Jornadas de Automática

Estimation of joint angles in upper limbs using a smartphone

Carlos Montero[a,*], Eva Portillo[a], Asier Zubizarreta[a], Gil De Sousa[b], Itziar Cabanes[a]

[a] *Dept. of Automatic Control and Systems Engineering, Faculty of Engineering Bilbao (EHU), Plaza Ingeniero Torres Quevedo, 1, 48508, Bilbao, España.*
[b] *Gorliz Hospital, Paseo Urezarantza 1, 48630 Gorliz, España.*

To cite this article: Montero, C, Portillo, E, Zubizarreta, A, De Sousa, Gil, Cabanes, I. 2026. Estimation of joint angles in the upper limbs using a smartphone. Jornadas de Automática, 47.
https://doi.org/10.17979/ja-cea.2026.47.13777

Resumen

La evaluación clínica de la espasticidad en pacientes que han sufrido ictus depende de la valoración subjetiva de los especialistas. Con el objetivo de cuantificar objetivamente la movilidad articular, este trabajo valida el modelo BlazePose, basado en IA e integrado en el framework de MediaPipe, para estimar cuatro ángulos articulares del miembro superior. La precisión de este modelo se compara con el sistema de captura de movimiento OptiTrack. Para ello, se ha diseñado un protocolo experimental y se ha evaluado la abducción y flexión de hombro, la flexión de codo y la flexión de muñeca. Los resultados muestran que los tres primeros ángulos alcanzan un coeficiente de correlación $r \geq 0.94$ respecto al sistema de referencia. La abducción de hombro obtiene el mejor resultado, con un error medio absoluto de $8,3$º mientras que el codo muestra un sesgo sistemático de +18.9º. La conclusión principal del trabajo es que la viabilidad clínica de esta solución está condicionada a protocolos de calibración específicos.

Palabras clave: Medición y procesamiento de señales biomédicas, Imagen biomédica y médica, procesamiento de imágenes, visualización, Ingeniería de rehabilitación y prestación de asistencia sanitaria, Percepción y detección en robótica, Aprendizaje automático y aprendizaje profundo para la identificación de sistemas

Estimation of joint angles in the upper limbs using a smartphone.

Abstract

The clinical assessment of spasticity in stroke patients relies on the subjective judgement of specialists. With the aim of objectively quantifying joint mobility, this study validates the BlazePose model—an AI-based model integrated into the MediaPipe framework—to estimate four joint angles of the upper limb. The accuracy of this model is compared with the OptiTrack motion capture system. To this end, an experimental protocol was designed and shoulder abduction and flexion, elbow flexion and wrist flexion were assessed. The results show that the first three angles achieve a correlation coefficient $r \geq 0.94$ with respect to the reference system. Shoulder abduction yields the best result, with a mean absolute error of 8.3º, whilst the elbow shows a systematic bias of +18.9º. The main conclusion of the study is that the clinical viability of this solution is contingent upon specific calibration protocols.

Keywords: Biomedical signal measurement and processing, Biomedical and medical imaging, image processing, visualization, Rehabilitation engineering and healthcare delivery, Robot perception and sensing, Machine and deep learning for system identification

*Autor para correspondencia: carlos.montero@ehu.eus

1. Introduction

A stroke, defined as a sudden disruption of an individual's normal physiological state, is the second leading cause of death among women and the leading cause of disability in adults in Europe (Ministerio de Sanidad, 2024). At a physical level, stroke survivors tend to develop spasticity. This condition affects both the upper and lower limbs and is characterised by continuous, involuntary muscle tension and/or stiffness. Spasticity is not understood merely as an isolated phenomenon. It is the most striking manifestation of a broader complex of symptoms that affects the patient's reasoning, mood and other vital faculties.

Prevalence data are significant: 43% of the stroke population is affected by spasticity (Doussoulin et al., 2020) and, in Spain in particular, it is estimated that this percentage rises to 54% after three months (Béseler Soto et al., 2020). Nevertheless, the shortage of specialist therapists, subjective assessment and the lack of technological innovation contribute to the underdiagnosis of this disorder. To address these challenges, a new methodology for efficient and reliable assessment of spasticity in the upper and lower limbs is required.

