

Лк №1
07.02.05

Путятин Евгений Петрович

Методы и системы распознавания образов

список литературы:

1. Анисимов, Злобин, Курганов. Распознавание и цифровая обработка изображений. М.: Высш.школа
2. Васильев В.И. Распознающие системы. справочник. Киев: Наукова думка, 1983 г, 424 с.
3. Пред, Цифровая обработка изображений. Тома 1 и 2. М.: Мир, 1982 г.
4. Путятин Е.П., Овелин С.И. Обработка изображений в робототехнике. М.: Машиностроение, 1993 г, 320с.
5. Гарелик А.Л., Скрипник В.А. Методы распознавания. М.: Высш.школа. 1984 г. 208 с.
6. Гарелик А.Л., Гуревич, Скрипник В.А. Современное состояние проблемы распознавания. М.: Радио и связь, 1985 г.
7. Путятин Е.П., Гороховатский, Матат. Методи та алгоритми комп'ютерного зору. Навчальний посібник. Сміт. Харків, 2006 р.
8. Шлезингер М.И. Главач. 10 лекций по статистическому и структурному распознаванию. Киев, Наукова думка, 2004 г, 535 с.
9. Путятин, Матат. Pattern recognition and image processing.
10. Шапиро, Стокман. Компьютерное зрение. М.: Бином, 2006 г, 752 с.
11. Гонсалес, Гудз. Цифровая обработка изображений. М.: Техносфера, 2005 г, 1070 с.
12. Форсайт, Понт. Компьютерное зрение. Современный подход.

предметом наших исследований будут изображения.

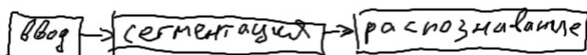
для нас изображение - это $V(x,y)$.

(общий случай - $V(x,y,z,t,\lambda)$)

основные задачи, которые связаны с обработкой изображений:

1. техническое зрение роботов
2. медицина (магнитно-резонансные томографы, УЗИ)
3. сельское хозяйство (анализ семян по размеру, и т.п.)
4. дистанционный анализ изображений со спутника

общая схема обработки изображения:



системы ввода информации:

1. цифровые камеры, видеокамеры, сканеры
2. УЗИ
3. радиолокация
4. магнитно-резонансная томография
5. графические планшеты
6. дигитайзеры - спец.устройства для ввода карт, чертежей и других линейных изображений

Лк №2
07.02.08

Обработка и распознавание документов на бумажных носителях

- 1) содержательная постановка задачи распознавания
- 2) сканеры
- 3) работа со сканером
- 4) особенности программы FineReader
- 5) особенности обработки документов

1) содержательная постановка задачи распознавания

имеется некоторая совокупность объектов и явлений, которая классифицирована в соответствии с какими-то правилами (разбита на классы). составлен алфавит классов. классы отличаются друг от друга совокупностью признаков. говорят о словаре признаков. нужно сопоставить входные изображения/объекты/явления какому-то классу. для этого измеряют признаки входящих объектов, возможно, их перерабатывают, обчисляют, и по определенному правилу, которое называется решающим правилом, сравнивают входящие признаки с эталонными признаками классов. отсюда вытекает, что классы предварительно должны быть изучены. для каждого класса должен быть набор признаков, который называют эталонными признаками.

пример:

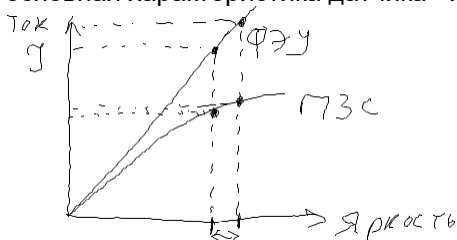
А А П

первое требование при построении системы распознавания образов - это анализ эталонов, причем каждый класс должен отличаться от другого существенным образом.

другой пример - система ПВО, она должна распознавать разные самолеты.

2) сканеры

основная характеристика датчика - характеристика свет-сигнал:



у ФЭУ эта характеристика линейная, т.е. лучше, чем у ПЗС (CCD).

также одна из характеристик - разрешение, измеряется в точках на дюйм (dpi) и линиях на дюйм (lpi).

3) работа со сканером

система TWAIN - универсальный стандарт связи всех сканеров с компьютером, независимо от типа компьютера и сканера.

включается сканер командой prescan. открывается диалоговое окно, свое для каждой модели сканера. в нем можно задавать параметры сканирования (черно-белое или цветное сканирование, количество цветов, разрешение...), можно регулировать яркость, контрастность, подавить всяческие помехи с помощью фильтров и т.д. цель режима prescan - чтобы отсканировать документ как можно более точно и без ошибок.

4) особенности программы FineReader

FineReader - программа для распознавания отсканированных изображений.

этапы распознавания:

1. сегментация - разделение рисунка на сегменты (рисунки, абзацы, строки и отдельные буквы).
2. распознавание.

таким образом, есть 5 режимов работы со сканером:

1. сканирование
2. сегментирование
3. распознавание
4. проверка
5. сохранение

анкеты, бюллетени для голосования, опросные листы - особенность всех их - наличие постоянной и переменной части. данные о таком сканировании обычно хранят в формате, который ориентируется на последующую обработку, например, средствами электронных таблиц (Excel).

классификация систем и методов распознавания образов

1) математическая постановка задачи сегментации

необходимо создать математическую модель входного сигнала.

на вход поступает матрица $B(i,j)$, где $i=1,2,\dots,m$; $j=1,2,\dots,n$ - строки и столбцы.

$$B(i,j) = H_1(i,j) + H_2(i,j) + \dots + H_s(i,j) + H_\Phi(i,j)$$

- т.е. сцена - это совокупность отдельных изображений, которые не пересекаются.

$H_\Phi(i,j)$ - фон.

$$D_1 \cup D_2 \cup \dots \cup D_s \cup D_\Phi = D$$

количество объектов, которые необходимо различать, (s) зависит от поставленной задачи.

$$D_i \cap D_j = \emptyset, \quad i \neq j$$

$$H_k(i,j) = \begin{cases} B(i,j), & (i,j) \in D_k \\ 0, & (i,j) \notin D_k \end{cases}$$

$$H_\Phi(i,j) = \begin{cases} B(i,j), & (i,j) \in D_\Phi \\ 0, & (i,j) \notin D_\Phi \end{cases}$$

$$k=1,2,\dots,s$$

сегментация - это построение отображения $\pi: D \rightarrow \{0,1,2,\dots,s\}$ (другими словами, присвоение отдельного номера каждому из отдельных объектов)

$$\pi(i,j) = l, \text{ если } (i,j) \in D_l, \quad l=1,2,\dots,s$$

$$\pi(i,j) = 0, \text{ если } (i,j) \in D_\Phi$$

практически всегда присутствует априорная информация о допустимом диапазоне яркостей, о связности объектов (например, буква i - не связная) и т.д.

переход от общего изображения к совокупности изображений объектов и фона, включая определение числа объектов s , называется сегментацией.

2) общая классификация систем распознавания

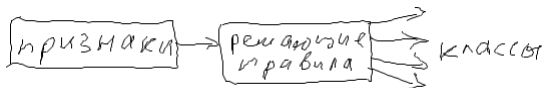
по сложности системы распознавания делятся на простые и сложные.

