

Análisis y Visualización Interactiva de Datos Aplicado al Estudio del Movimiento Humano por Videos de Actividad

Rodriguez M. Wilmar, Automation Engineering and Industrial Computing Master's degree, Universidad de Oviedo, 2020

*Contact Author: uo278841@uniovi.es

Resumen—El estudio de los movimientos humanos es de relevante importancia en la medicina e ingeniería moderna, con aplicaciones en el estudio de la interacción hombre-maquina, en robótica o investigaciones biomédicas; este trabajo propone un algoritmo completo y métodos para el acercamiento al análisis del movimiento articular y biomecánico, usando herramientas de visión artificial, análisis de imagen y el procesamiento digital de señales, a partir de videos de una o varias cámaras, los resultados preliminares muestran que la técnica es válida y de desempeño comparable con el uso de sensores MEMS, mientras mantiene un nivel de abstracción bajo e interactivo de las señales e indicadores mas importantes. (Abstract)

Keywords— *biomecánica, deep learning, movimiento humano, movimiento articular, visualización de datos, visión artificial.*

INTRODUCCION

El estudio de los movimientos humanos, al igual que sucede con otros acercamientos de la ciencia de datos, se realiza generalmente en 3 grandes etapas [1]: extracción de características, correspondencia y análisis de alto nivel; en cada etapa los problemas son vastos, debido a ello, la extracción de características se detalla en este estudio con un nivel mas granular.

Para la etapa de detección, se usan algoritmos de aprendizaje profundo DL (Deep learning) para marcar articulaciones, que gracias al buen desarrollo actual de librerías y hardware, así como la validada experiencia en innumerables proyectos industriales usando DL, forman una base firme, sustentable, y con procesamiento alcanzable por ordenadores de escritorio o portátiles convencionales. No obstante, la detección es la etapa que mayor demanda de procesamiento tiene en todo el algoritmo.

El resto del algoritmo tiene un enfoque de análisis unipersonal (no se contemplan actividades en grupo o multiples instancais de personas), que podrían complicar el algoritmo. Con ello se prosigue al análisis dinámico, la extracción de señales secundarias o complementarias y el tratamiento de video y de visualizaciones correspondientes a las actividades humanas, y termina con una propuesta de visualización interactiva usando la herramienta de Bokeh server. El desarrollo de la aplicación general y librerías auxiliares se ha realizado con el uso de Python 3.7, Numpy 1.18, SciPy 1.4, Pandas 1.0.3, Pytorch 1.4, OpenCV 4.02, OpenVINO 2020.1 y Bokeh 2.0.2.

I. DETECCION DE ARTICULACIONES

En la fase de detección se realiza el análisis de video extrayendo fotogramas *frames*, para su análisis independiente, este estudio se basará en los hallazgos realizados por dos grupos de trabajo en la detección de articulaciones y sus correspondientes modelos de redes neuronales de aprendizaje profundo (*Deep Learning*) DL para la detección de articulaciones. Esta detección es diferente respecto del acercamiento general de detección humana por cuadros de

enmarcado (*bounding boxes*), usado en aplicaciones de recuento de visitantes en retail o seguimiento en seguridad corporativa, entre otros usos. Uno de los modelos más rápidos disponibles en la actualidad [3] es el realizado por Z. Cao y T. Simon [3][2] usando una metodología de detección de partes corporales, articulaciones y un algoritmo de tipo agrupamiento bottom-up para enlazar cuerpos en una imagen; Las características importantes de este modelo y que serán explotadas son: la capacidad de detectar múltiples cuerpos en una imagen, la segmentación de partes del cuerpo y la velocidad de detección del algoritmo.

Un segundo modelo, propuesto por D. Osokin [4], utiliza la arquitectura de maquinas convolucionales, un diseño de red neuronal con un acercamiento top-down multifase y con información del resto del cuerpo, que permite una mejora de fiabilidad importante al diferenciar con tasas de error del 5% los lados de izquierda y derecha de cada articulación; este modelo es caracterizado por el establecimiento de mapas de probabilidad, precisión a priori mayor al modelo de Z. Cao y una mejor escalabilidad a la alta resolución, sin embargo como inconvenientes tiene una menor velocidad de detección, mayor consumo de procesamiento, y limitación de la detección a un solo individuo.

