

# Un Estudio sobre la Identificación de Personas basada en su Movimiento al Caminar (*Gait*)

Ángel Sánchez, Juan José Pantrigo, Alberto Rubio and Jesús Virseda

Departamento de Ciencias de la Computación  
Universidad Rey Juan Carlos, C/Tulipán, s/n,  
28933 Móstoles, Madrid, Spain

{angel.sanchez, juanjose.pantrigo, a.rubio, j.virseda}@urjc.es

**Abstract.** Este trabajo presenta un prototipo de sistema basado en el conjunto de la postura y el movimiento al caminar (en inglés, *gait*) con el objetivo de reconocer personas. Esta modalidad biométrica presenta las ventajas de ser poco invasiva y de resultar relativamente fácil capturar las secuencias de datos para la experimentación. En nuestro caso particular, se han obtenido resultados de reconocimiento muy satisfactorios usando secuencias de vídeo muy cortas (en promedio, 51 fotogramas por secuencia), con muy pocos individuos para la experimentación (sólo 6 personas) y también usando pocas características discriminantes (en total, 5).

## 1 Introducción

El reconocimiento de personas basado en su postura y forma de caminar (*gait*) es un método de identificación de individuos estudiada desde el siglo XIX por la Medicina y la Biomecánica [1]. Desde un punto de vista más general, el patrón de movimiento al caminar de una persona puede indicar algún tipo de patología. Además, este patrón podría ser analizado para el diseño de tipos de calzados y pavimentos deportivos. El análisis del movimiento humano (y del *gait*) desde la Visión Artificial es un área de investigación mucho más reciente. El concepto de biometría basada en *gait* aparece hacia 1994 [2]. Desde entonces esta modalidad conductual de reconocimiento biométrico recibido también una mayor atención desde la Visión Artificial. En 2006 aparece un libro de Nixon, Tan y Chellappa [1] que incluye los principales técnicas, sistemas, bases de datos y trabajos realizados hasta esa fecha dentro de la identificación humana basada en el *gait*. Un trabajo destacado es el de Sarkar y otros [3] que trata de medir la evolución y caracterizar las propiedades del *gait* usando un conjunto de 12 experimentos y una gran base de datos (este abordaje ha sido denominado por los autores de trabajo como *HumanID Gait Challenge Problem*). Otro trabajo a destacar es la propuesta recopilatoria de Boyd y Little [4] que describe los factores que afectan al reconocimiento basado en *gait*, los métodos de evaluación usados y también compara diferentes sistemas de reconocimiento basados en este tipo de biometría. Una propuesta reciente en esta área es la realizada por Boulgouris y

Chi [5], basada en el ajuste (por separado) entre las componentes de la siluetas corporales a reconocer. En dicho trabajo se identifica también la contribución de cada una de las componentes corporales consideradas (p. ej. cabeza, torso, brazos, piernas, etc) en el rendimiento del sistema de reconocimiento propuesto. Entre las bases de datos de secuencia de *gait* presentadas en la literatura, deben destacarse las siguientes por su tamaño y variabilidad: *CASIA Gait Database* [6], *Southampton Human ID at a Distance Gait Database* [1] y *UMD Surveillance Data* [7].

Nuestro trabajo presenta un estudio preliminar y un prototipo de sistema de reconocimiento biométrico basado en la postura y la forma de caminar. En este caso particular, y al tratarse de un primer abordaje al problema, se ha creado una base de datos propia que contiene exclusivamente secuencias de vídeo grabadas en interior y de corta longitud, con sólo 6 secuencias para la experimentación (una secuencia de vídeo por individuo) y pocas características discriminantes (en total, 5). Con estos datos se ha conseguido identificar satisfactoriamente al 100% de los individuos analizados.

## 2 Descripción de la solución propuesta

La Figura 1 describe gráficamente las etapas seguidas para resolver el problema de reconocimiento planteado. En las secciones sucesivas se explican con detalle cada una de las etapas consideradas.

## 3 Preproceso de los datos

La etapa de preproceso puede descomponerse en la secuencia de pasos intermedios que se describen a continuación.