по количеству входной информации системы делятся на системы без обучения, с обучением и с самообучением.

проблема самообучения в литературе еще носит название *кластер-анализа*, *анализа таксонов*.

признаки могут быть детерминированные и стохастические, соответственно, по характеру признаков системы также делятся на детерминированные и стохастические. кроме этого, если признаки - логические, то система называется логической. бывают и комбинированные системы.

систему распознавания можно представить в виде двух блоков:



3) основные требования к признакам распознавания

1. информативность - хороший признак для двух классов распознавания. для многих классов часто неопределенный и до конца не изучен.

2. объем вычислений - чем меньше, тем лучше.

3. помехоустойчивость - значение признака не должно сильно изменяться при действии случайных помех.

4. инвариантность к геометрическим преобразованиям (независимость от масштаба, угла поворота и т.д.)

5. реализуемость, включая экономическую, экологическую и т.п. составляющие

6. независимость - признаки должны быть независимы друг от друга

детерминированные системы распознавания

геометрические признаки

1) основные типы признаков

1. геометрические признаки - характеризуют форму объекта.
2. топологические признаки
3. яркостные признаки
4. спектральные признаки
5. структурные признаки
6. вероятностные и нечеткие признаки

2) метрические признаки

1. признаки расстояния

$$D_E(i, j), (h, k) = \sqrt{(i-h)^2 + (j-k)^2}$$

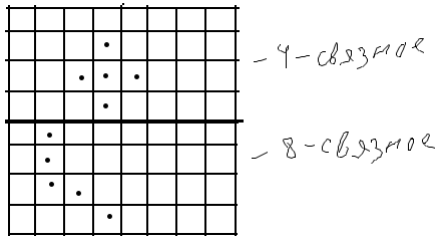
$$D_4(i, j), (h, k) = |i-h| + |j-k|$$

$$D_8(i, j), (h, k) = \max\{|i-h|, |j-k|\}$$

2. признаки соседства

если $D_4 = 1$, то пиксели считаются соседними в смысле четырехсвязности, если $D_8 = 1$ - в смысле 8-связности.

если множество 4-связное, то оно тем более 8-связное, но не наоборот.



A_1, A_2, \dots, A_n - последовательность соседних элементов. эта последовательность называется путем. связная область - это такая область, где все пути между произвольными точками этой области принадлежат этой области.

также существуют понятия границы и края (edge). край - это локальный признак, а граница - локальный (т.е. касается всего объекта).

3) инвариантные признаки

(инвариантные к смещениям, поворотам, масштабированию...)

пусть P - признак или множество признаков, M - множество признаков. отображение $\Theta: M \rightarrow P$ называется инвариантным относительно группы геометрических преобразований G , если для него справедливо $\Theta(gB) = \Theta(B)$, $g \in G$

рассмотрим признак - площадь фигуры (количество пикселей внутри фигуры). она инвариантна к смещениям, но не инвариантна к масштабированию. легко вычисляется.

следующий признак - длина контура. требует гораздо большего объема вычислений, чем площадь. инвариантна к смещениям и поворотам, но так же зависит от масштаба, как и признак площади.

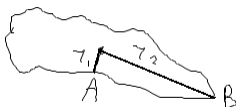
пусть $S(E)$ - площадь, $L(E)$ - длина контура. тогда можно ввести признак $V(E) = \frac{\sqrt{S(E)}}{L(E)}$ - он является

инвариантным к смещениям, поворотам и масштабированию.

также инвариантным является следующий признак:

$$V_1(E) = \frac{S(E)}{L^2(E)}$$

но его нельзя использовать вместе с предыдущим, т.к. эти два признака взаимосвязаны.

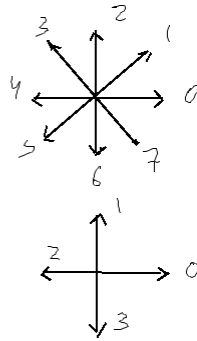
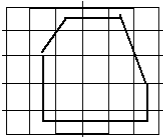


- в качестве признака можно использовать кратчайшее и наибольшее расстояния от центра масс фигуры. такой признак является инвариантным к смещениям и поворотам, но зависит от масштаба. зато

отношение этих расстояний от масштаба уже не зависит.

4) признаки формы

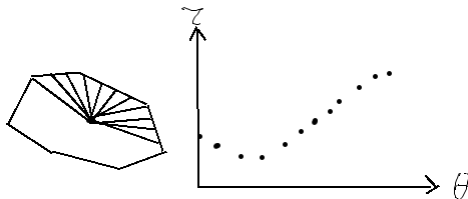
цепные коды:



(i^0, j^0) для четырехсвязного признака: 00000111

(i^0, j^0) для восьмисвязного признака: 00000222333

следующий способ описания границ - применение сигнатуры. сначала необходимо найти геометрический центр, из него построить лучи с некоторым интервалом (в градусах):



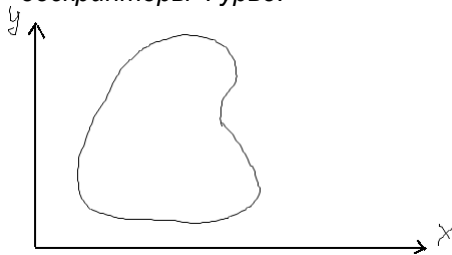
Θ - угол, r - длина вектора из центра до границы при угле Θ . таким образом, получаем вектор, зависимость $r(\Theta)$.

признаки формы еще называют дескрипторами границ.

длина контура - общее число пикселей вдоль границы. для цепных кодов она вычисляется легко, а для сигнатур - сложнее.

еще один из признаков - максимальное расстояние между точками границы. эксцентрицитет.

дескрипторы Фурье:



$k=0,1,2,\dots,K-1$

$$x(k) = x_k$$

$$y(k) = y_k$$

$$S(k) = x_k + jy_k$$

$$a(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) e^{-\frac{j2\pi uk}{K}}, \quad u = 0, 1, \dots, K-1$$

$a(u)$ - дескрипторы границы.

граница восстанавливается обратным преобразованием Фурье:

$$S(k) = \sum_{u=0}^{K-1} a(u) e^{\frac{j2\pi uk}{K}}, \quad k = 0, 1, \dots, K-1$$

можно за счет уменьшения точности многие $a(u)$, близкие к 0, приравнять к 0:

$$\tilde{S}(k) = \sum_{u=0}^{K-1} a(u) e^{\frac{j2\pi uk}{K}}, \quad a(u) \neq 0$$

это позволит заметно сократить время вычислений.

5) аффинные преобразования

$$x' = a_{11}x + a_{12}y + a_{13}$$

$$y' = a_{21}x + a_{22}y + a_{23}$$

смещения, повороты, масштабы (разные по разным осям), зеркальные отображения, косые сдвиги
свойства аффинных преобразований: прямая переходит в прямую; параллельные прямые переходят в параллельные прямые.

отношение трех точек:



$$(A, B, C) = \frac{|AC|}{|BC|} = \frac{|A'C'|}{|B'C'|} = \lambda \text{ - этот параметр не меняется при аффинных преобразованиях.}$$

центр фигуры также инвариантен к аффинным преобразованиям.

