



Cinvestav
Unidad Saltillo

Eigenhands For Recognition

Detección y reconocimiento de gestos de mano

Gabriela Hilario & Luis Alberto Pineda

17 de Abril de 2023



Descripción del proyecto

Detección y reconocimiento de gestos de mano en 2D, implementando un algoritmo basado en Principal Component Analysis (PCA), llamado *Eigenfaces for Recognition*, y descrito por Turk y Pentland en su artículo publicado en 1991 [3].

Objetivos

- Utilizar una base de datos disponible en la web o crear una (*Training*).
- Generar una base pequeña con gestos para las pruebas experimentales (*Test*).
- Implementar el algoritmo de *Eigenfaces* pero enfocado para gestos de mano.

Tabla de contenido

1 Introducción

- Introducción
- Materiales y Métodos
- Resultados
- Conclusiones
- References

Principal Component Analysis (PCA)

1 Introducción

El Análisis de Componentes Principales (PCA), es un método estadístico cuya utilidad radica en la reducción de la dimensionalidad o de variables dado un conjunto de datos.

Existe un *trade-off* [precisión - simplicidad] al momento de utilizar PCA.

Pasos del PCA - 1

1 Introducción

El análisis por PCA puede descomponer en los siguientes pasos:

- Estandarización de los datos. Esto evita que haya un **sesgo o bias**.

$$Z = \frac{\text{valor} - \text{media}}{\text{desviacion estandar}} \quad (1)$$

- Obtención de la matriz de covarianza. Se busca conocer cómo es que las variables **varian** entre ellas.

$$\begin{bmatrix} Cov(x, x) & Cov(x, y) & Cov(x, z) \\ Cov(y, x) & Cov(y, y) & Cov(y, z) \\ Cov(z, x) & Cov(z, y) & Cov(z, z) \end{bmatrix}$$

Pasos del PCA - 2

1 Introducción

- Obtención de los eigenvalores y eigenvectores de la matriz de covarianza, los cuales son los **componentes principales**. Cada componente principal contiene información **única**. Los componentes carecen de explicación o sentido.
- Construcción del vector característico. Se toma la decisión de descartar los **eigenvectores** menos relevantes.
- Reconformar la base de datos a lo largo de los **eigenvectores más relevantes**.

Principal Component Analysis (PCA)

1 Introducción

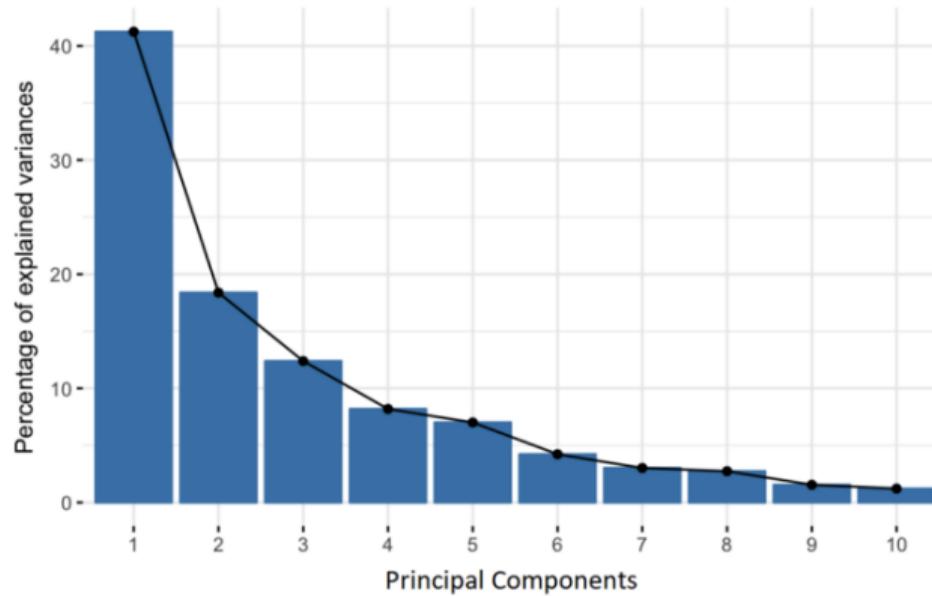


Figure: Comparación del Porcentaje de la varianza *vs* el número de componentes principales

Algoritmo de Eigenfaces

1 Introducción

Fue presentado por primera vez por Sirovich y Kirby en 1987 [2], y posteriormente formalizado por Turk y Pentland en 1991 [3].

