www.czasopisma.pan.pl PAN www.journals.pan.pl

BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 62, No. 3, 2014
DOI: 10.2478/bpasts-2014-0065

ELECTRONICS

Aspects of Microsoft Kinect sensor application to servomotor control

M. PARZYCH*, A. DABROWSKI, and D. CETNAROWICZ

Poznan University of Technology, Chair of Control and System Engineering, Division of Signal Processing and Electronic Systems,
24 Jana Pawla II, 60-965 Poznan, Poland

Abstract. This paper presents the design process of a gesture control system based on the Microsoft Kinect sensor. An environment enabling implementation of the integrated system using a variety of equipment and software was selected and prepared. A method for integrating the sensor with the Arduino environment has also been discussed. Algorithms for remote gesture control of the given servodrive angle and the position of the robot arm gripper were prepared. The results of several experiments, which were carried out in order to determine the optimal method for starting, controlling, and stopping the drive and for assessment of the accuracy of the proposed method for the arm control, are presented.

Key words: gesture-based control systems, MS Kinect, Arduino, servomotor control.

1. Introduction

Due to expansion of technologies enabling realization of various biometric systems [1], intelligent vision systems [2], and convenient human-machine interfaces, a growing interest in solutions can be seen that enable gesture control. Effective control of a machine with the use of gestures requires selection of appropriate gestures that allow transmission of the precision in the control signals. The natural language of communication consists of expressions, thus in the case of gestures it is also needed to look for a proper dictionary of them. The richer this dictionary, the greater the machine control capabilities. Each gesture should be unambiguously interpreted. Assuming that the final verification of the gesture control scenario requires experimentation, the authors propose a solution to the experimental prototype gesture control system.

The implementation of the system is based on a widely available Microsoft Kinect sensor and the Arduino platform, which allows fast prototyping of control systems. The results of the analysis of several scenarios of control, obtained with the proposed system, are presented in the paper.

2. Applications of gesture control systems

Gesture control systems can be classified basing on their various applications. One of the most popular applications includes human-computer (human-machine) interfaces, which allow hands-free operation of industrial systems, software and peripherals. Such solutions were proposed in [3, 4], and [5]. The authors of paper [6] proposed the use of the gesture control system to operate the television set, in [7] an attention was paid to the possibility of using gestures to control the interactive whiteboard, and in [8] gestures were used to support interactive desks. Moreover, gesture recognition can be used to control smart home systems, such as the lighting control [9]. The use of gestures to control mobile robots is also a popular approach [10]. Tracking and recognition of movements of the whole body can be used to control humanoid robots and this has been done by the authors of papers [11] and [12].

Another criterion for categorization of these control systems is the chosen method of implementation of the gesture recognition. One of them is the use of sensors recognizing hand movements. Sensors can be placed directly on the hand or glove equipment for gesture recognition. This solution was used by the authors of paper [9]. In [13] the information on the values of the accelerations measured by the accelerometers mounted on the wrist were used. However, the most common gesture recognition method is the use of cameras and machine vision algorithms [2, 3, 5, 7, 10]. Some authors expand the vision system by adding a second camera [4], or using depth cameras [6], such as, e.g., in the Kinect sensor [11, 12].

Possible gesture control applications depend on a number of various visual signs recognized by a system. Many signs can be expressed by fingers of a hand [14]. Signs can be expressed by arm also [15]. A range of possible gestures depends on sensor technology, which can be applied in a recognition system.

3. Hardware and software used in the implementation and testing

During the implementation of the gesture-based control system, hardware and software were selected, corresponding to different stages of the operation. The system diagram is shown in Fig. 1.

