POLITECNICO DI TORINO
Repository ISTITUZIONALE

A Model-Based Markerless Protocol for Clinical Gait Analysis Based on a Single RGB-Depth Camera:
Concurrent Validation on Patients With Cerebral Palsy

(Article begins on next page)

11 May 2024

RESEARCH ARTICLE

A Model-Based Markerless Protocol for Clinical Gait Analysis Based on a Single RGB-Depth Camera: Concurrent Validation on Patients With Cerebral Palsy

DILETTA BALTA [1], **GIULIO FIGARI**[2], **GABRIELE PAOLINI**[2], **EVELINA PANTZAR-CASTILLA** [3], **JACQUES RIAD**[4,5,6], **UGO DELLA CROCE** [7], **AND ANDREA CEREATTI** [1]

[1]Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy
[2]GPEM srl, 07041 Alghero, Italy
[3]Department of Orthopedics, Örebro University Hospital, 701 85 Örebro, Sweden
[4]Department of Orthopedics, Skaraborg Hospital Skövde, 405 30 Skövde, Sweden
[5]Institute of Clinical Sciences, 405 30 Skövde, Sweden
[6]Sahlgrenska Academy, University of Gothenburg, 413 90 Gothenburg, Sweden
[7]Department of Biomedical Sciences, University of Sassari, 07100 Sassari, Italy

Corresponding author: Diletta Balta (diletta.balta@polito.it)

ABSTRACT Clinical gait analysis is a diagnostic tool often used for identifying and quantifying gait alterations in cerebral palsy (CP) patients. To date, 3D clinical gait analysis protocols based on motion capture systems featuring multiple infrared cameras and retroreflective markers to be attached to the subject's skin are considered the gold standard. However, the need for fully dedicated personnel and space in addition to the inconvenient requirement of multiple markers attached on the patient's body limit their use in the clinical practice. To shorten the time necessary to setup the patient and to limit his/her discomfort motion tracking performed using markerless technologies may offer a promising alternative to marker-based motion capture. This study aims at proposing and validating on 18 CP patients, an original markerless clinical gait analysis protocol based on a single RGB-D camera. Accuracy and reliability of the spatial-temporal parameters and sagittal lower limb joint kinematics were assessed based on a 3D marker-based clinical gait analysis protocol. The smallest percent mean absolute errors were obtained for stride duration (2%), followed by the step and stride length (2.2% and 2.5%, respectively) and by gait speed (3.1%). The average angular offset values between the two protocols were 8° for the ankle, 6° for the knee and 7° for the hip joint. The smallest root mean square error values were found for the knee joint kinematics (3.2°), followed by the hip (3.5°) and the ankle (4.5°). Both protocols showed a good-to-excellent reliability. Thus, this study demonstrated the technical validity of a markerless single-camera protocol for clinical gait analysis in CP population. The dataset containing markerless data from 10 CP patients along with the MATLAB codes have been made available.

INDEX TERMS Markerless gait analysis, RGB-depth camera, cerebral palsy, spatial-temporal parameters, sagittal lower-limb joint kinematics.

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea F. Abate .

I. INTRODUCTION

Clinical gait analysis is fundamental to understand and interpret physio-pathological characteristics of human locomotion and its importance as a clinical diagnostic tool is widely

accepted. In particular, there is strong clinical evidence on its effectiveness in supporting the identification of optimal surgical procedures and the consequent rehabilitation pathways in children with bilateral cerebral palsy (CP) [1].

To date, standard 3D clinical gait analysis protocols are based on the use of infrared multi-camera system to reconstruct the trajectories of markers attached to the skin of the patient on specific locations [2], [3], [4] (MB). Unfortunately, MB protocols routine use is limited by several issues such as the need for highly qualified staff, its high price tag and skin markers acceptance which are very critical when dealing with younger patients.

To reduce the time associated to the subject preparation and discomfort, motion tracking performed using markerless technologies (MS) may offer a promising alternative to MB motion capture. Recently, different MS multi-camera solutions have been proposed for three-dimensional (3D) joint kinematics analysis [5], [6], [7], [8], [9], [10], [11], [12]. However, multi-camera set-up requires time for installation and for extrinsic cameras calibration and therefore it is not the ideal solution for ambulatory settings with no dedicated space.

Conversely, there are applications where a two-dimensional (2D) joint kinematic analysis is still clinically relevant (e.g. for screening purposes, to identify gait patterns, for follow-up over time and to evaluate treatment). For these purposes, system portability, affordability, and user-friendliness are essential requirements. Methods based on the use of a single camera with minimum set-up time would therefore be preferred. Lately, several manufacturers have been producing inexpensive tracking systems (200-400 €/$) featuring an RGB camera integrated with an infrared depth sensor (RGB-depth). By combining the information of the RGB image with depth data, these systems can be used to generate depth color images (2D+) without requiring a multi-camera set-up.

The single camera MS methods proposed in the literature can be grouped into three categories: i) black-box methods either based on software development kit (SDK) integrated with proprietary hardware ([13], [14], [15]) or commercial software (e.g. IPsoft iPi Biomech, MediaPipe Studio), ii) open source methods based on deep learning approaches ([5], [16], [17], [18]), and iii) replicable non-machine learning methods ([19], [20], [21], [22], [23], [24], [25], [26]).

Generally, black-box methods are conceived especially for animation or gaming purposes and are not compliant with clinical standards and terminology [27]. The major limitation of this category is related to its 'black box' functioning resulting in the inability of fine-tuning some model parameters for pathological data at the expense of external validity and performance [28]. In addition, body tracking SDKs are developed for specific hardware solutions and therefore difficult to generalize.

The majority of the open-source methods based on deep learning approaches are often trained on synthetic generic movement data (e.g. AlphaPose, OpenPose) ([5], [16], [17], [18]), not necessarily gait, and training relies on reference data not based on clinical gait analysis standards [29] (e.g. clear and anatomical/functional rules for joint centers definition [30]). Furthermore, original training data sets do not include people with impaired gait and, therefore, methods performance is not optimized and clinical validity is not established.

Thirdly, replicable non machine learning methods have the advantage that they do not require a specific training set although they need to be optimized for the specific problem and their performance is not expected to improve with the dataset size. The most common factors limiting the clinical applicability of most past studies included the use of color filter and homogeneous background for subject segmentation ([19], [20]), single joint analysis ([20], [21], [22], [23]), lack of technical validation against gold standards and on pathological populations ([19], [23], [24], [25], [26]).

The aim of the present study is to propose and validate on 18 CP patients, an original MS clinical gait analysis protocol based on a single RGB-D camera. Accuracy and reliability of the sagittal lower limb joint kinematics and spatial-temporal parameters were assessed based on a 3D MB clinical gait analysis protocol.

II. MATERIAL AND METHODS

A. SUBJECTS
Gait data were collected from 18 participants, 4 females and 14 males, age between 6.5 and 28 years old (mean 15 y.o.). Most participants showed bilateral CP (11), some showed unilateral CP (3), some suffered from dyskinetic CP (3), and one from ataxic CP. In the Gross Motor Function Classification System (GMFCS), six of them were classified at level I, eleven at level II, and one at level III. The study was approved by the regional ethical review board in Gothenburg, Sweden (approval number 660-15).

B. EXPERIMENTAL PROTOCOL
Instrumentation - An RGB-depth camera (Kinect 2 for Xbox One, Microsoft, RGB images: 1920 × 1080 pixels at 30 fps, FOV = 84.1° × 53.8°; Depth images: 512 × 424 pixels at 30 fps, FOV = 70.6° × 60°) was positioned laterally at a 2.5-meter distance from the center of the walkway and a 5-meter distance from the background. The total Kinect capture volume was 7.08 m (length) × 5.77 m (height) × 5 m (width). The image coordinate system (I) of the video camera was aligned to the sagittal plane (x_I, y_I) identified by the direction of progression and the vertical direction. To prevent blurred images due to automatic exposure, two additional cool-white LED lamps (3360 Lux/m) were used (Fig.1).

Subject preparation – Each subject was asked to wear colored ankle socks (red for the right and blue for the left) and underwear. External anatomical landmarks including the lateral malleolus (LM), lateral epicondyles (LE), great trochanter (GT), anterior superior iliac spine (ASIS) and
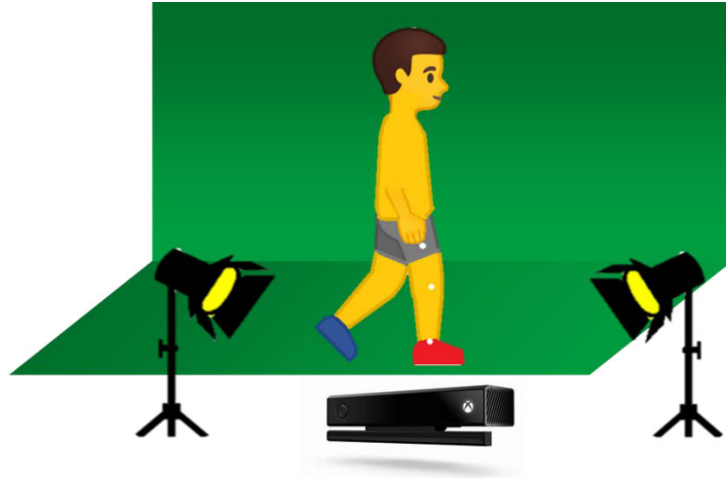
FIGURE 1. Experimental setup.

posterior superior iliac spine (PSIS) were identified by palpation by an expert operator and marked with a black felt pen.

Data collection – Two static lateral views (right and left side) of the subject while standing upright were captured at the beginning of the experimental session. Participants were then asked to walk at a comfortable self-selected speed along a straight 10-meter walkway. Ten gait trials per subject were recorded including five right and left full gait cycles. The dataset containing MS data from 10 CP patients has been uploaded on IEEE DataPort (RGB-Depth_CP_patients_POLITO_dataset | IEEE DataPort (ieee-dataport.org)).

Validation – a 12-camera stereo-photogrammetric system (Oqus 400 Qualisys medical AB, Gothenburg, Sweden) was used to collect 3D reference data at 100 fps. The capture volume was of 14 m × 3 m × 8 m and completely included the Kinect volume. Thirty-eight retro-reflective spherical markers (14 mm diameter) were attached to the subjects according to the modified Helen-Heyes protocol [2]. Calculations of 3D reference joint angles were performed using the Visual 3D software (C Motion Inc., USA).

C. IMAGE PRE-PROCESSING

Calibration refinement and camera lens correction was implemented using the Heikkilä undistortion algorithm ([31], [32]). A matching operation was carried out by using intrinsic and extrinsic parameters obtained from the calibration refinement of both RGB and Depth sensor to overlap RGB and Depth images of the same size (N_{row} = 1080, N_{col} = 1536).

D. METHOD DESCRIPTION

The proposed method consisted of four main stages: gait cycle identification, subject segmentation, subject-specific models calibration and joint center trajectories estimation (Fig.2). MATLAB codes have been made available on Github (https://github.com/dilettabalta/ModelBased_MarkerlessProtocol.git).

1) GAIT CYCLE IDENTIFICATION

From each gait trial, the most central gait cycle was selected and analyzed based on the identification of initial foot contacts. To this purpose, a specific algorithm was developed to account for different types of foot-ground contacts commonly encountered in subjects with CP as shown in Fig. 3 [33].

Specifically, for each video frame, a binary segmentation mask $^{I}M_{foot}$ expressed in the image coordinate system I, was obtained for each foot by applying a segmentation technique based on color filters [34]. An ellipse was fitted on each $^{I}M_{foot}$ (Fig.4a). Then, a foot coordinate system (f) was defined with the axes coincident to the inertial ellipsoid principal axes and the origin coinciding with the centroid. The transformation matrix $^{I}T_{f}$ from f to I was computed by simple geometrical rules and applied to transform $^{I}M_{foot}$ in the f ($^{f}M_{foot}$). From $^{f}M_{foot}$, the point Q ($^{f}Q = [Q_{xf}, Q_{yf}]$) with the highest y-coordinate was identified, the foot points included between Q_{yf} and ($Q_{yf} - \epsilon$) were isolated and the least square fitting line was computed to approximate the sole of the foot (Fig.4b). A similar procedure was implemented to reconstruct the line approximating the posterior contour of the foot starting from point R with the lowest x-coordinate (Fig. 4b).

