EPiC Series in Engineering
Volume 1, 2018, Pages 41–49
ReVo 2017: Laval Virtual ReVolution 2017
"Transhumanism++"

LightAir: Augmented Reality System for Human-Drone Interaction using Foot Gestures and Projected Image

Dzmitry Tsetserukou, Mikhail Matrosov, Olga Volkova, Evgeny Tsykunov
Space CREI, Skolkovo Institute of Science and Technology, Moscow, Russia
d.tsetserukou@skoltech.ru, m.matrosov@skoltech.ru,
o.volkova@skoltech.ru, e.tsykunov@skoltech.ru

Abstract

We propose a new paradigm of human-drone interaction through projecting image on the ground and foot gestures. The proposed technology allowed creating a new type of tangible interaction with drone, e.g., DroneBall game for augmented sport and FlyMap to let a drone know where to fly.

We developed LightAir system that makes possible information sharing, GPS-navigating, controlling and playing with drones in a tangible way. In contrast to the hand gestures, that are common for smartphones, we came up with the idea of foot gestures and projected image for intelligent human-drone interaction. Such gestures make communication with the drone intuitive, natural, and safe. To our knowledge, it is the world's first system that provides the human-drone bilateral tangible interaction. The paper also introduces the novel concept of Drone Haptics.

1 Introduction

Drone technology progress is extremely fast and the reason behind this is the significant interest from electronic commerce companies for parcel delivery (Amazon, Alibaba, Google, etc.), filmmakers, oil and gas industry. Nowadays, a drone is not just a device that flies but rather the device that should communicate with the human, e.g., during the cargo delivery to the customer. We propose a new paradigm of human-drone interaction through projecting the image on the road in front of the user and foot gestures.

Radio controller with two joysticks is common tool to hover the drone. A group of researchers from ETH Zurich demonstrated the control of quadcopter using Kinect sensor [9]. Hand gesture recognition lets the operator to take off, land, flip, and hover the copter. A wearable gestural interface SixthSense [3] was designed to allow user to interact with the digital world through hand gestures. In [6], robot with LiDARs scans the environment and transmits the information about the obstacles to the operator through the tactile patterns. The speed and direction of the mobile robot is controlled by the torso posture of the human and, hence, the hands can be used for gesture-based interaction with the remote environment. More recently, authors [4] proposed human-drone interaction with human hand tracking through camera processing the feature points. Such system enhances the level of interactivity. However, human interacting with a drone expects not only that it follows the commands but also it exchanges the digital information with us.

We have developed LightAir system that makes possible the information exchanging and sharing, GPS-navigating, controlling, and playing with drones in a tangible way (see Figure 1). To our knowledge, it is the world's first system that provides the bilateral digital human-drone communication.

