❖ *Тема 1. Обработка на изображения и разпознаване на образи. Основни понятия. Разпознаване на двумерни образи. Интензитет и цвят. Видове приложни задачи. Разпознаване на текст и графика. Компютърен анализ на документи. Биометрика и разпознаване.*

Дефиниция. *Образът* е отражение на обект от реалния свят върху физически носител.

Дефиниция. *Интензитет* наричаме свойството на обект да отразява светлината в определен спектър. В зависимост от това в коя част на спектъра попада отразената светлина, ние възприемаме светлината като цвят. Видимият за човешкото око спектър обхваща интервал с дължина на вълната около 390 до 700 nm.

Дефиниция. *Обработката на изображения* е всяка форма на обработка на сигнали, при която когато постъпи входно изображение, изходът от обработката трябва да бъде или образ или множество от характеристики или параметри, свързани с входното изображение.

Дефиниция. *Разпознаването на образи* е подразделение на обработката на изображения, чиято цел е да определи по данните за изображението (с които разполага) дали изображението съдържа обект или определена характеристика.

За човек е лесно да определи това, но се оказва, че не е толкова лека задача за компютъра (който се явява нашето измервателно устройство). Съществуват много методи за решаването на този проблем и те се справят добре за определени обекти, като например за геометрични обекти, човешки лица, ръкописни символи, печатни символи и др.

Нека имаме множество от образи **I** , които са от един и същ тип.

I = {образ1, ..., образN} и множество от реални обекти O, с които работи нашето измервателно устройство. На всеки обект от множеството O съответства образ от множеството I. Тоест на лице е изоморфизъм.

Образите са набор от признаци на реалните обекти. Можем да класифицираме обектите по определено тяхно качество или признак в определени групи. Например, можем да ги разделим в група на лоши и добри спрямо определен техен признак. Група от множеството на обектите съответства на група в множеството на образите. Така получаваме разбиване на множеството I на групи, които наричаме класове. Разпознаването на образи е задача за класификация.Така, ако получим като вход образ, трябва да определим към кой клас принадлежи.

Изображението е статичен обект, носещ информация за реалния обект. Двумерните изображения се представят като двумерна матрица, чиито клетки наричаме пиксели.

➢ Системи за оптично разпознаване на образи (OCR - Optical Character Recognition)

OCR системите се използват широко за разпознаване на символи (графични знаци, йероглифи, букви, цифри, завъртулки и т.н.) Целта им е да превръщат ръкописен, принтиран или печатен текст в машинно кодиран текст. Компютъризираните текстове могат да се изполват за съхраняване в компактен вид, по- удобни са за електронно търсене, визуализиране и изполване при машинни обработки като Text-To-Speech например.

Когато OCR системата получи сканирана страница с текст, разпознаването не може да бъде направено веднага. Необходимо е да се извърши предварителна обработка и да се декомпозира задачата на подзадачи. След като се обходят всички пиксели на изображението първо трябва да се отделят редовете, после думите и накрая отделните символи.

Следващата стъпка е да се определи какъв е символът - дали е цифра, определен знак, буква и т.н. Ако е буква, от кой език е?

Буквите формират множество от класове. Ако разглеждаме буквите от английската азбука, тогава мощността на множеството на класовете е 26, за българския език броя класове ще бъде 30, за китайския – над 3000.

За да се класифицира даден символ към кой клас принадлежи – се проверява в таблица от предварително зададени символи, например ASCII.

При намиране на съответствие в таблицата, символът се смята за разпознат. Ако входният текст е печатен, OCR системата ще трябва да вземе предвид шрифта и големината на текста при разпознаването и все пак това е решим проблем, т.к. има бази данни с шрифтове, с които да се сравнява.

Ако текстът е ръкописен, задачата се усложнява от редица характеристики: наклонът на текста, почерк, ширината на написаното и др. Задачата за разпознаване става по-трудна за OCR системата, т.к. се увеличават графичните признаци, които трябва да се вземат под внимание при сегментирането на отделните букви.

- **Компютърен Анализ на Документи** – използва се при кодиране на документи за компютърна обработка

При анализа на ръкописен текст, чрез почерка или подписа може да се определи съответната личност –автор на текста. Това е изключително важно в някои институции, където е необходимо да се удостовери авторът - например в банките, когато става въпрос за електронни транзакции. Затова OCR системите са полезни в този случай и намират приложение при компютърния анализ на документи.

