



Application of Augmented Reality in Mobile Robot Teleoperation

Tomáš Kot^(✉) ID, Petr Novák, and Ján Babjak ID

Department of Robotics, Faculty of Mechanical Engineering, VŠB-Technical University Ostrava,
17. listopadu 15, 708 33 Ostrava, Czech Republic
{tomas.kot,petr.novak,jan.babjak}@vsb.cz

Abstract. The paper deals with the utilisation of augmented reality on the operator control panel of a teleoperated mobile robot. The work is a continuation of the previous research, which resulted in the creation and successful practical implementation of a virtual operator station based on the virtual reality head-mounted display (HMD) Oculus Rift. The new approach suggests using the new Microsoft Hololens augmented reality headset to add the virtual control elements directly to the real world, which has some very important benefits – especially the fact that the operator is not visually isolated from his surroundings. The device is introduced in the beginning of the article and then follows the description of all tasks required to create the augmented reality operator station. Mentioned are also other possible ways of using augmented reality to assist the operator of a mobile robot.

Keywords: Mobile robot · Teleoperation · HMD · Oculus Rift · Hololens
Augmented reality · Virtual reality

1 Introduction

Mobile robots controlled remotely by a human operator (teleoperated mobile robots) are widely used in many fields, including for example manipulation with dangerous objects (explosives, contaminated objects, chemical substances), exploration of unknown spaces, military applications etc. The operator must typically rely purely on information from the sensory subsystem of the mobile robot, especially the camera(s) mounted on the robot as the primary source of feedback.

Complex or delicate remote manipulation tasks can be very difficult with limited visual feedback of the robot chassis, the manipulator arm, the surroundings of the robot and the object of manipulation. It is not always feasible to equip the robot with a huge number of cameras, but a pair of stereovision cameras can help with the very important sense of depth in comparison with the still more typical simple 2D view [1–3]. The problem, however, is how to clearly and comfortably display the stereoscopic scene to the operator. One of the possible ways is to use a *head-mounted display (HMD)* [4] that presents different images for each eye, thus recreating the stereoscopic 3D impression.

In the last years, there has been a big evolution in the area of virtual reality headsets, started by the innovative HMD device *Oculus Rift* [5] and followed by other similar

rival products (*HTC Vive* [6], *Sony PlayStation VR* [7]). The VR devices are now incomparably better and cheaper than a few years ago and they are thus finding still new areas of application – besides the obvious entertainment industry or science visualizations, VR is used for example in military command and control systems as a new way of visualization of battlefield [8, 9] or in mobile robot remote control [10, 11].

1.1 Previous Work – Virtual Operator Station

One research project on the Department of Robotics (VSB-Technical University of Ostrava, Czech Republic) focused on improvements of existing control applications of teleoperated mobile robots and the result of this project was development and implementation of a new approach – the *virtual operator station* [12, 13].

The virtual operator station is displayed in virtual reality (VR) using the Oculus Rift HMD. The operator wears this device and the virtual reality places him in an empty black room with several virtual screens around him and a 3D model of the robot, see Fig. 1.

