HUMANOID ROBOT PROGRAMMING THROUGH FACE EXPRESSIONS

ALVARO URIBE-QUEVEDO, HERNANDO LEON
Industrial Engineering, Military Nueva Granada University, Cra 11#101-80
Bogota, Colombia, Zip: 110111.

BYRON PEREZ-GUTIERREZ
Mechatronics Engineering, Military Nueva Granada University, Cra 11#101-80
Bogota, Colombia, Zip: 110111.

Accessibility plays a fundamental role for interacting with several current technologic devices, advances in this area are focused in features for natural interfaces that benefits regular and handicapped users. Robot motion and programming requires configuring a set of actions to be executed by each joint, in humanoid didactic kits these tasks are achieved from preloaded commands, or sliding bars for positioning each joint. Current user interfaces based on image processing allows complementing robot motion and programming through gestures without requiring mouse or keyboard thus, making more appealing the experience for students by offering a more natural form of interaction through face or hand gestures. This project implemented a low-cost framework for programming a humanoid robot with 14 degrees of freedom through facial and hand gestures as means to facilitate accessibility to robot motion and programming without the keyboard or mouse. Results proved that moving and programming sequences with gestures is attractive no new users, however, actions based on eyebrow movement such as click or double click, along error in the expression detection affected user interactions, however, interactions through hand motion was more comfortable and was only affected by the sensors precision.

1. Introduction

Didactic humanoid robotics is characterized for its simply motion and programming which has been a success for attracting non-specialists to robotics [1]. Many available kits on the market offer various types of robots whose features vary from Degrees of Freedom (DOF), modularity and ease of programming [2] [3] [4]. Interaction with these robots is commonly done through keyboards, mouse or remote controls with preprogrammed functions for moving each joint, however, newer form of user interfaces are already being used as alternatives to traditional means of input, image processing,

neuroheadsets or even electromechanical sensors are ways to overcome some problems associated with occupational health and accessibility [5].
Current technological advancements in image processing and computing power have yielded to noninvasive gesture interactions using depth of field cameras (Primesense and Kinect) for tracking users [6] [7]. Ease of access has expanded through computers, tablets and gaming benefiting users with some sort of disability [8], in rehabilitation [9], in computer interaction [5], in navigation [10] and in educational aids [11]. Particularly in robot programming, the keyboard and mouse are important elements; input data often requires alphanumerical entries along with navigation through the user graphic interface (GUI).

In robotics several systems have been using the Kinect, Yanik et al. used a growing neural gas algorithm for improving robot response from recognized gestures [12]. Intuitive robot operation has also become a field of interest, Marinho et al. proposed a control system based on the dual quaternion framework, for operating a robotic arm performing pick and place tasks based on hand tracking [13]. This trend in using Kinect in robotics shows how the gap for a natural interaction o even human motion mapping to several types of robots are currently taking place [14] [15].

This work presents the implementation of a face and hand gesture-based tool for programming a humanoid robot as an alternative to keyboard, keypad or even body gestures. The goal of this project is to offer a programming alternative to traditional inputs for students learning robotics. The paper is organized in sections as follows, in Section 2 a brief review of face expression characteristics and how the Kinect tracks them; in Section 3 the humanoid robot characteristics is presented; in Section 4 the system architecture, development and integration is covered; in Section 5 the results and adjustments are presented; and finally in Section 6 the conclusions are discussed along future works.

2. Gesture tracking

Gesture tracking allows natural forms of interactions as it takes advantages of our ergonomics, face and hand gestures can be programmed as alternative inputs. The human face can express numerous emotions that are possible thanks to the muscles, ligaments and tendons allowing changes in the forehead, eyebrows, nose, eyelids, libs and chin [16]. Facial expression can be involuntary or controlled across different scenarios, for our implementation, the focus is centered on voluntary muscle motion given that, a particular set of muscle motion will determine what command is executed.

The Kinect sensor allows tracking faces and skeletons through several APIs [17] [18]. The tracking is done in real time; the sensor algorithms calculate the head and body position and facial expressions. The origin of the tracking coordinate system is centered in the sensor; in the case of face tracking 83 facial points are detected and mapped them over a grid as presented in Figure 1; in the case of skeletal tracking only arms are of interest, the sensor tracks the shoulders, elbows and neck as presented in Figure 1.

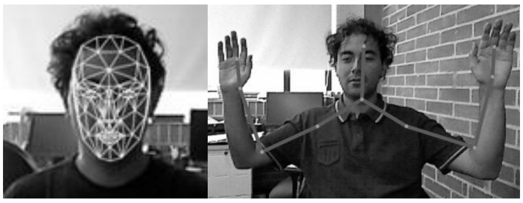


Figure 1: 3D mesh from the tracked face expressionless and seated skeleton tracking.