Procedimiento General

Cada cara se almacena en un vector de dimensión N^2 . Y el Análisis de los Componentes Principales (PCA) se utiliza para encontrar un subespacio de dimensión M cuyos vectores de la base corresponden a las direcciones de máxima varianza en el espacio original de la imagen. Este nuevo subespacio es normalmente de dimensión más baja ($M \ll N^2$).

Etapas del algoritmo

1 Introducción

(I) Initialization

Training set
 $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M$



Compute eigenfaces
 $u_l = \sum_{k=1}^M v_{lk} \Phi_k$



Calculate the class
vector Ω_k for each
known class
 $\Omega^T = [\omega_1, \omega_2, \dots, \omega_M]$

(II) Classification

Calculate the pattern
vector Ω , for each
new image



Calculate the
euclidean distances
to each known class
 $\epsilon_k = \|\Omega - \Omega_k\|^2$



Classify the input
image according to
the smallest euclidean
distance ϵ_k

Figure: Etapas del algoritmo eigenfaces.

Tabla de contenido

2 Materiales y Métodos

- Introducción
- Materiales y Métodos
- Resultados
- Conclusiones
- References

Materiales

2 Materiales y Métodos

- Computador
- Cámara/sensor

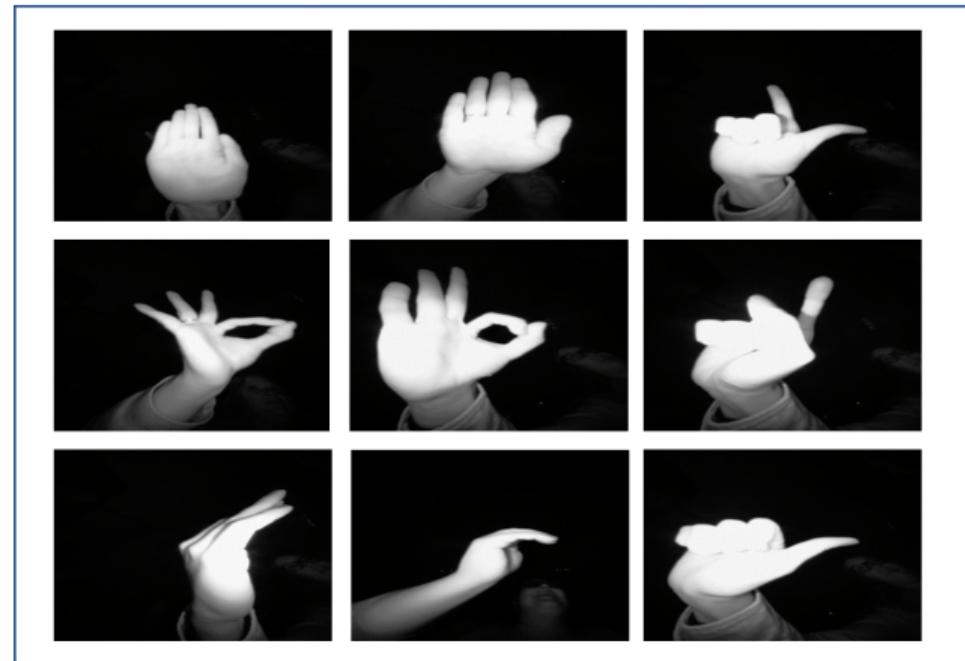


Figure: Infrared and depth sensor.

Adquisición de la BD

2 Materiales y Métodos

- Inicialmente se estuvo trabajando con una base de datos de gestos manuales, alojada en la plataforma kaggle.

