

Proceedings

Collection of Kinematic and Kinetic Data of Young and Adult, Male and Female Subjects Performing Periodic and Transient Gait Tasks for Gait Pattern Recognition [†]

Paolo Mistretta ^{1,*}, Cecilia Marchesini ¹, Andrea Volpini ¹, Luca Tagliapietra ², Tommaso Sciarra ³, Aldo Lazich ^{3,4}, Salvatore Forte ⁵, Mauro De Matteis ⁵, Emanuele Menegatti ² and Nicola Petrone ¹

¹ Department of Industrial Engineering, University of Padua, 35121 Padua, Italy;

der.suchende@hotmail.it (C.M.); volpini.andrea1994.av@gmail.com (A.V.); nicola.petrone@unipd.it (N.P.)

² Department of Information Engineering, University of Padua, 35121 Padua, Italy;

tagliapietra.work@gmail.com (L.T.); emg@dei.unipd.it (E.M.)

³ Joint Veteran Center, Scientific Department, Army Medical Center, 00184 Rome, Italy;

sciarratommaso@hotmail.com (T.S.); lazich@diag.uniroma1.it (A.L.)

⁴ DIAG, Sapienza University of Rome, 00185 Rome, Italy

⁵ MES S.p.A. Meccanica per l'Elettronica e Servomeccanismi, 00131 Rome, Italy; s.forte@mesroma.it (S.F.);

m.dematteis@mesroma.it (M.D.M.)

* Correspondence: paolo.mistretta.1992@gmail.com; Tel.: +39-349-084-9997

† Presented at the 13th conference of the International Sports Engineering Association, Online, 22–26 June 2020.

Published: 15 June 2020

Abstract: The aim of the study was to develop a database of biomechanical data for multiple gait tasks. This database will be used to create a real-time gait pattern classifier that will be implemented in a new-generation active knee prosthesis. With this intent, we collected kinematic and kinetic data of 40 subjects performing 16 gait tasks, categorized as periodic and transient motions. We analyzed four distinct sub-populations, differentiated by age and gender. As the classifier will be based also on inertial data, we chose to synthesize these signals within the motion capture environment. To assess the effects of gender and age we performed a correlation analysis on the signals used as input of the classifier. The results obtained indicate that there is no need to differentiate into four distinct classes for the development of the classifier. Sample data of the dataset are made publicly available.

Keywords: database; populations; multiple gait-task; classifier; correlation; virtual IMU

1. Introduction

In this paper we present the data collection and analysis of a biomechanical database comprehensive of multiple gait tasks. This database will be used to develop a real-time classifier implemented in a new-generation active knee prosthesis, called GARP (*Ginocchio Artificiale a Rigenerazione di Potenza*, i.e., artificial knee with power regeneration).

The classifier will implement a machine learning-based algorithm called random forest, similar to other works present in literature [1,2]. The classifier outcome will be based on data coming from a rotary encoder on the knee joint and a six-axis IMU (inertial measurement unit) located inside the prosthetic element corresponding to the biological shank. To implement this control system, a comprehensive database is needed both to train the classifier and to develop the control logic.

We chose to create the database instead of using data from literature for multiple reasons.

First, even if many databases can be found in literature [3–7], only a few of them give full access to all data, as they typically present only means and standard deviations. Typical classifiers, such as the random forest, instead need to be trained using data from as many single trials as possible.

Secondly, we wanted to have a database containing certain specific gait tasks, and we were not able to find any database containing all those sought in the same dataset. As observed by different authors [3,4], it is not advisable to merge data from multiple sources, as some inconsistencies can arise due to the use by different research groups of different biomechanical protocols, experimental conditions, subject characteristics and so forth.

Finally, we chose to collect data without using an instrumented treadmill, as we did not want to bias the measure of the IMU due to the relative movement of the belt with respect to the ground. Further, the clinical validity of data collected on a treadmill is still under discussion [8].

Given these considerations, we decided to create our own comprehensive dataset. To do so, we built instrumented stairs and ramps, and we collected kinematic and kinetic data of 40 subjects, divided into four classes evenly distributed, based on age and gender. Subjects were asked to perform 16 activities of daily living, such as level walking, climbing stairs and sitting.

