



Article

# IMPAct: A Holistic Framework for Mixed Reality Robotic User Interface Classification and Design

Dennis Krupke <sup>1,2,\*</sup> , Jianwei Zhang <sup>2</sup> and Frank Steinicke <sup>1</sup>

<sup>1</sup> Human-Computer Interaction (HCI), Department Informatics, University of Hamburg, 22527 Hamburg, Germany; steinicke@informatik.uni-hamburg.de

<sup>2</sup> Technical Aspects of Multimodal Systems (TAMS), Department Informatics, University of Hamburg, 22527 Hamburg, Germany; zhang@informatik.uni-hamburg.de

\* Correspondence: krupke@informatik.uni-hamburg.de

Received: 28 February 2019; Accepted: 5 April 2019; Published: 11 April 2019



**Abstract:** The number of scientific publications combining robotic user interfaces and mixed reality highly increased during the 21st Century. Counting the number of yearly added publications containing the keywords “mixed reality” and “robot” listed on Google Scholar indicates exponential growth. The interdisciplinary nature of mixed reality robotic user interfaces (MRRUI) makes them very interesting and powerful, but also very challenging to design and analyze. Many single aspects have already been successfully provided with theoretical structure, but to the best of our knowledge, there is no contribution combining everything into an MRRUI taxonomy. In this article, we present the results of an extensive investigation of relevant aspects from prominent classifications and taxonomies in the scientific literature. During a card sorting experiment with professionals from the field of human–computer interaction, these aspects were clustered into named groups for providing a new structure. Further categorization of these groups into four different categories was obvious and revealed a memorable structure. Thus, this article provides a framework of objective, technical factors, which finds its application in a precise description of MRRUIs. An example shows the effective use of the proposed framework for precise system description, therefore contributing to a better understanding, design, and comparison of MRRUIs in this growing field of research.

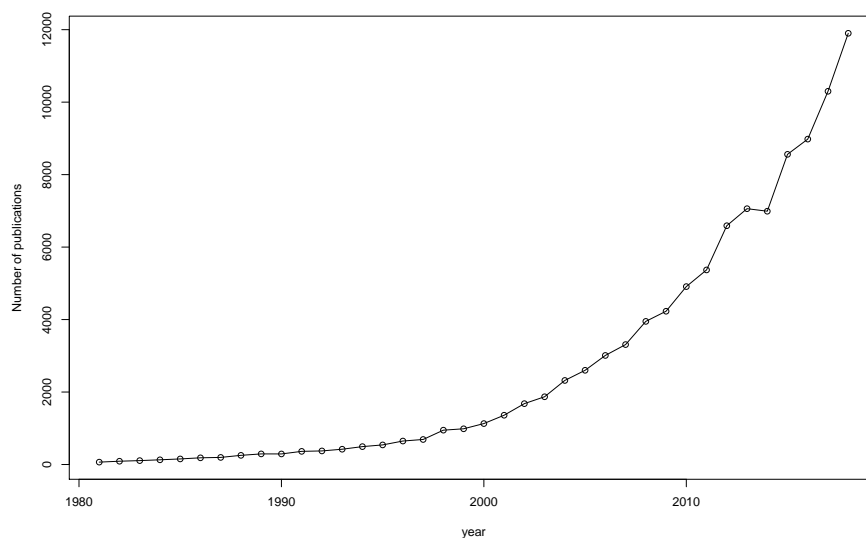
**Keywords:** mixed reality; robotic user interface taxonomy; human–robot interaction; robotic user interface; holistic user interface design; user interface analysis

## 1. Introduction

Currently, robots are becoming a part of our daily life in households, for example as vacuum cleaners, lawn mowers, and personal drones, but also in the area of intelligent toys. Industrial machines have started to collaborate with human co-workers, and assistive robots find application in health-care. The development of accessible and secure robots is a major challenge for current and future applications. Burdea [1] already pointed out how virtual reality (VR) and robotics are beneficial to each other. Intuitive, natural, and easy-to-use user interfaces for human–robot interaction have the potential to fulfill the needs in this context. While a fully-automated robotic companion, which is capable of performing almost every task for humans, still remains a vision of the future, it is essential to also focus research on the interaction between humans and robots. Mixed reality (MR) technology provides novel vistas, which has enormous potential to improve the interaction with robots.

*“The implementation of a mobile and easy deployable tracking system may trigger the use of trackers in the area of HRI. Once this is done, the robot community may benefit from the accumulated knowledge of the VR community on using this input device.” [2]*

The enormous increase in research interest of MR combined with robotics is demonstrated by counting the number of new scientific publications during the past few years (cf. Figure 1). Many contributions involving robotic user interfaces (RUI) and the field of MR are already published, and the number of new publications per year regarding mixed reality robotic user interfaces (MRRUI) shows a trend of exponential growth. To the best of our knowledge, there is no classification integrating all relevant aspects of this highly multi-disciplinary field of research available in the literature. However, a taxonomy of MRRUIs would represent a valuable tool for researchers, developers, and system designers for a better understanding, more detailed descriptions, and easier discriminations of newly-developed user interfaces (UI). Therefore, we propose a holistic approach that incorporates the relevant aspects of MRRUIs for system design, while supporting a comprehensible overview and understanding the interconnections and mutual influences of the different aspects.



**Figure 1.** Yearly number of new publications listed on Google Scholar containing the keywords “mixed reality” and “robots” (date of acquisition: 13 March 2019).

The main aim of this contribution is to provide an appropriate structure for the classification of robotic user interfaces involving mixed reality. The authors believe that good and successful design starts with understanding all relevant fields by performing a holistic analysis. The proposed structure is generated from unfolding prominent and relevant taxonomies, continua, and classifications from the most important aspects into a list of highly-relevant factors. Only factors being under direct control of system designers are taken into account; thus making it possible to select or determine their actual values without the need to be concerned with the mental models of probable users of the system. This is beneficial for both purposes: system design and classification. Due to the strong interdisciplinary nature of the topic of MRRUI, a new taxonomy is needed that summarizes all relevant factors and provides the necessary connections, explaining the interplay between them.

In this article, the “IMPAct framework” is introduced for the classification of mixed reality-based robotic user interfaces. As explained in Section 3.2, it identifies different factors from the categories Interaction, Mediation, Perception, and Acting as active contributors to actual MRRUI implementations. The factors are grouped into different categories in order to provide a memorable structure for application. Its purpose is to categorize existing MR-based robotic systems and to provide the process of designing new solutions for specific problems, in which interaction with robots in MR is involved, with a guideline to include all relevant aspects into the decision process. Single factors are explained as an aid for their application. The authors selected the included factors carefully to avoid dependency on the mental models of individuals. Thus, only technical, measurable, or identifiable factors were taken into account. As a result, the IMPAct framework serves as an effective tool for discriminating different

MR-based robotic systems with a focus on the UI. Additionally, it can be consulted for designing or improving MRRUI by making use of the holistic view in order to create effective and specialized UIs.

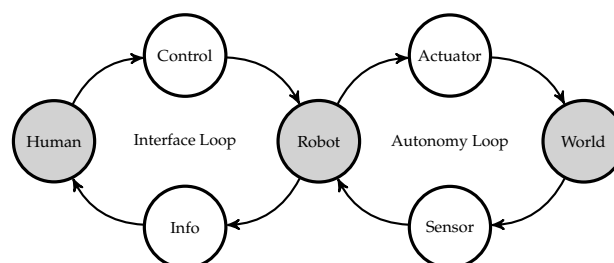
The remainder of this article is as follows: Section 2 introduces the general structure of human–robot interaction, starting with teleoperation, and extends the model to a more detailed version with the aim to describe MR systems for human–robot interaction more specifically. The section concludes with an overview of prominent theory papers regarding the relevant fields of 3D interaction, mixed reality (MR), robot autonomy, interaction with virtual environments (VE), and many more. In Section 3, the experimental procedure is explained and the results are summarized. The extraction of relevant factors from theoretical papers, the card sorting experiment, and the second clustering into major categories are described in Section 3.1. Section 3.2 presents the results from the card sorting experiment and the following refinement in tabular and graphical representations. A validation of the results is presented in Section 4 by applying the taxonomy to an MRRUI described in a publication. The validation is discussed in detail for every separate group and also provides additional information beyond the given information in the tables. Section 5 summarizes the article and discusses the results and their limitations and future work. The Appendix contains the complete table of factors for reference (cf. Table A1) including the factor’s publication source. Table A2 represents the short profile of the system categorized in Section 4, and Table A3 is the empty template, ready for use to classify other MRRUI.