The mid-rear foot (MRF) position in f ($^{f}MRF = [MRF_{xf}, MRF_{yf}]$) was identified as the intersection between the sole and posterior lines of the foot, while the forefoot (FF) position in f ($^{f}FF = [FF_{xf}, F F_{yf}]$) was identified as the point with the highest x-component (Fig.4). The ^{f}MRF and ^{f}FF were then transformed in I based on $^{I}T_{f}$.

The foot points MRF and FF were assumed to be in contact with the ground when the vertical velocity along y_{I} and horizontal velocity along x_{I} were below a given threshold Th (Fig.5a). Then, the foot initial contact (IC) was identified as the first instant of time characterized by zero-velocity between MRF and FF.
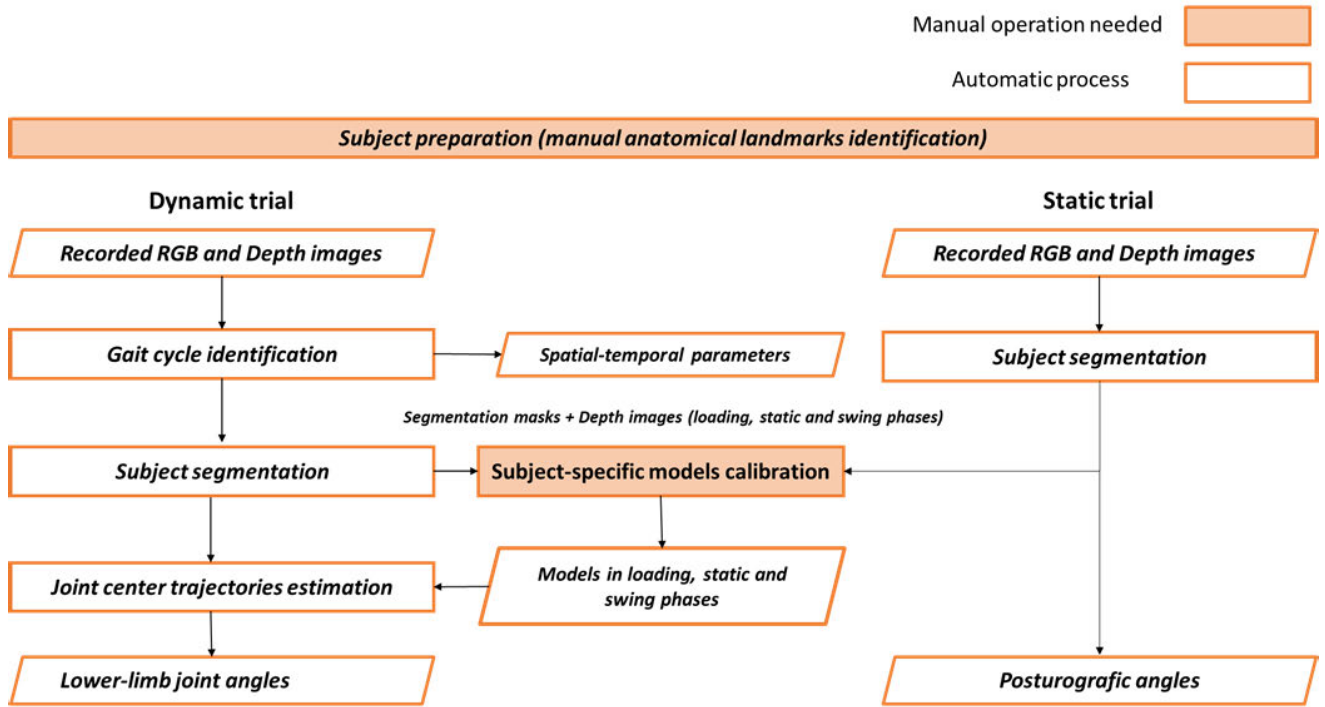
FIGURE 2. Block-diagram of the proposed MS protocol.
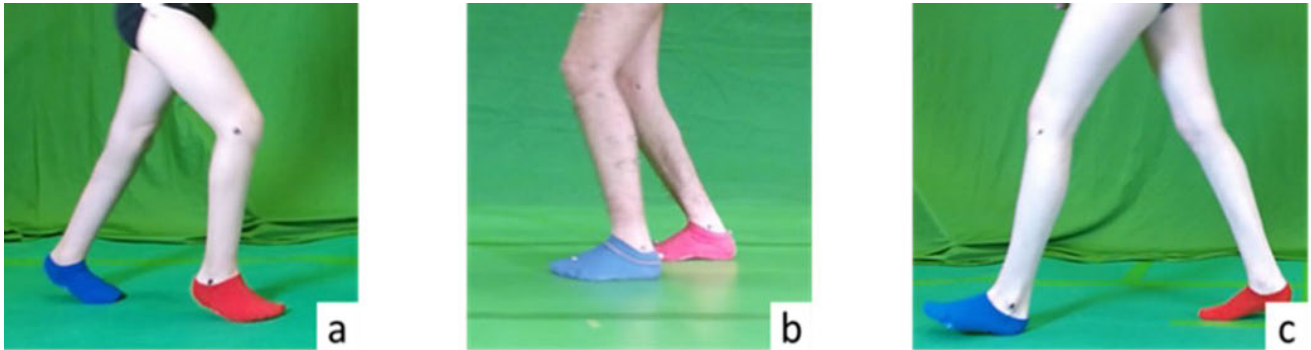
FIGURE 3. Different types of foot contacts: a) fore-foot contact in equinus gait, b) Foot-flat contact in individuals who walk in a crouch gait with excessive knee flexion, c) rear-foot contact in patients classified at a low level of the GMFCS.

The gait cycle was identified by two consecutive ICs of the foreground foot, whereas the step was identified by the first IC of the foreground and the subsequent IC of the background foot.

Spatial-temporal parameters (stride and step length and stride duration) were calculated based on the positions of the relevant foot points (MKF/FF) at the IC instants (Fig.5b). Finally, gait speed was computed by considering pixel/meter conversion factor (see appendix A for a detailed description).

2) SUBJECT SEGMENTATION

For each frame, a preliminary background subtraction was performed between the RGB image $^I\boldsymbol{I}$, containing the subject, and the image background $^I\boldsymbol{B}$ to obtain the difference image

$^I\boldsymbol{D}$ as follows:

$$^I D\,(x,y,c) = \left|{}^I I\,(x,y,c) - {}^I B\,(x,y,c)\right|$$

where $I_D(x,y,c)$, $^I I(x,y,c)$, and $I_B(x,y,c)$ are the generic pixels expressed in i and $c = [r,g,b]$ is the color channel vector.

The resulting difference image $^I\boldsymbol{D}$ was converted to grayscale $^I\boldsymbol{D}_{gray}$ by computing the norm of color channel of each pixel:

$$^I D_{gray}\,(x,y) = \sqrt{{}^I D\,(x,y,r)^2 + {}^I D\,(x,y,g) + {}^I D\,(x,y,b)^2}$$

The subject was separated from the background by applying a proper threshold on the grey levels of the image pixels. The threshold level was set to the weighted mean of the grayscale
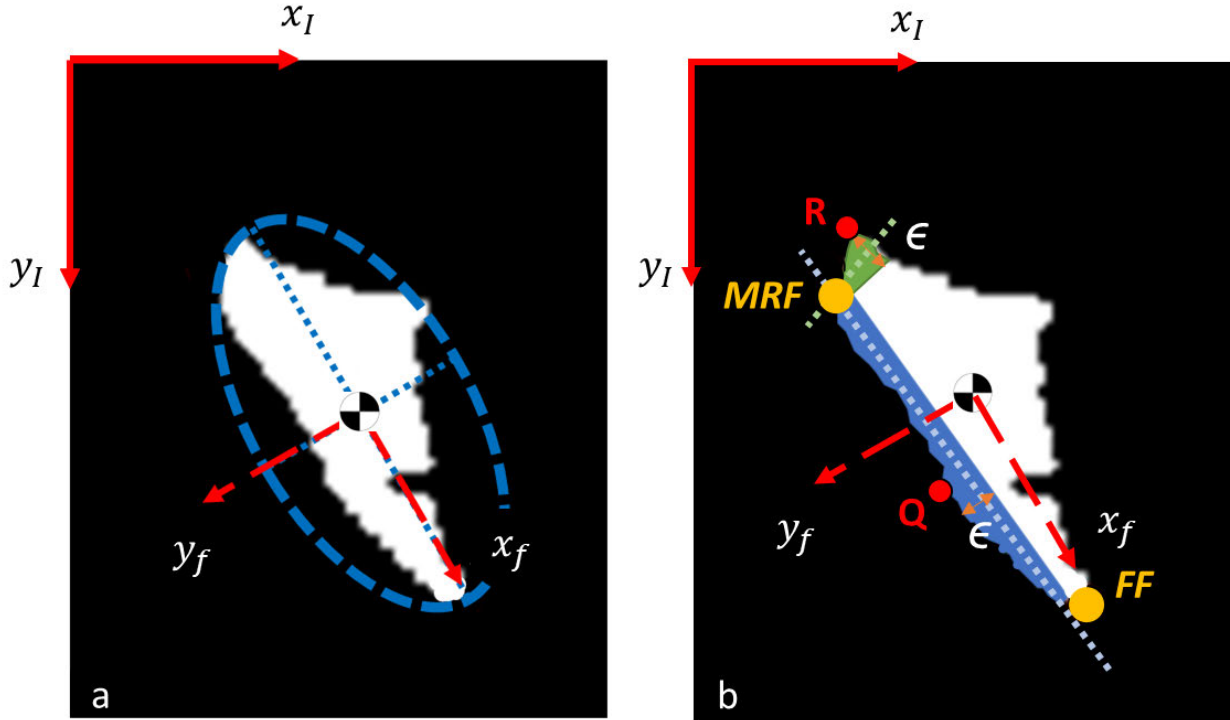
FIGURE 4. Identification of the MRF and FF. a). An ellipse was fitted on each foot; the centroid and the principal axes (x_f, y_f) were identified. b). The intersection between the sole of the foot (light blue area) and the posterior area (light green area) of the foot was identified as the mid-rear foot (MRF) while the extremity of the foot along the x- axis was identified as the forefoot (FF).

histogram [35]:

$$Th = \frac{\sum_{i=0}^{255} w_i \cdot g_i}{\sum_{i=0}^{255} w_i}$$

where w_i is the histogram count (occurrence) for the i-th grayscale level (g_i: 0,...,255).

The segmentation mask $^I\boldsymbol{M}_{sub}$ was obtained from the $^I\boldsymbol{D}_{gray}$ as follows:

$$^I\boldsymbol{M}_{sub}(x, y) = \begin{cases} 1, & |^I\boldsymbol{D}_{gray}(x, y) \geq Th \\ 0, & |otherwise \end{cases}$$

Undesired residual small regions due to noise or time-variant shadows were removed under the assumption that the subject is associated to the largest connected area.

The feet segmentation was then refined implementing a color filter technique exploiting the use of colored socks to avoid inaccuracies due to the presence of shadows during foot approach to the ground.

3) MULTI-SEGMENTAL MODEL DEFINITION

A 2D subject-specific kinematic lower limb model was introduced to estimate lower limb joint angles. The model consisted in four body segments (foot, shank, thigh and pelvis) connected by revolute joints (ankle, knee and hip joints) for a total of 6 degree of freedom (DoF). The foot

segment was assumed to be the parent segment and its motion was characterized by two translational and one rotational DoFs. The ankle joint (AJC) was centered with the lateral malleolus (LM), the knee joint center (KJC) with the lateral epicondyles (LE) and the hip joint (HJC) with the great trochanter (GT).

