Sit-to-Stand Measurement for In-Home Monitoring Using Voxel Analysis

Tanvi Banerjee, *Student Member, IEEE*, Marjorie Skubic, *Member, IEEE*, James M. Keller, *Fellow, IEEE*, and Carmen Abbott

***Abstract*—We present algorithms to segment the activities of sitting and standing, and identify the regions of sit-to-stand (STS) transitions in a given image sequence. As a means of fall risk assessment, we propose methods to measure STS time using the 3-D modeling of a human body in voxel space as well as ellipse fitting algorithms and image features to capture orientation of the body. The proposed algorithms were tested on ten older adults with ages ranging from 83 to 97. Two techniques in combination yielded the best results, namely the voxel height in conjunction with the ellipse fit. Accurate STS time was computed on various STSs and verified using a marker-based motion capture system. This application can be used as part of a continuous video monitoring system in the homes of older adults and can provide valuable information to help detect fall risk and enable early interventions.**

***Index Terms*—Activity recognition, ellipse fit, eldercare technology, sit-to-stand (STS), voxel.**

I. Introduction

FALLS are a major cause of injuries among older adults. Fall risk assessments are performed to measure the physical decline of older adults since a change in the functional decline over time could indicate a higher risk of falling. In our research on fall risk assessment, we have seen the importance of regularly monitoring the physical activity of elderly persons as an indication of their functional decline in order to offer early intervention and maintain independence [1], [2]. Tinetti [3] also emphasized the importance of continuous mobility assessment to identify high risk mobility conditions (e.g., falling). Among fall risk parameters, STS analysis has been considered an important component, specifically for the Timed Up and Go (TUG) test [4] as well as the five-times-sit-to-stand test (FTSTS) [5]. The TUG test has been shown to be a sensitive and specific indicator of fall risk among older adults [6]. This test measures the time taken by a participant to rise from a chair, walk 3 m, turn around, walk back, and sit in the chair. In [7], the authors implemented the TUG, FTSTS, and the one-leg balance test on a population of 1618 community dwellers of age 65 and above. They were successfully able to classify older patients in low, moderate, or high risk groups of recurrent falls using these assessments and the FTSTS showed the maximum differentiating capability among these clinical tests. Apart from these two assessments, the short physical performance battery test (SPPB) was also tested on older residents age 70 and above and found the scores to be a strong indicator of subsequent disability. In [8], the authors indicate that a drop in the physical level of older adults significantly increases the fall risk. Again, the SPPB uses STS and stand-to-sit as a part of the evaluation of physical functionality. All these studies highlight the importance of the measurement of STS time and indicate a strong need for a low-cost passive sensor system that can automatically measure the STS time.

This paper proposes a unique markerless technique to measure STS time using inexpensive cameras. Section II discusses different approaches in activity segmentation; Section III describes the background subtraction technique we implemented, algorithms used for preprocessing and the three methods used for measuring STS time. We also discuss the ground truth used for comparing the results we obtained in the laboratory and at the senior housing facility TigerPlace. Section IV presents the experimental setup and the results, and finally, Section V discusses result analysis and concludes our paper.

II. Related Work

The STS time measurement, the primary focus of this paper, has been used by physical therapists since the introduction of the TUG that used this measure as an important component of its balance assessment for frail older adults [4]. Whitney [5] emphasized the FTSTS as an important parameter of gauging balance disorder. She tested 93 subjects, of which 65% were correctly identified with physical dysfunction. Whitney also showed that if the test was conducted for adults of age 60 or younger, the continuous physical assessment was able to accurately discriminate between their previous balance control and their current disorders. Researchers [9], [10] have shown that the inability to perform the basic sit to stand movement can lead to institutionalization, impaired activities of daily living (ADL) functioning, and impaired mobility. The importance of STS assessment was also emphasized in the work of Kerr *et al*. [11], who used the assessment as a means of screening elder adults at risk of falling (EARF), and then, providing strategies for preventing falls to those identified as at risk. Although physical therapists routinely conduct fall risk assessments at TigerPlace, we have found that continuous monitoring in the home helps detect early health changes that would otherwise remain unnoticed [12]. This

Manuscript received June 15, 2013; revised September 24, 2013; accepted September 30, 2013. Date of publication October 3, 2013; date of current version June 30, 2014. This work was supported in part by the U.S. National Science Foundation under Grant IIS-0703692.

T. Banerjee, M. Skubic and J. M. Keller are with the Electrical and Computer Engineering Department, University of Missouri, Columbia, MO 65211, USA (e-mails: tsbycd@ mizzou.edu; skubicm@missouri.edu; kellerj@missouri.edu).

C. Abbott is with the School of Health Professions, Physical Therapy, University of Missouri, Columbia, MO 65211, USA (e-mail: AbbottC@missouri.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JBHI.2013.2284404

inspired us to develop algorithms to measure the transition time using vision sensors in an automated manner.

Among the studies involving visual sensors, Allin *et al*. [13] focused on two parameters related to STS as a measure of physical capability (use of hands/arms and position of feet). Using three cameras, 3-D features such as distance between feet, head, and body centroid were constructed. These features were tracked using ellipsoid tracking of the individual positions of the head, torso, and feet using the Weka Machine Toolkit for classification [14], and strong correlations were achieved for five participants between the measured rise time and the Berg balance score. However, the body parts for each subject had to be manually labeled for at least one image for the system to learn the color information for the individual. A unique method of extracting the silhouette of the human body using the snake algorithm was implemented by Goffredo *et al*. [15] to get posture information. The trajectories of the marker-less pivotal limb joints of the human body were tracked and used to obtain a comparison of the STSs of different subjects as well as of the same subject over different time periods to indicate physical deterioration over time. Pehlivan *et al*. [16] used circular features as pose descriptors from volumetric image data; features included are the number of circles, area of outer circles, etc., from each layer. Nearest neighbor was utilized for classification for activity identification including STS motion. Another interesting work studied activities including STS in Alzheimer's patients [17]. They used *a priori* information such as the location of furniture and objects in the room to identify the possible activities as well as the 3-D position, height, and width of silhouettes to track their activities.

