ISSN 2229-5518

Real-time gesture control of a CNC machine tool with the use Microsoft Kinect sensor

Karol Miądlicki[1], Mirosław Pajor[2]

Abstract— The paper presents a control system of the AVIA VC 760 milling machine. Its control system is based on gesture monitoring and recognition using the Microsoft Kinect Motion Controller. The system was implemented in the Matlab/Simulink environment with Kinect for Windows Runtime and Kinect SDK & Developer Toolkit library. To investigate the system, three sets of gestures were developed for the purposes of the study (basic commads and movement control). The motion of the operator's/controller's hand can naturally and intuitively control the table and the spindle. The influence of light, distance from the robot, the speed of gestures performed by the operator on the system was tested. The paper discusses the concept and structure of the system and test results

Index Terms— CNC control, gesture control, machine tool, machine vision, Microsoft Kinect, infrared optics, real-time processing

1 Introduction

While attempts to modify currently used control systems are being made, some try to design virtual environments to use them for remote control purposes in modern factories [1]. They are meant to provide easy and remote operator/machine interaction methods. Gesture control systems might be of particular interest for the industry. A proper selection of gestures, used to give precise commands, enables effective and intuitive interaction with a machine without having to be close to it or to learn how to operate a complicated control panel. Gesture control systems can be used for mobile [2], humanoid [1, 3], flying robots [4] or industrial manipulators [5, 6]. Attempts are made to use them in surgical robots [7] for controlling data displayed on a monitor without having to touch the keyboard or the mouse [8] which would ensure maintaining the sterile field within the operating theatre. In another attempt [9], a virtual environment was used to program a CNC machine using gestures.

Sensors and vision systems are the two main methods used in gesture control. The former relies on accelerometers, gyroscopes and resistors mounted directly on a hand or as an element of a glove or a skeleton [10] used for gesture detection. However, the latter method based on machine vision is more common. Depth data is of particular importance for a comprehensive analysis of the environment surrounding a machine, gesture recognition and the positioning of a workpiece which is being worked on or which is being picked and/or moved. Optical methods for depth sensing do have their drawbacks: they require substantial computational time (stereoscopy), do not provide adequate speed of data processing (laser methods), are expensive (Time of Flight) and are sensitive to external stimuli (structural light) [11].

1. *mgr inż. ZUT Institute of Mechanical Technology, West Pomeranian University of Technology, Szczecin, karol.miadlicki@zut.edu.pl*

2. *dr hab. Inż. Prof. ZUT Institute of Mechanical Technology, West Pomeranian University of Technology, Szczecin, miroslaw.pajor@zut.edu.pl,*

2 Case of study

2.1 Current system of CNC machine control

One of the best sought for skills on the job market is the expertise of controlling CNC machine tools and programming machining operations. Recent years have seen progress in the efficiency and accuracy of machining. A number of CNC control algorithms have been developed. Machine vision systems were introduced to improve precision and to support the reference positioning of the workpiece. However, the methods of communicating with the machine through the operator interface have not been developing at the same pace. Despite developments in the graphic display of CNC interfaces or the introduction of touch panels, current CNC control systems available on the market are not very intuitive. Manual interfaces, in particular, still rely on complicated control panels and require the knowledge of how to use particulate control systems (Siemens, Heidenhein, Fanuc). Consequently, it is necessary to learn a set of procedures, key sequences and to memorise the functions of diodes, switches and what actions are triggered by what procedures. Since machines, CNC drivers and operator panels are often differently configured, it only adds to the current difficulties. As a result, the interface between the operator and the machine has become increasingly complex [12]. Only qualified operators can be employed to operate CNC machines. They must go through advanced training to improve their efficiency and to minimize the risk of error.

