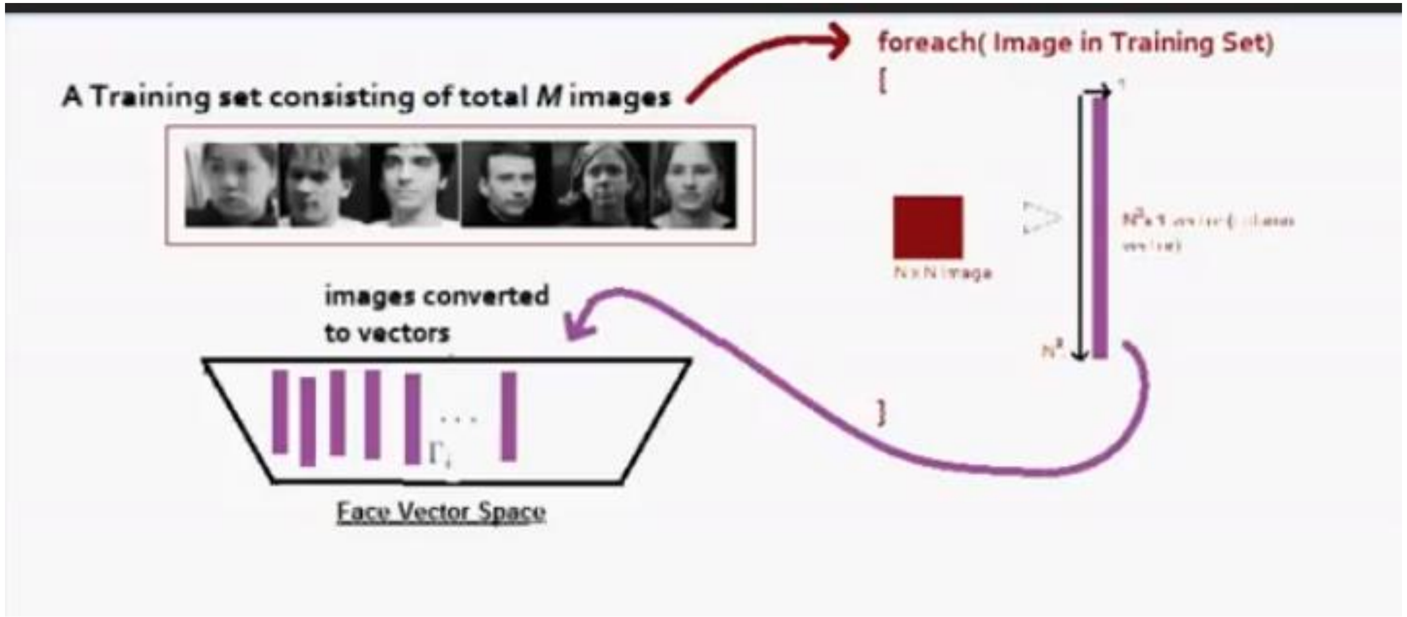
CMP 0575 - PCA

Esteban Flores

Intro

- **Conjunto de datos:** 197 imágenes (256x256 RGB).
- **Lenguaje utilizado:** Python
- **Librería utilizada:** Sci-kit