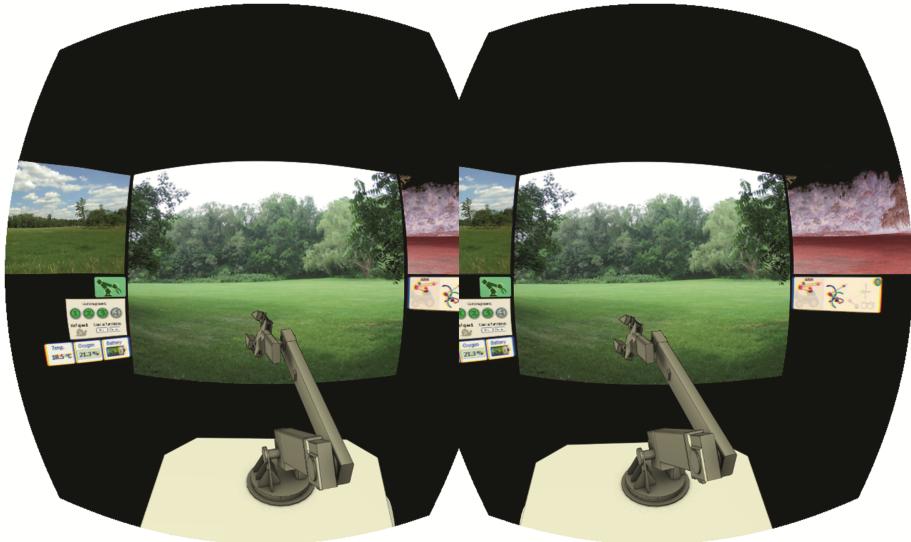


Fig. 1. Virtual operator station in Oculus Rift (2D composite view of the two displays)

The screens show images from the cameras mounted on the robot. It is also possible to display stereo-vision images on these screens – such a screen then works similarly to a 3D television or cinema – and can have much bigger size than physically possible or affordable.

The 3D model of the robot is also very important because by quickly looking at it, the operator gets a very clear, illustrative and accurate overview of the mechanical position of the manipulation arm in relation to the chassis. Displayed are also other information about the robot (sensor readings, battery status, operation modes and configurations etc.).

The virtual operator station was fully implemented for the first developer version of Oculus Rift (DK1), then adapted to the second improved version (DK2) and finally also for the final customer version. It was applied on several mobile robots developed by the Department of Robotics, especially the robot TAROS V2 (Tactical Robotic System) – unmanned military robotic mobile system solved in the frame of *CAFR (Center for Advanced Field Robotics)* [14].

1.2 Detected Problems

Extensive testing and practical use of the virtual operator station verified its advantages but also showed some disadvantages and problems.

One of the disadvantages was brought with later versions of the Oculus Rift control software and drivers. While the first versions were not strict about connection of the video output from the computer to the HMD, later versions require connection to a HDMI port of a dedicated graphics card [15], which is a serious problem for most notebooks, even powerful ones with a separate graphics card (the HDMI port is shared between the power-saving integrated graphics chip on the motherboard and the high-performance GPU and Oculus Rift cannot recognize it). The device targets more on desktop PCs, or expects notebooks especially certified for VR (only very few – and very expensive – were available). This is a problem for mobile robot teleoperation because it is usually not acceptable to carry a full desktop PC together with the mobile robot.

It also turned out to be a bigger problem than expected that the operator of the mobile robot wearing the HMD device has no visual feedback from his surroundings, which can be even dangerous to him; and he is tied to the control computer with a cable (Oculus Rift is not wireless). It is much better when the operator is sitting (which can be problematic sometimes in field conditions), as this prevents him from moving around, pulling the cables, colliding with some objects around him or even falling somewhere.

2 Augmented Reality

In contrast with *Virtual Reality (VR)* [16], which puts the user into a completely virtual world (hiding the real world around him), *Augmented Reality (AR)* allows the user to see the real world around him and puts some virtual objects into it (or modifies existing objects) [17]. The first functional AR systems were developed for military (air force) use; semi-transparent head-up displays (HUDs) have been used in aeroplane cockpits for many years.

Using AR instead of VR has the potential to remove some of the drawbacks of the virtual operator station displayed in VR, so it was decided to try to move the virtual

operator station from VR to AR. Augmented reality can also bring some additional features to mobile robot teleoperation and assist the operator in new ways.

2.1 Augmented Reality Headsets

Augmented reality is often presented with handheld devices like tablets or smartphones, where the virtual objects are added to the camera images of the real world displayed on the screen of the device [18]. Spatial alignment of the virtual objects is done typically by image processing and usually requires special 2D markers placed in the real world. This type of AR is, of course, unusable for the creation of an AR virtual operator station because there would be no practical advantages.

