Contour Tracking of Human Exercises

Timothy C. Havens, Gregory L. Alexander, Carmen Abbott,
James M. Keller, Marjorie Skubic, and Marilyn Rantz

***Abstract*—We developed a novel markerless motion capture system and explored its use in documenting elder exercise routines in a health club. This system uses image contour tracking and swarm intelligence methods to track the location of the spine and shoulders during three exercises — treadmill, exercise bike, and overhead lateral pull-down. Preliminary results of our qualitative study demonstrate that our system is capable of providing important feedback about the posture and stability of elders while they are performing exercises. Study participants indicated that feedback from this system would add value to their exercise routines.**

I. Introduction

EVERYONE can benefit from some type of exercise including, and perhaps most importantly, older adults. However, it has been shown that only 30% of older adults aged 65 and older include a regular exercise routine in their daily activity [1]. Sedentary elders who begin an exercise program ultimately benefit from improved quality of life and reduced health care expenditures [2]. Additional benefits of a daily exercise routine for elders include prevention of falls, alleviation of depression, improved cognitive function, improved bone density, improved cardiovascular function — the list of benefits is virtually endless [2–7]. In [7], the authors discovered that exercise is an under prescribed therapeutic intervention due to misconceptions by elders, their caregivers, and their health care providers about exercise safety. For the reasons stated above, improvements in exercise safety for older adults could have a significant impact on the overall health of elders.

To address the issues described above, we conducted a pilot study that examined the human factors issues of a novel technology interface designed to capture range of motion and provide feedback to elderly people using exercise equipment. Human factors is defined as the study of how humans accomplish work related tasks in the context of human-machine system operation and how behavioral and non-behavioral variables affect that accomplishment [8, 9]. In health care, human factors researchers attempt to understand the interrelationships between humans and the tools they use, the environments in which they live and work, and the tasks they perform [10–12]. Thus, the goal of a human factors approach is to optimize the interactions between technology and the human in order to minimize human error and maximize human-system efficiency, human well-being, and quality of life [6]. This paper describes the technology used in our human factors research. This technology is a novel exercise-feedback computer interface that combines image segmentation, contour tracking, motion capture, and swarm intelligence.

Section II introduces previous work in using contour tracking for human motion analysis. Section III gives a detailed description of our system while Section IV describes the results of using our system with older adults. We summarize in Section V.

II. Related Work

Human motion analysis is a well researched topic and is pertinent to many fields including sports medicine, nursing, physical therapy and rehabilitation, and surveillance. There have been special issues of journals and tracks in computer-vision conferences completely dedicated to human motion analysis in video.

References [13–16] provide a good background on human motion analysis techniques. Reference [17] uses silhouette-based features to recognize falls in monocular video of elders in the home environment. In [18], the author proposes a method to analyze human pose during exercise. This method has the strength that it is generalizable to any pose; however, as the author points out, it is very error prone. Also, the assumption is made that the background subtraction (silhouette extraction) is near ideal. Achieving an ideal silhouette in a gym environment is virtually impossible as we show throughout this paper. Another assumption that is made in [18] is that the subject is facing the camera and upright. We wish to measure the angle of the spine, as seen from the side view in both upright (treadmill) and sitting (overhead pull-down) poses; hence, this assumption makes this method undesirable for use in our research.

Active contours, called snakes [19], have been used to track face features (e.g. eyebrows and mouth) and humans in video. Although this method is effective for tracking the features of faces and humans in video, it does not give us the capability of measuring the posture information we require. For instance, measuring the angle of the spine with respect to

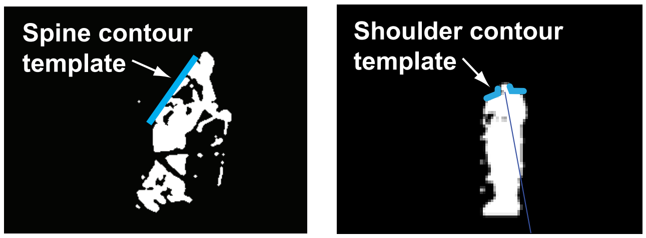
This work was supported by the Center for Eldercare and Rehabilitation Technology, Building Interdisciplinary Geriatric Health Care Research Centers Initiative, RAND/Hartford Foundation, PI: Marilyn Rantz (Award #: 9920070003), by the National Science Foundation ITR Technology Interventions for Elders with Mobility and Cognitive Impairments project, PI: Marjorie Skubic (Award #: IIS-0428420), and the Agency for Healthcare Research and Quality, PI: Gregory Alexander (Award #: K08HS016862). The content is soley the responsibility of the authors and does not necessarily represent the official views of the granting agencies.

