One IMU Is Sufficient: A Study Evaluating Effects of Dual-Tasks on Gait in Elderly People

Rolf Adelsberger[1], Nathan Theill[2], Vera Schumacher[2], Bert Arnrich[1], and Gerhard Tröster[1]

[1] Federal Institute of Technology Zurich, ETHZ, Switzerland
[2] University of Zurich, Zurich, Switzerland
{rolf.adelsberger,bert.arnrich,gerhard.troester}@ife.ee.ethz.ch, n.theill@inapic.uzh.ch,v.schumacher@psychologie.uzh.ch

Abstract. In industrialized countries the share of elderly subjects is increasing. Hence, diseases or symptoms associated with aging are more common than they were in the past. As a consequence, more effort is invested into research analyzing the effects of aging on the motion and cognition. However, economical and flexible methods to measure motion and its cross-effects with cognition are still missing. Therefore, we developed a new approach which neither requires a specific location, large infrastructural requirements, nor does it require large investments. We base our setting on match-box sized inertial measurement units (IMUs) attached to the participants' legs. 47 elderly subjects participated in our study where we analyzed the interplay between cognitive load and gait features. We show that it is feasible to automatically detect episodes of interest, e.g. straight path, during walking periods of a subject only using IMU data. Our approach detects the steps autonomously and calculates gait features without supervision. The results demonstrate that cognitive load induces a significant increase ($p = 0.007$) in step-duration variability from $16ms$ (baseline) to $21ms$ (load). Our findings demonstrate that IMUs are a proved alternative to static setups that usually require a non-trivial infrastructure, e.g. optical movement tracking.

Keywords: wearable computing, gait analysis, elderly people, risk of falling, imu, sensors.

1 Introduction

Increasing age might affect people in motoric skills as well as in cognitive performance. In general, the ability to sit, stand, walk and to perform activities of daily living (ADL) can be condensed in the term *mobility*. Mobility contributes the lion's share to an elderly persons' independence and as such is a combination of mental resources and their physical expressions. Limited motoric or mental capabilities result in a lowered mobility and with reduced mobility the risk of falling (RoF) increases [1]. If we can objectively measure mobility of a person,

B. Godara and K.S. Nikita (Eds.): MobiHealth 2012, LNICST 61, pp. 51–60, 2013.

there might be a model to predict her RoF. This is our main motivation: To estimate automatically the mobility of elderly people with future applications for safety in mind (e.g. reducing RoF).

In this paper we focus on gait features as they are by nature closely linked to RoF. It is known that gait features are affected by cognitive load levels of a subject. Especially for elderly people the *threshold level* where gait-feature-changes are noticeable is low [2] (compared to younger individuals' levels), sometimes as low as a task of subtracting numbers.

We demonstrate that a sensor-based automatic acquisition and analysis setup is a efficient alternative to the currently used methods. To this purpose we are looking at the step duration and its dynamics in situations with and without cognitive load. We present the analysis and results of a study with elderly people (aged 65+) and analyze the changes of gait features between a baseline setting (i.e. common walk) and a setting where the subjects were under elevated cognitive load while walking. We compare state-of-the-art (SoA) to our approach and show that we can detect the differences between situations with elevated cognitive load and situations without.

Furthermore, our longer-term goal is it to contribute to a transparent estimation system - not requiring any special action by the subjects - to make statements about a human's relative mental load level and the consequences for her motoric performance. We position our work as an initial contribution to that goal.

2 Related Work

2.1 Tests Not Using Electronic Devices

In general, in geriatrics the term mobility refers to a person's aptitude of performing a physical task in her everyday life. The definition of mobility is usually tailored to a specific target group, e.g. hospitalized patients or subjects at home and to a specific environment, e.g. medical care facility, home etc. [3, 4].

One of the most often used mobility indices (MI) is the timed-up-and-go test (TUG) first introduced by Podsialdo et al. [5] which is analyzed in more detail by Thrane et al. in [6]. TUG is often used as an indicator for RoF of a person. It measures the time a person needs to rise from a chair and walk a given distance. The *Short Physical Performance Battery* (SPPB) [7] focuses on the lower extremities and their functionality. SPPB can be divided into three sections: Balance Tests, Gait Speed Test and Chair Stand Tests (similar to TUG). The *Motor Assessment Scale* (MAS) [8] analyzes 8 motor functions. In particular, it also assesses transition movements (standing up), static tasks (standing still) and dynamic tasks (walking).

