

Red Neuronal Lagrangiana para Modelar Bipedestación Sagital Usando Visión de un Celular

Francisco Pérez-Encina
Dept. Posgrado
Tecnológico Nacional de México,
Instituto Tecnológico de Nuevo León
Guadalupe, México. Email:
mg15480429@nuevoleon.tecnm.mx

Dulce Martinez-Peon
Dept. Eléctrica y Electrónica
Tecnológico Nacional de México,
Instituto Tecnológico de Nuevo León
Guadalupe, México. Email:
dulce.mp@nuevoleon.tecnm.mx

Vicente Parra-Vega
Posgrado de Robótica y Manufactura
Avanzada, Centro de Investigación y Estudios
Avanzados (Cinvestav), Unidad Saltillo
México. Email: vparra@cinvestav.mx

Resumen—Los protocolos de entrenamiento o rehabilitación motora requieren de modelos biomecánicos que dependen de múltiples variables y parámetros. Existen plataformas completas para tal propósito las cuales requieren de una gran instrumentación y poder de procesamiento, adicional a su alto costo, lo que los hace imprácticos e inasequibles a la mayoría de usuarios clínicos. Para el caso de bipedestación dinámica, o la marcha, en el plano sagital, es crucial tanto el sistema de visión utilizado como el modelo cinemático de cadera y rodilla, además del algoritmo de procesamiento que entrega la medición o estimación de los ángulos involucrados. En este artículo, se obtienen la velocidad y la posición de los miembros inferiores de un sujeto usando un celular de gama baja, las librerías de MediaPipe, la implementación de una Red Neuronal Lagrangiana (RNL) y un derivador robusto, Levant, para la velocidad angular. La RNL resuelve la cinemática diferencial inversa como si fuera un robot redundante de 7 GdL. Resultados muestran un RMSE de 0,024 con una desviación estándar de $\sigma = 19,68$ para la cadera, mientras que se logra un RMSE de 0,0032 y $\sigma = 6,19$ para la rodilla. Este sistema tiene la potencial aplicación práctica como laboratorio de marcha portátil y sin entrenamiento de datos.

Palabras clave: Bipedestación dinámica, Sistema de visión monocular, Red Neuronal Lagrangiana.

I. INTRODUCCIÓN

Los sistemas de captura de movimiento, o MOCAP por su nombre en inglés (motion capture systems) [1], han impactado industrias diversas dado que son sistemas de medición no invasivos que proveen medición o estimación de variables del objeto o sujeto medido [2]. En el caso de la biomecánica, el interés es el estudio del sujeto quasi-estático o en movimiento, lo que típicamente involucra medir los ángulos de las articulaciones, que no son otra cosa más que el ángulo relativo entre la proyección de los segmentos óseos en cada articulación. Posteriormente, se requieren modelos biomecánicos de la estructura ósea bajo estudio para inferir un fenómeno en biomecánica. Con ello, los MOCAP son poderosas herramientas multicámaras estéreo para el estudio preciso de la biomecánica, [3], [4], incluyendo áreas clínicas, diversos diagnósticos, tratamientos y rehabilitación de diversas enfermedades, como Parkinson o infartos cerebrovasculares que afectan el movimiento de las extremidades. Cuando el interés son los miembros inferiores, el estudio de pararse y de caminar se conoce como bipedestación dinámica, [5], al

involucrar las dos piernas. En esta área, se utilizan adicionalmente las llamadas placas de fuerza sobre las que el paciente camina. Con ello, el MOCAP y las placas de fuerza cuantifican de manera aproximada ciertos parámetros del equilibrio y balanceo del cuerpo en una variedad de patrones biomecánicos. Los vectores de fuerza resultantes son a su vez dependiente del modelo biomecánico y sus variables angulares, por lo que lo anterior es determinante para la interpretación del caminado, [6]. Los MOCAP profesionales como Vicon, [7] y Optitrack, [8], los cuales utilizan múltiples cámaras infrarrojas estéreo, requieren de toda una infraestructura física y humana para funcionar adecuadamente, lo que los hace imprácticos y alcanzables solamente a grandes instalaciones hospitalarias y proyectos que subvencionaban grandes recursos. Estos sistemas requieren accesorios específicos como trajes, marcadores y costoso hardware y software de análisis de movimiento de marcadores deformables (que no mantienen la distancia entre ellos, al ser colocados sobre el traje o piel del sujeto). Si bien Vicon se ha convertido en un estándar de la comunidad médica e investigación clínica, tiene la enorme desventaja de ser muy especializado y costoso, situación que exige a los investigadores la creación de soluciones alternas, viables y factibles.

