Sit to Stand Detection and Analysis

Sonya Allin and Alex Mihailidis

Department of Occupational Therapy
University of Toronto
s.allin@utoronto.ca and alex.mihailidis@utoronto.ca

Abstract

The ability to rise from a chair is a prerequisite for independent ambulation. Difficulty rising from a chair, moreover, is an indicator of balance deficits likely to result in a fall. In this paper we present preliminary work to affordably detect sit to stand strategies associated with balance impairment using web cameras. The long term goal is to create systems that can monitor functional movements that are common at home in a way that reflects changes in stability and impairment.

Introduction

More than 2 million people over the age of 65 experience difficulty rising from a chair [1]. The ability to rise from a chair, however, is vital to functional mobility in the home. In addition, difficulty rising from a chair is associated with an increased likelihood of sustaining a fall [2], and this likelihood increases as people age.

In our research, we seek to affordably measure changes in elders' ability to perform activities like rise from a chair. We want to do this to: (1) enable early detection of mobility problems associated with instability, (2) alert therapists when there is a need for adjusted seating arrangements in the home, and (3) detect when an individual adopts a movement strategy that may inhibit his or her rehabilitation.

In this paper we present preliminary work to automatically and unobtrusively detect sit to stand (STS) strategies known to reflect functional disability. Our long term goal is to create systems that can continuously monitor behaviors like sit-to-stands in locations where such behaviors most routinely take place. To facilitate monitoring at home, we focus on technologies that are affordable, that can operate without markers, and can perform quickly and robustly. We are currently using low cost commercial webcams to identity STS strategies that reflect instability as well as functional change.

Defining Sit to Stand Strategies

Strategies known to impact the ability to rise from a chair include foot positioning, movement of the torso, and swinging or pushing with the arms against the chair while rising. Some of these strategies are encoded into popular assessments of balance like the Berg Balance Scale [3]; detecting changes in the STS strategy, then, directly reflects changes in clinically defined balance health.

Here, we present preliminary work to detect the following STS strategies:

1. **Use of the hands or arms while rising.** Prior research has revealed a hierarchy of arm related movement strategies that are commonly used as STS difficulty increases and/or functional ability decreases [4]. While severely disabled individuals cannot rise from a chair, those at the lower end of the functional spectrum may use hands to push from the seat. Those who are more functional may swing the arms while rising or not require arms for support.
2. **Positioning of the feet while rising. Many** elderly subjects are inclined to place their feet forward while rising from a chair; a major reason for this is may be to alleviate discomfort experienced as a result of arthritis [5]. Far foot placement, however, requires increased momentum at the hip while rising, which, in turn, creates a potential for instability [6]. Moving the center of mass over the feet before leaving the seat, by contrast, reduces the difficulty of transitions [7]; this strategy is commonly adopted as well.

Measuring Sit to Stand Strategies

In our current implementations, detecting an STS strategy with web cameras involves extracting features from input images and relating them to canonical poses or motions that represent the strategy. The image features we are currently exploring are two dimensional and three dimensional in nature. In two dimensional images, we are looking at features of foreground silhouettes. Silhouettes have been used extensively to recognize canonical poses or actions; specific features of silhouettes commonly used for pose detection include Hu Moments [8], Feret diameters [9] and wavelet responses [10].

We are also, however, exploring three dimensional features. In our application, three dimensional reconstructions are made by triangulating features of silhouettes, like centroids and extrema, across multiple synchronized camera views. Similar blob based stereo reconstructions have been used for sign language recognition [11] and gesture recognition in interactive games [12]. Three dimensional reconstructions have an advantage over 2D ones in that they more directly relate to kinematic and dynamic analyses of STS stability, like [6].

Using Web Cameras to Predict Strategies

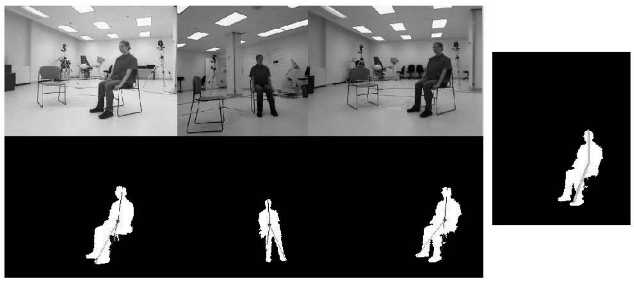

Figure 1. STS recognition system. Three calibrated Logitech webcams are used to capture synchronized video of individuals as they rise from a chair. At the top are input images from all cameras; at the bottom are extracted silhouettes annotated with 2D estimates of the head, feet and torso locations. 2D features are triangulated across views to create 3D reconstructions of the torso and lower body. An example reconstruction is shown at the far right; this reconstruction has been back projected on top of a silhouette from one view.