## 2. Related Work

In the following, the main components involved in MRRUI are identified and relevant taxonomies, classifications and definitions from scientific publications are listed and discussed very briefly.

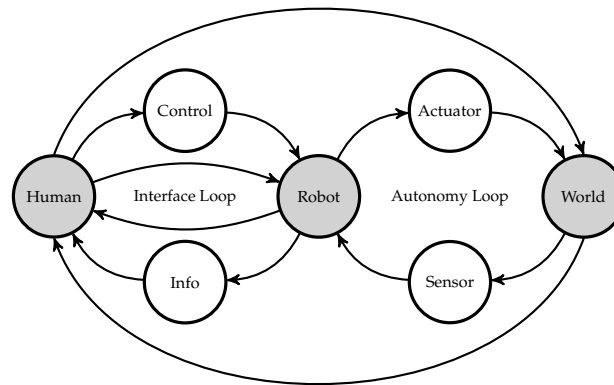
### 2.1. The Structure of Information Flow in Human–Robot Interaction

The main components in a human–robot system and how they are connected were already described in Crandall et al. [3]. Figure 2 shows an adapted version. This basic scheme of a robot remote operation identifies *human*, *robot*, *world* as the essential entities to perform a human–robot interaction task. Required components on the side of the interface between human and robot are responsible for transferring *control* data to the robot and some *information* from the robot to the human as feedback. In such setups, the robot is able to manipulate the world and to interact with the world physically, or at least is co-located with other real-world entities. To fulfil this task, the robot needs to gather information about the world it is acting on by utilizing *sensors*. It can influence the actual state of the world by utilizing *actuators*. To a certain degree, the robot can act autonomously. The knowledge of the robot about the world and its own state is accessible for the human operator by the user interface in general and provides the operator with some information helpful for making decisions about the next control.



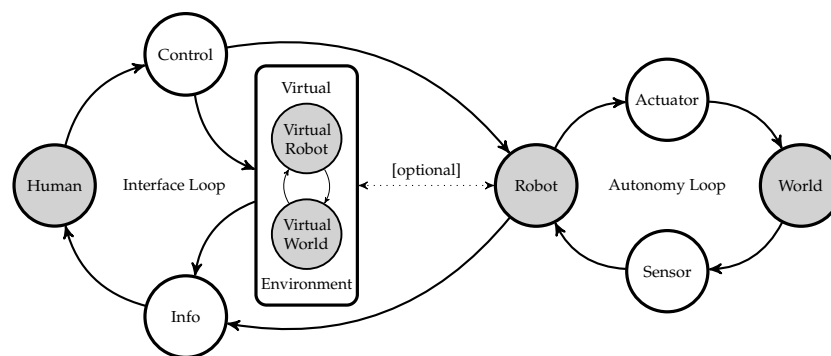
**Figure 2.** Remote robot operation (adapted from Crandall et al. [3]).

In a scenario where human and robot are sharing the same physical space (cf. Figure 3), there can be additional direct and physical interactions of the human and the robot, and the human and the world. Both of these schemes explain how humans can interact with a distant or nearby robot in order to manipulate the real world, but there is no detail about how VEs are probably embedded into this.



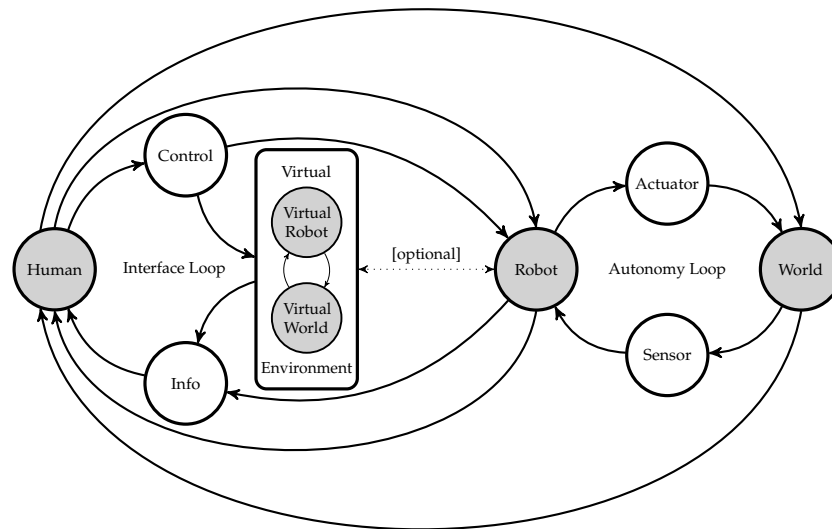
**Figure 3.** Co-located robot operation (inspired by the remote robot operation by Crandall et al. [3]).

Enhancing the basic scheme of remote operation with the component of a *virtual environment* containing a *virtual robot* model and a model of a *virtual world* (cf. Figure 4) reveals new potential interactions. One conclusion is based on the fact that even though a remote operation scenario is described, the operator has the possibility to influence directly the virtual world by virtual physical interaction or implicitly by utilizing some linkage between the VE and the real robot. This is the case if we use VR as an additional component in our human–robot system. By including a virtual model of the robot and the world in the system, the same control means, used for the real robot, are applicable to the VE. However, interaction with components of a VE removes the physical boundaries of our real world. The implementation of VR-based user interfaces means for the system designer that he/she has to deal with a much higher complexity in creating an appropriate means of interaction. Extending the scheme of VR robot operation by integrating MR technology means for the operator that he/she is not separated from the real world anymore. The multitude of interconnections further grows, and direct interaction with the real world is additionally possible.



**Figure 4.** Virtual reality robot operation (inspired by the remote robot operation by Crandall et al. [3]).

Enhancements of the classical co-located robot operation scheme with a VE by using MR technology (cf. Figure 5) couples the operator directly with the real robot and the real world, while all relationships of the VR robot operation are maintained. We can conclude that mixed reality robot operation is the most flexible and complex case we can currently imagine to implement a robotic user interface. Especially how the interaction with the VE and the VE itself is connected to the real robot is a challenging question for MRRUI design and implementation.



**Figure 5.** Mixed reality robot operation (inspired by the remote robot operation by Crandall et al. [3]).

## 2.2. On Classifying Mixed Reality Robotic User Interfaces

Much effort went into understanding and classifying a multitude of aspects relevant to the large fields of human–computer interaction (HCI) and human–robot interaction (HRI). These taxonomies, continua, rules, and guidelines each serve a special purpose. Despite contributing valuable terms and knowledge, combining them together into a kind of composed taxonomy reveals the issue of overlap and conflicting models and involves the risk of misleading from the core concept of the actual focus of MRRUI analytics.

Robinett [4] discusses the topic of *synthetic experiences*. He clarifies many aspects of how modalities are captured and transferred from the world to the human. The other direction, from the human to the world, resembles how to manipulate the world using technical devices. The bidirectional mediation, discussed here, is highly relevant for robotic UIs involving MR.

A taxonomy with VR as one of the extrema is the *AIP-Cube* proposed by Zeltzer [5]. The authors determined three categories, *Autonomy*, *Interaction*, and *Presence*, as important and combined these as orthogonal axes in a cube-like taxonomy. Especially, autonomy is very relevant to robots in general, but in the context of VEs, the question arises where autonomy should be: at the robot side, at the human side, or at the VE side. If we agree not to specify the location of an MRRUI in detail, but to define the UI as the interplay of different aspects and components involved in enabling the interaction, we can find a subset of suitable elements for classifying MRRUIs.