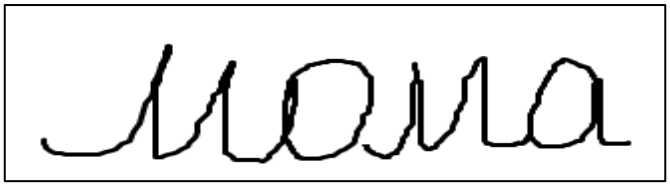
Алгоритъм за разпознаване на полиграфски текстове от OCR:

1) Сканиране на изображението
2) Декомпозиране, състоящо се в сегментиране на отделните текстови редове, думи и накрая символи
3) Същинско разпознаване на символа
4) Композиране на решенията
5) Край

Графеми наричаме най-малките семантично различими единици при писмения език. Пример за графеми са отделните „ченгелчета“ и „бастунчета“, които изграждат буквите.

Първия етап от алгоритъма за разпознаване е предварителна обработка на изображението, включваща сканиране. След сегментирането до думи, следва разбиване на графеми, за да се определят буквите.

Нека имаме сканирано изображение на български текст:



Броя класове от букви е 30 и нека сме разпознали първа буква „М“. OCR не знае, коя е следващата и ще направи всички възможни сравнения – тоест наивно ще проверява за принадлежност към всеки от 30-те класа, поради информационното излишество. Същинското разпознаване може да допусне грешка поради излишъка/шума. Работата може малко да се опрости като се направи post-processing.

Известно е, че подреждането на отделните букви в думите се осъществява по синтактични правила. Така, ако е разпозната първа буква „М“ на думата, ще се знае, че след нея може да има само ограничено подмножество на буквите в българския език. Така се намаля броя класове. Ако се изполва речник от действително ползваните думи в българския език, ще останат по-малко на брой смислени резултати, което е още едно подобрение. И така се продължава докато се разпознае и последната буква.

❖ Биометрика

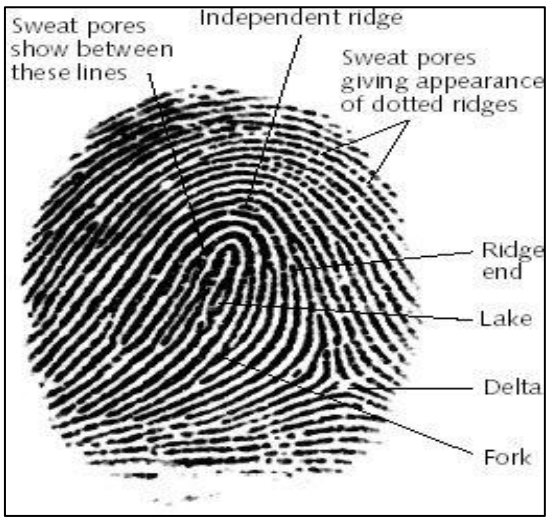
Биометричното разпознаване е система за разпознаване на хора на базата на специфични физиологически или поведенчески черти, които притежават. Физиологическите характеристики са по-надеждни.

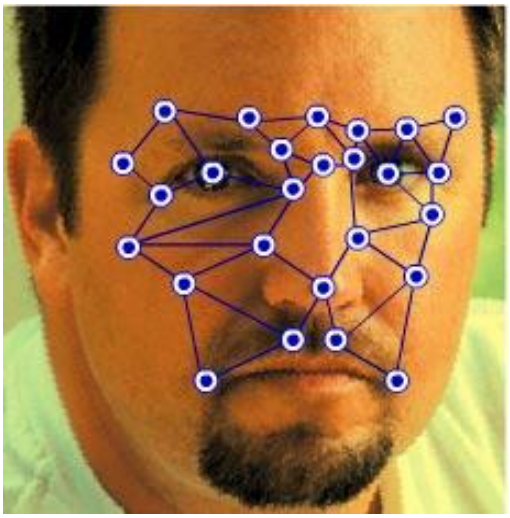
Обикновено има 2 основни цели: потвърждаване и идентифициране на даден човек.Идентифицирането включва сравнение на биометричната информация с шаблони на всички хора в базата от данни. При потвърждаването сравнението става само с ограничен брой шаблони (тези за които се твърди , че е личността).

Разпознаването в този случай може да се извърши по различни характеристики : по лице, уши, отпечатък от ръка, пръстов отпечатък, походка, ирис, глас, подпис и др.

За най-надеждно се смята разпознаването по лице, пръстов отпечатък и ДНК.

Пръстовия отпечатък включва гребени и бразди, разположени на върха на всеки пръст. Разпознаването по този признак за идентифициране изисква директен контакт на пръста с четящо устройство. Сравняват се определени сегменти от направения отпечатък и даден шаблон.





Разпознаването по лице се базира на местоположение и форма на очите, веждите, носа, устните и брадичката и връзките между тях.

ДНК анализа се смята за най-точната биометрика. Това е код, уникален за всеки човек. Изключения може да има само при едноядчните близнаци. Разпознаването по ДНК включва сложни химични взаимодействия, извършвани от специалисти, като се сравнява по всяка двойка хромозоми.