*e-mail: marianna.parzych@put.poznan.pl

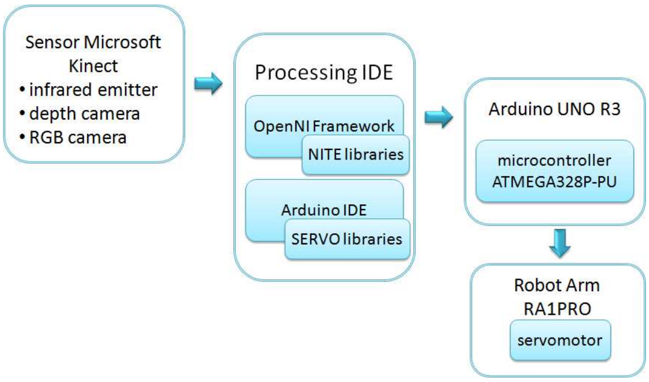

Fig. 1. A block diagram of the system

This system is based on the first generation of the Microsoft Kinect sensor. This is a device that provides functionality of a 3D camera and it is equipped with a microphones grid, that allows for location of a sound source. The Kinect sensor consists of an RGB camera, depth sensor, and four microphones. An RGB camera provides three channel data in a 1280×960 pixels resolution at 12 frame per second or a 640×480 pixels resolution at 30 frames per second [16]. The depth sensor consists of an infrared laser emitter and a monochrome CMOS (complementary metal-oxide-semiconductor) sensor, which captures video data. The infrared (IR) emitter emits infrared light beams grid that is captured by the IR sensor in order to measure the depth. Based on the distance between the sensed grid points, the depth map with a 640×480 pixels resolution and 11-bit depth (which provides 2 048 levels of precision) is calculated. The Kinect depth sensor range is minimum 800 mm and maximum 4000 mm. It is also possible to use the Near Mode where the range is minimum 500 mm to maximum 3000 mm [17]. Limited range and interference caused by intense sunlight are the main disadvantages of this device, but they are not relevant to the solution proposed in this paper.

The four microphones grid that comes with the device operates with each channel, processing 16-bit audio at the sampling rate of 16 kS/s. These microphones are equipped with built-in noise filters. In addition, the Kinect sensor is equipped with an automatic tilt controller that can be used to automatically adjust the position of the sensor according to the size of the tracked objects. The Kinect construction, described in [18], is shown in Fig. 2.

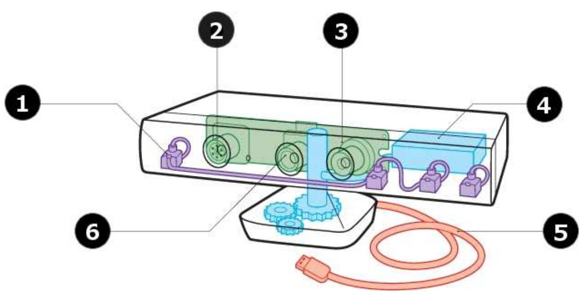

Fig. 2. Construction of Kinect sensor: microphones (1), infrared emitter (2), depth sensor (3), tilt controller (4), cable (5), RGB camera (6)

The last element of the system is an educational robot Arexx Robot Arm Pro. This robot is controlled by the ATMEGA64 microcontroller. The I/O interfaces are compatible with the I2C bus system. The USB interface supports the uploader software. The robot arm length is 390 mm, base diameter is 210 mm and the whole device is 460 mm high. The robot is powered by 6–12 V power supply. A six servomotors system is responsible for the movement of the robot. The Arexx Robot Arm Pro, distributed by [19], is shown in Fig. 3.