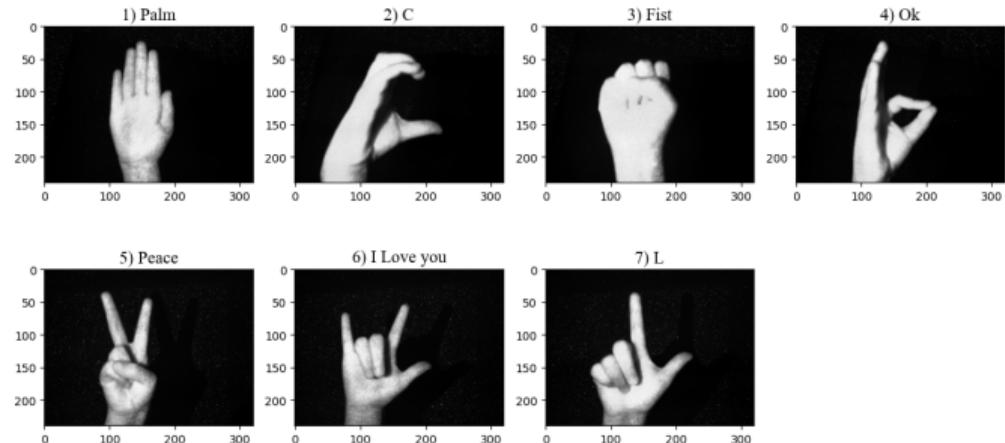


Adquisición de la BD

2 Materiales y Métodos

- Se optó por crear una base de datos propia, tanto para el *training set* como para el *test set*.

Gestos para el set de entrenamiento



Preprocesamiento

2 Materiales y Métodos

Es necesario el **estandarizar** las imágenes antes de que sean ingresadas al sistema. Para ello, se realizó el siguiente preprocesamiento, mostrado en la figura 4.

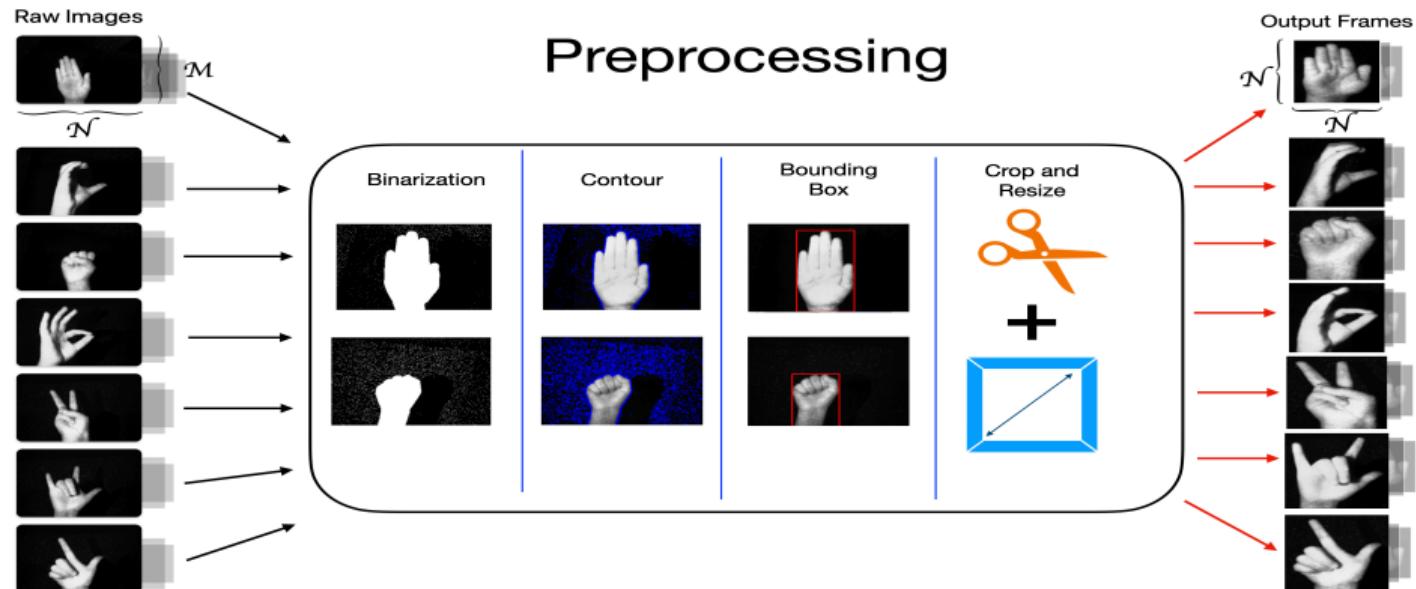


Figure: Preprocesamiento de las imágenes para conformar la base de datos de sistema

Procedimiento del PCA - 1

2 Materiales y Métodos

Comenzamos transformando nuestra imagen de entrada de dimensiones $N \times N$ a un vector columna de dimensiones $1 \times N^2$. Ver figura 5.