T.C. Havens, J.M. Keller, and M. Skubic are with the Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211, USA (email: havenst@gmail.com, {kellerj, skubicm}@missouri.edu).

G.L. Alexander and M. Rantz are with the Sinclair School of Nursing, University of Missouri, Columbia, MO 65211, USA (email: {alexanderg, rantzm}@missouri.edu)

C. Abbott is with the Department of Physical Therapy, University of Missouri, Columbia, MO 65211, USA (email: abbottc@missouri.edu)

978-1-4244-2775-8/09/$25.00 ©2009 IEEE



(a) Spine contour tracking (b) Shoulder contour tracking

Fig. 1: Examples show spine and shoulder tracking of treadmill exercise with contour templates shown in blue.

the horizontal with an active contour is not straight forward as active contours are not rigid, static features. However, one of the features of the rigid, static contour templates in our method is the angle with respect to the horizontal. Hence, we use a contour tracking method based on the edge distance transform [20] with rigid, static templates.

III. Exercise-Feedback System

Our method tracks body contours in the video of exercising humans. The two contours we are interested in are the edge of the back (spine) as seen from the side view and the shoulders as seen from the rear or front view. Fig. 1 shows these two contours on example video frames of a research participant walking on a treadmill. Fig. 2 illustrates our approach in a block diagram. We designed our approach to be both robust and flexible.

The environment in which we are performing this research study is a public gym; hence, our ability to control experimental conditions, such as lighting conditions, background environment, and subject clothing, is very limited. As a result, we chose simple, safe, and proven methods to perform the operations in our algorithm.

First, the silhouette of the human in each video frame was computed. We used a statistics-based background subtraction algorithm that is adapted from [21]. Second, the chamfer distance transform of each silhouette frame was computed, as in [22]. The chamfer distance transform provides an error surface upon which we can fit a contour template. We used *Roach Infestation Optimization* (RIO) [23] to find the best position of the contour template, which, ideally, is located on the body contour of interest, either the back or spine. RIO is inspired by the *Particle Swarm Optimization* (PSO) [24] algorithm. In [23] we showed that RIO is superior to PSO in finding the global optima of highly modal objective functions. RIO works well for the problem we present in this paper.

The best position of the contour template is defined by a temporal fitness function that accounts for exercise dynamics and template translation and rotation. We now describe in more detail each element shown in the block diagram in Fig. 2.

A. Human Silhouette Extraction

Silhouette extraction or background subtraction is a problem that is very pertinent to many fields of research, such