An important element of novelty of this database is that we collected data of two categories of motion, namely periodic and transient tasks. Periodic tasks are motions that are repeated cyclically, whereas transient ones represent the passage from one periodic gait task to the next one.

After collecting the data, we performed an analysis of correlation on the data that will be used to train the classifier of the knee prosthesis. The goal of the analysis was to assess if it would be more appropriate to treat the classes as separate ones when training the classifier, or if we could consider all subjects as belonging to a unique population.

Sample data of the dataset are made publicly available (see Supplementary Materials).

2. Materials and Methods

2.1. Subjects

Forty healthy subjects participated in the data collection. They were selected in order to form 4 evenly distributed sub-populations based on age and gender. The classes were labelled as young male (YM; 20–35 years old; age: 25.6 ± 1.8 y; height: 1.86 ± 0.06 m, weight: 76.9 ± 8.3 kg), young female (YF; 20–35 years old; age: 25.4 ± 2.9 y; height: 1.69 ± 0.06 m, weight: 58.3 ± 4.0 kg), adult male (AM; 40–60 years old; age: 48.2 ± 7.1 y; height: 1.77 ± 0.08 m, weight: 74.8 ± 9.8 kg) and adult female (AF; 40–60 years old; age: 49.3 ± 6.1 y; height: 1.62 ± 0.06 m, weight: 55.8 ± 8.2 kg).

2.2. Instrumentation

Kinematics data were collected using an 8-camera optoelectronic motion capture system (Smart-DX 6000, BTS Bioengineering), capturing images at 250 Hz. Kinetics signals were provided by force platforms (BTS P6000D, BTS Bioengineering) with a sampling rate of 1000 Hz.

We built instrumented wooden stairs and ramps (Figure 1). For the design of the stairs we took as a reference the regulation of Cybathlon [9] (step height: 170 mm; step width: 280 mm). The inclination of the ramp was set to be 10° .

In order to collect kinetic data, force platforms were included in the design of both the stairs and the ramp. Thanks to the modular design of the stairs, it was possible to move the force platforms into different positions within the structure (Figure 1a).

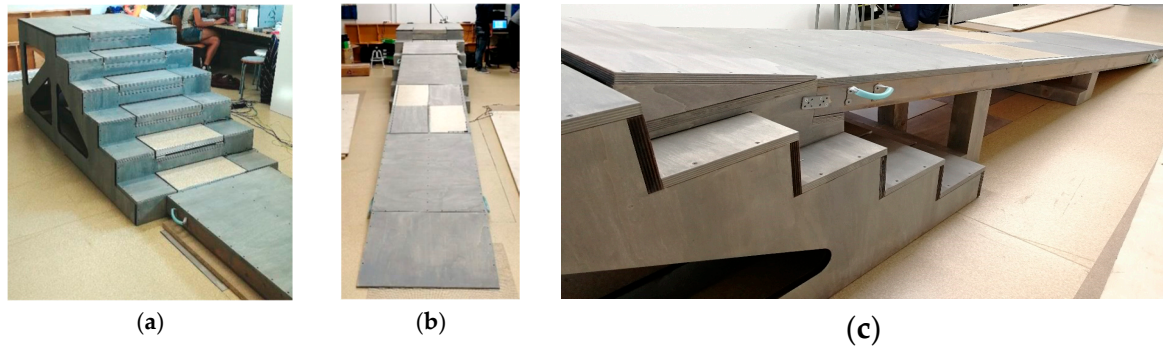


Figure 1. Instrumented stairs and ramp. (a) Stairs. In the photo the force platforms (beige) are located on the first two steps of the flight of stairs; (b,c) Ramp. The ramp is attached to the structure of the stairs.

2.3. Processing Protocol and Virtual IMU

We used a modified version of the built-in processing protocol of the Smart-DX software (Smart Clinic, BTS Bioengineering) called Helen Hayes MM, derived from the studies of Kadaba and Davis [10,11]. It uses 22 markers placed in specific anatomical reference locations.