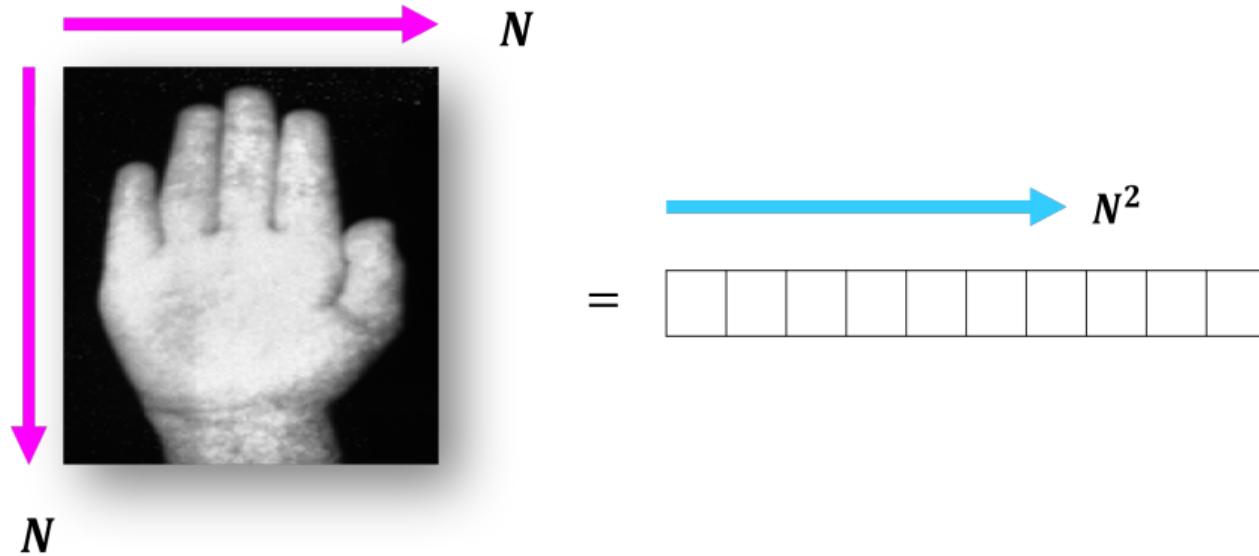


Figure: Vector imagen.

Procedimiento del PCA - 2

2 Materiales y Métodos

Ya con todas las imágenes del data-set aplanadas, formamos la siguiente matriz, la cuál tiene dimensiones de $M \times N^2$ (M Imágenes de entrada por N^2 número de píxeles). A esta le llamamos conjunto de entrenamiento $\Gamma_1, \Gamma_2, \dots, \Gamma_M$.

$$\begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_M \end{bmatrix} = \left(\begin{array}{cccccccc} \boxed{} & \boxed{} \\ \boxed{} & \boxed{} \\ \boxed{} & \boxed{} \\ \vdots & & & & & & & \\ \vdots & & & & & & & \\ \boxed{} & \boxed{} \end{array} \right)_{M \times N^2}$$

Figure: Conjunto de entrenamiento.

Fase 1 - Mano promedio

2 Materiales y Métodos

Obtenemos la mano promedio de acuerdo a la ecuación (2).

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad (2)$$

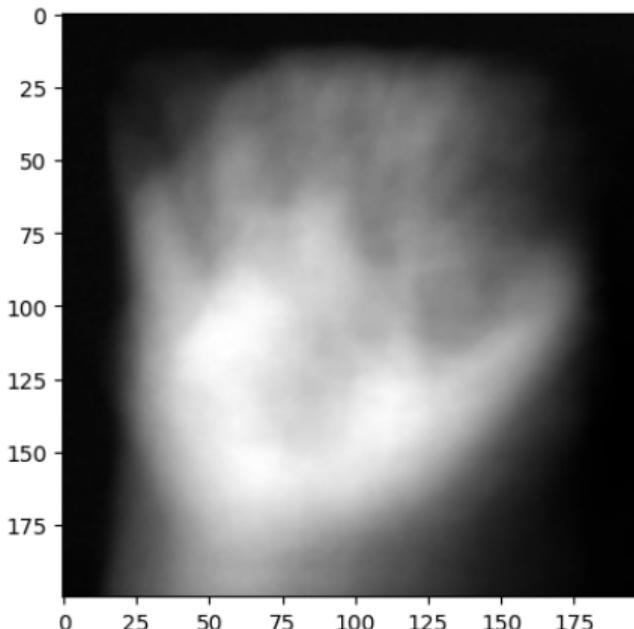


Figure: Mano promedio

Fase 1 - Desviaciones estándar

2 Materiales y Métodos

Ya con la mano promedio, obtenemos las desviaciones estándar que presenta cada imagen de entrada contra la mano promedio Ψ :

