Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 54 (2021) 197–202

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

10th CIRP Sponsored Conference on Digital Enterprise Technologies (DET 2021) – Digital Technologies as Enablers of Industrial Competitiveness and Sustainability

# Multi-modal interfaces for natural Human-Robot Interaction

Dionisis Andronas[a], George Apostolopoulos[a], Nikos Fourtakas[a], Sotiris Makris[a],*

[a]*Laboratory for Manufacturing Systems and Automation, Department of Mechanical Engineering and Aeronautics, University of Patras, Patras 26504, Greece*
* Corresponding author. Tel.: +30-261-091-0160; fax: +30-261-099-7314. E-mail address: makris@lms.mech.upatras.gr

## Abstract

Recent years, manufacturing aims on increasing flexibility while maintaining productivity for satisfying emerging market needs for higher product customization. Human Robot Collaboration (HRC) is able to bring about this balance by combining the benefits of manual assembly and robotic automation. When introducing a hybrid concept, safety and human acceptance are of vital importance for achieving implementation. Fenceless coexistence may lead to discomfort of operators especially in cases where close Human Robot Interaction (HRI) occurs. This work aims at designing and implementing a natural Human-System and System-Human interaction framework that enables seamless interaction between operators and their "robot colleagues". This natural interaction will strengthen hybrid implementation through increased: a) operator's and system's awareness, b) operator's trust to the system, and through the decrease of: a) human errors and b) safety incidents. The overall architecture of the proposed system makes it scalable, flexible, and applicable in different collaborative scenarios by enabling the connectivity of multiple interfaces with customizable environments according to operator's needs. The performance of the system is evaluated on a scenario originating from the automotive industry proving that an intuitive interaction framework can increase acceptance and performance of both robots and operators.

*Keywords:* Human Robot Collaboration; Human Robot Interaction; Interface; Augmented Reality; Robotic cells

## 1. Main text

In recent years, industry aims at increasing productivity and flexibility of the production system, for keeping up with modern market requirements [1]. Harsh competitiveness and smaller product lifecycles stress existing manufacturing systems forcing businesses to find alternative production schemes for adapting to production changes [2]. Human-Robot Collaboration (HRC) has offered a number of advantages in manufacturing by combining the intelligence of human operators with the tireless accuracy and repeatability of robots [3]. The primary target of HRC is to increase flexibility and reduce operator physical strain by employing robots as coworkers without physical barriers. In such a setup numerous interaction may between the operator and the robot may be encountered in a plethora of occurrences. Therefore, it is required to establish a controllable and efficient communication between those "colleagues" for synchronizing their actions and establishing awareness on both sides. For industrial implementation of hybrid systems, operator's safety and sense of safety will always be the baseline for achieving acceptance. So far, these attributes are achieved by the use of certified safety systems that control hazardous situations, however imposing limitations to human actions would lead to reduced flexibility which contradicts the main objective of HRC. Moreover, safety systems or traditional Human Machines Interfaces (HMIs) are not anthropocentric, thus they do not provide intuitively information to the operators. This impacts the performance of the robotic cell due to stochastic interruptions that lead to frequent stops or operating robots at a reduced speed. Reducing

10.1016/j.promfg.2021.07.030

the number of those interruptions can be achieved by providing more natural communication and cognition of the robotic “colleague’s” intentions so that a smoother interaction can be achieved.

The aim of this paper is the introduction of a set of multimodal interfaces for Human-Robot and Robot-Human (HR&RH) communication. The proposed interfaces provide a two-way communication between the robot and the operator. The system’s approach is founded on most important human factors (i.e. task switching”, “stress quantification”, “concentration or sustained attention” and “situation awareness”) as they have been classified [4]. The main focus is on maximizing operator’s sense of comfort, productivity and acceptance through state-of-the-art HMIs, interconnected on a reconfigurable platform. The following sections are organized as follows: Section 2 discusses about current state of the art on HMIs. Section 3 provides the description of the approach that was used and analyses the key elements that a hybrid workstation communication interface needs to consider. The implementation of the proposed interface is presented in Section 4, whereas an industrial case study, that is used for assessing the performance of the system is analyzed in Section 5. Finally, the paper concludes with results and future steps based on the assessment’s findings.