Due to practicality reasons, we chose to synthesize the signals of linear acceleration and angular velocity coming from the IMU that was to be mounted on the prosthesis. In order to do so, we created a system of reference integral with the leg within the motion capture environment, coincident both in location and in orientation with the IMU itself. Subsequently, we derived twice the position and once the spatial orientation of the system of reference created. Finally, we added the gravity vector, properly reoriented, to the 3D acceleration obtained previously.

To assess the validity of this method we performed a test in-vitro, that consisted in comparing data coming from the physical and the virtual IMU sensors. To do so we attached an IMU on a T-shaped marker triad (Figure 2a), and then we collected data from both systems while moving them together, performing semi-random rapid movement (range of linear acceleration: ± 40 m/s², angular velocity: ± 800 deg/s) (Figure 2b). The results of the validation are presented in the section Results.

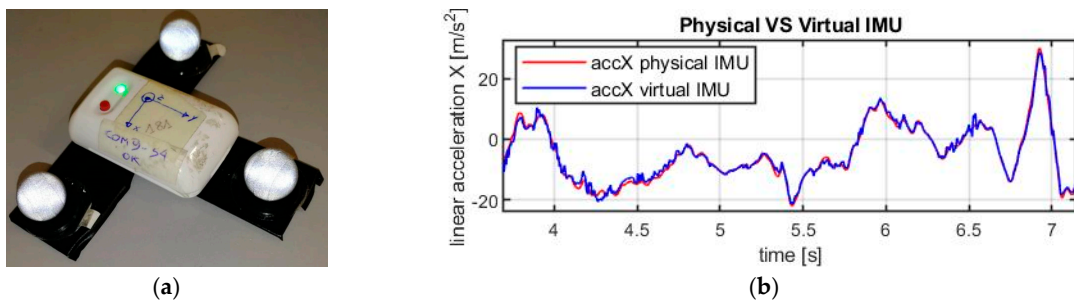


Figure 2. Assessment of the validity of the virtual IMU (inertial measurement unit) method. (a) Triad of reflective markers placed corresponding to the axes of the IMU (in white); (b) Example of comparison of data of linear acceleration coming from the physical (red) and virtual (blue) IMU sensor.

2.4. Data Collection

We collected lower limb kinematic and kinetic data of 40 subjects, performing 5 repetitions of 16 different gait tasks at self-selected speed, for a total of 3200 sets of signals. During data processing, some of them were discarded, mainly due to missing markers not noticed during data collection.

The 16 gait tasks were categorized as periodic and transient, as shown in Table 1. In periodic gait tasks the biomechanical pattern is repeated periodically, whereas transients represent the movement used to pass from one periodic state to the next. We chose to collect data of only 8 transient tasks over a total of 56 partial permutations ideally possible, because a huge number would make the data collection impractical due to time and fatigue constraints.

Table 1. Gait tasks collected in the database, together with their code.

Type	Description	Trial code	Type	Description	Trial Code
Periodic	Level walk neutral	P_NW	Transient	Sit to stand	T_SI_ST
	Level walk slow	P_SW		Stand to sit	T_ST_SI
	Uphill walk	P_UW		Stair up to stand	T_SU_ST
	Downhill walk	P_DW		Stand to stair down	T_ST_SD
	Stair up	P_SU		Stand to stair up	T_ST_SU
	Stair down	P_SD		Stair down to stand	T_SD_ST
	Standing	P_ST		Level walk to stair up	T_NW_SU
	Sitting	P_SI		Stair down to level walk	T_SD_NW

2.5. Correlation Analysis

We wanted to assess the best strategy to train the random forest gait-task classifier. We considered two options, the first being to keep the four sub-populations analyzed as distinct, and the second was to consider all the subjects belonging to the same unique class, hence maintaining a good number of signals per gait task.

We focused our analysis on the signals that are more representative of the motion in the sagittal plane (i.e., knee flexion angle, linear accelerations X, Z and IMU angular velocity Y), as the others present a higher noise-to-signal ratio.

We computed the Pearson correlation coefficient of all the combinations of signals corresponding to a specific gait task. We called “overall correlation coefficient” (OCC) the mean of the coefficient of correlations corresponding to the combination of all the available trials of a specific gait task. This metric is a measure of the overall repeatability of the signal under analysis. We called “class correlation coefficient” (CCC) the mean of all the coefficients of correlations of the combination of trials corresponding only to a specific sub-population (YM, YF, AM or AF) for a gait task. Using the latter formulation, the coefficient of correlations between waveforms belonging to different classes was not taken into account. The mean of the four CCCs obtained, called “mean class correlation coefficient” (mCCC), represents the repeatability of the signal within sub-population.