Paso 1

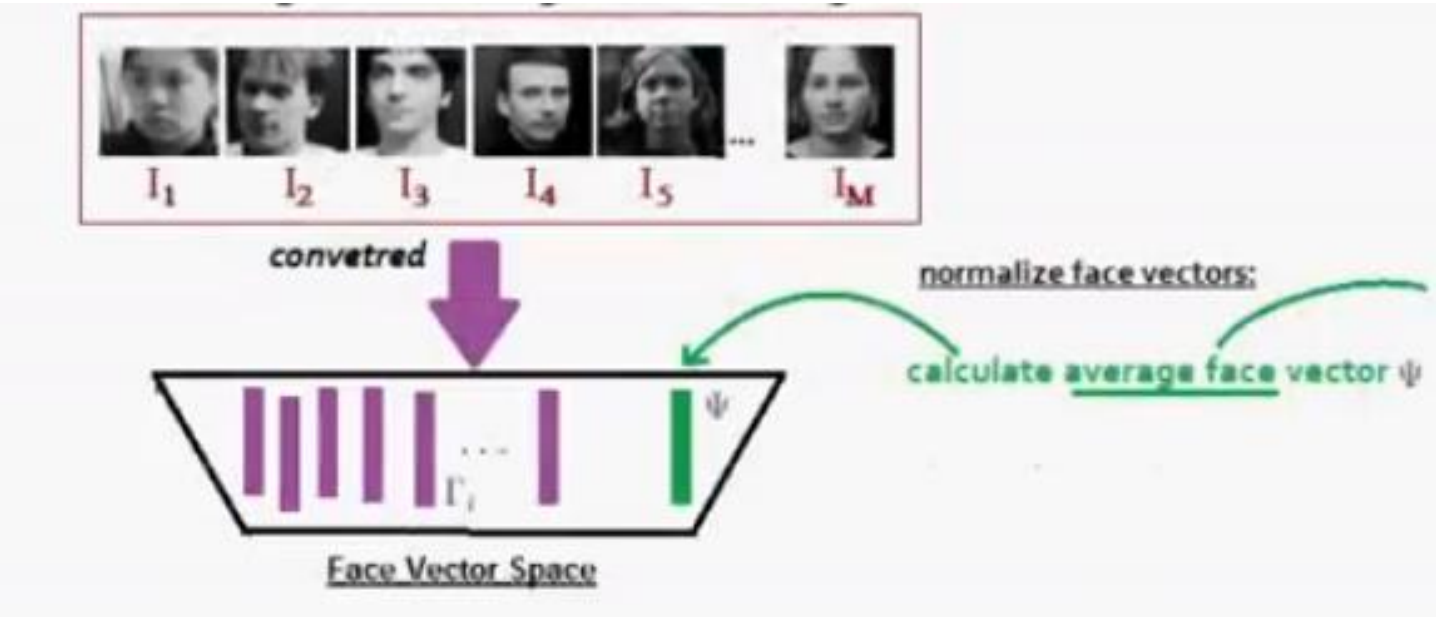


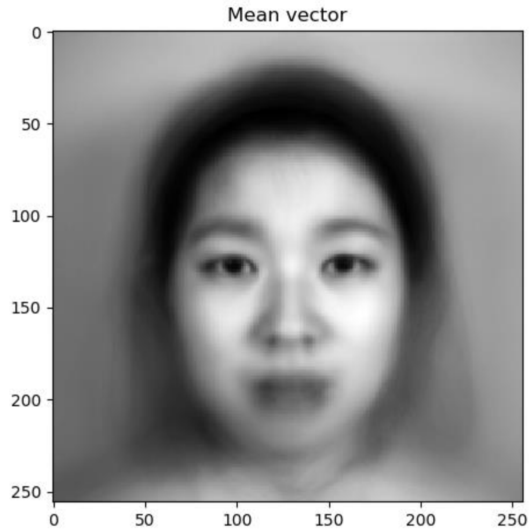
256 X 256 X 3 = 196 608

256 X 256 X 1 = 65 536

• Obtener la imagen promedio

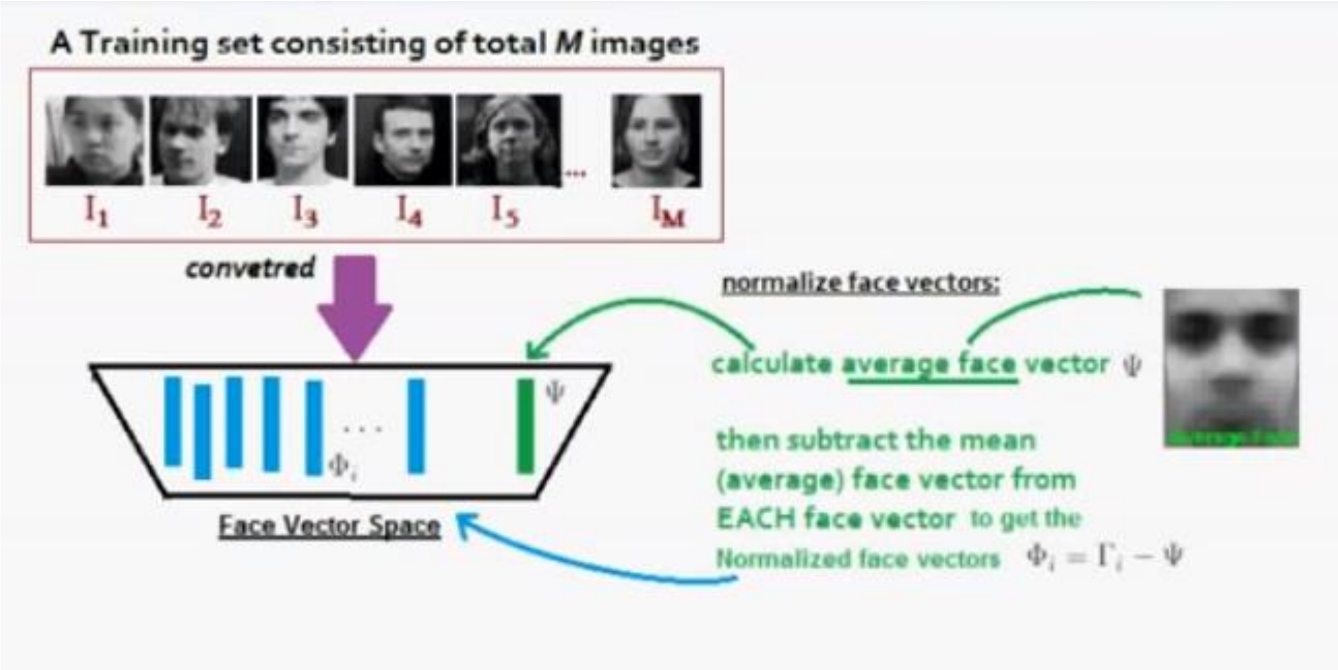
Paso 2





Representa las características promedio del dataset

Paso 3



A Training set consisting of total M images
I1
I2
I3
I4
I5
IM
convetred
normalize face vectors:
calculate average face vector Ψ
then subtract the mean
(average) face vector from
EACH face vector to get the
