

## ЛЕКЦИЯ 5. ВЫДЕЛЕНИЕ И АНАЛИЗ ПРИЗНАКОВ

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

Демидов Д.В.

### План лекции

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

# Признаки для ч/б изображений

Скалярные признаки. Веса.

### Вес чёрного, доля чёрного

□ Вес чёрного:

$$weight = \sum_{x} \sum_{y} f(x, y)$$

□ Удельный вес чёрного:

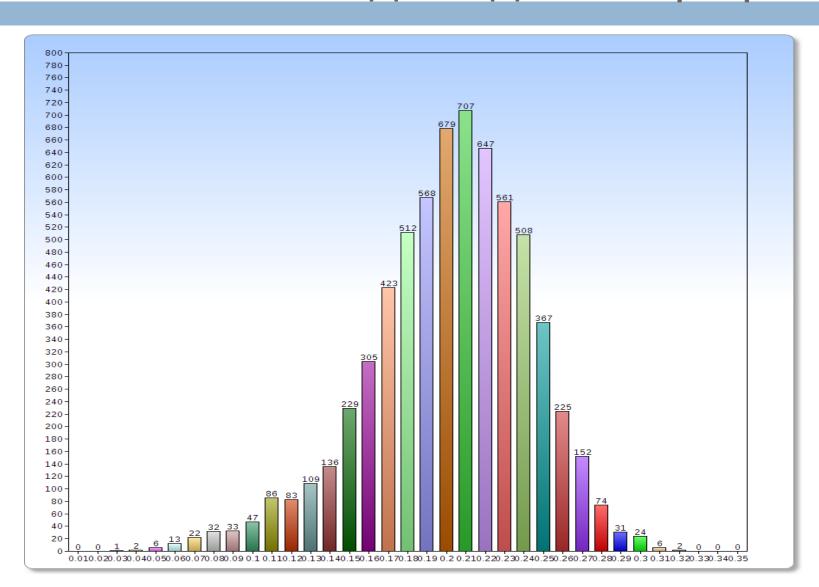
$$weight_{rel} = \frac{\sum_{x} \sum_{y} f(x, y)}{S}$$

□ 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



### Изолированные пиксели

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

1 0 1 1 1 1

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

 $\begin{array}{cccc} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}$ 

 Удельный вес изолированных пикселей может сказать о степени зашумлённости изображения.

# Доля пикселей с k соседями, k1-k2 соседями

- □ Можно посчитать удельный вес пикселей, имеющих k соседей того же цвета.
- □ Можно посчитать удельный вес пикселей, имеющих  $k1 \le k \le k2$  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

Доля чёрных крестиков	0 1 0
<u> Доличерных крестиков</u>	1 1 1
	0 1 0
<ul><li>Доля белых крестиков</li></ul>	1 0 1
	0 0 0
	1 0 1
<ul><li>Доля косых чёрных крестиков</li></ul>	1 0 1
(очень похожи на белые крестики)	0 1 0
	1 0 1
<ul><li>Доля косых белых крестиков</li></ul>	0 1 0
(очень похожи на чёрные	1 0 1
крестики)	0 1 0

□ Сумма долей, другие фигуры...

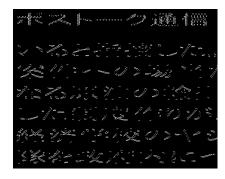
### Признаки для выявления шума

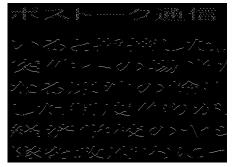
- Логические и морфологические
   преобразования изменяют изображение.
- Признаки, рассчитанные для изменённого изображения, несут другую информацию.
- Пусть А матрица яркостей исходного изображения. Обозначим через А] матрицу А без последнего столбца, а через [А матрицу А без первого столбца.

# Примеры преобразований для оценки шума

- Вычислим матрицы:
  - Noise = A] xor [A
  - T = A] or [A
  - R = (not Noise) or (not dilate (not T))
  - Noise2 = A] xor (T and R)
  - Noise3 = Noise xor (not R)
- □ Доли белого в Noise, Noise2, Noise3 по-разному говорят об уровне шума в А.

