Mapping Kinect-based In-Home Gait Speed to TUG Time: A Methodology to Facilitate Clinical Interpretation

Erik E. Stone and Marjorie Skubic
Center for Eldercare and Rehabilitation Technology
Department of Electrical and Computer Engineering
University of Missouri
Columbia, MO, USA
ees6c6@mizzou.edu, skubicm@missouri.edu

***Abstract*— A methodology for mapping in-home gait speed (IGS), measured unobtrusively and continuously in the homes of older adults, to Timed-Up-and-Go (TUG) time is presented. A Kinect-based gait system was used to collect in-home gait data on 15 older adults over time periods of up to 16 months. Concurrently, the participants completed a monthly clinician administered fall risk assessment protocol that included TUG and habitual gait speed (HGS) tests. A theoretical analysis of expected performance is presented, and the performance of the IGS-based TUG estimates is compared against that of estimates based on HGS measured at the same time as the TUG. Results indicate that the IGS-based estimates are as accurate as the HGS-based estimates as compared to the observed TUG times. After filtering the TUG times to reduce noise, the IGS-based estimates are more accurate. The mapping of in-home sensor data to well studied domains facilitates clinical interpretation of the in-home data.**

Keywords – gait; Kinect; fall risk; TUG; Timed-Up-and-Go;

I. INTRODUCTION

RESEARCH has shown that the parameters which describe locomotion are indispensible in the diagnosis of frailty and fall risk [1]. Additionally, studies have indicated that changes in gait parameters may be predictive of future falls and adverse events in older adults [2-6], and that scores on certain mobility tests are good indicators of fall risk [7-9]. Studies have also shown that interventions to prevent falls among seniors, such as household modifications and exercise routines, could significantly reduce falls and be highly cost effective [10]. Despite this, these gait parameters and mobility tests are generally assessed infrequently, if at all, through observation by a clinician with a stop watch or using expensive equipment in a physical performance lab. Furthermore, these sparse, infrequent evaluations may not be representative of a person's true functional ability [11].

In [12-14], an unobtrusive, continuous gait monitoring system based on the Microsoft Kinect sensor was developed and evaluated that could measure the gait of older adults, in their homes, during normal daily activity. However, the output of the system, measures of in-home gait speed, stride time, stride length, etc., is not easily interpretable, as such parameters have never before been available. Thus, either a large scale study to directly relate these in-home gait parameters to fall risk and/or health status is needed, or a methodology to relate the parameters to existing well studied and understood domains needs to be developed.

For this work, the Kinect-based in-home gait system was deployed in the homes of 15 older adults for time periods of up to 16 months. While the systems were installed, the participants also completed monthly fall risk assessment protocols consisting of standard mobility tests, such as the Timed-Up-and-Go (TUG) [15], Habitual Gait Speed (HGS) [16], and Short Physical Performance Battery (SPPB) [17]. The TUG test has been widely studied and shown to be a good measure of functional ability as well as an indicator of fall risk in the elderly [8-9, 15]. As such, mapping of the in-home gait data to this well understood domain would facilitate interpretation of the data by a clinician.

This paper presents a methodology for and results from estimating TUG time from in-home gait data, specifically walking speed, collected by the Kinect-based systems. The purpose is not to measure TUG time directly, but to map the in-home gait data to a domain that clinicians understand; and to assess the accuracy of the mapping with regards to what could theoretically be expected. Although this paper focuses on mapping in-home gait speed to TUG time, the approach could be used with any data source, and any well studied domain.

Section II of this paper discusses related work in the area of in-home gait and functional ability assessment. Section III contains a brief overview of the Kinect-based in-home gait system. Section IV describes the methodology used to map in-home walking speed to TUG time, including a theoretical analysis of expected performance for a system that was modeling the true TUG of the individuals. Section V contains the results of estimating TUG time from in-home gait speed and compares it to estimates made using HGS measured by a clinician at the same time as the TUG. Finally, a discussion of the results and their implications is given in Section VI.

II. RELATED WORK

A number of researchers have looked at assessing mobility and/or measuring TUG using wearable sensors [8, 18-21]. In [18], two sensor units, each with an accelerometer and a gyroscope, were used to assess individuals while they performed the TUG test and a methodology was developed to identity each phase of the TUG. Good agreement was found to measurements made by therapists. In [8], patients were assessed using two kinematic sensors while performing the TUG test and significant discrimination was provided by 29 out of 44 derived metrics between patients with and without a history of falls. Finally, in [19], researches used seven inertial sensors attached to the body to segment the TUG test and capture additional spatial and temporal metrics. These additional metrics were able to differentiate a group of patients in the early stages of Parkinson's disease, whereas TUG time alone could not.

In [21], researchers developed an approach for characterizing the mobility of elderly subjects in unsupervised environments using an accelerometer and found good correlation between extracted signals and a clinical fall risk gold standard. However, the approach requires monitoring participants during a specific, directed routine of physical tasks, not normal everyday activity. Furthermore, as with other wearable technologies, improper device usage and/or placement, along with the inability of users to safely perform the required tasks in unsupervised settings were identified as clear limitations for long term monitoring.

Although wearable systems are useful in supervised, clinical settings where there is a need to reliably, quantitatively assess a patient's mobility, they are not practical for long term, continuous monitoring in the home. The need for users to be actively involved, changing or charging batteries for example, often leads to poor compliance over time. Studies have shown older adults to prefer ambient, non-wearable sensors [22] for in-home monitoring.