Normalized face vectors Φi = Γi − Ψ
Face Vector Space

Se calcula los eigenvectors

- Para realizar esto es necesario calcular la matriz de covarianza C

$$C = AA^T$$

Donde

$$A = [\Phi_1 \Phi_2 \dots \Phi_M]$$

Pero A tiene dimensión $N^2 \times M$

Por lo que C tendría dimensión

$$N^2 \times M \;\cdot M \times N^2 = N^4$$

En nuestro caso, 65536 X 65536,

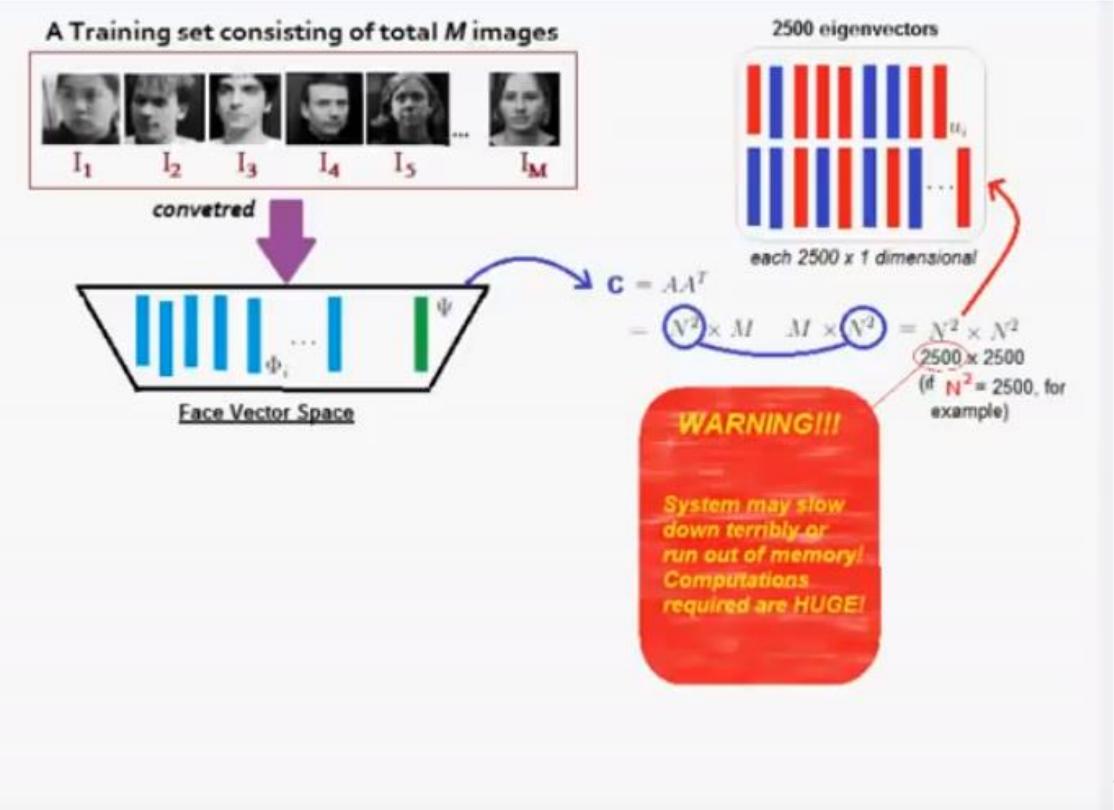
Una matriz de este tamaño generaría 65536 eigenvectors de dimensión 65536 x 1

