Design of a Low-Cost Social Robot: Towards Personalized Human-Robot Interaction

Christian G. Puehn, Tao Liu, Yixin Feng, Kenneth Hornfeck, and Kiju Lee

Case Western Reserve University
Cleveland, Ohio 44106, USA
kiju.lee@case.edu

Abstract. This paper presents a low-cost social robot, called *Philos*, and human-robot interaction (HRI) design. The system is accompanied with a user interface that allows customization of interactive functions and real-time monitoring. The robot features eight degrees of freedom that can generate various gestures and facial expressions. HRI is realized by two elements, internal characteristics of the robot and external vision/touch inputs provided by the users. Internal characteristics determine the predefined personality of Philos among the five: Friendly, Hyperactive, Shy, Cold, or Sensitive, and set the behavioral control parameters accordingly. Vision-based interaction includes face tracking, face recognition, and motion tracking. Embedded touch sensors detect physical touch-based interaction. Behavioral parameters are updated in real time based on the user inputs, and therefore Philos can engage each user in personalized interaction via uniquely defined behavioral responses. The cost of Philos is estimated to be relatively low compared to other commercially available robots promising a broad range of potential applications for domestic and professional use.

Keywords: Human-Robot Interaction, Social Robot, Face Tracking, Face Recognition, Behavioral Control.

1 Introduction

Social robots are designed to entertain, assist, or provide service to humans though vision, touch, and sound-based interaction. Therefore, human-robot interaction (HRI) often resembles the way humans interact with each other. Recently, social robots have been receiving growing interest for their great potential as a long-term health care solution. For example, a social robot can serve as a companion for older people by helping them maintain independent living [1]. In addition, recent studies have demonstrated potential uses of social robots in behavioral training for children with developmental disabilities [2], [3], [4].

Over the past several decades, a number of socially interactive robots have been developed, covering a range of design and functionality objectives. The Huggable, a robot with the outer appearance of a teddy bear, focuses on implementing a sophisticated touch–sensitive skin, allowing for therapeutic interaction

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2014, Part III, LNCS 8515, pp. 704–713, 2014.

through physical touch inputs and responses [5]. Sparky and Feelix are mobile robots with actuated faces, each with 4 degrees of freedom [6], [7]. Kismet is an anthromorphic head with 21 degrees of freedom that can produce complex facial expressions in response to user inputs [8]. Sage is also a social robot that serves as a robotic tour guide while adjusting behavioral parameters over time based on external interaction [9]. Olivia, is another robotic tour guide that can inform and entertain visitors [10]. Targeting one of the public health epidemics, Autom^{TM} demonstrates its use as a weight loss coach [11].

NAO is a commercial robotic platform that is often employed for various research and education applications. For example, some recent studies employed NAO in social training for children with autism spectrum disorders. NAO has a combination of lights, vocal cues, and motions to interact with children. While simple behaviors are autonomously generated, more complex motions were controlled by researchers monitoring the process. In addition, NAO has the capability to record video and detect touches on its head. By utilizing an intuitive graphical user interface (GUI), NAO allows clinicians to interact wirelessly with the users and has shown success in clinical studies involving children with autism spectrum disorders [12]. Paro and NeCoRo are designed to provide companionship to older people [13], [14], [15]. iCat is another commercially available platform that can recognize objects and faces, recognize speech and sound, and generate various facial expressions [16]. Similar to Paro, iCat was also tested for its potential benefits to the elderly population. The results showed that older people are more comfortable and more expressive with a more sociable robot than with a less social one. In the design of iCat, all emotion expressions are enabled through facial movement and voice generation. However, one existing problem is that while interacting with iCat, there will be no direct body contact between iCat and the human user, which may limit the range and type of interaction. While many existing social robots have proven their effectiveness in entertaining, assisting, and providing service to human users, the cost and maintenance of such robots may discourage many from considering the purchase. The commercial price of Paro is about $6,000 and Nao costs over $15,000. Furthermore, personalized HRI is still a challenging problem to be addressed.

