

# Reconeixement de les persones pels trets de la cara

Francesc Serratosa

PID\_00195433



*Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>*

# Índex

<b>Introducció</b> .....	5
<b>Objectius</b> .....	7
<b>1. Etapes dels sistemes de reconeixement de persones pels trets de la cara</b> .....	9
<b>2. Detecció de les cares</b> .....	11
<b>3. Normalització de la cara</b> .....	18
3.1. Normalització de les dimensions de la subfinestra .....	18
3.2. Normalització de la intensitat de la subfinestra .....	18
<b>4. Extracció de les característiques</b> .....	20
<b>5. Verificació d'una cara</b> .....	26
<b>6. Identificació d'una cara</b> .....	27
<b>7. Generació de cares</b> .....	28
<b>Resum</b> .....	32
<b>Activitats</b> .....	33
<b>Abreviatures</b> .....	34
<b>Bibliografia</b> .....	35
<b>Annex</b> .....	36



## Introducció

El reconeixement de les persones per la cara és el mètode d'identificació de les persones més usat pels humans. Un altre mètode natural és l'olor, que és un dels mètodes més usat per molts animals, però aquesta tècnica de reconeixement va deixar de ser usada pels humans fa molts milers d'anys. El reconeixement de les persones per la cara és una tasca que portem a terme constantment i sense un esforç aparent. No obstant això, en la intel·ligència artificial continua sent una tasca de gran esforç computacional i de recursos. L'aparició de dispositius de baix cost i amb molta capacitat de càlcul, com portàtils o telèfons intel·ligents, ha despertat un interès enorme en el processament d'imatges i vídeos en moltes aplicacions, entre les quals hi ha la identificació mitjançant la biometria i el seguiment de persones o de les seves cares. Una aplicació típica és el seguiment automàtic de la cara d'un conferenciant.

La recerca en el reconeixement de cares està motivada no sols pel desafiament de la metodologia en si, sinó també per nombroses aplicacions pràctiques en què la identificació de les persones és necessària. El reconeixement de cares és un dels mètodes biomètrics més usats i ha esdevingut cada cop més i més important a causa dels avenços tan ràpids en les tecnologies de les càmeres digitals, Internet, dispositius mòbils i la demanda cada cop més gran de mesures de seguretat. La capacitat de càlcul del processador d'una càmera de fer fotos de baix cost és superior als ordinadors de sobretaula de principis de la dècada de 1980. Així, han pogut aparèixer els sistemes de detecció de cares en càmeres de baix cost.

Els principals avantatges del reconeixement de les persones per la cara són dos:

- 1) És el mètode més natural, al qual estem acostumats, i per tant és menys intrusiu.
- 2) És el mètode més fàcil d'usar de cara a l'usuari; simplement una càmera ha de captar la cara amb poca o cap col·laboració de l'usuari.

A més, és un sistema biomètric que pot treballar en un procés de verificació (o autenticació) o en un procés d'identificació (o reconeixement). Han aparegut aplicacions que duen a terme automàticament aquests dos processos (per exemple, Picassa o iPhoto). El seu objectiu és anar cercant cares conegudes en fotos personals i així poder-nos mostrar en quines fotos apareix una persona concreta. Hi ha un procés d'aprenentatge en què l'aplicació selecciona les cares i pregunta de quina persona són. Com més cares verifiquem, més aprèn el sistema i per tant, més cares troba i més en classifica en noms concrets. Un dels aspectes pràctics interessants és que l'aprenentatge es pot fer en qualsevol moment, és a dir, que sempre som a temps de corregir cares mal classificades o

classificar cares sobre les quals el sistema dubta i ens demana la nostra ajuda. A més, quan hi anem afegint fotos noves, el sistema va detectant automàticament les cares noves que puguin aparèixer.

L'ordre en què es comenten les tècniques i els processos en aquest mòdul té certa relació amb el flux de la informació en el procés general de la comparació de dues cares o la cerca d'una cara en una base de dades. Comentarem aquests mètodes de la manera següent:

- En l'apartat "Etapas dels sistemes de reconeixement de persones pels trets de la cara" comentem l'existència de les cinc etapes fonamentals d'un sistema de reconeixement de cares.
- Després, detallarem cadascuna de les etapes. D'aquesta manera, en l'apartat "Detecció de les cares" expliquem com es pot detectar una cara dins d'una imatge, és a dir, com es pot saber que en alguna part o algunes parts de la imatge hi ha una cara o diverses cares.
- En l'apartat "Normalització de la cara" expliquem un procés de normalització de la part de la imatge en què creiem que hi ha una cara. Aquesta normalització té l'objectiu de tenir la informació de la cara independentment de la posició de la persona i de la llum.
- Un cop tenim la imatge normalitzada, en l'apartat "Extracció de les característiques" n'extraïem les característiques principals. En l'apartat "Verificació d'una cara" ja tenim el procés específic dels sistemes biomètrics que es basa a verificar si dues imatges de cares pertanyen a la mateixa persona.
- Com a continuació natural, en l'apartat "Identificació d'una cara" s'explica com s'ha d'identificar una persona per la seva cara.
- Finalment, en l'apartat "Generació de cares" s'explica com s'ha de generar una cara de manera sintètica a partir d'un conjunt de cares. Aquesta tècnica és interessant quan no tenim la imatge completa de la cara d'una persona i necessitem tenir-ne una bona aproximació.

## Objectius

Els objectius bàsics d'aquest mòdul són els següents:

- 1.** Conèixer les etapes bàsiques dels sistemes de reconeixement de les persones basats en la cara.
- 2.** Conèixer les tècniques de la visió per computador aplicades a l'extracció de les característiques principals de la cara.
- 3.** Representar una cara en un registre d'una base de dades.
- 4.** Comparar dues cares.
- 5.** Generar una cara de manera sintètica usant un conjunt de cares.





## 1. Etapes dels sistemes de reconeixement de persones pels trets de la cara

El primer sistema de reconeixement de persones per la cara fet públic va ser la tesi doctoral presentada per **Kanade** l'any 1974. Des d'aquell primer sistema fins ara, el rendiment dels sistemes de reconeixement de cares ha millorat significativament. A més, ara aquest procés es pot dur a terme a temps real amb imatges o vídeos capturats en situacions més o menys favorables. No obstant això, encara que la millora en aquests sistemes ha estat encoratjadora, aquesta tasca ha resultat ser un esforç difícil, especialment en situacions no controlades en què hi pot haver canvis d'il·luminació, expressions, punts de vista o oclusions. Els humans i alguns animals tenen una capacitat de reconeixement molt superior als sistemes artificials. Ara per ara, no se sap quan un sistema artificial serà millor que un de natural.

Avui en dia, les etapes principals dels sistemes de reconeixement de persones per la cara no ha variat gaire respecte de la primera solució proposada per Kanade. El que ha variat és la solució aplicada dins d'aquestes etapes. Tots els sistemes acostumen a estar formats per cinc grans mòduls que descriurem a continuació.

1) **Detecció i seguiment de la cara**<sup>1</sup>: extreu la zona en què hi pot haver la cara dins de la imatge. En cas que treballem directament amb un seqüència de vídeo, també haurem de fer el seguiment de la cara, és a dir, anar sabent en quina zona de la imatge es troba la cara, imatge per imatge. No comentarem res més sobre els sistemes de seguiment de cares<sup>2</sup>; simplement esmentar que l'avantatge de treballar amb vídeo és que tenim una informació passada (posició, velocitat, acceleració, orientació, mida...) que usem per a predir aquestes dades en imatges futures.

<sup>(1)</sup>En anglès, *face detection*.

<sup>(2)</sup>En anglès, *face tracking*.

2) **Normalització de la cara**<sup>3</sup>: l'objectiu és normalitzar la imatge de la cara respecte de les propietats geomètriques, tal com la mida i la posició, usant transformacions geomètriques o morfològiques. També es normalitza respecte de la il·luminació. L'objectiu d'aquestes transformacions és que la imatge resultant sigui tan invariant com sigui possible a tots els agents externs (focus, il·luminació, posició relativa d'on és la càmera respecte d'on és la cara, zoom de la càmera...). Es detecten els components facials com els ulls, les orelles, el nas o la boca i s'usen per a fer les transformacions necessàries.

<sup>(3)</sup>En anglès, *face alignment*.

**3) Extracció de les característiques<sup>4</sup>:** l'objectiu és obtenir informació efectiva de la imatge de la cara que sigui útil per a distingir entre dues cares diferents però en canvi que sigui invariant a dues imatges de la mateixa persona. A més, s'ha d'aconseguir que sigui estable respecte als agents externs.