треугольники остаются треугольниками, параллелограммы остаются параллелограммами, но ромб, квадрат и прямоугольник становятся неразличимыми между собой.

Лк №5
07.03.02

ТОПОЛОГИЧЕСКИЕ И ЯРКОСТНЫЕ ПРИЗНАКИ

- 1) инвариантные признаки проективных преобразований
- 2) топологические признаки
- 3) основные признаки яркости
- 4) гистограммы изображений как признаки
- 5) признаки на основе моментов изображений

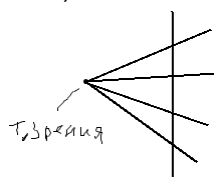
отношение трех точек:

$$(A, B; C) = \frac{|AC|}{|BC|} = \frac{|A'C'|}{|B'C'|} = \lambda$$

1) инвариантные признаки проективных преобразований

все точки, которые лежат вдоль луча, идущего из точки зрения, проектируются в одну точку для наблюдателя.

при проективных преобразованиях говорят об однородных координатах (изображение является плоским).



$$x' = \frac{x}{z}, \quad y' = \frac{y}{z}$$

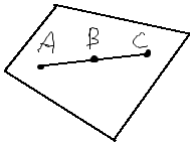
инварианты для проективных преобразований:

1. точки, лежащие на прямой, при проективных преобразованиях также будут лежать на прямой (т.е. прямая переходит в прямую).

2. точки, лежащие на кривой второго порядка (эллипс, парабола, гипербола), останутся на кривых второго порядка.

3. инвариантом является сложное отношение четырех точек на прямой (сложное отношение - это отношение простых отношений):

$$(A, B; C, D) = \frac{(A, B; C)}{(A, B; D)} = \lambda$$



метрические признаки - это по сути признаки описания формы объекта (силуэта).

наука, позволяющая определять признаки типа крайних точек, скелетов, внутренних характерных точек и т.д. - называется математическая морфология.

2) топологические признаки

отличительное свойство их инвариантов - при топологических преобразованиях связность не меняется. отверстия остаются отверстиями. число отверстий является инвариантом. зато, например, треугольники не всегда будут оставаться треугольниками.

признаки:

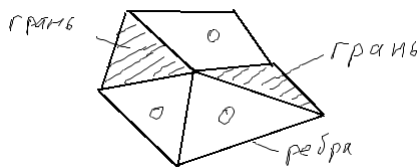
H - число отверстий.

C - число связных компонент.

признак $E = C - H$ является инвариантом при топологических преобразованиях, и называется числом Эйлера.

для буквы А число Эйлера равно $1-1=0$. для буквы В число Эйлера равно $1-2=-1$.

это является хорошим признаком при распознавании букв.



- для этого многоугольника $C=1$ (через центральную точку есть связь между гранями), $H=3$ (три отверстия). т.е. $E=1-3=-2$.

пусть V - число вершин, Q - число ребер, F - число граней.

имеет место следующая инвариантная зависимость:

$$E = C - H = V - Q + F$$

$$V - Q + F = 7 - 11 + 2 = -2$$

еще на практике используются такие топологические признаки как площадь (например, площадь фигуры в конкретной задаче может быть ограничено некоторым числом, - если она превышает это число, ее отбрасывают), длина контура и т.п.

3) основные признаки яркости

1. средняя яркость изображения B_{cp} - инвариантна к поворотам и сдвигам.

$$B_{cp} = \frac{1}{N} \sum_{i,j} B(i, j) - \text{суммируются значения пикселей в каждой точке и делятся на их число.}$$

2. B_{max} , B_{min} - соответственно максимальное и минимальное значения яркости, но они плохо защищены от помех

3. $\mu(B) = \sum_{i,j} |B(i, j) - B_{cp}|$ - разброс относительно среднего значения яркости.

$$\mu(B) = \frac{\mu(B)}{\|B\|}$$

4. пусть $B(i, j) = B_0(i, j) + C$ - яркость может меняться. тогда инвариантом будет например $B_{max} - B_{min}$.

5. пусть $B(i, j) = kB_0(i, j)$, тогда инвариантом будет $\frac{B_{max}}{B_{min}}$ и например $\frac{B_{max}}{B_{cp}}$.

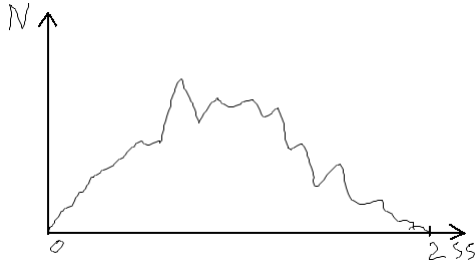
6. пусть $B(i, j) = kB_0(i, j) + C$. тогда инвариант - $p = \frac{B_{max} - B_{min}}{B_{max} + B_{min} - 2B_{cp}}$

4) гистограммы изображений как признаки

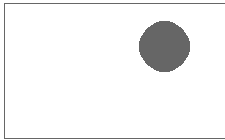
гистограмма - это широко распространенная характеристика, которая часто используется при распознавании образов.

гистограмма $H(B)$ - это отображение значений яркости в их количества, т.е.

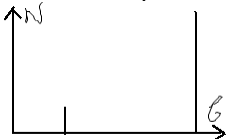
$$H(B) = \{(i, j) \mid B(i, j) = b\}$$



важное применение гистограммы при распознавании - это нахождение порога яркости, отделяющее объекты от фона.



для вышеприведенного рисунка гистограмма будет выглядеть следующим образом:



в общем случае гистограмма - многомодальная функция.

5) признаки на основе моментов изображений

определение моментов:

$$M_{pq} = \iint_D B(x, y) x^p y^q dx dy$$

$$M_{pq} = \sum_i \sum_j B(i, j) i^p j^q$$

некоторые из этих моментов имеют установившийся физический смысл.

1. если $p, q=0$ (называется - энергия, или момент нулевого порядка), то получаем сумму от $B(i, j)$

2. $p=1, q=0$, или $p=0, q=1$. это моменты первого порядка.

3. центр тяжести изображения:

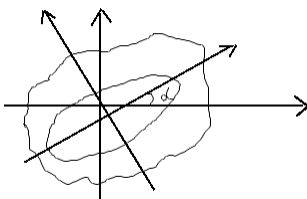
$$i_0 = \frac{\sum_{i,j} B(i, j) i}{\sum_{i,j} B(i, j)}, \quad j_0 = \frac{\sum_{i,j} B(i, j) j}{\sum_{i,j} B(i, j)}$$

4. $p=2, q=0$ или $p=0, q=2$. - моменты инерции (моменты второго порядка).

$$M_{20} = \sum_{i,j} B(i, j) i^2$$

$$M_{02} = \sum_{i,j} B(i, j) j^2$$

они определяют так называемые главные оси инерции.



твердое тело характеризуется некоторым эллипсом инерции, где оси инерции и определяются этими моментами. угол наклона главной оси:

$$\alpha = \frac{1}{2} \operatorname{arctg} \frac{2M_{11}}{M_{20} - M_{02}}$$

$$M_{11} = \sum_{i,j} B(i,j)ij \text{ - центробежный момент.}$$

также используются моменты третьего порядка: $M_{30}, M_{03}, M_{21}, M_{12}$

все приведенные моменты используются как признаки при распознавании образов.