## 2. Human System Interaction – a literature view

HR&RH interfaces has been an area of growing interest for the last decade. Improvements in bilateral communication of users and robots or machines have created new horizons in systems’ capabilities and performance. Overall, interfaces allow operators to control resources or perform operation functions. Those actions are supported by graphical representation of the process being controlled, sensor values and output states [5].

In the context of HRC, different approaches on developing interfaces have been implemented in an effort to achieve seamless human-system communication. However, despite the advances in computing and hardware performance, there are still limitations on managing complex interactions and behaviors [6]. This directs to solutions that utilize additional data or sensors to overcome handicaps. A systematic approach on how multiple communication modals can be used depending the operation context has been proposed [7]. Other frameworks like [8] focus on the extensibility of multi-modal interfaces by adding input modalities or changing classification strategies.

Studies have also elaborated on evaluating the impact of common modalities on mental workload. Results indicate that there is a preference on verbal communication, however additional studies on different scenarios are obligatory for strict conclusions [9]. Driven by industrial requirements, easiness and intuitiveness of deployed interfaces is crucial, as operators need time for their adoption. Adoption can be supported by implementing solutions that do not occupy operator’s arms and affecting productivity. In this regard, Augmented Reality (AR) headset-based solutions tend to become more popular as HMIs [10]. Last but not least, other methods indicate different means of interaction based on gestures [11] or vocal commands [12], [13] for effortless control of the system or robots. Finally, as wearable devices become popular in everyday applications, a series of applications are based on smartphones, tablets, smartwatches etc. [14].

## 3. Approach for Natural Human Robot and Robot Human Interaction

The work of this manuscript focuses on the design and development of a multi-modal interface that: a) shares valuable information regarding the manufacturing process and robot actions to the operator, and b) allows operator to effectively interact with the robot without compromising manual operation productivity. The exchanged information strengthens operator’s trust to the system, sense of comfort and confidence as well as contributes to the reduction of errors. Given the improvisation and decision making of human operators, intuitive stream of information is rather important. The system needs to be capable of providing useful information to operators allowing them to make the right decisions (i.e. perform correct task), in the right way (i.e. using correct parts, procedures and tools) and the safest context (i.e. avoiding collision with robot). On the other hand, the designed interface can be a valuable tool for monitoring operator’s actions. The same monitoring modules can also be used for translating operators’ commands into executable actions by the robot in cases of direct control. The schematic of Fig. 1 illustrates a common hybrid production interaction stream.

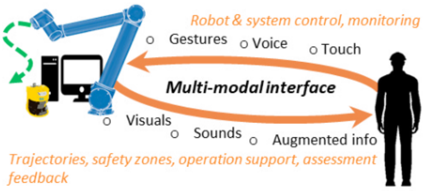


Fig. 1. Human System and System Human Interaction

For implementing each mode of interaction, at least one of the front-end system devices requires features that exploit human’s senses (e.g. vision, hearing, etc.) and interaction means (e.g. voice, gestures, etc.) according to the intended functionalities.

### *3.1. Overall architecture*

The proposed multi-modal interface architecture aims on achieving flexibility and scalability as far as the implemented devices are concerned. New devices can be added without degrading the overall system’s efficiency and performance. This scheme allows operators to select the most appropriate front-end devices according to their needs or device autonomy limitations. Once a mode of interaction or a device is unavailable, smooth operability is ensured by the alternative modes of interaction offered by other connected devices.

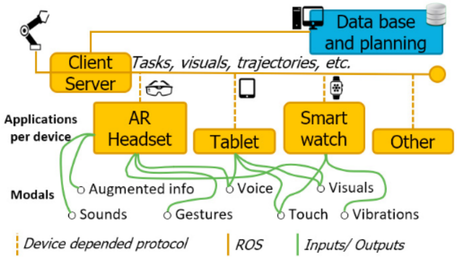


Fig. 2. Multi-modal interface overall architecture

Referring to the architecture (Fig. 2), all devices are connected to a main client-server. This module acts like a parser and it is responsible for transferring system's information from the data base to the end-devices and vice versa. For HR interaction (Fig. 3), hosted applications translate inputs in valuable information and communicate it to the client server for further uses. For RH interaction (Fig. 4), information is streamed to all connected front-end devices, and hosted applications are responsible for its communication to the users through the appropriate modals.

