Markerless Human Motion Capture-Based Exercise Feedback System to Increase Efficacy and Safety of Elder Exercise Routines

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

***Abstract*—We developed a novel markerless motion capture system and explored its use in documenting elder exercise routines in a health club. We present two case studies of elders using a treadmill from our pilot work. Results demonstrate that our system is capable of providing important feedback about the posture and stability of elders while they are performing exercises. Our participants indicated that feedback from this system would add value to their exercise routines.**

I. Introduction

This paper presents a novel exercise feedback system to increase exercise effectiveness and safety for older adults. This system displays both dynamic and summative feedback regarding exercise posture and body position via a custom human-computer interface. We outline our design of the system and present two case studies as examples to the efficacy of this system

II. Background and significance

Older adults make up the fastest growing segment of the American society. Elderly people participate in limited exercise; leading to decreased fitness, deconditioning of muscles, and reduced energy levels [1]. Conversely, regular physical activity is strongly associated with better physical and psychological health outcomes; thus, the promotion of physical activity is recognized as a high public priority [2]. By maintaining or initiating a lifestyle of physical activity including exercise, older persons can improve their cardiovascular, metabolic and skeletal muscle function. Human motion capture and analysis [3] has important implications to help elders understand more about their body posture and motion during exercise routines leading to greater efficacy and safety of exercise regimens.

Human motion analysis is characterized by the capture and evaluation of continual repetition of motion patterns such as walking or cycling. Measured parameters for human motion include stride length, cadence and walking speed; these have been analyzed and could have important implications for the biomechanical examination of hemiparetic patients and patients with hip and knee arthritis [4]. We describe two case studies from a pilot project incorporating a human motion analysis system to examine repetitive motions performed by seniors >65 years old who were performing exercises (treadmill, stationary bicycle, and over lateral pull down machine) at a health club. In this paper we describe only the treadmill results on these two subjects. Processed images were shown to research participants during key informant interviews to explore the usefulness of the images to inform elders about their exercise routines.

III. Methods

A. Recruitment and Selection of Participants

Our subjects were recruited during a local health fair for elders living in the community. The subjects were required to complete a short health questionnaire prior to participating in the study; if any health concerns were identified (e.g. previous heart condition, high blood pressure) they were required to obtain permission from their physician before they could participate. Once the subjects were cleared by their physician, informed consent, approved through the university's IRB, was obtained from each participant.

B. Tracking Contours on Images

Our method tracks body contours in the video of exercising humans. The two contours we are interested in are 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. We did control the speed and incline of the treadmill to limit risk for our elderly subjects; each subject was asked to walk on the treadmill at the lowest setting of 0.8 mph with a 0% incline for approximately 30 seconds while images were taken.

Manuscript received January 28, 2008. This work was supported in part by the Center for Eldercare and Rehabilitation Technology, Building Interdisciplinary Geriatric Health Care Research Centers Initiative. RAND/Hartford Foundation PI: Marilyn Rantz (Award #: 9920070003). Further support was provided by the US Administration on Aging Technology to Enhance Aging in Place project, PI: Marilyn Rantz. (Award #: 90AM3013), the National Science Foundations 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: Greg Alexander (Award #: K08HS016862). The content is solely the responsibility of the authors and does not necessarily represent the official views of the granting agencies.

G. L. Alexander is with the Sinclair School of Nursing, University of Missouri, Columbia MO, USA 65211. Correspondence may be sent to phone: 573-882-9346; e-mail:alexanderg@missouri.edu).

