ORIGINAL ARTICLE

Real-time limb tracking in single depth images based on circle matching and line fitting

Michael Tschiedel[1,2] · Michael Friedrich Russold[1] · Eugenijus Kaniusas[2] · Markus Vincze[3]

Abstract

Modern lower limb prostheses neither measure nor incorporate healthy residual leg information for intent recognition or device control. In order to increase robustness and reduce misclassification of devices like these, we propose a vision-based solution for real-time 3D human contralateral limb tracking (CoLiTrack). An inertial measurement unit and a depth camera are placed on the side of the prosthesis. The system is capable of estimating the shank axis of the healthy leg. Initially, the 3D input is transformed into a stabilized coordinate system. By splitting the subsequent shank estimation problem into two less computationally intensive steps, the computation time is significantly reduced: First, an iterative closest point algorithm is applied to fit circular models against 2D projections. Second, the random sample consensus method is used to determine the final shank axis. In our study, three experiments were conducted to validate the static, the dynamic and the real-world performance of our CoLiTrack approach. The shank angle can be tracked at 20 Hz for one sixth of the entire human gait cycle with an angle estimation error below $2.8 \pm 2.1°$. Our promising results demonstrate the robustness of the novel CoLiTrack approach to make "next-generation prostheses" more user-friendly, functional and safe.

Keywords Shank modeling · Body tracking · Depth image · Gait analysis

1 Introduction

Extraction of 3D human limb parameters from depth images is a research topic that has recently attracted the attention of the scientific community. Driven by the availability of low-cost depth cameras, human pose estimation strategies [1–3], telerehabilitation concepts [4,5] and patient interaction monitoring approaches [6,7] are evolving continuously. For lower limb prostheses control, it can be advantageous to gather information about the healthy residual leg. Humans combine proprioception with visual information to navigate different terrains smoothly. In contrast, state-of-the-art commercial prosthetic devices only use device-embedded sensors and finite-state controllers to adapt to the patient's intent [8]. They do not collect information about the state of the other

leg, which could improve overall system performance. A systematic analysis of different bilateral lower limb signals for predicting locomotion activities was performed by Hu et al. [9,10] in 2018. It came to the conclusion that only one additional contralateral leg parameter could reduce error rates of intent recognition significantly. However, the additional effort of instrumenting the contralateral shank can be inconvenient and impractical for amputees.

In this paper, we propose a novel contralateral limb tracking approach named CoLiTrack, which utilizes unilaterally worn depth cameras. Placing the camera on the ipsilateral (prosthetic) side eliminates the need for an additional sensor on the contralateral shank and allows an easier integration into future products. The key to our method is that it separates the complex modeling problem into two less computationally intensive parts, in order to perform real-time shank axis estimation. Initially, layers from the point cloud input (depicting the residual leg) are projected in 2D, before fitting predefined circular models. Next, a projected line using the center points of the circles is used to estimate the shank axis. This reduces computing time considerably, allowing for accurate estimation of the shank axis fast enough for real-time applications. Furthermore, in comparison with traditional methods,

✉ Michael Tschiedel
   michael@tschiedel.at

1 Department of Global Research, Ottobock Healthcare Products GmbH, Vienna, Austria

2 Institute of Electrodynamics, Microwave and Circuit Engineering, TU Wien, Vienna, Austria

3 Automation and Control Institute, TU Wien, Vienna, Austria

CoLiTrack is unsupervised and has no training-related bias. Experimental results produced by CoLiTrack demonstrate a robust and accurate solution to improve human–device interaction in real time. Although our method was developed with lower limb prostheses in mind, the proposed system could be applied to a wide variety of applications in many areas, including health care, gaming and human–device interaction.

The remainder of the paper is organized as follows: Sect. 2 presents the related work most relevant to the context of this paper. Details of the proposed method for real-time contralateral limb tracking are given in Sect. 3. Section 4 provides the experimental results, which are then discussed in Sect. 5. The conclusion of this paper is presented in Sect. 6.

2 Related work

In this section, we focus primarily on existing work on environment recognition as a means for improving prosthetic control. We present relevant principles of human gait and known approaches to modeling the leg.

2.1 Principles of human gait

Humans use upright gait on two legs for efficient locomotion. Biomechanics of gait have been studied in detail [11]; however, spinal and brain control of human walking cannot yet be fully explained. Walking is described as a repetitive sequence to move forward; a single sequence is called a gait cycle, which begins with the heel strike (the heel touching the ground) and continues until the heel strike of the same foot. For approximately 60% of the gait cycle the foot stays on the ground (stance phase), before the foot is lifted off the ground at toe off and swings freely in the air (swing phase) for the rest of the gait cycle. Numerous papers have been published on markerless motion capture [12–15] in the past, and it is beyond the scope of this paper to provide a complete overview of this literature. For our approach, it is important to mention that the shank kinematics are generally independent of the ankle kinematics. The ankle adjusts automatically to different heel heights, for example, as a result of wearing different types of shoes. The shank kinematics, however, remain the same [16,17]. The shank angle α is therefore a very informative parameter representing the status of the contralateral shank and can play a crucial role in movement-dependent control applications. A typical sagittal plane shank motion during free walking for one entire gait cycle is shown in Fig. 1. The sagittal plane passes across the body, defined by the Y/Z plane in the world coordinate system, as indicated in Fig. 2. The shank angle is defined to be zero, when it is perpendicular to the ground, and positive in the case of a counterclockwise direction around the reference axis.
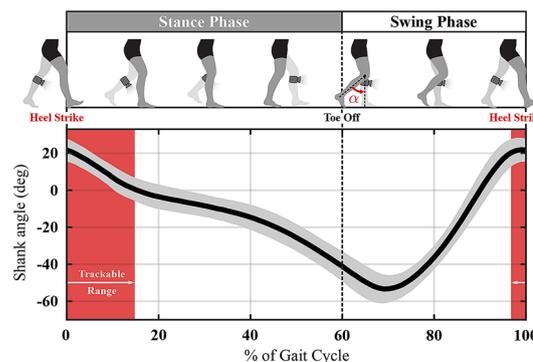

