Proceedings of the Annual Symposium of the Institute of Solid Mechanics
and Session of the Commission of Acoustics, SISOM 2016 Bucharest 12-13 May

ADVANCED CONTROL SOLUTION FOR AN ANTHROPOMORPHIC GRIPPER

Catalin Moldovan, Ionel STARETU

Transilvania University of Brasov, Brasov, email: istaretu@yahoo.com

This paper presents an advanced control solution, used a Kinect sensor, for an anthropomorphic gripper with five fingers intended to be used in industrial robots equipment assemblies used for achieving low and medium complexity. A control mechanism refers to the way of controlling the movements of a gripper using input data captured from a computerized system of human hand gestures detection. A haptic system used to control an anthropomorphic gripper sends data related to the physical properties of an object that is intended to be gripped by the human operator. Currently, there are two techniques used for capturing 3D data from physical environments. They are classified by the capture procedure or by the type of sensors used in active and passive techniques. In the research in the field of digital image processing, recognition of 3D objects involves recognition and detection of 3D objects in an image or a frame, part of a video stream. The Kinect Microsoft sensor combines a set of hardware and software mechanisms that make up a digital 3D representation of a physical environment.
Keywords: anthropomorphic gripper, haptic interface, Kinect sensor, control system.

1. INTRODUCTION

A control mechanism refers to the way of controlling the movements of a gripper using input data captured from a computerized system of human hand gestures detection. Thus, a gripping simulation model can be studied from the perspective of the behaviour of a dynamic system with inputs. It can be described as a control algorithm with close loop in which the elements of the control system and the sensor represent the human hand respectively the visual acuity (see Figure 1). In this case, the error can be measured in terms of visual accuracy, the human user being able to say if the anthropomorphic gripper caught an object or not. The main function of such a system is to control a virtual or physical gripper by reducing estimations made by the mechanism used for gestures capture.

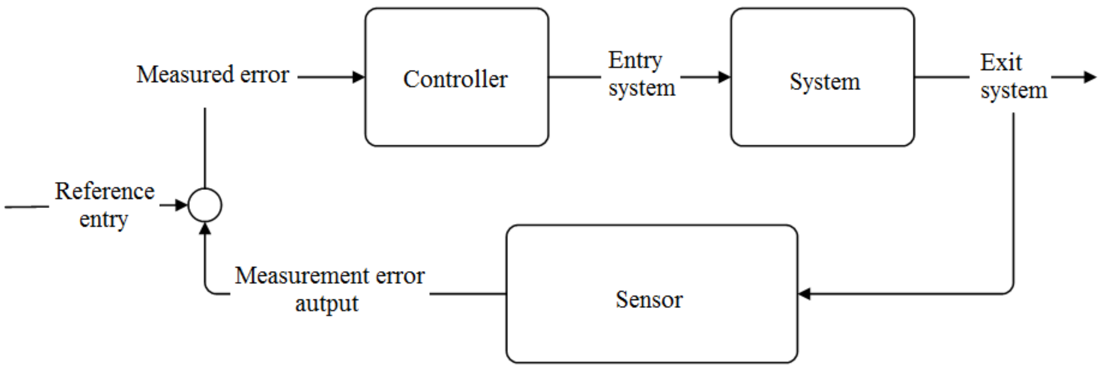

Figure 1. Control system

The effectiveness of these types of systems can be easily demonstrated by the study of existing specialty literature in this scientific field and of the various applications created for this purpose. These include control of vehicles [1] industrial automation, robots in general [2], [3], [4], [5] and robotic gripping in particular [6], [7].

2. ANTHROPOMORPHIC GRIPPER CONTROL USING HAPTIC SYSTEMS

A haptic system used to control an anthropomorphic gripper sends data related to the physical properties of an object that is intended to be gripped by the human operator. The user and the haptic interface should collaborate to enable gripping in the virtual environment. This aspect involves handling the haptic device, which sends sensory information to the user, stimulating his/her tactile, visual or kinematic systems.