2.2 New control concepts

With the advent of readily available computers with huge computing capacity, the development of virtual reality technology and 3D imaging, the application of gestures and human senses to remotely control various machines (haptic feedback) seems to be inevitable [13]. This is evidenced by the current trends on the market of consumer and industrial electronics. While traditional control methods (remote control units, buttons, keys and joysticks) seem to be disappearing,

ISSN 2229-5518

new control solutions allowing the use of gestures, voice [2] and eyes [14] are more frequently used in game consoles, smartphones and TV sets [15]. The next step in the development of control methods is going to be full integration of a user with a device through ordinary conversation. This is already being tested on the latest smartphones and computers using voice assistants: Cortana (Microsoft), Siri (Apple) and Google Now (Google Inc.) [16]. It is only a question of time when the new systems can be implemented into industrial machines.

With the advent of readily available computers with huge computing capacity, the use of virtual reality and 3D environments, the application of gestures and human senses to remotely control various machines (haptic feedback) seem to be inevitable. Attempts have been made to modify existing control methods and to develop virtual environments with the view of executing remote control of machinery in modern factories [17]. Such sys-tems should enable innovative, simple, remote ways of operator-machine interaction. Programming, robot- and machine tool positioning should be performed remotely by the operator, who is able to control several machines while sitting at the control centre. A remote control system would help to fit more machines in existing shop floors. Remote control would make safety zones redundant since all the machine control operations would be performed remotely [18].

2.3 Concept of the proposed control system

The paper presents the control system of the AVIA VC, 760 millign machine, robot, based on gesture monitoring and recognition. Its concept is based on tracking operator's gestures using a Microsoft Kinect Motion Controller. Track point position data acquired by the system is processed by the algorithms of control and gesture recognition. The operator's hand, used as the controller, provides a natural and intuitive control interface which imitates day-to-day operations of picking and placing objects. The system was im-plemented in the Matlab/Simulink environment, using "Kinect for Windows Runtime" and Kinect SDK & Developer Toolkit library.

3 Gesture sensor – Microsoft Kinect

3.1 Relase history

The Microsoft Kinect sensor was unveiled as Project Natal by Microsoft on 1st June, 2009 at the E3 fair. This novel solution enabled its users to inter-act with their console without having to use pads or other control systems, common at that time. The system implemented a new interface using hand gestures, body movements and spoken commands.

The Microsoft Kinect sensor is used in robotics, medicine and IT as it can measure depth indoors more quickly and precisely (for distances up to 2-3 meters) than other methods, e.g. stereoscopy. The sensor is used for remote control, environment scanning and collision avoidance in mobile, humanoid, flying robots and industrial manipulators [19]. Attempts were made to use the sensor to control surgical robots [7] and to control data displayed on the screen without the need to touch the keyboard or the mouse [8] thus maintaining aseptic conditions in the operating theatre. The controller can also be used as part of a physical rehabilitation system. The ability to measure depth and track the position of track points on a human body offered by Microsoft Kinect is used in image processing, including 3D scanners [20], human detection [21] and emotion recognition [22].

3.1 Hardware description

Kinect enables user skeletal tracking using the depth camera and specific data on human kinematics. The points spread on the body (rigging) are processed. The images of rigs are used to teach the algorithm of human motion generation. To enable body recognition for it to be used as the controller, data fed by sensors are processed by the software in several steps. The moment a user appears in front of the Kinect controller, a 3D cloud representing the user is generated using depth data. The model is later fitted to a given typology of the initial skeleton. Using human body kinematics data, Kinect learns to recognize different parts of the user's body and assigns specific weights to each of them depending on how certain the algorithm is of having properly recognized the part. Once the process is finished, Kinect finds the most probable skeleton position against the body and connects the body parts according to their weights [19].

