Research Article

Toward Monitoring and Increasing Exercise Adherence in Older Adults by Robotic Intervention: A Proof of Concept Study

Prathik Gadde, Hadi Kharrazi, Himalaya Patel, and Karl F. MacDorman

School of Informatics, Indiana University, 535 West Michigan Street, Indianapolis, IN 46202, USA

Correspondence should be addressed to Karl F. MacDorman, kmacdorm@indiana.edu

Socially assistive robots have the potential to improve the quality of life of older adults by encouraging and guiding their performance of rehabilitation exercises while offering cognitive stimulation and companionship. This study focuses on the early stages of developing and testing an interactive personal trainer robot to monitor and increase exercise adherence in older adults. The robot physically demonstrates exercises for the user to follow and monitors the user's progress using a vision-processing unit that detects face and hand movements. When the user successfully completes a move, the robot gives positive feedback and begins the next repetition. The results of usability testing with 10 participants support the feasibility of this approach. Further extensions are planned to evaluate a complete exercise program for improving older adults' physical range of motion in a controlled experiment with three conditions: a personal trainer robot, a personal trainer on-screen character, and a pencil-and-paper exercise plan.

1. Introduction

1.1. Benefits of Humanoid Robots to Older Adults. The proportion of adults aged 65 or older has been steadily increasing for more than a century in most developed countries. In the USA, it has increased from 4.1% (3.1 mil.) in 1900 to 8.1% (12.6 mil.) in 1950 to 12.4% (34.6 mil.) in 2000 and is projected to reach 20.6% (82 mil.) in 2050 [1]. This steep increase raises the concern of where older adults will live. A 1992 study found most of them prefer "aging in place," that is, remaining in their homes with little or no supervision [2, 3]. Although aging in place has some advantages, like increased autonomy and maintaining familiar surroundings, one potential disadvantage is fewer opportunities for receiving encouragement to engage in physical activity.

Physical activity may delay the onset of physical deficits contributing to frailty. These deficits include decreased skeletal muscle strength, gait speed, musculoskeletal flexibility, range of joint motion, postural stability (including balance, coordination, and reaction time), and cardiovascular responsiveness [4]. These conditions result in significant functional limitations. For example, 15% of people aged 75 to 84 are unable to climb stairs, and a substantial proportion of otherwise healthy older adults have limitations in gait speed that prevent them from crossing an intersection quickly enough to comply with traffic signals [4, 5]. Increased physical activity, such as through a daily exercise program, has been found to reduce physical ailments and improve strength and mobility [5–7]. However, exercise programs are beneficial only when followed regularly and consistently; living conditions and other factors may impede program adherence [8–11]. For instance, older adults living at home typically have reduced access to healthcare and health interventions as compared with those living under long-term care at a nursing home [12]. Furthermore, employing live-in-care staff can be expensive. Without the supervision and encouragement of nursing staff, stay-at-home older adults are at an increased risk of not adhering to an exercise program [13].

Another potential disadvantage of aging in place as compared with group living facilities is fewer opportunities for regular, meaningful interpersonal relationships. Companionship has important health benefits for older adults who age in place. For example, living together with another person significantly decreases feelings of loneliness in older

adults [14]. When human companionship is unavailable, an animal or robot companion can reduce feelings of isolation, in part by giving a sense of physical and social presence [15–17]. An animal companion positively impacts the health of socially isolated individuals and can produce long-term positive effects on the health and behavior of older adults [18, 19]. When a companion animal is not feasible, a robot may be substituted; this has the benefits of lower recurring costs, a lower burden of care and responsibility, and greater hygiene. Robots have elicited similar palliative outcomes when substituted for a companion animal [20–22].

However, few studies have been conducted on how to use robots or robot therapy to improve the mental and physical health or quality of life of older adults [20–24]. To close this research gap, the field of socially assistive robotics has emerged. Socially assistive robots have been used in certain contexts to aid recovery through social interactions [25]. The robot's interaction style may be informed by the human user's personality [26], movement, or physical orientation [27].

Socially assistive robots can be relatively inexpensive and simple to use. Paro [15, 28], for example, has been used in nursing homes in Japan, the United States, and Europe for companionship and to stimulate social interaction among patients [17, 20, 22]. Paro, which looks like a baby harp seal, is designed to provide therapy for older adults with dementia. Its sensors enable it to respond to both touch and speech in a manner resembling a domesticated animal companion.

Although Paro's cognitive capacities are extremely limited relative to those of people, animal pets, and even other robots, nonverbal cues such as looking toward the person speaking or responding to being petted can convey a sense of physical and social presence that in turn reduces loneliness and encourages the sharing of feelings [17]. Paro and other similar robots may provide comfort by giving the impression that "somebody is there." They succeed to the extent that they are able to "press our Darwinian buttons" by mimicking largely unconscious human and other animal behavior that elicits in their users prosocial behavior, such as the human desire to nurture and be nurtured [16]. Paro's success as a companion robot may result from its anthropomorphic appearance, and especially the inclusion of eyes, a conclusion supported by the findings of several unrelated experiments [29–31]. In particular, these findings support the theory that human beings have inherited an automatic, unconscious neural mechanism that conferred on their ancestors' selective advantage by increasing prosocial behavior when being observed. Thus, interactive technologies can be engineered to exploit unconscious mechanisms to promote adherence to a physical exercise program or to any other kind of activity supported by social expectations.

An advantage to humanoid robot companions is that not only can they be endowed with social intelligence but their appearance also affords the automatic perception of them as socially intelligent. Thus, when robots look, act, or are presented as humanlike, social entities, they are more likely to elicit in us the same responses that other human beings elicit [32]. This effect has been measured by the human interaction partner's conscious behavior, unconscious behavior (e.g., gaze) [33], attributions of thoughts, feelings, and intentions, and adherence to advice [34, 35]. The anthropomorphic physical embodiment of a humanoid robot could have a significant effect on patients' adherence to a physician-prescribed exercise program. Shinozawa et al. [34] found that participants are more likely to follow a robot's recommendation than that of an on-screen character. Kidd and Breazeal [35] found that participants track their exercise and calorie consumption for almost twice as long with a robot as with a computer or with paper and pencil. They also develop a closer relationship with the robot.