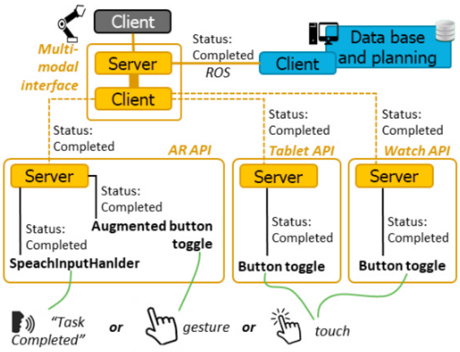


Fig. 3. HR Communication (e.g. feedback that manual task is completed)

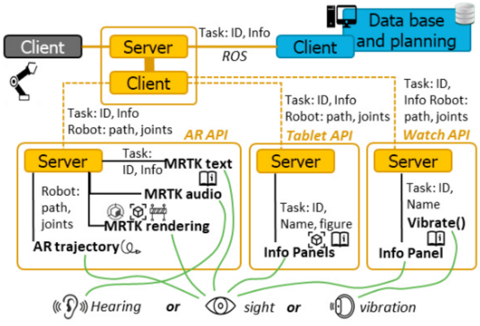


Fig. 4. RH Communication (e.g. task information, robot trajectories)

In the proposed work, a distinct flexible architecture is formed where numerous devices can be connected to the main client-server through the most suitable communication protocols (e.g. ROS, TCP/IP, WebSocket, etc.), without affecting the communication between other devices. This framework creates a scalable system where more interaction modes can be introduced in accordance with the implemented devices' specifications. In case modals requiring large computational resources (i.e. beyond devices' capabilities) are selected, then processing is delegated to computers whereas end-devices are used for fused outputs.

### 3.2. Personalization through customizable interfaces

Each person is a unique entity with own characteristics preferences and priorities. Therefore, it seems more than necessary that operators should be able to handle interfaces and select the amount and the content of the information they receive. By creating individual customization profiles, every user can build the most suitable working environment. Thus, they are not getting slowed down or misled during operation thus maintaining production rate at a high level. The type of information that is required per operator depends on working experience, trust to the system and assembly operation complexity. Referring to the schematic of Fig. 5, covering a wide range of operators in a wide range of scenarios requires multiple information schemes and subsequently alternative devices able to communicate such information. Personalization of interfaces through profile customization supports the usage of same applications and same set of modes for different operators and different manufacturing scenarios. Customization options consist of parameters regarding the level of detail of streamed information, positioning of notification panels as well as activation of visualization features. Those parameters are stored in the system's main data base and are retrieved upon user's successful login. The same login process also promotes system's security by excluding unauthorized personnel from system control and data communication.

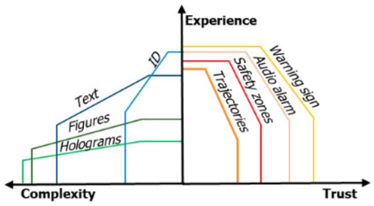


Fig. 5. Type of information required based on operator's experience, trust to the system, and assembly operation(s) complexity

### 3.3. Operator support with intuitive instructions

Fluctuations in KPIs and product quality are caused by wrong task execution or non-proper assembly operations. The probability of such event increases with process complexity, frequency of production changes and inexperience of operators. Moreover, parallel task execution with an expressionless robot colleague may provoke misinterpretations due to current state unawareness. The proposed multi-modal interface addresses those issues through intuitive presentation of executed tasks. This presentation consists of executed task identification on every front-end device. Depending the connected device's capabilities, notifications involve task naming, assembly instructions via plain text, 2D informative visuals, 3D holograms or even physical object tracking.

