

REVISIÓN DE LAS TÉCNICAS BÁSICAS PARA EL RECONOCIMIENTO DE ROSTROS

JHONNY LEANDRO RESTREPO VARGAS
COD. 1093221265

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
FACULTAD DE TECNOLOGÍA
PROGRAMA DE INGENIERÍA MECATRONICA
PEREIRA - RISARALDA

REVISIÓN DE LAS TÉCNICAS BÁSICAS PARA EL RECONOCIMIENTO
DE ROSTROS

JHONNY LEANDRO RESTREPO VARGAS
COD. 1093221265

DIRECTOR
PhD(c). JIMMY ALEXANDER CORTÉS OSORIO

INGENIERÍA MECATRONICA CICLOS PROPEDÉUTICOS
CICLO DE TECNOLOGÍA

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
FACULTAD DE TECNOLOGÍA
PROGRAMA DE INGENIERÍA MECATRONICA
PEREIRA - RISARALDA

Dedicado a

A mi familia por la confianza depositada en este proceso.

Agradecimientos

Las metas alcanzadas en cada etapa de la vida, son el resultado del esfuerzo propio y la ayuda incondicional de quienes creen en los proyectos ajenos, por ello extendo mi gratitud a mi director de proyecto Jimmy Alexander Cortés Osorio, quién siempre ha creído en mí, brindándome conocimiento y sabiduría. A mi familia, compañeros y amigos, por su presencia y aportes en cada triunfo y desfallecimiento.

Avanzar es un proceso a través del cual nos encontramos como seres capaces y al mismo tiempo, la vida y Dios nos permiten disfrutar de personas valiosas. Gracias por estar siempre allí.

Introducción

Los sistemas de reconocimiento de identidad de una persona han evolucionado considerablemente en los últimos años, en la actualidad la biometría informática es la aplicación de las técnicas matemáticas y estadísticas sobre las diferencias físicas de un individuo, que son únicos e irrepetibles.

Después de saber los beneficios que tiene biometría informática, se empezaron a desarrollar métodos con los cuales se aprovechan dichas diferencias para la elaboración de métodos de reconocimiento en sistemas de seguridad (en empresas, aeropuertos, entidades carcelarias y entidades bancarias, etc.).

En este proyecto se dará a conocer de forma general las diferentes técnicas de la biometría, enfatizado en el reconocimiento de personas por medio de sus patrones faciales. Se hablara sobre las técnicas más utilizadas en el reconocimiento de rostros con la finalidad de encontrar una técnica de reconocimiento facial viable.

Índice general

Agradecimientos.....	5
Introducción.....	7
1. Introducción al reconocimiento e identificación de personas	1
1.1. Origen de la biometría.....	15
1.2. Métodos biométricos.....	16
1.2.1. Biometría de la mano.....	17
1.2.2. Biometría del iris.....	18
1.2.3. Biometría de la voz.....	19
1.2.4. Biometría de la retina	20
1.2.5. Biometría de la huella dactilar.....	21
1.2.6. Biometría del rostro.....	21
2. Biometría del rostro	23
2.1. Métodos del reconocimiento biométrico del rostro.....	23
2.1.1. Principal Component Analysis (PCA).....	23
2.1.2. Pasos para el reconocimiento de rostros por Eigenfaces.....	26
2.1.3. Pasos para el reconocimiento de rostros por Fisherfaces.....	27
2.2.1. Elastic Bunch Graph Matching (EBGM)	29
2.2.2. Modelado y ubicación de los puntos de referencia facial.....	30
2.2.3. Histogramas de Gradientes orientados (HOG) descriptor.....	32
3. Diseño	34
3.1 Captura.....	34
3.1.1 Cámaras Web digitales.....	34
3.2 Base de datos	36
3.2.1 The Color FERET Database, USA	36
3.2.2 The Yale Face Database.....	36
3.2.3 The ORL Database of Faces.....	36
3.3 Software.....	37

3.3.1	Matlab.....	37
3.3.2	Microsoft Visual Studio.....	37
3.4	Revisión y propuesta de las técnicas de reconocimiento.....	38
3.5	Hardware.....	38
3.5.1	PcDuino.....	38
3.5.2	Raspberry PI.....	39
4	Resultados y conclusiones	40
4.1.	Resultados	40
4.2.	Conclusiones.....	40

Índice de cuadros

1.1	Conjunto de datos de dos dimensiones y la media respectiva [Los autores].	24
1.2	Cálculo de la varianza [Los autores]	25
1.3	Tabla comparativa de cámaras digitales web [Los autores]	35
1.4	Tabla comparativa entre los modelos de Raspberry Pi [Los autores]	39

Índice de figuras

1.	Huella dactilar [Los autores]	16
2.	Impresión de la palma [Los autores].....	16
3.	Partes externas del ojo [Los autores].....	17
4.	Niveles de frecuencia en la voz [Los autores]	18
5.	Ubicación de la retina [Los autores].....	19
6.	Venas de la retina [4]	19
7.	Partes de la huella dactilar [Los autores]	20
8.	Huella dactilar [Los autores]	20
9.	Identificación biométrica basada en los rasgos faciales [Los autores].....	21
10.	Grafo frontal del rostro con 25 puntos de referencia [Los autores].....	30
11.	Enrejados espaciales de un descriptor HOG [Los autores].....	32
12.	PcDuino [Los autores].....	39

Capítulo 1

Introducción al reconocimiento e identificación de personas

En este proceso investigativo se realizará una revisión de los métodos más eficientes y eficaces para implementar un sistema de reconocimiento facial biométrico. Para tal motivo, se identificará el origen de la biometría y los métodos de reconocimiento biométrico.

1.1 Origen de la biometría

Los sistemas de reconocimiento e identificación de personas por medio de una computadora forman parte del área denominada biometría. La palabra biometría se deriva del griego (Bio y metria) Bio (vida) Metria (medida) la cual estudia métodos de reconocimiento únicos, es decir características físicas en las personas que serían propias de ellas e intransferibles, estas características son utilizadas para identificar a las personas.

Los principios básicos de la biometría para identificar y verificar a las personas, fueron empleados por primera vez a mediados del siglo XIV, en china por los comerciantes que estampaban las huellas de la palma de las manos en papel con tinta como método de identificación. Luego en el siglo XIX, investigadores en criminología intentaron relacionar características físicas con tendencias criminales, el sistema funcionaba con medidas precisas de ciertas longitudes y anchos de la cabeza y del cuerpo, así como registros de marcas individuales como tatuajes y cicatrices. Los resultados no eran concluyentes por los cambios en los individuos que afectaban las medidas, pero la idea de medir las características físicas de un individuo parecía efectiva al igual la identificación de huellas digitales que años después se convirtió en la metodología internacional de identificación utilizada por las fuerzas policiales de todo el mundo. Muy similar al sistema que utilizaban los chinos años atrás. [1]

En el siglo XX, las metodologías biométricas como la verificación de huellas digitales eran constantemente mejoradas y refinadas al punto de convertirse en métodos confiables. En la actualidad la biometría no sólo se centra en la identificación por medio de la huella dactilar sino que ya se emplean diferentes técnicas de reconocimiento teniendo en cuenta un rasgo o factor físico único de la persona como por ejemplo, la cara, palma, huella digital, retina iris, venas, voz y demás técnicas.