Fig. 1 Sagittal plane shank motion. Heel strike is the first inertial contact, when the foot touches the ground. At about 60% of the full gait cycle, toe off is the end of the stance phase initiating the swing phase. Then, the leg swings freely in the air until the next heel strike, before repeating the cycle. The plot shows nominal shank angle α relative to vertical during free walking with standard deviation in a light gray band, data from [11]. The trackable range of our CoLiTrack approach determined during the dynamic experiment is shown in red
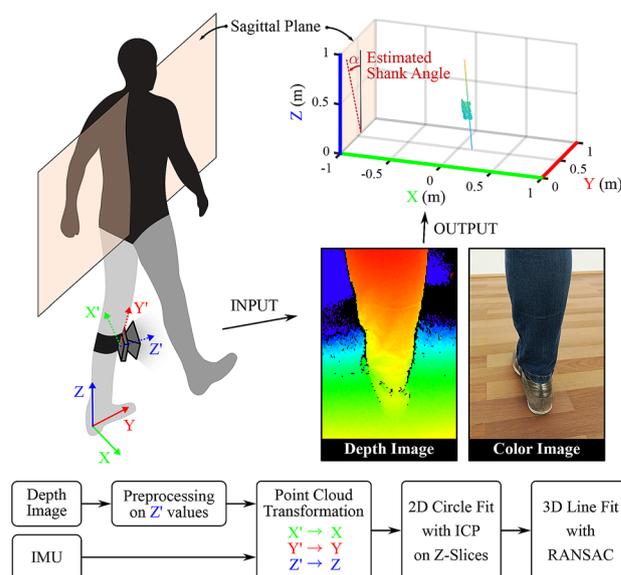

Fig. 2 Sensor configuration and overall process of CoLiTrack. An IMU and a depth camera were fused to estimate the axis of the uninstrumented contralateral shank. (X',Y', Z') represents the camera coordinate system and (X,Y, Z) the stabilized world coordinate system. Mounting is depicted in detail in Fig. 3. The color of the *Depth Image* corresponds to the distance between an object and the camera: red parts are close to the sensor, blue parts are further away. The *Color Image* was taken with a mobile phone and is for demonstration purpose

2.2 Environment detection for prostheses

One of the most important aims of lower limb prosthetic systems is the imitation of the physiological gait pattern [8]. Modern prosthetic systems can replace missing body parts to a high degree and improve patients' independence and mobility [18]. Numerous studies have been carried out to determine the best strategies to control locomotion in pros-

thetic devices. Most of these studies used neuromuscular or mechanical signals of the prosthetic leg, either individually or in combination [19–21]. Inertial measurement units (IMU) are also commonly used to estimate the position and orientation of the prosthetic leg in a world reference system [8]. In contrast, environmental sensor technologies are not yet commercially used in prosthetic devices, although we have identified a clear research tendency toward the recognition of surrounding objects and terrains, as shown in our survey [22]. The work found was divided into two categories: explicit environmental sensing—direct estimation of terrain features—and implicit environmental sensing—creating an understanding of the locomotion mode by measuring the state of the patient's residual body. Within the first group, distance and depth-based sensors were used to estimate the mode of locomotion. Sensors were placed on different body segments, such as the shank, thigh, trunk or even the upper body, scanning the environment in front of the patient. Intent recognition was performed by means of geometry-based decision trees [23], finite-state or support vector machines [24,25] and neural networks [26,27]. One group investigated whether the patient was approaching stairs [28], and another estimated soil properties with a color camera was mounted on the foot [29]. Toe clearance, which is an important parameter to prevent stumbling or falling, was estimated as well [30], but not evaluated in a prosthetic setup.

The second category, implicit environmental sensing, incorporates the state of the amputee's body, which we consider to be more promising. This is based on the fact that the amputee voluntarily decides where to go or what to do. The strong physical inter-joint coordination between human limbs [31] could be used to improve prosthesis control. Concepts of primitive "echo-control" strategies, which try to estimate the state of the missing limb depending on the state of the residual sound side, have been investigated for more than 40 years [32–35]. For example, stepping over unknown obstacles becomes possible without explicit classification of the environment [36]. However, errors occur especially at the beginning and at the end of an activity, when the limbs do not necessarily "echo" each other. As an improvement, "whole-body" approaches [37–39] with distributed IMUs and pressure insoles can distinguish between a limited number of modes of locomotion in real time, but need numerous additional sensors worn by the user.

For our work, the concept proposed by Hu et al. [40] is the most relevant one: It predicts bilateral gait events from unilaterally worn sensors. A thigh-mounted depth sensor and an IMU are fused to extract the angle between the ground and the shank of the contralateral leg in its field of view. Then, classifiers are used to predict ipsilateral toe off and contralateral heel contact, representing the beginning and end of the human gait double support phase. Although their methodology was sound, and the results suggested that depth

vision could improve device control, several limitations were mentioned. To begin with, the evaluation was carried out with only a single participant. Moreover, initiation and termination steps were excluded due to different kinematics compared to steady-state steps. No tests were performed regarding the influence of different environments, neither to investigate the robustness against reflectance and clutter, nor to analyze the effect of unknown objects in the field of view. Additionally, the implementation was not optimized for timing, resulting in a high computation time of 1.16 ± 0.56s, which prevented any online (real-time) evaluation.

2.3 Approximation of shank axis

The shank of a human can be modeled relatively well by the primitive shape of a cylinder. The challenge is, however, to fit such a model to the incomplete and deformed point cloud captured with a depth camera. On the one hand, simple least square methods [41] fail due to outliers, and methods using surface normals [42] cannot be applied due to noise from pleats on the clothes. On the other hand, complex object fitting approaches [43] are unsuitable, since the high processing time prevents any real-time evaluation. To address these challenges, we separate the complex shank estimation problem into two less computationally intensive steps: First, predefined circular models are matched in 2D. Second, a projected line using the center points of the circles is used to estimate the shank axis, as shown in Fig. 4. In general, detecting circles and lines is a fundamental task in computer vision and has been widely studied and developed in a variety of ways. Well-known techniques, such as the Hough transform or neural network approaches, are used to detect circle-like foreign objects in chest X-ray images [44–46] or line-like lanes for autonomous driving systems [47,48], to name recent applications.

3 Proposed method

Our contralateral limb tracking method named CoLiTrack consists of four main functions for estimating the axis and, therefore, the angle of the shank. The overall process is illustrated in Fig. 2: In a first step, the depth values were preprocessed in the camera coordinate system. Next, the information from the IMU was used to transform the filtered point cloud data into a ground coordinate system. The transformed points were projected onto 2D planes for fitting circle models with the iterative closest point (ICP) algorithm. Based on these results, the axis was estimated with a 3D line fit using the random sample consensus (RANSAC) method.
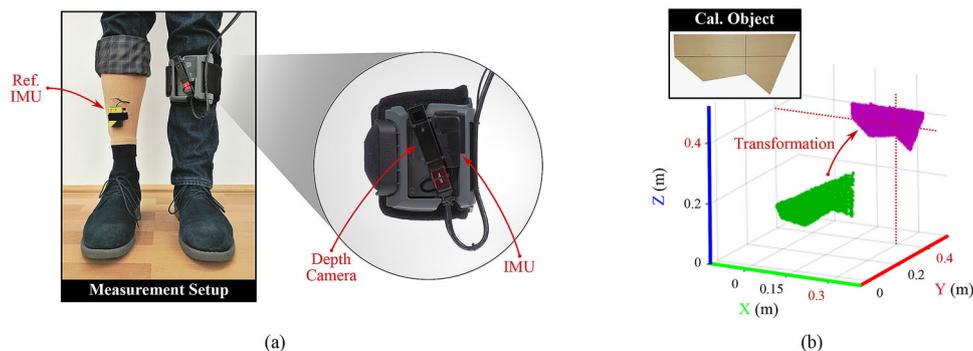
(a)

(b)

Fig. 3 Measurement setup **a** worn by a participant and the calibration step **b** of the depth camera. IMU and depth camera were fixed on a wearable support, attached to the shank of the participant. The second IMU, mounted within a modified support stocking on the contralateral side, served as reference; trousers were rolled up only for the photo.