オストーク猟揺 いると指摘した。シペ 案件シーの場当、 タグシーの 湯 なる原色油の利倫目 グきるか した部別度作のが 条巻 沙客 当年本交 クラーマミニ 参巻 シターディング 事家を10女房自内にこっ







# Признаки для ч/б изображений

Скалярные признаки. Моменты

### Статистические моменты

- 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

- □ р, q = 0,1,...,∞ порядки момента;
- М 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

Центральные моменты являются функцией расстояния точки от центра тяжести символа:

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

где х и у с чертой – координаты центра тяжести.

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

### Вес на основе моментов

□ 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

□ Абсолютные координаты центра тяжести соответствуют парам p=1, q=0 для х и p=0, q=1 для у, отнесённым к нулевому моменту:

$$\overline{x} = \frac{\sum_{x}^{M} \sum_{y}^{N} x \cdot f(x, y)}{weight} \qquad \overline{y} = \frac{\sum_{x}^{M} \sum_{y}^{N} y \cdot f(x, y)}{weight}$$

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

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

### Осевые моменты инерции

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

$$I\overline{x} = \sum_{x}^{M} \sum_{y}^{N} (y - \overline{y})^{2} \cdot f(x, y) \qquad I\overline{x}_{rel} = \frac{I\overline{x}}{M^{2} + N^{2}}$$

$$I\overline{y} = \sum_{x}^{M} \sum_{y}^{N} (x - \overline{x})^{2} \cdot f(x, y) \qquad I\overline{y}_{rel} = \frac{I\overline{y}}{M^{2} + N^{2}}$$

$$Icm 45 = \frac{1}{2} \sum_{x}^{M} \sum_{y}^{N} (y - \overline{y} - x + \overline{x})^{2} \cdot f(x, y) \qquad Icm 45_{rel} = \frac{Icm 45}{M^{2} + N^{2}}$$

$$Icm 135 = \frac{1}{2} \sum_{x}^{M} \sum_{y}^{N} (y - \overline{y} + x - \overline{x})^{2} \cdot f(x, y) \qquad Icm 135_{rel} = \frac{Icm 135}{M^{2} + N^{2}}$$

### Применение моментов

- 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 а 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.

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

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

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

## Признаки для ч/б изображений

Векторные признаки. Профили

### Профили

- Профиль (проекция) представляет собой сумму яркостей пикселей, подсчитанную вдоль какоголибо 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)

- □ 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

#### 1. 概要

#### 1. 沿革

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、 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 るロシア第二の都市であり、行政、経済、学術文化の中心地となっています。 (「型なる」の意) ぬきで、単に「ペテルブルク」と呼ぶこともあります。 (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習地で 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 太量の農奴を動員して、ここに新都市の建設を開始。織しい気候や過重な 農奴は数知れず、そのため当地は「姥の上に築かれた都市」と呼ばれます。 スクワから当地に首都が移されます。以来、四欧各国から優秀な建築家を 「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 トルなら、中興の祖は女帝エカテリーナ二世です。ピョートルが理想としたの。 建築様式でいえば奇想にみちたパロック様式ですが、エカテリーナは一転、 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、 げをしたのは、19 世紀前半のニコライ一曲です。厳しい専制政治で評判の 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 →型堂やエルミタージュ宮殿新館など、この時代の建物は豪奢な中にも独特。

約 : 無純にわたりロシアの首都であった当市も、革命の物年、1918年 3 月にその原を再びモスクフに譲りました。同境への近さから、ソビエト新政権が懸念を抱いたためといわれます。第 : 次大機では900 日もの間ドイツ軍に包囲され (1941~1943年)、記録的な客報もあって 彼 中民にも 及大な職死者、病死者を出しました。 優繁や空爆により市街もなかば難け事時になりましたが、吸後に復旧がなされ、見しい歴史都市としてよみがえっています。 現代、市の中心部にある建築は、大半が 18~19 世紀のものです。

