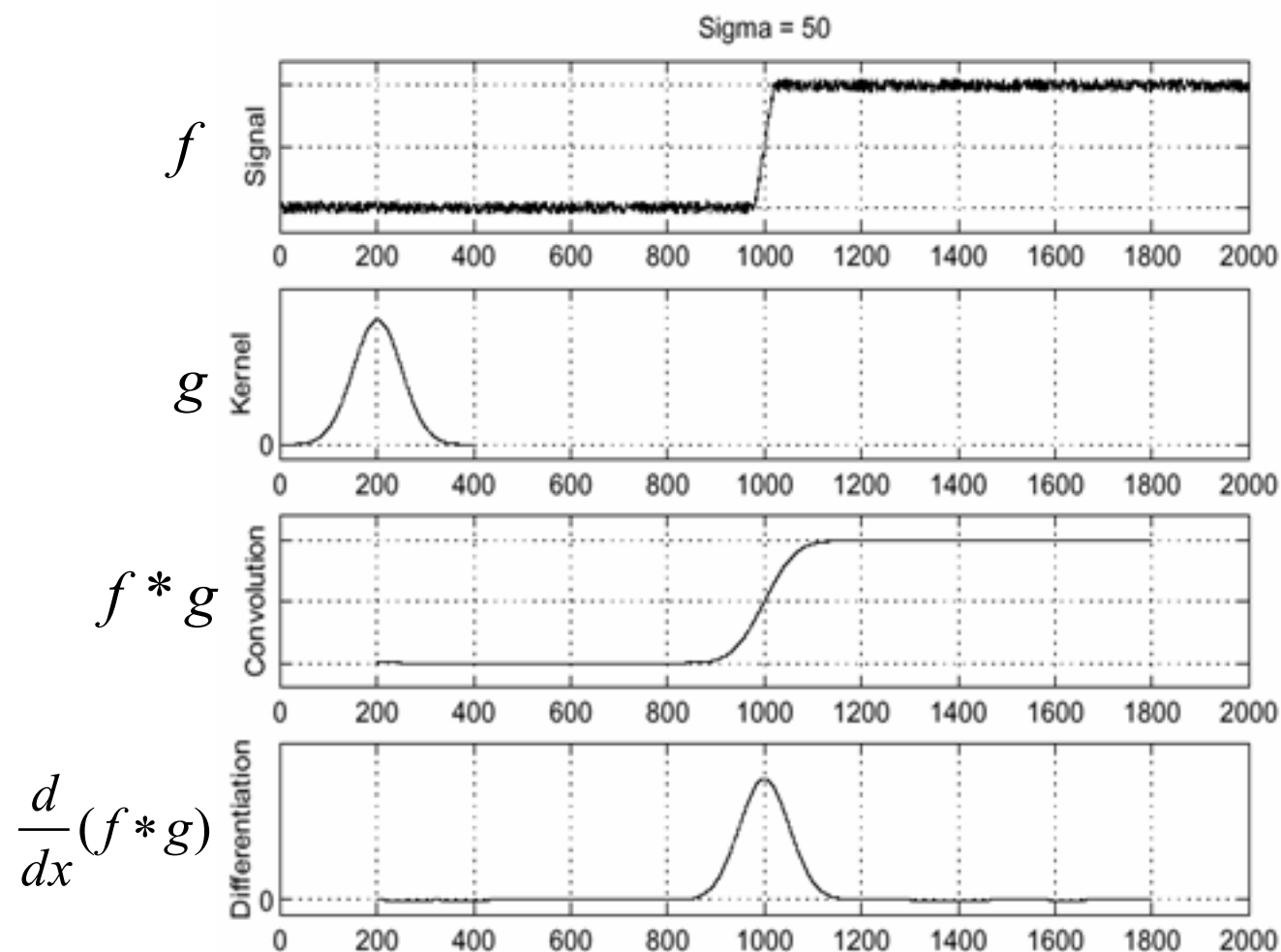


Влияние шума

- Разностные производные очень чувствительны к шуму
 - Зашумленные пиксели отличаются от соседей
 - Чем сильнее шум, тем выше отклик
- Сглаживание
 - Сглаживание делает все пиксели (зашумленные?) чуть более похожими на соседей



Предобработка (сглаживание)

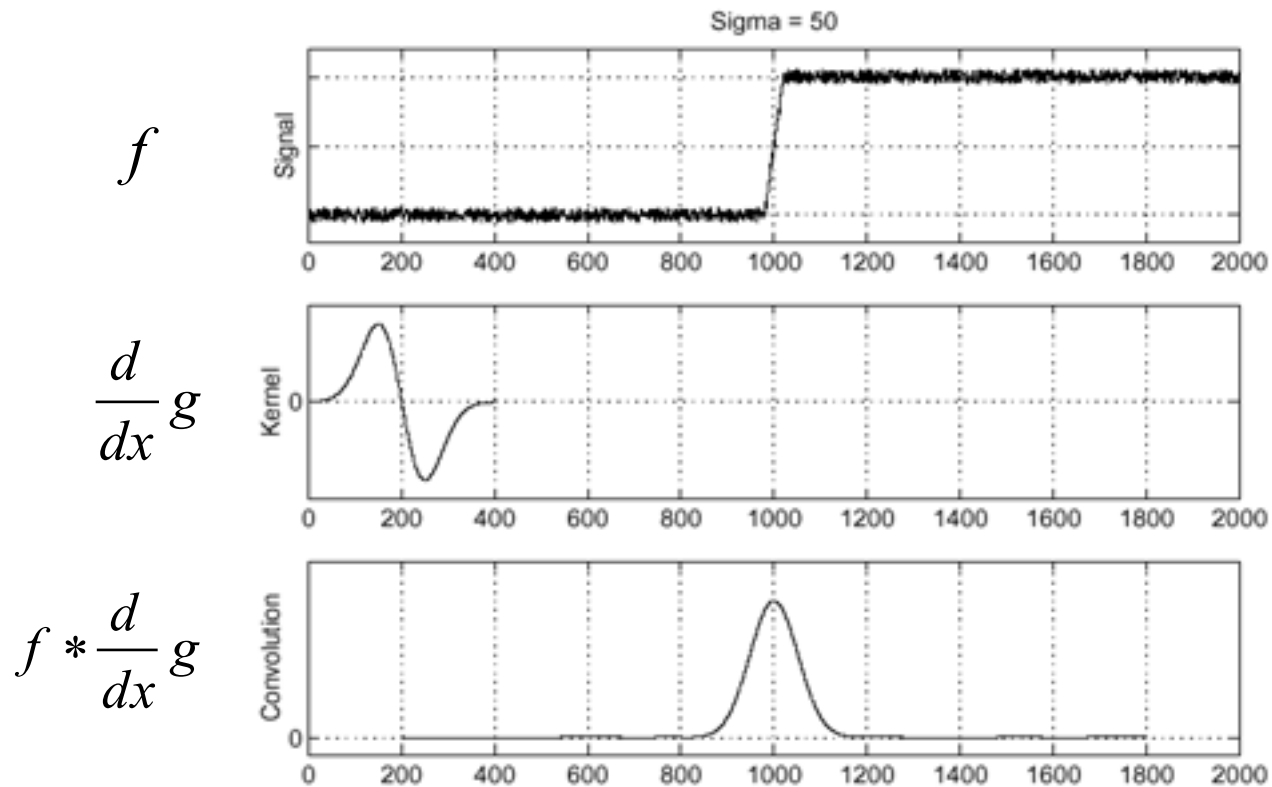


- Для поиска краев ищем пики в: $\frac{d}{dx}(f * g)$



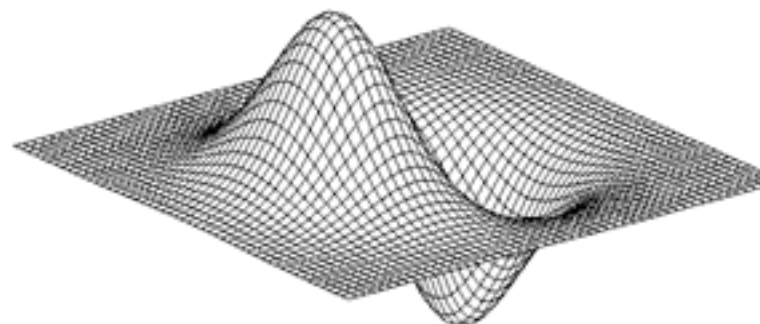
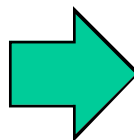
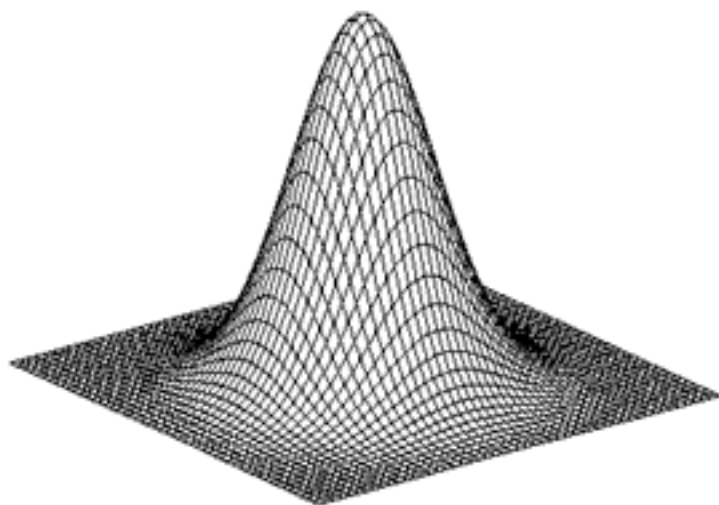
Свойства свертки

- Операции свертки и дифференцирования ассоциативны:
$$\frac{d}{dx}(f * g) = f * \frac{d}{dx}g$$
- Это экономит 1 операцию:

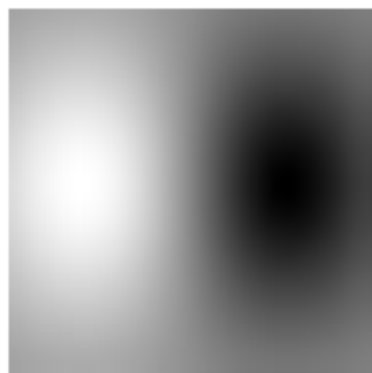
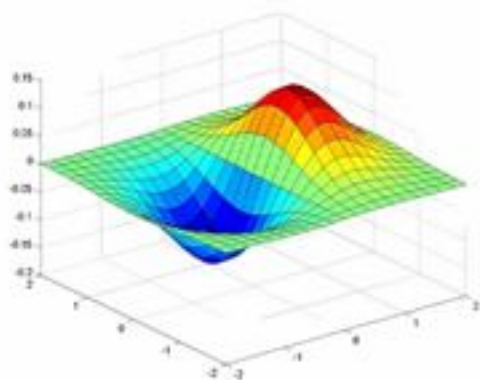




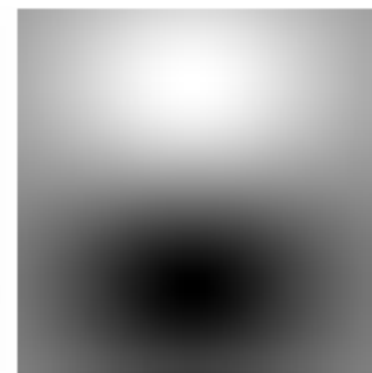
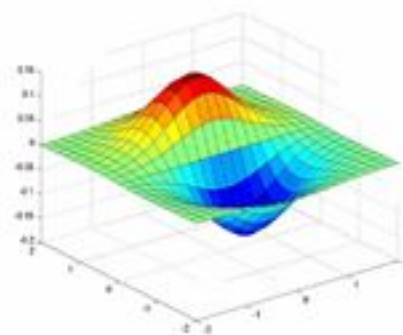
Производная фильтра Гаусса



По y



По y:





Известные фильтры

Несколько фильтров, по разному оценивающие производные по направлению:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Робертса

Превитт

Собеля

Превитт и Собель чуть-чуть сглаживают шум



Карта силы краев

Примеры:



Робертса



Превитт

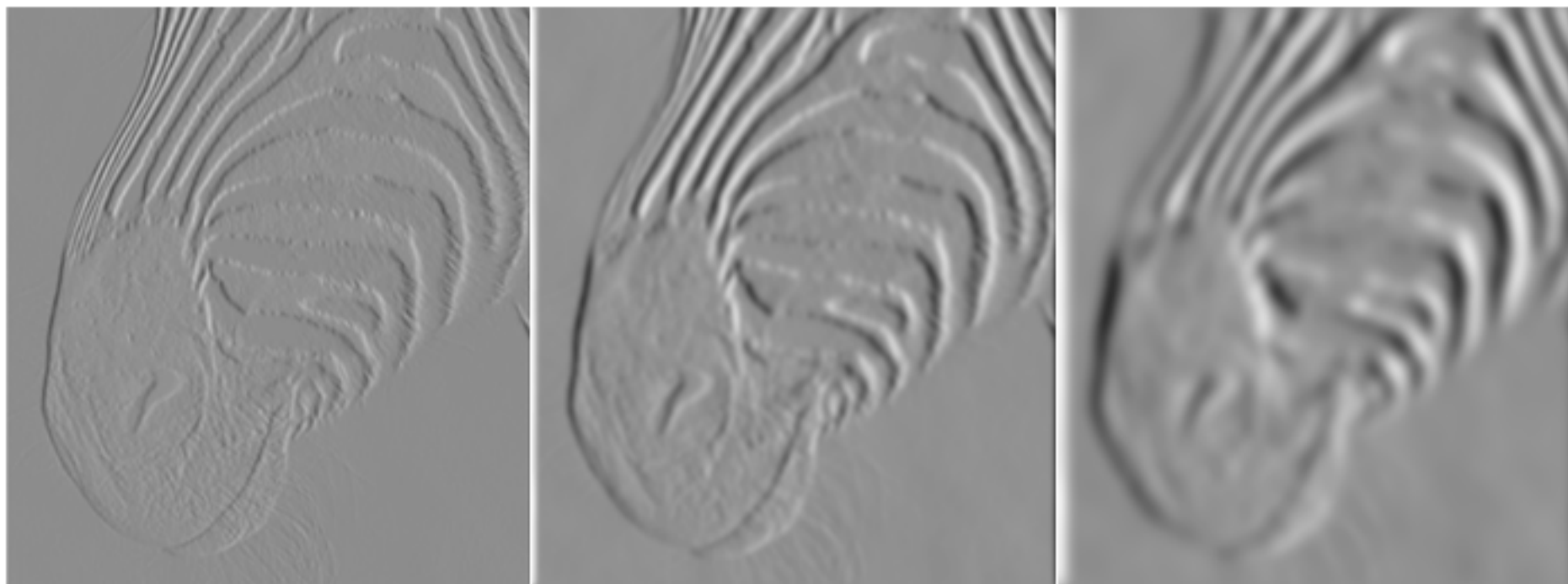


Собея



Сглаживание и локализация

Применим сглаженные производные разного размера:



1 pixel

3 pixels

7 pixels

Сглаженные производные подавляют шум, но размывают края. Плюс края находятся на разных «масштабах»



Выделение краев

- Вычисление градиента – не идеальный метод для поиска краёв.



Исходное изображение



Карта силы краев

- Чего не хватает?
 - Точности – края «толстые» и размытые
 - Информации о связности



Детектор Canny

1. Свертка изображения с ядром – производной от фильтра гаусса
 2. Поиск силы и направления градиента
 3. Выделение локальных максимумов (Non-maximum suppression)
 - Утоньшение полос в несколько пикселей до одного пикселя
 4. Связывание краев и обрезание по порогу (гистерезис)
 - Определяем два порога: нижний и верхний
 - Верхний порог используем для инициализации кривых
 - Нижний порог используем для продолжения кривых
-
- MATLAB: `edge(image, 'canny')`

J. Canny, [*A Computational Approach To Edge Detection*](#), IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986.

Source: D. Lowe, L. Fei-Fei

Пример



- Исходное изображение (Lena)



Пример



Норма градиента

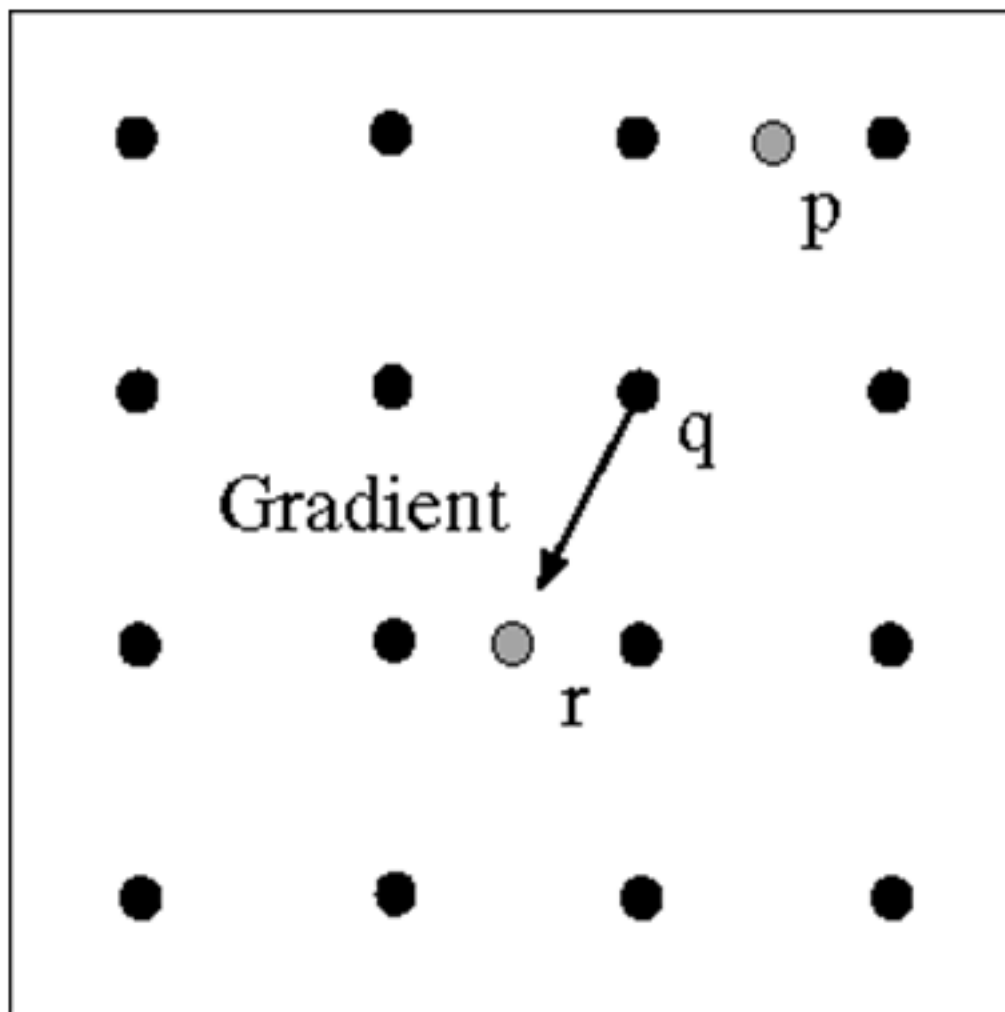
Пример



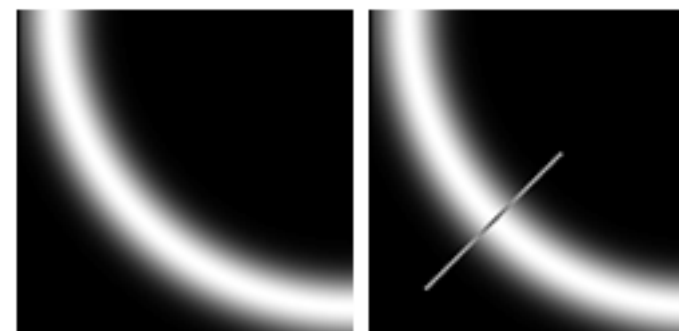
Отсечение по порогу



Поиск локальных максимумов

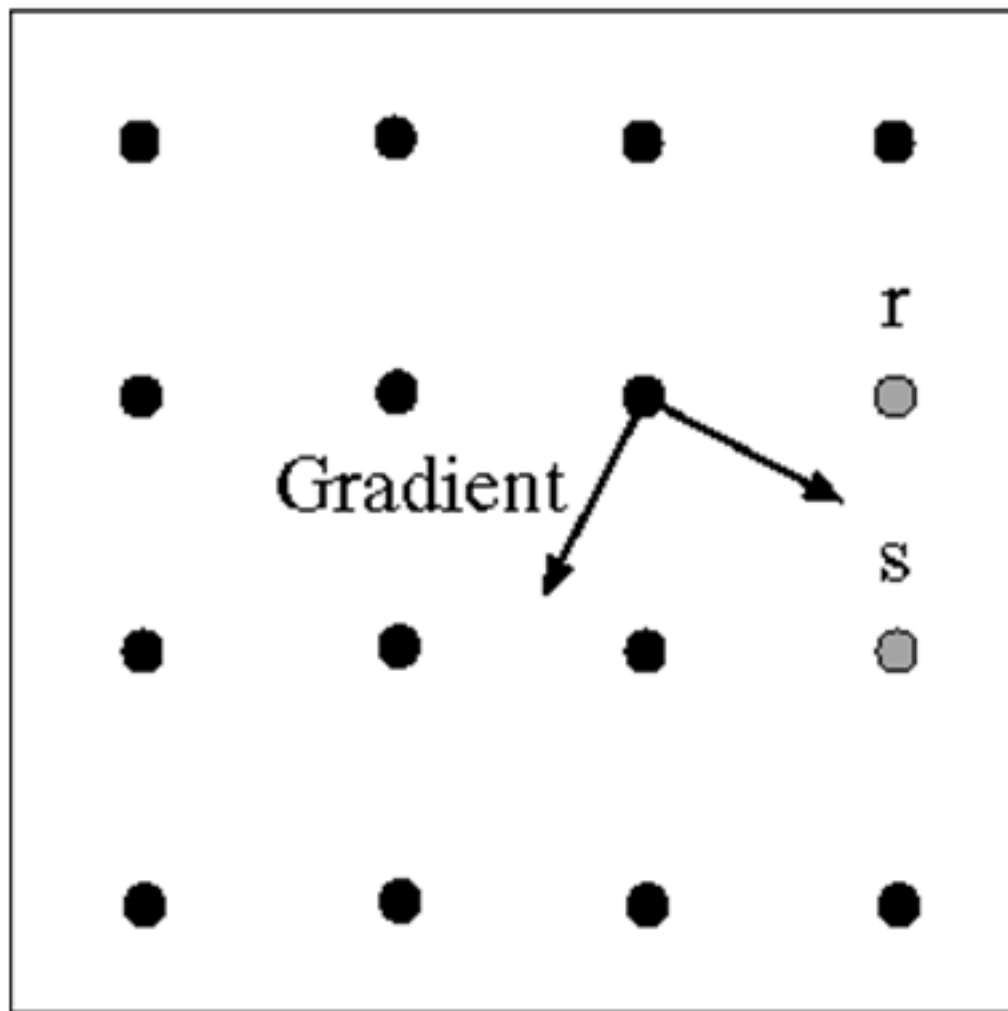


Максимум достигается в q , если значение больше p и r . Значения в p и r интерполируем.

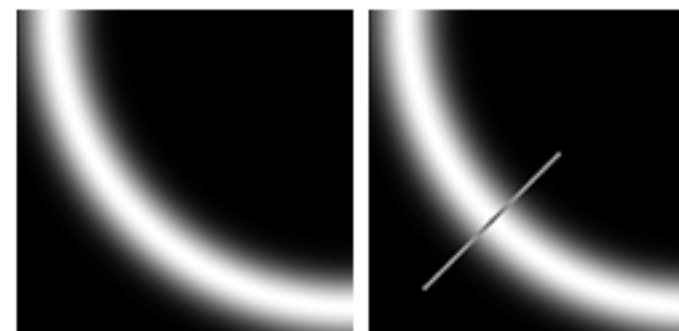




Связывание точек



Пусть отмеченная точка – край. Строим касательную к границе (нормаль к направлению градиента) и используем ее для предсказания новой точки (это либо s либо r).