1.2. Affordable Interactive Exercise Systems. Despite the development of technologies for rapidly and robustly detecting human faces and hands, only recently have these technologies been applied to monitoring exercise performance and providing feedback [36]. Systems have been developed that demonstrate exercises [37] or provide feedback and encouragement for performing stroke rehabilitation exercises [38, 39] or completing mental and physical button-pressing tasks [40]. In 2011, Respondesign MayaFit Virtual Fitness Trainer [41], implemented on the PrimeSense OpenNI Framework, combines exercise adherence monitoring with an animated on-screen human-looking character to guide healthy individuals through a personalized sequence of exercise movements, monitor their progress, and provide feedback. (MayaFit uses the same three-dimensional motion capture technology as Microsoft's Kinect [42].) However, these are examples of specialized hardware and software. It should be possible to encourage exercise with more affordable, mass-produced devices [43].

In summary, older adults, especially those aging in place, are subject to physical and mental problems that drastically diminish their quality of life [44]. To reduce these problems, an interactive system could instruct, monitor, and encourage older adults during the performance of physician-prescribed exercises. Such a system would offer a combination of distinct advantages as compared with the usual paper-and-pencil-based materials or an automatic telephone reminder system. The interactive system performs the exercises in front of the participant; the system provides continuous instant feedback and encouragement during the exercises; the system provides a more affordable substitute for a human personal trainer; the system provides exercise guidance at flexible times; the system can increase adherence by presenting itself as a humanlike, social entity; the system can report the results back to the physician.

Interactive technologies can present their humanlike agency through the virtual embodiment of an on-screen character or through the physical embodiment of a humanoid robot. Each approach has its advantages. The advantages of an on-screen character, which requires only a computer, a video camera with a fixed focal length (i.e., a webcam), and software, are numerous: low purchase and maintenance costs, high portability (for notebook computer models), high reliability (as compared with robots, which are animated by motors that can jam and break), and the absence of safety risks related to physical contact (e.g., fingers pinched

by a robot joint) [45]. The advantages of an interactive robot are likewise numerous: the robot has heightened sociality because of its enduring anthropomorphic and physical presence, which is likely to increase adherence to treatment including exercise [34, 35]; even simple robots can provide a sense of companionship [16, 17]; robots often have mobility, which enables them to navigate their environments autonomously [23, 46], and thus can be designed to accompany their owners on walks; it is easier to see and understand exercises performed in three dimensions by a robot than in two dimensions by a character on a screen because the former affords depth perception by binocular disparity and movement parallax.

2. Prototype Interaction Design and System Design and Implementation

The long-term goal of this research is to increase adherence among adults aged 65 or older to a physician-prescribed exercise program through their interaction with a personal trainer robot in their own homes. This technology is intended for older adults who have a sedentary lifestyle. The fully developed system should be inexpensive, communicate with older adults in real time, and report the results to their physicians through a hospital webserver.

The humanoid robot in this study has been designed to start an exercise session with users and help them adhere to a predetermined schedule. When the user is ready, the robot demonstrates the first prescribed exercise by moving its body parts. If the user performs the exercise correctly, the robot praises the user and begins the next repetition. To communicate with the user, the robot uses synthesized speech in addition to the gestures that depict each exercise movement. The robot also recognizes hand and head movements to monitor the user's progress through the exercise set and to estimate the user's activity level (Figure 1).

2.1. Software Components of Personal Trainer Robot Prototype. The software is composed of a vision-processing unit and an exercise adherence unit, which communicate with each other to determine and carry out the next move of the robot. The interactive prototype detects the users' physical presence and recognizes the users' gestures (i.e., the exercise movements). It determines whether the user has successfully performed the exercise move as demonstrated by the robot. Next, the exercise adherence unit obtains the tracked head and upper arm positions from the vision-processing unit and delivers the appropriate voice commands accordingly (Figure 2).

2.1.1. Vision-Processing Unit. The vision-processing unit uses a wide-angle USB video camera. The camera captures the video, which is then processed in its native size. The system sets two different regions of interest (ROIs) to indicate the most likely location for features: one for the face and the other for the hands. The face and hand detection are then performed in their respective ROIs.

The user is encouraged to sit at a certain distance from the camera so as to be positioned at the center of the video
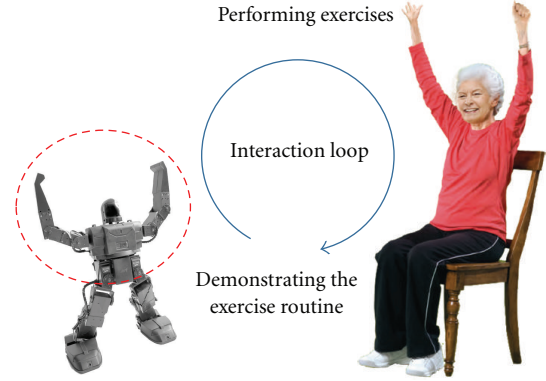

FIGURE 1: An interaction loop links the older adult user and the personal trainer robot. Within this loop, the robot leads the user through each exercise move by physically demonstrating it and giving encouragement and feedback on the user's adherence through synthesized speech.
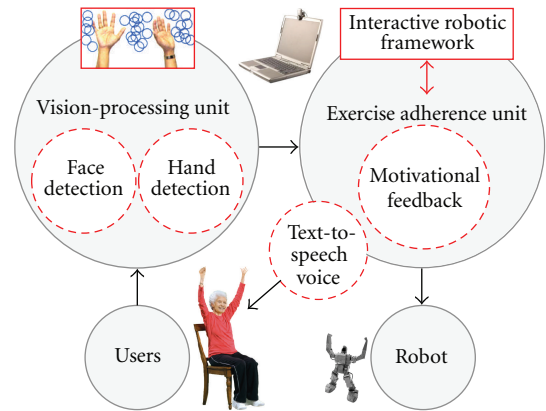

FIGURE 2: Information flows from the vision-processing unit, which monitors the user's movements, to the exercise adherence unit. The exercise adherence unit controls the robot's performance of each exercise move and, after the user completes (or fails to complete) the move, reports the adherence assessment. The interactive framework used in this study is shown in Figure 7.

frame. The optimal distance is determined by the user's height. Initially, the user is encouraged to perform trial exercise moves in front of the camera to check whether the user in the raised-hands position fits within the video frame. Lighting conditions consistent with a well-lit room should be maintained throughout the process so that enough light falls on the face and hands for accurate detection.

2.1.2. Face Detection and Tracking. Face detection is a part of object recognition research [47], and much work has been conducted on it since the inception of vision processing. Turk and Pentland [48] developed an automatic recognition system based on eigenfaces that compares the features of novel faces to already known faces. Liu [49] applied Bayesian discriminating features to frontal face detection, while Mohan et al. [50] devised an example-based algorithm. Viola and Jones [51] used machine-learning techniques and Haar-like features for rapid, accurate face detection and later added a set of tilted Haar-like features to enhance detection [52].