### 3.4. Safety through robot-centric information

Novel production systems are able to react on dynamic production or environment events. For robots, adaptability can be achieved through autonomous behavior in terms of motions, tool changing, etc. Planning a collision-free trajectory is fundamental for autonomous robots. Many researches nowadays have shown that dynamic robot trajectories have significant impact on physiological parameters [15]. Fenceless coexistence with a robot, that can update its routine, can generate trust and fear issues to operators. The proposed interface depicts the planned trajectories allowing users to understand robot actions. The coordinates generated by path planning modules are streamed to the multimodal interface that presents them through lines. Awareness of operator is also increased through notification panels with warning signs, sound alerts or vibrations signaling robot movements. The same objective is also served through communication of safety related information. In many manufacturing systems, safeguard is achieved by optoelectronic devices. Given the non-visible nature of safety zones their positioning and dimensioning in the collaborative area may be unknow by operators. The proposed interface visualizes augmented safety zones with classical green, yellow, red (GYR) zone configurations so as operator does not proceed to unnecessary infringements.

### 3.5. Human System Interaction through novel means

Monitoring of operators is quite a challenge due to their improvisation and variety of actions when executing a task. Despite advances in deep learning for human task identification, the most stable method is keeping operator in the loop. Thus, the proposed interface provides a number of means for controlling resources or providing feedback to the system. Devices that support gesture recognition, body segment tracking, voice recognition, acceleration and force monitoring, or touch inputs are used for enabling “task completed” feedback or even “start”, “pause” or “stop” of processes or robots. The main advantage of such interaction means lays on the easiness of their usage by giving feedback without pausing manual operations or loosing environment alertness.

### 3.6. Ergonomics improvement through real-time feedback

Poor ergonomics can be a factor of fatigue which affects performance and product quality. When ergonomic evaluation through perception data is available, having a system able to share valuable information to the operator is an opportunity for ergonomics optimization. Through the devices, operators are communicated live ergonomic evaluation feedback that helps them have a thorough understanding of actions that can cause musculoskeletal issues. Ergonomic evaluation results are transmitted to users via: a) GYR light (with colors indicating the severity of postures), b) detailed human skeleton with GYR colored segments, and c) video assistance. To finish, ergonomic results are stored and can be accessed after the end of the shift offering operators a chance for elaborating further.

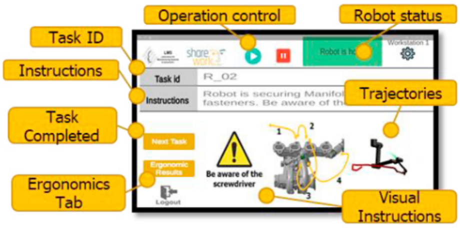
## 4. Implementation

This section describes the implementation of the multi-modal interface and it is structured in four subsections for facilitating reader. Each one analyses one key element of the implemented system presenting how modals were used.

### 4.1. Client-server for information flow

The proposed scalable and multi-protocol architecture allows plug-n-play connectivity of the devices (Section 3.1). The client-server’s communication with the rest of the hybrid’s system modules is established through ROS. Without compromising the generality of the system, ROS was also used for the communication of messages to/from the end-devices. Data is exchanged through topics, actions and services depending their complexity and necessity of feedback.

### 4.2. Tablet application



Task id R_02

Instructions Robot is securing Manifol fasteners. Be aware of th

Fig. 6. Tablet interface operation scene with all features enabled

On the basis that traditionally HMIs were classic stationary workstation screens, an application was designed and implemented on android tablet. Subsequently, either as a fixed screen or as a portable device, tablet gives access on valuable information besides operation control. Screen does not support spatial visualization, however it gives a very straightforward environment for interacting with. Inputs as well as operator reporting actions via touch are feasible (i.e. for “task completed”). The interface, developed in Unity, consists of scenes with a) task information fields (id, name, description), b) assembly instructions through visuals, c) robot trajectory depiction and d) ergonomic results (through body segment coloring or traffic light). The level and the type of visualized information is adjustable through customization options. As tablets have touch-based keyboards and present great autonomy, easy recharge and portability, additional functionalities are also supported. For example, authorization process (logging or register) is swiftly performed through such devices. For implementing the interface depicted on Fig. 6, a market available Samsung Galaxy Tab 4 tablet was used.