Пример

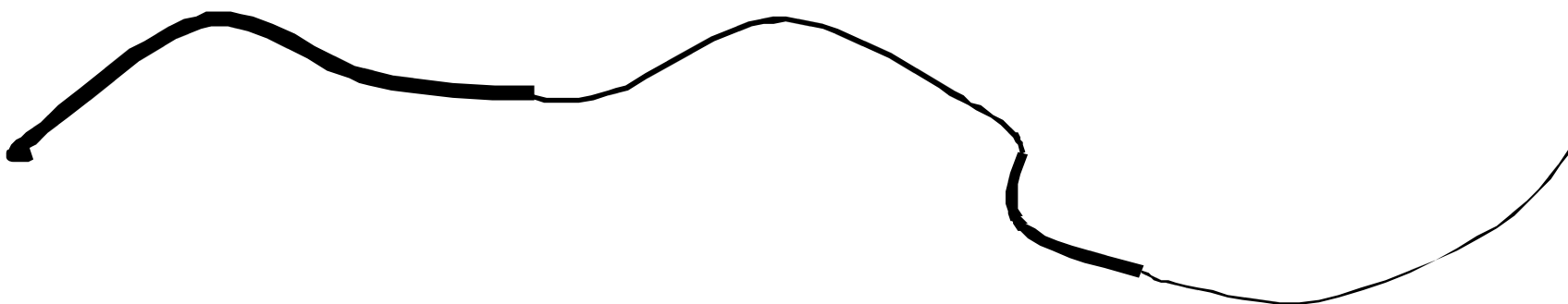


Утоньшение
(non-maximum suppression)



Отсечение по порогу

- Проверяем точку, чтобы значение градиента было выше порога
 - Используем **гистерезис**
 - Большой порог для начала построения кривой и низкий порог для продолжения края (связывания)





Эффект гистерезиса



Исходное изображение



Высокий порог
(сильные края)



Низкий порог
(слабые края)



Порог по гистерезису



Влияние σ (Размер ядра размытия)



original



Canny with $\sigma = 1$



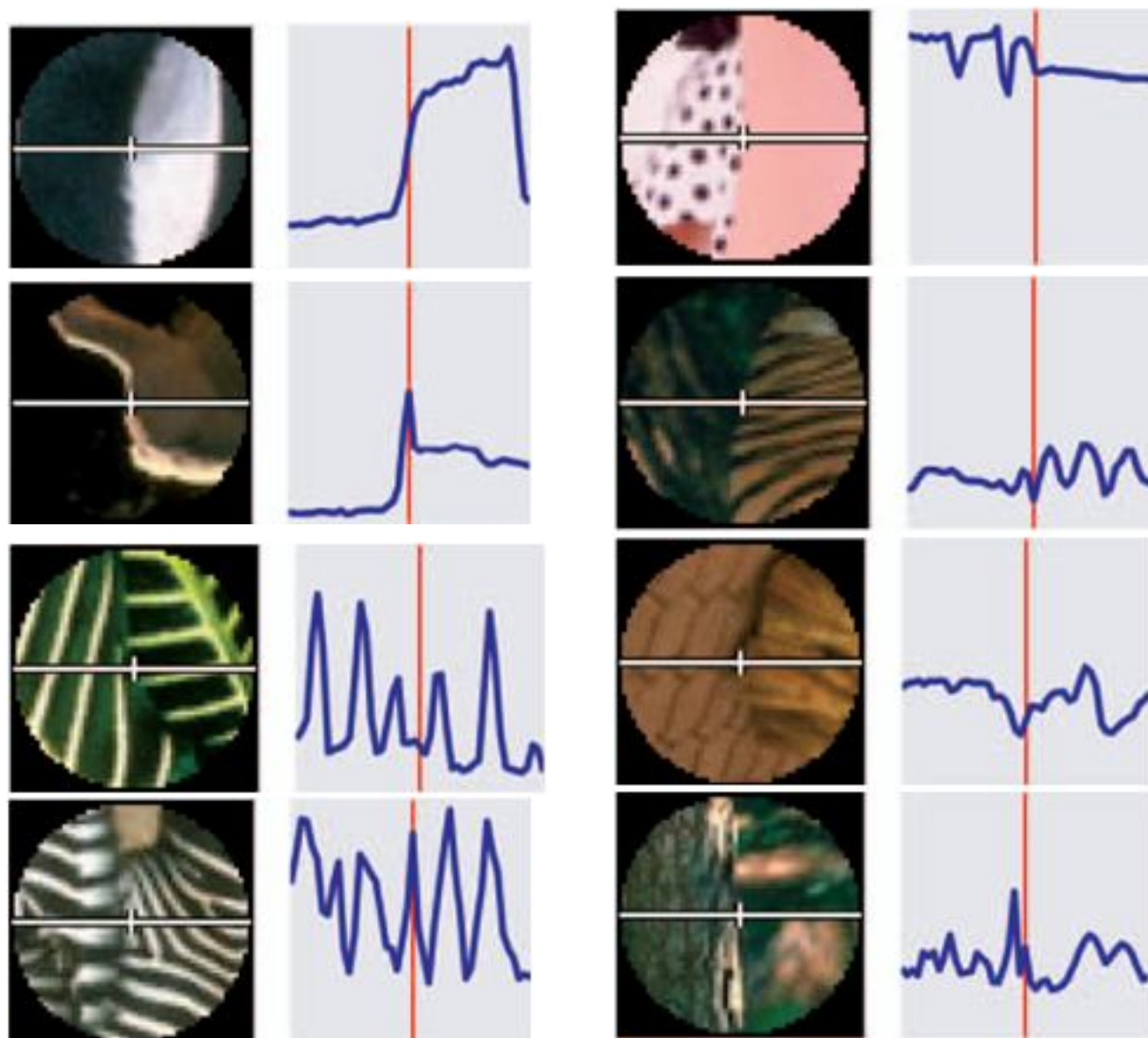
Canny with $\sigma = 2$

Выбор σ зависит от задачи

- большое σ - поиск крупных границ
- маленькое σ - выделение мелких деталей



Ограничения детектора Canny



Source: Martin et al. 2003

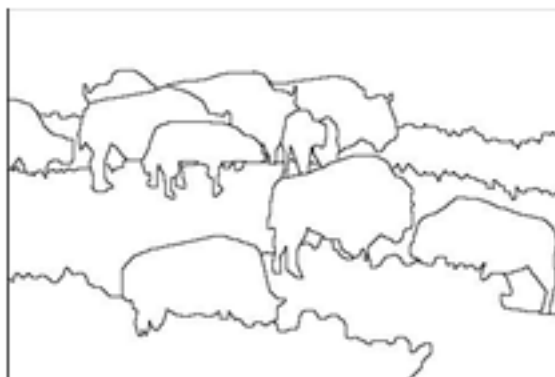


Поиск краев – это только начало...

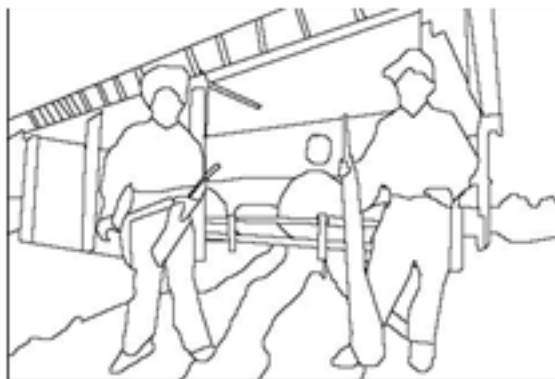
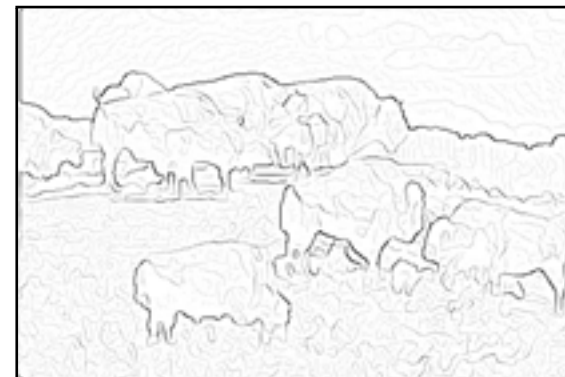
изображение



разметка вручную



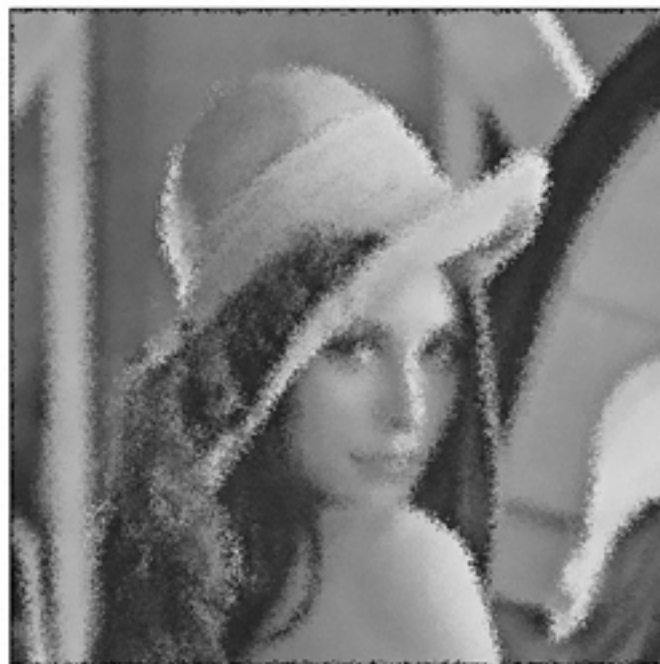
норма градиента



- Berkeley segmentation database:
<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>



«Эффект стекла»



$$x(k; l) = k + (\text{rand}(1, 1) - 0.5) * 10;$$
$$y(k; l) = l + (\text{rand}(1, 1) - 0.5) * 10;$$



Спецэффекты

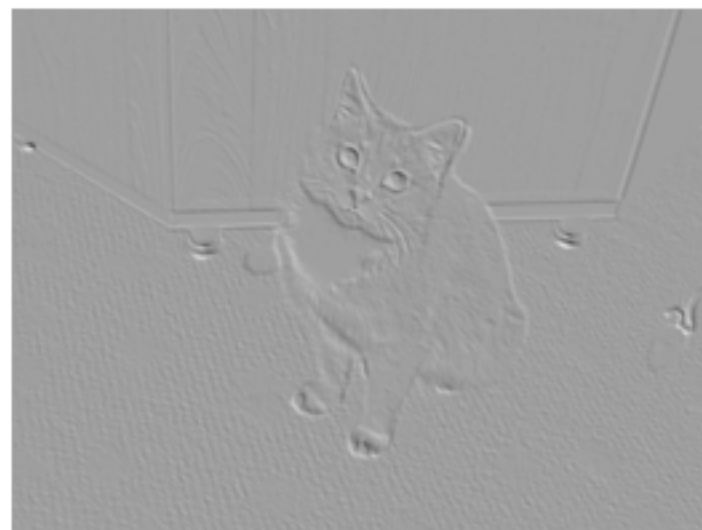
- Рассмотрим
 - Тиснение
 - Негатив
 - «Свелящиеся» края
 - Геометрические эффекты
 - Перенос/поворот
 - Искажение
 - «Эффект стекла»



Тиснение

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Фильтр + сдвиг яркости, нормировка...





Цифровой негатив



$$R' = 255 - R; \quad G' = 255 - G; \quad B' = 255 - B;$$



Светящиеся края



Медианный фильтр + выделение краев + фильтр «максимума»



«Волны»



Волны 1:

$$x(k; l) = k + 20\sin(2\pi l / 128); y(k; l) = l;$$

Волны 2:

$$x(k; l) = k + 20\sin(2\pi k / 30); y(k; l) = l;$$



«Эффект стекла»



$$x(k; l) = k + (\text{rand}(1, 1) - 0.5) * 10;$$
$$y(k; l) = l + (\text{rand}(1, 1) - 0.5) * 10;$$



Из чего состоит изображение?





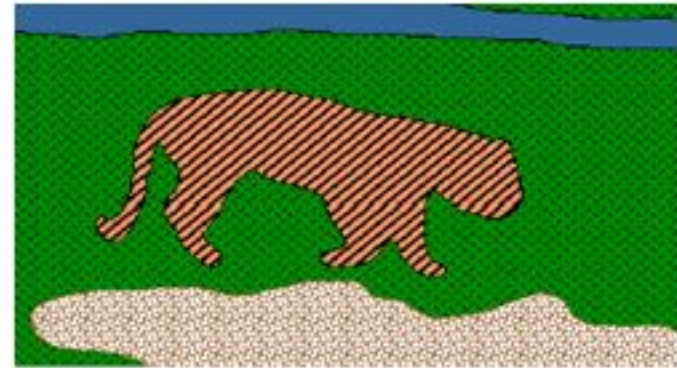
Из «кусков» - отдельных объектов





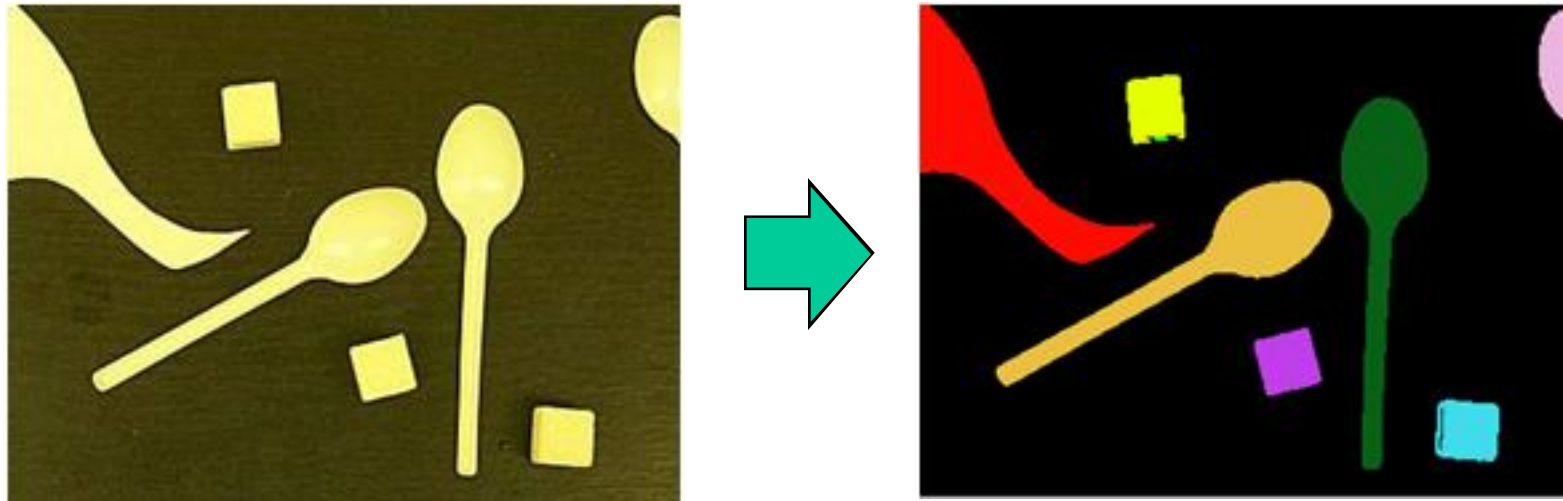
Сегментация

- Сегментация - это способ разделения сцены на «куски», с которыми проще работать
- (Тесселяция) Разбиение изображения на неперекрывающиеся области, покрывающие все изображение и однородные по некоторым признакам
- Можно и по другому сегментировать изображение
 - Пересекающиеся области
 - Иерархическое представление