<sup>(4)</sup>En anglès, *face extraction*.

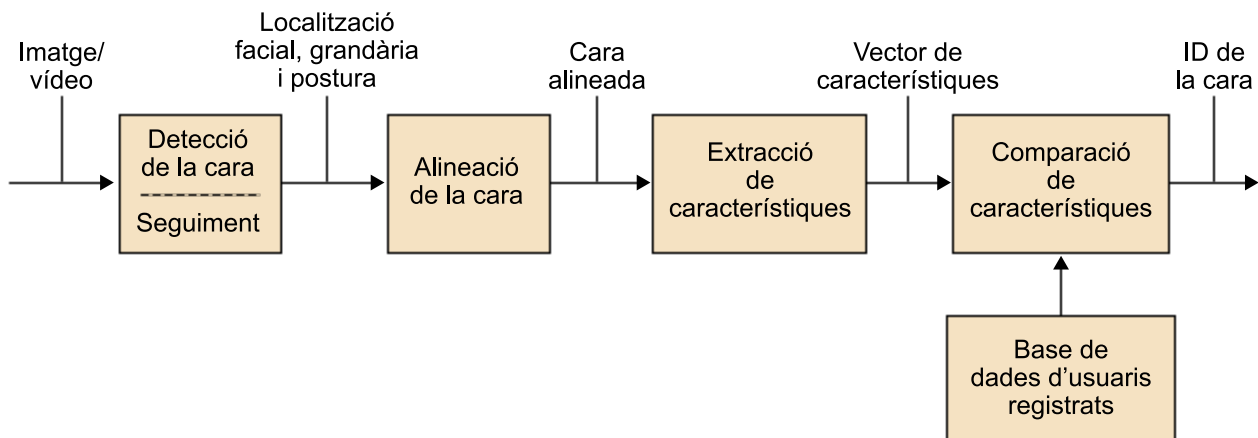
**4) Comparació de cares<sup>5</sup>:** l'objectiu d'aquest mòdul depèn de si estem en un procés d'identificació o de verificació. En un procés d'identificació, compara les característiques extretes d'una imatge nova amb les característiques de diverses imatges de la base de dades i retorna la més semblant (o conjunt de més semblants). En un procés de verificació, compara les característiques de dues imatges i retorna si són la mateixa persona o no. També pot retornar una distància o probabilitat que siguin la mateixa persona.

<sup>(5)</sup>En anglès, *feature matching*.

**5) Base de dades<sup>6</sup>:** aquest mòdul emmagatzema les característiques de totes les imatges que ens han servit per a matricular les persones. En els processos d'identificació, va oferint característiques de cares perquè siguin comparades. Acostumen a tenir uns mecanismes interns anomenats *indexació de bases de dades*, que serveixen perquè no sigui necessari recórrer tota la base de dades i així reduir el temps de cerca.

<sup>(6)</sup>En anglès, *database of unrolled users*.

Figura 1



## 2. Detecció de les cares

La detecció de la cara és el primer pas del procés d'identificació de les cares. La seva fiabilitat té una influència cabdal en tot el sistema de reconeixement, ja que si la cara no es detecta, llavors no s'aplica cap altre procés en aquella part de la imatge. Donada una sola imatge o una seqüència de vídeo, un detector de cares ideal hauria de tenir la capacitat de localitzar totes les cares, malgrat la posició, escalat, edat, expressió... A més, hauria de ser independent d'il·luminacions estranyes i del contingut de la imatge o vídeo. Fixeu-vos que detectar una cara vol dir que el sistema sàpiga que en aquells píxels de la imatge hi ha una cara, però encara no sap quina persona és.

S'han dissenyat diversos mètodes per localitzar cares en una imatge i també per al seguiment i detecció de cares en seqüències de vídeo.

La localització de la cara pel **mètode basat en l'aparició**<sup>7</sup> es basa en la idea de cercar si la cara apareix en qualsevol punt de la imatge. El procés es porta a terme de la manera següent: la imatge d'entrada s'escaneja a totes les posicions possibles i en escalats diferents per una subfinestra. Aquesta subfinestra normalment és un quadrat de diferents dimensions. La detecció de la cara es planteja com un sistema que analitza només la subfinestra i sap classificar la zona de la imatge que cau dins de la subfinestra en només dues classes. Una classe és "cara" i l'altra classe és "no-cara". Els píxels externs a la subfinestra no afecten aquesta decisió. La subfinestra serveix per a aïllar-nos de la resta de la imatge. El classificador que decideix si és una cara o no és una cara ha estat ensenyat amb molts exemples de subfinestres que són cares i molts exemples de subfinestra que no són cares. Com veurem a continuació, aquests exemples serveixen per a aportar informació estadística. Si les mides de la subfinestra són, per exemple, de  $20 \times 20$  píxels, llavors detectarem cares a la imatge que ocupin  $20 \times 20$  píxels. I per entrenar el classificador, hauréem de passar imatges de  $20 \times 20$  que siguin cares i imatges de  $20 \times 20$  que no siguin cares.

En la figura 2 es veuen deu imatges de  $20 \times 20$  píxels usades per a aprendre la classe "cara" i deu imatges de  $20 \times 20$  píxels usades per a aprendre la classe "no-cara".

Figura 2. Imatges usades per a aprendre la classe "cara" i imatges usades per a aprendre la classes "no-cara"



### Picassa i iPhone

En les aplicacions Picassa o iPhone, si la cara no és detectada, no hi ha manera de seleccionar manualment on hi ha una cara.

### Reflexió

Aquí comentarem un dels mètodes més usats per a reconèixer cares en imatges estàtiques per la seva senzillesa i eficàcia. No farem cap comentari respecte dels sistemes de seguiment de cares en vídeos.

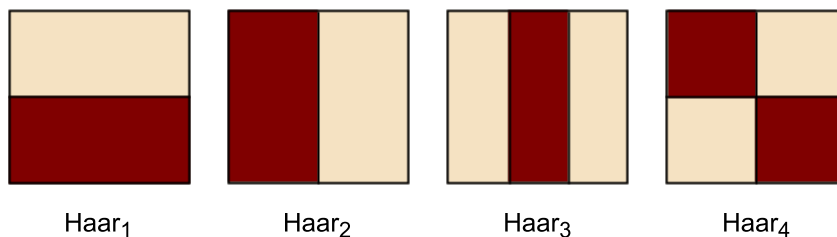
<sup>(7)</sup>En anglès, *appearance-based detection*.

Aquí presentarem el classificador més senzill que s'ha dissenyat, s'anomena **classificador de Bayes**. Hi ha un altre classificador també molt usat anomenat **Adaboost** que normalment funciona una mica millor. Aquest altre classificador complica una mica la idea inicial de Bayes en afegir-hi ponderacions o pesos. Les mostres que anem deduint de manera iterativa que estan mal classificades es tenen molt en compte, ja que el pes amb què se les pondera va augmentant. Aquests dos classificadors de "cara" i "no-cara" aprenen perquè en les subfinestres que contenen una cara apareix una correlació entre els seus píxels. Per exemple, sempre que una zona està molt il·luminada, també ho està aquesta altra zona, o al revés, sempre que està molt il·luminada, detectem que aquesta altra zona és molt fosca. D'altra banda, se suposa que aquesta correlació no es detecta en les subfinestres on hi ha qualsevol altre objecte o part d'objecte que no sigui una cara.

Aquests classificadors no reben directament la subfinestra de la imatge, sinó un conjunt de característiques (normalment és un vector de nombres reals) extretes d'aquesta subfinestra de la imatge. Hi ha diversos mètodes per a generar aquestes característiques. Els més usats són les **característiques de Haar** proposades per **Viola i Jones** l'any 2001. Un altre mètode és el d'extreure les característiques amb els **filtres de Gabor**. A continuació, descriurem les característiques de Haar.

Les característiques de Haar (figura 3) estan formades per quatre imatges de dimensions diverses amb formes rectangulars bicolors.

Figura 3. Les quatre característiques bàsiques de Haar



$$F_i(I) = a^i$$

Quan s'ubica una característica de Haar sobre una subfinestra, els píxels de la subfinestra que cauen sobre la part blanca se sumen i els píxels de la subfinestra que cauen sobre la part negra es resten. El valor final d'aplicar un filtre concret en una posició concreta és el valor resultant de sumar les parts negres i restar-hi les parts blanques.