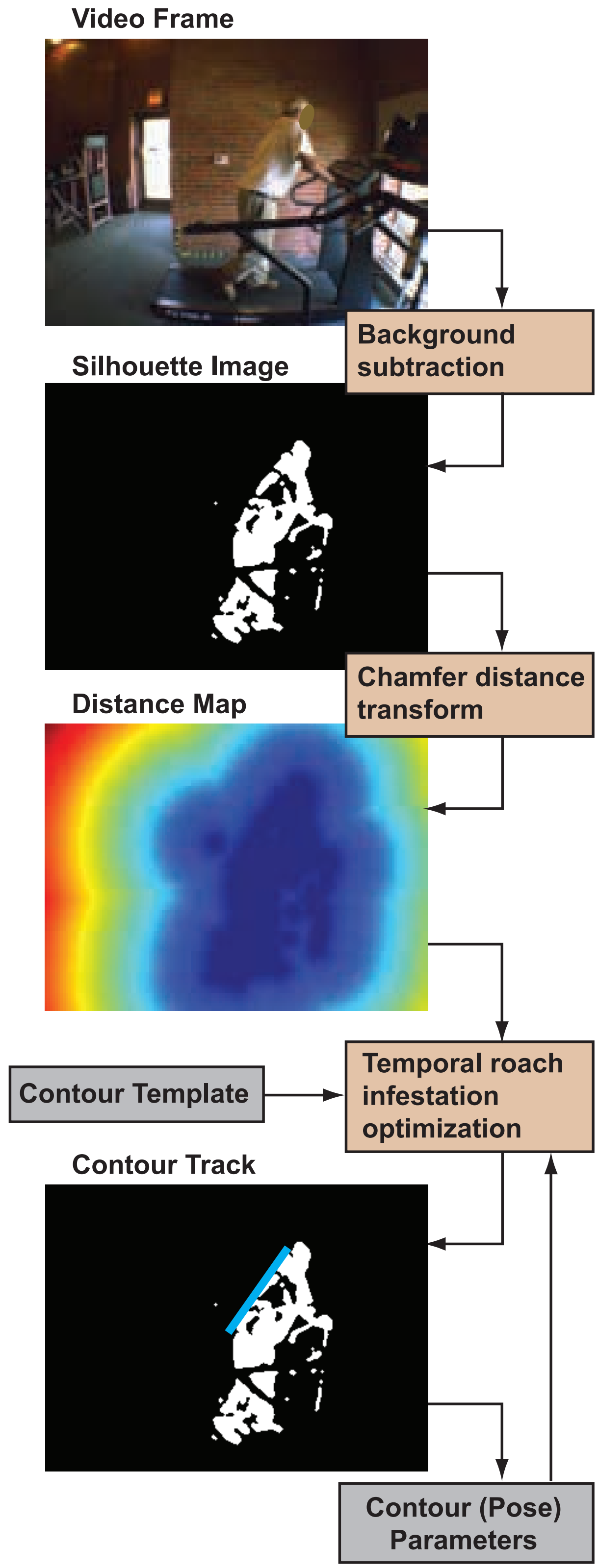
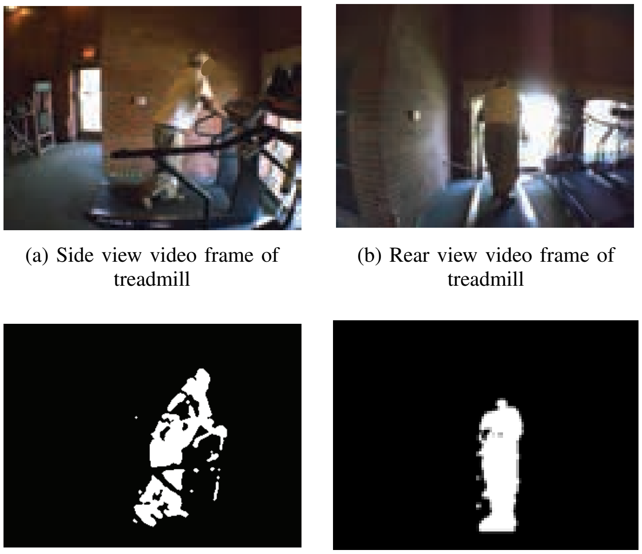


Fig. 2: Block diagram of exercise-feedback system components — spine tracking on side view of treadmill exercise.



(a) Side view video frame of treadmill

(b) Rear view video frame of treadmill

(c) Side view silhouette image of treadmill

(d) Rear view silhouette image of treadmill

Fig. 3: Silhouette extraction examples of treadmill exercise.

as surveillance, activity recognition, and computer vision. However, this problem has many difficult facets including dynamic lightings conditions and backgrounds, poor scene illumination, inferior cameras, and highly variable foregrounds. It is beyond the scope of this paper to address these matters; however, we emphasize that extracting "good" silhouettes is essential to our algorithm. The silhouette extraction algorithm we use is adapted from [21]. The red-green-blue (RGB) digital image (one video frame) was converted to a hue-saturation-value (HSV) color space. Then a statistical background representation was formed from approximately 100 background video frames (no human is in view). This statistical background representation was used to subtract the background from each video frame of the participant exercising, leaving only the image pixels that correspond to the image of the participant. Figs. 3(a,b) show the silhouettes computed from the corresponding video frames shown in Figs. 3(c,d). We denote the silhouette image of video frame f as $S(i,j,f)$, where $S(i,j,f)=1$ indicates the ith row, jth column pixel is a foreground pixel and $S(i,j,f)=0$ indicates a background pixel.