Результат сегментации



- Как мы будем записывать результат сегментации?
- Сделаем карту разметки – изображение, в каждом пикселе которого номер сегмента, которому принадлежит этот пиксель
- Визуализировать удобно каждый сегмент своим цветом



Простейшая сегментация

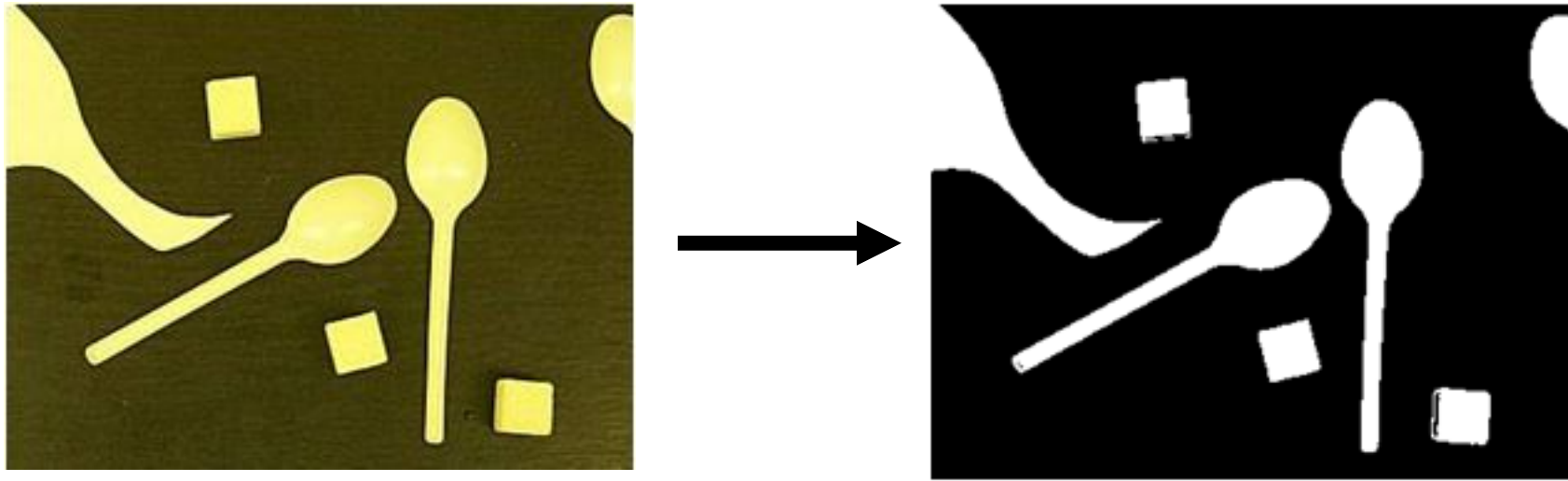
Чем отличаются объекты на этом изображении?



- Все объекты яркие, фон тёмный
- Для сегментации такого изображения нам достаточно:
 - пороговая бинаризация
 - обработки шума
 - выделения связанных компонент



Пороговая бинаризация



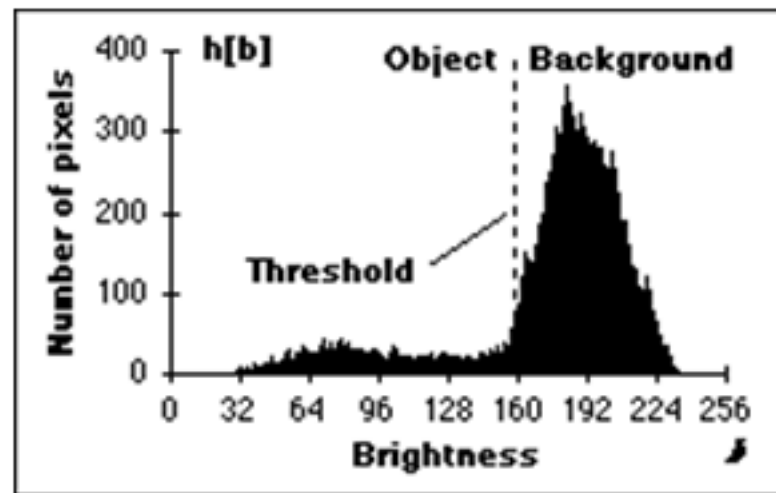
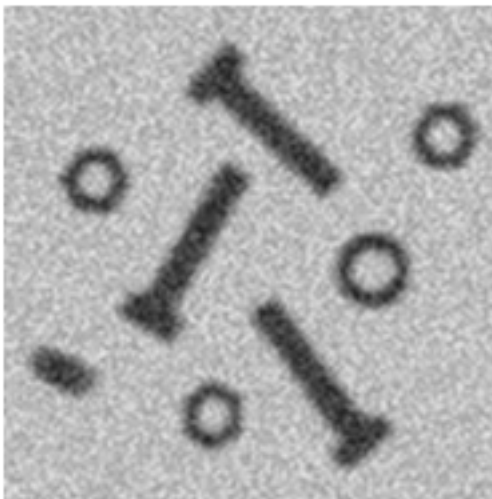
- Пороговая фильтрация (thresholding)
 - Пиксели, которых выше/ниже некоторого порога, заданного «извне», помечаются 1
 - Ниже порога помечаются 0
- Бинарное изображение – пиксели которого могут принимать только значения 0 и 1
- Бинаризация - построение бинарного изображения по полутоновому / цветному



Пороговая фильтрация

Более интересный способ – определение порога автоматически, по характеристикам изображения

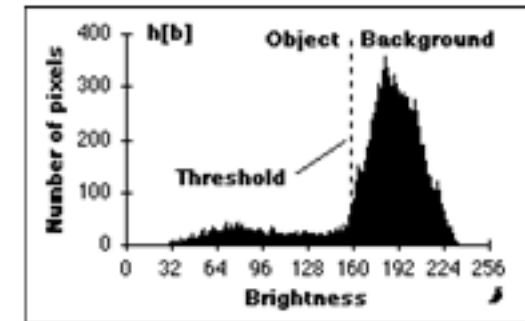
- Анализ гистограммы



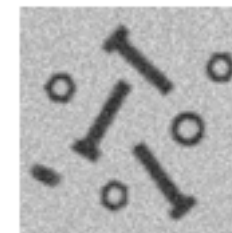


Анализ гистограммы

- Анализ симметричного пика гистограммы
- Применяется когда фон изображения дает отчетливый и доминирующий пик гистограммы, симметричный относительно своего центра.



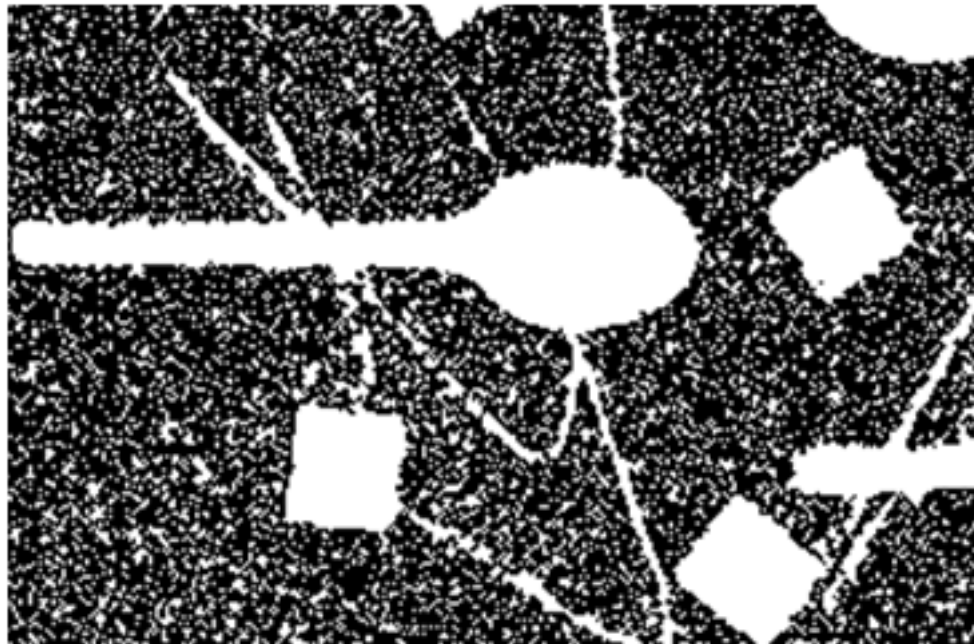
1. Сгладить гистограмму;
2. Найти ячейку гистограммы h_{\max} с максимальным значением;
3. На стороне гистограммы не относящейся к объекту (на примере – справа от пика фона) найти яркость h_p , количество пикселей с яркостью $\geq h_p$ равняется $p\%$ (например 5%) от пикселей яркости которых $\geq h_{\max}$;
4. Пересчитать порог $T = h_{\max} - (h_p - h_{\max})$;





Шум в бинарных изображениях

Пример бинарного изображения с сильным шумом

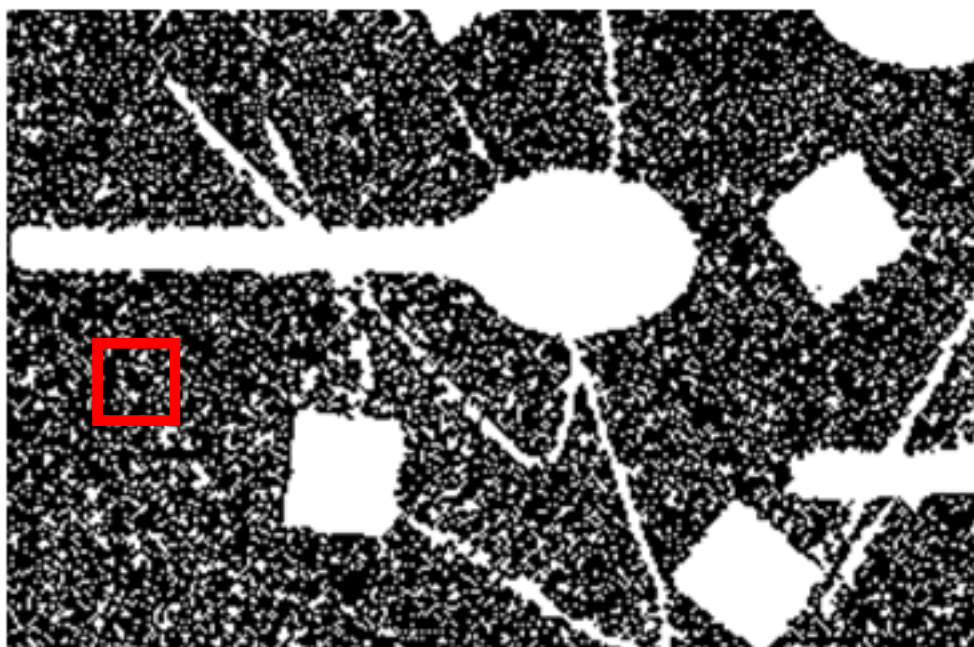


Часто возникает из-за невозможности полностью подавить шум в изображениях, недостаточной контрастности объектов и т.д.



Шум в бинарных изображениях

- ⑩ По одному пикселю невозможно определить – шум или объект?
- ⑩ Нужно рассматривать окрестность пикселя!