Cada tipus es pot concretar en una figura segons els paràmetres següents:

- Ubicació de la característica de Haar<sub>i</sub> en la imatge  $(x, y)$ . Normalment, el centre de coordenades dels píxels de la imatge original, i també la subfinestra on s'aplica Haar, està localitzat a l'extrem inferior esquerre i les coordenades comencen pel punt  $(0, 0)$ .

#### Reflexió

El mètode que utilitza els filtres de Gabor no s'explica en aquest material didàctic.

- Diferents dimensions horitzontals i verticals ( $h, w$ ) de la característica de Haar <sub>$i$</sub> .

És a dir, que donada una subfinestra aplicarem  $Haar_i(x, y, h, w)$  i retornarà un valor enter.

### Exemple

Donada aquesta petita imatge original de  $6 \times 6$ :

23	56	67	78	56	32
23	55	12	90	234	105
245	211	178	190	176	87
23	54	78	189	232	190
23	123	239	123	56	32
203	23	29	12	116	132

$Haar_2(1, 1, 4, 4)$  retorna el valor  $-(55 + 12 + 211 + 178 + 54 + 78 + 123 + 239) + (90 + 234 + 190 + 176 + 189 + 232 + 123 + 56) = -950 + 1.290 = 340$

Per tant, el vector d'entrades del classificador de Bayes està format pels quatre tipus de característiques en posicions diferents i dimensions diferents. Hi ha moltes combinacions possibles, per això, s'ha de decidir un subconjunt.

### Exemple

En subfinestres de  $24 \times 24$ , es pot obtenir un vector de vuitanta valors. És a dir, vint combinacions diferents per a cada tipus de Haar. Si fem filtres de Haar de  $6 \times 6$ , llavors en podem posar setze. I si també considerem filtres de Haar de  $12 \times 12$ , llavors en podem posar quatre. Lògicament, es poden posar moltes altres combinacions, per exemple, i acceptar solapament entre filtres o que no hagin de ser quadrats.

Un cop tenim convertida la subfinestra en un vector de  $n$  característiques  $x$ , explicarem com muntem un classificador de Bayes. Primer de tot, hem de deduir quina és la probabilitat que el valor de la  $i$ -èsima característica concreta sigui de la classe "cara"  $P_i^c(x_i)$  i també que sigui de la classe "no-cara"  $P_i^{nc}(x_i)$ . Un cop tinguem aquestes probabilitats, assumirem que són independents entre elles i llavors considerem que una mostra  $x$  és de classe "cara" o de classe "no-cara" segons el producte de probabilitats de totes les característiques de la manera següent:

La mostra  $x$  és "cara" si  $\prod_{i=1}^n P_i^c(x_i) > \prod_{i=1}^n P_i^{nc}(x_i)$ . Altrament, és "no-cara".

Ara ens queda comentar com calculem les probabilitats  $P_i^c(x_i)$  i  $P_i^{nc}(x_i)$ . En el mètode de Bayes, aquestes probabilitats es diuen **probabilitats a posteriori** i es calculen a partir d'altres probabilitats marginals i condicionals. La probabilitat *a posteriori* de la classe "cara" es defineix:

$$P_i^c(x_i) = \frac{P(x_i | \text{"cara"}) \cdot P(\text{"cara"})}{P(x_i)} \quad (1)$$

De manera similar, la probabilitat *a posteriori* de la classe "no-cara" es defineix:

$$P_i^{nc}(x_i) = \frac{P(x_i | \text{"no cara"}) \cdot P(\text{"no cara"})}{P(x_i)} \quad (2)$$

Les probabilitats condicionals  $P(x_i | \text{"cara"})$  i  $P(x_i | \text{"no cara"})$  s'anomenen **probabilitats a priori** i són la funció de possibilitat (*likelihood*) de la classe "cara" i de la classe "no-cara", respecte del valor  $x_i$  de la  $i$ -èsima característica. Les probabilitats de les classes  $P(\text{"cara"})$  i  $P(\text{"no cara"})$  són la probabilitat *a priori* de la classe "cara" i de la classe "no-cara". La probabilitat que aquesta mostra  $x_i$  aparegui,  $P(x_i)$ , es calcula amb el teorema de la probabilitat conjunta:

$$P(x_i) = P(x_i | \text{"cara"}) \cdot P(\text{"cara"}) + P(x_i | \text{"no cara"}) \cdot P(\text{"no cara"}) \quad (3)$$

Només ens queda definir les probabilitats de les classes i les probabilitats *a priori*. Normalment, es considera que hi ha la mateixa probabilitat que a la subfinestra aparegui una cara o que no hi aparegui. Per aquest motiu, es pot assumir que:

$$P(\text{"cara"}) = P(\text{"no cara"}) = 1/2 \quad (4)$$

Amb la definició d'aquestes probabilitats, podem redefinir la fórmula general de la probabilitat *a posteriori* per a les dues classes de la manera següent:

$$P_i^c(x_i) = \frac{P(x_i | \text{"cara"})}{P(x_i | \text{"cara"}) + P(x_i | \text{"no cara"})} \quad (5)$$

i

$$P_i^{nc}(x_i) = \frac{P(x_i | \text{"no cara"})}{P(x_i | \text{"cara"}) + P(x_i | \text{"no cara"})} \quad (6)$$

Aquesta funció ens indica que per a considerar que un element sigui de tipus "cara" (o "no-cara") és important que la seva probabilitat *a priori* sigui alta, però també és important que la probabilitat *a priori* de l'altra classe sigui baixa.

Ara definirem les probabilitats *a priori*. Primer comencem per la probabilitat de ser classe tipus "cara". Donat el conjunt de  $k$  subfinestres que sabem que hi ha una cara, calcularem els vectors de les seves  $n$  característiques i obtenim el conjunt de vectors  $\{a^1, a^2, \dots, a^k\}$ , en què:

$$a^j = [a_1^j, \dots, a_n^j] \quad (7)$$

L'element  $a_i^j$  és el valor de la característica  $i$ -èsima de la mostra  $j$ -èsima que sabem que hi ha una cara. Aquesta característica es calcula amb una funció Haar aplicada a la mostra  $j$ -èsima amb uns paràmetres  $(x, y, h, w)$  concrets. Amb aquests vectors de característiques calcularem el valor mitjà de cada característica:

$$\mu_i^c = \frac{1}{k} \sum_{j=1}^k a_i^j \quad (8)$$

i també la seva desviació estàndard:

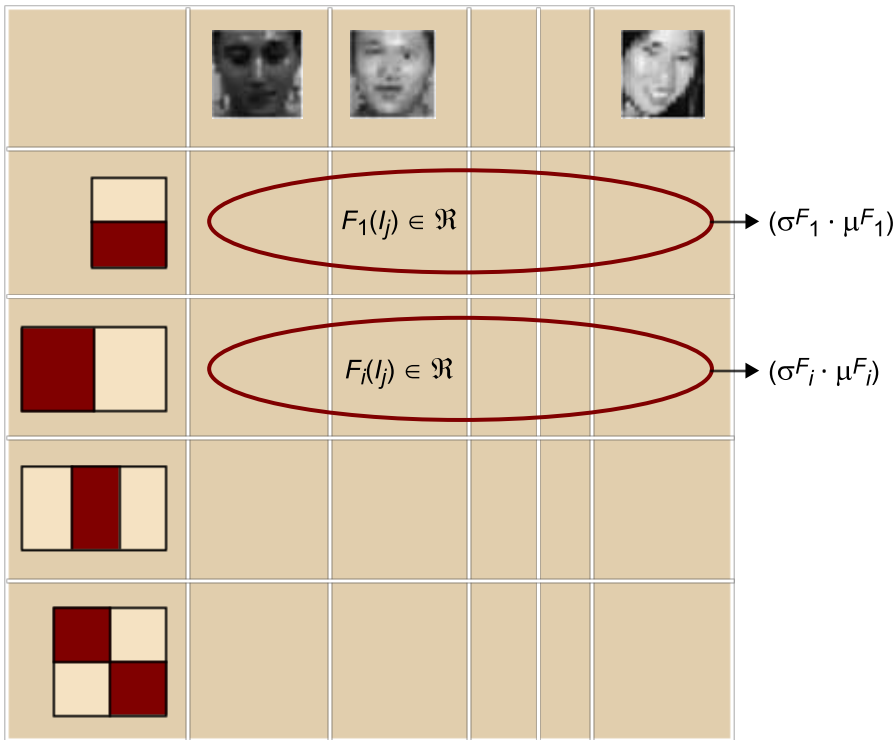
$$\sigma_i^c = \sqrt{\frac{1}{k-1} \sum_{j=1}^k (a_i^j - \mu_i^c)^2} \quad (9)$$

Llavors, assumim que aquestes característiques es poden aproximar per una funció normal i obtenim la probabilitat *a priori* de ser cara donada la mostra següent:

$$P(x_i | \text{"cara"}) = \frac{1}{\sigma_i^c \sqrt{2\pi}} e^{-\frac{(a_i^j - \mu_i^c)^2}{2(\sigma_i^c)^2}} \quad (10)$$

La figura 4 mostra de manera gràfica l'obtenció de la desviació estàndard i la mitjana de cada característica donades les imatges de la classe "cara". En la matriu hi ha tantes columnes com imatges de cares i tantes files com filtres Haar s'hagin aplicat depenent dels cinc paràmetres de la funció Haar $_i(x, y, h, w)$ . En la figura es mostren quatre possibles valors del paràmetre  $i$ , i no es concreten els altres quatre paràmetres.