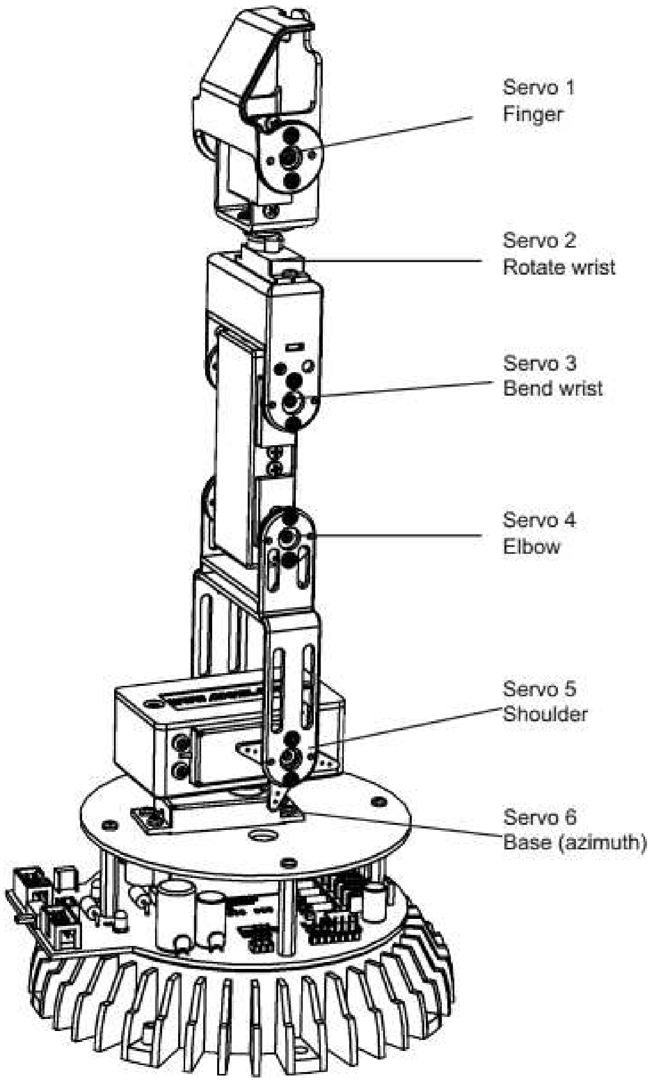

Fig. 3. Arexx Robot Arm Pro in calibration position

A typical development environment dedicated to the implementation of algorithms based on the information provided by the Kinect sensor is the MS Kinect Software Development Kit. However, this software has not been used during our research because of a difficult access to the detailed documentation. This environment does not support moving applications on non-Windows platforms (Linux, Mac), and the license does not allow for modifications of libraries and sharing resources.

For the implementation of the gesture control system the Processing IDE [20] – a multiplatform environment enabling integration of different devices, was selected. This environment is equipped with an OpenNI (*open natural interaction*) framework [21] and NITE library [22]. The OpenNI is a tool for acquiring the depth map and the RGB picture from the Kinect cameras and for implementation of algorithms for video. The NITE library extends the functionality of the OpenNI with the hand and body tracking.

In the developed system the Processing IDE is integrated with the Arduino IDE environment, which makes it possible to operate the developed drive with a microcontroller ATMEGA328P-PU.

4. Implementation

The developed program for the servo control via the Kinect sensor can be divided into three stages: initialization of the control by the defined start-gesture, controlling the servo by the position of the hand in the picture, and the control closing (the stop gesture). The block diagram of the whole operation process is shown in Fig. 4.

As a start/stop gesture three different gestures were tested: "Wave", "RaiseHand" and "Click". They are illustrated in Fig. 5. Two ways of controlling the rotation of the servo drive were tested: circular and horizontal movements of the hand. Depending on the position of the hand, the given angle of the servo drive is calculated in the range from 0° to 180°. Both tested control gestures are presented in Fig. 6.

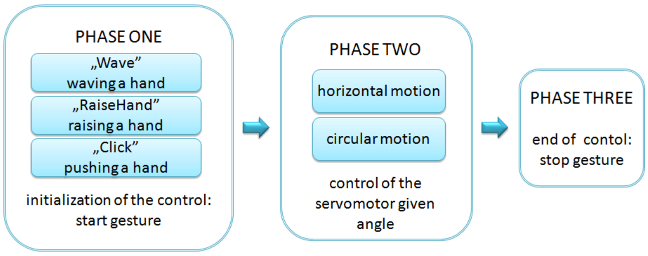

Fig. 4. Operation steps of the developed application

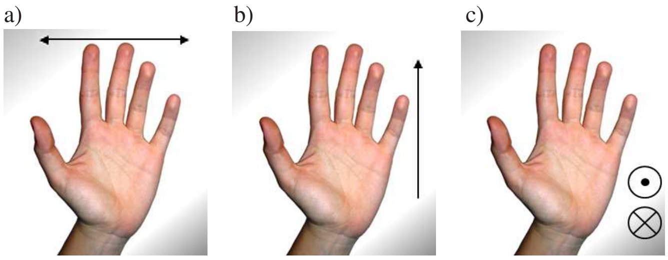

Fig. 5. Selected start/stop gestures: a) "Wave", b) "RaiseHand", c) "Click"

