

ЛЕКЦИЯ 5. ВЫДЕЛЕНИЕ И АНАЛИЗ ПРИЗНАКОВ

Демидов Д.В.

Обработка аудиовизуальной информации.
Бакалавры, 6 семестр. Магистры, 9 семестр

План лекции

2

- Анализ изображений текста.
- Выделение признаков символов.
- Вычисление моментов.
- Анализ профилей, сегментация.
- Алгоритмы Data mining для классификации символов

3

Признаки для ч/б изображений

Скалярные признаки. Веса.

Вес чёрного, доля чёрного

4

- Вес чёрного:

$$weight = \sum_x \sum_y f(x, y)$$

- Удельный вес чёрного:

$$weight_{rel} = \frac{\sum_x \sum_y f(x, y)}{S}$$

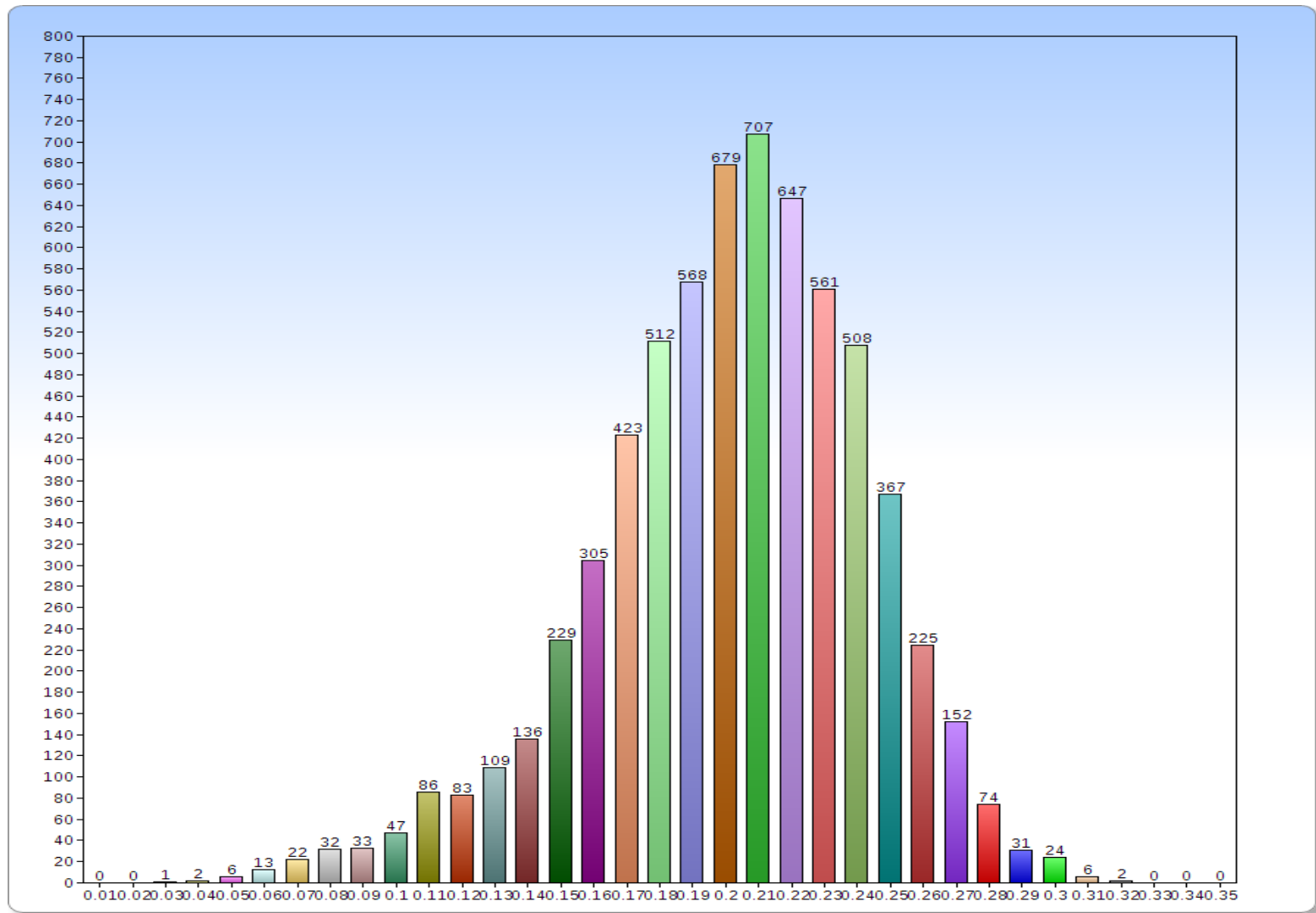
- Вес белого:

$$weight^{f=0} = S - weight^{f=1} = \sum_x \sum_y (1 - f(x, y))$$

- Самый простой интегральный признак.
- В макромасштабе может сказать о характере изображения (график, текст, фото)
- В микромасштабе позволяет различать символы по весу.

Пример гистограммы удельных весов символов для одного шрифта

5



Изолированные пиксели

6

- Изолированный белый пиксель (все соседи чёрные) – простейший вид зашумлённости типа «соль».

```
1 1 1
1 0 1
1 1 1
```

- Изолированный чёрный пиксель (все соседи белые) – простейший вид зашумлённости типа «перец».

```
0 0 0
0 1 0
0 0 0
```

- Удельный вес изолированных пикселей может сказать о степени зашумлённости изображения.

Доля пикселей с k соседями, k_1 - k_2 соседями

7

- Можно посчитать удельный вес пикселей, имеющих k соседей того же цвета.
- Можно посчитать удельный вес пикселей, имеющих $k_1 \leq k \leq k_2$ соседей того же цвета. $k \in [1..8]$
- Эти признаки могут сказать о повышенном шуме «соль»/«перец», когда искажений столько, что помимо изолированных пикселей появляются пары, тройки, четверки смежных искажённых пикселей.
- На микроуровне по этим долям можно оценить степень фрагментированности символов, зная доли для нефрагментированного текста.

Особые фигуры

8

- Доля чёрных крестиков

0	1	0
1	1	1
0	1	0
- Доля белых крестиков

1	0	1
0	0	0
1	0	1
- Доля косых чёрных крестиков
(очень похожи на белые крестики)

1	0	1
0	1	0
1	0	1
- Доля косых белых крестиков
(очень похожи на чёрные крестики)

0	1	0
1	0	1
0	1	0
- Сумма долей, другие фигуры...

Признаки для выявления шума

9

- Логические и морфологические преобразования изменяют изображение.
- Признаки, рассчитанные для изменённого изображения, несут другую информацию.
- Пусть A – матрица яркостей исходного изображения. Обозначим через $A]$ матрицу A без последнего столбца, а через $[A$ – матрицу A без первого столбца.

Примеры преобразований для оценки шума

10

- Вычислим матрицы:
 - ▣ $\text{Noise} = A \text{ xor } [A]$
 - ▣ $T = A \text{ or } [A]$
 - ▣ $R = (\text{not Noise}) \text{ or } (\text{not dilate}(\text{not } T))$
 - ▣ $\text{Noise2} = A \text{ xor } (T \text{ and } R)$
 - ▣ $\text{Noise3} = \text{Noise} \text{ xor } (\text{not } R)$
- Доли белого в Noise, Noise2, Noise3 по-разному говорят об уровне шума в A.

ポストオーク通信

いると指摘した。
案件への場合
なる原油の輸出
した制度作りが
経済学校のハー
隊を政府内に

ポストオーク通信

いると指摘した。
案件への場合
なる原油の輸出
した制度作りが
経済学校のハー
隊を政府内に

ポストオーク通信

いると指摘した。
案件への場合
なる原油の輸出
した制度作りが
経済学校のハー
隊を政府内に

ポストオーク通信

いると指摘した。
案件への場合
なる原油の輸出
した制度作りが
経済学校のハー
隊を政府内に

11

Признаки для ч/б изображений

Скалярные признаки. Моменты

Статистические моменты

12

- Рассчитываются для совокупности чёрных точек относительно некоторого выбранного центра.
- Наиболее общеупотребительными для распознавания символов являются построчные, центральные и нормированные моменты.
- Для цифрового изображения, хранящегося в двумерном массиве, построчные моменты являются функциями координат каждой точки изображения следующего вида:

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$

где

- $p, q = 0, 1, \dots, \infty$ - порядки момента;
- M и N являются размерами изображения по горизонтали и вертикали
- $f(x, y)$ является яркостью пикселя в точке (x, y) на изображении.