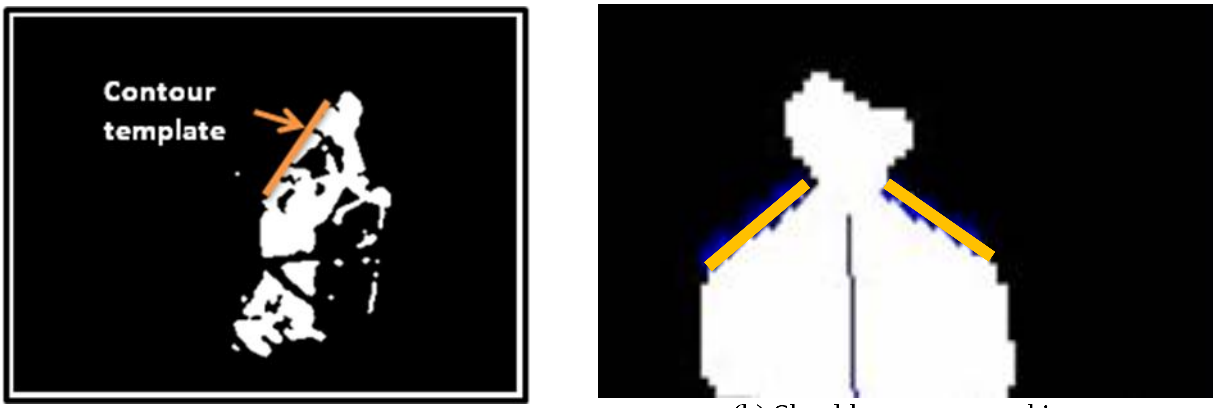
T. C. Havens is with the Department of Electrical and Computer Engineering, University of Missouri, Columbia MO, USA 65211 (e-mail: havenst@gmail.com).

M. Skubic is with the Department of Electrical and Computer Engineering, University of Missouri, Columbia MO USA 65211; (e-mail: skubicm@missouri.edu).

M. Rantz is with the Sinclair School of Nursing, University of Missouri, Columbia MO, USA 65211. (email: rantzm@missouri.edu)

J. M. Keller is with the Department of Electrical and Computer Engineering, University of Missouri, Columbia MO, USA 65211. (email: kellerj@missouri.edu)

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



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

Fig. 1 Contour tracking examples showing (a) spine tracking and (b) shoulder tracking