In addition, accelerometers positioned on the body have been used to measure STS-related parameters. In [18], a single waist-mounted triaxial accelerometer was used to classify activities such as sitting, lying, standing, running, and transition activities like STS and falling. Experimental results showed that the successful detection rate for all activities was about 96%. Weiss *et. al* [19] used accelerometers to distinguish between subjects with Parkinson's Disease and normal subjects. The second half of the STS task; actively rising from the chair, was used as the distinguishing point since it was especially difficult for the patients as compared with the control group. In [20], an array of 132 embedded fiber optic pressure sensors was placed under a hospital bed mattress and sampled at 10 Hz. The pressure images and regions of interest were formed based on their location and the weighted centroid, and the total pressure measurement over time was used to compute STS time.

Many of the aforementioned approaches involve the use of wearable sensors [18]–[20], which indicate the need for patients to constantly wear the devices for monitoring. Among the visual segmentation techniques, some of the drawbacks include the use of a three-camera system, which increases complexity and the activity segmentation more computationally expensive [13], or a computationally intensive algorithm for measuring STS time [13]. Our goal is to create algorithms that are suitable for continuous monitoring in the unstructured home setting. Thus, we have investigated fast, noniterative algorithms for segmenting movements and have tested the performance under different viewing angles and in different settings. The ultimate goal is to accurately measure the STS times in an unstructured environment from a video sequence, to facilitate continuous monitoring in the home.

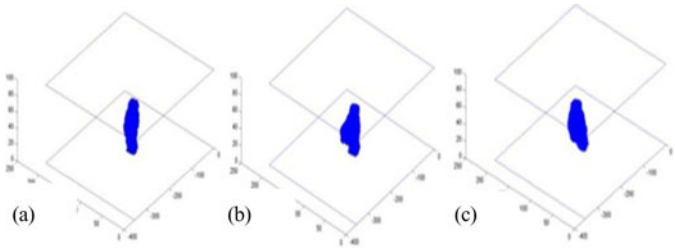

Fig. 1. Voxel person models (a) standing, (b) transitioning (stand-to-sit), and (c) sitting.

III. Methodology

In this section, we propose methods for capturing STS transitions in unstructured environments with no constraints on the viewing angle. We begin by providing an overview of silhouette extraction and 3-D model construction. Then, three methods are presented for segmenting the sitting and upright states, from which the STS time can be computed.

A. Background Subtraction and Voxel Model Construction

Silhouettes of moving persons are extracted from image sequences using a mixture of Gaussians with color and texture features [21]. Here, the color components are expressed in Hue, Saturation, and Value format; these results are fused with the texture results using the Yager union. Morphological operations are then carried out to give the final silhouettes. For our application, around ten images are used to build the background. The resultant background is then subtracted from each image to yield the foreground. In the home setting, silhouettes are extracted using dynamic background updates as described in [22]. Research has shown that the use of silhouettes not only defines the region of interest but also helps protect privacy when monitoring an older adult in the normal daily living environment. Our research shows that elderly residents do not consider the use of silhouette imagery to be a privacy invasion [2].

Two orthogonal cameras with a frame rate of 5 per second were utilized to capture the image sequences. The silhouettes from both cameras were then combined with the points of intersection, creating a 3-D model in voxel space. This creates a model that is independent of the viewing angle. Voxel models have been used in several applications such as real-time level set matching [23], head tracking [24], and modeling specular surfaces [25]. We have termed the 3-D person created in voxel space a "voxel person" [26]. The voxel resolution used here is a 1 in cube (2.54 cm). Examples of the voxel person model in different states are shown in Fig. 1.

B. State Segmentation Using VH

For segmenting the upright and the sit regions in an image sequence, the VH of a person is investigated as an initial

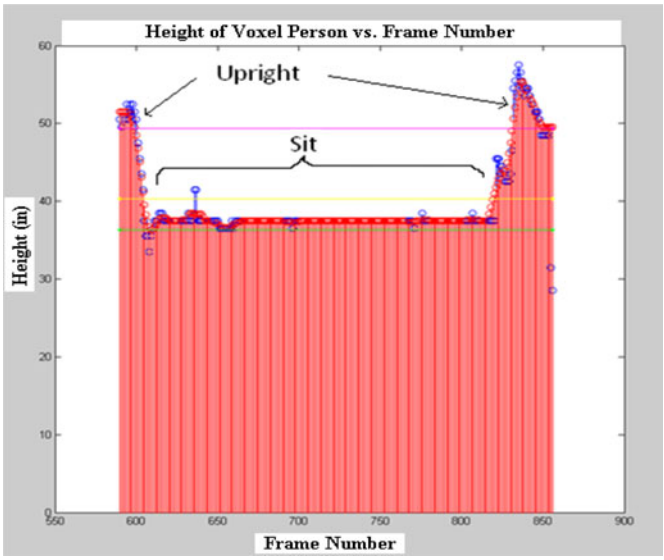

Fig. 2. Height graph of a sequence indicating the upright and sit regions. The height before filtering is shown in blue and after filtering in red.

approach. Fig. 2 shows the height of a sequence of a person sitting down and getting up from a chair.