Presence, as defined by Slater and Wilbur [6] in the *FIVE framework* (“A Framework for Immersive Virtual Environments”), depends on individual mental models of different humans, and thus, the presence level of a certain implementation cannot be specified. Instead, *immersion* is found to be an appropriate technical factor, which can be determined objectively and also contributes to the resulting level of presence.

Strongly related to immersion is the definition of the *reality-virtuality continuum* by Milgram et al. [7] and display classes in the context of the continuum, thus contributing to the level of immersion [7–9]. Major structural properties of MRRUI are specified by these topics.

One very large field that is highly relevant to MR technology is *3D interaction*. Bowman et al. [10] provided a very general discrimination of 3D user interfaces (3DUI) into three different applications, *selection and manipulation*, *travel*, and *system state*. The classification of *hand gestures* by

Sturman et al. [11] provided a very complete taxonomy, which takes the dynamics into account, as well as up to six degrees of freedom (DoF). The proposed classification is also applicable for gestures, other than hand gestures. Interaction techniques using up to 6 DoF and their relevant properties were issued by Zhai and Milgram [12] with the *6-DOF-Input-Taxonomy*. Here, a strong focus lied on the mapping properties of the implementation of an actual input method. The *interactive virtual environments (IVE) Interaction Taxonomy* proposed by Lindeman [13] formalizes interaction methods with arbitrary DoF as a combination of an *action type* and a *parameter manipulation type*.

*Robot classification* is a difficult task, due to the polymorphic characteristic of robots. Some books, serving as an introduction to robotics, list a classification based on their intelligence level, defined by the Japanese Industrial Robot Association (JIRA) [14,15]. Another classification from Association Francaise de Robotique (AFR) describes their capabilities of interaction. Murphy and Arkin [16] introduced classical paradigms of robot behavior. However, there is not only one general definition of a robot. Across robotics' history, many accepted definitions appeared, but in the end, it depends on the specific community that one has selected. Therefore, the question of how a user interface for robots is defined is almost untouched in the literature. In the paper of [15,17], the term *robotic user interface (RUI)* seeds to be used for the first time. In this paper, the authors defined four categories for classifying RUIs with regard to HRI-relevant factors. The authors took into account the *level of autonomy*, the *purpose* of the robot, the *level of anthropomorphism*, and the control paradigm resulting in a certain *type of communication*.

Regarding *autonomy*, there is a long history of research methods addressing questions of design and analysis what should be automated and how far. The oldest and most prominent conceptualization goes back to 1978 by Sheridan and Verplank [18]. The authors introduced *levels of autonomy (LOA)*, a classification for rating the interaction with a computer along with a discrete scale of autonomy. A refined version of LOA, published by Parasuraman et al. [19], offers a more detailed level of granularity by dividing tasks, which can be automated, into four different stages of *acquisition, analysis, decision, and action*. The next step involves *dynamic assignments of autonomy levels*, which are not fixed by design, but depend on the current state of the world [20]. Miller and Parasuraman [21] further improved the model by explaining how tasks at the four different processing stages consist of several *subtasks*, each with its own LOA. Thus, the resulting autonomy level at each stage is the result of a process of aggregation. Discussions of autonomy were not especially focused on issues of intelligent robots; furthermore, they were targeted to questions of industrial automation, e.g., the question of how to replace human workers by machines on specific tasks. This resulted in taxonomies that did not include specific demands and abilities of different kinds of robots. Beer et al. [22] proposed *levels of robot autonomy (LORA)*. Schilling et al. [23] provided a very recent perspective on shared autonomy by taking *multiple dimensions* of relevance for robot interaction into account. A very interesting explanation of the term *autonomy and its connection to intelligence and capabilities* was provided by Gunderson and Gunderson [24]. The human-centered perspective of the effects of interaction with autonomous systems has been explained in detail by several authors [25–27].

Important aspects in human–robot collaboration are the *roles* of the interaction partners, the *structure of the interaction*, and how the initiative and autonomy are distributed among the interacting partners [2,28].

### 3. Materials and Methods

As demonstrated by Adamides et al. [29], it is possible to develop a taxonomy from extensive literature investigation of scientific publications, followed by a session of card sorting in order to cluster the resulting data. Further processing of the intermediate results can lead to a reasonable structure for the classification of an actual topic. Our aim was to create an objective taxonomy for specifying technical factors of MRRUIs. Thus, every aspect involving qualitative evaluation of the system is neglected in this work. Even though we believe that qualitative measures of MRRUIs are of importance, we argue for integrating these factors into another work about MRRUI evaluation.



### 3.1. Procedure

In order to assess important factors for MRRUI classification, we collected high impact publications containing classifications from the fields of *3D interaction*, *mixed reality*, *user interfaces*, *immersive virtual environments*, *robot autonomy*, *automation*, and *many more*. The decision to include or neglect a certain publication was governed by its citation count or the reputation of the main authors. In general, the most important authors were found by reading survey and review papers, followed by a combination of systematic and cumulative literature investigation on Google Scholar, Researchgate, the IEEE Digital Library, and the ACM Digital Library. Then, we extracted factors, which directly impact the nature of the underlying MRRUI, from the investigated papers by unrolling the described taxonomies, definitions, and classifications. Typically, many authors describe classifications of a certain aspect using a discrete list or a continuum, which allows relative ordering of actual instances. In this article, a single property of this kind is called a *factor*. After determining factors in this way, we tried to find aspects not covered so far and identified three more factors that were important in our previous work: *“interaction reality”*, *“location”*, and *“system extent”*. The resulting list contained 62 different, mostly technical factors (cf. Table A1), all being useful to discriminate different approaches of robotic user interface design involving MR. Caused by the technical nature of these factors, their actual values are easy to determine without the need to be experienced with the actual system and without the interference of individual mental models of different specific users or operators. The achieved objectivity of the selected factors makes them useful for other researchers and designers to be utilized in classification tasks.

After extracting these factors, we applied *open card sorting* in order to cluster these into named groups. Due to the high number of relevant factors, we were interested especially in these groups in order to be able to create a more granular taxonomy for a better understanding of relevant categories to the special case of MRRUI design. The card sorting was prepared by noting down every factor from Table A1 to the front side of the cards. The description of the each factor (cf. Table A1) was written on the backsides of the cards. To avoid strong bias by the authors, the participants were instructed not to look at the backside of the cards unless every card was assigned to a named category. During the experiment, 9 experts (3 female, 6 male) from HCI research with their focus on virtual reality, augmented reality (AR), and interaction with 3D user interfaces were recruited to perform open card sorting on a shared set of cards in an open group form. Participants were allowed to enter and leave the session as they liked during a 3-h time slot. The active times of the participants were protocolled. The session took 2 h and 40 min, and as a result, the 62 factors were clustered into 13 groups in an iterative approach. Each participant spent on average 35 min on the task ( $SD \approx 18.498$ ). The protocol revealed a total temporal effort of 5 h and 17 min. During the session, audio recordings were made, and the participants were instructed to think aloud. This was intended to identify misunderstandings of the tasks and to better understand the decisions of the participants. Participants were allowed to note down comments on the cards themselves and to create new cards or to remove already existing cards if an explanation was provided. No comments were made, and no card was removed. One card, DoF, was duplicated by the participants and integrated into two different categories, *interaction parameters* and *input*.

Followed by the card sorting session, the resulting groups were further clustered by the authors into four major categories in order to create a better memorable structure. The session protocol was considered during the decision process to reduce the influence of the authors' peculiar view. This step of further categorization is important to increase the likelihood that the resulting taxonomy is adopted by other researchers. The resulting categories were further discussed with two researchers from our department, who did not coauthor this contribution, but took part in the card sorting experiment, until the result revealed a reasonable structure without logical flaws (cf. Table 1).

