In *IEEE Engineering in Medicine and Biology Conference,* 2010.

Model-Based Inference of Cognitive Processes from Unobtrusive Gait Velocity Measurements

Daniel. Austin *Student Member, IEEE*, Todd Leen, Tamara L. Hayes *Member, IEEE*, Jeff Kaye *Member, IEEE* , Holly Jimison *Member, IEEE*, Misha Pavel, *Member, IEEE*

***Abstract*—In this paper we describe a preliminary modeling and analysis of a unique data set comprising unobtrusive and continuous measurements of gait velocity in the elder participants' residences. The data have been collected as a part of a longitudinal study aimed at early detection of cognitive decline. We motivate these analyses by first presenting evidence that suggests significant relationship between gait parameters and cognitive functions. We then describe a simple, model-based approach to the analysis of gait velocity using a weighted correlation function estimates. One of the main challenges is due to the fact that the daily estimates of the gait parameters vary with the number of walks. We illustrate the importance of using weighted as opposed to unweighted estimates on a sample of different houses. The correlation functions appear to capture behavioral differences that can be related to the cognitive functioning of the participants.**

I. Introduction

EARLY detection of subtle changes in physical or cognitive functions of individual elders is an essential component of future systems supporting elders living independently in their own homes. Detection of changes in physical and mental health states are also key enablers of effective intervention, mediation, and maintenance. For example, monitoring elders' activities can be used effectively and efficiently in conjunction with coaching systems [1, 2].

Recent results in aging and neurodegenerative diseases suggest that aspects of gait in general, and gait velocity in particular, are sensitive measures of mobility as well as cognitive state of the individual. There are several clinical studies that suggest a statistically significant relationship between gait velocity and cognitive state of elder patients [3, 4]. For example, in one longitudinal study, elders walking slower would be diagnosed sooner with mild cognitive impairment (MCI) than those walking faster [3].

A more direct relationship between cognitive function and gait velocity has been observed in terms of mutual interference between walking and a number of cognitive activities. This relationship has been investigated in the framework of dual tasks [5]. In a dual task experiment, participants are asked to walk while performing a concurrent cognitive task, for example counting backward from 100 by 3. Although the exact results vary across studies and subject populations, there is clearly identifiable interference. When a human participant performs a concurrent cognitive task, their gait velocity is reduced.

To take advantage of both the long-term and short-term effects of cognitive function on gait, a number of studies, including several in the Oregon Center for Aging and Technology (ORCATECH) are beginning to assess gait velocity on a frequent (continuous) basis. For example, the largest ORCATECH monitoring study involves over 230 participants whose activities have been monitored continuously and unobtrusively in their homes for a couple of years. In this study, the participants' homes are equipped with a set of passive infrared motion sensors that enable unobtrusive, continuous monitoring of the participants' movements within their house. In addition, several of the motion sensors were installed in such a way that enables us to estimate the gait velocity from the sensor firing patterns [6]. This approach provides gait velocity measurement every time the occupant passes through the line of sensors, resulting in stochastic sequences comprising thousands of measurements during each year of observations.

The purpose of this paper is to describe a preliminary analysis of these gait velocity data collected by the ORCATECH over more than two years. The main challenge for this analysis is the complexity of the processes that underlie the variability of gait velocity and therefore the sophistication of the models used to characterize the data. Perhaps the main cause of the high degree of complexity is the fact that these data represent a mixture of diverse processes that influence the gait velocity measurements on a wide range of different time scales. On the coarsest time scale of months and years, the measurement reflect relatively slow and subtle changes associated with normal aging or the expected progression of chronic, neurodegenerative diseases. Shorter range processes include temporary changes due to episodes such as diseases, colds, and depression. At the

This study was funded by NIA grants P30AG024978, P30AG08017, and R01AG024059.

Daniel Austin (e-mail: austidan@bme.ogi.edu), Todd Leen (e-mail: tleen@bme.ogi.edu), and Tamara L. Hayes (e-mail: hayest@bme.ogi.edu) are with the Department of Biomedical Engineering, Oregon Health & Science University. 3303 SW Bond Avenue, Portland, OR 97239 USA

Holly Jimison (e-mail: jimisonh@ohsu.edu) is with the Department of Medical Informatics, Clinical Epidemiology, and the Oregon Center for Aging and Technology (ORCATECH), Oregon Health & Science University. 3181 SW Sam Jackson Park Rd, Portland, OR 97239 USA

Jeff Kaye (e-mail: kaye@ohsu.edu) is with the Department of Neurology and ORCATECH, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239 USA.

