

ЛЕКЦИЯ 5. ВЫДЕЛЕНИЕ И АНАЛИЗ ПРИЗНАКОВ

Демидов Д.В.

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

План лекции

2

- Выделение признаков
- Вычисление моментов
- Анализ профилей, сегментация
- Фазовые пространства
- Утончение изображений
- Меры близости

3

Признаки для ч/б изображений

Скалярные признаки. Веса.

Вес чёрного и белого, доля чёрного и белого

4

- Для двоичной функции яркости $f(x, y) \in \{0, 1\}$

- Вес белого:

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

- Доля белого (удельный вес):

$$weight_{rel}^{f=1} = \frac{1}{M \cdot N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

- Доля чёрного:

$$weight_{rel}^{f=0} = 1 - weight_{rel}^{f=1}$$

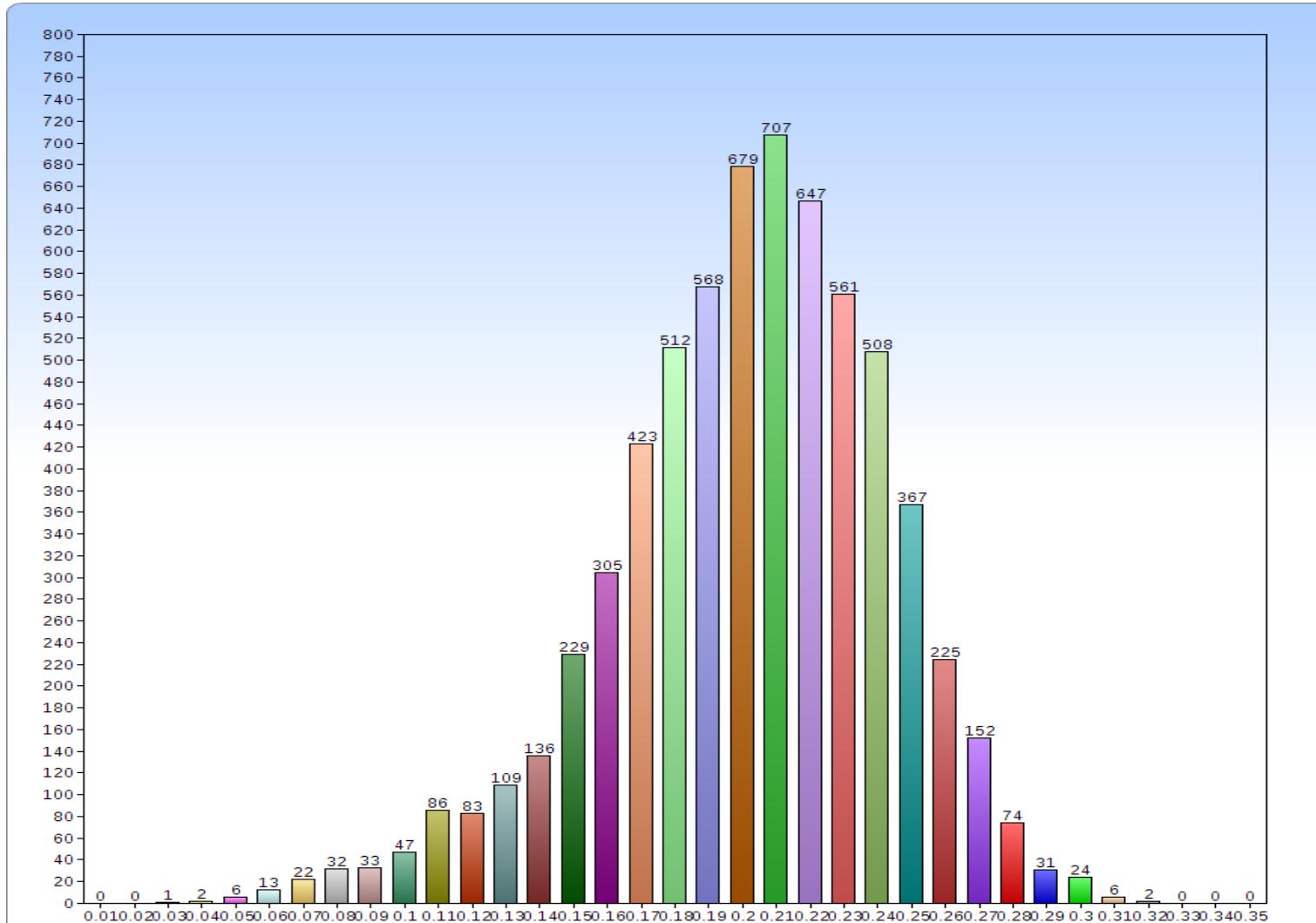
- Самый простой интегральный признак.

- 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. $k \in [1..8]$
- Можно посчитать удельный вес пикселей, имеющих $k_1 \leq k \leq k_2$ соседей того же цвета.
- 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 фрагментированости 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 креcтики)

1	0	1
0	1	0
1	0	1

□ Доля косых белых крестиков (очень похожи на чёрные креcтики)

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.

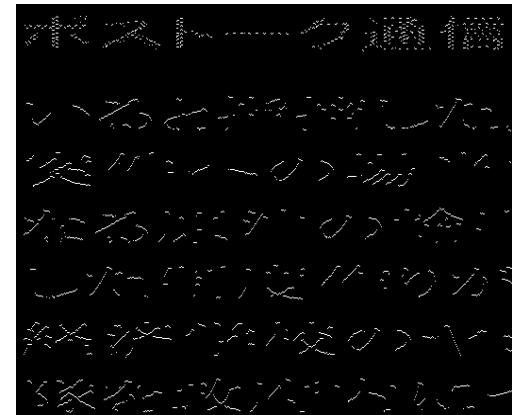
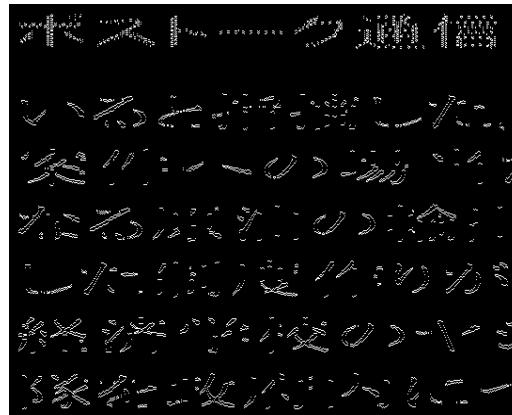
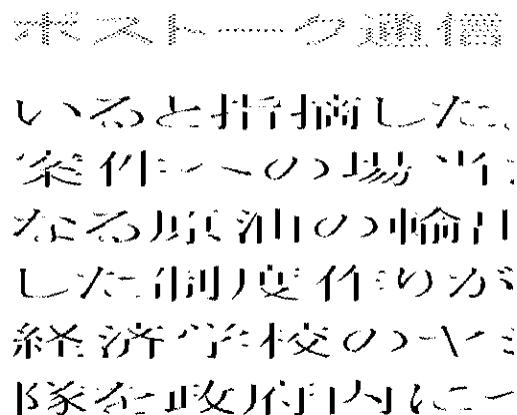
Примеры преобразований для оценки шума