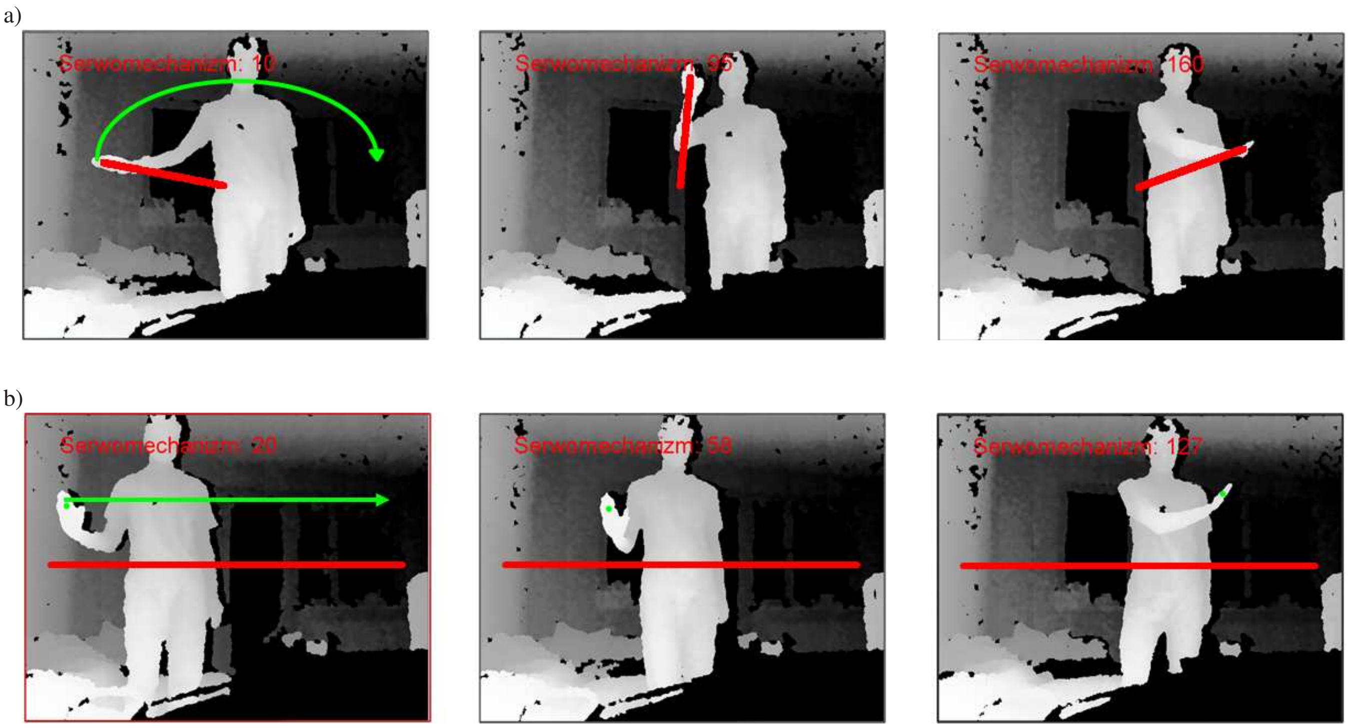

Fig. 6. Selected control methods: a) circular motion, b) horizontal motion

5. Experimets

5.1. Single servomotor control with the sequence of gestures. During experiments the gesture sequences consisting of three phases were studied: the starting gesture, the servodrive angle control gestures, and the ending gesture. When the system recognizes the starting gesture, it proceeds to the second phase, i.e., to interpreting the hand movements as signals for the control of the servo rotation. All movements must be precise, because residues of the movement elements in the first phase can be interpreted as a part of the second phase movements – Fig. 7. The aim of this study was to select such gestures for the start and the stop that minimize the presence of residues of the first phase movements in the second phase. Figure 8 shows fluctuations in the given angle, which result from residues of the starting gesture. The unwanted fluctuations are depicted in Fig. 8 by shadows. The better the control scenario, the lower these fluctuations.

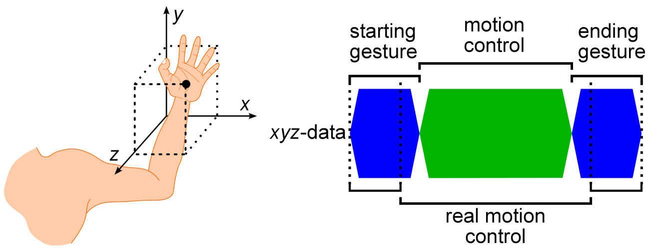

Fig. 7. The effect of asynchronous interpretation of input signals – ending part of starting gesture is interpreted as a motion control

Using the "Wave" gesture in the first phase (Fig. 8a and 8b), significant oscillations resulting from the interpretation of the waving a hand as the second phase of the control sequence are observed. On the contrary, using the "RaiseHand" gesture in the first phase and the horizontal motion control in the second phase (Fig. 8c), the resulting oscillations are small and thus acceptable. Further reduction of these oscillations, if necessary, can be achieved with filtering of the executive signals. This can be done in the digital domain [23] or even in the analog domain [24, 25].