B. Contour Tracking

We adapt the chamfer distance transform, described in [20, 22], and define

$$C(r,c,f) = \max\left\{ \min_{\{\forall i,\forall j: S(i,j,f)=0\}} \left[(r-i)^2+(c-j)^2\right] \right. \\ \left. \min_{\{\forall i,\forall j: S(i,j,f)=1\}} \left[(r-i)^2+(c-j)^2\right] \right\}. \quad (1)$$

Essentially, Eq.(1) calculates the minimum squared distance between each pixel location and the edge of the human silhouette. We compute Eq.(1) for each pixel in the image and this distance transform map can be used to determine the best location of the contour template. The template error score is

$$M(\vec{x}_f, f) = \sum_{\{(r,c)\in\text{template pixels}\}} C(r,c,f), \quad (2)$$

where $\vec{x}_f = (x_f, y_f, \theta_f)$ are the contour template parameters, and f is the video frame. We then map Eq.(2) onto the interval $[0,1]$ with a simple mapping function of the form

$$M_{[0,1]}(\vec{x}_f, f) = \min\left\{M(\vec{x}_f, f)/1000, 1\right\}. \quad (3)$$

This mapping is performed so that we can use standard fuzzy operators to combine this objective function with a membership function that limits the search space (see Section III-C).

The best contour template location is defined as

$$\vec{x}_{\text{best},f} = \arg\min_{\vec{x}_f} M_{[0,1]}(\vec{x}_f, f). \quad (4)$$

The template is a discrete list of pixel coordinates, which define the template shape. For example, a linear (line) template, such as that used to track the spine, could be defined as

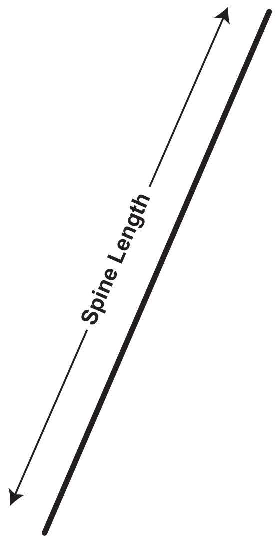
$$\mathbf{T} = \{[0\ 0]^T, [0\ 1]^T, [0\ 2]^T\},$$

where, in this example, T is a vertical line, three pixels long. This template formulation is very general and can represent any types of shapes, including lines, curves, and broken shapes. We use a straight line to model the contour of the spine, see Fig. 4(a), and two sloping lines with attached vertical lines to model the contour of the shoulders, see Fig. 4(b). As Fig. 4 shows, the templates we used for this study are customizable for each participant. However, if one wished to use our technique to track other body contours, a template could easily be designed. Eq.(2) is an error score of the fit of the contour template, for a given parameter vector $\vec{x}_f$, to the edge of the silhouette in frame f. The coordinates (r,c), over which the summation in Eq.(2) is computed, are found by the linear transformation

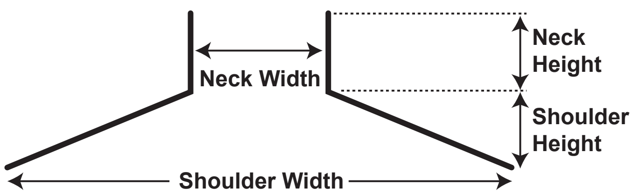
$$\begin{bmatrix} c \\ r \end{bmatrix}_i = \begin{bmatrix} \cos\theta_f & -\sin\theta_f \\ \sin\theta_f & \cos\theta_f \end{bmatrix} [\mathbf{T}_i] + \begin{bmatrix} x_f \\ y_f \end{bmatrix}, \quad (5)$$

where $[\mathbf{T}_i] = [t_c\ t_r]_i^T$ is the coordinate of the ith pixel in the contour template $\mathbf{T}$. In our algorithm we define the center of the template as the origin, but this is arbitrary.

Fig. 5 illustrates the value of the template error scores M and $M_{[0,1]}$ for two examples. Both examples use the same silhouette and contour template, only the parameter vector $\vec{x}_f$ is changed. As Fig. 5(a) shows, the best location of the contour template — on the spine — results in the lower error score. Assuming that the contour template is defined properly and the silhouette image is ideal, the human contours can be tracked in video by solving Eq.(4) for each successive video frame.

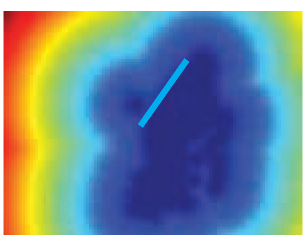


(a) Spine template

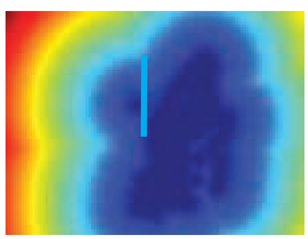


(b) Shoulders template

Fig. 4: Templates used to track (a) spine and (b) shoulders.