4) ANATOMICAL CALIBRATION AND BODY SEGMENT TEMPLATES DEFINITION

The body segments' templates and the relevant coordinate systems were calibrated on the static upright standing acquisition (image "0") by manually selecting the anatomical landmarks (LM, LE, GT, ASIS, PSIS) to obtain their position vectors in I ($^I\boldsymbol{LM}_0$, $^I\boldsymbol{LE}_0$, $^I\boldsymbol{GT}_0$, $^I\boldsymbol{ASIS}_0$, $^I\boldsymbol{PSIS}_0$). The identification of the point MRF and FF ($^I\boldsymbol{MRF}_0$ and $^I\boldsymbol{FF}_0$) was obtained following the procedure presented in par.I. To account for potential right/left asymmetries, the subject-specific model was defined for both sides.

a: FOOT TEMPLATE

From $^I\boldsymbol{M}_{foot}$, the mid-rear foot portion was extracted to define a template $^I\boldsymbol{TMP}_{foot}$ where the value of its generic pixel $^I\boldsymbol{TMP}_{foot}(x, y)$ in the I was obtained as:

$$^I\boldsymbol{TMP}_{foot}(x, y)$$
$$= \begin{cases} 1, & |^I\boldsymbol{M}_{foot}(x, y) = 1 \cap MRF_{xi} < x < MRF_{xi} + 0.9l_f \\ 0, & |otherwise \end{cases}$$
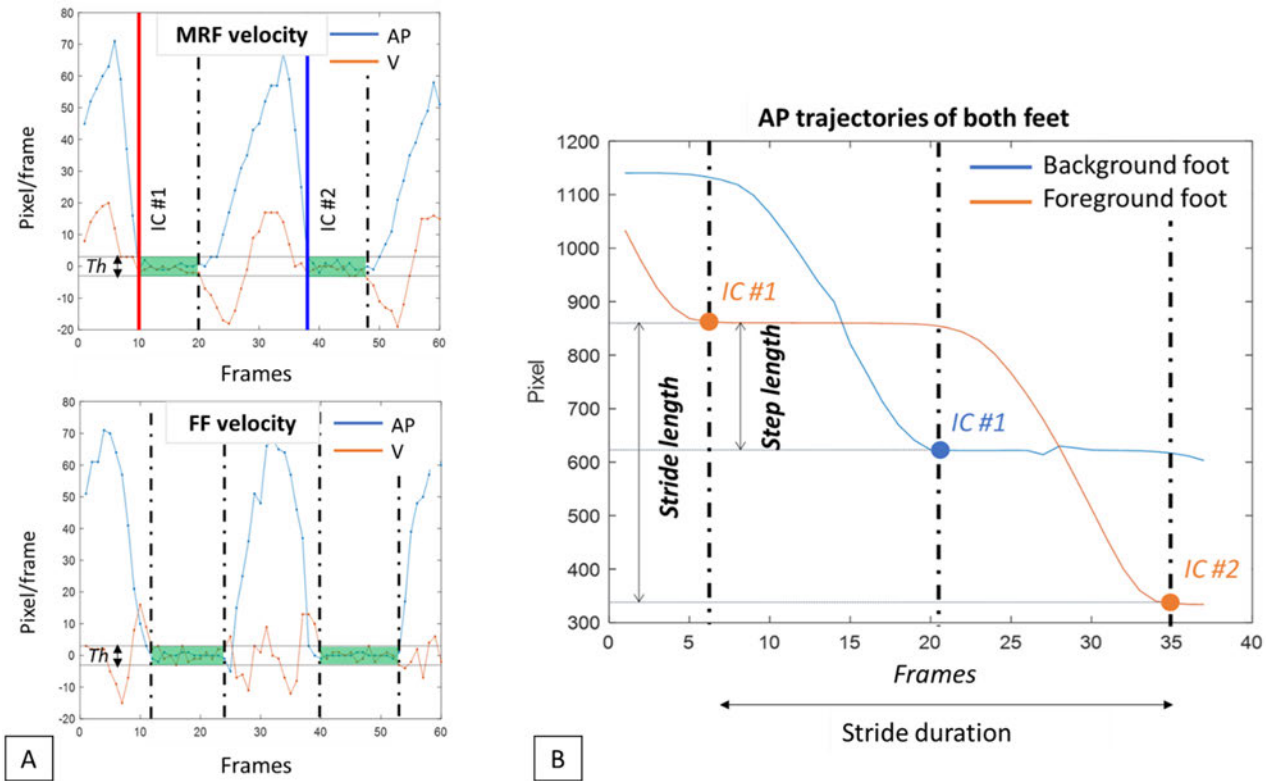
FIGURE 5. Computation of stride length and duration, step length and gait speed. A) Velocity of MRF and FF coordinates. The red line in bold defines the first initial contact (IC #1) while the blue one represents the following initial contact (IC #2). The green areas represent the intervals in which MRF and FF are assumed to be in contact with the ground (stationary condition). B) The stride length is the distance between two consecutive initial contacts of the foreground foot (IC #1 and IC#2 in orange). Step length is the distance between initial contact of the foreground foot (IC #1 in orange) and the initial contact (IC #1 in blue) of the contralateral one.
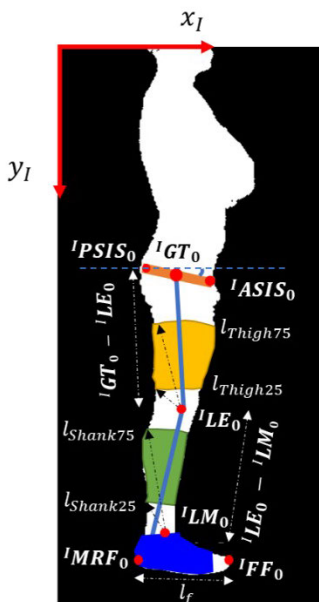
FIGURE 6. Body segment templates definition for the right side.

where $^{I}M_{foot}(x, y)$ is a generic pixel of $^{I}M_{foot}$ expressed in the I, l_f is the distance between $^{I}MRF_0$ and $^{I}FF_0$ (Fig.6).

The foot coordinate system f_0 was defined as described in par. I. and the transformation matrix $^{I}T_{f0}$ from f_0 to I determined and applied to transform $^{I}TMP_{foot}$ in the f_0 ($^{f0}TMP_{foot}$).
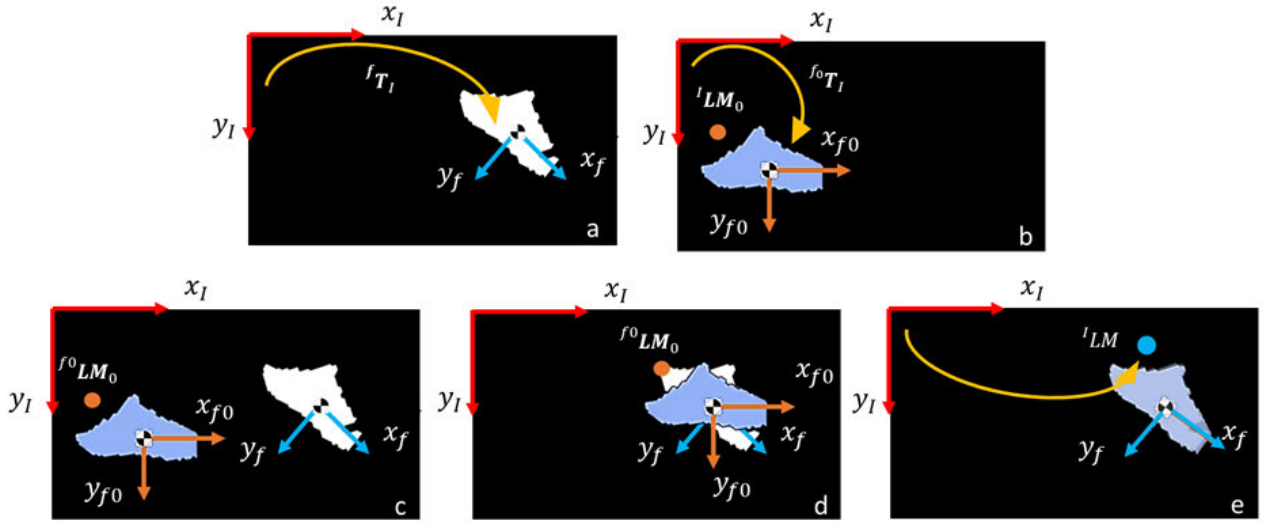
b: SHANK TEMPLATE

The central shank portion was extracted as the region included in the anulus centered in $^{I}LM_0$ and defined by the radius $l_{shank25}$ and the radius $l_{shank75}$ equal to the 25% and the 75% of the distance between $^{I}LM_0$ and $^{I}LE_0$, respectively (Fig.6).

Then, the generic pixel $^{I}TMP_{shank}(x, y)$ of $^{I}TMP_{shank}$ in I was obtained as:

$$^{I}TMP_{shank}(x, y)$$
$$= \left\{ \begin{array}{l} 1, \ |^{I}M_{sub}(x, y) = 1 \cap l_{shank25} < \sqrt{x^2 + y^2} < l_{shank75} \\ 0, \ |otherwise \end{array} \right\}$$

An ellipse was fitted on $^{I}TMP_{shank}$. Then, a shank coordinate system ($s0$) was defined with the axes coincident to the inertial ellipsoid principal axes and the origin coinciding with the centroid. The transformation matrix $^{I}T_{s0}$ from $s0$ to I was computed by simple geometrical rules and applied to transform $^{I}TMP_{shank}$ in the $s0$ ($^{s0}TMP_{shank}$).

FIGURE 7. Ankle joint center estimation. a) $^IM_{foot}$ and its relevant ^IT_f b) $^ITMP_{foot}$ with $^ILM_0$. c) $^ITMP_{foot}$ and $^IM_{foot}$ in the common I. d) The origins of f and f_0 were made to coincide and e) the $^{f_0}TMP_{foot}$ was matched with the $^fM_{foot}$ and the relevant matrix $^{f_0}T_f$ determined.

c: THIGH TEMPLATE

The central thigh portion was extracted as the region included in the anulus centered in $^ILE_0$ and defined by the radius $l_{thigh\,25}$ and the radius $l_{thigh\,75}$ equal to the 25% and the 75% of the distance between $^ILE_0$ and $^IGT_0$, respectively. (Fig.6).

Then, the generic pixel $^ITMP_{thigh}(x, y)$ of $^ITMP_{thigh}$ in I was obtained as:

$$^ITMP_{thigh}(x, y)$$
$$= \begin{Bmatrix} 1, & \left|^IM_{sub}(x, y) = 1 \cap l_{thigh25} < \sqrt{x^2 + y^2} < l_{thigh75} \right. \\ 0, & \left|otherwise \right. \end{Bmatrix}$$

Similarly to the $^ITMP_{shank}$, the thigh coordinate system t_0 and the transformation matrix $^I_0T_{t_0}$.from t_0 to I were defined and applied to transform $^ITMP_{thigh}$ in the t_0 ($^{t_0}TMP_{thigh}$).

d: PELVIS

The pelvis inclination, with respect to the x_I, was determined during the static upright standing acquisition based on the positions of the $^IASIS_0$ and $^IPSIS_0$ (Fig.6).

5) JOINT CENTERS TRAJECTORIES ESTIMATION

For each frame of the gait cycle, the joint center positions were identified following a bottom-up tracking approach from the foot to the pelvis.

a: ANKLE JOINT CENTER (AJC) ESTIMATION

The foreground foot was extracted from the RGB image based on color filters ($^IM_{foot}$) and the posterior part of the foot was isolated following the same procedure presented in the par. IV. The foot coordinate system f and the relevant transformation matrix ^IT_f was defined as described in par. I (Fig.7a).

After having expressed $^ITMP_{foot}$ (Fig.7b) and $^IM_{foot}$ in the common I (Fig.7c), the origins of f and f_0 were first

made to coincide (Fig.7d) and then, using an iterative closest point (ICP) technique [36], the $^{f_0}TMP_{foot}$ was matched with the $^fM_{foot}$ and the relevant matrix $^{f_0}T_f$ (4×4), determined (Fig.7e).

Finally, the ^IAJC position in I, coincident to^ILM, was obtained for each frame based on the position of LM in the template $^ILM_0$ by applying the three subsequent transformations:

$$^IAJC \equiv^I LM =^I T_f^f T_{f_0}^{f_0} T_I^I LM_0$$

### b: KNEE JOINT CENTER (KJC) ESTIMATION

The separation between the foreground and background shanks was carried out using two alternative strategies depending on whether there was or was not overlap between foreground and background shanks.