In [23], researchers developed a technique to assess walking speed in home environments using an array of passive infrared (PIR) sensors mounted on the ceiling in a hallway or above a natural walking path. Using data from a one month period centered on a participant's first annual physical evaluation, the researchers found statistically significant associations between in-home walking speed and various mobility assessments [11], including the motor section of the Unified Parkinson's Disease Rating Scale, stopwatch timed gait speed, and the Tinetti balance scale. However, the PIR-based system does have limitations. For example, distinguishing between residents in multi-resident homes is problematic given the single available feature of walking speed, and finer grained information, such as stride time and stride length, may prove critical to a complete assessment of fall risk and functional mobility.

Finally, in [24], researchers developed a methodology for assessing mobility in unsupervised environments by segmenting traditional assessments into basic movement components, and measuring those basic components during normal everyday activity. Specifically, a laser range finder integrated into a chair was used to measure the basic components of the TUG test: standing, walking, turning, and sitting. In a small field trial, good agreement was shown between TUG time measured using their aTUG apparatus and TUG time measured using a stopwatch on a set of six TUG tests.

III. In-Home Gait System

In [12], the Microsoft Kinect sensor was evaluated for the purpose of passive, in-home gait measurement in a lab setting. This evaluation consisted of developing algorithms for measuring multiple gait parameters from the Kinect depth imagery and comparing them to measurements obtained from a Vicon marker-based motion capture system. The results of this study showed good agreement between measurements from the Kinect and those from the Vicon, along with good reliability of the measurements from the Kinect.

In [13, 14], Kinect-based systems using these algorithms

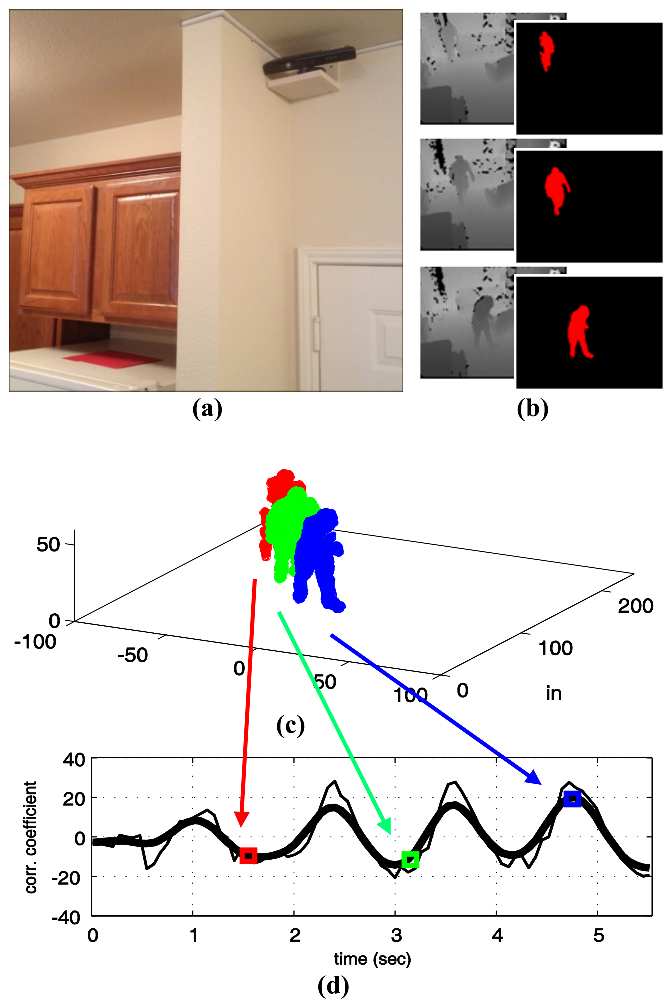

Fig. 1 (a) Kinect and computer (inside cabinet) as deployed in apartments. (b) Example depth images and extracted foreground during a walk in an apartment. (c) Three-dimensional model of person obtained at selected frames using extracted foreground. (d) Plot of correlation coefficient time series of normalized ground plane projections during walk (thin is raw, thick is filtered); used to identify when steps occur. Local maxima correspond to left steps, while local minima correspond to right steps. Algorithm details and parameter definitions can be found in [xx].

were deployed in the apartments of older adults in an independent living facility, and a methodology for modeling the gait of the residents and tracking changes over time was developed. A brief description of basic system operation and resident gait modeling follows.

A. System Operation

Fig. 1 shows the Kinect sensor as mounted in one of the apartments. The Kinect is placed on a small shelf a few inches below the ceiling (height 2.75 meters), above the front door. A computer is placed in a cabinet above the refrigerator. Foreground objects, represented as a set of 3D points, are identified from each frame using a dynamic background subtraction technique. Next, a tracking algorithm is used to track extracted 3D objects across multiple frames. Walks are identified from the path histories of the tracked objects. A set of criteria including path straightness, speed, duration, and distance are used to identify suitable walks. This is done online in real-time. Current minimum requirements for a suitable walk are a relatively straight path of at least 1.2 meters, with a continuous minimum speed of 12.7 centimeters per second.

Walking speed and height are extracted for every identified walk as they are computed using the centroid and maximum value in the vertical direction from each frame, respectively. However, due to issues such as occlusion of the legs and bad segmentation, stride parameters are only extracted for walks for which at least five steps could be identified which met three screening criteria used to eliminate invalid sequences. The reader is referred to [12-14] for a more detailed explanation of the algorithms.

B. Resident Modeling and Gait Assessment

The output of the Kinect-based gait systems is a dataset in which each entry corresponds to a walk that occurred in the apartment. As the systems are deployed in real-world environments, this dataset will include walks from all residents of the apartment, as well as any visitors. Thus, it is necessary to identifying which walks are from each specific resident before the gait of the resident can be assessed.