Objetivo PCA: representar cada imagen como una combinación lineal de los k eigenvectors

Se calcula los eigenvectors

$$K < M$$

En nuestro caso M = 197



Se calcula los eigenvectors

- **Por ello se calcula C de dimensión reducida**

$$C = A^T A$$

$$M \times N^2 \cdot N^2 \times M = M^2$$

Ahora, 197 X 197; es decir 197 eigenvectors de dimensión 197 X 1.

Podemos ver la diferencia con 65536 eigenvectors de dimensión 65536 x 1

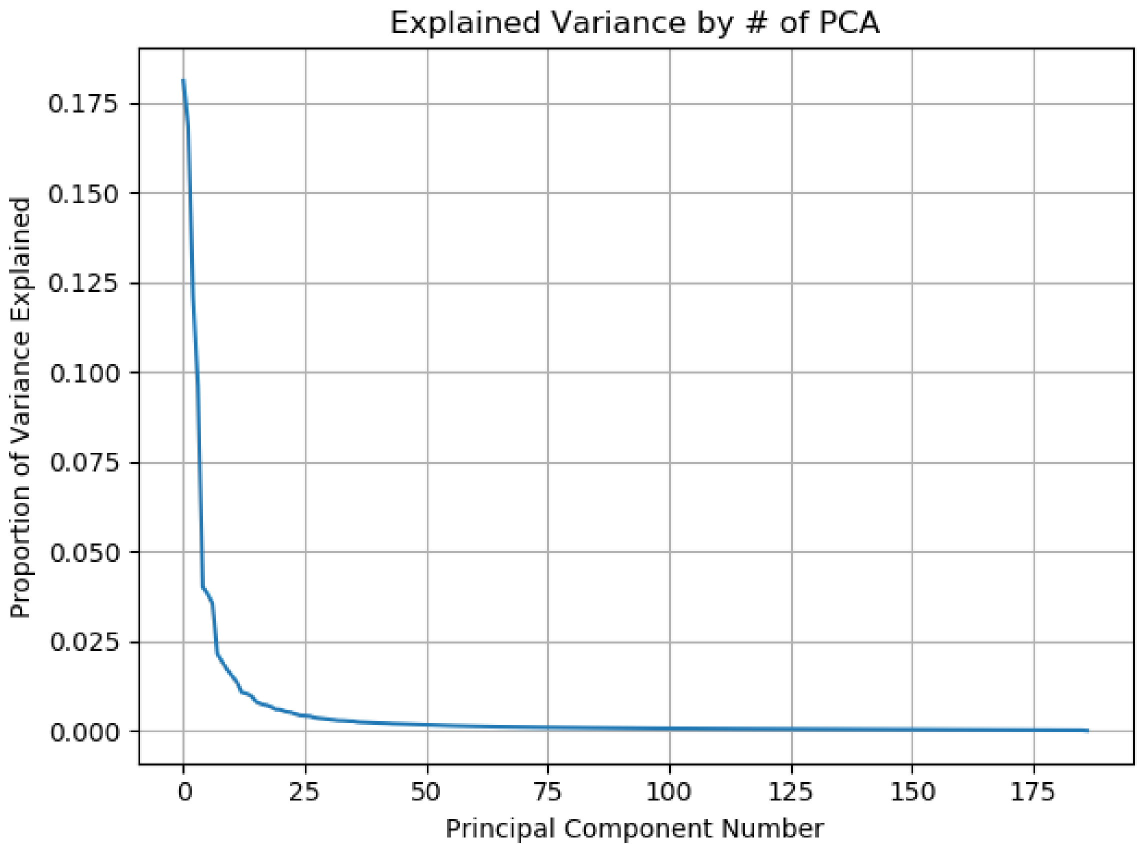
Seleccionar K eigenvectors (eigenfaces)

Los que capturen más varianza

Variance

- The **total variance** is the sum of variances of all individual principal components
- The fraction of **variance explained** by a principal component is the ratio between the variance of that principal component and the total variance.