(a) Contour parameters $\vec{x}^{(1)} = \{171.6, 100.0, 30.5\}$, Error $M(\vec{x}^{(1)}) = 79.8$, $M_{[0,1]}(\vec{x}^{(1)}) = 0.01$



(b) Contour parameters $\vec{x}^{(2)} = \{150, 100, 0\}$, Error $M(\vec{x}^{(2)}) = 767.8$, $M_{[0,1]}(\vec{x}^{(2)}) = 0.89$

Fig. 5: Example values of contour scoring function Eq.(2) for spine contour tracking of treadmill exercise.

C. Temporal Contour Search

Under ideal circumstances where a "good" silhouette image can be computed, solving Eq.(4) would be sufficient for tracking the contours on the human. However, we conducted this research in a gym-environment; hence, "good" silhouettes were not always achieved. We added a temporal term to Eq.(4) that limited the candidate contour locations $\vec{x}_f$ to those that only changed slightly from the previous frame. In other words, because we are tracking human motion, we can assume that the spine or shoulder contours only move a small amount between video frames (video was taken at 7.5 frames-per-second). For each video frame, the error function that must be minimized is the fuzzy union of $M_{[0,1]}$ and R

$$E(\vec{x}_f, \vec{x}^*_{f-1}, f) = \max\{R(\theta_f, \theta^*_{f-1}), M_{[0,1]}(\vec{x}_f, f)\} \quad (6)$$

where θ^*_{f-1} is the previous frame's best rotation parameter solution and $R(\theta_f, \theta^*_{f-1})$ is the membership in "rotated more than expected for one video frame". The temporal damping function is designed such that large changes in the template parameters produce high membership. We use the following formulation for the membership R

$$R(\theta_f, \theta^*_{f-1}) = \begin{cases} 0, & \bar{\theta} \le a \\ 2\left(\frac{\bar{\theta}-a}{b-a}\right)^2, & a \le \bar{\theta} \le \frac{a+b}{2} \\ 1 - 2\left(b - \frac{\bar{\theta}}{b-a}\right)^2, & \frac{a+b}{2} \le \bar{\theta} \le b \\ 1, & b \le \bar{\theta} \end{cases}, \quad (7)$$

where $\bar{\theta} = |\theta_f - \theta^*_{f-1}|$ and, a and b set the inflection points of the spline. Values of the inflection points that we have found effective for our study are $a = 5$ and $b = 10$. In essence, a sets the maximum expected change in rotation between frames, while b sets the point at which candidate solutions are severely punished. Values that we found effective for our study are $a = 5$ degrees and $b = 10$ degrees. Hence, it is clear, by comparing Eqs. (3) and (7), that R will dominate the value of E in Eq.(6) for changes in rotation angle greater than a. E reduces to $M_{[0,1]}$ for $\bar{\theta} \le a$.

RIO [23] is used to optimize E for each video frame in the following way:

1) The roaches are initialized randomly within a predefined bounding box around the contour of interest — the spine or the shoulders — and within a predefined parameter space;
2) The RIO algorithm searches for the best set of parameters $\vec{x}_f$ that minimize E;
3) Advance the video frame and return to step 1.

D. Interface

The interface we developed provides feedback to the study participants. The layout of our interface is shown in Fig. 6. The upper left shows the silhouette extraction result, the upper right shows the tracking contour on the silhouette. The lower left view is an exploded view of the tracking area. This view provides the participant with a more detailed view on how their spine or shoulders look as compared to the contour tracking reference. Finally, the angle of the spine or shoulders is graphed on the lower right. The solid blue line shows the angle at each video frame, while the dotted red line is a running average of the angle. Thus, information on both the movement (solid blue line) and overall posture (dotted red line) is shown on the graph. Section IV discusses the participants' views on the exercise feedback interface.

IV. Results

Our pilot study consisted of qualitative study of key informant interviews of 38 older adult participants aged 65 years and older. We recorded images of each participant doing each of the three exercises and then showed each participant the

