

AUTENTICUZ: SISTEMA DE RECONOCIMIENTO FACIAL PARA CONTROL DE ACCESO AUTOMÁTICO.

Alejandro Villalba, Juan Manuel Artacho, Diego Sanchez, Emiliano Bernués

Departamento de Ingeniería Electrónica y Comunicaciones

Instituto de Investigación de Ingeniería de Aragón

Universidad de Zaragoza

Email de contacto : ebr@posta.unizar.es

Abstract-In this paper, we present a face recognizing system for automatic access control. This system identifies the user by mean of an identification card, afterwards, a frontal photography of the user is taken and processed to obtain the AAM (Active Appearance Models) parameters of the image. Once the system has loaded the parameters, it compares them with the parameters of the users which were previously stored in the database. The decision of the identity of the user is based in different methods, but best results were obtained using Support Vector Machines (SVM).

I INTRODUCCIÓN

En este artículo se presenta el desarrollo de una aplicación para control automático de acceso basado en reconocimiento facial. Los rasgos faciales a reconocer, son una característica biométrica [1] única de cada individuo, como también lo son las huellas digitales o la forma del iris.

Al igual que un sistema biométrico general, nuestro sistema actúa de la siguiente manera: tomamos una muestra de la característica biométrica a reconocer, que en nuestro caso será una imagen. La entrada del sistema, es sometida a un procesado que extrae sus rasgos característicos.

Una vez obtenidos estos rasgos característicos, se comparan con los previamente almacenados y el resultado de la comparación verifica o descarta la identidad de la persona. Así, es necesaria una etapa previa de creación de la base de datos. A esta etapa la llamaremos fase de registro y se realizará 'off-line' y previamente al funcionamiento del sistema en la fase de identificación.

En nuestro sistema la extracción de parámetros biométricos, consistió en el cálculo de los parámetros ASM[2] (Active Shape Models) y AAM[3] (Active Appearance Models), los cuales se explican en el apartado II.

Una vez obtenidas las características biométricas, se procede a la comparación de éstas, sin embargo, existen muchas posibles formas de comparación de ellas. Es este artículo se presentarán los resultados obtenidos utilizando

tanto el decisor basado en distancia Euclídea como los decisores basados en SVM, que se detallan en el apartado III.

II MODELOS ESTADÍSTICOS DE FORMA ASM Y AAM.

La técnica AAM (Active Appearance Models) consiste en la construcción un modelo estadístico de forma y de textura de un objeto a partir de las propiedades geométricas y de la textura observadas en un conjunto de imágenes de entrenamiento. Este conjunto contiene imágenes de objetos de una misma clase inicial (en nuestro caso los rostros) marcados con una serie de puntos denominados 'landmarks'.

El modelo estadístico está compuesto por el contorno y textura medias del conjunto de entrenamiento, junto con las direcciones de variación respecto a esos valores medios. Estas direcciones de variación se obtienen de la matriz de covarianza creada con las de las imágenes de entrenamiento marcadas.

El modelo estadístico permite generar ejemplares válidos del objeto a ajustar sobre otro objeto nuevo que se encuentre en una nueva imagen. Es por tanto, una técnica top-down: primero se genera el ejemplar del modelo y después se deforma según las direcciones de variación para ajustarlo a la imagen.

El proceso completo de análisis de una imagen mediante AAM, está constituido por dos fases principales:

- Fase de construcción del modelo AAM (entrenamiento)
- Fase de ajuste del modelo AAM (reconocimiento).

La fase de construcción del modelo abarca diferentes tareas: la elección de las imágenes de entrenamiento, elección de puntos de control ('landmarks') adecuados, marcado de las imágenes, alineamiento y normalización del conjunto de entrenamiento, cálculo del modelo estadístico de forma y cálculo del modelo estadístico de textura. Esta etapa se realiza una sola vez. Una vez se ha obtenido el modelo, se puede utilizar tantas veces como sea necesario.