### – Captura y procesamiento inicial del vídeo

Para este trabajo se han usado secuencias de vídeo propias, por ello es importante una adecuada captura de los datos, ya que el resto del proceso dependerá de la fiabilidad y facilidades que nos proporcione este paso. Se ha trabajado con secuencias de vídeo cortas (en promedio de 51 fotogramas por captura de individuo con, al menos, un ciclo completo de paso) y las grabaciones se han realizado en un estudio fotográfico mediante un esquema de iluminación basado en contraluz, con un fondo blanco y un foco dirigido hacia el individuo caminando, con la idea de obtener su silueta en negro y el fondo totalmente blanco. El procedimiento de captura favorece a la segmentación por umbralización para separar en cada fotograma el sujeto del fondo. Cada secuencia de vídeo capturada se tiene en formato AVI y se ha extraído la componente azul del vídeo, utilizándose como imagen en escala de grises, ya que esta componente del espacio de color es la que mejor representa el contraste de luces. Un ejemplo de imagen en escala de grises obtenida por el procedimiento explicado aparece en la Figura 2.a.

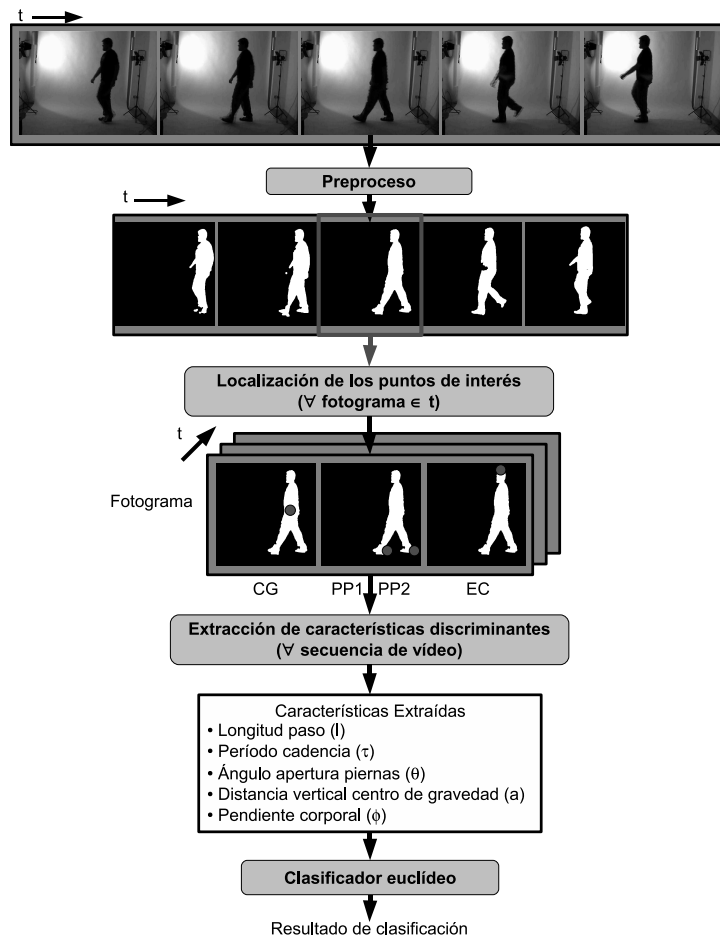


Fig. 1. Arquitectura del sistema propuesto, basado en *gait*.

#### – Umbralización adaptativa de los fotogramas

A pesar de la calidad de las tomas realizadas, y debido a la autorregulación por parte del diafragma de la cámara, la intensidad lumínica en la secuencia del vídeo no es uniforme para toda la escena, por lo que no podemos establecer un umbral fijo de binarización desde el comienzo. Por este motivo, y gracias al contraste y equilibrio que hay entre las zonas claras y oscuras en cada fotograma, se puede utilizar la media del valor de todos los píxeles para fijar un buen umbral para la binarización de los fotogramas. Esta idea, sin embargo, no ha producido resultados correctos ya que, en general, la proporción de zonas oscuras (que originalmente representan individuos y que se quieren separar del fondo) es sólo de un tercio. Por lo tanto, multiplicando el valor medio de los niveles de intensidad en la imagen por un tercio,



**Fig. 2.** (a) Imagen en escala de grises obtenida de la componente azul del fotograma y (b) resultado de la binarización adaptativa para un fotograma ejemplo.

obtenemos el correcto valor del umbral. El resultado de esta umbralización para un fotograma ejemplo aparece en la Figura 2.b.