Average filtering with a window size of 5 is used to smooth the VH over time as shown in Fig. 2. This technique implemented used the height of the voxel person alone for the measurement of the STS time. As shown in Fig. 2, by merely observing the height of the voxel person in 3-D space, it is possible to estimate upright and sitting positions.

C. State Segmentation Using Voxel Height and Orientation (VHO)

As a second approach for segmenting upright and sitting regions, the orientation of the silhouette is captured using image moments. The transition region was roughly identified using the average filtered VH described in the previous section. After the initial segmentation, the results were further refined using the orientation information.

The silhouette orientation is computed using the following equation:

$$\theta = \tan^{-1}\left(\frac{2 * \mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{1}$$

where

$$\mu_{pq} = \sum_{x=1}^{M}\sum_{y=1}^{N}(x-\bar{x})^p * (y-\bar{y})^q * f(x,y)\,. \tag{2}$$

Here, μ_{pq} are the central image moments at the image centroid $(\bar{x}, \bar{y})$ with $f(x, y)$ as the image intensity value at coordinate (x, y). For our application, the second-order image moments are used to compute the orientation of the silhouette that represents the shape and the angle of the silhouette computed with respect to the normal to the floor plane.

In order to choose the camera view containing more information, the orientation values for the segmented silhouette sequence are computed from both views. Then, the view containing the maximum range of values is chosen for further analysis. Fig. 3 illustrates the orientation (in degrees) in a STS sequence from both camera views. The graphs indicate the results after averaging the values using an averaging filter with window size 5 so as to mitigate the noise present in the silhouettes. In this particular sequence, a person was initially sitting down, and then, he got up from the chair. For the side view, the plateau-like region in the graph indicates the region of the image sequence where the person was sitting. As this is the orientation of the entire body of the person as he is sitting down, the angle is quite large with respect to the normal to the ground. As he gets up, this angle keeps decreasing until it reaches to almost 0° when the person is completely erect. The horizontal line in the graph indicates the threshold to obtain the location of the first frame from which STS begins. A threshold of 4° from the maximum sitting angle yields satisfactory results on the image sequences for which this algorithm has been tested. For the front view, the range of values is very low compared to the side view. This is further apparent from the silhouettes shown in row 3 in Fig. 3. Our automated view selection chooses the side view as it has the higher range or variability in the orientation values.

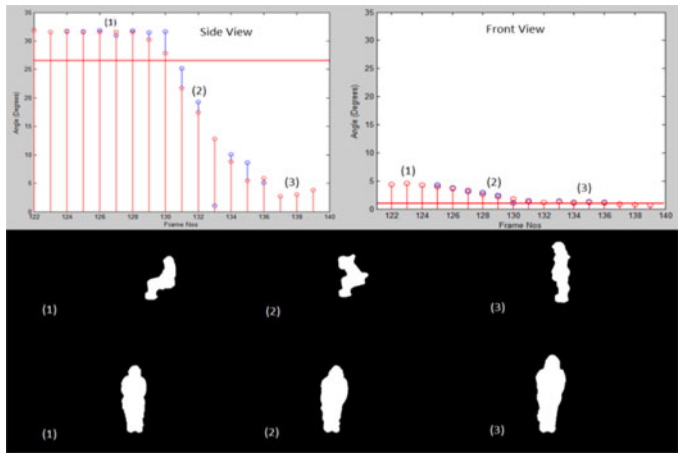

Fig. 3. Graph of orientation (in degrees) versus the frame number of a sequence using side view as well as front view. The filtered results are shown in red. Row 2 contains the silhouettes as can be seen by the camera from the side view. Row 3 contains the front view silhouettes. The corresponding silhouettes are marked in the graphs for both views.

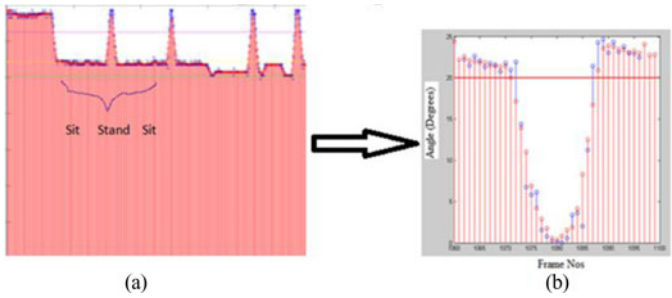

Fig. 4. (a) Height (inches) of a person in a sequence (viewing angle 90° with camera 1) indicating a person performing 4 STS; (b) result of orientation (in degrees) after using the averaging filter on the sequence marked in 4(a). The filtered results are shown in red.

Fig. 4(a) shows the height of the person in a sequence showing a person walking into a room and performing four STS motions

in a chair, and then, ending in the sitting position. In this example, the chair is positioned at approximately 90° with respect to camera 1. The region of interest here consists of the two consecutive sit regions highlighted (sit–stand–sit) which was then used as inputs to compute the orientation. Once the smooth VH information was extracted, the transition region highlighted in Fig. 4(a) was then analyzed using the orientation feature to refine the results. The results obtained from orientation are displayed in Fig. 4(b).

As can be seen from Fig. 4(b), the STS region is defined more clearly after adding the orientation feature to the existing VH information. The red line shows the thresholded STS region with the intersection points indicating the beginning and end of the STS.

