



INSTITUTO TECNOLÓGICO DE COSTA RICA

ÁREA DE INGENIERÍA EN COMPUTADORES

PROCESAMIENTO Y ANÁLISIS DE IMÁGENES DIGITALES

Identificación de Rostros

Estudiantes:

Arturo CÓRDOBA V.
Fabián GONZÁLEZ A.
Erick CARBALLO P.
Sebastián MORA R.

Profesor:

PhD. Juan Pablo SOTO

25 de mayo de 2020

1. Identificación de Rostros

Este documento presenta la solución al problema donde se requiere realizar el reconocimiento de cada una de las imágenes de rostros que se encuentran en un archivo de nombre *Comparar*, implementando el algoritmo explicado en el artículo científico *Facial Recognition with Singular Value Decomposition*, el cual implementa un sistema para el reconocimiento de rostros faciales, utilizando el concepto de espacios vectoriales de rostros y la descomposición en valores singulares.

En primer lugar, se debe generar un espacio de rostros faciales utilizando la base de datos de imágenes que se encuentran en el archivo *Database*, que contiene un conjunto de imágenes de rostros de la base de datos de ORL. Esta base de datos está formada por 360 imágenes de caras de 40 personas, es decir, 9 imágenes de caras por persona y cada imagen es de tamaño 92×112 .

Para ello, se construye un conjunto S con las 9 imágenes de caras de las 40 personas expresado de la siguiente manera:

$$S = [f_1, f_2, \dots, f_{360}]$$

Donde cada f_i representa las imágenes de rostros por persona en forma vectorial. Es decir, S resultaría en una matriz de tamaño 10304×360 , donde cada columna es una de las imágenes vectorizadas de la base de datos. Luego se procede a calcular el rostro promedio \bar{f} del conjunto S de la siguiente manera:

$$\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i$$

Donde $N = 360$, que representa el número de columnas del conjunto S y el resultado de este rostro promedio se puede observar en la siguiente figura:



Figura 1: Rostro Promedio

Después de calcular el rostro promedio, se resta \bar{f} a cada una de las imágenes de los rostros originales para obtener la siguiente matriz:

$$A = [a_1, a_2, \dots, a_{360}]$$

Donde cada $a_i = f_i - \bar{f}$, $i = 1, 2, \dots, 360$. Se asume un rango r de A , donde $r \leq 360 \ll 10304$. Una vez obtenida la matriz A , se calcula la descomposición en valores singulares (SVD) de la matriz A de la siguiente manera:

$$A = U \Sigma V^T$$

Donde Σ es una matriz diagonal de tamaño 10304×360 :

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{r+1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma_N \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Donde $i = 1, 2, \dots, 360$ y σ_i son los valores singulares (SV) de la matriz A . La matriz V es una matriz ortogonal de tamaño 360×360 y la matriz U es una matriz ortogonal de tamaño 10304×10304 expresadas de la siguiente manera:

$$V = [v_1, v_2, \dots, v_r, v_{r+1}, \dots, v_{360}]$$

$$U = [u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_{10304}]$$

Luego de realizar la descomposición en valores singulares, para cada persona se calcula el vector de coordenadas x_i , que se utiliza para encontrar cual de los rostros del conjunto describe mejor el rostro f de la siguiente forma:

$$x_i = [u_1, u_2, \dots, u_r]^T (f_i - \bar{f})$$

Donde cada u_i , $i = 1, 2, \dots, r$, es conocido como un *rostro base*. Se debe escoger un límite ε_1 que defina la distancia máxima permitida del espacios de caras, que se definió como $\varepsilon_1 = 10$. También se debe determinar un límite ε_0 que define la distancia máxima permitida para cada rostro conocido en el conjunto S , que se calculó a partir de la imágenes de la base de datos de la siguiente manera:

$$\varepsilon_0 = \|x - x_0\|_2 = [(x - x_0)^T (x - x_0)]^{1/2}$$

Donde su resultado es $\varepsilon_0 = 3,8780$. Para realizar una comparación e identificar una cara ingresada contra las caras almacenadas en la base de datos, se debe calcular su propio vector de coordenadas x , el vector de proyección f_p y la distancia ε_f al espacio de caras, con las siguientes fórmulas:

$$x = [u_1, u_2, \dots, u_r]^T (f - \bar{f})$$

$$f_p = [u_1, u_2, \dots, u_r] \cdot x$$

$$\varepsilon_f = \|(f - \bar{f}) - f_p\|_2 = [(f - \bar{f} - f_p)^T (f - \bar{f} - f_p)]^{1/2}$$