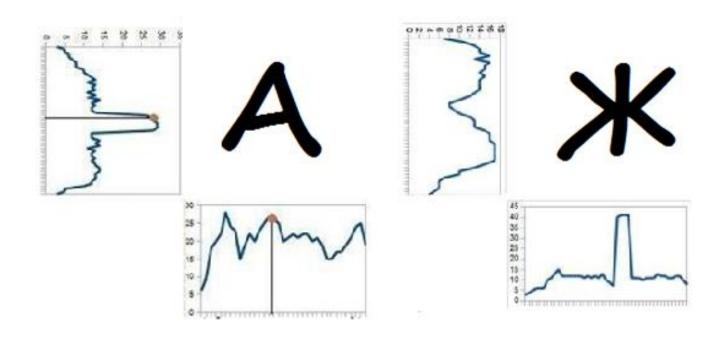
報序の一次繋に減ちた歴史をたどった当市は、市の名前自体、何度も改称されています。
1914年第一次大阪の発売とともにドイン風の「ペテルブルク」は、ロシア風に「ペトログ
フード」と改められました(どちらも意味は同じで、「ビョートル大帝の情」、「彼使ペテロの守護する情」、「不の情」といった意味が掛け合わされています)、半命を経て、1924年にレーニンがビぐなると「レニングラード」(レーニンの情)と改名。ソ連邦の解体前夜、1991年9月には、住民投票によって再びペテルブルクとなりました。

市内に対文豪プーシキン、ドストエフスキー、ネクラーソフ、ゾシチェンコ、アフマー トワなどのHIPも保存されており、ロシア文学委好家は必見。また、2008年5月に旋任し

# Профили (3)

- 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



Снизу вертикальный, слева горизонтальный профиль.

### Косые профили

- 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.

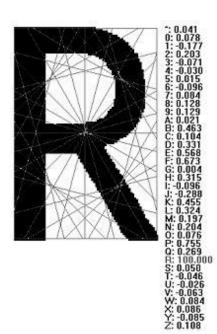
### Признаки для ч/б изображений

Векторные признаки. Метод пересечений и метод зон

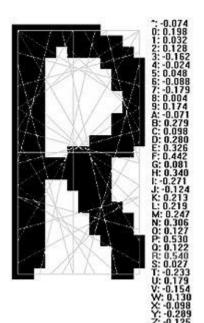
### Метод пересечений

- □ 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



MINdistance: 99.245



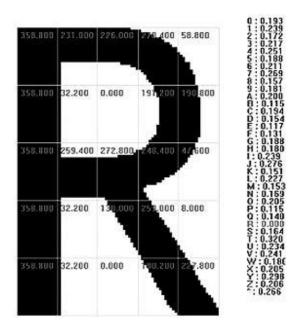
MINdistance: 0.010

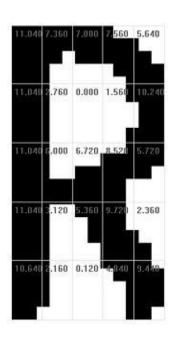
- □ Слева эталонное изображение
- □ Справа реальное изображение
- Рядом вектора близости с другими образцами

### Метод зон

- 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







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

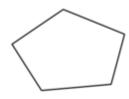
## Признаки для ч/б изображений

n-мерные признаки. Анализ фазовых пространств

## Фазовое пространство (r, φ)

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

### Пример пространства Хафа

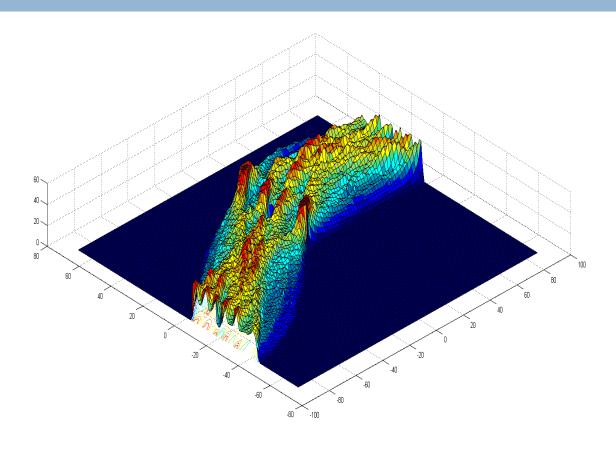


Изображение пятиугольника



2D-визуализация в координатах (R, phi)