The current study builds on previous research by using classifiers with an extended set of Haar-like features, including edge, line, and center-surround features for rapid face detection and localization [51, 52]. The Haar-like features test for a face in the face ROI of each frame. A cascade of classifiers was employed to increase the detection rate. Once detected, the face is extracted within a bounded rectangular region, and the centroid of the rectangle is computed and continuously tracked.

Haar face and hand classifiers were created through a training process using the OpenCV library. Training a Haar classifier is a CPU time- and memory-intensive process that requires sample images of the object of interest (positive images) and of other objects (negative images). To improve classification accuracy, a 20-stage cascade of Haar classifiers was trained with the Gentle AdaBoost algorithm [53] using 5,000 positive and 6,500 negative images. The training required 18 days and was performed on a PC with an Intel Pentium 4 (2.2 GHz) processor and 2 GB of RAM.

To improve the system's classification accuracy, the user is positioned at the center of the video frame. Because users' heights in this study varied between 165 and 188 cm, they were first asked to sit in front of the camera so that the approximate regions of the face could be calculated and an ROI could be obtained. Based on these calculations, the face ROI origin point was set 200 pixels to the right and 100 pixels down from the top-left corner of the original frame (i.e., $x = 200$ and $y = 100$). An ROI 340 pixels wide by 300 pixels tall was then created with its top left corner at the face ROI origin point. Presetting the ROI increases the efficiency and the accuracy of the system because it provided a smaller area in which the face is likely to be found. The region was cropped and displayed in a separate window. The face detection system searched the entire region for a face by repeatedly applying the cascade of Haar classifiers in the Haar-like feature space. After locating a rectangular area containing a face, the system returns the coordinates of its centroid and four corners and tracks the face in that area. Thus, the movements of the face are continuously monitored, and the coordinates are stored in memory for future reference. The calculation of the centroid enabled the tracking of even small movements of the head, thereby increasing the sensitivity and effectiveness of the system.

2.1.3. Upper-Limb Motion Detection and Tracking. The Haar-like features and Haar classifiers that were used for face detection were initially applied to hand detection. However, because hand gestures are often more complex than those made by the head and face (i.e., because the hands can be twisted into more physically distinct configurations), the Haar classifiers were less effective for hand detection than for face detection and, subsequently, the results could not be used during real-time video streaming. A more feasible method of observing hand movements was using motion detection (Figure 3). The method was reused for detecting motion in the other parts of the upper limbs, including the forearms.

Motion Detection. The frames used for vision processing are captured from the camera. From each of these frames, the subframe defined by the ROIs is extracted and processed. An example of one such subframe on which the image processing is performed is shown in Figure 3(a). Subframes will contain some noise from the camera sensor, which should be reduced to avoid false positives (Figure 3(b)). Applying to the subframe a simple blur reduces this noise. Once noise is removed, the presence of motion is detected by calculating the absolute difference in corresponding pixel values from two consecutive subframes. A new image with the calculated difference is created, and the difference image is converted to 8-bit grayscale, so that it is easy to apply filters (Figure 3(c)). A binary threshold filter is then applied to the subframe (Figure 3(d)). The presence of motion is defined as a sufficient difference in pixel luminance values between two consecutive subframes [54]. In this difference image, white pixels indicate the presence of motion.

When the position of the object changes between subframes, it produces a shift in darker and lighter pixels: darker when an object at the location in the first frame disappears in the second frame, and lighter when an empty location in the first frame contains an object in the second frame. When the difference image is converted to grayscale, it becomes easier to find the differences between the two source images.

To provide visual feedback of upper-limb motion detection, the ROI is overlaid with equally sized circles, each of which is bounded by a square. The circles scatter away from areas where motion is present. For example, if motion occurs at the bottom of the ROI, the circles will scatter towards the middle and top. Whenever the total number of changed pixels in each bounding square exceeds the predefined value of 100, motion is considered present in the bounding square.

Upper-Limb Tracking. The absence of circles in two large areas of the ROI indicates the location of the upper limbs. The upper limbs are tracked by calculating the difference between the centroids of the two corresponding rectangular areas in consecutive frames.

Determining the Upper-Limb Region of Interest. For detecting motion in the upper limbs, an ROI was obtained by a method similar to the one used for calculating the ROI for face detection, namely, by observing several users making the motion, which in this case was raising their hands. An assumption was made that the location of the hands would provide sufficient information about the location of the corresponding forearms. On that assumption, the ROI for the hand region was obtained by observing the raised hands of several users. The hands ROI origin was set 100 pixels to the right of the top-left corner of the frame. An ROI subframe 400 pixels wide and 200 pixels tall was then created with its top left corner at the hands ROI origin point.

2.1.4. Exercise Adherence Unit. The exercise adherence unit monitors data from the vision-processing unit and plays the appropriate voice commands.

The communication of the robot is controlled by a program developed using the robot's software development kit
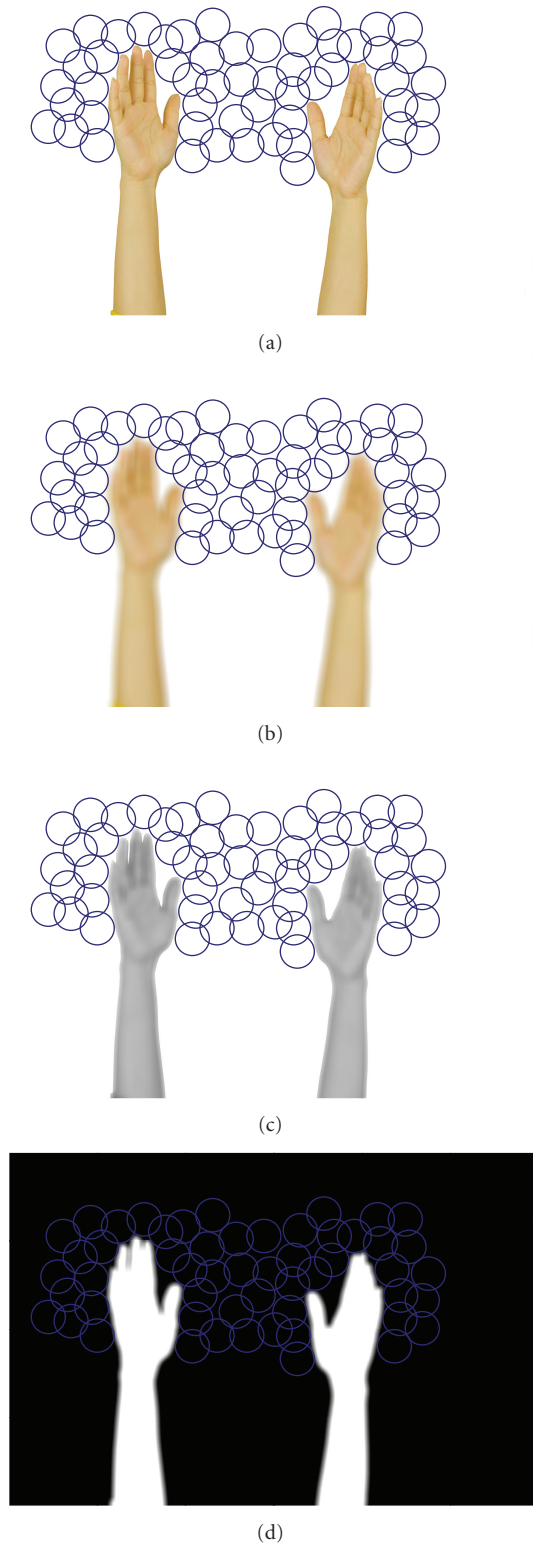
(a)