3. Humanoid robot

The robot has 14 servo motors for 2 DOF in the arms and 2 DOF in the legs, thus mimicking the human form. Figure2 presents the configuration of the humanoid which allows it for performing tasks such as raising the arms, walk, sit and stand up. The structure was constructed using aluminium for reducing the weight and maintain rigidity, due to the fact that the motions to be executed are slow, dynamic effects are not considered and the structure’s goal is to support servos and allow the robot to move as previously stated. The distribution of elements is symmetric so the weight is uniformly resulting in a gait suitable device that does not require advanced control for walking.

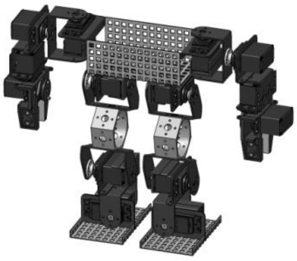


Figure 2: Designed humanoid robot

The main control system is connected to the controllers of the servos through a RS485 bus. The programming is achieved through its proprietary software communicating to the computer through the RS232 serial port. The motion sequences and code is generated from moving each servo to the desired position thus, creating a list sequential order for executing. The data is managed and configured using an Atmega128 microcontroller with a boot-loader which allows user to change the code and access directly to the servo controller parameters. The microcontroller also allows retrieving the output signals from the Humanoid Robot servos; these provide information regarding the actual servos angular position, angular velocity, DC current and voltage. The robot is powered by high-current Lithium-polymer rechargeable batteries, which are located in its pelvis, the cells last for about 30 minutes of operation offering and acceptable autonomy for a learning environment.

The robot can run also autonomously, this type of operation deals with all kinds of movements, storage functions, online posture adjustment and feedbacks. The robot has a remote control which manages several numbers of programs sequentially. In addition, to the required aesthetics of the robot, the method of actuation of joints also plays a large part in determining mechanical design.

4. System development and integration

The proposed system architecture consists of the user, the Kinect, a computer with the servo programming tools and the humanoid robot. Mouse interactions are mapped to *X Y* motions related to pitch and yaw head rotations and hand motions, click actions are configured to be accomplished by pulling up the eyebrows or pushing closer the hand to the camera. These events were selected as sufficient after analyzing the graphic user interface for programming the humanoid robot which offers the selection of configurable commands through buttons, sliders for configuring the position, direction and servo identification, and check buttons for gain, performance and time response aspects. Even though the Kinect has an operational field of view covering 2 m from the sensor, after tracking several persons, 0.5 m was chosen to be an appropriate distance as closer positions to the camera or incorrect posture generate tracking error information given the sensor limitations as pitch optimal tracking is less than 10º, roll less than 45º and yaw less than 30º [19], when using the seated tracking mode the detection presented problems when the user distance was below 0.5 m.

Tracking information was analyzed for defining which face feature would allow mapping click actions without causing discomfort or fatigue. The SDK allows through its shape units (SU) and animation units (UI) to identify expressions based on the *Candide 3* model [20], however, minor movements of the eyebrows, lips and mouth can be falsely detected just by inclining the head over any axis, which would lead to an unsatisfactory interaction. After detecting these non-present expressions, the next step was to determine which of the facial features would provide comfortable motion for mapping the click action. Within a social environment, talking is very common, so considering lip and mouth movements would be counterproductive, even during work, several involuntary mouth motions is present, smiling and yawing are some examples. By disregarding the lips and the mouth, the only significant feature left are the eyebrows, which only by extreme surprise are fully extended upward on the forehead. Eyebrow tracking is significantly better recognized and was chosen for act as a mouse button with a configured range of action proportional to tracked face; it is detected by the AU4 equal to -1 which recognizes the eyebrows completely rose. The system configuration is presented in Figure 3 where the head and hand motion is mapped for horizontal and vertical mouse scrolling.