10

□ Вычислим матрицы:

- $\text{Noise} = A] \text{xor } [A$
- $T = A] \text{ or } [A$
- $R = (\text{not Noise}) \text{ or } (\text{not dilate (not T)})$
- $\text{Noise2} = A] \text{xor (T and R)}$
- $\text{Noise3} = \text{Noise xor (not R)}$

□ Доли белого в Noise, Noise2, Noise3 по-разному говорят об уровне шума в A.



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:

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

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:

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

где x и y с чертой — координаты центра тяжести.

- Нормированные центральные моменты получаются в результате деления центральных моментов на моменты нулевого порядка.
- Центральные и нормированные моменты предпочтительнее из-за их большей инвариантности к преобразованиям изображений: сдвигу, повороту, масштабированию, отражению и др.

Вес на основе моментов

14

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

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

- Удельный вес соответствует нулевому моменту отнесённому к площади изображения:

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

Центр тяжести (центр масс)

15

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

$$\bar{x} = \frac{1}{weight} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x \cdot f(x, y) \quad \bar{y} = \frac{1}{weight} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} y \cdot f(x, y)$$

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

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

Оевые моменты инерции

16

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

$$I\bar{x} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (y - \bar{y})^2 \cdot f(x, y) \quad I\bar{x}_{rel} = \frac{I\bar{x}}{M^2 N^2}$$

$$I\bar{y} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^2 \cdot f(x, y) \quad I\bar{y}_{rel} = \frac{I\bar{y}}{M^2 N^2}$$

$$Icm45 = \frac{1}{2} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (y - \bar{y} - x + \bar{x})^2 \cdot f(x, y) \quad Icm45_{rel} = \frac{Icm45}{M^2 N^2}$$

$$Icm135 = \frac{1}{2} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (y - \bar{y} + x - \bar{x})^2 \cdot f(x, y) \quad Icm135_{rel} = \frac{Icm135}{M^2 N^2}$$

Моменты высших порядков

17

- На основе моментов можно вычислить такие признаки как:

- эллиптические параметры: наклон главной оси a и эксцентриситет (соотношение малой и главной полуосей);

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

- скошенность (асимметрию) как нормированный момент 3-го порядка;

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

- коэффициент эксцесса (степень островершинности, крутизны) на основе нормированного момента 4-го порядка;

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

- моменты более высоких порядков.

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 горизонтали 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:

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

Пример для японского

- *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

Крупинки тонера
сыпались по горизонтали влево —
горизонтальный профиль

Крупинки тонера
сыпались вертикально вниз —
вертикальный профиль

1. 概要

I. 沿革

北緯約 60 度、ネヴァ川といく筋もの運河のはしる「水の都」サンクトペテルブルクは、帝政ロシアの首都として、また大革命勃発の地として歴史的に有名ですが、今もなお人口 450 万余を擁するロシア第三の都市であり、行政、経済、学術文化的中心地となっています。

「サンクト」(聖なる)の意)ぬきで、単に「ペテルブルク」と呼ぶこともあります。ピョートル一世(大帝)により 1703 年に建都。元々はフィン人漁民しか住まない沼地でしたが、近代ロシア国家の建設のためには海への出口と「西欧への窓」が不可欠と考えたピョートルは、大量的農奴を動員して、ここに新都市の建設を開始。厳しい気候や過重な労働で倒れる農奴は数知れず、そのため当地は「屍の上に築かれた都市」と呼ばれます。

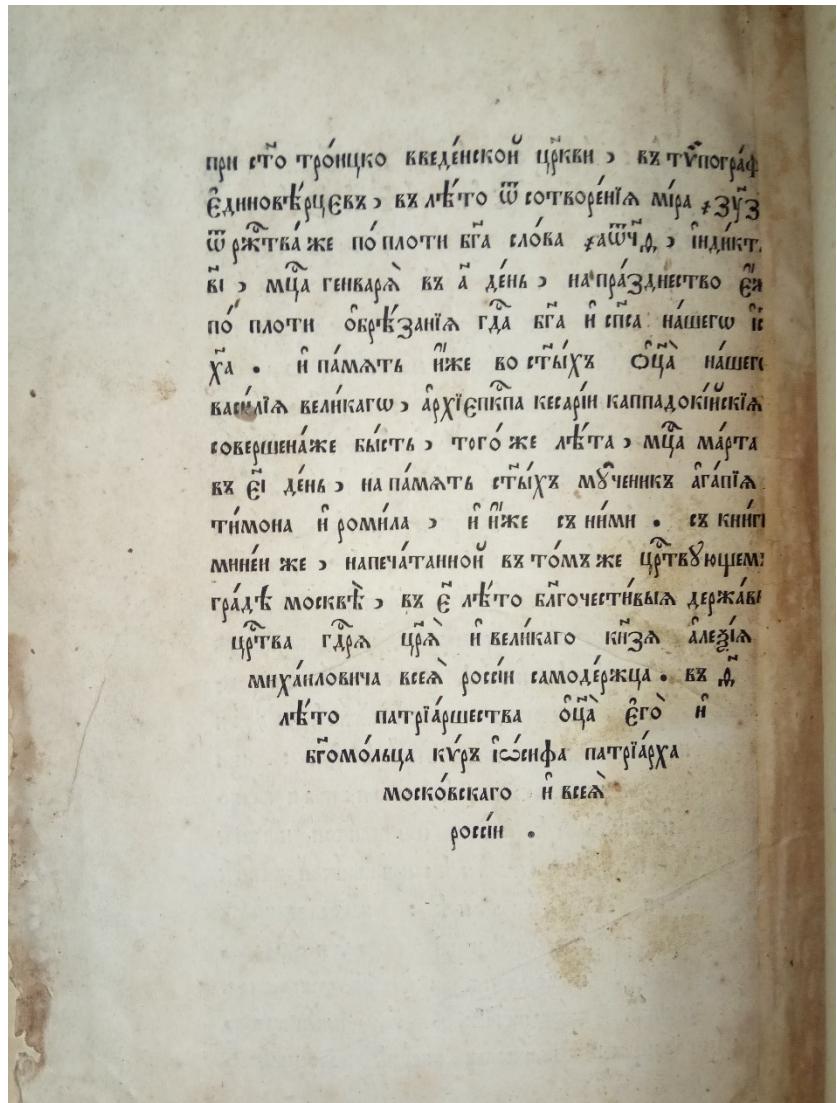
1712 年にはモスクワから当地に首都が移されます。以来、西欧各国から優秀な建築家を招いた結果、ロシアには珍しくヨーロッパ的雰囲気の漂う都になりました。当市の生みの親がピョートルなら、中興の祖は女帝エカテリーナ一世です。ピョートルが理想としたのはオランダ風、建築様式でいえば奇想にみちたバロック様式ですが、エカテリーナは一転、パリをモデルに優美なクラシック様式の街づくりを行いました。彼らの事業をひきつぎ、都市整備の仕上げをしたのは、19世紀前半のニコライ一世です。厳しい專制政治で評判の悪い君主ではありますが、当市建設にかけては相当の功績がありました。彼の性格を反映してか、イサク教堂やエルミタージュ宮殿新館など、この時代の建物は豪華な中に獨特の厳しい面立ちをしています。

