

Isabella Salgado 201730418
Johan Hernández 201729696
Juan Pablo Naranjo 201730006

Tarea 9

Punto 1

a

En primer lugar, se tomaron todos los datos para calcular la media de Procrustes, la cual se muestra en la figura 1.

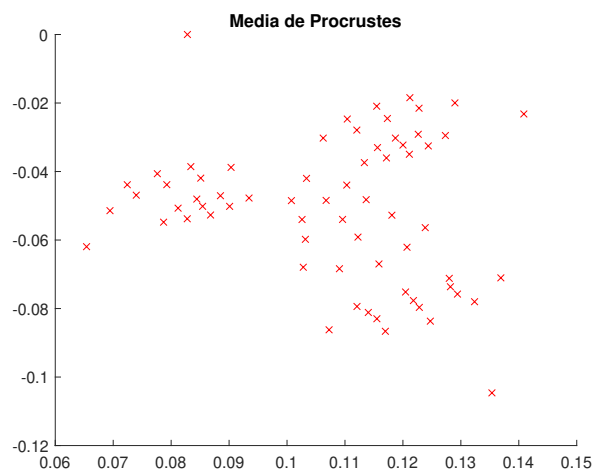


Figura 1: Media de procrustes con todos los datos

Ahora bien, se alinearon los landmarks a la media de Procrustes tal como se muestra a continuación:

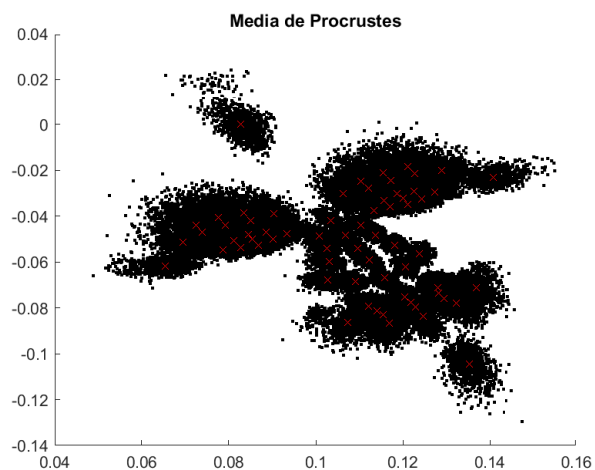


Figura 2: Landmarks alineados a la media de procrustes

b

Con los landmarks alineados, se creó un vector de características de tamaño 134 por cada expresión facial, donde la primera característica corresponde a las componente horizontal del landmark 1 y la segunda a la componente vertical del mismo landmark. Se continuó construyendo el vector intercalando

componente horizontal y vertical de cada landmark para la expresión facial en cuestión. Finalmente, se obtuvieron 974 vectores de 134 posiciones, que corresponden a 974 expresiones faciales.

c

Con todos los datos, se encontraron los vectores principales y se tomaron los tres primeros. Cada uno de ellos se graficó como si fuera una expresión facial, teniendo en cuenta que el primer número de cada vector correspondía a la componente horizontal, el segundo a la vertical, el tercero nuevamente a la horizontal y así sucesivamente. De igual manera, se realizó la gráfica utilizando los índices de los landmarks asociados a las diferentes características de la expresión facial: ojos, cejas, nariz, labios y contorno facial.