Si $\varepsilon_f > \varepsilon_1$, la imagen ingresada no es una cara. Si $\varepsilon_f < \varepsilon_1$, se calcula la distancia ε_i , para todas las caras de las personas conocidas. Si todos los $\varepsilon_i > \varepsilon_0$, la imagen ingresada se puede considerar como un rostro desconocido y puede ser utilizado más para empezar una nueva cara. Si $\varepsilon_f < \varepsilon_1$ y algunos $\varepsilon_i > \varepsilon_0$, se clasifica la imagen ingresada como la persona conocida asociada con el mínimo ε_i de los resultados de la fórmula:

$$\varepsilon_i = \|x - x_i\|_2 = [(x - x_i)^T (x - x_i)]^{1/2}$$

2. Pseudocódigo

En el algoritmo 1 se muestra el pseudocódigo de los pasos explicados en el artículo científico *Facial Recognition with Singular Value Decomposition*. Como entrada se recibe la matriz $S \in \mathbb{R}^{M \times N}$, donde $M = m \times n$, la cual contiene los datos de entrenamiento, cada columna es una imagen vectorizada $f_i \in \mathbb{R}^{M \times 1}$ cuyas dimensiones originales fueron $m \times n$. También recibe como entrada una imagen vectorizada $f \in \mathbb{R}^{M \times 1}$, la cual será comparada con las imágenes presentes en S . Como salida se tiene una imagen vectorizada $Y \in \mathbb{R}^{M \times 1}$ la cual es la coincidencia encontrada por el algoritmo.

Algorithm 1 Facial Recognition with Singular Value Decomposition

Entradas: $S = [f_1, f_2, \dots, f_N] \in \mathbb{R}^{M \times N}$, $f \in \mathbb{R}^{M \times 1}$

Salida: $Y \in \mathbb{R}^{M \times 1}$

```
1:  $\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i$  Cálculo del rostro promedio
2: for  $i = 1, 2, \dots, N$  do
3:    $f_i = S(:, i)$  Obtener cada imagen
4:    $a_i = f_i - \bar{f}$  Restarle el rostro promedio a cada imagen
5: end for
6:  $A = [a_1, a_2, \dots, a_N] \in \mathbb{R}^{M \times N}$ 
7:  $A = U \Sigma V^T$  Descomposición SVD de A
8:  $r = \text{rank}(A)$ 
9: for  $i = 1, 2, \dots, N$  do
10:    $x_i = U_r^T \cdot A(:, i)$ 
11: end for
12:  $x = [x_1, x_2, \dots, x_N]$ 
13: Elegir un  $\varepsilon_i$  y un  $\varepsilon_0$ 
14: for nueva imagen  $f$  ingresada do
15:   Calcular  $x_f$  para  $f$ 
16:    $f_p = U_r \cdot x_f$ 
17:    $\varepsilon_f = \|x - x_f\|_2$ 
18:   if  $\varepsilon_f > \varepsilon_i$  then
19:      $f$  no es un rostro
20:   else
21:     for  $i = 1, 2, \dots, N$  do
22:        $\varepsilon_i = \|x_f - x_i\|_2$  Calcular la distancia a cada imagen
23:     end for
24:      $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N]$ 
25:      $pos =$  Obtener la posición del menor elemento en  $\varepsilon$ 
26:     return  $Y = S(:, pos)$ 
27:   end if
28: end for
```

3. Resultados obtenidos

Para probar el algoritmo de reconocimiento de rostros se ingresaron cada una de las imágenes del archivo *Comparar*, que contiene 40 imágenes, de tamaño de 92×112 cada una, de las caras de las 40 personas que se utilizaron para la base de datos. Se debe destacar que estas imágenes son diferentes a las imágenes del archivo *Database*. Se obtuvieron los siguientes resultados para cada una de la imágenes ingresadas:

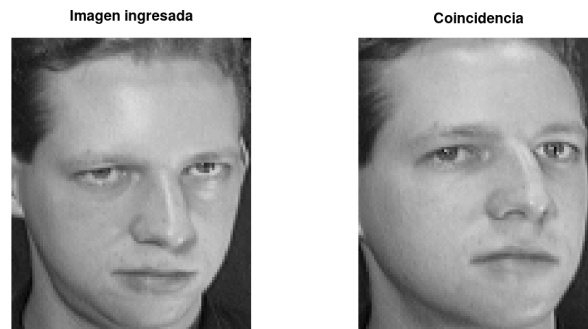


Figura 2: Comparación 1

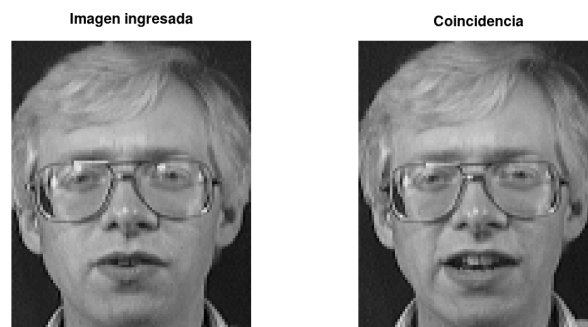


Figura 3: Comparación 2

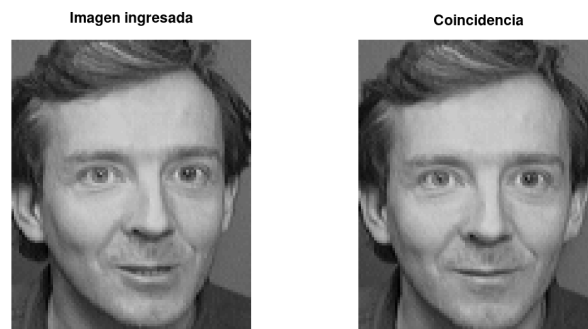


Figura 4: Comparación 3

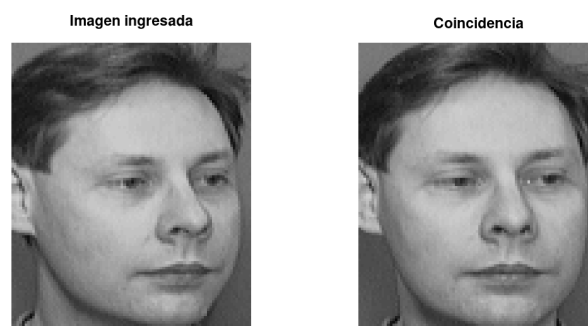


Figura 5: Comparación 4

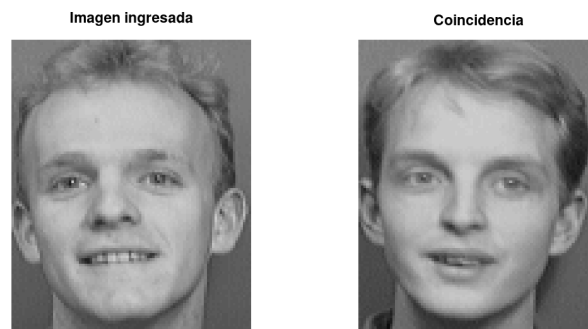


Figura 6: Comparación 5

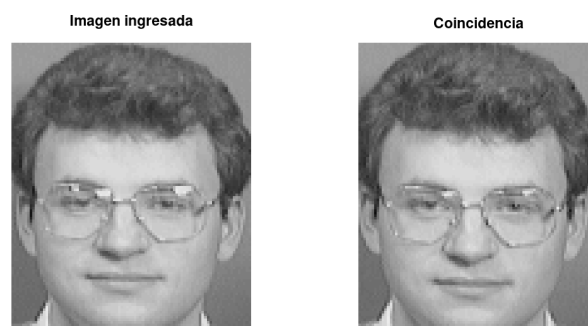


Figura 7: Comparación 6

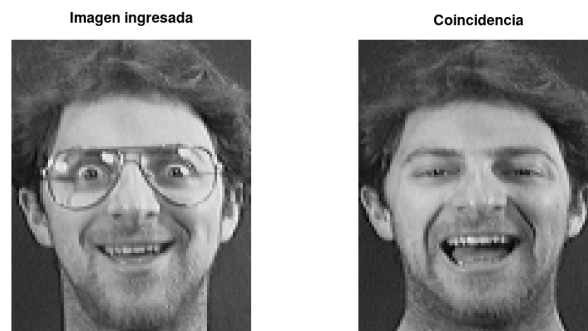


Figura 8: Comparación 7