Movements in extreme positions of the circular motion, if used in the second phase, are similar to the "RaiseHand" starting gesture. The result is an unacceptable shift shown in Fig. 8d, which disqualifies this control scenario.

The use of the gesture "Click" (Fig. 8e and 8f) is free from the defects observed in the above-mentioned scenarios. For both, the horizontal and the circular motion control, to calculate the required servodrive angle, two-dimensional coordinates are used and the "Click" gesture takes place in a direction perpendicular to the control motion, regardless of the control method. Thus basing on the conducted analysis, the "Click" gesture should be considered as the optimal gesture for starting and stopping the control. Indeed, with the "Click" start gesture the observed movement of the servomotor is smooth and precise regardless of the control gestures in the second phase. However, during the tests the circular motion control occurred to be more convenient and more intuitive than the horizontal motion control, as the circular motion control allows to control the speed of movements of the servodrive not only with the speed of the hand movement but also through selection of the radius of the circle. This is an additional advantage of this control scenario.

5.2. Hand tracking accuracy. It is planed to expand the developed system to control a robot arm with six degrees of freedom. This involves controlling the position of the robot arm gripper with the position of the operator's hand. In order to determine feasibility of this concept, a series of tests of the hand tracking precision was performed.

During the tests the operator moved his hand along the track in three directions: horizontal, vertical, and depth. The aim of the experiments was to determine the undesired deviations of the hand position during the control process of the servomotor. The goal was to achieve as small deviation as possible in a given scenario. The adopted coordinate system is shown in Fig. 9. Three kinds of moves were tested: along x-axis, along y-axis, and z-axis. In each case the operator tried to keep the remaining coordinates constant i.e. y, z, x, z, and x, y, respectively. The results are shown in Fig. 10. The maximum deviation of x while moving along the vertical track (Fig. 10c) is 25 mm and while moving into the depth (Fig. 10e) amounts up to 46 mm. Taking into account that the full range of motion is 1.5 meters, the deviation is 1.6% and 3%, respectively. Overall deterioration in performance was caused by natural vibrations, wavings of the hand in the horizontal direction and accuracy of the Kinect sensor. The maximum deviation of y while moving along the horizontal track (Fig. 10a) and while moving into the depth (Fig. 10f) is 20 mm (1.3%). The maximum deviation of z while moving along the horizontal track (Fig. 10b) and the vertical track (Fig. 10d) is 30 mm (2%).

The presented study showed that the Kinect sensor can not be used to control high-precision servodrive arrangements, needed in such applications as, e.g., the teleoperation. However, the Kinect control is feasible for applications with modest precision requirements, such as the use of the controlled arm to carry and move objects. The average achieved deviation of the three coordinates is 22 mm, which is a sufficient accuracy for the considered application.