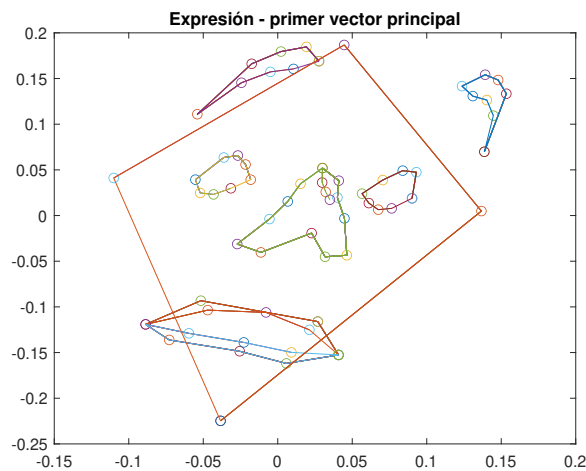


Figura 3: Expresión facial con el primer vector principal

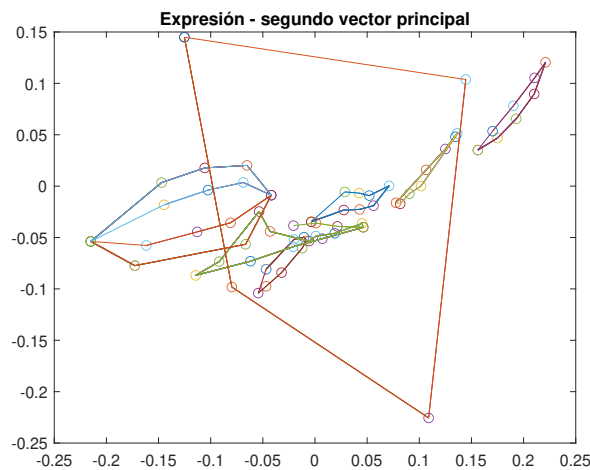


Figura 4: Expresión facial con el segundo vector principal

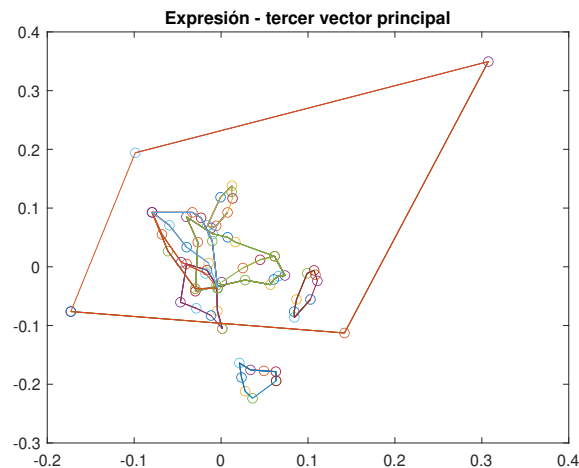


Figura 5: Expresión facial con el tercer vector principal

De las figuras anteriores, se visualiza una clara expresión facial en la figura 3 que podría corresponder a miedo. Por el contrario, en las figuras 4 y 5 se muestran las diferentes características faciales pero no están ubicadas en el espacio correcto por lo que no se muestra una expresión evidente. Sin embargo, en estas dos figuras el contorno facial si es evidente. Se puede decir que el primer vector principal representa la base que captura la máxima variabilidad de todas las expresiones faciales, mientras que los siguientes dos vectores (segundo y tercero) son los vectores que le siguen, es decir, las bases que capturan la segunda y tercera máxima variabilidad de todas las expresiones faciales.

d

En la siguiente figura se muestra una separación de los datos por colores indicando el tipo de expresión facial. Se puede evidenciar una clara separación o distinción de la expresión de miedo, la cual es muy característica. Por otro lado, los datos de disgusto y de furia se ven mezclados, lo cual es coherente teniendo en cuenta que este tipo de expresiones se suelen confundir. Así mismo, los datos de neutral se encuentran aproximadamente en el centro de los demás datos excepto los de miedo, lo que demuestra que las demás expresiones se podrían confundir con las de neutral. Esto anterior es cierto y se evidenció en la tarea 7.

