doi:10.3311/BMEZalaZONE2022-007

Teleoperation with a Real-time Digital Twin

Viktor Remeli[1a], Ádám Boronyák[1], Viktor Tihanyi[1]

[1] Department of Automotive Technologies, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics

[a] Corresponding author: remeli.viktor@kjk.bme.hu

Abstract

One of the state-of-the-art trends in intelligent transportation systems research is the development of real-time digital twin technologies and their applications for increased safety and autonomy on the roads. The areas of deployment also include test tracks, logistic yards, factories, hospitals, airports, railways, and any such constrained and managed environment where transportation automation may be of benefit. While applying the same principles as for open road networks, managed environments allow the efficient deployment of optimally distributed sensory infrastructure while also observing certain simplifying assumptions (constraints on the expected scenarios), both factors leading to earlier adoption and ROI. In the case of fleets or swarms of automated vehicles, certain cost savings can be enacted by recovering the environmental model exclusively from infrastructure sensors. However, such a setup is never exempt from the necessity of occasional manual intervention performed by a teleoperator. In our current paper we explore system and HMI designs that would facilitate safe and efficient on-demand teleoperation functionality in such a system.

Keywords: *digital twin, ITS, real-time, teleoperation*

1 Real Time Digital Twin

There is an increasing demand for functionalities and services that can be derived from the real time digital twin of highly automated environments. A highly automated environment consists of a distributed system of heterogeneous platforms that must work together towards common and individual goals. Smart roads and intelligent and connected vehicles of the future comprise such an environment. Logistic, agricultural, and military uses also come to mind as related areas with some overlapping requirements. Advances in highly automated manufacturing colloquially known as industry 4.0 [1] point in a similar direction: an Internet of Things (IoT) network of subsystems must establish efficient cooperation even though aspects of sensing, decision making, and control are highly distributed.

While the employed solutions and technologies differ depending on the application, real time digital twin systems can be most beneficial when the task to be solved involves a distributed environment perception problem. The objective of the digital twin is to maintain a real time, consolidated and consistent digital replica of the environment [2]. This joint model should be more reliable and complete than any model the individual perception systems could provide by themselves. A clear case is provided by objects that cannot be detected from a given platform due to cover or the lack of appropriate sensors. Another, slightly related case is the need for coordination of platforms in certain tasks, or their remote management.

The digital twin can be established in the cloud or on the platforms themselves. A logically centralized digital twin can be physically centralized, or physically distributed. We can also imagine cases where several competing or complementary digital twins might be employed, as in the case of peer-to-peer communicating platforms. In this paper we will explore the possibilities relating to the cloud-based, logically centralized digital twin. The exact deployment of the data processing steps (platform-side, edge-side and central) will be of less consequence to our discussion. For more details on the implementation of intelligent transportation and automotive digital twin systems of the mentioned type, refer to previous work, notably including [2] and [3].

The specific problem we want to address within the context of the digital twin is the task of manual teleoperation of the participating platforms. This kind of task is expected to come up in nearly any industrial use-case or scenario, as the business workflow must be ensured even in case of environmental disturbances or platform malfunctions.

A cost-effective method of mitigating a large proportion of such situations is remote human teleoperation of the otherwise autonomous system. Another field of interest is telemanipulated robotics that focuses on remotely solving dexterous tasks with humanoid machines and robotic arms [4], operating in harsh environments like elevated radiation risk [5], or solving complex tasks that require specific and remote expertise like surgeries, see [6].

In this paper we explore how such tasks can be approached when a real-time digital twin of the operation environment is readily available.

2 Teleoperation workflow

Teleoperation can proceed according to different workflows and can require different approaches strongly dependent on the constraints of the communication medium – latency and bandwidth [7]. Below we describe some high-level ideas for a general approach that is suitable for most use cases involving a number of teleoperable platforms and a – usually much smaller – number of operators.

The teleoperation task in general begins with a *triggering event*, which first stops certain aspects of the ongoing automation, second it generates a *notification* towards the teleoperating personnel. The operator then decides whether to solve the situation using manual *remote-control* subsystems or to further escalate it. After the situation has been solved, the control must be *handed back* to the automation system. Thus, the simplified sequence of phases in a typical teleoperation workflow are:

- ***Handover phase***
- ***Teleoperation phase***
- ***Handback phase***

2.1 Handover phase

The handover phase consists of three discernible (but not necessarily sequential) steps for the platform:

- it begins with the *detection* of the trigger event,
- it continues with an *emergency maneuver* intended to stop affected aspects of the automation,
- and it finishes by issuing an *event update* towards some (centralized) event broker.

The platform steps are then followed by steps taken by the notification subsystem and the operator:

- incoming events generate necessary *notifications* towards affected parties (operators),
- operators confirm *receipt and acknowledgement* of notification,
- operator initiates coordinated *handover* of control to the operator.