$$\Phi = \Gamma_i - \Psi \quad (3)$$

Cada uno de los vectores resultantes de realizar la diferencia, es almacenado en la matriz A :

$$A = [\Phi_1, \Phi_2, \dots, \Phi_M] \quad (4)$$

Ya con la matriz A construida, obtenemos la matriz L con dimensiones $M \times M$:

$$L = A^T A \quad (5)$$

Fase 1 - Obtención de los eigenvalores y eigenvectores

2 Materiales y Métodos

Encontramos los **eigenvalores** μ_i y los **eigenvectores** v_i de la matriz L . Tales que satisfagan la siguiente relación:

$$Av_i = \mu_i v_i \tag{6}$$

Fase 1 - Suma acumulada de PCA

2 Materiales y Métodos

Ya que contamos con los **eigenvalores** μ_i y los **eigenvectores** v_i de la matriz L , se encuentra aquellos que sean los más representativos. Para esto, se hace uso de la **suma acumulada de los componentes principales**. Esta suma se muestra en la siguiente figura 9.

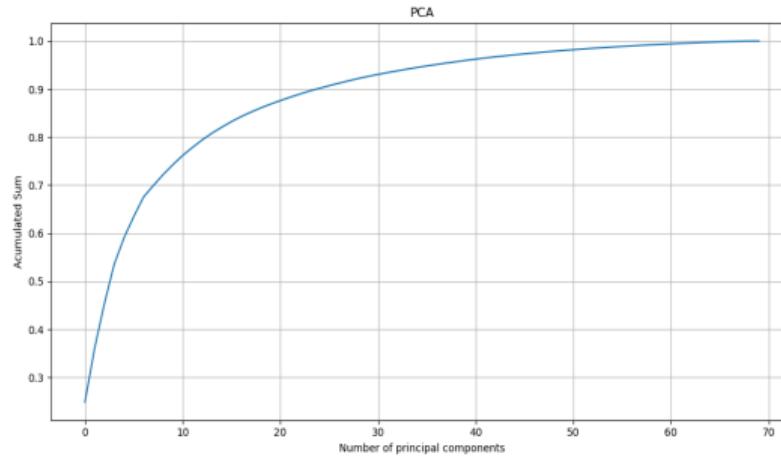


Figure: Suma acumulada de los componentes principales de la matriz L .

Fase 1 - Subespacio $\mathbb{R}^M \rightarrow \mathbb{R}^{M'}$

2 Materiales y Métodos

Reducción de dimensionalidad:

Los vectores v_i asociados a los λ_i más representativos se utilizan para formar un subespacio de dimensión $M'xN^2$, de acuerdo a (7).

$$u_l = \sum_{k=1}^{M'} v_{lk} \Phi_k, \quad l = 1, \dots, M' \quad (7)$$

Para este proyecto se utilizaron $M' = 30$

Fase 1 - Eigenhands

2 Materiales y Métodos

A continuación se muestran las 6 primeras **Eigenhands** más representativas.

