

1. Основни понятия. Класификация на системите за компютърно зрение.

Човешкият труд се разделя на физически и умствен.

Енергетични машини : ползват се за автоматизация на физическия труд.

Информационни машини : използват се за автоматизация на умствения труд.

Роботехнически системи : идват от Робот. Робота е конюнктурно понятие свързано с изкуствения интелект (ИИ).

Изкуствен интелект (Нилс Нилсон) : Известно е пространство решения на дадена задача, необходимо е да се намери поне едно от тези решения. Най-добре е да се намери най-ефективното решение. Развиват се методите за търсене на решения. Предефиниране на определението на Нилсон :

Създаване на програми, които реализират функции на човешкия разум.

Развиват се експертни системи на основата на това определение.

Една система е ИИ, ако отговаря на следните 5 критерия :

1) Изгражда ли системата модел на околната среда?

- 2) Използва ли системата този модел на околната среда за изграждане на план на действие в тази среда?
- 3) При изпълнението на плана за действие анализира ли системата алтернативни възможности?
- 4) Може ли системата да предефинира плана за действие, ако той води към непредсказуеми състояния за околната среда?
- 5) Използва ли системата минал опит за разширение и подобряване на модела на околната среда?

Основното тук е изграждане на модела на околната среда. Целта на КЗ е да се изгради модела за наблюдаваната околна среда. Ние трябва да извлечем информация за тази околна среда, за да постигнем тази цел.

Човек възприема най-много (90%) от входящата информация чрез зрението си. На обонянието и вкуса се падат по около (4%). Останалите проценти са за осезанията.

Околна среда : съвкупност от обекти, попадащи в полето на наблюдение(възприемане) от съответната видеовъзприемаща апаратура.

Визуална сцена : двумерна картина на съвкупността от двумерни или тримерни обекти попадащи в полето на наблюдение на видеозаснемащата апаратура.

Класификация на СКЗ според

- 1) Област на приложение : визуален контрол, медицина, авионаблюдение,

космически наблюдения.

- 2) Сложност : прости, със средни възможности, с големи възможности.
- 3) Размерност : двумерни и тримерни.
- 4) Тип на изображението : двоично, полутоново, цветно.
- 5) Сложност на задачата : центрирани детайли, детайли в разширен безпорядък, детайли в бункер.
- 6) Използвани средства : телевизионни камери (видикони и твърдотелни камери), мултиспектрални скенери, компютърни системи.
- 7) Обработка на изображението : последователна, паралелна.

Двумерни СКЗ : когато става дума за обработка на двумерни обекти.

Класификация по типа на изображението :

- 1) двоично изображение : използват се само 2 нива – черно (1)/бяло (0).
- 2) полутоново изображение : използват се повече от 2 нива. Обикновено 256 нива на сивото. Това изображение дава по-добра представа за формата на обектите във визуалната сцена.

3) цветно : използват се 3 цветови канала (син, зелен и червен).

Степента на сложност нараства както с преминаване от двоично, през полутоново до цветно изображение, така и при преминаване от двумерни към тримерни сцени. Разделянето е на малка сложност, средна сложност и голяма сложност (цветно тримерно изображение).

Класификация според сложността на задачата :

- 1) Центрирани детайли : имаме конвейер с детайли. Детайлите не се препокриват.
- 2) Детайли в равнинен безпорядък : На конвейера детайлите могат да се закриват един друг.
- 3) Детайли в бункер : системата трябва да вземе конкретен детайл от бункера.

Класификация според използваните средства:

Телевизионните камери могат да бъдат черно-бели, полутонови и цветни.

Мултиплексорни скенери : измерват дължината на електромагнитното колебание.

Зрение : процеса на определяне какво се съдържа в наблюдаваната околна среда и къде се разполагат обектите в тази наблюдавана среда.

Трябва да се получи изображение на тази среда. То трябва да се представи по определен начин в системата. На базата на това изображение се определят обектите. Системата трябва да бъде обучена да разпознава обекти.

Трябва да се определят отношенията между обектите. Определя се центъра на тежестта на всеки разпознат обект и се определят техните отношения. Може да има неразпознати обекти.

Целият този процес изисква съхраняването на голям обем информация.

Система за КЗ : представлява видеовъзприемаща (или система за възприемане на видеоинформация), интерфейс между тази система и компютърната система.

От друга страна за с-ма за КЗ са необходими методи, алгоритми и програми за обработка на изображения;определяне на признаци на обектите; построяване на формално описание на наблюдаваната сцена; разпознаване на обектите; изграждане на околната среда

2. Физически основи на КЗ

Всяко тяло, което има температура над абсолютната нула генерира електромагнитно излъчване. Излъчването се движи със скоростта на светлината. С увеличаване на вътрешната температура на тялото се увеличава и количеството на излъчваната енергия, както и излъчването преминава от по-дълги към по-къси вълни.

Ако се вземе по абциса дължината на вълната то имаме следните области (започва се от най-късите вълни):

- 1) Гама лъчи

- 2) X лъчи (рентгеново излъчване)
- 3) Ултравioletови лъчи
- 4) Видима светлина
- 5) Инфрaчервена (отразена и термална) светлина
- 6) Микровълни
- 7) Радиовълни

Видимата област на електромагнитния спектър е в рамките 400-700 нанометра. За нас основен източник на електромагнитно излъчване е Слънцето. То излъчва в много честотни диапазони : от 1 пикометър до 30 микрометра. Максимум на излъчване е около 500нанометра. Температурата на Слънцето е 6000K.

Земята също излъчва електромагнитна енергия при средна температура 300K. Нейният максимум обаче не е в във видимата част на спектъра, а е на 9,7nm.

Видимата област се разделя на 3 подобласти :

- 1) синьо (blue) – 400-500nm

2) зелено (green) – 500-600nm

3) червено (red) - 600-700nm

Човешкото зрение може да се разглежда като дистанционно наблюдение (измерване от разстояние). Ако при него имам естествен източник на електромагнитни колебания, то тогава системата е *пасивна*. Ако източника на светлина е изкуствен – системата е *активна*.

Светлината се дели на :

1) ахроматична : не съдържа цвят, изменя се от черно до бяло през нива на сивото. Основна нейна характеристика е интензитета на отразената светлина.

2) хроматична : характеризира се със цвят. Анализа на цветни изображения става на базата на трите основни цвята (RGB). Всеки цвят може да се получи като комбинация от компоненти по всеки един от тези основни цветове.

Трите основни цвята образуват тримерно пространство на цветовете. Ако стойностите са дискретни и ограничени, пространството е куб. Началото на координатната система е (0,0,0). Колинеарните вектори в това пространство представят един и същ цвят, но се различават по яркостта.

За да получим всеки един от тези 3 цвята от едно хроматично изображение се използват филтри.

Интегрални уравнения за реализиране на филтрите :

$$R = \int E(\lambda) R'(\lambda) d\lambda$$

$$G = \int E(\lambda) G'(\lambda) d\lambda$$

$$B = \int E(\lambda) B'(\lambda) d\lambda$$

$R'(\lambda)$, $G'(\lambda)$, $B'(\lambda)$ са функции на разпределение за съответните основни цветове.

Във всяка от трите части на видимата част на спектъра, кривата на спектралният състав има изпъкнала форма (това е фигура). Максимумите в отделните части са определени да са : 435,8nm (blue), 546,1nm(green), 700nm (red)

Има и други системи за анализ на цветове и изображения.

Начина на възприемане на човешкото око се доближава повече до системата HSI(Hue, Saturation, Intensity)

- 1) цветови отенък (Hue)
- 2) наситеност (Saturation)
- 3) яркост (Intensity) – пропорционална на интензитета на отразената светлина.

Преминаването от едната система в другата става лесно чрез използването на следните формули :

$$I = R + G + B ; H = (G - B) / (I - 3B) ; S = (I - 3B) / I$$

Описание на различните части от електромагнитният спектър :

- 1) Гама лъчи (
- 2) Рентгенови лъчи (0.03 – 30 nm) : Напълно се поглъща от горните слоеве на атмосферата.
- 3) Ултравиолетови лъчи (0.003 – 0.4 μm) : Вълните с дължина
- 4) Фотограска УВ област (0.3 - 0.4 μm) : Минава през атмосферата. Улавя се със фоточувствителен филм. Атмосферното разсейване е много силно.
- 5) Видима област (0.4 – 0.7 μm) : Изобразяване с филми и фотодетектори.
- 6) Инфрачервена област (0.7 – 100 μm) : При взаимодействие с материята (отражение) променя дължината на вълната си. Съществуват атмосферни прозорци на пропускане.
- 7) Отражена ИЧ (0.7 – 3 μm) : Съдържа информация за топлинните свойства на материята. Лентата (0.7 – 0.9 μm) се улавя с филм и се нарича фотографска ИЧ област.
- 8) Термална ИЧ (3 – 5 μm), (8 – 14 μm) : Принципни атмосферни прозорци.
- 9) Микровълни (0.1 – 30 cm) : По-дългите вълни проникват през облаците. Изображение може да бъде получено по активен начин.

10) Радиообласт (> 30 cm) : Най-дългите вълни в електромагнитният спектър.

3. Апаратен модел на СКЗ

Апаратен модел за възприемане на чернобяло и полутоново изображение.

Изображението се заснема с телевизионна камера. Получените данни се подават на входа на АЦП откъдето излиза кадровата памет. Тази кадрова памет влиза в оперативната памет на СКЗ. АЦП и кадровата памет се управляват от УУ. Всеки от тези блокове се реализира автономно.

Телевизионните камери биват : видикони и твърдотелни

Принцип на работа на **видикон** :

Видикона има отпред светлочувствителен слой върху, който се изобразява сцената. Под този слой се намира слой от свързани със земя кондензатори и активни съпротивления. В задната част на видикона се намира устройство, което генерира електромагнитен лъч. Този лъч се насочва от друго устройство. Лъча обхожда всеки пиксел от предната част на видикона. Когато той попадне върху даден пиксел, последният отдава натрупаният си заряд и така се отчита интензитета на светлината попаднал в пиксела. Съвкупността на интензитетите във всяка точка ни дава цялата картина ($I = f(x, y)$). Изменението на тази функция е пропорционално на падащата светлина.

Принцип на работа на **твърдотелна камера (CCD – charge couple device)** :

Изградено от фоточувствителни елементи, които представляват MOS (metal, oxide, semiconductor) транзистори. Всеки фоточувствителен елемент се изгражда като двойка

(по двете координатни оси). Съответно имаме регистър по X и регистър по Y. Цялата камера представлява матрица от зарядно свързани кондензатори. Когато трябва да се прочете определен ред, той се активира и натрупаните заряди постъпват в регистър по Y. Когато трябва да се прочете колона се използва регистъра по X. След прочитане в регистъра тези стойностите на интензитетите се подлагат на АЦ преобразуване.

Предимства на **CCD** пред видикон камера :

- 1) Фоточувствителните елементи са с добра геометрия.
- 2) Отношението сигнал/шум при CCD е по-добро отколкото при видиконите.
- 3) Захранващото напрежение при CCD е по-малко и следователно понижена консумация на енергия. Това комбинирано с много по-малките размери дава възможност за вграждане в малки системи.

Пространствената дискретизация зависи от разделителната способност. Тя може да е 256x256, 512x512 и др.

Ако се използват **m** нива на сивото и изображението е с размерност **NxN**, то обемът памет заеман от един кадър е

$$H = N^2 \cdot m \cdot \log_2 m \text{ (bit)}$$

При цветно изображение имаме 3 канала за всеки от трите основни цвята. Съответно сложността на задачата се увеличава 3 пъти. От цветната телевизионна камера сигнала се предава на 3 филтъра (за всеки цвят). От всеки от тези филтри сигнала се предава за дискретизация на АЦП (отделен за всяка верига). След което данните се записват в

кадровата памет (отделна за всеки цвят).

Мултиспектрален скенер, елементи :

- спектромер : определя дължината на електромагнитно излъчване
- детектор : определя нивото на електромагнитната енергия в определен интервал.

Имаме толкова ленти, колкото са ни нужни за анализа на изображението.

На цифрова обработка се подлага винаги дискретизираната функция след АЦП.

4. Информационен модел на СКЗ

Информационният модел дава етапите през, които се преминава входното изображение, за да се получи модел на околната среда. Етапите за обработка са следните.

- 1) Квантуване. Филтрация. (вход : $f_1(x,y)$, изход : $g(x,y)$)
- 2) Отделяне на ръбове и граници (вход : $g(x,y)$, изход : $b(x,y)$)

- 3) Отделяне на признаци. Сегментация.
- 4) Построяване на описание.
- 5) Интерпретация. Разпознаване.
- 6) Изграждане на модел на околната среда.

Тези стъпки могат да се разделят по два начина на нива. Едната система е с 2 нива (ниско/високо), другата система е с 3 нива (ниско/средно/високо). При първата, ниското ниво са стъпки 1,2 и 3, останалите са във високото. При втората система разделянето е на по 2 стъпки в ниво в същата посока както при първата система.

На входа на информационният модел постъпва функция $f_1(x,y)$, която се квантува до функция $g(x,y)$. Върху тази функция (дискретизирана) се прилага филтрация с цел подобряване на първоначалното изображение.

Сегментация : определяне на формата на различните обекти във сцената. $b(x, y)$ е двоично изображение, т.е. стойностите на функцията са само 1 или 0. Границите на ръбовете са означени с 1, а всичко останало е 0 (процес на отделяне на ръбовете).

Интерпретация, разпознаване : Системата трябва да е предварително обучена, за да може да прави разпознаване.

Накрая, ако не са разпознати някои обекти, те ще участват в модела, но ще трябва да се дообучи системата за да ги разпознава следващ път. Моделът се създава на база отношението между обектите в сцената.

На основание на обработката се получават БД и БЗ.

Бази данни. Бази знания. Управление.

Информацията, която се получава в компютъра се дели на :

- декларативна : това са данните, които се обработват

- процедурна : множество програми, по които става обработката.

Особености на една СКЗ по отношение на данните :

- 1) Обемът данни, които се обработва в една СКЗ е много голям. Следователно капацитета на БД е много голям.

- 2) Има йерархичност на данните.

Ако информационните единици в БД придобият следните свойства те стават БЗ :

- 1) Вътрешна интерпретируемост : към ИЕ можем да се обърнем с ключ. С вторични ключове можем да се обърнем към определени полета на ИЕ.

- 2) Структурност : едни ИЕ да могат да се включват в други ИЕ. (част – цяло, обект – клас)

3) Свързаност на ИЕ : определя конкретни отношения за ИЕ (“пред”, “над”, “причина – следствие”, “аргумент – функция”). Тези отношения се отнасят към декларативната информация. “аргумент – функция” се отнася към процедурната информация.

4) Семантична свързаност : отнася се за конкретни предметни области. Близки ИЕ се анализират с цел формиране на някакво понятие.

5) Активност : анализирайки, конкретните данни и отношенията между тях, могат да се предизвикват конкретни действия.

Представянето на знанията в СКЗ трябва да отговаря на следните условия :

1) Трябва да можем да представяме в системата различни видове знания.

2) Системата трябва да води до ефективно съхранение на данните.

3) Лесна манипулация с елементите на знанието.

4) Знанието трябва да бъде така представено, че да бъде лесно разбираемо за човека.

Начини за представяне на знанието :

1) Използват се портретни модели (iconic) – изходното полутоново или цветно изображение. Тези портретни модели са много чувствителни по отношение на :
а)точката на наблюдение б)осветеността. Тези портретни модели изискват голям обем памет.

2) Графови модели : възли, които съответстват на отделни елементи на сцената (стени, обекти) и клонове (ребра) свързващи възлите. Ребрата указват връзките между елементите в сцената. Графовите модели са по-малка чувствителност към осветеността и ориентацията на обектите в сцената.

3) Семантични мрежи : развитие на граф модела. Това е ориентиран граф, възлите на който показват елементи на сцената, а ребрата именувани (смислови) отношения.

Фрейми (Мински) : Фреймите са друг начин за представяне на декларативно процедурно значение. Фрейма представлява структура от данни за определяне на стереотипни ситуации. Като под ситуация се разбира визуален образ, събитие. Фреймовото представяне на знанието представлява ориентиран граф, като възлите на този граф представляват елементите на наблюдаваната сцена. Ребрата на графа дават отношенията между елементите. Всички възли на ориентираният граф на горни нива са определени, а на най-ниското ниво възлите са празни. Празните възли се запълват в зависимост от конкретната ситуация. Всеки възел може да бъде обръщане към подфрейм, при което обръщане могат да се активират процедури за изчисление на информация за конкретният възел или за даване на конкретна информация за него.

От всички представяния на знанието, най-универсално е фреймовото.

Управление на обработката на информацията :

1) “отдолу – нагоре” (управление на данните) : на входа на всеки етап постъпват данни, на изхода излизат резултати, които активират изпълняването на следващият етап и т.н.

2) “отгоре – надолу “ (управление на моделите) : прави се хипотеза, че при дадена обработка като краен резултат ще се получи резултат съответстващ на założеният. Недостатък е възможността да се направят голямо количество тестове преди да се достигне желаният резултат. Използва се често за анализ на тримерни сцени.

3) Обратка връзка : анализират се резултатите на изхода на всеки етап. Ако те са удовлетворителни се минава към следващият етап. Ако не са удовлетворителни се връщаме на входа на този етап и модифицираме обработката в някаква посока с цел получаване на различен резултат.

5. Филтрация на полутоново изображение по средна стойност. Анизотропна и рекурентна филтрация.

Хистограма на изображението : дава броят на пикселите имащи един и същ интензитет.

Хистограмата много често се използва за определяне на праг(ове) при обработка на изображение : за бинаризиране, за проверка на качеството на апаратурата за възприемане.

Филтрация по средна стойност : Използва се малко, локално, пространство наречено апертура. Апертурата обикновено е с размери 3x3 или 5x5.

$u(i, j)$ – входно цифрово изображение . $i = 1 \dots N, j = 1 \dots M$

Приемаме, че i, j е центъра на апертурата. Тогава за 3x3 имаме координати : $(i-1, j-1), (i, j-1), (i+1, j-1) \dots (i-1, j+1), (i, j+1), (i+1, j+1)$.

Сумират се стойностите на интензитите на всички пиксели попадащи в апертурата. Получената сума се дели на броят на пикселите и така се получава средна стойност.

$$S = \sum g(i+r, j+k)$$

Съществуват 2 вида филтрация по средна стойност :

1) анизотропна

2) рекурентна

При анизотропната филтрация се използват два масива. Входното изображение се намира в първият масив, а стойностите на пикселите изчислени по метода на филтрация по средна стойност се записват във вторият масив.

При рекурентната филтрация се използва един масив. Когато се изчисли нова стойност, на базата на метода, тя се записва на мястото на оригиналната. Както може да се забележи така се постига рекурентност. Рекурентната филтрация е по-удобна в смисъл на това, че е с понижение изисквания за количеството нужна памет.

Начина на сумиране по-горе може да се изрази, чрез матрица, всички елементи на която са единици. Разбира се съществуват и матрици, при които елементите на са единици, а други числа, като деленето става на база сумата на тези числа. По този начин определени позиции в апертурата получават по-голяма тежест при изчислението. Пример за такива матрици са : (111 121 111) и (121 242 121). Апертурата взета от изображението се умножава по матрицата, но не като матрици, а чрез друга операция наречена конволюция. Тя има свойствата асоциативност и комутативност.

Недостатък на филтрите по средна стойност е размиването на преходите в изображенията, но от друга страна се размиват и шумовете.

6. Медианна филтрация

При медианната филтрация на входното изображение, целта е да се премахнат случайни импулсни шумове. Най-често използваната апертура е 3x3, по-рядко се използва 5x5.

Принципа на работа е следният :

Обработка се изображението пиксел по пиксел. За всяко положение на апертурата се сортират стойностите на интензитетите на пикселите обхванати от апертурата. За стойност на пиксела в средата на апертурата се използва медианата на стойностите – стойността на позиция 5 (при 3x3), 13 при (5x5) от сортираните стойности.

По този начин импулсните шумове се премахват лесно тъй като те са с висока/ниска стойност и отиват в края/началото на сортираната редица. За съжаление чрез медианна филтрация ръбовете в изходното изображение се нащърбват.

Алгоритъма е бавен, защото за всеки пиксел се извършва сортиране. Едно подобрение е да се сортира, докато се намери петия по големина елемент и тогава да се спре.

7. Гаусови филтри.

На базата на Гаусовата функция за разпределение се конструират матрици, наречени Гаусови. Те обикновено са двумерни.

$$G(x) = e^{-x^2 / 2\sigma^2}, G(y) = e^{-y^2 / 2\sigma^2}; G(x, y) = e^{-x^2 - y^2 / 2\sigma^2}$$

σ - стандартно отклонение на Гаусовото разпределение. Сигма зависи от размера на апертурата (n). $n = 2.2^{1/2} \cdot \sigma$. Ако първият коефициент е 3, резултатите са още по-добри.

За конкретна стойност на n се изчислява σ .

13. Основни принципи и методи на разпознаването на образи

Етапи :

- 1) отделяне на признаци
- 2) описание на обектите
- 3) обучение на системата за разпознаване на образите, на базата на конкретни критерии и знание
- 4) етап на разпознаване

Видове обучение :

- 1) С учител : учителя дава правилата, алгоритъма за класификация на обектите по класове. Той също посочва и имената на класовете.

2) Без учител : на базата на множество образи, които постъпват на входа на системата (обучаващо множество), тя търси характерни особености, свойства. И на тяхна база да раздели входното множество на класове. За класовете назначава номера.

След като е преминал етапа на обучение на системата за тестово множество от образи се прави проверка. Ако резултати не са добри се провежда дообучение.

Образ : описанието, на който и да е обект от даден клас.

Клас : множество от образи имащи едни и същи свойства.

Признак : общо свойство за обектите от даден клас.

Класификация на образи : отнасянето на даден образ към даден клас на базата на приети критерии и знание.

Разпознаването на образи се разделя на 2 етапа :

- 1) Класификация на образите по класове
- 2) Идентификация на класовете, т.е. даване на имена.

Съществуват 3 принципа и 3 метода за разпознаване на образи:

1-ви принцип : Използване на еталони. За даден клас w_i се определя еталон e_i . Един клас може да има множество еталони, като те се номерират с горен индекс. Образът постъпващ на входа на системата се сравнява с всеки един от еталоните. Недостатък на този принцип е чувствителността към шумове във входните изображения.

2-ри принцип : Използват се множество общи признаци. За даден клас w_i се определят признаци (x_1, \dots, x_n) . Всеки клас се характеризира с n общи признака. За входните образи се определят множество от признаци и те се сравняват с множеството от признаци на всеки един клас.

3-ти принцип : Използване на клъстърни. На i -тия клас съответства конкретен клъстър K_i . А всеки клъстър се определя от множеството образи $w_i \rightarrow (a_1, \dots, a_n)$. Всеки клъстър има център. За образа постъпващ на входа на системата се изчислява разстоянието до центровете на всички клъстърни. Образът е причислява към клъстърния център е най-близо.

1-ви метод : Евристични методи : дават знанието на добър експерт за класификация на обектите за конкретна задача от дадена област. Използват се *1-ви и 2-ри* принцип.

2-ри метод : Математически методи : посочват конкретни формализми за класификация на образите. *Използват се и трите метода за разпознаване.*

3-ти метод : Структурно – лингвистични методи (синтактични) : за образите се отделят структурни признаци на базата, на които се прави описание на образа. Описанието се анализира с цел извличане на конкретна информация или разпознаване на входния образ. Използва се *2-ри принцип.*

14. Класификация на образи с използване на минимално разстояние.

Класификацията на образи с използване на минимално описание по своята същност е математически метод за разпознаване на образи.

Сходството между 2 образа зависи от “разстоянието” между тях. Разстоянието се означава с **d**.

$$d(P_i, P_j) = \| P_i - P_j \| \geq 0$$

Разстоянието между образите има следните свойства :

- 1) $d(P_i, P_j) = 0$ - пълно подобие
- 2) $d(P_i, P_j) = d(P_j, P_i)$ - симетричност
- 3) $d(P_i, P_k) + d(P_k, P_j) \geq d(P_i, P_j)$ - неравенство на триъгълника

Минимално разстояние за класификация : използване на еталонният принцип. За всеки образ се изчислява разстоянието до еталоните.

$$d(P_i, P_j) = \| P - L_i \| = (\sum_{k=1..n} (P_k - L_k)^2)^{1/2} - \text{Евклидово разстояние}$$

$d_{\min} = \text{Min}(d_1, \dots, d_n)$ – този образ се отнася за съответният клас.

Ако имаме двумерно пространство се търсят минимални разстояния в двумерното пространство.

15. Клъстеризация

Постановка на задачата : имаме множество от образи $\{a_1, \dots, a_n\}$. Всеки образ се характеризира от признака $\{x_1, \dots, x_n\}$.

x_1, \dots, x_n

}. Необходимо е входното множество от образи да се раздели в m клъстър(а) (групи) така, че всеки образ от даден клъстър да е по-близо до образите в този клъстър отколкото до всеки образ от останалите клъстъри. За реализацията на клъстеризацията се използва алгоритъм за **ефективна клъстеризация**, който се реализира като показател за качеството на клъстеризация. Този показател се изчислява по следната формула :

$J = \sum_{j=1..m} \sum_{a \in K_j} \|a - Z_j\|^2$

$$J = \sum_{j=1..m} \sum_{a \in K_j} \|a - Z_j\|^2$$

m – брой на клъстърите, които сме задали

K_j – клъстър съдържащ подмножество от образи

Z_j – вектор от средните стойности на елементите в клъстър K_j .

r – брой на образите в клъстър

J - показва сумата от разликите на образите до центровете на техните клъстъри.

При минимизация на тази сума се постига ефективна клъстеризация.

Алгоритъм за клъстеризация :

1) Определят се m центъра на клъстерите : $Z_1 \dots Z_m$. Обикновено се вземат първите m образа.

2) $i = 1$ - параметър на цикъла (итерации)

3) Определят се най-близките образи до центровете, т.е. формират се m клъстър K₁, ..., K_m

- това са клъстърите формирани на i-тата итерация.

4) За клъстърите получени в точка 3 се определят нови коиригирани центрове. Изхожда се от условието, че сумата от квадратите на разстоянията между новия център на зададеният клъстър и образите в него ще бъде минимална.

$$J_i = \sum_{K_j^i} \| a - Z_j^{i+1} \|^2, j = 1, \dots, m$$

5) Ако се получат едни и същи центрове за клъстърите в две последователни итерации се преминава към 6) иначе се връщаме към 3

6) Край

Задачата за клъстеризация се използва в медицината и други области за взимане на решения.

Алгоритъма за клъстеризация по същество представлява обучение без учител.