For calibration, the camera was placed in the origin of the world coordinate system. The true (pink) position of the calibration object was known. The transformation matrix could be calculated out of the captured (green) image

3.1 Configuration of CoLiTrack

In this study, a 3D time-of-flight depth camera (CamBoard pico flexx, Pmd Tech, Germany) and an IMU (BNO055, Bosch Sensortec, Germany) were used to estimate parameters of the contralateral leg. In addition, a second IMU was mounted on the contralateral shank to serve as reference signal. The sensor configuration is depicted in Fig. 3a. The IMU and the depth camera were fixed together on a wearable support for two reasons: firstly, to combine camera orientation information with vision information. Secondly, this wearable support allowed for fast and easy positioning on the participant's shank. With a resolution of $171 x 224$ pixels, the depth camera could measure the 3D position of object points (point cloud) relative to the camera origin. The IMU was able to estimate the orientation and acceleration of the wearable support relative to the world reference system. By fusing orientation information and point cloud data, scene information could be stabilized. However, after the camera was mounted, its initial rotation needed to be corrected. For this, the wearable support carrying the IMU and the depth camera was placed in an upright position with the known calibration object in its field of view. The object was mounted in a predefined position, as shown in Fig. 3b in pink. The required rigid transformation can be calculated from a scene capture, which is depicted in green. The calculation was done by using the coherent point drift algorithm [49] to assign correspondences between two sets of points, provided within the MATLAB Computer Vision Toolbox. The determined transformation parameters were stored and applied every time before continuing with the evaluation algorithm. The experimental data acquisition and analysis were conducted using MATLAB R2018b and a wrapper library provided by the camera manufacturer, running on a laptop with an Intel Core i5-8250U and 8 GB memory size.

3.2 Preprocessing of depth data

The captured depth information was preprocessed in the camera coordinate system. Pixels without any depth information were removed and then blurred with a Gaussian blur to reduce depth noise. If the contralateral leg was in the camera's field of view, the nearest detected point relative to the camera belonged to the sound leg and was, therefore, selected as "point of interest" (POI). Next, the confidence map, which was also retrieved from the depth sensor, was used to identify all points belonging to the contralateral leg. This parameter represented the confidence of a measured distance (Z') for every pixel (X', Y') of the input—the closer a point, the higher the corresponding confidence. The confidence of the POI was taken as reference, points above a certain threshold were selected as belonging to the contralateral leg. Finally, the selected point cloud was downsampled using a 1 cm grid filter for computational efficiency.

3.3 Point cloud transformation

As the shank of the human was swinging within a gait cycle, the roll angle (rotation around the X-axis) and the pitch angle (rotation around the Y-axis) were estimated with the help of the ipsilateral IMU. By applying the Euler angle rotation matrix, the retrieved point cloud in camera coordinates was transformed into a stabilized world coordinate system (ipsilateral IMU reference frame). The origin was fixed on the foot calibrated to the sagittal plane, as shown in Fig. 2. Due to the instability of the yaw angle, this parameter was not used, and a rotation around the Z-axis remained unaccounted for.
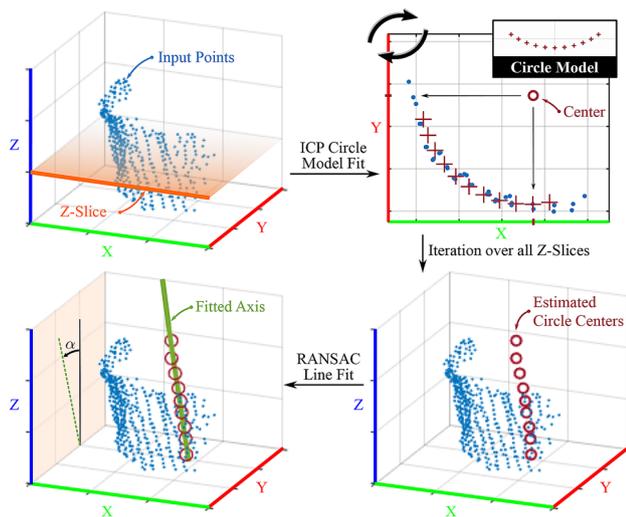
Fig. 4 Visualization of the CoLiTrack algorithm. The point cloud after preprocessing and transforming was taken as input. Z-slices were extracted, and a 10-point circle model was fitted iteratively by using the ICP algorithm in 2D (X/Y plane). All model circle center points were then used to estimate the shank axis correctly by using the RANSAC method in 3D. Finally, the contralateral shank angle α was calculated with respect to the sagittal plane (Y/Z plane)

3.4 Circle fit with ICP

After transforming the point cloud input, Z-layers were projected in 2D. Then, a predefined circle model of 10 points evenly distributed over 90° in a radius of 5 cm was fitted, as shown in Fig. 4. The Z-layer height was set to 3 cm with 1 cm overlapping on both sides. The model fit itself was realized with the ICP algorithm [50,51]. The basic concept of this algorithm was as follows: The Z-layer projection was kept fixed, while the circle model was rigidly transformed to match the input in the best way possible. By doing this, shape and size were preserved. The algorithm iteratively revised the transformation to minimize the sum of squared differences between the coordinates of the matched pairs. If the error metric fell below a certain threshold, this was a criterion for stopping the iterations. This was done for all Z-layers, which resulted in center points with their associated heights. Figure 4 depicts this ICP and the subsequent RANSAC fitting process.

3.5 Line fit with RANSAC

By using the RANSAC method [52], the newly calculated center points served to estimate the shank axis correctly. The principle of this iterative approach was to estimate parameters of a predefined mathematical model—in our case, a line representing the shank axis—from the input data. Therefore, two circle-center points were randomly taken to generate a hypothesis, which was then verified against all the other

points. If the distance between a point and the current hypothesis lay below a certain threshold, the point was marked as "inlier." Otherwise it was marked as "outlier." The algorithm was repeated, until the obtained hypothesis exceeded a certain ratio. Finally, only the points marked as inliers were used to calculate the optimal line, whereas outliers had no further influence on the result. For our specific application, we used a threshold of 3 cm, a probability of 0.99 and 1000 as maximum number of random trials. As shown in Fig. 4, the number of input points was limited by the previous Z-slice height and the total object height, which varied between 5 cm and 30 cm. Finally, the estimated axis was used to calculate the contralateral shank-to-vertical angle α with respect to the sagittal plane. The relative position information between the depth camera and the contralateral leg was not used in our application.