Центральные и осевые моменты

13

- Центральные моменты являются функцией расстояния точки от центра тяжести символа:

$$\mu_{pq} = \sum \sum (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

где x и y с чертой – координаты центра тяжести.

- Нормированные центральные моменты получаются в результате деления центральных моментов на моменты нулевого порядка.

Вес на основе моментов

14

- Вес чёрного соответствует нулевому моменту ($p=0, q=0$):

$$weight = \mu_{0,0} = \sum_{x=1}^M \sum_{y=1}^N f(x, y)$$

- Удельный вес соответствует нулевому моменту отнесённому к площади изображения:

$$weight_{rel} = \frac{weight}{M \cdot N} \in [0; 1]$$

Центр тяжести

15

- Абсолютные координаты центра тяжести соответствуют парам $p=1$, $q=0$ для x и $p=0$, $q=1$ для y , отнесённым к нулевому моменту:

$$\bar{x} = \frac{\sum_x \sum_y x \cdot f(x, y)}{weight} \quad \bar{y} = \frac{\sum_x \sum_y y \cdot f(x, y)}{weight}$$

- Относительные координаты центра тяжести в отрезке $[0; 1]$:

$$\bar{x}_{rel} = \frac{\bar{x} - 1}{M - 1} \in [0; 1] \quad \bar{y}_{rel} = \frac{\bar{y} - 1}{N - 1} \in [0; 1]$$

Осевые моменты инерции

16

- Осевые моменты инерции для четырёх центральных осей: горизонтальной, вертикальной и двух диагональных

$$I\bar{x} = \sum_x^M \sum_y^N (y - \bar{y})^2 \cdot f(x, y)$$

$$I\bar{x}_{rel} = \frac{I\bar{x}}{M^2 + N^2}$$

$$I\bar{y} = \sum_x^M \sum_y^N (x - \bar{x})^2 \cdot f(x, y)$$

$$I\bar{y}_{rel} = \frac{I\bar{y}}{M^2 + N^2}$$

$$I_{cm45} = \frac{1}{2} \sum_x^M \sum_y^N (y - \bar{y} - x + \bar{x})^2 \cdot f(x, y)$$

$$I_{cm45}_{rel} = \frac{I_{cm45}}{M^2 + N^2}$$

$$I_{cm135} = \frac{1}{2} \sum_x^M \sum_y^N (y - \bar{y} + x - \bar{x})^2 \cdot f(x, y)$$

$$I_{cm135}_{rel} = \frac{I_{cm135}}{M^2 + N^2}$$

Применение моментов

17

- Строковые моменты, как правило, обеспечивают низкий уровень распознавания.
- Центральные и нормированные моменты предпочтительнее из-за их большей инвариантности к преобразованиям изображений.
- На основе моментов можно вычислить такие признаки как:
 - вес символа (число черных пикселей в бинарном изображении);
 - центр масс на основе момента 1-го порядка;
 - эллиптические параметры: наклон главной оси α и эксцентриситет (соотношение малой и главной полуосей);
 - скошенность символа (асимметрию) как нормированный момент 3-го порядка;
 - коэффициент эксцесса символа (степень островершинности, крутизны) на основе нормированного момента 4-го порядка;
 - моменты более высоких порядков.

$$e = \sqrt{1 - \frac{b^2}{a^2}}$$

$$A_3 = \frac{\mu_3}{\sigma_3}$$

$$\gamma_2 = \frac{\mu_4}{\sigma_4} - 3$$

18

Признаки для ч/б изображений

Векторные признаки. Профили

Профили

19

- Профиль (проекция) представляет собой сумму яркостей пикселей, подсчитанную вдоль какого-либо направления.
- Горизонтальный профиль по уровню Y рассчитывается как сумма яркостей пикселей, лежащих на высоте Y

$$Proj_Y [x] = \sum_{x=0}^{W-1} I(x, y)$$

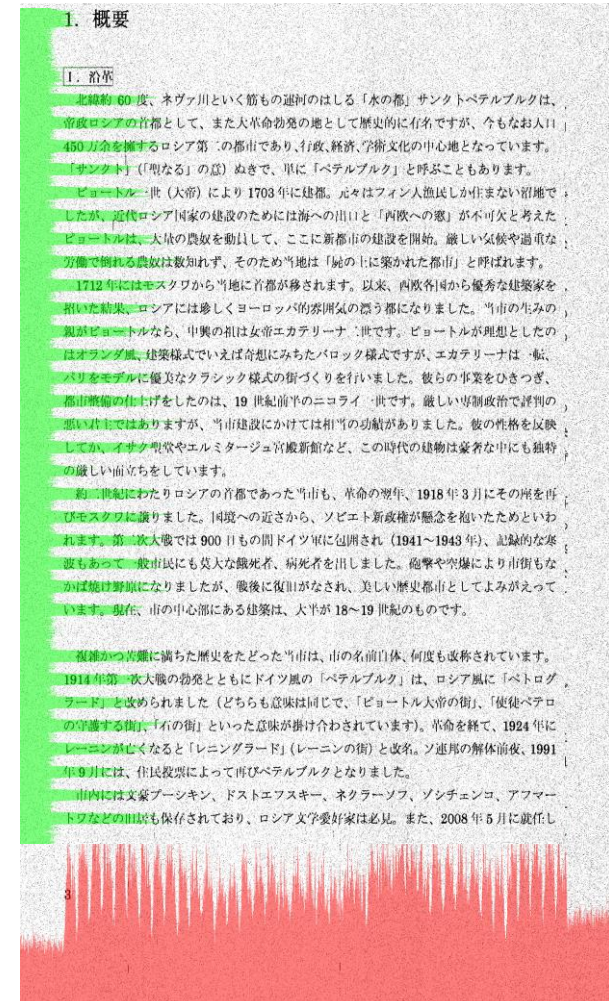
- Вертикальный профиль по уровню X рассчитывается как сумма яркостей пикселей, лежащих на вертикали X :

$$Proj_X [y] = \sum_{y=0}^{H-1} I(x, y)$$

Профили (2)

20

- Горизонтальный профиль – совокупность горизонтальных профилей по всем уровням Y от 0 до $H-1$
- Вертикальный профиль – совокупность вертикальных профилей по всем уровням X от 0 до $W-1$



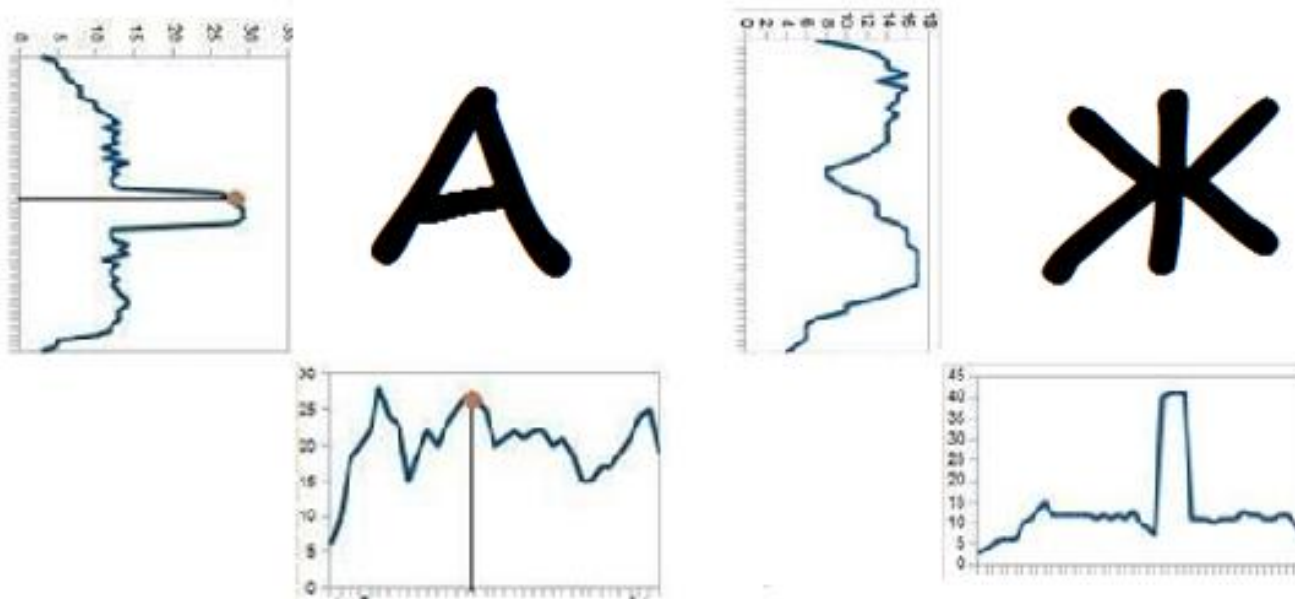
Профили (3)

21

- Для **изображений текста** один из профилей будет иметь вид «расчёски», где ширина зубца соответствует высоте строки.
- Второй профиль не имеет чёткой структуры.
- Ширина основания профиля (минимальный уровень) указывает на общий уровень шума. Это подсказка для автоматического подбора фильтра.
- В **масштабе страницы** профили могут сказать об ориентации текста: отличить горизонтальное письмо от вертикального.
- В **масштабе символа** несут характерную информацию о нём.

