

“ANÁLISIS DE MODO DE CAMINAR DE UNA PERSONA EN TIEMPO REAL UTILIZANDO UN ALGORITMO DE ESQUELETIZACIÓN”

AUTORES:

Pablo Crovetto, jcrovetto@uni.pe
 Daniel Palomino, dpalominop@uni.pe
 Santiago Cortijo, scortijoa@uni.pe.

ASESOR:

Dr. Javier Solano, jsolano@uni.edu.pe

*Centro de Tecnologías de Información y Comunicaciones,
 UNIVERSIDAD NACIONAL DE INGENIERÍA - PERÚ*

Resumen— Mediante una imagen en tiempo real se detecta a una persona caminando de perfil y utilizando un ordenador, el objetivo es afirmar autónomamente si ésta se encuentra caminando, corriendo o cojeando.

El primer procedimiento para conseguir eso se denomina "Foreground detection". Consideramos como foreground la mínima región de la imagen que contiene íntegramente a la persona caminando. Se aplica un detector de movimiento utilizando un filtro exponencial, obteniéndose como resultado una imagen binaria que distingue foreground volviéndose este nuestra región de interés (RDI). El procedimiento siguiente se centra en la eliminación de ruido. La salida binaria obtenida presentará ruido que se elimina mediante los procedimientos de erosión y de dilatación en la imagen.

El tercer procedimiento que se lleva a cabo es el de "esqueletización". Una vez filtrado el ruido de la RDI, se procederá a calcular el contorno de esta en coordenadas polares (con origen en el centroide de la RDI). Para suavizar las irregularidades de esta curva polar, se le tomará Transformada de Fourier filtrando las altas frecuencias. Luego se trazarán segmentos de recta desde el centroide hasta los puntos más alejados del contorno (los cuales corresponden teóricamente a las extremidades de la persona en la toma de video), formándose así la llamada "esqueletización-estrella". Se medirán los ángulos del segmento que va hacia la cabeza y de uno de los segmentos que corresponde a una pierna, ambos respecto de la normal. Finalmente se procede con la detección del ciclo. Los ángulos medidos serán transformados también en el dominio de la frecuencia. Conociendo la velocidad, el grado de inclinación del sujeto respecto a la cámara, y los ángulos de inclinación de sus pies, se determinará si la persona frente a la cámara camina, corre o cojea.

Palabras clave— Análisis de modo de caminar, Detección de Foreground, Esqueletización Estrella, Transformada de Fourier, Detección de ciclo.

Abstract— By real-time imaging a person walking seen from the side is detected through a computer, the goal is to assert autonomously if it is walking, running or limping. The first procedure to achieve that is called "Foreground detection." We consider "foreground" as the minimum image region that contains entirely to the person walking. A motion detector is applied using an exponential filter, resulting in a binary image that distinguishes the "foreground" which is our region of interest (ROI). The following procedure focuses on the elimination of noise. The binary output noise obtained shall be removed by the procedures of erosion and dilation of the image.

The third procedure that is performed is the "skeletonization". When ROI's noise is filtered, the contour is calculated using polar coordinates (with origin at the centroid of the ROI). In order to smooth the irregularities of the polar curve, it will take Fourier transform for high frequency filtering. Then, line segments will be drawn from the centroid to the furthest points of the contour (which theoretically correspond to the extremities of the person on the video), forming the "star-skeletonization". After that takes place

measurement the angles of the segment that goes to the head and one of the segments corresponding to one leg, both with respect to the normal.

Finally detection of the cycle is done. The measured angles will be transformed also in the frequency domain. Knowing the velocity, the inclination of the person from the camera and tilt angles of his feet; it will be determined whether the person facing the camera walks, runs or limps.

1. Introducción

El objetivo de la presente investigación es implementar un meta-algoritmo para detectar y analizar el modo de andar de una persona vista de perfil desde una cámara web, lograr esto en tiempo real y en una resolución alta de imagen (640x480 píxels a más).

La importancia de radica en su posible adaptación para aplicaciones medicas en detección no-supervisada de defectos de caminar en personas utilizando un ordenador y una cámara, lo cual sería práctico y barato respecto a otros sistemas que ya existen para hacer lo mismo.