(b)

(c)

(d)

FIGURE 3: During motion detection, subframes within the region of interest are captured (a), blurred (b), and converted to shades of gray (c). In the areas of the subframe indicated by circles, the pixels changing between consecutive captures (d) comprise a direct measurement of motion. Such pixel changes are repeatedly summed and compared with a predefined threshold value. At a given time in a given area, if the threshold value is exceeded, motion is detected.

(SDK). The voice commands delivered by the robot during interaction were synthesized using AT & T Labs' Natural Voices text-to-speech software [55]. These voice commands are incorporated into the software system and synchronized with the exercise routines. A connection is made such that the exercise adherence unit communicates directly with the vision-processing unit to interact with the user. The exercise adherence unit receives information about the presence of a person from the vision-processing unit (Figure 4). If it detects a user, the robot greets the user and requests consent to start the exercise routine. The user indicates readiness by waving one hand overhead. The robot then starts the interaction cycle by demonstrating the first of the recommended physical exercise routines. The robot vocally announces the first exercise movement and then demonstrates it by moving its body parts. Next, the robot waits for the user to imitate the movement. The robot detects the movement, analyzes its timing and form, and judges whether the user's action is correct. After a successful attempt, the robot praises the user; otherwise, it repeats the movement and instructions. This continues until the user performs the exercise correctly. At the end of the interaction cycle, the robot gives verbal feedback on the user's performance during the exercise routines. The robot then provides a goodbye message and ends the session.

The exercise adherence unit demonstrates exercises by specifying parameters to the robot's servomotors for each exercise move: the desired angle for every joint and both the velocity and number of displacement steps with which the joint should move to that angle. Two exercise moves were used in this study: the overhead arm raise and the head turn.

Detecting an Overhead Arm Raise. To calibrate the system, the robot asks the user to raise the hands as far as possible. During the overhead arm raise, one repetition is counted if the user raises the hands at least 90% of that extent. The success rate is measured by the number of attempts divided by the number of trials. Range of motion is informally quantified as the percentage of the maximal arm raise averaged across all trials.

Detecting a Head Turn. The extent of a head turn is estimated from the deviation of the face's centroid from its head-on position divided by a constant and then taking the arcsine. The constant is determined empirically after setting up the system. The exercise adherence unit counts one head turn if the head is turned at least 45 degrees. If the user is unable to make a 45-degree head turn after three attempts, the threshold for a head turn is lowered to the mean of the maximal angle during the three attempts. The success rate is measured by the number of attempts divided by the number of trials. Range of motion is approximated as the mean degrees the head is turned across all trials, regardless of whether they were successful.

2.2. Hardware Components of Personal Trainer Robot Prototype. The interactive system has three hardware components: the robot, the webcam, and the controller computer.
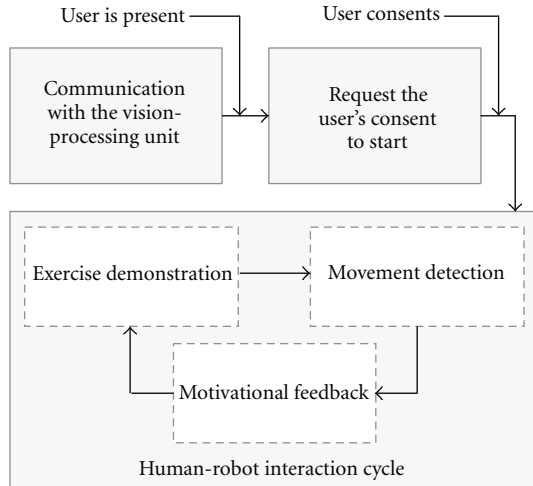
Figure 4: In the exercise adherence interaction framework, the interaction is controlled mainly by the user's consent to participate. Consent is requested both before the first exercise and after each completed exercise.
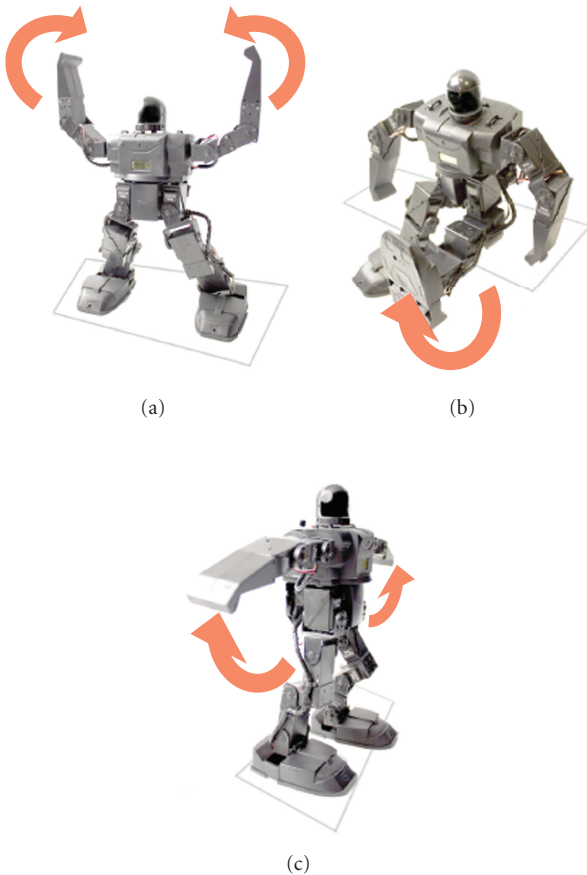

(a)

(b)

(c)

Figure 5: The set of possible exercises is determined by the physical capabilities of the humanoid robot. The robot's 20 servo motors enable the turning movements of the head, waist, and thighs and joint movements of the limbs.