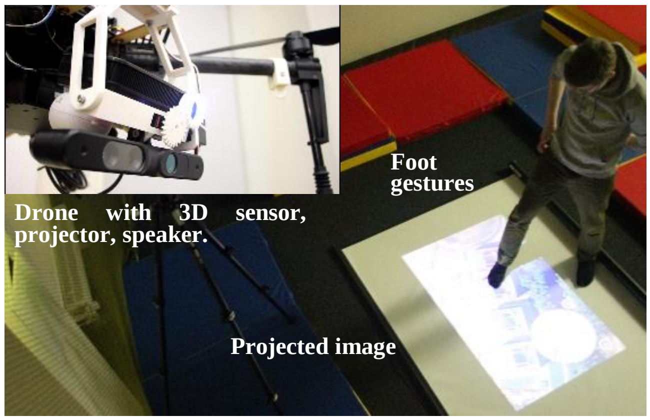


Figure 1: LightAir system for human-drone communication

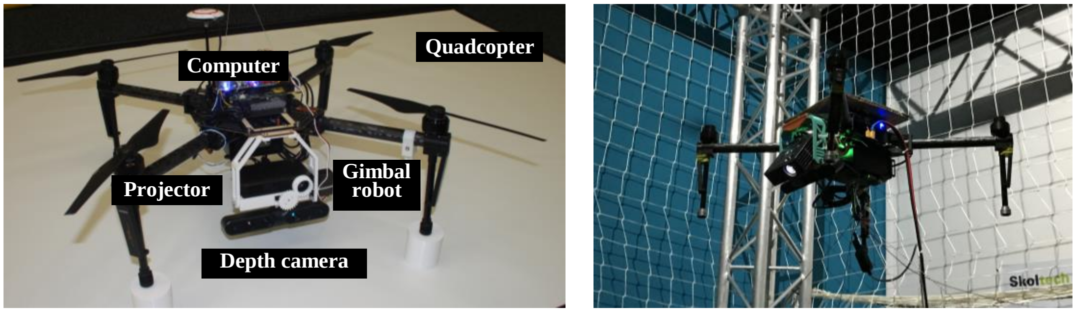
In contrast to the hand gestures, that are common for smartphones, we came up with the idea of foot gestures and projected image for the tangible interaction. Such gestures make communication with the drone intuitive, natural, and safe that is especially crucial for mid-size drones. Our experiments with LightAir revealed that drone could be an excellent communication tool between human beings and flying robot through the multimedia and graphic interface.

2 Principle and Technologies

The LightAir is equipped with several modules that are responsible for control and navigation: Intel NUC Kit QS77, pocket projector Dell M115HD, 3D sensor Asus, and speaker. Quadcopter DJI Matrice 100 of max. 1 kg payload hosts all components and flies with them robustly. Intel onboard computer processes a large amount of sensory data, such as 3D point cloud, RGB video streams, radio signals. The designed and 3D-printed gimbal with 2-DoF moves depth camera and pico-projector jointly (Figure 2).

The main program sends cues to ATMega microcontroller to activate the servomotors of the gimbal robot. The gimbal features driver system for automatic adjustment of the projector focus. Program calculates the center of the virtual screen and projector focuses at that point.

The depth camera Asus Xtion Pro determines 3D scene and generates point cloud (Figure 3). The floor is detected from the point cloud with RANSAC plane detection algorithm. Then we define the orthonormal basis of the virtual screen with the Gram-Schmidt process. The foot location on the ground is calculated from the points of the 3D cloud that arranged in the predefined region above the floor.



a) b)

Figure 2: a) Projector and depth camera are installed on the quadcopter DJI Matrice 100, b) Flying LightAir scans the gestures and projects the image in front of a user

We have implemented the multi-touch foot input. Each press of the foot is tracked with the Hungarian optimization algorithm. We proposed the algorithm for the detection of foot orientation. The points of cloud from the depth camera are clustered into two groups representing heel and tip of the foot. Line connecting the centers of gravity of each region defines the foot orientation. We process the resulting vector using the adaptive low-pass filter.

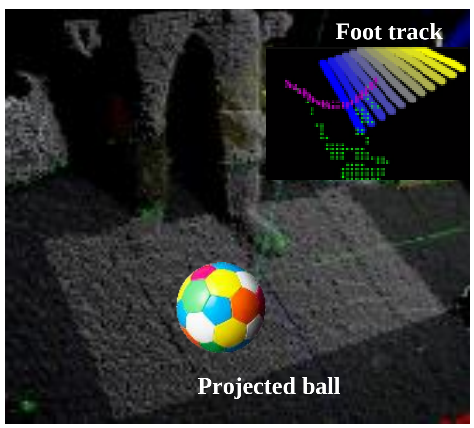


Figure 3: 3D point cloud processing for DroneBall application

Audio processing was realized by streaming audio data directly to the PulseAudio server. The program creates a new thread that connects to the server and streams sound for each key pressed. This way we can simultaneously play as many sounds as we want.

3 Drone Flight Stabilization with Computer Vision

The motion of any type of multicopter is supervised by a flight controller. Flight controllers are capable of the basic attitude stabilization of the vehicle (pitch, roll, yaw axes) for the given position of the quadcopter. In the outdoor navigation of drones, global navigation satellite system (GNSS, e.g., GPS, GLONASS) estimates the position of the flying robot. We employ the computer vision algorithms for external position estimation because GNSS is not available for the indoor navigation.