1.2 Métodos biométricos

1.2.1 Biometría de la mano

El sistema de reconocimiento de la mano es uno de los sistemas más utilizados ya que es muy similar al reconocimiento dactilar. En este reconocimiento biométrico de la mano se puede realizar por dos o tres dimensiones, los sistemas de dos dimensiones tiene características relevantes como las líneas principales de la mano y las arrugas, estas son casi tan distintivas como las huellas dactilares. [2]

Para obtener los datos biométricos es necesario tener una buena definición de la imagen para obtener las terminaciones y las bifurcaciones existentes ya que representan un 75% de las minucias, es decir, patrones biométricos los cuales ayudaran a un mejor reconocimiento. [Véase en la figura 1,2]



Figura 1. Huella dactilar
Fuente. Autor

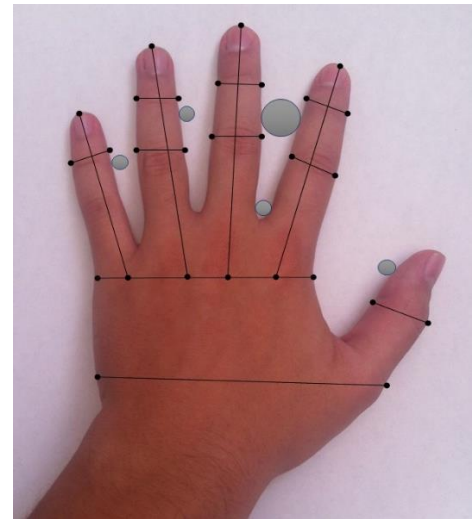


Figura 2. Impresión de la palma
Fuente. Autor.

Los sistemas de tres dimensiones son método geométrico que toma medidas como la longitud de los dedos y la separación entre ellos así como las distancias entre puntos claves de la mano, haciendo de este un sistema más preciso.[3]

Estas características se transforman en una serie de patrones numéricos que pueden ser comparados. Su principal aplicación es la verificación del usuario con una probabilidad de error aceptable en la mayoría de los casos.

1.2.2 Biometría del iris

La identificación basada en el reconocimiento del iris humano se viene utilizando hace unos años para la autenticación de usuarios, se debe tener en cuenta que el iris no cambia a lo largo de la vida del individuo de esta manera se obtiene un gran porcentaje de confiabilidad y exactitud a la hora de reconocer o identificar a una persona. [4]

El iris se encuentra situado en la parte interna del ojo detrás de la córnea y el líquido acuoso, su principal característica es la textura lo suficientemente compleja y distintiva para cada persona y cada ojo como para aplicar un método biométrico de reconocimiento. [Véase en la figura 3]

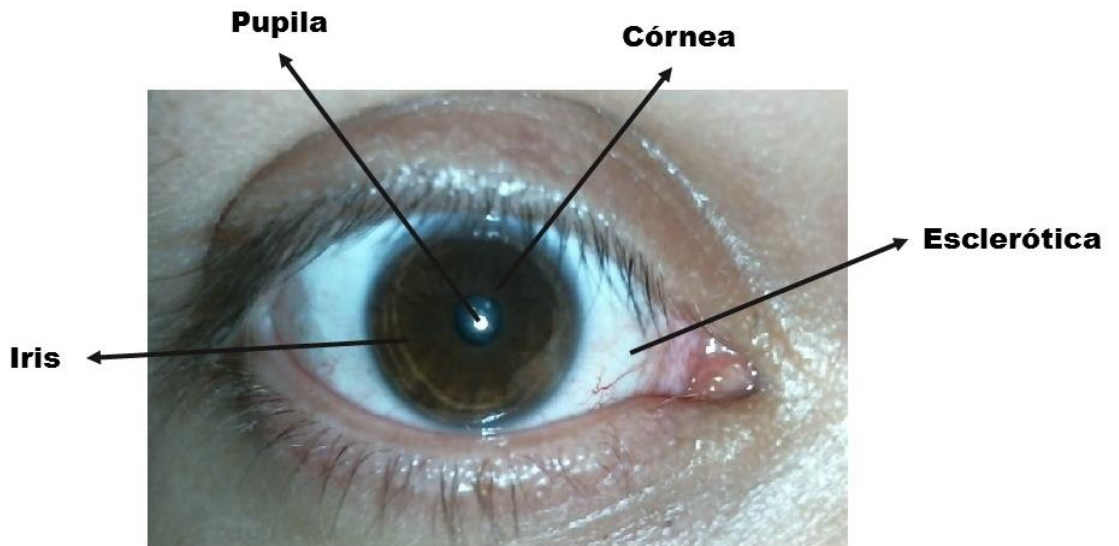


Figura 3. Partes externas del ojo
Fuente. Autor

El proceso básicamente consta de adquirir la imagen y realizar un tratamiento relativamente sencillo para eliminar ruidos, iluminación y otros defectos que presente la imagen capturada. [2]

Este identificador tiene algunos factores que han afectado su eficaz desarrollo ya que se requiere cooperación por parte del usuario para realizar la captura de la imagen.

1.2.3 Biometría de la voz

La biometría de la voz es un sistema de última generación con otras características que las personas utilizan comúnmente para identificar a los demás. Existen personas que poseen voz similares a otras, pero nunca van a ser iguales si se evalúa la señal que genera la voz.

La idea principal de este sistema de reconocimiento, no es intentar reconocer lo que el usuario dice, sino identificar una serie de sonidos emitidos como los tonos bajos y agudos, vibración de la laringe, tonos nasales y de la garganta, estos factores pueden verse afectados por patrones como enfermedades, cambios de estados de ánimo y de edad que representa un factor de cambio con el tiempo. [5] [Véase en la figura 4]

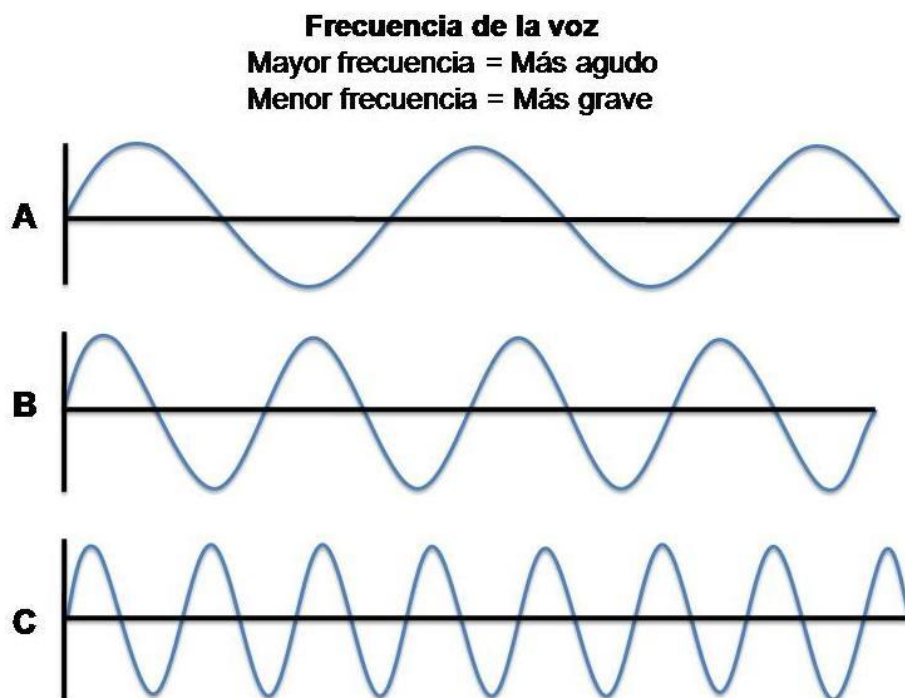


Figura 4. Niveles de frecuencia en la voz
Fuente. Autor

Sin embargo se sigue usando por su efectividad cuando se trata de reconocimiento no presencial como lo es la telefonía.