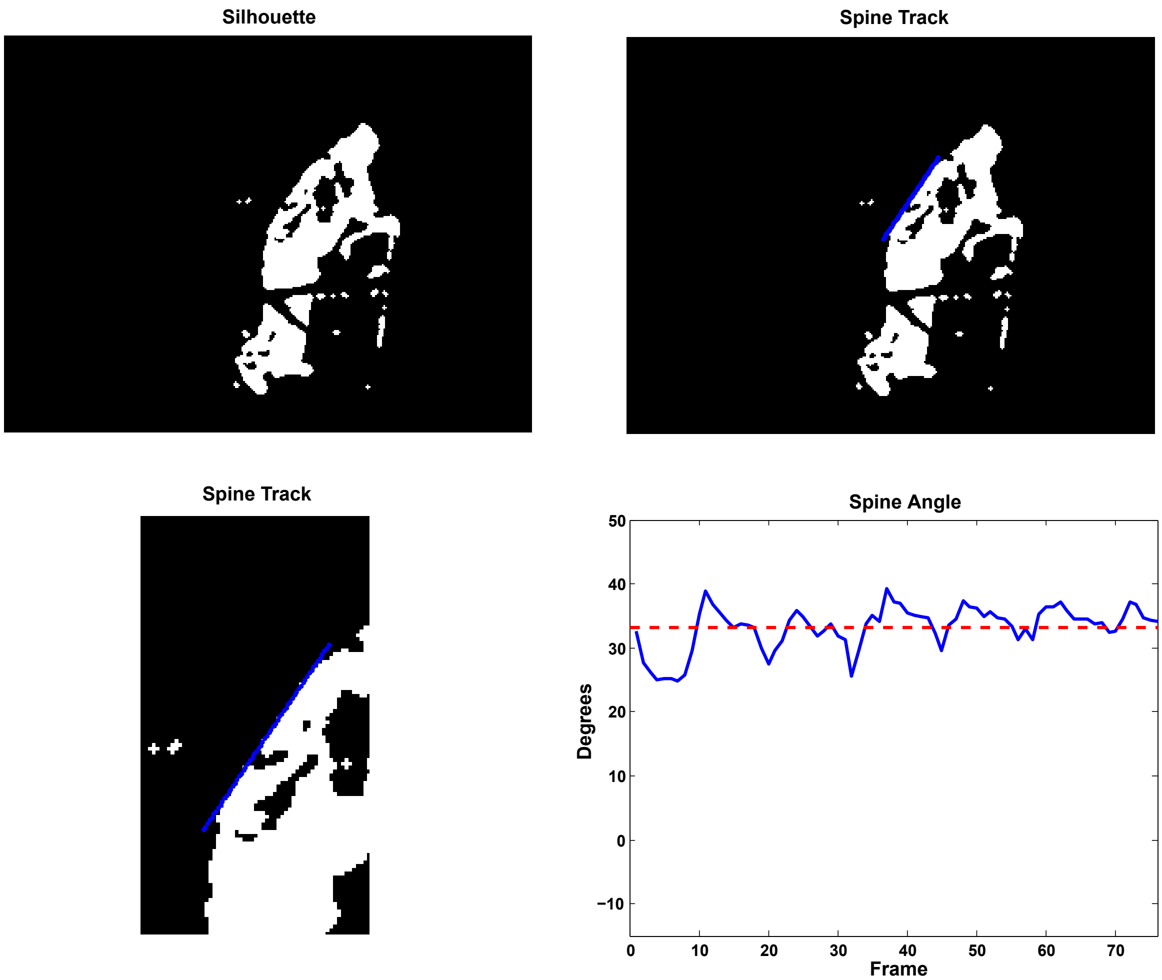


Fig. 6: Layout of contour tracking interface — upper left shows silhouettes, upper right shows tracking result on silhouettes, lower right shows exploded view of tracking area, and lower right is a plot of the angle versus video frame.

results of the exercise-feedback system. Structured key informant interviews were then conducted to gain feedback about how the interface could be developed further to support the participants during their exercise routines. These interviews were recorded and transcribed and phenomena that emerged and reappeared across all interviews and observations were identified. A case study of two participants from our study is described in [25].

In brief, research participants were interested in seeing their images after they performed the exercises. All participants were most interested in how their posture appeared during the period of exercise. Participants expressed that processed images assisted them to visualize how they interacted with the exercise equipment, if they were using good technique to perform the activities, and if they had any unusual movements while performing the desired tasks. For example, one participant stated,

> Well, it seems to me that [the images tells you how to] use the body the way you're supposed to use it to maintain good leg support and arm support. I do sway back and forth, but I don't think you can do anything other than that when your body is moving like it is below the trunk.

Many of the key informants interviewed discussed their fear of losing their balance and falling while walking; they indicated that the images provided them the ability to see if they were maintaining good balance over the core of their body, which is important in preventing falls. Some participants indicated the images would provide some added value to their exercise, making them feel safer, less likely to be injured, and less likely to fall. Others participants indicated that the images were useful but could not really take the place of a trainer who could help interpret what they need to do to be most successful in reaching their exercise goals.