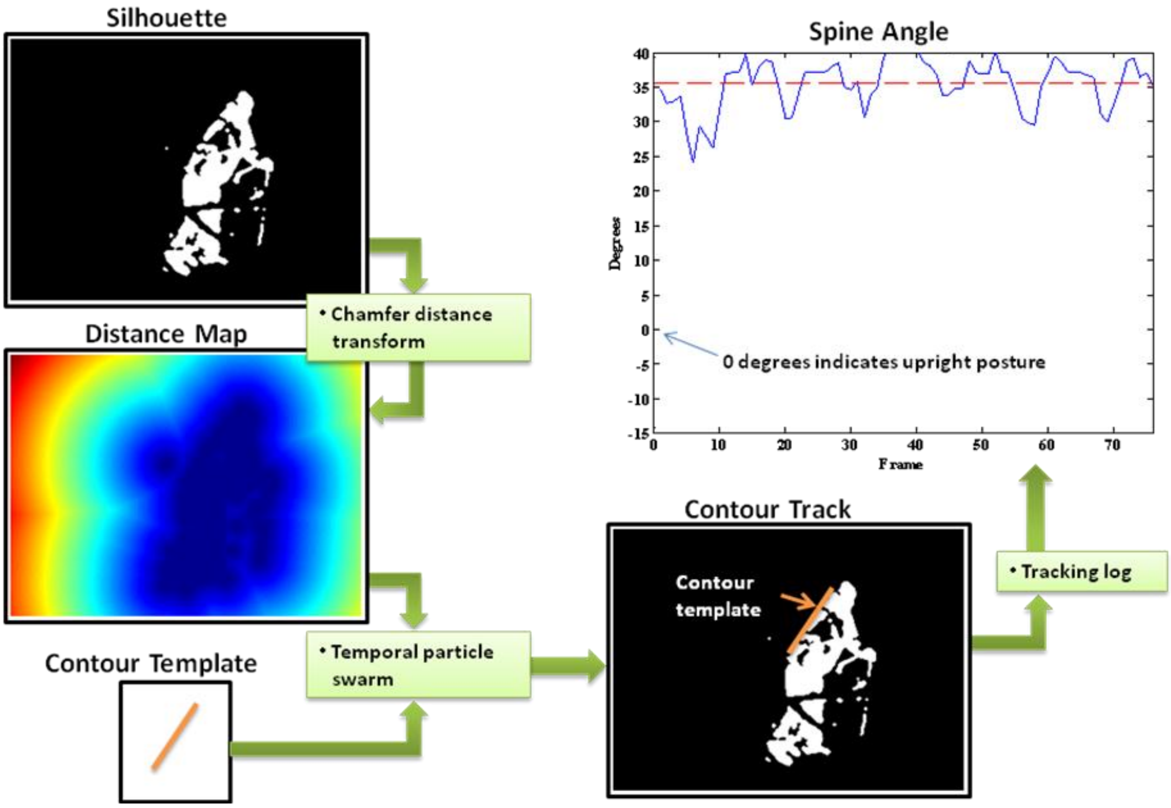
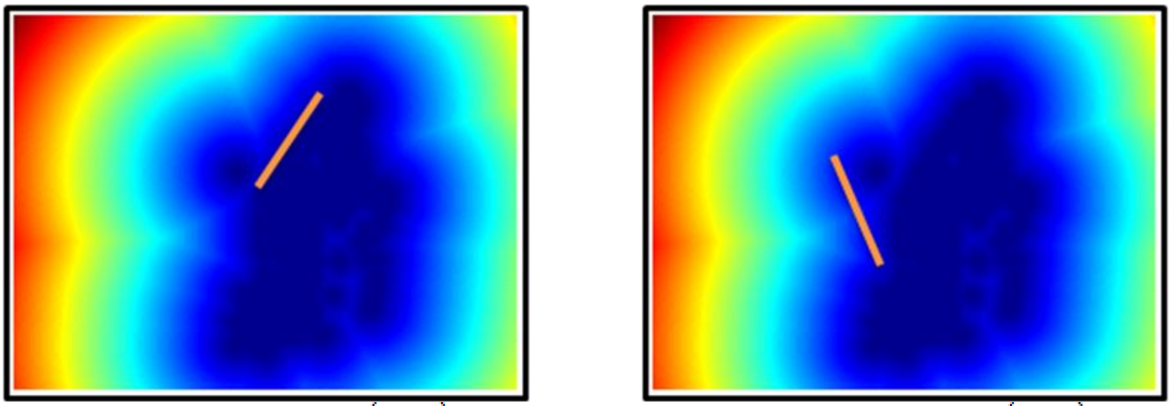


Fig. 2 Body contour tracking system block diagram

First, the silhouette of the human in each video frame was computed. We used a statistics-based background subtraction algorithm that included shadow suppression. Our silhouette extraction approach is adapted from [5]. Second, the chamfer distance transform of each silhouette frame was computed, as in [6]. The chamfer distance transform provides an error surface upon which we can fit a contour template. We used a particle swarm optimization algorithm, as in [7], to find the best position of the contour template, which, ideally, is located on the body contour of interest, either the back or spine. 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.

C. *Human Silhouette Extraction*

Silhouette extraction or background subtraction is a problem that is very pertinent to many fields of research, such 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 [5]. 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



(a)Contour match score $M(X_f,f) = 22$ (b) Contour match score $M(X_f,f) = 187$

Fig. 3 Example illustrates results of chamfer distance-based contour matching for two candidate contour template locations – candidate (a) is located on spine and has a better score than candidate (b) [note: the silhouette region is masked in these images]

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 pixels that correspond to the image of the participant. The algorithm in [5] also has a statistics-based shadow detection method that reduces the effects of shadows on detection of human silhouettes. Figs. 4(b,e) show the silhouettes computed from the corresponding video frames shown in Figs. 4(a,d). We denote the silhouette image of video frame f as $S(i,j,f)$, where $S(i,j,f) = 1$ indicates a foreground pixel and $S(i,j,f) = 0$ indicates a background pixel.