2.2 Tests with Electronic Devices

Webster et al. [9] introduced the GAITRite sensor system used for the evaluation of walking performance[1]. This sensor system consists of a pressure sensitive mat

[1] GAITRite Gold, CIR Systems, Easton, PA.

in various sizes. The largest model is about $1m$ wide and $7.5m$ long, allowing for the analysis of step length, step width and frequency. Webster et al. compared the system's performance to a state-of-the-art optical 3D motion tracking system, e.g. VICON. Van Iersel et al. [10] investigated the effect of cognitive dual tasking on balance of older adults. They used the GAITRite system for data acquisition and extracted spatial features of gait (e.g. stride length) and temporal gait features (time variability). Hollman et al. [11] performed a study incorporating older and younger subjects. In that study they analyzed the differences of dual-task walking between the two age groups. Kuys et al. [12] used a system to evaluate spatio-temporal gait features of stroke patients: the researchers used the data from the system to compute the MAS gait score of the patients. In [13] Bamberg et al. present a sensor system that provides three pressure measuring points as well as orientation data of the feet using intertial measurement units (IMU). All system components were integrated in a shoe. The authors used that system to analyze heel-strike and toe-off events during gait periods as well as the feet orientation. Within a sport focused setting Strohrmann et al. [14] used IMUs attached to the legs to analyze the running behavior of healthy younger people.

2.3 Evaluating Cognitive Features

Theill et al. [15] suggest that performing simple mathematical calculations[2] suffices to generate sufficient cognitive load to induce a measurable physical response of a subject. In their paper they present a precise method for measuring situations with cognitive loads of varying degree. Schaefer et al. [16] demonstrate that elderly subjects, when put under cognitive load, express an increased variance in step frequency and might even show difficulties maintaining balance. They used a variant of the N-Back test [17] to induce elevated cognitive load levels in their subjects. Schaefer et al. contributed to the motivation of evaluation training impact on cognitive performance and motoric fitness.

Cinaz et al. developed in [18] a system to estimate mental workload using the heart rate variability. They were able to train a classifier separating the instances of low mental workload from samples with higher mental workload. Cinaz et al. showed that the links between cognitive load and physical expression are abundant and feasible to measure.

3 Experiment

3.1 Hypotheses and Approach

We aimed at demonstrating the feasibility of using sensor data from IMUs to automatically detect periods of regular walk. During those intervals distinguished between situations without and situations with cognitive load. For this purpose we have setup a study where elderly people were asked to perform a simple

[2] e.g. starting from 50 subtract consecutively 2.

walking task once with, and once without a cognitive task in parallel. In the following sections we are going to provide the details.

3.2 Participants

Elderly people at the age 65 or older were recruited for the study. The inclusion criteria for participants was an age within the range $[\geq 65, 85 \leq]$. The applicants for participation had to pass a cognitive screening test [19] in order to be included in the study. To assess their overall motor activity we asked them to perform TUG. We were interested in healthy subjects with no evident disabilities. Subjects included in our study tested normal in the cognitive test and in the motoric evaluation, TUG. Out of 63 participants 47 individuals (32 female and 16 male) successfully completed the study and we could use their data for our evaluation[3]. The subjects' demographics are listed in Table 1 .

Table 1. Age distribution of the 47 subjects

	Min	**Max**	**Mean**	**Std.Dev.**
overall	65	84	71.77	4.89
female	65	81	71.71	4.70
male	65	84	71.88	5.33