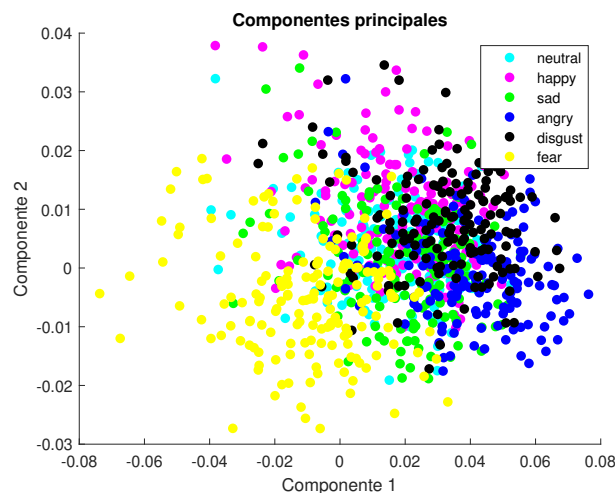


Figura 6: Análisis de componentes principales

Punto 2

a

La imagen original usada para este punto se puede apreciar en la figura 7.



Figura 7: Imagen Original

Esta imagen está representada como un arreglo de tres dimensiones de tamaño $600 \times 800 \times 3$, el primer paso de este trabajo consistió vectorizar este arreglo, con lo que, se obtuvo una matriz con el siguiente tamaño:

$$(600 \cdot 800) \times 3 = 480000 \times 3$$

donde cada columna indica una componente en el espacio RGB.

b

Luego de obtener la matriz de tamaño 480000×3 se procedió a graficarla en un espacio en 3D, donde cada eje corresponde a una de las componentes RGB. En la figura 8 se puede observar la imagen original, pero, en el espacio RGB vista desde tres posiciones diferentes.

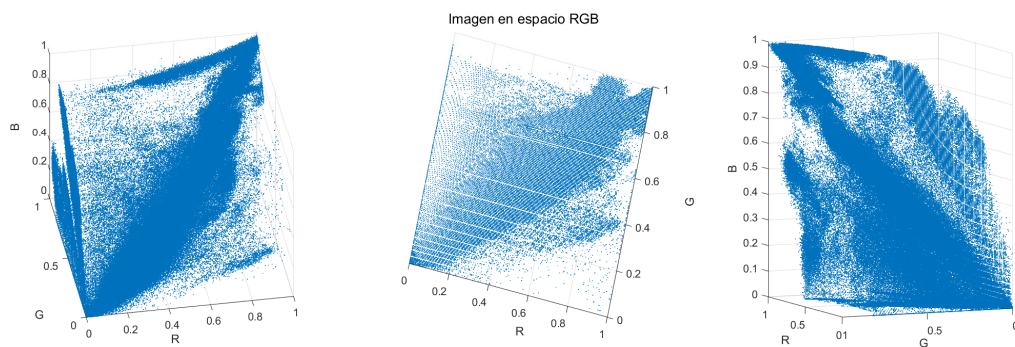


Figura 8: Imagen en Espacio RGB

C

Para la realización de este punto se empleó el algoritmo EM utilizado en el punto 1 de la tarea 7 del curso Análisis inteligente de señales y sistemas. En este desarrollo se utilizaron los siguientes parámetros:

- Número de Gaussianas: $K = 16$
- Valor para la regularización: $\lambda = 0.007$
- Valor de condición de parada - Cambio entre media actual y media anterior: $\epsilon = 0.001$
- Valor de condición de parada - Máximo número de iteraciones: $\text{cont} = 350$

En cuanto a las condiciones iniciales; el vector de medias se creó como un vector aleatorios con valores entre 0 y 1, las matrices de covarianza se inicializaron como la identidad y el vector de α se inicializó con un valor uniforme de $\frac{1}{16}$ para todas sus posiciones.

Empleando el algoritmo EM se obtuvieron como resultados las medias, matrices de covarianza y el vector de α , con esto se procedió a graficar la posición de las medias en el espacio RGB donde se ubica la imagen original, este gráfico se puede observar en la figura 9, donde se pueden apreciar tres perspectivas diferentes.

