EAN 2017
55th conference on experimental stress analysis

3D Vision Methods for Enhanced Human-robot Interaction

Tölgyessy Michal[1,a], Dekan Martin[1,b] Rodina Jozef[1,c]

[1]Faculty of Electrical Engineering and Information Technology STU in Bratislava, Slovakia

[a]michal.tolgyessy@stuba.sk, [b]martin.dekan@stuba.sk, [c]jozef.rodina@stuba.sk

Keywords: HRI, robot motion, gesture recognition

Abstract. This article presents an overview of several original human-robot interaction methods. All of them are gesture based. Human motion and hand position is captured by a 3D sensor. Presented methods have been experimentally verified on real mobile robot platforms.

Introduction

Robots are becoming a more and more visible part of human lives. UAV drones with onboard visual systems make real-time videos at music festivals, or capture weddings and other social events from above. Many people own autonomous vacuum cleaners. This trend stresses more and more the need for natural human-machine interaction. In the field of robotics the term HRI (Human-Robot Interaction) became a standard name. Since it is a vast field different researchers concentrate on different topics. Some focus on gesture control of UAV's [1][2]. Others on pointing to physical objects and evaluating robot perception [3][4]. Gesture based control has even been implemented in communication with robotic swarms [5][6].
Research work presented in this paper focuses primarily on hand-gesture based human-robot communication. Perception is based on 3D visual sensing. Sensor used for experimental evaluation is the Kinect [7] [8] [9] [10] [11] [12]. Two of the presented methods rely on 3D joint positions. To extract 3D data from depth stream mainly the work of Shotton et. al. was used [13].

1. A plane-based gesture algorithm for mobile robot control

Algorithm is based on position of human operator's hands in relation to his or her torso. These positions define discrete movement commands for a mobile robot that are being continuously sent over a wireless network. Firstly, a set of points to define the *torso plane* are chosen. In the example in Fig. 1 they are right shoulder (A), left shoulder (B) and right hip (C). Position of hands (H_1 and H_2) is then analyzed in relation to the ABC plane.

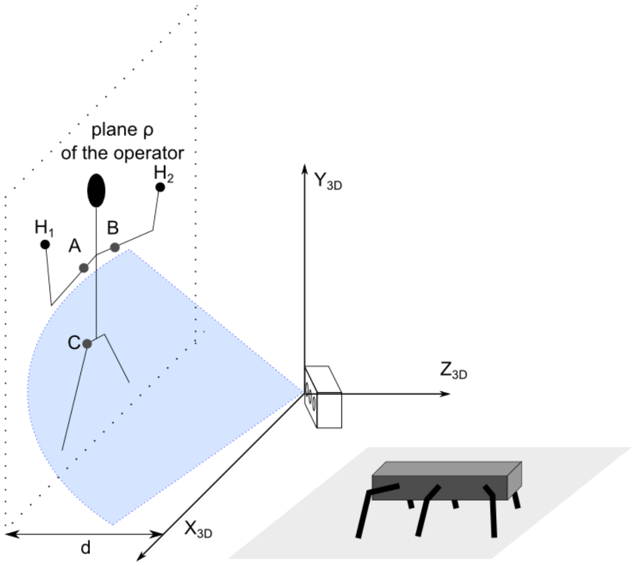

Fig 1. Obtaining the torso plane for robot control

To ensure the ability to work with different gestures 3 zones are introduced. Each hand can be present at any given time in only one of these zones (Fig. 2).