Figura 4. Extracció de la desviació estàndard i la mitjana donades les imatges de cares i cada combinació del filtre de Haar



El procés per a calcular la probabilitat de classe “no-cara” és similar al descrit amb la classe “cara”. De la mateixa manera, calcularem els vectors  $\{b^1, b^2, \dots, b^k\}$  que representen el mateix i es calculen de la mateixa manera que els vectors  $\{a^1, a^2, \dots, a^k\}$  però les mostres inicials són de subfinestres de la classe “no-cara”. Amb aquests vectors de característiques calcularem el valor mitjà de cada característica:

$$\mu_i^{nc} = \frac{1}{k} \sum_{j=1}^k b_i^j \quad (11)$$

i també la seva desviació estàndard:

$$\sigma_i^{nc} = \sqrt{\frac{1}{k-1} \sum_{j=1}^k (b_i^j - \mu_i^{nc})^2} \quad (12)$$

Llavors, assumim que aquestes característiques es poden aproximar per una funció normal i obtenim la probabilitat de no ser cara donada la mostra:

$$P(x_i | \text{"no cara"}) = \frac{1}{\sigma_i^{nc} \sqrt{2\pi}} e^{-\frac{(b_i^j - \mu_i^{nc})^2}{2(\sigma_i^{nc})^2}} \quad (13)$$



La figura 5 mostra una fotografia en què s'han detectat automàticament les cares. Fixeu-vos que apareixen dos tipus d'errors de detecció que són semblants als errors comentats en el mòdul "Avaluació dels sistemes biomètrics en aplicacions reals". Algunes cares no han estat detectades: **error de no-detecció**<sup>8</sup> i algunes parts que no són cares han estat considerades cares: **error de detecció**<sup>9</sup>. Tanmateix, el nombre d'acceptacions correctes<sup>10</sup> és molt superior als errors.

<sup>(8)</sup>En anglès, *false rejection*.

<sup>(9)</sup>En anglès, *false acceptance*.

<sup>(10)</sup>En anglès, *correct acceptance*.

Figura 5. Fotografia en què s'han detectat les cares de manera automàtica



### 3. Normalització de la cara

Com hem remarcat, els mètodes basats en l'aparició operen en subfinestres de mida fixa. Per aquest motiu, és necessari ajustar la mida de totes les finestres a una mida estàndard, per exemple  $20 \times 20$ . Això es deu al fet que els mètodes que comentarem més endavant per al reconeixement de cares necessiten tenir un nombre de píxels concret. A més a més, el mètode de Haar pretén ser independent als canvis d'il·luminació. Per això, és necessari un procés posterior de normalització de les subfinestres en què s'ha detectat la cara segons la seva intensitat. Si aquest procés no es fes, el reconeixement de cares es podria dur a terme però no funcionaria tan bé. Aquestes operacions són globals des del punt de vista que tots els píxels de la subfinestra es modifiquen.

#### 3.1. Normalització de les dimensions de la subfinestra

Normalment, les subfinestres són quadrades ja que les cares es poden enquadrar bastant bé en un quadrat. Suposem que la subfinestra que ha detectat la cara té unes dimensions de  $(l_1, l_1)$  i nosaltres volem que totes les subfinestres en què hi ha possibles cares tingui unes dimensions de  $(l_2, l_2)$ . Cada píxel de la subfinestra nova adoptarà el valor:

$$I_2(x, y) = I_1\left(\left\lfloor x \cdot \frac{l_1}{l_2} \right\rfloor, \left\lfloor y \cdot \frac{l_1}{l_2} \right\rfloor\right) \quad (14)$$

Les dimensions normalitzades acostumen a ser les de la finestra més petita. Això és així perquè si no ens trobaríem que hauríem d'omplir alguns píxels de la nova subfinestra de manera artificial, ja que alguns píxels de la finestra nova no serien omplerts per cap píxel de la finestra original. Un altre mètode més intel·ligent per a omplir les noves posicions es basa a trobar la mitjana ponderada dels píxels del voltant. Aquest mètode és el més usat quan es modifiquen les dimensions d'imatges encara que té l'inconvenient que té un efecte de filtratge.

#### 3.2. Normalització de la intensitat de la subfinestra

Hi ha diverses tècniques per a normalitzar la imatge des d'un punt de vista de la intensitat. Aquí comentarem el mètode basat en l'**equalització de l'histograma**. Aquest mètode es basa a augmentar el contrast de la imatge ajustant els valors de l'histograma. Amb aquest ajustament, les intensitats dels píxels es distribueixen més equitativament en l'histograma. Primer haurem de calcular la intensitat de píxel mínima,  $v_{\min}$ , i màxima,  $v_{\max}$ . És a dir, els valors dels píxels de partir dels quals cap píxel de la imatge no té un valor inferior

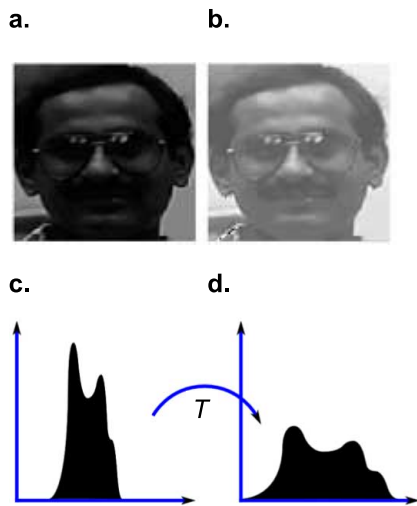
o superior (de vegades, en lloc de “cap” es considera un nombre petit). Llavors, aplicarem una transformació global a tots els píxels de la imatge segons l’equació:

$$I_2(x, y) = \frac{I_1(x, y) - V_{min}}{V_{max} - V_{min}} \cdot MAX \quad (15)$$

en què  $MAX$  representa el valor màxim possible del píxel (habitualment és de 255).

Aquesta transformació de la imatge implica que l’histograma resultant omple tot el rang possible de valors. La figura 6 mostra una imatge original i la mateixa imatge l’histograma de la qual s’ha equalitzat. També es mostra l’histograma original i l’histograma equalitzat.

Figura 6. Imatge equalitzada



a. Imatge original. b. Imatge tractada. c. Histograma abans del procés. d. Histograma equalitzat

## 4. Extracció de les característiques

Una **imatge** es pot veure com un punt en un espai discret que té tantes dimensions com píxels té la imatge i el nombre de valors de cada dimensió és la profunditat del píxel (normalment  $256 \times 3$  canals).

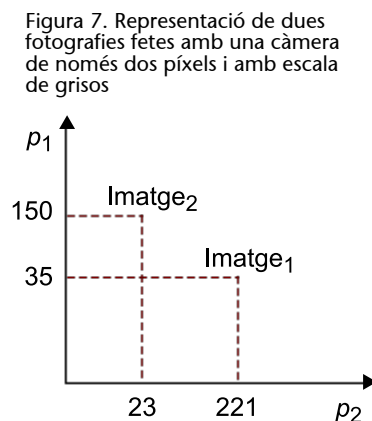
Fixeu-vos que aquest espai és enorme. Suposem una càmera web de baixa resolució ( $640 \times 480$ ). Això vol dir que el nostre espai té 307.200 dimensions. El mètode que aquí comentarem considera aquesta observació i intentarem reduir el nombre de dimensions d'aquest espai. Normalment, es porta a terme una reducció dràstica ja que s'acostuma a passar de 307.200 dimensions a unes 7 dimensions. Aquesta reducció es pot dur a terme per dos motius principals:

- 1) Només una petita zona d'aquest espai enorme és la que conté imatges amb cares.
- 2) Es duu a terme una transformació de l'espai i aquest espai nou on anem a parar és l'espai de la covariància entre els píxels de les imatges.