Los dos modelos tienen como salida o bien la segmentación (áreas probables de tener una articulación – fig 1) o mapas de calor de probabilidad (fig 2). La naturaleza de ambos métodos es la de una imagen de salida, en el caso inicial como imagen binaria y en el segundo como un mapa de calor.



Fig. 1. Imagen de ejemplo de detección de áreas de articulaciones



Fig. 2. Imagen de ejemplo de mapa de calor en codo izquierdo

Usando la información en imágenes de las articulaciones, se pasa a la detección de puntos, un paso en el cual se busca llevar las áreas de detección o los mapas de probabilidad a un punto único de coordenadas (x,y) en el plano de imagen. Según la fuente de datos pueden usarse diferentes acercamientos a este problema como la detección de máximos usada en la transformada de Hugh (opción seleccionada para el algoritmo), la búsqueda de centroides geométricos, calculo de centroides ponderado, entre otros.



Fig. 3. Ejemplo de detección del punto de codo izquierdo

A partir de los puntos década articulación y su identificación se puede construir un modelo de coordenadas *skeleton*, en el espacio de coordenadas de la imagen, en este punto se pasa de un problema de procesamiento de imágenes a la resolución de problemas geométricos con dichos puntos. Una imagen se convierte en una instancia de *skeleton* con sus coordenadas y conexiones.

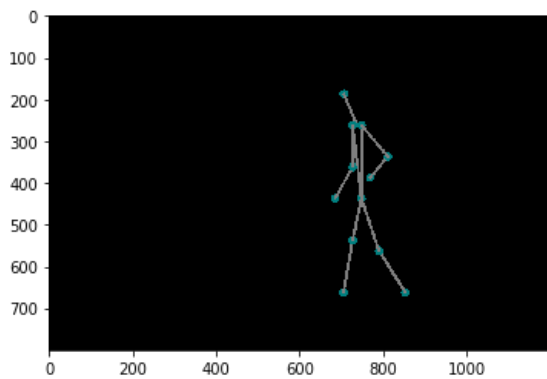


Fig. 4. Imagen reconstrucción de skeleton

II. EXTRACCIÓN DE CARACTERÍSTICAS

Los análisis dinámicos han sido ampliamente estudiados en facultades de ingeniería biomédica, centros deportivos, hospitales y centros de rehabilitación esta sección que ver con el estudio de los movimientos y la biomecánica propiamente dicha, usando herramientas del procesamiento digital de señales y de imágenes, en este último particular, se han hecho

en trabajos como el de M. Evans [3]. El objetivo en términos prácticos es construir una tabla de ubicaciones para cada instancia de *skeleton*, lo que traduce un video en una tabla de datos de ubicaciones de cada articulación, cuyas consultas corresponderían con señales de desplazamiento.

A. Las señales de ubicación

El desplazamiento del cuerpo corresponde al análisis mas simple, basta con buscar un punto representativo del *skeleton* en el sistema de coordenadas y reconstruir su señal directa de la tabla, para la ubicación del punto puede optarse por varios acercamientos como el cálculo del centroide geométrico de cuerpo, la ubicación del centro de masa [5] o el uso de una articulación como punto de referencia, cada uno de los acercamientos tiene consecuencias en la carga de procesamiento o en la interpretación de los resultados. En el ámbito de este estudio para favorecer la simplicidad del algoritmo, se asumirá que el cuerpo se referencia a partir de la ubicación del pecho, una coordenada de fácil reconstrucción geométrica a partir de las coordenadas de los hombros y la cabeza.