Augmented reality glasses, like *Vuzix AR3000* [19], *Sony SmartEyeglass* [20], *Epson Moverio BT-200* [21] or *Google Glass*, have limited use and are designed typically for displaying of simple overlays with various information (text messages, news, photos etc.).

There are only a few commercially available HMDs for AR that provide fully immersive augmented reality. We can name for example the *Meta 2 Augmented Reality Development Kit* [22], which is not in a customer version yet but is available for developer use. It offers a 90-degrees field of view, 2560×1440 resolution, hand interaction and positional tracking sensors and a 720 p front-facing camera. It requires a wired connection to an external PC, the device itself is only a display with sensors.

Another AR headset is *HoloLens* from Microsoft [23]. This HMD was available to us thanks to a project grant, so the new version of virtual operator station was first implemented using HoloLens. The first prototype was created for the mobile robot Hercules [24], which has all the necessary features (teleoperation; a mobile chassis; a complex manipulation arm; and cameras including stereovision).

2.2 Microsoft HoloLens

The HoloLens HMD is a fully stand-alone device that provides immersive AR without connection to a computer. The device contains a 32-bit Intel CPU and a special HPU (Holographic Processing Unit), 2 GB RAM + 1 GB RAM for HPU, 64 GB flash memory, and runs on Windows 10. The hardware further contains [25]:

- 2.3 megapixel widescreen stereoscopic see-through display,
- spatial sound technology (built-in speakers),
- inertial measurement unit (accelerometer, gyroscope, magnetometer),
- 4 environment understanding cameras,
- 1 depth camera ($120 \times 120^\circ$),
- a 2 MP photo/HD video camera,
- 4 microphones,
- ambient light sensor.

The extensive sensory system of the device with movement prediction and image stabilization provides excellent stability of the holograms in the real world. Interaction

with the user can be done by gaze tracking, gesture input, voice commands or standard blue-tooth input devices (keyboards, gamepads etc.) (Fig. 2).



Fig. 2. Microsoft HoloLens

The *Software Developer Kit (SDK)* available for HoloLens supports the *Unity engine* [26], but applications can also be created at a lower level, using *DirectX 11*. The AR applications are developed as *Windows 10 UWP (Universal Windows Platform)* applications using Visual Studio. Unity development provides greater abstraction over the hardware and simplifies the process, while DirectX 11 programming provides more flexibility and performance.

The virtual operator station was created without the help of Unity. The rendering process differs from standard DirectX 11 applications in the following details:

- An instance of the *HolographicSpace* class must be created (this class controls full-screen rendering, provides camera data and other important features).
- Holographic cameras must be created and configured (the cameras represent the user's eyes in the virtual space).
- Frames of reference must be set (one frame is attached to the head, other frames are stationary in the environment).
- During rendering, once per frame, the application must acquire a *HolographicFrame* which contains predicted camera position and orientation and other necessary information.
- Rendering must be done for each camera (each eye), using the camera parameters provided by the SDK.
- Specific to HoloLens applications is also the way of processing gaze and gesture inputs.

3 Virtual Operator Station in AR

The basic idea of the virtual operator station remains the same as in the previous research. The main difference is that it is not displayed in a purely virtual space; it is augmented into the real world instead.

3.1 Contents of the Virtual Station

The virtual operator station contains three basic types of objects:

- flat screens (planes) with camera images (one for every camera mounted on the robot),
- 2D icons or simple 3D models with text information showing the status of the robot (sensors readings, battery level, mode of operation etc.),
- 3D interactive model of the robot showing the real actual position of the manipulator arm and all other movable parts of the robot.

In VR, all these objects were arranged in a specific way around the operator in the virtual space. This changes in AR – the individual components of the operator station can now be fully and easily customized by the user. A specific virtual object can be selected by the *Gaze gesture* (looking at the object) and performing the *click gesture* (index finger movement). This object can then be freely positioned in the space and placed anywhere by performing another click gesture.