2.2.1. Humanoid Robot. RoboPhilo, a programmable humanoid [56] robot, has been used in this study. It has 24 available servo channels with up to eight input-output interfaces. It has 20 servomotors that enable the turning movements of the head, waist, and thighs and joint movements of the limbs. It can be connected directly to a PC via an RS-232 serial connection and can be programmed using its SDK. It can be controlled directly by using the infrared remote or autonomously by using the SDK. The robot can be programmed with various exercise movements (e.g., **Figure 5(a)**, 5(b), and 5(c)). At US\$500, RoboPhilo is relatively inexpensive given the number and mobility of its joints. At the time of this writing, a complete system including computer, webcam, and SDK could be purchased for about US\$900.

2.2.2. Webcam. The vision-processing unit can use a video camera that is either built in or externally attached to the PC. The testbed uses a Logitech Webcam Pro 9000 USB camera with a 72° diagonal field of view and a maximum resolution of 1600×1200 pixels. At that resolution, the video frame rate was 10 fps. The large field of view simultaneously captures both the head and hand regions, and the higher pixel count allows greater precision during motion tracking.

2.2.3. Controller PC. The vision-processing unit, exercise adherence unit, and robot were controlled by a PC running a 32-bit version of Microsoft Windows XP. The PC had an Intel Core 2 Duo processor, 4 GB of RAM, and a PCI-based RS-232 serial port. (A USB–RS-232 adapter could not be used because it increased the startup time to 15 s, which exceeds the 10 s interval during which RoboPhilo must receive its initial response from the PC to avoid halting.)

3. Methods

The quality of the human-robot interaction of the prototype system was assessed in a diagnostic usability test. The testing of the interaction was important at this initial stage to ascertain the participants' enthusiasm and interest when interacting with the robot and their perception of the robot as a potential trainer. This information is essential to assessing the system's user friendliness and commercial viability.

3.1. Procedure. The usability test was conducted in a room with large windows, overhead fluorescent lights, and lightly textured off-white walls. The setup conditions enabled the system to detect the face and hands of the participant easily. The testing equipment included the robot hardware unit, video camera, and computer. A chair was positioned against the wall and facing the robot and the camera. It was positioned so that the participant would always fit within the video frame (**Figure 6**). Other objects in the camera's field of view were removed to reduce false positives.

The participants were instructed to sit on the chair and were introduced to the purpose of the study, the capabilities

Figure 6: The first author demonstrates a typical position of a participant in the study.

and limitations of the robot, and the session's flow of interaction. They were cautioned about unexpected and out-of-sequence occurrences owing to the fact that the testbed was still in the development and testing stage. The participants were then asked to watch the robot and repeat its actions using only hand and head movements. The robot demonstrated two basic moves to the participants: the overhead arm raise and the head turn. The interactions were observed, and feedback was sought from the participants. The interactions were also video recorded for later analysis. Figure 7 details the step-by-step procedure of the interaction cycle.

4. Results and Discussion

4.1. Qualitative Results of the Usability Test. The interaction sequence for seven of the ten participants was conducted without any technical glitches (Table 1). However, for three of the participants, the sequence was interrupted. Two participants failed to comply with the instructions: they either moved their hands too quickly or did not turn their head in the same direction the robot turned its head. The robot failed to detect the raised hands of the third participant because the hands were outside of the captured frame. Although the interaction typically took seven to eight minutes, one interaction took 12 minutes.

Nine of the ten participants assumed that the robot could listen, understand, and process what they were saying and respond accordingly. They were reminded not to talk to the robot, but to communicate with it using hand or head movements. Their attempts at verbal communication indicate that most users expect an interactive robot to listen and reply to them in a way that is uniquely appropriate to the direction and nature of the conversation and situation [56]. The fulfillment of these user expectations is an area for future research.

4.2. Technical Refinements. As expected, the robot was unable to change its behavior and act according to the situation when the sequence was altered or disturbed. For example, in the case of the three participants with whom the interaction was not smooth, the robot either continued with its programmed responses regardless of the participants'

reactions or it abruptly ended the sequence. This inconsistent behavior occurred because the robot has been programmed to detect threshold values for each of the exercise moves. The robot waits for the participant to reach before considering the participant's attempt as successful. However, in the case of an unsuccessful attempt by the participant, the vision-processing unit reported a random value, resulting either in the robot skipping essential steps in the sequence or in the termination of the entire sequence. These two errors were observed with the first two participants, but were fixed immediately thereafter. The errors did not occur with the remaining eight participants.

The trained Haar classifier detected multiple faces at the same time. This was problematic because in some cases, multiple faces were detected when there was only one face. Thus, enabling the detection of multiple faces resulted in more frequent false positives. This occurs because the classifier identifies as many sets of coordinates as the number of faces it detects, which results in the termination of the vision-processing unit. Although multiple face detection was intentionally implemented, we did not anticipate the problem of cycle termination. The bug was fixed by restricting the classifier to detect only one face in the frame and to reject all subsequent potential faces.

It was also observed that the robot terminated the interaction if the participants' response came too slowly after each exercise demonstration instead of waiting for the participant to respond. The bug was fixed after the usability test. Finally, the participants reported feeling that the robot terminated the program too abruptly after the exercise routines were completed. The participants reported that they would have preferred to receive more feedback about their performance before the robot conveyed the goodbye message. The exercise adherence unit was modified to report the percentage of repetitions completed and the average extent in percentages of the overhead arm raise and head turn.