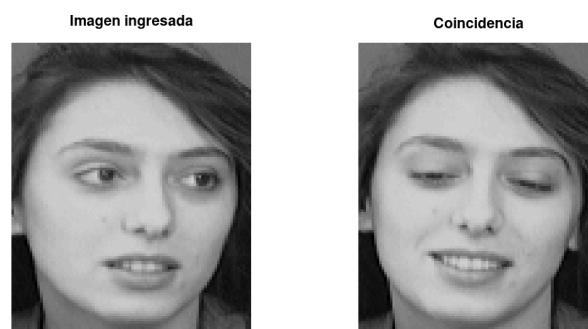


Figura 9: Comparación 8

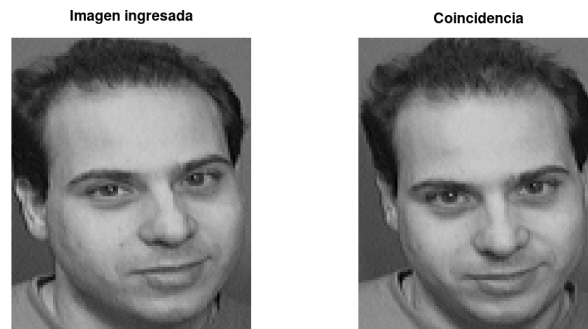


Figura 10: Comparación 9

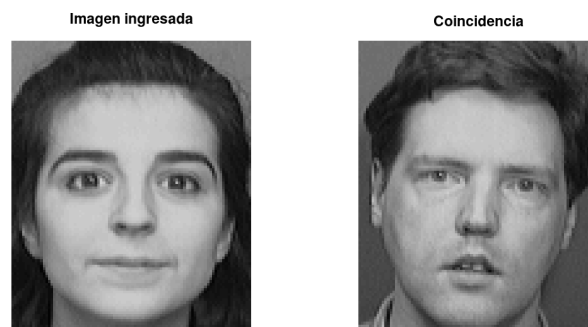


Figura 11: Comparación 10

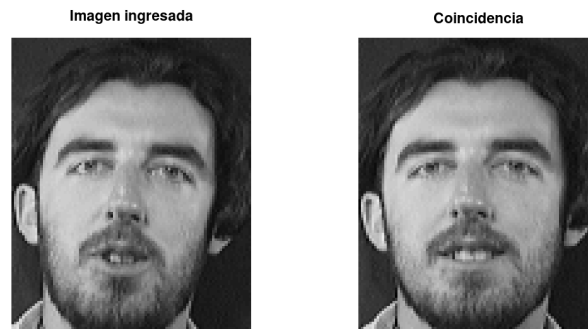


Figura 12: Comparación 11

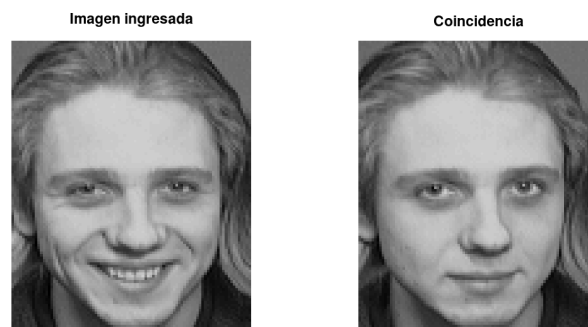


Figura 13: Comparación 12

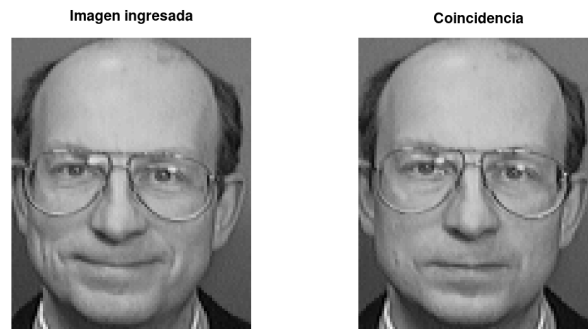


Figura 14: Comparación 13

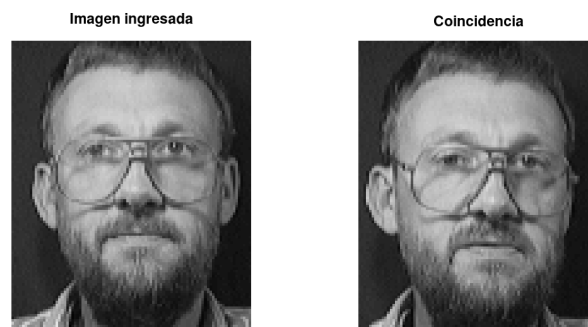