4 Experiments and results

We conducted three experiments to examine the *static performance*, the *dynamic performance* and the *real-world performance* of our novel approach. Generally, we analyzed the performance of our CoLiTrack method as follows: For each step, recorded gait data were separated, based on the local positive peak of the shank angle captured with the (second) reference IMU (α_{REF}). As depicted in Fig. 1, this corresponds with the beginning or, respectively, the ending of one entire gait cycle. Individual steps were then interpolated from 0 to 100%. Furthermore, gait initiation and termination steps were excluded, due to their kinematics differing from steady-state walking. The deviation between the depth camera-based estimation (α_{CLT}) and the reference IMU was calculated as shank angle error $|\alpha_{REF} - \alpha_{CLT}|$ for each percent of the gait cycle. Where possible, mean and standard deviation over the entire gait cycle were calculated and reported.

4.1 Static performance

The goal of the static performance experiment was to evaluate the performance of the proposed algorithm over the entire gait cycle. As the depth camera had a very narrow field of view (62° horizontal x 45° vertical, taken from the product's datasheet), parameters of the contralateral leg were "trackable" only for a small part of the total gait cycle, as it was out of view for the rest of the time. In order to evaluate our approach over the entire gait cycle, the wearable support, containing the depth camera and the IMU, was placed in front of a commercial treadmill, so that the participant's contralateral shank was constantly in view. The incline of the treadmill was set to 1% for all experiments, which is considered to be the same resistance level as an outdoor surface without incline [53]. Walking speed was defined as slow (0.5 km/h),
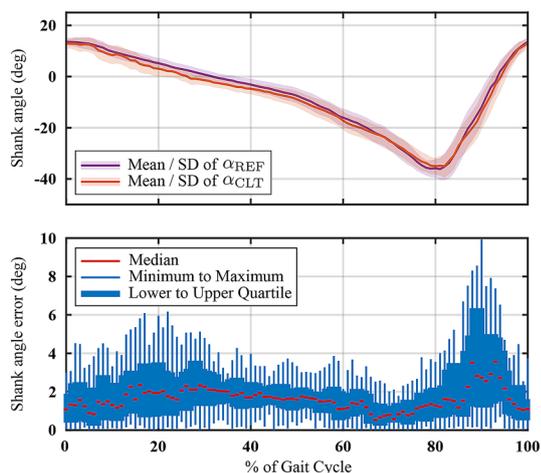
Fig. 5 Static performance visualization of CoLiTrack at medium walking speed. Top diagram depicts the mean and the standard deviation of the α_{CLT} and α_{REF} for one participant (N=1) and 30 steps (n=30). Bottom diagram shows the corresponding shank angle error $|\alpha_{REF} - \alpha_{CLT}|$ as box-plots depicting the *minimum to maximum*, the *lower to upper quartile* and the *median* error for each percent of the gait cycle. The mean error at medium walking speed was calculated to be $1.8 \pm 1.4°$

Table 1 Static performance of CoLiTrack

N=1 / n=30	% of Gait cycle		Shank angle error (deg)	
Speed	Min/max	$\overline{x} / \sigma$	Min/max	$\overline{x} / \sigma$
Slow	100/100	100/0	$1.4e^{-4}$/09.3	1.4/1.2
Medium	100/100	100/0	$4.4e^{-4}$/11.3	1.8/1.4
High	100/100	100/0	$3.2e^{-5}$/14.1	**3.4/1.9**

medium (1.0 km/h) and high (1.5 km/h), corresponding to the gait speed determined for amputees with a low mobility grade [54], which is slower than healthy subjects' walking speed. The second (reference) IMU was mounted in a modified support stocking on the participant's contralateral (right) shank, as shown in Fig. 3a. In this experiment, the point cloud transformation step was not necessary because the camera did not move. However, the input scene was cropped to the area of the (contralateral) leg of interest.

This experiment was carried out with one participant (N=1) at all three walking speeds (slow, medium and high). A total of 30 steps (n=30) was extracted from each walking speed level. As shown in Fig. 5, the CoLiTrack estimations corresponded closely to the reference measurements. The individual box-plots—median, lower and upper quartile, as well as the minimum to maximum range—depict the corresponding tracking error for each percent of the gait cycle. The mean error for medium walking speed was about $1.8\pm1.4°$. Results for the other walking speeds were summarized in Table 1. The highest mean error of $3.4\pm1.9°$ was measured at high walking speed.

4.2 Dynamic performance

The goal of the *dynamic performance* experiment was to determine how the accuracy and the tracking range of our CoLiTrack method might vary across different walking speeds. Therefore, our system was tested by five participants on a treadmill. Parameters were set as previously in the static test—walking speed levels of slow, medium and high at an incline of 1%. The wearable support (depth camera and IMU combination) was mounted on the ipsilateral (left) shank of the participant, and a second (reference) IMU was installed in a modified support stocking on the contralateral (right) shank, as shown in Fig. 3a. To begin, offset calibration was done for each participant in an upright standing position without movement. These parameters were stored and used for all subsequent experiments.

This experiment was carried out with five participants (N=5) at all three walking speeds (slow, medium and high). Instrumentation and calibration took less than 10 minutes for all of them. For the statistical evaluation, a total of 150 steps were combined, 30 steps per speed level (n=30) from each participant (N=5). The dynamic test revealed a trackable range of about one sixth of the total gait cycle, coinciding with the end of the swing phase, when the heel strike initiates the next step, as indicated in Fig. 1. The results from the dynamic test at medium walking speed are shown in Fig. 6. With our CoLiTrack method, it was possible to estimate the contralateral shank angle from about 97–14% of the gait cycle. For some steps, tracking was even possible for longer periods—up to 28% of the entire cycle, as shown in the magnified area of the plot. Although the maximum estimation error for all individual steps at medium walking speed was 13.7°, the mean estimation error was only $2.4\pm2.0°$. The highest mean error of $2.8\pm2.1°$ was measured at slow walking speed. Table 2 reports the results for the other walking speeds, showing similar values for all three speed levels. Our claim that the proposed method works independently from the walking speed was thus confirmed.

The averaged computation time for the overall process of CoLiTrack was 50 ms: Data read-in from depth camera and IMUs took about 10 ms, preprocessing about 25 ms and fitting, finally, the remaining 15 ms. The processing time of RANSAC, however, was almost negligible. Therefore, processing speed can be as high as 20 frames/s.