This paper presents Philos, a low-cost social robot for use in a broad range of applications that involve personalized human-robot interaction. The estimated commercial price of Philos is less than $3,000, where all associated software can be available for free when used for research or educational purposes. Philos can interact with users via touch, face detection/recognition, and motion detection. Philos is actuated by eight servo motors that can generate various gestures and simple facial expressions using moving eyebrows. The behavioral control of Philos is based on two elements: 1) internal characteristics of the robot and 2) external vision and touch inputs provided by the users. Internal characteristics determine the robot's initial personality and set the behavioral control parameters accordingly. Therefore, Philos can engage each user in personalized interaction via uniquely defined behavioral responses tailored for each user.

2 Hardware and Software Design

2.1 Hardware Design and Control Scheme

Philos is capable of performing a wide range of simple behavioral motions, including nodding or shaking its head, waving, flapping its arms, and moving the eyebrows through actuation of eight servo motors: two for each arm, two servos enabling the head to pan and tilt, and two for eyebrows [17]. Moving eyebrows allow Philos to generate simple facial expressions, representing three emotional statuses: positive, neutral, and negative as shown in Fig 1. The servos are controlled by an mbed ARM® core microcontroller through serial communication. Philos utilizes 14 force-sensitive resistors (FSR) that cover its chest, head, hands, and feet. These FSRs allow Philos to detect where the robot is touched as well as determine whether it is an aggressive or gentle touch. Philos also has a small speaker installed in its body chassis. This speaker is controlled via the mbed device to playback prerecorded sound clips including human voices, music, or penguin sound. The robot also has two cameras on its head for face detection, face recognition, and motion detection. The exterior covering of Philos is designed to resemble a penguin. Inside the plush outer surface, there is a thin plastic shell to protect the inner components of the robot and to attach the FSRs so that they read more accurately.

Philos utilizes two microcontrollers working in tandem. Philos uses an mbed ARM® core microcontroller for behavioral control and voice generation. A Raspberry Pi is used to read the raw image data and to apply the face and motion detection algorithms. These microcontrollers were chosen due to their respective strengths in terms of processing speed and cost. The Raspberry Pi has fairly high processor speed for an embedded controller and uses a Linux based operating system enabling the use of OpenCV. FSRs are low-cost and widely available sensors can effectively detect touch and its magnitude up to 100N. Each sensor has a surface area of 1.5×1.5 inch^2. The mbed controller receives and processes analog data from the FSR clusters and the noise from the circuit is accounted for by implementing a sampling rate to the FSRs. The servos are controlled via pulse-width modulation (PWM) in order to move the eyebrows of Philos to express emotions. The movement of the arms and head of Philos is controlled by six AX-12 servo motors via half-duplex serial communication. This communication protocol allows

Fig. 1. Internal view of Philos and three emotional statuses (positive, neutral, and negative) represented by the eyebrow angles and hand gestures

the servo motors to be connected in series and for multiple servos to be controlled with a single command. This particular servo greatly reduces the cost of using multiple servos as only two of the GPIO pins must be dedicated to AX-12 control.

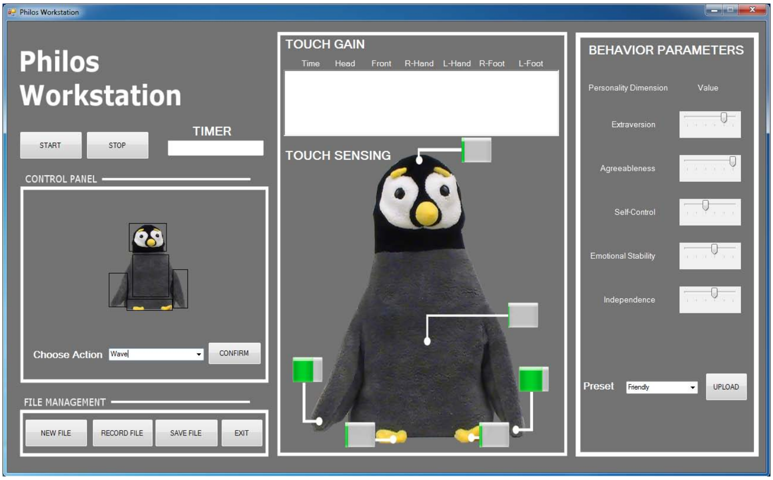

Fig. 2. Philos workstation for programming behavior parameters and real-time monitoring of HRI