# Пример пространства Хафа



В виде 3D-карты высот

# Выявление таблиц с помощью преобразования Хафа

ポストーク節位

#### ベドモスチ紙 《 BEAOMOCTU 》

英Financial Times。米The Wall Street Journalと提供するロシアの行力経済紙。

	ВЕДОНОСТИ
1033 111	「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」
	エキルベー等の依頼をつけ、部式方市場の管理組織である。市場会議「影中生・可能・キルベー (8193) 門への投資資源を分析した。全発派はこれの活の部門には発生等がよっす方が、同省の日標は2020年までは4.5%にすることで、東半から2020年までは63%に対した。大人を検集しる回来計画はある。この日標が改訂には同年までに計1100万キョウラーの発電協議の対策が42要差が42要差が42要差が42要差が10万ル・フルノキョウテー度を大で4.5%に10万ル・クルノキョウテー度を表して4.5%に10万ル・フルノキョウテー度を表して4.5%に10万ル・の火力発的が14円協会となる。ぎょは62条は13年で42年で43条によるいイブリテートによる統領を扱の専門家の結論だ。一方で、2月に発売は10万ル・フルテータの合発を表して4.5%に10万円で5.5
10/12/11	「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」
	■17世末間中でインの所後異性アンシーを運発するIlluring Vestokが、外質系ではロンテ酸人となる原用 資金目鏡は、のBuring Vestok Private Equity Found LP Ve表流では、Buring Vestokは未要機能行うのエ シープンザー・不用時のXulver更が返せ着の、人人な別、検索人、ドヤンデックス等に投資にている。新ファ ンドにはまれは、東直域、アジア等の構造には100異智量が参加、関に対なトディーのタスプンス執行等に 投資を開始した。Culveryでのように、中国の経済成長中を誇るロシアを投資魅力が大きいと見る外国 人投資室によりい、特に投資先生して有望なのに前費財、医療、自動車、自分野がたい。一方、政府系 のロロシアは接受性を作用用質金を制度、から様実的には100度。に対けておろうまだ。
1031311	「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」
	ロシアの企業グルー・パエフィ (新西看していた)無額 ケマニ 清穀物ケ・ミナルの株式が、ウケリイナの Record Tabiling by スイスの食物食料ケレットツの子会対 Remaison By (訪問 田野守する合き企業主義) 6500 が、で売用された。Remaisの9月27日付業 長によると、合金企業主義物ケ・ミナルの1409名株式を 取料した。ケ・ミナルの食物食料料金力は「田部300 か、で、里海に油によロシアの大水湿に埋ケ・モナル では187日に大きに火ケ・ご添の表計処理能力は対1間900 が、だが、現在はその5年1度繋の5段線が、 物輸出のロー・ク型は188%の積料シェアをよめるケ・ご、またが、フルエッチスの200 か、アルッティア( 200 か、ひといった他港と異なり、鉄道 女線が鉄液されていたいことがシェア鉱大のネック。しかし穀物輸出が 用が注目を集める中、他に出音可能な進出で、ファンフを見っまりるのは極めて繋した。(本記をおりに関連に (4)
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」
	モスタリ市のトラティトニコで南市長住た会業展開限制。第13シジオモスタリのただまの及談で、市立病 院の施設をエンセッション代制業務委託に11年半準備を進めていることを明まかました。市内男グ所の市 立高院の方式の新院がセンセッションに用されるかは未育だが、前市民に上れば、第63市立新院には 民間医療機関 度州医療センターのMC が入れる加を表明しているという。発後30年は立たの政権の定 されているという部の指定が、声自ははこことを、1900年クタトを支出される原面・新院業務の約 60%は強制加入保険(OMS)の保険証書を持った患者や、投急機造された患者の音楽に割り当でかれ、毎 りは任意加入保険(OMS)の保険証書を持った患者や、自費負担で会計を行う患者と当ていれる。コン セッションの会祭には11アルファ保険に居り、部保険会社に関心を示している。医療輸品のメロンセッショ ンは、ノボンにルスタ州やウタルスタン共和中で跳にている。
1011511	「米国製電子部品の密輸出にスパイ容疑がかけられる」
	小型電子が出る本来申から不正に輸出したとして、ロシア名をもつけるの容疑者が未可洗当局に出席された。シェナス州のアーク・ロレクレニックス社は輸出の際、販売先は信号機ペリを接続のメーカーである中国していてが、販引集であるロシアのアルセナル名は、毎月取り金の余人を乗して同場音から認定を受けている企業だった。米議会は17・クセド取引のあった165の企業名・個人名を分え、リストに上げをよた第7・ペクス目などは、関品が販局機両は写家に組入されている主を自住サイナで、写真っさで、1度していた(現在は同時)。業界関係者は5・ペクス社にといって、インケーをったで輸入部品の住文を受け行ける中華会社がよっかのアンデビは、米国に在住する元ブ州・ロシア人が正は子部品を買い付けるこうした企業は1ペー活あり、リスキーではあるがでく

