Improving the Understanding of a Remote Environment by Immersive Man-Machine Interaction

11

David Böken, Michael Schluse and Jürgen Rossmann

Abstract

In a changing world, the way we interact with machines must change as well. Teleoperation becomes more important. This poses its own set of challenges. To solve these a new Human-Machine interface must be developed. By developing this HMI around the concept of immersion, these challenges can be solved. This new kind of HMI can be applied to different fields. Examples using forestry or remote robot operations are demonstrated.

Key words:

Man-machine interaction • Human-machine interface • Remote operation • Intuitive visualization

D. Böken (✉) · M. Schluse · J. Rossmann

Institute for Man-Machine Interaction RWTH Aachen University, Ahornstrasse 55, 52074 Aachen, Germany

E-mail: boeken@mmi.rwth-aachen.de

URL: <https://www.mmi.rwth-aachen.de/en/>

M. Schluse

E-mail: schluse@mmi.rwth-aachen.de

J. Rossmann

E-mail: rossmann@mmi.rwth-aachen.de

© The Author(s) 2022

T. Schüppstuhl et al. (eds.), *Annals of Scientific Society for Assembly, Handling and Industrial Robotics 2021*,

https://doi.org/10.1007/978-3-030-74032-0_11

1 Introduction

Industry 4.0 combined with the ideas of the Internet of Things (IoT) is changing the way machines and entire factories are operated. Everything is interconnected, is getting more automated, and thus is becoming more complex [1]. The way humans interact with machines is changing. As with remote control, humans don't have to be in the same room as the physical machine, human control can be centralized, control is facilitated and eventually, the number of human operators can be reduced.

Current Human-Machine Interfaces (HMI) are not yet well prepared for this claim. A well suited, modern HMI should provide an intuitively comprehensible overview and intuitively operable interaction metaphors to interact with the physical system.

Modern technologies can be used to create intuitive HMIs for these use cases. We present two examples leveraging such technologies for the benefit of the user.

In the first example, modern simulation and visualization techniques are used to enable the intuitive understanding of a forest and of the effects, different harvesting measures will have. We will show that based on the HMI-principles of "feedback" and "affordance", basic features without futuristic technologies already make a big difference.

In the second example, Virtual Reality (VR) technologies are applied to create an intuitive user interface concept for a remotely controlled robot, see Fig. 1. It connects the Human on the left with a robot on the right.

Abb. 1 Structure for connecting a real robot to a virtual reality headset

2 Current Developments in Man-Machine Interaction

2.1 Remote Operation

The way machines are operated today is changing [1]. While previously the operator would stand next to the machine or operate it from the next room, nowadays the physical distance between man and machine as well as the cognitively relevant distance increases. This poses several challenges.

The further away the operator is, the longer it takes for the information to reach him and for his input to be relayed back to the machine. Depending on the application, this latency can be critical.

Another problem lies within the amount of information being relayed via the existing man-machine interface. Current machine interfaces rely mostly on the concept of providing the user with information the machine can interpret itself. But the user also pulls information from different sources.

Take a CNC-Router for example. While the router's interface provides the user with an operational status such as its progress it normally does not provide additional information like vibration or the noise level of the machine. This information is vital as trained operators rely on this additional information. If the machine is too loud or vibrates too much it could be an indicator that the stress on the machine is too high and settings need to be adjusted in order to avoid damaging the machine or the workpiece.

A different problem occurs when considering information that may not be intuitive for a human. This problem is not particular to remote operation but through detaching the operator from his target system, it becomes more pronounced. Previously the human could look at the machine itself to try and make sense of its' data. Without that possibility, he needs to understand the situation solely relying on the data. An example is 2D laser scans. When controlling a mobile robot platform it is easy to show a camera view with which an operator can survey the surrounding. But 2D laser data is often presented as a 2D view from the top. So the operator needs to transfer that knowledge from a 2D view into his 3D awareness. Often he needs to look at several monitors to get all the information and most have different ways of providing that information. This is a challenge for the operator and imposes additional mental strain.

2.2 Reliability

Today's machines are getting more and more complex. In itself, this is not a problem [2]. The problem arises through the reliability of these systems. Grieves [2] defines four categories of system behavior. Predicted Desirable (PD), Predicted Undesirable (PU), Unpredicted Desirable (UD), and Unpredicted Undesirable (UU). The first three classes (PD, PU, and UD) are not problematic, the last class is (UU). With rising complexity, the risk of a failure

to be catastrophic also rises [3]. In his book “Normal Accidents”, Perrow describes the idea of Normal Accidents, or “system accidents”. Small unpredicted events or failures can cascade through a system unpredictably and cause large events with severe consequences. One prominent example of a small event causing a big failure is the “Space Shuttle Challenger disaster” [3]. Here a simple failure of an O-Ring caused the loss of vehicle and crew.