Among the range of low-cost technologies, human pose estimation (HPE) using computer vision is of particular interest. It is a markerless promising technique as it performs detection and tracking of position and orientation of human body parts in images or videos using computer vision and artificial intelligence.

Among the most widely used pose estimation models—BlazePose, OpenPose and AlphaPose—BlazePose is the only one capable of operating in real time on mobile devices (Opiña and Fajardo, 2024), whilst OpenPose remains the most widely used (36% compared to 16% for BlazePose) (di Biase et al., 2025). BlazePose, integrated within the MediaPipe framework, has been validated against a universal goniometer and a digital angle ruler as a reliable and valid tool for measuring shoulder abduction, adduction, flexion, and extension range of motion (Latreche et al., 2023), and in 2D space it has proven sensitive in detecting motor recovery in stroke patients (Wagh et al., 2025).

Despite the promising results, several challenges remain. In post-stroke patients performing a standardised drinking task, raw MediaPipe keypoints failed to generalise across unseen participants for compensatory shoulder movement detection, achieving only 50% inter-participant accuracy; incorporating custom kinematic features raised this figure to 70% (Unger et al., 2025). Occlusion, depth ambiguity and the need for labelled data are the main current limitations (Opiña and Fajardo, 2024). Camera placement also poses a significant challenge: a 45º ipsilateral view yields better results than a frontal view for detecting compensatory movements in stroke patients (Unger et al., 2025). Furthermore, the moderate inter-therapist reliability observed in certain upper limb tasks reinforces the need for objective tools to standardise the assessment of spasticity (Sauerzopf et al., 2024). However, its implementation in clinical settings is still in its early stages. Thus, their accuracy must be validated against gold-standard motion capture systems (MCS).

In this context, the present work introduces a pilot study to validate the potential of HPE algorithms implemented in mobile devices for characterising the range of motion of stroke patients, with the aim of providing preliminary technical evidence on spasticity-related data under controlled conditions. To this end, the data obtained by HPE algorithms is compared and validated with 3D Optical MCS systems, considered the gold standard in the literature. Although spasticity affects both upper and lower limbs, this study focuses on the left upper limb as a first step. The remainder of this paper is organised as follows: Section 2 reviews the HPE algorithms and justifies the selection of BlazePose; Section 3 describes the validation methodology; Section 4 analyses the results; and Section 5 draws the conclusions.

2. Selection of HPE Algorithms

HPE models follow two main architectural paradigms for multi-person detection: bottom-up and top-down (Zheng et al., 2023). The bottom-up approach first detects all keypoints present in the image—regardless on how many people there are—and then groups them to reconstruct the skeleton of each individual. OpenPose or MoveNet algorithms fall into this group. The computational complexity of this strategy scales sublinearly with the number of people, making it particularly efficient in densely populated scenes. The top-down approach, implemented in models such as PoseNet or BlazePose, first locates each person using a person detector (usually an object detection model) and then applies a pose estimator independently to each cropped region. This produces more accurate estimates per individual, although the computational cost increases linearly with the number of people detected. Table 1 summarises these models.

Table 1: Basic specifications of OpenPose, PoseNet, MoveNet and BlazePose.

	OpenPose	**PoseNet**	**MoveNet**	**BlazePose**
Arch.	B-up	B-up	B-up	T-down
Platform	S/D	W/M	W/M/E	M/W/E
Keypoints	137	17	17	33
Detection	Multi	Multi	S/Multi	Single
3D Est.	No	No	No	Yes

B-up: Bottom-up, T-down: Top-down. S: Server, D: Desktop, W: Web, M: Mobile, E: Edge. Arch.: Architecture, Est.: Estimation.

Of the models presented, BlazePose shows a higher potential given its direct relevance to this study. The selection responds to the following criteria: (a) the target platform, as the proposed methodology for spasticity diagnosis is intended to be deployed on a smartphone; (b) the integration of BlazePose within the MediaPipe framework, which provides a higher number of keypoints and supports 3D pose estimation, features that are expected to be advantageous for the purposes of this study.