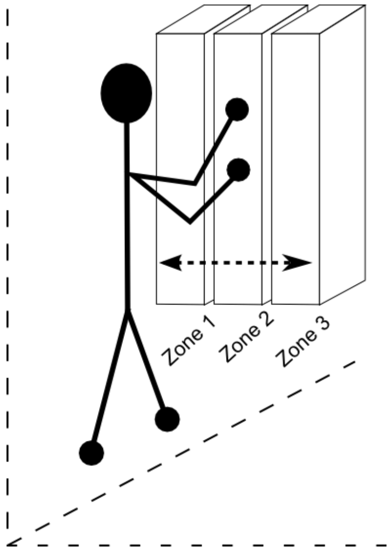

Fig. 2 Zones relative to the torso plane of the human operator

To determine the zone in which a hand is currently present it is necessary to execute these steps:

1. Compute vectors **AB** and **AC** from operator’s body joints
2. Compute the normal vector **n** of the torso **ABC** plane from cross product **ABxAC**
3. Obtain the parameters of *ax + by + cz + d = 0* by applying **n**
4. Distance between hand and torso plane is computed thusly:

$$D(H,\rho) = \frac{ax_H+by_H+cz_H+d}{\sqrt{a^2+b^2+c^2}} \quad (1)$$

Commands for the robot are defined by presence of hands in particular zones:

- STOP – both right and left hand in Zone 2

- ROTATION RIGHT – right hand in Zone 1, left hand in Zone 3
- ROTATION LEFT – right hand in Zone 3, left hand in Zone 1
- MOVE FORWARDS – both right and left hand in Zone 3
- MOVE BACKWARDS – both right and left hand in Zone 1

Experiment

To evaluate this gesture recognition method a candy carrier robot was controlled in a laboratory as well as crowded environment (Fig. exp1).

Fig. exp1 Experimental evaluation on a candy carrier robot

2. Controlling a mobile robot via pointing gesture

Second of the presented algorithms allows a human operator to point to a spot in the ground and guide the robot to this spot. To achieve this, 3D sensor is mounted directly on the robot. It detects position of elbow (A) and wrist (B) of the operator and computes a line defined by these points. Subsequently, robot finds the intersection of this line with the plane (ground) where both robot and human are present. An illustration of this scenario is in Fig. 3.

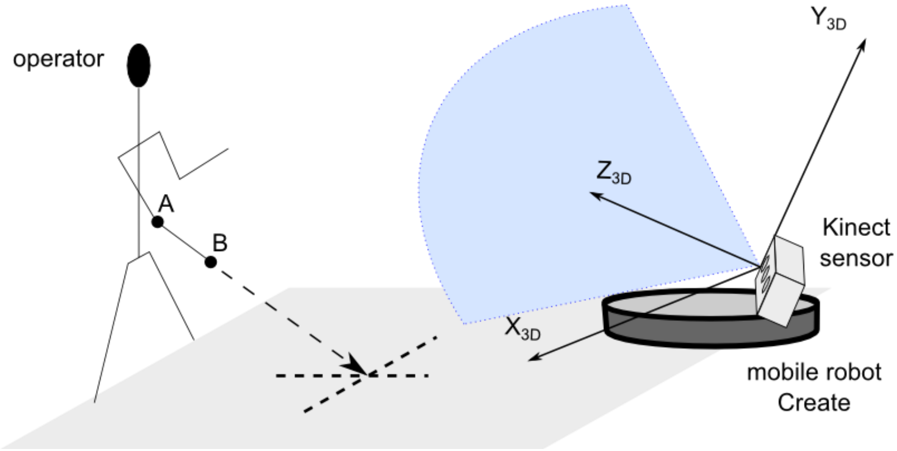

Fig. 3 Human operator points to the spot, where robot Create is supposed to move

Movement is begun when operator's hand is raised. A flowchart of the whole process is in Fig. 4.