Este análisis se hace por diversos métodos, como la cronofotografía (múltiples fotos en el tiempo), con cámaras de video (que puede ser con cámaras Infrarrojas), y con sensores activos de movimiento.



Figura 1: Laboratorio de Analisis de Modo de Andar – Inglaterra. fuente: Dr. D. Gordon E. Robertson

2. Materiales y métodos

2.1. **Recopilación de Data.**
 Nos ayudamos con una cámara web para realizar diversos videos de personas caminando, corriendo

y cojeando; todos ellos vistos de perfil. Además, para poder validar nuestros experimentos, grabamos también videos de personas con marcas de color verde en sus extremidades para comparar el resultado de estas con los del filtro de detección a utilizarse. A estos videos con las marcas verdes se les llamará **“data de validación”**.



Figura 2. Se aprecia las pruebas que se hicieron para recompilar data para analizar. A la derecha se aprecia una prueba con marcadores de color verde (data para validación).



Figura 3. Se utilizó una cámara Web, la cual se fijó en una posición elevada.

2.2 Procesamiento de la “data de Validación”.

Para la data de validación implementamos se implementó un filtro de color verde, y un contador de regiones conexas. Mediante ello obtenemos las coordenadas de cada extremidad de la persona caminando, y podemos realizar el análisis de modo de caminar y contrastarlo con el que obtendremos sin la ayuda visual de color verde. Todo el proceso de procesamiento de imágenes se llevo a cabo con la librería OpenCV 2.1 (ver [1])

2.3 Procesamiento de la Data en tiempo real.

En la toma de video a analizar, lo primero que se busca es captar la región de interés, y ya que estamos analizando modo de caminar nos valemos de un detector de “foreground” el cual detecta movimientos significativos de en la toma de video (se debe tener presente que la cámara esta fija), para ello se implementó el siguiente filtro exponencial discreto:

$$y(t) = (1 - \alpha) \cdot y(t - 1) + \alpha \cdot x(t)$$

Donde $y(t)$ es la salida del filtro, $x(t)$ es su entrada, $\alpha \in < 0; 1 >$

Este filtro funciona para series de tiempo discretas. La serie de tiempo que se va a analizar con este filtro es el valor de intensidad de un píxel en la toma de video. El filtró será aplicado a cada píxel de la imagen. Este proceso supone un coste computacional excepcionalmente alto para cualquier ordenador actual, por lo cual se implementó en CUDA (que son librerías de programación paralela, para tarjetas de video NVIDIA, [2]).

Cada vez que un nuevo valor de intensidad de píxel genere una desviación muy seria, entonces se considerara a este como “Píxel de Foreground” (en la práctica, cada vez que un nuevo la intensidad de un nuevo píxel se salga del radio de 2 desviaciones estándar respecto al promedio en el tiempo del valor de intensidad promedio de dicho píxel)

Conformamos una imagen binaria, la cual nos indicará si un píxel es “foreground” o “background”. Sobre esta imagen se aplica un proceso de Erosión y otro de dilatación con el fin de eliminar ruidos (huecos en la imagen de la persona caminando, y manchas en el exterior de esta misma). Luego se le aplica un filtro de Canny para obtener sus bordes. Se calcula el centroide del contorno de la persona moviéndose. A este procedimiento se le conoce como “Esqueletización Estrella” [5]



Figura 3: Resultado optimo de una detección de foreground, con eliminación de ruido(imagen tomada de [4])

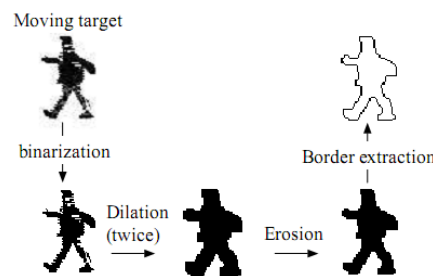


Figura 4. Bosquejo del procedimiento para la esqueletización estrella (imagen tomada de [5])

Se calcula el centroide del contorno de la persona caminante:

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = \frac{\sum_{i=1}^N \begin{bmatrix} X_i \\ Y_i \end{bmatrix}}{N}$$

Donde X_i, Y_i son los puntos de este contorno, y N es el número total de puntos de contorno.