– **Sustracción de fondo aplicada a cada fotograma**

Para la perfecta localización del sujeto, es necesaria la eliminación del fondo del fotograma, ya que las irregularidades que introduce sobre el sujeto pueden complicar la obtención posterior de puntos de interés, añadiendo ruido e imperfecciones. La sustracción de fondo se realiza para cada fotograma del vídeo con el fondo de la escena (es decir, usando una misma imagen sin el sujeto caminando).

– **Eliminación de ruido**

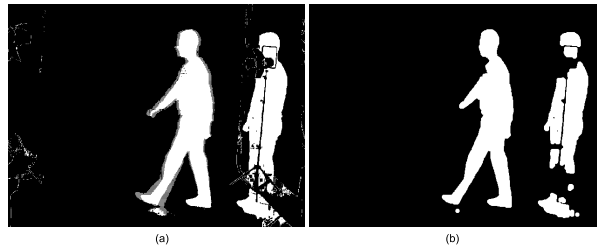
En esta etapa se trata de “limpiar” la silueta del sujeto, que contiene una serie de líneas horizontales resultantes del entrelazado introducido durante la conversión del formato propio de la cámara al formato AVI. Además, aparece un ruido en el fondo correspondiente a unos puntos blancos donde estaban situados los focos. Este ruido se debe a una inevitable vibración de la cámara a causa del propio paso del sujeto y a variaciones de intensidad durante la secuencia de vídeo. Para eliminar este ruido, se ha aplicado una sencilla operación de apertura morfológica. El resultado de la sustracción de fondo junto con la eliminación de ruido se muestran en las Figuras 3.a y 3.b, respectivamente.

– **Recorte y selección de fotogramas**

Para conseguir secuencias donde aparezcan en todos los fotogramas la silueta completa del individuo caminando, se eliminan manualmente algunos fotogramas al principio y al final de cada secuencia grabada. Estas nuevas secuencias contienen, al menos, dos o tres pasos (por lo que se dispone, al menos, de un ciclo completo de paso por vídeo). Un fotograma correctamente recortado y completo dentro del ciclo de un paso aparece en la Figura 4.a.

## 4 Obtención de puntos de interés

La etapa de detección de puntos de interés permitirá posteriormente calcular las características discriminantes consideradas. Los puntos de interés necesarios y su proceso de detección se describen a continuación.



**Fig. 3.** (a) Resultado de la substracción de fondo y (b) de la eliminación de ruido para un fotograma ejemplo



**Fig. 4.** (a) Silueta recortada para un fotograma ejemplo y (b) localización del centro de gravedad, marcado en el centro de la cadera.

– **Centro de gravedad (punto CG)**

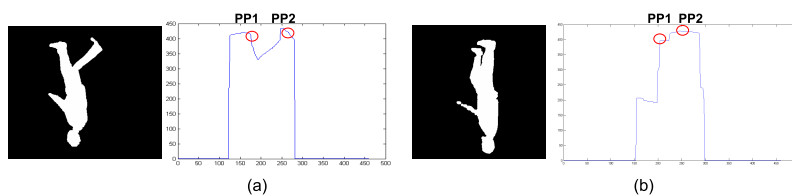
Este punto permite localizar al sujeto dentro de cada fotograma y nos indica dónde está el centro de la cadera del mismo en cada instante. Su cómputo se realiza a partir de la media de las abscisas y de las ordenadas, respectivamente, de todos los puntos de la silueta del sujeto (píxeles blancos). El resultado de la localización de este punto para un fotograma ejemplo aparece en la Figura 4.b.