To discriminate between overlap/non overlap conditions, a circle centered in ^ILM with radius equal to the distance between $^ILM_0$ and $^ILE_0$ was drawn. If there was no overlap, the segmentation mask $^IM_{sub}$ were grouped in two separated regions, and the foreground shank, being closer to the camera, coincided with the largest area (Fig.8a).

Conversely, when there was overlap, a single connected region was found, and auxiliary depth sensor data were used to separate foreground and background shanks (Fig. 8b). To this purpose, the histogram of depth values within the region was computed and the Otsu method [37] was applied for a binary classification (class 0: foreground shank, class 1: background shank) based on the minimization of the variance between classes.

The central portion of the foreground shank ($^IM_{shank}$) was extracted as the region included in the anulus centered in $^ILM_0$ and defined by the radius $l_{shank25}$ and the radius $l_{shank75}$.
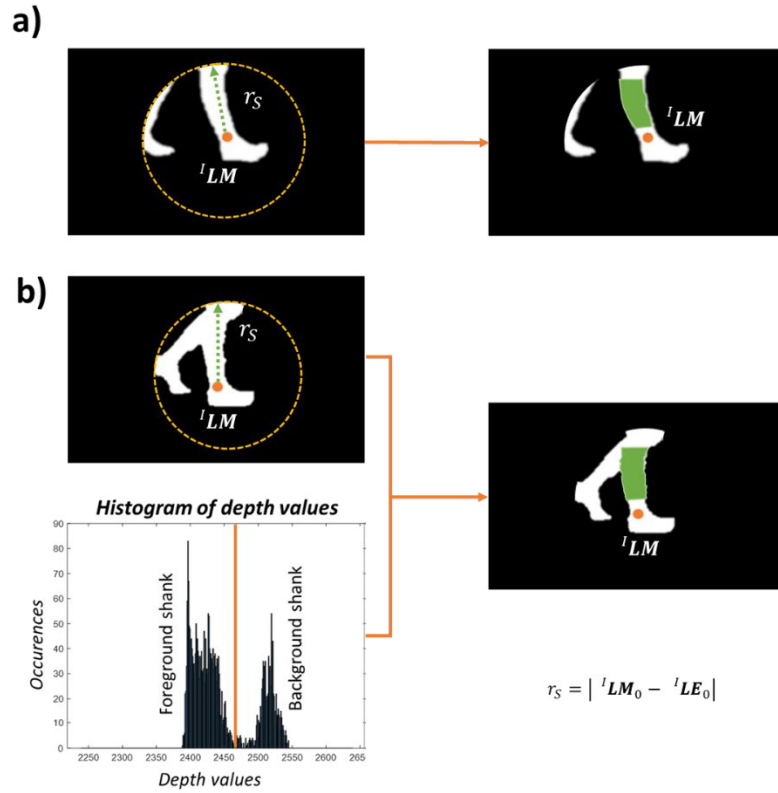
FIGURE 8. Separation between foreground and background shanks. A circle centered in $^I LM$ with radius equal to the euclidean distance between $^I LM_0$ and $^I LE_0$ was drawn. A) No-overlap between the two shanks. On the left, Two separated regions were found. On the right, the foreground shank $^I M_{shank}$. was identified B) Overlap between the two shanks. On the left, a single connected region was found, the histogram of depth values inside the region was computed and the Otsu method was implemented for separating the two shanks. On the right, the foreground shank $^I M_{shank}$ was identified.

The shank coordinate system s was defined with the axes coincident to the inertial ellipsoid principal axes of the $^I M_{shank}$ and the transformation matrix $^I T_s$. from s to I computed. Similarly to the ankle joint center estimation, the origin of s and s_0 was made to coincide and then, using an ICP technique [36], the $^{s_0} TMP_{shank}$ was matched with the $^s M_{shank}$ and the relevant matrix $^{s_0} T_s$ (4×4), determined.

Finally, the $^I KJC$ position in I, made to coincide with $^I LE$, was obtained for each frame based on the position of LE in the shank template $^I LE_0$ by applying the three subsequent transformations:

$$^I KJC \equiv {}^I LE = {}^I T_s {}^s T_{s_0} {}^{s_0} T_I {}^I LE_0$$

c: HIP JOINT CENTER (HJC) ESTIMATION
To separate foreground thigh from the background thigh and the hand during arm oscillation, two alternative procedures were implemented depending on whether the foreground hand was superimposed to the foreground thigh or not. Preliminarily, a circle centered in $^I LE$ with radius equal to the distance between $^I LE_0$ and $^I GT_0$ was drawn and the envelope

of the histogram of depth values of the pixels within this circle was computed and the maxima were identified.

In case of foreground hand superimposition, three peaks, corresponding to the foreground hand (class 0), foreground thigh (class 1) and background thigh (class 2), were found on the histogram envelope (Fig.9a). Then, the Otsu method [37] was applied for a three-classes classification. Alternatively (Fig.9b), a binary classification was implemented (class 0: foreground thigh, class 1: background thigh).

The central portion of the foreground thigh ($^I M_{thigh}$) was extracted as the region included in the anulus centered in $^I LE$ and defined by the radius $l_{thigh25}$ and the radius $l_{thigh75}$.

The thigh coordinate system t was defined with the axes coincident to the inertial ellipsoid principal axes of the $^I M_{thigh}$ and the transformation matrix $^I T_t$. from t to I computed. The origin of t and t_0 was made to coincide and then, using an ICP technique [36], the $^{t_0} TMP_{thigh}$ was matched with the $^t M_{thigh}$ and the relevant matrix $^{t_0} T_t$ (4×4), determined.

Finally, the $^I HJC$ position in I, made to coincide with $^I GT$, was obtained for each frame based on the position of GT in the shank template $^I GT_0$ by applying the three subsequent
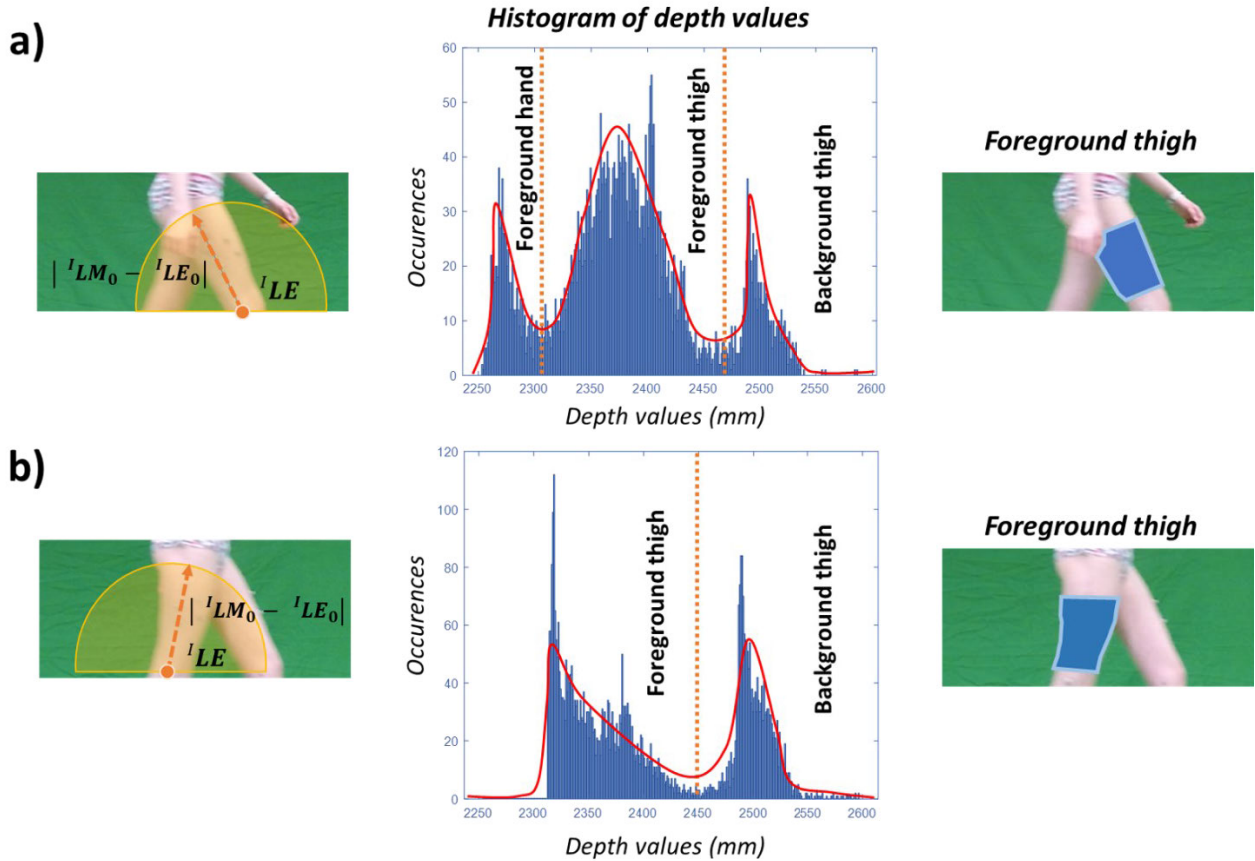
FIGURE 9. Separation between foreground and background thighs. A circle centered in $^I LE$ with radius equal to the euclidean distance between $^I LE_0$ and $^I GT_0$ was drawn, the histogram of depth values inside the region was computed and the envelope was calculated. a) foreground hand superimposed to the foreground thigh. The envelope was characterized by three peaks and the Otsu method was implemented for a three-classes classification. b) No foreground hand inside the circular region. Only two peaks were present in the envelope and the Otsu method was implemented for binary classification.

transformations:

$$^I HJC \equiv {}^I GT = {}^I T_t \, {}^t T_{t_0} \, {}^{t_0} T_I \, {}^I GT_0$$

6) SUBJECT-SPECIFIC MODELS CALIBRATION

It must be highlighted that within the recorded gait cycle, size and shape of the lower limb body segments vary due to soft tissue deformation [38], changes in the subject position relative to the camera field of view, and potential out-of-plane movements, thus limiting the effectiveness of the matching procedure between the body segment templates and the segmented body segment masks. To overcome these limitations, a multiple calibration procedure [39] was implemented based on three sets of body segment templates, the first defined from the standing posture (Fig 10a), the second and the third from frames selected during the loading and the swing phases of the gait cycle, respectively (Fig.10b, and 10c). The procedure for the identification of joint centers trajectories as described in par.V, was then repeated using the additional templates, thus obtaining three different trajectories for each center.

7) JOINT KINEMATICS ESTIMATION

Joint kinematics was determined based on the segment inclination as defined by the lines connecting the joint centers. For the ankle, the plantar-dorsi flexion angle was determined as the angle between $^I AJC$-$^I TOE$ and the $^I AJC$-$^I KJC$ vectors, the knee joint flexion-extension angle was determined as the angle between the $^I AJC$-$^I KJC$ and $^I KJC$-$^I HJC$ vectors. The-hip joint flexion-extension angle was determined as the angle between the $^I KJC$-$^I HJC$ vector and the time-invariant direction identified by the $^I ASIS_0$-$^I PSIS_0$ vector (pelvic tilt) during the standing posture. For each joint, three kinematic curves were obtained based on the three sets of body segment templates. These curves were then combined into a single curve by a nonlinear sinusoid weight function [40] based on the percentage of the gait cycle.