The participant shown in Fig. 3 has issues both with posture, hunching, and gait (a significant limp). In contrast to the participant shown in Fig. 3, the participant shown in Fig. 7 has good mobility and little to no afflictions that affect gait and posture. Figure 8(a) shows the spine angle plot of the participant shown in Fig. 3 and Fig. 8(b) shows the spine angle plot of the participant shown in Fig. 7. As these

(a) Side view video frame of treadmill exercise

(b) Rear view video frame of treadmill exercise

Fig. 7: Sample video frames of participant with good mobility on treadmill.

plots show, the contour tracking method is able to detect not only the overall difference in the participants' postures — as represented by the overall deviation from 0 degrees — but the difference in their gaits are shown by the differences in the patterns shown in the two plots.

V. Conclusions and Future Work

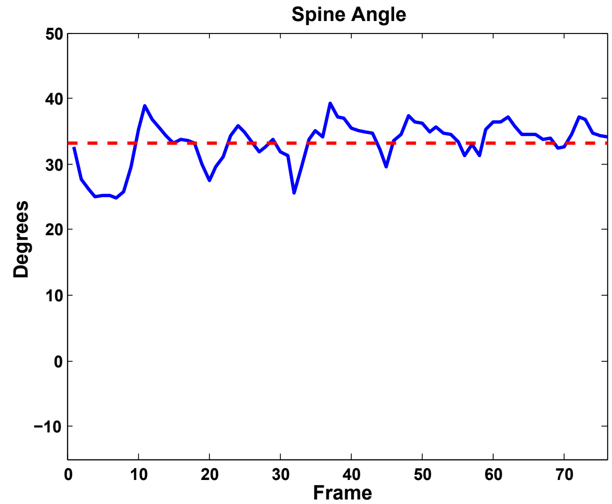
Our exercise feedback interface has broad application in fields where measuring human body movement is important — e.g. physical therapy, sports medicine, and nursing. We applied our methods to eldercare, specifically to improve the safety and effectiveness of exercise. Our study included key informant interviews of 38 older adult participants and our preliminary analysis of a small sample of these interviews indicated that our interface was both effective in showing older adults information on how they move while they exercise and, also, in showing older adults areas in which they could improve their exercise.

The interface is based on a contour tracking method that tracks the movement of the spine and shoulder. Our technique is generalizable, such that health professionals could choose to track other parts of the body that can be represented by a rigid contour template.

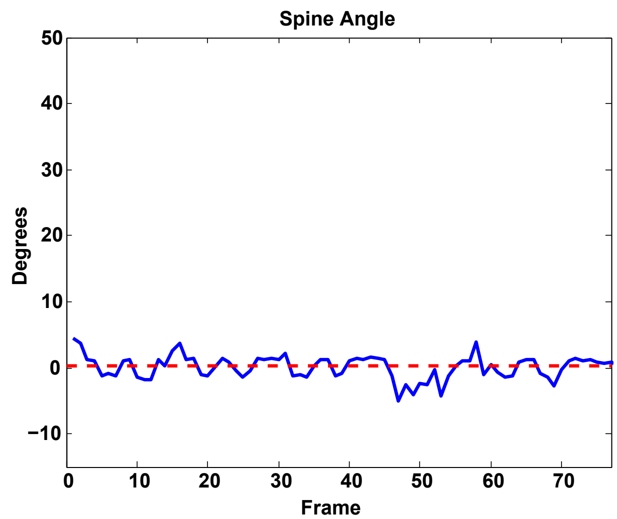
Finally, we are adapting the contour tracking methods for use in a home environment. A large part of our ongoing eldercare research is on the use of technology to help older adults maintain independence. We are investigating the use of silhouette-based techniques to detect falls, assess mobility, and perform activity analysis. These techniques are important to effectively and inexpensively address the needs of our aging population.

Acknowledgment

The authors would like to acknowledge the staff at The Health Connection for allowing us to use their exercise facility for this research.



(a) Spine angle plot of participant with limp



(b) Spine angle plot of participant with good mobility

Fig. 8: Spine angle plots that show how contour tracking captures posture information.

References

[1] "Physical activity and health: a report of the Surgeon General," US Dept of Health and Human Services, Center for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, President's Council on Physical Fitness and Sports, Atlanta, GA and Pittsburgh, PA, Tech. Rep., 1996.