The Kinect sensor consists of six modules: a multi-array microphone, an infrared emitter, a depth camera, an angularity

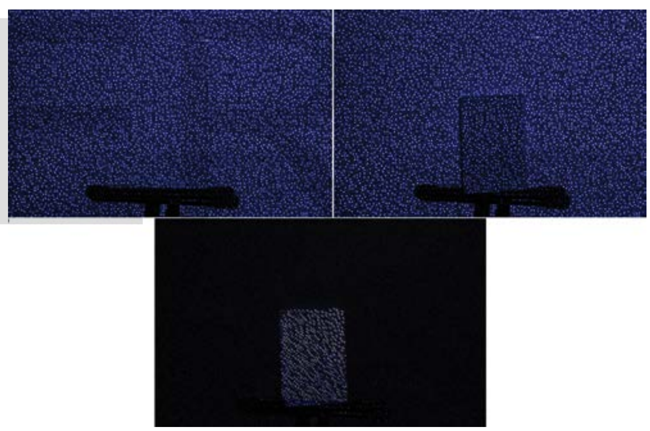

Fig. 1. the points generated by the infrared emitter [23]

controller, connection lead, and an RGB camera. The Kinect sensor has a range of 1.2 - 3.5 meters in the normal mode and 0.8 - 2.5 in the near mode. The lens has an angular field of

Fig. 2. Microsoft Kinect sensor [23]

1. Infrared optics – A projector and sensor map over 48 points on human body
2. RGB camera – The camera combines with the 3D map to create the image
3. Motorized tilt – Mechanical gears at the base
4. Multi- array microphone – Four microphones cancel out ambient noise

ISSN 2229-5518

view of 43.5° vertically, with tilt adjustment of ±27°, and 57° horizontally. The RGB camera features a CMOS sensor and enables user facial recognition and image processing. It operates in two modes: 640x480 at 30 fps and 1280x960 at 12 fps.

3.3 Depth calculation

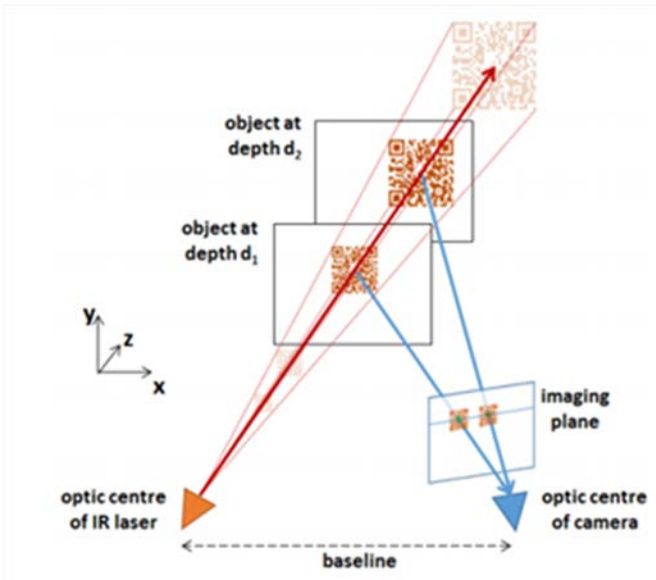

Fig. 3. Calculating the depth of using triangulation [24]

For depth data input, the sensor interprets data stream from the depth camera. The infrared emitter projects a grid of pseudorandom points over a large area. The depth camera (IR CMOS) is capable to detect position with a resolution of 80x60, 320x240 and 640x480, at 30 fps. Next, a triangulation based range sensor measures the distance to a track point. Although traditional triangulation requires two cameras, Kinect uses only one depth camera and the infrared emitter which functions as the other camera. Since it remembers the structure of emitted points, it can be compared with an image provided by the IR CMOS. As the device knows the position of concrete points, a frame of reference points can be inferred. Then, the depth camera records the translated image and uses it to perform discrepancy check of the points. This is how the distance between an object and the sensor is calculated. If the images overlap, depth calculation is easy and can be quickly performed by the processor.

Fig. 4. Stages figure recognition by Kinect sensor [24]

Kinect allows you to track the skeleton user through a combination of camera depth and detailed analysis of the kinematics of the human body. The analysis and determination are subjected to special points (rigs) on the human body. Images were then marked off points are used for the learning algorithm generates the data for human motion. In order to detect the human body and use it as a controller, the data from the sensors are processed by a few steps.