The digital twin can be especially useful in detecting the trigger event, in coordinating the emergency maneuver, and in serving as the event, notification and handover broker. We will explore these possibilities step by step.

2.1.1 Trigger event detection, maneuver, and update

Detection of the triggering event is more likely and more accurate if a platform can rely on sensors of other platforms as well. This can also be cheaper or even unavoidable if outside (off-platform) sensors are in some way more suited for the task, e.g. a loitering drone (single or swarm) over the field of operations. In this case, the central system performs the fusion of the incoming data streams to produce a unified environment model. Real-time requirements, especially regarding raw data processing and communication channels, are crucial in this regard. Such requirements will finally decide whether processing is done on-platform and whether raw data is communicated. It is often the case that only processed metadata is sent to the cloud. Therefore, central fusion usually occurs at object-level. For example, the current spatial state of intelligent platforms and other important participants or obstacles is estimated in real time based on incoming detection or track streams. The detection of a trigger event pertaining to one or more platforms must then occur also in real time, and preferably on the integrated environment model that is considered most robust.

In fact, trigger events can be detected in three main ways: (a) a platform locally detects a trigger event pertaining to itself, (b) a platform locally detects a trigger event pertaining to other platforms (as well), and finally, (c) the central system detects an event that affects one or more platforms. In case (a), the emergency maneuver should be initiated immediately, and the event update sent to the central broker. In case (b), the event update must be sent to the cloud: the affected platforms are immediately notified by the central system to perform the proper maneuvers.

This can also be achieved in a peer-to-peer (P2P) fashion. The triggered platforms must take care to deduplicate events which might point to the same situation only detected from different broadcasting sensory platforms. When case (b) is implemented with central processing, however, the deduplication is a form of data fusion that happens centrally. Case (c) assumes that fusion of perception sources happens on a lower level, prior to the event detection.

In all cases (b) and (c), and sometimes even (a), event data is backpropagated to the appropriate platforms without delay allowing them to perform the emergency maneuver in time. The exact maneuver decisions and controls are calculated either on the platform itself, or in the cloud. How this is handled depends not only on the compute capacity of the setup, but also on whether the set of environment data necessary for maneuvering is more readily available in the cloud (via collective perception) or on the platform (via unique sensors that are not connected to the central system). The best of both worlds is achieved when the platform receives a fused collective perception data stream from other sensing platforms and installations, while also having access to its own sensors.

2.1.2 Notification and handover

Every time a trigger event occurs for a platform, an update is dispatched to the central broker and awaits further processing to notify operators. A subscribe-publish scheme can be implemented to allow different stakeholders to be notified of different events via different channels, and even using differing schedules (e.g. some may need immediate SMS notification while some may just want a weekly summary in their email). For audit purposes, it might be important to develop ways to confirm the date and time of notification receipt and acknowledgement. Since the requirements for the notification system are fairly general, certain software tools like [8] may already support much of the required functionality out of the box.

The central system can also act as an arbiter for the handover, especially when multiple operators might want to take control of the same system at the same time. The management of the teleoperation state of the platforms can be handled centrally.

2.2 Teleoperation phase

Using the digital twin approach to teleoperation, the advantages over classical manual operation go beyond the reduction of the need for personnel to be present physically (passively on alert or actively operating) at various locations, sometimes in rapid succession. Notably, extended perception, virtual reality and central management aspects contribute to unique and crucial capabilities that cannot be delivered by traditional setups.

2.2.1 Extended perception

The digital twin allows for extended environment perception using multiple sensors and sensors of different platforms. This can have several advantages, starting with the obvious capacity to teleoperate based on raw data from many disparate sources. Relevant camera images (or other sensor data) can even be injected dynamically into the teleoperation user interface according to the state (e.g. position and trajectory) of the controlled platform and its relation to all the surrounding intelligent platforms and installations.

We note that the conventional kind of teleoperation that relies on a single sensor data stream (video camera without a digital twin) usually reduces the situational awareness of the operator, as it usually lacks both spatial and modal aspects of the data – e.g. a fixed field of view and no audio. Extended perception allows more precise and safer remote manual control not only due to better situational awareness of the teleoperator, which is the major factor, but also as it enhances the possibility for further electronic assistance mechanisms like collision avoidance (emergency braking), and other ADAS (automated driving assistance) and related functions.