La señal de ubicación del cuerpo tomado a partir de una imagen (abstraído como un punto en el espacio), tiene dos componentes de acuerdo con (1), correspondientes a la posición horizontal (x) y vertical (y). En sistemas con múltiples cámaras calibradas, este problema se convierte en uno de tres componentes, correspondientes con la ubicación instantánea en las coordenadas (x, y, z) del espacio de movimiento.

$$posJ_i = \begin{bmatrix} xf_i \\ yf_i \\ 1 \end{bmatrix} \approx \begin{bmatrix} ux \\ uy \\ u \end{bmatrix} = M_{3 \times 4} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

B. El vector de desplazamiento

Habiendo obtenido las señales de ubicación, tanto del cuerpo como de cada una de las articulaciones, se observa una curva de desplazamiento con datos preliminares interesantes, por ejemplo para el cuerpo (fig 5), esta curva denota que es posible filtrar y posteriormente derivar numéricamente para calcular la velocidad de desplazamiento, para ello se usa la aproximación numérica hacia adelante y hacia atrás (2) para hallar la curva de velocidad, el resultado también será en dos dimensiones para imágenes simples (x,y)

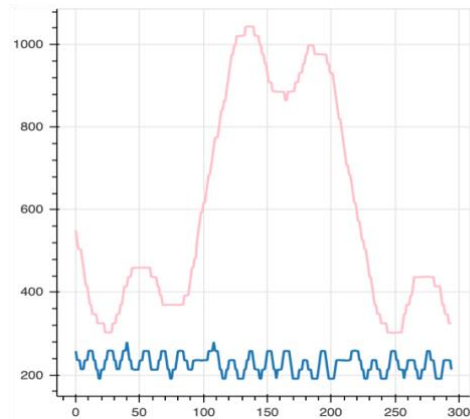


Fig. 5. Grafica (pixels/frame) de ubicación del cuerpo (eje x azul; eje y rosa)

$$f'(a) \approx \frac{1}{2} \left(\frac{f(a+h) - f(a)}{h} + \frac{f(a) - f(a-h)}{h} \right) = \frac{f(a+h) - f(a-h)}{2h} \quad (2)$$

Una propiedad importante emerge del proceso de derivación es el vector de desplazamiento, que tiene como componentes la dirección de movimiento en cada una de las direcciones, y que tiene un módulo y ángulo desde el origen, que pueden ser calculados fácilmente por geometría básica; en el caso de dos dimensiones, el módulo se simplifica al hallado de la hipotenusa entre las magnitudes de desplazamiento en cada punto. En fig 6 se muestra la gráfica con los desplazamientos en x (rosa), y (azul) y módulo (verde) para el estudio de movimiento de ejemplo usado anteriormente (fig 5).

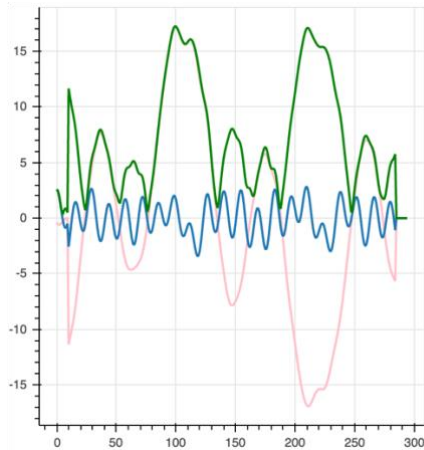


Fig. 6. Gráfica de velocidad (px/frame/frame) en ejes x e y, en verde módulo del vector de desplazamiento

La magnitud de desplazamiento del cuerpo está directamente relacionada con el índice de actividad y el trabajo realizado a priori por cualquier actividad humana estudiada, por lo que es un indicador global idóneo para mantener en cualquier estudio de movimiento.

C. El desplazamiento Relativo

El desplazamiento del cuerpo se ha obtenido a partir del movimiento global, sin embargo, muy poca información se obtiene al medir este indicador en articulaciones periféricas como las manos o los pies, debido a que contienen implícitamente la actuación de las otras articulaciones o del movimiento absoluto del cuerpo, lo cual actúa como ruido en dichas señales.

En términos simples, el trabajo efectivo realizado por una parte periférica del cuerpo esta definido por el movimiento realizado desde la articulación de nivel inmediatamente superior, así, el trabajo realizado por el antebrazo se define por el movimiento de la muñeca respecto del codo, y el trabajo del brazo se define por el movimiento del codo respecto del hombro. Esta trivialización, sin embargo, es puramente abstracta, por ejemplo, el trabajo fisiológico de los músculos del brazo puede derivarse también de movimientos de la muñeca sin existir movimiento del codo; este concepto de cooperación extendida no es tenido en cuenta de forma directa en el ámbito de este estudio, aunque una extensión multicámara y con información de las manos y pies completa si implícitamente incluye las magnitudes de ese tipo de trabajo.