BlazePose GHUM 3D models are designed for real-time human pose estimation on mobile devices. They detect 33 key body landmarks (x, y, z coordinates) and provide visibility values and presence for each point, as shown in Figure 1(b). There are three variants depending on the desired balance between speed and accuracy: lite, full and heavy (see Table 2). In any variant, the input of the model is a RGB image resized to 256 × 256 pixels. The output consists of a vector containing the image-relative normalised coordinates, the hip-relative depth coordinates (on a synthetic scale) and the probability of presence and visibility.

Table 2: BlazePose model performance comparison (reference: Pixel 3).

	Size	FPS (CPU)	FPS (GPU)
Lite	3 MB	~ 44	~ 49
Full	6 MB	~ 18	~ 40
Heavy	26 MB	~ 4	~ 19

3. Validation methodology

To assess the accuracy of the proposed smartphone-based pipeline, a structured validation protocol was designed. It is organised as follows. Section 3.1 describes the OptiTrack reference system and marker set. Section 3.2 defines the comparison metrics between both systems. Section 3.3 details the experimental design and camera viewpoints evaluated. Finally, Section 3.4 specifies how joint angles are computed from each system.

3.1. Reference system

In order to validate the selected HPE algorithm, its results will be compared with the gold standard in human motion capture. The reference system used is OptiTrack, a MCS based on retroreflective markers. Of the various models available, the Baseline model—comprising 41 markers—was used for upper limb capture (Figure 1(a)), and the software used for data processing is Motive. Among the MCS used as a benchmark in clinical and biomechanical research, OptiTrack has demonstrated accuracy comparable to that of Vicon, with only marginal errors. Reproducibility also depends on the documentation of key parameters such as sampling rate, camera resolution and marker size (Nagymate and Kiss, 2018). In the context of markerless systems, approximately 23% of studies validated their results by comparing them with reference systems (Jeyasingh-Jacob et al., 2024), underscoring the need for standardised validation protocols (Knippenberg et al., 2017). Single-camera markerless systems show clinical potential but insufficient performance for complex biomechanics without prior rigorous validation (Scott et al., 2022). Of particular relevance to the present work, the validation of MediaPipe against OptiTrack for estimating range of motion (ROM) has shown minimal angular differences and 92% accuracy in posture estimation (Urriola et al., 2023).

The working hypothesis is that, under controlled recording conditions, the joint angles estimated from the points inferred by BlazePose show clinically acceptable agreement with those measured by OptiTrack.

Figure 1 illustrates the two capture models used: the Baseline marker model (Figure 1(a)) and the BlazePose model (Figure 1(b)). A notable difference between the two systems lies in the anatomical placement of their respective markers. In OptiTrack, joint positions are not defined by single points but are inferred from clusters of retroreflective markers placed on specific body landmarks: the shoulder center is computed as the midpoint between the superior and the posterior acromion, the elbow corresponds to the lateral epicondyle, and the wrist center is derived from the midpoint of radial styloid and ulnar styloid. The hand marker is placed on the dorsum over the fifth metacarpal. In contrast, BlazePose approximates the joint centers directly. This distinction is significant from a biomechanical perspective, as the position may vary depending on body shape, clothing or occlusion, introducing a systematic bias specific to each individual that must be taken into account during validation. The specific landmarks used by each system for the four angles evaluated in this study are summarised in Table 3, where all markers and landmarks correspond to the left upper limb (L).

3.2. Comparison metric

Given the above, the chosen comparison metric is the joint angle, as it is the biomechanical parameter of greatest clinical relevance and the one that best encapsulates the function of the movement. Unlike the Cartesian coordinates of the markers—whose interpretation depends on the global reference frame and the subject's position in space—the joint angle remains invariant with respect to translations and rotations of the body. The angles selected for validation are calculated according to the degrees of freedom of the three joints of the upper limb: shoulder, elbow and wrist. The aim is to decompose each movement into the anatomical planes using the X, Y, Z coordinates, so that each degree of freedom is described by an angle with clear functional significance.

3.3. Experimental design

A pilot capture session was carried out at the Biomedical Engineering Laboratory of the Bilbao School of Engineering. The study was approved by the Ethics Committee of the University of the Basque Country (UPV/EHU) under approval code M10_2025_122. Written informed consent was obtained from the participant prior to data collection. One healthy male participant (age: 23 years) took part in it.