D. Chamfer Distance Transform

We adapt the standard chamfer distance transform and define

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

For each pixel in the image, this distance transform calculates the minimum squared distance to the edge of the silhouette S. We can now use C to determine the best location for the contour template by summing the pixels of C that correspond to the contour template pixels,

$$M(X_f,f) = \sum_{\substack{\text{template}\\ \text{pixels}}} C(r,c,f), \quad (2)$$

where $X_f = \{x_f, y_f, \theta_f\}$ is the contour template parameters, and f is the video frame. This summation is an error score of the contour template fit to the edge of the silhouette. Fig. 3 illustrates this score for two example contour template locations in the chamfer distance transform C. As this Fig. shows, the presumably best location of the contour template – on the spine – results in the lower error score. The rows r and columns c of the pixels of C that are summed to compute (2) are found by computing

$$\begin{bmatrix} c \\ r \end{bmatrix}_i = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} t_x \\ t_y \end{bmatrix}_i + \begin{bmatrix} x \\ y \end{bmatrix}, \forall i \quad (3)$$

where $[t_x \quad t_y]_i^T$ is the coordinates of the ith template pixel. The template is defined as i pixel locations on a grid. Hence, the template is very general can model anything from straight lines to non-linear shapes. We use a straight line to model the contour of the spine (see Fig. 1a) and two sloping lines to model the contour of the shoulders (see Fig. 1b). In order to track the spine or shoulder contours, the best location of the contour template must be found (or searched for) in each video frame.

E. Temporal Particle Swarm-Based Contour Search

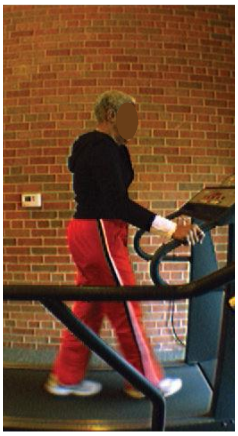
The error function that must be minimized is

$$E(X_f, X_{f-1}^*, f) = R(X_f, X_{f-1}^*) \times M(X_f, f), \quad (4)$$

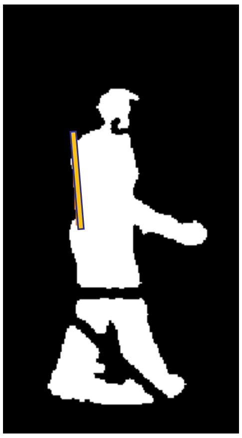
where X_{f-1}^* is the previous frame's parameter solution, $R(X_f, X_{f-1}^*)$ is the temporal damping function, and $M(X_f, f)$ is the contour matching function in (2). The temporal damping function $R(X_f, X_{f-1}^*)$ is designed such that large changes in the contour template parameters $X_f = \{x_f, y_f, \theta_f\}$ produce a higher error value. The contour template parameters are directly related to the movement of the spine or shoulders; hence, during normal movement – such as walking on a treadmill – these parameters should only change slightly between video frames. In the future, we hope to incorporate a more realistic dynamic model. Currently, we compare the current template parameters X_f to the previous frame's solution X_{f-1}^* with

$$R(X_f, X_{f-1}^*) = 1 \Big/ \left[\exp\left(-\frac{|\theta_f - \theta_{f-1}^*|^2}{2\sigma_\theta^2}\right)\right], \quad (5)$$

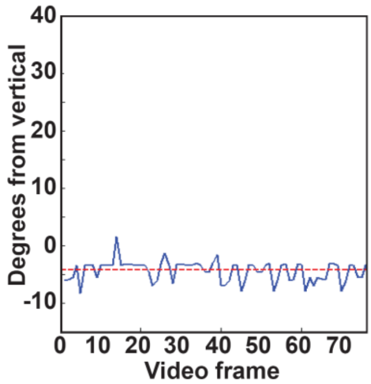
where σ_θ determines the amount of damping as a function of change between the previous angle θ_{f-1}^* and the candidate angle θ_f. We use $\sigma_\theta = 10$ degrees in the examples shown in the paper. The particle swarm algorithm searches for the contour template parameters X_f that minimizes $E(X_f, X_{f-1}^*, f)$. We use a standard particle swarm optimization as described in [7].