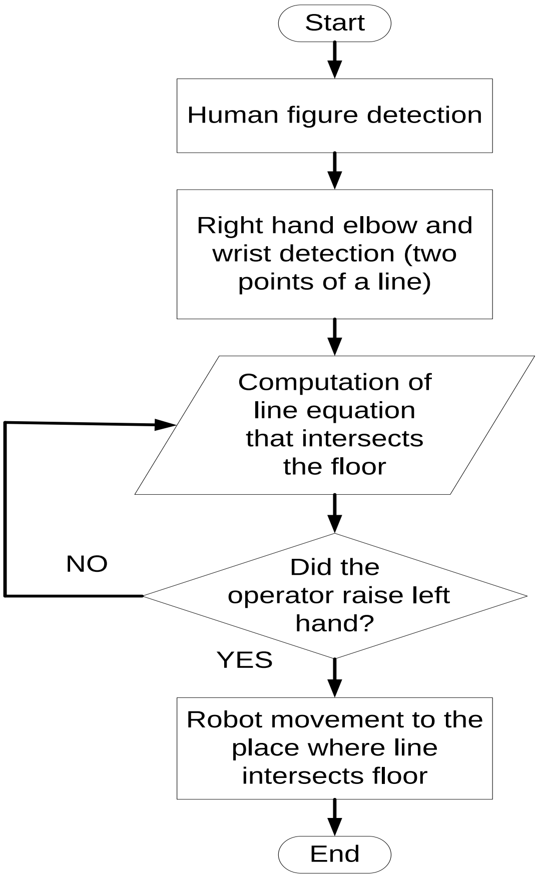
Start
Human figure detection
Right hand elbow and wrist detection (two points of a line)
Computation of line equation that intersects the floor
NO
Did the operator raise left hand?
YES
Robot movement to the place where line intersects floor
End

Fig. 4 State machine flowchart of the method

To compute necessary vectors standard methods of analytical geometry are used. For this, fundamental geometrical features need to be extracted (Fig. 5).

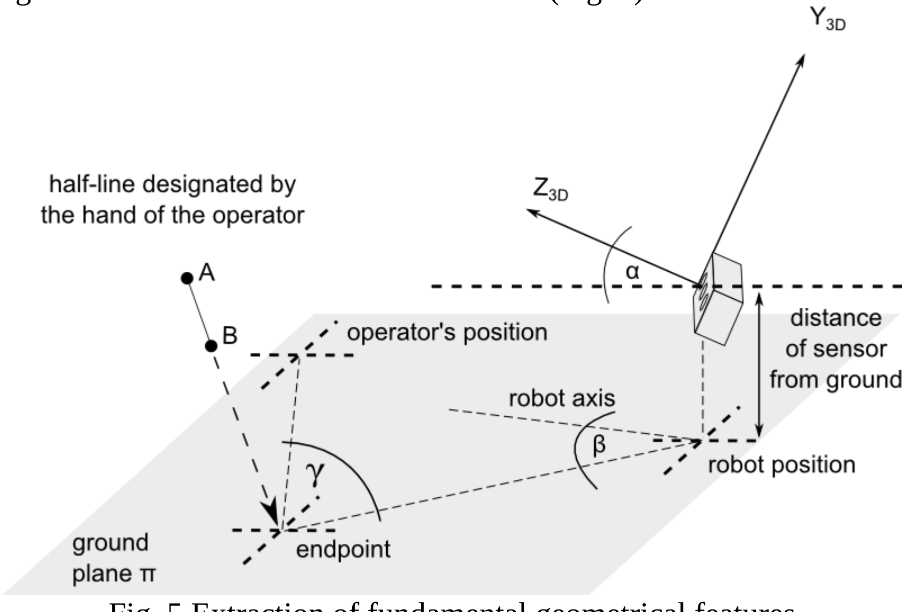
Y3D
half-line designated by the hand of the operator
Z3D
α
A
B
operator's position
distance of sensor from ground
robot axis
β
γ
robot position
ground plane π
endpoint

Fig. 5 Extraction of fundamental geometrical features

3D sensor mounted on robot is not centered and it is also rotated, therefore transformation between coordinate system of the robot (R) and Kinect (K) is needed (Fig. 6a, Fig 6b).