### 4.3. Augmented Reality application

The implemented AR application overlays information and digital content on the physical world for intuitive guidance. Like aforementioned application, the AR interface communicates information for operation support, safety and robot actions but it is also used for system control (Fig. 7). During operation, inside operator's point of view, augmented content can include: a) task information fields, b) 3D holograms visualizing how parts should be moved for every assembly step c) real time visualization of the robot trajectories, d) ergonomics feedback (GYR light or manikin segments), e) safety zones and f) augmented information on tracked physical objects. Moreover, audio notifications are also available for robot warnings. Object recognition is used for tracking objects at the workstation and highlighting them during a related task. As for robot trajectories, a line, that is formed by prerecorded operations or real-time generated waypoints from the motion planner (e.g. MoveIt), is erased as robot performs movements by echoing current position from robot's controller.

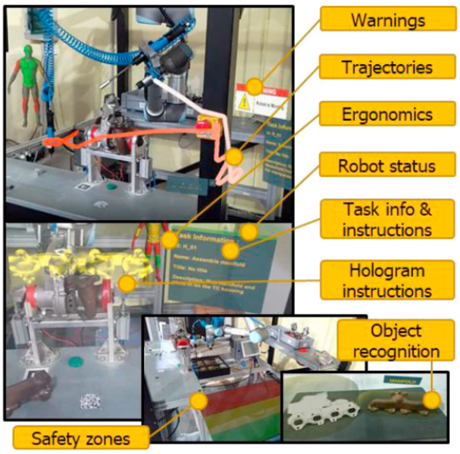


Fig. 7. Augmented reality application with all features

For human robot interaction, the selected device (Microsoft HoloLens2) supports a number of modals including gestures, voice recognition, hearing and sight through visuals. The operator can interact with voice commands or gestures on augmented buttons and objects. In aspects of customization, the operator is able personalize the interface by enabling or disabling features. Moreover, the panels can be grasped, dragged and dropped at positions that increase comfort. All options reduce the risk of implementing a confusing interface that disrupts operators, blocks their point of view or reduces their awareness through excessive information flow.

### *4.4. Smartwatch application*

Fig. 8. Smartwatch interface

Aiming of providing operator fundamental information about the process, that is available continuously without bulky equipment, a smartwatch application on LG W200E was designed and developed (Fig. 8). Due to the limitations of the screen area, the application only broadcasts the task's ID and name whereas a "task completed" button is also available for monitoring.

## 5. Case Study

The proposed multi-modal interface has been implemented on a case study deriving from the automotive industry. More specifically the performance of the system was evaluated on a hybrid workstation were a human operator and a robot collaborate for assembling a powertrain component. The assembly operation contains a sequence of handling and fastening tasks. As extensively described in [16], the robot arm is responsible for manipulating heavy objects or performing fastening operations with high repeatability. Contrarily, assembly operations requiring dexterity and improvisation are assigned to the operator. The execution of both parallel and sequential tasks, in a fenceless environment, makes the case a fertile ground for testing in depth the capabilities and features of the created multi-modal interface.

Fig. 9. Multi-modal interface overall architecture

The applications, as presented in Section 4, were deployed to the front-end devices of the multimodal system. Tasks, operator profiles, instructions etc. were communicated by the hybrid's system data base and planning system whereas the visualization of robot trajectories was correlated to the robot controller and path planning modules. After authorization, the operator was able to control the system and interact with the robot via gestures, touch or voice commands and receive visual or audio information. Each operator that tested the interface was able to customize it for meeting personal needs.

## 6. Results

A series of experiments involving multiple assembly cycles, by operators was performed. The objective of those tests was to estimate the performance of the interface, identify handicaps and verify if natural communication was achieved. The results indicated that the proposed system is flexible and robust. This was mainly related to the customization features and the connectivity of various devices. AR headset, smartwatch and tablet presented optimum performance and usability. Interaction through touch, gestures and voice commands (at low to medium noise environment) was rather helpful given the limited effort required. Focusing on operator support, task instructions were very intuitive and helpful, as users were able to get familiar with the assembly process quite easily. However, the object recognition through AR headset was not fully stable and stressed hardware's resources. Regarding robot related information, the visualization of trajectories and safety zones resulted in increased awareness and zone intrusions were limited.