Portarem a terme la reducció amb el mètode d'**anàlisi dels components principals (PCA)** que s'explica més endavant.

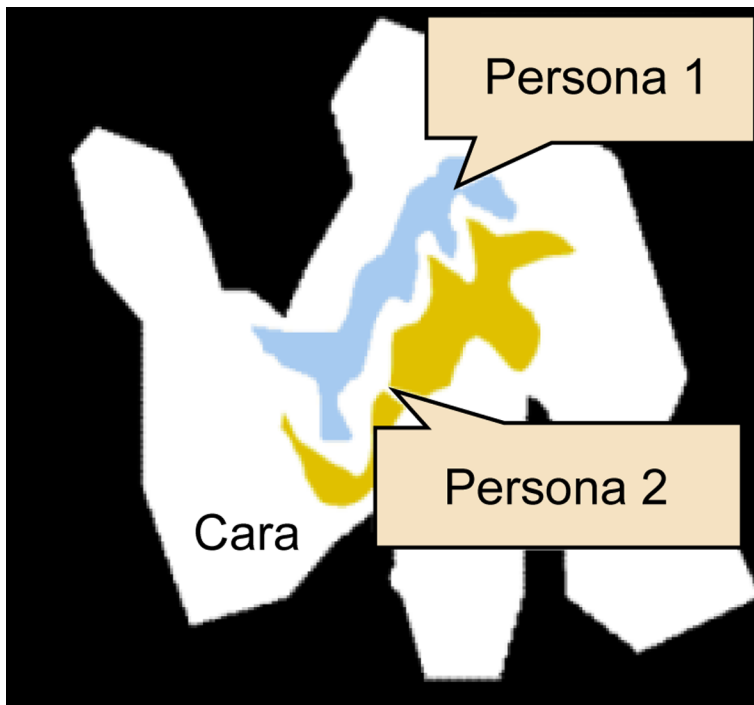
### Exemple de reducció

Imaginem-nos una càmera de resolució molt, molt baixa. Tan baixa que les imatges que fa només tenen dos píxels. En aquest cas, les imatges es poden definir en un espai de dues dimensions. La dimensió  $x_1$  representa el píxel  $p_1$  i la dimensió  $x_2$  representa el píxel  $p_2$ . Com que aquesta càmera de dos píxels fa només fotografies en escala de grisos, ens trobem que la profunditat dels píxels és de 256 i, llavors, les dimensions d'aquest nou espai discret és de 256 posicions. Suposem que fem dues fotos amb aquesta càmera:  $\text{imatge}_1 = [35 | 221]$ . Això vol dir que el valor del píxel  $p_1 = 35$  i el valor del píxel  $p_2 = 221$ . I la segona foto és  $\text{imatge}_2 = [150 | 23]$ . Si representem aquestes fotografies en el nou espai tindrem (figura 7):



En aquest nou espai podem analitzar quines fotografies han estat fetes a cares, quines no i, fins i tot, podem agrupar les fotos de cares depenent de la persona. D'aquesta manera, ens podria quedar un espai com el que mostra la figura 8.

Figura 8. Espai bidimensional en què es representa la zona on hi ha fotografies de cares i, dins d'aquesta zona, on hi ha les cares de dues persones



Com que en aquest espai és difícil separar de manera lineal les zones on hi ha cares de les persones, transformem aquest espai en un espai en què es considera la variabilitat entre els píxels, és a dir, la variabilitat entre les coordenades d'aquest espai nou.

**L'anàlisi dels components principals (PCA)** és una tècnica de reducció de les dimensions de l'espai basada a obtenir només un nombre imposat com a paràmetre d'entrada de components principals de dades multidimensionals.

#### Vegeu també

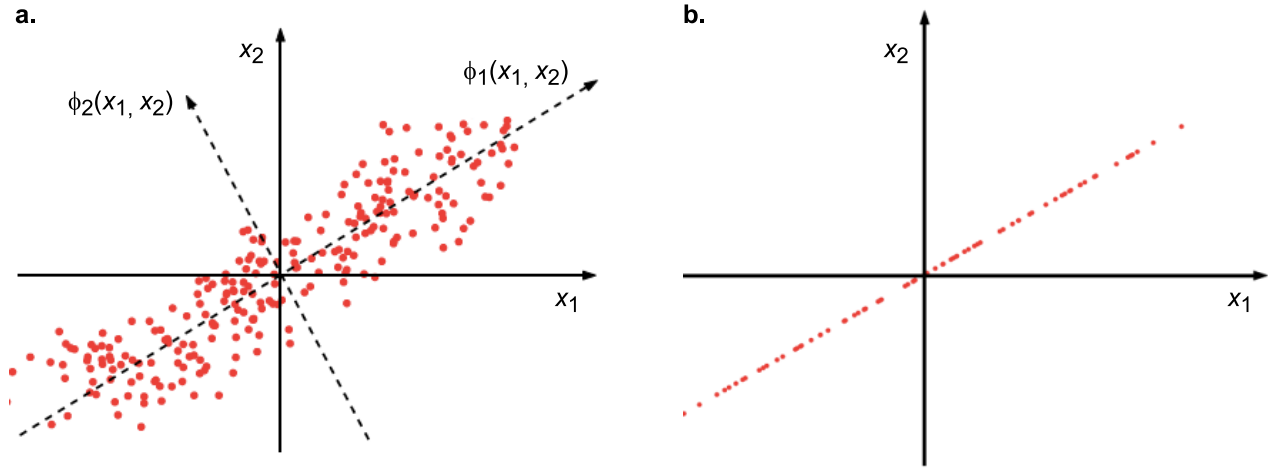
En l'annex d'aquest mòdul hi ha dos enllaços a vídeos que expliquen en detall aquesta tècnica.

El primer component principal és la combinació lineal de les dimensions originals que té la màxima variància. L' $n$ -èsim component principal és la combinació lineal amb la màxima variància i amb la restricció que ha de ser ortogonal als  $n - 1$  primers components principals.

La idea de la PCA s'il·lustra en la figura 9. Els eixos  $x_1$  i  $x_2$  representen la covariància entre dos píxels.  $\varnothing^1(x_1, x_2)$  correspon a la direcció de la màxima variància i, per tant, s'escull com a primer component principal. En el cas de dues dimensions, el segon component principal,  $\varnothing^2(x_1, x_2)$ , es determina unívocament per les restriccions d'ortogonalitat. En espais de més dimensions, la selecció de components continuaria, guiada per les variàncies de les projeccions.

ons. Si es volgués reduir l'espai a una sola dimensió, llavors posaríem tots els punts a la línia formada per l'eix  $\phi^1(x_1, x_2)$  i imposaríem  $\phi^2(x_1, x_2) = 0$  en aquest nou espai  $(\phi^1, \phi^2)$  com mostra la figura 9.

Figura 9

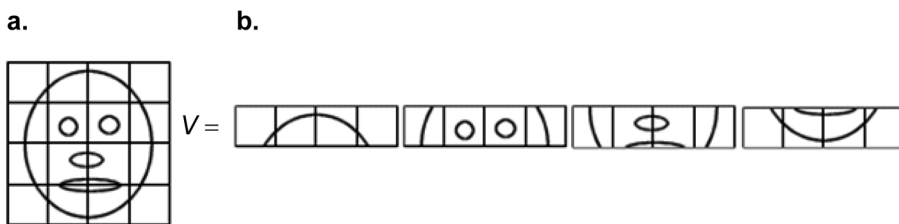


a. Representació dels dos eixos principals. b. Reducció a una sola dimensió

Per a trobar aquests nous eixos, usarem la tècnica de la **transformació de Karhunen-Loève**. No obstant això, per tal d'usar aquesta tècnica, hem de suposar que la mitjana de les dades és zero. Aquesta imposició no és cap inconvenient, ja que l'únic que ens exigeix és que inicialment restem la mitjana de les cares a totes les cares.

Suposem que les imatges de les cares tenen  $m \times n$  píxels. Llavors, descrivim cada imatge com un vector  $v$  de  $N = m \cdot n$  elements, simplement concatenant les columnes de la imatge per formar el vector.  $v_i$  representa el valor de la posició  $i$ -èsima del vector  $v$ . En la figura 10 mostrem una cara i el vector transposat que es genera (normalment, els vectors es representen de manera vertical).