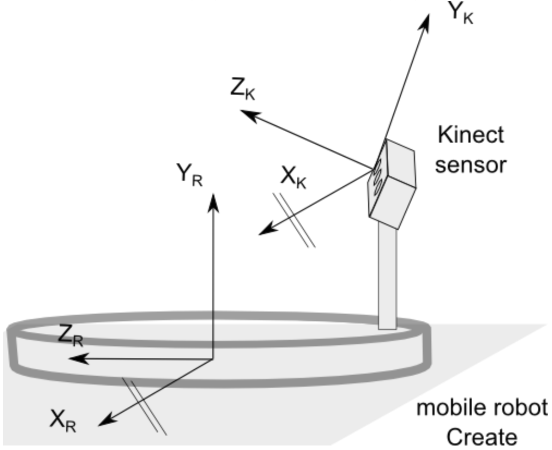

Fig. 6a Illustration of the two coordinate systems

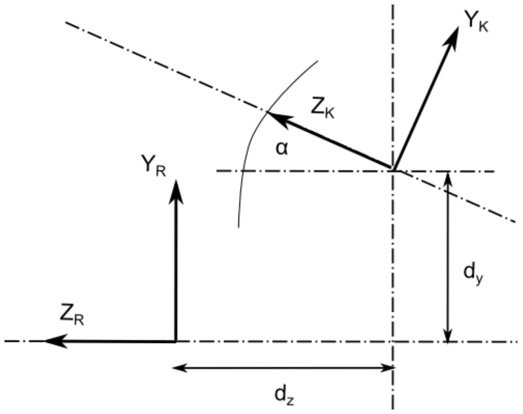

Fig. 6b Transformation between the coordinate system of the sensor and the coordinate system of the robot

Transformation between the two coordinate systems is defined by:

$$\begin{pmatrix} X_R \\ Y_R \\ Z_R \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & d_y \\ 0 & -\sin\alpha & \cos\alpha & d_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X_K \\ Y_K \\ Z_K \\ 1 \end{pmatrix}$$

Final vector features are illustrated in Fig. 7. Vector **v** is the basis for distance and rotations of robot movement.

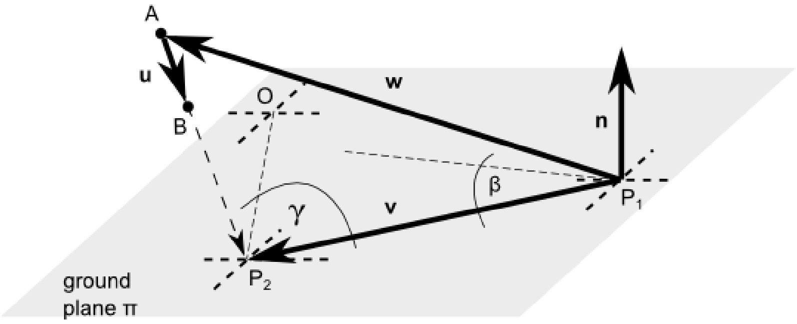

Fig. 7 Vector description of the main features

Experiment

To test this method experimental evaluation was done in laboratory environment. Several participants were stood in front of the robot and pointed to a desired spot in the ground. From experiment data probability density models were computed and evaluated. Illustration of experiment scenario is in Fig. exp2.