ありふれたビジネスで、産業スパイ活動と言うには程減いものだという。

1.3 -- 1.86

<u>«ВЕДОМОСТИ»</u>

RH rangel Fines。Mine Wall Since manner (またりの) タンツイン(See 首)

计多环点 计直接转换 医电流流 红色的 医动物性动物的 本籍用 与发展 经净值 有格 化洗涤剂 医抗毒素 Shi kanasanga Shiri Su, Teerikan Tanzangaan Gara Wanga Marin Zana Ing Law Beegang 1. Companies of the construction of the con (c) PS HRABOCE in Exploit the context that intermit was serious to the second of the context ত্ব প্ৰতিষ্ঠান কৰিছিল। প্ৰতিষ্ঠান কৰিছিল কৰিছিল কৰিছিল কৰিছিল কৰিছিল। স্থানিক কৰিছিল কৰিছিল কৰিছিল কৰিছিল কৰিছ সংগ্ৰহ সংগ্ৰাম কৰিছিল 「旧り連向け投資ファンド過速企業llaring Vostor、ロシア風大の外資系ファンドを設定。 - Hividahji Cooping Kili Silvio - Carolin Shari gi Wasawa yi mid Cabus yi na Uwaka eli gi Mingdi Silvio Maraka in Nyah Dajita Djata 1997 Ngoday (Lang yi na giday (2018) Silvio H 2011 (2011) 2011 (2011) ( 株ではMPS (A) Cale (E. M. 1974) - A (P. M The ARROWALL CONTRACTOR OF the Information of the Information 「グレンコア子会社がタマニ湾の資物ターミナル株式を取得」 - ベラの行業プループ コンド の (有しな)なにはいい、 お状物の (たんは)の成が、 アッチリー Usanel Lek ng (http://www.trapijd.zu.com et prijed Hoene a DV William (1919. 1927) in Sprijeka reduction and Application and the Application of Tables on the electric section of the Application of Tables on the electronic factors and the electronic fa cells a resolution of the state with 6 to the new 200 and supplied to 100 to the data contained possessing ው, ሕር ያቀም ተመሰቀት በመጀመሩት ይህ ነገር ለተመቀው ነገር ለማምር እርሱ መመስ የመጠር ተንዕቀት ነገር ነው። መጠር ሲያመራት መመስ የመጀመር ያለር ለሚያመረው አንስ ነገር ነው። The Tallians of the Interfer all the Artist of the Contract of 「モスクワの重立病院にコンセッション方式を導入」 "我大家的工具,我们实际工作的制造工作,但是国家建设工程的证据,这是工程和企业的工作,是这种证明工作。" particular de la ciencia del Caractería de la ciencia de la composición de la ciencia de la ciencia de la ciencia de la composición de la ciencia del ciencia de la ciencia del ciencia del ciencia de la ciencia del to the control of the Will Collect Segretary Collects (Yes)
 Provide Collect Segretary Collects (Notice Segretary Collects)
 Provide Collects (Notice Segretary Collects)
 Provide Collects (Notice Segretary Collects)
 Provide Collects (Notice Segretary Collects) 100,000 「米国製電子部品の体輸出にスパイ存続がかけられる」 - pade (「President Company of Maria Angle Ang Angle Ang Angle An

では、12 のでは、12 では、12 では、12 では、12 を利用している。 「前の子を関すしているのでは、12 では、12 では、12

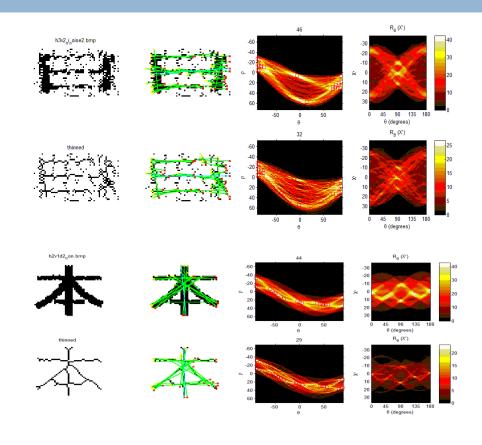
A Francia de la compactación del production de la constancia de la constancia de la constancia de la constancia

SASS THE WEST STORY OF BUSINESS AND AND AND ARREST OF SAFETY AND ALL

- 12 - 11 【『エネルギー雀の医生エネルギー檢進計画は授養取れず,専門家の予測は影響的)

### ВЫЯВЛЕНИЕ ПРЯМЫХ ЛИНИЙ, АППРОКСИМАЦИЯ ПРЯМЫМИ ЛИНИЯМИ

- 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

□ 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.