Figura 10. Cara i vector transposat



a. Imatge de  $16 \times 16$  píxels. b. Vector transposat que representa la imatge

També suposem que tenim  $M$  imatges diferents de la mateixa cara o persona. La cara  $j$ -èsima es descriu pel vector  $v^j$ . La posició  $i$ -èsima del vector  $v^j$  es descriu amb el símbol  $v_i^j$ . El primer que hem de fer és tornar a calcular els vectors de les cares de manera que la mitjana dels nous vectors  $x^j$  sigui zero. Cadascuna de les posicions del vector mitjana es defineix:

$$\bar{v}_i = \frac{1}{M} \sum_{j=1}^M v_i^j \quad (16)$$

Aquest vector representa la cara mitjana, és a dir, com si les imatges fossin transparències i les poséssim totes en un munt. Llavors, els vectors normalitzats s'obtenen sostraint-ne la mitjana:

$$x_i^j = v_i^j - \bar{v}_i \quad (17)$$

per a tot  $1 \leq i \leq N$  i per a tot  $1 \leq j \leq M$ . La variable  $x_i^j$  representa el valor normalitzat de la posició  $i$ -èsima de la imatge  $j$ -èsima.

Amb aquest conjunt de vectors normalitzats formarem una matriu  $X$  de  $N$  files i  $M$  columnes:

$$X = \begin{bmatrix} x_1^1 & \cdots & x_1^M \\ \vdots & \ddots & \vdots \\ x_N^1 & \cdots & x_N^M \end{bmatrix} \quad (18)$$

Volem estudiar la variabilitat entre els píxels que componen el conjunt de les imatges de la cara d'una mateixa persona. Per exemple, estudiar si sempre que un píxel augmenta de color, també augmenta el color d'un altre píxel. En cas que la imatge no sigui de la mateixa persona, assumim que aquesta variació concreta no serà tan important i així detectarem que no és la mateixa persona. Com que el que volem és estudiar la variabilitat entre les dades, amb la matriu  $X$  calcularem la matriu de covariància  $\Sigma$  de  $N \times N$  dimensions de la manera següent:

$$\Sigma = \frac{1}{M} \sum_{j=1}^M x^j \cdot x^{jT} \quad (19)$$

En aquesta equació, hem assumit que els vectors que representen les cares són verticals i  $T$  representa la funció transposada, és a dir, que genera un vector horitzontal. Fixeu-vos que el còmput de la covariància és molt senzill perquè les dades han estat normalitzades prèviament i la mitjana ja ha estat sostreta.

Si  $\sum_{i,j}$  representa la posició  $(i, j)$  de la matriu de covariància, llavors, el seu valor es calcula com a:

$$\sum_{i,j} = \frac{1}{M} \sum_{k=1}^M x_i^k x_j^k \quad (20)$$

per a tot  $i, j$  d'1 a  $N$ .

Un cop calculada la matriu de covariància, calcularem els vectors propis  $\varnothing = [\varnothing^1, \dots, \varnothing^N]$  i els valors propis  $\lambda = [\lambda^1, \dots, \lambda^N]$ . Cada vector propi  $\varnothing^i$  té  $N$  elements. Fixeu-vos que els vectors propis són els nous eixos representats en la figura 7. I els valors propis representen la variabilitat de les dades projectades en l'espai de vectors propis. Els vectors propis i valors propis s'ordenen de manera que:

$$\lambda^1 \geq \lambda^2 \geq \dots \geq \lambda^N \quad (21)$$

És a dir, que el vector propi  $\varnothing^i$  correspon a l' $i$ -èsim valor propi  $\lambda^i$  més gran.

Finalment, la cara de la persona  $p$ -èsima, que a la base de dades original era representada per  $M$  imatges de  $N = m \times n$  píxels, aquí es representa amb una matriu i dos vectors. La matriu de  $N \times K$  elements són els  $K$  primers vectors propis  $\varnothing^p = [\varnothing^{1^p}, \dots, \varnothing^{K^p}]$ . I el primer vector està format pels  $K$  valors propis  $\lambda^p = [\lambda^{1^p}, \dots, \lambda^{K^p}]$ . Si  $K$  és exactament  $N$ , llavors no hi haurem aplicat cap tipus de reducció de les dades. Com més petit sigui  $K$ , menys ocupa la representació de cada persona però més probabilitat hi haurà que el sistema de reconeixement s'equivoqui. I el segon vector és la mitjana de les cares usades per a generar el model  $\bar{v}^p = [\bar{v}^{1^p}, \dots, \bar{v}^{N^p}]$ . Aquest vector sempre és de  $N$  elements. La representació de la persona  $p$  quedarà, doncs:

$$1) \varnothing^p = \begin{bmatrix} \varnothing_1^{1^p} & \dots & \varnothing_1^{K^p} \\ \vdots & \ddots & \vdots \\ \varnothing_N^{1^p} & \dots & \varnothing_N^{K^p} \end{bmatrix} \text{ en què } \varnothing_i^{q^p} \text{ representa l'element } i \text{ del vector propi } q \text{ de}$$

la persona  $p$  amb una compressió de  $K$  vectors propis, i és  $K \leq M$ .

2)  $\lambda^p = [\lambda^{1^p}, \dots, \lambda^{K^p}]$  en què  $\lambda^{q^p}$  representa el valor propi  $q$ -èsim més gran de la persona  $p$ . El vector propi  $q$  es relaciona amb el vector propi  $q$ .

3)  $\bar{v}^p = [\bar{v}^{1^p}, \dots, \bar{v}^{N^p}]$  en què  $\bar{v}^{i^p}$  representa la mitjana del píxel  $i$ -èsim de totes les imatges usades per a modelar la persona  $p$ .

Resum dels índexs:

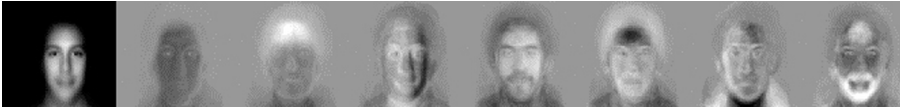
- $i$ : posició de l'element dins dels vectors.  $i \in [1, \dots, N]$
- $j$ : índex que identifica una cara d'una persona concreta.  $j \in [1, \dots, M]$
- $p$ : índex que identifica una persona.  $p \in [1, \dots, Z]$
- $q$ : índex que identifica un vector propi i un valor propi.  $q \in [1, \dots, K]$

La figura 11 mostra la cara mitjana de la persona  $p$ ,  $\bar{v}^p$ , a l'esquerra. I els set primers vectors propis  $\varnothing^{1^p}, \dots, \varnothing^{7^p}$  de la persona  $p$ . Els vectors propis s'acostumen a anomenar *eigenfaces*<sup>11</sup> i es tradueix com a *eigencares*.

<sup>(11)</sup> *Eigenvector* és un vector propi i *face* vol dir 'cara'.



Figura 11. Representació d'una persona a partir de la seva mitjana i els set vectors propis



## 5. Verificació d'una cara

Suposem que tenim una imatge nova d'una cara en format de vector  $v$  i la representació de la persona  $p$  expressada de la manera que hem descrit en l'apartat anterior. Volem saber si aquesta imatge és de la persona  $j$ .

El procés de verificació farà els passos següents:

1) Normalitzar les dades. Obtenim el vector  $x$  a partir del vector imatge  $v$  del qual hem extret el vector mitjana  $\bar{v}^p$  de la persona  $p$ . Per tant:

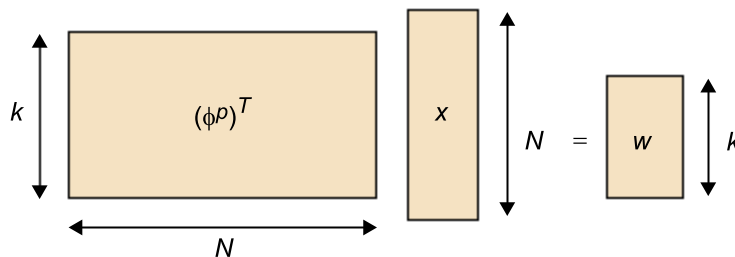
$$x = v - \bar{v}^p \quad (22)$$

2) Apliquem un canvi de base a aquest vector  $x$  en què la nova base són els vectors propis  $\phi^p$  i obtenim la mateixa dada  $w$  però en les noves coordenades. El vector  $w$  és de  $k$  elements i es calcula:

$$w = (\phi^p)^T \cdot x \quad (23)$$

El superíndex  $T$  representa la matriu transposada. Gràficament en la figura 12.