также говорят об *инвариантных моментах* - т.е. таких, которые не зависят от смещений изображения, пропорциональных изменениях яркости и т.д.

$$\mu_{11} = M_{11} - \bar{j}m_{10}$$

$$\mu_{20} = M_{20} - \bar{i}m_{10}$$

$$\mu_{02} = M_{02} - \bar{j}m_{01}$$

- вычитаются центры тяжести

Лк №6

07.03.09

момент в общем виде:

$$m_{pq} = \iint_{-\infty} B(x,y)x^p y^q dx dy$$

$$m_{pq} = \sum_{i,j=1}^N j^p i^q B(i,j)$$

центральные моменты:

$$\mu_{pq} = \sum_{i,j=1}^N \left(j - \frac{m_{10}}{m_{00}}\right)^p \left(i - \frac{m_{01}}{m_{00}}\right)^q B(i,j)$$

группа центральных моментов характерна тем, что их значение не зависит от положения изображения в пространстве (т.е. они инвариантны по отношению к смещениям).

$$\mu_{00} = m_{00}; \mu_{10} = \mu_{01} = 0$$

центральные моменты инерции:

$$\mu_{20} = m_{20} - \frac{m_{10}^2}{m_{00}}$$

$$\mu_{02} = m_{02} - \frac{m_{01}^2}{m_{00}}$$

и т.д.

на базе центральных моментов вводится нормированный момент $\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^z}$, $z = \frac{p+q}{2} + 1$,

для $(p+q) \geq 2$ - он не зависит от масштаба.

на этой базе предлагаются моменты, инвариантные к поворотам, смещениям и масштабированию:

$$I_1 = \eta_{20} + \eta_{02}$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

и т.д.

- соответственно инварианты первого, второго и третьего порядков. все они вводятся для $(p+q) \leq 3$

система инвариантных моментов Цернике более устойчива для ошибок дискретизации; например, $z_1 = 3[2(\eta_{20} + \eta_{02}) - 1]$.

спектральные признаки

общий вид преобразований:

$$g(m_1, m_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2) W(n_1, n_2, m_1, m_2)$$

где m_1, m_2 - столбцы и строки, W - ядро преобразования.
с математической точки зрения это называется "свертка".

если положить, что $W(n_1, n_2, m_1, m_2) = \Psi(n_1, n_2)$ (где $\Psi(n_1, n_2)$ - собственные векторы), то это - разложение Карунена-Лозва. оно будет оптимальным. мощность (энергия) изображения концентрируется в минимально возможном числе признаков. т.е. подвергнув входное изображение разложению Карунена-Лозва, мы получим наибольшее количество информации при минимальном количестве признаков.

ввиду сложности вычислений (отсутствия быстрых алгоритмов) коэффициентов Карунена-Лозва а также невозможности точного знания корреляционной функции, описывающей поле, и, следовательно, собственных векторов. однако если изображение стационарно и размеры его достаточно велики, то хорошим приближением к преобразованию Карунена-Лозва является спектр Фурье (косинусные преобразования).

$$W(n_1, n_2, m_1, m_2) = e^{-j2\pi \left(\frac{n_1 m_1}{N_1} - \frac{n_2 m_2}{N_2} \right)}$$

получаем $g(m_1, m_2)$, которая позволяет получить спектр изображения, характеризующий значения амплитуд пространственных частот, которые входят в разложение.

входное изображение, состоящее из множества пикселей, называется сигнальным пространством.

коэффициенты Фурье используются как признаки. существование быстрых алгоритмов их вычисления является большим преимуществом. недостатком является то, что коэффициенты Фурье и Карунена-Лозва не инвариантны к смещениям, поворотам и масштабированию.

$g(m_1, m_2)$ является инвариантным к смещениям.

косинусные преобразования (нету sin):

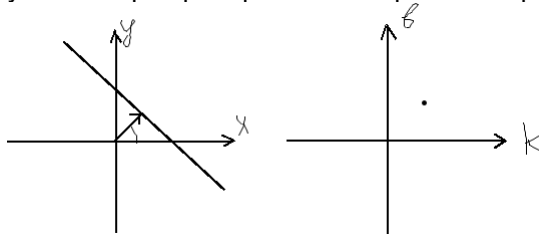
$$W(n_1, n_2, m_1, m_2) = \cos\left(2\pi \frac{(2n_1 + 1)m_1}{N_1}\right) \cos\left(2\pi \frac{(2n_2 + 1)m_2}{N_2}\right)$$

преобразования радона:

$$W(n_1, n_2, m_1, m_2) = \delta(m_2 - n_1 \cos m_1 - n_2 \sin m_1)$$

пусть $m_2 = n_1 \cos m_1 + n_2 \sin m_1$ - уравнение прямой.

$y=kx+b$ - при преобразовании прямая сворачивается в точку, в новой системе координат kb .



для этого преобразования также существуют быстрые преобразования, как и для Фурье.

вместо большого сигнального пространства благодаря спектральным преобразованиям получают ряд признаков, работать с которым гораздо удобнее.

вероятностные признаки

они характеризуются значением вероятности (как часто появляется) и среднеквадратическим отклонением.

метод Монте-Карло. - например, можно провести многократные испытания по "бросанию" прямой на какую-либо фигуру. в результате по длинам отрезков, попавших на фигуру, можно оценить размеры и площадь фигуры.

на практике редко используются, в виду большой вычислительной сложности

признаки цветных изображений

цветное изображение - это три монохроматических. можно обрабатывать каждый канал отдельно, но тогда количество признаков утраивается.

можно сказать, что человек воспринимает информацию в системе HSL (Hue, Saturation, Lightness).

вывод: изображения, которые представляются в сигнальном пространстве, можно перевести в набор признаков, и заменить распознавание изображений распознаванием этого набора признаков, что

значительно легче.

Лк №7 07.03.16

к/р на след. лекции (23 марта) - один вопрос; на второй полупаре; 30 минут.

меры сходства в пространстве признаков

$\omega_l(P_1, P_2, \dots, P_r)$ - вектор признаков, который описывает некоторый объект.

$\omega_0(P_1^0, P_2^0, \dots, P_r^0)$ - вектор признаков, описывающий другой, эталонный объект.

если объекты одинаковы, то будут совпадать и признаки. но если есть помехи, признаки могут отличаться.

чтобы решить вопрос о распознавании объектов, нужно уметь сравнивать векторы признаков, описывающие эти объекты.

$k = 1, 2, \dots, s$ - конечное множество классов. в каждом классе есть эталон, т.е. всего s эталонов. простейшее правило распознавания, очевидно, следующее: вектор признаков входного изображения сравнивается поочередно со всеми эталонными векторами. объект должен быть отнесен к такому классу, где сходство будет наибольшим.