Using input and output sensors, a virtual environment may be generated and the objects can be handled directly, through the virtual reality engine, or by the human user. Thus, a general representation of a haptic system of robotic gripping could be considered in the system shown, at concept level, in Figure 2. In this figure, a human operator interacts with a virtual simulator transferring hand movements and gripping force to the virtual environment and he/she can back the force return from the virtual environment through the haptic device.

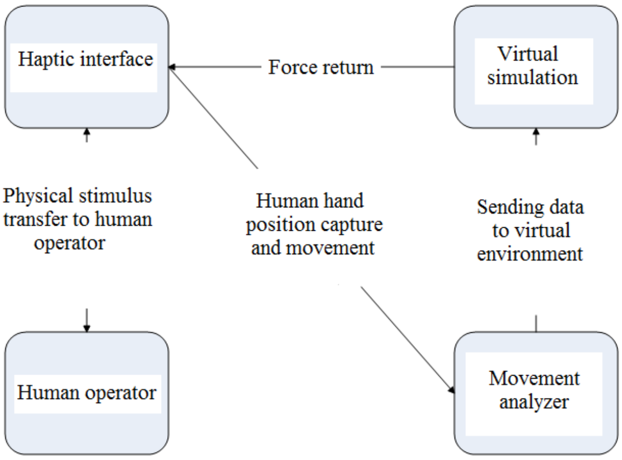

Figure **Error! No text of specified style in document.**. A representation of a generic haptic system and its components

This section describes how to shape, conceptually, the components of a command and control haptic system, of an anthropomorphic gripper, mechanical or virtual. Furthermore, we describe the components of a control haptic system of an anthropomorphic gripper.

Human operator

The human operator is always part of a motion transmission system to a virtual reality system. The human operator can move the haptic device, and its movement can activate a virtual gripper, which, in turn, can interact with different objects in the virtual environment. Each object, and the environment itself, can respond differently to the human operator, based on the physical properties of an object or based on the environment properties. The answer lies in sending return data to the haptic device, which can influence further decisions of the operator.

Using advanced haptic devices, equipped with complex sensors, we can conduct a really immersive and interactive experiment.

Haptic interface

Any haptic interface is made of a mechanic device typically operated by a human user, to communicate with a virtual reality generation system. In addition, a haptic device can send different return data to the human operator.

Such haptic devices are different from case to case, for example, in the robotic gripping study, the data gloves are used successfully. These haptic interfaces capabilities may be enhanced adding specific components to capture aspects of the human hand motion, and of return transmission to their user. In this regard, at present, there are many sensors that can be used with different purposes, e.g.: pressure sensor, heat sensor, and so on, which can measure the position and the force sent from the user to the virtual environment, and sensors that can react to stimuli sent to the environment generated.

The data gloves on the market today fall into two categories taking into account the way in which contact is transferred to the user, which are ways based on force return; ways based on tactile return. The two devices are different at constructive level, from the point of view of costs and functionality.

The devices with force reaction have an exoskeleton attached through which contact data between the virtual hand and a virtual object in the virtual reality system can be sent to the real hand.

Tactile feedback devices have vibration sensors attached, placed at fingertips and palm level. The intensity level of the vibration is variable and can be related to the hand movement in virtual reality or the way of gripping elastic objects in virtual reality.

In terms of costs, tactile return gloves are used more often, because they are cheaper and the device is simpler. However, if we want a complex gripping scenario, in which the force reaction must be precise (e.g.: medical application), then we prefer force feedback devices. Force feedback devices are used when at the contact with a virtual object we want the hand or the finger that touches or virtual rigid object to stop its forward movement applying a counter force through the exoskeleton at the level of the gloves.

Motion analyzer

The motion analyzer is an interface between the haptic device and the virtual environment. On the one hand, it is used to analyze data from the haptic device and to translate them in command and control data for an anthropomorphic gripper movement, and on the other hand, to send return data from the virtual environment to the human operator.

Virtual reality simulators and their evaluation criteria

The virtual reality software system is necessary to create a simulation, in real time, of a gripping operation at a very low cost of implementation. A virtual reality simulator provides a visual interface with the virtual environments. It offers the opportunity to the human operator to explore, to incorporate dynamic simulations and to interact with different virtual objects.