約 2 世紀にわたりロシアの首都であった当市も、革命の煙草、1918 年 3 月にその座をモスクワに譲りました。国境への近さから、ソビエト新政権が懸念を抱いたためといわれます。第二次大戦では 900 日もの間ドイツ軍に包囲され(1941~1943 年)、記録的な寒波もあって一般市民にも莫大な犠牲者、病死者を出しました。砲撃や空爆により市街もなかなか焼け野原になりましたが、戦後に復旧がなされ、美しい歴史都市としてよみがえっています。現在、市の中心部にある建築は、大半が 18~19 世紀のものです。

複雑かつ苦難に満ちた歴史をたどった当市は、市の名前自体、何度も改称されています。1914 年第一次大戦の勃発とともにドイツ風の「ペテルブルク」は、ロシア風に「ペトログラード」と改められました(どちらも意味は同じで、「ピョートル大帝の街」、「使徒ペテロの守護する街」、「石の街」といった意味が掛け合われています)。革命を経て、1924 年にレーニンが亡くなると「レニングラード」(レーニンの街)と改名。ソ連邦の解体前夜、1991 年 9 月には、住民投票によって再びペテルブルクとなりました。

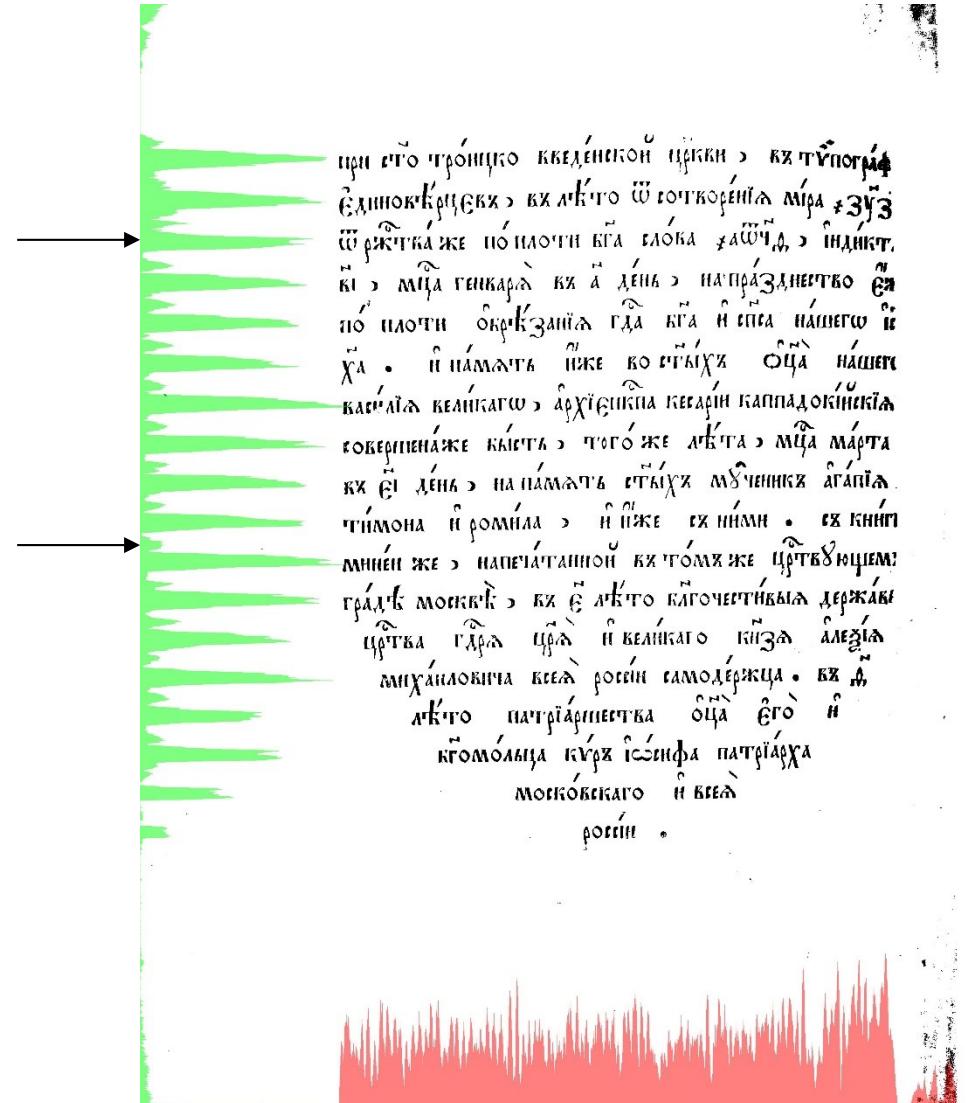
市内には文豪ブーシキン、ドストエフスキイ、ネ克拉ーソフ、ソシシェンコ、アフマトワなどの巨匠も保存されており、ロシア文学爱好者は必見。また、2008 年 5 月に就任し

Пример для славянского



Большие пики
образованы
символами в
строках

Малые пики
образованы
надстрочными
символами



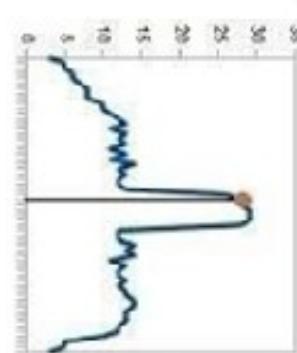
Профили (3)

22

- 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

23



A



Ж



Снизу вертикальный, слева горизонтальный профиль.