El desplazamiento relativo se define fácilmente a partir de los vectores de desplazamiento absolutos ya medidos, de acuerdo con (3), basados en el concepto vectorial de cambio de base. Esta operación da como resultado la señal de movimiento relativo de cada articulación, es de fácil

comprensión y de aplicación inmediata en muchos estudios de descripción de movimiento que detallen el desempeño de cada articulación.

$$fX_{muñeca} = |V_{muñeca}(x, y)| - |V_{codo}(x, y)| \quad (3)$$

D. Indicadores resumen

A partir de las señales analizadas y calculadas, se obtiene un primer grupo de indicadores de análisis temporal para cada articulación, que podrían usarse para el estudio de la clasificación de movimientos o como base estadística para evaluar el desempeño de ciertas actividades rutinarias. Los cuatro indicadores son: velocidad máxima, velocidad promedio, desplazamiento neto, desplazamiento relativo.

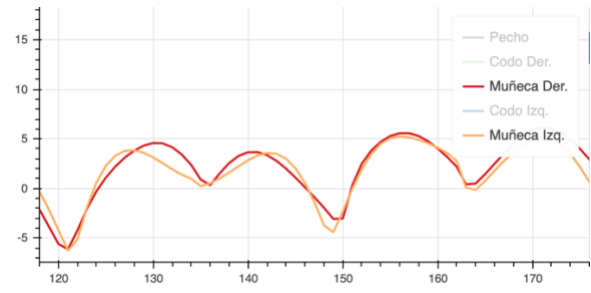


Fig. 7. Gráfica de ejemplo de velocidad relativa en muñecas

III. ANÁLISIS DE ALTO NIVEL

El análisis de alto nivel en un primer acercamiento incluye la aplicación de técnicas del procesamiento digital de señales a las señales ya calculadas, algunos ejemplos de estos tratamientos de alto nivel que den información importante sobre la dinámica del movimiento son:

- La estimación de la aceleración de movimiento articular
- Estimación de ángulos de apertura y cierre de articulaciones
- Cálculo del trabajo realizado por los músculos de la articulación
- Cálculo de velocidad angular
- Análisis frecuencial del movimiento articular
- Índice de desplazamiento global

Se detallarán a continuación solamente acercamientos al análisis frecuencial y al índice de desplazamiento global, dejando a estudios posteriores el cálculo y verificación cruzada de técnicas para llegar a estos indicadores.

A. Índice de desplazamiento global

Corresponde al desplazamiento total de cada articulación y podría ser importante en estudios de postura, extensión máxima, optimización de movimientos, etc. Su cálculo proviene de la integral del vector de desplazamiento relativo de cada articulación que se simplifica a (4) debido a la naturaleza discreta de la señal.

$$X_{muñeca} = \sum_{i=0}^n |V_{muñeca}(x, y)| - \sum_{i=0}^n |V_{codo}(x, y)| \quad (4)$$

La forma mas fácil de mostrar este resultado es a través de un mapa de calor, que además de mostrar el valor (en escala cromática) del desplazamiento, encasilla el valor en una coordenada fácil de localizar para el lector.

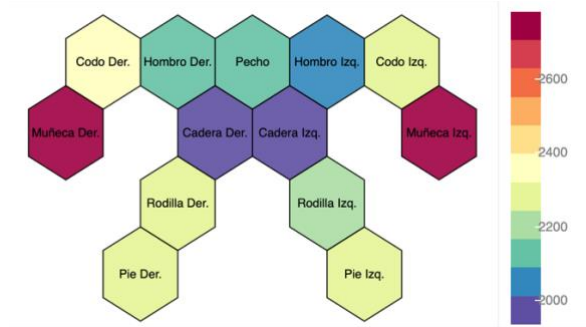


Fig. 8. Mapa de calor de desplazamiento absoluto articular

B. Análisis frecuencial

Con las señales del vector de desplazamiento se puede calcular la potencia en cada banda de frecuencia a partir de la transformada de Fourier con el algoritmo de FFT como se muestra en la fig 9.