E. PERFORMANCE ASSESSMENT AND STATISTICAL ANALYSIS

The accuracy of the gait events identification was evaluated by computing the time difference in terms of the mean
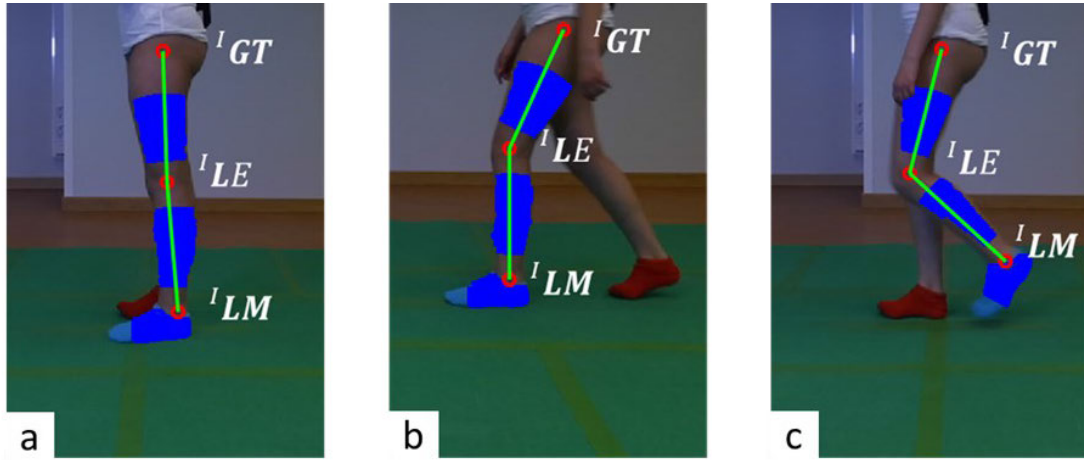
FIGURE 10. Set of body segment templates definition during static phase (a), loading phase (b) and swing phase (c).
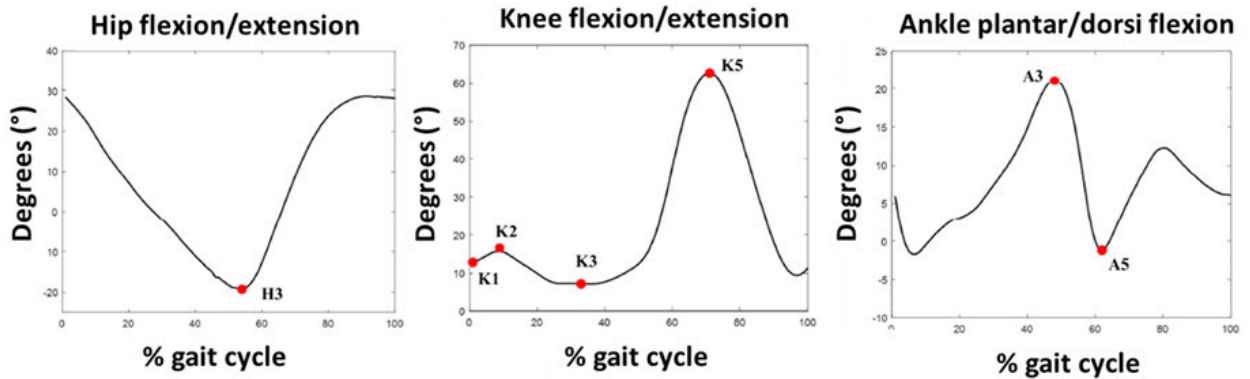

FIGURE 11. Key gait features extracted from sagittal lower limb joint kinematics.

absolute error (MAE) and mean error (ME) between the gait events found by visual inspection from the RGB images and those estimated by the automatic MS method over trials and subjects.

The spatial-temporal gait parameters estimated were assessed in terms of MAE, MAE%, ME, ME% with respect to the estimates provided by the 3D MB protocol over trials and subjects.

Before comparison, both the MS and MB kinematic curves were filtered using a fourth order Butterworth filter (cut off frequency at 7 Hz) and were time-normalized to the gait cycle (1-100%) [41].

For each subject, gait trial and joint, the performance of the proposed MS method were assessed in terms of offset and waveform similarity [42]. The offset was computed as the absolute difference between the mean value of the MS ($\overline{MS}$) and MB kinematic curves ($\overline{MB}$) within a gait cycle:

$$Offset_{s,t,j} = \left| \overline{MB_{s,t,j}} - \overline{MS_{s,t,j}} \right|$$

For each joint, the latter values were then averaged across trials and subjects:

$$Offset_j = \frac{1}{N_S} \sum_{s=1}^{N_S} \frac{1}{N_T} \sum_{t=1}^{N_T} Offset_{s,t,j}$$

where $N_T = 10$ is the number of trials and $N_S = 18$ is the number of subjects.

For each subject, gait trial and joint, the waveform similarity was evaluated as the root mean square error (RMSE) of the MS joint kinematic curves with respect to the MB joint kinematic curves, after removing their mean values [42]:

$$RMSE_{s,t,j} = RMS\left(\left(MB_{s,t,j} - \overline{MB_{s,t,j}} \right) - \left(MS_{s,t,j} - \overline{MS_{s,t,j}} \right) \right)$$

For each joint, the latter values were then averaged across trials and subjects:

$$RMSE_j = \frac{1}{N_S} \sum_{s=1}^{N_S} \frac{1}{N_T} \sum_{t=1}^{N_T} RMSE_{s,t,j}$$

where $N_T = 10$ is the number of trials and $N_S = 18$ is the number of subjects.

In addition, a set of clinically relevant key gait features were extracted according to [43] from the MB and MS sagittal lower limb joint kinematics after offset removal (Fig.11):

K1: the knee flexion at the initial contact (0% of the gait cycle);

K2: the knee maximum flexion during the loading response (0 - 40% of the gait cycle);

K3: the knee maximum extension during the stance phase (25- 75% of the gait cycle);

K5: the knee maximum flexion during the swing phase (50 - 100% of the gait cycle);

A3: the ankle maximum dorsiflexion during the stance phase (25 - 75% of the gait cycle);

A5: the ankle maximum dorsiflexion during the swing phase (50- 100% of the gait cycle);

H3: the hip maximum extension during the stance phase (25 - 75% of the gait cycle).

For each key gait feature, the MAE and the ME were computed with respect to the MB estimates along with 95% confidence intervals (95% CI) were computed. Normality of the key gait features distributions was assessed by applying Shapiro–Wilk test. After verifying normality, a two-sample t-test with a significant level of 95% was implemented to quantify differences between the MS and MB methods. A p-value of ME and MAE less than 0.05 were considered statistically significant.

For each gait feature and each method (MS and MB), the reliability was evaluated with intraclass correlation based on absolute agreement and 2 way random effects (ICC$(2,k)$) computed based on the formulas reported in [44] based on the data collected over subjects (n = 18) for the different gait cycles (k = 10).

ICC values less than 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability [44].

Spearman's correlation coefficient (R) was used to correlate the differences between MS and MB systems. The estimated values, derived from R, can be interpreted as follows: values below 0.19 indicate a negligible relationship, values between 0.2 and 0.29 suggest a weak relationship, values between 0.3 and 0.39 indicate a moderate relationship, values between 0.4 and 0.69 imply a strong relationship, and values greater than 0.70 signify a very strong relationship. [45]

III. RESULTS

Results for gait events identification, spatial-temporal gait parameters and lower limb joint kinematics, are reported in Table 1. Regarding spatial-temporal parameters, the smallest errors in terms of MAE % values were obtained for stride duration (2%), followed by the step and stride length (2.2% and 2.5%, respectively) and by gait speed (3.1%).

The RMSE values computed for lower-limb joint kinematics ranged between 3.2° and 4.5°, smallest RMSE values were found for the knee joint kinematics (3.2°), followed by the hip (3.5°) and the ankle (4.5°).

Results related to the extracted key gait features in terms of ME, MAE and their 95% CI are summarized in Table 2. Overall, MAE values were significantly different from zero and ranged from 3.1° to 5.9°.

Results for ICC$(2,k)$ and R for both MS and MB protocols, are reported in Table 3. Both MS and MB measurements

TABLE 1. Mean absolute error and mean error between the visual inspection on the RGB images and the automatic identification of initial contacts for both feet in seconds. Gait spatial parameters. Mean absolute error (MAE), mean error(ME) of the gait speed, stride length, stride duration and step length. Lower limb joint kinematics. The average root-mean-square errors (RMSE) value between the joint kinematics curves estimated by the MS method and the MB system are computed over the gait cycle and averaged across trials and subjects.

Gait events	ME	MAE
IC identification	4 ± 20	20 ± 20
Spatial-temporal parameters	ME (ME%)	MAE (MAE%)
Gait speed (m/s)	-0.01 (1.5 ± 3)	0.02 (3.1 ± 1.7)
Stride length (cm)	-0.7 (0.8 ± 2.5)	2.2 (2.5 ± 0.9)
Stride duration (ms)	-6.4 (0.7 ± 2.5)	20 (2.0 ± 1.6)
Step length (cm)	-0.1 (0.3 ± 2.6)	1.2 (2.2 ± 1.1)
Joint angles	Offset (°)	RMSE(°)
Ankle	8	4.5
Knee	6	3.2
Hip	7	3.5

TABLE 2. Mean error (ME) and mean absolute error (MAE) between MS and MB protocol in the estimation of gait features. 95 % CI: 95% of confidence interval.

Gait features (°)	ME (95% CI)	MAE (95% CI)
A3	3.4 (-3.9, -2.9)*	3.5 (3.1, 4.1) *
A5	-2.2 (2.4, 1.8)*	3.1 (2.6, 3.6) *
K1	1.7 (-1.9, -1.4)*	3.7 (3.2, 4.3) *
K2	4.5 (-5.2, -3.8)*	4.8 (4.1, 5.6) *
K3	-3.5 (4.0, 3.0)*	3.8 (3.3, 4.5) *
K5	5.5 (-6.3, - 4.7)*	5.9 (5.0, 6.8) *
H3	-2.7 (3.1, 2.3)*	3.7 (3.2, 4.3)*

TABLE 3. Comparison of estimates from the MS and MB protocol for each gait feature. MS ICC(2,k): intraclass correlation for MS protocol, MB ICC(2,k): intraclass correlation for MB protocol, R: Spearman's correlation coefficient.

| Gait features (°)	MB ICC(2,k)	MB ICC(2,k)	R
A3	0.88	0.80	0.66
A5	0.80	0.81	0.66
K1	0.91	0.94	0.85
K2	0.90	0.92	0.85
K3	0.93	0.94	0.94
K5	0.93	0.94	0.90
H3	0.94	0.94	0.86

revealed excellent reliability for K1, K2, K3, K5 and H3 (ICC = 0.90-0.94) while for A3 and A5 both protocols showed a good reliability (ICC = 0.80-0.88). Correlation between MB and MS kinematics ranged from very strong for all knee and hip gait features (≥ 0.85) and strong (= 0.66) for ankle kinematics features.

An ensemble view of the normalized joint kinematics curves, averaged over trials and subjects, are reported in Fig.12.