(a)Particpant 1 video frame



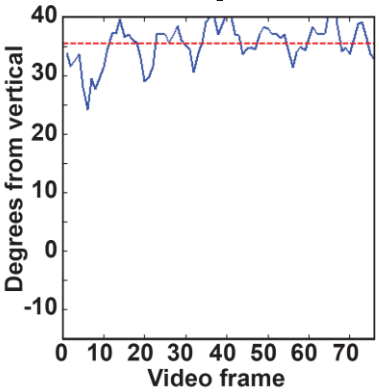
(b) Participant 1 silhouette



(c)Participant 1 spine angle plot

(d) Participant 2 video frame

(e)Participant 2 silhouette



(f) Participant 2 spine angle plot

Fig. 4 Spine tracking comparison of two research participants – tracking shows that Participant 1 has better posture and a smoother gait

F. Examples

Fig. 4 illustrates two examples of the feedback produced by our system. Figs. 4(a,d) show a video frame of two participants while Figs. 4(b,e) display the corresponding silhouette with the spine tracking reference overlaid. Finally, Figs. 4(c,f) are plots of the spine angle (positive angle is leaning forward) versus video frame (at 7.5 frames/second).

The interface layout that is shown to research participants during the key informant interviews is illustrated in Figs. 5 and 6. The top left view is of just the silhouette, the top right view shows the contour template on the silhouette image. The bottom left view in Figs.5 and 6 show a zoomed-in view of the contour template region (in this case, the spine), while the logged angle of the spine is displayed as a line graph in the bottom right view. Also shown on the graph is a running average, denoted by a dashed-red horizontal line. Each view in the interface shown in Figs. 5 and 6 is a synchronized video. Hence, the interface shows both instantaneous feedback – in the form of the silhouette videos – and historical numerical feedback – in the form of the line graph.

IV. Results

The images captured during the treadmill exercise routines of our two research subjects demonstrate the motion detection system was able capture both overall differences in posture and temporal differences between the subjects. Figs. 4(b,e) illustrate the differences in postural attitude between the two subjects. Participant 2 has a much more pronounced forward pitch as he was walking on the treadmill compared to participant 1 who appears more upright with shoulders centered over her torso.

The severity of participant 2's forward pitch during his treadmill activity is also shown graphically in Fig. 4(f) as a 35 degree angle from vertical. It can be seen by the casual observer that participant 2 appears to put a great deal of stress on his upper extremities, bracing himself for each stride of his gait in order to maintain balance. In contrast, participant 1, in Fig. 4(c), has nearly upright posture, measuring nearly vertical (0 degrees). The interface shows she has a slightly negative backward lean, which we attribute to a hooded jacket that she was wearing (see Fig. 4(a)). This clothing impacted the tracking of her back angle slightly.

Our system was also able to capture participant 2's pronounced limp as evidenced by the periodic pattern in Fig. 4(f). His limp was partially a result of bad arthritic conditions in both his right hip and knee. Also, he was wearing a prosthesis on his right shoe to overcome a condition that caused shortening of his right leg, resulting in a pronounced limp while ambulating. As a result of this limp, the baseline gait pattern associated with his walking on the treadmill demonstrated a variable pattern with wide fluctuations as he pitched forward and backward during ambulation. This periodic pattern is consistent with his limp, as confirmed by visual inspection of his video sequences.

Participant 1 exhibited a much more stable ambulation pattern at her baseline. She appeared to be very confident in her stride and gait and she appeared to have little difficulty performing the exercise. This was demonstrated when she let go of the treadmill handles, continuing to walk at the same speed without any deviation from her previous baseline measures.

Posterior images of the subjects during the treadmill activities are shown in Fig. 6. These images demonstrate, through the shoulder contour images, the degree of side to side balance each participant has during ambulation. Again, participant 1 appears to have a very balanced posture from side to side with a slightly lower left shoulder posture compared to the right. She was able to maintain this position fairly well as illustrated by the blue line which remains consistent around the baseline.