Computer vision algorithm requires a high computational power, therefore we are using ground-based server that communicates with the flight controller using wired/wireless link. The Logitech C930e camera is mounted on the ceiling and connected to the ground-based computer. In contrast to the system when camera fixed to the drone, immobility of the ceiling camera facing downwards allows obtaining the high quality images using cheap device.

An augmented reality (AR) marker is mounted on top of the drone. This configuration lets the AR marker to be always inside of the field of view (FOW) of the camera for the both cases, while taking off and landing (Figure 4). The resolution of the camera of 960x1600 pixels ensures obtaining the coordinates of the AR marker with approximately 7 Hz frequency using Intel® Core™ i7, 8Gb of RAM computer.

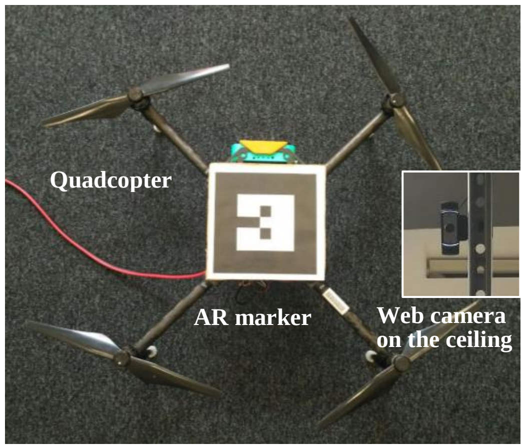


Figure 4: AR marker is placed on the top of the drone. Web camera sees the marker from the height of ceiling

Wi-Fi/wired USB-UART communication link connects the flight controller Pixhawk with ground-based computer. Several experiments were conducted to find out the PID gains generating the minimal error for position and speed of the drone. The adjusted PID gains are transferred to the flight controller through ground control station. Figure 5 illustrates the positions of the drone using Rviz visualization tool of Robot Operating System (ROS) [5].

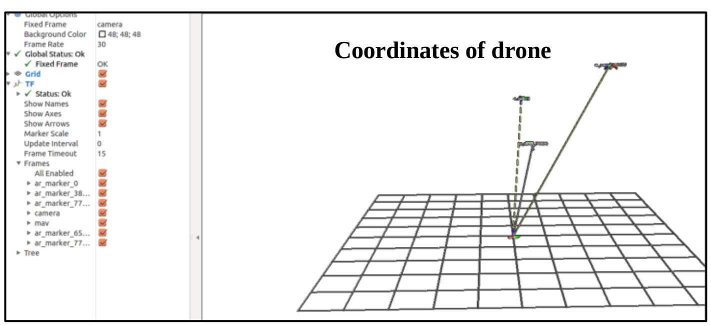


Figure 5: Visualization of the drone position in 3D space

The preliminary experiment revealed that the cumulative average error of the drone in each direction is about 10 cm around the center position. It is accurate enough for continuous human-drone interaction. The precision of the drone hovering can be further improved with more accurate and expensive system, e.g., Natural Point OptiTrack.

We have to consider the limitation of the drone flight time that typically equals 20-25 min. Quadcopter DJI Matrice 100 consumes as high as 10-12A under 24V. Therefore, we have attached external wired power supply to the drone to make the daylong demonstration possible.

The ROS is used for the high level data processing and communication between the flight controller and computer vision algorithm. The color images from the camera with resolution of 960x1600 are processed by ar_track_alvar package, which is the ROS wrapper for Alvar (AR marker tracking library). Image processing results in position and orientation of the AR tag. After frame transformation, coordinates of the marker are conveyed to the PX4 flight stack using MAVROS. MAVROS is the ROS package that helps to interface the ROS environment of the computer with Pixhawk flight controller. The MAVLink protocol is applied for the data transfer. This approach helps to use ROS as the main development environment, which is very powerful tool for the complex robotic system design. PX4 flight stack has the function of the vision-data-based position estimation. The vision data are loaded to the local position estimator (LPE), and after fusing with data from inertial measurement unit (IMU) and other sensors we calculate the position of drone. The estimation of quadcopter heading (Yaw axis) is also based on visual information as soon as drone compass suffers from strong interference while placing indoor.