1.2.4 Biometría de la retina

La retina posee características únicas en cada individuo que son utilizadas para la identificación satisfactoria de personas. Por lo tanto, es un método biométrico confiable. Se encuentra ubicado en la capa más interna del ojo donde realmente se lleva a cabo la visión, forma parte del sistema nervioso central y está conectada directamente al cerebro a través del nervio óptico. [Véase en la figura 5]

El reconocimiento biométrico de la retina se basa en la identificación del patrón creado por los vasos sanguíneos ubicados debajo de la superficie de la retina. Esta información es altamente distintiva ya que no existen dos patrones iguales, ni siquiera en hermanos gemelos. [Véase en la figura 6]

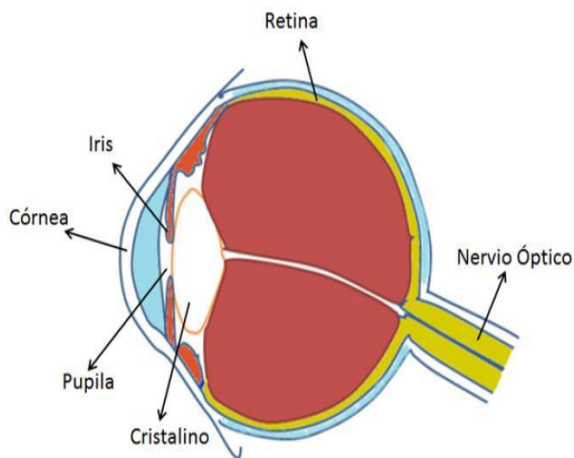


Figura 5. Ubicación de la retina
Fuente. Autor

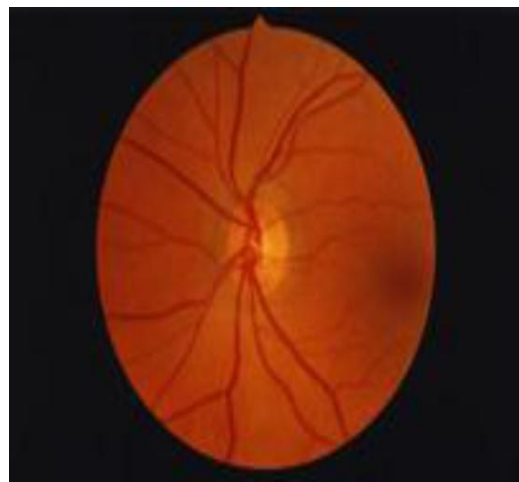


Figura 6. Venas de la retina
Fuente. Ver Referencia 4

En este método se escanea la retina emitiendo una luz de baja intensidad (puede ser infrarroja) para desvelar los nodos y ramas del área retina y así poder capturar la imagen para el reconocimiento biométrico. El escaneo de retina puede ser sumamente preciso pero requiere que el usuario mire fijamente y muy cerca un dispositivo y mantenga su vista fija en un punto determinado.

A pesar de su confiabilidad como sistema de reconocimiento en muchos casos no se puede hacer que los sujetos a identificar cumplan con el procedimiento, ya que

el haz de luz que se utiliza para tomar la foto es demasiado incomodo, además requiere de un equipo bastante costoso. [4]

1.2.5 Biometría de la huella dactilar

La identificación por medio de la huella dactilar es la técnica más utilizada a lo largo de la historia. Pero sólo hasta los años 70 se pudo establecer un sistema estable y confiable, sin embargo la fiabilidad del reconocimiento depende de la calidad de imagen si se tiene una imagen de baja calidad se debe trabajar con algoritmos más complejos para el reconocimiento biométrico.

Este sistema contiene un patrón único de canales, crestas y figuras que se forma en la yema del dedo de cada individuo a partir de la sexta semana de vida intrauterina y no varían sus características a lo largo de los años. La única forma de que esta característica se vea afacetada es por una causa radical (herida, quemadura, etc.). En la yema de los dedos o se esté sufriendo de algún tipo patología de la piel. [4]

La identificación con huellas dactilares está basada principalmente en las minucias (la ubicación y la terminación de las crestas y valles, bifurcaciones, delta, etc.), las cuales podemos identificar. [6] [Véase en la figura 7]



Figura 7. Partes de la huella dactilar
Fuente. Autor



Figura 8. Huella dactilar
Fuente. Autor

Lo poco invasivo este sistema lo ha convertido en un método altamente aceptado por el público a la hora de adquirir los datos de la imagen para la base de datos. Gracias a esta ventaja se convierte en uno de los métodos con más aplicaciones en el mundo real.

1.2.6 Biometría del rostro

Este sistema de reconocimiento facial es la forma más común hoy en día, generalmente se basa en la captura de patrones geométricos del rostro, ubicando puntos claves de la cara los cuales se almacenan y se comparan en una base de datos para determinar la identidad de la persona. [4]



Figura 9. Identificación biométrica basada en los rasgos faciales
Fuente. Autor

La identificación se puede realizar de diferentes formas por medio de métodos locales, holísticos y mixtos:

- Los rasgos locales son aquellos en los que se tiene como prioridad puntos clave como la de los ojos, cejas, nariz y boca para medir distancia entre ellas y lograr así un reconocimiento.
- Los rasgos holísticos son aquellos en los que se toma información de toda la cara para el reconocimiento.

- Los rasgos mixtos toman los dos métodos anteriores en uno solo para tener un reconocimiento más exacto.

Los sistemas actuales comúnmente imponen restricciones en como las imágenes faciales debe ser obtenidas generalmente con un fondo simple o una iluminación especial. El proceso generalmente es realizado de manera que una vez generada la imagen de la cara se procede a realizar una comparación con las imágenes del sistema o de entrenamiento. [2]

Capítulo 2

Biometría del rostro

Una gran variedad de métodos de reconocimiento de rostros se han propuesto durante los últimos años, a continuación se describen las técnicas más empleadas a la hora de desarrollar un sistema biométrico facial.

2.1 Métodos de reconocimiento biométrico del rostro

2.1.1 Principal Component Analysis (PCA)

El análisis de componentes principales es una de las técnicas más utilizada en el reconocimiento de rostros, es un método de análisis estadístico multivariante que se clasifica entre los métodos de simplificación o reducción de dimensión de variables. El PCA también permite minimizar los datos de manera que se reduzca el número de dimensiones sin tener pérdida de información.

Se debe tener en cuenta los siguientes procesos matemáticos para la obtención del PCA.