Obviously, these systems are different, from case to case, based on the final outcome that must be measured. Initially, the main features define for comparison, were interactivity and immersion but, with the evolution of research in this area, other criteria were listed. For example, in [8] they study current criteria used for virtual simulators analysis, and a new criterion is added, environment perception. It is a qualitative criterion that refers to how the human operator perceives the environment generated.

Other research, such as [9], [10] and [11] add new comparison elements to the environments generated: fidelity, active participation and real time response.

Systems analysis and methods used to capture depth data

Digital images creation represents the process of data acquisition from the physical environment using different data capture sensors. Digitizing physical environment is not limited to image capture, but it is also extended to areas like: digital processing, image compression, storage or application of artificial intelligence algorithms to images.

Initially, digital images creation was limited to data capture from the physical environment using RGB sensors, but with the evolution at microprocessors and computing power, depth data capture became a necessity and, with the sensors evolution, it became a reality. Obviously, in parallel with the development of advanced sensors to capture depth data software applications were developed to take advantage of the research progress in the area of sensors. Among the first researches with promising results, we can mention ways of depth data capture using triangulation techniques detailed in the papers of [12] or [13]. Other approaches such as measuring reaction time from sensor to object and back, defined first in [14] were successful, but because sensors are expensive, this approach was not available to the public.

At present 3D data capture used in research is based on mixed approach, namely, considering on the one hand progress in the image processing area and on the other hand progress in the sensors technology. Thus, accessible devices were launched on the market, both from the cost point of view, and highly accurate, like Microsoft Kinect and Motion Leap.

A presentation of the latest approaches of research for robotic handling systems analysis using anthropomorphic grippers can be found in the work of [15] and [16] there is a presentation of existing methods of mobile robots control based on digital image processing algorithms.

3. SYSTEMATIZATION OF METHODS USED TO CAPTURE 3D IMAGE

Currently, there are two techniques used for capturing 3D data from physical environments. They are classified by the capture procedure or by the type of sensors used in active and passive techniques.

The passive procedure refers to the use of methods for examining an image from two different angles. The depth calculation is carried out based of analyzing points from the two different angles, using geometric algorithms.

The active procedure refers to the use of light projection (time of flight) or light patterns (structured light) onto a specific area, subsequently, measuring the speed of light coming back to the sensor, or the pattern distortion in the area for the depth calculation [17]. Structured light represents a type of active triangulation [13]. The method used to measure depth through this technique consists in the projection of a predefined pattern onto an area, and subsequently applying triangulation equation between the pattern captured and the reference pattern. A detection system based on the structured light concept is made of a light emitter that can project a pattern and an RGB sensor (CCD or CMOS) used to detect image. Flight time measurement is another method to calculate depth in images and consists in using a source of light projected onto a surface.

The distance from the emitter to the object depends on time [14]. Distance is calculated by the time difference between the light pulse emission and the detection of the reflected light by the sensor.

4. KINECT MICROSOFT DEVICE

Microsoft marketed the Kinect Microsoft device in late 2010. It was originally used as an accessory for the XBox console. Since its inception, the Kinect Microsoft device has proved to be very successful. In the first year, sales of about 10 million units were estimated [18].

Initially, the Microsoft Kinect device was used exclusively for the games industry, but, subsequently it was remarked by researchers who noticed that the sensor can be used as a very accessible alternative, from the point of view of the price, in the process of 3D mapping of space. Soon, PrimSense (OpenNI), which used the sensor for any purpose, not only for games, developed a SDK. Obviously, in time, OpenNI alternatives were developed in parallel, and in 2011, Microsoft launches the first SDK officially, with new opportunities for the sensor use.

One of these opportunities is the use of the sensor in the semantic analysis of images. This opportunity creates new areas of research like applying semanic analysis algorithms to images in robots and automatesystems.

Initially, automate robots that used Microsoft Kinect sensors lacked retroaction, but this drawback was easily overcome, through a solution that uses 3D environment mapping process and the defining of areas inaccessible to the robot and retroaction transmission to the control system for information [19]. This method is used successfully in Stowers book [20], who shows how robots can be programmed to fly autonomously without touching anything.