HoloLens automatically makes a 3D map of the surroundings and the resulting 3D mesh can be used to detect objects in the real world and align holograms to them – for example, place a hologram on a table or on a wall. The mesh can also be used to hide holograms behind real objects. The AR virtual operator station utilizes this functionality and allows the operator to attach the virtual objects to all detected flat surfaces. It is thus possible to place the 3D model of the robot on a table and the screens with camera images either also on a table (horizontally) or on a wall (vertically).

This is, of course, possible only when the operator is in an indoor environment; in field conditions, this can be done for example in a truck car or in a tent (military applications). In open space outdoor environments (or indoor without an appropriate wall), the operator station is rendered floating in the air. The holograms in HoloLens are fixed in the real world and do not move with the operator (unless it is explicitly programmed this way). If the operator wants to move elsewhere, he can perform a double-click gesture and the whole operator station automatically moves in front of him.

Examples of two possible configurations of the virtual operator station are shown in Figs. 3 and 4. The operator station in both cases contains images from two cameras (a front camera mounted on the arm and a rear camera facing backwards), two panels with status icons and information text, and a 3D model of the robot Hercules representing the real actual state of the robot.



Fig. 3. View of one configuration of the virtual operator station (picture was taken using the *mixed reality capture* function of HoloLens)

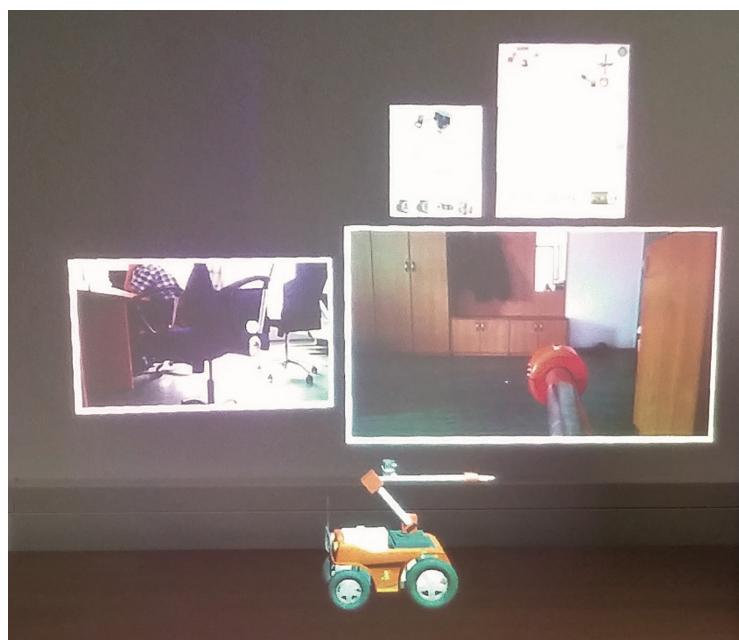


Fig. 4. An example of a different configuration of the components of the virtual operator station (picture was taken directly through the glasses by a digital camera)

3.2 Communication with the Robot

The HoloLens hardware specifications include Wi-Fi 802.11ac and Bluetooth 4.1 LE. Both these wireless standards can be used for communication either directly with the mobile robot or with some external communication device (the HMD device contains only a built-in Wi-Fi antenna, so the reach may be limited and a repeater may be necessary).

The SW control system of the mobile robot Hercules contains two parts – a server application running on a small PC located inside the robot chassis (this SW controls all electric motors and reads data from all sensors); and a client application running on a touch-screen notebook located in the operator's case (this SW displays camera images and other information to the operator and gets input from him).

These two applications communicate via Wi-Fi, using the UDP protocol and custom packets. The Windows 10 operating system installed on HoloLens is capable of UDP communication, so it was possible to transfer the client application to the HMD device (the graphical user interface had to be completely rewritten) (Fig. 5).