La fase de ajuste toma de punto de partida el modelo ya construido e intenta ajustarlo sobre el objeto de una imagen nueva. Para llevar a cabo esto, la forma o contorno medio de nuestro modelo será proyectado sobre la imagen. Mediante

sucesivas iteraciones, se va deformando modificando los parámetros dentro de los límites establecidos, de manera que nunca deje de ser un ejemplo válido. Cuando acaba este proceso de ajuste, el contorno resultante será una realización del modelo con unos parámetros únicos y que se ajustará lo más posible a la imagen origen.

El uso de la técnica AAM, para realizar el proceso de ajuste sobre un objeto de una imagen nueva, requiere que el contorno del modelo se halle próximo al contorno del objeto de la imagen nueva. La colocación del contorno del modelo, de la que partimos para realizar el proceso iterativo de ajuste al objeto de la imagen nueva, recibe el nombre de 'inicialización del sistema AAM'. En nuestro caso, la inicialización del sistema AAM que se considera es la salida de la técnica ASM (Active Shape Model) [2]. Esta última técnica proporciona unos resultados inferiores a los del AAM, pero sin embargo no requiere de una inicialización precisa, por lo que se ajusta a los requisitos de la inicialización del sistema AAM.

La diferencia entre estas dos técnicas (ASM y AAM) radica en el proceso de ajuste de su respectivo modelo a un objeto de una nueva imagen. Mientras el sistema ASM utiliza, exclusivamente, la textura próxima al contorno del objeto de una imagen, el sistema AAM se basa en la textura total del objeto y, obviamente, esto cambia el método de ajuste.

En el proceso de ajuste AAM, es necesaria una función de ajuste que transforme el conjunto de puntos en los que vamos a muestrear de la forma media, en el conjunto de puntos a muestrear en cada una de las imágenes del conjunto de entrenamiento.

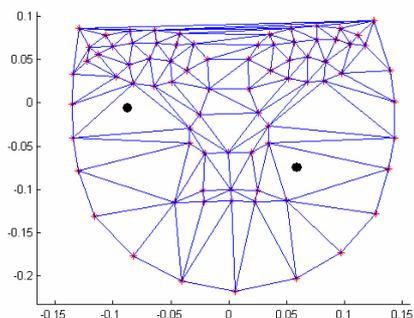


Figura.1. Triangulación de Delanuy.

La función f utilizada en esta transformación es la llamada 'Piece-wise Affine', basada en la triangulación de Delanuy[3] (Figura 1y 2)

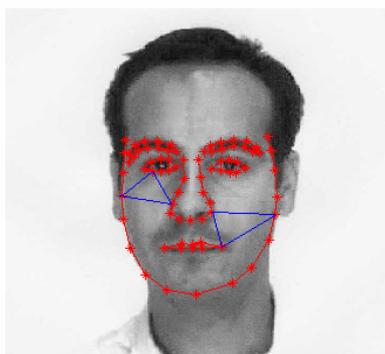


Figura 2. Ajuste de coordenadas AAM.

III SUPPORT VECTOR MACHINES (SVM)

En nuestro sistema hemos utilizado un sistema de decisión SVM multiclase tanto lineal como no lineal, pero para comprender este sistema, analizaremos previamente el SVM binario [4].

III.1 SVM LINEAL BINARIO.

El SVM, es un método de clasificación binaria, esto quiere decir que en principio, sólo nos será posible clasificar cada una de las muestras como perteneciente a una de las dos clases existentes, que etiquetaremos como $\{-1,1\}$.

Este sistema, consta de una fase de entrenamiento de la máquina SVM y de una fase de reconocimiento que se basa en los datos calculados en entrenamiento. Para dicho entrenamiento, el sistema se basa en un conjunto de muestras de entrenamiento x_i , a los que corresponde una clasificación

y_i (1 ó -1). Una vez tenemos el conjunto de entrenamiento, el método se basa en maximizar el margen o la distancia entre las muestras más cercanas de ambas clases, para que la separación entre clases sea máxima.