1. на входе - один объект.

$$\omega_l \Leftrightarrow \begin{cases} \omega_0^1 \\ \omega_0^2 \\ \vdots \\ \omega_0^s \end{cases}$$

2. на входе - несколько объектов. мы хотим узнать, есть ли среди них объект, принадлежащий определенному классу.

$$\begin{matrix} w_i^1 \\ w_i^2 \\ \vdots \\ w_i^s \end{matrix} \Leftrightarrow \{w_i^0\}$$

3. общая задача распознавания (первые две были частными): много объектов и много эталонов.

$$\begin{matrix} w_i^1 \\ w_i^2 \\ \vdots \\ w_i^s \end{matrix} \Leftrightarrow \begin{cases} \omega_0^1 \\ \omega_0^2 \\ \vdots \\ \omega_0^s \end{cases}$$

$$\arg \left\{ \text{opt}_{k=1,2,\dots,s} \rho(\{w_l\}, \{\omega_0^k\}) \right\} = k^*$$

(для первой задачи)

под ρ понимается либо мера сходства, либо мера различия, в зависимости от задачи.

меры сходства между векторами признаков:

1. расстояние Махаланобиса

в общем случае признаки P_i могут быть статистически зависимыми, и отнесение объекта к одному из классов может определяться некоторыми весовыми коэффициентами.

$$\rho(\omega_l, \omega_0) = \sqrt{(\omega_l - \omega_0)^T \Lambda^T \Sigma^{-1} \Lambda (\omega_l - \omega_0)}$$

Σ - ковариационная матрица

Λ - матрица весовых коэффициентов

$(w_l - w_0)$ - вектор, составляющие которого - разница между отдельными составляющими каждого из векторов ω_l и ω_0

2. метрика Минковского

$$\rho(\omega_1, \omega_0) = \left[\sum_{i=1}^r |P_i - P_i^0|^m \right]^{1/m}$$

предполагается, что признаки взаимно независимы, имеют одинаковую дисперсию и одинаково важны. благодаря тому, что признаки взаимно независимы, ковариационная матрица не используется.

3. евклидова метрика

$$\rho(\omega_1, \omega_0) = \left[\sum_{i=1}^r |P_i - P_i^0|^2 \right]^{1/2}$$

$$4. \rho(\omega_1, \omega_0) = \sum_{i=1}^r |P_i - P_i^0|$$

$$5. \text{ метрика с весами: } \rho(\omega_1, \omega_0) = \left[\sum_{i=1}^r \gamma_i |P_i - P_i^0|^2 \right]^{1/2}$$

для каждого класса вводится свой вес γ_i - для случая, когда признаки неравноценны.

6. для бинарных признаков отдельно вводится метрика Хэмминга:

$$\rho(\omega_1, \omega_0) = \sum_{i=1}^r |P_i - P_i^0|, \quad P_i = \begin{cases} 0 \\ 1 \end{cases}$$

основные признаки здесь - признаки формы (например, есть угол или нет угла).

7. метрика Чебышева:

$$\rho(\omega_1, \omega_0) = \max |P_i - P_i^0|$$

- если максимальное различие превышает некоторое значение, то к заданному классу объект отнести нельзя

8. функция Камберра??....

$$\rho(\omega_1, \omega_0) = \sum_{i=1}^r \frac{|P_i - P_i^0|}{|P_i + P_i^0|}$$

$$9. \rho = \frac{\sum_{i=1}^r P_i P_i^0}{\sqrt{\sum_{i=1}^r (P_i)^2 \sum_{i=1}^r (P_i^0)^2}} - \text{коэффициент корреляции}$$

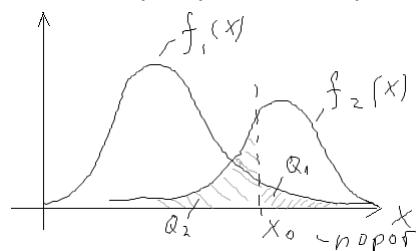
мера сходства назначается постановщиком задачи.

классы: Ω_1, Ω_2

1. вероятность появления объекта из определенного класса - $P(\Omega_1), P(\Omega_2)$

2. мат. ожидание и среднеквадратическое отклонение (для нормального закона распределения). для других законов распределения - тоже свои параметры.

плотность распределения признака из первого класса ($f_1(x)$) и признака из второго класса ($f_2(x)$):



x_0 - порог.

1. $x > x_0 \Rightarrow \Omega_2$

2. $x \leq x_0 \Rightarrow \Omega_1$

если объект на самом деле относится к классу Ω_1 , а его отнесли к классу Ω_2 , то совершается ошибка первого рода. ее вероятность -

$$Q_1 = \int_{x_0}^{\infty} f_1(x) dx = 1 - F\left(\frac{x_0 - m_1}{\sigma_1}\right). \text{ ошибочно принята гипотеза } H_2$$

ошибки второго рода относятся к ситуации, когда справедлива гипотеза H_2 , а принята гипотеза H_1 (т.е. объект - из класса Ω_2 , а его отнесли к классу Ω_1). ее вероятность -

$$Q_2 = \int_{-\infty}^{x_0} f_2(x) dx = F\left(\frac{x_0 - m_2}{\sigma_2}\right)$$

пример:

пусть гипотеза H_1 - наличие помехи, отсутствие цели. H_2 - гипотеза о том, что есть цель. ошибка первого рода - когда мы говорим что есть цель, при том что ее нету. второго рода - когда есть цель, но мы принимаем ее за помеху.

вероятности правильного обнаружения:

$$D_2 = \int_{x_0}^{\infty} f_2(x) dx = 1 - Q_2$$

$$D_1 = \int_{-\infty}^{x_0} f_1(x) dx = 1 - Q_1$$

Лк №8

07.03.23

возникает вопрос об оптимальном выборе величины x_0 . существует несколько подходов для построения критерия оптимальности. самый распространенный - критерий Байеса или критерий среднего риска. для его построения вводятся величины затрат за правильное принятие решения и за ошибочное принятие решения, которые представляются матрицей:

$$C^* = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix}$$

c_{11} - плата за правильное определение шумов; c_{22} - плата за правильное определение цели.

c_{12} - плата за ошибку первого рода (мы сказали, что есть цель, а ее нет)

c_{21} - плата за ошибку второго рода.

Лк №9

07.03.30

метод минимизации среднего риска (критерий Байеса)

$$C^* = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix}$$

средние потери $R = c_{11}D_1P(\Omega_1) + c_{12}Q_1P(\Omega_1) + c_{22}D_2P(\Omega_2) + c_{21}Q_2P(\Omega_2)$ - сумма потерь при всех четырех возможных вариантах. задача состоит в том, чтобы минимизировать эти потери.

$$Q_1 = \int_{x_0}^{\infty} f_1(x) dx = 1 - F\left(\frac{x_0 - m_1}{\sigma_1}\right)$$

$$Q_2 = \int_{-\infty}^{x_0} f_2(x) dx = F\left(\frac{x_0 - m_2}{\sigma_2}\right)$$

$$D_1 = \int_{-\infty}^{x_0} f_1(x) dx = 1 - Q_1$$

$$D_2 = \int_{x_0}^{\infty} f_2(x) dx = 1 - Q_2$$

делаем подстановку:

$$R = P(\Omega_1) \left[c_{11} \int_{-\infty}^{x_0} f_1(x) dx + c_{12} \int_{x_0}^{\infty} f_1(x) dx \right] + P(\Omega_2) \left[c_{22} \int_{x_0}^{\infty} f_2(x) dx + c_{21} \int_{-\infty}^{x_0} f_2(x) dx \right]$$

$$\left. \frac{dR}{dx} \right|_{x=x_0} = P(\Omega_1)[c_{11}f_1(x_0) - c_{12}f_1(x_0)] + P(\Omega_2)[c_{21}f_2(x_0) - c_{22}f_2(x_0)] = 0$$

отсюда вытекает так называемое *отношение правдоподобия*:

$$\frac{f_2(x_0)}{f_1(x_0)} = \frac{P(\Omega_1)(c_{12} - c_{11})}{P(\Omega_2)(c_{21} - c_{22})} = \lambda_0$$

для объекта вычисляется $\lambda(x) = \frac{f_2(x)}{f_1(x)}$. если $\lambda(x) \leq \lambda_0$, то причисляем объект к классу Ω_1 , иначе - к

классу Ω_2 .