Подавление и устранение шума

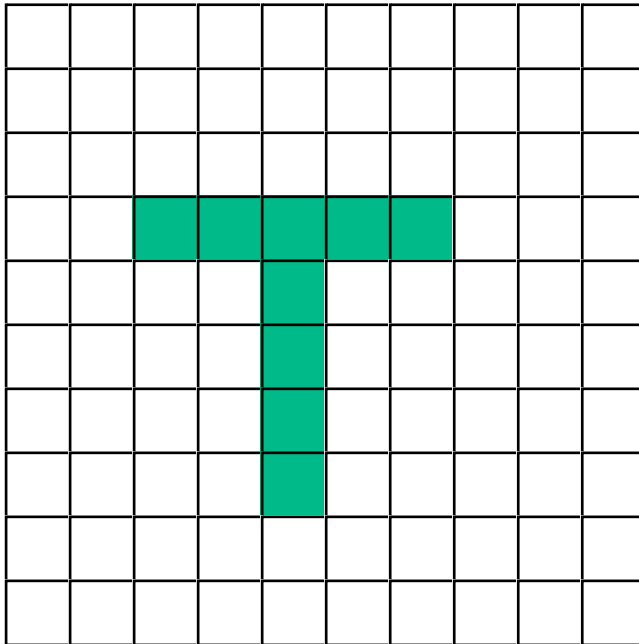
Широко известный способ - устранение шума с помощью операций математической морфологии:

- Сужение (erosion)
- Расширение (dilation)
- Закрытие (closing)
- Раскрытие (opening)

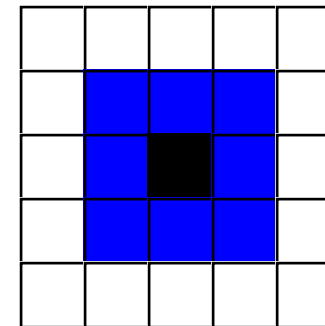


Математическая морфология

A



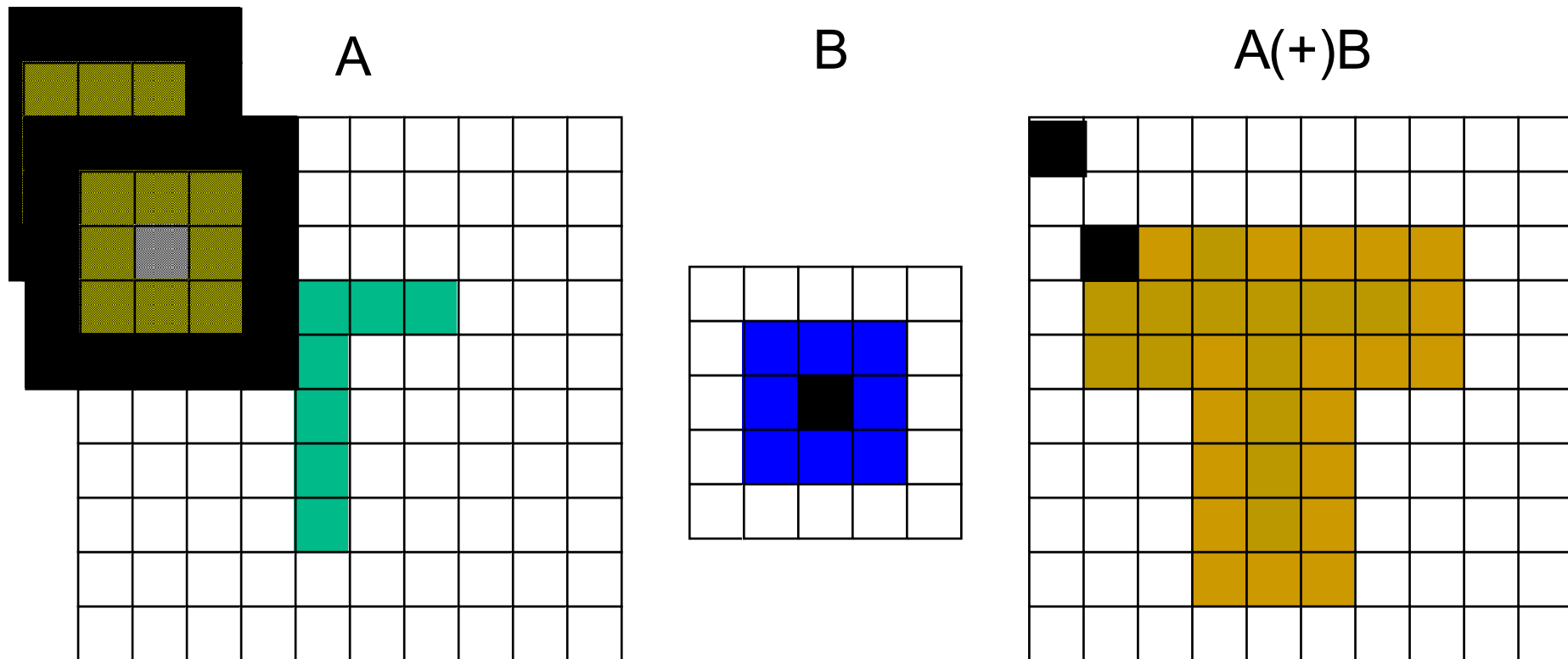
B



- Множество A обычно является объектом обработки
- Множество B (называемое структурным элементом) – инструмент обработки



Расширение в дискретном случае



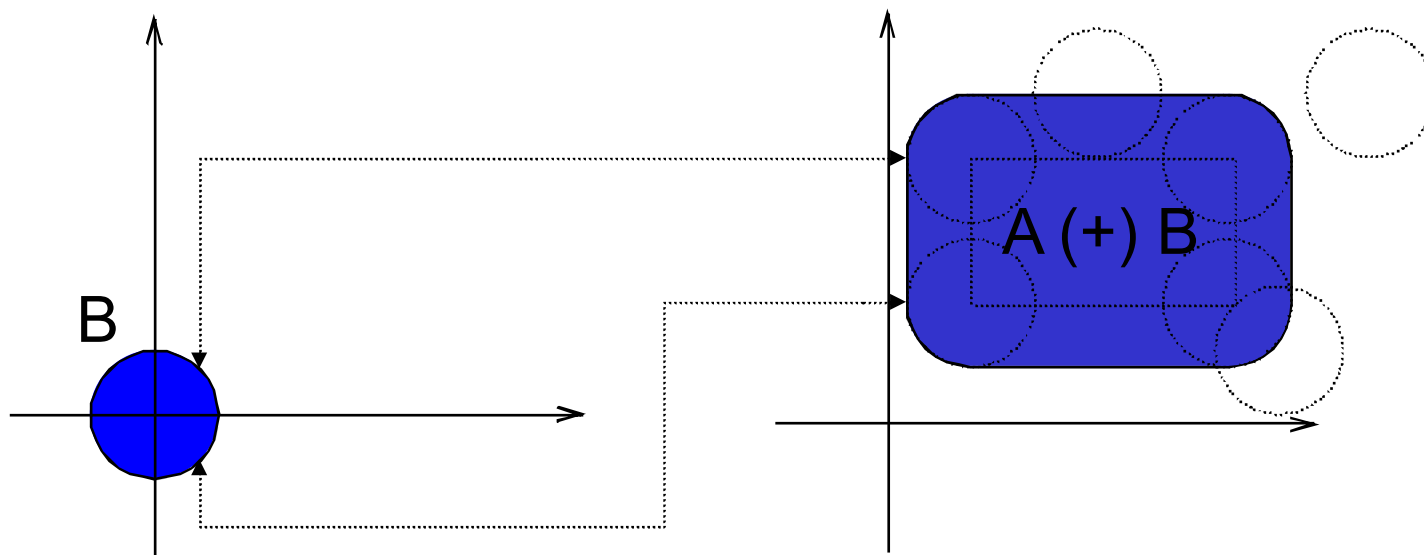
Операция «расширение» - аналог логического «или»



Расширение

Расширение (dilation)

$$A (+) B = \{t \in \mathbb{R}^2: t = a + b, a \in A, b \in B\}$$

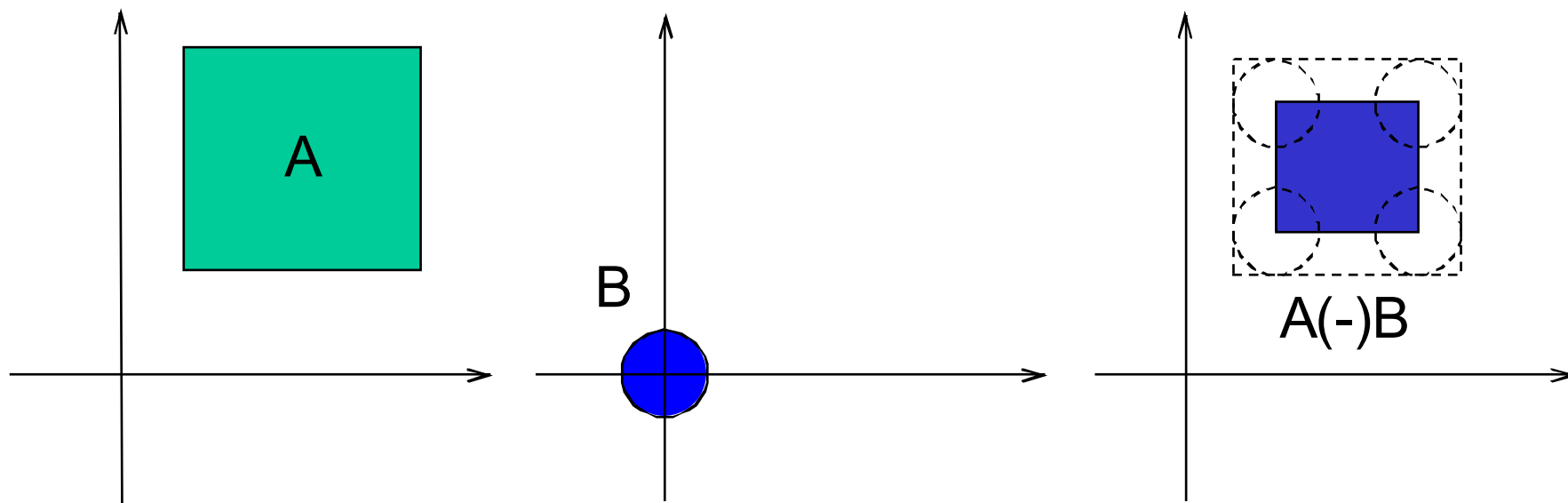




Сужение

Сужение (erosion)

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





Результат операции сужения



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



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



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



Важное замечание

Результат морфологических операций во многом определяется применяемым структурным элементом. Выбирая различный структурный элемент можно решать разные задачи обработки изображений:

- Шумоподавление
- Выделение границ объекта
- Выделение скелета объекта
- Выделение сломанных зубьев на изображении шестерни



Операции раскрытия и закрытия

Морфологическое раскрытие (opening)

- **$\text{open}(A, B) = (A (-) B) (+) B$**

Морфологическое закрытие (closing)

- **$\text{close}(A, B) = (A (+) B) (-) B$**

Попробуйте догадаться, что эти операции делают?



Применение открытия

Применим операцию открытия к изображению с сильным шумом:



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



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



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



Сужение vs Открытие



Сужение



Открытие



Дефекты бинаризации

Пример бинарного изображения с дефектами
распознаваемых объектов





Применение закрытия

Применим операцию закрытия к изображению с дефектами объектов:



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



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



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



Не лучший пример для морфологии

Не во всех случаях математическая морфология так легко убирает дефекты, как хотелось бы...

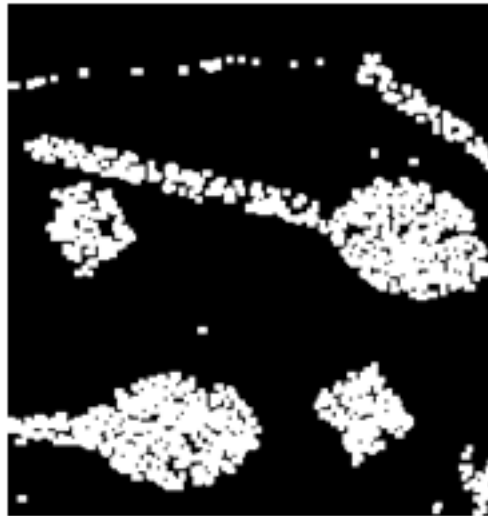




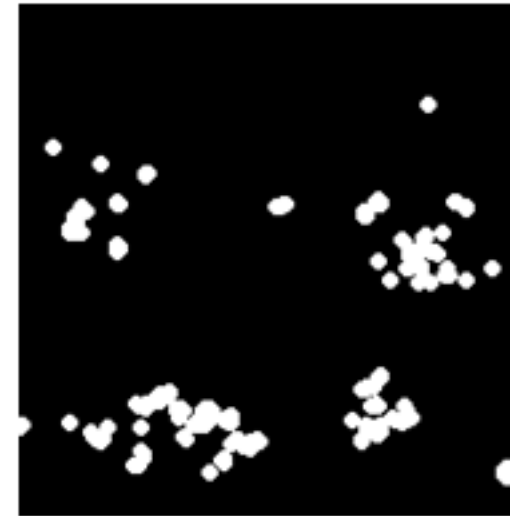
Применения операции открытия



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



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



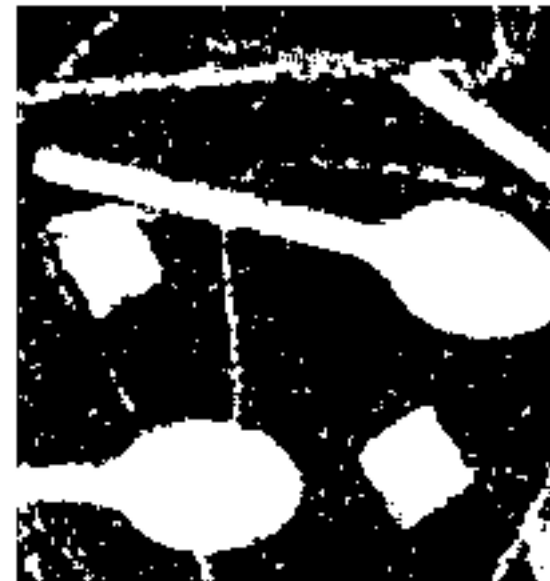
$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Часто помогает медианная фильтрация!