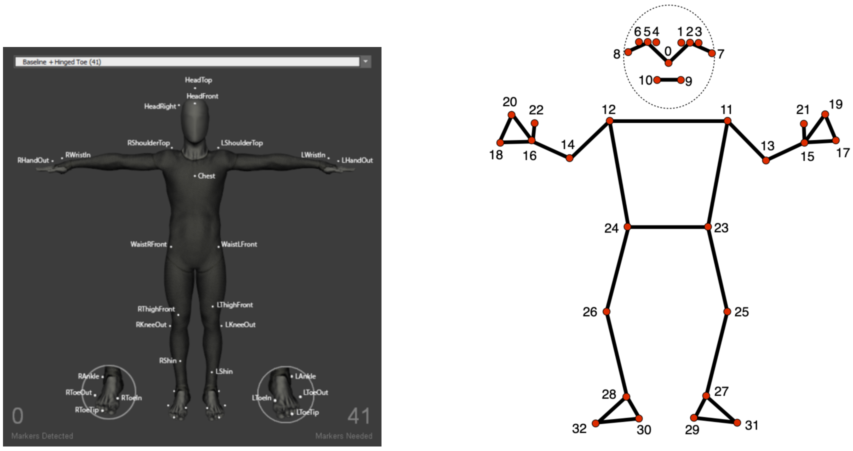
The four movements were selected on the basis of technical feasibility: each involves a single degree of freedom, a wide angular range, and landmarks that remain visible throughout the motion. Upper limb spasticity patterns are a combination of these 4 basic movements. These upper limb tasks, all performed with the left arm, were: (1) *pure abduction*, consisting of raising the extended arm laterally up to 90°; (2) *pure flexion*, consisting of raising the extended arm forward up to 90°; (3) *elbow flexion/extension*, pulling the arm up against the trunk, perform full flexion and extension of the elbow; (4) *wrist flexion/extension*, recorded from a lateral viewpoint, consisting of repeatedly bending the wrist upward and downward through its full range of motion. Every task was repeated 1 time.

OptiTrack and the smartphone camera (Google Pixel 8) simultaneously captured motion data at 60 fps from the optimal viewpoint for each movement. The capture viewpoint was selected per movement: frontal for pure abduction and pure flexion, and lateral for elbow and wrist flexion/extension. The synchronisation of both systems is established by a clap: the moment of minimum distance between the wrists, detected in both time series. The synchronization signal enables temporal alignment between both time series, facilitating their comparison.

Table 3: Anatomical landmarks used by each capture system to compute the joint angles evaluated in this study.

Joint angle	Role	OptiTrack marker	BlazePose landmark
Elbow flex/ext	Shoulder[a,b,c]	*midpoint*(LShoulderTop, LShoulderBack)*	11
	Elbow[a,b]	LElbowOut	13
	Wrist[c]	*midpoint*(LWristIn, LWristOut)*	15
Shoulder abduction	Trunk[b]	LShoulderTop and Back, WaistLFront and RFront	11, 12, 23, 24
Wrist flex/ext	Hand long. axis	*midpoint*(LWristIn, LWristOut) → LHandOut	*midpoint*(17, 19)* → 15
	Hand trans. axis	LWristIn → LWristOut	17 → 19

*Computed as the geometric midpoint of the two landmarks; [a] used in shoulder abduction, [b] used in shoulder flexion, [c] used in wrist flexion; L: left.



(a) OptiTrack marker model.

(b) BlazePose 33-landmark model.

Figure 1: Capture models used in the study. *Left*: *Baseline* skeleton processed with Motive. *Right*: body landmarks inferred by BlazePose.

3.4. Joint angle computation

The summary of the estimated joint angles is detailed in Table 4. Each joint angle is computed from the three-dimensional positions of the set of anatomical landmarks common to both systems as detailed in Figure 2.

Table 4: Joint angles computed for the upper limb validation study.