### 3.2. Card Sorting Results

Based on a literature investigation, a card sorting experiment with HCI professionals was performed. Further analysis of the card sorting results revealed a general structure for the classification of MRRUIs with four major categories, as shown in Figure 6. The detailed results of the two clustering steps are summarized in Table 1, listing each of the factors according to their associated group and category. The main categories provided a plausible and memorable structure with the aim to improve the workflow during classification or design tasks. The groups of the next detail-level resulted directly from the iterative approach of the card sorting experiment, in which all identified factors from the literature investigation (cf. Table A1) were clustered into groups. Many of these sources were listed in the related work (cf. Section 2).

Table A3 summarizes all factors on a single page, which serves as a template for profiling arbitrary MRRUI designs. Figure 6 together with Table 1, containing the results of the clustering, Table A1, the alphabetical list for reference and explanation, and the empty template in Appendix A (cf. Table A3) serve as a framework for classification and design of MRRUI. As demonstrated in the following section, using the IMPAct framework enables very precise and extensive analysis and description of relevant systems. Since we included only objective factors, mostly technical, which were under the control of the system designer, we can conclude that the results represent a useful framework for initiating diversity and effectiveness in MRRUI research and development.

**Table 1.** Results of the card sorting.

Category	Group	Factors
Interaction	Interaction Parameter	Action Type, AFRInteraction Level of the Robot, Control Alignment, Control Order, Directness, DoF, Interaction Reality, Interaction Type, Mapping Relationship, Parameter Manipulation Type, Sensing Mode, Structure of Interaction
	Paradigm	Human–Robot Communication, Intelligent Robot Control Paradigm, Robot Control Concept, Role of Interaction, RUI-Type, UI-Type
Mediation	Image/Display/Vision	Display Centricity, Display Class (Milgram), (Display) Space, Image Scaling, Vividness
	Input	Action Measurement Type, Directness of Sensation, DoF, Gesture Type, Sensor Type
	Output	Actuator Type, Display Type
Perception	Embodiment	Extent of Body Matching, Proprioceptive Matching
	Immersion/User Perception	EPM, Integration of VE, Level of MR, Reproduction Fidelity
	Modalities	Causality of Modalities, Extensiveness, Inclusiveness, Model of Modality-Source, Superposition of Modalities, User Experience
	Space and Time	EWK, Location, Real-time Level, Surrounding System Extent, Time (interaction), Time (system purpose)
Acting	Autonomy	Level of Capabilities, Level of Intelligence, Plot, Refined LOA, Initiative, JIRAIntelligence Level of the Robot, Timely Aspects of Actors' Autonomy
	Behavior	Robot Behavior Level, Robot Communication Capability, Transparency of System State
	Robot Appearance	Level of Robot Anthropomorphism
	Application	Robot Purpose, Robot Type





**Figure 6.** The overall structure of the mixed reality robotic user interface taxonomy. MRRUIs are directly shaped by the aspects of *interaction*, *mediation*, *perception*, and *acting*. These groups including their relevant subcategories and containing factors generate the IMPAct framework for MRRUI classification and design.

#### 4. Validation

In this section, the results are applied to an MRRUI from a recent publication.

Example: Application of the Framework for System Classification.

In the following, an MRRUI for a pick-and-place task [30] is classified using the IMPAct framework. A detailed analysis of the technical properties in the categories *interaction*, *mediation*, *perception*, and *acting*, including their subgroups, is listed and, if necessary, further explained to enable readers to have a complete understanding of the setup and implementation given.

##### 4.1. Interaction

Interaction is represented by the two groups *interaction parameter* (cf. Table 2) and *paradigm* (cf. Table 3). The actual values of the classification of the analyzed system are noted in the tables. Further explanations are given in the following paragraph.

Some characteristics of the interaction depend on the interaction level of the robot. The AFR describes Class D as an extension of Class C, adding the capability to the robot of acquiring information from its environment. Class C itself is defined as “programmable, servo controlled robots with continuous or point-to-point trajectories” [14]. The degree of spatial matching, expressed by the factor *Directness*, is probably lowered by sensor error. This involves detection of the position of the targets for grasping by the robotic system and the registration process within the implementation of marker detection at the

main UI device for aligning the virtual world components along their actual alter egos. The interaction method for selecting targets needs the human to be at an appropriate position and to look at a certain point at the surface of the target. One could argue that this method involves 6 DoF, 3 for defining a position in the real world of the human's head and another 3 for rotating it as desired. Mathematically, the *rolling* rotation around the forward vector of the head is not used, so 5 DoF is correct, as well. The interaction with real-world objects is mediated through the robotic hardware after selection using virtual counterparts of the actual targets. Thus, the *interaction reality* is rated as virtual. Nevertheless, there is a kind of passive interaction, as well. By utilizing a see-through HMD, the actual targets are observable in the real environment during displaying of planned trajectories, which are visualized by actuating the virtual model in a loop with the very same trajectories. Even if *selection* is the most prominent active interaction concept, here *system control* is identifiable as well, when voice commands are used to start a specific action. Then, the state machine of the system moves over to a different state, changing the way of interaction in the next state; e.g., when confirming a picking trajectory with a voice command, the system proceeds to the *execution* state, and after finishing the grasping, it is ready to receive a command for placing the grasped object on the table. The current system design in the experiment, described by Krupke et al. [30], is for single users, single robots, and performing a single pick-and-place task. However, the implementation would also allow multiple operators to be logged into the system, simultaneously.

Table 2. Interaction parameter.

Factor	Value
Action Type	discrete
AFR Interaction Level of the Robot	type D
Control Alignment	1-to-1
Control Order	0
Directness	almost isomorph
DoF	5
Interaction Reality	virtual
Interaction Type	selection
Mapping Relationship	position
Parameter Manipulation Type	direct
Sensing Mode	isometric
Structure of Interaction	SO, SR, ST

Table 3. Paradigm.

Factor	Value
Human–Robot Communication	two monologues
Intelligent Robot Control Paradigm	hierarchical at user level
Robot Control Concept	feedback control loop
Role of Interaction	supervisor and manager
RUI-Type	interactive
UI-Type	SUI

The robot itself had integrated joint controllers, which were accessed by the dedicated industrial computer, provided by the manufacturer of the robot. The control computer sent goal positions to the joint controllers. The low-level joint controllers itself used a *feedback control loop* to keep joints at a desired position and to alert about position mismatch, or unreasonable high currents, or other issues to the control computer. During virtual or actual robot movements, the operator served as a *supervisor*, but during the planning steps of *selecting a pick position* or *selecting a place position*, he/she fulfills the role of a *leader*. The RUI-type is classifiable as *interactive*, but with a tendency towards very high-level programming. Regarding the history of user interfaces, the system may be called a *spatial* or *supernatural* user interface; spatial because the user is situated in the very same place of the operation

and benefits from exploring the RE by natural walking and looking around; supernatural because he/she can see the probable future. Despite the focus of interaction, there was also classical graphical user interface (GUI) elements realized in the implementation. A head-up display showed the system state in textual form and gave feedback about the success of a given command. Regarding the way commands were given by the operator, the term *natural user interface* (NUI) is applicable because voice commands were triggering the main functions of the system.

#### 4.2. Mediation

Mediation consists of three groups *image/display/vision* (cf. Table 4), *input* (cf. Table 5), and *output* (cf. Table 6). The actual values of the various factors are collected in the tables. The following paragraph contains further explanations.

**Table 4.** Image/display/vision.

Factor	Value
Display Centricity	egocentric
Display Class (Milgram)	class 3
(Display) Space	registered
Image Scaling	1:1
Vividness	mid-to-low

The level of *vividness* is mainly limited by the hardware and the selected software used for rendering. Since the first version of the Microsoft HoloLens with a small field of view was utilized, as well as the general quality of the see-through display, regarding resolution, color range and brightness were taken into account. In addition, the quality of the real environment (RE) content by looking through the head-mounted display (HMD) was negatively influenced, causing overall mid-to-low vividness.