4.3 Real-world performance

Finally, the goal of the third experiment was to validate the real-world performance of our CoLiTrack approach, as this is a crucial factor for a successful implementation in a future product. Therefore, an online walking test was administered to evaluate the algorithm behavior with unknown objects in the camera's field of view and to qualify the estimation per-
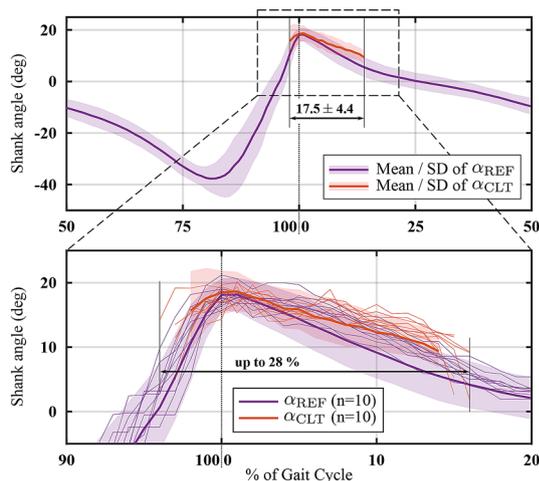
Fig. 6 Dynamic performance visualization of CoLiTrack at medium walking speed. Top diagram depicts the mean and the standard deviation of the α_{CLT} and α_{REF} for all five participants (N=5) and 30 steps (n=30) each. The x-axis was shifted for a better visualization. Bottom diagram magnifies the total tracking area and shows 10 randomly taken curves out of all 150 steps. The mean error at medium walking speed was calculated to be 2.4°

Table 2 Dynamic performance of CoLiTrack

N=5 / n=30	% of Gait cycle		Shank angle error (deg)	
Speed	Min/max	$\overline{x}$ / σ	Min/max	$\overline{x}$ / σ
Slow	7/29	16.7/4.6	$4.7e^{-3}$/16.9	**2.8/2.1**
Medium	6/28	17.5/4.4	$1.6e^{-3}$/13.7	2.4/2.0
High	9/26	18.4/3.8	$1.0e^{-3}$/14.1	2.4/1.9

formance in other types of terrains, such as up/down ramps or stairs. The instrumentation for the real-world test was identical to the dynamic experiment (wearable support on the left leg and reference IMU on the right leg), apart from the laptop for data acquisition, which was carried in a backpack on the paticipant's back.

This experiment was carried out with one participant (N=1) at a self-selected walking speed, starting with normal walking on level ground, before going into other terrains, as shown in Fig. 7. Level-ground walking led mostly to results similar to the treadmill evaluation. Gait initiation and termination steps were successfully tracked, too. However, due to the variability of the walking speed, there was no statistical evaluation carried out. If the participant walked too fast, tracking failed due to the limited update rate of a maximum of 20 frames/s.

As long as the contralateral leg was the closest object in the depth camera's field of view, unknown other obstacles were successfully suppressed. If the leg was out of view, nearby objects such as banisters or even another person standing in front of the participant occasionally led to misclassifications.
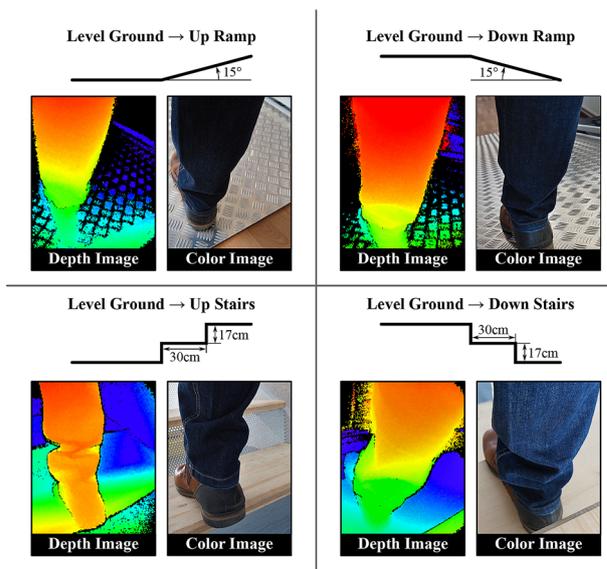

Fig. 7 Visualization of terrain changes from level ground into up/down ramp and stairs. *Depth Images* were taken straight out of the depth camera without any preprocessing. *Color Images* were taken with a mobile phone and are for comparison only

Ground reflectance or clutter, however, had no influence on the estimation performance.

As the camera positioning was optimized for level-ground walking, going into other terrains increased the error. Transitioning from level ground to inclines had almost no influence on the estimation performance, while going from level ground down a ramp reduced it. Still, trackability for ramps worked better than for stairs: In the case of stairs, estimation mostly failed, both upwards and downwards. Although the contralateral leg was still in the camera's field of view, as shown in Fig. 7, wrinkles in the shoe area of the trousers were the main cause for preventing a successful evaluation. In comparison, when walking on level ground, more of the proximal part of the shank was in view, where clothing was normally less wrinkled, as depicted in Fig. 2.

5 Discussion

5.1 Advantages of the proposed method

This paper introduced a robust contralateral limb tracking strategy for enhancing lower limb device control. To the best of the authors' knowledge, this is the first concept capable of detecting contralateral shank parameters only from unilaterally worn sensors and in real time. Several qualities of our proposed CoLiTrack might underline its effectiveness for enhancing "next-generation prostheses."

Firstly, the low computing time of only 50 ms is the most important achievement of our concept to be mentioned. This

allows an online evaluation up to 20 frames/s—fast enough to be implemented in a prosthetic device. Compared to the previous research by Hu et al. [40], this is a decrease by a factor of 23. Furthermore, we do not need any prior network training, as we use direct estimation strategies, guaranteeing that the presented CoLiTrack system is user-independent.

In addition to high processing speed, the estimation accuracy was also high. During the static experiment, the presented algorithm successfully tracked the leg over the entire gait cycle with a maximum mean error of 3.4±1.9°. Higher errors were found in the area between 85 and 95% of the duration of the gait cycle (Fig. 5). They are not caused by estimation errors, but can rather be explained by imperfectly synchronized data read-in procedures. The slight increase in errors at the beginning of the gait cycle between 5 to 30%, however, might be explained by a relative movement of the trousers with respect to the shank. In the dynamic experiment, tracking was possible for about one sixth of the full gait cycle with a mean estimation error below 2.8±2.1°. Given that a joint angle difference of more than 5° is considered a clinically significant difference for gait analysis [55], we claim the accuracy of our CoLiTrack approach to be sufficient. In comparison, results from instrumented crutches [56] as well as from a smart walker (rollator + depth camera) [57] using principal component analysis to estimate the shank angle showed deviations of up to 10°. Moreover, our real-world experiment demonstrated that neither ground reflectance nor clutter has an influence on CoLiTrack performance. As long as the contralateral leg is in the field of view of the camera, other (unknown) objects are eliminated, which underlines the efficiency of our preprocessing approach.