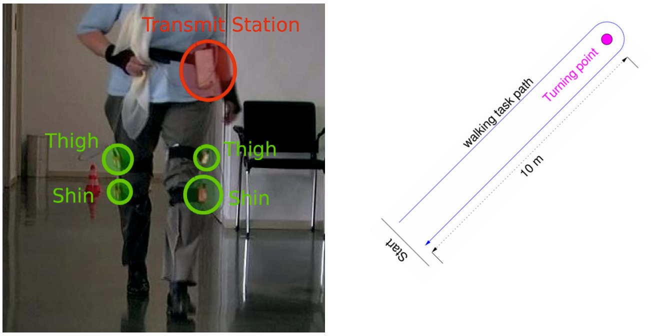
3.3 Measurement Setup

For the testings we equipped the subjects with sensor devices. In order to track gait features of our subjects we used four IMUs by XSens [20]. With Velcro straps we attached on each shin and each thigh a sensor having the x-axis pointing towards ground (cf. Figure 1a). Four sensors at these locations allowed us to track more features than just step durations: angles between shin and thigh, leg orientation etc. The devices were tethered; the data was sent to and power comes from a gateway device (transmit station) that was worn by a belt around the subjects' waist. We configured the devices to report raw acceleration, rotation rate values, but also Euler angles which reflect the orientation in space relative to the earth.

At 50Hz motion data was streamed via Bluetooth to a standard notebook where we stored it for later analysis. Higher sampling was not possible due to bandwidth limitation of the Bluetooth system.

The subjects were additionally recorded on video for a validity check of our automatic feature calculations. Data analysis was performed offline using MATLAB®.

[3] For the first 15 subjects our measurement setup suffered a technical problem and the recordings failed. One subject did not perform a testing at all. Our set therefore contains data from 47 subjects.



(a) Subject with four sensors and one gateway.

(b) The path for tasks 1 and 2.

Fig. 1. Test procedure illustrations

3.4 Test Procedure

At the beginning of the test session the mental test [19] (Mini Mental State Evaluation, MMSE) was presented to the subjects in direct interaction with an expert. This study controller (a psychologist) asked the questions and noted the answers of the subjects on the evaluation sheets. We required the subjects to score above 26 to be included in our study.

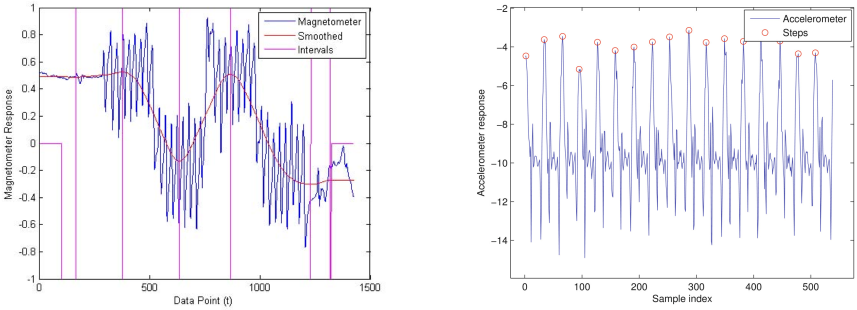
Motoric testing was performed with an instance of TUG: subjects were asked to sit on a regular chair. The controller then asked the subjects to stand up. Time until completion was measured starting from the issue of the command until the subject was in an upright position. TUG scores above 10 seconds are considered as noticeable [7].

During testings, subjects were two times required to walk down a aisle of length 10m, turn around and walk back again (baseline testing, task 1). In the second part of the testing we asked them to do the walk as before but now while subtracting from a random number provided by us (arbitrarily chosen from the set $[501, 502, 503]$) at each step a specific number, i.e. 7 (task 2).

4 Methods

We primarily wanted to compare statistics of the data from the baseline task to data from the cognitive-load task. To this purpose we firstly needed to detect the intervals in our data that were of interest, e.g. periods of straight walking, segmented by turning points. These intervals were detectable using the magnetometer data. In Figure 2a an axis of a magnetometer is plotted in blue. In the high frequency spectrum the steps are visible, the four changes of the mean value correlate with the walking direction of the subject. We calculated the turning points by first applying a smoothing, e.g. low-pass, filter to the data (red curve). Then, we used the sign of the slope (magenta) of this curve as the limits for the intervals. Finally, we declared the turning point as the mid-point of a decreasing interval. To maintain comparability to related work we analyzed the walking interval for the first $20m$, $(2 \times 10m)$. Therefore, only data until the second turning point (start position) was used. The turning points were detected by our algorithm without false positives. At each turning point we disregarded two steps before and one step afterward since we were only interested in statistics from straight walk.