This should be taken into account when creating a new HMI for remotely operated systems. A new HMI should not be more complex as it needs to be, and if possible complexity should be reduced. At the same time, it has to enable the operator to manage such complex machines without himself becoming part of the complexity.

3 Immersive HMI

To meet these requirements a new way of interaction between a human and the operated machine must be developed. One that does not rely on the human being present at the scene of operation, but one that does not necessarily introduce more complexity into the interaction.

Recent studies have shown, that the sense of vision is the most important in English speaking cultures [4]. While the study suggests that this old hypothesis does not appear to be true for every culture, it still is true for the English one.

The findings of the study suggest that a system should primarily rely on the human sense of vision.

3.1 The Concept

The new HMI needs to address the problems described in Sect. 2. The idea is to “virtually” bring people to the area of interest by using an intuitive visual representation of information collected from the target area. This immerses the operator into the situation of the target system. Using the real sensor data and presenting it to the operator helps him to get a sense of the situation.

By chaining different systems the potential of an “Unpredicted Undesirable” event is raised. By keeping it low, the probability of “Normal Accidents” can be reduced. One possibility to mitigate this problem might be to present the user with all the raw data that is processed for him. The idea is that by making the raw information accessible, the user can verify that the system is working as intended. He can see what is really happening and not

only what the system is interpreting. Otherwise, the user would just be another chain in the link of systems and can himself cause a “Normal Accident” based on the wrong information preprocessed by the target system.

This concept does not apply to all types of information. Some information can only be understood by the user if they are preprocessed for him. In those cases, other methods to avoid an “Unpredicted Undesirable” event need to be developed.

At the same time, the system needs to provide a benefit over existing solutions. This is done by giving the user access to as much information as available. Not just as numbers with a label but by actually putting them into a context which he can immediately understand.

Imagine an IoT dashboard within a 3D representation not just with pictograms but actually showing what their values are and showing where the data was collected.

4 Application: Forestry

This concept can be deployed in several different fields. The first example will be a more basic one which shows where this concept is already achievable. This example is taken from forestry. To be more specific privately owned small forest property.

In Germany, there are many small privately-owned forest property. Due to the demographic change, a lot of those properties are inherited by younger generations. Oftentimes they have migrated to cities and lack a personal connection to the forest and knowledge about it.

This results in forest parcels sitting dormant in the forest without being used or cultivated. This leads to problems.

One problem lies withing the economic usage of the surrounding forest. If some owners want to use their forest, they can only do so in an ecological way by combining their property and treating it together, but if someone is not interested in his forest where will be spots which can not be used and this forestry will be more difficult.

4.1 Visualizing the Information

One idea to mobilize the younger generation into getting more involved with their forest is to give them easy access to what it means to own part of a forest and what you can do with it. They often don't know what it means to own a forest. Part of that is that they do not know what they own. They never visited the forest and caused by their disinterest, do not plan to change that.

Abb. 2 App view showing forest at different growth state

The idea is that by using an intuitive visualization you can immerse the user into what he personally owns. This helps to better transfer the knowledge of what they own, and what they can and should do with it. Only through that understanding will the owner be able to make an educated decision based on all facts and not just disinterest.

Through providing the user with a 3D representation the new forest owner is more likely to identify himself with his property and increase interest. See Fig. 2 for an example of a 3D representation of a forest.

Actual engagement can be encouraged by simulating different forest treatments and presenting the results to the user in an intuitive way. This is feedback to the user. He can set different parameters for the simulation and look at the result. This engagement helps fortify the relation to the forest and its possibilities.

Even though all these materials are generated from flyover imaging or satellite pictures they still provide a much higher value especially to the uneducated user who maybe never thought about any properties a forest might have. The user does not have to think about the information he is seeing because it is mostly just like a view he would have in real life. This uses the principle of “affordance”. The user might not know what a specific amount of wood is, in numbers. But showing him the trees instead gives him a better sense of the situation.

By not displaying graphs about the forest growths, but instead showing a slide show or a video of the forest growing, the facts are conveyed to the viewer in a simpler and more accessible way.