Figura 15: Comparación 14

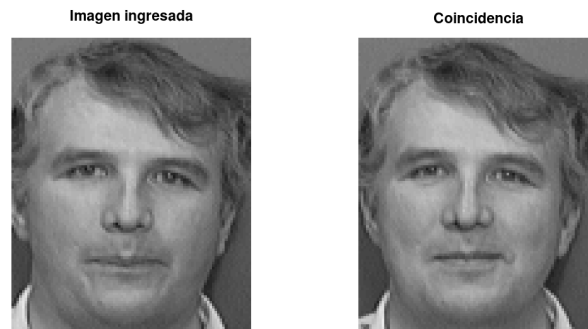


Figura 16: Comparación 15

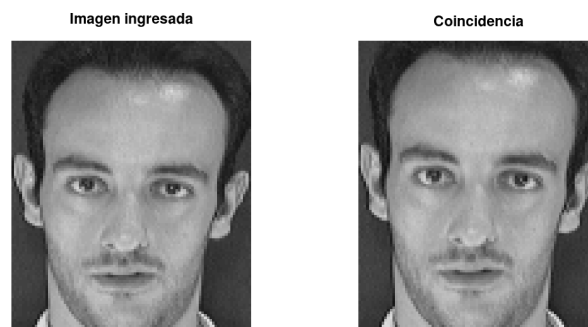


Figura 17: Comparación 16

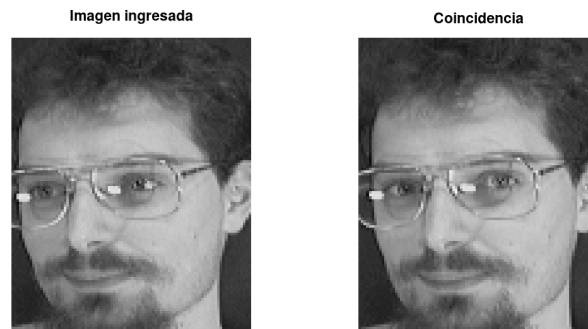


Figura 18: Comparación 17

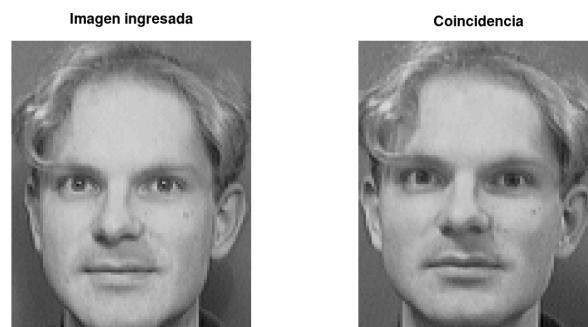


Figura 19: Comparación 18

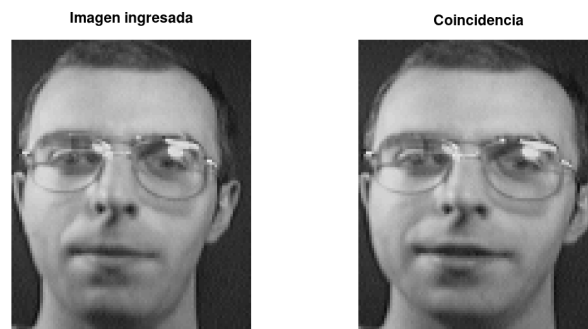


Figura 20: Comparación 19

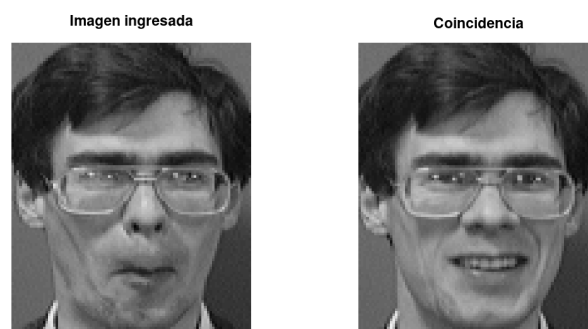


Figura 21: Comparación 20

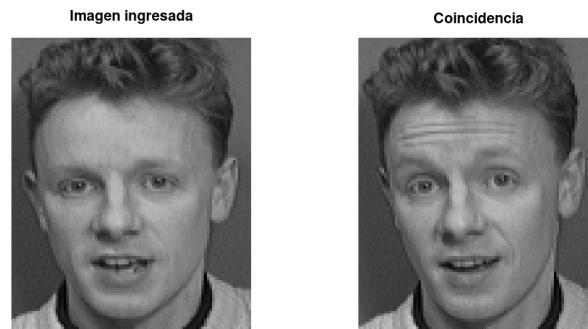