We repeated this procedure for all the 16 gait tasks analyzed, obtaining a vector of OCC and mCCC. We chose to exclude from this analysis the seating and standing tasks, due to their inherent poor self-correlation because of their small range of motion.

3. Results

3.1. Virtual IMU Validation

We compared data measured by a physical IMU with the corresponding virtual one. We confirmed the validity of the method, as the correlation between signals was always above 0.985 for all the 6 axes, even when performing highly dynamic movements (Figure 2b).

3.2. Database

We found good correspondence between the data we collected and that found in literature in terms of overall trends [3–7]. However, we chose not to perform any further statistical comparison between databases. In fact, it would be not trivial to distinguish if a possible difference in the signals has to be attributed to a real difference in the gait biomechanics, or it is imputable to the different processing protocol used. Some representative kinematics signals are reported in Figure 3.

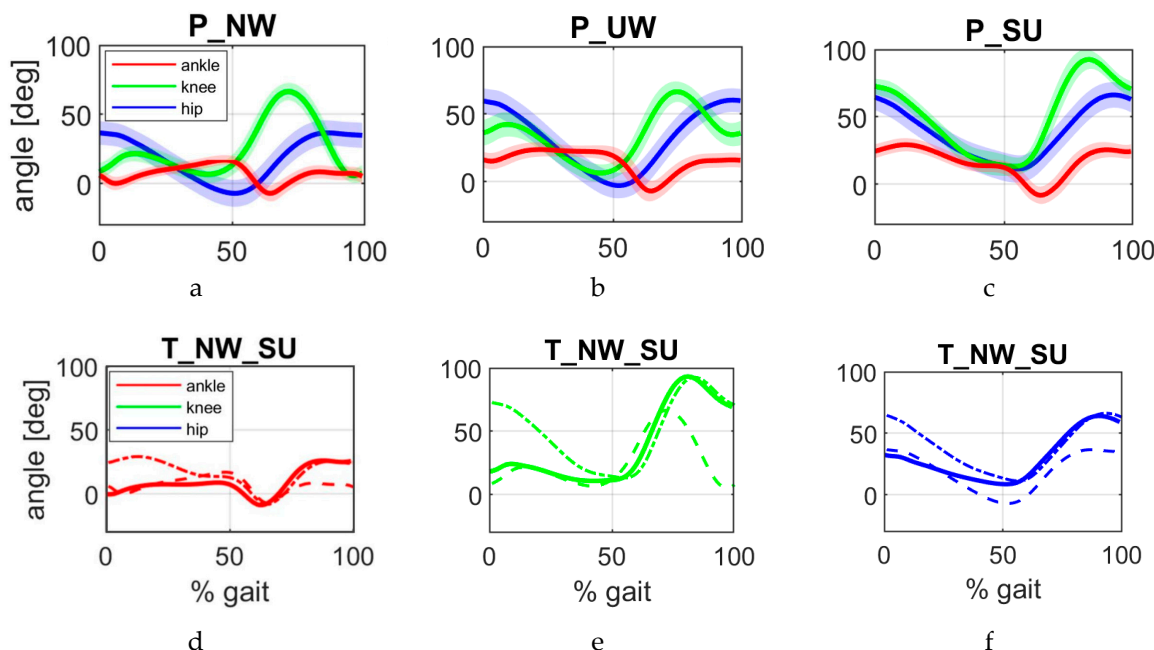


Figure 3. Ankle (red), knee (green) and hip (blue) flexion angles over a gait cycle (solid line: mean; shaded area: mean \pm 1 SD). See Table 1 to interpret the trial code. (a–c) Examples of periodic gait tasks. (d–f) Single joint analysis of the transient T_NW_SU (solid line: mean; dashed line: starting periodic task; dash-dot line: ending periodic task). (d) Ankle. (e) Knee. (f) Hip.