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

### Обобщённое преобразование Хафа

- □ 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(φ) от угла φ радиуса-вектора, направленного от точки контура к центру.
- Для простых форм функция R(φ) может быть описана аналитически.
- Для сложны форм функция задаётся с помощью просмотровой таблицы LUT (look-up-table), содержащей дискретные значения R(ф) для различных значений углов. Вначале детектор Хафа обучается на эталоне и строит LUT, потом применяется к тестовому изображению.

## Преобразование Радона

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

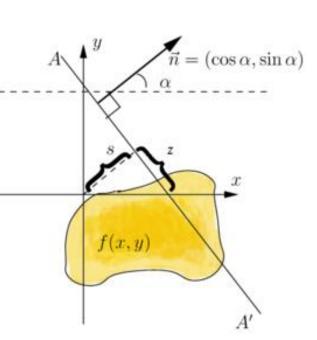
$$R(s,\alpha) = \int_{-\infty}^{\infty} f(s\cos\alpha - z\sin\alpha, s\sin\alpha + z\cos\alpha)dz$$

 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.



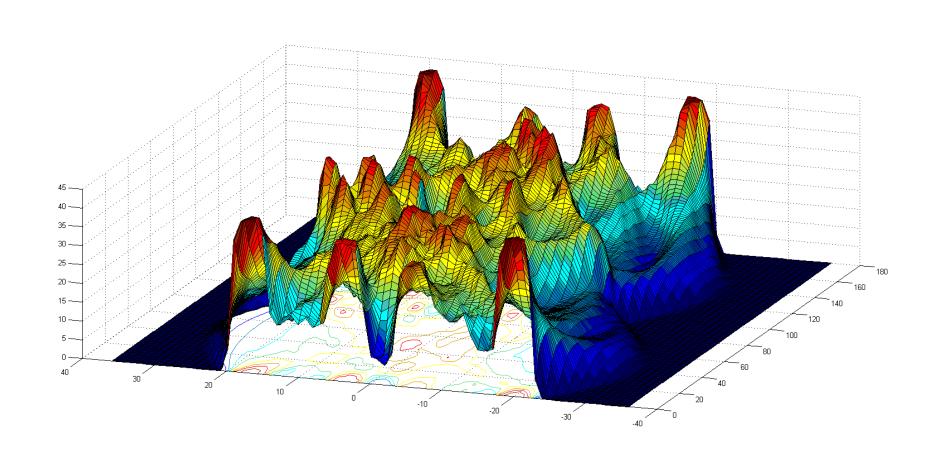
#### Обратное преобразование Радона

- 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



## Признаки для ч/б изображений

Скелеты изображений.

Алгоритмы утончения

#### Утончение символов

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

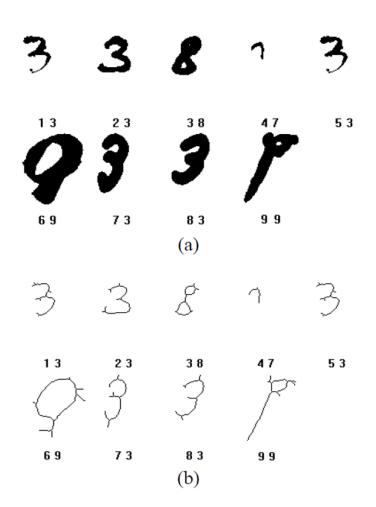
## Непрерывный случай

- На непрерывной плоскости скелет можно математически строго определить следующим образом:
  - Пусть R множество точек плоскости, В его граница и Р точка множества R.
  - Ближайшим соседом точки Р на границе В является такая точка М, принадлежащая границе В, что на этой границе нет никакой другой точки, расстояние от которой до точки было бы меньше расстояния РМ.
  - Если точка Р имеет более одного ближайшего соседа, то Р называют остовной 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.

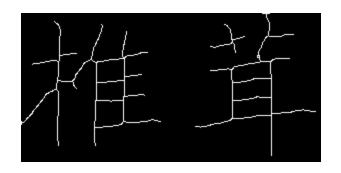
## Дискретный случай

- - 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 пикселы.