Then, when the user is in front of the Kinect controller, using depth information is generated 3D surface forming a "cloud" user mapping points. For such a design model is assigned an initial skeleton. Then, using data on kinematics of the human body, Kinect learns to recognize parts of our body and assigns them weights depending on how the algorithm is confident recognized parts. When it is completed Kinect finds the most likely position of the frame relative to the body and associated with each part of the body in accordance with their weights (Fig. 4). All points, which allows the sensor to track Microsoft's Kinect SDK version 1.7 are shown in the figure (Fig. 5)

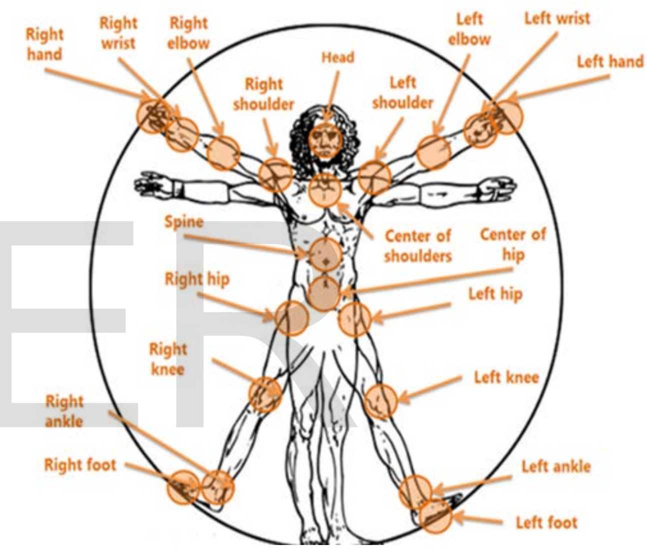

Fig. 5. Tracked points implemented in Kinect SDK 1.7 [23]

3.4 Microsoft Kinect SDK

Microsoft's Kinect controller offers two versions of the SDK. Official coming from the manufacturer, the Kinect for Windows SDK and toolset Windows Developer Toolkit and developed by PrimeSense OpenNI SDK. At the time of writing the two development kits offer similar capabilities, with a slight predominance of official tool.

The Microsoft Kinect SDK includes device drivers, development tools (NUI API) and documentation. This ensures access to complex capabilities of the device in a simple manner, i.e. By tools and libraries that allow developers to create code responds to events, captured from the outside world. This allows you to easily control the behavior of the program. Kinect Sensor with Kinect SDK program allows you to interact with the programmer, as shown in Figure 6. NUI (Natural User Interface) - is an interface between the user and the application, which allows them to interact with one another, without the need for additional input devices. Is intended to provide users from using the application in a natural way for

ISSN 2229-5518

human beings. NUI API capabilities implemented in Kinect SDK 1.7: access to the Kinect sensor connected to a computer, access to data from the RGB camera and depth in the software environment, providing tools to the so-called. skeletal tracking, which is a smooth response to behavior caused by the movement of human skeleton, the ability to scan 3D objects using the "Fusion" and recognize simple gestures. In addition, the installation of additional Windows Developer Toolkit package, it is possible to work with MATLAB / Simulink and OpenCV libraries. Access to a single sensor Kinect at a time, can have only one application.

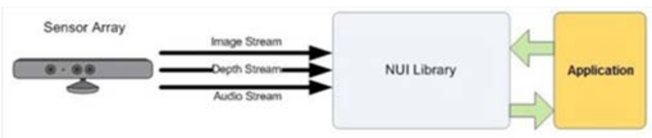

Fig. 6. Communication scheme of the Kinect sensor and SDK

4 Test Stand

4.1 Milling machine - Avia VC 760

VC 760 belongs to the modern machine tools and best in the series machining of small and medium-sized parts. Direct drive without belts from the motors to ball screws allows for positioning accuracy and repeatability. Vertical machining center VC series is built on the principle of cross-table, moving in X and Y and headstock, moving the guides columns in the Z-axis machine uses fully automatic lubrication system. Oil is delivered as a function of time and distance, all the units of the machine, requiring control, such as tracks or ball screws.