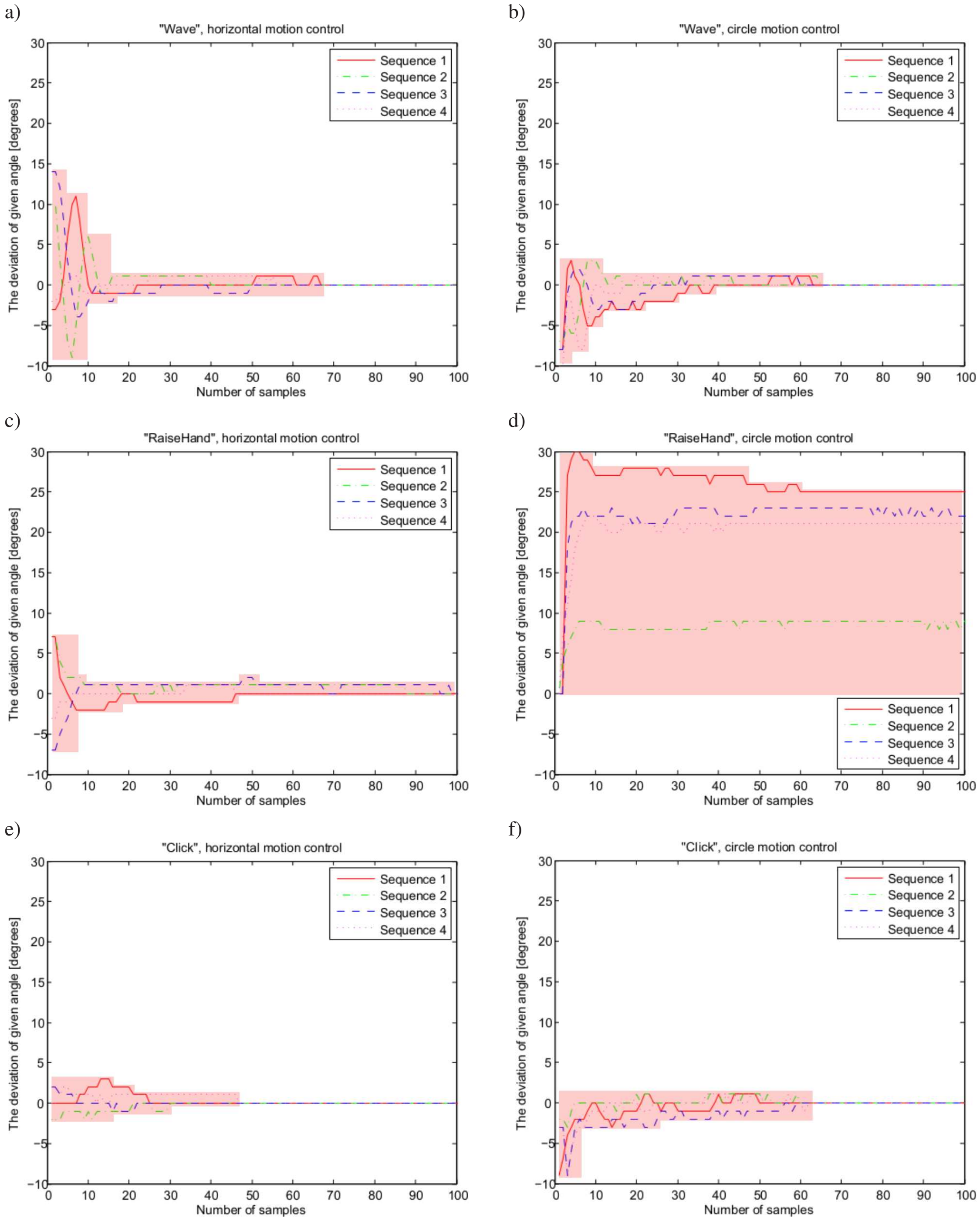

Fig. 8. Fluctuations of given angle for different control methods (sequence repeated 4 times): horizontal motion control with "Wave", "RaiseHand", "Click", circle motion control with "Wave", "RaiseHand", "Click"