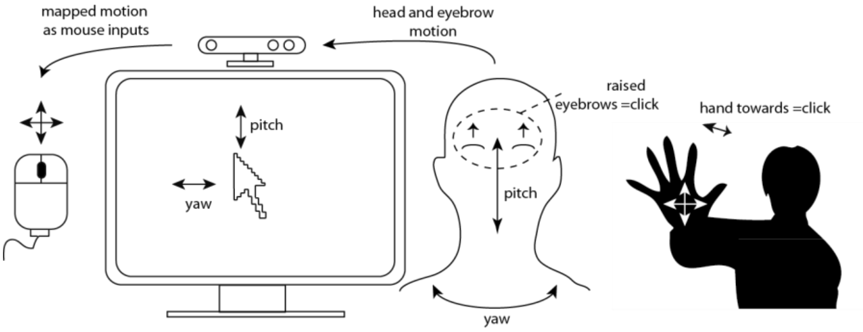
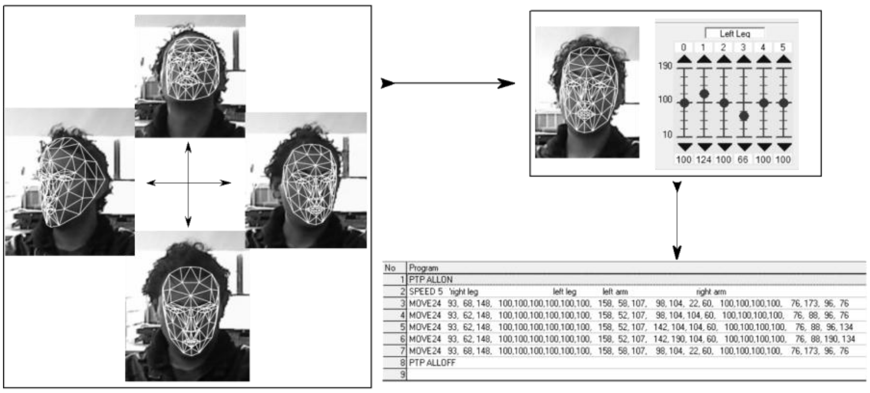


Figure 3: Motion mapping interaction diagram

5. Results

The implemented gesture interaction was tested for validating its suitability, comfort and potential as an alternative for moving and programming a humanoid robot through its interface, the set up goal is to move and program arm motion for rising them up, at this stage, neither efficacy nor efficiency were analyzed. Mouse scrolling across the windows was successfully achieved; however, when tracking the face a scaling factor was required for reaching all screen corners without having to perform full head rotations that were not recognized by the sensors, this limitation was not encountered when using the hand. Given the simplicity of the GUI for programming the robot, the left mouse button action mapped to eyebrow movement was sufficient for the selecting ratio buttons, checkboxes and slider controls. Mouse cursor control based on face tracking resulted satisfactory; users moved the cursor around the screen comfortably, although continuous clicking resulted in eyebrow fatigue given the number of DOF of the humanoid robot; however this difficulty was overcome when using the hand tracking. Tracked positions, mouse control over robot GUI programming, and robot upper member motion are presented in Figure 4 and Figure 5, were each servo's motion is configured and the code generated.



No	Program
1	PTP ALLON
2	SPEED 5 'right leg left leg left arm right arm
3	MOVE24 93, 68, 148, 100,100,100,100,100,100, 158, 58, 107, 98, 104, 22, 60, 100,100,100,100, 76, 173, 96, 76
4	MOVE24 93, 62, 148, 100,100,100,100,100,100, 158, 52, 107, 98, 104, 104, 60, 100,100,100,100, 76, 88, 96, 76
5	MOVE24 93, 62, 148, 100,100,100,100,100,100, 158, 52, 107, 142, 104, 104, 60, 100,100,100,100, 76, 88, 96, 134
6	MOVE24 93, 62, 148, 100,100,100,100,100,100, 158, 52, 107, 142, 190, 104, 60, 100,100,100,100, 76, 88, 190, 134
7	MOVE24 93, 68, 148, 100,100,100,100,100,100, 158, 58, 107, 98, 104, 22, 60, 100,100,100,100, 76, 173, 96, 76
8	PTP ALLOFF
9	

Figure 4: Programming sequences selected using head motion.

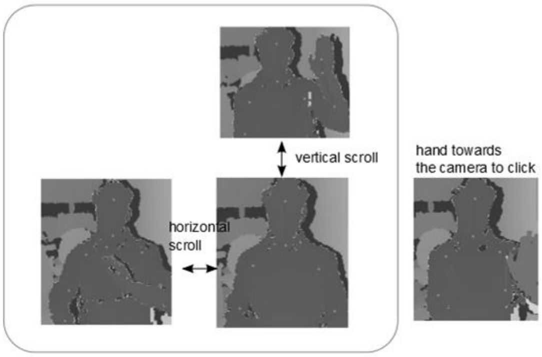


Figure 5: Programming sequences selected using hand motion

The resulted motion sequence in both scenarios was successful and its representation is presented in Figure 6.