Fig. 7. AVIA VC760 machine tool without cabinet

Vertical AVIA centers are equipped with fast, electric storage tools. Drum magazine a 24 tools with rotary changer allows very quick tool changes (2.8 sec.) And is the best solution for batch processing wherever required a large number of tools and fast tool exchange.

Fig. 8. AVIA VC760 machine tool with cabinet [25]

4.2 Test equipment

A VC 760 milling machine with an open control system (O.C.E.A.N.) developed at the Centre for Mechatronics of the West Pomeranian University of Technology, Szczecin, was used for the implementation and tests of the proposed system. The machine has two kinematic chains of body element movement: object and tool branches. It has screw couplings in 3 axes, with power applied directly from the motor. The test stand is presented in Fig. 9

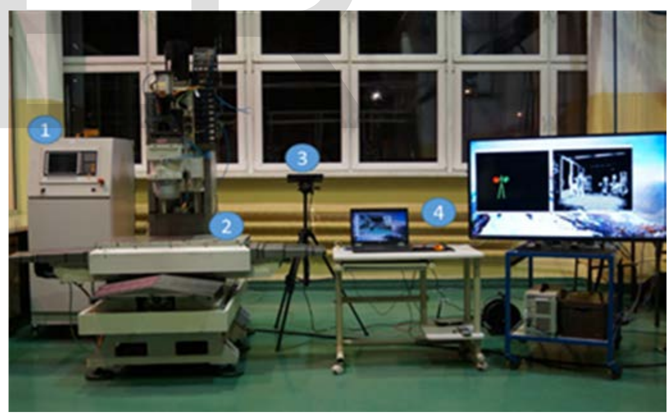

Fig. 9. Test stand

It consists of:

1. O.C.E.A.N. control system, based on Acopos intelligent servo drives manufactured by B&R
2. VC 760 milling machine
3. Microsoft Kinect sensor
4. PC used for image processing and communication with the CNC machine.

4.3 Communication

The Kinect controller communicates with the milling machine according to a scheme presented in Fig. 10. The control of the VC 760 milling machine using an OPC server in the PLC driver (manufactured by B&R) required writing a program. It was written in SFC programming language. Position detection can be done in two ways:

ISSN 2229-5518

- Absolute - the table or the workpiece are moved directly to a position defined by (x, y, z) coordinates
- Incremental - the table or the workpiece are moved by a pre-defined vector from the current position (Δx, Δy, Δz)

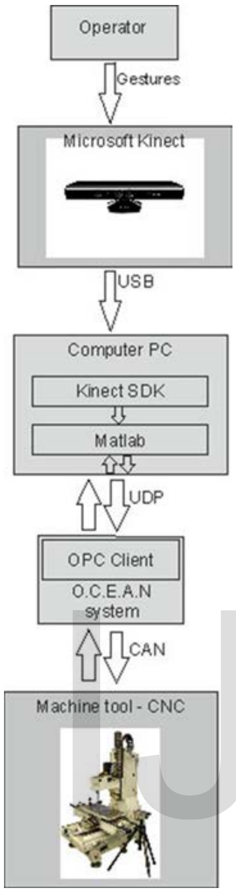

Fig. 10. The communication sheme gesture control system

The second programme, responsible for the Microsoft Kinect sensor operation, gesture recognition and feeding coordinates to the CNC driver, was written in MATLAB environment using "Kinect for Windows Runtime" and Kinect SDK & Developer Toolkit library. The programme consists of 5 modules responsible for:

– Communication with Kinect - receiving data about the position of tracked points.

– Filtering of received coordinates - smoothing tracked points using a simple average filter with a step of 5 samples. The filter turned out to be the best out of all 19 filters presented in [24].