Примеры профилей символов

22



Снизу вертикальный, слева горизонтальный профиль.

Косые профили

23

- Проекции на прямую, проходящую под некоторым углом.
- Могут пригодиться для определения угла поворота текста: для этого нужно построить несколько косых профилей с углами от -15 до $+15$ градусов и найти самую лучшую «расческу».
- Профили под углами 45 и 135 градусов могут пригодиться как дополнительные признаки символов при классификации.

Признаки для ч/б изображений

Векторные признаки. Метод пересечений и метод зон

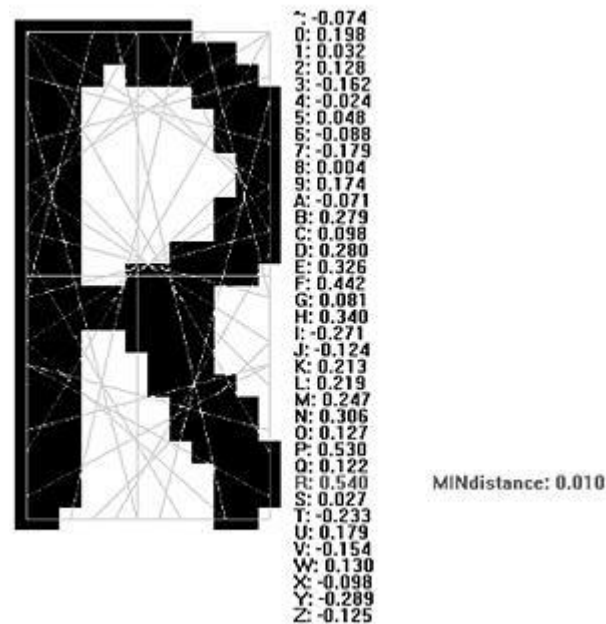
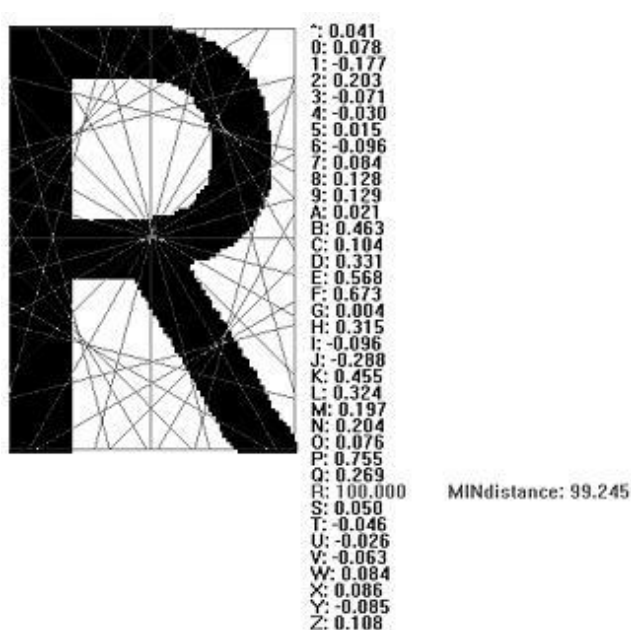
Метод пересечений

25

- Задаётся система секущих прямых.
- Рассчитывается количество пересечений изображения с каждой из прямых.
- Получается вектор, длина которого равна количеству прямых, а значение элемента вектора равно количеству пересечений исходного изображения соответствующей прямой.
- Для сравнения векторов вводится мера близости. Например, Евклидово расстояние, Манхеттенское, Левенштейна (терпения, друзья).

Метод пересечений. Пример

26



- ❑ Слева эталонное изображение
- ❑ Справа реальное изображение
- ❑ Рядом вектора близости с другими образцами

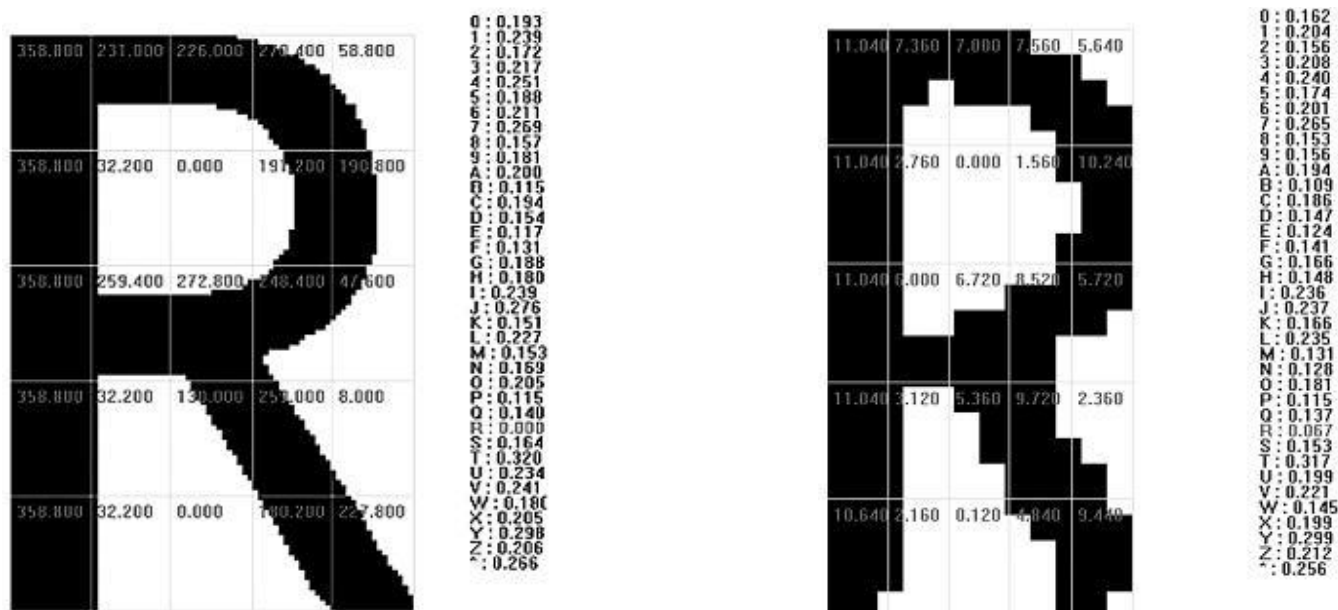
Метод зон

27

- Метод зон предполагает разделение площади рамки, объемлющей символ, на области и последующее использование плотностей точек в различных областях в качестве набора характерных признаков.
- Получается вектор, длина которого равна количеству зон, а значение элемента вектора равно пиксельному весу каждой зоны.

Метод зон. Пример

28



- ❑ Слева эталонное изображение
- ❑ Справа реальное изображение
- ❑ Рядом вектора близости с другими образцами

Признаки для ч/б изображений

n-мерные признаки. Анализ фазовых пространств

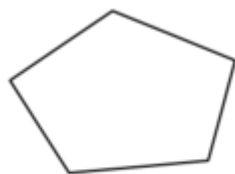
Фазовое пространство (r, φ)

30

- Пусть
 - ▣ r – расстояние прямой от центра координат
 - ▣ φ – угол наклона прямой
 - ▣ $A(r, \varphi)$ – аккумуляторная функция (аккумулятор «голосов»), вычисляемая как количество точек на прямой (r, φ) .
- Локальные максимумы $A(r, \varphi)$ соответствуют наиболее длинным прямым линиям исходного изображения.
- Фазовое пространство строится с помощью преобразования Хафа (Hough Transform).