Косые профили

24

- 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

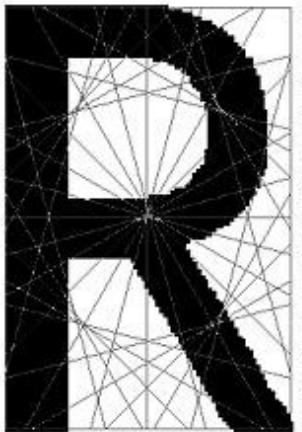
Метод пересечений

26

- 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 (терпение, друзья).

Метод пересечений. Пример

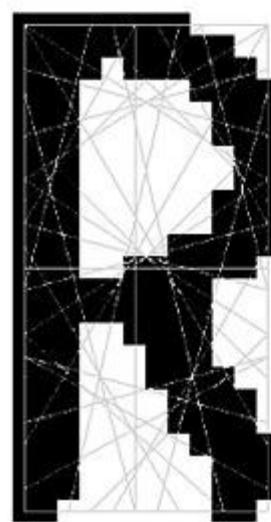
27



^: 0.041
0: 0.078
1: -0.177
2: 0.203
3: -0.071
4: -0.030
5: 0.015
6: -0.096
7: 0.084
8: 0.128
9: 0.129
A: 0.021
B: 0.463
C: 0.104
D: 0.331
E: 0.568
F: 0.673
G: 0.004
H: 0.315
I: -0.096
J: -0.288
K: 0.455
L: 0.324
M: 0.197
N: 0.204
O: 0.076
P: 0.755
Q: 0.269
R: 100.000
S: 0.050
T: -0.046
U: -0.026
V: -0.063
W: 0.084
X: 0.086
Y: -0.085
Z: 0.108

MINdistance: 99.245

Эталонное изображение



^: -0.074
0: 0.198
1: 0.032
2: 0.128
3: 0.162
4: -0.024
5: 0.048
6: -0.088
7: -0.179
8: 0.004
9: 0.174
A: -0.071
B: 0.279
C: 0.098
D: 0.280
E: 0.326
F: 0.442
G: 0.081
H: 0.340
I: -0.271
J: -0.124
K: 0.213
L: 0.219
M: 0.247
N: 0.306
O: 0.127
P: 0.530
Q: 0.122
R: 0.540
S: 0.027
T: -0.233
U: 0.179
V: -0.154
W: 0.130
X: -0.098
Y: -0.289
Z: -0.125

MINdistance: 0.010

Вектора близости с
другими образцами

Реальное изображение

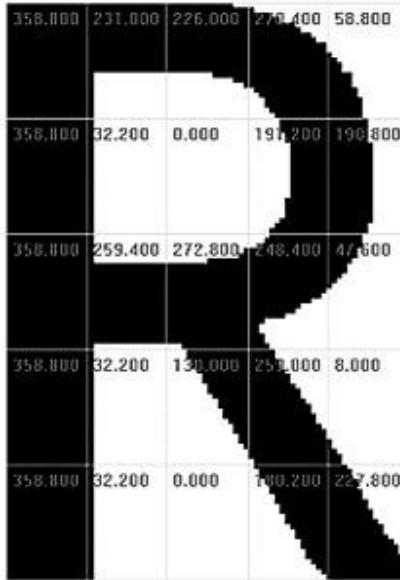
Метод зон

28

- 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

29



Эталонное изображение

0 : 0.193
1 : 0.239
2 : 0.172
3 : 0.217
4 : 0.251
5 : 0.188
6 : 0.211
7 : 0.269
8 : 0.157
9 : 0.181
A : 0.200
B : 0.115
C : 0.194
D : 0.154
E : 0.117
F : 0.131
G : 0.188
H : 0.180
I : 0.239
J : 0.276
K : 0.151
L : 0.227
M : 0.153
N : 0.169
O : 0.205
P : 0.115
Q : 0.140
R : 0.000
S : 0.164
T : 0.320
U : 0.234
V : 0.241
W : 0.160
X : 0.205
Y : 0.298
Z : 0.206
_ : 0.266

Вектора близости с
другими образцами



Реальное изображение

0 : 0.162
1 : 0.204
2 : 0.156
3 : 0.208
4 : 0.240
5 : 0.174
6 : 0.201
7 : 0.265
8 : 0.153
9 : 0.156
A : 0.194
B : 0.109
C : 0.186
D : 0.147
E : 0.124
F : 0.141
G : 0.166
H : 0.148
I : 0.236
J : 0.237
K : 0.166
L : 0.235
M : 0.131
N : 0.128
O : 0.181
P : 0.115
Q : 0.137
R : 0.067
S : 0.153
T : 0.317
U : 0.199
V : 0.221
W : 0.145
X : 0.199
Y : 0.299
Z : 0.212
_ : 0.256

Признаки для ч/б изображений

п-мерные признаки. Анализ фазовых пространств

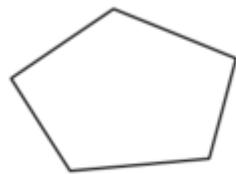
Фазовое пространство (r, φ)

31

- Пусть
 - r — расстояние прямой от центра координат
 - φ — угол наклона прямой
 - $A(r, \varphi)$ — аккумуляторная функция (аккумулятор «голосов»), вычисляемая как количество точек заданной яркости на прямой (r, φ) .
- Локальные максимумы $A(r, \varphi)$ соответствуют прямым, на которых лежит наибольшее число точек.
- Фазовое пространство строится с помощью преобразования Хафа (Hough Transform).