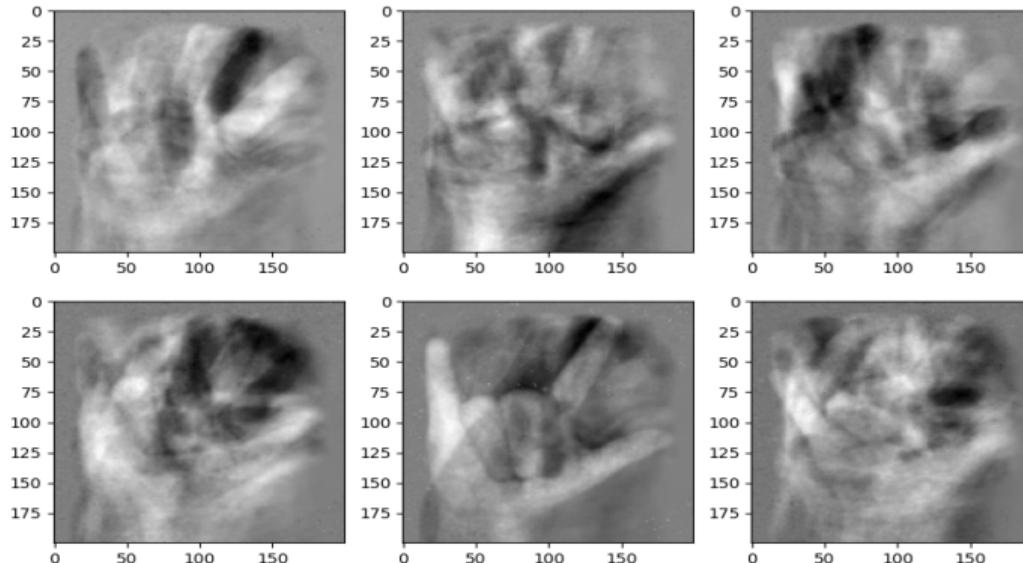


Figure: 6 Eigenhands más representativas.

Fase 1 - Hand space

2 Materiales y Métodos

Las eigenhands calculadas a partir de los vectores propios u_i abarcan un conjunto básico (de dimensión $M'xN^2$), al que llamamos "hand space".

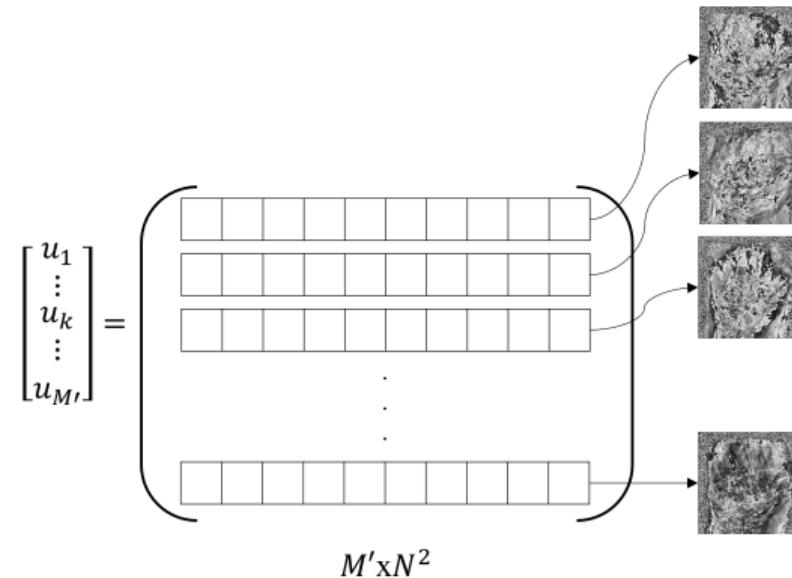


Figure: Hand space.

Fase 1- Patrones de pesos

2 Materiales y Métodos

Es necesario encontrar los patrones Ω_K que describen a la k-ésima clase de gestos.

$$\Omega^T = [\omega_1, \omega_2, \dots, \omega_{M'}] \quad (8)$$

Los pesos que forman al vector Ω se calculan mediante la siguiente operación

$$\omega_k = u_k^T (\Gamma - \Psi) \quad (9)$$

Las clases de los gestos se calculan promediando los patrones de pesos Ω de un pequeño número de imágenes de cada gesto.

Fase 1 - k-ésima clase

2 Materiales y Métodos

Patrones Ω_K que describen a la k-ésima clase (gesto).

$$\begin{aligned}\Omega_1 &= [\omega_1, \omega_2, \dots, \omega_{M'}] \\ \Omega_2 &= [\omega_1, \omega_2, \dots, \omega_{M'}] \\ &\vdots\end{aligned}\tag{10}$$

$$\Omega_n = [\omega_1, \omega_2, \dots, \omega_{M'}]$$

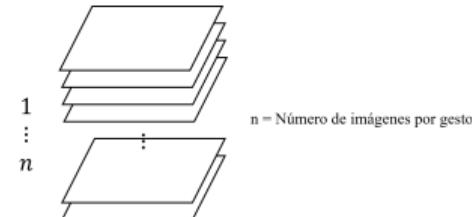


Figure: Imágenes por gesto.