Fig. 5. Communication with the robot – from the existing operator station or from HoloLens

Camera images in the existing control system are transferred to the operator station in a simple way – as a sequence of JPEG images for each frame of every active camera. This method was chosen a long time ago because of its very low latency. Compared to modern video compression methods (especially H.264), it requires much more bandwidth. This problem was partially reduced by thorough optimization of the transfer – transferred are only bodies on the JPEG files (not headers) and only the primary camera is transferred in high quality. In case of lower Wi-Fi signal, the quality of the images is automatically gradually lowered.

This simple method of video encoding was advantageous for migration to HoloLens because just a quite simple JPEG decoding algorithm had to be implemented instead of a video stream reception and decoding. The decoded JPEG images are easily mapped as textures to the DirectX polygons representing virtual video screens in AR.

3.3 Stereovision Cameras

Various methods of stereovision image display to the user in VR were tested in the previous project [12, 13] and the best way proved to be to map the stereo-image on a plane in VR, simulating a 3D television or cinema.

This is easily applicable also in AR – the virtual 3D television screen can be displayed as a hologram. The main difference between a simple 2D camera image display (see previous chapters) and a stereoscopic image is that two camera images (left and right) with identical compression quality have to be transferred from the robot and the holographic plane then has to be rendered to each eye with a different texture.

Convergence also has to be dealt with – either by *Horizontal Image Translation (HIT)* and cameras with parallel optical axes as suggested for the VR; or by using a mechanism with 1 degree of freedom and the ability to converge optical axes of the cameras on a chosen focus point [27].

4 Evaluation of Basic Types of Operator Stations

The acquired experiences with various types of mobile robot operator stations showed their advantages and disadvantages. The following comparison is made only for operator stations that meet the following requirements:

- usable both indoors and in field conditions (portable),
- capable of Wi-Fi communication with the robot,
- can get input commands from the operator,
- can display the necessary status and sensors data from the robot, including a 3D model of the robot,
- can display camera images in acceptable size and quality,
- can display stereovision images.

Some of the requirements are not solvable by devices like smartphones, tablets or similar systems. Included in this comparison thus will be the following three types of operator stations:

- a notebook or a PC (with a 3D monitor to fulfil the stereovision requirement),
- virtual operator station based on VR (Oculus Rift),
- virtual operator station based on AR (HoloLens).

4.1 Criteria Rating

Rated for each type of operator station will be the following criteria:

- costs (approximate minimal required costs),
- size and portability,
- convenience during use (how comfortable it is to use the station),
- interaction with the surroundings during use (awareness of the surroundings),
- disturbing effect of the surroundings (especially ambient light),
- stereovision display (how clear and good is the 3D immersion),

- legibility of additional information (text, icons),
- the ability of multiple people to cooperate on teleoperation of the robot.

Each criterion gets the mark from 1 (worst) to 5 (best).

It is important to mention that the total numbers do not directly determine the “best” or “worst” variant because the table does not incorporate any kind of criteria weight or importance. This may differ – for some applications, for example, price may be a very limiting factor and size may not be an issue; for other applications, it may be exactly the opposite. The numbers from Table 1 are meant to be taken rather as seven separate comparisons of the three types of operator stations, not as a general objective evaluation.

Table 1. Criteria rating of the three types of operator stations (1 = worst, 5 = best)

	PC/Notebook	Oculus rift	Hololens
Stereovision immersion	1	5	4
Costs	3	2	1
Size & portability	1	2	5
Comfort	4	1	5
Surroundings awareness	5	1	4
Ambient light interference	3	5	1
Text legibility	5	2	2
Multi-person cooperation	5	3	4
Total	27	21	26

4.2 Comments on Some of the Criteria

Stereovision immersion is an important factor in our research because stereoscopic images can help with complex manipulation tasks and lose their meaning without an acceptable way of 3D rendering. The PC/Notebook variant requires some kind of 3D LCD display technology, typically with some active or passive glasses, and the immersion is only very basic. On the contrary – the virtual operator station (both in VR and AR) excels in this field and this was one of the main reasons for its development.