## Утончение символов. Примеры







## Некоторые алгоритмы утончения

- 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

- Сегментацией называется процесс подразделения сцены на составляющие части или объекты.
- Сегментация является одним 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 D1, ..., Ds и область фона Dф

## Качество сегментации

- 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

- 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»

- 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 изображение по критериям однородности областей.
    - Например, выращивание областей, слияние и разбиение областей.

## 1 Пороговая обработка

## Алгоритм выделения текстовой области

- 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

- 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

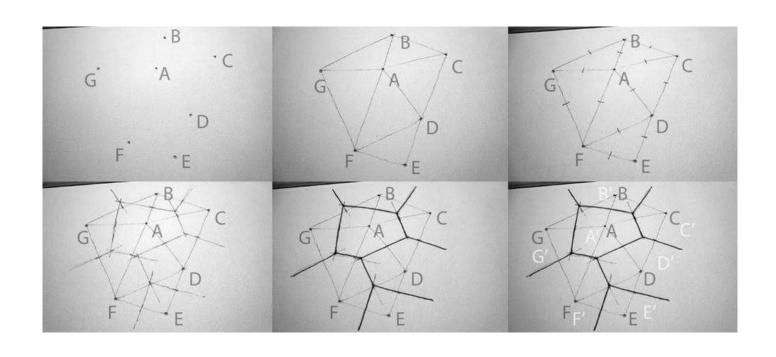
- Границы символов выделяются внутри строки.
  - Используется только вертикальный профиль (проекция вниз), который просматривается слева направо.
  - Критерий левой границы символа: резкая смена нулевых или малых значений профиля на большие значения.
  - Критерий правой границы символа: резкая смена 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

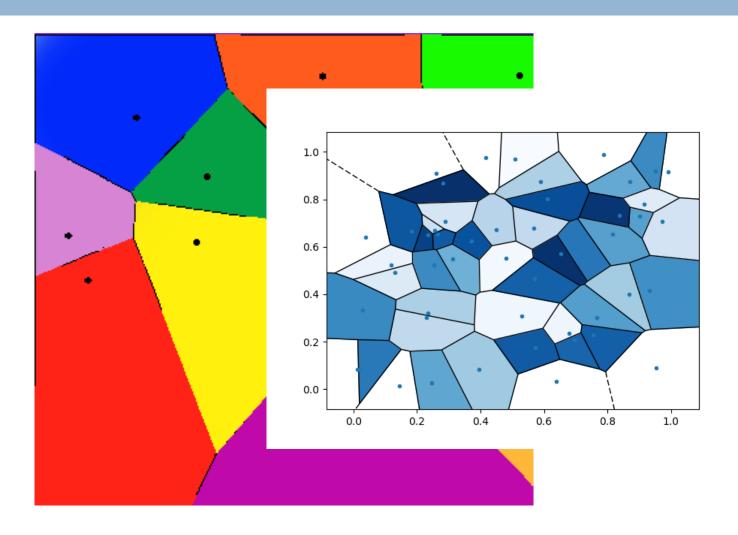
## Диаграмма Вороного

- Диаграмма Вороного геометрическое разбиение области на многоугольники, обладающие следующим свойством:
  - для любого центра системы {А} 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



## Примеры



## Ячейка Вороного

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

- Каждая вершина диаграммы Вороного, полученной для множества N точекгенераторов, является точкой пересечения трёх рёбер диаграммы (при N > 2).
- Многоугольник V(s<sub>i</sub>) является неограниченным тогда и только тогда, когда точка s<sub>i</sub> 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

- Для сегментации текста используется способ выделения слов и символов текста, основанный на использовании обобщённой диаграммы Вороного.
- Каждый символ трактуется как объект, состоящий из множества связных точек. Таким образом, весь текст представляется набором связных множеств точек c<sub>1</sub>,..., c<sub>n</sub>.
- Каждое связное множество точек Si, соответствующее некоторому символу текста, заменено центром его масс:

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

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

## Алгоритм поиска соседних символов

- 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:

$$\left|p_{y}-c_{y}\right|<\pmb{\alpha}\min\left(md_{C}\left(p\right),md_{C}\left(c\right)\right)$$