❖ *Тема 10. Вероятностни и статистически методи за разпознаване. Подход на Бейс. Гаусово разпределение и разстояние по Мехаланобис. Методи за разпознаване по разстояние*

В тази тема ще бъде разгледано приложението на вероятностните и статистическите подходи при определянето на принадлежност на образец x към даден клас C.

Имаме множество W -> E^n (n-мерно Евклидово пространство)

Разбиваме множеството W на класове: $\cup_{i=1}^{N} Ci = W$: $C_i \cap C_j = \emptyset$

P ще наричаме вероятностна функция такава, че P: W-> [0..1] и ако отговаря на няколко аксиоми:

A1) $0 \leq P(x) \leq 1$, за всяко x от W

A2) P(W) = 1

A3) P(A+B) = P(A)+P(B), $A \cap B = \emptyset$

$P(\cup A_i) = \sum_{i=1}^{N} P(A_i)$

A4) $A \cap B \neq \emptyset$ => P(A, B) = P(A)+P(B)+P(A∩B)

Събитие A: x да е n-мерен вектор от света W

Събитие B : избор на класа на x

Тоест идеята е при постъпване на образец x, нашето разпознаващо устройство да определи към кой клас C_i принадлежи.

Използваме правилото на Бейс за разпознаването:

$P(C_i \mid x) = P(x \mid C_i) * P(C_i) / P(x)$,

Където $P(C_i)$ – априорна вероятност за появата на образец от клас C_i ,

$P(x \mid C_i)$ – вероятност за поява на образец x, ако той е от C_i ,

а $P(C_{i|}\, x)$ е вероятността x да е от C_i .

Според правилото на Бейс за намирането на максималната апостериорна вероятност x да е C_i, трябва y=Argmax{ $P(C_{i|}\, x)$ }, т.е $P(C_{i|}\, x) \geq P(C_{j|}\, x)$, $j \neq i$, j = 1..N

- **Основна схема на вероятностния класификатор**

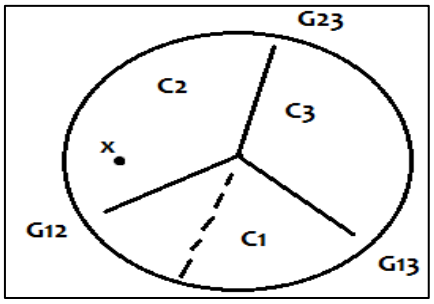
За всеки образец x можем да намерим максималната вероятност $P(C_{i|}\, x)$, тоест ще получим най-добра оценка за принадлежността на x към клас.

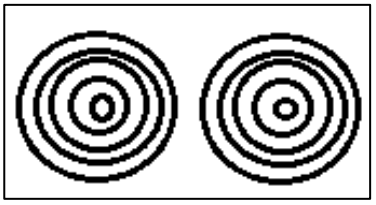
Нека приемем, че всеки един от класовете е Гаусово разпределен, тогава съпоставяме на всеки клас C_i ->функция $g_i(x)$ указваща дали входния образец се намира в Гаусиан i.

Нека имаме разделящи функции между отделните класове

$g_i(x)$ {>,=,<} $g_j(x) \Leftrightarrow G_{ij}(x)=g_i(x)-g_j(x)=P(x|\ C_i)$ - $P(x|\ C_j)$ {>,=,<}0

Ако $G_{23}(x) > 0$ => по-вероятно е x да е в C_2, в противен случай C_3 е по-вероятния клас(Фиг. 1)





Фигура 1 Фигура 2

Ако разглеждаме гаусианите отгоре, те ще приличат на изолинии, отговарящи на определена вероятност (Фиг.2). Според предположението ни класовете са нормално разпределени около центъра.

$$P(x|\ C_i) \sim N(\mu, \sum, x) = [\ (2\pi)^{n/2}\ (\sum)]^{-1} * \exp(-\frac{1}{2}(\mu-x)^T \sum^{-1}(\mu-x)),$$

където n – определя размерност на пространството, а $\sum = \| \sigma^{ij} \|$ е корелационна матрица (n x n) и

$(\sigma^{ij})^2 = \sum (\mu_i - x_i)*(\mu_j - x_j)$ за x_j от X_j , x_i от Xi , $x_i \neq x_j$

Пресечната линия $G^x = N(\mu_1, \sigma, x) - N(\mu_2, \sigma, x)$ за x от Евклидова равнина. В едномерния случай G^x ще бъде пресечна точка, за двумерния – квадратна крива, а ако е ротационен гаусиана (има форма на овал), тогава G^x ще е симетралата.

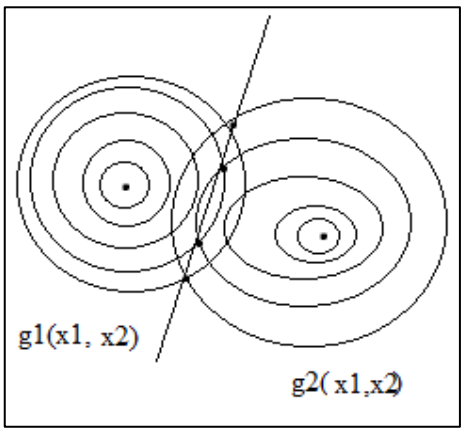
$\sum = \| \sigma^{ij} \|$ ковариационната матрица е такава, че $\sigma^{ij} = E[(\mu - x) (\mu - x)^T P(x)]$

и ако тя е диагонална $\sum = \| \sigma_{11}^2 ..... 0 \|$

$\| 0 ....\sigma_{nn}^2 \|$

тогава разпределението става ротационно симетрично във всички n направления и сеченията са окръжности.

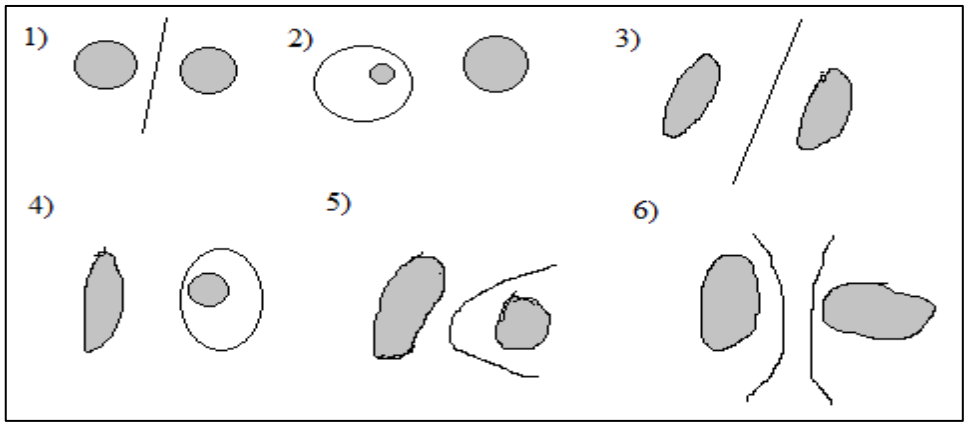
Нека сега разгледаме 2 класа/гаусиана за n=2. (Фиг. 3 отдолу)



Вида на дисперсиите σ_1 и σ_2 определя вида на разделящата ги повърхност. Ако $\sigma_1 = \sigma_2$ => всички пресечни точки лежат на 1 права и тя е симетралата.

На Фиг. 4 могат да се видят някои случаи на разделяне, в зависимост от отклоненията на 2-та класа (които са защриховани).

Следващата стъпка за решаването на нашата класификационна задача е да се определи от коя страна на разделящата повърхност попада входния образец и кой клас е по-сходен с него. Вероятността той да попадне в класа на прилежащата полуравнина е по-голяма.



Фиг.4

Равните вероятности от 2-те разпределения отговарят на еднакви разстояния до центровете на класовете им, измерени спрямо Евклидовото разстояние. На практика може да се окаже, че въпреки еднаквата отдалеченост, образеца се намира по-близо до единия клас. Причината за това е вида на разпределението. Например, ако точките от 1-я клас са групирани в овал, а точките от 2-рия равномерно по всички направления в идеален кръг, тогава въпреки, че образеца е на равно разстояние от центровете на 2-та класа, той би трябвало да е по-сходен с точките от 2-рия клас (точките от 1-вия клас се разпределят повече на височина отколкото на ширина). И така се оказва, че Евклидовото разстояние не ни върши работа. Затова се налага въвеждането на ново разстояние, което да вземе предвид този факт.

- **Разстояние на Мехаланобис**

$M(x)$: $r_i(x) = (\overrightarrow{\mu_i} - \vec{x}) * \sum^{-1} * (\overrightarrow{\mu_i} - \vec{x})^T$

$P_i(x) = C_i * \exp^{-r_i(x)}$

Когато класовете са с еднаква ковариационна матрица, тогава разстоянието по Мехаланобис = Евклидовото разстояние.

M(x) прави изкривяване на Евклидовото разстояние, за да отчете вероятностите. С негова помощ пространството може да се разбие дори при несиметрични разпределения.

M(x) не е метрика, тоест при него има понятие разстояние до центрове, но не и разстояние между две точки.