Пример пространства Хафа

32



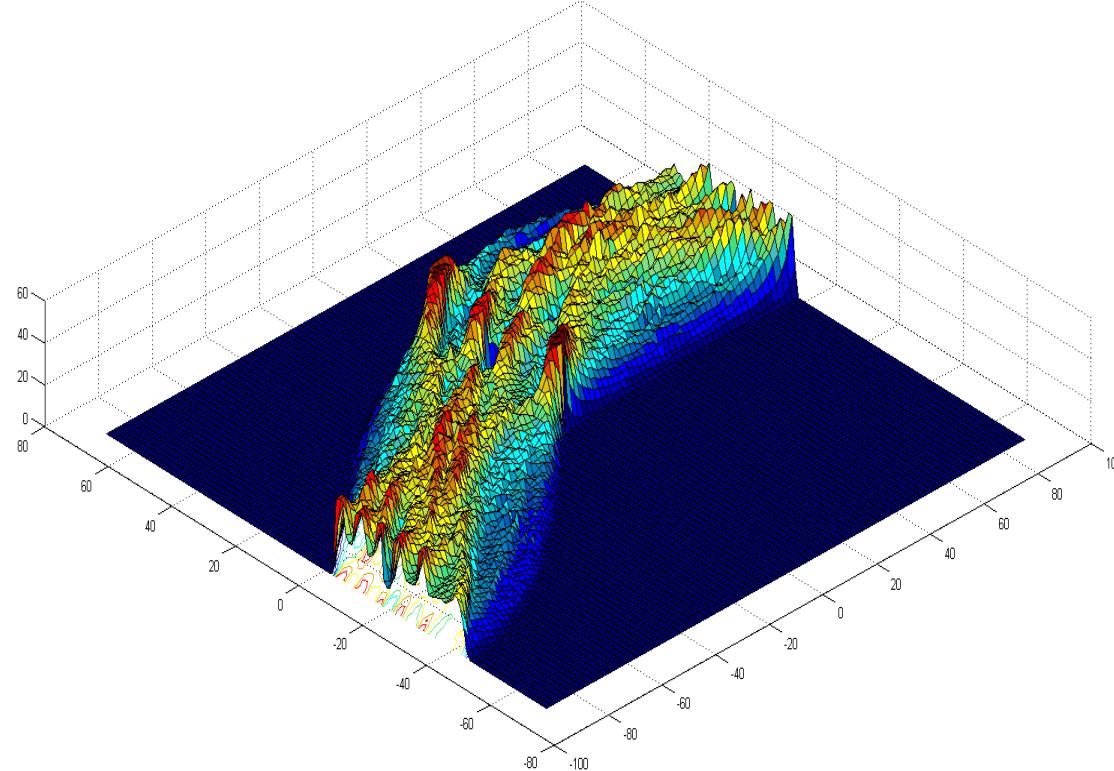
Изображение
пятиугольника



2D-визуализация в
координатах (R, ϕ)

Пример пространства Хафа

33



В виде 3D-карты высот

Выявление таблиц с помощью преобразования Хафа

34

ポストカード

ペドモスチ紙 『ВЕДОМОСТИ』	
英Financial Times、米The Wall Street Journalと提携するロシアの有力経済紙。	
10月1日	「エネルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的」
10月2日	「旧ソ連向け投資ファンド運営企業Baring Vostok、ロシア最大の外資系ファンドを設立」
10月3日	「グレンコア子会社がタマニ港の穀物ターミナル株式を取得」
10月4日	「モスクワの市立病院にコンセッション方式を導入」
10月5日	「米国製電子部品の密輸出にスパイ容疑がかけられる」

エネルギー省の依然を受け、前回も市場の管理組織である市場会議が再生エネルギー部門への投資効果を分析した。全発電量に占める可変合意出力は19%だったが、同省の目標は2020年までに50%にすることだ。来年まで2020年比で同じ部門への6900億ルーブルを投じる国家計画がある。これで目標達成にかかる年までに計1100万吉ルーブルの発電設備の建設が必要だが、小水力を除く既存設備の平均建設費は10万ルーブル／キロワットで最大でも6万5000ルーブルの火力発電炉の初期高さとなる。されば既存設備には從来型重油炉が準備充電して必要な、既存の計画は採算が合っていないのが市場会議前多岐の専門家の結論だ。一方で、2月に發表されたロシアの会員企業によるハイブリッド太陽光発電所の建設（本誌93号に掲載記事）、既に施工用具がRIS事業もある。

旧ソ連向け投資ファンド運営企業Baring Vostok、外資系でロシア最大となる渋川資金15億円のBaring Vostok Private Equity Fund LP.V.を設立した。Baring Vostokは米投資銀行のバンク・シティ出身のCalvary氏が設立者の一人となり、横田大手ヤング・クックス等に投資している。新ファンドは中国、中東、アジアから新たに10の投資家が参加し、既にボストン・マッケンゼン銀行等に投資を開始した。Calvary氏の案では、年4%の経済成長率を基点にシナフの投資能力が大きくなる外国投資家が多い。特に投資先として有望なのは石油精製、医療、自動車、IT分野だ。一方、政府系のロシア直接投資基金は渋川資金を20倍（か）持続的に100億円に引き上げる予定だ。

ロシアの企業グループ「Kornit」が所有していた黒海タマニ港穀物ターミナルの株式が、ウクライナの「Kernel Holding」とスイスの穀物商社グレンコアの子会社Renault BVが競争買収する合意を企業に2億6500万ドルで見附された。Kornitの9月27日付登表によると、合意企業は穀物ターミナルの100%株式を取得した。ターミナルの積荷精査能力は年間300万台で、黒海に面したシナフの大水深岸壁ターミナルで年144箱日に大さくベターブル港の設計処理能力は年間900万台だ。現在は16隻の第1支線船の稼働率で40%前後で、ターミナルの積荷シェアを占めるタマニ港だが、ロシアイスク（1200万台）やトルコアゼル（300万台）といった他港と競り、現直支線が競合化していることがシナフ最大のネック。しかし、穀物輸出が注目を集め、前回の出資可能な額はインフレを見込まれるのに極めて懸念しい。（本誌96号に掲載記事）

モスクワの市立病院にコンセッション方式を導入

モスクワ州のモトニコフ湖市長は社会開発相閣僚のミハイル・シクリンの訪問中の挨拶で、市立病院の施設をコンセッション方式で運営を進めていることを明かにした。市内97箇所の市立病院うち10箇所の病院がコンセッションで運営されるのは市立だが、湖市長によれば、第63市立病院はEU民間医療機関「欧洲医療センター（OMC）」が入り参加を表明しているといつ。実際2010年に市立病院の約40%は強制加入保険（OMS）の保険証を持った患者や被扶養者で運営されているが、そのうち約30%は任意加入保険（MS）の保険証を持つ患者や、自費負担で会計を行なう患者にてて行われる。コンセッションの外算にヨリアルファ保険などだ。一部保険会社も関心を示している。医療衛生のコンセッションは、ノルヒビルク州ベータクルスタン共和国で既に実施されている。