D. State Segmentation Using Voxel Height and Ellipse Fit (VHE)

The third technique investigated for segmenting upright and sitting regions uses curve fitting to identify an ellipse that best fits the silhouette. Again, the VH is used as a first stage in segmenting the regions. Similar to the VHO technique, the results are further refined by evaluating the ratio of the major axis length and minor axis length from silhouettes extracted from both angles. Here, the optimization technique proposed by Fitzgibbon and Fischer [27] is used for the ellipse fit. This idea is based on least-squares optimization in reference to the general equation of a quadratic curve (3).

$$F(x,y) = ax^2 + bxy + cy^2 + dx + ey + f = 0. \quad (3)$$

Subject to the constraint (for ellipse)

$$b^2 - 4ac < 0. \quad (4)$$

where $\bar{a} = [a, b, c, d, e, f]$ are the coefficients of the ellipse and $\bar{x} = [x^2, xy, y^2, x, y, 1]$ is the vector of coordinates.

Equation (3) can be rewritten as

$$F_a(x) = \bar{a} * \bar{x} = 0. \quad (5)$$

Thus, the fitting of an ellipse on a set of points is given by

$$\min_a \sum_{i=1}^{N} F_a(x_i, y_i)^2 = \min_a \sum_{i=1}^{N} (\overline{x_i}.\bar{a})^2. \quad (6)$$

In general, Fitzgibbon showed that convergence was slow for constraints like (4). Thus, an additional constraint was specified as

$$4ac - b^2 = 1. \quad (7)$$

The problem can be reformulated as

$$\min_a ||Da||^2 \text{ subject to constraint } \overline{a}^T C\bar{a} = 1 \quad (8)$$

where the Design Matrix of size $N \times 6$ was given by

$$D = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_i^2 & x_i y_i & x_i^2 & x_i & y_i & 1 \\ x_N^2 & x_N y_N & x_N^2 & x_N & y_N & 1 \end{pmatrix}.$$

The details of the constraint matrix C are explained in our implementation in [28]. The final system of equations is then solved by

$$||Da||^2 = \bar{a}^T D^T \bar{a} D = \bar{a}^T S\bar{a} = \bar{a}^T \lambda C\bar{a} = \lambda. \quad (9)$$

This technique was utilized by Halir and Flusser [29] who addressed the drawbacks in [27] by considering the possibility of the scatter matrix S being singular. The details of our implementation are described in [28]. This method was an improvement in terms of numerical stability and preventing local optimization. The advantage of this ellipse fit algorithm is that it is noniterative and extremely fast, making it suitable for real-time applications. Fig. 5 shows examples of ellipse fitting on the set of points located at the edge of the silhouettes obtained using the Canny edge detector [30].

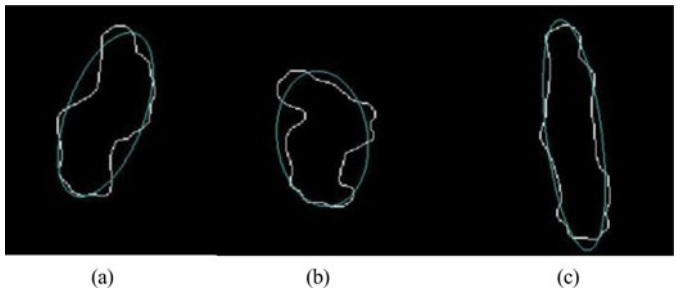

Fig. 5. Edge of the silhouette extracted using Canny edge detector and ellipse fit of a person (a) sitting, (b) getting up, and (c) standing.

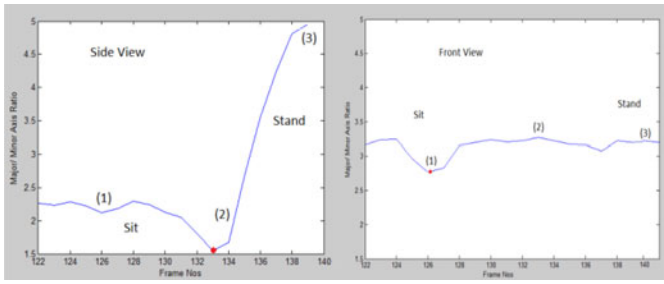

Fig. 6. Ellipse major axis to minor axis ratio for side view as well as front view of the sequence from Fig. 3. The red points indicate the beginning of the detected STS. The frames from Fig. 3 are marked in the graphs.

As can be seen in Fig. 5(b), the fitted ellipse approaches a circle as the person is transitioning from sitting to standing. The ratio of the major axis length to the minor axis length is the least at that point. Using the same example as shown in Fig. 3, the major axis to minor axis ratio graph for both camera views is shown in Fig. 6. There is a noticeable dip in the side view that occurs when the person begins the process of getting up (silhouette marked 2 in Fig. 3). This position corresponds to the silhouette shown in Fig. 5(b) where the ratio is minimized. Here, these minimum ratio points are used to segment the sitting and standing regions. Again, we see that the range of values is much lower for the front view as compared to the side view so the side view sequence is selected for refining our results. It can also be seen that the side view correctly identifies the beginning of the STS, whereas the front view locates the minima at the sit region which is incorrect.

The algorithms described in Sections III-B, -C, and -D were used for capturing STS transitions in the experiments described in Section IV. Once the transition frames were identified, the

Fig. 7. Elderly participant with markers on the head, shoulder, two on the back and feet (highlighted in red) for the VICON motion capture system.

time was measured by computing the difference between the time stamps of the beginning and end of the transition sequences. The results were further validated using ground truth described below in Section III-E.

E. Vicon System: Ground Truth for Laboratory Experiments

The Vicon Nexus motion capture system (Version 1.4, Vicon Motion Systems, Inc., Centennial, CO, USA) was used as ground truth in the laboratory experiments to identify the activities as well as measure the STS times. Reflective markers are placed at key points on the subject's body. These markers are detected by the MX cameras and their 3-D locations are determined over time which gives an accurate description of the activities performed. For our experiments, reflective markers were placed on the top of the head, shoulders, on top of the back in line with the shoulders, in the middle of the back, and on the feet of the subjects. This can be seen in Fig. 7. These markers allowed the Nexus software to detect the activities of the person, while he/she was walking, standing, sitting, or getting up from the chair within the field of view of the camera system.