Así pues, el objetivo del decisor por SVM lineal en el caso separable, consistirá en encontrar aquel hiperplano que maximice el margen entre las dos clases. Esta condición puede expresarse con las inecuaciones siguientes:

$$\underline{x}_i \cdot \underline{w} + b \geq 1 \text{ para } y_i = +1 \quad (1)$$

$$\underline{x}_i \cdot \underline{w} + b \leq -1 \text{ para } y_i = -1 \quad (2)$$

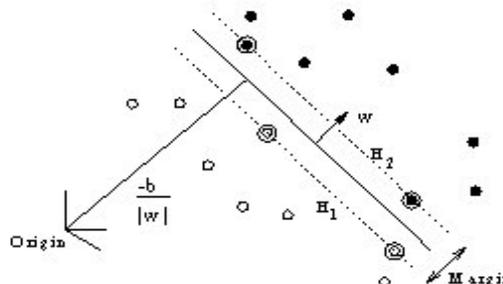


Figura.3. SVM lineal.

Donde w es el vector normal a la superficie de separación interclases.

Ahora, tomemos en consideración los puntos para los que en las inecuaciones (1) y (2), se cumple la condición de igualdad (afirmar que dichos puntos existen, es equivalente a escalar los valores de b y $\|w\|$, de forma que la distancia punto-hiperplano sea la exigida). Como vemos en la figura, la distancia entre hiperplanos (margen a maximizar) $abs(H_1 - H_2)$ será $2/\|w\|$. Con lo que nuestro objetivo será minimizar $\|w\|^2$.

Utilizando la formulación Langragiana, la relación entre las muestras aparecerá sólo en forma de productos escalares.

$$Lp \equiv \frac{1}{2} \|\underline{w}\|^2 - \sum_{i=1}^l \alpha_i y_i (\underline{x}_i \cdot \underline{w} + b) + \sum_{i=1}^l \alpha_i \quad (3)$$

El problema se reduce a minimizar Lp , sujeto a que $\alpha_i \geq 0$ para todo i , y además las derivadas de Lp , respecto a cada uno de los α_i se anulan.

Nos encontramos ante un problema convexo cuadrático [5], cuya resolución mediante métodos numéricos, permitirá obtener los valores de los multiplicadores de lagrange.

Una vez hemos entrenado la máquina SVM, podremos pasar a la fase de decisión, es decir a la clasificación de una nueva muestra como perteneciente a una de las dos clases posibles $\{1, -1\}$. Puesto que los puntos del hiperplano central H cumplen que : $\underline{x} \cdot \underline{w} + b = 0$, clasificaremos las muestras, según estén a un lado o a otro del hiperplano, o lo que es lo mismo, según el signo de la expresión:

$$\text{sgn}(\underline{x} \cdot \underline{w} + b) \quad (4)$$

Para el caso de sistemas no separables, añadiremos una penalización por los errores cometidos a la función objetivo[4].

III.2 SVM NO LINEAL BINARIO.

En muchos de los problemas reales, las muestras no pueden separarse en clases mediante hiperplanos, sino que las fronteras entre ellas, deberían tener formas no lineales [4]. La extensión al problema no lineal, se consigue gracias a que en la función objetivo y en la función de clasificación, las muestras del conjunto de entrenamiento aparecen sólo en forma de productos escalares. Este hecho, permite realizar una correspondencia previa de los datos originales $\underline{x}_i \in \mathfrak{R}^d$, con otros elementos pertenecientes a otro espacio vectorial \mathfrak{N} que podría ser de dimensión infinita.

En este caso, el entrenamiento de la SVM, dependería de los datos en función del producto escalar de sus respectivas asignaciones en el espacio vectorial \mathfrak{N} : $\Phi(\underline{x}_i) \cdot \Phi(\underline{x}_j)$.

En ocasiones, es posible obtener una función kernel K , tal que :

$$K(\underline{x}_i, \underline{x}_j) = \Phi(\underline{x}_i) \Phi(\underline{x}_j) \quad (5)$$

Si encontramos esta función K , entonces podremos realizar el entrenamiento sin necesidad de realizar los productos escalares. Además, puesto que, como hemos dicho, el proceso de entrenamiento sólo depende de los datos por medio de estos productos escalares, no necesitaremos ni siquiera realizar físicamente la asignación de los datos al nuevo espacio vectorial, por lo que se evita conocer la asignación.