Joint	Motion
Shoulder	Flexion / Extension
	Abduction / Adduction
Elbow	Flexion / Extension
Wrist	Flexion / Extension

- Elbow flexion/extension is defined as the angle between the humerus vector and the forearm vector, with neutral at 180° in full extension (Figure 2(a)):

$$\theta_{\mathrm{flex\_elb}} = \operatorname{atan2}(\|\mathbf{v}_{\mathrm{h}} \times \mathbf{v}_{\mathrm{f}}\|,\ \mathbf{v}_{\mathrm{h}} \cdot \mathbf{v}_{\mathrm{f}})\,. \tag{1}$$

- Shoulder abduction is computed by projecting the arm vector $\mathbf{v}$ and the trunk downward reference $\mathbf{e}_{\mathrm{t}}$ onto the subject's anatomical frontal plane, whose unit normal $\mathbf{n}_{\mathrm{f}}$ points in the anteroposterior direction. $\mathbf{n}_{\mathrm{f}}$ is computed as $\mathbf{n}_f = (\mathbf{v}_s \times \mathbf{v}_h)/\|\mathbf{v}_s \times \mathbf{v}_h\|$, where $\mathbf{v}_s$ is the vector connecting the left and right shoulders and $\mathbf{v}_h$ is the vector connecting the left and right hips. The projections are: $\mathbf{v}_{\mathrm{proj}} = \mathbf{v} - (\mathbf{v} \cdot \mathbf{n}_{\mathrm{f}})\,\mathbf{n}_{\mathrm{f}}$ and $\mathbf{v}_{\mathrm{ref}} = \mathbf{e}_{\mathrm{t}} - (\mathbf{e}_{\mathrm{t}} \cdot \mathbf{n}_{\mathrm{f}})\,\mathbf{n}_{\mathrm{f}}$, and the abduction angle is then:

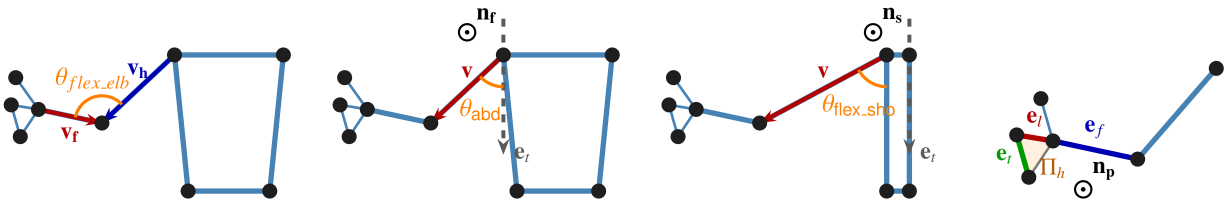
$$\theta_{\mathrm{abd}} = \operatorname{atan2}\Big(\|\mathbf{v}_{\mathrm{ref}} \times \mathbf{v}_{\mathrm{proj}}\|,\ \mathbf{v}_{\mathrm{ref}} \cdot \mathbf{v}_{\mathrm{proj}}\Big)\,, \tag{2}$$

with 0º corresponding to the arm in a neutral position at the side of the body.

- Shoulder flexion/extension is computed analogously to shoulder abduction, but the projection plane is changed: instead of removing the anteroposterior component, the mediolateral component is eliminated. Concretely, $\mathbf{n}_{\mathrm{f}}$ is replaced by the trunk lateral unit vector $\mathbf{n}_{\mathrm{s}}$, so that $\mathbf{v}_{\mathrm{proj}}$ and $\mathbf{v}_{\mathrm{ref}}$ lie in the subject's anatomical sagittal plane. The sagittal plane normal $\mathbf{n}_s$ is obtained as $\mathbf{n}_s = (\mathbf{v}_t \times \mathbf{n}_f)/\|\mathbf{v}_t \times \mathbf{n}_f\|$, where $\mathbf{v}_t$ is the trunk downward vector. The angle $\theta_{\mathrm{flex\_sho}}$ is then obtained with the same expression as in Eq. (2), with 0° corresponding to the arm hanging neutrally at the side of the body and approximately 180° to full anterior flexion.