Next we detected the steps using SoA [14] on the accelerometer signals. The steps manifest themselves as peaks in the accelerometer signal. In Figure 2b a fragment of a data set is shown: the accelerometer signal is in blue, the step locations are marked with red circles. Our focus lied on the variances of step



(a) Intervall detection (x-axis)

(b) Step detection

Fig. 2. Analyzing gait data

duration for task 1 and task 2. Hence, the time delays between individual steps served as input for our further analysis: For each subject we calculated the step durations for the baseline task (task 1) and for the cognitively loaded task (task 2) in milliseconds. Next, we calculated for each task $t = \{1, 2\}$ for each subject i the mean μ_t^i, the median m_t^i and standard deviation σ_t^i (or variance, resp.) of the step durations. We denote the collection of all μ_k^j of all subjects for task k as the vector $\bar{\mu}_k$. The definitions for $\bar{\sigma}_k$ and $\bar{m}_k$ are analogous. Since we are interested in individual changes we calculated the difference of μ^i and σ^i between the two tasks for each subject i: $\tilde{\mu}^i$ and $\tilde{\sigma}^i$, resp. $\tilde{\mu}^i := \mu_2^i - \mu_1^i$, $\tilde{\sigma}^i := \sigma_2^i - \sigma_1^i$. So, $\tilde{\mu}^3$ is the positive or negative change of the mean value of the step durations for subject 3.

In our analysis we looked at the set of means for both tasks, $\bar{\mu}_1$, and $\bar{\mu}_2$, resp. We also considered the sets of standard deviation for both tasks, $\bar{\sigma}_1$ and $\bar{\sigma}_2$. Finally, we also analyzed the set of individual progresses, $\bar{\tilde{\mu}}$ and $\bar{\tilde{\sigma}}$.

In order to make a statement about the development of gait features between task 1 and task 2 we needed to compare the variances of step duration of the first task to those variances from the second task. A requirement for a valid comparison is the two sets originate from the same distribution, e.g. a Normal distribution. The distribution parameters for each of the sets might be different.

We used the Lilliefors test [21] based on the Kolmogorov-Smirnov test [22] to verify task-1 data and task-2 data are from the same distribution family. Lilliefors' test performs better for smaller sample sizes than the Kolmogorov-Smirnov test.

For visualization the QQ-Plot [23] allows for graphical comparison of two distributions: it draws the quantiles of two empirical distributions against each other.

Finally, the variances of the two data sets were analyzed with a person-independent analysis of variance (ANOVA[4]).

5 Results

The validation of equality of distribution between the two sets yielded a positive result: in Figure 3a we show that the baseline data set and the cognitive load set originate from the same distribution family. The blue points represent the quantiles of the distributions. The abscissa represents the distribution for task 1 the ordinate is for task 2. Indicated in red is the linear interpolation line for the two sets. As can be seen the two sets relate in a linear manner to each other.

The evaluation of the Lilliefors test [21] for either set accepted the null hypothesis of the data originating from a normally distributed population with a confidence level $\alpha = 0.05$.

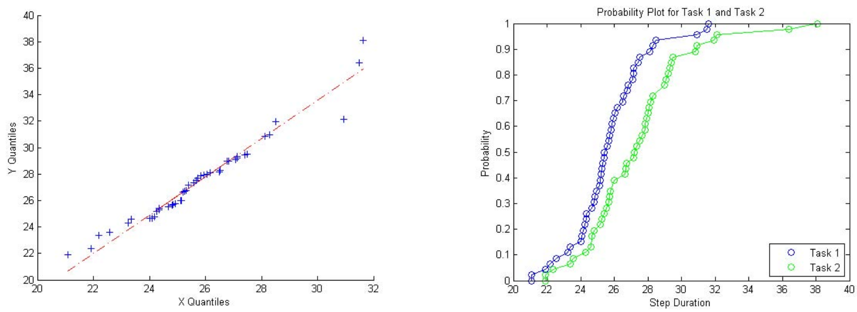
During evaluation we noticed that the sensor at the shin positions produced the best signal-to-noise (SNR) ratio. The shin sensors were less susceptible to *motion noise* that may be introduced by low-friction clothing (like synthetic trousers) and the sensor devices moving uncontrolled relatively to the leg or textile. We believe this result is caused by a looser attachment of the thigh sensors as a consequence of the thigh being by nature more sensitive to pressure than the shin. A tight Velcro was considered uncomfortable at that position. Too tight strappings might even have had an impact on the gait pattern. The shin, however, is not that sensitive and the muscular tissue does not perform large movements. Due to this reasons we decided to evaluate for each subject data solely from one (e.g. the left) shin sensor. A manual verification of the peak positions proved that the step detection algorithm worked with 100% accuracy.