5. Conclusion

By 2050, the number of Americans who are 65 or older is expected to more than double, reaching 82 millions. These older adults constitute the most sedentary segment of the US population, and they suffer from the most chronic conditions that are preventable through exercise. Although the most convenient and effective means of increasing exercise adherence among older adults living at home involve one-on-one monitoring and encouragement, the required human resources are in short supply and the costs are prohibitive. Medically, treating chronic conditions that are preventable through exercise incurs high economic costs for the afflicted individuals, their families, and for Medicare or their private health insurers. The loss of mobility can necessitate additional costs associated with nursing care, either at home or at a nursing facility. In addition, conditions preventable through exercise can incur high personal and social costs, including physical pain and suffering and social isolation caused by loss of mobility.
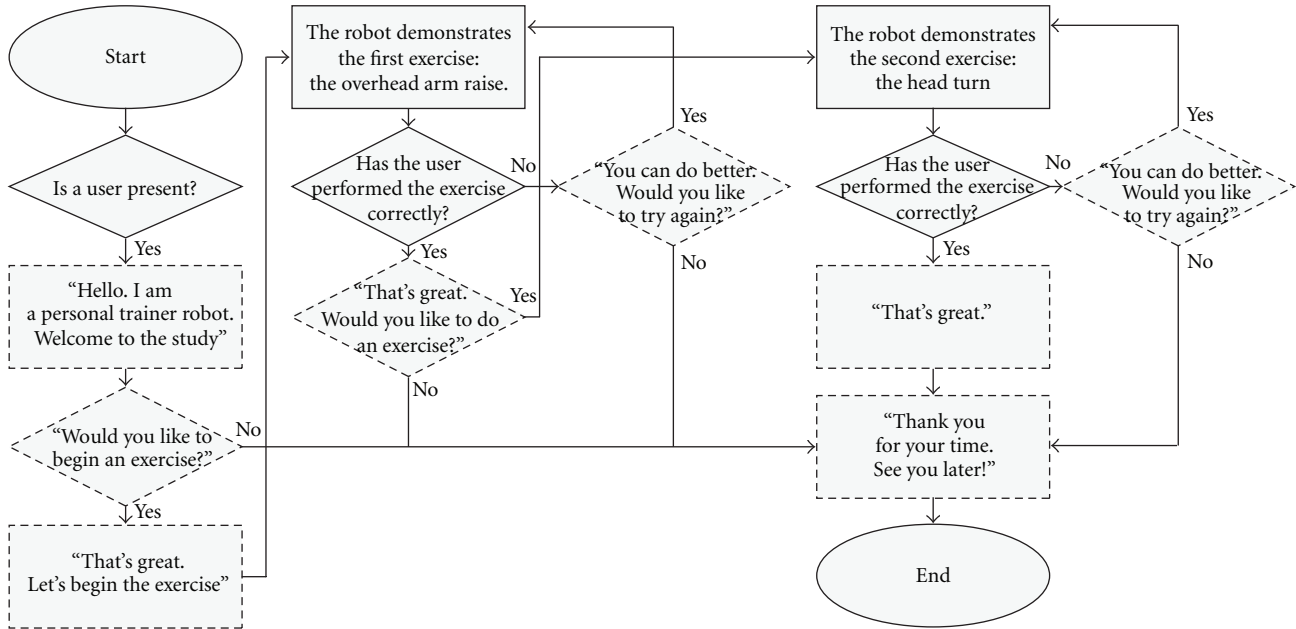
FIGURE 7: The initial interactive robotic framework is represented by a flowchart. When the vision-processing unit detects the presence of a user, it signals this information to the exercise adherence unit, which begins the interaction. If the user consents to begin an exercise, the robot first demonstrates it. Based on information from the vision-processing unit, the exercise adherence unit determines whether the user has completed an exercise correctly and, if the user has not, give the user an opportunity to be guided through the exercise again.

TABLE 1: Usability test.

Participants	Successful	User error	Tech. error
10	7	2	1

Presently, no widespread systems exist to measure and increase adherence to a physician-prescribed exercise program with the potential advantages of a humanoid robot, including enhanced sociality and companionship and the ability to lead or follow a person by their own locomotion. To address these issues and as a replacement for a human personal trainer, we have proposed an interactive framework for a personal trainer robot to remind users to perform their exercises, to demonstrate the exercises and provide instruction, to monitor performance progress in real time, and to provide feedback and encouragement. If an interaction framework embedded in a robot can increase exercise adherence, this could greatly reduce healthcare and nursing costs for the elderly and, by incorporating some of the abilities of a personal trainer, provide a high return on investment as compared with other interventions. It could also enable physicians to reliably and systematically monitor patients' adherence to a prescribed in-home exercise program by uploading data to a hospital webserver.

Using the robot as a personal trainer to empower patients in making a behavioral change is a relatively new area to explore. In this study, we conducted a usability test on an incomplete prototype system. The participants' initial response to the personal trainer robot was very positive. They were receptive and responded favorably to interacting with it. Valuable feedback was obtained through their interactions,

which led to the implementation of changes to improve the functionality and usability of the robot.

5.1. Technical Limitations. One of the major limitations of the study is that the vision-processing software requires good lighting conditions to detect head and hand movements accurately. Poor lighting may result in false negatives, and excessive lighting may result in false positives. Another limitation is that to obtain accurate results, only one person can be in front of the camera during the detection process. A third limitation is that the vision-processing unit must be reset between users to adjust for differences in height. Finally, the software does not yet fully allow for tracking actions outside of the sequence.

5.2. Directions for Future Research. Because this is a proof of concept study, it has considerable scope for extension. To make the robot a feasible option to help older adults increase their physical activity, future research should include programming the robot to incorporate a complete exercise program, such as the moves recommended by the US National Institute of Aging [57]. To validate that this approach can increase both the exercise adherence and physical range of motion of older adults as compared with alternative interventions, an experiment is planned using this demographic with three conditions: a personal trainer robot, a personal trainer on-screen character, and a pencil-and-paper exercise plan.

During usability testing, users suggested an idea to improve the interaction: the incorporation of more realistic and timely feedback through voice commands or comments

after each step in the exercise routine and after both successful and unsuccessful attempts by the user. These voice messages could be tailored to the particular motivational needs of the user to increase adherence to physical activity. A validated methodology for determining what messages are appropriate for a particular individual is to apply a model of behavior change, such as the theory of planned behavior (TPB) [11, 58].

According to TPB, it is possible to change a person's behavior by changing that person's beliefs about behavioral outcomes, the normative expectations of others, and controlling factors, such as facilitating conditions or barriers [59]. These in turn elicit positive or negative attitudes toward the behavior and responses to social pressure. TPB has been found effective in changing the behavior of patients with diabetes, inflammatory bowel disease, and obesity in interactive games [60–62]. The behavior change model could be used by the personal trainer robot to give encouragement to a user that addresses that individual's particular concerns and priorities. This may prove critical to maintaining the user's motivation during long-term interventions. This interactive robot testbed can also be used in health games and by the health games community.