Пример пространства Хафа

31



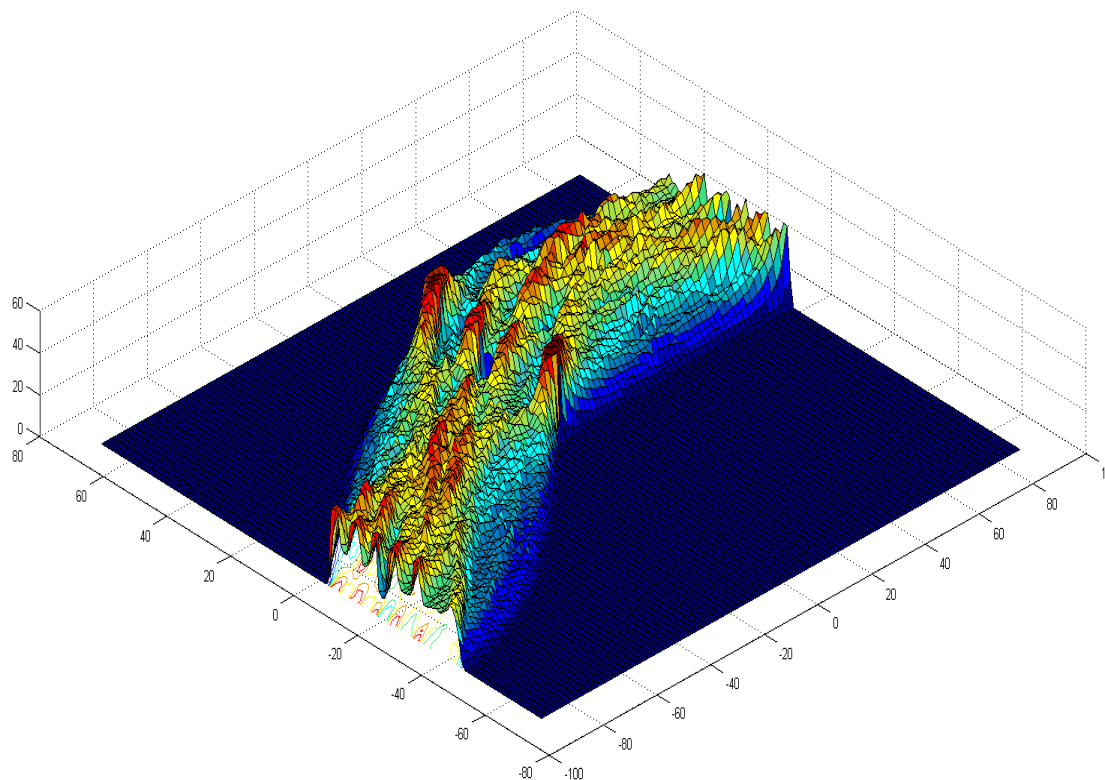
Изображение
пятиугольника



2D-визуализация в
координатах (R, ϕ)

Пример пространства Хафа

32

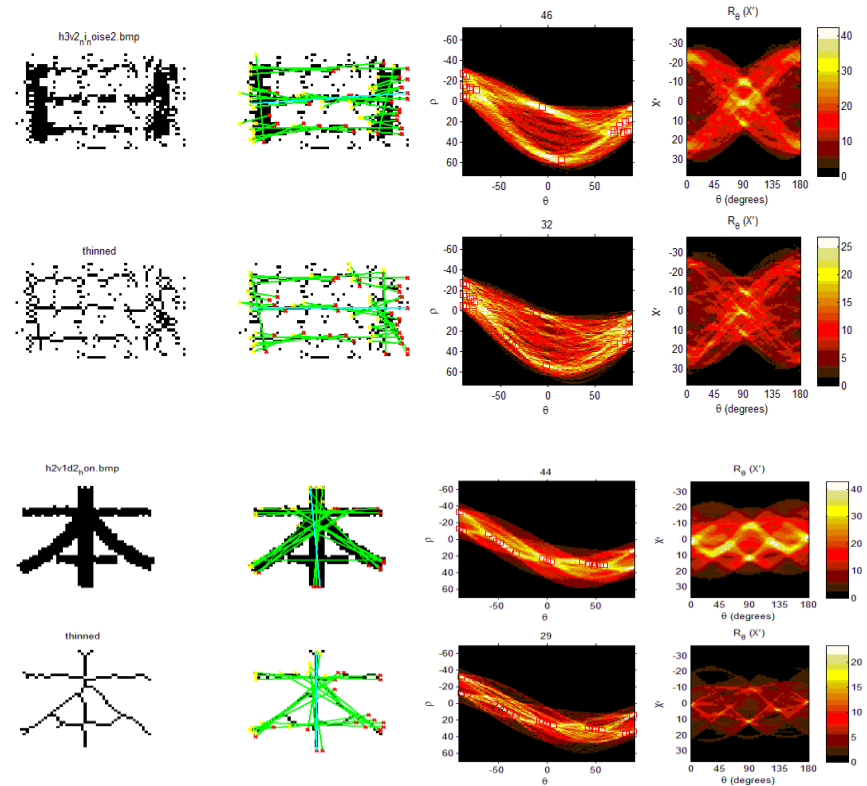


В виде 3D-карты высот

ВЫЯВЛЕНИЕ ПРЯМЫХ ЛИНИЙ, АППРОКСИМАЦИЯ ПРЯМЫМИ ЛИНИЯМИ

34

- Самые яркие 50 точек соответствуют 50-ти наиболее длинным линиям, которые хорошо аппроксимируют исходное изображение.
- Однако, начало и конец линии нужно подбирать.



Поиск окружностей с помощью преобразования Хафа

35

- Поиск окружностей заданного радиуса R можно проводить в параметрическом пространстве (x, y) с аккумуляторной функцией $A(x, y)$, вычисляемой как число голосов, удовлетворяющих уравнению:

$$(x - x_0)^2 + (y - y_0)^2 = R^2$$

- Максимум аккумулятора соответствует положению центра окружности на изображении.
- Если радиус окружности является неизвестным или переменным, необходимо добавить R в качестве дополнительной переменной в параметрическое пространство-аккумулятор $A(x, y, R)$.
- Если важны только сами центры, а радиусы не важны, то можно не увеличивать размерность пространства, но вместо голоса-точки использовать луч в сторону центра.

$$(x - x_0)^2 + (y - y_0)^2 = R^2, R \in [0, R_{\max}]$$

Обобщённое преобразование Хафа

36

- GHT (Generalized Hough Transform) предложено Баллардом для обнаружения кривых линий заданного типа.
- в отличие от задач обнаружения окружности, существенно то, что расстояние R от текущего пиксела границы искомого контура до ее центра больше не константа, а является функцией $R(\varphi)$ от угла φ радиуса-вектора, направленного от точки контура к центру.
- Для простых форм функция $R(\varphi)$ может быть описана аналитически.
- Для сложных форм функция задаётся с помощью просмотровой таблицы LUT (look-up-table), содержащей дискретные значения $R(\varphi)$ для различных значений углов. Вначале детектор Хафа обучается на эталоне и строит LUT, потом применяется к тестовому изображению.

Преобразование Радона

37

- Интегральное преобразование, обладающее свойством обратимости (в отличие от преобразования Хафа).
- Определение:

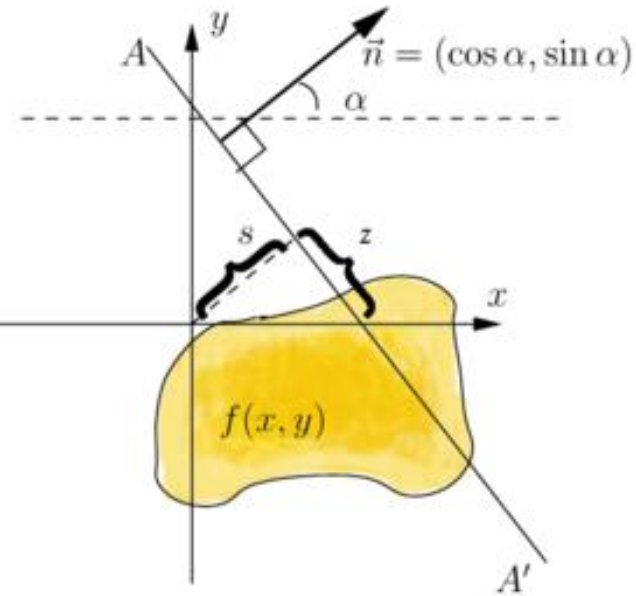
$$R(s, \alpha) = \int_{-\infty}^{\infty} f(s \cos \alpha - z \sin \alpha, s \sin \alpha + z \cos \alpha) dz$$

- Геометрический смысл: интеграл от функции $f(x, y)$ вдоль прямой, перпендикулярной вектору

$$\vec{n} = (\cos \alpha, \sin \alpha)$$

и проходящей на расстоянии s от начала координат.