- Paso 1: Obtener un conjunto de N datos inventados de sólo dos dimensiones X y Y cómo se puede visualizar en el cuadro 1 para analizar cada paso del PCA.
- Paso 2: Calcular la media del conjunto de N datos, se debe tener en cuenta que la media no dice mucho de los datos solo muestra el punto medio de los datos analizados, la ecuación es la siguiente.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

X	Y	(X - \bar{X})	(Y - \bar{Y})
3.1	3.9	0.9	1.5
2.4	2.2	0.2	-0.2
0.3	1.5	-1.9	-0.9
1.1	0.6	-1.1	-1.8
4.5	3.2	2.3	0.8
3.2	2.6	1.0	0.2
2.4	4.1	0.2	1.7
3.1	2.4	0.9	0.0
2.7	2.1	0.5	-0.3
0.2	0.5	-2.0	-1.9
0.8	2.3	-1.4	-0.1
2.6	3.4	0.4	1.0
$\Sigma =$	26.4	28.8	0.0
Media =	2.2	2.4	

Cuadro 1. Conjunto de datos de dos dimensiones y la media respectiva.

- Paso 3: Calcular la varianza de un conjunto con N datos, la varianza es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística, esta medida nos permite identificar la diferencia promedio que hay entre cada uno de los valores respecto a su punto central. Su fórmula es la siguiente.

$$Var(X) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n - 1)}$$

- Paso 4: Halla la covarianza que es una medida de dispersión que permite analizar si existe relación lineal entre dos o más variables estadísticas, la covarianza está dada de la siguiente forma.

$$Cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)}$$

	$(X - \bar{X}) * (y - \bar{y})$	$(X - \bar{X}) * (X - \bar{X})$	$(y - \bar{y}) * (y - \bar{y})$
	1.35	0.81	2.25
	-0.04	0.04	0.04
	1.71	3.61	0.81
	1.98	1.21	3.24
	1.84	5.29	0.64
	0.20	1.00	0.04
	0.34	0.04	2.89
	0.00	0.81	0.00
	-0.15	0.25	0.09
	3.80	4.00	3.61
	0.14	1.96	0.01
	0.40	0.16	1.00
Σ =	11.57	19.18	14.62
Cov =	1.05	1.74	1.33

Cuadro 2. Cálculo de la varianza.

- Paso 5: Hallar la matriz de covarianza o matriz de dispersión teniendo en cuenta los datos de la ecuación anterior.

$$Cov = \begin{Bmatrix} Cov(x, x) & Cov(x, y) \\ Cov(y, x) & Cov(y, y) \end{Bmatrix}$$

La matriz de covarianza con los datos implementados es:

$$Cov = \begin{bmatrix} 1.74 & 1.05 \\ 1.05 & 1.33 \end{bmatrix}$$

También se pueden encontrar una matriz de covarianza de n x n dimensiones la cual podemos representar de la siguiente manera.

$$\Sigma = \begin{Bmatrix} Cov(1,1) & Cov(1,2) & \dots & Cov(1,n) \\ Cov(2,1) & Cov(2,2) & \dots & Cov(2,n) \\ \dots & \dots & \dots & \dots \\ Cov(n,1) & Cov(n,2) & \dots & Cov(n,n) \end{Bmatrix}$$

- Paso 6: calcular los vectores propios y valores propios de la matriz de covarianza, como la matriz de covarianza es cuadrada se puede realizar este paso.

$$Eigenvalues = \begin{bmatrix} 0,4652 \\ 2,6048 \end{bmatrix}$$

$$\text{Eigenvectors} = \begin{bmatrix} 0,6358 & -0,7719 \\ -0,7719 & -0,6358 \end{bmatrix}$$

2.1.2 Pasos para el reconocimiento de rostros por Eigenfaces

Teniendo en cuenta los procesos matemáticos anteriormente mencionados procedemos a realizar los pasos pertinentes para el desarrollo del reconocimiento de rostros por medio del Eigenfaces, siguiendo los pasos que se señalan a continuación.

Este método es el principio de Análisis de componentes principales (PCA). Se debe tener una base de datos con N imágenes de entrenamiento, cada imagen con la misma n-dimensión. Este método se basa en realizar una proyección lineal que maximiza la dispersión de todas las imágenes del conjunto, se considera que para procesarlas se pone la imagen completa en forma de vector. Se asume que cada una de las imágenes pertenece a una de las c clases.

$$X = \{x_1, x_2, \dots, x_c\} \quad \text{Con } x_i \in \mathbb{R}^n$$

De este modo se considera una transformación lineal que lleva el espacio de imágenes original de n-dimensiones al espacio de características de dimensión m, donde $m < n$. los nuevos vectores de características $y_k \in \mathbb{R}^m$ son definidos por la siguiente transformación lineal.

$$y_k = W^T \cdot x_k \quad k = 1, 2, \dots, N \quad (1)$$

Donde $W \in \mathbb{R}^{n \times m}$ es una matriz con columnas ortonormales. Se define además la matriz de dispersión total de ST se define de la siguiente manera.

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T$$

En la anterior ecuación podemos observar que n es el número de imágenes de entrenamiento y $\mu \in \mathbb{R}^n$ es la imagen media de todas las muestras, después de aplicar la transformación lineal W^T la distribución de los vectores de característica $\{y_1, y_2, \dots, y_N\}$ es $W^T S_T W$. En PCA, la proyección W_{opt} se elige para maximizar la varianza de la matriz de dispersión total de las muestras proyectadas, es decir la solución a este problema es una matriz compuesta por los vectores propios en orden descendente.

$$W_{opt} = \arg \underbrace{\max}_W |W^T S_T W|$$

$$= [w_1, w_2, \dots, w_m]$$

Donde $\{w_i | i = 1, 2, \dots, m\}$ es el conjunto de vectores propios de n-dimensionales de S_T correspondiente a los m valores propios más grandes. Para reducir la dimensionalidad de los vectores propios asociados a los menores valores propios (en valor absoluto), pues contienen la menor cantidad de información. La cantidad a descartar no es fija, sino depende de las características de las imágenes que se están utilizando: en *fisherfaces* se define un máximo como se verá a continuación. [7]

2.1.3 Pasos para el reconocimiento de rostros por Fisherfaces

El método de *Fisherfaces* es un análisis de discriminantes lineales (LDA), particularmente se busca aprovechar la información, normalmente disponible sobre la clasificación de las imágenes de entrenamiento, para buscar una proyección que maximice la separación entre imágenes de diferentes personas (o clases) y minimice la distancia entre imágenes de una misma clase, así logra concentrar las imágenes mejorando, en forma importante la tasa de reconocimiento.