$$\Omega_k = \frac{\Omega_1 + \Omega_2 + \dots + \Omega_n}{n} \tag{11}$$

Fase 2 - Ω para cada nueva image

2 Materiales y Métodos

Para determinar a qué clase pertenece una nueva imagen de entrada se calcula su patrón de pesos Ω ,

$$\Omega^T = [\omega_1, \omega_2, \dots, \omega_{M'}] \quad (12)$$

para luego calcular la distancia euclíadiana entre Ω y cada patrón de clase Ω_k

$$\epsilon_k = ||\Omega - \Omega_k||^2 \quad (13)$$

Fase 2 - Clasificar con la mínima ϵ_k

2 Materiales y Métodos

La imagen de entrada se clasifica como perteneciente a la clase que tenga una menor distancia euclíadiana ϵ_k

$$\epsilon_k = \|\Omega - \Omega_k\|^2 \tag{14}$$

Tabla de contenido

3 Resultados

- Introducción
- Materiales y Métodos
- Resultados
- Conclusiones
- References

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	1,765.23
C	8,965.39
Fist	3,463.38
Ok	8,155.25
Peace	6,431.8
I love you	7,620.04
L	7,851.43

Table: Distancias ϵ_k



Figure: Imagen de entrada 1

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	8,857.44
C	3,916.78
Fist	8,732.62
Ok	5,174.77
Peace	5,918.08
I love you	4,644.84
L	4,988.04

Table: Distancias ϵ_k .



Figure: Imagen de entrada 2

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	3,005.77
C	8,465.81
Fist	3,018.81
Ok	7,784.55
Peace	6,935.6
I love you	7,843.9
L	8,295.76

Table: Distancias ϵ_k .



Figure: Imagen de entrada 3

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	4,931.07
C	3,599.23
Fist	5,952.86
Ok	2,650.07
Peace	5,106.42
I love you	3,643.67
L	3,966.60

Table: Distancias ϵ_k .



Figure: Imagen de entrada 4

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	4,455.62
C	7,809.5
Fist	4,914.63
Ok	7,040.54
Peace	3,081.97
I love you	5,660.82
L	5,739.24

Table: Distancias ϵ_k .



Figure: Imagen de entrada 5

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	5, 658.64
C	4, 849.28
Fist	6, 366.77
Ok	4, 161.36
Peace	4, 155.54
I love you	1, 512.36
L	2, 969.92

Table: Distancias ϵ_k .



Figure: Imagen de entrada 6

Experimentos

3 Resultados

Gesto	ϵ_k
Palm	5,072.34
C	4,386.48
Fist	6,338.76
Ok	3,252.06
Peace	4,155.84
I love you	2,661.54
L	1,876.78

Table: Distancias ϵ_k .



Figure: Imagen de entrada 7

Experimentos - kaggle data-set

3 Resultados

Gesto	ϵ_k
Palm	5,442.51
<i>C</i>	2,950.36
Fist	5,314.97
Ok	3,610.58
Peace	4,634.85
<i>I love you</i>	4,88.71
L	4,688.54

Table: Distancias ϵ_k .



Figure: Imagen de entrada 8

Experimentos - kaggle data-set

3 Resultados

Gesto	ϵ_k
Palm	4,313.39
C	7,089.99
Fist	2,454.74
Ok	6,995.88
Peace	5,825.65
I love you	7,177.62
L	7,874.09

Table: Distancias ϵ_k .

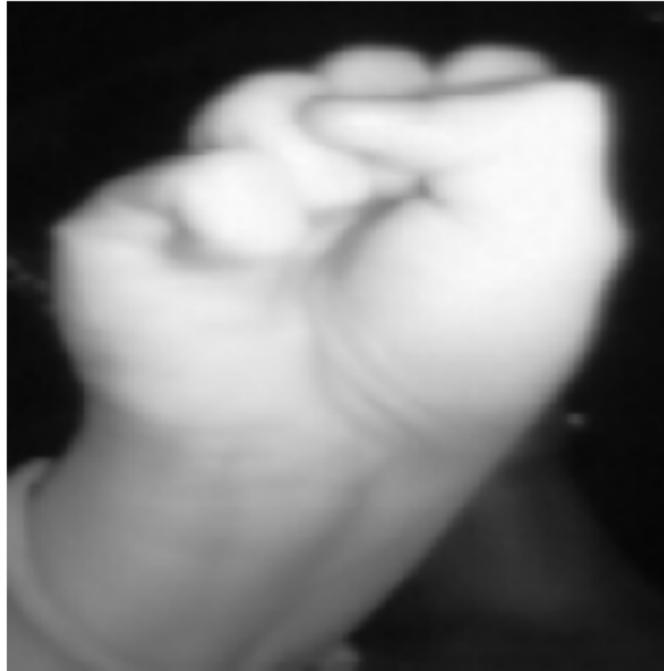


Figure: Imagen de entrada 9

Experimentos - kaggle data-set

3 Resultados

Gesto	ϵ_k
Palm	6,047.11
C	4,409.60
Fist	5,077.87
Ok	4,531.17
Peace	2,846.53
I love you	4,660.70
L	4,987.18

Table: Distancias ϵ_k .