En cuanto a la función de decisión deberíamos aplicar la función $\text{sgn}(\underline{x} \cdot \underline{w} + b)$, pero en este caso $\underline{w} \in \mathfrak{N}$, y por tanto no podemos tener acceso directo a él, por lo que utilizaremos la expresión equivalente, donde se ha sustituido \underline{w} por su expresión en función de los multiplicadores de lagrange obtenidos.

$$\text{sgn}\left\{\sum_{i=1}^{N_s} \alpha_i y_i k(\underline{s}_i, \underline{x})\right\} \quad (6)$$

Algunas de las funciones kernel que utilizamos para este propósito son:

Gausiana:

$$K(\underline{x}_i, \underline{x}_j) = \exp\left(-\frac{\|\underline{x}_i - \underline{x}_j\|^2}{2\sigma^2}\right) \quad (7)$$

Polinomial:

$$K(\underline{x}_i, \underline{x}_j) = (\underline{x}_i \cdot \underline{x}_j + 1)^p \quad (8)$$

III.3 SVM MULTICLASE.

La extensión al problema multiclase en nuestro sistema, se implementó de dos maneras distintas: SVM uno vs resto (1 vs R) y SVM- DAG (Directed Acyclic Graph).

III.3.1 SVM UNO VS RESTO.

Este sistema consiste en realizar comparaciones binarias sucesivas entre cada una de las clases y el resto de ellas. De esta forma, es de suponer que si una de las clases fuera la correcta, sólo esa comparación nos daría un resultado positivo. Esta extensión del problema binario, presenta la ventaja de permitir el resultado de “no perteneciente a ninguna de las clases”, resultado que se produciría si ninguna de las clasificaciones “uno respecto a los demás”, resultara positiva. También existe la posibilidad de que, debido a un error, varias de las comparaciones fueran positivas, y por tanto habría que realizar algún tipo de “desempate” entre las categorías en las que resultó positiva la comparación con las demás.

Con este sistema necesitaremos n máquinas y en cada una de ellas utilizaremos todas nuestras muestras.

III.3.2 SVM DAG (Directed Acyclic Graph).

En este caso, para extender el problema binario, construiremos un gráfico de manera que cada nodo del gráfico representará una máquina SVM. Cada uno estos nodos dispondrá de dos ramas, una por cada posible decisión tomada por el SVM. De esta manera, el gráfico completo irá tomando una estructura piramidal.

Con esta estructura, iremos descartando una de las posibles categorías en cada uno de los niveles del gráfico. En este caso utilizaremos $n(n-1)/2$ máquinas, pero en cada una sólo utilizaremos las muestras de dos clases, con lo que cada una de las máquinas será entrenada más rápidamente que en el caso anterior.

IV RESULTADOS.

Con el sistema implementado y utilizando distintos decisores, se realizaron diferentes ensayos con 3 grupos distintos de imágenes, obteniéndose los siguientes resultados.

IV.1 PRIMER CONJUNTO DE IMÁGENES.

En las primeras pruebas llevadas a cabo en este trabajo se han utilizado un conjunto de imágenes obtenidas de [6]. Dicho conjunto esta formado por imágenes de formato RAW y tamaño 768x576. Estas imágenes corresponden a 75 individuos varones y a 58 mujeres, disponiéndose de 4 fotografías de cada uno de ellos. Estas 4 imágenes se corresponden con 4 expresiones faciales: neutral, sonriente, seria y bostezando.

	Nº de Imágenes incluidas en el modelo	Nº de Imágenes por persona en el modelo.	Nº de Imágenes por persona en los decisores SVM.
CASO 1	39	3	3
CASO 2	40	1	3

Tabla.1. Características del primer conjunto de ensayo.

	EUCLÍDEO	EUCLÍDEO NORMALIZADO	SVM	
			DAG	1-VS-R
CASO 1	58,12%	56,40%	-	-
CASO 2	39,16%	37,50%	-	-

Tabla.2. Resumen de resultados con el primer conjunto.

Con este conjunto de imágenes de pruebas en el que sólo disponemos de cuatro imágenes por persona, no es posible utilizar decisores basados en SVM. Por otra parte, los resultados de los decisores Euclídeos no eran demasiado altos (50%-60%), por lo que decidimos utilizar un número mayor de imágenes para poder utilizar los decisores SVM. El decisor Euclideo normalizado, trabaja para la decisión con los parámetros del modelo AAM normalizados a media cero y varianza unidad.