Misha Pavel (corresponding author) is with the BME Department and ORCATECH, Oregon Health & Science University. 3303 SW Bond Avenue, Portland, OR 97239 USA (e-mail: pavel@bme.ogi.edu; phone: 503-418-9302; fax: 503-418-9311).

finest grain of analysis the individual walk instances are affected by a variety of instantaneous environmental and situational factors such as the specific reason for any given walk, footwear, or the nature of potential concurrent tasks, to name a few examples.

The analysis in this paper is focused on the processes at the more coarse time scales ranging from weeks to years. To this end we assume that the very short term variability is a stochastic process that can be characterized as an ergodic, independent, identically distributed random process.

II. Data Analysis

In this section we describe a probabilistic model for how walking events are observed from an unobtrusive in-home sensor system. In addition, we construct a stochastic process of estimated average daily gait velocity useful for analysis of longitudinal behavior from the daily measurements. We discuss the issue of parameter estimation for this process, and why weighted parameter estimation is necessary with this type of data due to the data collection procedure. We illustrate this by showing how to estimate the weighted autocorrelation function of average daily gait velocity, which we compare to the un-weighted autocorrelation function in a later section.

A. A Probabilistic Model for and Description of the Data

In a prior paper [6] we discuss the estimation of gait velocity from a “sensor line” of passive infrared motion sensors placed in the dwellings of seniors living in homes and retirement communities in the Portland (OR, USA) metropolitan area. The data collected with this system is event driven in the sense that a resident must walk through a sensor line in order for a walking event to be detected and the gait velocity for this event to be estimated. As a result, events may occur and be collected anytime during the day. Additionally, the number of events collected in a given day is determined by the placement of the sensor line in the home and how many times the resident walks through that area, which likely changes from day to day.

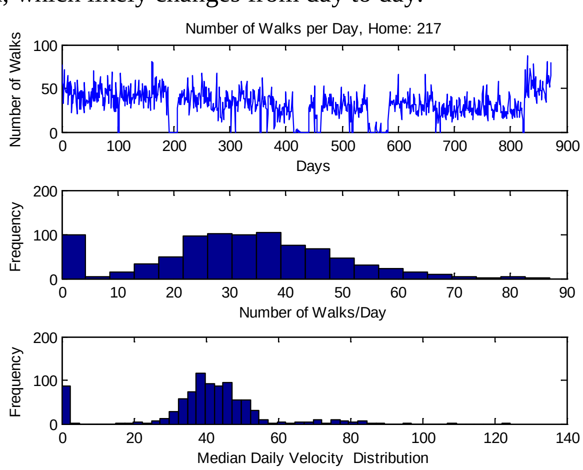

Figure 1. Number of walks per day and the corresponding distribution of daily averages of gait velocities for home ID 217.

In order to address these sources of variability, we propose the following model for the set of collected gait velocities for a given day, *d*:

$$\{v_i(d)\}_{i=1,\ldots,N(d)} = \{X(d+t_i)\}_{i=1,\ldots,N(d)} \qquad (1)$$

with $\{t_i\}_{i=1,\ldots,N(d)}$ a set of independent random time-of-day variables drawn from the same unknown distribution and $N(d)$ a random variable indicating the number of events collected for day *d*. So, for a given day we have a set of estimated gait velocities where both the number of events and the time of events will differ across different days.

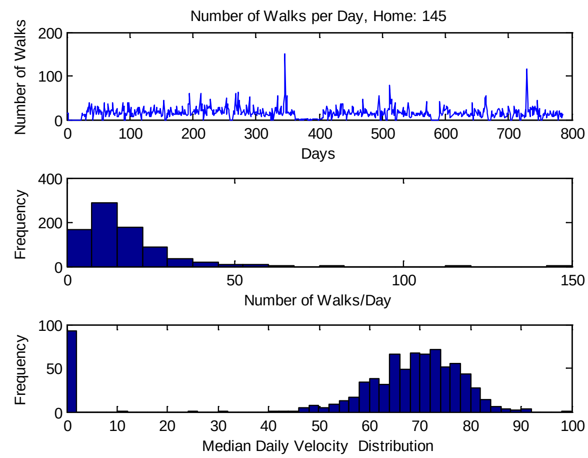

Figure 2. Number of walks per day and the corresponding distribution of daily averages of gait velocities for house ID 145

Prior to proceeding with the analysis of the proposed model, we examined the distribution of the velocity measurements under the assumption of ergodicity. Fig. 1 shows an example of the probability histogram of the measurements from two different houses. As in many cases, the distribution of the measurements is unimodal and can be easily characterized by its mean and variance.