Although HoloLens is very expensive at the moment (\$3,000), it is the only hardware required. Oculus Rift must be supplemented with a powerful PC and the first option requires a 3D screen. This also affected the given marks in the size and portability criterion.

As far as comfort is concerned, Oculus Rift is very problematic, because it is heavy on the head and must be connected by a relatively short cable to the computer. HoloLens is graded slightly higher than PC/Notebook – although it also has to be worn on the head, it feels more comfortable than Oculus Rift and it has the advantage that the operator can freely move around, sit or stand and is not tied to one place. And the PC/Notebook variant in fact also requires some kind of glasses (although just light-weight and simple) to be worn for the 3D monitor.

Awareness of the surroundings and the negative impact of ambient light are very clear criteria with expected results. Oculus Rift completely shields the user from the surroundings in both the extreme negative and extreme positive way.

Cooperation. Cooperation of multiple people on teleoperation of a mobile robot is easiest using a simple computer because the computer screen can be directly watched by many people (each of them needs the glasses for stereovision).

Oculus Rift is solely a one-person experience. It, however, requires a computer with a screen anyway and it is very easy to use this screen as a secondary output, showing either a direct copy of the image in the VR (Fig. 1) or a standard user interface adapted for the 2D view. In this case, however, only one person gets the benefits of VR (including the stereovision).

HoloLens is also a one-person device, but multiple HoloLens HMDs can be synchronized together to share the augmented world and every user can watch the virtual operator station from his own perspective, or every user can see his own configuration of the virtual station. Data from the robot have to be transferred to every HoloLens (broadcast), or one HoloLens can serve as a proxy and distribute the data – this, however, increases the lag. Input commands can come only from one person (the commander), or it is possible to distribute responsibilities among multiple people (Fig. 6).

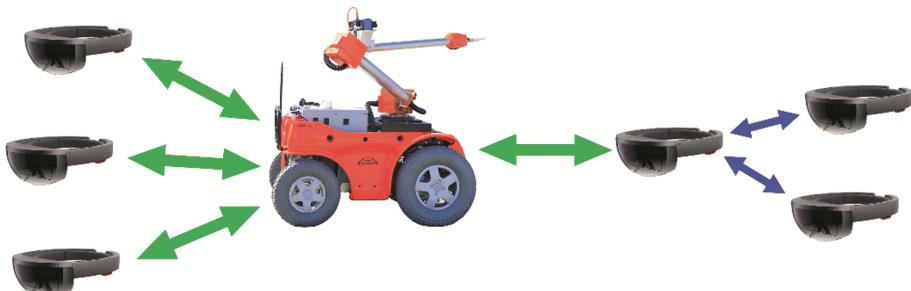


Fig. 6. Connecting more HoloLens devices to a robot directly (left) or through a proxy (right)

5 Additional Uses of Augmented Reality

The virtual operator station based on augmented reality can be supplemented with additional features that are possible in AR.

A typical use of AR, in this case, could be for example display of service information for maintenance [28]. Another scenario is a display of sensor readings and other important data directly on the real physical robot, as can be seen in Fig. 7. The information is displayed as holograms (text, icons, animations...) attached to a non-rendered virtual model of the robot which is aligned with the real robot. This process is not directly supported by HoloLens and it would require some external API for image processing that would find special marks placed on the robot. HoloLens can provide the camera image for this because it is equipped with a forward-looking camera.