2.2 Software Interface

The Graphical User Interface (GUI) allows the user to personalize the robot and monitor real-time interaction data (Fig. 2). The GUI is designed to enable non-technical users to easily reprogram the robot when desired. Real-time monitoring data is also realized in the GUI by displaying current interaction data. In addition to numerical data, a graphical representation of the force applied on different parts of the body is overlaid on a picture of the robot in the main section of the GUI. The data collected during interaction can also be exported to a text document with time stamps if further analysis is desired. Reprogramming, manual control, and data collection are enabled by wireless Zigbee technology using a USB dongle connecting an XBee with the computer.

3 Human-Robot Interaction Design

There are two elements that influence the behavioral characteristics of Philos: 1) the internally defined personality and 2) external inputs provided by human users to Philos. An operator can initially specify a personality type for Philos which will generate a unique set of behavior parameters. Philos behavioral responses are also affected by external user inputs provided through touch-based and vision-based interactions enabled by onboard sensors.

3.1 Generation of Internal Characteristics

Methods: Based on the big five dimensions of human personality as defined by several psychologists [18], we consider five predefined personality types that are adapted for a sociable robot application: Friendly, Hyperactive, Sensitive, Shy, and Cold, as described below:

- **Friendly:** The robot tends to seek out interaction. It is prone to positive increases in behavior, but resistant to negative changes.
- **Hyperactive:** The robot aggressively seeks out interaction. It is not exceptionally prone to changes in behavior.
- **Shy:** The robot avoids interaction and is not prone to behavior changes.
- **Cold:** The robot avoids interaction. It tends to respond negatively to external inputs and is resistant to behavior changes.
- **Sensitive:** The robot neither avoids nor aggressively seeks interaction. It is very prone to behavior changes and is easily affected by user inputs.

Each personality type is classified by a predefined set of values assigned to each of the following personality dimensions: Extraversion (EXT), Agreeableness (AGR), Self-Control (SC), Emotional Stability (ES), and Independence (IND). The values range from 1 to 5 where a value of 1 means the personality dimension is weakly displayed and 5 means the dimension is strongly displayed. These internal characteristics generate the following behavioral parameters of Philos:

- **Room scan frequency** (f_{scan}): The frequency at which the robot will scan the room for faces when it has not recently detected one.
- **Face track probability** (p_{track}): The likelihood that the robot will follow a subjects face after the face has been detected.
- **Frequency of idle state activity** (f_{idle}): The frequency at which the robot considers itself idle during a period of no human interaction, and will exhibit some action to draw attention to itself.
- **Range of idle state activity** (d_{idle}): The number of behaviors the robot may exhibit when it has been idle for a period of time, which is defined by the idle behavior frequency.
- **Level of positive behavioral response** (r_p): A higher value indicates a higher probability that the robot will respond positively to external inputs provided by a user and will be more inclined to seek out interaction.
- **Behavioral change factors** (c_{inc}, c_{dec}): The factors that determine the magnitude that the above parameters will either increase due to positive external inputs or decrease due to negative ones.