B. Daily Averages and the Process Average

In order to facilitate longitudinal analysis of gait velocity (and to have a uniformly sampled time series), we calculate the average for each set of velocity estimates for a given day,

$$\hat{\mu}(d) = \frac{1}{N(d)} \sum_{i=1}^{N(d)} X(d+t_i) \qquad (2)$$

along with the estimation error:

$$\hat{\sigma}^2(d) = \frac{s^2(d)}{N(d)} \qquad (3)$$

where $s^2(d)$ is the sample variance for day *d*. We then form the random process $\{\hat{\mu}(d)\}_{d\in D}$ with D a collection of sequential days for which there are measurements.

We assume that this process is second order stationary (preliminary analyses of the data omitted here for brevity of presentation indicate that this is not an unreasonable assumption) and would like to estimate the process mean. Since the number of samples for each day is different, it is inappropriate to apply the standard estimator of the mean

In *IEEE Engineering in Medicine and Biology Conference*, 2010.

$\hat{\mu}_p = \frac{1}{|D|}\sum_{d\in D}\hat{\mu}(d)$ (with $|D|$ the cardinality of D) as this gives equal weight to each time sample of the random process. However, since we have an estimate of the variability of each estimated time sample, we can use the best linear unbiased estimator (BLUE) [7] of the mean as

$$\hat{\mu}_p = \frac{1}{\sum_{d\in D} 1/\hat{\sigma}^2(d)} \sum_{d\in D} \frac{\hat{\mu}(d)}{\hat{\sigma}^2(d)} \qquad (4)$$

where the subscript p indicates the process mean.

C. Process Autocorrelation

For the same reason the standard mean estimator is inappropriate for use with this type of data, the standard autocorrelation estimate (s_p^2 is the sample variance of the process):

$$\hat{R}(k) = \frac{1}{(|D|-k)s_p^2} \sum_{j=1}^{|D|-k} (\mu(j)-\mu_p)(\mu(j+k)-\mu_p) \qquad (5)$$

is also inappropriate. Again, however, we can find an appropriately weighted estimate

$$\hat{R}(k) = \left(\frac{1}{\sum_{j=1}^{|D|-k} \frac{1}{\hat{\sigma}^2(j)\hat{\sigma}^2(j+k)}} \right) \times \left(\sum_{j=1}^{|D|-k} \frac{(\mu(j)-\mu_p)(\mu(j+k)-\mu_p)}{\hat{\sigma}^2(j)\hat{\sigma}^2(j+k)} \right) \qquad (6)$$

by using the estimation error in (3) combined with the fact that the variance of a product of zero mean independent random variables is the product of the variances[8]. Note that this estimate is biased since we have not adjusted the weighting for the number of parameters that are being estimated, but it is asymptotically unbiased as $|D| \to \infty$.

III. Results

Figs. 1 and 2 show raw (non-averaged) gait velocity time series for a few sample months for two homes (top plots), and weighted (black line) and unweighted (gray line) estimated average daily gait velocity autocorrelation sequences (bottom plots) for lags of up to 20 days. For home 145, we used 687 samples (days) of average daily velocity to estimate the autocorrelation sequence whereas 737 samples (days) were used for home 217. As can be seen for both homes, the unweighted autocorrelation estimates tend to make the average daily gait velocities appear uncorrelated or weakly correlated. While the weighted autocorrelation function for home 145 is somewhat similar to the unweighted autocorrelation for this home, the weighted autocorrelation for home 217 shows a persistent, slowly decaying correlation that appears to have been averaged out of the unweighted estimate.