米国製電子部品の密輸出にスパイ容疑がかけられる

小型電子部品を米国から不正に輸出したとして、ロシア有名モードルの容疑者が米司法当局に起訴された。テクサス州のアーヴィング市に本拠地の輸出業者、販売元は日本機械や中華機器の一つであると申請しているが、取引先であるシナフのアルセナル社は、米国製電子部品の納入企業として国防省からの認証を受けている企業だった。米議会EPA・タバコと取引のあったE65の企業名・個人名を公表。リストに上位に載った米国製電子部品の輸出業者で、年貿易額で1億500万ドル超だ。インダーネットで輸入部品の仕入れを受ける中堅企業がいたのが特徴だ。自國企業の資本が持っているシナフでは、米国に在する元ロシア人で高級人材を買付けたところした企業はいくつもあり、リスクの少ない企業がござるみればビジネスで、産業スパイ活動も、ついに上程跡が見えてきた。

ペドモスチ紙 『ВЕДОМОСТИ』	
英Financial Times、米The Wall Street Journalと提携するロシアの有力経済紙。	
10月1日	「エネルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的」
10月2日	「エカルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的」
10月3日	「旧ソ連向け投資ファンド運営企業Baring Vostok、ロシア最大の外資系ファンドを設立」
10月4日	「グレンコア子会社がタマニ港の穀物ターミナル株式を取得」
10月5日	「モスクワの市立病院にコンセッション方式を導入」
10月6日	「米国製電子部品の密輸出にスパイ容疑がかけられる」

エネルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的

エネルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的

旧ソ連向け投資ファンド運営企業Baring Vostok、ロシア最大の外資系ファンドを設立

エカルギー省の再生エネルギー推進計画は採算取れず：専門家の予想は悲観的

グレンコア子会社がタマニ港の穀物ターミナル株式を取得

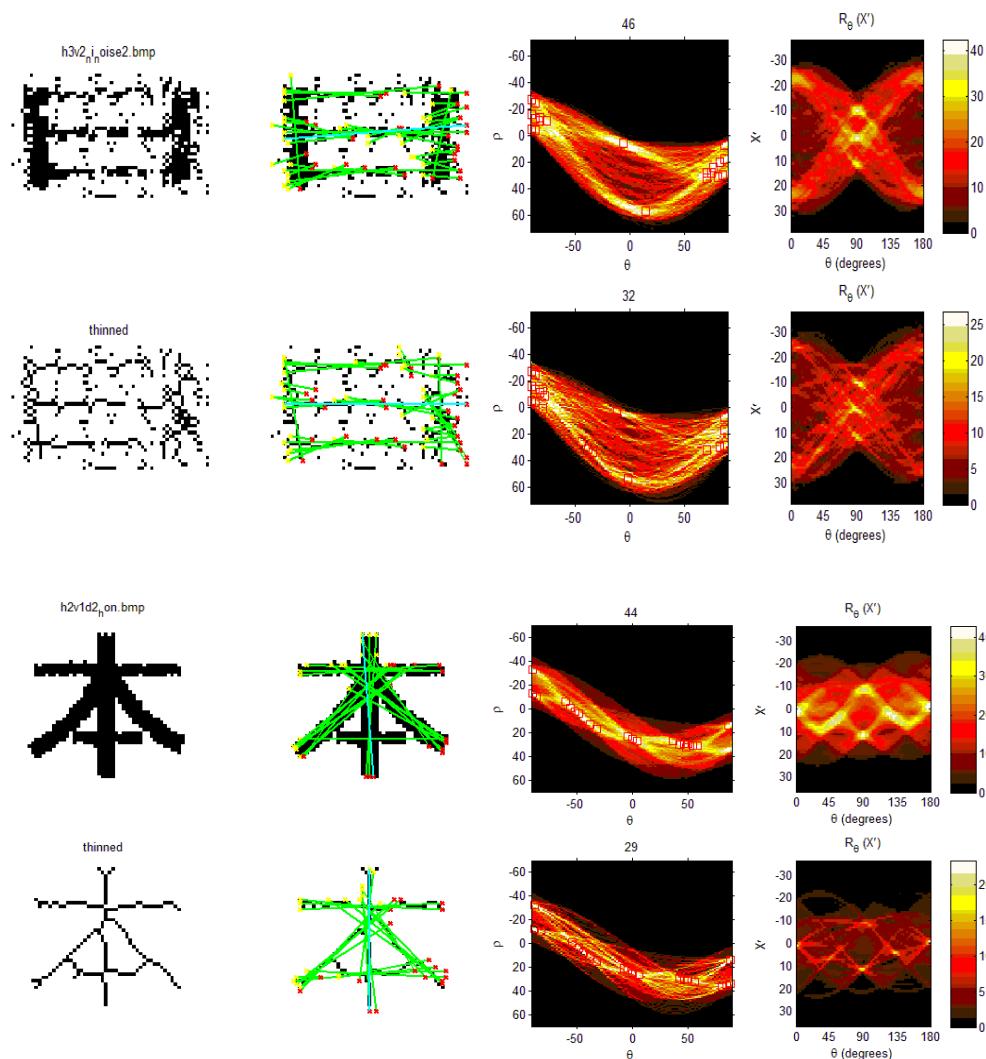
モスクワの市立病院にコンセッション方式を導入

米国製電子部品の密輸出にスパイ容疑がかけられる

Выявление прямых линий, аппроксимация прямыми линиями

35

- 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

36

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

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

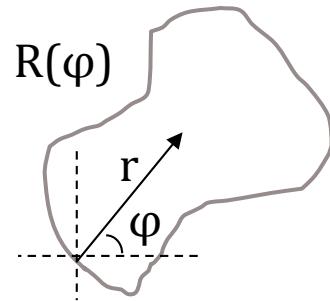
- Максимум аккумулятора соответствует положению центра окружности на изображении.
- Если радиус окружности заранее не задан, можно добавить R как третье измерение параметрического пространства $A(x, y, R)$.
- Если важны только сами центры, а радиусы не важны, то можно не увеличивать размерность пространства, 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

37

- GHT (Generalized Hough Transform) предложено Баллардом для обнаружения кривых линий заданного типа.
- Рассматривается расстояние R от контура фигуры до центра. В отличие от окружностей это не константа, а функция $R(\varphi)$ от угла поворота радиус-вектора, направленного от точки контура к центру.
- Для простых форм функция $R(\varphi)$ может быть описана аналитически.
- Для сложных форм функция задаётся с помощью таблицы поиска 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 предъявляемому изображению.