The use of Microsoft Kinect sensor is not limited to the automation of robotic systems. In addition, this sensor can be used to solve the problem of human hand detection and of recognizing its gestures from the perspective of analyzing the depth image. For this operation, analysis and understanding Microsoft Kinect sensor is a first step.

Constructively, Kinect sensor is made up of (see Figure 3): infrared sensor: emitter and receiver. The emitter projects a light pattern onto a surface, which is subsequently captured by the receiver; RGB camera: which stores data on three (RGB) channels at a resolution of 1280x960 and a frequency of 30 Hz. Kinect sensor visibility field, as specified in the Microsoft documentation [18] is 43 degrees vertically and 57 degrees horizontally. The sensor can monitor people with an accuracy of 1 cm at a distance of 2 m; a system of four microphones to capture sound from different positions; an engine used to lift the sensor without physical interaction between user and sensor; an accelerometer to detect the sensor current tilt reported to the horizon line.

The visibility field and the resolution change with the distance between the object and the Microsoft Kinect sensor [21], thus, the visibility field increases linearly with the distance and the resolution decreases along directions x and y. when distance increases. The visibility field and resolution in pixels can be calculated based on formulas adapted from [22]:

$$CampVizibilitate = 2 * \tan\frac{\theta}{2} * dist \quad , \tag{1}$$

$$Rsl = \frac{Pixel}{cm^2} = \frac{\alpha}{CampVizibilitate} \quad . \tag{2}$$

where: *dist*: is the distance between user and sensor, *Rsl*: is the camera resolution calculated and α : is the maximum resolution of the RGB camera, 1280 horizontally and 960 vertically [18].

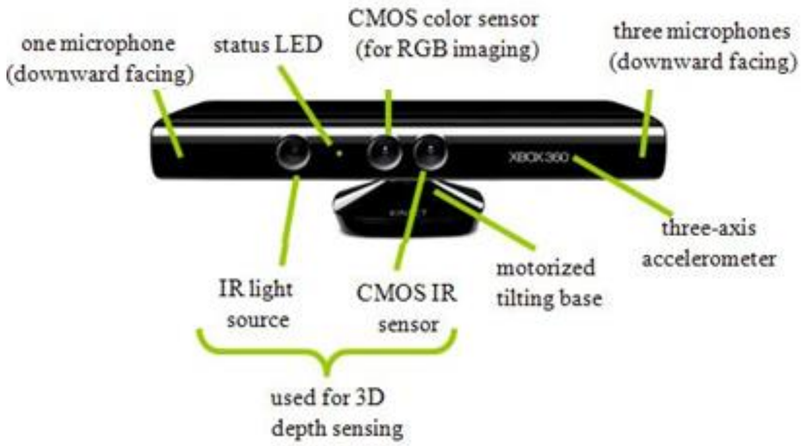

Figure 3. Elements of Kinect Microsoft device

Functionally, the monochrome sensor CMOS and the "depth sensor" analyzes the image captured and create a 3D map of the visibility field. The combination between the monochrome sensor CMOS and the depth sensor can capture image and motion in any conditions of light [21]. The depth sensor is adjustable, the SDK with the Microsoft Kinect sensor is able to self-calibrate based of the physical environment or on the obstacles in the physical environment etc. The microphones system is used to detect a voice location and to cover noise. All these sensors provide multiple capabilities of body recognition in 3D and of body motion, face, and voice recognition. The process of capturing a depth image consists in simultaneous capture of two images, RGB image, using CMOS RGB sensor, and depth image, captured by the monochrome sensor CMOS (see Figure 4).

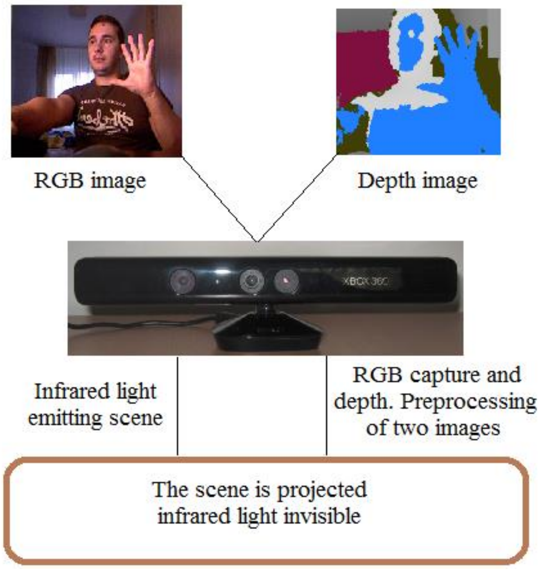

Figure 4. CMOS depth sensor. Invizible light IR is emitted and monitored using CMOS image sensor. The image processor generates depth image.

5. MICROSOFT KINECT SENSOR USE TO CAPTURE HUMAN HAND MOTION

In the research in the field of digital image processing, recognition of 3D objects involves recognition and detection of 3D objects in an image or a frame, part of a video stream. This recognition is possible in real time or it can be performed in a video stream previously captured and saved in the memory. The Kinect Microsoft sensor combines a set of hardware and software mechanisms that make up a digital 3D representation of a physical environment.

Algorithms created to recognize 3D objects, based on data captured by the Microsoft Kinect sensor analyze in parallel two types of data, namely: RGB image and data related to image depth. In this paper, we developed an application through which we try to overcome capture limits with web cam, namely depth capture for axes Z and Y. For the development of the application, we used Windows SDK and Natural User Interface library (NUI).

In Figure 5 there is the system proposed from conceptual point of view.

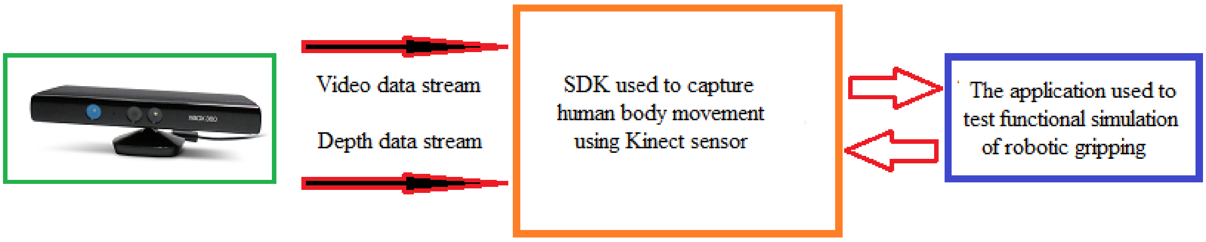

Figure **Error! No text of specified style in document.**. Interaction between Kinect sensor and the application that tests functional simulation

To develop the 3D data capture and control system of a virtual gripper, we considered four steps:

-Initialization: when the driver using Kinect sensor is loaded in the memory and initialized;

-Detection: the system must detect human hand and recognize gestures;

-Interpreting and recognition: each gesture of the human hand must be interpreted correctly;

-Visualizing: the gesture recognized must be transmitted to a virtual simulator for the control of an anthropomorphic gripper.

6. ANALYSIS OF DEPTH MEASURING PROCESS

Microsoft Kinect device uses a mapping process of the 3D space, adding depth data to each captured pixel. The value attached to each pixel represents the distance from the sensor to the object in front of the sensor in the sensor orientation direction [23]. To estimate each pixel depth, the Kinect sensor uses the concept previously presented, of structured light. Thus, the infrared sensor, emits a pattern known in the environment, then, based on the data captured by the monochrome sensor, the internal algorithm of the Microsoft Kinect sensor attaches to each pixel, a value (see Figure 6).

Where dE represents the distance estimated by the Kinect sensor, between object and sensor; dR represents the real distance between the Kinect sesor and the object. Based on the method presented, which estimates depth, we may conclude that the device becomes an efficient way of capturing coordinates (x, y, z) of any 3D object, but there is as well a difference between the estimated distance and the real one. Because for the purpose of this paper we seek as high precision as possible, we analyzed as well other depth data capturing methods.