For the experiments described here, the height of the person (obtained from the marker) is used to get information about the person regarding his or her state of motion, i.e., whether he is sitting, upright, walking, or in transition state (sit-to-stand or stand-to-sit).

F. Stop Watch: Ground Truth for TigerPlace

Physical therapists use the stop watch to capture STS times. For comparison, it was used as a secondary measure of ground truth in the laboratory experiments. To capture time using the stopwatch, our physical therapy expert manually observed a video sequence of the subject performing the STS motions. She went through the sequence several times and identified the image frames that defined the beginning and end of the STS. This method is more accurate than using a manual stopwatch since sometimes there is no immediate response from the subject as he or she rises to perform a STS and this delay also gets counted as part of the STS time. Also, there is a chance of manual error in case the therapist does not stop the timer at the same instant when the person touches down on the chair seat as he or she sits back in the chair. These potential errors make observing the video sequence to determine the start and end frames of the STS and use the time stamps of these selected frames a more accurate measure of ground truth. In our laboratory experiments, a physical therapist validated the use of a stop watch as a ground truth measure by measuring the STS time using silhouettes of the participants. Since we could not use the Vicon system as ground truth at TigerPlace, we were able to successfully match the results from the stop watch and the Vicon system results in laboratory settings in order to justify our use of a stop watch for our experiments at TigerPlace. We describe the experimental setup further in Sections IV-A and -B.

TABLE I
AVERAGE STS TIME DIFFERENCE FOR FIVE SUBJECTS WITH CHAIR AT 90° WITH RESPECT TO CAMERA 1 USING THE SW, VH, VHO, AND VHE COMPARED WITH THE VICON SYSTEM IN SECONDS

Subjects	SW – Vicon(s)	VH – Vicon(s)	VHE – Vicon(s)	VHO – Vicon(s)
1	0.02	-0.41	-0.18	-0.30
2	0.04	-0.37	-0.19	-0.23
3	0.03	-0.32	-0.09	-0.28
4	0.01	-0.39	-0.15	-0.22
5	-0.02	-0.39	-0.20	-0.29
Avg	**0.02**	**-0.38**	**-0.16**	**-0.26**

IV. EXPERIMENTAL SETUP, RESULTS, AND ANALYSIS

A. Testing in Laboratory

Preliminary experiments were conducted on five healthy participants with the video sequence captured at a rate of five frames per second using a two-camera system. The Vicon system was used for validation along with the stopwatch. The chair used had a standard seat height (approximate 46 cm) as recommended for the STS component in the Berg balance scale test [4]. To test the robustness of the different approaches, experiments were conducted with the chair placed at different angles with respect to the camera. The results using this setup are shown in Table I and Fig. 8.

Table I shows the average STS time differences for five subjects between the different techniques, namely, the stop watch (SW), computing STS time using the VH alone, using VH in conjunction with the ellipse fit technique (VHE) and using the VH along with the orientation method (VHO). The average for all the different types of STSs was taken for each participant. The angle of the chair was kept fixed at 90° with respect to Camera 1. Each subject performed six different types of STSs: slouching forward, slouching to the left and right sides, getting up from the chair with feet away from the chair, a normal sit to stand, and with the assistance of hands. Each type of STS was performed twice. The differences between the times measured using these methods were then compared by subtracting them from the ground truth, in this case the Vicon system.

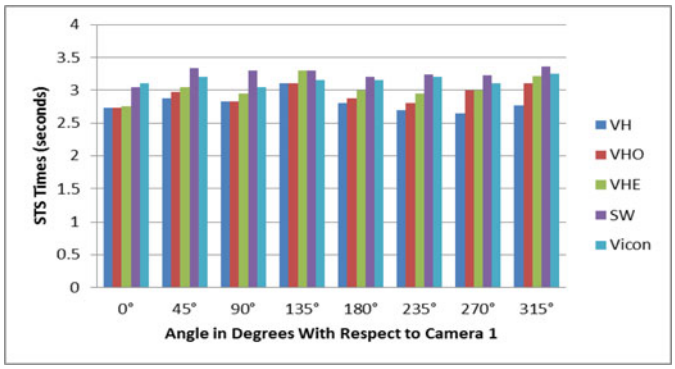

Fig. 8. Comparison of all the techniques to measure the STS time with chair at various angles with respect to Camera 1. Note 0° means the chair is at 0° with respect to Camera 1, i.e., front view.

TABLE II
PERCENTAGE ACCURACY USING THE THREE TECHNIQUES VH, VHO, AND VHE

	Perpendicular Camera View	Variable Angle Chair Location
VH	89.8 %	86.7 %
VHO	95.4 %	93.5 %
VHE	97.3 %	96.6 %

As can be seen, the VHE method yields the lowest difference with respect to the Vicon system. After testing our algorithms on the five subjects with different types of STS, we tested for angle variability. Fig. 8 shows the measured STS time with the chair positioned at varying angles from the cameras (here, the angles are computed with respect to a specific camera termed Camera 1 for clarification) using the three techniques (VH, VHO, and VHE) compared with the Vicon and SW. Each subject performed two normal STSs in this experiment.

Table II shows the individual activity recognition accuracy for the activities of Sit, Upright, and Transition (sit-to-stand and stand-to-sit).

As can be seen, the VHE method shows the best agreement with the Vicon system. The VHO results appear slightly less accurate than VHE and the VH results indicate that the VH alone cannot measure accurate transition times.