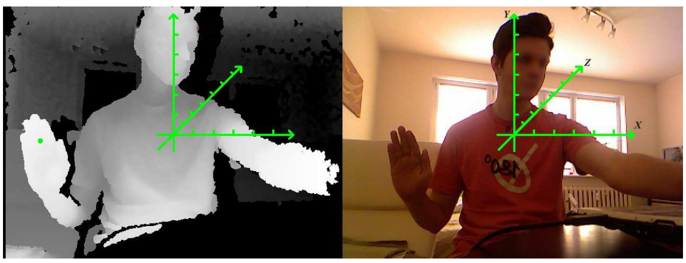

Fig. 9. Adopted coordinate system

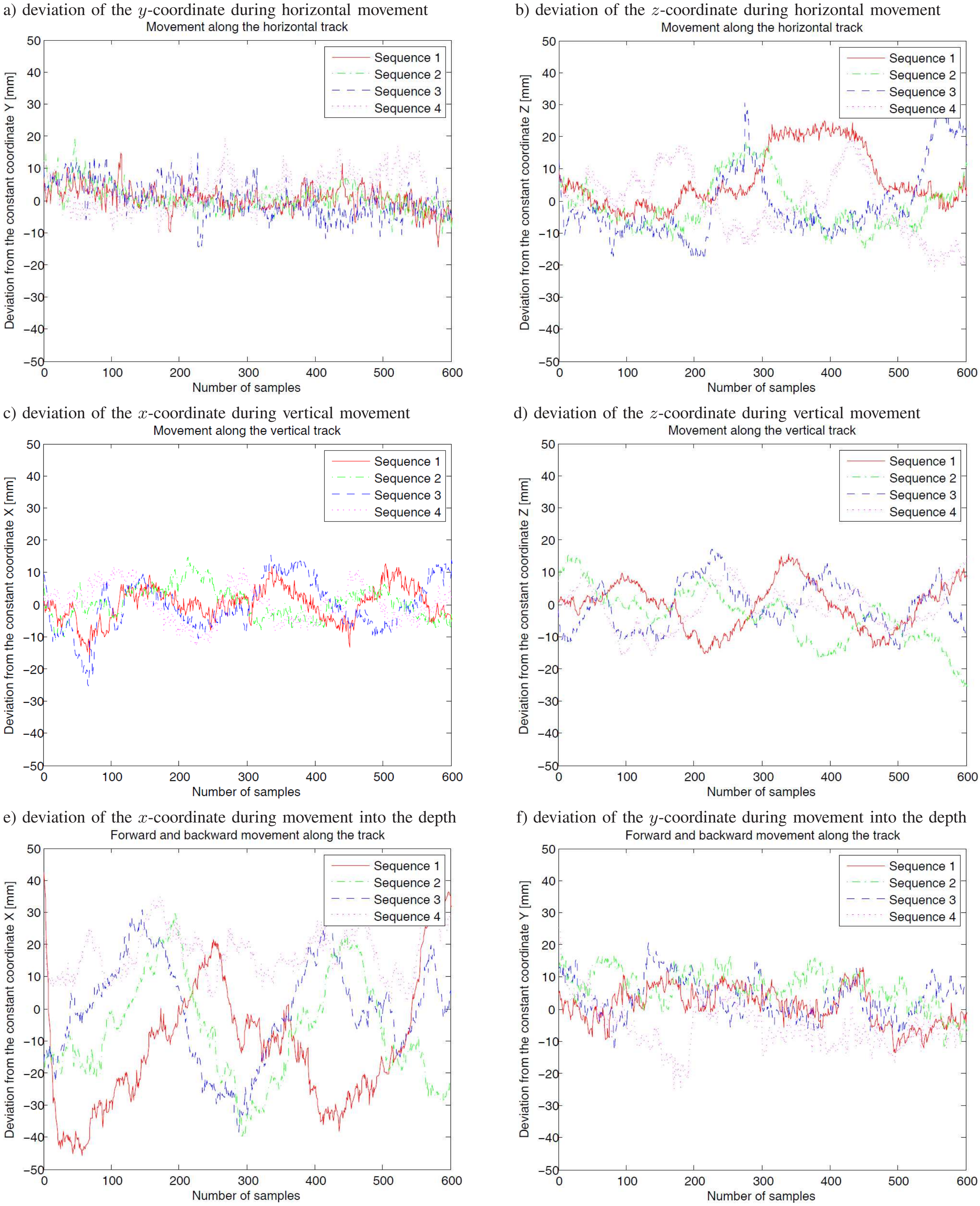

Fig. 10. Deviation of coordinates of tracked hand while moving in three directions: horizontal, vertical, and in depth

6. Future work

In the future it is planned to expand the system to control the robot arm. It will be based on the tracking of the operator's hand and exploitation of coordinates to control the robot gripper. Location of the remaining degrees of freedom of the robot will be calculated using the inverse kinematics.

The performed study showed that the Kinect sensor has a sufficient accuracy for the considered servodrive control.

7. Conclusions

The authors designed a servomotor control system and studied the impact of the start, stop, and motion control gestures on the precision of the drive in order to find the optimal scenario of gestures.

It was found that the circular motion meets the control precision requirements in the best way and it provides an easy, intuitive, and convenient method for the servodrive control including control of the rotation speed.

The gesture "Click" was found to be the optimum starting and stopping method because of its negligible impact on reduction of the precision of the servodrive motion control.

Acknowledgements. The paper was prepared with the DS-PB 09-93/DS-PB 0412 (2014) funds.

REFERENCES

[1] T. Marciniak, A. Dabrowski, A. Chmielewska, and A.A. Krzykowska, "Selection of parameters in iris recognition system", *Multimedia Tools and Applications* 68 (1), 193–208 (2014).

[2] P. Pawlowski, K. Borowczyk, T. Marciniak, and A. Dabrowski, "Real-time object tracking using motorized camera", *Proc. 13th IEEE Conference on Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)* 1, 110–115 (2009).

[3] S. Kumar and J. Segen, "Gesture based 3D man-machine interaction using a single camera", *Proc. Int. Conf. on Multimedia Computing and Systems* 1, 630–635 (1999).

[4] A. Prieto, F. Bellas, R.J. Duro, and F. Lopez-Pena, "An adaptive visual gesture based interface for human machine interaction in inteligent workspaces", *Proc. Int. Conf. on Virtual Environments, Human Computer Interfaces and Measurement Systems* 1, 43–48 (2006).