Cognitive load had a significant impact on the gait features. This effect on the variance of step duration can be seen in Figure 3b where we draw probability plots for the two sets. In the plot data points (e.g. step durations) are plotted against their probability. The blue points mark the data points from the first task, the data of the second task is in green.

We performed a person-independent ANOVA test: In Table 2a we list the results of our analysis. In each line we report the mean value for the features $\bar{\mu}_k$, $\bar{\sigma}_k$ and the median, m_k resp., introduced in Section 4 . Additional to the mean values of the features we provide also the standard deviation. All values are in milliseconds. The F column are the F-numbers from ANOVA[5]. The $p-$values in

[4] Using a one-tailed significance level of p = .05.

[5] F-statistics from the ANOVA test: $F = \frac{\text{between-group variability}}{\text{within-group variability}}$.



(a) QQ-Plot of mean values for baseline task and cognitive load task

(b) Cumulative Probability plot for baseline task and cognitive load label

Fig. 3. Comparing tasks

Table 2. Results of analysis

(a) ANOVA

Feature	**Task 1** (ms)	**Task 2** (ms)	**F**	**p**
$\bar{\sigma}_k$	16.84 ± 5.17	21.64 ± 10.62	6.62	0.007
$\bar{\mathrm{m}}_k$	513.04 ± 45.41	547.39 ± 63.82	8.85	0.0038
$\bar{\mu}_k$	514.93 ± 44.24	549.01 ± 62.88	9.04	0.0034

(b) Mean changes.

Feature	**Value** (ms)
$\bar{\bar{\sigma}}$	4.8065
$\bar{\bar{\mu}}$	34.0807

the last column in Table 2a indicate that all three features differ significantly between the two testings.

Table 2b depicts the mean changes on gait features induced by the cognitive task. The standard deviation between the baseline task (task 1) and the cognitive load task has increased by $4.8ms$ (mean). The mean of the mean values increased by $34.1ms$. The variability of the step duration changes from baseline to cognitive loaded situations and the mean step duration increases. Our findings are comparable to previous findings of related work [10, 11].

6 Conclusion and Future Work

In this paper we have taken the first step towards an autonomous mobility assessment system by automatically analyzing the correlations of gait features and cognitive load with an IMU setup. We have shown that by using step duration it is possible to distinguish between situations of cognitive load and those without. We have also provided a proof-of-concept for the feasibility of performing the analysis automatically. In our paper we successfully demonstrated that with a minimal setup of one single inertial measurement sensor it is feasible to conduct studies equivalent to SoA, but requiring substantially less infrastructure, e.g. no cameras, no human resources etc. at arbitrary locations.

We are going to base our future work on the results and findings of this study. We further believe that for future gait analysis it is possible to reduce the hardware requirements even more. We envision a single-sensor setup - untethered - providing us with gait features like the ones used here but also with additional ones. We believe that there are training effects over longer periods of time: redoing task 2 several times over a longer time span might reduce the effect of the cognitive load. We want to measure this progression in the future. Also, generalizing our setup even more in order to allow for many more movement features is planned for the future.

We envision a automatic system to assess mobility: an unobtrusive self-contained sensor system that constantly monitors the movement of its wearer. This paper represents a part of the whole, but for the future we want to add additional modalities and more importantly at some point leave the lab setting and go into real life environments.