Esto ha ocasionado la creación de soluciones tecnológicas asequibles como cámaras estéreo de bajas prestaciones y bajo costo, como el Kinect v2 de Microsoft, [9]. Comparando el caminado de un voluntario en condiciones de iluminación controlada, en un rango de 200 a 300 luxes, [10] se concluyeron limitaciones importantes del Kinect V2, aunque requiriendo más investigación para determinar el potencial de este sistema. Múltiples sistemas similares existen hoy en el mercado, [3], para detectar de manera aproximada, y limitada, la biomecánica del cuerpo humano, incluyendo la Astra Pro de Orbec, [11]. Esto ha impulsado la investigación de software potenciado con algoritmos de inteligencia artificial, tales como OpenPose y MediaPipe, [12], [13], que con datos de menor resolución y repetibilidad detectan y estiman variables biomecánicas utilizando visión monocular de un dispositivo móvil y procesándolos en PC para reconstruir visión estéreo que infiere en la coordenada de profundidad y, por lo tanto, el mapeo completo 3D de cada pixel. Sin embargo, esto

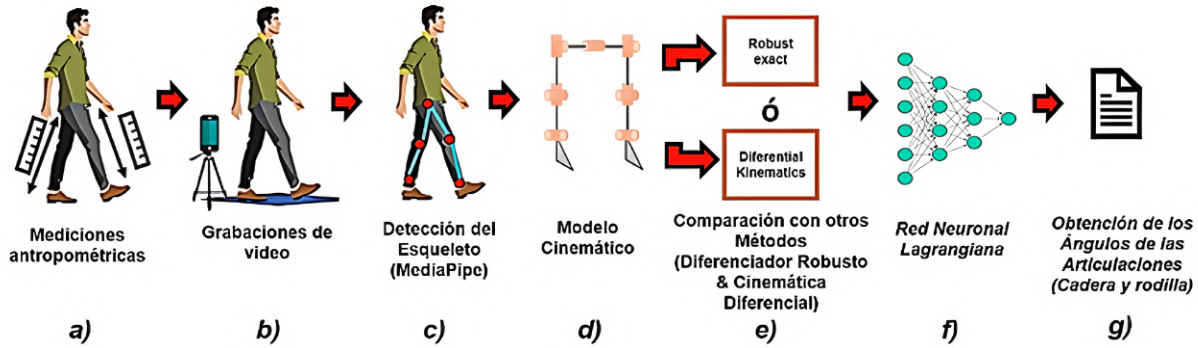


Figura 1. Metodología empleada para la obtención de los ángulos de la cadera y rodilla.

aún están en etapa de desarrollo aplicaciones específicas basadas en MediaPipe para el análisis de la caminata, que resultan útiles para la detección de anomalías en los patrones de caminado en base en un análisis geométrico, [14]. La inferencia de profundidad de MediaPipe es a su vez una desventaja y una ventaja que hay que sintetizar caso por caso. Por un lado, es claro que visión monocular con un celular es 2D, por lo que inferir 3D en realidad termina siendo 2.5D, proveyendo el plano tridimensional inexacto, con cierto error, pero explota la increíble penetración de los celulares en el mercado. Azhand y colaboradores propusieron el modelo VNect con métodos de aprendizaje de máquina, o ML por su acrónimo en inglés de Machine Learning, para estimar 3D a partir de datos 2D. Tanto OpenPose como MediaPipe incluyen librerías para reconstruir de manera aproximada y simplificada el esqueleto articulado, proporcionando la posición cartesiana de las articulaciones; sin embargo, es necesario programar adicionalmente modelos cinemáticos para calcular variables angulares de las articulaciones, esenciales para analizar la caminata o pose de extremidades. Aprendizaje profundo, o DL del acrónimo Deep Learning, se ha propuesto para tal efecto, [15], con la desventaja que DL requiere intenso entrenamiento supervisado. Por lo tanto, el problema de calcular las coordenadas Cartesianas tridimensionales de los ejes de articulación y los ángulos correspondientes de extremidades inferiores, con tecnologías asequibles, sigue siendo un problema abierto con herramientas sin entrenamiento.

Contribución En este artículo proponemos una plataforma para determinar ángulos de un modelo simplificado de miembro inferior, implementable con una tecnología asequible de un celular de gama baja; el propósito es obtener los ángulos de la cadena cinemática articulada que representa las extremidades inferiores durante la marcha. Para ello, proponemos la implementación de un algoritmo RNL y un modelo cinemático directo que corresponde al modelo usual biomecánico de los miembros inferiores. La entrada al esquema son posiciones y velocidades Cartesianas 3D que sintetiza MediaPipe sin entrenamiento y se propone un derivador robusto para calcular la velocidad, [16]. El derivador robusto converge asintóticamente,

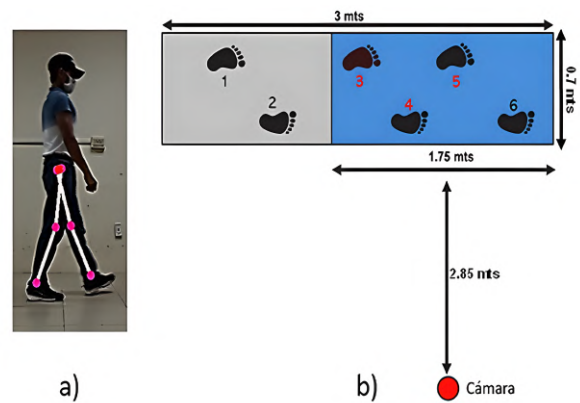


Figura 2. a) Detección de los miembros inferiores en el plano sagital por MediaPipe. b) Área de pruebas de la caminata y distribución del equipo.

por lo que podemos validar asintóticamente la solución de cinemática inversa directa y diferencial con la RNL, lo cual representa los ángulos de las extremidades inferiores. Para comparar nuestros resultados, proveemos datos sintéticos al algoritmo, con lo cual se muestra la utilidad práctica de nuestra propuesta.