4 Drone Haptics Concept and Its Application for LightAir

In this section we introduce a new concept of Drone Haptics. With LightAir system person interacts with quadcopter using foot gestures. He can zoom in, zoom out, click, and slide the image. However, because of swing motion of drone near the hovering position, the screen in front of the user moves in some extent. Hence, it is often the case of mistouch on the button, or lag of the action stop. Tactile feedback can improve the effectiveness of the foot gestures. In the future version of LightAir we propose to apply a foot worn tactile display to present the button click sensation, surface texture, object shapes, and stiffness. The user study methodology for experiments on the object shape and stiffness will be designed based on the research presented in [8].

Additionally, tactile stimuli can also provide the foot navigation patterns to make the control of gestures more accurate. For example, they can be used to notify the user where to move the foot: right, left, front, back, up, down to accomplish the correct action. The foot worn tactile device will be based on the LinkTouch tactile display with inverted five-bar linkage mechanism [7] aimed at presentation of 2-DoF force feedback, vibration, texture, and object softness. Sense of touch can potentially improve the navigation of drone or drone swarm based on the results presented in [6]. The mid-air touch can be delivered to the user from drone using ultrasound tactile display installed on quadcopter [1].

Drone Haptics is the emerging area of research which focuses on the study and design of devices and systems that communicate the touch sensation from drone (or swarm of drones) to the human, allow drone (or swarm of drones) to physically interact with human using the sense of touch, and provide drone landing on an uneven terrain using force-sensitive robotic landing gear.

The touch sensation of contact with flying robot can be generated through the mediated tactile human-drone interaction or direct physical interaction, e.g., using robot arm attached to the drone. The former case assumes application of wearable haptic displays or ultrasound speakers on the drone. In the latter one, person can pass the parcel to robotic arm or pull the arm to navigate the drone to the landing area. Physical interaction of the drone with ground surface can be implemented with specially designed robotic landing mechanism. Drone will be capable of landing on uneven, moving, and unstable surface.

5 Applications and Conclusions

The proposed technology allowed developing a new type of interaction with drone, e.g., DroneBall game for augmented sport and FlyMap to let drone know where to fly.

DronePiano emulates a piano keyboard (Figure 6). Keys react on the user's foot as if they were real buttons: even partial appearance of the foot on the key invokes the reaction. DronePiano makes possible to press several keys simultaneously and each keypress produces its own sound. Unique music played by drone is generated by the user.