IV. DISCUSSIONS

The aim of the study was to present and evaluate accuracy and reliability of a clinical markerless gait analysis protocol
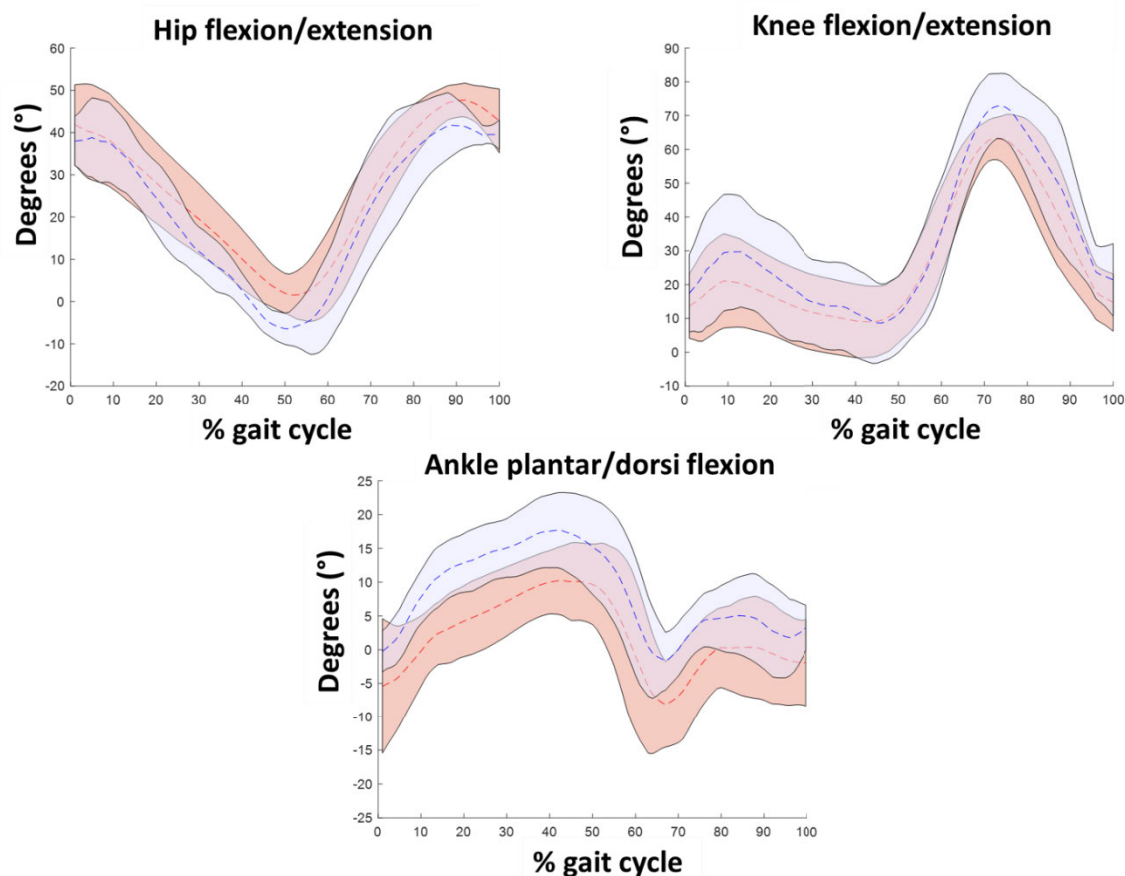
FIGURE 12. Sagittal lower limb joint kinematics (hip, knee and ankle) averaged over subjects and trials (average: dashed lines; SD: shaded area; red = MB system; blue = MS system).

based on the use of a single RGB-depth camera for estimating spatial-temporal parameters and sagittal lower limb joint kinematics on patients with cerebral palsy. The protocol was specifically devised to provide a quantitative tool which could be easily implemented for screening and monitoring disease-related motor progression in patients with CP.

The proposed protocol includes several fundamental improvements with respect to previous research [20]. First, an automatic thresholding segmentation algorithm was proposed which does not require the use of any homogeneous background, thus improving the system clinical applicability and portability. Second, potential issues associated with left-right confusion [46] and skeleton tracking in presence of foreground-background segments overlap [20] were addressed by introducing a robust separation approach which relies on the maximization of the variance between classes based on depth data. Third, the analysis of the clinical concurrent validity of the method was not limited to the knee kinematics as in [20], but also extended to the ankle and hip kinematics and spatial-temporal parameters. The abovementioned improvements, besides increasing protocol robustness to variations in the experimental conditions, also contributed to an increase of about 35% of the accuracy of the knee joint gait features [20].

It is worth noting that the method for the detection of the gait events was specifically conceived for taking into account the different types of foot contact normally observed in CP patients (i.e. heel foot, flat foot and toe-ground contacts). The proposed method relies on the orientation of the foot model with respect to the ground and it represents a novelty with respect to other studies based on the 3D coordinates of ankle joint center only, thus neglecting the foot contact mechanism ([14], [19], [47], [48], [49], [50], [51]).

A. SPATIAL-TEMPORAL PARAMETERS
The proposed MS protocol showed a very good accuracy compliant with clinical requirements ([52], [53]) with MAE values equal to 1.2 cm for step length, 20 ms for stride duration, 2.5 cm for stride length and 0.02 m/s for the gait speed. The latter errors found on CP patients were comparable to those reported in previous single-camera studies but obtained on healthy subjects ([14], [17], [19], [22], [50], [51]).

To the best of the authors knowledge, this is the first study specifically validating the spatio-temporal parameters in children with CP against a MB protocol. In the literature, there are only a few single camera-based methods validated on post stroke ([49], [54]) and parkinsonian patients [47], and they showed lower performance. In particular,

Ferraris et al. [54] and Cimolin et al. [47] have assessed the errors associated to the spatio-temporal parameters estimation using the Kinect v2 body tracking SDK on eleven post-stroke and ten parkinsonian subjects, respectively. Both studies found ME values equal to 0.02 m/s for gait speed and 2 cm for step length, consistently larger than those found with the proposed MS protocol (0.01 m/s for gait speed and 0.06 cm for step length). In a recent study, Lonini et al. [49] evaluated the performance of DeepLabCut software for the analysis of the gait of ten post-stroke patients using a single RGB camera, reporting a high error variability for gait speed ($\pm$ 0.11 m/s in terms of ME).

B. LOWER-LIMB JOINT KINEMATICS

When comparing the 2D joint kinematics estimated by the proposed MS method against the 3D joint kinematics provided by the reference MB protocol, it is convenient to discriminate between the effects associated to the use of different anatomical axes definitions and angular conventions from the actual estimation errors [42]. While the adoption of different anatomical axes definition mainly reflects on an offset between curves, the errors in the reconstruction of the joint center trajectories would affect the waveform similarity and it can be quantified by the RMSE after offset removal.

The average angular offset values were 8° for the ankle, 6° for the knee and 7° for the hip joint. The ankle offset can be partially explained considering that the foot antero-posterior axis in MS protocol is computed as the principal axis of the best-fitting inertial ellipsoid whereas in the MB protocol from the position of the markers attached to the second metatarsal joint and the calcaneus. Similarly, the offset at the knee joint can be ascribed to the different definition for the HJC identifications implemented in the MS and MB protocol. In fact, while in the MS protocol the HJC coincides with the position of GT, in the MB protocol is determined as the geometrical center of the acetabulum and it is determined based on anthropometric regression equation [2]. The offset at the hip joint is associated to the fact that in the MS protocol, the pelvis inclination is assumed to be constant during the gait cycle and coinciding to the pelvic tilt during the static upright standing acquisition.

In terms of waveform similarity, the most accurate joint angle was obtained for the knee joint (RMSE = 3.2°), followed by the hip joint (RMSE = 3.5°) and the ankle joint (RMSE = 4.5°). It is important to highlight that, from a clinical perspective, errors between 2° and 5° are likely to be regarded as reasonable but may require consideration in data interpretation [55]. The largest errors affecting the ankle kinematics are mainly due to the auto-exposure of the camera which can cause blurred images of the foot and distal part of the shank during fastest movement such as the swing phase.

In the last years, several single-camera MS methods were proposed for gait analysis, however, in many cases, a direct comparative evaluation with the proposed method was not possible because: (*i*) the kinematic outputs were not validated against a clinically-accepted gold standard ([56], [57], [58], [59]) (*ii*) the method performance was only validated in terms of accuracy in tracking joint centers ([46], [60]), (*iii*) the objective was to classify motor activities or to detect gait abnormalities ([54], [61], [62], [63], [64], [65]).

To the best of authors knowledge, the only MS study involving CP children was presented by Nguyen et al. [46] that evaluated the concurrent validity of the built-in body tracking SDK (Kinect v2) against an MB gait protocol on 10 CP children (GMFS I-II) based on a frontal view. However, the reported errors were large for all the joints (RMSE = 11.2° for the hip, 10.3° for the knee, and 7.5° for the ankle).

In addition, there are a few MS studies that have been only applied and evaluated on normal gait ([17], [19], [66]).

In Yeung et al. [66], the effect of five camera viewing angles on the estimates of kinematics curves on healthy subjects by using body tracking SDK of Kinect v2 was investigated. They found that Kinect v2 performed better at frontal camera viewing angle showing a RMSE of 8° for the hip flexion/extension angle, 11.4° for the sagittal knee, and 17.4° for the ankle plantar/dorsi flexion angle.

Yamamoto et al. have tested the performance of Open-Pose [17] on healthy subjects reporting comparable results in terms of reliability for knee and hip kinematics (ICC ranging from 0.60 to 0.98) but very poor reliability for ankle kinematics represented by an ICC of 0.1 for the ankle maximum dorsiflexion during the stance phase and the swing phase in contrast to the ICC of 0.66 obtained from our MS protocol.

Castelli et al. [19] reported a RMSE of 4.8 ° for the hip, 3.6 ° for the knee, and 3° for the ankle on 10 healthy adults which are comparable with the values obtained using our method on CP children.

Interestingly, the proposed MS method provide a level of accuracy similar to that obtained by popular observation-based clinical gait assessment tools such as the Salford Gait Tool based on the manual identification of the anatomical landmarks on each recorded image [67].

In particular, Larsen et al. [67] have analyzed the accuracy of the Salford Gait Tool against MB protocol on 10 adult CP patients. They showed comparable errors in terms of ME on some key gait features showing that the proposed method obtained comparable performance of observation-based clinical gait assessment with less efforts required from the clinicians since our protocol includes manual intervention exclusively for calibrating the three templates.

V. CONCLUSION

The proposed MS protocol was designed to satisfy as priority its usability in clinical setting in terms of set-up and cost. For this reason, it was decided to use a single consumer-grade RGB-D camera. Clearly, this choice inevitably limited the kinematic analysis to the joint movement in the sagittal plane and hence to the description of the flexion-extension lower limb joint angles.

Furthermore, the movement of the subject was reconstructed based on a 2D multi-segmental model defined from a single 2D RGB image as the information provided by the depth sensor were only used to extract the foreground body segments. However, the projection of human 3D bodies motion to a 2D space necessarily leads to errors and ambiguities which could be only partially compensated using the multiple anatomical calibration proposed.

Another critical factor was represented by the quality of the image recorded by the specific RGB-D camera. In fact, due to automatic exposure time implemented by the Kinect v2, the recorded images resulted blurred when capturing fast moving body parts, and it negatively affected the results of the template/mask matching. The latter problem could be easily solved by selecting a camera which allows to control the exposure parameters.

Finally, it should be acknowledged that the proposed protocol is not fully automatic as it is required a preliminary identification of the external anatomical landmarks for the subject-specific model definition. Nonetheless, it is important to consider that gait analysis in CP patients is generally preceded by a clinical examination during which the clinicians assess range of joint motion and spasticity and may easily perform the identification of few anatomical landmarks.

In conclusion, the present study demonstrated the technical validity of a MS single-camera protocol for clinical gait analysis in CP population. Results showed a good accuracy in the joint kinematics estimation and good to excellent reliability for the extraction of a complete set of clinically relevant key gait features.

APPENDIX

In this appendix, the procedure followed to calculate the conversion factor (m/pixel) for the computation of spatial-temporal parameters was presented.

To convert the spatial parameters in meters, it is necessary to determine the pixel-to-meter conversion ratio. It should be noted that this conversion factor varies with the distance from the camera since the higher the distance, the higher the conversion factor is. To determine the conversion factor, a series of preliminary acquisitions were performed by positioning an object of known size at various distances from the camera. For each distance, the conversion factor was calculated as:

$$Conversion\,factor\,(m/pixel) = \frac{Size\,(m)}{Size\,(pixel)}$$

The conversion factor values were fitted though a linear fitting technique to best approximate them. Consequently, this procedure is able to obtain a linear model that associates each distance to its corresponding conversion factor.

The spatial-temporal parameters in meter were calculated applying the conversion factor to those values in pixel.

REFERENCES

[1] M. G. Benedetti et al., "SIAMOC position paper on gait analysis in clinical practice: General requirements, methods and appropriateness. Results of an Italian consensus conference," *Gait Posture*, vol. 58, pp. 252–260, Oct. 2017, doi: 10.1016/j.gaitpost.2017.08.003.

[2] R. B. Davis, S. Õunpuu, D. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Hum. Movement Sci.*, vol. 10, no. 5, pp. 575–587, 1991. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016794579190046Z and https://doi.org/10.1016/0167-9457(91)90046-Z

[3] A. Leardini, Z. Sawacha, G. Paolini, S. Ingrosso, R. Nativo, and M. G. Benedetti, "A new anatomically based protocol for gait analysis in children," *Gait Posture*, vol. 26, no. 4, pp. 560–571, Oct. 2007, doi: 10.1016/j.gaitpost.2006.12.018.