Participant 2 has a more noticeable droop in his left shoulder of nearly 10 degrees compared to the right side. His baseline is more inconsistent, demonstrating again his pronounced limp.

Although results are preliminary, our two subjects indicated during key informant interviews that this sort of system might be useful as a reminder to stand up straighter during exercises, to increase or decrease speed if their positioning was not good, and to demonstrate periodic improvement during their exercises. Both subjects stated that the actual silhouette images themselves were more informative than the line graphs; although, both subjects appeared to use both interface elements to assess their exercise activities.

V. Discussion

Through our pilot work using a markerless human motion detection system we have been able to detect and evaluate elders who were performing exercises on a treadmill. We were able to detect significant differences in postural attitude of elders from two different angles (side and back views) while they were using the treadmill in a health club setting.

Some limitations that we experienced while conducting our experiments in the field included the different daily lighting conditions created by large floor to ceiling windows directly in front of treadmills, which were part of the structure of the health fitness complex. We also encountered some difficulties with recording images when someone else was exercising in the background. These sorts of complications were very informative for us as we worked with other subjects and will also provide important clues as to how these systems might be used in real-world settings in the future.

Our two subjects saw added value in using this motion detection system during their exercise routines. The markerless motion capture system has potential use for the clinical as well. Quantitative information can be measured from the kinematic data to use as a baseline measurement for later comparisons throughout a rehabilitation program. The system could also be used to periodically track progress and help with patient education regarding body mechanics that need to be performed in order to prevent further postural abnormalities and consequent adaptations.

VI. Conclusion

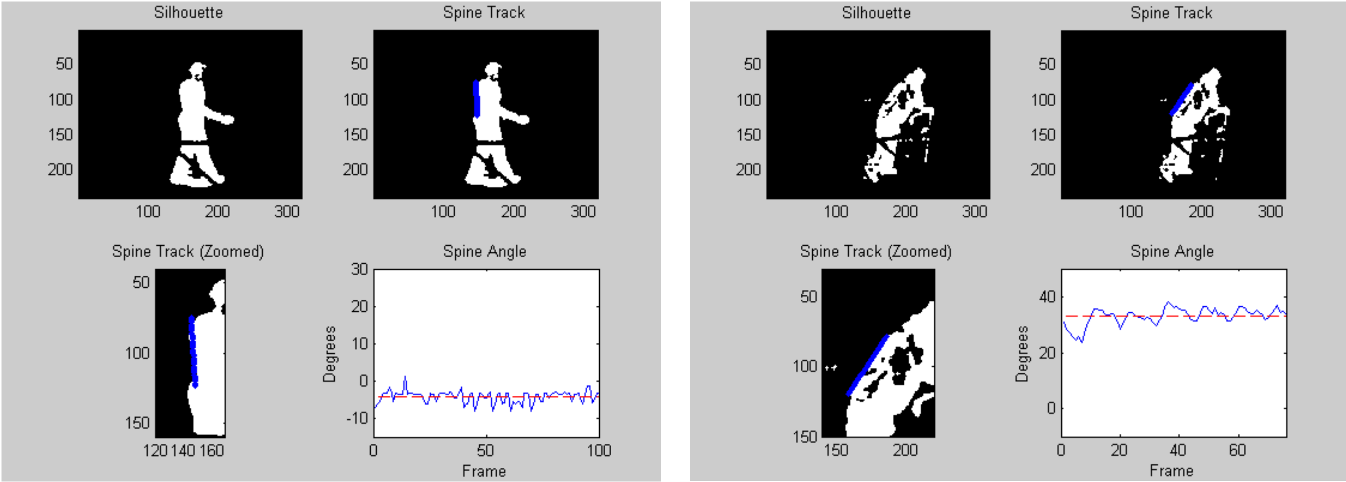
Exercise has important implications for postural balance and strengthening of people who exercise. Continuing active lifestyles that include exercise is extremely important for our increasing elder population to prevent deconditioning of muscles and joints, falls, and other adverse events. New technologies, such as markerless motion detection systems, that evaluate exercise regimens could better provide feedback to elders about their exercise regimens, add value, promote healthier lifestyles, an prevent exercise-related injuries.