Este método selecciona el W en la ecuación (1) de tal manera que el cociente entre la distribución entre clases (imágenes de la misma persona) y la distribución intra-clases sea la máxima. Para esto se define la matriz S_B de distribución entre clases como:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T$$

Y la matriz S_W de distribución intra-clases

$$S_W = \sum_{i=1}^c \sum_{x_k \in X_i} N_i (\mu_k - \mu_i) (x_k - \mu_i)^T$$

En la anterior ecuación se muestra que μ_i es la imagen media de la clase X_i y N_i es el número de muestras de imágenes en la clase X_i . Si la matriz S_W es no singular, la proyección W_{opt} se elige como la matriz con columnas ortonormales que maximiza la relación entre el determinante de la matriz de dispersión entre la clase de las imágenes proyectadas y el determinante de la matriz de dispersión intra-clases de las imágenes proyectadas, esto es

$$W_{opt} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} \quad (2)$$

$$= [w_1, w_2, \dots, w_n]$$

Donde $\{w_i | i = 1, 2, \dots, m\}$ es el conjunto de vectores propios generalizados de S_B y S_W correspondientes a los m mayores valores propios generalizados en $\{\lambda_i | i = 1, 2, \dots, m\}$ es decir

$$S_B w_i = \lambda_i S_W w_i \quad i = 1, 2, \dots, m.$$

Se observa que en la mayoría de $c - 1$ valores propios distintos de cero, por lo tanto el limite superior en m es $c - 1$, donde c es el numero de clases vease en la ecuacion (2).

El problema de reconocimiento de rostros, se tiene que la matriz $S_W \in \mathbb{R}_{n \times m}$ es siempre singular dado que el rango de S_W es a lo sumo $N - c$, y en general, el numero de imágenes de entrenamiento N es mucho menor que el numero de pixiles en cada imagen n . Por lo tanto esto significa que es posible seleccionar una matriz W de tal manera que la distribucion intra-clase de las imágenes proyectadas pueda ser exactamente cero.

Como alternativa a la complicacion de S_W singular, el metodo de *Fisherfaces* evita este problema mediante la proyeccion del conjunto de imágenes a un espacio dimensional inferior de manera que la matriz resultante intra-clases S_W es no singular.

Utilizando PCA se realiza la reduccion de dimensiones del espacio de características de $N - c$ y luego aplicar FLD estandar definido por (2) para reducir la dimension a $c - 1$. De esta manera W_{opt} esta dada por

$$W_{pca}^T = W_{fld}^T W_{pca}^T$$

Donde

$$W_{pca} = \arg \max_W |W^T S_T W|$$

$$W_{fda} = \arg \max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$

Se debe de tener en cuenta que la optimización de W_{pca} se realiza sobre $n \times (N - c)$ matrices son columnas ortonormales, mientras la optimización para W_{fld} se realiza sobre $(N - c) \times m$ matrices con columnas ortonormales. [7]

2.2.1 Elastic Bunch Graph Matching (EBGM)

El algoritmo de EBGM fue diseñado para realizar tareas de reconocimiento facial, utilizando solamente algunos puntos de referencia faciales como los ojos, la boca, la punta de la nariz, etc. [10]

HOG-EBGM es un enfoque basado en EBGM en la que normalmente los Chorros de Gabor son usados como características para la localización y la relación de los puntos de referencia, para superar parte de la varianza generada por cambios en la expresión y la iluminación se sustituyen por Histograma de los Gradiente Orientados (HOG); como descriptor también se ha demostrado ser robusto para la ubicación de la gráfica y para tareas de reconocimiento.

HOG-EBGM se puede descomponer en tres pasos: Normalización de la imagen, la creación de los gráficos y finalmente la adecuación gráfica.

El primer paso es la normalización en las imágenes de la cara, reduce la variabilidad producida por los cambios en la iluminación, escala y rotación, se necesitan las coordenadas de los ojos para el proceso de normalización de las imágenes de la cara.

El siguiente paso crea un gráfico de cara después de la detección de puntos de referencia faciales, tal como se describe a continuación, el éxito del algoritmo del reconocimiento depende de una buena selección de puntos de referencia faciales: los puntos de referencia faciales deben ser muy distintivos entre diferentes personas y fácil de detectar en un sistema totalmente automático.

En HOG-EBGM la gráfica sigue la estructura de la cara propuesta con 25 puntos de referencia faciales. [Véase en la figura 10]

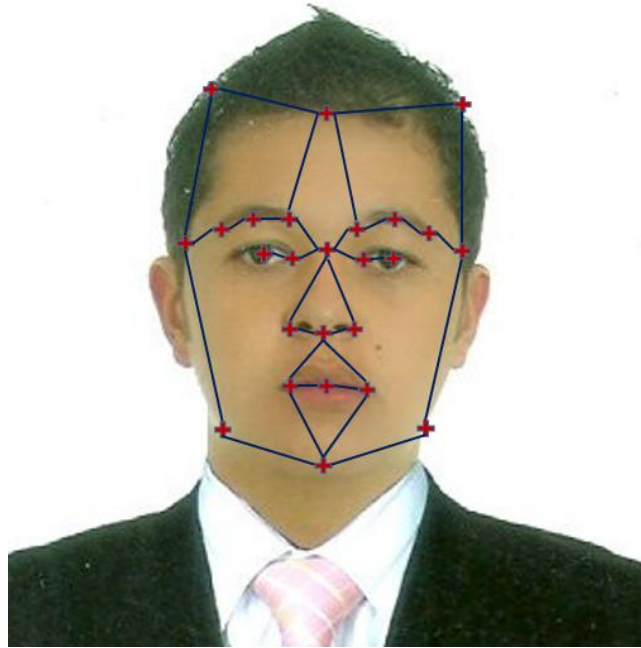


Figura 10. Grafo frontal del rostro con 25 puntos de referencia
Fuente. Autor

Por último, vamos a notar que no todas las zonas de la cara contribuyen por igual al reconocimiento facial. Varias estudios [8] han demostrado que el área alrededor de los ojos y la nariz son muy importantes para el reconocimiento, y así más puntos de referencia se colocan allí.

2.2.2 Modelado y ubicación de los puntos de referencia faciales

En HOG-EBGM cada cara es descrita por una gráfica (FG), integrado por puntos de referencia faciales ubicados estratégicamente y sus correspondientes descriptores:

$$FG = \{X_i \mid J_i = HOG(X_i), 1 \leq i \leq 25\}$$

Cuando creamos un nuevo FG, necesitamos un conjunto de modelos para localizar automáticamente puntos de referencia faciales, estos conjuntos de modelos se deben dar cuenta de los cambios de expresión, tipos de cabellos, iluminación, etc. Construimos un modelo diferente para cada señal mediante un conjunto de descriptores HOG extraídos de un conjunto de imagen de entrenamiento N_f . Deje $fbg_i(k), 1 < i < 25, 1 < k < N_f$ sea HOG el descriptor de i -th punto significativo de la k -th del entrenamiento de la cara (Cabe recordar que cada FG contiene 25 puntos de referencia.) definimos gráfico de la cara como (FBG):

$$FBG = \{fbg_i(k), 1 \leq i \leq 25, 1 < k < N_f\}$$

La ubicación de los descriptores $fbg_i(k)$ en una imagen de entrenamiento está marcada manualmente. Sin embargo, los puntos de referencia facial de imágenes de prueba se encuentran automáticamente. Esto se hace en un proceso interactivo en el que un nuevo punto de referencia utiliza la información de las señales detectadas anteriormente para reducir el área de búsqueda. El proceso para detectar la $i - th$ ($i > 2$) del punto de interés facial es: [10]