**Table 5.** Input.

Factor	Value
Action Measurement Type	HoloLens inside-out tracking
Directness of Sensation	partially directly viewed
DoF	3 + 1
Gesture Type	static posture
Sensor Type	IMU, structured light IR camera, BW camera, joint encoder, current measurement

The *action measurement* of the operator was mainly contributed by the integrated self-localization feature of the Microsoft HoloLens. Since relevant information is three-dimensional coordinates in the reference frame of the real world, marker detection served the task of providing a reference anchor for the transformation of local device coordinates to RE coordinates. The *directness of sensation* can principally be regarded as direct. The RE was directly viewed, and the virtual components were supposed to be anchored at corresponding points of their alter ego. Without sensor errors, this would be perfectly direct as well, but in our implementation, there was some minor perceptible error. From the robot middleware's view, the system input was just a single 3 *DoF* position within the reference frame of the target object and its name. On the lower level, during user input, walking was used to reach a certain position in the RE, then the posture of the body might be altered, and finally, the head was turned around pitch and yaw axes to move the cursor along the surface of the target object until the desired position was reached. Typically, the system used static posture as the *gesture type* for defining pick and place positions and confirmed the current selection with a voice command. The second condition used the pointing gesture of the index finger; thus, a motionless finger posture was consulted in this case.

**Table 6.** Output.

Factor	Value
Actuator Type	robot arm and gripper
Display Type	see-through HMD

### 4.3. Perception

In the category perception, four groups, *embodiment* (cf. Table 7), *immersion/user perception* (cf. Table 8), *modalities* (cf. Table 9), and *space and time* (cf. Table 10), were analyzed. Relevant factors and their actual values are collected in the tables. Further explanations are provided in this section.

**Table 7.** Embodiment.

Factor	Value
Extent of Body Matching	no VB
Extent of Proprioceptive Matching	virtual cursor and viewing direction

In the analyzed system, there was no virtual body (VB) involved. One could argue that a cursor, which was augmented by the view, matched at least a little the definition of a virtual body. In this case, it should be mentioned that there was no matching between the moving bodies, cursor, and head. However, the proprioceptive matching was appropriate in terms of visualizing the viewing direction and its spatial cues when projecting the cursor on top of the surface of viewed objects.

**Table 8.** Immersion/user perception.

Factor	Value
Extend Presence Metaphor	strong tendency towards real-time imaging
Integration of VE	almost completely integrated
Level of MR	AR with tendency to RE
Reproduction Fidelity	stereoscopic real-time rendering, but with simple shaders

The *extend presence metaphor* (EPM) is rated very high since the human was mostly looking at the RE, but through a see-through HMD. This fact can potentially lower the EPM in comparison to looking directly at the RE. Only some elements like the robot, a table, and some objects for grasping were virtual. However, these elements were anchored with reference to RE coordinates in a very stable way, resulting in a very well-performed integration to the RE. The *integration of VE* by superimposition of the virtual robot on top of the actual robot caused a logical mismatch to natural perception, since human perception usually does not allow the existence of two objects at the very same place. Regarding *level of MR*, due to the fact that there were only a few objects having virtual counterparts, we can conclude that the RE was very dominant. Artificial-looking objects, which were perfectly integrated into the real environment, caused a moderate *reproduction fidelity* since their level of detail was quite low and simple shaders were utilized.

**Table 9.** Modalities.

Factor	Value
Causality of Modalities	partially transmitted and simulated
Extensiveness	only vision
Inclusiveness	almost none
Model of Modality-Source	computed
Superposition of Modalities	merged
User Experience	mixture of transmitted, simulated, and supervised robot

The *causality of modalities* was based on a mixture of data from reading an actual posture of the robot, then transmitting, and finally applying them to the virtual model, which was almost perfectly matching with the real robot. Some displayed trajectories were computed and represented only options for the future, but can become real. Regarding *inclusiveness*, it should be mentioned that sometimes, very bright virtual content on the see-through display can occlude or distract from parts of the real world. Except for this side-effect in the analyzed system, operators were not shut-off from the real world. The *user experience* was classified by three different concepts of data flow (cf. Figures 7–9), which were all present during system use and describe different aspect of the system. *Transmitted experience* (cf. Figure 7) represents especially how information like the current pose of the robot was transferred to the human. *Simulated experience* (cf. Figure 8) describes what happens when a new target position is selected and the generated trajectory from the robot middleware is sent to the HMD. *Robot supervised by human* (cf. Figure 9) represents the case when the robot started to manipulate the world and the human was interacting with a VE.

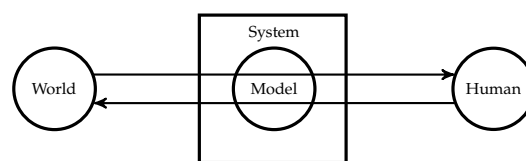


Figure 7. Data flow in a “transmitted experience” (adapted from Robinett [4]).

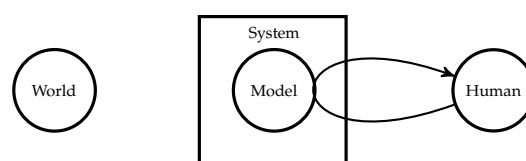


Figure 8. Data flow in a “simulated experience” (adapted from Robinett [4]).

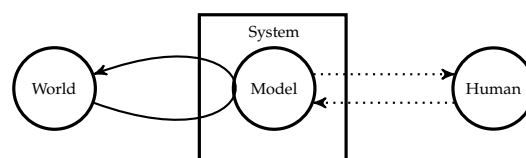


Figure 9. Data flow during “robot supervision” (adapted from Robinett [4]).

Table 10. Space and time.

Factor	Value
Extent of World Knowledge	objects of interest are modeled with normal to high accuracy
Location	in locu
Real-time Level	delayed up to half a second
Surrounding	only 60° FoV
System Extent	network-based in one room
Time (interaction)	frozen
Time (purpose)	real-time

Regarding *location*, it should be mentioned that the analyzed system is intended to be used side-by-side with the actual robot, but technically, it also works as a remote system. Unfortunately, then the cues provided from the RE are missing, and the potential of recognizing problematic situations with probable collisions is reduced. Even if the *real-time level* of the system is classifiable as real time, it is not designed for continuous real-time control from the interaction point of view. Evaluating the factor *surrounding* reveals that, despite the low FoV, the tetherless operation and the inside-out 6 DoF

tracking of the HMD resulted in a spatial experience. The VE can be viewed from arbitrary positions and explored by walking around. Thus, even by the visual experience of looking through a looking glass, spatial presence was generated and improved the resulting experience of the surrounding factor in comparison to a simple 2D video glass with the same FoV. For the *system extent*, it should be mentioned that the HMD was self-contained, making it a head-mounted computer (HMC). It was connected by WLAN with the computer controlling the robot and, if necessary, extendable by additional network-based processing nodes.

#### 4.4. Acting

Acting consists of the four groups *autonomy* (cf. Table 11), *behavior* (cf. Table 12), *robot appearance* (cf. Table 13), and *application* (cf. Table 14). The actual values of the classification of the analyzed system are noted in the tables. Further explanations are given in the following paragraph.

**Table 11.** Autonomy.

Factor	Value
Level of Capabilities	generate trajectories, self-localize
Level of Intelligence	collision avoidance, extract grasp points
Plot	virtual robot shows probable solution
Refined LOA	acquisition: 10; analysis: 3; decision: 1–2; action: 5
Initiative	fixed at the human
JIRA Intelligence Level of the Robot	class 5
Timely Aspects of Actors' Autonomy	static

The *level of intelligence* was divided among two devices. The robot control computer calculated collision-free trajectories according to its dynamic planning world. The input and output device of the operator, the Microsoft HoloLens, assisted in selection of grasp points, according to its programming.

**Table 12.** Behavior.

Factor	Value
Robot Behavior Level	remote controlled
Robot Communication Capability	reactive
Transparency of System State	visible

The *transparency of the system state* was limited by the implementation. All modeled states were displayed during system use in the head-up display element of the UI.

**Table 13.** Robot appearance.

Factor	Value
Level of Robot Anthropomorphism	very low

The robot incorporated only a tendency towards *anthropomorphism* since the gripper was three-fingered and had some similarities to a humanoid hand, e.g., bending of joints is only possible in one direction when starting from a configuration in which all fingers are completely straight.