Медианный фильтр

Фильтр с окрестностью 3x3



Теперь можем с помощью морфологии убрать оставшиеся точки, тонкие линии и т.д.



Что дальше?



Получили бинарное
изображение



Нужна карта разметки



Выделение связных областей

- Определение связной области:
 - Множество пикселей, у каждого пикселя которого есть хотя бы один сосед, принадлежащий данному множеству.



Соседи пикселей:

	1	
2	*	3
	4	

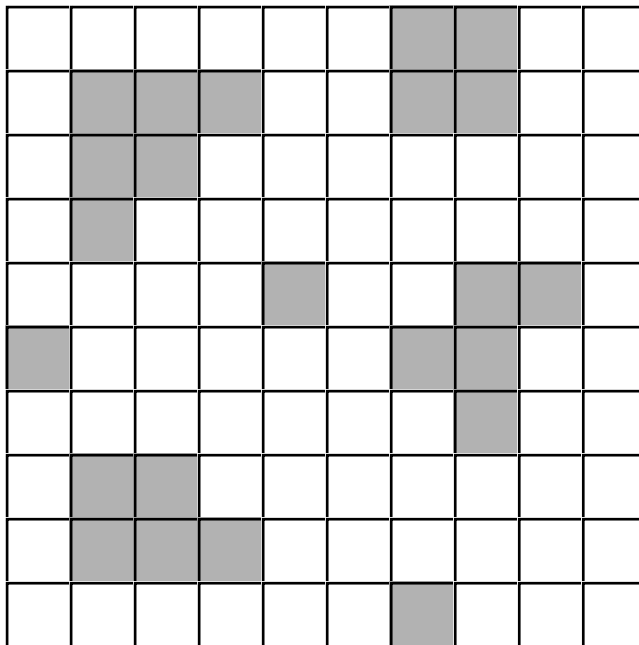
4-СВЯЗНОСТЬ

1	2	3
4	*	5
6	7	8

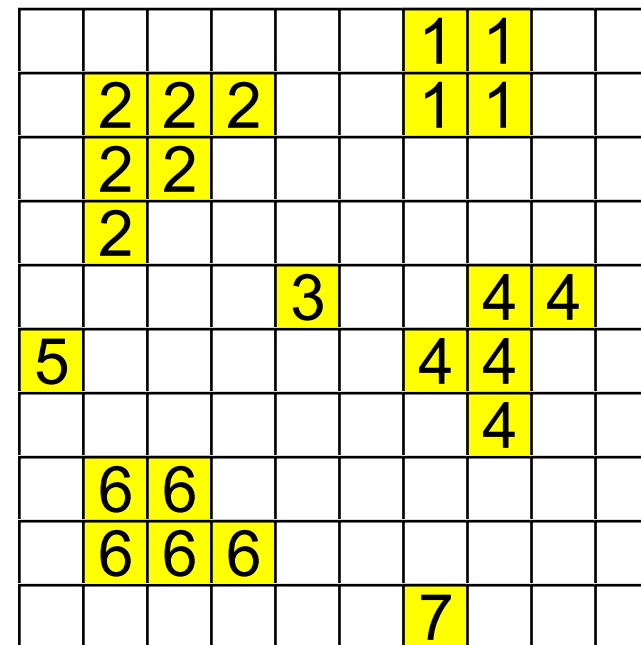
8-СВЯЗНОСТЬ



Разметка связанных областей



Бинарное изображение



Размеченное изображение



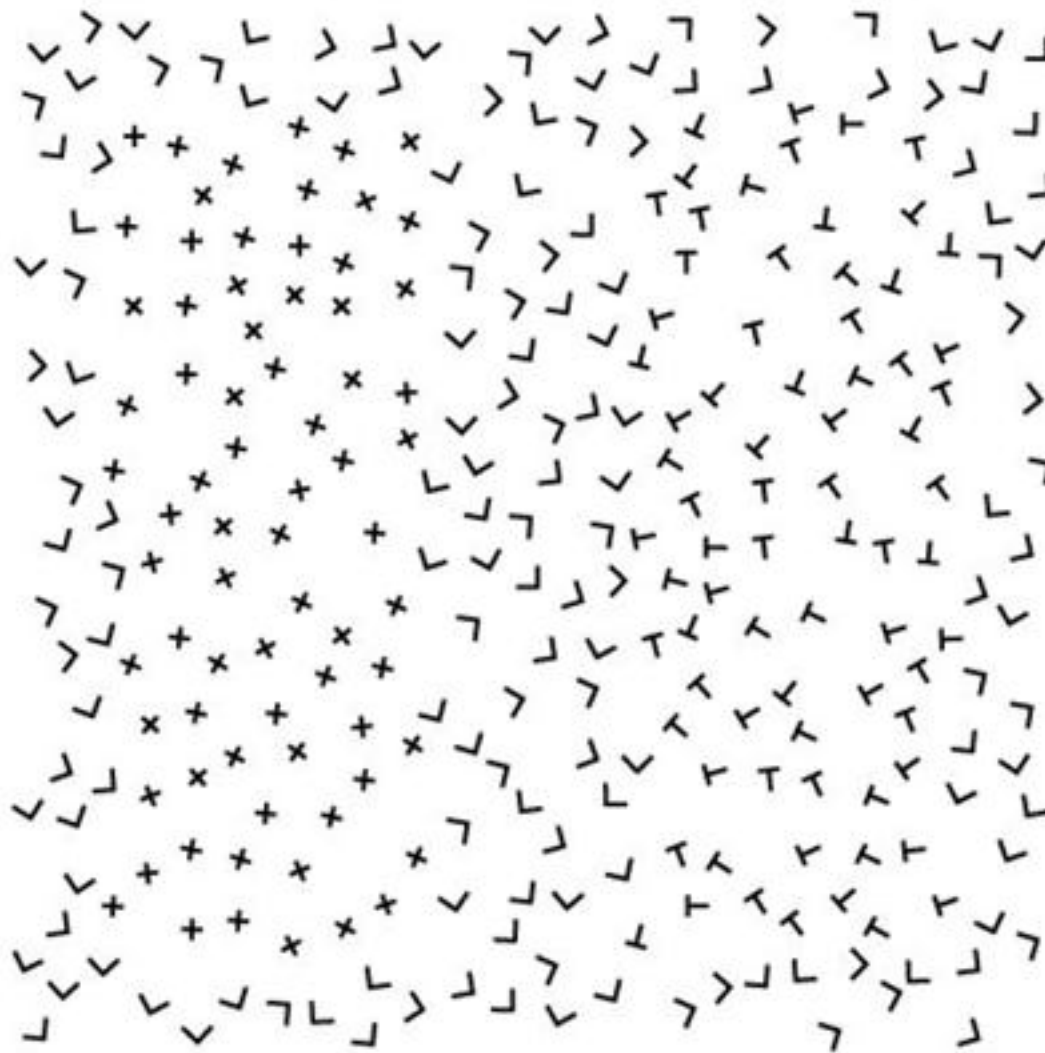
Признаки изображения

Какие признаки мы можем использовать для сравнения пикселей и регионов?

- Яркость
- Цвет
- ?



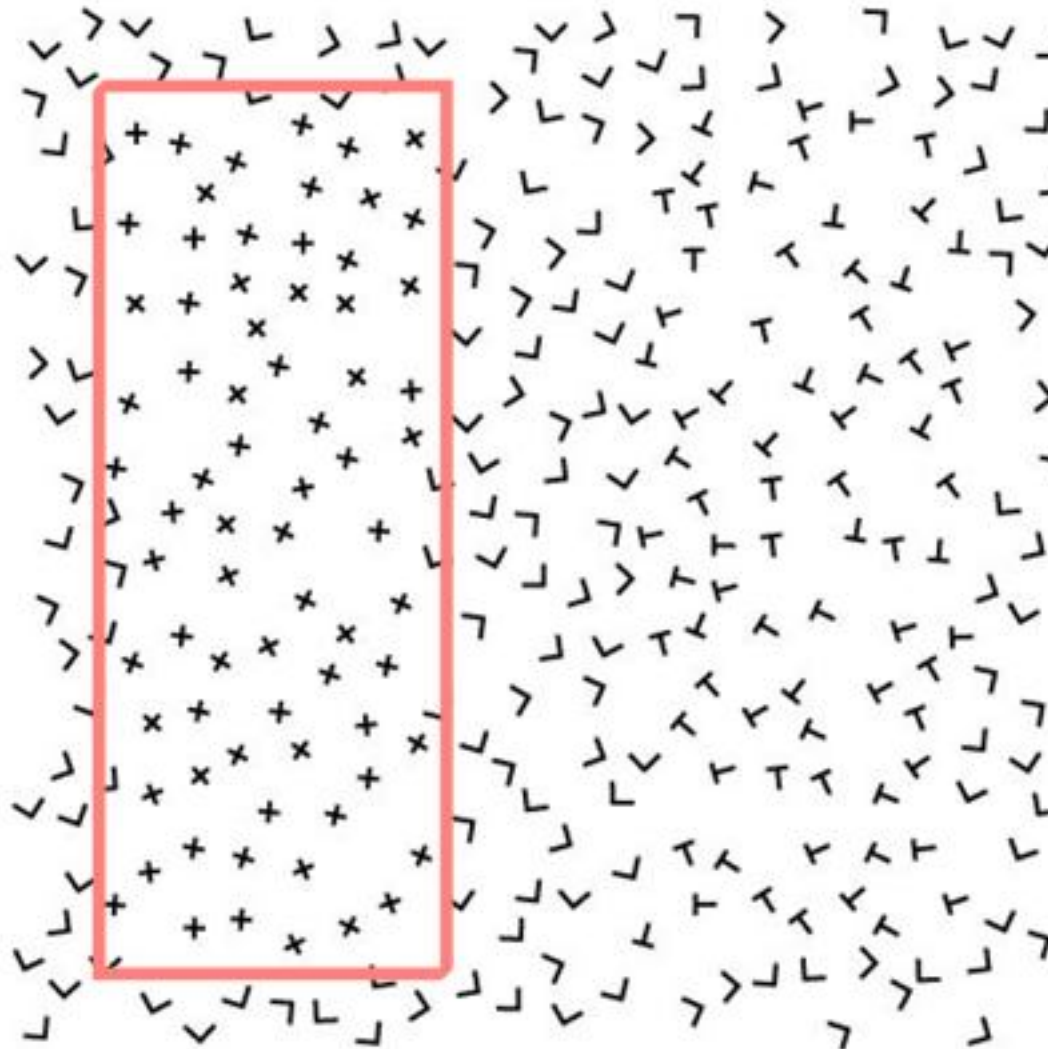
Пример



Видите отдельные области?