Figure 3 shows the structure of the system required to visualize the forest. To reach as much people as possible a mobile app is used as a gateway for the users to access and visualize their forest. As mobile devices are not capable of executing all tasks necessary to generate a 3D representation of the forest, this task is outsourced to a server. Interaction with the forest is handled in a way that is compatible to current structures. The app takes instructions from the user and generates tasks for the worker inside the forest to execute.

Abb. 3 Structure of a system presenting the real forest to the user

5 Application: Mobile Robotics

Another example where this principle can be applied is remote-controlled robots. While the above example is more of a current use case, this one is more futuristic. Inspiration for this can be taken from different sources. For example from “The Matrix” movies, specifically their flight control room, see [5].

Another source of inspiration for this futuristic interface can be taken from “Bret Victor” and his “Seeing Spaces” [6]. His Vision describes the next generation of “Maker-Spaces” or Workshops. See [7] He describes tools to allow someone to get an insight into a system. He proposes to record and archive lots of data from a system, including a lot of internals and making them accessible. By visualizing them and providing means of “seeing” them inside his “seeing space”, the human can get a better idea of why a system is behaving like it is.

When dealing with teleoperated robots one of the difficult tasks is to provide the operator with an understanding of the operated robots surrounding. While cameras and 2D Maps from the top provide an insight into the situation of the robot, the operator needs to stitch these information together in his mind to extend his understanding of the situation.

This can be improved by using a system that helps the operator to immerse into the situation of the operated robot. To achieve this a system needs different components. A 3D representation is the best fit for such a task. But instead of using a computer or a smartphone, Virtual Reality (VR) is the best fit. Because a real-world is represented, the operator just needs to look around and can have most information put into context for him.

But visualization is just one aspect of a remote control system. Actually providing input to it, is another.

(a) Virtual robot shown in virtual surrounding representation, Live generated using SLAM

(b) Stereo image of operator pointing at target

Abb. 4 Different visualizations

Just like Fig. 3 for the forestry application Fig. 1 shows the structure used to connect a robot to a Virtual Reality (VR) system. As the VR hardware is already powerful, there is no need for an additional server in between the user and remote side. The data from the hardware system is taken and visualized to the user. The input from the user is taken, combined with the information from the robot, and translated into control commands the hardware can understand.

5.1 3D Visualization

As visualization the information provided by the robot is taken and displayed around the user. Lidar data for example is displayed as walls on the ground around the robot, see Fig. 4a. Other information like Camera pictures are integrated into the world and rendered at the position where the image is captured so the operator has the context where the information is coming from. Other information like the current target of the robot or the path the robot will take are overlaid on the floor within the map of walls.

5.2 Input Devices

Due to the challenging way of making interactions with a Virtual Reality (VR) intuitive, different kinds of hardware are needed.

For 3D visualization, an HTC Vive was chosen. Instead of using the provided controllers, which have to be held to provide input, a mixture of a LeapMotion sensor and a Kinect V2 was chosen.

The LeapMotion sensor was mounted to the VR headset, while the Kinect was positioned in front of the user. The LeapMotion was used to track the individual fingers, while the Kinect was chosen to track the Body Position including Hands. The Kinect tracking data made it possible to survey the Hands while they were not in front of the LeapMotion Sensor.

The information from the different devices are combined to provide better coverage for different types of input from the user.

5.3 Interaction

To simplify the interaction and to provide better immersion, the tracking information from the user is shown in the 3D representation. Thus the user can identify when there is a problem with the VR Hardware and can act to avoid causing further problems.

The actual interaction is separated into several parts. One part is navigating in and manipulation of the visualization, the other is operating the target system. A third way of interacting is interaction with a window style of UI elements displayed floating in front of the user, either fixed in the world or fixed to the user, depending on what information is shown or requested from the user.

Interactions with the 3D representation revolve around the way information is displayed to the user. You can move the map around you, scale it, and manipulate what is shown. This can help the user to get a different perspective of a situation. He can also zoom into the scene to inspect details. Trough zooming out he can get a better overview and view the surroundings.

The motion the user uses for these actions is comparable to a drag and drop action. By closing his hands (recognized by the Kinect Sensor) the user starts the action. By opening them again the user ends it. The actions are translated into the world 1 to 1. So “grabbing” something and it around is translated into a map movement. Moving hands closer together is translated into a Zoom gesture. The same goes for moving hands further apart. The user can select different modifiers for this operation. If he wants to move faster or more precisely he can change the factor which links his actions to the movement of the Virtual Reality (VR). The same goes for zooming and rotation actions.