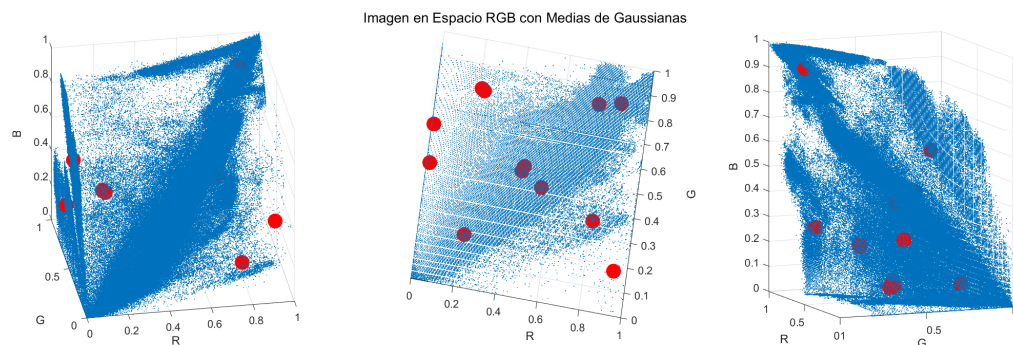


Figura 9: Imagen en Espacio RGB con Medias Gaussianas

Analizando la figura 9 se puede observar que no aparecen 16 medias Gaussianas, esto sucede porque hubo algunas medias Gaussianas que quedaron muy cerca entre ellas, lo que implica que no es posible apreciar todas las medias a simple vista. Esto ocurre porque este algoritmo es muy sensible a las condiciones iniciales, por lo que, si dos condiciones iniciales son muy cercanas entre ellas es probable que terminen encontrando el mismo máximo local y por ello terminen en la misma posición.

Dado que ya se tienen las medias y covarianzas de cada una de las 16 Gaussianas, el siguiente paso consistió en calcular la probabilidad de cada punto de la imagen original con estas 16 Gaussianas y, finalmente, asignarle la media con la que tuviera mayor probabilidad. En la figura 10 se pueden observar todos los puntos de la imagen original en un espacio RGB desde tres perspectivas diferentes, pero, con un color que corresponde a la media Gaussianas con la que está relacionado.

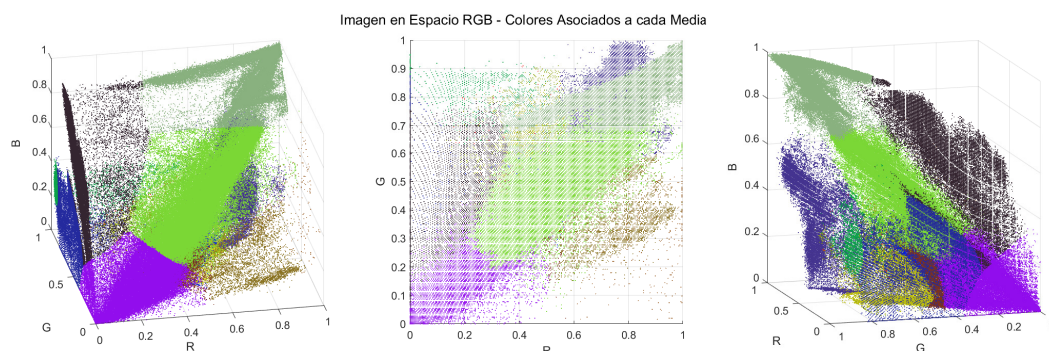


Figura 10: Imagen en Espacio RGB con Colores Asociados a cada Media Gaussianas

d

En este literal se reasignó cada punto de la imagen original a su media correspondiente, que se calculó en el literal inmediatamente anterior del presente trabajo, de esta forma, los 480000 puntos se redujeron a 16, pero solo en valor, es decir, se sigue teniendo una matriz de 480000×3 solo que ahora el valor de cada punto va a ser una de las medias obtenidas anteriormente. En la figura 11 se pueden observar todos los 480000 puntos reasignados. Como se mencionó anteriormente, no aparecen 16 puntos debido a que algunas medias Gaussianas quedaron muy cercanas entre ellas, por lo tanto, no es posible notarlas a simple vista.