Preliminary Testing of the Algorithm: To evaluate the effects of internal parameters on Philos' behavior and user interaction on Philos' behavior, a simple laboratory test was conducted. Behavioral dimension values, (EXT, AGR, SC, ES, IND), for each personality are defined as: Friendly (4, 5, 4, 3, 3), Hyperactive (5, 4, 2, 2, 5), Shy (1, 2, 3, 4, 1), Cold (2, 1, 5, 5, 2), and Sensitive (3, 3, 1, 1, 4). Holding r_p constant, 200 "gentle" and 200 "harsh" touch inputs were provided

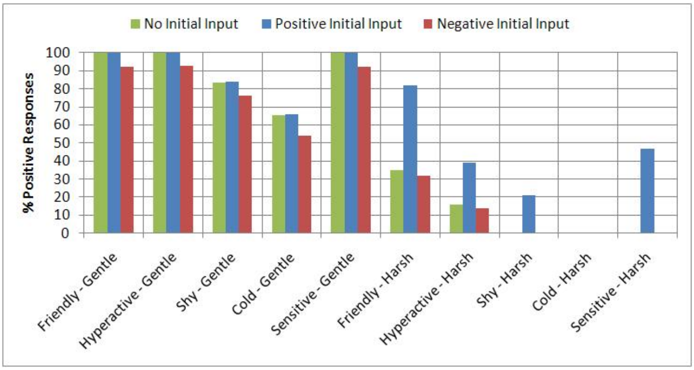

Fig. 3. Percentage of positive and negative responses as affected by user input and programmed personality

based on the predefined threshold for the force sensors. The number of positive responses enacted by Philos was recorded. In order to determine the probability that a positive response would occur after a series of either positive or negative user inputs, we conducted the following test. For each predefined personality, ten touch inputs were provided while r_p was allowed to change. All ten were either "gentle" or "harsh." After the first ten, an additional 200 inputs were given, the first 100 "gentle" and the next 100 "harsh", while r_p was held constant. The results of the tests where the first 10 inputs were positive and the tests where the first 10 inputs were negative are also both plotted in Fig. 3.

The data presented in Fig. 3 shows how Philos' behavior is affected by either positive or negative user input. Personalities with a high value for AGR (i.e. Friendly, Hyperactive) are more significantly affected by positive inputs than negative. The opposite is true for personalities with a low AGR value (i.e. Shy, Cold). Furthermore, the lower the value of ES, the more drastically the behavior will change. This explains why the effects of positive input cause Philos' behavior to change a similar amount for both the Sensitive and Friendly personalities, even though AGR is higher for the Friendly personality type.

3.2 User-Based Interaction

Touch-Based Interaction: Touch based interaction is realized through clusters of FSRs that cover the hard shell of Philos. These clusters are located on the body, the hands, the feet, and the top of the head of Philos. The FSR clusters are created by connecting multiple FSRs in parallel. Touches are first categorized as either being harsh or gentle. The threshold values for gentle and harsh touches can be prespecified or determined by initial parameter training.

Real-Time Face Tracking: Initial face detection uses the AdaBoost Classifier, Haar classifiers, and skin color based algorithm [19], [20], [21]. However, these

Algorithm 1. Face Tracking

loop Command from controller
  % Use frame difference to find edges of moving object
  $F_{i-1} \leftarrow$ *Capture Frame* at t_{i-1}; $F_i \leftarrow$ *Capture Frame* at t_i
  $M_i = F_i - F_{i-1}$ where moving objects in the M frame are white on black
  if Average(Center(M_i)) != [0, 0] **then**
    Move toward the Center(M_i)
  end if
  Preprocessing: RGB to GRAY, Equalization
  Optimize the potential searching area (Algorithm 2)
  Face Detection using Haar cascades and Skin color filter
end loop

Algorithm 2. Optimize Potential Searching Area

loop Camera moved to follow the object
  $V_i = (P_i - P_{i-1})/\Delta t$
  $P_{i+1} = P_i + V_i \times \Delta t$
  if Face is not found **then**
    Increase Searching Area by 10% until the face is found
  else
    Pan Camera Searching
  end if
end loop

methods are often not suitable for embedded, real-time tracking. To reduce the detection time, two strategies are applied: 1) reducing the searching area within the image and 2) optimizing the potential area by projecting the face location in the future frame. First, the original image in the RGB space is transformed into a gray scale and then smoothened by the equalization process. Secondly, the potential face area that projects the future face location is estimated assuming that the user's face will move without erratically changing direction and speed using the results from face detection in current and previous frames. Typically, a *potential face area* indicating possible face locations in the next time frame is slightly larger than the detected face window. Depending on the speed of face movement, the size of the potential face area is dynamically determined as described in Algorithm 2. For example, if the face moves quickly within the image in two consecutive frames, the potential area for searching is increased accordingly. Otherwise, if the face moves slowly, the potential area is reduced. Increasing the potential area by 20% of the actual face size and decreasing the searching area by 10% ensures that if the face either moves closer or farther from the camera it will not be outside the potential area, and accounts for the face moving left or right in relation to the camera. In this way, we can significantly lower the searching time. Table 1 compares computational times depending on