Figura 22: Comparación 21

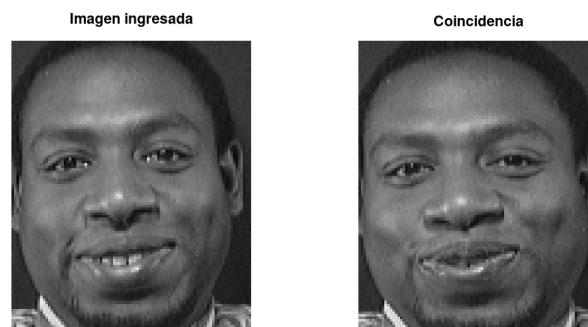


Figura 23: Comparación 22

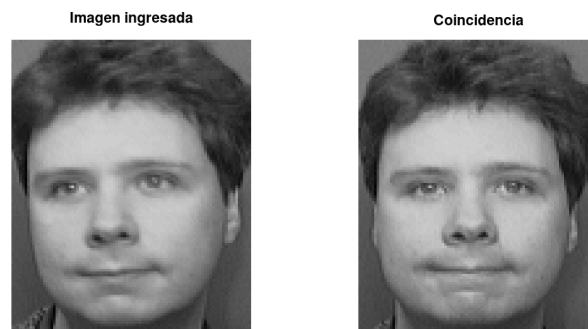


Figura 24: Comparación 23

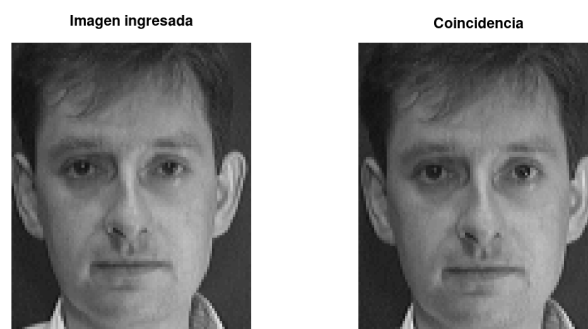


Figura 25: Comparación 24

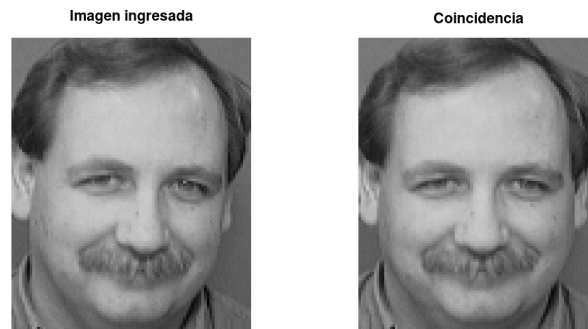


Figura 26: Comparación 25

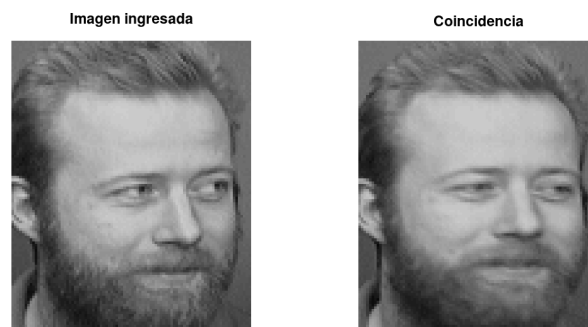


Figura 27: Comparación 26

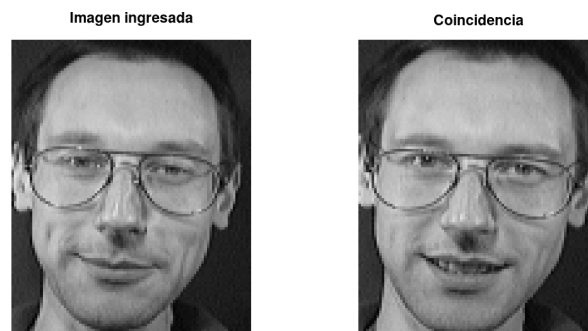


Figura 28: Comparación 27