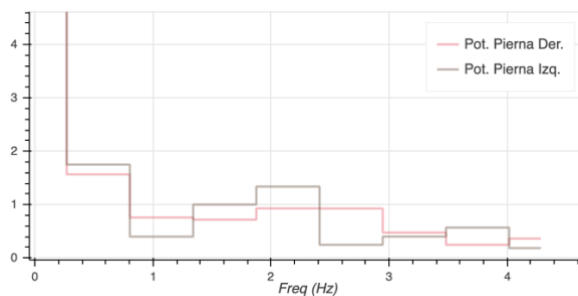


Fig. 9. Gráfica de ejemplo de potencia FFT por rangos de frecuencia

Para una mejor visualización de los indicadores de frecuencia se propone un mapa de calor (heatmap) de todas las articulaciones, en el cual se puede fácilmente analizar la densidad de movimiento de cada articulación por componentes frecuenciales.

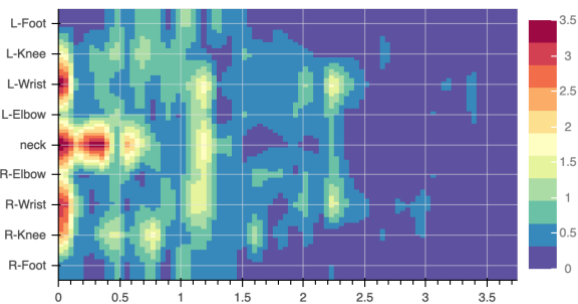


Fig. 10. Mapa de calor de potencias frecuenciales para todo el cuerpo

DISCUSIÓN

Los resultados finales del algoritmo muestran posible el cálculo de sistemas dinámicos completos con datos de velocidad y movimiento, al igual que la toma de datos a partir de sensores cinemáticos MEMS, lo cual hace compatible ambos acercamientos o podría usarse como información complementaria para la reconstrucción de estudios mas complejos y precisos.

Los errores descubiertos en las etapas de toma de datos son mejorables realizando un postprocesado de datos incluyendo filtrado de señales, aplicación de modelos multiescala o uso de modelos complementarios como sensores virtuales, estudios posteriores podrían utilizar los modelos de DL propuestos para evaluar su efectividad o realizar comparaciones con modelos desarrollados a posteriori.

La capacidad de procesamiento importante que requieren los modelos de DL, en particular el modelo de D. Osokin [4], de maquinas convolucionales, en conjunto con el múltiple procesamiento de imagen que debe realizarse a cada *frame*, hace que el sistema tenga una gran demanda de recursos de procesamiento, que pueden ser un problema para el procesamiento en tiempo real. En la práctica, aunque se han probado modelos simplificados de ambas redes con velocidades de hasta 190fps [3][4] al desarrollarse el algoritmo completo de este estudio no se pueden alcanzar mas de 4fps en videos de baja resolución usando un procesador Intel® M5.

REFERENCIAS

- [1] Vasconcelos, Maria & Tavares, Joao. (2008). Human Motion Analysis: Methodologies and Applications. Conference Paper https://www.researchgate.net/publication/37650298_Human_Motion_Analysis_Methodologies_and_Applications
- [2] Z. Cao, T. Simon, S. Wei and Y. Sheikh, "Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 1302-1310, doi: 10.1109/CVPR.2017.143
- [3] M. Evans, S. Colyer, D. Cosker and A. Salo, "Foot Contact Timings and Step Length for Sprint Training," 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, 2018, pp. 1652-1660, doi: 10.1109/WACV.2018.00184.
- [4] Osokin, Daniil. (2019). Global Context for Convolutional Pose Machines, Intel Laboratories Russia, jun, 2019.
- [5] F. Martínez, F. Gómez, and E. Romero. "Análisis De Vídeo Para Estimación Del Movimiento Humano: Una Revisión." Revista Med, (2009) 17(1), pp. 95-106.