Figura 12



3) Calcularem la distància euclidiana entre els pesos obtinguts  $w$  i els valors propis de la persona  $p$ ,  $\lambda^p$ . Formalment:

$$d^p(v) = \sqrt{\sum_{q=1}^k (\lambda^{q^p} - w^q)^2} \quad (24)$$

4) Finalment, si el valor obtingut és inferior a un llindar  $T$  imposat per l'administrador del sistema, considerarem que la cara  $v$  pertany a la persona  $p$ . Si  $d^p(v) < T$  llavors retorna SÍ. Altrament, retorna NO.

## 6. Identificació d'una cara

Com s'ha explicat en el mòdul “La biometria per a la identificació de les persones”, per identificar una persona, s'han de comparar les seves dades amb les dades de la base de dades. Ja hem dit que hi ha mètodes que permeten evitar algunes comparacions i no explorar tota la base de dades. No és l'objectiu d'aquesta assignatura parlar de temes de bases de dades, per això, suposarem que realment s'exploren tots els registres de la base de dades.

Suposem que tenim una imatge nova d'una cara en format de vector  $v$  que volem identificar i també suposem que tenim una base de dades de cares amb  $Z$  persones i que de cada persona hem extret les dades explicades en l'apartat “Extracció de les característiques”. El procés d'identificació farà els passos següents:

1) Per a totes les persones  $p$  inscrites a la base de dades, calcularem la distància entre la cara representada pel vector  $v$  i la representació de les persones,  $d^p(v)$ . Obtindrem un vector de distàncies:

$$D = [d^1(v), \dots, d^p(v), \dots, d^z(v)] \quad (25)$$

2) Reordenarem el vector  $D$  de menys a més i generem el vector:

$$D' = [d^{1'}(j, v), \dots, d^{e'}(i, v), \dots, d^{z'}(t, v)] \quad (26)$$

Mantindrem la informació de cada distància a quina persona corresponia, ja que:

$$d^{e'}(i, v) = d^i(v) \quad (27)$$

en què la distància  $d^i(v)$  és l' $e$ -èsima més petita.

3) Retornarem la identificació de les  $E$  primeres persones amb distància més petita. En alguns casos, només es vol saber la persona que més s'assembla, llavors  $E = 1$ . Però en algunes aplicacions, és habitual no ser tan restrictiu, i  $E$  pot variar, per exemple  $E = 10$ .

### Vegeu també

La representació de les persones s'ha explicat en l'apartat 5 d'aquest mòdul.

## 7. Generació de cares

Donada una fotografia d'una persona, la podem projectar sobre els models de components principals i després recuperar aquesta nova persona amb la informació de covariància de l'altra persona. La imatge obtinguda s'acostuma a anomenar *imatge recuperada*  $R$ . El procés s'anomena *generació de cares*. Aquest procés es compon de dues etapes:

- 1) Projectem la cara a l'espai de covariàncies.
- 2) Tornem a projectar les dades obtingudes a l'espai de les imatges.

Aquest procediment s'usa en dues situacions diferents. En la primera situació, volem generar una nova cara amb la informació de la persona que vulguem. És similar a imposar els trets característics d'una persona a una altra persona. La segona situació apareix quan tenim imatges amb molt de soroll o amb oclusions parcials.

Imagineu-vos que hem generat el model d'una persona amb diverses fotos "correctes" en què els ulls sempre estan oberts i miren a la càmera. Però la foto que tenim de la persona té els ulls tancats. Aquest procés pot servir perquè puguem veure la foto nova però amb els ulls oberts.

Si  $v$  és el vector que representa la imatge de la persona  $p^1$  i volem generar una imatge nova de la cara  $R$  ( $R$  és un vector de  $N$  posicions) com si fos la persona  $p^2$ , els passos són els següents:

- 1) Normalitzar les dades. Obtenim el vector  $x$  a partir del vector imatge  $v$  de la persona  $p^1$  del qual hem extret el vector mitjana  $\bar{v}^{p^1}$  de la persona  $p^1$ .
- 2) Apliquem un canvi de base a aquest vector  $x$ , en què la nova base són els vectors propis  $\varnothing^{p^1}$ . Fins a aquest punt, l'únic que hem fet és convertir la imatge de la cara  $v$  a l'espai de les covariàncies. Es calcula:

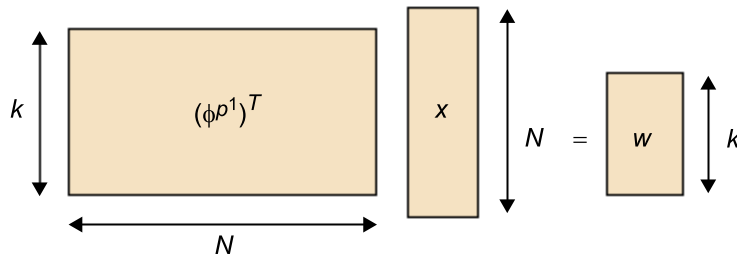
$$w = (\varnothing^{p^1})^T \cdot x \quad (28)$$

El superíndex  $T$  representa la matriu transposada. Gràficament en la figura 13.

### Vegeu també

Aquest pas s'ha explicat en apartats anteriors.

Figura 13

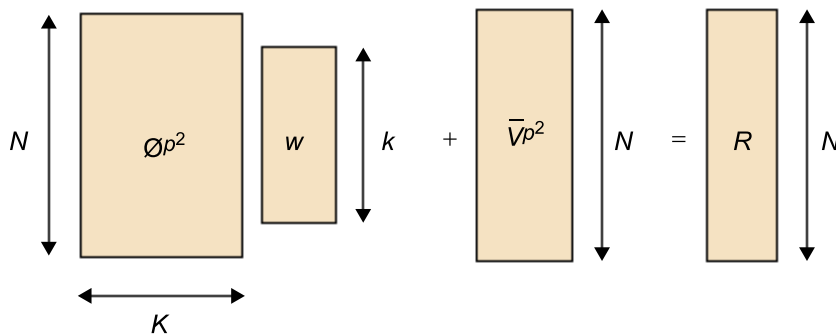


3) Ara, el que farem és obtenir el vector de la imatge recuperada  $R$  tornant a la base original però a partir dels vectors propis de la persona  $p^2$  i sumant la mitjana de la persona  $p^2$ . És a dir:

$$R = \varnothing^{p^2} \cdot w + \bar{v}^{p^2} \quad (29)$$

Gràficament en la figura 14.

Figura 14



**Nota**

Aquest pas es pot dur a terme perquè la matriu inversa de la matriu transposada d'una matriu ortogonal és la mateixa matriu:

$$M = (M)^T \quad (30)$$

Els superíndex  $T$  i  $'$  representen les matrius transposada i inversa, respectivament.

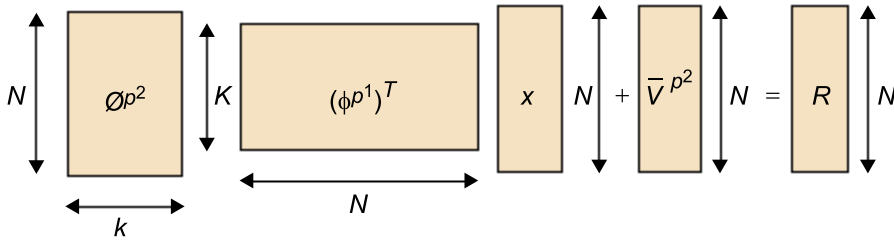
En cas que  $p^1 = p^2$  i quan no s'ha portat a terme reducció de vectors propis,  $k = N$ , llavors el resultat final és exactament la imatge original. En cas que  $k$  sigui inferior a  $N$ , llavors hi ha pèrdua d'informació, ja que la base dels vectors propis és de dimensió més petita. En aquest cas, la imatge original no és la mateixa que la recuperada. Però aquest fet, de vegades, és positiu, perquè serveix per a extreure la informació més "important", la que té més covariància.

Si agrupem tots els passos en una sola equació matemàtica, tenim:

$$R = \varnothing^{p^1} \cdot (\varnothing^{p^2})^T \cdot x + \bar{v}^{p^1} \quad (31)$$

Gràficament en la figura 15.

Figura 15



En cas que volguéssim recuperar una imatge d'una mateixa persona, llavors  $p = p^1 = p^2$ ; per tant:

$$R = \varnothing^p \cdot (\varnothing^p)^T \cdot x + \bar{v}^p \quad (32)$$

Les primeres imatges de la figura 16 mostren la cara d'una persona el punt de vista de la qual ha anat canviant. L'última imatge és la imatge mitjana. Fixeu-vos que en totes les imatges se'n poden veure els ulls.