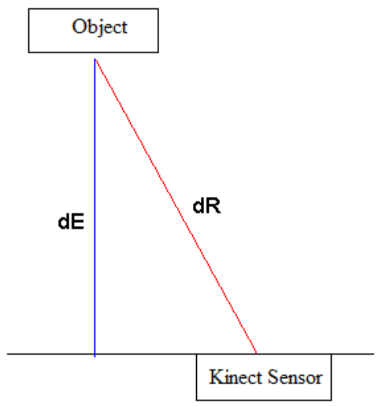

Figure 6. Estimated distance (dE) and real distance (dR) between Microsoft Kinect device and a random object. Image adapted after [23]

7. CONCLUSIONS

The methods presented in this paper are considered successful because they satisfy functional and non-functional requirements. The systm is explored in real time and it is also accessible – the system does not need expensive equipment. Any type of user can be helpful for research or for the control of an anthropomorphic gripper. In this paper we tested interaction between man and virtual environment, using visual methods of human hand capture, and a complete system of command and control of a red anthropomorphic gripper was created, going through a virtual stage. In Fig. 7 there is the scheme of the system proposed based on the Kinect sensor.

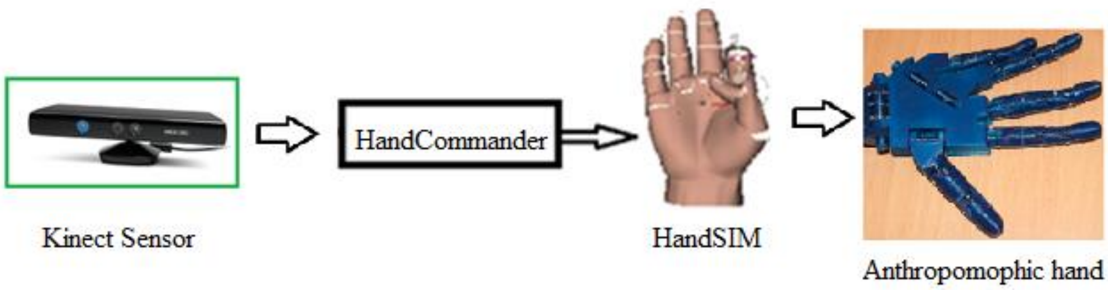

Figure 7. The system proposed for the command and control of an anthropomorphic gripper.

The communication protocol is used afterwards by the rest of the applications of gestures capture to send data to the virtual environment and then to the real one, the five-fingered anthropomorphic gripper. The software modules obtained are HandProcessor, HandCommander and HandSim (see Figure 7).

ACKNOWLEDGEMENT. We express our gratitude to the Company Cloos Germany and its Romanian representative Robcon which partly sustained these works. Their know how and support offered us the possibility to present in this Conference the results of our programming and research activities.

REFERENCES

1. MAJIDI, A., POURGHASSEM, H., NEJATI, M., *Vehicle Stop Detection Algorithm Based on Motion Analysis in Access Control System Application*, Proceeding CICN '12 Proceedings of the 2012 Fourth International Conference on Computational Intelligence and Communication Networks. IEEE Computer Society, Washington DC, USA, 690-694, 2012.
2. HAMETNER, R., SCHITTER, G., VOIGT, A., ZOITL, A., *Implementation guidelines for closed loop control algorithms on PLCs*, IEEE International Conference on Industrial Technology, 2013.