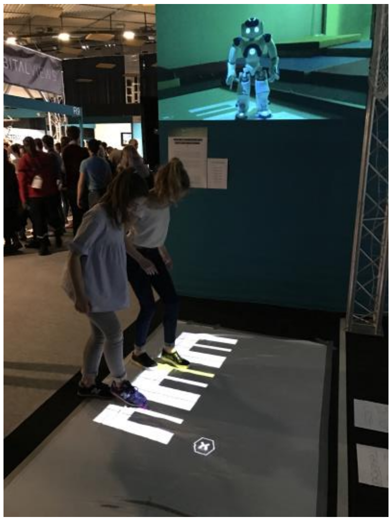


Figure 6: DronePiano application

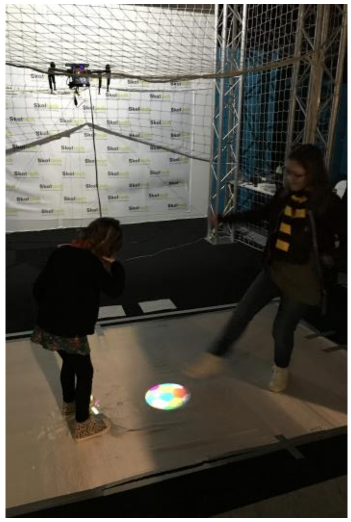


Figure 7: DroneBall application

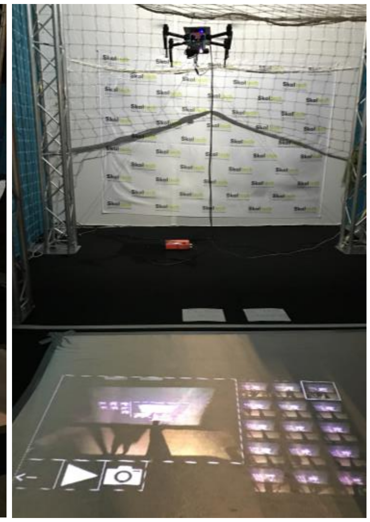


Figure 8: AirVideo application

DroneBall: augmented football game allows users to interact with each other by kicking the ball (Figure 7). Application detects the foot direction and ball responds respectively. High sensitivity makes application realistic and could be expanded in the future to the real-mode game of the player and drone. Drone will fly to the direction of the kicked ball to catch it virtually. To navigate drone autonomously we are planning to setup motion-capturing system OptiTrack Prime 13. The safety of the user during demo is secured by the net cage. The extension of the application can be the *TransDrone* game, the multi-user drone-augmented game when people can send the drone to the friend located over the long distance by hitting the projected ball. The second party should capture the ball by navigating the projected goalkeeper to catch the virtual ball delivered by drone. Such game can improve the social connectivity among people living in the countryside. While playing the game, users can not only enjoy the time but also send the gifts over the distance to the friends.

AirVideo: the copter has an onboard depth sensor with camera which is used for foot gesture recognition of the user. We created the application which shows a video stream of what drone sees. It also can grab the photos and store them. The interface contains two major elements, i.e., video stream preview and a bank of thumbnails, where you can select a photo to show (Figure 8).

Since the whole image is projected onto the floor, it must be distorted due to the perspective transformation. Same principle applies to the video stream. An interesting feature is that the perspective transform itself is calculated using the same video stream (and depth stream) from the ASUS Xtion camera. The algorithm embeds the input data into the output video stream in the same way as for any other element of interface (buttons, text, icons, etc.). The program draws whether video or objects of graphical user interface on the virtual screen and then it is transformed to the perspective view. This fact demonstrates the generalized approach to creating a human-drone interface.

In the future we are planning to implement the real-time video editor. The video captured by drone along with the video editing tools will be projected in front of the user. The person will manipulate the buttons, e.g., trim, zoom, move, and stretch, to create the desired video clip which can be immediately uploaded on his/her social network page or cloud storage. It will be much more convenient and time saving, as no file copying from drone is required anymore.

FlyMap: using image of the real-time map, drone projects it on the surface (Figure 9). The user can choose and press any point on the map and system zooms it. In the outdoor application GPS (Global Positioning System)-sensor shows the current drone position. Application calculates the approximate time along with the flight path from the given position to the selected one. The extended future version can be similar to Uber or Lyft service, but for parcel delivery system in the city. The user of *DroneNet* selects the available nearest drone, specifies the address of the target person, pays for the service, and sends the drone to the destination.

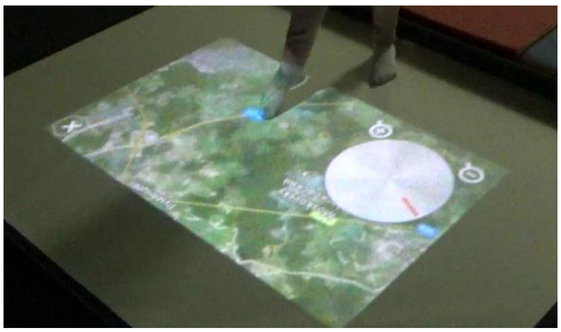


Figure 9: FlyMap application: user can zoom in/out, and move the map to show a drone the coordinates where to