если $f_1(x)$ и $f_2(x)$ подчиняются нормальному закону распределения, то задача решается непосредственным нахождением оптимального значения x_0 :

$$x_0 = \frac{\sigma_2^2 m_1 - \sigma_1^2 m_2 \pm \sigma_1 \sigma_2 \sqrt{(m_2 - m_1)^2 + (\sigma_2^2 - \sigma_1^2) \ln \lambda_0 \frac{\sigma_2^2}{\sigma_1^2}}}{\sigma_2^2 - \sigma_1^2}$$

если $\sigma_1 = \sigma_2 = \sigma$, тогда $x_0 = \frac{m_1 + m_2}{2} + \frac{\sigma^2}{m_2 - m_1} \ln \lambda_0$

если кроме того еще предположить, что плата за правильное распознавание отсутствует, т.е. потери

$$c_{11} = c_{22} = 0, \text{ то } x_0 = \frac{m_1 + m_2}{2}.$$

случай с m классами, вместо двух:

$$\Omega = 1, 2, \dots, m$$

признак каждого класса характеризуется N компонентами: $X = (x_1, x_2, \dots, x_N)$

вводим отношение правдоподобия между классами Ω_k, Ω_l :

$$\lambda_{k,l} = \frac{f_k(X)}{f_l(X)}, \quad k, l = 1, 2, \dots, m$$

матрица платежей:

$$C = \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \end{vmatrix}$$

каждый элемент матрицы имеет свой смысл, например, c_{m2} - объект класса m приняли за объект класса 2.

по аналогии как для двух классов, выписываем средний риск:

$$R = \sum_{k=1}^m \sum_{l=1}^m P(\Omega_l) c_{kl} \int_{D_k} f_l(X) dX$$

$P(\Omega_l)$ - вероятность появления объекта из класса l .

область D_k - многомерная область, разделяющая класс k от соседних классов.

берем производную от функции потерь:

$$P(\Omega_k) f_k(X) (c_{kl} - c_{kk}) - P(\Omega_l) f_l(X) (c_{lk} - c_{ll}) = 0, \quad k, l = 1, 2, \dots, m$$

- получили систему нелинейных уравнений. решая ее, получим оптимальное значение области D_k .

задача упрощается для нормального закона распределения:

m_i - мат. ожидание; K - ковариационная матрица.

имеется ряд критериев, которые упрощают задачу за счет пренебрежения теми или иными параметрами.

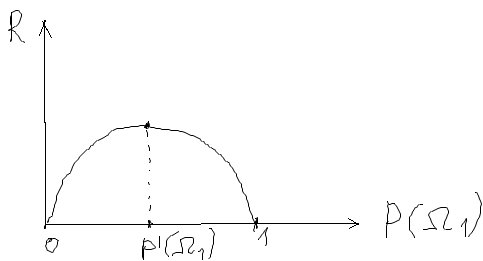
первый из них - *минимаксный критерий* (лучший вариант из худших). пусть $c_{11} = c_{22} = 0$. обозначим

$$c_{12} = c_1, \quad c_{21} = c_2.$$

$$P(\Omega_1) = 1 - P(\Omega_2)$$

$$R = P(\Omega_1) c_1 \int_{x_0}^{\infty} f_1(x) dx + [1 - P(\Omega_1)] c_2 \int_{-\infty}^{x_0} f_2(x) dx$$

построим график зависимости R от $P(\Omega_1)$:



в минимаксном методе ориентируются на наихудший случай $P'(\Omega_1)$, и минимизируют потери в наихудшем случае.

$$\text{отношение правдоподобия будет выглядеть как } \lambda'_0 = \frac{c_1}{c_2} \left(\frac{P'(\Omega_1)}{1 - P'(\Omega_1)} \right)$$

$$\lambda(x) \geq \lambda'_0 \Rightarrow \Omega_2$$

$$\lambda(x) < \lambda'_0 \Rightarrow \Omega_1$$

Лк №10

07.04.06

алгоритмы распознавания в условиях неполной информации

1) критерий Неймана-Пирсона

предполагается, что неизвестны не только $P(\Omega_1)$ или $P(\Omega_2)$, но неизвестна и платежная матрица

$\|c_{ij}\|$.

задаются допустимые вероятности ошибки первого рода (применительно к радиолокации - вероятность ложной тревоги. - чтобы она не возникала чаще чем задано):

$$Q_1 = \int_{x_0}^{\infty} f_1(x) dx \leq A$$

ставится задача минимизировать вероятность ошибки второго рода (применительно к радиолокации - вероятность пропуска цели):

$$\min_{x_0} Q_2 = \min_{x_0} \int_{-\infty}^{x_0} f_2(x) dx$$

решение задачи соответствует x_0 , найденному из уравнения $\int_{x_0}^{\infty} f_1(x) dx = A$

2) алгоритм последовательных решений

раньше мы считали, что решение о принадлежности входного объекта ω к классу Ω_i , $i = 1, 2, \dots, m$ принималось после измерения вектора признаков $X(x_1, x_2, \dots, x_N)$. на практике возникает ситуация, когда получение измерений связано, например, с риском для здоровья, или измерения дорогие. тогда возникает вопрос принятия решений по неполному количеству измерений. пусть $i = 1, 2$. должны быть заданы плотности распределения для каждого класса f_i :

$$f_i(x_i), f_i(x_1, x_2), \dots, f_i(x_1, x_2, \dots, x_N)$$

допустим, что проведено $n < N$ измерений, так что получили в результате x_1, x_2, \dots, x_n .

$$\lambda_n = \frac{f_2(x_1, x_2, \dots, x_n)}{f_1(x_1, x_2, \dots, x_n)}$$
 - находим отношение правдоподобия, и сравниваем его с двумя числами -

$[A_n, B_n]$.

алгоритм распознавания мог бы выглядеть следующим образом:

если $\lambda_n < A_n$, то объект принадлежит классу Ω_1 , если $\lambda_n \geq B_n$, то объект принадлежит классу Ω_2 , а если λ_n находится между A_n и B_n , то необходимо дальнейшее исследование.