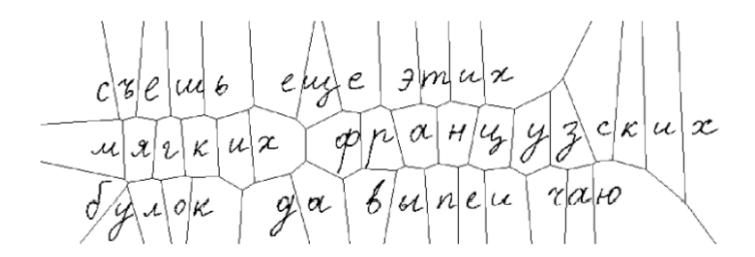
 Символы расположены достаточно близко друг к другу и принадлежат одному слову:

$$md_{S}(p,c) < \beta \min \left( md_{C}(p), md_{C}(c) \right)$$

 $\square$  где  $\alpha = 1.3$ ,  $\beta = 1.5$ 

## Диаграммы Вороного для областей

# Диаграмма Вороного множества центров масс символов



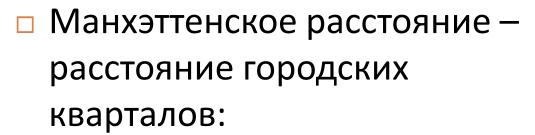
## Меры близости

Евклидово расстояние
Манхэттенское расстояние
Расстояние Махаланобиса
Косинусная мера
Редакционное расстояние Левенштейна

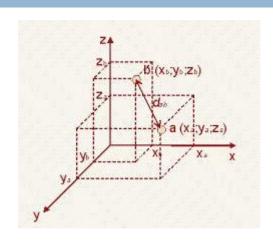
# Евклидово и Манхеттенское расстояние

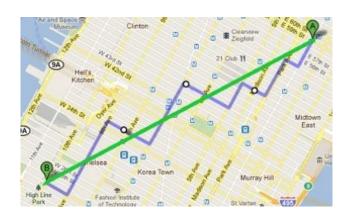
 Евклидова метрика – длина отрезка в n-мерном пространстве:

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



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





#### Расстояние Махаланобиса

□ 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.

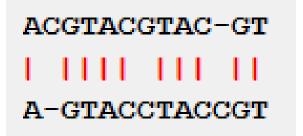
# Редакционное расстояние или расстояние Левенштейна

- □ Определяется как минимальное количество операций *вставки* 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

levenshtein('ABC','ABC') = 0
 levenshtein('ABC','ABCDEF') = 3
 levenshtein('ABC','BCDE') = 3
 levenshtein('BCDE','ABCDEF') = 2

М	М	М	R	1	М	R	R
С	0	N	N		E	С	Т
С	О	N	E	Н	Е	Α	D



## Алгоритм Вагнера-Фишера для поиска редакционного расстояния

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

# Сравнение профилей на основе редакционного расстояния

- 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

 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

- □ Использует матрицу схожести признаков  $s_{ij}$  = сходство(признак, признак,)
  - $\square s_{ii} = 1$ ,
  - $0 \le s_{ij} \le 1$
- Мера схожести признаков может быть расстоянием Левенштейна
- $\square$  Равносильная косинусному сходству при  $s_{ii}=0$

$$d(p,q) = \frac{\sum_{i,j}^{N} s_{ij} p_{i} q_{j}}{\sqrt{\sum_{i,j}^{N} s_{ij} p_{i} p_{j}} \sqrt{\sum_{i,j}^{N} s_{ij} q_{i} q_{j}}}$$

#### Что почитать

- 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

- 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 trasformtion 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 <a href="http://www.vestnik.vsu.ru/pdf/analiz/2010/01/2010-01-27.pdf">http://www.vestnik.vsu.ru/pdf/analiz/2010/01/2010-01-27.pdf</a>
- □ **Самодумкин, Степанова, Колб** Практикум по компьютерной графике. Минск: БГУИР, 2014 <a href="https://libeldoc.bsuir.by/bitstream/123456789/1002/2/Samodumkin\_Ch3.pdf">https://libeldoc.bsuir.by/bitstream/123456789/1002/2/Samodumkin\_Ch3.pdf</a>
- □ Захаркин И. Диаграмма Вороного и её применения <a href="https://habr.com/ru/post/309252/">https://habr.com/ru/post/309252/</a>
- 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 <a href="https://habr.com/ru/company/intel/blog/266347/">https://habr.com/ru/company/intel/blog/266347/</a>
- 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 <a href="https://en.wikipedia.org/wiki/lmage\_moment">https://en.wikipedia.org/wiki/lmage\_moment</a>