Table 1. Detection time with various sizes of the searching area applied. PA: Potential Area, SA: Searching Area

Searching area	Original	120% PA & 80% SA	120% PA & 90% SA
50*50	985ms	433ms	376ms
80*80	1303ms	553ms	462ms
160*160	2997ms	739ms	507ms
400*400	7814ms	1190ms	696ms

the size of the searching area. Algorithm 1 and 2 shows our overall face tracking strategies.

Face Recognition: Face recognition is enabled by the Principal Component Analysis (PCA) and adaptive learning algorithm for gradually improved performance presented in [21]. To first obtain a standardized image of a person, the image is rescaled to 320×240 pixels and an ellipse shade is applied to eliminate the hair. Ten images for each person are taken and an average image is calculated. By comparing the difference between a user and these average images it is possible to determine if the person is in the database and if they are not in the database, he or she can be added to it.

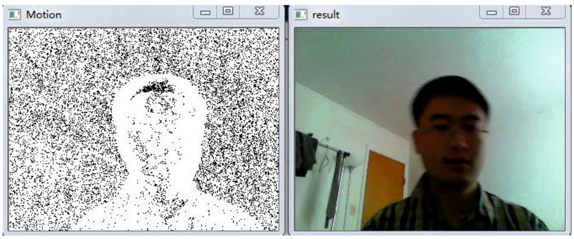

Fig. 4. The processed image and motion detection result

Motion Detection: Motion detection aims at face detection of moving objects/persons at a relatively low resolution. For example, if someone moves toward Philos from a distance, the robot may not be able to obtain images that are suitable for face detection. However, the motions can be detected and tracked. In order to accomplish this, two consecutive frames of the image are compared to calculate the differences between the two frames [22]. The images are first converted into a gray scale to further reduce the processing time and THEN compared for each pixel. If two pixels in two images show the same value, it is marked as 0. If not, it is marked as 1. Then the center of the areas marked with 1's is calculated. Fig. 4 shows this process.

4 Conclusion

In this paper, we presented Philos, a social robot for personalized social interaction. Personalization is realized by predefined internal characteristics of the robot and HRI based on external user inputs. Data collection and processing is performed on board. Philos can provide users a low-cost platform for various education and research applications with its estimated mass-production cost of about $3,000. Building on the current prototype, we are developing the next generation of Philos with improved hardware and user interface design. Furthermore, speech recognition is one of the important areas of exploration while it is omitted in the current version of Philos.

Acknowledgement. This work was supported by the Clinical and Translational Science Collaborative (CTSC) of Cleveland.