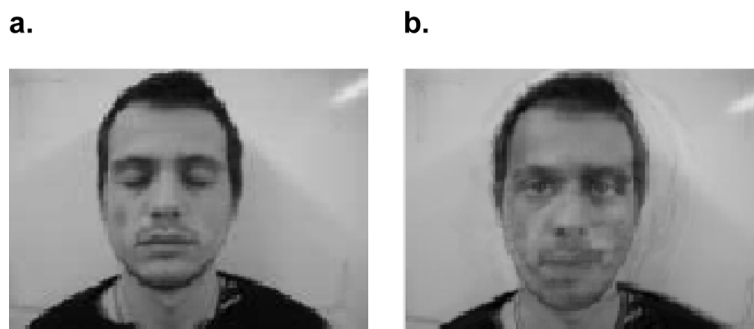
Figura 16



a. Diverses imatges d'una persona. b. Imatge mitjana d'una persona

La imatge esquerra de la figura 17 mostra una imatge nova. Amb aquesta imatge, que té els ulls tancats, volem generar una imatge amb els ulls oberts. Usant la informació de les altres imatges i aplicant el sistema de generació d'imatges obtenim la imatge dreta de la figura 17.

Figura 17



a. Imatge usada per a la generació automàtica. b. Imatge resultat

La figura 18 mostra les *eigenfaces* ordenades segons la magnitud del vector propi. Fixeu-vos que la primera *eigenface* destaca el moviment de rotació del cap (píxels més clars).

Figura 18. *Eigenfaces* ordenades segons el valor propi



## Resum

En aquest mòdul hem explicat algunes tècniques relacionades amb la identificació de les persones per les seves cares. La cara és el tret biomètric més usat pels humans per a identificar les persones, per això és extensament acceptat per la societat. Aquest fet ha influenciat de manera notable la posada en marxa d'aplicacions civils basades en el reconeixement de la cara. A més, el sensor per a capturar cares no és res més que una màquina de fer fotografies i tothom en té una i tothom està generant fotografies amb aquest dispositiu. Fixeu-vos que aquest és el motiu pel qual els sistemes d'identificació de persones com els que usen Picassa o iPhone tenen sentit. Si la gent no anés generant milers de fotografies i no les anés emmagatzemant, no tindríem el desig de cercar persones en les fotografies.

És important remarcar que la identificació de persones per les cares és un mètode no intrusiu i passiu. Per tant, no es demana a l'usuari que toqui res i s'evita així el rebuig de la societat a tocar aparells que d'altres toquen i s'eviten també possibles contagis. A més, el fet de ser passiu és usat per a identificar persones sense la seva col·laboració o, fins i tot, sense que se n'adonin.

No sols hem parlat d'identificar persones, sinó que gràcies a l'anàlisi dels components principals, hem aconseguit generar imatges noves amb cares. Aquesta és una aplicació útil per a reduir soroll en imatges o completar imatges de cares parcialment ocultes.

Finalment, per a explicar el procés d'identificació o verificació de cares ens ha fet falta explicar el classificador de Bayes, les característiques de Haar, i també l'anàlisi de components principals. Aquestes tècniques són àmpliament usades en molts altres sistemes i aplicacions.



## Activitats

1. Descriviu les etapes del sistema de reconeixement de persones per la cara. Dibuixeu-ne l'esquema complet.
2. Poseu en funcionament el sistema de reconeixement de cares de Picassa, i si teniu un Mac, el de l'iPhone.
3. Com es diu el mètode que expliquem per a detectar cares? Per què creieu que es diu així? Creieu que és un mètode intel·ligent o és un mètode que usa "la força bruta"?
4. Com ho faríeu per definir un sistema que només detectés cares rient? O només cares de dones? O només cares d'una persona concreta?
5. Creieu que podríem definir altres característiques de Haar? Quina és la idea bàsica d'aquestes característiques? Com influencien els canvis d'il·luminació a aquestes característiques?
6. Imagineu-vos que volem fer un classificador de cares de dones i homes. Tenim dues opcions, una de les quals és fer dos classificadors. El primer detecta la classe "cara dona" i la classe "altres imatges", i el segon detecta la classe "cara home" i la classe "altres imatges". En cas que els dos classificadors prenguin la decisió que hi ha una cara en el mateix punt s'ha de prendre una decisió de quin classificador es considera el correcte. El segon mètode és crear un classificador directament que generi tres classes: "cara dona", "cara home" i "altres imatges". Podeu definir les probabilitats d'aquest classificador?
7. En la definició de les probabilitats hem suposat que es poden modelar usant una normal. Si fos el cas que no volguéssim aquesta modelització, com podríem definir aquestes probabilitats de manera que no fos necessari suposar cap tipus de model?
8. Detecteu acceptacions falses i rebutjos falsos en la imatge de la figura 3. En algunes acceptacions falses no es pot intuir la forma de cara. Per què creieu que pot haver generat aquesta acceptació falsa? I en el cas dels rebutjos falsos, per què creieu que no han detectat la cara?
9. Descriviu un algorisme en alt nivell que normalitzi tant les mides com la intensitat, donats els paràmetres necessaris.

## Abreviatures

**PCA** *Principal component analysis* ('anàlisi dels components principals')

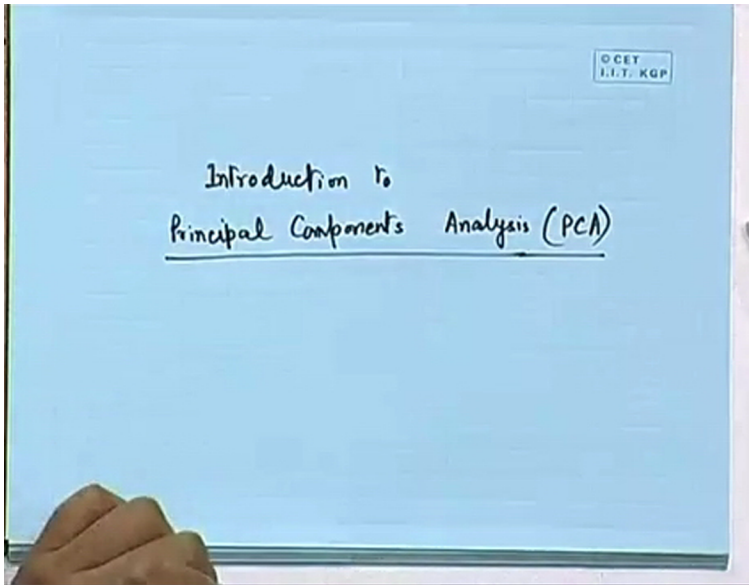
## **Bibliografia**

**Jain, Anil; Flynn, Patrick; Ros, Arun** (editors) (2008). *Handbook of biometrics*. Editorial Springer.

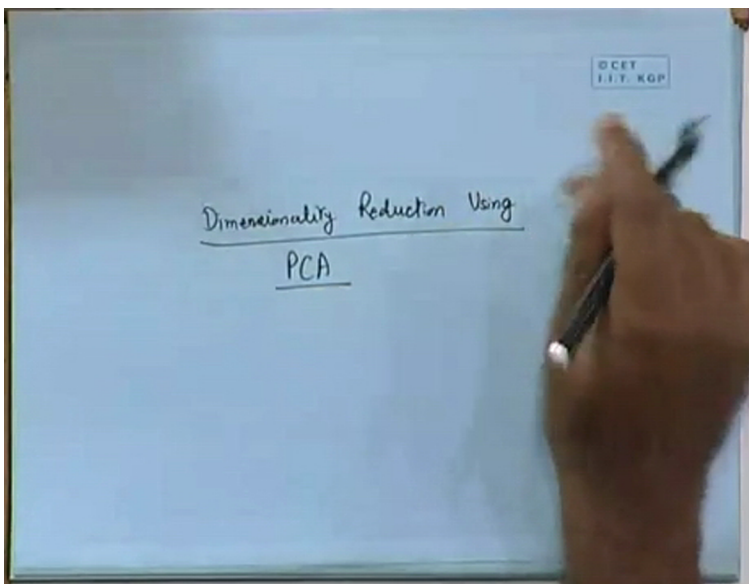
**Li i altres** (2005). *Handbook of face recognition*. Editorial Springer.

## Annex

La tècnica de l'anàlisi dels components principals (PCA) o transformació de Karhunen-Loève s'aplica en camps molts diversos. Té una base matemàtica molt forta que no hem analitzat en detall. En els vídeos següents s'explica la base matemàtica d'aquesta tècnica.



"Lec-32 Introduction to Principal Components and Analysis". *Neural Network and Applications*



"Lec-33 Dimensionality reduction Using PCA". *Network and Applications*