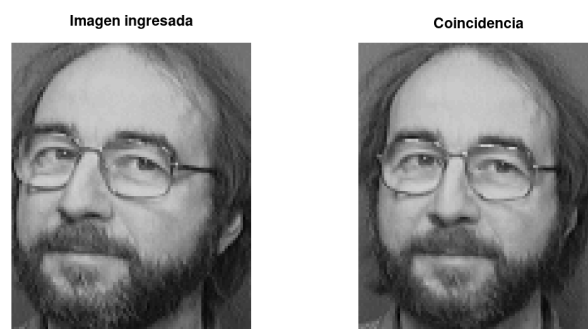


Figura 29: Comparación 28

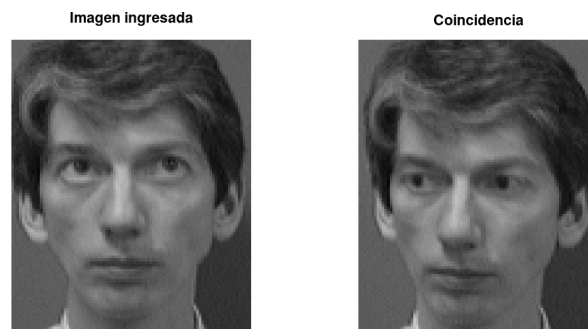


Figura 30: Comparación 29

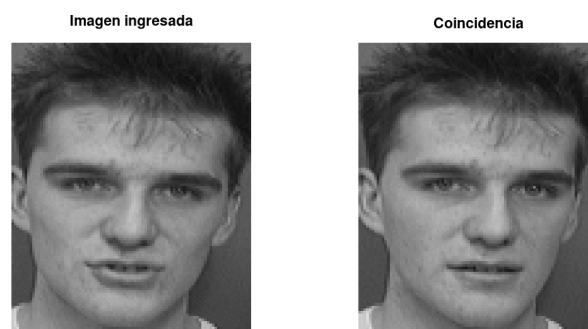


Figura 31: Comparación 30

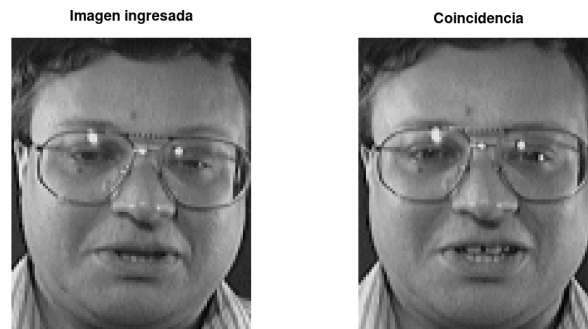


Figura 32: Comparación 31

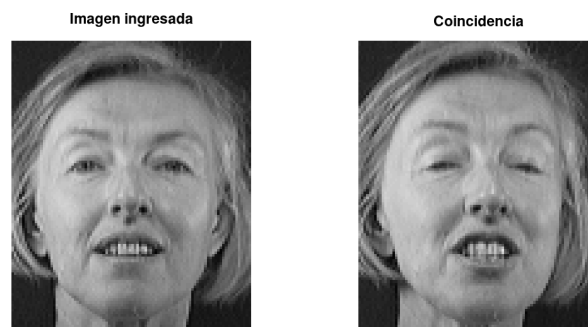


Figura 33: Comparación 32

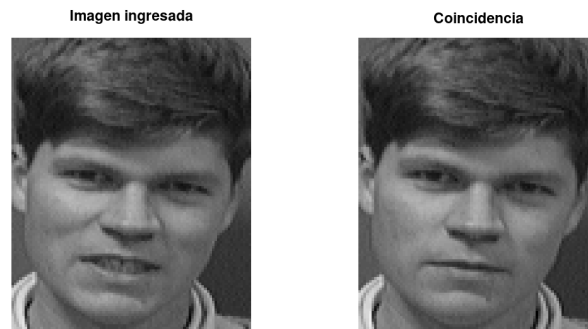


Figura 34: Comparación 33

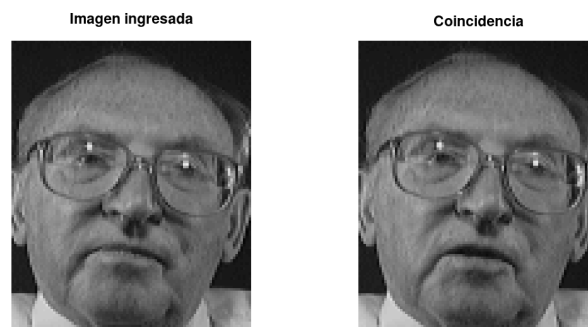