**Table 14.** Application.

Factor	Value
Robot Purpose	tool
Robot Type	industrial



In the proposed setup, an industrial robot was used as a tool for manipulation tasks in a tabletop scenario.

## 5. Discussion

This section discusses the experimental results and explains their potential impact on MR robotics. Furthermore, the limitations of this contribution and future ideas are explained.

### 5.1. Conclusions and Discussion

In this article, the IMPAct framework for classification and analytical design of mixed reality robotic user interfaces was presented and applied to a robotic pick-and-place system utilizing the Microsoft HoloLens. Relevant factors were carefully investigated from the scientific literature. Prominent taxonomies from relevant fields were decomposed to find important factors that directly influence the system and can be determined regarding their actual values. Since many of these factors represent a position on a continuum and are not exactly measurable, it is, in general, a difficult task. Nevertheless, it was demonstrated that even with these inaccuracies, very detailed descriptions of existing systems are possible. The framework can be regarded as an important and necessary step towards holistic system design and description of MRRUIs.

The authors believe that using the proposed IMPAct framework for exhaustive system descriptions helps to remove ambiguity, as demonstrated in Section 4. By taking an exhaustive list of relevant factors into account, differences from other systems, as well as unique features, become clear. Additionally, the framework provides a methodology for improving actual system designs by providing a standardized template, which enforces intentional decision making.

### 5.2. Limitations and Future Work

Even though the authors performed an extensive literature investigation, it is likely that some relevant aspects are missing in the list of factors and thus are not regarded in the proposed framework. Especially, if certain topics are further developed and new taxonomies arise, the proposed taxonomy should be adapted to these changes and then incorporate new relevant factors of the given topic. In general, the authors believe that a description of MRRUI, covering as many relevant aspects as possible, is desirable. The proposed framework would further benefit from efficient means to assess the values of the listed factors in a more comfortable and time-saving way. An implementation as a wizard-based GUI tool with prepared options, tooltips, and web links for explanations of the single factors would reduce the current workload.

The proposed work was limited to objective and technical measures, which are assessable without any practice with the analyzed or described MRRUI. Qualitative aspects were neglected in this contribution, but are valuable feedback for further improvements in system design. An extension of the IMPAct framework would explain how the objective factors are related to perceptive and qualitative factors. Then, the framework would provide means for holistic re-design as a part of an iterative design process of MRRUI. Furthermore, some robotic web archives are currently arising (cf. <https://robots.ieee.org/robots/>). Currently, there is no such site for MRRUI. The proposed framework within this article could serve as a basis for creating an MRRUI archive.

**Author Contributions:** Conceptualization, D.K. and F.S.; methodology, D.K.; software, D.K.; validation, D.K.; formal analysis, D.K.; investigation, D.K.; data curation, D.K.; writing—original draft preparation, D.K.; writing—review and editing, D.K. and F.S.; visualization, D.K.; supervision, F.S. and J.Z.; project administration, D.K. and F.S.; funding acquisition, F.S. and J.Z.

**Funding:** This research was partially funded by the German Research Foundation (DFG) and the National Science Foundation of China (NSFC) in project Crossmodal Learning, TRR-169.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

3DUI	3D user interface
AIP	autonomy, interaction and presence
AR	augmented reality
DoF	degrees of freedom
EBM	extent of body matching
EPM	extend of presence metaphor
EWK	extend of world knowledge
FIVE	framework for immersive virtual environments
FoV	field of view
GUI	graphical user interface
HCI	human–computer interaction
HMC	head-mounted computer
HMD	head-mounted display
HRI	human–robot interaction
IMPAct	interaction, mediation, perception and acting
IMU	inertial measurement unit
IVE	interactive virtual environment
LOA	levels of autonomy
LORA	levels of robot autonomy
MR	mixed reality
MRRUI	mixed reality robotic user interface
NUI	natural user interface
PC	personal computer
RE	real environment
RUI	robotic user interface
SA	situation awareness
SoA	sense of agency
SoO	sense of ownership
SUI	spatial/supernatural user interface
UI	user interface
VB	virtual body
VE	virtual environment
VR	virtual reality

## Appendix A

**Table A1.** This is the alphabetically-ordered list of the extracted relevant factors from the literature investigation including their possible values as lists or continua. Additionally, the descriptions placed on the backside of the cards from the card sorting experiment are listed.

Factor	Values	Description
Action Measurement Type [4]	tracker, glove, joystick, force feedback arm, ...	What technology is used to measure action of the operator as input to the system?
Action Type [13]	discrete–continuous	If you have to do something permanently or just from time to time.
Actuator Type [4]	robot arm, STM tip, aircraft flaps, ...	What kinds of actuators are controlled within the system by means of the operator?
AFR Interaction Level of the Robot [14]	Type A–Type D	Which interaction level fits the robot, involved the system, at most?
Causality of Modalities [4]	simulated, recorded, transmitted	Where do modalities (perceptible sensations) have their origin.
Control Alignment [9]	1-to-1, positional offset, rotational offset, pose offset	How are input means from in the real world connected to a linked virtual world?
Control Order [9]	Order of Mapping Function (0–n)	What is the complexity of the relationship between user input variables and controlled variables?
Directness [9]	degree of mapping to display space/isomorphism	What is the degree of spatial matching between display (not only visual) content and linked modalities in the real world?
Directness of Sensation [9]	directly-perceived-real-world – scanned-and-resynthesized	How is the real world perceived through technical devices?
Display Centricity [7,9]	egocentric–exocentric	What is the spatial relationship between contents of a display (not only visual) and the user?
Display Class (Milgram) [7]	class 1–class 7	How is the operators display classified according to Milgram?
Display Space [4]	registered, remote, miniaturized/enlarged, distorted	How are different spaces aligned (virtual, real, ...)?
Display Type [4]	HMD, screen, speaker, ...	What kinds of displays are utilized?
DOF (Degrees of Freedom)—parameters manipulated [13]	0–n	This is about the number of values altered during interaction, not the robot structure.
DOF (Degrees of Freedom)—input device based [13]	0–n	This is about the number of values assessed from means of input.
Extensiveness [6]	single modality–multi-modal	How many different modalities are presented to the operator by the system?

Table A1. Cont.

Factor	Values	Description
Extent of Body Matching [6]	no correlation–perfect matching	In case of the existence of a virtual body: How strong is the virtual body matching with corresponding real body?
Extent of Proprioceptive Matching [6]	no correlation–perfect matching	In case of the existence of a virtual body: How strong can the proprioceptive matching be?
Extent of Presence Metaphor (EPM) [7,8]	WoW/monoscopic imaging vs. HMD/real-time imaging	How strong is the display technology likely to induce sense of presence. Or better: How immersive is the technology?
Extent of World Knowledge (EWK)/where and what [7,8]	unmodeled–modeled	How much of the real world is known in detail?
Gesture Type [11]	static posture, static oriented posture, dynamic/moving gesture, moving oriented gesture	If so, what category of gesture is used?
Human-Robot Communication [27]	single dialog–two (or more) monologues	Is the communication between human and robot based on equal rights?
Image Scaling [9]	(1:1–1:k)/ (orthoscopic/conformal mapping vs. scaled images)	How is the size of rendered modalities in comparison to real entities?
Inclusiveness [6]	none–complete	To which degree are you shut off from the modalities of the real world around you?
Initiative [31]	fixed–mixed	How is the initiative of acting or communicating between human, robot and the rest of the system? Fixed at a single actor or distributed between several partners?
Integration of VE [12]	separated–integrated	Spatio-temporal relationship between real environment and virtual environment.
Intelligent Robot Control Paradigm [15]	hierarchical, reactive, hybrid	How is the intelligent behavior of the robot organized?
Interaction Reality <sup>1</sup>	physical–virtual	Is the interaction mainly with real or virtual objects?
Interaction Type [10]	Selection and Manipulation, Travel, System Control	Which of Bowman’s 3D interaction category is mainly present?
JIRA Intelligence Level of the Robot [14]	Class 1–Class 6	Which class of intelligence fits the robot, involved in the system, at most?
Level of Capabilities (cf. [24])	System skills are not enough to solve/act-out any task–system is capable of acting-out at least any task perfectly	How (potentially) skilled is the system independent from the operators skills?