- La estimación inicial de la ubicación de los puntos de referencia, X_i^s basado en la medida de los desplazamientos entre los puntos significativos $i - th$ y los puntos significativos $j - th$ ($j < i$), utilizando información de la geometría de los gráficos FBG.
- Calcular el descriptor HOG en la ubicación anterior, HOG (X_i^s).
- Comparar HOG (X_i^s) con todos los descriptores HOG de los modelos fbg_i ,

$$k_{min} = \frac{min}{k} \|HOG(X_i^s) - fbg_i(k)\|$$

- Definir un área como S_i alrededor de X_i^s . La extensión del área de búsqueda depende del punto significativo particular como se muestra en la Figura 10. Nosotros establecimos empíricamente las áreas de búsquedas teniendo en cuenta la dispersión de la ubicación de los puntos faciales de cada $fbg_i(k)$, $0 \leq k \leq N_f$ de la FBG.
- Precisar la estimación inicial del punto significativo $i - th$ utilizando el descriptor $fbg_i(k_{min})$:

$$X_i = \min_{X \in S_i} \|HOG(X) - fbg_i(k_{min})\|$$

En este trabajo utilizamos la distancia euclidiana para comparar descriptores de HOG. También para comparar dos graficas faciales FG^k and FG^l de dos caras diferentes, solo resumimos las distancias entre dos puntos significativos:

$$D = |FG^k, FG^l| = \sum_{i=1}^{25} \|J_i^k - J_i^l\|$$

2.2.3 Histogramas de Gradientes orientados (HOG) descriptor

Los descriptores HOG utilizan la información de los gradientes locales para generar un histograma en el que los componentes forman una red tridimensional con $N_p = 4$ componentes para cada dirección espacial y $N_0 = 8$ componentes para la orientación con un total de $N_p^2 N_0 = 128$ componentes. En nuestro trabajo, cada componente espacial es un cuadrado de 5 x 5 píxeles. Este tamaño se elige de acuerdo a la distancia entre los ojos de las caras normalizadas, que en nuestro trabajo es de 40 píxeles.

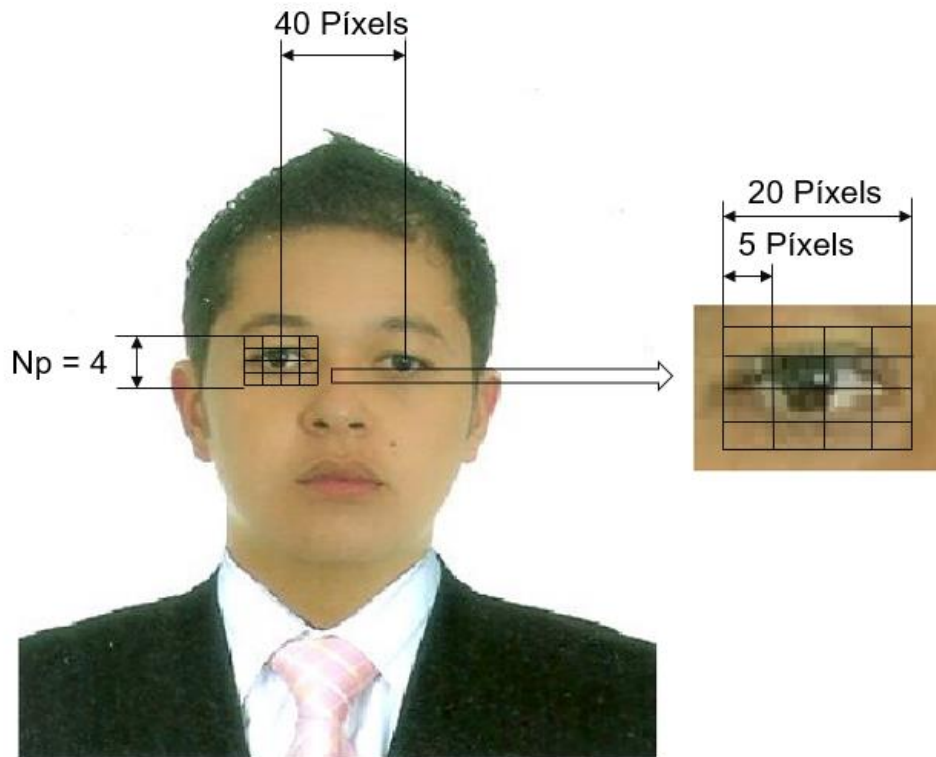


Figura 11. Enrejados espaciales de un descriptor HOG

Fuente. Autor

La contribución de cada gradiente de píxeles para el histograma se pondera por el módulo del gradiente y una ventana gaussiana centrada en el punto clave con media desviación estándar de la extensión del rango de componentes espaciales (10 píxeles). Asimismo, la contribución trilinealmente se interpola con los componentes de los alrededores. [11]

La ventana Gaussiana y la interpolación trilineal aumenta la robustez del descriptor contra los pequeños desplazamientos de la ubicación del punto clave. Finalmente,

para aumentar la invariancia a la iluminación no lineal cambia descriptores HOG por la normalizada.

Capítulo 3

Diseño

Teniendo en cuenta que en este proyecto se realizó una investigación sobre las diferentes técnicas de reconocimiento, en este capítulo se hablará sobre los componentes necesarios para tener un buen desempeño en el reconocimiento de rostros que son los siguientes.

3.1 Captura

Para obtener una correcta identificación de personas inicialmente se debe contar con una buena captura o adquisición de la imagen, teniendo en cuenta que las anteriores técnicas se basan en imágenes fijas del individuo que se desea reconocer. Dentro de la captura de la imagen se encuentra la selección de la cámara digital y las características de iluminación a continuación se presentaran algunas características de cámaras digitales.

3.1.1 Cámaras Web digitales

Como la adquisición o captura de imágenes debe ser de buena calidad se realizó un cuadro de características de varias cámaras digitales con buen rendimiento y optima eficiencia para el reconocimiento de rostros, lo podemos ver en el cuadro 3.

Logitech HD Webcam	
C270	C310
Videoconferencias HD (1280 x 720 píxeles) con el sistema recomendado	Videoconferencias HD (1280 x 720 píxeles) con el sistema recomendado
Captura de vídeo: Hasta 1280 x 720 píxeles	Captura de vídeo HD: Hasta 1280 x 720 píxeles
Tecnología Logitech Fluid Crysta	Tecnología Logitech Fluid Crystal
-----	-----
Fotos: Hasta 3.0 megapíxeles	Fotos hasta 5 megapíxeles
Micrófono integrado con reducción de ruido	Micrófono integrado con reducción de ruido
Certificación USB 2.0 de alta velocidad	Certificación USB 2.0 de alta velocidad
C525	C615
Videoconferencias HD (1280 x 720 píxeles) con sistema recomendado	Captura de vídeo Full HD 1080p (1920 x 1080 píxeles) con sistema recomendado
Captura de vídeo HD: Hasta 1280 x 720 píxeles	Captura de vídeo: Hasta 1280 x 720 píxeles
Tecnología logitech Fluid Cristal	Tecnología Logitech Fluid Crystal
Enfoque automático	
Fotos hasta 8 megapíxeles	fotos hasta 8 megapíxeles
Micrófonos integrados con reducción de ruido automático	Micrófono integrado con reducción de ruido automático
Certificación USB 2,0 de alta velocidad	Certificación USB 2,0 de alta velocidad

Cuadro 3. Tabla comparativa de cámaras digitales web
Fuente. Autor

3.2 Base de datos

Las técnicas de reconocimiento de rostros revisadas (Ver capítulo 2) trabajan con imágenes fijas, según investigaciones son las que mejores resultados han mostrado [14] y disponen de una gran cantidad de diversas base de datos de prueba

disponibles en línea y algunas son gratuitas [13], a continuación vamos a describir las más utilizadas en la actualidad.

3.2.1 The Color FERET Database, USA

La colección de bases de datos fue un esfuerzo de colaboración entre el Dr. Wechsler y el Dr. Phillips. Las imágenes se recogieron en un entorno semi-controlada. Para mantener un grado de consistencia en toda la base de datos, se utilizó la misma configuración física en cada sesión fotográfica. Debido a que el equipo tuvo que ser vuelto a montar para cada sesión, hubo alguna variación menor en imágenes recogidas en diferentes fechas. La base de datos FERET se recogió en 15 sesiones entre agosto de 1993 y julio de 1996. La base de datos contiene 1564 conjuntos de imágenes para un total de 14.126 imágenes que incluyen 1199 personas y 365 conjuntos duplicados de las imágenes. Un conjunto duplicado es un segundo conjunto de imágenes de una persona que ya están en la base de datos y por lo general se tomó en un día diferente. Para algunas personas, habían transcurrido más de dos años entre sus primeras y últimas sesiones, con algunos temas que se fotografiaron varias veces. Este lapso de tiempo fue importante porque permitió a los investigadores para estudiar, por primera vez, los cambios en la apariencia de un sujeto que se producen más de un año. [13]

3.2.2 The Yale Face Database

Contiene 165 imágenes en escala de grises en formato GIF de 15 individuos. Hay 11 imágenes por sujeto, una por la expresión facial diferente o de configuración: centro-luz, feliz, luz izquierda, sin gafas, haga la luz normal, triste, somnoliento, sorprendido, y un guiño. [13]

3.2.3 The ORL Database of Faces

La “Base de Datos de Rostros ORL”, contiene un conjunto de imágenes de 40 personas, con 10 fotos de la cara de cada una de ellas, tomadas entre abril de 1992 y abril de 1994 en el Laboratorio AT&T de Cambridge. En algunas personas, las imágenes fueron tomadas en diferentes momentos, la variación de la iluminación, las expresiones faciales (ojos abiertos / cerrados, sonriendo / No sonrientes) y los detalles faciales (con gafas / sin gafas). Todas las imágenes fueron tomadas en contraste de un fondo homogéneo oscuro con las personas en posición vertical y frontal. Las imágenes de esta base están en niveles de gris, y se modificaron para dejarlas con un tamaño de 100x100 pixels. [13]

3.3 Software

3.3.1 Matlab

MATLAB (abreviatura de “Matrix Laboratory”, o laboratorio de matrices) es un lenguaje de alto nivel de Mathworks con un entorno de desarrollo integrado propio que permite realizar funciones de cálculo numérico, visualización y programación en análisis de datos, desarrollo de algoritmos y creación de modelos o aplicaciones. En MATLAB la programación sigue una estructura similar a los lenguajes C, C++ o Java, y es ampliamente utilizado en sectores tan dispares como el industrial, el económico-financiero, el educativo, etc. Además, el lenguaje del cálculo técnico de MATLAB se utiliza en gran cantidad de universidades y centros de investigación de todo el mundo.

Sus funciones básicas son la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware.

Una de las principales ventajas de MATLAB es que tiene la capacidad de ampliar sus funcionalidades mediante la adición de “toolboxes” por instalación de paquetes propios de la compañía Mathworks. Algunas de estas “toolboxes” son: visión por computador, adquisición de datos, redes neuronales, robótica, adquisición y procesamiento de imagen.

En este proyecto se ha propuesto Matlab para el desarrollo del código y adquisición de datos para el reconocimiento rostros ya que es un software muy potente y empleado en los grandes centros de investigación a nivel mundial.

3.3.2 Microsoft Visual Studio

Microsoft Visual Studio es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) para sistemas operativos Windows. Soporta múltiples lenguajes de programación tales como C++, C#, Visual Basic .NET, F#, Java, Python, Ruby, PHP; al igual que entornos de desarrollo web como ASP.NET MVC, Django, etc., a lo cual sumarle las nuevas capacidades online bajo Windows Azure en forma del editor Mónaco.

Visual Studio permite a los desarrolladores crear sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma .NET (a partir de la versión .NET 2002). Así se pueden crear aplicaciones que se comuniquen entre estaciones de trabajo, páginas web, dispositivos móviles, dispositivos embebidos, consolas, etc.

3.4 Revisión y propuesta de las técnicas de reconocimiento

A partir de la revisión teórica y práctica realizada a las diferentes técnicas de reconocimiento de rostros (ver capítulo 2) se deduce que se pueden utilizar los métodos de Eigenfaces y Fisherfaces, porque de acuerdo a lo manifestado en el artículo [Huiyuan Wang y otros] Eigenfaces tiene un porcentaje de credibilidad del 97 % con la base de datos UMIST y con la base de datos Ar un 78%. El método Fisherface tiene una credibilidad del 98% con la base de datos UMIST y con la base de datos Ar un 89%. [15]

3.5 Hardware

Una vez mencionados los diferentes entornos de programación posibles a trabajar, se mencionan los soportes sobre el cual podremos ejecutar los códigos y realizar la función de captura de imágenes y adquisición de datos.

Como las placas a plantear deben ser de bajo costo, pasaremos a estudiar diferentes placas de desarrollo existente en el mercado.

3.5.1 PcDuino

Es una placa de dimensiones 121 mm X 65 mm de alto rendimiento que soporta sistemas operativos completos como Ubuntu y Android. Se puede programar en C, C++, Java, Python, Arduino, entre otros lenguajes.

El PcDuino 3B tiene una serie de opciones incluyendo un receptor IR, un host SATA, un puerto USB-OTG, una interfaz LCD por puerto LVDS, una interfaz cámara MIPI, un puerto de salida audio de 3.5mm, puerto WiFi, puerto Ethernet, unos headers para conectar una batería Li-po, por lo tanto sólo necesita 5v de alimentación, un teclado y un ratón para funcionar.

El PcDuino 3B utiliza un procesador dual Core A20 lo que lo transforma en el mini-pc más potente actualmente existente.

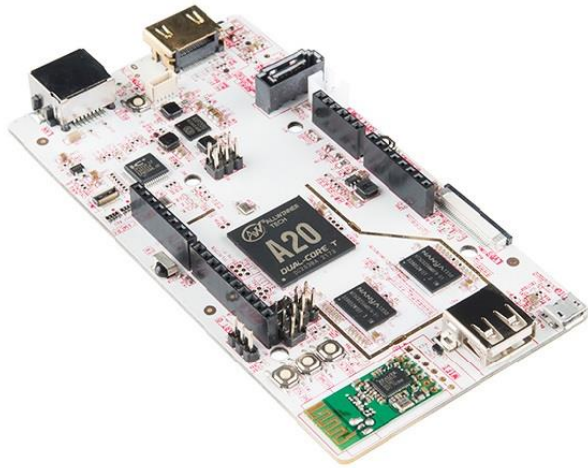



Figura 12. PcDuino
Fuente. Ver referencia 12

3.5.2 Raspberry Pi

Es un ordenador de placa reducido extremadamente sencillo, creada por la fundación Raspberry Pi del Reino Unido con fines educativos, en la actualidad existen 5 modelos diferentes cuyas características se especifican en el siguiente cuadro.