- Wrist flexion and extension are calculated relative to the plane of the hand, defined by two vectors: the longitudinal axis $\mathbf{e}_{\mathrm{l}}$ and the transverse axis $\mathbf{e}_{\mathrm{t}}$. The cross-product of these two vectors defines the normal $\mathbf{n}_{\mathrm{p}}$ to that plane, pointing towards the back of the hand. The angle of flexion is then obtained as the arcsine of the dot product between the forearm axis $\mathbf{e}_{\mathrm{f}}$ and the palmar normal (see Equation 3). Positive values correspond to palmar flexion; negative values to dorsal extension.

$$\theta_{\mathrm{flex\_wr}} = \arcsin\Big(\mathbf{e}_{\mathrm{f}} \cdot \mathbf{n}_{\mathrm{p}}\Big)\,. \tag{3}$$



(a) Elbow flexion/extension. (b) Shoulder abduction. (c) Shoulder flexion/extension. (d) Wrist flexion/extension.

Figure 2: Geometric definition of the four joint angles computed. Highlighted elements indicate the points, vectors and reference planes involved.

Table 5: Validation results for the best-performing session of each joint angle.

Joint angle	**Bias (°)**	**MAE (°)**	**RMSE (°)**	r	**LOA (°)**
Elbow flex/ext (L)	+18.9	18.9	21.2	0.98	[−0.2, 38.0]
Shoulder abduction (L)	+7.8	8.3	11.4	0.98	[−8.4, 24.0]
Shoulder flexion (L)	+15.4	17.4	20.7	0.94	[−12.0, 42.7]
Wrist flex/ext (L)	−7.0	21.2	29.4	0.22	[−63.1, 49.1]

L: left upper limb.

4. Analysis of results

Table 5 summarises the accuracy metrics obtained for each joint angle during its best-performing session, i.e., the session recorded from the optimal viewing plane for that movement. The first three joint angles achieved a Pearson correlation coefficient of $r \geq 0.94$ against the OptiTrack reference system, confirming that the smartphone-based pipeline reliably captures the temporal dynamics of upper-limb motion. Shoulder abduction yielded the strongest agreement: a systematic bias of +7.8, MAE of 8.3°, RMSE of 11.4°, and 95% limits of agreement (LOA) of $[-8.4°, 24.0°]$. Elbow flexion/extension produced a larger systematic overestimation (+18.9°, $RMSE = 21.2°$) and LOA span (38.2°), though the correlation remained high ($r = 0.98$).

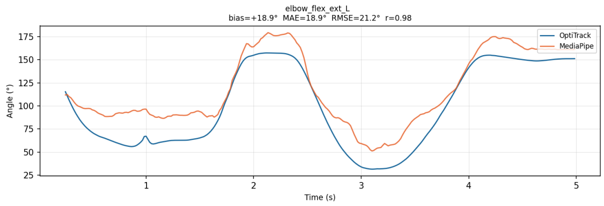
Shoulder flexion showed the lowest correlation ($r = 0.94$), a bias of +15.4°, and LOA range of 54.7°. The LOA relative to shoulder abduction likely reflects a larger systematic offset in this movement rather than increased random variability, consistent with the high correlation coefficient retained despite the elevated bias. Beyond viewpoint selection, part of the residual error observed across these joint angles is attributable to anatomical mismatches between the landmark definitions of the two systems rather than to estimation inaccuracy per se.

Figure 3 shows the time-series comparison between the reference OptiTrack signal and the smartphone-derived estimate for each of the four joint angles during their corresponding best-performing session. It should be noted in Figure 3(a) that the bias is not perfectly constant throughout the time series. Certain intervals—particularly between 1.6s–2s, 2.1s–2.5s and 3.7s–4s— correspond to abrupt changes in arm orientation relative to the camera, introducing transient depth ambiguity. This suggests that the systematic offset is partly conditioned by the selected viewpoint.