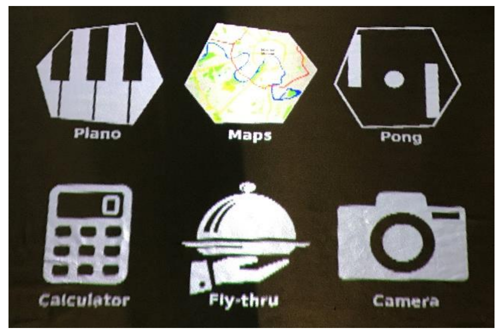


Figure 10: DroneStore system with six applications the coordinates where to fly

DroneStore: currently it includes six application and user launches the desirable one by clicking on it (Figure 10). Actually, the number of apps potentially can be considerably larger. Therefore, we are going to build the digital distribution platform DroneStore where user can browse and download the desired application on their platform. Software development kit for designers of drone applications should be provided as well.

FlyCulator: large-scale interactive calculator with 18 buttons appears in front of the user (Figure 11). Person touches the buttons to perform the arithmetic operations. To reduce the false triggering of buttons, keyboard of the FlyCulator was intentionally reduced to a single-touch interface. The similarity with typical hardware device makes it intuitive. The application can be useful for Math education of children outdoors. For example, let us imagine the interactive garden, dubbed *DroneGarden*, where children count the trees, flowers and make the arithmetic on ground using LightAir. Drone projects the information about particular species in the case of successful calculations.

Figure 11: FlyCulator application: user accomplishes math calculus stepping on the appropriate digits and operators

Figure 12: Graphical user interface of DronePay application: user provides the CVV number for confirmation

DronePay is the unique system that supports the potential online payment in midair for the drone-delivered goods (Figure 12). In the general scenario, customer selects the desired food from the menu, adds it to the cart, and proceeds with the checkout. There is an option to pay by cash or credit card. In the latter case, use shows the credit card in front of the quadcopter, the drone takes the picture of it, recognizes the card number, etc., and asks to enter the Card Verification Value (CVV) number for confirmation.

Another interesting application of LightAir could be the smart house control, *HouseOnAir* system. One can imagine that instead of digits drone projects the images of a car door, gate, home door, lawn irrigation, outdoor/indoor light, washing machine, fridge, etc. Householder is capable of controlling the infrastructure of the house and garden while being outdoors. He/she just clicks the corresponding icon and drone activates the device through Wi-Fi. The future *DroneCare* application detects the physical condition of the injured person and provides the instruction for the witnesses how to save the life of the victim before the ambulance arrives. Additionally, we can envision the augmented interaction with a drone while parcel delivering to check the specification and to run the virtual gadget before unfolding (Figure 13).

The LightAir system with DroneBall, DronePiano, and FlyMap applications was demonstrated at Siggraph 2016, Emerging Technologies, where the project received the Laval Virtual Award [2]. Hundreds of participants experienced the novel approach to interaction with drones and found it very interesting and engaging. The demonstration at Laval Virtual 2017 was enriched with new applications, e.g., FlyCulator, DronePay, AirVideo, and autonomous drone hovering.