Acknowledgement

The authors would like acknowledge the staff at The Health Connection at the University of Missouri School of Health Professions for allowing us to use their exercise facilities for this research and for all their gracious assistance.

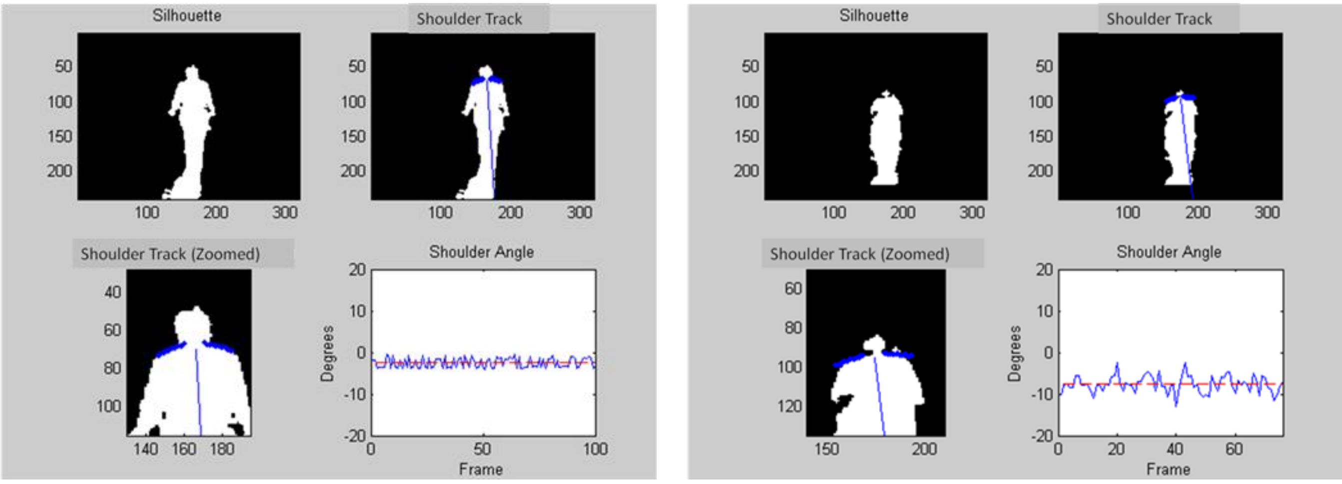
References

[1] Fletcher BJ, Gulanick M, Braun LT. Physical activity and exercise for elders with cardiovascular disease. Medsurg Nursing 2005;14(2):101-9.
[2] Humpel N, Owen N, Leslie E. Environmental factors associated with adults' participation in physical activity: A review. American Journal of Preventative Medicine 2002;22(3):188-99.
[3] Moeslund T, Granum E. A Survey of Computer Vision-Based Human Motion Capture. Computer Vision and Image Understanding 2001;81:231-68.
[4] Paroczai R, Kocsis L. Analysis of human walking and running parameters as a function of speed. Technology and Health Care 2006;14:251-60.
[5] Wildenauer P, Blauensteiner A, Kampel M. Motion detection using and improved colour model. Advances in Visual Computing.Berlin/Heidelberg: Springer; 2006. p. 607-16.
[6] Rosenfeld A, Pfaltz JL. Distance functions on digital pictures. Pattern Recognition 1968;1(1):33-61.
[7] Kennedy J, Eberhart RC, Shi Y. Swarm Intelligence. San Francisco, CA: Morgan Kaufmann; 2001.



(a)Participant 1 (b) Participant 2

Fig. 5 Examples of spine tracking interface layout that is shown to research participants



(a)Participant 1 (b) Participant 2

Fig. 6 Examples of shoulder tracking interface layout that is shown to research participants