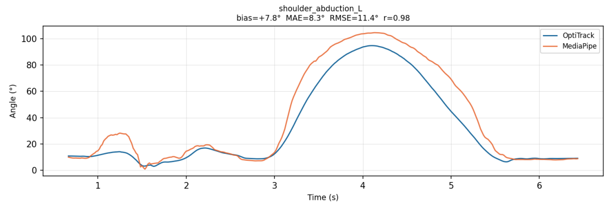
On the other hand, wrist flexion/extension yielded substantially weaker results than other joint angles: $r = 0.22$, MAE = 21.2°, RMSE = 29.4°, and the widest LOA span of all movements (112.2°), indicating that the pipeline fails to track this movement reliably. The poor agreement is reflected in the time-series signal, where the smartphone-derived estimate does not follow the reference waveform. A landmark-level analysis reveals the origin of this failure. The X and Y components of the relevant landmarks track reasonably well, but the depth component (Z) breaks down entirely, yielding near-zero or negative correlations ($r = -0.09$ for the elbow, $r = 0.32$ for the wrist). This landmark mismatch is further confirmed at the level of the vector used in the angle computation: the unit normal to the hand plane—constructed from the longitudinal and transverse vectors—. It differs by 0.5 rad ($\approx 28.6°$) between systems, which represents a large deviation for a unit vector constrained to the surface of a sphere. As both systems define hand geometry using anatomically inconsistent points, the resulting vectors diverge independently of camera placement, and this discrepancy cannot be recovered within a single-camera framework using a full-body pose estimator. A dedicated hand pose estimator such as MediaPipe Hands would supply anatomically consistent distal landmarks and is therefore identified as the most promising avenue for improving estimation.

5. Conclusions

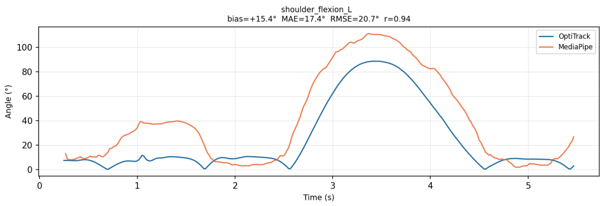
This pilot study demonstrates that, when the smartphone camera is placed at the optimal viewing plane for a given joint movement, the proposed markerless pipeline achieves clinically meaningful tracking of upper-limb kinematics. Shoulder abduction achieved an MAE of 8.3°, elbow flexion/extension had a systematic bias of +18.9°, and shoulder flexion obtained a MAE of 17.4°. These joints exceeded the $r \geq 0.94$ threshold, confirming that the pipeline captures the underlying movement pattern reliably. However, the LOA spans alongside the systematic biases observed indicate that the dominant source of error is a fixed, systematic offset rather than random noise. The remaining challenge, therefore, lies in how the optimal viewing plane is identified automatically in practice. That problem could be addressed in future work through systematic strategies such as sweeping the camera through candidate viewpoints or training a data-driven model to predict the most informative. Wrist flexion, by contrast, revealed the limits of



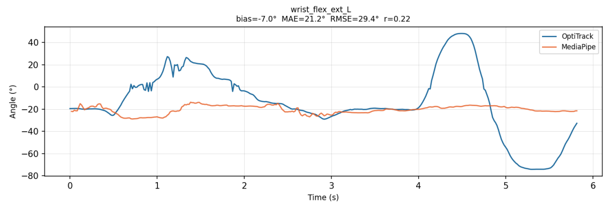
(a) Elbow flexion/extension session.



(b) Shoulder abduction session.



(c) Shoulder flexion session.



(d) Wrist flexion session.

Figure 3: Joint angle time series for the best-performing session of each movement.

full-body pose estimation for distal joints: the structural inconsistency between landmark definitions across systems produces plane-normal errors of $\approx$ 28.6º that are irrecoverable within this framework, and replacing the full-body estimator with a dedicated hand pose model is identified as the necessary next step.

Acknowledgments

This work has been supported by grant ref. PID2024-155343OB-I00 funded by MCIU/AEI/10.13039/501100011033/FEDER, UE and grant IT1836-26 funded by the Basque Government.

References

Béseler Soto, M. R., Montes García, J., Máñez Añón, I., 2020. Espasticidad tras ictus: ¿la edad es un factor de riesgo? Estudio observacional de la espasticidad en pacientes neurovasculares en una serie retrospectiva de dos centros. Revista Española de Geriatría y Gerontología 55 (5), 258–265.
DOI: `10.1016/j.regg.2020.04.004`

di Biase, L., Pecoraro, P. M., Bugamelli, F., 2025. Ai video analysis in parkinson's disease: A systematic review of the most accurate computer vision tools for diagnosis, symptom monitoring, and therapy management. Sensors 25 (20).