на каждом шаге должны быть известны вероятности ошибок первого и второго рода - Q_{1n}, Q_{2n} , $n = 1, 2, \dots$. по аналогии с предыдущими лекциями, можно вычислить

$$A_n = \frac{1 - Q_{1n}}{Q_{2n}}, \quad B_n = \frac{Q_{1n}}{1 - Q_{2n}}.$$

3) обучение систем распознавания образов

если классы $\Omega_1, \Omega_2, \dots, \Omega_m$ нам заданы, а признаков $X_i = (x_1, x_2, \dots, x_N)$, $i = 1, 2, \dots, m$ не хватает, и по ним нельзя точно определить, к какому классу относится входной объект. в этой ситуации требуется обучение системы распознавания образов. на начальном этапе оно осуществляется с помощью учителя: на вход подаются объекты и учитель указывает, к какому классу они относятся. если вероятность $P(\Omega_i)$ и плотность распределения $f_i(X)$ заданы надежно, то распознавание тоже будет надежным.

Лк №11

07.04.13

решающие правила:

1) Байеса

$$P(s_i/\bar{x}) = \frac{P_0(s_i)P(\bar{x}/s_i)}{\sum_{j=1}^m P_0(s_j)P(\bar{x}/s_j)} \quad (1)$$

P_0 - априорная вероятность появления объекта с s_i класса. $P_0(s_i) > 0$ $\sum P_0(s_i) = 1$

$$\arg \{ \max_i P(s_i/\bar{x}) \} = j \Rightarrow s_j$$

$$\begin{matrix} c_{11} & c_{12} & \dots & c_{1m} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mm} \\ c_{ii} = 0 & c_{ij} = 1 & & \end{matrix}$$

Байесова стратегия обеспечивает минимальные вероятности ошибок распознавания.

$$P_0(s_1)P(\bar{x}/s_1) > P_0(s_0)P(\bar{x}/s_2)$$

$$\omega \Rightarrow s_1$$

$$\omega \Rightarrow s_2$$

$$\underbrace{P_0(s_j)P(\bar{x}/s_j)}_a > \underbrace{P_0(s_i)P(\bar{x}/s_i)}_b$$

$$\omega \Rightarrow s_j$$

$\alpha(i, j) = a - b$ - дискрим. функция.

$$\alpha(i, j) > 0 \Rightarrow s_j, \quad j \neq i$$

правило минимальных расстояний:

определяется эталон - как усредненный вектор признаков. входной объект приписывается к тому классу, чей эталон ближе.

$$d_{ij}(\bar{x}) = d_i(\bar{x}) - d_j(\bar{x}) = \bar{x}^T(m_i - m_j) - \frac{1}{2}(m_i - m_j)^T(m_i - m_j) = 0, \quad i \neq j$$

2) по признаку ближайшего соседа



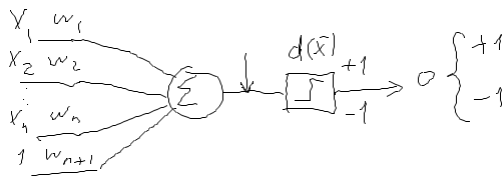
обводим гиперсферу и подсчитываем количество объектов из других классов (n-мерное пространство). каких объектов попало больше в гиперсферу определяет принадлежность к классу. недостаток: необходимость хранить большое множество данных.

3) корреляционное решающее правило

подходит для поиска отдельных изображений, заданных эталонами в поле зрения. макс.значению корреляции соответствует найденный объект. функции построены на основе персептрона.

$$d(\bar{x}) = \sum_{i=1}^n w_i x_i + w_{n+1}, \quad \bar{x}(x_1, x_2, \dots, x_n)$$

W - веса.



$$O = \begin{cases} +1 & d(x) > 0 \Rightarrow \omega_1 \\ -1 & d(x) < 0 \Rightarrow \omega_2 \end{cases}$$

$$d(x) = 0$$

$$O = \begin{cases} +1, & \sum w_i x_i > -w_{n+1} \\ -1, & \sum w_i x_i < -w_{n+1} \end{cases}$$

переобозначим $d(y) = \sum_{i=1}^{n-1} w_i y_i = WY$
 $y_i = x_i$

статистические методы, рассмотренные выше, предполагают знания о плотностях распределения признаков и о вероятностях поступления объектов.

рассмотрим 2 класса: Ω_1, Ω_2 . обучение осуществляет учитель. имеется две выборки. выбирают начальный вектор весовых коэффициентов $w = 1$

1. $Y(k) \in \omega, \quad w^T Y(k) \leq 0$

$w(k+1) = w(k) + cY(k)$

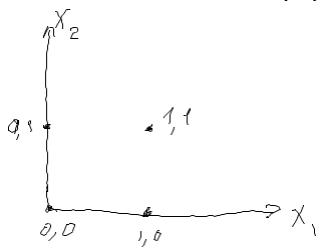
2. $Y(k) \in \omega_2, \quad w^T Y(k) \geq 0$

$w(k+1) = w(k) - cY(k)$

3. распознавание произошло правильно. $w(k+1) = w(k)$

доказано, что такой алгоритм сходится за конечное число шагов.

задача: необходимо сформировать выборку.



$Y^T = \{ \underbrace{(0,0,1), (0,1,1)}_{\omega_1}, \underbrace{(1,0,1), (1,1,1)}_{\omega_2} \}$

сформируем решающее правило

$Y(1) = (0,0,1) \in \omega_1$

$d(y) = w^T Y(1) = (0,0,0) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$

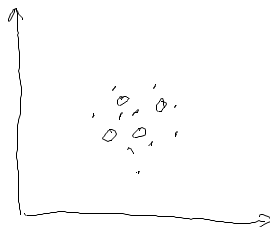
$w(2) = w(1) + 1 \cdot Y(1) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$d(Y) = w^T(2) Y(2) = (0,0,1) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1$

$w(3) = w(2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$d(Y) = -2y_1 + 1 = -2x_1 + 1; \quad -2x_1 + 1 = 0 \quad x_1 = 1/2$

рассмотрим линейно неразделимые 2 класса



алгоритм Уидроу-Хоффа

на каждом шаге обучения минимизируется ошибка между фактической и ожидаемой реакциями.

$$I(w) = 1/2(r - w^T Y)^2$$

r - желаемая реакция

$$+1 - \omega_1$$

$$-1 - \omega_2$$

алгоритм состоит в том, чтобы путем последовательных приращений к вектору w корректировать невязку $I(w)$ в сторону спуска (противоположную градиенту). тогда классификация будет безошибочной.

Лк №12

07.04.27

нейронные сети

1) назначение и исторические моменты

2) архитектура сети

3) обучение

4) особенности практического применения

1) назначение и исторические моменты

эффективными становятся методы, в которых дискриминантные функции строятся непосредственно в процессе обучения. наиболее яркими представителями таких систем являются нейронные сети (по сути они же - нейроконтроллеры, нейроморфные системы, многослойные самонастраивающиеся сети, модели распределенной параллельной обработки).