[2] B. Fletcher, M. Gulanick, and L. Braun, "Physical activity and exercise for elders with cardiovascular disease," *Medsurg Nursing*, vol. 14, no. 2, pp. 101–109, 2005.

[3] J. Schnelle, P. MacRae, J. Ouslander, S. Simmons, and M. Nitta, "Functional incidental training (ftt), mobility performance and incontinence care with nursing home residents," *J. Am. Geriatric Soc.*, vol. 43, pp. 1356–1362, 1995.

[4] N. Fisher, D. Pendergast, and E. Calkins, "Muscle reha-

bilitation in impaired elderly nursing home residents," *Arch. Phys. Med. Rehabil.*, vol. 72, pp. 181–185, 1991.

[5] M. McMurdo and L. Rennie, "A controlled trial of exercise by residents of old people's homes," *Age Ageing*, vol. 22, pp. 11–15, 1993.

[6] L. Sauvage Jr., B. Myklebust, and J. Crow-Pan, et al., "A clinical trial of strengthening and aerobic exercise to improve gait and balance in elderly male nursing home residents," *Am. J. Phys. Med. Rehabil.*, vol. 71, pp. 333–342, 1992.

[7] J. Heath and M. Stuart, "Prescribing exercise for frail elders," *J. Am. Board Fam. Prac.*, May-June 2002.

[8] *Conceptual Aspects of Human Factors*. Washington, DC: The John Hopkins University Press, 1989.

[9] G. Salvendy, *The Handbook of Human Factors and Ergonomics*, 2nd ed. New York, NY: John Wiley and Sons, 1997.

[10] "Human factors: Imperative concepts for information systems in critical care," *AACN Clinical Issues*, vol. 14, pp. 310–319, 2003.

[11] M. Weinger, C. Pantiskas, M. Wiklund, and P. Carstensen, "Incorporating human factors into the design of medical devices," *J. Am. Med. Assoc.*, vol. 280, p. 1484, 1998.

[12] N. Staggers, "Human factors: The missing element in computer technology," *Computers in Nursing*, vol. 9, pp. 47–49, 1991.

[13] J. K. Aggarwal and Q. Cai, "Human motion analysis: A review," *Computer Vision and Image Understanding*, vol. 73, pp. 90–102, 1999.

[14] L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis," *Pattern Recognition*, vol. 36, no. 3, pp. 586–601, March 2003.

[15] R. Collins, A. Lipton, H. Fujiyoshi, and T. Kanade, "Algorithms for cooperative multi-sensor surveillance," *Proc. of the IEEE*, vol. 89, no. 10, 2001.

[16] S. Mubarak, "Understanding human behavior from motion imagery," *Machine Vision and Applications*, vol. 14, no. 4, pp. 210–214, September 2003.

[17] D. Anderson, J. Keller, M. Skubic, X. Chen, and Z. He, "Recognizing falls from silhouettes," in *Proc. IEEE EMBS*, New York, NY, August 2006, pp. 6388–6391.

[18] F. Zijlstra, "Silhouette-based human pose analysis for feedback during physical exercises," in *6th Twente Student Conf. on IT*, Netherlands, February 2007.

[19] M. Pardas and E. Sayrol, "A new approach to tracking with active contours," in *Int. Conf. Image Proc.*, vol. 2, Vancouver, BC, September 2000, pp. 259–262.

[20] A. Thayananthan, P. Torr, and R. Cipolla, "Likelihood models for template matching," in *British Machine Vision Conference*, 2004, pp. 949–958.

[21] P. Wildenauer, A. Blauensteiner, and M. Kampel, "Motion detection using an improved colour model," *Advances in Visual Computing*, pp. 607–616, 2006.

[22] A. Rosenfeld and J. Pfaltz, "Distance functions on digital pictures," *Pattern Recognition*, vol. 1, no. 1, pp. 33–61, 1968.

[23] T. Havens, C. Spain, N. Salmon, and J. Keller, "Roach infestation optimization," in *Proc. SIS*, St. Louis, MO, September 2008.

[24] J. Kennedy and R. Eberhardt, "Particle swarm optimization," in *Proceedings of the IEEE Int. Conf. on Neural Networks*, Piscataway, NJ, 1995, pp. 1942–1948.

[25] G. Alexander, T. Havens, M. Skubic, M. Rantz, J. Keller, and C. Abbott, "Markerless human motion capture-based exercise feedback system to increase efficacy and safety of elder exercise routines," *Gerontechnology 2008*, vol. 7, no. 2, p. 67, May 2008.