[4] M. P. Kadaba, H. K. Ramakrishnan, and M. E. Wootten, "Measurement of lower extremity kinematics during level walking," *J. Orthopaedic Res.*, vol. 8, no. 3, pp. 383–392, May 1990.

[5] E. D'Antonio, J. Taborri, I. Mileti, S. Rossi, and F. Patané, "Validation of a 3D markerless system for gait analysis based on OpenPose and two RGB webcams," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17064–17075, Aug. 2021, doi: 10.1109/JSEN.2021.3081188.

[6] *Move 4D*. Accessed: Aug. 7, 2023. [Online]. Available: https://www.move4d.net/

[7] *Theia Markerless*. Accessed: Aug. 7, 2023. [Online]. Available: https://www.theiamarkerless.ca/

[8] M. Moro, G. Marchesi, F. Hesse, F. Odone, and M. Casadio, "Markerless vs. marker-based gait analysis: A proof of concept study," *Sensors*, vol. 22, no. 5, p. 2011, Mar. 2022, doi: 10.3390/s22052011.

[9] J. Colombel, D. Daney, V. Bonnet, and F. Charpillet, "Markerless 3D human pose tracking in the wild with fusion of multiple depth cameras: Comparative experimental study with Kinect 2 and 3," in *Activity and Behavior Computing, Smart Innovation, Systems and Technologies*. Springer, 2020.

[10] E. Ceseracciu, Z. Sawacha, and C. Cobelli, "Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: Proof of concept," *PLoS ONE*, vol. 9, no. 3, Mar. 2014, Art. no. e87640, doi: 10.1371/journal.pone.0087640.

[11] M. Sandau, H. Koblauch, T. B. Moeslund, H. Aanæs, T. Alkjær, and E. B. Simonsen, "Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane," *Med. Eng. Phys.*, vol. 36, no. 9, pp. 1168–1175, Sep. 2014, doi: 10.1016/j.medengphy.2014.07.007.

[12] L. Mündermann, S. Corazza, and T. P. Andriacchi, "The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications," *J. Neuroeng. Rehabil.*, vol. 3, no. 1, Mar. 15, 2006, doi: 10.1186/1743-0003-3-6.

[13] L. Romeo, R. Marani, M. Malosio, A. G. Perri, and T. D'Orazio, "Performance analysis of body tracking with the Microsoft azure Kinect," in *Proc. 29th Medit. Conf. Control Autom. (MED)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Jun. 2021, pp. 572–577, doi: 10.1109/MED51440.2021.9480177.

[14] R. A. Clark, K. J. Bower, B. F. Mentiplay, K. Paterson, and Y.-H. Pua, "Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables," *J. Biomech.*, vol. 46, no. 15, pp. 2722–2725, Oct. 2013, doi: 10.1016/j.jbiomech.2013.08.011.

[15] Y. Ma, K. Mithraratne, N. Wilson, X. Wang, Y. Ma, and Y. Zhang, "The validity and reliability of a Kinect v2-based gait analysis system for children with cerebral palsy," *Sensors*, vol. 19, no. 7, p. 1660, Apr. 2019, doi: 10.3390/s19071660.

[16] X. Lv, S. Wang, T. Chen, J. Zhao, D. Chen, M. Xiao, X. Zhao, and H. Wei, "Human gait analysis method based on sample entropy fusion AlphaPose algorithm," in *Proc. 33rd Chin. Control Decis. Conf. (CCDC)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, May 2021, pp. 1543–1547, doi: 10.1109/CCDC52312.2021.9602427.

[17] M. Yamamoto, K. Shimatani, M. Hasegawa, Y. Kurita, Y. Ishige, and H. Takemura, "Accuracy of temporo-spatial and lower limb joint kinematics parameters using OpenPose for various gait patterns with orthosis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 2666–2675, 2021, doi: 10.1109/TNSRE.2021.3135879.

[18] J. Stenum, C. Rossi, and R. T. Roemmich, "Two-dimensional video-based analysis of human gait using pose estimation," *PLOS Comput. Biol.*, vol. 17, no. 4, Apr. 2021, Art. no. e1008935, doi: 10.1371/journal.pcbi.1008935.

[19] A. Castelli, G. Paolini, A. Cereatti, and U. Della Croce, "A 2D markerless gait analysis methodology: Validation on healthy subjects," *Comput. Math. Methods Med.*, vol. 2015, pp. 1–11, 2015, doi: 10.1155/2015/186780.

[20] E. Pantzar-Castilla, A. Cereatti, G. Figari, N. Valeri, G. Paolini, U. D. Croce, A. Magnuson, and J. Riad, "Knee joint sagittal plane movement in cerebral palsy: A comparative study of 2-dimensional markerless video and 3-dimensional gait analysis," *Acta Orthopaedica*, vol. 89, no. 6, pp. 656–661, Nov. 2018, doi: 10.1080/17453674.2018.1525195.

[21] E. Surer, A. Cereatti, E. Grosso, and U. D. Croce, "A markerless estimation of the ankle–foot complex 2D kinematics during stance," *Gait Posture*, vol. 33, no. 4, pp. 532–537, Apr. 2011, doi: 10.1016/j.gaitpost.2011.01.003.

[22] M. Hatamzadeh, L. Busé, F. Chorin, P. Alliez, J.-D. Favreau, and R. Zory, "A kinematic-geometric model based on ankles' depth trajectory in frontal plane for gait analysis using a single RGB-D camera," *J. Biomechanics*, vol. 145, Dec. 2022, Art. no. 111358, doi: 10.1016/j.jbiomech.2022.111358.

[23] A. Leu, D. Ristic-Durrant, and A. Gräser, "A robust markerless vision-based human gait analysis system," in *Proc. 6th IEEE Int. Symp. Appl. Comput. Intell. Informat. (SACI)*, May 2011, pp. 415–420, doi: 10.1109/SACI.2011.5873039.

[24] M. Goffredo, J. N. Carter, and M. S. Nixon, "2D markerless gait analysis," in *Proc. 4th Eur. Conf. Int. Fed. Medical Biol. Eng. (ECIFMBE)*, Antwerp, Belgium. Berlin, Germany: Springer, Nov. 2009.

[25] D. Balta, M. Salvi, F. Molinari, G. Figari, G. Paolini, U. D. Croce, and A. Cereatti, "A two-dimensional clinical gait analysis protocol based on markerless recordings from a single RGB-depth camera," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2020, pp. 1–6, doi: 10.1109/MeMeA49120.2020.9137183.

[26] J. Saboune and F. Charpillet, "Markerless human motion capture for gait analysis," *ArXiv*, vol. abs/cs/0510063, 2005. [Online]. Available: https://api.semanticscholar.org/CorpusID:2069744

[27] R. A. Clark, B. F. Mentiplay, E. Hough, and Y. H. Pua, "Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives," *Gait Posture*, vol. 68, pp. 193–200, Feb. 2019, doi: 10.1016/j.gaitpost.2018.11.029.

[28] L. Büker, V. Quinten, M. Hackbarth, S. Hellmers, R. Diekmann, and A. Hein, "How the processing mode influences azure Kinect body tracking results," *Sensors*, vol. 23, no. 2, p. 878, Jan. 2023, doi: 10.3390/s23020878.

[29] G. Wu, S. Siegler, P. Allard, C. Kirtley, A. Leardini, D. Rosenbaum, M. Whittle, D. D D'Lima, L. Cristofolini, H. Witte, O. Schmid, and I. Stokes, "ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—Part I: Ankle, hip, and spine," *J. Biomech.*, vol. 35, no. 4, pp. 543–548, 2002, doi: 10.1016/S0021-9290(01)00222-6.

[30] A. Cereatti, F. Margheritini, M. Donati, and A. Cappozzo, "Is the human acetabulofemoral joint spherical?" *J. Bone Joint Surg. Brit.*, vol. 92-B, no. 2, pp. 311–314, Feb. 2010, doi: 10.1302/0301-620x.92b2.22625.

[31] D. Herrera C., J. Kannala, and J. Heikkilä, "Joint depth and color camera calibration with distortion correction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 2058–2064, Oct. 2012, doi: 10.1109/TPAMI.2012.125.

[32] J.-Y. Bouguet, "Camera calibration toolbox for MATLAB," Comput. Vis. California Inst. Technol., Camera Calibration Toolbox MATLAB. Accessed: Aug. 7, 2023. [Online]. Available: https://data.caltech.edu/records/jx9cx-fdh55

[33] J. Rodda and H. K. Graham, "Classification of gait patterns in spastic hemiplegia and spastic diplegia: A basis for a management algorithm," *Eur. J. Neurol.*, vol. 8, no. s5, pp. 98–108, Nov. 2001, doi: 10.1046/j.1468-1331.2001.00042.x.

[34] H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognit.*, vol. 34, no. 12, pp. 2259–2281, Dec. 2001, doi: 10.1016/S0031-3203(00)00149-7.

[35] M. Salvi and F. Molinari, "Multi-tissue and multi-scale approach for nuclei segmentation in H&E stained images," *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 89, Dec. 2018, doi: 10.1186/s12938-018-0518-0.

[36] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992, doi: 10.1109/34.121791.

[37] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[38] A. Cereatti, T. Bonci, M. Akbarshahi, K. Aminian, A. Barré, M. Begon, D. L. Benoit, C. Charbonnier, F. Dal Maso, S. Fantozzi, C.-C. Lin, T.-W. Lu, M. G. Pandy, R. Stagni, A. J. van den Bogert, and V. Camomilla, "Standardization proposal of soft tissue artefact description for data sharing in human motion measurements," *J. Biomech.*, vol. 62, pp. 5–13, Sep. 2017, doi: 10.1016/j.jbiomech.2017.02.004.

[39] A. Cappello, A. Cappozzo, P. F. La Palombara, L. Lucchetti, and A. Leardini, "Multiple anatomical landmark calibration for optimal bone pose estimation," *Hum. Movement Sci.*, vol. 16, no. 2, pp. 259–274, 1997, doi: 10.1016/S0167-9457(96)00055-3.

[40] A. Cereatti, C. Rosso, A. Nazarian, J. P. DeAngelis, A. J. Ramappa, and U. D. Croce, "Scapular motion tracking using acromion skin marker cluster: In vitro accuracy assessment," *J. Med. Biol. Eng.*, vol. 35, no. 1, pp. 94–103, Feb. 2015, doi: 10.1007/s40846-015-0010-2.

[41] E. Bergamini, G. Ligorio, A. Summa, G. Vannozzi, A. Cappozzo, and A. Sabatini, "Estimating orientation using magnetic and inertial sensors and different sensor fusion approaches: Accuracy assessment in manual and locomotion tasks," *Sensors*, vol. 14, no. 10, pp. 18625–18649, Oct. 2014, doi: 10.3390/s141018625.

[42] P. Picerno, A. Cereatti, and A. Cappozzo, "Joint kinematics estimate using wearable inertial and magnetic sensing modules," *Gait Posture*, vol. 28, no. 4, pp. 588–595, Nov. 2008, doi: 10.1016/j.gaitpost.2008.04.003.

[43] M. G. Benedetti, F. Catani, A. Leardini, E. Pignotti, and S. Giannini, "Data management in gait analysis for clinical applications," *Clin. Biomech.*, vol. 13, no. 3, pp. 204–215, 1998, doi: 10.1016/S0268-0033(97)00041-7.

[44] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *J. Chiropractic Med.*, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012.

[45] C. P. Dancey and J. Reidy, *Statistics Without Maths for Psychology*. London, U.K.: Pearson, 2007.

[46] M.-H. Nguyen, C.-C. Hsiao, W.-H. Cheng, and C.-C. Huang, "Practical 3D human skeleton tracking based on multi-view and multi-Kinect fusion," *Multimedia Syst.*, vol. 28, no. 2, pp. 529–552, Apr. 2022, doi: 10.1007/s00530-021-00846-x.

[47] V. Cimolin, L. Vismara, C. Ferraris, G. Amprimo, G. Pettiti, R. Lopez, M. Galli, R. Cremascoli, S. Sinagra, A. Mauro, and L. Priano, "Computation of gait parameters in post stroke and Parkinson's disease: A comparative study using RGB-D sensors and optoelectronic systems," *Sensors*, vol. 22, no. 3, p. 824, Jan. 2022, doi: 10.3390/s22030824.