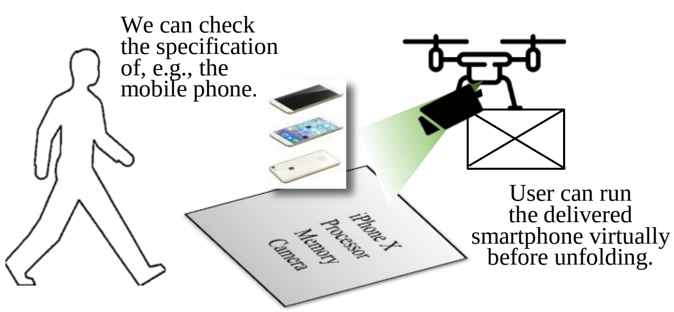


Figure 13: Concept of application for augmented parcel delivery

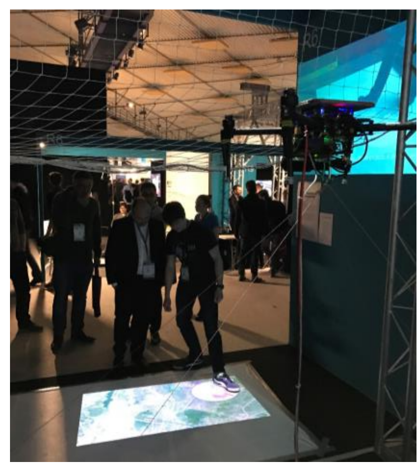


Figure 14: Demonstration of LightAir at Laval Virtual 2017

The future research will be devoted to the user study of human-drone interaction in order to evaluate the novel interface, intuitiveness of foot gestures, and usability of the proposed applications. It should be emphasized that proposed interface is the universal platform for any type of drone. Interface of LightAir can be used not only by humans but also by humanoid robots as they have similar kinematics and gestures.

The developed technology potentially can have a big impact on the multi-modal communication and interaction with flying robots. LightAir suggests much more intimate, interactive, and immersive interaction with drones leveraging image projection and foot gesture recognition to make it augmented.

Acknowledgements

The authors acknowledge the support by Unmanned Vehicle Program, Center for Entrepreneurship and Innovation, Skoltech.

References

[1] T. Hoshi, M. Takahashi, K. Nakatsuma, and H. Shinoda. Touchable holography. In *Proceedings of SIGGRAPH 2009 Emerging Technologies*, *ACM Press/ ACM SIGGRAPH*, *New York*, *USA*, *SIGGRAPH' 09*, Article No. 23.

[2] M. Matrosov, O. Volkova, and D. Tsetserukou. LightAir: a novel system for tangible communication with quadcopters using foot gestures and projected images. In *Proceedings of SIGGRAPH 2016 Emerging Technologies*, *ACM Press/ ACM SIGGRAPH*, *New York*, *USA*, *SIGGRAPH' 16,* Article No. 16.

[3] P. Mistry and P. Maes. SixthSense: a wearable gestural interface. In *Proceedings of SIGGRAPH Asia 2009 Emerging Technologies*, *ACM Press/ ACM SIGGRAPH*, *New York*, *USA*, *SIGGRAPH ASIA' 09*, Article No. 11.

[4] K. Miyoshi, R. Konomura, and K. Hori. Above Your Hand: direct and natural interaction with aerial robot. In *Proceedings of SIGGRAPH 2014 Emerging technologies, ACM Press/ ACM SIGGRAPH, New York, USA, SIGGRAPH' 14*, Article No. 8.

[5] M. Quigley, B. Gerkey, and W. D. Smart. Programming robots with ROS: a practical introduction to the robot operating system, *O'Reilly Media, Sebastopol*, 2015.

[6] J. Sugiyama, D. Tsetserukou, and J. Miura. NAVIgoid: robot navigation with haptic vision. In *Proceedings of SIGGRAPH Asia 2011 Emerging Technologies*, *ACM Press/ ACM SIGGRAPH*, *New York, USA, SIGGRAPH ASIA' 11*, Article No. 9.

[7] D. Tsetserukou, S. Hosokawa, and K. Terashima. LinkTouch: a wearable haptic device with five-bar linkage mechanism for presentation of two-DOF force feedback at the fingerpad. In *Proceedings of Haptics Symposium 2014, IEEE Press*, pp. 307-312.

[8] D. Tsetserukou and S. Tachi. Efficient object exploration and object presentation in telerobotic system TeleTA with tactile feedback. In *Proceedings of WorldHaptics 2009, IEEE Press*, pp. 97-102.

[9] M. Waibel. Controlling a quadcopter using Kinect. IEEE Spectrum. https://spectrum.ieee.org/automaton/robotics/robotics-software/quadrotor-interaction, July 2011.