The difference in the course of temporal persistence observed in the weighted correlation functions in Figs. 3 and 4 may provide an interesting insight into the detailed aspects of the participants' behaviors. The correlation function in Fig. 3 represents mostly short-term processes without much predictive power from day to day – perhaps with the exception of a subtle periodic effect with a period of one week. In contrast, the correlation in Fig. 4 which decays much slower suggests the existence of long term interactions implying that there are aspects of the behaviors that are predictable over long periods of time.

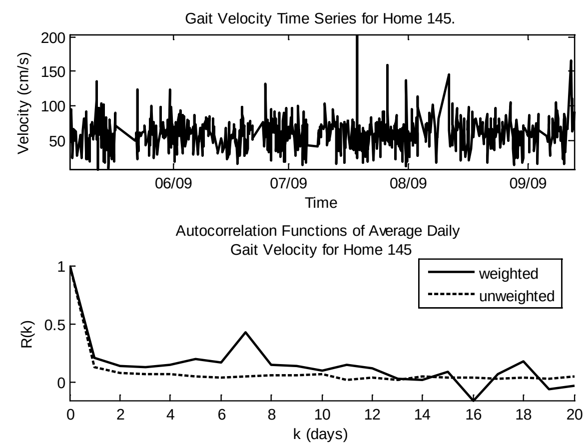

Figure 3: Gait velocity time series (top plot) and weighted (bottom plot, solid line) and unweighted (bottom plot, dashed line) autocorrelation for the resident in home ID 145.

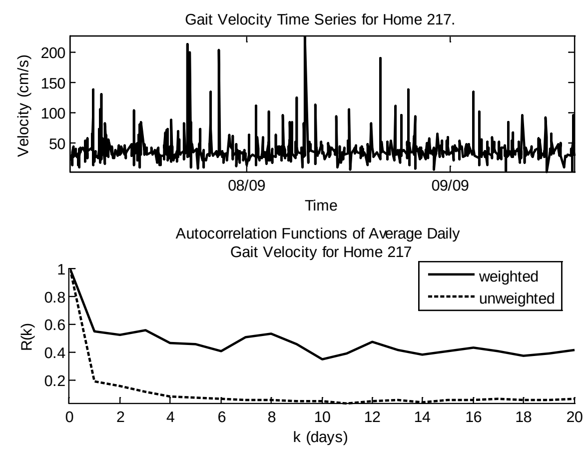

Figure 4: Gait velocity time series (top plot) and weighted (bottom plot, solid line) and unweighted (bottom plot, dashed line) autocorrelation for the resident in home ID 217.

Another way to examine more closely the patterns that lead to the different correlation functions is to estimate the power spectral density of the respective stochastic processes. The resulting power spectrum estimates are shown in Fig. 5 and Fig. 6, where the calculation was made as the Fast Fourier Transform of the weighted autocorrelation sequence for 64 lags. In a similar manner to the correlation functions, the power spectral density representations suggest two different behavioral models for these two participants.

In the power spectral density representing home number 145 we see power spread across multiple frequencies including significant power at the frequencies which represent weekly and monthly periodicities, and a peak at what is approximately a 3 month cycle, which may indicate strong seasonal patterns. This means that this participant is likely to behave differently from day to day due to multiple periodicities in the data. In contrast, the power spectrum associated with the participant in home number 217 has significantly the majority of the power concentrated in the lowest frequencies suggesting behavioral patterns with long periodicities and little change from day to day.

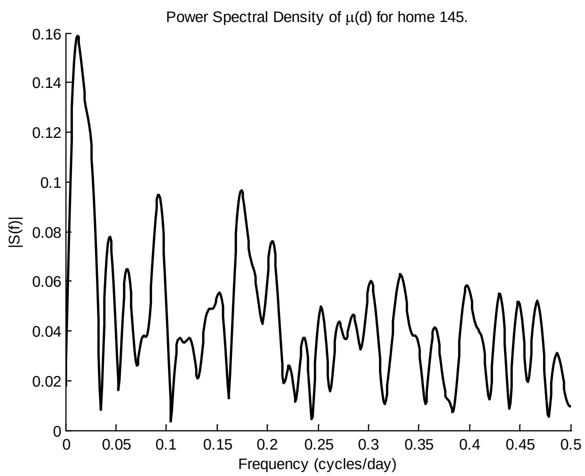

Figure 5: Power spectral density of the sequence of daily velocity estimates for the home ID 145.