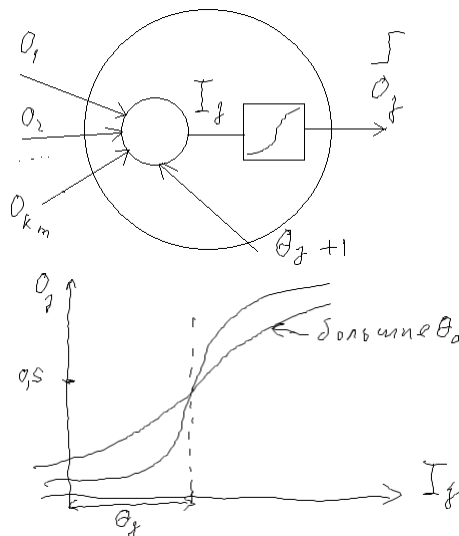
история ведет начало с 1943 года, с работы Мак-Каллаха и Питса, которые впервые предложили модель нейрона в виде двоичного порогового устройства.

2) архитектура сети

узел нейронной сети - это отдельный перцептрон.

закон построения сети: выход каждого узла одного слоя соединяется со входом каждого узла следующего слоя.

выходной слой содержит столько узлов, сколько классов выделено для распознавания.



сеть распознает вектор признаков $X(x_1, x_2, \dots, x_n)$. $X \in \omega_i$, если сигнал на выходе узла ω_i последнего слоя - высокий, а все остальные - низкие.

эту функцию можно записать аналитически:

$$h_j(I_j) = \frac{1}{1 + e^{-\frac{I_j + \Theta_j}{\Theta_0}}} \quad (1)$$

Θ_0 - параметр крутизны сигмоидальной функции

$$I_j = \sum_{k=1}^{N_k} w_{jk} O_k, \quad j = 1, 2, \dots, N_j \quad (2)$$

$$O_k = h_k(I_k)$$

для одного слоя - $N_j \times N_k$ коэффициентов. если это число умножить на количество слоев, получим общее количество коэффициентов.

с учетом сказанного, можно записать окончательное выражение для функции h_j :

$$h_j(I_j) = \frac{1}{1 + e^{-\frac{\left(\sum_{k=1}^{N_k} w_{jk} O_k + \Theta_j \right)}{\Theta_0}}} \quad (3)$$

суммарная ошибка выходного слоя:

$$E_Q = \frac{1}{2} \sum_{q=1}^{N_Q} (r_q - O_q)^2 \quad (4)$$

r_q - желаемая реакция, O_q - то, что получилось на самом деле.

суть алгоритма состоит в разработке обучающего правила, подобного Δ -правилу с предыдущей лекции. это правило должно корректировать веса в каждом слое таким образом, чтобы устремить значение функции ошибки (4) к минимуму. по правилу градиентного спуска будем корректировать веса пропорционально частным производным функции ошибки по этим весам. вводим коэффициент коррекции $\alpha > 0$.

$$\Delta w_{qp} = -\alpha \frac{\partial E_Q}{\partial w_{qp}} \quad (5)$$

$$O_q = f(I_q)$$

$$\frac{\partial E_Q}{\partial w_{qp}} = \frac{\partial E_Q}{\partial I_q} \frac{\partial I_q}{\partial w_{qp}}$$

$$\frac{\partial I_q}{\partial w_{qp}} = \frac{\partial}{\partial w_{qp}} \sum_{p=1}^{N_p} w_{qp} O_p = O_p \quad (6)$$

$$\Delta w_{qp} = -\alpha \delta_q O_p \quad (7)$$

$$\delta_q = \frac{\partial E_Q}{\partial I_q} = (r_q - O_q) h'_q(I_q)$$

$$\Delta w_{qp} = \alpha (r_q - O_q) h'_q(I_q) O_p = \alpha \delta_q O_p \quad (9)$$

все члены в выражении (9) или известны, или могут быть получены из наблюдения за нейронной сетью.

для обучения предъявляют один и тот же образ обучающей выборки на вход сети... минимизируется ошибка итерационным путем.

$$\Delta w_{pj} = \alpha (r_p - O_p) h'_p(I_p) O_j = \alpha \delta_p O_j \quad (10)$$

величины O_p и I_p могут быть получены в результате наблюдения. $\delta_p = (r_p - O_p) h'_p(I_p)$ - составляющая ошибки.

мы не знаем, какая должна быть реакция во всех промежуточных слоях.

$$\delta_p = -\frac{\partial E_p}{\partial O_p} \frac{\partial O_p}{\partial I_p}$$

производную от ошибки по выходу $\frac{\partial E_p}{\partial O_p}$ можно представить как производную сложной функции,

зависящую от ошибок выходного слоя:

$$\frac{\partial E_p}{\partial O_p} = -\sum_{q=1}^{N_q} \frac{\partial E_p}{\partial I_q} \frac{\partial I_q}{\partial O_p} = \dots = \sum_{q=1}^{N_q} \delta_q w_{qp}$$

мы полагаем, что $E_p = E_q$ (ошибка для текущего слоя совпадает с ошибкой для предыдущего).

$$\delta_p = h'_p(I_p) \sum_{q=1}^{N_q} \delta_q w_{qp}$$

δ_p можно найти, а соответственно и Δw_{pj} . т.е. мы нашли по методу градиента, как нужно изменять коэффициенты на предыдущем слое. аналогично работаем со слоем j , и т.д. таким образом, мы минимизируем ошибку в выходном слое, распространяя минимум ошибки на предыдущие слоя.

Лк №13

07.05.11

$$E_q = \frac{1}{2} \sum_{q=1}^{N_Q} (r - O_q)^2$$

N_Q - количество узлов в выходном слое

объект считается распознанным, если высокий сигнал появляется на одном из выходных слоев, а на остальных слоях - низкий сигнал.

$$W_{qp}^{(k+1)} = W_{qp}^{(k)} + \Delta W_{qp}$$

$$\Delta W_{qp} = \alpha \delta_q O_p$$

$$\delta_q = (r_q - O_q) h'_q(I_q)$$

$$W_{pj}^{(k+1)} = W_{pj}^{(k)} + \Delta W_{pj}$$

$$\Delta W_{pj} = \alpha \delta_j O_k$$

$$\delta_p = h'_p(I_p) \sum_{q=1}^{N_Q} \delta_q W_{qp}$$

итоговый алгоритм для произвольного j -го слоя:

δ_j вычисляется по двум ветвям: первая ветвь - для последнего слоя, а вторая - для предпоследнего p -го слоя, но суть остается точно такой же для всех слоев.

$$1. \delta = (r_j - O_j) h'_j(I_j)$$

$$2. \delta_j = h'_j(I_j) \sum_{p=1}^{N_p} \delta_p W_{jp} \quad j = 1, 2, \dots, N_j$$

замечание:

если в активирующей функции (3) положить, что $\Theta_0 = 1$, то записи упрощаются:

$$h'_j(I_j) = O_j(1 - O_j)$$

$$1. \delta_j = (r_j - O_j) O_j (r - O_j)$$

$$2. \delta_j = O_j (1 - O_j) \sum_{p=1}^{N_p} \delta_p W_{jp}$$

выводы:

нейронная сеть применяется когда нету никакой статистики (функции распределения, вероятности), но особенно - когда классы входных объектов неразделимы

Converted on 11.05.2007, 22:28 using TeXconvert and DocMerge scripts v2.0 (c) 2006-2007 Chervov Dmitry aka Deathdemon >:)
Последние конспекты всегда можно найти по адресу <http://deathdemon.int.ru/lectures/>