Table A1. Cont.

Factor	Values	Description
Level of Intelligence (cf. [24])	System never generates a theoretic solution towards a goal–system always finds optimal solution	How intelligent is the system without the operator?
Level of Mixed Reality [7]	RE–AR–AV–VE	Where is the current system positioned on Milgram’s Reality-Virtuality Continuum?
Level of Robot Anthropomorphism [17]	none–non-distinguishable from human	How human-like is the robot?
Location <sup>2</sup>	in locu vs. remote	What is the mental distance between the virtual and the real world components?
Mapping Relationship [12]	position–rate	How input is mapped to output mathematically.
Model of Modality-Source [4]	scanned, constructed, computed, edited	How are modalities (perceptible sensations) generated?
Parameter Manipulation Type [13]	direct–indirect	In which way is a value changed.
Plot [6]	none–perfect integration into mental models	Is there a plot and how good does it fit the situation?
Real-time Level [14–16,32]	no delay–delayed	How close to real time is the system behavior?
Refined LOA in steps acquisition, analysis, decision and action (or sense, plan and act) [18,19,21]	1–10	Which level is the autonomy of the system in the single steps of executing a task: “analysis”, “decision”, “action”?
Reproduction Fidelity (RF) [7,8]	monoscopic video/wireframes vs. 3D HDTV/real-time high-fidelity	How is the quality of the presented modalities to be rated?
Robot Behavior Level [17]	remote controlled–autonomous	How can the behavior of the robot be rated? Is it doing something on its own? Does it have some software induced abilities?
Robot Communication Capability [17]	reactive–dialog	How is the robot communicating with the operator? Only on demand, or does it have needs to communicate?
Robot Control Concept [14,15]	open loop, feedback control loop, feedforward control, adaptive control	How is the robot controlled in general?
Robot Purpose [17]	toy–tool	What is the typical purpose of the current robot?
Robot Type [14–16,32]	Industrial, Mobile, Service, Social, ...	What kind of robot is used in terms of common sense?
Role of Interaction [33]	supervisor, operator, mechanic, bystander, team mate, manager/leader	Which social role inherits the operator? (Scholtz)

Table A1. Cont.

Factor	Values	Description
RUI-Type [14–17,32]	programming, command-line, real-time, interactive	How is the robot itself mostly interfaced?
Sensing Mode [12]	isotonic–elastic–isometric	Mapping from signal source to sensor reaction.
Sensor Type [4]	Photomultiplier, STM, ultrasonic scanner, ...	What kind of sensors are used in the system for generating information for the operator?
Structure of Interaction [2]	[SO, SR, ST], [SO, SR, MT], ..., [MO, MR, MT] (S = single, M = multiple, O = operator, R = robot, T = Tasks)	What is the typical combination in the system?
Superposition of Modalities [4]	merged, isolated	How are technologically mediated modalities presented in comparison to modalities which are part of the real environment?
Surrounding [6]	small 2d image (WoW)–large 2d–curved–cylindrical–spherical	How much is the display everywhere around you?
System Extent <sup>3</sup>	self-contained / all-in-one–network-based across planets	How much is the system, especially the user interface distributed among different devices?
Time (interaction) [4]	1-to-1, accelerated/retarded, frozen, distorted	How is the interaction time related to execution time?
Time (purpose) [14–16,32]	real time–recorded for later	How much time passes between the user input and the manipulation of the real world by the robot by intention.
Timely Aspects of Actors' Autonomy [27]	static–dynamic	Is the level of autonomy regarding different aspects of the system changing over time (depending on some other factors) or is it static and defined by design?
Transparency of System State [34–36]	invisible–all states are displayed	Does the operator know the system state during operation?
UI-Type [10,37]	Batch, CLI, GUI, NUI, SUI, ...	What stereotype of user interface describes the present system at most?
User Experience [4]	recorded, transmitted, simulated, supervised robot	How originate the modalities the user is experiencing?
Vividness [6]	very low–very high	How good is the rendering quality?

<sup>1</sup> This factor was added by the authors. After reading many publications about interactive virtual environments (IVE), it was inevitable to add this. <sup>2</sup> This factor was added by the authors. The spatial distance between the virtual and the real environment was only implicitly part of the discussions in the literature. <sup>3</sup> Added by the authors. The general system design, based on physical components, was not found explicitly in the literature of mixed reality and robots.



**Table A2.** Summary of Section 4 using the IMPAct template.

INTERACTION	Interaction Parameter	Action Type AFR Interaction Level Control Alignment Control Order Directness DoF Interaction Reality Interaction Type Mapping Relationship Parameter Manipulation Type Sensing Mode Structure of Interaction	discrete Type D 1-to-1 0 almost isomorph 5 virtual selection position direct isometric SO,SR,ST
	Paradigm	Human-Robot Communication Intelligent Robot Control Paradigm Robot Control Concept Role of Interaction RUI-Type UI-Type	two monologues hierarchical at user level feedback control loop supervisor & manager interactive SUI
MEDIATION	Image Display Vision	Display Centricity Display Class (Milgram) (Display) Space Image Scaling Vividness	egocentric class 3 registered 1:1 mid-to-low
	Input	Action Measurement Type Directness of Sensation DoF Gesture Type Sensor Type	HoloLens inside-out tracking partially directly viewed 3 + 1 static posture IMU, structured light IR camera, bw camera, joint encoder, current measurement
	Output	Actuator Type Display Type	robot arm & gripper see-through HMD
PERCEPTION	Embodiment	Extent of Body Matching Extent of Proprioceptive Matching	no VB virtual cursor and viewing direction
	Immersion	Extend of Presence Metaphor Integration of VE Level of MR Reproduction Fidelity	strong tendency towards real-time imaging almost completely integrated AR with tendency to RE stereoscopic realtime rendering but with simple shaders
	Modalities	Causality of Modalities Extensiveness Inclusiveness Model of Modality-Source Superposition of Modalities User Experience	partially transmitted and simulated only vision almost none computed merged mixture of transmitted, simulated and supervised robot
	Space & Time	Extent of World Knowledge Location Realtime Level Surrounding System Extent Time (interaction) Time (purpose)	objects of interest are modeled with normal to high accuracy in locu delayed up to half a second only 60° FoV network-based in one room frozen realtime
ACTING	Autonomy	Level of Capabilities Level of Intelligence Plot Refined LOA Initiative JIRA Intelligence Level of the Robot Timely Aspects	generate trajectories, self localize collision avoidance, extract grasp points virtual robot shows probable solution acquisition: 10; analysis: 3; decision: 1-2; action: 5 fixed at the human Class 5 static
	Behavior	Robot Behavior Level Robot Communication Capability Transparency of System State	remote controlled reactive visible
	RA	Level of Robot Anthropomorphism	very low
	Application	Robot Purpose Robot Type	tool industrial

**Table A3.** Empty IMPAct template.

INTERACTION	Interaction Parameter	Action Type AFR Interaction Level Control Alignment Control Order Directness DoF Interaction Reality Interaction Type Mapping Relationship Parameter Manipulation Type Sensing Mode Structure of Interaction
	Paradigm	Human-Robot Communication Intelligent Robot Control Paradigm Robot Control Concept Role of Interaction RUI-Type UI-Type
MEDIATION	Image Display Vision	Display Centricity Display Class (Milgram) (Display) Space Image Scaling Vividness
	Input	Action Measurement Type Directness of Sensation DoF Gesture Type Sensor Type
	Output	Actuator Type Display Type
PERCEPTION	Embodiment	Extent of Body Matching Extent of Proprioceptive Matching
	Immersion	Extend of Presence Metaphor Integration of VE Level of MR Reproduction Fidelity
	Modalities	Causality of Modalities Extensiveness Inclusiveness Model of Modality-Source Superposition of Modalities User Experience
	Space & Time	Extent of World Knowledge Location Realtime Level Surrounding System Extent Time (interaction) Time (purpose)
ACTING	Autonomy	Level of Capabilities Level of Intelligence Plot Refined LOA Initiative JIRA Intelligence Level of the Robot Timely Aspects
	Behavior	Robot Behavior Level Robot Communication Capability Transparency of System State
	RA	Level of Robot Anthropomorphism
	Application	Robot Purpose Robot Type