3. XU, J., HUANG, D., WANG, W., *Ballistic learning control: formulation, analysis and convergence*, Journal of Control Theory and Applications, **11**/3, 325-335, 2012.
4. KARRAS, G., BECHLIOULIS, C., ABDELLA, H. K., LARKWORTHY, T., KYRIAKOPOULOS, K., LANE, D., *A Robust Sonar Servo Control Scheme for Wall-Following using an Autonomous Underwater Vehicle,* IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 2013.
5. MARANTOS, P., DRITSAS, L., KYRIAKOPOULOS, K. J., *Robust Attitude Control for an Unmanned Helicopter in near-hover flights*, European Control Conference, Zurich, Switzerland, 2013.
6. LIAROKAPIS, M., ARTEMIADIS, P., KYRIAKOPOULOS, K. J., *Telemanipulation with the DLR/HIT II Robot Hand Using Cyberglove II and a Low Cost Force Feedback Device*, 21st Mediterranean Conference on Control and Automation MED'13, Chania, Crete, Greece, 2013.
7. LIAROKAPIS, M., ARTEMIADIS, P., KYRIAKOPOULOS K. J., *Quantifying Anthropomorphism of Robot Hands*, IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, 2013.
8. BURDEA, G., COIFFET, P., *Virtual Reality Technology* , 2nd edition , John Wiley and Sons, New Jersey, USA, 2003.
9. HEDBERG, J., ALEXANDER, S., *Virtual Reality in Education: Defining Researchable Issues*, Educational Media International, **31**/4, 214-220, 1994.
10. WHITELOCK, D., BRNA, P., HOLLAND, S., *What is the value of virtual reality for conceptual learning? Towards a theoretical framework*, Proceedings of European Conference on AI in Education, Leeds, UK, 1996.
11. BLACKLEDGE, J., BARRETT, M., *Evaluation of a Prototype Desktop Virtual Reality Model Developed to Enhance Electrical Safety and Design in the Built Environment*, ISAST Trans. on Computing and Intelligent Systems, **3**/3, 1 - 10, 2012.
12. MAHDY, Y. B., HUSSAIN, K. F., ABDEL-MAJID, M. A., *Projector Calibration Using Passive Stereo and Triangulation* , International Journal of Future Computer and Communication, **2**/5, 385-390, 2013.
13. CHEN, Q., LI, D., TANG, C., *Compressive Structured Light for Recovering Inhomogeneous Participating Media*, IEEE Transaction on Pattern analysis and machine intelligence, **35**/3, 2175-2188, 2013.
14. GUILHAUS, M., *SPECIAL FEATURE: TUTORIAL Principles and Instrumentation in Time-of-fight Mass Spectrometry Physical and Instrumental Concepts*, Journal of mass spectrometry, **30**, 1519-1532, 1995.
15. YOUSEFA, H., BOUKALLELA, M., ALTHOEFERB, K., *Tactile sensing for dexterous in-hand manipulation in robotics—A review*, Sensors and Actuators A: Physical, **167**/2, 171-187, 2011.
16. PIETERS, R. S., *Direct Methods for Vision-Based Robot Control: Application and Implementation*, PhD thesis, Eindhoven University of Technology, The Netherlands, 2013.
17. BUTLER, A., IZADI, S., HILLIGES, O., MOLYNEAUX, D., HODGES, S., KIM, D., *Shake'n'Sense: Reducing Interference for Overlapping Structured Light Depth Cameras*, Proceeding CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, USA, 1933-1936, 2012.
18. WWW01, http://msdn.microsoft.com/en-us/library/hh438998.aspx, Accessed on 2012.
19. RYDEN, F., CHIZECK, H., KOSARI, S. N., KING, H., HANNAFORD, B., *Using kinect and a haptic interface for implementation of real-time virtual fixtures*, Proceedings of the 2nd Workshop on RGB-D: Advanced Reasoning with Depth Cameras (in conjunction with RSS 2011), Los Angeles, USA, 2011.
20. STOWERS, J., HAYES, M., BAINBRIDGE-SMITH, A., *Altitude control of a quadrotor helicopter using depth map from microsoft kinect sensor*, IEEE International Conference on In Mechatronics (ICM), 358-362, 2011.
21. ANDERSEN, M.R., JENSEN, T., LISOUSKI, P., MORTENSEN, A. K., HANSEN, M. K., GREGERSEN, T., AHRENDT, P., *Kinect depth sensor evaluation for computer vision applications*, Electrical and Computer Engineering Technical Report ECE-TR-6, 2012.
22. KVALBEIN, M., *The use of a 3D sensor (Kinect™) for robot motion compensation. The applicability in relation to medical applications*, Master thesis, 2012.
23. DRAELOS, M. T., *The Kinect Up Close: Modifications for Short-Range Depth Imaging*, IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Hamburg, Germany, 251-256, 2012.