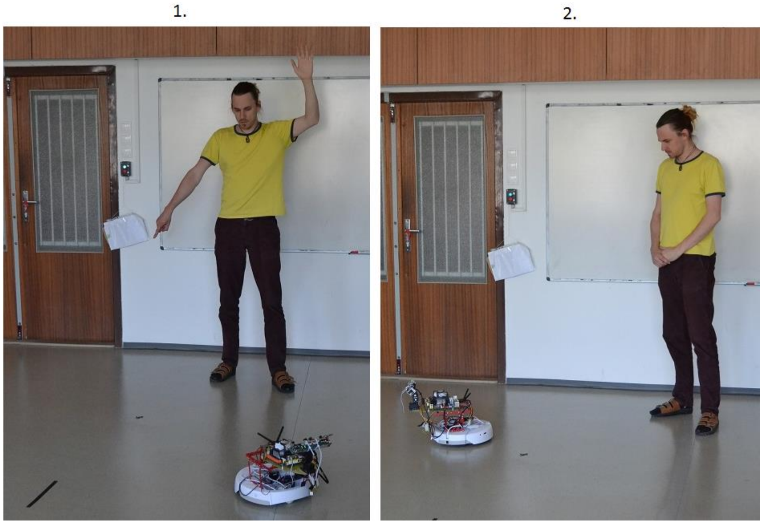

Fig. exp2 Laboratory experiment

3. Controlling a group of robots to perform a common task by gestures only

Basic idea of third algorithm is to use two hands of a human operator for every action issued to a group of robots. One hand of the operator is used to select a robot from the group. The other is used to control motion of the selected robot. Each robot is assigned a number – index that is used for its identification (ID). Number of raised fingers on operator’s right hand matches the ID of the currently controlled robot. Therefore, if the operator wishes to control a hexapod robot with ID 3, he needs to raise 3 arbitrary fingers on his right hand. A healthy human has 5 fingers, so it is possible to pick from a group of 2-5 robots, theoretically 2-6 robots if a fist with no fingers raised was used, too.

Left hand is left to manage the motion control process itself. To define several motion types we compare the position of both hands in 3D space. The workspace of left hand is divided into 3 zones (Fig. 8). If both hands are close enough (overhead viewpoint) the hand controlling the robot is considered to be present in Zone 2. If it crosses a defined threshold in front of the right hand it’s in Zone 3, if it’s put behind right hand it’s in Zone 1. This way 3 motion types for the robot can be selected, however this can be insufficient. That is why position of the thumb of left hand is also considered to increase the number of motion types.

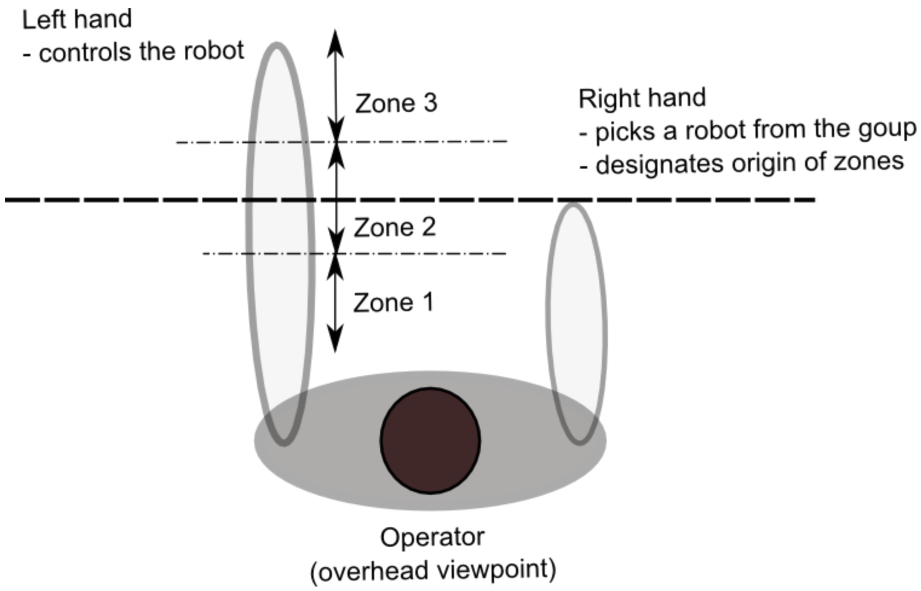

Fig. 8 Overhead viewpoint of the human operator

Fig. 9 shows an example of possible gestures of the two operator's hands. First case shows a robot with ID 3 rotating left (thumb position), second case shows a robot with ID 5 rotating right and last case a robot with ID 2 being in the state of halt.