[48] C. Ferraris, V. Cimolin, L. Vismara, V. Votta, G. Amprimo, R. Cremascoli, M. Galli, R. Nerino, A. Mauro, and L. Priano, "Monitoring of gait parameters in post-stroke individuals: A feasibility study using RGB-D sensors," *Sensors*, vol. 21, no. 17, p. 5945, Sep. 2021, doi: 10.3390/s21175945.

[49] L. Lonini, Y. Moon, K. Embry, R. J. Cotton, K. McKenzie, S. Jenz, and A. Jayaraman, "Video-based pose estimation for gait analysis in stroke survivors during clinical assessments: A proof-of-concept study," *Digit. Biomarkers*, vol. 6, no. 1, pp. 9–18, Jan. 2022, doi: 10.1159/000520732.

[50] J. Bertram, T. Krüger, H. M. Röhling, A. Jelusic, S. Mansow-Model, R. Schniepp, M. Wuehr, and K. Otte, "Accuracy and repeatability of the Microsoft azure Kinect for clinical measurement of motor function," *PLoS ONE*, vol. 18, no. 1, Jan. 2023, Art. no. e0279697, doi: 10.1371/journal.pone.0279697.

[51] J. A. Albert, V. Owolabi, A. Gebel, C. M. Brahms, U. Granacher, and B. Arnrich, "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," *Sensors*, vol. 20, no. 18, pp. 1–22, Sep. 2020, doi: 10.3390/s20185104.

[52] S. Perera, S. H. Mody, R. C. Woodman, and S. A. Studenski, "Meaningful change and responsiveness in common physical performance measures in older adults," *J. Amer. Geriatrics Soc.*, vol. 54, no. 5, pp. 743–749, May 2006, doi: 10.1111/j.1532-5415.2006.00701.x.

[53] J. S. Brach, S. Perera, S. Studenski, M. Katz, C. Hall, and J. Verghese, "Meaningful change in measures of gait variability in older adults," *Gait Posture*, vol. 31, no. 2, pp. 175–179, Feb. 2010, doi: 10.1016/j.gaitpost.2009.10.002.

[54] C. Ferraris, G. Amprimo, G. Pettiti, G. Masi, and L. Priano, "Automatic detector of gait alterations using RGB-D sensor and supervised classifiers: A preliminary study," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, Jun. 2022, pp. 1–6, doi: 10.1109/ISCC55528.2022.9912923.

[55] J. L. McGinley, R. Baker, R. Wolfe, and M. E. Morris, "The reliability of three-dimensional kinematic gait measurements: A systematic review," *Gait Posture*, vol. 29, no. 3, pp. 360–369, Apr. 2009, doi: 10.1016/j.gaitpost.2008.09.003.

[56] J. Latorre, R. Llorens, C. Colomer, and M. Alcañiz, "Reliability and comparison of Kinect-based methods for estimating spatiotemporal gait parameters of healthy and post-stroke individuals," *J. Biomech.*, vol. 72, pp. 268–273, Apr. 2018, doi: 10.1016/j.jbiomech.2018.03.008.

[57] C. Ferraris, G. Amprimo, G. Masi, L. Vismara, R. Cremascoli, S. Sinagra, G. Pettiti, A. Mauro, and L. Priano, "Evaluation of arm swing features and asymmetry during gait in Parkinson's disease using the azure Kinect sensor," *Sensors*, vol. 22, no. 16, p. 6282, Aug. 2022, doi: 10.3390/s22166282.

[58] J. Latorre, C. Colomer, M. Alcañiz, and R. Llorens, "Gait analysis with the Kinect v2: Normative study with healthy individuals and comprehensive study of its sensitivity, validity, and reliability in individuals with stroke," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, pp. 1–11, Jul. 2019, doi: 10.1186/s12984-019-0568-y.

[59] G. Amprimo, G. Pettiti, L. Priano, A. Mauro, and C. Ferraris, "Kinect-based solution for the home monitoring of gait and balance in elderly people with and without neurological diseases," in *Proc. Italian Workshop Artif. Intell. Ageing Soc. (AIxAS)*, Nov. 2021, pp. 70–80.

[60] N. Hesse, S. Baumgartner, A. Gut, and H. J. A. van Hedel, "Concurrent validity of a custom method for markerless 3D full-body motion tracking of children and young adults based on a single RGB-D camera," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1943–1951, 2023, doi: 10.1109/TNSRE.2023.3251440.

[61] S.-W. Chen, S.-H. Lin, L.-D. Liao, H.-Y. Lai, Y.-C. Pei, T.-S. Kuo, C.-T. Lin, J.-Y. Chang, Y.-Y. Chen, Y.-C. Lo, S.-Y. Chen, R. Wu, and S. Tsang, "Quantification and recognition of parkinsonian gait from monocular video imaging using kernel-based principal component analysis," *Biomed. Eng. OnLine*, vol. 10, no. 1, p. 99, Nov. 2011, doi: 10.1186/1475-925X-10-99.

[62] T. Li, J. Chen, C. Hu, Y. Ma, Z. Wu, W. Wan, Y. Huang, F. Jia, C. Gong, S. Wan, and L. Li, "Automatic timed up-and-go sub-task segmentation for Parkinson's disease patients using video-based activity classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2189–2199, Nov. 2018, doi: 10.1109/TNSRE.2018.2875738.

[63] R. A. Clark, S. Vernon, B. F. Mentiplay, K. J. Miller, J. L. McGinley, Y. Pua, K. Paterson, and K. J. Bower, "Instrumenting gait assessment using the Kinect in people living with stroke: Reliability and association with balance tests," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 15, Feb. 2015, doi: 10.1186/s12984-015-0006-8.

[64] N. Kojovic, S. Natraj, S. P. Mohanty, T. Maillart, and M. Schaer, "Using 2D video-based pose estimation for automated prediction of autism spectrum disorders in young children," *Sci. Rep.*, vol. 11, no. 1, Jul. 2021, Art. no. 15069, doi: 10.1038/s41598-021-94378-z.

[65] M. Stricker, D. Hinde, A. Rolland, N. Salzman, A. Watson, and T. G. Almonroeder, "Quantifying step length using two-dimensional video in individuals with Parkinson's disease," *Physiotherapy Theory Pract.*, vol. 37, no. 1, pp. 252–255, Jan. 2021, doi: 10.1080/09593985.2019.1594472.

[66] L.-F. Yeung, Z. Yang, K. C.-C. Cheng, D. Du, and R. K.-Y. Tong, "Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2," *Gait Posture*, vol. 87, pp. 19–26, Jun. 2021, doi: 10.1016/j.gaitpost.2021.04.005.

[67] K. L. Larsen, G. Maanum, K. F. Frøslie, and R. Jahnsen, "Ambulant adults with spastic cerebral palsy: The validity of lower limb joint angle measurements from sagittal video recordings," *Gait Posture*, vol. 35, no. 2, pp. 186–191, Feb. 2012, doi: 10.1016/j.gaitpost.2011.09.004.
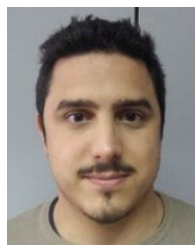
GIULIO FIGARI was born in Genoa, in April 1990. He received the degree in biomedical engineering in Genoa, in 2013, while attending the ISICT Course of Excellence, and the master's degree in bioengineering, in 2015. He started working under a research grant for the University of Sassari on a project in collaboration with GPEM, in 2015. In 2016, he was hired by GPEM, where he still holds the position of a research and development and support manager.

GABRIELE PAOLINI received the master's degree in electronic engineering from the University of Bologna, in 2003. During the same year, he was a Consultant with the Motion Analysis Laboratory, Istituti Ortopedici Rizzoli, Bologna. In 2004, he assumed the role of a Researcher with Centro Protesi I.N.A.I.L. From 2004 to 2006, he held the positions of a lab manager and a researcher consultant with the University of Virginia. From 2006 to 2007, he was a Senior Customer Support Engineer with Vicon Motion System. From 2012 to 2016, he was a Research Fellow with the University of Sassari. In 2012, he founded GPEM srl.

EVELINA PANTZAR-CASTILLA received the Ph.D. degree in the field of gait analysis in children with cerebral palsy, in 2021. She has been an Orthopaedic Surgeon with the Department of Orthopaedics, Örebro University Hospital, since 2012. Since 2019, her specialization has been paediatric orthopaedic.

DILETTA BALTA was born in Grottaglie, Taranto, in October 1995. She received the B.S. and M.S. degrees in biomedical engineering from Politecnico di Torino, Italy, in 2017 and 2019, respectively, where she is currently pursuing the Ph.D. degree in bioengineering and medical-surgical sciences. In 2020, she was a Research Assistant with the University of Sassari, where she contributed to the development of modeling techniques of the lower limb joint based on RGB-depth data. Her research interest includes the use of RGB-D sensors in human movement analysis for the estimation of the upper and lower limb joint angles.
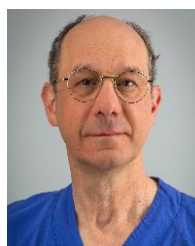
JACQUES RIAD received the M.S. degree from the Medical School, University of Uppsala, Linköping, Sweden and the Ph.D. degree, in 2011.

He worked with Uppsala University Hospital, Stockholm University Hospital, and Umeå University Hospital. He was a Clinical Fellow with the Women´s and Children´s Hospital, Adelaide, Australia, and a Research Fellow with the duPont Hospital for Children, Wilmington Delaware, USA. Since 2005, he has been a Senior Consultant with the Skaraborg Hospital Skövde. Since 2012, he has been an Associate Professor with the University of Gothenburg, Gothenburg, Sweden.

UGO DELLA CROCE received the Ph.D. degree in bioengineering from the University of Bologna, in 1996. Then, he moved to work with the Spaulding Rehabilitation Hospital, Boston, as a Postdoctoral Researcher, contributing to the development of the Motion Analysis Laboratory. While pursuing his academic career with the University of Sassari, Italy, where he became an Assistant Professor, in 1997. He spent a year with the School of Medicine, University of Virginia, to establish a brand new gait analysis laboratory with the Physical Medicine and Rehabilitation Department. Since 2006, he has been a Full Professor with the University of Sassari, while maintaining his academic appointments with the Harvard Medical School and the School of Medicine, University of Virginia. In 2017, he was selected as a Science Attaché with the Embassy of Italy, Washington, DC, USA, with the responsibility of enhancing bilateral research in the biomedical field. He has authored more than 100 publications. His research interests include the development of methods to improve quantitative human movement measurements based on various sensing technologies, including video, dynamometry and inertial sensing, and clinical applications of human movement analysis, including studies regarding the elderly, patients with Parkinson disease, multiple sclerosis, burns, chorea, stroke, and the development and validation of technology (prosthetics, orthoses, and exoskeletons) and protocols (treadmill, robotics, and/or virtual reality-based), for motor rehabilitation programs. His research interests also include biomechanics, including posturography, electromyography, and image analysis applications for motion measurements. He was the President of the Italian Society of Clinical Movement Analysis (SIAMOC), from 2013 to 2015. He was also invited by European and North American institutions to hold about 20 lectures and seminars on human movement analysis.

ANDREA CEREATTI received the M.S. degree (cum laude) in mechanical engineering, in 2002, and the Ph.D. degree in bioengineering, in 2006. From 2016 to 2018, he was an Adjunct Professor of bioengineering with Politecnico di Torino, Italy. From 2019 to 2020, he was an Associate professor with the University of Sassari, Italy. He is currently a Full Professor with Politecnico di Torino. He has coauthored more than 100 publications and is the inventor of two patents. His research interests include high-resolution joint kinematics estimation, wearable sensors for locomotor capacity and performance assessment, and innovative approaches for neuro-muscular rehabilitation. He is a member of GNB and the Italian National Group of Bioengineering. He has served as the Board of Directors of the Italian Society of Clinical Movement Analysis, from 2009 to 2013, and the 3-D Analysis of Human Movement Technical Group of the ISB, from 2014 to 2018.

· · ·