II. METODOLOGÍA

La Figura 1 muestra las siete fases que comprenden la metodología desarrollada.

1. Obtención de las medidas antropométricas del sujeto, en particular las medidas de la cadera y la altura de la cadera al tobillo. Posteriormente, se preparó un área como la que se observa en la Figura 2.
2. Las grabaciones de los vídeos. La cámara se mantuvo a 0.65 metros colocada verticalmente en un tripié sin obstáculos y con la iluminación ambiental controlada ($364 \text{ luxes} \pm 3.39$). El sujeto realizó seis pasos de izquierda a derecha frente a la cámara en el plano sagital, donde los pasos del tres al seis (región azul), son los que son grabados, ver Figura 3.

3. Procesamiento en MediaPipe que detecta el esqueleto simplificado del sujeto, proporcionando coordenadas Cartesianas 3D de cada articulación de huesos adyacentes de extremidades inferiores.
4. Cálculo de la cinemática inversa que proporciona la pose angular (ángulos relativos en los ejes de rotación).
5. Cálculo de la velocidad utilizando el derivador robusto (el cual no se ha reportado en la literatura de esta temática), [16]. Cuando se utilizan datos sintéticos, en este paso se resuelve la cinemática inversa diferencia, siempre que todo sea conocido teóricamente (eso significan los datos sintéticos).
6. Se ejecuta la RNL cuya entrada es el vector de pose absoluto a un referencial base al calcular los ángulos de la cadera y la rodilla, con base en las medidas antropométricas del sujeto y el mapeo de las velocidades mediante el Jacobiano geométrico. Se repite en cada nueva imagen, ver el algoritmo RNL del Cuadro 1.
7. Documentación anonimizada del sujeto, ciclos de marcha y parámetros.

II-A. Materiales

Para las grabaciones de vídeo se utilizó un celular de gama baja como es Xiaomi Redmi 9, el cual cuenta con una cámara de 13 Mpx, una resolución de vídeo de 1080 x 1920, 30 fps y 4GB de RAM. Mientras que para el procesamiento de los datos se utilizó una computadora Dell inspiron 14 con un procesador intel Core i5 de 7ma generación y 8 GB de RAM.

II-B. Diseño del Experimento

Un voluntario (sujeto) sano participó en los experimentos, quien no contaba con antecedentes de ninguna discapacidad

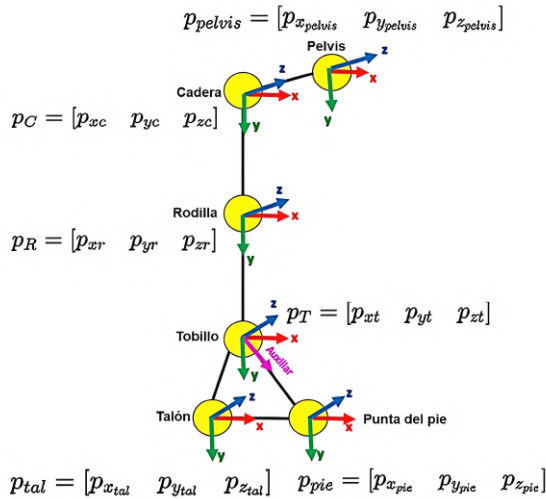


Figura 3. Diagrama de las articulaciones de la pierna derecha. Donde p_C , p_R y p_T son los vectores que contienen las posiciones espaciales de la cadera, rodilla y tobillo, respectivamente.

motora ni neurológica, cumpliéndose además con la Declaración de Helsinki y habiendo firmado el Consentimiento Informado, se le instruyó en dar seis pasos empezando por el pie izquierdo. Se solicitó en cada repetición que el talón lo colocará en una marca en el suelo, si esto no se llevaba de esta forma se descartaba esa repetición. La Figura 2, muestra la cantidad de pasos que el sujeto realizó, siendo los primeros dos (área gris), los pasos que el sujeto dio para entrar en cadencia, los cuatro últimos (área azul) son los que fueron grabados y son los utilizados en el procesamiento posterior. El proceso se deja de grabar al finalizar al sexto paso y el sujeto regresa al inicio para una nueva repetición, hasta completar 10 repeticiones.

II-C. Método

Los vídeos de la caminata se postprocesan para extraer las imágenes correspondientes a la ejecución de los pasos 3, 4 y 5 (ver Figura 2), como entrada a la librería de MediaPipe, que entrega las posiciones Cartesianas 3D de cada articulación. Por ejemplo, para el caso de la cadera, MediaPipe proporciona su posición respecto al marco de referencia de la pelvis, como se muestra en la Figura 3. Para fines ilustrativos de este proyecto, se consideran solamente los ángulos de la cadera y rodilla de la pierna derecha, correspondientes al el paso 4.