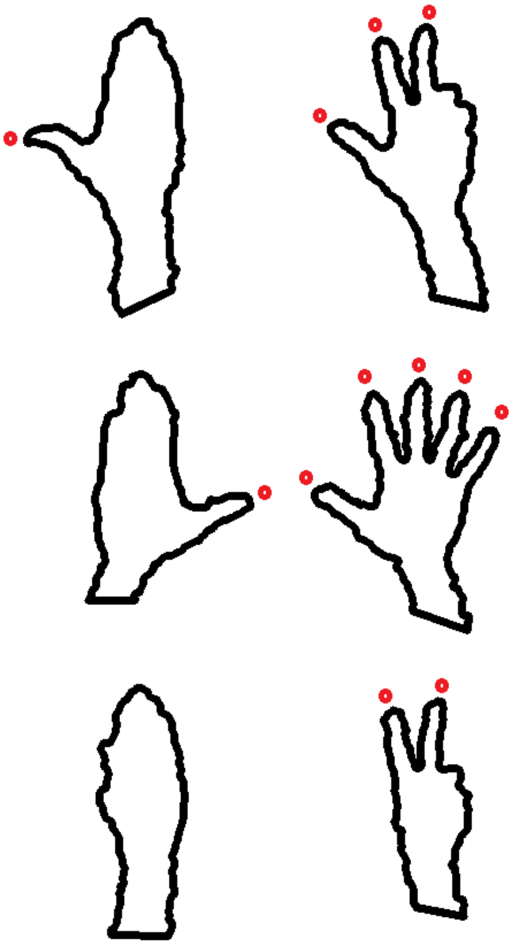

Fig. 9 Gesture examples

Adding the zone concept to examples from Fig. 9 leads to 9 motion types. All of them are listed in the following table.

Left hand thumb position	Left hand zone	Robot index	Motion type
Left	3	1-5	Left forward circular motion
Right	3	1-5	Right forward circular motion
Middle	3	1-5	Forward motion
Left	2	1-5	Rotation left
Right	2	1-5	Rotation right
Middle	2	1-5	Stop
Left	1	1-5	Left backward circular motion
Right	1	1-5	Right backward circular motion
Middle	1	1-5	Backward motion

To segment fingertips, contours of operator's hands are found. Within these contours all sharp vertexes are searched for and subsequently, fingers are classified. Fingertip vertexes are the sharpest. Therefore, to test the sharpness vectors **w** and **v** from nearby contour points are formed and their angle is computed based on simple dot product (Fig. 10).

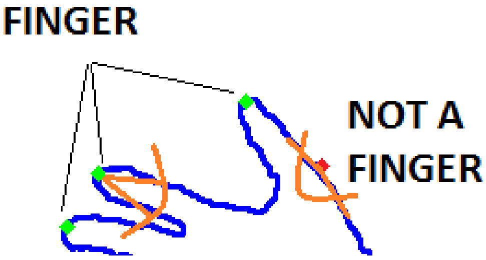

Fig. 10 Finger detection

Experiment

To test this method 3 different robots were used in laboratory environment. Hexapod, iRobot Create and a ball robot were controlled by one operator using wireless technology as shown in Fig. exp3.