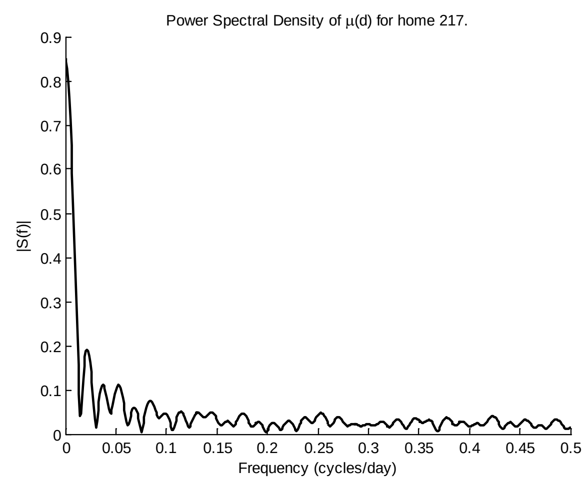

Figure 6 Power spectral density of the sequence of daily velocity estimates for the home ID 217.

These differences in the underlying behavioral models motivated an examination of the neuropsychological assessments obtained as a parallel, but independent part of the study for these two participants. The resulting scores indicate that the participant in home 145 has neuropsychological scores within normal range, namely the average value of 28 on the Mini-Mental state examination with maximum score of 30. In contrast, the participant in home 217 received an average score of 24 and a further clinical dementia rating (CDR) of 0.5 corresponding to the designation of mild cognitive impairment. Thus the unobtrusive measurements appear to be consistent with the neuropsychological test results.

IV. Conclusion

The work described in this paper represents a preliminary model-based analysis of the vast amounts of data collected in longitudinal studies by ORCATECH. The results are promising in the ability to relate cognitive functions to the observed data. Even very simple analysis suggests that the unobtrusive assessment of the gait velocity may provide important information about the participants' activities as well as about their cognitive functionality.

Future work will focus on a more in-depth analysis of the sequences and in particular the deviations from ergodicity. One question that might be worth pursuing is to what extent are behaviors and their reflection in the measurements of gait velocity consistent with fractal processes. This issue is closely related to the multiple scales of the temporal processes underlying behaviors.

In addition, future work will include algorithmic improvements such as calculating confidence intervals to assess the estimation performance of the weighted autocorrelation estimates and identifying possible cognitive or physical pathologies associated with the amount of velocity correlation present across short or long time scales. In this paper we used only the most superficial features of the observed data for the preliminary analyses.

V. Acknowledgements

The authors would like to thank Nicole Larimer for her assistance in recruiting subjects for this study, and Nora Mattek for assistance with the neuropsychological data.

VI. References

[1] H. B. Jimison and M. Pavel, "Integrating Computer-Based Health Coaching into Elder Home Care," in *Technology and Aging*, A. Mihailidis, J. Boger, H. Kautz, and L. Normie, Eds. Amsterdam, The Netherlands: IOS Press, 2008.

[2] H. B. Jimison, M. Pavel, N. Larimer, and P. Mullen, "A General Architecture for Computer Based Health Coaching," presented at International Conference on Technology & Aging, Toronto, 2007.

[3] R. Camicioli, D. Howieson, B. Oken, G. Sexton, and J. Kaye, "Motor slowing precedes cognitive impairment in the oldest old," *Neurology*, vol. 50, pp. 1496-8, 1998.

[4] J. Verghese, R. B. Lipton, C. B. Hall, G. Kuslansky, M. J. Katz, and H. Buschke, "Abnormality of gait as a predictor of non-Alzheimer's dementia," *N Engl J Med*, vol. 347, pp. 1761-8, 2002.

[5] M. Woollacott and A. Shumway-Cook, "Attention and the control of posture and gait: a review of an emerging area of research," *Gait Posture*, vol. 16, pp. 1-14, 2002.

[6] S. Hagler, D. Austin, T. Hayes, J. Kaye, and M. Pavel, "Unobtrusive and Ubiquitous In-Home Monitoring: A Methodology for Continuous Assessment of Gait Velocity in Elders," *IEEE Trans Biomed Eng*, 2009.

[7] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Control*, 2 ed: Prentice Hall, 1995.

[8] D. Blumenfeld, *Operations Research Calculations Handbook*, 2 ed: CRC Press, 2010.