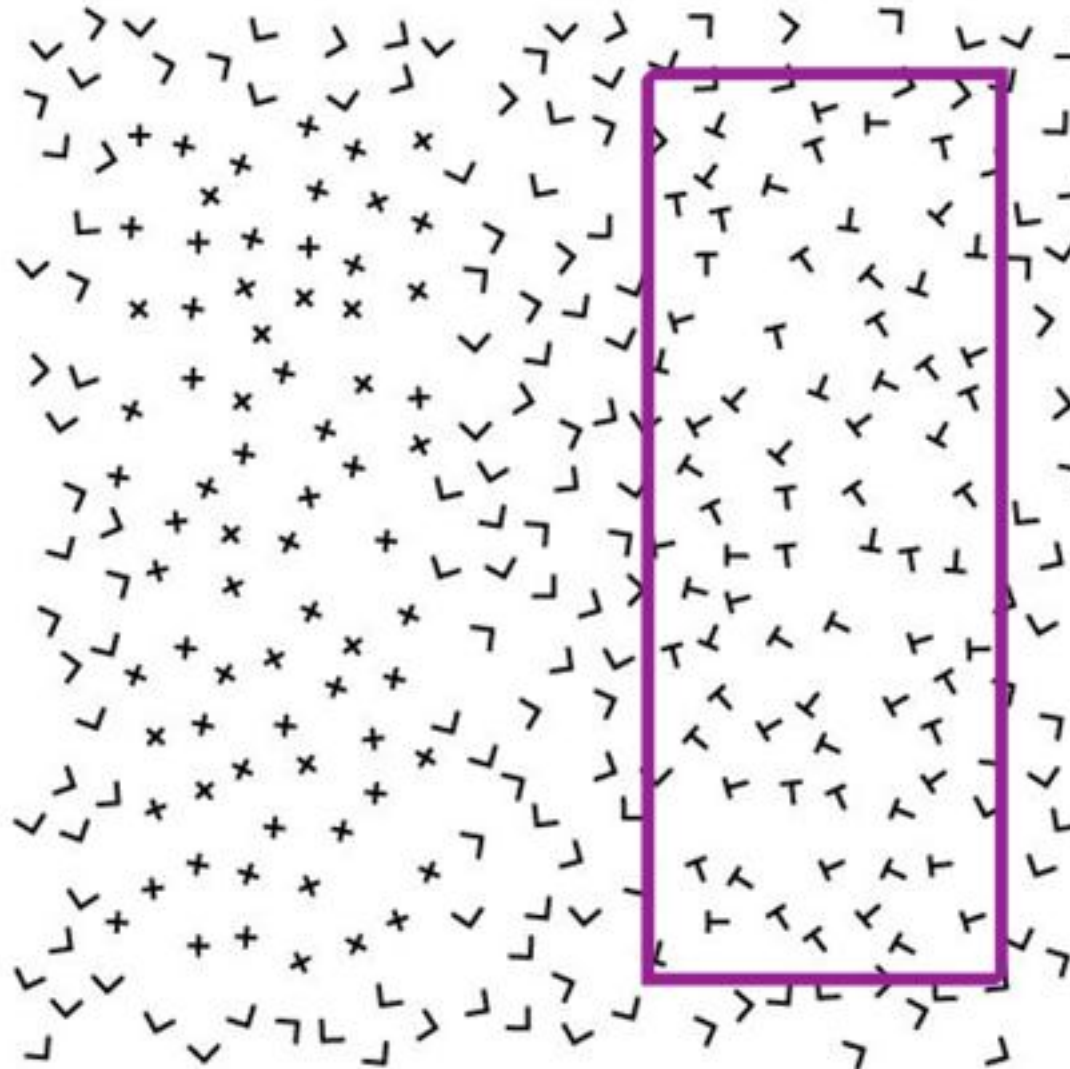


«Текстура»



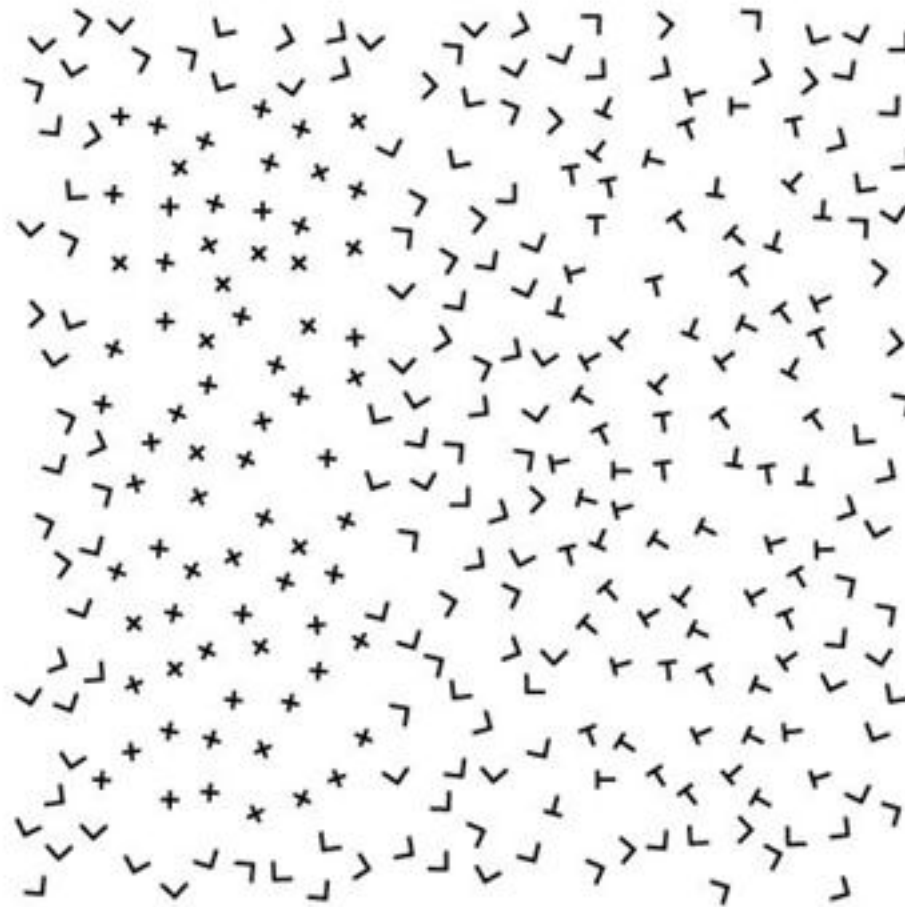


«Текстура»





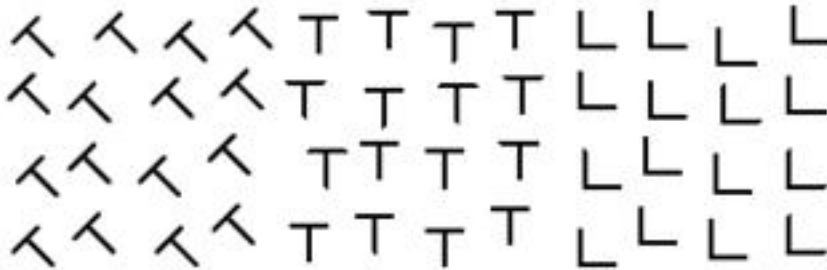
«Текстура»



Типичный пример текстурного шаблона для исследований психофизиологического восприятия изображений



Текстура



(a)

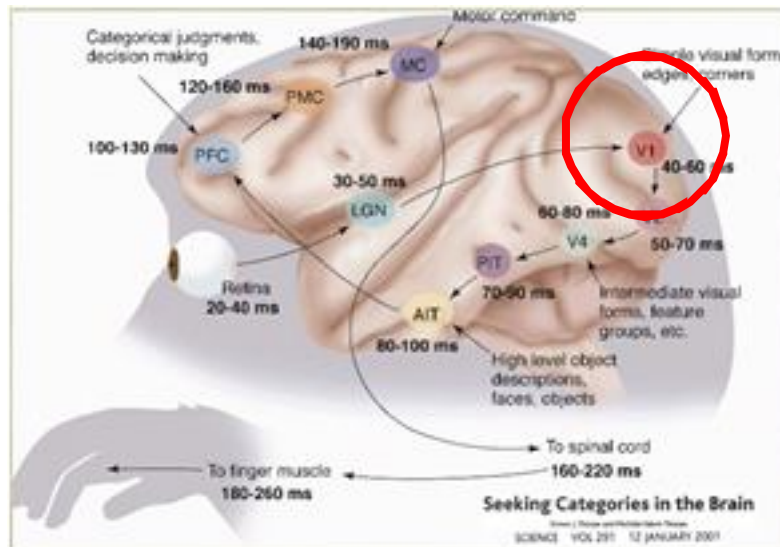


(b)

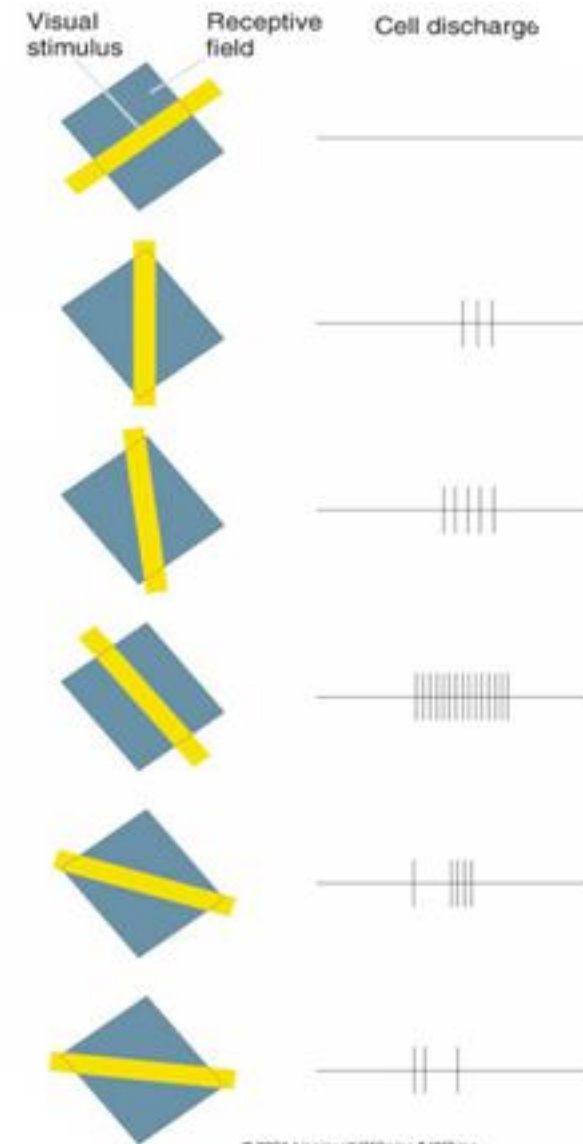
- Это типичные примеры текстурных шаблонов для исследований психофизиологического восприятия изображений
- Человек явно использует не только яркость и цвет, но и ориентацию краёв (градиентов изображения), их распределение, для анализа изображений



«Простые клетки» V1

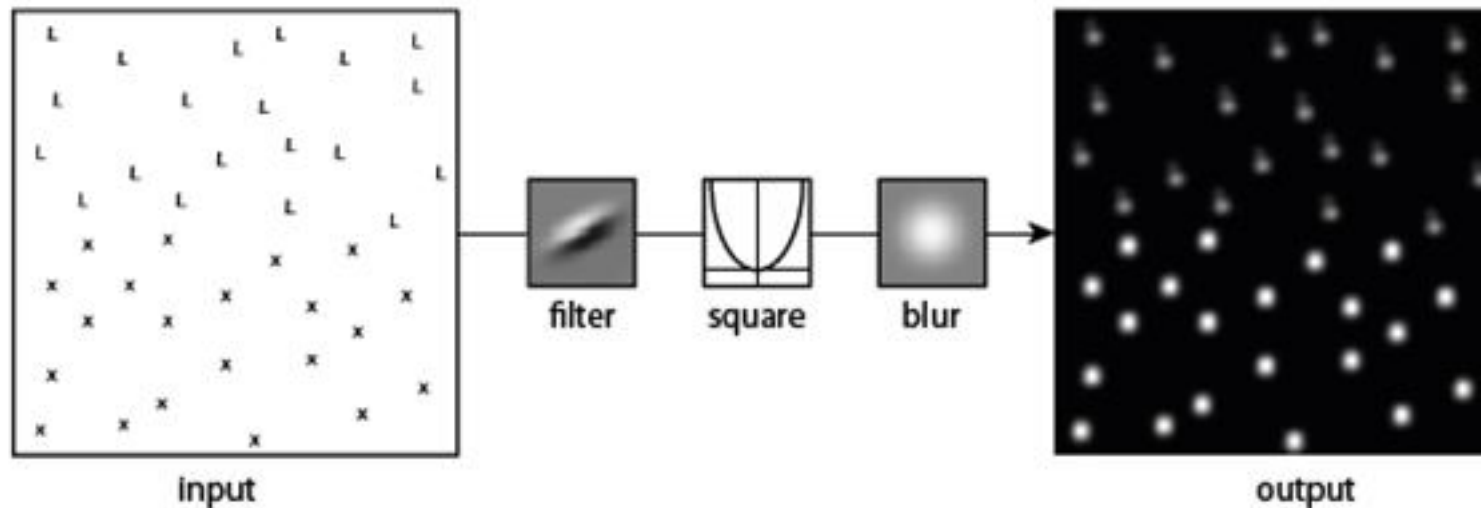


- В первичной визуальной коре головного мозга есть клетки, чувствительные к краям определенной ориентации
- Для каждой области есть набор таких клеток, чувствительные к краям разной ориентации





Анализ текстуры



- Выберем фильтр, чувствительный к краю определенной ориентации
- Результат фильтрации сгладим
- Будут «подсвечены» области, содержащие текстуру с краями заданной ориентации

Pietro Perona and Jitendra Malik «Detecting and Localizing edges composed of steps, peaks and roofs», ICCV 1990



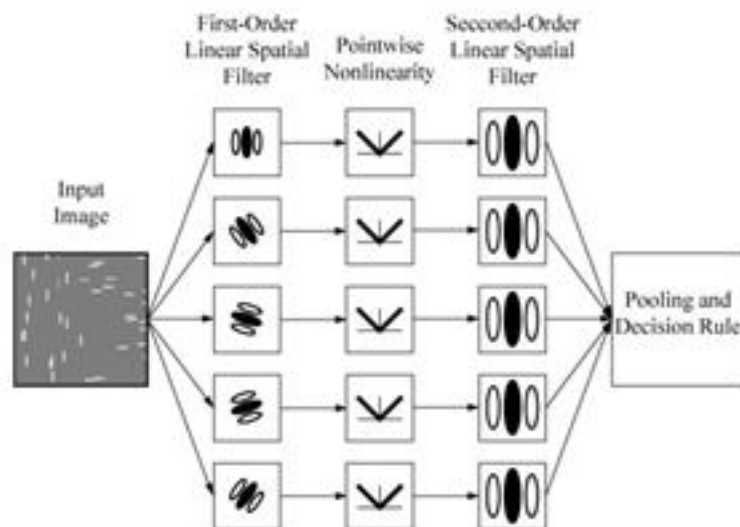
Банки фильтров

- Возьмём теперь несколько фильтров разного масштаба и ориентации
- Такой набор называют «банк фильтров»
- Каждый пиксель изображения после обработки банком фильтров даёт вектор признаков
- Этот вектор признаков эффективно описывает локальную текстуру окрестности пикселя
- Активно используется в сегментации, распознавании изображений и т.д.