II-C1. Reconstrucción del esqueleto: MediaPipe detecta 33 puntos del cuerpo humano en imágenes y vídeos, proporcionando una estimación de las coordenadas cartesianas del esqueleto, [13]. MediaPipe reporta que su algoritmo supervisado fue entrenada con una base de datos de más de 200,000 imágenes etiquetadas de diferentes personas realizando diferentes actividades físicas, [17]. Así, MediaPipe entrega las posiciones de la cadera p_C , rodilla p_R y tobillo p_T (ver Figura 3), siendo el tobillo la articulación considerada como el efector final, en analogía de un robot. Sin embargo, el vector de coordenadas Cartesianas $X \in R^3$ requiere no solo de la posición espacial, sino también de la orientación. Para esto se modela un marco de referencia para el tobillo de la siguiente manera: Para el eje de dicho marco, se calcula un vector y_T , el cual se obtiene mediante el cálculo de la distancia Euclideana $d_1 = p_T - p_R$ entre la posición del tobillo $p_T = [p_{xt}, p_{yt}, p_{zt}]^T$ y la rodilla $p_R = [p_{xr}, p_{yr}, p_{zr}]^T$. Para el eje z, se obtiene el vector z , para ello se obtuvo un vector auxiliar u_{aux} conformado por la línea del empeine que va a la punta del pie, calculándose la distancia $d_2 = p_T - p_{pie}$ entre el tobillo y de la punta del pie $p_{pie} = [p_{x_{pie}}, p_{y_{pie}}, p_{z_{pie}}]^T$, posteriormente, se normalizan las diferencias entre ejes y la distancia para hacerlos unitarios. Una vez teniendo el vector auxiliar, se calcularon los vectores x y z , mediante producto cruz del vector y , utilizando el vector u_{aux} . Para la orientación, se calcularon los tres ángulos de Euler correspondientes a los vectores directores de rotación x, y, z , llamados respectivamente alabeo, cabeceo y guiñada (pitch, roll, yaw en inglés), a partir de las matrices de rotación, [18], obteniéndose:

$$\theta = \arctan \left(\frac{-r_{31}}{\sqrt{r_{11}^2 + r_{21}^2}} \right) \quad (1)$$

$$\phi = \arctan 2 \left(\frac{r_{32}}{c_\theta}, \frac{r_{33}}{c_\theta} \right) \quad (2)$$

$$\psi = \arctan 2 \left(\frac{r_{21}}{c_\theta}, \frac{r_{11}}{c_\theta} \right) \quad (3)$$

$$r = [p_{xt} \ p_{yt} \ p_{zt} \ \theta \ \phi \ \psi]^T \quad (4)$$

donde $r = [p_t, \alpha]^T$ es el vector correspondiente a la pose compuesta por la posición $p_t = [p_{xt}, p_{yt}, p_{zt}]^T$ y orientación $\alpha = [\theta, \phi, \psi]^T$ del tobillo respecto al marco de referencia de la pelvis.

II-C2. Modelo Biomecánico: Se estableció un modelo biomecánico constituido por catorce grados de libertad (GDL). Siete para la pierna derecha, considerando las articulaciones de la cadera (3), rodilla (1) y tobillo (3), como se muestra en la Figura 4.

II-C3. Matrices de Transformación Homogénea A_j^i : La cinemática directa del modelo fue resuelta mediante la metodología de Denavit Hartenberg, [18], obteniéndose:

$$A_0^1 = \begin{bmatrix} c_{\theta_1} & -s_{\theta_1}c_{\alpha_1} & s_{\theta_1}s_{\alpha_1} & a_1c_{\theta_1} \\ s_{\theta_1} & c_{\theta_1}c_{\alpha_1} & -c_{\theta_1}s_{\alpha_1} & a_1s_{\theta_1} \\ 0 & s_{\alpha_1} & c_{\alpha_1} & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_1^2 = \begin{bmatrix} c_{\theta_2} & -s_{\theta_2}c_{\alpha_2} & s_{\theta_2}s_{\alpha_2} & a_2c_{\theta_2} \\ s_{\theta_2} & c_{\theta_2}c_{\alpha_2} & -c_{\theta_2}s_{\alpha_2} & a_2s_{\theta_2} \\ 0 & s_{\alpha_2} & c_{\alpha_2} & d_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_2^3 = \begin{bmatrix} c_{\theta_3} & -s_{\theta_3}c_{\alpha_3} & s_{\theta_3}s_{\alpha_3} & a_3c_{\theta_3} \\ s_{\theta_3} & c_{\theta_3}c_{\alpha_3} & -c_{\theta_3}s_{\alpha_3} & a_3s_{\theta_3} \\ 0 & s_{\alpha_3} & c_{\alpha_3} & d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_3^4 = \begin{bmatrix} c_{\theta_4} & -s_{\theta_4}c_{\alpha_4} & s_{\theta_4}s_{\alpha_4} & a_4c_{\theta_4} \\ s_{\theta_4} & c_{\theta_4}c_{\alpha_4} & -c_{\theta_4}s_{\alpha_4} & a_4s_{\theta_4} \\ 0 & s_{\alpha_4} & c_{\alpha_4} & d_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_4^5 = \begin{bmatrix} c_{\theta_5} & -s_{\theta_5}c_{\alpha_5} & s_{\theta_5}s_{\alpha_5} & a_5c_{\theta_5} \\ s_{\theta_5} & c_{\theta_5}c_{\alpha_5} & -c_{\theta_5}s_{\alpha_5} & a_5s_{\theta_5} \\ 0 & s_{\alpha_5} & c_{\alpha_5} & d_5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_5^6 = \begin{bmatrix} c_{\theta_6} & -s_{\theta_6}c_{\alpha_6} & s_{\theta_6}s_{\alpha_6} & a_6c_{\theta_6} \\ s_{\theta_6} & c_{\theta_6}c_{\alpha_6} & -c_{\theta_6}s_{\alpha_6} & a_6s_{\theta_6} \\ 0 & s_{\alpha_6} & c_{\alpha_6} & d_6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A_6^7 = \begin{bmatrix} c_{\theta_7} & -s_{\theta_7}c_{\alpha_7} & s_{\theta_7}s_{\alpha_7} & a_7c_{\theta_7} \\ s_{\theta_7} & c_{\theta_7}c_{\alpha_7} & -c_{\theta_7}s_{\alpha_7} & a_7s_{\theta_7} \\ 0 & s_{\alpha_7} & c_{\alpha_7} & d_7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La transforma-