[5] Y.V. Parkale, "Gesture based operating system control", *Proc. Int. Conf. on Advanced Computing and Communication Technologies* 1, 318–323 (2012).

[6] D. Ionescu, B. Ionescu, C. Gadea, and S. Islam, "An intelligent gesture interface for controlling TV sets and set-top boxes", *Proc. Symp. on Applied Computational Intelligence and Informatics* 1, 159–164 (2011).

[7] M. Lech and B. Kostek, "Gesture-based computer control system applied to the interactive whiteboard", *Proc. Int. Conf. on Industrial Technology* 1, 75–78 (2012).

[8] M. Nishino, Y. Nakanishi, M. Sugi, Y. Soo, J. Ota, and T. Arai, "Design of gesture interface for deskwork support system", *Proc. Int. Joint Conf. ICROS-SICE* 1, 2260–2267 (2009).

[9] B. Mrazovac, M.Z. Bjelica, D. Simic, S. Tikvic, and I. Papp, "Gesture based hardware interface for RF lightning control", *Proc. Int. Symp. on Intelligent Systems and Informatics* 1, 309–314 (2011).

[10] M. Manigandan and I.M. Jackin, "Wireless vision based mobile robot control using hand gesture recognition through perceptual color space", *Proc. Int. Conf. on Advances in Computer Entertainment Technology* 1, 95–99 (2010).

[11] A. Wakabayashi, S. Motomura, and S. Kato, "Communicative humanoid robot control systems reflecting human body movement", *Proc. Int. Symp. on Micro-NanoMechatronics and Human Science* 1, 122–127 (2011).

[12] W. Song, X. Guo, F. Jiang, S. Jang, G. Jiang, and Y. Shi, "Teleoperation humanoid robot control system based on kinect sensor", *Proc. Int. Conf. on Intelligent Human-Machine Systems and Cybernetics* 2, 264–267 (2012).

[13] N. Helmi and M. Helmi, "Applying a neuro-fuzzy classifier for gesture-based control using a single wrist-mounted accelerometer", *Proc. Int. Symp. on Computational Intelligence in Robotics and Automation* 1, 216–221 (2009).

[14] A. Wojciechowski, "Hand's poses recognition as a mean of communication within natural user interfaces", *Bull. Pol. Ac.: Tech.* 60 (2), 955–971 (2012).

[15] A. Babiarz, R. Bieda, K. Jaskot, and J. Klamka, "The dynamics of the human arm with an observer for the capture of body motion parameters", *Bull. Pol. Ac.: Tech.* 61 (4), 955–971 (2013).

[16] *Kinect for Windows: Product Features* (access on 29.09.2013) http://www.microsoft.com/en-us/kinectforwindows/discover/features.aspx (2013).

[17] *Microsoft Developer Network: Kinect Sensor* (access on 29.09.2013) http://msdn.microsoft.com/en-us/library/hh438998.aspx (2013).

[18] T. Kowalczyk, *Kinect SDK – Introduction, Microsoft Knowledge Base*, (access on 29.09.2013), http://msdn.microsoft.com/pl-pl/library/kinect-sdk-wprowadzenie (2013).

[19] *Educational Robot: Robot Arm Pro. Mounting instructions: Model RA1-PRO, Arexx Engineering* (access on 29.09.2013) http://www.arexx.com/robot_arm/html/en/index.htm (2013).

[20] *The Documentation of Processing Programming Environment and Libraries* (11.06.2013), http://processing.org/reference/ (2013).

[21] The documentation of OpenNI Framework (11.06.2013) http://www.openni.org/reference-guide/ (2013).

[22] *NITE Libraries Official Website* (11.06.2013) http://www.primesense.com/solutions/nite-middleware/

[23] P. Korohoda and A. Dabrowski, "Generalized convolution as a tool for the multi-dimensional filtering tasks", *Multidimensional Systems and Signal Processing* 19 (3–4), 361–377 (2008).

[24] A. Dabrowski, A. Menzi, and G.S. Moschytz, "Design of switched-capacitor FIR filters with application to a low-power MFSK receiver", *IEE Proc.-G Circuits, Devices, and Systems* 139 (4), 450–466 (1992).

[25] A. Dabrowski, R. Dlugosz, and P. Pawlowski, "Integrated CMOS GSM baseband channel selecting filters realized using switched capacitor finite impulse response technique", *Microelectronics Reliability* 46 (5–6), 949–958 (2006).