Finally, through the integration of depth camera and IMU into a compact wearable support frame, instrumentation and calibration of the system on the participants leg take less than 10 minutes. Since the kit is worn on the ipsilateral (prosthesis side) leg, the additional sensors could be embedded directly into a lower limb prosthetic device in the future.

5.2 Limitation and future work

Although the proposed method can estimate the contralateral shank axis accurately and with only a short time delay, there are some limitations, which need to be addressed.

Firstly, the trackable gait cycle range needs to be increased. So far, our system successfully estimates the contralateral shank axis in the range of one sixth of the full gait cycle, limited by the field of view of the camera. Several depth cameras can be combined into a camera array, increasing the field of view and, thus, the trackable range. Furthermore, the optimal position and orientation of the system for other terrains, such as stairs or ramps, need to be defined, in order to capture those trouser areas that are less wrinkled. In addition,

the preprocessing step of the depth data could be extended to make it more robust against unknown objects. Simple solutions, for example, could be to remove depth areas below a minimum size or to select the region of interest based on the last valid position of the shank. In this case, it would not be necessary to implement more computationally intensive approaches of background subtraction, in order to increase the robustness of the algorithm. Nevertheless, in its current form, the system is capable of tracking the contralateral shank angle within its area of coverage. This information can be used for the control of the next step. Additionally, it seems possible to determine the heel strike with the help of the shank's angular velocity, as suggested in [58]. A positive velocity indicates the swing phase, a negative velocity indicates the stance phase. This would allow to derive the timing of the heel strike, representing the beginning of the human gait double support phase, as depicted in Fig. 1. Furthermore, the distance between the contralateral shank and the ipsilateral sensor—inherently detected by a depth camera, but not utilized here—can be used to calculate spatial parameters, such as step length [11].

Additionally, even though the system in its current form is real time capable with an update rate of up to 20 frames/s, practice has shown that higher walking speeds can lead to misclassifications. Considering a walking speed of 3.6 km/h for very active amputees [54] and assuming a step length (heel strike to heel strike of the same leg) of 1 m, we can expect one gait cycle per second. If we further resume that CoLiTrack covers approximately one sixth of the gait cycle and is sampling with an update rate of 20 frames/s, we may calculate that no more than 3–4 images can be captured per gait cycle. Although our approach does not rely on time history, we estimate that a minimum of 10 images is required for use in real life in order to derive heel strike timing as mentioned above. This can be achieved by increasing the update rate to at least 60 Hz. On the one hand, this is naturally constrained by the frame rate of the camera. For example, the one used in our study is limited to 45 frames/s according to the manufacturer's data sheet. On the other hand, calculation time needs to be reduced even further. This can be done by converting MATLAB code into C++ programs and running it on dedicated vision processors.

Finally, the depth camera used in our work is also restricted by the requirement of having an unobstructed field of view. Although sensor positioning would allow an implementation directly into a prosthetic device, clothing or prosthetic covers cannot be worn above it. Instead, high-frequency super near-field radar sensors are able to "look through" them. For example, Google's radar-based gesture-sensing technology (Project Soli) allows a touchless interaction with their new smartphone Pixel 4 [59]. Therefore, the use of these sensors should be considered in future work.

6 Conclusion

In this work, we present a robust real-time leg tracking method called CoLiTrack for improved lower limb prosthetic device control. A depth camera and an IMU are placed on the ipsilateral (prosthetic) leg, capable of estimating the axis of the contralateral (healthy residual) leg. We conducted three experiments to validate our proposed CoLiTrack algorithm. The evaluation of *static performance* demonstrated a trackability of the shank axis throughout the entire gait cycle with a mean error of less than $3.4\pm1.9°$ for one participant. *Dynamic performance* was evaluated with five participants wearing the sensor kit while walking on a treadmill at three different speeds. This resulted in a mean trackable range of one sixth of the entire gait cycle, since the leg is out of the camera's field of view for the remaining time. The overall processing time of the presented CoLiTrack system took less than 50 ms, and the mean estimation errors for all walking speed levels were below $2.8\pm2.1°$. Finally, the *real-world performance* testing with one participant demonstrated robustness against ground reflectance or clutter, but showed the limitations of the approach in terms of walking speed and terrain variations. Our immediate plan is to enlarge the trackable range by increasing the field of view, as well as to reduce the processing time even further, in order to use the residual healthy leg information in movement-dependent control applications of "next-generation prostheses."

Declarations

References

1. Wu, J., Hu, D., Xiang, F., Yuan, X., Su, J.: 3D human pose estimation by depth map. Vis. Comput. (2020). https://doi.org/10.1007/s00371-019-01740-4

2. Zhang, Y., Tan, F., Wang, S., Yin, B.: 3D human body skeleton extraction from consecutive surfaces using a spatial-temporal consistency model. Vis. Comput. (2020). https://doi.org/10.1007/s00371-020-01851-3

3. Liu, Z., Zhu, J., Bu, J., Chen, C.: A survey of human pose estimation: the body parts parsing based methods. J. Vis. Commun. Image Represent. **32**, 10–19 (2015)

4. Antón, D., Goñi, A., Illarramendi, A., Torres-Unda, J.J., Seco, J.: KiReS: A Kinect-based telerehabilitation system. In 2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (2013). https://doi.org/10.1109/HealthCom.2013.6720717

5. Naeemabadi, M., Dinesen, B., Andersen, O., Najafi, S., Hansen, J.: Evaluating accuracy and usability of microsoft kinect sensors and wearable sensor for tele knee rehabilitation after knee operation. In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (2018). https://doi.org/10.5220/0006578201280135

6. Gavrilova, M.L., Wang, Y., Ahmed, F., Paul, P.P.: Kinect sensor gesture and activity recognition: new applications for consumer cognitive systems. IEEE Consum. Electr. Mag. (2018). https://doi.org/10.1109/MCE.2017.2755498

7. Saini, R., Kumar, P., Kaur, B., Roy, P.P., Dogra, D.P., Santosh, K.C.: Kinect sensor-based interaction monitoring system using the BLSTM neural network in healthcare. Int. J. Mach. Learn. Cybern. (2019). https://doi.org/10.1007/s13042-018-0887-5

8. Fluit, R., Prinsen, E.C., Wang, S., van der Kooij, H.: A comparison of control strategies in commercial and research knee prostheses. IEEE Trans. Biomed. Eng. (2020). https://doi.org/10.1109/TBME.2019.2912466

9. Hu, B., Rouse, E., Hargrove, L.: Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals. Front. Robot. AI (2018). https://doi.org/10.3389/frobt.2018.00014