Table 1. SUS Analysis Results

| User | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Score | 85 | 82.5 | 72.5 | 77.5 | 75 | 72.5 | 75 | 77.5 | 72.5 |
| User | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| Score | 82.5 | 70 | 85 | 70 | 80 | 77.5 | 82.5 | 90 | 82.5 |
| Total Score = 78.33 | | | | | | | | | |

Questionnaires, based on System Usability Scale [17], were provided to a sample of 18 users. Having overall average score over 78% (Table 1), the proposed multimodal interface proves to be a valuable tool that participants would like to use in an industrial environment. Moreover, users expressed that: a) personalization of the system is very convenient, b) using of the system does not require advanced skills, c) it is very easy to get familiar with controls and information and d) the multi-modal interface helped them in aspects of operation training, through error elimination as well as acclimatization with robots at fenceless environments.

## 7. Conclusions and Future Steps

The proposed interaction framework proved to be an effective tool for operators. The flexible architecture allowing rapid device connectivity and robust information stream. Connecting or excluding devices does not affect system's performance instead it allows users to customize the system upon their needs. Operator acceptance is increased through customization of each interface's environment and availability of multiple modals for same purposes. The autonomy and fortitude of the system against device failures, that could lead to downtimes, is strengthened by its flexible and scalable architecture. The implemented devices, namely: AR headset, tablet and smartwatch, supported high quality interaction with the robot by using natural means of communication.

Future work involves the development of additional front-end applications or optimization of the existing ones as well as their connection to the overall multi-modal interface. In addition, object detection using the wearable headset will be optimized for smoother and more reliable augmented information on physical objects. Moreover, the flexibility and the scalability of the proposed HR&RH interaction framework will be demonstrated in alternative industrial scenarios highlighting additional handicaps that need to be addressed. Last but not least, a more extensive methodology for assessment of physiological aspects besides acceptance will be implemented. Finally, this work acts as a baseline for future interface implementations that aim to bring operators closer to robots in industrial environments through non ad hoc solutions. The benefits of adopting such interfaces in manufacturing start from improvement of KPIs, in terms of values and consistency, and end-up to huge financial business benefits.

## Acknowledgements

This work has been partially funded by the EC research project "SHAREWORK - Safe and effective HumAn-Robot coopEration toWards a better cOmpetiveness on current automation lacK manufacturing processes" (Grant Agreement: 820807).

## References

[1] Chryssolouris G. Manufacturing Systems: Theory and Practice. Springer. 2006.
[2] Makris S. Cooperating Robots for Flexible Manufacturing. Springer International Publishing: Cham 2021.
[3] Khojasteh Y. Production management : advanced tools, models, and applications for pull systems. 2017.
[4] Paletta L., Pszeida M., Ganster H., Fuhrmann F., Weiss W., Ladstatter S., Dini A., Murg S., Mayer H., Brijacak I., Reiterer B. Gaze-based Human Factors Measurements for the Evaluation of Intuitive Human-Robot Collaboration in Real-time. IEEE International Conference on Emerging Technologies and Factory Automation, ETFA. Institute of Electrical and Electronics Engineers Inc.: 2019.
[5] Heyer C. Human-robot interaction and future industrial robotics applications. IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings. 2010.
[6] Tsarouchi P., Makris S., Chryssolouris G. Human–robot interaction review and challenges on task planning and programming. International Journal of Computer Integrated Manufacturing. 2016, 29, p. 916–931.
[7] Kardos C., Kemény Z., Kovács A., Pataki B. E., Váncza J. Context-dependent multimodal communication in human-robot collaboration. Procedia CIRP. Elsevier B.V.: 2018.
[8] Rossi S., Leone E., Fiore M., Finzi A., Cutugno F. An extensible architecture for robust multimodal human-robot communication. IEEE International Conference on Intelligent Robots and Systems. 2013.
[9] Abich J., Daniel ·, Barber J., Barber D. J. The impact of human-robot multimodal communication on mental workload, usability preference, and expectations of robot behavior. J Multimodal User Interfaces. 2017, 11, p. 211–225.
[10] Ong S. K., Yuan M. L., