

15. Метод на собствените лица (собствените обекти) за разпознаване на физиономии (обекти) [39]. Линеен дискриминант на Фишер при разпознаването на лица [40].

(Виж ПР-част1 60-63 страници, както и Lecture_Notes_Besnik – Chapter_XX.docx)

На база метода на главните компоненти за редуциране на признаковото пространство е разработен значим метод за разпознаване на лица.

Лицата са:

- 1) Тримерни
- 2) Изменчиви – гримаси, с времето
- 3) Изменчиви – бради, коси, обици, очила

Различни подходи при разпознаването на лица:

- 1) Метод на точното еластично сравняване – почти не се използва за лица (само за кукли, изкуствени фигури и др.)
- 2) Геометрия на формата – тримерно възстановяване на формата на лицето (3D модел, текстури, цветове), трябва да се симулират мускули, гримаси... Има нужда от създаването на Аватар. Твърде сложен.
- 3) Метод на Невронните мрежи
- 4) Евристични подходи
- 5) Вероятности – по профил и анфас (генерира всички проекции в профил по анфас)
- 6) Метод на собствените лица(Eigen Faces)
Илюстрира приложението на метода на главните компоненти.

Класовете са дефинирани като клас от снимките на един човек. Нека в тестовите данни всеки клас да има по един представител и снимките да са $N \times N = N^2$. Сканираме ги и за всяка снимка получаваме по един N^2 -мерен вектор. Представяме така получените вектори като едно N^2 -мерно Евклидово пространство.

Нека имаме M на брой входни картинки.

Сега да намерим главните компоненти.

$$\lambda_1 > \lambda_2 > \dots > \lambda_M > \dots > \lambda_{N^2}$$

В $E^{(N^2)}$ имаме само M точки, които дефинират подпространство с размерност $M-1$, защото останалите собствени вектори са линейно зависими от другите.

Затоа вместо да изчисляваме $\Sigma_{[N^2 \times N^2]}^{(\text{лицата})}$ се опитваме да намерим линейно независимите собствени вектори. След това работим с матрицата именно на тези вектори - $\Sigma_{[M \times M]}^{\text{друга}}$. Намираме собствените вектори.

И така всяка една картинка се представя в M -мерно пространство и търсим по близост с M на брой числа.

Трябва да отбележим, че всички лица се нормализират в началото – по разстояние между двете очи и разстоянието уста – очи.

Анализ на главните компоненти

$$E(x) = \sum_{i=1}^N p_i x_i = \frac{1}{N} \sum x_i, \quad p_i = p(x_i) = \frac{1}{N} - const$$

$$var(x) = E \left\{ (x - E(x))^T (x - E(x)) \right\}$$

$$\sigma_x = \sqrt{var(x)}$$

$$cov(x, y) = E \left\{ (x - E(x))^T (y - E(y)) \right\}$$

$$c_{ij} = E \left\{ (x_i - E(x_i))^T (y_j - E(y_j)) \right\}$$

$$c_{ij} < 0 - x \text{ и } y \text{ имат различни посоки}$$

$$c_{ij} > 0 - x \text{ и } y \text{ имат еднакви посоки}$$

Ако е 0, то x и y са неколинеарни.

Нормирана ковариация:

$$p_{xy} = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

1) Векторите на лицата са от типа:

$$x_1 = \begin{bmatrix} x_{11} \\ \vdots \\ x_{N^2 1} \end{bmatrix} \in A = |x_1, x_2, \dots, x_{N^2}| = |a_{ij}|_{N^2 \times M}, \text{ където } i = 1 \div N^2, j = 1 \div M.$$

2) Среден вектор:

$$m = \begin{bmatrix} m_1 \\ \vdots \\ m_{N^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{M} \sum_{i=1}^M a_{1i} \\ \vdots \\ \frac{1}{M} \sum_{i=1}^M a_{N^2 i} \end{bmatrix}$$

3) Сега ще приложим нормализация на средния вектор от лице

$$A' = |a_{ij} - m_i|_{N^2 \times M}$$

4) Дисперсия на датата за лицата

$$var(a_{ki}) = E(a_{ki} - E(a_{ki}))^2, \text{ където } a_{ki} \in A.$$

$$cov(a_{ki}, a_{lj}) = E(a_{ki} - E(a_{ki}))E(a_{lj} - E(a_{lj})) = |c_{kl}| = \begin{cases} cov(A_k, A_l), \text{ при } k \neq l \\ var(A_k), \text{ при } k = l \end{cases}$$

$$C = A^T A_{(N^2 \times N^2)}$$

5) Търсим собствените вектори и собствените стойности

$$C = \lambda u \Rightarrow A^T A = \lambda u.$$

6) Редуцираме пространството

$$L_{(M \times M)} = AA^T$$

$$LA = AA^T A = \lambda(Au) = \lambda \bar{u}$$

$\Rightarrow \bar{u}$ – линейно не зависими, собствени вектори на L – M-броя

$\Rightarrow \lambda$ – собствените стойности на L – същите като на C.

Изчисляваме сега собствените вектори на L:

$\bar{u}_{M \times N^2} = A_{M \times N^2} u_{N^2}$ – което са координатите на лицето в подпространството $E^{(M)}$. (Проекцията на лицата в това евклидово пространство се наричат собствени лица).

7) Собствените лица:

Собствените лица са линейна функция на $\bar{u}$.

$$w = \Pr(x \rightarrow E^{(M)})$$

8) Разпознаване

Идва ново лице и се проверява проекцията му в $E^{(M)}$ до кой клас е най-близо.

$\varepsilon_0 \geq r_i > d(\bar{x}, x_i) = \min_{j=1 \div M} (d(\bar{x}, x_j))$ - в този случай лицето е разпознато, ако обаче разстоянието до най-близкия клас е по-голямо от някакъв определен радиус на класа – лицето не се разпознава. (непознато лице)

Линеен дискриминант на Фишер при разпознаването на лица

Много би било хубаво ако можехме, когато имаме нужда то редуциране на измеренията, използвайки линейна проекция, да запазим линейната граница между класовете. Това би било особено полезно когато искаме да постигне независимост от осветеността. За целта можем да използваме информацията от обучаващото множество. Линейния дискриминант на Фишер спомага за това да оконтурим обектите в рамките на един клас. Както знаем идеята е да максимизираме отношението между-класова плътност и вътре-класова плътност, където:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T$$

$$S_W = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i) (x_k - \mu_i)^T$$

където μ е осредненото изображение на класа X_i , N_i е броя обучаващи примери в класа X_i и c е броя класове.

Ако S_W не е единична матрица, оптималната проекция би била:

$$W_{opt} = \operatorname{argmax}_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, \dots, w_m].$$

Тук w_i е множеството от обобщените собствени вектори на S_B и S_W отговарящи на m -те най-големи собствени стойности - λ_i . Важно да отбележим, че имаме максимум $c-1$ ненулеви обобщени собствени вектора.

Проблеми би при този метод при разпознаване на лица, защото обикновено S_W е единична, поради факта, че $N \gg c$, защото броя на картинките е обикновено много по-малък от броя пиксели. Заради това се прави комбиниран подход между Метод на главните компоненти и линейния дискриминант на Фишер – който се нарича Fisher Faces. При него първо се прави метод на главните компоненти за редуциране на пространството до N -с и после се прави метода на Фишер за редуциране до $c-1$:

$$W_{opt}^T = W_{fld}^T W_{pca}^T$$

$$W_{pca} = \operatorname{argmax}_W |W^T C W|$$

$$W_{fld} = \operatorname{argmax}_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$