Figure: Imagen de entrada 10

EXAMPLE DEMO

3 Resultados

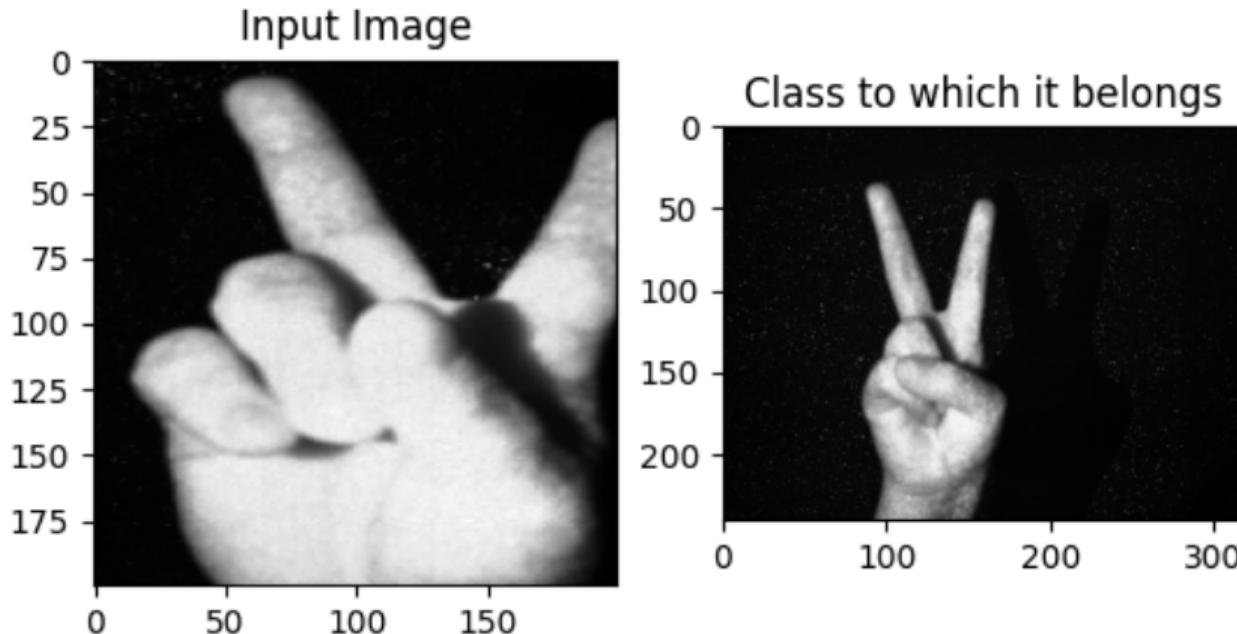


Figure: Prueba 1

Tabla de contenido

4 Conclusiones

► Introducción

► Materiales y Métodos

► Resultados

► Conclusiones

► References

Conclusiones

4 Conclusiones

- PCA limitado
- La importancia de la BD
- Camara/sensor
- Preprocesamiento
- Es dificil clasificar gestos similares
- Poco tiempo de computo

Agradecimientos

4 Conclusiones

- Dr. Keny Ordaz Hernández por facilitarnos el sensor.
- Dr. Mario Castelán por sus asesorías con respecto al PCA.

Tabla de contenido

5 References

- Introducción
- Materiales y Métodos
- Resultados
- Conclusiones
- References

References

5 References

- [1] Structure - the world's leading healthcare 3d scanning platform.
Consultado el 10 de abril de 2023.
- [2] Lawrence Sirovich and Michael Kirby.
Low-dimensional procedure for the characterization of human faces.
Josa a, 4(3):519–524, 1987.
- [3] Matthew Turk and Alex Pentland.
Eigenfaces for recognition.
Journal of cognitive neuroscience, 3(1):71–86, 1991.



Cinvestav
Unidad Saltillo

Q&A

Gracias por su atención!

*El trabajo puede ser consultado dentro del siguiente repositorio
de Github:*

Hand Gesture For Recognition