Another potential improvement is to train more efficient hand detection classifiers. For this version, the motion detection algorithm followed a low-level procedure. However, training Haar classifiers is a more robust method for detecting movement relative to motion detection, because it eliminates false positives from the motion of irrelevant objects. Using Haar classifiers is a high-level approach. The human hand can assume a number of positions and subsequently training classifiers for detecting hand movement are not straightforward. One solution, then, is to train multiple classifiers for the hand and to use them simultaneously [63]. Although using multiple classifiers might adversely affect processing speed, it may be suitable for use with older adults who have a restricted range of motion and slower hand movements, which will increase the detection rate and accuracy of motion tracking.

Another future direction is the development of an animated character for this system that will act as a trainer and substitute for the robot. This will enable the system to be distributed online rather than physically. An animated humanoid character, especially one with facial expressivity, may increase exercise adherence more than the paper-and-pencil method while costing less for the user than a humanoid robot. Whether the effectiveness of an animated character can rival that of a humanoid robot is a question for future research to address.

Acknowledgments

References

[1] W. He, M. Sengupta, V. A. Velkoff, and K. A. DeBarros, "U.S. Census Bureau," 2005, http://www.census.gov/prod/2006 pubs/p23-209.pdf.

[2] J. J. Calahan, "Aging in place," *Generations*, vol. 16, no. 2, p. 5, 1992.

[3] K. D. Marek and M. J. Rantz, "Aging in place: a new model for long-term care," *Nursing Administration Quarterly*, vol. 24, no. 3, pp. 1–11, 2000.

[4] N. G. Kutner, M. G. Ory, D. I. Baker, K. B. Schechtman, M. C. Hornbrook, and C. D. Mulrow, "Measuring the quality of life of the elderly in health promotion intervention clinical trials," *Public Health Reports*, vol. 107, no. 5, pp. 530–539, 1992.

[5] C. Vallbona and S. B. Baker, "Physical fitness prospects in the elderly," *Archives of Physical Medicine and Rehabilitation*, vol. 65, no. 4, pp. 194–200, 1984.

[6] M. P. Buman, E. B. Hekler, W. L. Haskell et al., "Objective light-intensity physical activity associations with rated health in older adults," *American Journal of Epidemiology*, vol. 172, no. 10, pp. 1155–1165, 2010.

[7] D. E. R. Warburton, C. W. Nicol, and S. S. D. Bredin, "Health benefits of physical activity: the evidence," *Canadian Medical Association Journal*, vol. 174, no. 6, pp. 801–809, 2006.

[8] S. N. Culos-Reed, W. J. Rejeski, E. McAuley, J. K. Ockene, and D. L. Roter, "Predictors of adherence to behavior change interventions in the elderly," *Controlled Clinical Trials*, vol. 21, no. 5, pp. 200S–205S, 2000.

[9] A. C. King, W. J. Rejeski, and D. M. Buchner, "Physical activity interventions targeting older adults: a critical review and recommendations," *American Journal of Preventive Medicine*, vol. 15, no. 4, pp. 316–333, 1998.

[10] M. A. Napolitano, G. D. Papandonatos, B. A. Lewis et al., "Mediators of physical activity behavior change: a multivariate approach," *Health Psychology*, vol. 27, no. 4, pp. 409–418, 2008.

[11] M. S. Hagger, N. Chatzisarantis, and S. J. H. Biddle, "A meta-analytic review of theories of reasoned action and planned behavior in physical activity," *Journal of Sports and Exercise Psychology*, vol. 24, no. 1, pp. 3–32, 2002.

[12] W. D. Spector, J. Fleishman, L. Pezzin, and B. Spillman, "The characteristics of adult long-term care users: comparing the non-elderly and the elderly," *Association for Health Services Research Meeting*, vol. 16, p. 240, 1999.

[13] R. J. Nied and B. Franklin, "Promoting and prescribing exercise for the elderly," *American Family Physician*, vol. 65, no. 3, pp. 419–428, 2002.

[14] K. Holmén, K. Ericsson, and B. Winblad, "Social and emotional loneliness among non-demented and demented elderly people," *Archives of Gerontology and Geriatrics*, vol. 31, no. 3, pp. 177–192, 2000.

[15] Y. Tao, H. Wei, T. Wang, and D. Chen, "A multi-layered interaction architecture for elderly companion robot," *Lecture Notes in Artificial Intelligence*, vol. 5314, no. 1, pp. 538–546, 2008.

[16] S. Turkle, "Authenticity in the age of digital companions," *Interaction Studies*, vol. 8, no. 3, pp. 501–517, 2007.

[17] S. Turkle, W. Taggart, C. D. Kidd, and O. Dasté, "Relational artifacts with children and elders: the complexities of cyber-companionship," *Connection Science*, vol. 18, no. 4, pp. 347–361, 2006.

[18] E. Friedmann, A. H. Katcher, J. J. Lynch, and S. A. Thomas, "Animal companions and one-year survival of patients after discharge from a coronary care unit," *Public Health Reports*, vol. 95, no. 4, pp. 307–312, 1980.

[19] J. Serpell, "Beneficial effects of pet ownership on some aspects of human health and behaviour," *Journal of the Royal Society of Medicine*, vol. 84, no. 12, pp. 717–720, 1991.

[20] K. Wada, T. Shibatal, T. Musha, and S. Kimura, "Effects of robot therapy for demented patients evaluated by EEG," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS '05)*, pp. 1552–1557, 2005.

[21] K. Wada, T. Shibata, T. Saito, and K. Tanie, "Effects of robot-assisted activity for elderly people and nurses at a day service center," in *Proceedings of the IEEE*, pp. 1780–1788, November 2004.

[22] K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie, "Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2785–2790, April 2005.

[23] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: challenges and results," *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 271–281, 2003.

[24] J. Fasola and M. J. Matarić, "Robot exercise instructor: a socially assistive robot system to monitor and encourage physical exercise for the elderly," in *Proceedings of the 19th IEEE International Workshop on Robot and Human Interactive Communication*, Viareggio, Italy, 2010.

[25] D. Feil-Seifer, K. Skinner, and M. J. Matarić, "Benchmarks for evaluating socially assistive robotics," *Interaction Studies*, vol. 8, no. 3, pp. 423–439, 2007.

[26] A. Tapus, C. Tapus, and M. J. Mataric, "User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation," *Intelligent Service Robotics*, vol. 1, pp. 169–183, 2008, Special issue on multidisciplinary collaboration for socially assistive robotics.