da homogénea A_0^7 del efector final (referencial 7 o σ_7) hasta

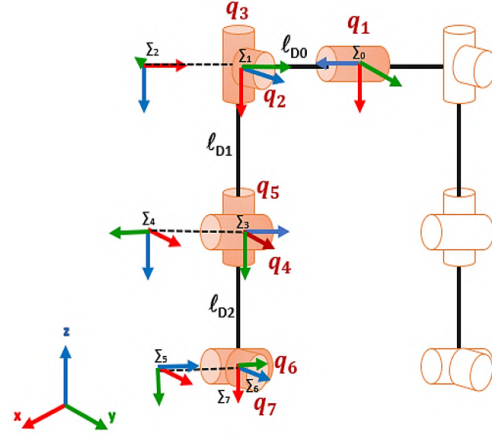


Figura 4. Abstracción del modelo biomecánico para la pierna derecha representado aquí como una cadena cinemática articulada de 7 grados de libertad, siendo un total de 14 para ambas piernas

la base (referencial 0, o σ_0) para la pierna derecha se establece de la siguiente manera:

$$A_0^7 = A_0^1 A_1^2 A_2^3 A_3^4 A_4^5 A_5^6 A_6^7 \quad (5)$$

El modelo biomecánico está entonces representado como una cinemática directa articulada por ejes rotacionales, con lo que se puede calcular el vector r del efector final para generar datos sintéticos, y validar el funcionamiento de la RNL conociendo el valor de cada q_i . Los parámetros correspondientes a las medidas antropométricas del sujeto fueron de $l_{D0} = 0,18 \text{ m}$, $l_{D1} = 0,43 \text{ m}$, y $l_{D2} = 0,42 \text{ m}$.

II-C4. Cálculo de Ángulos de las Articulaciones: Una vez calculadas las posiciones y rotaciones angulares r , éstas se derivan usando el derivador de Levant, el cual garantiza convergencia robusta hacia la derivada analítica \dot{r} , representando un enorme ventaja con respecto a la mal llamada "derivada sucia de Euler", comúnmente utilizada en biomecánica. Una vez obtenida la derivada, ésta se envía como entrada a la RNL para resolver la cinemática inversa, aplicable dado que RNL está diseñada para calcular la cinemática inversa para sistemas redundantes, [19], en el presente caso tenemos 7 grados de libertad y el sistema se mueve en el espacio por lo que se requieren seis grados de libertad, 3 de orientación y 3 de posición, por lo que es redundante.

Para el cálculo del Jacobiano geométrico, donde z_i son los vectores unitarios de rotación, con origen en $O = (O_1, \dots, O_n)$, se utiliza [18]:

$$J = \begin{bmatrix} z_{i-1} \times (O_n - O_{i-1}) \\ z_{i-1} \end{bmatrix} \quad (6)$$

que representa al operador $J_\theta \in \mathbb{R}^{6 \times n}$ el cual retroalimenta a la red aproximando los valores del vector de velocidad. Los vectores $u \in \mathbb{R}^{6 \times n}$, y $v \in \mathbb{R}^{6 \times n}$ son las capas ocultas y de salida de la RNL, respectivamente. La matriz $W \in \mathbb{R}^{n \times n}$ es una matriz identidad [20]. En el Cuadro I se puede observar

```

Entradas:  $I_{D_0}, I_{D_1}, I_{D_2}, \theta_0, \dot{r}$ ;
// Variables de posición y condiciones iniciales //
Calcular  $\dot{r}$ ; // Diferenciador de Levant //
Durante  $t \leftarrow 1$  // Tiempo de simulación en segundos //
    Calcular  $J_\theta$ ; // Jacobiano geométrico //
     $\dot{u} \leftarrow J_\theta - \dot{r}$ ;
     $u \leftarrow \int \dot{u} dt$ ;
     $\dot{v} \leftarrow (-J_\theta^T)(u) - W$ ; // W es una matriz de identidad //
     $v \leftarrow \int \dot{v} dt$ ;
Salida:  $\theta \leftarrow \int v dt$ ;
Fin
    
```

Cuadro I

METACÓDIGO DEL PROCESO DE OBTENCIÓN DE LOS ÁNGULOS θ DE CADERA Y RODILLA A PARTIR DE COORDENADAS CARTESIANAS r RESOLVIENDO LA RNL, CUYA ENTRADA ES LA DERIVADA \dot{r} .

el metacódigo del programa, donde el tiempo de simulación se estableció de acuerdo al tiempo que el sujeto tarda en dar todos los pasos requeridos, con una condición inicial para el ángulo de la cadera en 45° y para la rodilla en 0° , y $\theta \in \mathbb{R}^{n \times 1}$.