2.2.2 Visualization

Since some (or all) data may also be available in processed form, the concept of extended perception can be visualized in an augmented reality continuum ranging from a collage of video streams (entirely real) up to a 3D simulation (entirely virtual – so called synthetic vision systems [9]), including all stages in between (e.g. point clouds and bounding boxes projected on images and other attention assistance and information conveying methods much like heads-up displays on modern aircraft). Certainly, in visualizing such complex environments, augmented reality (AR) glasses can be of great help, for instance see [10]. The ergonomic design of such interfaces is of notable importance, as they must support the recommendation, choice, and visual integration of the most relevant and useful information streams. A typical problem of data integration and joint visualization may arise around time synchronization of different data streams due to a range of factors including differing latencies experienced on various channels.

2.2.3 Central management

Intelligent coordination of automatically and remotely controlled operations quickly becomes a requirement in any system that has a large number of platforms and/or teleoperators. Besides unified handling of authentication, authorization and other security related issues, resource allocation and deconfliction are also among the obvious benefits of central management, another one is auditing. A further benefit can be that the heterogeneity of the platforms can be encapsulated and made transparent to the user, who only must know how to access the central teleoperation interface.

It can be argued that the necessity for centralization could become even more pronounced if the controlled system is a complex machine that requires the simultaneous coordination of several operators (e.g. certain military platforms), or conversely, if several simple machines have to coordinate in a common environment. The same is true for coordinating manual control in semi-automatic platforms and in environments with platforms of various levels of automation.

2.3 Handback phase

The central system setup can aid the operator in finding the proper time, location, mode, etc. for the re-automation or handback of a given aspect of control, with particular regard to other platforms' state and the mission objectives. The central system can also oversee the timed handback at a later or more opportune moment without the need for the operator's intervention or presence. The mission of the re-automated platform can be a default mission, one defined by the operator via the central system, or one chosen by an algorithm that decides using the central environment and system state model.

3 Conclusion

We have examined the theoretical possibilities of teleoperation in highly automated environments with a centralized real-time digital twin. We conclude that depending on the problem type and complexity, there can be significant potential benefits to be explored and exploited using the digital twin setup. The benefits derive from two main properties: the *common environment model* allows better event detection, maneuvering and teleoperation via extended perception, while the *central state management* of the system helps with resource allocation, notification and auditing.

Acknowledgement

The research reported in this paper and carried out at the Budapest University of Technology and Economics has been supported by the National Research Development and Innovation Fund (TKP2020 National Challenges Subprogram, Grant No. BME-NC) based on the charter of bolster issued by the National Research Development and Innovation Office under the auspices of the Ministry for Innovation and Technology.

References

[1] Lasi, H., Fettke, P., Kemper, H. G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. Business & information systems engineering, 6(4), 239-242.

[2] Tihanyi, V., Rövid, A., Remeli, V., Vincze, Z., Csonthó, M., Pethő, Z., ... & Szalay, Z. (2021). Towards Cooperative Perception Services for ITS: Digital Twin in the Automotive Edge Cloud. Energies, 14(18), 5930.

[3] Krämmer, A., Schöller, C., Gulati, D., & Knoll, A. (2019). Providentia – a large scale sensing system for the assistance of autonomous vehicles. In Robotics: Science and Systems (RSS), Workshop on Scene and Situation Understanding for Autonomous Driving.

[4] Jang I., Carrasco J., Weightman A., Lennox B. (2019) Intuitive Bare-Hand Teleoperation of a Robotic Manipulator Using Virtual Reality and Leap Motion. In: Althoefer K., Konstantinova J., Zhang K. (eds) Towards Autonomous Robotic Systems. TAROS 2019. Lecture Notes in Computer Science, vol 11650. Springer, Cham. https://doi.org/10.1007/978-3-030-25332-5_25

[5] Allspaw, J., Roche, J., Yannuzzi, M., Yanco, H. A., & Lemiesz, N. (2018, July). Remotely Teleoperating a Humanoid Robot to Perform Fine Motor Tasks with Virtual Reality-18446. WM Symposia, Inc., PO Box 27646, 85285-7646 Tempe, AZ (United States).

[6] Meng, C., Wang, T., Chou, W., Luan, S., Zhang, Y., & Tian, Z. (2004, April). Remote surgery case: robot-assisted teleneurosurgery. In IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 (Vol. 1, pp. 819-823). IEEE.

[7] Lichiardopol, S. (2007). A survey on teleoperation. Technische Universitat Eindhoven, DCT report, 20, 40-60.

[8] Amazon Simple Notification Service https://aws.amazon.com/sns

[9] Alexander, A. L., Wickens, C. D., & Hardy, T. J. (2005). Synthetic vision systems: The effects of guidance symbology, display size, and field of view. Human Factors, 47(4), 693-707.

[10] Szalai, M., Remeli, V., Tihanyi, V. (2022). Development of 3D Visualization for Displaying Real-time Sensor Information from Multiple Sources. First Conference of ZalaZONE Related R&I Activities of Budapest University of Technology and Economics 2022