Doussoulin, A., Rivas, C., Bacco, J., Sepúlveda, P., Carvallo, G., Gajardo, C., Soto, A., Rivas, R., nov 2020. Prevalence of spasticity and postural patterns in the upper extremity post stroke. Journal of Stroke and Cerebrovascular Diseases 29 (11), 105253.
DOI: `10.1016/j.jstrokecerebrovasdis.2020.105253`

Jeyasingh-Jacob, J., Crook-Rumsey, M., Shah, H., Joseph, T., Abulikemu, S., Daniels, S., Sharp, D. J., Haar, S., Aug 2024. Markerless motion capture to quantify functional performance in neurodegeneration: Systematic review. JMIR Aging 7, e52582.
DOI: `10.2196/52582`

Knippenberg, E., Verbrugghe, J., Lamers, I., Palmaers, S., Timmermans, A., Spooren, A., jun 2017. Markerless motion capture systems as training device in neurological rehabilitation: a systematic review of their use, application, target population and efficacy. Journal of NeuroEngineering and Rehabilitation 14 (1), 61.
DOI: `10.1186/s12984-017-0270-x`

Latreche, A., Kelaiaia, R., Chemori, A., Kerboua, A., 2023. Reliability and validity analysis of mediapipe-based measurement system for some human rehabilitation motions. Measurement 214, 112826.
DOI: `https://doi.org/10.1016/j.measurement.2023.112826`

Ministerio de Sanidad, 2024. Estrategia en ictus del sistema nacional de salud. actualización 2024. Tech. rep., Ministerio de Sanidad, Gobierno de España, actualización aprobada por el Consejo Interterritorial del Sistema Nacional de Salud el 5 de abril de 2024. NIPO en línea: 133-24-118-7.
URL: `https://cpage.mpr.gob.es`

Nagymate, G., Kiss, R., 07 2018. Application of optitrack motion capture systems in human movement analysis a systematic literature review. Recent Innovations in Mechatronics 5.
DOI: `10.17667/riim.2018.1/13`

Opiña, P., Fajardo, A., 04 2024. Real-time pose estimation for mobile devices: A review. International Journal of Multidisciplinary Research and Analysis 07.
DOI: `10.47191/ijmra/v7-i04-44`

Sauerzopf, L., Panduro, C., Luft, A., Kühnis, B., Gavagnin, E., Unger, T., Awai, C., Schönhammer, J., Degenfellner, J., Spiess, M., 12 2024. Evaluating inter- and intra-rater reliability in assessing upper limb compensatory movements post-stroke: creating a ground truth through video analysis? Journal of NeuroEngineering and Rehabilitation 21.
DOI: `10.1186/s12984-024-01506-7`

Scott, B., Seyres, M., Philp, F., Chadwick, E. K., Blana, D., may 2022. Healthcare applications of single camera markerless motion capture: a scoping review. PeerJ 10, e13517.
DOI: `10.7717/peerj.13517`

Unger, T., Kühnis, B., Sauerzopf, L., Spiess, M. R., de Spindler, A., Luft, A. R., Easthope Awai, C., Schönhammer, J. G., Gavagnin, E., 2025. Using deep learning to detect upper limb compensation in individuals post-stroke using consumer-grade webcams—a feasibility study. Frontiers in Medicine Volume 12 - 2025.

Urriola, G. J. M., Trejos, C. I. P., Ferre, M., 2023. Comparison of intrusive and non-intrusive motion capture technologies for upper body kinematic analysis. Proceedings of the 21th LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2023): "Leadership in Education and Innovation in Engineering in the Framework of Global Transformations: Integration and Alliances for Integral Development".
URL: `https://api.semanticscholar.org/CorpusID:260985344`

Wagh, V., Scott, M. W., Andrushko, J. W., Jones, C. B., Larssen, B. C., Boyd, L. A., Kraeutner, S. N., nov 2025. Using MediaPipe to track upper-limb reaching movements after stroke: a proof-of-principle study. Journal of NeuroEngineering and Rehabilitation 22 (1), 268.
DOI: `10.1186/s12984-025-01808-4`

Zheng, C., Wu, W., Chen, C., Yang, T., Zhu, S., Shen, J., Kehtarnavaz, N., Shah, M., 2023. Deep learning-based human pose estimation: A survey 56 (1).