- В дискретном двумерном случае соответствует преобразованию Хафа.



Обратное преобразование Радона

38

- Одномерное преобразование Фурье от преобразования Радона для функции $f(x, y)$ есть двумерное преобразование Фурье от функции $f(x, y)$
- Существование обратного преобразования Фурье говорит о существовании обратного преобразования Радона

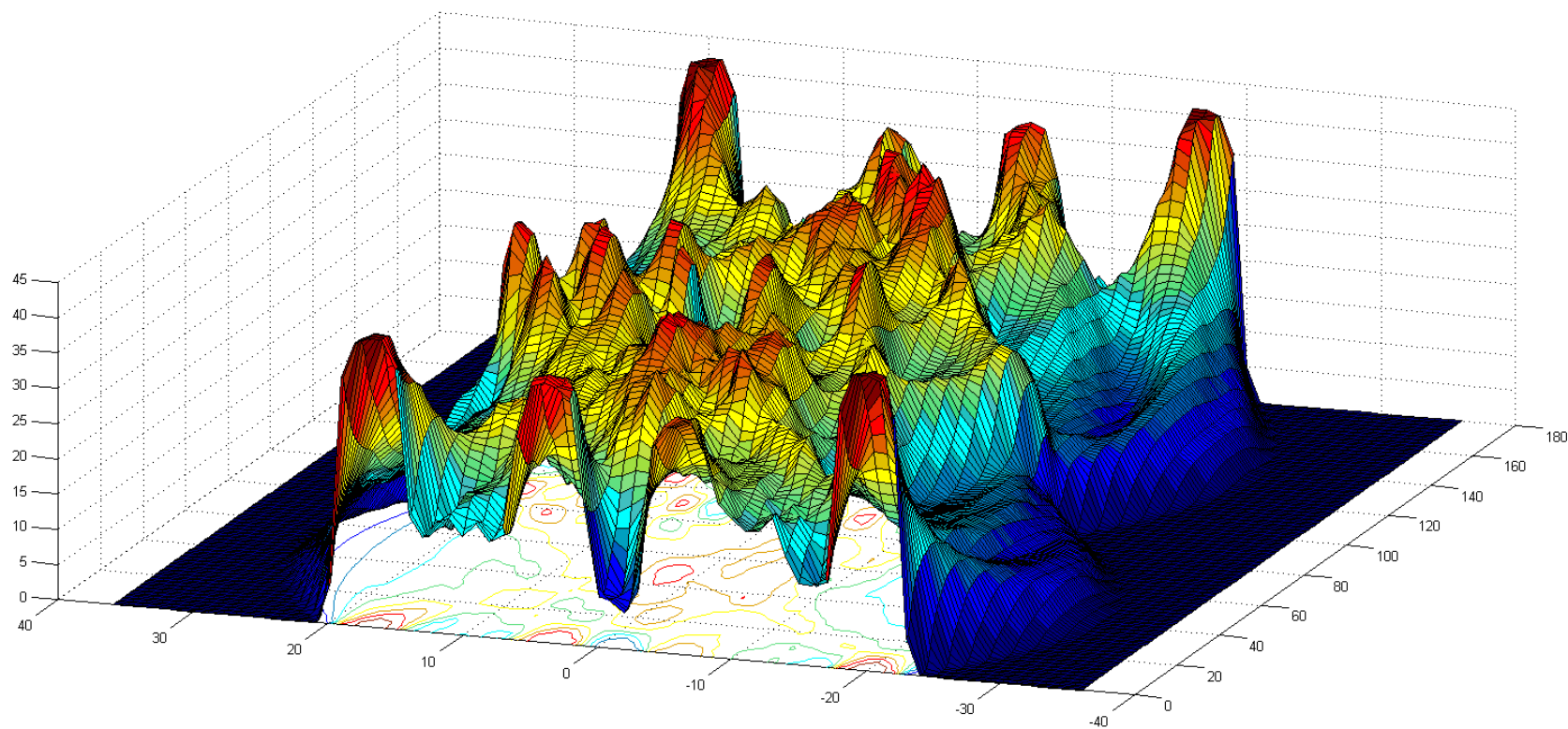
$$f(x, y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^\infty e^{i\omega(x\cos\alpha + y\sin\alpha)} \tilde{R}(\omega, \alpha) \omega d\omega d\alpha$$

$$\tilde{R}(\omega, \alpha) = \int_{-\infty}^{\infty} R(s, \alpha) e^{-i\omega s} ds$$

- Таким образом, по значениям R можно восстановить значения f .

Пример карты высот

39



Признаки для ч/б изображений

Скелеты изображений.

Алгоритмы утончения

Утончение символов

41

- Алгоритмы приблизительной скелетизации бинарных изображений часто называют алгоритмами *утончения* (*утоньшения*).
- Дискретные скелеты называют *остовами*.

Непрерывный случай

42

- На непрерывной плоскости скелет можно математически строго определить следующим образом:
 - ▣ Пусть R - множество точек плоскости, B - его граница и P - точка множества R .
 - ▣ Ближайшим соседом точки P на границе B является такая точка M , принадлежащая границе B , что на этой границе нет никакой другой точки, расстояние от которой до точки было бы меньше расстояния PM .
 - ▣ Если точка P имеет более одного ближайшего соседа, то P называют *остовой* точкой множества R . Объединение всех остовных точек называется *остовом*, или *серединной осью* множества R .
 - ▣ Из этого следует, что остовные точки являются центрами окружностей, полностью покрываемых множеством R , причем не существует окружностей с тем же центром и большим радиусом, покрываемых множеством R .

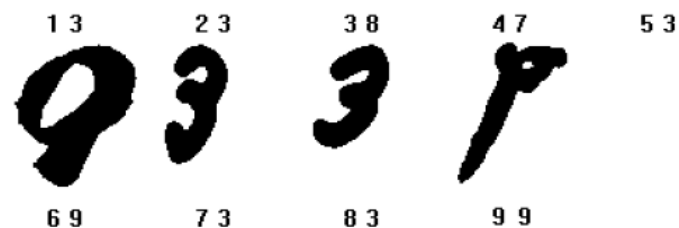
Дискретный случай

43

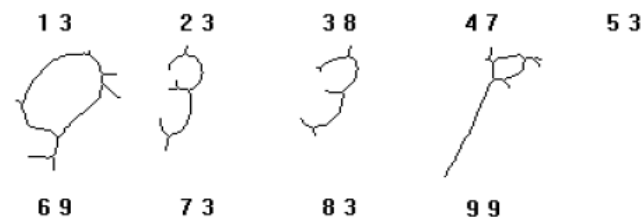
- На дискретной сетке остовом множества пикселей R называется множество, формируемое следующим образом.
 - Сначала определяются пиксели остова и пиксели контура, принадлежащие множеству R .
 - После этого все пиксели контура, не являющиеся остовными, удаляются и полученное в результате этой процедуры множество заменяет множество R .
 - Этот процесс повторяется до тех пор, пока не будет сформировано множество, включающее только остовные пиксели.