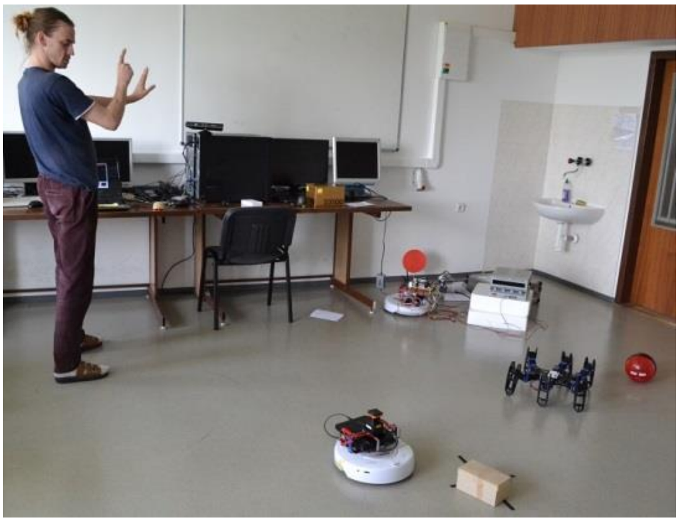

Fig. exp3 Experiment in laboratory environment

Conclusions

We have successfully implemented three human-robot interaction methods. All of them are based on 3D vision. Methods were tested by human operators on real mobile robotic systems. We believe our research contributes to the more and more spreading field of HRI.

Acknowledgments

This work was supported by national grants Stimuly Micsrostep 0691/2015 and VEGA 1/0752/17.

References

[1] G. Costante, E. Bellocchio, P. Valigi, E. Ricci, “Personalizing vision-based gestural interfaces for HRI with UAVs: A transfer learning approach,” in Proc. IEEE International Conference on Intelligent Robots and Systems, 2014, pp. 3319 – 3326.

[2] V.M. Monajjemi, J. Wawerla, R. Vaughan, G. Mori, “HRI in the sky: Creating and commanding teams of UAVs with a vision-mediated gestural interface,” in Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, pp. 617-623.

[3] C. P. Quintero, R. T. Fomena, A. Shademan, N. Wolleb, T. Dick, and M. Jagersand, “SEPO: Selecting by pointing as an intuitive human-robot command interface,” in Proc. IEEE International Conference on Robotics and Automation, 2013, pp. 1166-1171.

[4] C. Høilund, V. Krüger, and T. B. Moeslund, “Evaluation of human body tracking system for gesture-based programming of industrial robots,” in Proc. IEEE Conference on Industrial Electronics and Applications, 2012, pp. 477-480.

[5] A. Giusti, J. Nagi, L.M. Gambardella, S. Bonardi and G.A. Di Caro, "Human-swarm interaction through distributed cooperative gesture recognition," in Proc. 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2012, 401–401.

[6] J. Alonso-Mora, R. Siegwart, B. Perun and P. Beardsley, "Human-Robot Swarm Interaction for Entertainment," in Proc. 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2014, 98–98.

[7] Freedman et. al., “Depth Mapping Using Projected Patterns,” U.S. Patent 2010/0118123 A1, May 13, 2010.

[8] A. Reichinger. (2011, April). Kinect Pattern Uncovered. [Online]. Available: http://azttm.wordpress.com/2011/04/03/kinect-pattern-uncovered/

[9] B. Klug. (2010, Dec.). Microsoft Kinect: The AnandTech Review. [Online]. Available: http://www.anandtech.com/show/4057/microsoft-kinect-the-anandtech-review/2

[10] K. Konolige, P. Mihelich. (2012, Dec.). Kinect Operation. [Online]. Available: http://wiki.ros.org/kinect_calibration/technical

[11] M. R. Andersen, T. Jensen, P. Lisouski, A. K. Mortensen, M. K. Hansen, T. Gregersen, and P. Ahrendt, “Kinect Depth Sensor Evaluation for Computer Vision Applications,” Department of Engineering, Aarhus University. Denmark. 37 pp. - Technical report ECE-TR-6, Feb. 2012.

[12] J. Kramer, N. Burrus, F. Echtler, D. C. Herrera, and M. Parker, “Hardware,” in Hacking the Kinect, Apress, 2012, pp. 11–13.

[13] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, A. Blake, “Efficient human pose estimation from single depth images,” in Proc. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, pp. 2821-2840.