– **Plantas de los pies (puntos PP1 y PP2)**

Para la detección de las posiciones aproximadas de cada pie, se obtiene la imagen especular con respecto al eje de abscisas del fotograma considerado. A continuación, se trazan, por cada una de las columnas de la imagen, líneas descendentes en vertical hasta tocar algún píxel de la silueta invertida del sujeto, obteniéndose un tipo de histograma como el de la Figura 5.a para el ejemplo considerado. De esta manera, un pie se encuentra como el primer máximo relativo que supere un cierto umbral experimentalmente calculado, buscando desde la parte izquierda del histograma. Análogamente, el otro pie se localiza de forma similar pero comenzando ahora la búsqueda desde la parte derecha del histograma.

El problema que presenta esta solución es, básicamente, que ante la presencia de falsos máximos relativos producidos por las manos en algunas ocasiones (véase la Figura 5.b), existe una probabilidad alta de error en la localización

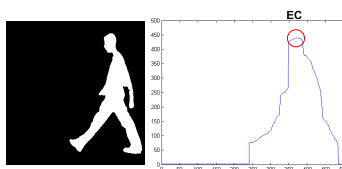
de las plantas de los pies. Sin embargo, se puede refinar la búsqueda, teniendo en cuenta el valor y la posición relativa de los máximos locales para “filtrar” las posiciones correctas de dichos máximos y resolver situaciones como la de la Figura 5.b. En el caso de que los pies estén juntos, el método determinará que las posiciones de ambos pies coinciden el mismo punto, es decir en el único máximo que habrá en toda la imagen. Finalmente, se deshace la simetría especular para obtener la posición correcta de cada uno de los puntos ( $PP1$  y  $PP2$ ) en cada fotograma.



**Fig. 5.** (a) Contorno invertido del sujeto y búsqueda de los pies y (b) análisis para evitar falsos máximos.

– **Extremo superior de la cabeza (punto  $EC$ )**

Se aplica un proceso similar al realizado para determinar la posición de las plantas de los pies. En este caso sólo se busca un único punto y, por una mayor sencillez, no se calcula la imagen especular de la silueta del individuo. El máximo absoluto que se obtiene sobre el histograma calculado determina el punto extremo superior de la cabeza ( $EC$ ). Un ejemplo de la localización de este punto aparece en la Figura 6.



**Fig. 6.** Localización del punto  $EC$  en la cabeza.

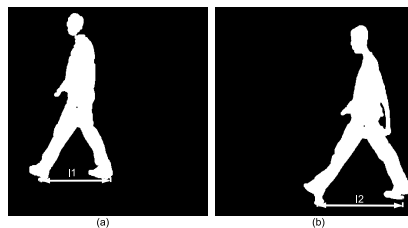
## 5 Obtención de características discriminantes

A partir de la localización automática de los cuatro puntos ( $CG$ ,  $PP1$ ,  $PP2$ ,  $EC$ ) extraídos de la silueta del individuo caminando, se procede a calcular un conjunto

de características discriminantes. Para dichas características se elegirá su valor máximo en cada secuencia de vídeo. Todas las características discriminantes consideradas se han calculado para resultar invariantes a escala y a traslación.

– **Longitud máxima (o envergadura) del paso**

Representa la distancia máxima, en el eje de abscisas, de separación de los pies para una secuencia de vídeo que contiene, al menos, un ciclo de paso. Con el fin de obtener invarianza a escalados, esta característica, expresada en número de píxeles, se normaliza dividiendo por la altura de la silueta del individuo en el primer fotograma. El resultado será un valor en el intervalo  $[0..1]$ . La Figura 7 ilustra esta característica con un par de ejemplos.