3.3. Correlation Analysis

Representative OCC and mCCC values are reported in Table 2.

Table 2. Overall correlation coefficient (OCC) and mean class correlation coefficient (mCCC) values of representative gait tasks.

Signal	P_NW		P_UW		P_SU		T_NW_SU	
	OCC	mCCC	OCC	mCCC	OCC	mCCC	OCC	mCCC
Knee angle	0.936	0.944	0.901	0.911	0.966	0.971	0.973	0.974
IMU linear acc X	0.831	0.845	0.720	0.747	0.831	0.849	0.861	0.864
IMU linear acc Z	0.707	0.732	0.634	0.659	0.641	0.670	0.582	0.604
IMU angular vel Y	0.926	0.940	0.886	0.906	0.880	0.894	0.874	0.881

We performed a paired t-test between the OCC and mCCC vectors, which refused the null hypothesis for all the four types of signals analyzed ($p_{knee} = 0.887$, $p_{linAccX} = 0.719$, $p_{linAccZ} = 0.699$, $p_{angVelY} = 0.726$). Therefore, the two vectors can be considered as belonging to the same population. For this reason, we chose not to distinguish between classes when training the classifier.

4. Discussion

In the present paper we reported several achievements obtained during the course of this work.

First of all, we were able to build instrumented stairs and ramps (Figure 1) that allowed us to collect kinematic and kinetic data of subjects performing a wide variety of activities of daily living.

Secondly, we developed a method that allowed us to synthesize six-axis IMU signals from motion capture data. We assessed the validity of the method by comparing signals coming from a physical and a virtual IMU (Figure 2). As the correlation of corresponding signals was above 0.985 for all the six axes of the IMU, we can consider this method valid.

Third, we performed a correlation analysis on the data that will be used in the gait classifier. It proved that, for the purpose of our study and in the above-mentioned conditions, we can rightfully train the classifier considering a unique population, composed by all the subjects belonging to the four

sub-populations analyzed. Moreover, by following this strategy, the classifier will show more robustness in classifying more scattered signals.

Finally, in this work we were able to compile a kinematics and kinetics database comprehensive of multiple gait tasks, that presents several features that make it particularly valuable.

First of all, it contains all the data of every single trial. Conversely, typical biomechanical databases available in literature report only means and standard deviations of the metrics studied. This aspect is of great importance, particularly when applying a machine learning algorithm, that typically requires a huge number of samples to be trained properly.

Secondly, it contains a huge number of kinematic and kinetic sets of trial data, namely 2605. This is due to the huge number of subjects enrolled, and the large variety of gait tasks analyzed.

Third, it contains some data relative to transient movement between periodic gait tasks. We hypothesize that this feature has the potential to improve gait classification reliability, by making possible the categorization and classification of these types of movements. In fact, once the classifier recognizes a specific transient motion, and by knowing the previous periodic gait task, it could hypothetically be able to “predict” the consecutive periodic gait task. As shown in Figure 3, the kinematic pattern of transient movements starts from the kinematic pattern of the previous periodic task and ends following the pattern of the consecutive periodic task. To our knowledge, this is the first work that proves this behavior.

Fourth, as every single set of data contains the position and orientation of the system of reference integral with every single lower limb, we are able to generate other virtual IMUs using the virtual IMU method. This potentially allows us to optimize the position, orientation and number of IMUs needed to maximize the classifier’s reliability.

A limitation of this work is that the database contains data of healthy subjects, who can have different biomechanical patterns compared to amputees [12,13] depending on many factors, such as amputation level or ankle-foot prosthesis model used. This choice was taken due to the difficulty of enrolling a large number of subjects with amputation. Nevertheless, a classifier trained on this dataset could represent an initial starting point that could be refined using subject-specific data.

Sample data of our dataset are made publicly available together with some more technical details about the processing protocol (see Supplementary Materials).

5. Conclusions

In this work we compiled a kinematic and kinetic database suitable to be used to create a gait task classifier. It presents several elements that make it valuable, such as the number of subjects enrolled, the number of gait tasks analyzed, the presence of periodic and transient gait tasks and the possibility to generate six-axis virtual IMU signals starting from kinematic data.