Seleccionar K eigenvectors (eigenfaces)



Seleccionar K eigenvectors (eigenfaces)

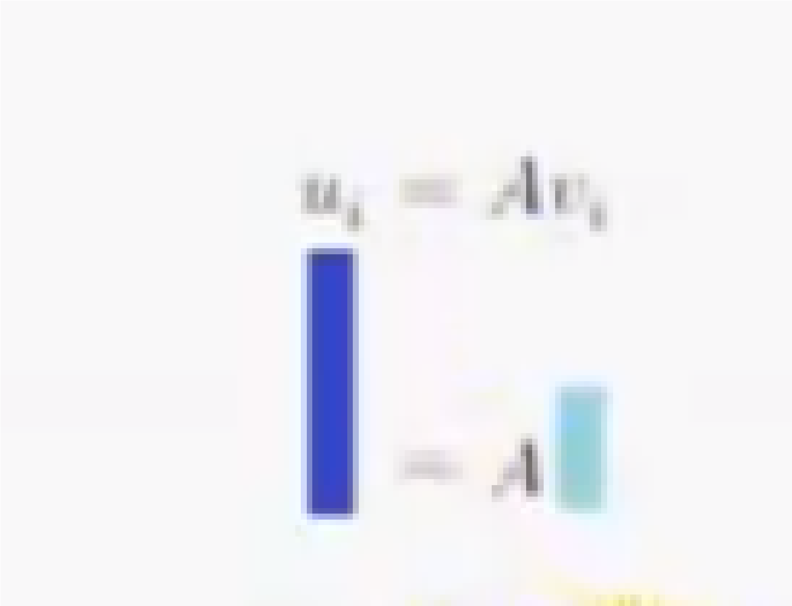
1 PC explain (0.181) variance. SUM: (0.181)
2 PC explain (0.168) variance. SUM: (0.349)
3 PC explain (0.121) variance. SUM: (0.47)
4 PC explain (0.096) variance. SUM: (0.566)
5 PC explain (0.04) variance. SUM: (0.607)
6 PC explain (0.038) variance. SUM: (0.645)
7 PC explain (0.035) variance. SUM: (0.68)
8 PC explain (0.021) variance. SUM: (0.701)
9 PC explain (0.019) variance. SUM: (0.721)
10 PC explain (0.017) variance. SUM: (0.738)
11 PC explain (0.015) variance. SUM: (0.753)
12 PC explain (0.014) variance. SUM: (0.767)
13 PC explain (0.011) variance. SUM: (0.777)
14 PC explain (0.01) variance. SUM: (0.788)
15 PC explain (0.01) variance. SUM: (0.797)
16 PC explain (0.008) variance. SUM: (0.805)
17 PC explain (0.007) variance. SUM: (0.813)
18 PC explain (0.007) variance. SUM: (0.82)
19 PC explain (0.007) variance. SUM: (0.827)
20 PC explain (0.006) variance. SUM: (0.833)
21 PC explain (0.006) variance. SUM: (0.838)
22 PC explain (0.005) variance. SUM: (0.844)
23 PC explain (0.005) variance. SUM: (0.849)
24 PC explain (0.005) variance. SUM: (0.854)
25 PC explain (0.004) variance. SUM: (0.858)
26 PC explain (0.004) variance. SUM: (0.862)
27 PC explain (0.004) variance. SUM: (0.866)
28 PC explain (0.004) variance. SUM: (0.87)
29 PC explain (0.003) variance. SUM: (0.873)
30 PC explain (0.003) variance. SUM: (0.876)
31 PC explain (0.003) variance. SUM: (0.88)
32 PC explain (0.003) variance. SUM: (0.883)
33 PC explain (0.003) variance. SUM: (0.885)
34 PC explain (0.003) variance. SUM: (0.888)
35 PC explain (0.003) variance. SUM: (0.891)
36 PC explain (0.003) variance. SUM: (0.894)
37 PC explain (0.002) variance. SUM: (0.896)
38 PC explain (0.002) variance. SUM: (0.898)
39 PC explain (0.002) variance. SUM: (0.901)
40 PC explain (0.002) variance. SUM: (0.903)
41 PC explain (0.002) variance. SUM: (0.905)
42 PC explain (0.002) variance. SUM: (0.907)
43 PC explain (0.002) variance. SUM: (0.909)
44 PC explain (0.002) variance. SUM: (0.911)
45 PC explain (0.002) variance. SUM: (0.913)
46 PC explain (0.002) variance. SUM: (0.915)
47 PC explain (0.002) variance. SUM: (0.917)
48 PC explain (0.002) variance. SUM: (0.918)
49 PC explain (0.002) variance. SUM: (0.92)
50 PC explain (0.002) variance. SUM: (0.922)
51 PC explain (0.002) variance. SUM: (0.923)
52 PC explain (0.002) variance. SUM: (0.925)
53 PC explain (0.001) variance. SUM: (0.926)
54 PC explain (0.001) variance. SUM: (0.928)
55 PC explain (0.001) variance. SUM: (0.929)
56 PC explain (0.001) variance. SUM: (0.93)
57 PC explain (0.001) variance. SUM: (0.932)
58 PC explain (0.001) variance. SUM: (0.933)
59 PC explain (0.001) variance. SUM: (0.935)
60 PC explain (0.001) variance. SUM: (0.936)
61 PC explain (0.001) variance. SUM: (0.937)
62 PC explain (0.001) variance. SUM: (0.938)
63 PC explain (0.001) variance. SUM: (0.939)
64 PC explain (0.001) variance. SUM: (0.941)
65 PC explain (0.001) variance. SUM: (0.942)
66 PC explain (0.001) variance. SUM: (0.943)
67 PC explain (0.001) variance. SUM: (0.944)
68 PC explain (0.001) variance. SUM: (0.945)
69 PC explain (0.001) variance. SUM: (0.946)
70 PC explain (0.001) variance. SUM: (0.947)
71 PC explain (0.001) variance. SUM: (0.948)
72 PC explain (0.001) variance. SUM: (0.949)
73 PC explain (0.001) variance. SUM: (0.95)