III. RESULTADOS

Para demostrar el funcionamiento de la red, se realizaron pruebas con datos sintéticos de posición a los que posteriormente, para obtener las velocidades, se les derivó utilizando el derivador de Levant. Se compararon utilizando la cinemática diferencial. Finalmente, una vez probada la red, se realizaron pruebas experimentales con un sujeto y se realizaron comparativas entre los tres métodos.

III-A. Prueba de Funcionamiento de la RNL

Para comprobar los resultados obtenidos por la RNL, se calculó una trayectoria en la que la cadera se desplazó de 45° a -22° , similar al desplazamiento planar de un péndulo que se balancea de derecha a izquierda. Para la rodilla se consideró una trayectoria de los 0° a los -22° , [21]. Estos datos fueron introducidos en el modelo biomecánico para calcular el vector de entrada de la RNL, con la cinemática directa y los ángulos de Euler. En la Figura 6 se muestra la salida de la RNL, donde q_1 , q_2 y q_3 son los tres grados de libertad de la cadera, q_4 y q_5 con los dos grados de libertad de la rodilla, q_6 y q_7 pertenecen al tobillo. También se realizó una prueba con los datos sintéticos, pero sin el uso del derivador de Levant, calculando el vector de velocidades deseadas mediante el Jacobiano geométrico y un vector de velocidades [22]. En la Figura 7 se muestra la salida de la RNL, con datos sintéticos usando la cinemática diferencial.

En la Figura 8 se muestra la detección de las articulaciones de los miembros inferiores del sujeto durante la etapa experimental. Se obtuvieron los datos de las coordenadas cartesianas mediante la implementación de MediaPipe para la pierna derecha para ser procesados para la entrada de la RNL. En la Figura 9, se muestran los ángulos de la cadera y

rodilla. La línea discontinua muestra la trayectoria calculada por la red para la cadera q_1 , que como se estableció para la validación de datos, esta articulación realizó una trayectoria de 45° a -22° . La línea negra punteada muestra la trayectoria de la rodilla q_2 , que como se esperaba, inicio sin flexión en 0° , pero con el paso del tiempo eventualmente se flexionó hasta aproximarse a los -22° . Para el resto de GDL, se logra apreciar que algunas divergen con el paso del tiempo, como el caso de q_3 y q_2 . Esta característica también se presentó para los datos sintéticos cuando se derivaron empleando Levant, a diferencia de los datos sintéticos que no fueron derivados por Levant, que se obtuvieron directamente con la cinemática diferencial (Cuadro II). En este cuadro se tiene la comparación de los resultados de los vídeos en conjunto con los dos métodos que se utilizaron para la validación de los datos. Se muestran los valores RMSE y las desviaciones estándar de los tres métodos.

GDL	Método					
	DL _e		DC _d		ML _e	
	RMSE	D.E	RMSE	D.E	RMSE	D.E
q ₁	0,0259	20,10	0,0137	19,40	0,0245	19,68
q ₂	0,0036	1,37	0,0025	0,88	0,0034	1,32
q ₃	0,0228	1,31	0,0236	0,88	0,0033	1,27
q ₄	0,0317	6,99	0,0321	6,74	0,0032	6,19
q ₅	0,0013	0,47	0,0026	0,94	0,0012	0,42
q ₆	0,0019	0,64	0,0004	0,10	0,0018	0,63
q ₇	0,0019	0,64	0,0004	0,10	0,0018	0,63

Cuadro II

VALORES RMSE Y DESVIACIONES ESTÁNDAR DE CADA GDL UTILIZANDO LOS TRES MÉTODOS: 1.- DATOS SINTÉTICOS Y DERIVADOR DE LEVANT (DL_e), 2.- DATOS SINTÉTICOS Y CINEMÁTICA DIFERENCIAL (DC_d), Y 3.- MEDIAPIPE Y LEVANT (ML_e)

III-B. Discusiones

A partir del estudio preliminar con un solo sujeto, podemos establecer, también preliminarmente, lo siguiente.

La RNL computa los datos de las coordenadas cartesianas empleando medidas antropométricas del sujeto sin un entrenamiento previo, por lo que es práctico y no requiere pasos adicionales para poner en marcha el sistema, además que no implica usar trajes con marcadores o equipo invasivo, entre otros elementos.

En el cuadro II, se observa que para q_1 , la desviación estándar y RMSE de los datos de MediaPipe (ML_e), resultado menor a la de los datos sintéticos con Levant (DL_e), pero no mejor que la prueba con los datos sintéticos y la cinemática diferencial (DC_d), siendo incluso menor. En caso de q_4 se obtuvieron mejores resultados para el método de MediaPipe, siendo el que obtuvo menor desviación de todos los métodos, y siendo el segundo mejor en RMSE con muy poca diferencia ante el método (DL_e).