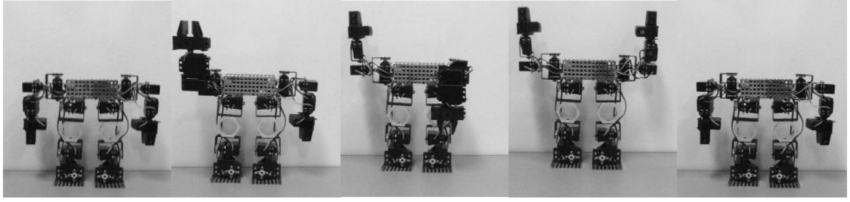


Figure 6: head motion and facial expression based robot programing

6. Conclusion

Alternative user interfaces are offering new ways of interaction that can be exploited in different scenarios other than entertainment, the availability of SDKs offer the opportunity for developing and applying these user interfaces accordingly to cases that can benefit from natural interactions. The implementation presented in this work allowed programming a humanoid robot without needing traditional input devices such as the keyboard and the mouse. In the same manner that repetitive task should be avoided when using the mouse and keyboard, head tracking and clicking using pitch and yaw along eyebrows pull up, can cause discomfort for long periods of using. Even though, for simply programming the humanoid robot none of the users expressed discomfort while moving the servos, however concerns were expressed because using the head as a pointer was a new experience. These inconvenient were addressed when using the hand tracking feature implemented, which was found more natural and no concerns were manifested. Future works will be focused on alternative interaction and voice commands for improving user interaction.

Acknowledgments

The authors thank the Military Nueva Granada University for their support under project ING1188, and undergraduate student Cesar Guerrero from the Multimedia Engineering program.

References

[1] P. Fiorini, "LEGO kits in the lab [robotics education]," *Robotics Automation Magazine, IEEE,* vol. 12, pp. 5-, Dec..

[2] C. N. Thai and M. Paulishen, "Using Robotis Bioloid systems for instructional Robotics," March.

[3] Kyosho, *Manoi,* 2013.

[4] LynxMotion, *LynxMotion Bipeds,* 13.

[5] S. Hakiel, "Delivering ease of use [software development]," *Computing Control Engineering Journal,* vol. 8, no. 2, pp. 81-87, april 1997.

[6] W.-N. Lie, H.-W. Shiu and C. Huang, "3D human pose tracking based on depth camera and dynamic programming optimization," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, 2012.

[7] Z. Zhang, "Microsoft Kinect Sensor and Its Effect," *MultiMedia, IEEE,* vol. 19, no. 2, pp. 4-10, feb. 2012.

[8] J. Abascal, "Human-computer interaction in assistive technology: from "Patchwork" to "Universal Design"," in *Systems, Man and Cybernetics, 2002 IEEE International Conference on*, 2002.

[9] A. da Gama, T. Chaves, L. Figueiredo and V. Teichrieb, "Poster: Improving motor rehabilitation process through a natural interaction based system using Kinect sensor," in *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, 2012.

[10] M. Tonnis, V. Broy and G. Klinker, "A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers," in *3D User Interfaces, 2006. 3DUI 2006. IEEE Symposium on*, 2006.

[11] A. Sherstyuk, D. Vincent, J. J. H. Lui, K. Connolly, K. L. Wang, S. Saiki and T. Cauclell, "Design and Development of a Pose-Based Command Language for Triage Training in Virtual Reality," in *3D User Interfaces, 2007. 3DUI '07. IEEE Symposium on*, 2007.

[12] P. Yanik, J. Manganelli, J. Merino, A. Threatt, J. Brooks, K. Green and I. Walker, "Use of kinect depth data and Growing Neural Gas for gesture based robot control," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*, 2012.

[13] M. Marinho, A. Geraldes, A. Bo and G. Borges, "Manipulator Control Based on the Dual Quaternion Framework for Intuitive Teleoperation Using Kinect," in *Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), 2012 Brazilian*, 2012.

[14] J. Ekelmann and B. Butka, "Kinect controlled electro-mechanical skeleton," in *Southeastcon, 2012 Proceedings of IEEE*, 2012.

[15] R. El-laithy, J. Huang and M. Yeh, "Study on the use of Microsoft Kinect for robotics applications," in *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*, 2012.

[16] W. F. Paul Ekman, Unmasking the face, Malor Books, 2003.

[17] OpenNi, *Open the standard framwork for 3D sensing,* 2013.

[18] E. Suma, B. Lange, A. Rizzo, D. Krum and M. Bolas, "FAAST: The Flexible Action and Articulated Skeleton Toolkit," in *Virtual Reality Conference (VR), 2011 IEEE*, 2011.

[19] Microsoft, *Use the power of Kinect to change de world,* 2013.

[20] J. Ahlberg, *Canide a parameterized face,* 2013.