B. Testing With Older Adults in TigerPlace

The goal of this study is to capture STS measures for older adults in their homes as a part of their normal routines to assess their physical functionality. In order to test this, we evaluated our algorithms at TigerPlace where participants enacted a scripted scenario performing everyday activities. This IRB approved study included ten older adults from ages of 83 to 97 with different health conditions such as orthopedic disorders, heart problems as well as some requiring assistance of a cane to walk around. In this two-person scenario, a "repairman" enters the scene. The activities performed by the participants include walking around, stretching their arm to reach for something, sitting on a chair and getting up, greeting the repairman, stepping over an object on the floor, activities which might occur daily in the lives of the residents at TigerPlace. Each participant repeats the scripted scene for reliability of results. The chair in the settings is approximately at right angle to one of the cameras and is facing the other camera in the scenario. Fig. 9 gives the results of implementing the STS techniques on the experiments conducted in TigerPlace.

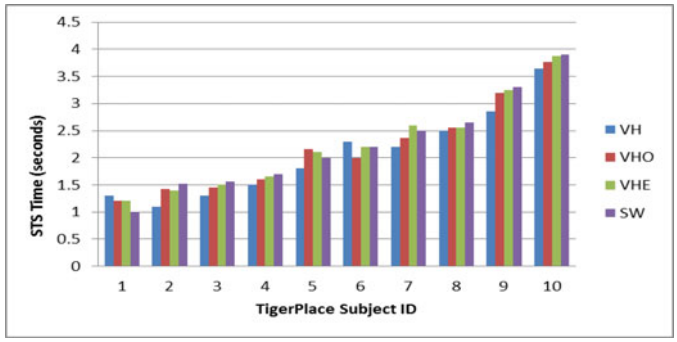

Fig. 9. Comparison of all the techniques to measure the STS time with older adults at TigerPlace. The X-axis represents the participant number and the Y-axis is the STS time in seconds.

TABLE III
PERCENTAGE ACCURACY USING THE THREE TECHNIQUES VH, VHO, AND VHE

%	VH	VHO	VHE
	88.3 %	95.5 %	96.9 %

TABLE IV
T-TEST RESULTS USING THE THREE TECHNIQUES VH, VHO, AND VHE AND COMPARISON WITH SW AND VICON UNDER BOTH SETTINGS

Setting	Model	Diff Mean	T
Lab	Vicon-SW	-0.14812	2.702
	Vicon-VH	0.33125	6.029*
	Vicon-VHO	0.23	4.133*
	Vicon-VHE	0.105	1.706
	SW-VH	0.479375	7.354*
	SW-VHO	0.378125	5.749*
	SW-VHE	0.253125	2.573
TP			
	SW-VH	0.17	0.441
	SW-VHO	0.094	0.244
	SW-VHE	0.011	0.027

* Significant at alpha = .01 with Bonferroni Correction applied.

The overall activity classification rates are given in Table III. The results are consistent with the results obtained from the laboratory experiments.

In order to compare the significant difference between the proposed methods under both laboratory settings and at TigerPlace, we reported the T-test results along with the Bonferroni correction in Table IV. The models with significant difference are highlighted with an asterix ($*$). We see that the SW and Vicon are not significantly different from each other which validates our use of SW as ground Truth in the repairman scenario at TigerPlace. Furthermore, the only technique which was not significantly different to the Vicon in the laboratory setting is

VHE. However, this changes for the results obtained at TigerPlace. Here, none of the techniques appear significantly different from SW.

V. Discussion and Conclusion

This paper describes three techniques to measure STS time using standard low-cost web cameras. These techniques were further validated with older adult participants at TigerPlace. In all, a total of 150 min of video data were collected with 15 subjects in laboratory as well as nonclinical unstructured settings at TigerPlace. Tables II and III indicate strong agreement between the stop watch measurements and the Vicon measurements. This validates our usage of the stopwatch as ground truth in the scenarios we tested at TigerPlace. It can also be seen from Figs. 8 and 9, and Tables II and III that the VHE method described in Section III-D is closest to the stop watch and Vicon measurement. The VHO technique described in Section III-C shows similar results to VHE, especially in the sequences with the lower STS time values. However, it is more sensitive to noise compared to VHE. In general, all techniques show promising results both in the laboratory as well as at TigerPlace; none appear significantly different from SW at TigerPlace. This is interesting because it implies that the VH alone can be used to approximate STS measurement in an in-home environment.

Rigorous testing was done for STSs at right angles to the camera view as well as for STSs at various angles to the field of view. For both the frame classification results as well as the STS time measurements, the technique using the VH along with the ellipse fitting (VHE) method yields the best results. An important observation that can be seen from Figs. 8 and 9 is the high variability in the STS time measurements between the residents at TigerPlace compared to the healthy participants in the laboratory. This corroborates with the findings of the researchers in [11] and shows a high potential in being able to identify the participants with a potential risk of falling from those who are not. The ten elderly residents who participated in the study were shown to have a range of walking speeds from 40–100 cm/s showing their different levels of physical functionality [31]. We were able to conduct passive opportunistic monitoring in a natural environment with a truer picture of the individual's capabilities. Using the techniques described, the rise time computation was effective in capturing the STS time. This is an alternative approach to activity recognition compared to [31] and [32].

An important challenge that needs to be addressed here is occlusion detection. Our system assumes that well-segmented image sequences are extracted as a first step. To address challenges in a realistic, dynamic environment, we are working on automatically detecting occlusions using foreground features. Depending on the degree of occlusion detected, we plan on using a fusion of the VH, VHO, and VHE methods. Future work also includes incorporating this measure into an in-home assessment that can also be used in conjunction with a multifactorial fall risk assessment (including Timed Up and Go, ten foot walk, tandem stand), which is periodically administered to track the status of functional mobility and increasing fall risk, which promises to be of importance in the future as a means of detecting early signs of decline. This in turn can expedite interventions to help older adults lead an independent lifestyle.