Утончение символов. Примеры

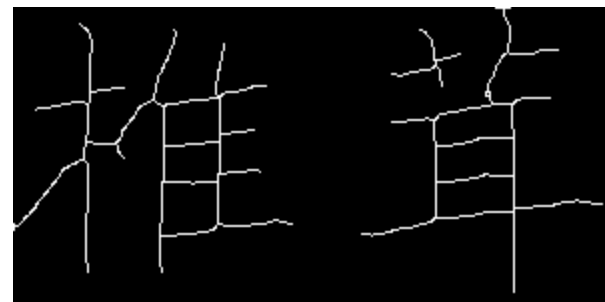
44



(a)



(b)



Некоторые алгоритмы утончения

45

- Harish Kumar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (5) , 2011, 2375-2379
 - ▣ FPTA - Fast Parallel Thinning Algorithm
 - ▣ GHPTA - Guo & Hall's parallel thinning Algorithm
 - ▣ RPTA - Robust Parallel Thinning Algorithm for binary images
 - ▣ PPTA - Preprocessing Thinning Algorithms for Handwritten Character Recognition
- Huang L., Wan G., Liu C. An improved parallel thinning algorithm // Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003). pp. 780-783, 2003
 - ▣ IPTA - Improved Parallel Thinning Algorithm
- Han N.H., La C.W., Rhee P.K. An Efficient Fully Parallel Thinning Algorithm // Proc. IEEE Int.Conf.Document Analysis and Recognition, Vol.1,pp.137-141(1997).
 - ▣ EFPTA - Efficient Fully Parallel Thinning Algorithm

Сегментация

Выделение текстовых областей, строк, символов
пороговой обработкой

Выращивание областей и диаграммы Вороного

Сегментация изображений

47

- Сегментацией называется процесс подразделения сцены на составляющие части или объекты.
- Сегментация является одним из основных элементов работы автоматизированной системы технического зрения, т.к. именно на этой стадии обработки объекты выделяются из сцены для дальнейшего распознавания и анализа
- Наиболее часто сегментацию проводят по яркости для одноцветного изображения и цветовым координатам для цветного изображения.
- Конечной целью сегментации изображений является разбиение поля зрения D на области объектов D_1, \dots, D_s и область фона D_ϕ

Качество сегментации

48

- Качество зависит от того, насколько учтена важная информация:
 - ▣ число объектов S ;
 - ▣ некоторые характеристики распределения яркости в областях объектов или фона, например экстремальные значения яркости, количество перепадов яркости;
 - ▣ оценки яркостного перепада при переходе из области фона в область объектов;
 - ▣ форма объекта;
 - ▣ информация о том, какую часть поля зрения занимает объединение областей объектов

Классификация по степени автоматизации

49

- Методы сегментации изображений делятся на два класса:
 - ▣ автоматические, то есть такие методы, которые не требуют взаимодействия с пользователем;
 - ▣ интерактивные (ручные) методы, использующие введенные пользовательские данные во время работы

Классификация по дихотомии «граница vs область»

50

- Алгоритмы сегментации также можно разделить на два типа:
 - ▣ основанные на разрыве функции яркости;
 - Изображение разбивается на области на основании некоторого изменения яркости, такого как, например, перепады яркости на изображении.
 - Например, *пороговая обработка* (пороговая классификация).
 - ▣ основанные на однородности функции яркости.
 - используется разбиение изображения по критериям однородности областей.
 - Например, *выращивание областей, слияние и разбиение областей*.

51

Пороговая обработка

Алгоритм выделения текстовой области

52

- Построить пару профилей изображения.
- Критерий начала зоны текста: при просмотре вертикального профиля от *начала* резкая смена нулевых или малых значений профиля на большие значения.
- Критерий окончания зоны текста: при просмотре вертикального профиля с *конца* резкая смена нулевых или малых значений профиля на большие значения.
- Внутри горизонтальной зоны выявить зону текста по вертикали, пользуясь теми же критериями, но для горизонтального профиля.
- Вернуть координаты найденной прямоугольной области.
- Обобщить алгоритм для поиска нескольких зон.

Алгоритм выделения строк

53

- Строки выделяются внутри текстовой области, выявленной предыдущим алгоритмом.
 - ▣ Используется только горизонтальный профиль (проекция влево), который просматривается слева направо.
 - ▣ Критерий верхней границы строки: резкая смена нулевых или малых значений профиля на большие значения.
 - ▣ Критерий нижней границы строки: резкая смена больших значений профиля на малые значения.
- Вернуть список пар высот, соответствующих найденным строкам.

Алгоритм сегментации символов

54

- Границы символов выделяются внутри строки.
 - Используется только вертикальный профиль (проекция вниз), который просматривается слева направо.
 - Критерий левой границы символа: резкая смена нулевых или малых значений профиля на большие значения.
 - Критерий правой границы символа: резкая смена больших значений профиля на 0 или 1.
- Удаление ложных границ:
 - Если левая и правая граница оказались ближе, чем ~ 5 пикселей, то удалить правую границу и следующую левую. Такое бывает с буквами Ы и Ю.
- Вернуть список пар границ, соответствующих символам.

Диаграммы Вороного

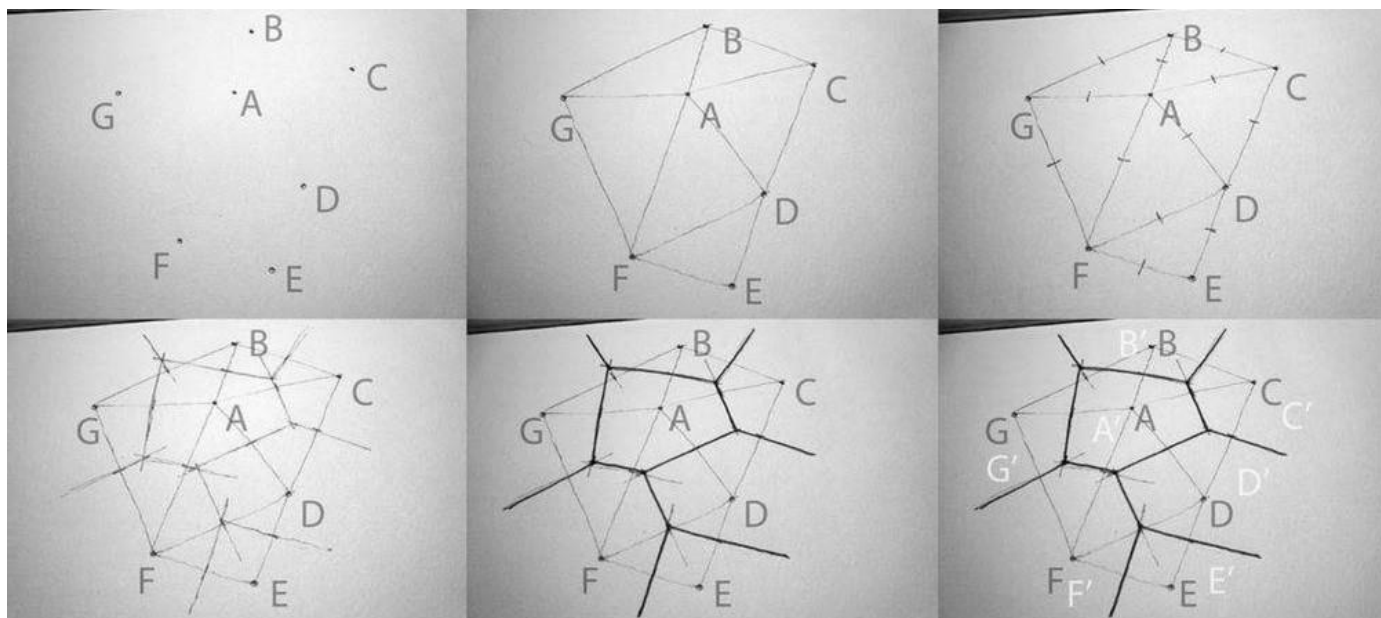
Диаграмма Вороного

56

- Диаграмма Вороного – геометрическое разбиение области на многоугольники, обладающие следующим свойством:
 - для любого центра системы $\{A\}$ можно указать область пространства, все точки которой ближе к данному центру, чем к любому другому центру системы.
 - Такая область называется многогранником Вороного (или областью Вороного)