In [14], an approach based on a fitting a Gaussian Mixture Model (GMM), $\lambda = \{\rho_r, \mu_r, \Sigma_r\}$, $r = 1, \ldots,$ K, with the number of distributions, K, equal to the number of residents in the apartment to the dataset was developed. This approach is based on the assumption that each resident will create a cluster, or mode, in the dataset representing their typical, in-home gait. Generally, walks from a two week to three month period are used to fit the GMM, depending on how many walks are identified in an apartment, and how well separated the modes of the residents are in the dataset.

The resulting distribution, or model, of each resident is then used to determine which walks are from that resident, over a given time period, typically three to fourteen days depending on what information is to be obtained (long term trends vs. short term fluctuations). By applying a sliding window, with a step size of one day, changes in the gait parameters of a resident can be tracked over time.

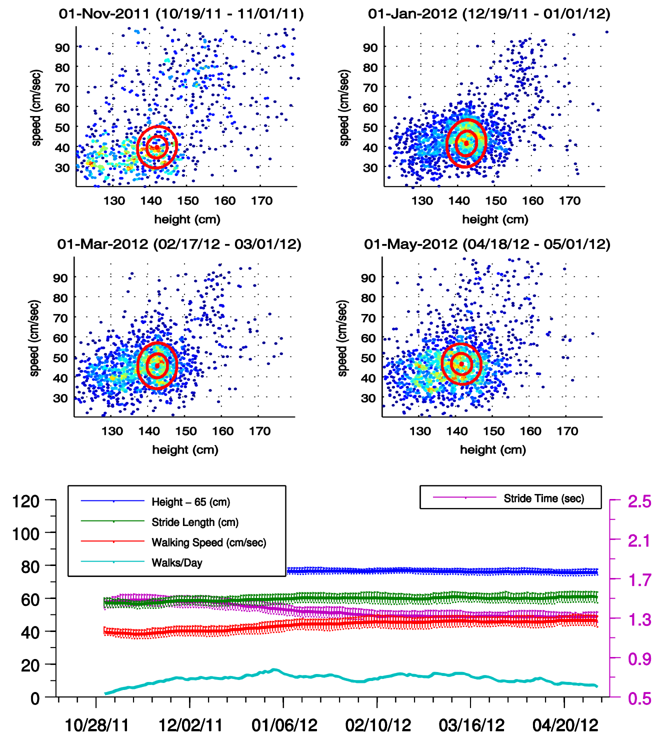

Fig. 2 Illustration of sliding window approach used to generate trends in gait parameters for an apartment. **Top:** Change in resident model over six month time period. Plots show the current model estimate (in 2D) and two weeks of data from the specified time period. **Bottom:** Parameter trend estimates over six month period. Error bars extend from the lower quartile to upper quartile. A two month window was used for the model estimation step, and a two week window was used for the parameter estimation step.

Fig. 2 illustrates this approach for an apartment with a single resident. This resident was admitted to the hospital needing femur surgery on Sep. 3, 2011, (before monitoring was active) and returned to her apartment after rehab on Oct. 25, 2011. Upon returning to her apartment, the resident continued intensive physical therapy while using an assistive walking device for a short period of time, before eventually making a full recovery. This period of recovery is captured in the gait parameter data as increasing walking speed and decreasing stride time.

IV. Methodology

A. Monthly Fall Risk Assessment Protocol

Timed-Up-and-Go (TUG) and Habitual Gait Speed (HGS) tests were administered monthly, when possible, to each of the participants included in the study, as part of a fall risk assessment protocol that also included other standard fall risk assessment instruments such as the Short Physical Performance Battery (SPPB). The TUG and HGS tests were conducted using a walking path of 3 meters, and the reported HGS is an average of two walks. The participants performed the tests unassisted and could decide whether or not use an assistive walking device if they had one.

Issues of participant unavailability and/or participants being physically unable to complete the assessments prevented collection of some tests for some residents. This difficulty in collecting data on even a monthly basis further illustrates the need for passive, continuous assessment.

B. TUG Test-Retest Variability

A number of studies have investigated the test-retest reliability of the Timed-Up-and-Go (TUG) test among various older adult populations and found a high level of intra-individual variation between sessions when the sessions are days or weeks apart [25, 26]; with variation increasing as TUG time increases. This variation depends, of course, on a number of factors including the individual being measured, the individual administering the measurements, and the inherent variability of the test itself. As such, although the TUG has been shown to be good for detecting changes or differences at the population level and useful as a screening tool for assessing fall risk [9], its usefulness for detecting changes in an individual from one session to another when measurements are taken days or weeks apart is limited; as even large changes may simply reflect normal variation, or noise, in the test itself.

In [26], three TUG tests were administered on different days (at approximately the same time of day) to 78 participants within a period of five to seven days. Using this data, the authors calculated 95 percent confidence bounds for an observed TUG measurement based on the expected, true TUG of an individual (taken as the mean of the tests). These bounds are shown in the graph at the top of Fig. 3. Although the population is described as dependent in the activities of daily living (ADL), whereas the population in this work was generally not, the observed TUG times more closely resemble those of the population in this work than TUG times from studies that used populations described simply as community dwelling older adults. Thus, although the time between assessments was significantly shorter than one month, the TUG variability observed is a reasonable basis from which to assess the data in this work. (Although the study in [26] was focused on determining the impact of cognitive state on TUG reliability, and so included a set of cognitively impaired individuals, it ultimately concluded there was no impact of cognitive state on TUG reliability.)

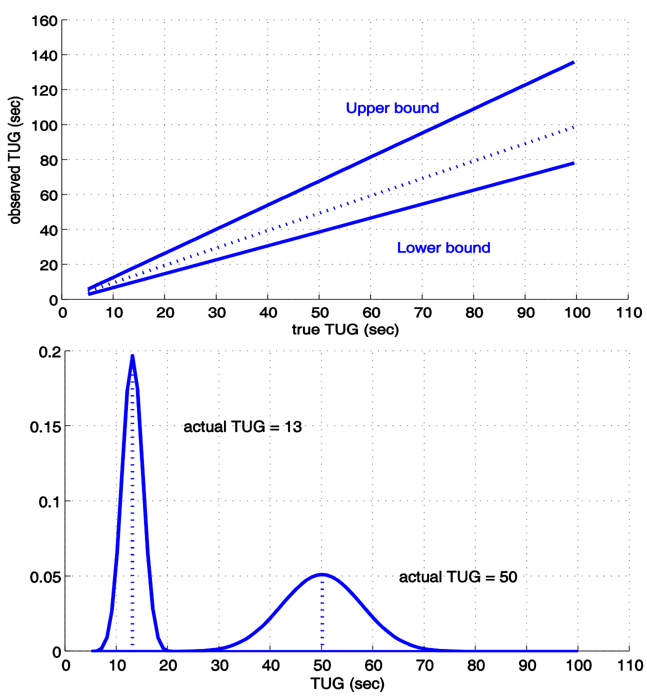

Fig. 3. ***Top:*** Estimated 95 percent confidence bounds for an observed TUG time given the true TUG time of an individual, as reported in [26]. ***Bottom:*** Illustration of the distribution of observed TUG time as a function of actual TUG time for two cases: $TUG_A = 13$ seconds and $TUG_A = 50$ seconds.

Given the high level of intra-individual variation in TUG times between sessions, determining the theoretical expected performance of a system that estimates TUG is vital to correctly evaluating the result. Based on the top graph in Fig. 3, a probability distribution of observed TUG (TUG_O) as a function of true TUG (TUG_A), can be approximated (with a slight bias) by the following normal distribution, shown in the bottom graph of Fig. 3:

$$TUG_O = N(TUG_A, (0.1557 * TUG_A)^2)$$

where the mean is simply the true TUG_A, and the standard deviation is equal to 15.57 percent of TUG_A. Given a probability distribution of TUG_O, it is possible to estimate the expected root mean square error (RMSE) of a set of observed $TUG_{O,i}$, $i = 1, \ldots, M$, times from M individuals compared to the true $TUG_{A,i}$ of the individuals.

First, expected RMSE, $RMSE_E$, of observed $TUG_{O,i}$ as compared to true $TUG_{A,i}$ can be written as:

$$RMSE_E = \sqrt{E\left[\left(TUG_{O,i} - TUG_{A,i}\right)^2\right]} = \sqrt{\frac{1}{N}\sum_i \sigma_i^2}$$
$$= \sqrt{\frac{1}{N}\sum_i (0.1557 TUG_{A,i})^2}$$

Next, the expected value of $(0.1557 TUG_{O,i})^2$ can be written as:

$$E\left[\left(0.1557 TUG_{O,i}\right)^2\right] = 0.1557^2 E\left[TUG_{O,i}{}^2\right]$$
$$= 0.1557^2(\mu^2 + \sigma^2)$$
$$= 0.1557^2\left(TUG_{A,i}{}^2 + \left(0.1557 TUG_{A,i}\right)^2\right)$$
$$= \left(0.1557 TUG_{A,i}\right)^2 (1 + 0.1557^2)$$

Finally, by substitution, $RMSE_E$ can be written in terms of $TUG_{O,i}$ as:

$$RMSE_E = \sqrt{\frac{1}{N}\sum_i \frac{E\left[\left(0.1557 TUG_{O,i}\right)^2\right]}{(1 + 0.1557^2)}}$$

$$= \sqrt{\frac{1}{N}\sum_{i}\frac{\left(0.1557TUG_{O,i}\right)^2}{(1+0.1557^2)}}$$

Thus, if a system that generates estimates of TUG time, $TUG_{E,i}$, were actually modeling the true $TUG_{A,i}$ of the individuals, then (given a sufficiently large dataset) the RMSE of the observed $TUG_{O,i}$ as compared to the estimated $TUG_{E,i}$ would equal the $RMSE_E$ computed above. If, however, the RMSE were significantly less than $RMSE_E$, it would imply that the system was modeling the observed $TUG_{O,i}$, and not the true $TUG_{A,i}$; a possible indication, in the case of a supervised model, of overfitting the training data. Finally, if the RMSE was significantly more than $RMSE_E$, it would imply that there was still room to improve the estimates such that they better modeled the true $TUG_{A,i}$ of the individuals.

Due to the fact that the computed $RMSE_E$ depends on the distribution of $TUG_{O,i}$, which is assumed based on data from another study with a slightly different population, an absolute comparison against this value is interesting and beneficial as an approximate reference point, but should be viewed cautiously. However, by filtering the observed $TUG_{O,i}$, another useful assessment of whether a system is modeling true $TUG_{A,i}$ can be made.

Based on standard techniques from signal averaging, the variance, $\sigma_{S_T}^2$, of a measurement, S_T, obtained by combining N measurements each subject to random noise with mean zero and standard deviation σ, using a filter, $f = [w_1, \dots, w_N], \sum_{i=1}^{N} w_i = 1$, is given by:

$$\sigma_{S_T}^2 = \sigma^2\left(\sum_{i=1}^{N} {w_i}^2\right)$$

This assumes that the signal being measured is stationary over the measurement window and that the measurement noise is independently distributed. Barring a catastrophic event such as a fall, functional ability (and, thus, true TUG) should change slowly over time; while the measurement noise from one session to another should not, in general, be related.

Using a Gaussian filter with N = 5 and σ = 1 (as an example), the variance of the individual observed $TUG_{O,i}$ measurements can be reduced as:

$$f = [w_1, \dots, w_N] \cong [0.054, 0.244, 0.403, 0.244, 0.054]$$

$$\sigma_{S_T}^2 = \sigma^2\left(\sum_{i=1}^{N} {w_i}^2\right) \cong 0.287\sigma^2$$

Substituting this into the previous derivation yields the following approximation for the expected RMSE, $RMSE_F$, of the filtered TUG times, $TUG_{F,i}$, as compared to the true $TUG_{A,i}$:

$$RMSE_F \cong \sqrt{\frac{1}{N}\sum_{i}\frac{(0.287)\left(0.1557TUG_{O,i}\right)^2}{\left(1+(0.1557^2)(0.287)\right)}}$$

$$RMSE_F \cong 0.54\, RMSE_E$$

Thus, if a system were modeling the true $TUG_{A,i}$, the RMSE of the estimated $TUG_{E,i}$ compared to the filtered $TUG_{F,i}$ should be significantly less than the RMSE of the estimated $TUG_{E,i}$ compared to the observed $TUG_{O,I}$. However, if the RMSE against the filtered $TUG_{F,i}$ was not significantly less than $RMSE_E$ this would imply the system is not modeling the true $TUG_{A,i}$.

In summary, the calculated $RMSE_E$ for a set of observed $TUG_{O,i}$ against the actual $TUG_{A,i}$ of the individuals gives a theoretical basis for interpreting the result of a system that attempts to estimate TUG for those individuals. Additionally, comparing the RMSE of the estimated $TUG_{E,i}$ against the observed $TUG_{O,i}$ to the RMSE against the filtered $TUG_{F,i}$ gives an additionally basis for assessing whether the system is estimating the true $TUG_{A,i}$, or simply reflecting the observed $TUG_{O,i}$.

C. Mapping Gait Speed to TUG Time

A simple non-linear neural network model, shown in Fig. 4, was created to map a single dimensional input, gait speed, to a single dimensional output, TUG time. The model uses a hyperbolic tangent activation function in the single hidden neuron, and a linear activation function in the single output neuron. The model has a total of four trainable parameters. The Nelder-Mead Simplex Search method was used to train the model, given randomly initialized weights. Each training run included 50 random initializations, with the best of the 50 trained models being selected. The cost function was defined as the average square error on the training data.

The simplicity of the model, with only four trainable parameters, significantly reduces the chance of overfitting, while increasing the chance of good generalization, even with small datasets. At the same time, it maintains the flexibility needed to closely approximate the non-linear mapping of in-home gait speed to TUG time, as shown in Section V.

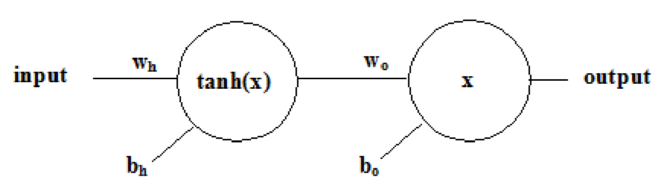

Fig. 4. Simple non-linear neural network model used to map gait speed to TUG time. The model uses a hyperbolic tangent activation function in the hidden neuron, and a linear activation function in the output neuron. The model has a total of four trainable parameters.

V. Results

The Kinect-based gait system was deployed for time periods ranging from 2 to 16 months in 14 apartments in an independent living facility for older adults. Three of the apartments had two residents, yielding a total of 17 study participants. Seven were male and ten were female, ages ranged from 68 to 98 years.

As part of a monthly fall risk assessment protocol, Timed-Up-and-Go (TUG) and Habitual Gait Speed (HGS) tests were administered, when possible, to each resident. This resulted in a total of 154 TUG and HGS measurement pairs. Due to the residents of one apartment being very similar in all physical characteristics measured by the Kinect-based gait system, combined with relatively few walks being identified in their apartment, their data had to be removed, as separate modes, and thus separate gait estimates, could not be identified for each resident. Consequently, the final dataset contained 122 TUG and HGS measurement pairs from 15 individuals. The distributions (mean and standard deviation) of the TUG and HGS measurements, respectively, were 19.5±7.8 sec, and 65.0±18.3 cm/sec.

The in-home gait speed (IGS) of each of the 15 residents was estimated from the Kinect-based gait systems using two weeks of data immediately preceding the date each fall risk assessment protocol was administered (including the day of). The resident models used to identify walks from the residents were based on two months of data immediately proceeding the date each fall risk assessment protocol was administered. The distribution of the IGS measurements was 48.7±12.3 cm/sec. (Although IGS measurements are only shown for the days when the fall risk assessments protocols were administered, the Kinect-based gait systems generate new estimates daily.)

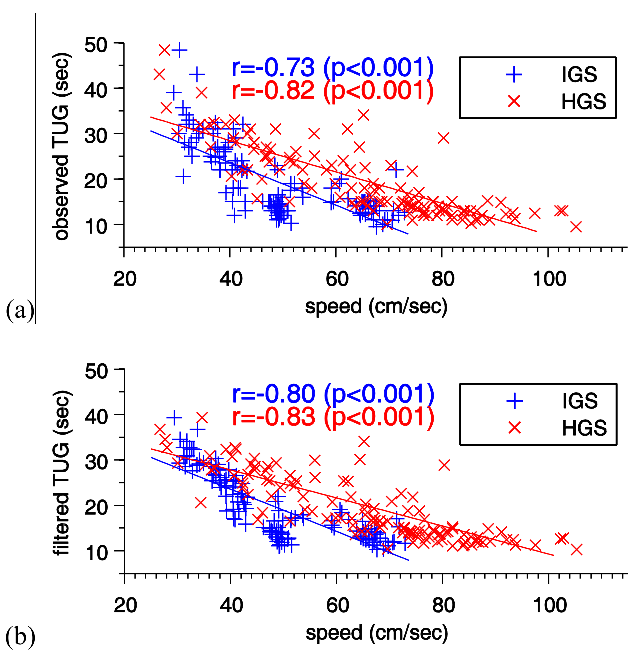

Fig. 5. Plots of observed, (a), and filtered, (b), Timed-Up-and-Go time (TUG$_O$ and TUG$_F$, respectively) as a function of in-home gait speed (IGS) measured using the Kinect-based gait system and habitual gait speed (HGS) measured during the same fall risk assessment protocols as the TUG. Pearson correlations are used to assess the quality of the linear fit.

A. Initial (Linear) Observation

Fig. 5 (a) shows observed TUG (TUG$_O$) recorded from the residents as a function of both IGS and HGS. Pearson correlations indicate that the relationship of HGS to TUG$_O$ is more linear than that of IGS to TUG$_O$. Fig. 5 (b) shows filtered TUG (TUG$_F$) from the residents as a function of both IGS and HGS, where a Gaussian filter (N=5, σ=1, replicated boundaries) has been applied to each resident's TUG$_O$. Pearson correlations indicate that the relationship of IGS to TUG$_F$ is more linear than to TUG$_O$; while the relationship of HGS to TUG$_F$ is essentially the same as to TUG$_O$.

These results suggest, based on the analysis in Section IV, there is a better linear relationship between IGS and the true TUG (TUG$_A$) of the residents, than between IGS and observed TUG$_O$. However, the similarity of the linear relationships of HGS to TUG$_F$ and TUG$_O$ suggest that the relationship of HGS to TUG$_A$ is not any stronger than to TUG$_O$.

B. Estimating TUG from Gait Speed

The simple non-linear network model described in Section IV was trained to map IGS or HGS to observed (not filtered) TUG$_O$. Specifically, leave-one-out cross-validation (LOOCV) was used as follows:

1) All the data from an individual was removed
2) A model was trained using the remaining data
3) The trained model was evaluated on the left out data

For each model run, the training data was normalized by subtracting the mean and dividing by the standard deviation. This normalization was then applied to the removed data. As the data set contained 15 individuals, 15 separate models were trained to map IGS to TUG$_O$, and 15 separate models were trained to map HGS to TUG$_O$.

Fig. 6 shows the mapping of IGS or HGS to TUG that was learned by each model, overlaid on the TUG$_O$ and TUG$_F$ data. Fig. 7 shows TUG$_O$, TUG$_F$, IGS estimated TUG (TUG$_{IGS}$), and HGS estimated TUG (TUG$_{HGS}$) for each individual. Table I contains measures of RMSE, normal error distribution, and Pearson correlation between the estimated values (TUG$_{IGS}$ and TUG$_{HGS}$) and the ground truth values (TUG$_O$ and TUG$_F$). Additionally, the $RMSE_E$ and $RMSE_F$ (described in Section IV) values computed for the data set are included in Table I for comparison purposes.

The small variation between model runs shown in Fig. 6 indicates no significant overfitting of the training data during LOOCV. In the case of both HGS and IGS, the relationship to TUG is clearly non-linear, with the relationship of IGS to TUG being more so than HGS to TUG. However, the simple non-linear model appears capable of closely approximating the relationship in both cases.

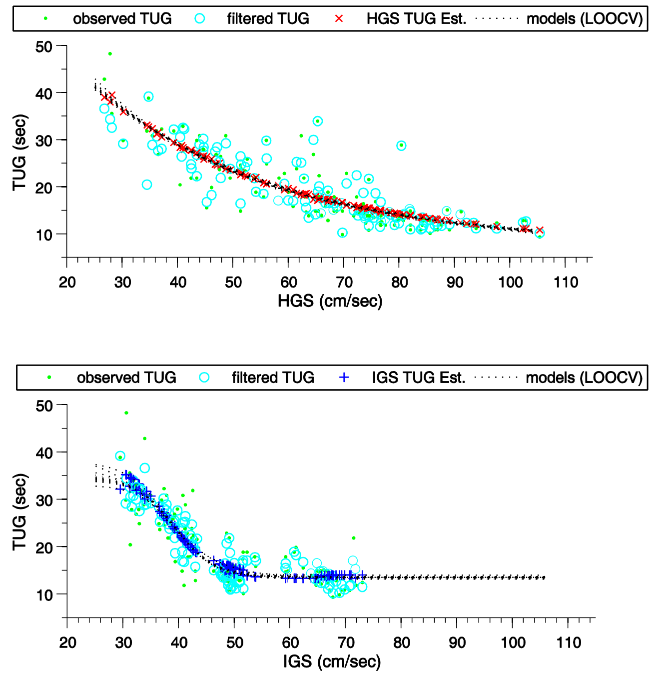

Fig. 6. Mappings of HGS, top, and IGS, bottom, to TUG time learned by each of the models trained during leave-one-out cross validation (LOOCV). Each model is represented by a black dotted line. Observed (TUG_O), filtered (TUG_F), and estimated (TUG_{HGS}/TUG_{IGS}) TUG are overlaid.

As compared against the theoretical performance level, $RMSE_{E,}$ it appears neither the TUG_{IGS} nor TUG_{HGS} estimates are perfectly modeling the true TUG of the individuals, as the RMSE calculated against the observed TUG_O is 4.38 and 4.20 seconds, respectively. However, the fact that TUG_{IGS} and TUG_{HGS} are, essentially, equal in their ability to estimate observed TUG_O (in terms of RMSE, error distribution, and Pearson correlation), suggests that unobtrusive, continuously measured IGS contains as much information about the TUG time an individual would be expected to receive if administered a TUG test as an HGS test would that was administered during the same session as the TUG test. This is compelling evidence that IGS captures key information about the functional ability of an individual.

In the case of TUG_{HGS}, RMSE measured against TUG_O and TUG_F is basically the same, 4.20 and 4.13 seconds, respectively. In the case of TUG_{IGS}, however, RMSE measured against TUG_F is significantly less than that measured against TUG_O, 3.05 vs. 4.38 seconds, respectively. This is compelling evidence that IGS is a better measure of an individual's true functional ability than HGS. Of course, this is somewhat expected, as IGS is based on tens or hundreds of walks captured during normal daily activity, whereas HGS is based on only two walks captured during an explicit performance evaluation.

VI. Discussion

Estimates of observed TUG time based on unobtrusive, continuously measured IGS were shown to be as accurate as estimates based on HGS measured during the same session

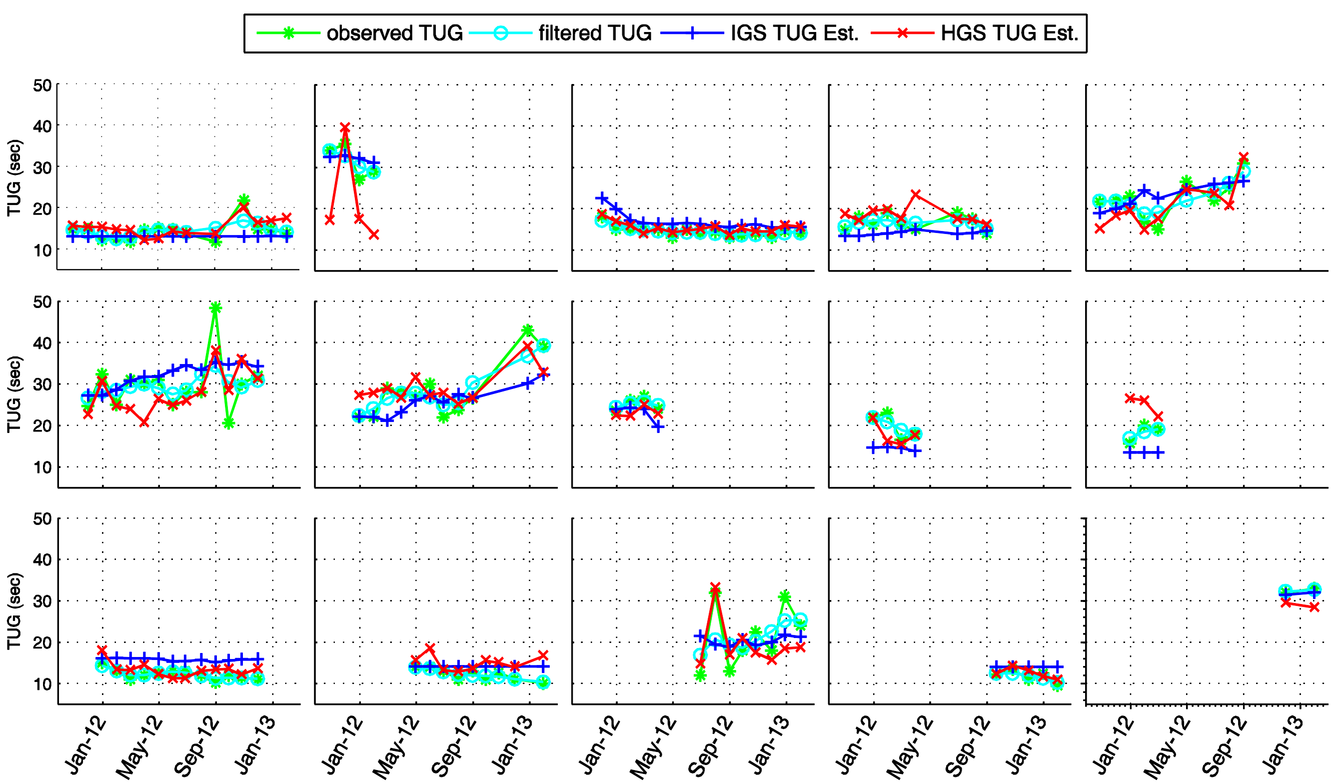

Fig. 7. Plots of observed (TUG_O), filtered (TUG_F), IGS estimated (TUG_{IGS}), and HGS estimated (TUG_{HGS}) TUG time for the 15 individuals included in the study. Each plot corresponds to a separate individual. TUG_{IGS} and TUG_{HGS} are estimates from leave-one-out cross-validation.

TABLE I
ESTIMATED TUG VS. GROUND TRUTH

N=122
15 unique individuals w/ leave-one-out cross validation

$RMSE_E$ for dataset: 3.24
$RMSE_F$ for dataset: 1.75

HGS ESTIMATED TUG (TUG_{HGS})

Ground Truth	RMSE (sec)	Error Dis. (sec)	Pearson Corr.
Observed TUG (TUG_O)	**4.20**	0.03±4.22	0.84 ($p<0.001$)
Filtered TUG (TUG_F)	**4.13**	0.05±4.15	**0.82** ($p<0.001$)

IGS ESTIMATED TUG (TUG_{IGS})

Ground Truth	RMSE (sec)	Error Dis. (sec)	Pearson Corr.
Observed TUG (TUG_O)	**4.38**	0.24±4.40	0.83 ($p<0.001$)
Filtered TUG (TUG_F)	**3.05**	0.26±3.05	**0.91** ($p<0.001$)

as the TUG. In addition, the IGS-based TUG estimates were shown to better model the actual TUG, and, thus, the true functional ability level, of the individuals (based on filtering their observed TUG) than the HGS-based estimates. Such a result is somewhat expected, as IGS is based on tens or hundreds of walks captured during normal daily activity, whereas HGS is based on only two walks assessed during explicit evaluation on the day of the TUG. It also implies that key information about an individual's functional ability, which has typically been assessed in clinical settings, can now be obtained on a continuous basis, in the home, during normal daily activity.

A theoretical analysis of TUG data from another study indicates there is likely room to improve the IGS-based TUG estimates. This improvement could possibly be achieved by incorporating additional parameters measured by the in-home gait system. However, the small size of the available dataset makes this problematic, as overfitting becomes a major issue as the model complexity increases. Techniques such as regularization could help, but a larger dataset is likely needed to truly evaluate the potential of the data captured by the Kinect-based in-home gait systems.

Future work will look to map the in-home gait data to the remaining traditional mobility assessments, such as the SPPB, that are included in the monthly protocols. The goal is to automatically generate, on a daily basis, the same fall risk report that is now collected on a monthly basis by a clinician, and, perhaps, have the report better reflect the true functional ability of the individual.

ACKNOWLEDGMENT

This work was supported in part by the U.S. National Science Foundation under grant CNS-0931607 and by the Agency for Healthcare Research and Quality under grant R01-HS018477.

REFERENCES

[1] M. Runge and G. Hunter, "Determinants of musculoskeletal frailty and the risk of falls in old age," *Journal of Musculoskeletal and Neuronal Interactions*, 6:167–173, 2006.

[2] BE. Maki, "Gait changes in older adults: predictors of falls or indicators of fear, J. of the Am Geriatrics Soc," 45(3): 313–20, 1997.

[3] JM Hausdorff, DA Rios, HK Edelberg, "Gait variability and fall risk in community-living older adults: a 1-year prospective study," Arch Phys Med Rehabilitation, 82:1050–6, 2001.

[4] Y. Barak et al., "Gait Characteristics of Elderly People with a History of Falls: A Dynamic Approach," Phys. Th., 86(11):1501-1510, 2006.

[5] M. Montero-Odasso, et al., "Gait velocity as a single predictor of adverse events in healthy seniors aged 75 years and older," J. of Gerontol. Series A: Biol. Sci. Med. Sci, 60(10):1304–1309, 2005.

[6] J. Verghese, et al., "Quantitative gait markers and incident fall risk in older adults," *J of Geron S. A: Bio Sci and Med Sci,* 64(8):896, 2009.

[7] R. Morris, et al., "A comparison of different balance tests in the prediction of falls in older women with vertebral fractures: a cohort study," *Age and ageing* 36(1):78-83, 2007.

[8] B.R. Greene, et al., "Quantitative Falls Risk Assessment Using the Timed Up and Go Test," IEEE Trans. on Biomed Engineering, 57(12):2918-2926, 2010.

[9] A. Shumway-Cook, S. Brauer, and M. Woollacott. "Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test," *Physical Therapy* 80, no. 9, 896-903, 2000.

[10] PG. Shekelle, et al., "Falls prevention interventions in the Medicare population," Rand Corp. for US Dept. of HHS, 2003.

[11] J. Kaye, et al., "One walk a year to 1000 within a year: Continuous in-home unobtrusive gait assessment of older adults," Gait and Posture, 35(2):197-202, 2012.

[12] E. Stone and M. Skubic, "Evaluation of an Inexpensive Depth Camera for In-Home Gait Assessment," *Journal of Ambient Intelligence and Smart Environments*, 3(4):349-361, 2011.

[13] E. Stone and M. Skubic, "Passive, In-Home Gait Measurement Using an Inexpensive Depth Camera: Initial Results," *Int. Conf. on Pervasive Comp. Tech. for Healthcare*, 2012.

[14] E. Stone and M. Skubic, "Unobtrusive, Continuous, In-Home Gait Measurement Using the Microsoft Kinect," in minor revision, *IEEE Transactions on Biomedical Engineering*, 2013.

[15] D. Podsiadlo and S. Richardson, "The timed 'Up & Go': a test of basic functional mobility for frail elderly person," *Journal of the American Geriatric Society*, 39:142–148, 1991.

[16] JM. Van Swearingen and JS. Brach, "Making geriatric assessment work: selecting useful measures," Phys Ther, 81(6):1233-52, 2001.

[17] JM. Guralnik, et al., "A short physical performance battery assessing lower extremity function: …," *J. of Gerontology, 49*(2):M85, 1994.

[18] Y. Higashi, et al., "Quantitative evaluation of movement using the timed up-and-go test," IEEE Eng in Med and Bio Magazine, 27(4):38-46, 2008.

[19] A. Salarian, et al., "iTUG, a sensitive and reliable measure of mobility," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18(3):303-310, 2010.

[20] C. N. Scanaill, et al., "A review of approaches to mobility telemonitoring of the elderly in their living environment," Ann Biomed Eng, 34(4):547-563, 2006.

[21] MR. Narayanan, et al., "Longitudinal falls-risk estimation using triaxial accelerometry," IEEE Trans on Biomed Eng, 57(3):534, 2010.

[22] G. Demiris, et al., "Older Adults' Attitudes Towards and Perceptions of 'Smarthome' Technologies: a Pilot Study," Medical Informatics and the Internet in Medicine, 29(2):87-94, 2004.

[23] S. Hagler, et al., "Unobtrusive and Ubiquitous In-Home Monitoring: A Methodology for Continuous Assessment of Gait Velocity in Elders," *IEEE Trans Biomed Engineering*, 2009.

[24] T. Frenken, et al., "Novel approach to unsupervised mobility assessment tests: Field trial for aTUG," *Int. Conf. on Pervasive Comp Tech for Healthcare*, 2012.

[25] R.W. Bohannon, and K. Schaubert, "Long-Term Reliability of the Timed Up-and-Go Test among Community-Dwelling Elders," J Phys Ther Sci, 17:93-96, 2005.

[26] E. Nordin, E. Rosendahl, and L. Lundin-Olsson, "Timed 'Up & Go' Test: Reliability in Older People Dependent in Activities of Daily Living – Focus on Cognitive State," Phys Therapy, 86:646-655. 2006.