Supplementary Materials: Sample data of our dataset and more technical information about the processing protocol are available online at <http://doi.org/10.5281/zenodo.3628229>.

Acknowledgments: This work was supported by MES S.p.A. and by Italian Ministry of Defence, Italian Joint Defence Veteran Center, Scientific Department, Army Medical Center.

Author Contributions: N.P., E.M., M.D.M., S.F., A.L. and T.S. contributed to the study design. N.P. and P.M. coordinated and supervised the study. P.M., C.M. and A.V. conducted the data collection. P.M. organized and analyzed the results. L.T. supported the organization of the database. P.M. drafted the manuscript.

Funding: This research was funded by MES S.p.A. and by Italian Ministry of Defence, Italian Joint Defence Veteran Center, Scientific Department, Army Medical Center, grant number N. 699 of 18/12/2017.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Feng, Z.; Mo, L.; Li, M. A Random Forest-based ensemble method for activity recognition. In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 5074–5077, doi:10.1109/EMBC.2015.7319532.
2. Fan, L.; Wang, Z.; Wang, H. Human activity recognition model based on decision tree. 2013 International Conference on Advanced Cloud and Big Data, Nanjing, China, 13–15 December 2014; pp. 64–68, doi:10.1109/CBD.2013.19.
3. Bovi, G.; Rabuffetti, M.; Mazzoleni, P.; Ferrarin, M. A multiple-task gait analysis approach: Kinematic, kinetic and EMG reference data for healthy young and adult subjects. *Gait Posture* **2011**, doi:10.1016/j.gaitpost.2010.08.009.
4. Schwartz, M.H.; Rozumalski, A.; Trost, J.P. The effect of walking speed on the gait of typically developing children. *J. Biomech.* **2008**, *41*, 1639–1650, doi:10.1016/j.jbiomech.2008.03.015.
5. Gates, D.H.; Lelas, J.; Croce, U.D.; Herr, H.; Bonato, P. Characterization of ankle function during stair ambulation. In Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, 26–30 August 2004; pp. 4248–4251, doi:10.1109/IEMBS.2004.1404184.
6. Scaglioni-Solano, P.; Aragón-Vargas, L.F. Age-related differences when walking downhill on different sloped terrains. *Gait Posture* **2015**, *41*, 153–158, doi:10.1016/j.gaitpost.2014.09.022.
7. Andrew, M.; Beatty, K.T.; Purcell, L.; Vickers, D. Gait dynamics on an inclined walkway. *J. Biomech.* **2006**, *39*, 2491–2502, doi:10.1016/j.jbiomech.2005.07.025.
8. Lee, S.J.; Hidler, J. Biomechanics of overground vs. treadmill walking in healthy individuals. *J. Appl. Physiol.* **2008**, doi:10.1152/jappphysiol.01380.2006.
9. Riener, R. The Cybathlon promotes the development of assistive technology for people with physical disabilities. *J. Neuro Eng. Rehabil.* **2016**, *13*, doi:10.1186/s12984-016-0157-2.
10. Kadaba, M.P.; Ramakrishnan, H.K.; Wootten, M.E. Measurement of Lower Extremity Kinematics During Level Walking. *J. Orthop. Res.* **1990**, doi:10.1002/jor.1100080310.
11. Davis, R.B.; Öunpuu, S.; Tyburski, D.; Gage, J.R. A gait analysis data collection and reduction technique. *Hum. Mov. Sci.* **1991**, *10*, 575–587, doi:10.1016/0167-9457(91)90046-Z.
12. Segal A.D.; Orendurff, M.S.; Klute G.K.; McDowell, M.L.; Pecoraro, J.A.; Shofer, J.; Czerniecki, J.M. Kinematic and kinetic comparisons of transfemoral amputee gait using C-Leg® and Mauch SNS® prosthetic knees. *JRRD* **2006**, doi:10.1682/jrrd.2005.09.0147.
13. Sup, F.; Varol H.A.; Goldfarb, M. Upslope walking with a powered knee and ankle prosthesis: Initial results with an amputee subject. *IEEE Trans. Neural. Syst. Rehabil. Eng.* **2011**, *19*, 71–78, doi:10.1109/TNSRE.2010.2087360.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).