[27] A. Panangadan, M. Matarić, and G. S. Sukhatme, "Tracking and modeling of human activity using laser rangefinders," *International Journal of Social Robotics*, vol. 2, no. 1, pp. 95–107, 2010.

[28] C. Breazeal and B. Scassellati, "How to build robots that make friends and influence people," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '99)*, pp. 858–863, October 1999.

[29] M. Bateson, D. Nettle, and G. Roberts, "Cues of being watched enhance cooperation in a real-world setting," *Biology Letters*, vol. 2, no. 3, pp. 412–414, 2006.

[30] T. C. Burnham and B. Hare, "Engineering human cooperation," *Human Nature*, vol. 18, no. 2, pp. 88–108, 2007.

[31] K. J. Haley and D. M. T. Fessler, "Nobody's watching?" *Evolution and Human Behavior*, vol. 26, no. 3, pp. 245–256, 2005.

[32] K. F. MacDorman and H. Ishiguro, "The uncanny advantage of using androids in social and cognitive science research," *Interaction Studies*, vol. 7, no. 3, pp. 297–337, 2006.

[33] K. F. MacDorman, T. Minato, M. Shimada, S. Itakura, S. Cowley, and H. Ishiguro, "Assessing human likeness by eye contact in an android testbed," in *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, Stresa, Italy, 2005.

[34] K. Shinozawa, F. Naya, J. Yamato, and K. Kogure, "Differences in effect of robot and screen agent recommendations on human decision-making," *International Journal of Human Computer Studies*, vol. 62, no. 2, pp. 267–279, 2005.

[35] C. D. Kidd and C. Breazeal, "Robots at home: understanding long-term human-robot interaction," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '08)*, pp. 3230–3235, Nice, France, September 2008.

[36] L. E. Sucar, R. Luis, R. Leder, J. Hernandez, and I. Sanchez, "Gesture therapy: a vision-based system for upper extremity stroke rehabilitation," in *Biomedical Engineering Systems and Technologies*, A. Fred, F. Joaquim, and H. Gamboa, Eds., pp. 531–540, Springer, New York, NY, USA, 2009.

[37] M. B.-V. Kuren, "Robotic solutions in pediatric rehabilitation," in *Rehabilitation Robotics*, pp. 1–12, I-Tech, Vienna, Austria, 2007.

[38] R. Colombo, F. Pisano, A. Mazzone et al., "Design strategies to improve patient motivation during robot-aided rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, no. 3, pp. 1–12, 2007.

[39] M. J. Matarić, J. Eriksson, D. J. Feil-Seifer, and C. J. Winstein, "Socially assistive robotics for post-stroke rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, no. 5, pp. 1–9, 2007.

[40] J. Fasola and M. J. Mataric, "Robot motivator: improving user performance on a physical/mental task," in *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 295–296, La Jolla, Calif, USA, 2009.

[41] P. A. Barnes, T. Spooner, J. L. Dhabolt, and J. Leighton, "Instructional gaming methods and apparatus," Patent A1 20090176581, 2009.

[42] J. Shotton, A. Fitzgibbon, M. Cook et al., "Real-time human pose recognition in parts from a single depth image," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, June 2011.

[43] A. M. Okamura, M. J. Matarić, and H. I. Christensen, "Medical and health-care robotics," *IEEE Robotics and Automation Magazine*, vol. 17, no. 3, Article ID 5569021, pp. 27–37, 2010.

[44] M. P. Cutchin, "The process of mediated aging-in-place: a theoretically and empirically based model," *Social Science and Medicine*, vol. 57, no. 6, pp. 1077–1090, 2003.

[45] D. J. Feil-Seifer and M. J. Mataric, "Ethical principles for socially assistive robotics," *Robotics and Automation Magazine*, vol. 18, no. 1, pp. 24–31, 2011, Special issue on roboethics.

[46] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, "Robovie: an interactive humanoid robot," *Industrial Robot*, vol. 28, no. 6, pp. 498–503, 2001.

[47] K. F. MacDorman, K. Tatani, Y. Miyazaki, and M. Koeda, "Proto-symbol emergence," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1619–1625, November 2000.

[48] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[49] C. Liu, "A Bayesian discriminating features method for face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 725–740, 2003.

[50] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 4, pp. 349–361, 2001.

[51] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 511–518, December 2001.

[52] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '02)*, pp. 900–903, September 2002.

[53] R. Lienhart, A. Kuranov, and V. Pisarevsky, "Empirical analysis of detection cascades of boosted classifiers for rapid object detection," in *Proceedings of the 25th DAGM Pattern Recognition Symposium*, vol. 2781, pp. 297–304, Magdeburg, Germany, 2003.

[54] P. Brox, I. Baturone, S. Sánchez-Solano, J. Gutiérrez-Ríos, and F. Fernández-Hernández, "Fuzzy motion adaptive algorithm for video de-interlacing," *Lecture Notes in Computer Science*, vol. 4252, pp. 363–370, 2006.

[55] AT&T Labs Inc. AT&T Labs Natural Voices Text-to-Speech Demo, 2010, http://www2.research.att.com/~ttsweb/tts/demo.php.

[56] P. Dourish, "What we talk about when we talk about context," *Personal and Ubiquitous Computing*, vol. 8, no. 1, pp. 19–30, 2004.

[57] National Institute of Aging. Go4Life, 2010, http://go4life.niapublications.org.

[58] I. Ajzen and M. Fishbein, *Understanding Attitudes and Predicting Social Behavior*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1980.

[59] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, 1991.

[60] H. Kharrazi, C. Watters, and S. Oore, "Improving behavioral stages in children by adaptive applications," *Journal on Information Technology in Healthcare*, vol. 6, no. 1, pp. 322–326, 2008.

[61] H. Kharrazi, "Improving healthy behaviors in type 1 diabetic patients by interactive frameworks," in *Proceedings of the American Medical Informatics Association Symposium*, vol. 2009, pp. 322–326, San Francisco, Calif, USA, 2009.

[62] H. Kharrazi and L. Vincz, "Using health information technology and behavior change model to increase exercise adherence in obesity," in *Proceedings of the AMA-IEEE Medical Technology Conference on Individualized Healthcare*, Washington, DC, USA, 2010.

[63] V. Spruyt, A. Ledda, and S. Geerts, "Real-time multi-colourspace hand segmentation," in *Proceedings of the 17th International Conference on Image Processing (ICIP '10)*, pp. 3117–3120, Hong Kong, 2010.