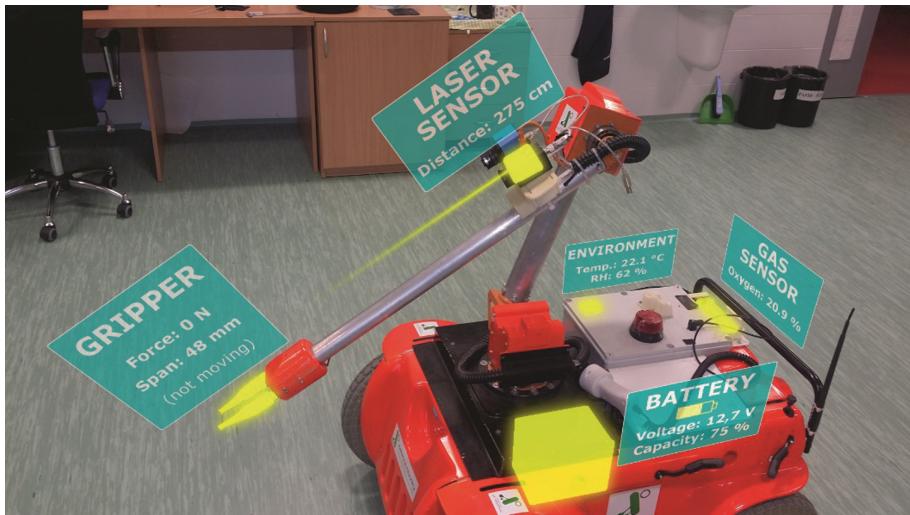


Fig. 7. Display of sensor data on the physical robot in AR (idea visualization)

6 Conclusion

The virtual operator station in AR has not been fully tested in real situations yet, unlike the previously designed operator station in VR. The system has been implemented in a usable prototype version and most features proved to be achievable with HoloLens. There are some problems still to solve; for example, it is problematic to find a gamepad controller with Bluetooth connection compatible directly with HoloLens.

The evaluation of the three different types of operator stations (a standard one with a computer and a screen; the virtual operator station in VR; and the virtual operator station in AR) pointed out some important advantages and disadvantages of the individual types. It is important to mention that there is not a clear general winner – the importance of individual criteria depends on the particular application.

The virtual operator station based on AR and rendered using HoloLens excels in stereovision immersion (the device contains a pair of stereoscopic see-through displays), size and portability (HoloLens weighs only 579 g and does not require any additional equipment), comfort of use (the HMD feels comfortable on the head and there are no cables), and also provides good awareness about the surroundings – which is an important improvement over the virtual operator station rendered using Oculus Rift.

Probably the biggest drawback is the very narrow angle of view, which is equivalent to watching a 15" screen in a standard viewing distance (this applies only to the augmented objects; the real world is not considerably obscured by the HMD). This greatly limits the amount of information the operator can see at one moment without having to move his head, and the virtual camera screens (the most important source of feedback) must be either quite small or moderately far away. Another very important limitation of HoloLens is related to its principle of additive rendering – the holograms

can only add light to the real-world background scene, not subtract from it. This means that a black colour in a hologram is actually completely transparent. In normal situations and especially in indoor environments, this is not a huge problem, as can be seen in Fig. 3 (the real view is much better than this image because it was taken by a camera with the lens outside of the ideal eye position). The device, however, cannot be properly used outdoors under a clear sky and some sort of physical shading must be used. The device is also considerably expensive.

It can be expected that other similar products – or a new version of HoloLens – come to the market shortly, with improved hardware parameters and lower prices. Then this concept could be used with even more advantages over the typical operator stations used nowadays.

Acknowledgements. This article has been elaborated in the framework of the research project “Augmented Reality Lab” under the programme “Support of Science and Research in the Moravian-Silesian Region 2016” (RCC/08/2016) and supported by the specific research project HS3541602 in cooperation with VOP CZ s. p. The article was also supported by the project “Research Centre of Advanced Mechatronic Systems” (Ministry of Education, Youth and Sports of the Czech Republic).