IV.2 SEGUNDO CONJUNTO DE IMÁGENES.

En este caso, se sintetizaron nuevas imágenes para cada individuo, 6 en concreto, con lo que tendremos 10 imágenes (con las 4 de que ya disponíamos). Consideraremos las nuevas imágenes como nuevas expresiones.

	Nº de imágenes incluidas en el modelo	Nº de expresiones por individuo	Nº de impostores	Nº de expresiones en el decisor SVM
CASO 1	40	4	0	9
CASO 2	64	4	8	9
CASO 3	40	4	5	9
CASO 4	60	1	50	9

Tabla.3 Características del segundo conjunto de ensayo

	EUCLÍDEO	SVM	
		DAG	1-VS-R
CASO 1	70%	76.67%	90%
CASO 2	90%	91.67%	90%
CASO 3	66.67%	83.33%	90%
CASO 4	72%	86%	88.89%

Tabla.4 Resumen de resultados con el segundo conjunto.

Una vez tenemos un número suficiente de imágenes, vemos que los resultados del SVM uno frente a los demás mejoran claramente respecto a los Euclídeos, sin embargo, habría que añadir que la probabilidad de falso positivo

aumenta, puesto que en ocasiones, este decisor devuelve más de un individuo como posible reconocido.

IV.3 TERCER CONJUNTO DE IMÁGENES.

En este apartado, utilizaremos expresiones reales, es decir, adquiridas por nuestra cámara y veremos que porcentaje de acierto alcanzamos en los distintos casos.

	Nº total de individuos.	Nº de Expresiones por individuo en el modelo AAM	Nº de individuos adquiridos con nuestra camara	Nº de Expresiones utilizadas en el SVM
CASO1	10	4	0	9
CASO2	11	4	1	9
CASO3	12	4	4	9

Tabla.5. Características del tercer conjunto de ensayo.

	EUCLÍDEO	SVM	
		DAG	1-VS-R
CASO 1	55.77%	<30%	<30%
CASO 2	66.67%	68.34%	88.88%
CASO 3	43.33%	47.03%	72.96%

Tabla.6. Resumen de resultados con el tercer conjunto.

Vemos que introduciendo imágenes reales, los resultados son peores, como era de esperar, puesto que influyen las imperfecciones en la captación de las imágenes. Aun así, los resultados son satisfactorios y pueden mejorarse con un sistema de adquisición más depurado.

V. CONCLUSIONES

Se ha presentado un sistema de verificación de identidad basado en biometría facial con un decisor basado en SVM y modelo estadístico de forma y textura AAM. Los resultados que se presentan son satisfactorios y se comprueba como con un sistema SVM bien dimensionado es posible una mejora notable de los resultados de verificación.

AGRADECIMIENTOS

Este proyecto ha sido financiado por la Universidad de Zaragoza. Referencia: UZ2001-TEC-01.

REFERENCIAS.

- [1] R. Chellappa, C. Wilson and A. Sirohey, "Human and machine recognition of faces: A survey", Proceedings IEEE, vol 83, nº 5, pp. 705-740, 1995.
- [2] T.F. Cootes y C.J. Taylor. "Statistical Models of Appearance for Computer Vision" (<http://www.isbe.man.ac.uk/~bim/refs.html>).
- [3] M. B. Stegmann, *Active Appearance Models: Theory, Extensions and Cases*, pp. 262, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 2000, (<http://www.imm.dtu.dk/~aam/>)
- [4] C. Cortes and Vladimir Vapnik: Support-Vector Networks. Machine Learning vol 20 nº 3: pp.273-297, 1995.
- [5] R.Fletcher. "Practical Methods of Optimization" 1987.
- [6] Ref.: A.M. Martinez and R. Benavente. Centro de Visión por Computador de la universidad autónoma de Barcelona *The AR Face Database*. CVC Technical Report #24, June 1998. (http://rv11.ecn.purdue.edu/~aleix/aleix_face_DB.html)