– Gesture recognition - we pre-programmed the following gestures: hands up/down, a hand moving up/down, a fist/an open hand. While algorithms implemented in Kinect SDK were used to detect the first two gestures, KHand library was used for the third gesture

– Position fitting and scaling - this part of the programme is responsible for a good fit of the robotic arm and the operator's hand. It is critical to safety as without the function, at first the robot would make an unknown move to assume the position signaled by the operator's hand.

– Communication with the CNC driver - sending, receiving and scaling of data about the robot position.

4.4 The set of control gestures

Three sets of gestures were developed for the purposes of the study. The first set includes four basic gestures: start, stop, confirm and change precision to be performed by the left hand.

The second set is used to control the table and the spindle - the machine follows the right hand gestures. Hand gestures performed in the air are followed by the table. Up and down gestures control the spindle. The velocity of movement is proportional to the velocity of gestures performed by the operator. The tracking of the operator's hand starts when he/she clenches his/her right hand into a fist and when the operator confirms it using the other hand. From now on, every hand gesture is followed by the respective axis of the machine (Fig. 11). The last set is used to choose the axis we want to control, by showing its number using fingers. Having selected the axis and then having confirmed it with 'confirm' gesture, it is possible to control the machine with the right hand clenched into a fist in the tracking mode.

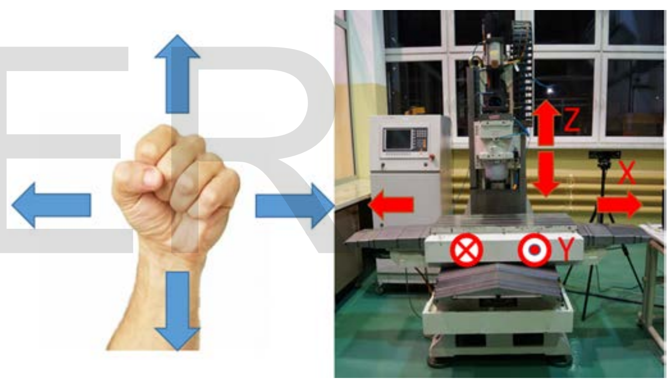

Fig. 11. Axis control in follow mode

5 SUMMARY

5.1 Test results

The proposed system was tested on a test stand in the Technology Hall of the Institute of Materials Science and Engineering, West Pomeranian University of Technology, Szczecin. Experiments revealed that the proposed control system for the VC 760 milling machine was much more intuitive than the standard control panel or the joystick. During the operation of the system, the background did not seem to affect its stability. The system could efficiently control the milling machine both in a homogeneous white static background and in its standard environment (people moving around, objects in the background). The system worked correctly in complete darkness but problem was the sensitivity of the system to light conditions. Since the sensor uses a structural light based method for depth determination, the moment a strong source of light is directed towards it, the sensor cannot operate properly. The

ISSN 2229-5518

readings of hands and other points have either random or maximum values rendering proper control of the system. At the current stage of research, precision and repeatability the robot moves were not measured. However, was observed that it depends mainly on the speed of operator's movements and light conditions. Background does not affect the operation of the system. The system allowed control robot over the static white background and as well as in a normal work environment (running in the background people and objects).