Seleccionar 16 eigenvectors (eigenfaces)

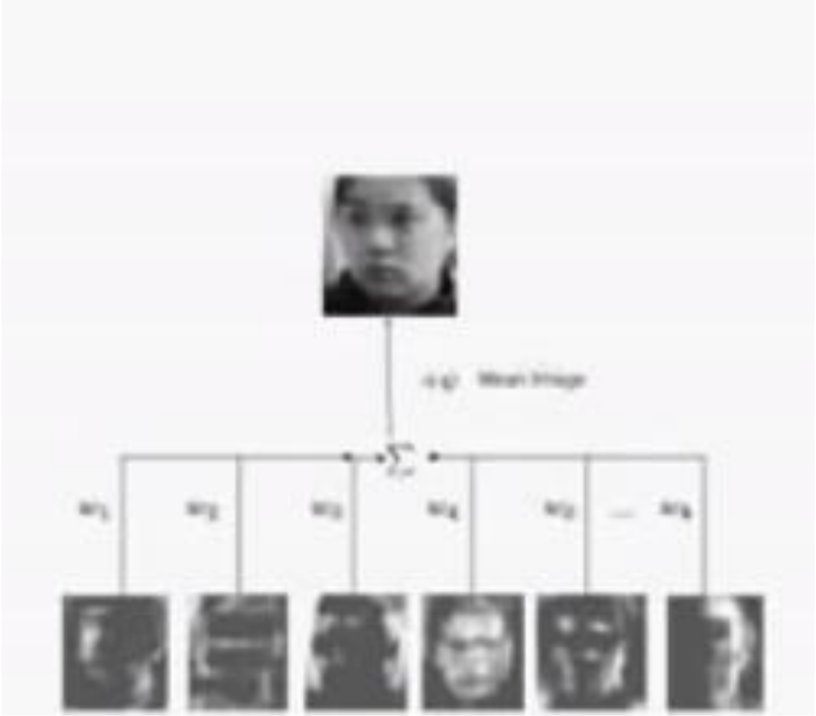
¿Cómo volver a la dimensionalidad original 256X256?

eigenfaces = pca.components_

Mapear



Entrenar



$$\Omega_i = \begin{bmatrix} w_1 \\ \vdots \\ w_k \end{bmatrix}$$

Finalmente medir la distancia entre dos vectores omega