Abb. 5 Virtual window UI element shown to the user in VR, every button and slider is controllable via the virtual hand

Interaction with the real system is operated by the same input capabilities. But instead of manipulating the 3D representation, the real system is manipulated and the results are shown to the user.

One possible action is pointing at the target and telling the robot to move there. The action of pointing is mostly known, but the problem is detecting when the user is pointing where the robot should go, or if he is still deciding. Another problem lies within the accuracy of determining what the user is pointing at. Probably everybody knows this problem, someone is pointing at something and it still turns into a guessing game sometimes. Only when you try to stand where the person pointing person is can you see what they point at. To avoid miscommunication between the system and the operator, a line is drawn where the system thinks the user is pointing. That way the operator only needs to move his index finger, and make sure the line goes where he wants it to go. See Fig. 4b.

Events or special ways of requiring input from the user are handled through popup windows in the 3D representation. This way of displaying or inquiring information from the user is familiar, or easy to understand. The user simply needs to press the buttons with his finger. Windows can be moved around and positioned where the user wants/needs them. See Fig. 5.

The concept of using windows as a UI element has another benefit on the development side. Developers adapting to this new style of HMI can use a proven way of interacting with the user and don't need to relearn the complete way of interaction. This can also reduce the possibility of mistakes. A new system always has a higher likelihood of the user misunderstanding what is wanted. The developer might think that it is clear what the system might want from the user, but while operating the user misinterprets this information and that might lead to a mistake. Using proven methods as a base for operations can function as a base for new interfaces.

(a) A constellation of two satellites next to each other

(b) UI presented to the user in front of a satellite, prompting him to select a specific component of the satellite

Abb. 6 View using Virtual Reality (VR) goggles, including a visualization of the arms and hands

Most times when a window System would be useful will be when the system has to operate outside of the normal limits. Here a common interaction way is most important which proven UI elements like buttons can provide.

5.4 Different Area of Application

Space is another field where the mobile robotics approach can be utilized. Visualizing the current state of a system using Virtual Reality (VR) can help the user understand the current state of a system. See Fig. 6a. This can help the operator to make educated decisions from the ground station. On the other hand, satellites are complex systems containing lots of parts. Lots of parts are not directly visible to the user, so their state needs to be observable differently. This is where the UI system can help. For example when trying to select a specific component within the satellite, to get further information or execute commands. The user can be presented with a UI listing of all components and can select the desired one. See Fig. 6b.

6 Conclusion

To solve the challenges presented by for example industry 4.0, a new HMI is needed. By using the concept of immersion with intuitive user interaction, this new HMI type can be created. The idea of immersion can be applied to different fields in varying degrees of complexity. Forestry and robotics were given as examples.

The more basic forestry example demonstrates what is already possible with widely available hardware, while the more futuristic robotics example provides a direction in which this concept might develop.

While some problems may be solved with this approach there are still more problems to consider, like suppressing “Unpredicted Undesirable” behavior.

By combining an immersive interface with a digital twin (DT), the DT can act as a mediator between the interface and the actual system [8]. This link can offer more benefits. It allows representing the complete system. Using this representation “Unpredicted Undesirable” behavior can be mitigated [2]. The DT can also provide a standard interface to connect to other DTs and provide a unified interface to immersive interact with multiple systems at the same time.

This leads to the conclusion that integrating digital twins is the next step to further push the concept of an immersive HMI.

Literatur

1. Dorst, W. (Ed.): Umsetzungsstrategie Industrie 4.0: Ergebnisbericht der Plattform Industrie 4.0. Bitkom Research GmbH. (2015)
2. Grieves, M., Vickers, J.: Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems (2017). https://doi.org/10.1007/978-3-319-38756-7_4
3. Perrow, C.: Normal accidents: living with high-risk technologies (1999). ISBN: 978-0-691-00412-9
4. Nadja Podbregar: Die Hierarchie der Sinne. (2018). <https://www.wissenschaft.de/gesellschaft-psychologie/die-hierarchie-der-sinne-2/>
5. Grime, T.: Matrix reloaded: virtual control design (2003). <http://www.kazumichi.com/Matrix-Reloaded-Virtual-Control-Design>
6. Victor, B.: Seeing spaces. In: Talk at EG Conference (2014)
7. Victor, B.: Seeing spaces (2014). <http://worrydream.com/SeeingSpaces/>
8. Cichon, T.: The digital twin: mediator for man-machine interaction (2019). <https://doi.org/10.18154/RWTH-2020-01645>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.