10. Hu, B., Rouse, E., Hargrove, L.: Fusion of bilateral lower-limb neuromechanical signals improves prediction of locomotor activities. Front. Robot. AI (2018). https://doi.org/10.3389/frobt.2018.00078

11. Perry, J., Burnfield, J.: Gait Analysis: Normal and Pathological Function, 2nd edn. Slack Incorporated, Thorofare, NJ, USA (2010)

12. Li, Q., Wang, Y., Sharf, A., Cao, Y., Tu, C., Chen. B., Yu, S.: Classification of gait anomalies from kinect. Vis. Comput. (2018). https://doi.org/10.1007/s00371-016-1330-0

13. Wang, K., Zhang, G., Yang, J., Bao, H.: Dynamic human body reconstruction and motion tracking with low-cost depth cameras. Vis. Comput. (2020). https://doi.org/10.1007/s00371-020-01826-4

14. Colyer, S.L., Evans, M., Cosker, D.P., Salo, A.I.T.: A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports Med. Open (2018). https://doi.org/10.1186/s40798-018-0139-y

15. Latorre, J., Colomer, C., Alcañiz, M., Llorens, R.: 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. Rehab. (2019). https://doi.org/10.1186/s12984-019-0568-y

16. Murray, M.P.: Gait as a total pattern of movement. Am. J. Phys. Med. **46**, 290–333 (1967)

17. Elaine, O.: The importance of being earnest about shank and thigh kinematics especially when using ankle-foot orthoses. Prosthet. Orthot. Int. (2010). https://doi.org/10.3109/03093646.2010.485597

18. Ballit, A., Mougharbel, I., Ghaziri, H., Dao, T.T.: Computer-aided parametric prosthetic socket design based on real-time soft tissue deformation and an inverse approach. Vis. Comput. (2021). https://doi.org/10.1007/s00371-021-02059-9

19. Hargrove, L.J., Huang, H., Schultz, A.E., Look, B.A., Lipschutz, R., Kuiken, T.A.: Toward the development of a neural interface for lower limb prosthesis control. In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2009). https://doi.org/10.1109/IEMBS.2009.5334303

20. Varol, H.A., Sup, F., Goldfarb, M.: Multiclass real-time intent recognition of a powered lower limb prosthesis. IEEE Trans. Biomed. Eng. (2010). https://doi.org/10.1109/TBME.2009.2034734

21. Young, A.J., Kuiken, T.A., Hargrove, L.J.: Analysis of using EMG and mechanical sensors to enhance intent recognition in powered lower limb prostheses. J. Neural Eng. (2014). https://doi.org/10.1088/1741-2560/11/5/056021

22. Tschiedel, M., Russold, M.F., Kaniusas, E.: Relying on more sense for enhancing lower limb prostheses control: a review. J. NeuroEng. Rehab. (2020). https://doi.org/10.1186/s12984-020-00726-x

23. Liu, M., Wang, D., Helen, H.: Development of an environment-aware locomotion mode recognition system for powered lower limb prostheses. IEEE Trans. Neural Syst. Rehab. Eng. (2016). https://doi.org/10.1109/TNSRE.2015.2420539

24. Yan, T., Sun, Y., Liu, T., Cheung, C.H., Meng, M.Q.H.: A locomotion recognition system using depth images. In 2018 IEEE International Conference on Robotics and Automation (ICRA) (2018). https://doi.org/10.1109/ICRA.2018.8460514

25. Massalin, Y., Abdrakhmanova, M., Varol, H.A.: User-independent intent recognition for lower limb prostheses using depth sensing. IEEE Trans. Biomed. Eng. **65**, 1759–1770 (2018)

26. Zhang, K., Xiong, C., Zhang, W., Liu, H., Lai, D., Rong, Y., Fu, C.: Environmental features recognition for lower limb prostheses toward predictive walking. IEEE Trans. Neural Syst. Rehab. Eng. (2019). https://doi.org/10.1109/TNSRE.2019.2895221

27. Laschowski, B., McNally, W., Wong, A., McPhee, J.: Preliminary design of an environment recognition system for controlling robotic lower-limb prostheses and exoskeletons. IN 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR) (2019). https://doi.org/10.1109/ICORR.2019.8779540

28. Krausz, N.E., Lenzi, T., Hargrove, L.J.: Depth sensing for improved control of lower limb prostheses. IEEE Trans. Biomed. Eng. **62**, 2576–2587 (2015)

29. Diaz, J.P., da Silva, R.L., Zhong, B., Huang, H., Lobaton, E.: Visual terrain identification and surface inclination estimation for improving human locomotion with a lower-limb prosthetic. In Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2018). https://doi.org/10.1109/embc.2018.8512614

30. Ishikawa, T., Murakami, T.: Real-time foot clearance and environment estimation based on foot-mounted wearable sensors, In IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society (2018). https://doi.org/10.1109/IECON.2018.8592894

31. St-Onge, N., Feldman, A.G.: Interjoint coordination in lower limbs during different movements in humans. Exp. Brain Res. **148**, 139–149 (2003)

32. Grimes, D.L., Flowers, W.C., Donath, M.: Feasibility of an active control scheme for above knee prostheses. J. Biomech. Eng. **99**, 215–221 (1977)

33. Borjian, R., Khamesee, M., Melek, W.: Feasibility study on echo control of a prosthetic knee: sensors and wireless communication. Microsyst. Technol. **16**, 257–265 (2010)

34. Vallery, H., Ekkelenkamp, R., Buss, M., van der Kooij, H.: Complementary limb motion estimation based on interjoint coordination: experimental evaluation. In 2007 IEEE 10th International Conference on Rehabilitation Robotics (2007). https://doi.org/10.1109/ICORR.2007.4428516

35. Bernal-Torres, M.G., Medellín-Castillo, H.I., Arellano-González, J.C.: Design and control of a new biomimetic transfemoral knee prosthesis using an echo-control scheme. J. Healthc. Eng. (2018). https://doi.org/10.1155/2018/8783642

36. Mendez, J., Hood, S., Gunnel, A., Lenzi, T.: Powered knee and ankle prosthesis with indirect volitional swing control enables level-ground walking and crossing over obstacles. Sci. Robot. (2020). https://doi.org/10.1126/scirobotics.aba6635

37. Ambrozic, L., Gorsic, M., Geeroms, J., Flynn, L., Molino Lova, R., Kamnik, R., Munih, M., Vitiello, N.: CYBERLEGs: a user-oriented robotic transfemoral prosthesis with whole-body awareness control. IEEE Robot. Autom. Mag. **21**, 82–93 (2014)

38. Goršič, M., Kamnik, R., Ambrožič, L., Vitiello, N., Lefeber, D., Pasquini, G., Munih, M.: Online phase detection using wearable sensors for walking with a robotic prosthesis. Sensors (Basel) (2014). https://doi.org/10.3390/s140202776