**Fig. 7.** Envergadura del paso. Dos ejemplos.

– **Periodo de cadencia del paso**

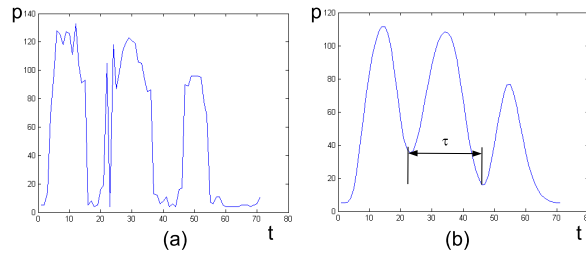
Si se analiza en una secuencia de vídeo, la distancia existente entre ambos pies en un semi-ciclo del paso (fotogramas que transcurren entre el apoyo consecutivo de cada uno de los pies en el suelo), los valores obtenidos describen un movimiento ondulatorio. Si se representa gráficamente esta distancia de separación de los pies (en píxeles) para un individuo concreto, entre fotogramas consecutivos, se obtiene una curva como aparece en la Figura 8.a. Al existir en esta gráfica pequeñas oscilaciones (causadas por la imperfección en la localización de los puntos), se ha suavizado la curva realizando la media entre el valor de cada punto de la gráfica y el siguiente, obteniéndose como resultado una nueva curva como la de la Figura 8.b. En ella, el valor  $\tau$  se calcula como el máximo entre dos mínimos locales de la gráfica y define el tiempo (o el número de fotogramas) de la cadencia en un paso del individuo.

– **Ángulo máximo de apertura de piernas al andar**

Para calcular esta característica angular, representada en radianes, se definen los dos vectores que unen el centro de gravedad del individuo con ambas plantas de sus pies, y se obtiene el ángulo  $\theta$  que forman ambos vectores. Dicho ángulo se calcula únicamente en el fotograma donde la envergadura del paso es máxima. La Figura 9.a ilustra esta característica.

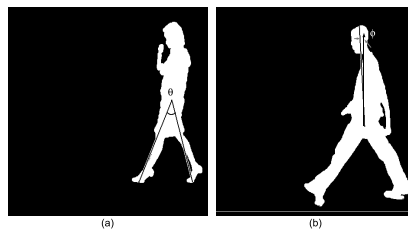
– **Pendiente de la postura corporal**

Esta característica angular ilustra la inclinación corporal una persona al andar (véase la Figura 9.b). Se define como el ángulo que forma el eje de



**Fig. 8.** (a) Variación de la distancia entre los pies en píxeles  $p$  en el eje de ordenadas a lo largo de los fotogramas (tiempo  $t$  en el eje de abcisas) y (b) suavizado de la curva de distancia (Nota: los valores de las distancias representadas son previos a la normalización).

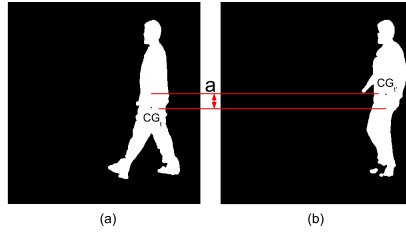
ordenadas con el vector que une el centro de gravedad  $CG$  con el punto más alto de la cabeza  $EC$  en la silueta del individuo. Este ángulo permanece más o menos constante durante los fotogramas de la secuencia de vídeo y caracteriza bien a cada persona, ya que ésta ha acostumbrado a su cuerpo a repartir las masas de una determinada forma para que se mantenga el equilibrio al andar. Este ángulo  $\Phi$ , en radianes, se ha normalizado dividiéndolo por su módulo.



**Fig. 9.** (a) Ángulo máximo formado por ambos pies con el centro de gravedad corporal y (b) Pendiente corporal de un individuo.

– **Variación máxima (en vertical) del centro de gravedad**

Como al caminar la cadera realiza un movimiento ondulatorio, el centro de gravedad del individuo también lo hará. Se ha calculado la variación en el eje de ordenadas de la posición del punto  $CG$  durante todos los fotogramas de cada secuencia de vídeo (ver Figura 10). El valor máximo  $a$  de esta variación, dada en píxeles, se ha normalizado de manera similar a la envergadura del paso (es decir, dividiendo por la altura en píxeles de la silueta del individuo en el primer fotograma), obteniéndose un resultado en el intervalo  $[0..1]$ .



**Fig. 10.** Amplitud del centro de gravedad.

## 6 Resultados de clasificación

Nuestra experimentación se ha realizado sobre un conjunto de seis individuos (cinco hombres y una mujer), grabándose una única secuencia de vídeo por persona. Los valores calculados para cada una de las cinco características discriminantes, descritas en la sección previa, y para cada uno de los seis individuos participantes en nuestros experimentos, aparecen en la Tabla 1.