	Comparativa entre Raspberry Pi				
	MODELO A	MODELO A+	MODELO B	MODELO B+	2 MODELO B
Soc	Broadcom BCM2835	Broadcom BCM2835	Broadcom BCM2835	Broadcom BCM2835	Broadcom BCM2836
CPU	700 MHz ARM1176JZF-S	700 MHz ARM1176JZF-S	700 MHz ARM1176JZF-S	700 MHz ARM1176JZF-S	900MHz Quad-core ARM Cortex-A7
GPU	VideoCore IV	VideoCore IV	VideoCore IV	VideoCore IV	VideoCore IV
RAM	256 Mb	256 Mb	512 Mb	512 Mb	1 Gb
USB	1	1	2	4	4
Video	RCA,HDMI	Jack, HDMI	RCA, HDMI	Jack, HDMI	Jack, HDMI
Audio	Jack, HDMI	Jack, HDMI	Jack, HDMI	Jack, HDMI	Jack, HDMI
Boot	SD	MicroSD	SD	MicroSD	MicroSD
Red	-	-	Ethernet 10/100	Ethernet 10/100	Ethernet 10/100
Consumo	300mA / 1,5w / 5v	400mA / 2w / 5v	700mA / 3,5w / 5v	500mA / 2,5w / 5v	800mA / 4w / 5v
Alimentación	MicroUSB / GPIO	MicroUSB / GPIO	MicroUSB / GPIO	MicroUSB / GPIO	MicroUSB / GPIO
Tamaño	85,6 x 53,98	65 x 56	85,6 x 56,98 mm	85 x 56 mm	85 x 56 mm

Cuadro 4. Tabla comparativa entre los modelos de Raspberry Pi

Fuente. Autor

Capítulo 4

Resultados y conclusiones

4.1 Resultados

- El software de Matlab, en cuanto a programación y simulación es eficiente, pero cuando se habla de comunicación, existen programas con mayores beneficios, de acuerdo a los parámetros de uso.
- Las metodologías de Eigenfaces y Fisherfaces son las que obtienen mejor rendimiento, como se dijo en el punto 3.4 donde se dio razón de la credibilidad de estas técnicas.
- Con base en la investigación realizada, sobre las placas de adquisición de datos, se propone usar para un proyecto a futuro la Raspberry Pi 2 modelo B, ya que como se puede observar en el cuadro [4], esta placa cuenta con unas características idóneas para la adquisición de datos. Es una opción viable, ya que se dispone de un dispositivo de entrada de imagen, con unas características favorables a los requerimientos.
- La Raspberry Pi 2 conserva la compatibilidad total con anteriores modelos, además con Windows 10 y su papel en IoT tiene ventajas destacables en cuanto a herramientas de desarrollo con el Software de Visual Studio.
- Se consultaron varias bases de datos para trabajar con ellas en reconocimiento de rostros entre ellas destacan The ORL Database of Faces, The Yale Face Database y The Color FERET Database, USA.

4.2 Conclusiones

Con base en la presente investigación sobre las técnicas de reconocimiento de rostros, las conclusiones obtenidas son

- Los métodos biométricos existentes en la actualidad, son de gran importancia para los sistemas de seguridad, a todo nivel (en empresas, aeropuertos, entidades carcelarias y entidades bancarias,

etc.), debido a la confiabilidad que brinda el realizar la identificación de personas a través de patrones faciales, ya que las diferencias individuales son únicas de una persona a otra.

- Uniendo la técnica Eigenfaces y Fisherfaces, en teoría se obtiene mejor rendimiento en el sistema de reconocimiento de rostros según las diferentes investigaciones revisadas en diversos artículos de autores que han profundizado en el tema. Cabe aclarar, es de suma importancia validar el comportamiento de la unión de estas dos técnicas en la práctica, con el fin de mejorar los niveles de efectividad.
- Para realizar un reconocimiento de rostros se propone utilizar el software de Visual Studio ya que este tiene ventajas de comunicación con las Raspberry pi 2 Modelo B, por este motivo facilita una mejor comunicación a la hora de realizar el reconocimiento de rostros.
- Después realizar pruebas con las bases de datos se concluye que es mejor realizar una base de datos propia, bajo condiciones controladas para el reconocimiento rostros.

Bibliografía

- [1] FBI, 'Face-Recognition', 2015. [Online]. Available: https://www.fbi.gov/about-us/cjis/fingerprints_biometrics/biometric-center-of-excellence/files. [Accessed: 21- Oct- 2015].
- [2] C. Travieso Gonzáles, M. del Pozo Baños and J. Ticay Rivas, 'Sistemas Biometricos', Doctorado, Universidad de las Palmas de Gran Canaria (ULPGC), Mayo 2011.
- [3] A. Eng and L. Wahsheh, 'Look Into My Eyes: A Survey of Biometric Security', *IEEE*, no. 10, p. 6, 2013.
- [4] F. Villalobos Castaldi, 'Uso de la red vascular de la retina como medio biométrico de identificación', Ingeniera, Instituto Politécnico Nacional, México, 2011.
- [5] P. Paredes Guaygua and E. Sánchez Mena, 'Sistema de control biométrico mediante reconocimiento de voz', Universidad Católica del Ecuador sede Ambato, 2004.
- [6] M. Aguilera, 'Reconocimiento Biométrico basado en Imágenes de huellas palmares', Universidad Autónoma de Madrid, 2012.
- [7] P. N., J. P. and D. Kriegman, 'Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection', *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 19, no. 7 pp. 711–720, 1997.
- [8] W. Zhao et al., "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, Dec. 2003.
- [9] F. Bhat and M. Arif Wani, 'Elastic Bunch Graph Matching Based Face Recognition Under Varying Lighting, Pose, and Expression Conditions', *International Journal of Advance Foundation And Research In Science & Engineering*, vol. 1, p. 9, 2014.
- [10] D. Monzo, A. Albiol, J. Sastre and A. Albiol, 'HOG-EBGM VS. GABOR EBGM', *IEEE*, p. 4, 2008.

[11] D. Monzo, A. Albiol, A. Albiol and J. Mossi, 'A Comparative Study of facial landmark localization methods for Face Recognition using HOG descriptors', *IEEE*, p. 4, 2010.

[12] Linksprite.com, 'LinkSprite | LinkSprite pcDuino', 2015. [Online]. Available: http://www.linksprite.com/?page_id=782. [Accessed: 21- Oct- 2015].

[13] Face-rec.org, 'Face Recognition Homepage', 2015. [Online]. Available: <http://www.face-rec.org/>. [Accessed: 14- Nov- 2015].

[14] 'Sistemas de reconocimiento basados en la imagen facil', *Facultad de ingenieria Físico-Mecánico, Universidad Industrial de Santander, Bucaramanga, Colombia*, vol. 83, p. 10, 2015.

[15] 'Face Recognition Combing Principal Component Analysis and Fractional-step Linear Discriminant analysis', *IEEE*, p. 4, 2015.