Our system includes three Logitech web cameras that record at a rate of 30 frames per second and are synchronized by means of a DirectShow filter. Video is encoded in MPEG4 format, at a resolution of 340 x 280 pixels, and is transferred to a laptop via USB.

To compute two and three dimensional image features, we first extract silhouettes of people from two dimensional views by means of foreground segmentation. From these two dimensional silhouettes, Hu moments [7] are computed, as these are shape descriptors that are robust to changes in scale and translation.

To create three dimensional reconstructions, we triangulate silhouette centroids across all camera views. This creates a single three dimensional point corresponding loosely to the position of a tracked individual's torso. We also triangulate points on contours of silhouettes that are maximally distant from centroids to estimate head and foot positions. Finally, we smooth estimates of all 3D positions with a Kalman filter. This is illustrated in Figure 1.

Methods

To determine the utility of two and three dimensional measurements made with web cameras, we performed a pilot experiment; in this experiment image statistics were related to various compensatory STS strategies.

Two healthy individuals were asked to sit and stand repeatedly from a pair of chairs while being recorded with the cameras. The two chairs were located two feet from one another and were oriented to form a 90 degree angle between them. Two chairs were used to test the ability of the system to detect sit to stands and STS strategies at various orientations and locations with respect to the cameras.

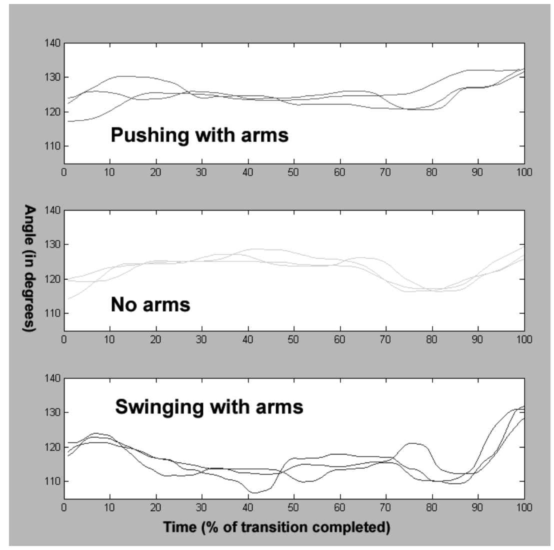

Figure 2. Estimated torso flexion (in degrees) during sit to stand transitions. At the top is estimated flexion when a subject rises repeatedly using his or her arms. In the middle is flexion when no arms are used and at the bottom is flexion when arms are swung. The time of all transitions have been normalized so they span 100 samples each.

Each trial required subjects to sit first on one chair, then the other. Each time a subject sat on a chair, however, he or she was asked to use one of a set of eight possible STS strategies defined by use of the arms or placement of the feet. Recorded arm strategies were as follows:

(1) no use of the arms while rising;
(2) use of the arms to push from the seat;
(3) swinging of the arms to generate momentum.

Foot strategies, by contrast, were as follows:

(1) with knees "normally" extended, at 90° angles;
(2) with knees under-extended, at 80° angles;
(3) with knees over-extended, at 100° angles.

All combinations of hand and arm positions were recorded save for the combination involving over-extended knees and no arm use; this was difficult to perform, even for young subjects. Three instances of each strategy were recorded and the order was randomized, yielding a total of 24 trails (and 48 STS actions) per subject.

Based the data, classifiers were designed to create robust mappings between image statistics recorded at each frame and phases of sit-to-stands (i.e. sitting, transitioning, or standing). In addition, images statistics measured at each frame during hand segmented STS transitions were mapped onto the eight STS strategies. Image statistics used to form mappings included 7 Hu moments and the following 3D features: distances between the torso, head and feet; distances between the torso, head and floor; the angle created by the head, torso and feet; the speed of the head; the angular speed of the torso; and the raw positions of the feet. Examples of recorded "torso angles" (i.e. angles between head, torso and feet) are shown in Figure 2.

The classification algorithm used was a J48 decision tree, which is an extension of the classic C4.5 algorithm of Quinlan [13] and has been built into the Weka Machine Learning Toolkit [14]. To build the decision tree, input video data was first hand labeled to indicate STS phases (i.e. when subjects were sitting in the chair, when they were transitioning to stand and when they had risen). In addition, data recorded during STS transitions was labeled to identify which STS strategy was used.