Puede observarse que los valores de las características permiten claramente diferenciar (o clasificar) a los individuos. Por ello, se ha usado un sencillo clasificador euclídeo basado en distancias entre clases con el fin de discriminar a los sujetos. Para nuestra base de datos se ha conseguido reconocer al 100% de los individuos.

**Table 1.** Valores de las características discriminantes para los individuos considerados.

Individuo	MaxEnv ( $l$ )	Periodo ( $\tau$ )	MaxAng ( $\theta$ )	AmpCG ( $a$ )	Pendiente ( $\phi$ )
1	0.327	19	0.7413	0.057	-0.3333
2	0.330	14	0.7298	0.037	0.3333
3	0.358	13	0.7759	0.041	0.4999
4	0.451	16	1.0057	0.049	0.0832
5	0.353	17	0.7139	0.057	0.4999
6	0.412	14	0.8918	0.062	0.1110

## 7 Conclusiones y trabajo futuro

Este trabajo presenta un prototipo inicial de sistema de reconocimiento biométrico basado en la postura y el movimiento corporal al caminar (*gait*). En general, el rendimiento de sistemas biométricos reales basados en *gait* está por debajo de lo requerido para su uso como modalidad biométrica [4]. Ello nos ha motivado para comenzar a explorar esta técnica de identificación personal. A pesar de la relativa simplicidad del prototipo desarrollado, consideramos que las características

extraídas en este estudio son fundamentales (y suficientes) para distinguir a dos personas caminando. También se ha observado que un mismo sujeto en condiciones anímicas distintas puede no resultar identificable como él/ella mismo/a. Para llegar a esta conclusión de forma experimental pedimos a un individuo que repitiese su toma pero simulando estar en un estado anímico diferente. Esta repetición produjo valores claramente diferenciados para las características discriminantes consideradas en el estudio.

Un trabajo futuro es ampliar la experimentación para un conjunto mucho mayor de individuos y grabando varias secuencias de vídeo por cada individuo. También es interesante el uso de bases de datos de *gait* más estandarizadas y extensas como por ejemplo la *CASIA Gait Database* [6], donde además las secuencias de vídeo corresponden a sujetos caminando en el exterior (en nuestro caso, todas las secuencias usadas han sido grabadas en interior). Ello permitiría comparar nuestros resultados con los de otros autores sobre los mismos conjuntos de datos. Además, se requiere ampliar el número de características extraídas así como estudiar su capacidad discriminante en la etapa de clasificación, donde se podrían usar otros clasificadores más complejos que requieren de un proceso de entrenamiento previo (como, por ejemplo, los SVM).

## Agradecimientos

Este trabajo ha sido subvencionado por el proyecto de investigación TIN2005-08943-C02-02 del Ministerio de Educación y Ciencia de España.

## References

1. Nixon, M.S., Tan, T.N., Chellappa, R.: Human Identification based on Gait. International series on Biometrics, Springer (2006)
2. Niyogi, S.A., Adelson, E.H.: Analyzing Gait with Spatiotemporal Surfaces, Proc. IEEE Workshop on Nonrigid and Articulated Motion (1994) 64-69
3. Sarkar, S. et al.: The human id gait challenge problem: data sets, performance, and analysis. IEEE Trans. Pattern Analysis and Machine Intelligence 27 (2005) 162-176.
4. Boyd, J.E., Little, J.J.: Biometric Gait Recognition. LNCS 3161, Springer (2005) 19-42
5. Boulgouris, N.V., Chi, Z.X.: Human gait recognition based on matching of body components. Pattern Recognition 40 (2007) 1763-1770
6. Yu, S., Tan, D., Tan, T.: A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. Proc. of the 18th International Conference on Pattern Recognition (ICPR06). Hong Kong, China (2006)
7. Kale, A., et al.: Identification of Humans using Gait, IEEE Trans. Image Processing (2004) 1163-1173