## References

1. Burdea, G.C. Invited review: The synergy between virtual reality and robotics. *IEEE Trans. Robot. Autom.* **1999**, *15*, 400–410. [\[CrossRef\]](#)
2. De Barros, P.G.; Lindeman, R.W. *A Survey of User Interfaces for Robot Teleoperation*; Computer Science Department, Worcester Polytechnic Institute: Worcester, MA, USA, 2009.
3. Crandall, J.W.; Goodrich, M.A.; Olsen, D.R.; Nielsen, C.W. Validating human–robot interaction schemes in multitasking environments. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2005**, *35*, 438–449. [\[CrossRef\]](#)
4. Robinett, W. Synthetic Experience: A Proposed Taxonomy. *Presence Teleoper. Virtual Environ.* **1992**, *1*, 229–247. [\[CrossRef\]](#)
5. Zeltzer, D. Autonomy, Interaction, and Presence. *Presence Teleoper. Virtual Environ.* **1992**, *1*, 127–132. [\[CrossRef\]](#)
6. Slater, M.; Wilbur, S. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence Teleoper. Virtual Environ.* **1997**, *6*, 603–616. [\[CrossRef\]](#)
7. Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. Augmented reality: A class of displays on the reality–virtuality continuum. *Proc. SPIE* **1995**, *2351*, 282–293.
8. Milgram, P.; Kishino, F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Inf. Syst.* **1994**, *77*, 1321–1329.
9. Milgram, P.; Colquhoun, H. A taxonomy of real and virtual world display integration. In *Mixed Reality: Merging Real and Virtual Worlds*; Springer-Verlag: Berlin/Heidelberg, Germany, 1999; Volume 1, pp. 1–26.
10. Bowman, D.; Kruijff, E.; LaViola, J.J., Jr.; Poupyrev, I.P. *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*; Addison-Wesley: Boston, MA, USA, 2004.
11. Sturman, D.J.; Zeltzer, D.; Pieper, S. Hands-on Interaction with Virtual Environments. In Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology, Williamsburg, VA, USA, 13–15 November 1989; ACM: New York, NY, USA, 1989; pp. 19–24. [\[CrossRef\]](#)
12. Zhai, S.; Milgram, P. Input Techniques for HCI in 3D Environments. In Proceedings of the Conference Companion on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994; ACM: New York, NY, USA, 1994; pp. 85–86. [\[CrossRef\]](#)
13. Lindeman, R.W. Bimanual Interaction, Passive-Haptic Feedback, 3D Widget Representation, and Simulated Surface Constraints for Interaction in Immersive Virtual Environments. Ph.D. Thesis, George Washington University, Washington, DC, USA, 1999.
14. Niku, S.B. *Introduction to Robotics: Analysis, Control, Applications*; Wiley India: Noida, India, 2011.
15. Mckerrow, P. *Introduction to Robotics*; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1991.
16. Murphy, R.R.; Arkin, R.C. *Introduction to AI Robotics*; MIT Press: Cambridge, MA, USA, 2000.
17. Bartneck, C.; Okada, M. Robotic user interfaces. In Proceedings of the Human and Computer Conference, Aizu, Japan, 5–10 August 2001; pp. 130–140.
18. Sheridan, T.B.; Verplank, W.L. *Human and Computer Control of Undersea Teleoperators*; Technical Report; Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab; Defense Technical Information Center: Fort Belvoir, VA, USA, 1978.
19. Parasuraman, R.; Sheridan, T.B.; Wickens, C.D. A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2000**, *30*, 286–297. [\[CrossRef\]](#)
20. Crandall, J.W.; Goodrich, M.A. Experiments in adjustable autonomy. In Proceedings of the 2001 IEEE International Conference on Systems, Man and Cybernetics, e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236), Tucson, AZ, USA, 7–10 October 2001; Volume 3, pp. 1624–1629. [\[CrossRef\]](#)
21. Miller, C.A.; Parasuraman, R. Beyond levels of automation: An architecture for more flexible human-automation collaboration. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* **2003**, *47*, 182–186. [\[CrossRef\]](#)
22. Beer, J.M.; Fisk, A.D.; Rogers, W.A. Toward a Framework for Levels of Robot Autonomy in Human-robot Interaction. *J. Hum. Robot Interact.* **2014**, *3*, 74–99. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Schilling, M.; Kopp, S.; Wachsmuth, S.; Wrede, B.; Ritter, H.; Brox, T.; Nebel, B.; Burgard, W. Towards a multidimensional perspective on shared autonomy. In Proceedings of the 2016 AAAI Fall Symposium Series, Arlington, VA, USA, 17–19 November 2016.

24. Gunderson, J.; Gunderson, L. Intelligence = autonomy = capability. In Proceedings of the Performance Metrics for Intelligent Systems, PERMIS, National Inst of Standards and Technology Gaithersburg MD, USA, 24–26 August 2004.
25. Bainbridge, L. Ironies of automation. In *Analysis, Design and Evaluation of Man–Machine Systems*; Elsevier: Amsterdam, The Netherlands, 1983; pp. 129–135.
26. Dekker, S.W.A.; Woods, D.D. MABA-MABA or Abracadabra? Progress on Human–Automation Co-ordination. *Cogn. Technol. Work* **2002**, *4*, 240–244. [[CrossRef](#)]
27. Calefato, C.; Montanari, R.; Tesauri, F. The adaptive automation design. In *Human Computer Interaction: New Developments*; InTech: London, UK, 2008.
28. Scholtz, J. Theory and evaluation of human robot interactions. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Big Island, HI, USA, 6–9 January 2003.
29. Adamides, G.; Christou, G.; Katsanos, C.; Xenos, M.; Hadzilacos, T. Usability Guidelines for the Design of Robot Teleoperation: A Taxonomy. *IEEE Trans. Hum. Mach. Syst.* **2015**, *45*, 256–262. [[CrossRef](#)]
30. Krupke, D.; Steinicke, F.; Lubos, P.; Jonetzko, Y.; Görner, M.; Zhang, J. Comparison of Multimodal Heading and Pointing Gestures for Co-Located Mixed Reality Human-Robot Interaction. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 1–9. [[CrossRef](#)]
31. Allen, J.; Guinn, C.; Horvitz, E. Mixed-initiative interaction. *IEEE Intell. Syst. Their Appl.* **1999**, *14*, 14–23. [[CrossRef](#)]
32. Craig, J.J. *Introduction to robotics: mechanics and control*; Pearson/Prentice Hall: Upper Saddle River, NJ, USA, 2005; Volume 3.
33. Scholtz, J.; Young, J.; Drury, J.L.; Yanco, H.A. Evaluation of human–robot interaction awareness in search and rescue. In Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, 26 April–1 May 2004; Volume 3, pp. 2327–2332. [[CrossRef](#)]
34. Keyes, B.; Micire, M.; Drury, J.L.; Yanco, H.A. Improving human–robot interaction through interface evolution. In *Human-Robot Interaction*; IntechOpen: London, UK, 2010.
35. Clarkson, E.; Arkin, R.C. Applying Heuristic Evaluation to Human-Robot Interaction Systems. In Proceedings of the Flairs Conference, Key West, FL, USA, 7–9 May 2007; pp. 44–49.
36. Elara, M.R.; Acosta Calderon, C.; Zhou, C.; Yue, P.K.; Hu, L. Using heuristic evaluation for human-humanoid robot interaction in the soccer robotics domain. In Proceedings of the IEEE-RAS International Conference on Humanoid Robots (Humanoids 2007), Pittsburgh, PA, USA, 29 November–1 December 2007.
37. Jacko, J.A. *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*; CRC Press: Boca Raton, FL, USA, 2012.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).