References

1. Scott, V., Votova, K., Scanlan, A., Close, J.: Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. Age and Ageing 36(2), 130–139 (2007)
2. Hausdorff, J.M., Schweiger, A., Herman, T., Yogev-Seligmann, G., Giladi, N.: Dual-task decrements in gait: contributing factors among healthy older adults. J. Gerontol. A Biol. Sci. Med. Sci. 63(12), 1335–1343 (2008)
3. de Morton, N.A., Davidson, M., Keating, J.L.: Reliability of the de morton mobility index (demmi) in an older acute medical population. Physiotherapy Research International: the Journal for Researchers and Clinicians in Physical Therapy (October 2010)
4. Kuys, S.S., Brauer, S.G.: Validation and reliability of the modified elderly mobility scale. Australasian Journal on Ageing 25(3), 140–144 (2006)
5. Richardson, S., Podsiadlo, D.: The timed 'up and go': A test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society 39(2), 142–148 (1991)
6. Thrane, G., Joakimsen, R.M., Thornquist, E.: The association between timed up and go test and history of falls: the tromso study. BMC Geriatr. 7, 1+ (2007)
7. Guralnik, J.M., Simonsick, E.M., Ferrucci, L., Glynn, R.J., Berkman, L.F., Blazer, D.G., Scherr, P.A., Wallace, R.B.: A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. Journal of Gerontology 49(2), M85–M94 (1994)
8. Carr, J.H., Shepherd, R.B., Nordholm, L., Lynne, D.: Investigation of a new motor assessment scale for stroke patients. Physical Therapy 65(2), 175–180 (1985)
9. Webster, K.E., Wittwer, J.E., Feller, J.A.: Validity of the gaitrite walkway system for the measurement of averaged and individual step parameters of gait. Gait & Posture 22(4), 317–321 (2005)
10. van Iersel, M.B., Ribbers, H., Munneke, M., Borm, G.F., Rikkert, M.G.O.: The effect of cognitive dual tasks on balance during walking in physically fit elderly people. Archives of Physical Medicine and Rehabilitation 88(2), 187–191 (2007)

11. Hollman, J.H., Kovash, F.M., Kubik, J.J., Linbo, R.A.: Age-related differences in spatiotemporal markers of gait stability during dual task walking. Gait & Posture 26(1), 113–119 (2007)
12. Kuys, S.S., Brauer, S.G., Ada, L.: Test-retest reliability of the GAITRite system in people with stroke undergoing rehabilitation.. Disabil. Rehabil. 33(19-20), 1848–1853 (2011)
13. Bamberg, S., Benbasat, A., Scarborough, D., Krebs, D., Paradiso, J.: Gait analysis using a shoe-integrated wireless sensor system. IEEE Transactions on Information Technology in Biomedicine 12, 413–423 (2008)
14. Strohrmann, C., Harms, H., Tröster, G., Hensler, S., Müller, R.: Out of the lab and into the woods: kinematic analysis in running using wearable sensors. In: Ubicomp, pp. 119–122 (2011)
15. Theill, N., Martin, M., Schumacher, V., Bridenbaugh, S.A., Kressig, R.W.: Simultaneously measuring gait and cognitive performance in cognitively healthy vs. cognitively impaired older adults: The basel motor-cognition dual task paradigm. Journal of the American Geriatrics Society 59, 1012–1018 (2011)
16. Schaefer, S., Schumacher, V.: The interplay between cognitive and motor functioning in healthy older adults: Findings from dual-task studies and suggestions for intervention. In: Gerontology, pp. 1–8 (2010)
17. Kirchner, W.K.: Age differences in short-term retention of rapidly changing information. Journal of Experimental Psychology 55(4), 352–358 (1958)
18. Cinaz, B., La Marca, R., Arnrich, B., Tröster, G.: Monitoring of mental workload levels. In: Proceedings of IADIS eHealth Conference, pp. 189–193 (2010)
19. Folstein, M.F., Folstein, S.E., McHugh, P.R.: “Mini-mental state”. a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 12, 189–198 (1975)
20. XSens, `http://www.xsens.com/`
21. Lilliefors, H.W.: On the kolmogorov-smirnov test for normality with mean and variance unknown. Journal of the American Statistical Association 62, 399–402 (1967)
22. Kolmogorov, A.N.: Sulla determinazione empirica di una legge di distribuzione. Giornale dell’Istituto Italiano degli Attuari 4, 83–91 (1933)
23. Wilk, M.B., Gnanadesikan, R.: Probability plotting methods for the analysis of data. Biometrika 55, 1–17 (1968)