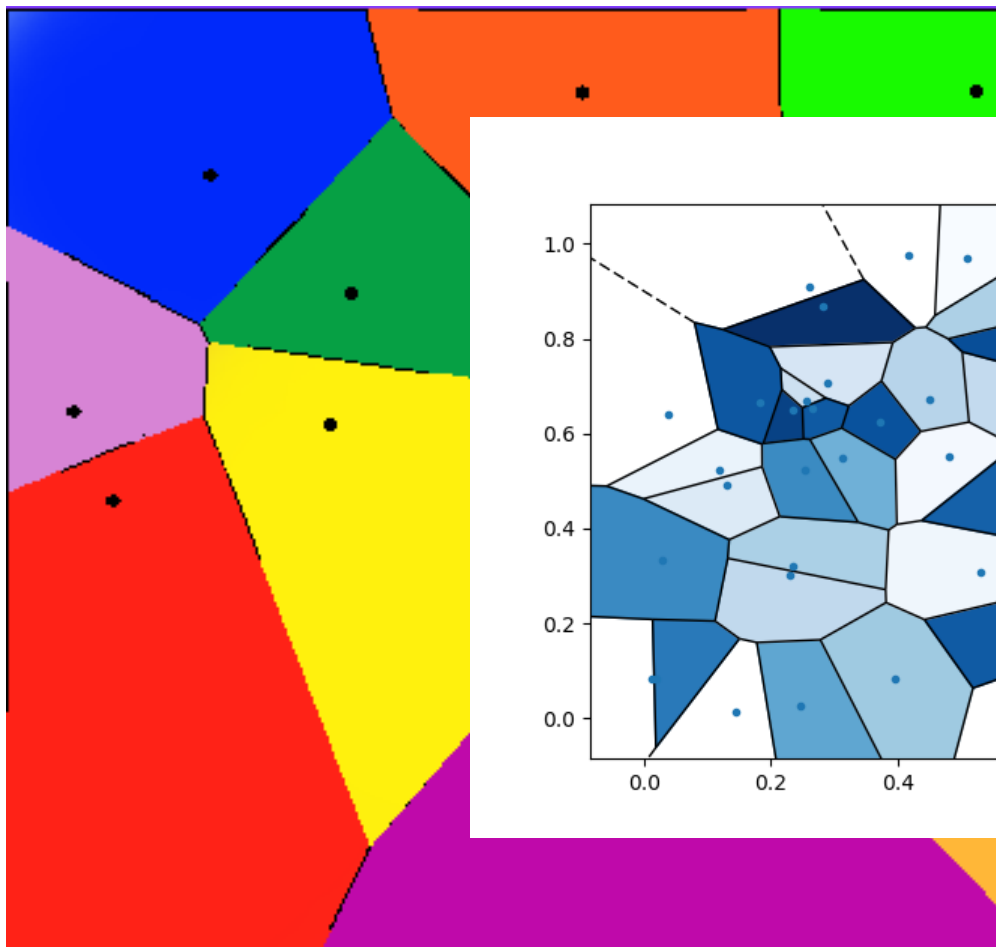
Построение диаграммы Вороного

57



Примеры

58



Ячейка Вороного

59

- Ячейка Вороного представляет собой выпуклый многоугольник, вершины многоугольников определяют вершины диаграммы Вороного, а соединяющие их отрезки – рёбра диаграммы Вороного. Таким образом, вся плоскость представляется объединением ячеек равноудалённых от точек-генераторов рёбер

Свойства диаграммы

60

- Каждая вершина диаграммы Вороного, полученной для множества N точек-генераторов, является точкой пересечения трёх рёбер диаграммы (при $N > 2$).
- Многоугольник $V(s_i)$ является неограниченным тогда и только тогда, когда точка s_i лежит на границе выпуклой оболочки множества S .
- Диаграмма Вороного, построенная для множества N точек имеет не более $2N - 5$ вершин и $3N - 6$ рёбер

Сегментация текста

61

- Для сегментации текста используется способ выделения слов и символов текста, основанный на использовании обобщённой диаграммы Вороного.
- Каждый символ трактуется как объект, состоящий из множества связанных точек. Таким образом, весь текст представляется набором связанных множеств точек c_1, \dots, c_n .
- Каждое связанное множество точек S_i , соответствующее некоторому символу текста, заменено центром его масс:

$$c_i = \frac{1}{|S_i|} \sum_{s \in S_i} s, \quad i = 1..N.$$

- Множество точек $C = \{c_i, i = 1..N\}$ рассматривается как множество генераторов диаграммы Вороного.

Алгоритм поиска соседних СИМВОЛОВ

62

- В качестве критерия принадлежности символов, заданных центрами масс $c = (c_x, c_y)$, и $p = (p_x, p_y)$, одному слову используются следующие условия:

- ▣ Символы принадлежат одной строке:

$$|p_y - c_y| < \alpha \min(md_C(p), md_C(c))$$

- ▣ Символы расположены достаточно близко друг к другу и принадлежат одному слову:

$$md_S(p, c) < \beta \min(md_C(p), md_C(c))$$

- ▣ где $\alpha = 1.3$, $\beta = 1.5$

Диаграммы Вороного для областей

63

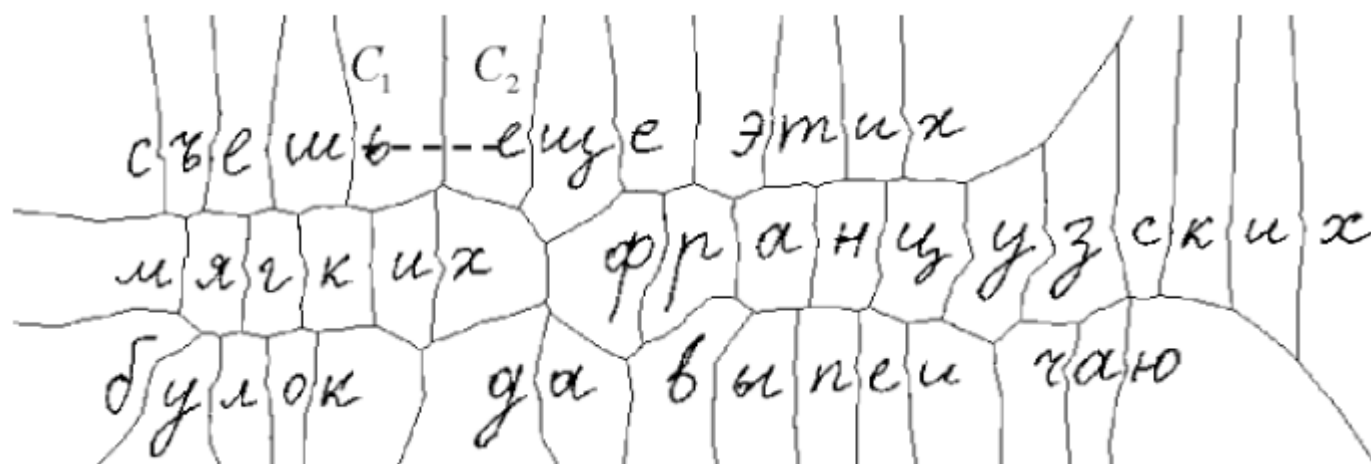
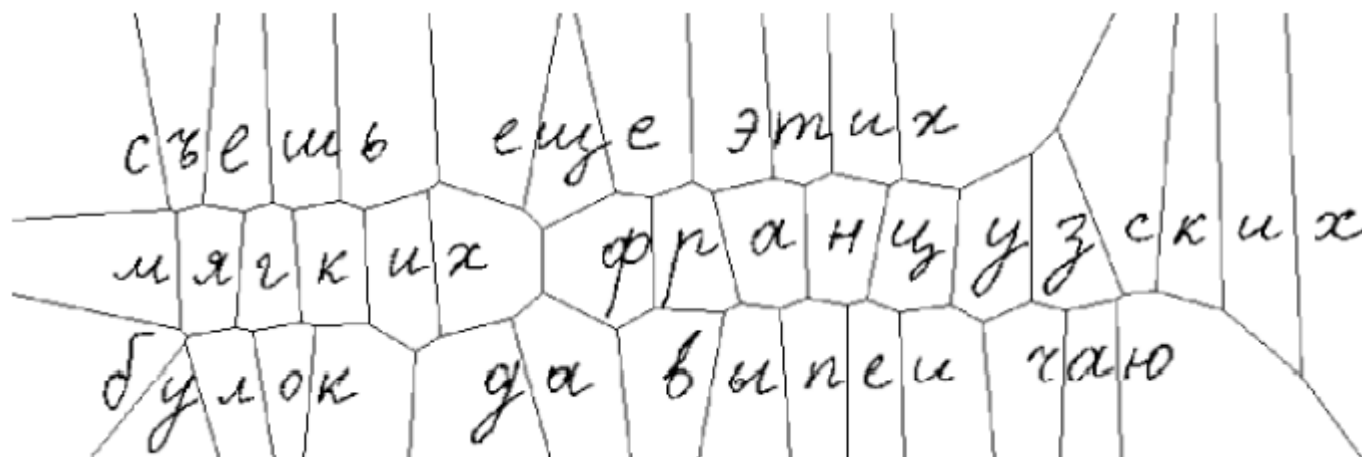