φ	$R(\varphi)$
0	45
10	42
20	39
30	39
40	43
50	38
...	
350	40

Преобразование Радона

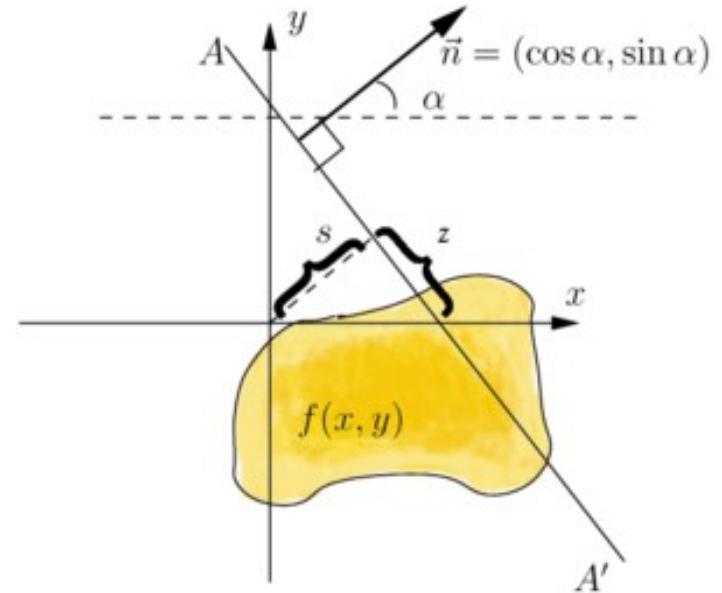
38

- 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

39

- 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

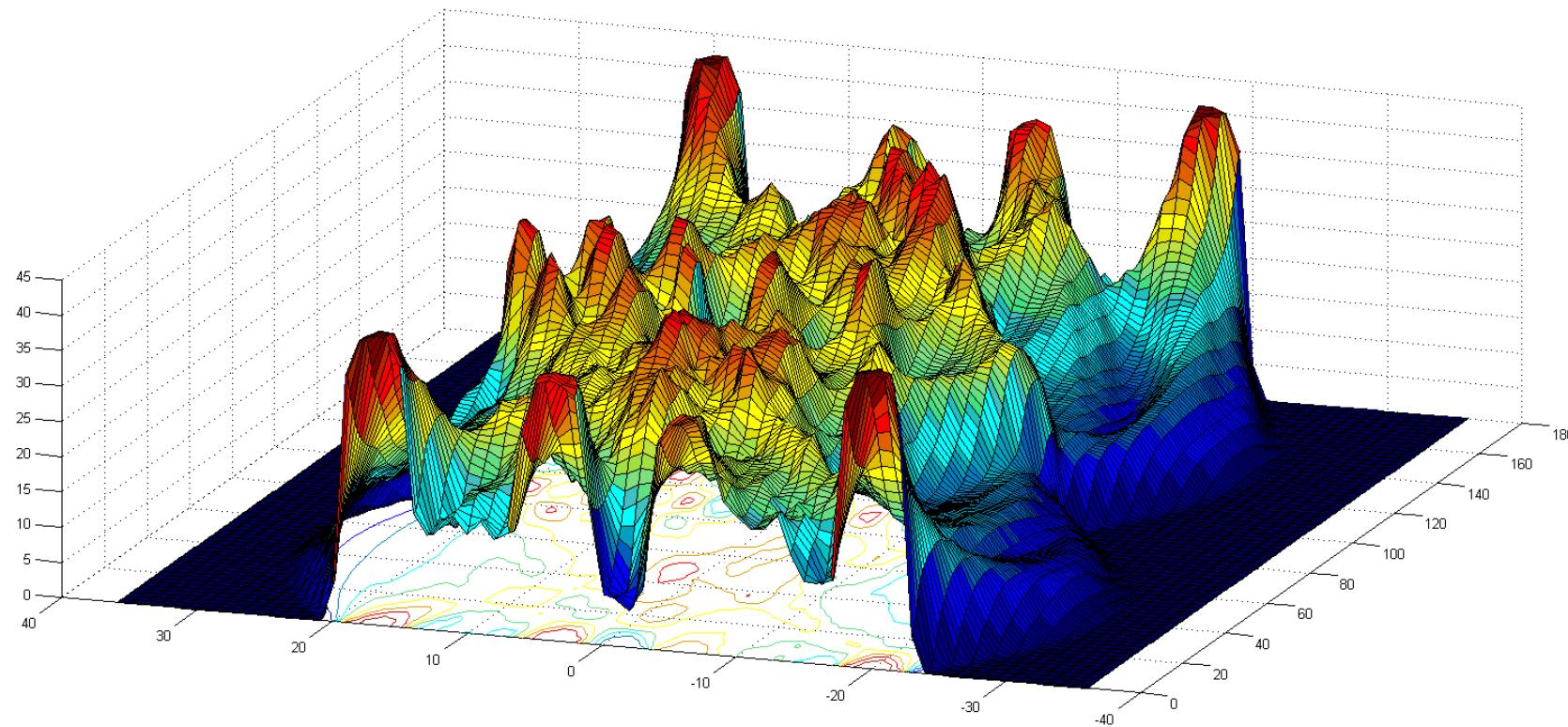
$$f(x, y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^\infty e^{i\omega(x\cos\alpha + y\sin\alpha)} \tilde{R}(\omega, \alpha) \omega d\omega d\alpha$$

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

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

Пример карты высот

40



Сегментация символов пороговой обработкой

Метод профилей

Алгоритм выделения текстовой области

42

- 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

43

- 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

44

- 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.