References

[1] G. Demiris, M. Skubic, M. Rantz, J. Keller, M. Aud, B. Hensel, and Z. He, "Smart home sensors for the elderly: a model for participatory formative evaluation," in *Proc. IEEE EMBS Int. Spec. Topic Conf. Inform. Technol. Biomed.*, 2006, pp. 1–4.

[2] G. Demiris, O. D. Parker, J. Giger, M. Skubic, and M. Rantz, "Older adults' privacy considerations for vision based recognition methods of eldercare applications," *Technol. Health Care*, vol. 17, no. 1, pp. 41–48, 2009.

[3] M. E. Tinetti, "Performance-oriented assessment of mobility problems in elderly patients," *J. Amer. Geriatric Soc.*, vol. 34, pp. 119–126, 1986.

[4] D. Podsiadlo and S. Richardson, "The timed "Up & Go": A test of basic functional mobility for frail elderly persons," *J. Amer. Geriatric Soc.*, vol. 39, pp. 142–148, 1991.

[5] S. L. Whitney, "Clinical measurement of Sit to Stand performance in people with balance disorders: Validity of data for the five times sit to stand test," *Phys. Therapy*, vol. 85, pp. 1034–1045, 2005.

[6] A. Shumway-Cook, S. Brauer, and M. Woollacott, "Predicting the probability for falls in community dwelling older adults using the Timed Up and Go test," *Phys. Therapy*, vol. 80, pp. 896–903, 2000.

[7] S. Buatois, D. Miljkovic, P. Manckoundia, R. Gueguen, P. Miget, G. Vancon, P. Perrin, and A. Benetos, "Five times sit to stand test is a predictor of recurrent falls in healthy community-living subjects aged 65 and older," *J. Amer. Geriatric Soc.*, vol. 56, pp. 1575–1577, Aug. 2008.

[8] J. M. Gurlanik, L. Ferrucci, and E. Simonsick, "Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability," *N. Engl. J. Med.*, vol. 332, pp. 556–561, 1995.

[9] S. L. Whitney, D. M. Wrisley, G. F. Marchetti, M. A. Gee, M. S. Redfern, and J. M. Furman, "Clinical measurement of sit-to-stand performance in people with balance disorders: Validity of data for the five-times-sit-to-stand test," *Phys. Therapy*, vol. 85, pp. 1034–1045, 2005.

[10] W. G. Janssej, H. B. Bussmann, and H. J. Stam, "Determinants of the sit-to-stand movement: A review," *Phys. Therapy*, vol. 82, pp. 866–879, 2002.

[11] A. Kerr, D. Rafferty, K. Kerr, and M. Durward, "Timing phases of the sit-to-walk movement: Validity of a clinical test," *Gait Posture*, vol. 26, no. 1, pp. 11–16, Jun. 2007.

[12] G. L. Alexander, M. Rantz, M. Skubic, R. J. Koopman, L. J. Phillips, R. D. Guevara, and S. J. Miller, "Evolution of an early illness warning system to monitor frail elders in independent living," *J. Healthcare Eng.*, vol. 2, no. 2, pp. 259–286, 2011.

[13] S. Allin and A. Mihailidis, "Low-cost, automated assessment of sit-tostand movement in "Natural" environments," in *Proc. 4th Eur. Conf. Int. Federation Med. Biolog. Eng.*, vol. 22, Berlin, Germany, 2009, pp. 76–79.

[14] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. San Mateo, CA, USA: Morgan Kaufmann, 2005.

[15] M. Goffredo, M. Schmid, S. Conforto, M. Carli, A. Neri, and T. D'Alessio, "Markerless human motion analysis in Gauss-Laguerre transform domain: An application to sit-to-stand in young and elderly people," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 2, pp. 207–216, Mar. 2009.

[16] S. Pehlivan and P. Duygulu, "A new pose-based representation for recognizing actions from multiple cameras," *Comput. Vis. Image Understanding*, vol. 115, no. 2, pp. 140–151, Feb. 2011.

[17] R. Romdhane, E. Mulin, A. Derreumeaux, N. Zouba, J. Piano, J. Lee, I. Leroi, P. Mallea, M. Thonnat, F. Bremond, and P. Robert, "Automatic video monitoring system for assessment of alzheimer's disease symptoms," *J. Nutrition, Health Aging*, 2011.

[18] D. W. Kang, J. S. Choi, and J. W. Lee, "Real-time elderly activity monitoring system based on a tri-axial accelerometer," *Disabil. Rehabil. Assist. Technol.*, vol. 5, pp. 247–253, 2010.

[19] A. Weiss, T. Herman, M. Plotnik, M. Brozgol, I. Maidan, N. Giladi, T. Gurevich, and J. Hausdorff, "Can an accelerometer enhance the utility of the timed up & go test when evaluating patients with Parkinson's disease?" *Med. Eng. Phys.*, vol. 32, no. 2, pp. 119–125, Mar. 2010.

[20] A. Arcel Arcelus, I. Veledar, R. Goubran, F. Knoefel, H. Sveistrup, and M. Bilodeau, "Measurements of sit-to-stand timing and symmetry from bed pressure sensors," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1732–1740, May 2011.