Диаграмма Вороного множества центров масс символов

64



Меры близости

Евклидово расстояние

Манхэттенское расстояние

Расстояние Махаланобиса

Косинусная мера

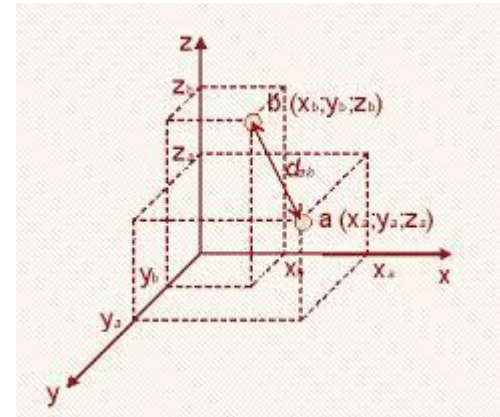
Редакционное расстояние Левенштейна

Евклидово и Манхеттенское расстояние

66

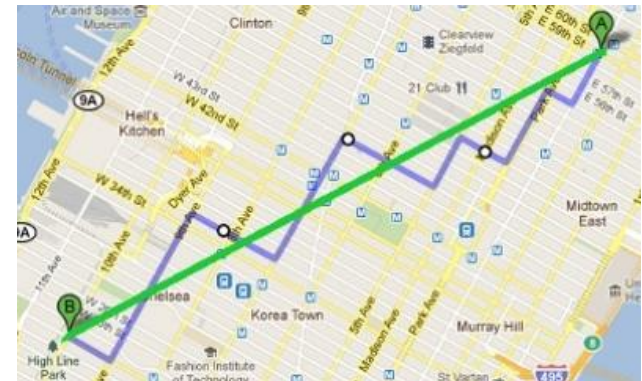
- Евклидова метрика – длина отрезка в n-мерном пространстве:

$$d(p, q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}$$



- Манхэттенское расстояние – расстояние городских кварталов:

$$d(p, q) = \sum_{k=1}^n |p_k - q_k|$$



Расстояние Махаланобиса

67

- Для двух векторов X , Y и матрицы ковариации S :

$$D(X, Y) = \sqrt{(X - Y)^T S^{-1} (X - Y)}$$

$$S = M \left[(X - M[X])(Y - M[Y])^T \right]$$

- С помощью расстояния Махаланобиса можно определять сходство неизвестной и известной выборки.
- Отличается от евклидового тем, что учитывает корреляции между переменными и инвариантно к масштабу.
- Если матрица ковариации является единичной матрицей, то расстояние Махаланобиса становится равным расстоянию Евклида.

Редакционное расстояние или расстояние Левенштейна

68

- Определяется как минимальное количество операций *вставки* одного символа, *удаления* одного символа и *замены* одного символа на другой, необходимых для превращения одной строки в другую.
- Применяется для сравнения строк (с произвольным алфавитом), которые могут не совпадать по длине. Например, в задачах исправления ошибок при поиске, распознавании, вводе текста.
- *Редакционным предписанием* называется последовательность действий, необходимых для получения из первой строки второй кратчайшим образом. Обычно действия обозначаются так: **I** (*insert*) - вставка, **D** (*delete*) - удаление, **R** (*replace*) — замена, **M** (*match*) — совпадение.
- Найти только расстояние Левенштейна — более простая задача, чем найти ещё и редакционное предписание.
- Операции вставки, удаления и замены могут иметь разную стоимость. Задача сведётся к нахождению совокупности операций с минимальной ценой.

Примеры

69

- $\text{levenshtein}('ABC', 'ABC') = 0$
 $\text{levenshtein}('ABC', 'ABCDEF') = 3$
 $\text{levenshtein}('ABC', 'BCDE') = 3$
 $\text{levenshtein}('BCDE', 'ABCDEF') = 2$

M	M	M	R	I	M	R	R
C	O	N	N		E	C	T
C	O	N	E	H	E	A	D

ACGTACGTAC-GT
| | | | | | | |
A-GTACCTACCGT

Алгоритм Вагнера-Фишера для поиска редакционного расстояния

70

- $D(0,0) = 0$
- для всех j от 1 до N
 - ▣ $D(0,j) = D(0,j-1) + \text{цена вставки символа } S2[j]$
- для всех i от 1 до M
 - ▣ $D(i,0) = D(i-1,0) + \text{цена удаления символа } S1[i]$
 - ▣ для всех j от 1 до N
 - $D(i,j) = \min\{$
 - $D(i-1, j) + \text{цена удаления символа } S1[i],$
 - $D(i, j-1) + \text{цена вставки символа } S2[j],$
 - $D(i-1, j-1) + \text{цена замены символа } S1[i] \text{ на символ } S2[j] \}$
- вернуть $D(M,N)$

Сравнение профилей на основе редакционного расстояния

71

- Вместо сравнения букв сравниваются числовые значения.
- Применяется адаптированный алгоритм Вагнера-Фишера с динамическим расчётом цен:
 - ▣ Цена вставки = цена удаления = `const`
 - ▣ Цена замены = модуль разности соответствующих значений профилей.
- Надо понимать, что профили должны быть построены для изображений одного размера, либо нормированы.

Косинусное сходство

72

- **Косинусное сходство** – мера сходства между векторами признаков, рассчитанная как косинус угла между векторами

$$d(p, q) = \cos(\theta) = \frac{pq}{\|p\| \|q\|} = \frac{\sum_{i=1}^n p_i q_i}{\sqrt{\sum_{i=1}^n (p_i)^2} \sqrt{\sum_{i=1}^n (q_i)^2}}$$

- Мера удобна в информационном поиске для разреженных векторов, т.к. учитываются только ненулевые компоненты

Мягкая косинусная мера

73

- Использует матрицу схожести признаков $s_{ij} = \text{сходство}(\text{признак}_i, \text{признак}_j)$
 - ▣ $s_{ii} = 1$,
 - ▣ $0 \leq s_{ij} \leq 1$
- Мера схожести признаков может быть расстоянием Левенштейна
- Равносильная косинусному сходству при $s_{ij} = 0$

$$d(p, q) = \frac{\sum_{i,j}^N s_{ij} p_i q_j}{\sqrt{\sum_{i,j}^N s_{ij} p_i p_j} \sqrt{\sum_{i,j}^N s_{ij} q_i q_j}}$$

Что почитать

74

- **Бондаренко А.В., Галактионов В.А., Горемычкин В.И., Ермаков А.В., Желтов С.Ю.** Исследование подходов к построению систем автоматического считывания символьной информации. Препринт ИПМ им. М.В.Келдыша РАН. Москва, 2003
- **Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В.** Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.
- **Местецкий Л.М.** Непрерывная морфология бинарных изображений. Фигуры. Скелеты. Циркуляры. – М.:Физматлит, 2009.
- **Дробков А.В., Семенов А.Б.** Обзор и анализ распознавателей рукопечатных символов // Математические методы распознавания образов: 15-я Всероссийская конференция, г.Петрозаводск, 11–17 сентября 2011 г.: Сборник докладов. М.: МАКС Пресс, 2011. С.350-353.

Дополнительная литература

75

- **Ballard D. H.** Generalizing the Hough Transform to detect arbitrary shapes // Pattern Recognition Vol. 13, No. 2, pp. 111-122. 1981
- **Duda R.O., Hart P.E.** Use of the hough transformation to detect lines and curves in pictures // Comm. ACM, Vol 15, No. 1, pp. 11-15 (January 1972).
- **С. А. Запрягаев, А. И. Сорокин** Сегментация рукописных и машинописных текстов методом диаграмм Вороного // Вестник ВГУ, серия: Системный анализ и информационные технологии, 2010, № 1. С.160-165 <http://www.vestnik.vsu.ru/pdf/analiz/2010/01/2010-01-27.pdf>
- **Самодумкин, Степанова, Колб** Практикум по компьютерной графике. Минск: БГУИР, 2014
https://libeldoc.bsuir.by/bitstream/123456789/1002/2/Samodumkin_Ch3.pdf
- **Захаркин И.** Диаграмма Вороного и её применения <https://habr.com/ru/post/309252/>
- **D. Comaniciu, P. Meer** Mean Shift: A Robust Approach Toward Feature Space Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, pp. 603–619.
- Обзор алгоритмов сегментации. Блог компании Intel
<https://habr.com/ru/company/intel/blog/266347/>
- **Zhe Wang, Yue Lu, Chew Lim Tan** - Word Extraction Using Area Voronoi Diagram
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.109.970&rep=rep1&type=pdf>
- Моменты изображения https://en.wikipedia.org/wiki/Image_moment