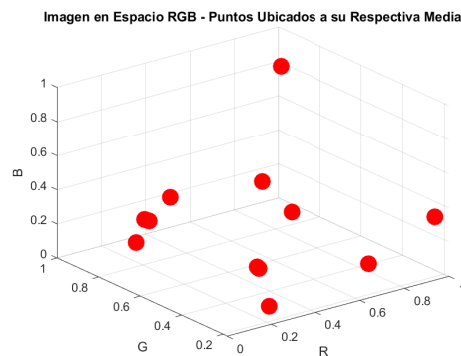


Figura 11: Imagen en Espacio RGB con Puntos Ubicados en su respectiva Media Gaussiana

e

Finalmente, se procedió a reorganizar la matriz de 480000×3 , presentada en el literal anterior, a un arreglo de $600 \times 800 \times 3$ con el fin de poder graficarlo como una imagen. En la figura 12 se puede observar tanto la imagen original como la imagen comprimida.

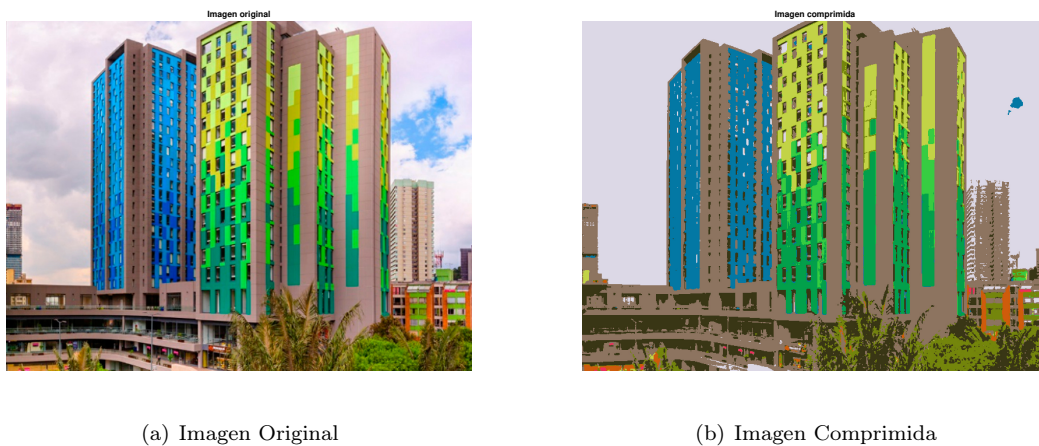


Figura 12: Imagen Original vs Imagen Comprimida

Analizando la figura 12 se puede notar que la imagen comprimida presenta una menor gama de colores y una menor definición, esto se puede notar en el cielo, en la imagen original se ven distintas tonalidades de grises, mientras que, en la imagen comprimida se observa un gris uniforme, esto sucede porque todos los colores que tenían esta tonalidad grisácea se encontraban cercanas en el espacio RGB, por lo que, al reasignarlas a su media Gaussiana correspondiente todas estas tonalidades se pierden y solo queda un único valor constante. Esto mismo se puede notar en el color azul del edificio, en la imagen original hay, al menos, dos tonalidades de azul diferentes, sin embargo, en la imagen comprimida ya solo aparece una única tonalidad de azul y, de hecho, se puede notar como esta tonalidad de azul no es exactamente igual a las tonalidades de azul presentes en la imagen original, sino, que parece como un punto intermedio entre las dos, lo que indica que realmente sí se encontró una media para esta tonalidad.

Finalmente, otro aspecto muy importante a señalar, es que el peso de las imágenes es el siguiente:

- Imagen original: 1652 KB
- Imagen comprimida: 145 KB

Esto es importante porque está demostrando que la imagen realmente sí se está comprimiendo, dado que reduce su peso en un 92 %.