IV. CONCLUSIONES

El uso de la RNL para la estimación 3D de las trayectorias angulares a partir de datos Cartesianos mejora el sistema de detección de MediaPipe con sensores no invasivos de visión monocular de bajo costo, en conjunto con el derivador de

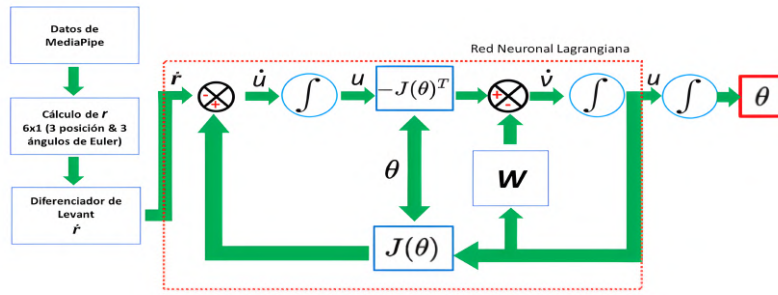


Figura 5. Diagrama general del proceso de obtención de los ángulos de cadera y rodilla.

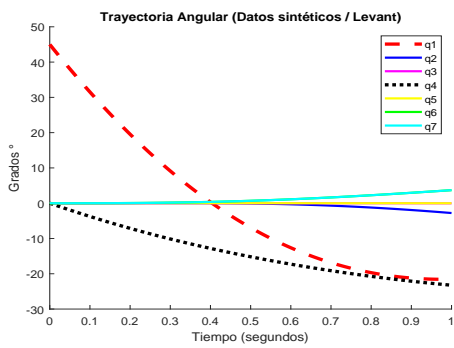


Figura 6. Trayectorias de la cadera q_1 (línea discontinua) y la rodilla q_4 (línea punteada) empleando datos sintéticos y el derivador de Levant.

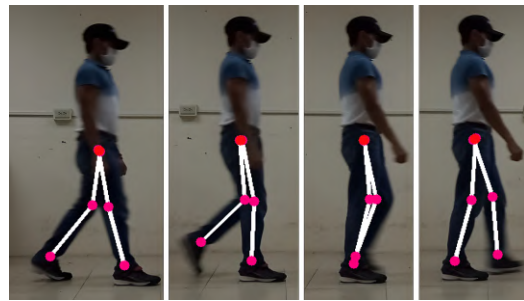


Figura 8. Prueba experimental para la detección de los miembros inferiores usando MediaPipe durante la ejecución de un paso

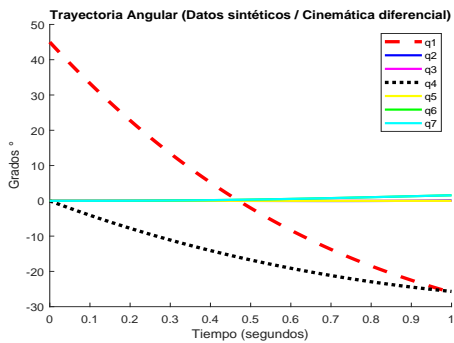


Figura 7. Gráficas de la estimación para las trayectorias de la cadera y rodilla empleando datos sintéticos y cinemática diferencial.

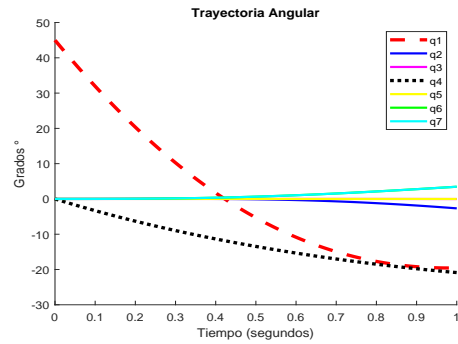


Figura 9. Gráfica de las trayectorias de la cadera q_1 (línea discontinua) y la rodilla q_4 (línea punteada)

Levant. Adicionalmente, sin calibración absoluta ni relativa, como aún lo requieren hasta los sistemas tipo Vicon y Optitrack. El sistema propuesto proporciona un vector de velocidad aproximado sin requerir de un sensor de velocidad físico.

Se validó el desempeño de la RNL mediante datos sintéticos que da soporte a la propuesta. A diferencia de algunas propuestas de la literatura, [14], [10], los resultados son en el

espacio 3D, lo cual resulta relevante en un entorno clínico o de entrenamiento. El costo computacional es bajo, primero porque el vídeo obtenido es de un celular de gama baja y es de una resolución baja, segundo, porque el algoritmo consiste de operaciones algebraicas gracias al uso del modelo biomecánico. Tercero, MediaPipe está optimizado, es decir, cuenta con un modelo para reconstruir el esqueleto, en contraste con [4], quienes utilizan una variante entrenada del modelo VNect. Todo esto hace que nuestra propuesta sea asequible a las necesidades clínicas que requieren datos de pacientes

durante rehabilitación motora ya sea durante la bipedestación o marcha. El usuario, ya sea el fisioterapeuta, el médico o un enfermero, solamente toman el video del paciente con un celular de gama baja y posteriormente este es analizado por nuestro algoritmo.