Меры близости

Евклидово и Манхэттенское расстояние

Расстояние Чебышёва и Махalanобиса

Косинусное сходство

Расстояние на основе корреляции

Редакционное расстояние Левенштейна

Евклидово и Манхэттенское расстояние

46

- Евклидова норма — длина отрезка в n-мерном пространстве:

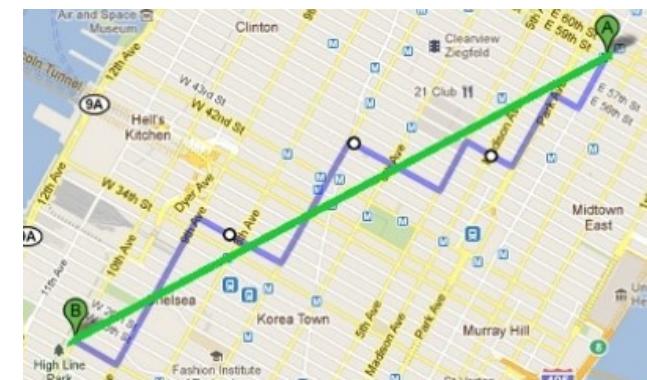
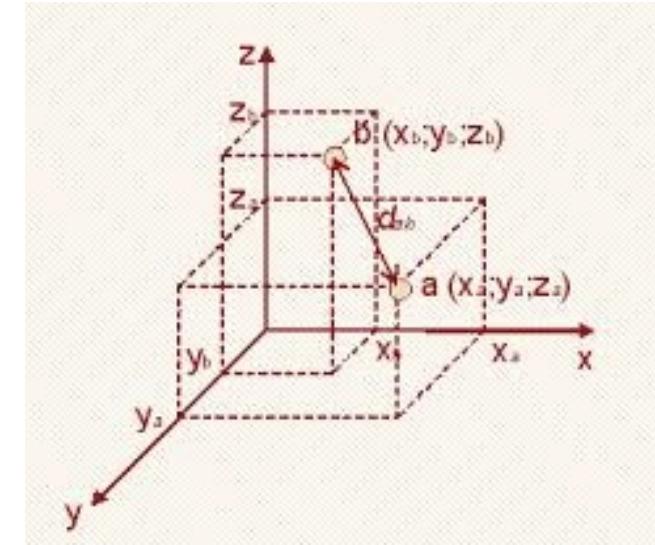
$$d(p, q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}$$

- Манхэттенская норма (расстояние городских кварталов) — сумма смещений по осям координат:

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

- Норма Чебышёва — максимальное из смещений по осям координат:

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



Расстояния по сетке координат

47

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

Манхеттенская норма

$\sqrt{18}$	$\sqrt{13}$	$\sqrt{10}$	3	$\sqrt{10}$	$\sqrt{13}$	$\sqrt{18}$
$\sqrt{13}$	$\sqrt{8}$	$\sqrt{5}$	2	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{13}$
$\sqrt{10}$	$\sqrt{5}$	$\sqrt{2}$	1	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{10}$
3	2	1	0	1	2	3
$\sqrt{10}$	$\sqrt{5}$	$\sqrt{2}$	1	$\sqrt{2}$	$\sqrt{5}$	$\sqrt{10}$
$\sqrt{13}$	$\sqrt{8}$	$\sqrt{5}$	2	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{13}$
$\sqrt{18}$	$\sqrt{13}$	$\sqrt{10}$	3	$\sqrt{10}$	$\sqrt{13}$	$\sqrt{18}$

Евклидова норма

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

Норма Чебышёва

- В целях визуализации значения норм округлены до большего целого

Расстояние Махalanобиса

48

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

$$\text{cov}(X, Y) = M \left[(X - M[X])(Y - M[Y]) \right]$$

- Расстояние Махалонобиса:

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

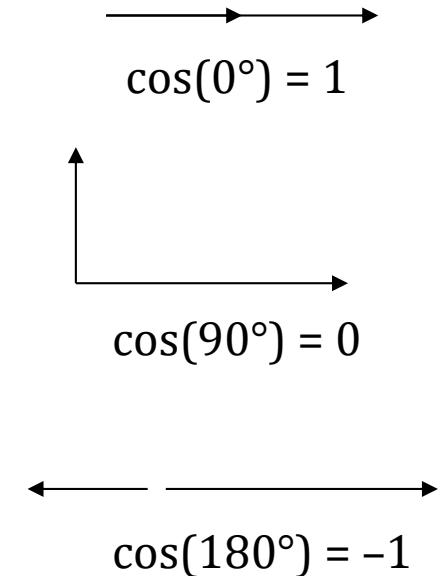
- Учитывает корреляции между переменными и инвариантно к масштабу. При единичной матрице ковариации совпадает с расстоянием Евклида.
- С помощью расстояния Махаланобиса можно определять сходство неизвестной и известной выборки.

Косинусное сходство

49

- **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. Кроме того, она нормирована и имеет область значений $[-1, +1]$.

Мягкая косинусная мера

50

- Использует матрицу схожести признаков s_{ij} = сходство(признак_i, признак_j)
 - $s_{ii} = 1$,
 - $0 \leq s_{ij} \leq 1$
- Мера схожести признаков может быть расстоянием Левенштейна
- Равносильна косинусному сходству при $s_{ij} = 0$ для $i \neq j$

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

Расстояние на основе корреляции

51

- Корреляция векторов r_{pq} размерности n :

$$r_{pq} = \frac{\sum_{k=1}^n (p_k - \bar{p})(q_k - \bar{q})}{\sqrt{\sum_{k=1}^n (p_k - \bar{p})^2 \sum_{k=1}^n (q_k - \bar{q})^2}}$$
$$\bar{p} = \frac{1}{n} \sum_{k=1}^n p_k$$
$$\bar{q} = \frac{1}{n} \sum_{k=1}^n q_k$$

- Значения r_{pq} лежат в отрезке $[0;1]$. Нулевая корреляция соответствует максимальному различию, а 1 — максимальному сходству.
- Расстояние на основе корреляции:

$$d(p, q) = 1 - r_{pq}$$

- Значения d также лежат в отрезке $[0;1]$.

Редакционное расстояние или расстояние Левенштейна

52

- Определяется как минимальное количество операций *вставки, удаления 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 сводится к нахождению совокупности операций с минимальной ценой.

Примеры

53

- `levenshtein('ABC', 'ABC') = 0`
- `levenshtein('ABC', 'ABCDEF') = 3`
- `levenshtein('ABC', 'BCDE') = 3`
- `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-GTACCTAACCGT

Алгоритм Вагнера-Фишера для поиска редакционного расстояния

54

- $D(0, 0) = 0$
 - 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

55

- 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.

Что почитать

56

- **Бондаренко А.В., Галактионов В.А., Горемычkin 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., Осоксов 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.
- **Гонсалес Р., Вудс Р.** Цифровая обработка изображений. Издание 3-е, исправленное и дополненное. Москва: Техносфера, 2012. — 1104 с.
- **Левенштейн В.И.** Двоичные коды с исправлением выпадений, вставок и замещений символов // Доклады Академии Наук СССР, 1965. Т. 163, №4: Стр. 845–848.
- **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).
- Моменты изображения https://en.wikipedia.org/wiki/Image_moment
- Image features https://ibsi.readthedocs.io/en/latest/03_Image_features.html