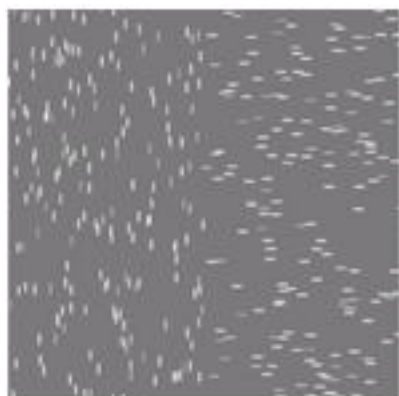




Модель «back pocket»



- Такие модели текстуры называют «back pocket»
- Модель «биологически ВОЗМОЖНА»
- По подобным моделям опубликовано много работ.



Ввод



После 1 этапа



После 2 этапа



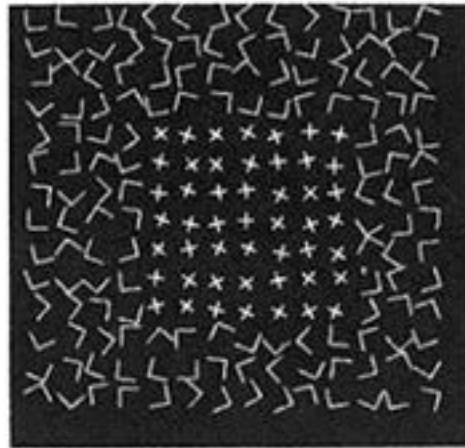
Выход



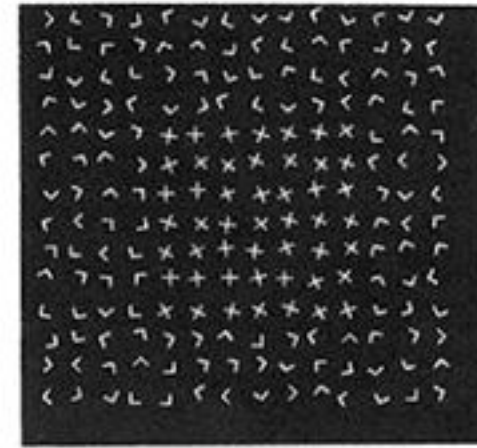
Пример, Bergen & Adelson (1988)



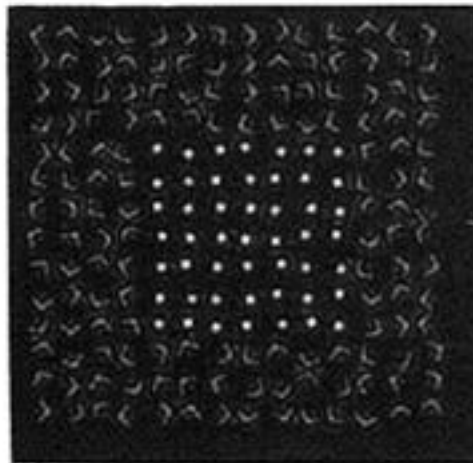
a



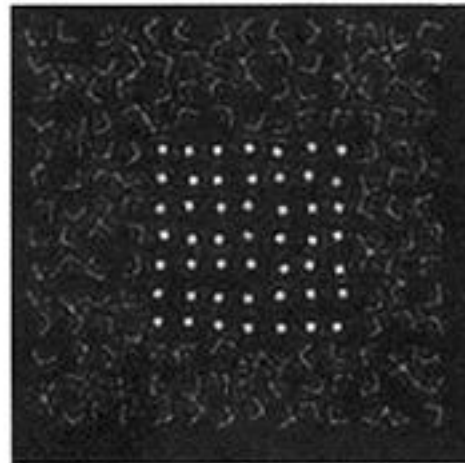
b



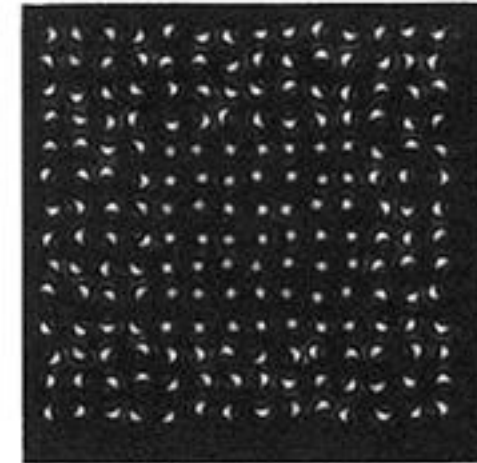
c



d



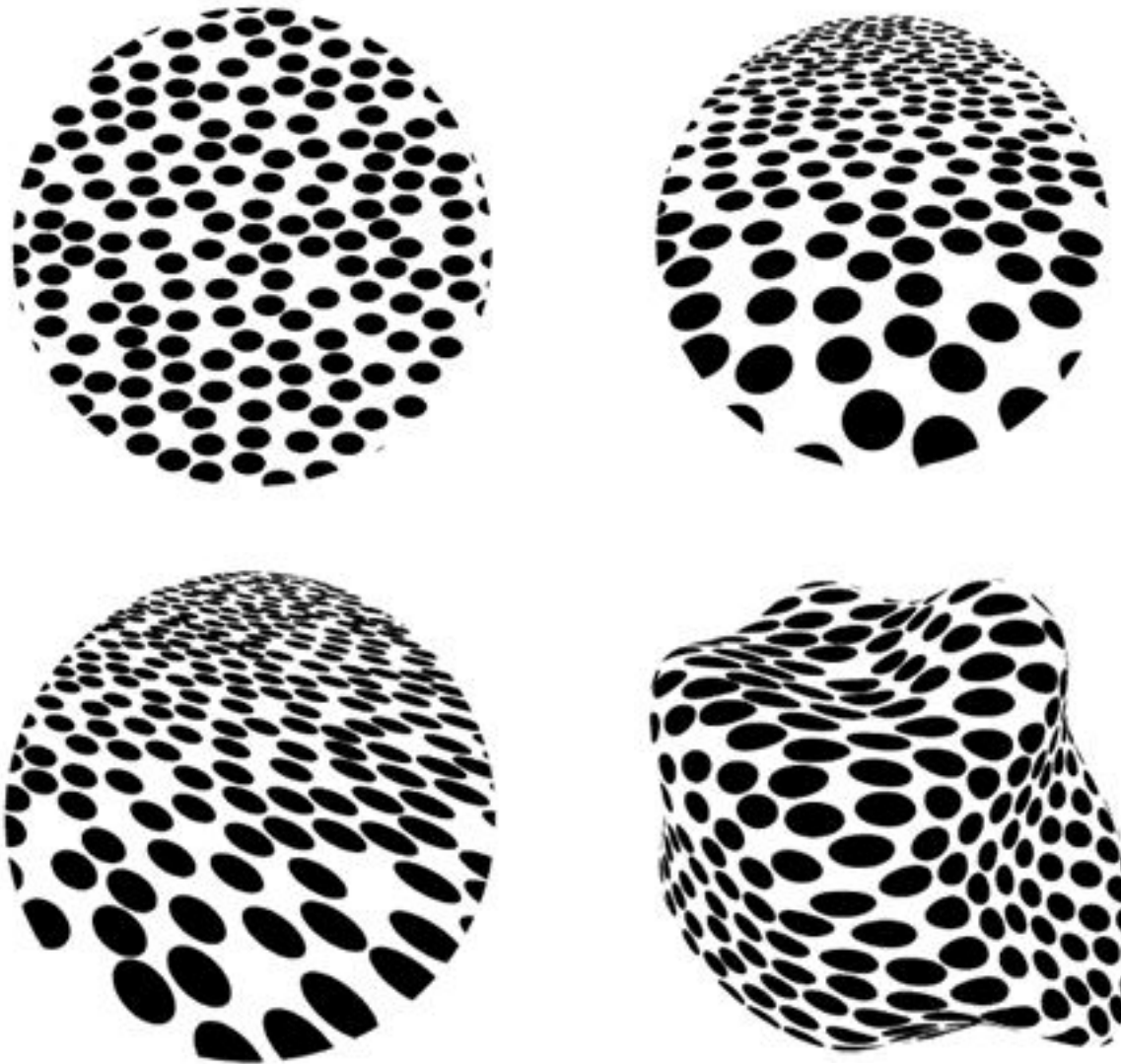
e



f



Психологическое свойство текстуры





Форма из текстуры

- Человек интуитивно считает текстуру **изотропной**, т.е. с постоянными свойствами на поверхности объекта
- Shape from texture: Исходя из предположения об изотропности шаблона текстуры, можно определить наклон поверхности

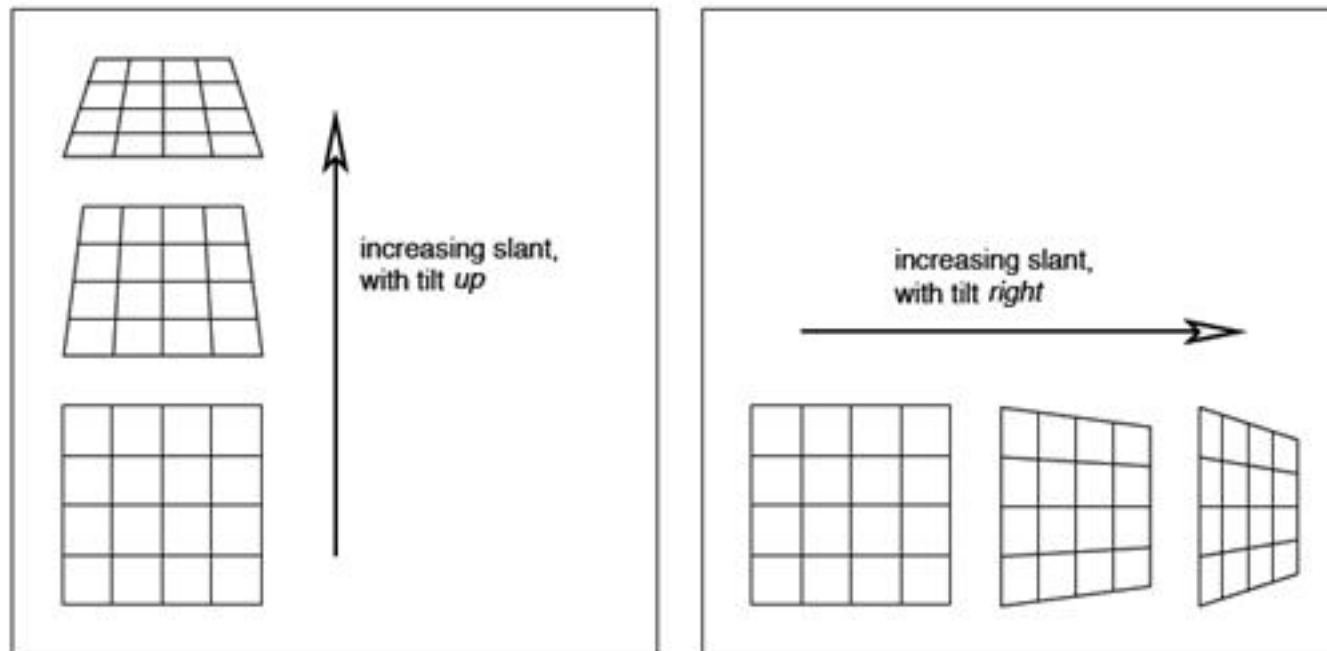


Figure 8.7. Surface orientation is often characterized in terms of *slant* and *tilt*.