Using 10 fold cross validation (90% training, 10% testing each time), the ability of image statistics to predict phases and sit to stand strategies at every frame was tested. We present results from one classifier built for both subjects, although reasonable results were obtained by training classifiers with one subject and testing on the other. Preliminary results are found below:

Table 1. Results using 2D features alone, from 1 view

Foot Posn	80 degrees	90 degrees	110 degrees
80 degrees	1367	434	112
90 degrees	509	1060	197
110 degrees	189	291	707

Total accuracy: 64%

Arm Usage	No arms	Push seat	Swing arms
No arms	784	296	223
Push seat	250	1321	248
Swing arms	221	340	1183

Total accuracy: 67%

STS Phase	Sitting	Transitioning	Upright
Sitting	8331	472	976
Transitioning	337	7440	369
Upright	1488	775	2603

Total accuracy: 81%

Table 2. Results using 3D features

Foot Posn	80 degrees	90 degrees	110 degrees
80 degrees	1817	79	17
90 degrees	89	1646	31
110 degrees	16	36	1135

Total accuracy: 94%

Arm Usage	No arms	Push seat	Swing arms
No arms	1136	100	67
Push seat	92	1636	91
Swing arms	79	95	1570

Total accuracy: 89%

STS Phase	Sitting	Transitioning	Upright
Sitting	9258	222	299
Transitioning	196	7626	324
Upright	360	388	4118

Total accuracy: 92%

Conclusion

We have presented preliminary data indicating that 2D features from silhouettes can reasonably be used to detect STS phases (i.e. Sits, Transitions, Stands) that take place at multiple orientations; coarse 3D reconstructions, moreover, predict STS strategies associated with disability. We expect results can improve significantly with temporal smoothing of classifications. Admittedly, results are limited as they relate to a small number of healthy individuals who are similar in size. In the future, we expect to deploy functional monitoring devices in the homes of elders to look for changes in behaviors, like sit to stands, which reflect instability or have the potential to adversely impact rehabilitation outcomes.

References

1. Dawson D, Hendershot G, Fulton J, 1987. *Aging in the Eighties. Functional Limitations of Individuals Age 65 and Over.* National Center for Health Statistics: Advance Data from Vital and Health Statistics, DHHS (PHS) 133: 87-1250.
2. Topper AK, Maki BE, Holliday PJ. 1993. *Are activity-based assessments of balance and gait in the elderly predictive of risk of falling and/or type of fall?* J Am Geriatr Soc 41: 479–87.
3. Berg K, Wood-Dauphinee S, Williams JI, Maki, B. 1992. *Measuring balance in the elderly: Validation of an instrument.* Can. J. Pub. Health 2:S7-11 .
4. Mazzà C, Benvenuti F, Bimbi C and Stanhope, S. 2004. *Association Between Subject Functional Status, Seat Height, and Movement Strategy in Sit-to-Stand Performance*. J Am Geriatr Soc 52(10): 1750-1754.
5. Pai, YC., Chang, H. J., Sinacore, J. M., Chang, R. W., Lewis, JL. 1994. *Alteration in multijoint dynamics in patients with bilateral knee osteoarthritis.* Arth & Rheum 37: 1297-1304.
6. Patton, JL., Lee, W. A., Pai, Y. C. 2000. *Relative stability improves with experience in a dynamic standing task.* Exp Brain Res, 135: 117-26.
7. Kawagoe, S, Tajima, N and Chosa, E. 2000. *Biomechanical analysis of effects of foot placement with varying chair height on the motion of standing up.* Journal of Orthopaedic Science 5(2): 124-133.
8. Hu, M. 1962. *Visual pattern recognition by moment invariants.* IEEE Trans Info Theory 8: 179–187.
9. Lee, T. and Mihailidis, A. 2005. *An intelligent emergency response system: preliminary development and testing of automated fall detection.* J Telemedicine and Telecare, 11:194-198.
10. Ren L, Sharknarovich G, Hodgins J, Pfister, H and Viola P. 2005. *Learning Silhouette Features for Control of Human Motion*, ACM Transactions on Graphics.
11. Starner, T. *Visual Recognition of American Sign Language Using Hidden Markov Models.* 1995. Automatic Face and Gesture Recognition.
12. Wren, C., Clarkson, B., Pentland, A. 2000. *Understanding Purposeful Human Motion.* Automatic Face and Gesture Recognition.
13. Quinlan, J. 1996. *Improved use of continuous attributes in c4.5.* J of Art Intel Res, 4:77-90.
14. Witten, I and Frank, E. 2005. *Data Mining: Practical machine learning tools and techniques*, 2nd Ed, Morgan Kaufmann, SF
15. Talis VL, Grishin AA, Solopova IA, Oskanyan TL, Belenky VE, Ivanenko YP. 2008. *Asymmetric leg loading during sit-to-stand, walking and quiet standing in patients after unilateral total hip replacement surgery.* Clin Biomech 23(4):424-33.