A partir de los resultados obtenidos vislumbramos como trabajo a futuro, analizar la estabilidad de la RNL para mejorar la sincronización con MediaPipe. Incluir el ángulo del tobillo y el ángulo transversal del eje de los dedos de los pies, se considera el método de Denavit Hartenberg modificado para modelar una estructura arborescente más compleja.

REFERENCIAS

- [1] Muhammad Yahya, Jawad Ali Shah, Kushsairy Abdul Kadir, Zulkhairi M. Yusof, Sheraz Khan, and Arif Warsi. Motion capture sensing techniques used in human upper limb motion: a review. 39(4):504–511. Publisher: Emerald Publishing Limited.
- [2] Matteo Menolotto, Dimitrios-Sokratis Komaris, Salvatore Tedesco, Brendan O’Flynn, and Michael Walsh. Motion capture technology in industrial applications: A systematic review. *Sensors*, 20(19), 2020.
- [3] Ling-Fung Yeung, Zhenqun Yang, Kenneth Chik-Chi Cheng, Dan Du, and Raymond Kai-Yu Tong. Effects of camera viewing angles on tracking kinematic gait patterns using azure kinect, kinect v2 and orbbeec astra pro v2. *Gait & Posture*, 87:19–26, 2021.
- [4] Arash Azhand, Sophie Rabe, Swantje Müller, Igor Sattler, and Anika Steinert. Algorithm based on one monocular video delivers highly valid and reliable gait parameters. *Scientific Reports*, vol. 11:14065, 2021.
- [5] Venkatesh Balasubramanian, K. Adalarasu, and Rahul Regulapati. Comparing dynamic and stationary standing postures in an assembly task. 39(5):649–654. Publisher: Elsevier.
- [6] T. Tanaka, N. Hashimoto, M. Nakata, T. Ito, S. Ino, and T. Ifukube. Analysis of toe pressures under the foot while dynamic standing on one foot in healthy subjects. 23(3):188–193.
- [7] Franziska Schlagenhaut, Siddarth Sreeram, and William E. Singhose. Comparison of kinect and vicon motion capture of upper-body joint angle tracking. *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, pages 674–679, 2018.
- [8] Sen Zhang, Boon Yuen Tan, Kee Hao Leo, and Ronny Quin Fai Tham. A hybrid human motion tracking system for virtual rehabilitation. pages 1993–1998, 2011.
- [9] Justin Amadeus Albert, Victor Owolabi, Arnd Gebel, Clemens Markus Brahm, Urs Granacher, and Bert Anrich. Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: A pilot study. 20(18):5104.
- [10] Aditya Viswakumar, Venkateswaran Rajagopalan, Tathagata Ray, and Chandu Parimi. Human gait analysis using openpose. pages 310–314, 2019.
- [11] Vinod Gutta, Edward D. Lemaire, Natalie Baddour, and Pascal Fallavollita. A comparison of depth sensors for 3d object surface reconstruction. *CMBES Proceedings*, 42, May 2019.
- [12] Thiago Buarque de Gusmao Lafayette, Alexandre Burle, Arthur de Almeida, Vinicius Lima Ventura, Vitor Carvalho, Alana Elza Fontes da Gama, Joao Marcelo Xavier Nataro Teixeira, and Veronica Teichrieb. The virtual kinect. pages 110–118, 2021.
- [13] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. BlazePose: On-device Real-time Body Pose tracking. *arXiv:2006.10204 [cs]*, pages 1–4, 2020. arXiv: 2006.10204.
- [14] Ardra Anilkumar, Athulya K. T., Sarath Sajan, and Sreeja K. A. Pose Estimated Yoga Monitoring System. (3882498), 2021.
- [15] Lukasz Kidziński, Bryan Yang, Jennifer Hicks, Apoorva Rajagopal, Scott Delp, and Michael Schwartz. Deep neural networks enable quantitative movement analysis using single-camera videos. *Nature Communications*, 11:4054, 08 2020.
- [16] Markus Reichhartinger and S.K. Spurgeon. A robust exact differentiator block for matlab/simulink. *ResearchGate*, pages 1–10, 2016.
- [17] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. Microsoft COCO: Common Objects in Context. *arXiv:1405.0312 [cs]*, pages 1–15, 2015. arXiv: 1405.0312.
- [18] Mark W Spong, Seth Hutchinson, and M Vidyasagar. Robot modeling and control. 2006.
- [19] Stefano Chiaverini, Giuseppe Oriolo, and Anthony A. Maciejewski. Redundant robots. pages 221–242.
- [20] Jun Wang, Qingni Hu, and Danchi Jiang. A lagrangian network for kinematic control of redundant manipulators. *Neural Networks, IEEE Transactions on*, 10:1123 – 1132, 1999.
- [21] Walter Pirker and Regina Katzenschlager. Gait disorders in adults and the elderly : A clinical guide. *Wiener Klinische Wochenschrift*, 129(3-4):81–95, 2017.
- [22] John J Craig. Robótica. 2006.