Mediante un sistema de coordenadas polares, con origen en el centroide hallado, se ubica la región de la cabeza, y la región del pie izquierdo (respecto de la perspectiva de la cámara). Se hallan 2 ángulos mediante los cuales se realizará el análisis de modo de caminar (ver figura 5.)

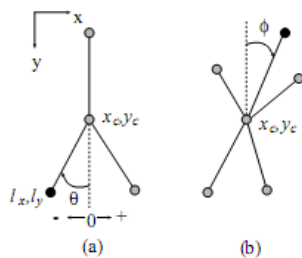


Figura 5.. Se medirá el ángulo de la normal respecto al segmento del centroide y la cabeza, además se medirá también el ángulo respecto a la normal del centroide y la pierna izquierda, (imagen tomada de [5])

A los ángulos medidos θ y Φ se les analizará en el dominio de la frecuencia, para detectar sus ciclos y para compararlos con la data tomada, para así poder distinguir el modo de caminar

3. Resultados y discusión

3.1 Resultados del procesamiento de la “Data de Validación”

Resultados del filtro de colores, en el cual se detecta, numera y calcula centroide de las regiones verdes en toma de video.



Figura 6: Se detectan y diferencian las partes verdes en la imagen de validación. En la parte izquierda se marcan con colores diferentes las regiones que nos interesan en las mediciones.

Se aprecia que el centroide está marcado, así como todas las extremidades. (ver figura 6.)

3.2 Resultados del procesamiento de la data en tiempo Real.

Se muestra el detector de “Foreground” y el cálculo del centroide en 2 tomas de video. (ver figuras 7 y 8)



Figura 7: Prueba del detector de “foreground” en un video de la “data de validacion”



Figura 8: Prueba del detector de Foreground en tiempo Real

El detector de “Foreground” es más que un simple detector de movimientos, ya que es capaz de no-filtrar movimientos insignificantes (como el cambio de brillo, o una mosca), y posee un coste computacional mucho más alto que un detector de movimiento común.

4. Conclusiones

Es factible analizar el modo de caminar de una persona en tiempo real y utilizando solamente un ordenador con una cámara web.

Las ventajas que presenta esto son la de menor coste del sistema respecto a los ya existentes, y mayor accesibilidad al público en general, pero a su vez presenta desventajas como la sensibilidad a la luz excesiva o muy tenue, lo cual obliga al acondicionamiento del ambiente en el cual se va a aplicar (aunque dicho acondicionamiento no es algo riguroso, solo consiste en conseguir un ambiente con una iluminación adecuada, en la que una persona camine, y fijar una cámara web desde una posición tal que tenga una vista panorámica del evento)

La presente investigación presenta potencial de aplicación en la medicina (en la detección no supervisada de defectos de caminar), así como en la deportiva (análisis de modo de correr en atletas, futbolistas, gimnastas, etc), entre otras.

5. Agradecimientos

Agradecemos al CTIC-UNI por brindarnos las facilidades necesarias para realizar tranquilamente nuestra investigación, así como al Dr. Javier Solano por su invaluable asesoría.

6. Referencias

Referencia a Libros:

- [1] Gary Bradski, Adrian Kaehler "Learning OpenCV" O'Reilly Media
- [2] NVIDIA "Nvidia CUDA Reference Manual v.3.0"

Referencia a Artículo:

- [3] Liang Wang, Weiming Hu, Tieniu Tan – "Recent Developments in Human Motion Analysis", Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China.
- [4] Sudeep Sarkar, P. Jonathon Phillips, Zongyi Liu, Isidro Robledo Vega, Patrick Grother, Kevin W. Bowyer "The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis" IEEE Transactions on Pattern Analysis and Machine Intelligence Vol.27, No 2, February 2005
- [5] Hironobu Fujiyoshi, Alan J. Lipton "Real-time human motion analysis by image skeletonization" The Robotics Institute. Carnegie Mellon University.