39. Parri, A., Martini, E., Geeroms, J., Flynn, L., Pasquini, G., Crea, S., Molino Lova, R., Lefeber, D., Kamnik, R., Munih, M., Vitiello, N.: Whole body awareness for controlling a robotic transfemoral prosthesis. Front. Neurorobot. (2017). https://doi.org/10.3389/fnbot.2017.00025

40. Hu, B.H., Krausz, N.E., Hargrove, L.J.: A novel method for bilateral gait segmentation using a single thigh-mounted depth sensor and IMU. In 2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob) (2018). https://doi.org/10.1109/BIOROB.2018.8487806

41. Stigler, S.M.: Gauss and the invention of least squares. Ann. Stat. (1981). https://doi.org/10.1214/aos/1176345451

42. Harms, H., Beck, J., Ziegler, J., Stiller, C.: Accuracy analysis of surface normal reconstruction in stereo vision. In 2014 IEEE Intelligent Vehicles Symposium Proceedings (2014). https://doi.org/10.1109/IVS.2014.6856436

43. Balaji, S.R., Karthikeyan, S.: A survey on moving object tracking using image processing. In 2017 11th International Conference on Intelligent Systems and Control (ISCO) (2017). https://doi.org/10.1109/ISCO.2017.7856037

44. Zohora F.T., Santosh, K.C.: Circular Foreign Object Detection in Chest X-ray Images. In: Santosh, K., Hangarge, M., Bevilacqua, V., Negi, A. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2016. Communications in Computer and Information Science, vol. 709. Springer, Singapore (2017)

45. Zohora, F.T., Antani, S., Santosh, K.C.: Circle-like foreign element detection in chest x-rays using normalized cross-correlation and unsupervised clustering. In Proceedings of the SPIE 10574, Medical Imaging 2018: Image Processing (2018). https://doi.org/10.1117/12.2293739

46. Santosh, K.C., Dhar, M.K., Rajbhandari, R., Neupane, A.: Deep neural network for foreign object detection in chest x-rays. In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS) (2020). https://doi.org/10.1109/CBMS49503.2020.00107

47. Yi, S.C., Chen, Y.C., Chang, C.H.: A lane detection approach based on intelligent vision. Comput. Electr. Eng. (2015). https://doi.org/10.1016/j.compeleceng.2015.01.002

48. Liang, D., Guo, Y.C., Zhang, S.K., Mu, T.J., Huang, X.: Lane detection: a survey with new results. J. Comput. Sci. Technol. (2020). https://doi.org/10.1007/s11390-020-0476-4

49. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Machi. Intell. (2010). https://doi.org/10.1109/TPAMI.2010.46

50. Bergström, P., Edlund, O.: Robust registration of point sets using iteratively reweighted least squares. Comput. Optim. Appl. **58**, 543–561 (2014)

51. Chang, W.C., Wu, C.H.: Candidate-based matching of 3-D point clouds with axially switching pose estimation. Vis. Comput. **36**, 593–607 (2020)

52. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM (1981). https://doi.org/10.1145/358669.358692

53. Jones, A., Doust, J.: A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. J. Sports Sci. (1996). https://doi.org/10.1080/02640419608727717

54. Batten, H.R., McPhail, S.M., Mandrusiak, A.M., Varghese, P.N., Kuys, S.S.: Gait speed as an indicator of prosthetic walking potential following lower limb amputation. Prosthet. Orthot. Int. (2019). https://doi.org/10.1177/0309364618792723

55. McGinley, J.L., Baker, R., Wolfe, R., Morris, M.E.: The reliability of three-dimensional kinematic gait measurements: a systematic review. Gait Posture (2009). https://doi.org/10.1016/j.gaitpost.2008.09.003

56. Pasinetti, S., Hassan, M.M., Eberhardt, J., Lancini, M., Docchio, F., Sansoni, G.: Performance analysis of the PMD camboard picoflexx time-of-flight camera for markerless motion capture applications. IEEE Trans. Instrum. Meas. **68**, 4456–4471 (2019)

57. Page, S., Martins, M.M., Saint-Bauzel, L., Santos, C.P., Pasqui, V.: Fast embedded feet pose estimation based on a depth camera for smart walker. In 2015 IEEE International Conference on Robotics and Automation (ICRA) (2015). https://doi.org/10.1109/ICRA.2015.7139781

58. Grimmer, M., Schmidt, K., Duarte, J.E., Neuner, L., Koginov, G., Riener, R.: Stance and swing detection based on the angular velocity of lower limb segments during walking. Front. Neurorobot. (2019). https://doi.org/10.3389/fnbot.2019.00057

59. Lien, J., Gillian, N., Karagozler, M.E., Amihood, P., Schwesig, C., Olson, E., Raja, H., Poupyrev, I.: Soli: ubiquitous gesture sensing with millimeter wave radar. ACM Trans. Graph. (2016). https://doi.org/10.1145/2897824.2925953

Michael Tschiedel was born in Vienna, Austria, in 1994. He received his B.Sc. degree in Electrical Engineering and his diploma (Dipl.-Ing.) in Biomedical Engineering from TU Wien in 2016 and 2018, respectively. He is currently an industry-related doctoral candidate at TU Wien and Ottobock. His research interests include computer vision and environmental sensing for prosthetic devices.

Michael Friedrich Russold was born in Vienna, Austria, in 1974. He obtained a masters degree (Dipl.-Ing.) from TU Wien. He then received a PhD degree from the University of Liverpool, UK. Afterwards, he worked as a post-doc at the University of Sydney, Australia, under a Marie Curie individual outgoing Fellowship. Michael Russold is author and co-author of more than 20 publications. Since 2008, Michael Russold works for Ottobock Healthcare Products GmbH, where he worked in product development and research. In the latter role, he has been involved in several international research projects, based around national and international funding.

Eugenijus Kaniusas is a professor in the field of biomedical theranostics and heads the research group Biomedical Sensing at TU Wien. His interests include auricular vagus nerve stimulation, prosthetics, perioperative monitoring, and wearable hardware/software concepts for diagnostic/therapeutic/theranostic devices. He gives numerous mandatory lectures at TU Wien, including Biophysics, Biomedical Sensors and Signals, Biomedical Instrumentation. Currently, he is the dean of academic affairs Biomedical Engineering at TU Wien.

Markus Vincze received his diploma in mechanical engineering from TU Wien in 1988 and a M.Sc. from Rensselaer Polytechnic Institute, USA, in 1990. He finished his PhD at TU Wien in 1993. With a grant from the Austrian Academy of Sciences, he worked at HelpMate Robotics Inc. and at the Vision Laboratory of Gregory Hager at Yale University. In 2004, he obtained his habilitation in robotics. Presently, he leads the "Vision for Robotics" team at TU Wien with the vision to make robots see. Markus' special interests are cognitive computer vision techniques for robotics solutions situated in real-world environments.