REFERENCES

[1] I. Almetwally, M. Mallem, Real-time tele-operation and tele-walking of humanoid Robot Nao using Kinect Depth Camera, 2013 10th IEEE International Conference on Networking, Sensing and Control (ICNSC), 2013, pp. 463-466.
[2] D. Katzmaier, Samsung Smart Interaction: Hands-on with voice and gesture control, www.cnet.com, 2012.
[3] W. Song, X. Guo, F. Jiang, S. Yang, G. Jiang, Y. Shi, Teleoperation Humanoid Robot Control System Based on Kinect Sensor, 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), IEEE, Nanchang, Jiangxi, 2012, pp. 264-267.
[4] A. Sanna, F. Lamberti, G. Paravati, F. Manuri, A Kinect-based natural interface for quadrotor control, Entertainment Computing, 4 (2013) 179-186.
[5] J. Gośliński, P. Owczarek, D. Rybarczyk, The use of Kinect sensor to control manipulator with electrohydraulic servodrives, Pomiary, Automatyka, Robotyka, 17 (2013) 481-486.
[6] S. Moe, I. Schjolberg, Real-time hand guiding of industrial manipulator in 5 DOF using Microsoft Kinect and accelerometer, 2013 IEEE RO-MAN, 2013, pp. 644-649.
[7] W. Liu, H. Ren, W. Zhang, S. Song, Cognitive tracking of surgical instruments based on stereo vision and depth sensing, 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, Shenzhen, 2013, pp. 316-321.
[8] L. Gallo, A. Placitelli, M. Ciampi, Controller-free exploration of medical image data: Experiencing the Kinect, 2011 24th International Symposium on Computer-Based Medical Systems (CBMS), 2011, pp. 1-6.
[9] K. Stateczny, M. Pajor, Project of a manipulation system for manual movement of CNC machine tool body units, Advances in Manufacturing Science, (2011) 33-41.
[10] M. Pajor, K. Stateczny, K. Pietrusewicz, Virtual reality applied for programming CNC machine tools, Control Engineering, 60 (2013) 50.
[11] M. Hansard, Time-of-flight cameras principles, methods and applications, Springer, New York, 2013.
[12] L. Shuo-Peng, J. Hsu, T. Chun-Chien, Development of the common human-machine interface for multi-axis machine tools, Advanced Intelligent Mechatronics (AIM), 2012 IEEE/ASME International Conference on, 2012, pp. 650-653.
[13] W. Ipke, Gesture in Embodied Communication and Human Computer Interaction, Springer-Verlag, Berlin, 2010.
[14] P. Corcoran, To Gaze with Undimmed Eyes on All Darkness [IP Corner], Consumer Electronics Magazine, IEEE, 4 (2015) 99-103.
[15] Y. Jing, C. Eun-Seok, C. Wook, B. Won-Chul, C. Sung-Jung, O. Jong-Koo, C. Joon-Kee, K. Dong-Yoon, A novel hand gesture input device based on inertial sensing technique, Industrial Electronics Society, 2004. IECON 2004. 30th Annual Conference of IEEE, 2004, pp. 2786-2791 Vol. 2783.
[16] G. Marshall, Cortana: everything you need to know about Microsoft's Siri rival, 2014.
[17] A.K. Bejczy, Virtual reality in robotics, IEEE Symposium on Emerging Technologies & Factory Automation, ETFA, 1996, pp. 7-15.
[18] Y. Xiang, E. Deines, C. Lauer, J.C. Aurich, A human-centered Virtual Factory, Management Science and Industrial Engineering (MSIE), 2011 International Conference on, IEEE Harbin, 2011, pp. 1138-1142.
[19] S. Shirwalkar, A. Singh, K. Sharma, N. Singh, Telemanipulation of an industrial robotic arm using gesture recognition with Kinect, 2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE), IEEE, Jabalpur, 2013, pp. 1-6.
[20] P. Rakprayoon, M. Ruchanurucks, A. Coundoul, Kinect-based obstacle detection for manipulator, 2011 IEEE/SICE International Symposium on System Integration (SII), IEEE, Kyoto, 2011, pp. 68-73.
[21] Y. Shen, Z. Hao, P. Wang, S. Ma, W. Liu, A Novel Human Detection Approach Based on Depth Map via Kinect, 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, Portland, OR, 2013, pp. 535-541.
[22] M. Oszust, M. Wysocki, Recognition of signed expressions observed by Kinect Sensor, 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), IEEE, Krakow, 2013, pp. 220-225.
[23] http://msdn.microsoft.com/en-us/library/jj131429.aspx , stan na 30.06.2014
[24] Teixeira J. M., Reis B., Macedo S., Kelner J., Open/Closed Hand Classification Using Kinect Data. In Proceedings of the 2012 14th Symposium on Virtual and Augmented Reality (SVR), 2012, pp. 18-25.
[25] www.avia.com , stan na 30.06.2015