References

1. Cybernet: Operator Control Unit. <http://www.cybernet.com/products/robotics.html>
2. Orpheus Robotic System Project. <http://www.orpheus-project.cz/>
3. Fong, T., Thorpe, C.: Vehicle teleoperation interfaces. *Auton. Robots.* **11**, 9–18 (2001). ISSN 0929-5593
4. Wikipedia: Head-mounted Display. http://en.wikipedia.org/wiki/Head-mounted_display
5. Oculus VR: Oculus Rift. <http://www.oculusvr.com/rift/>
6. HTC: Vive. <https://www.vive.com/us/>
7. Sony: PlayStation VR. <https://www.playstation.com/cs-cz/explore/playstation-vr/>
8. Františ, P., Hodický, J.: Human machine interface in command and control system. In: Proceedings of 2010 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems, VECIMS 2010, pp. 38–41 (2010). Art. no. 5609345
9. Františ, P., Hodický, J.: Virtual reality in presentation layer of C3I system. In: Proceedings of International Congress on Modelling and Simulation: Advances and Applications for Management and Decision Making, MODSIM 2005, pp. 3045–3050 (2005)
10. Chen, G.S., Chen, J.P.: Applying virtual reality to remote control of mobile robot. In: Proceedings of the 2nd International Conference on Intelligent Technologies and Engineering Systems (ICITES 2013), pp. 383–390 (2013)
11. Benoumeur, I., Zoubir, A., Reda, H.E.A.: Remote control of mobile robot using the virtual reality. *Int. J. Electr. Comput. Eng. (IJECE)* **5**(5), 1062–1074 (2015)
12. Kot, T., Novák, P.: Utilization of the Oculus Rift HMD in mobile robot teleoperation. *Appl. Mech. Mater.* **555**, 199–208 (2014). ISBN 978-3-03835-111-5
13. Kot, T., Novák, P., Babjak, J.: Virtual operator station for teleoperated mobile robots. In: International Workshop on Modelling and Simulation for Autonomous Systems, MESAS 2015, Prague, Czech Republic, 29–30 April 2015, pp. 144–153 (2015). ISBN 978-3-319-22383-4

14. CAFR. <http://www.cafr.cz/>
15. Oculus Support Centre: My Oculus Rift Headset isn't Connecting. <https://support.oculus.com/526448734199646>
16. Wikipedia: Virtual Reality. https://en.wikipedia.org/wiki/Virtual_reality
17. Wikipedia: Augmented Reality. https://en.wikipedia.org/wiki/Augmented_reality
18. CNET: Augmented Reality comes to Mobile Phones. <https://www.cnet.com/news/augmented-reality-comes-to-mobile-phones/>
19. Vuzix: AR3000 Series of Smart Glasses. <https://www.vuzix.com/Products/Series-3000-Smart-Glasses>
20. Sony: SmartEyeglass Developer Edition. <https://developer.sonymobile.com/products/smarteryeglass/>
21. Epson: Moverio BT-200. <https://www.epson.cz/products/see-through-mobile-viewer/moverio-bt-200>
22. Meta: Meta 2 AR. <http://www.metavision.com/>
23. Microsoft: Microsoft HoloLens. <http://www.metavision.com/>
24. Department of Robotics: Mobile Robot Hercules. <http://robot.vsb.cz/mobile-robots/hercules/>
25. Microsoft: HoloLens Hardware Details. https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details
26. Unity: Unity Game Engine. <https://unity3d.com/>
27. Aguiar, J., Pinto, A.M., Cruz, N.A., Matos, A.C.: The impact of convergence cameras in a stereoscopic system for AUVs. In: International Conference Image Analysis and Recognition, pp. 521–529 (2016). ISBN 978-3-319-41500-0
28. Hayes, J.: Is Augmented Reality a Breakthrough for Field Service Teams. <http://www.engineering.com/DesignSoftware/DesignSoftwareArticles/ArticleID/12434/>