References

1. Pearce, A.J., Adair, B., Miller, K., et al.: Robotics to enable older adults to remain living at home. J. Aging Res. 2012, 538169 (2012)
2. Robins, B., Dautenhahn, K., te Boekhorst, R., Billard, A.: Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Univ. Access Inf. Soc. 4, 105–120 (2005)
3. Colton, M.B., Ricks, D.J., Goodrich, M.A., Dariush, B., Fujimura, K., Fujik, M.: Toward therapist-in-the-loop assistive robotics for children with autism and specific language impairment. In: AISB Symposium: New Frontiers in Human-Robot Interaction (2009)
4. Dautenhahn, K., Nehaniv, C., Walters, M., Robins, B., Kose-Bagci, H., Mirza, N., Blow, M.: KASPAR: a minimally expressive humanoid robot for human-robot interaction research. Applied Bionics and Biomechanics 6(3) (2009)
5. Stiehl, W., Lieberman, J., Breazeal, C., Basel, L., Cooper, R., Knight, H., Lalla, L., Maymin, A., Purchase, S.: The huggable: A therapeutic robotic companion for relational, affective touch. In: Proceedings of the IEEE Consumer Communications and Networking Conference, Las Vegas, NV, p. 15 (2006)
6. Ferrari, E., Robins, B., Dautenhahn, K.: Therapeutic and educational objectives in Robot Assisted Play for children with autism. In: IEEE International Symposium on Robot and Human Interactive Communication, pp. 108–114 (2009)
7. Scheeff, M., Pinto, J., Rahardja, K., Snibbe, S., Tow, R.: Experiences with sparky, a social robot. In: Socially Intelligent Agents, pp. 173–180 (2002)
8. Breazeal, C.: Emotion and sociable humanoid robots. International Journal of Human-Computer Studies 59(1-2), 119–155 (2003)
9. Nourbakhsh, I.R., Bobenage, J., Grange, S., Lutz, R., Meyer, R., Soto, A.: An affective mobile robot educator with a full-time job. Artificial Intelligence 114(1), 95–124 (1999)
10. Niculescu, A., van Dijk, B., Nijholt, A., Limbu, D.K., Lan See, S., Li, H., Wong, A.: Socializing with Olivia, the youngest robot receptionist outside the lab. In: Ge, S.S., Li, H., Cabibihan, J.-J., Tan, Y.K. (eds.) ICSR 2010. LNCS, vol. 6414, pp. 50–62. Springer, Heidelberg (2010)

11. Kidd, C., Breazeal, C.: A robotic weight loss coach. In: Proceedings of the National Conference on Artificial Intelligence, pp. 1985–1986 (2007)
12. Shamsuddin, S., Yussof, H., Ismail, L., Hanapiah, F.A., Mohamed, S., Piah, H.A., Zahari, N.I.: Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO. In: 2012 IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA), pp. 188–193 (2012)
13. Wada, K., Shibata, T.: Social effects of robot therapy in a care house - change of social network of the residents for two months. In: IEEE International Conference on Robotics and Automation, Rome, Italy, pp. 1250–1255 (2007)
14. Kidd, C., Taggart, W., Turkle, S.: A sociable robot to encourage social interaction among the elderly. In: IEEE International Conference on Robotics and Automation, pp. 3972–3976 (2006)
15. Libin, E., Libin, A.: New diagnostic tool for robotic psychology and robotherapy studies. Cyber Psychology and Behavior 6(4), 369–374 (2003)
16. van Breemen, A., Yan, X., Meerbeek, B.: iCat: an animated user-interface robot with personality. In: Proceedings of the fourth International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 143–144 (2005)
17. Hornfeck, K., Zhang, Y., Lee, K.: Philos: A Sociable Robot for Human Robot Interactions and Wireless health Monitoring. In: Symposium on Applied Computing, 2012 (Extended Abstract) (2012)
18. Digman, J.: Personality structure: Emergence of the five-factor model. Annual Review of Psychology 41, 417–440 (1990)
19. Wu, X., Gong, H., Chen, P., Zhi, Z., Xu, Y.: Intelligent household surveillance robot. In: Robotics and Biomimetics, ROBIO 2008, pp. 1734–1739 (2008)
20. Liu, Q., Peng, G.Z.: A robust skin color based face detection algorithm. In: International Asia Conference on Control, Automation and Robotics, pp. 525–528 (March 2010)
21. Zhang, Y., Hornfeck, K., Lee, K.: Adaptive face recognition for low-cost, embedded human-robot interaction. In: Lee, S., Cho, H., Yoon, K.-J., Lee, J. (eds.) Intelligent Autonomous Systems 12. AISC, vol. 193, pp. 863–872. Springer, Heidelberg (2012)
22. Yang, J., Waibel, A.: A real-time face tracker. In: Proceedings of 3rd IEEE Workshop on Applications of Computer Vision, pp. 142–147 (1996)