Figura 35: Comparación 34



Figura 36: Comparación 35

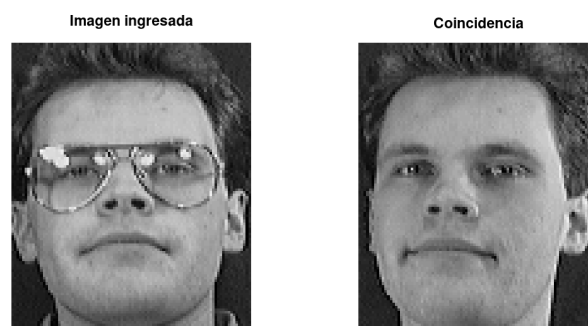


Figura 37: Comparación 36

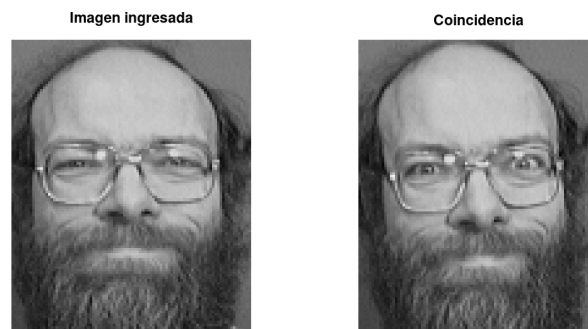


Figura 38: Comparación 37

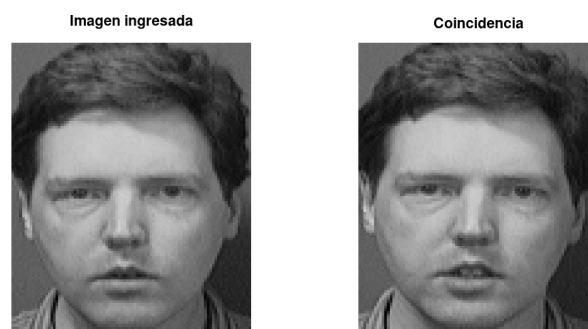


Figura 39: Comparación 38



Figura 40: Comparación 39

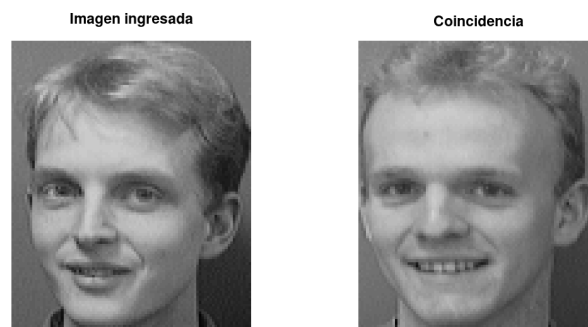


Figura 41: Comparación 40

4. Referencias Bibliográficas

- [1] G. Zeng, "Facial Recognition with Singular Value Decomposition", Arizona, 2007.