[21] R. H. Luke, D. Anderson, D. J. M. Keller, and M. Skubic, "Human segmentation from video in indoor environments using fused color and texture features," ECE Dept., University of Missouri, Columbia, MO, USA, Technical Report, 2008.
[22] E. E. Stone and M. Skubic, "Silhouette classification using pixel and voxel features for improved elder monitoring in dynamic environments," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, Seattel, WA, USA, Mar. 21–25, 2011, pp. 655–661.
[23] Y. Iwashita, R. Kurazume, T. Tsuji, K. Hara, and T. Hasegawa, "Fast implementation of level set method and its realtime applications," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2004, pp. 6302–6307.
[24] H. Kawanaka, H. Fujiyoshi, and Y. Iwahori, "Human head tracking in three dimensional voxel space," in *Proc. Int. Conf. Pattern Recog.*, Aug. 2006, Hong-Kong, pp. 826–829.
[25] T. Bonfort and P. Sturm, "Voxel carving for specular surfaces," in *Proc. Int. Conf. Comput. Vis.*, 2003, pp. 591–596.
[26] D. Anderson, R. H. Luke, M. Skubic, J. M. Keller, M. Rantz, and M. Aud, "Evaluation of a video based fall recognition system for elders using voxel space," presented at the 6th Int. Conf. Int. Soc. Gerontechnology, Pisa, Italy, Jun. 4–6, 2008.
[27] A. W. Fitzgibbon and R. B. Fischer, "A buyer's guide to conic fitting," in *Proc. Brit. Mach. Vis. Conf.*, Birmingham, U.K., 1995, pp. 265–271.
[28] T. Banerjee, "Activity segmentation with special emphasis on sit-to-stand analysis" Master's Thesis, University of Missouri—Columbia, Columbia, MO, USA, 2010.
[29] R. Halir and J. Flusser, "Numerically stable direct least squares fitting of ellipses," in *Proc. 6th Int. Conf. Comput. Graph. Vis.*, 1998, vol. 1, pp. 125–132.
[30] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–697, Nov. 1986.
[31] T. Banerjee, J. M. Keller, M. Skubic, and C. C. Abbott, "Sit-to-stand detection using fuzzy clustering techniques," *IEEE World Congress Comput. Intell.*, Barcelona, Spain, Jul. 18–23, 2010.
[32] T. Banerjee, J. M. Keller, M. Skubic, and C. C. Abbott, "Sit-To-stand detection using fuzzy clustering techniques," in *Proc. IEEE World Congr. Comput. Intell.*, Barcelona, Spain, Jul. 18–23, 2010.

Tanvi Banerjee (S'10) received the B.S. degree in electronics and telecommunications engineering from Pune University, Pune, India, in 2007 and the M.S. degree in electrical and computer engineering from the University of Missouri, Columbia, MO, USA, in 2010, where she is currently working toward the Ph.D. degree in electrical and computer engineering.

She is a Research Assistant and a Teaching Fellow in the Center for Eldercare and Rehabilitation Technology, University of Missouri. Her research interests include computer vision, and machine learning, and fuzzy logic.

Marjorie Skubic (S'90–M'91) received the Ph.D. degree in computer science from Texas A&M University, College Station, TX, USA, in 1997, where she specialized in distributed telerobotics and robot programming by demonstration.

She is currently a Professor in the Electrical and Computer Engineering Department, University of Missouri, Columbia, MO, USA, with a joint appointment in computer science. In addition to her academic experience, she has spent 14 years working in industry on real-time applications such as data acquisition and automation. Her current research interests include sensory perception, computational intelligence, spatial referencing interfaces, humanrobot interaction, and sensor networks for eldercare. In 2006, she established the Center for Eldercare and Rehabilitation Technology at the University of Missouri and serves as the Center Director for this interdisciplinary team.

James M. Keller (F'00) received the Ph.D. degree in Mathematics in 1978.

He holds the University of Missouri Curators' Professorship in the Electrical and Computer Engineering and Computer Science Departments, Columbia, MO, USA. He is also the R. L. Tatum Professor in the College of Engineering. His research interests include computational intelligence: fuzzy set theory and fuzzy logic, neural networks, and evolutionary computation with a focus on problems in computer vision, pattern recognition, and information fusion including bioinformatics, spatial reasoning in robotics, geospatial intelligence, sensor and information analysis in technology for eldercare, and landmine detection. His industrial and government funding sources include the Electronics and Space Corporation, Union Electric, Geo-Centers, National Science Foundation, the Administration on Aging, The National Institutes of Health, NASA/JSC, the Air Force Office of Scientific Research, the Army Research Office, the Office of Naval Research, the National Geospatial Intelligence Agency, the Leonard Wood Institute, and the Army Night Vision and Electronic Sensors Directorate. He has coauthored more than 400 technical publications.

Dr. Keller is a Fellow of the International Fuzzy Systems Association (IFSA), and a past President of the North American Fuzzy Information Processing Society (NAFIPS). He received the 2007 Fuzzy Systems Pioneer Award and the 2010 Meritorious Service Award from the IEEE Computational Intelligence Society. He finished a full six year term as Editor-in-Chief of the IEEE TRANSACTIONS ON FUZZY SYSTEMS, followed by being the Vice President for Publications of the IEEE Computational Intelligence Society from 2005 to 2008, and since then, an elected CIS Adcom member. He is the IEEE TAB Transactions Chair and a Member of the IEEE Publication Review and Advisory Committee. Among many conference duties over the years, he was the General Chair of the 1991 NAFIPS Workshop and the 2003 IEEE International Conference on Fuzzy Systems.

Carmen Abbott received the Ph.D degree in educational psychology, health education & wellness promotion from the University of Missouri, Columbia, MO, USA, in 2009.

She is currently a Clinical Professor in the Department of Physical Therapy, School of Health Professions, University of Missouri.

Dr. Abbott is a Member of the American Physical Therapy Association, and the Missouri Falls Free Coalition.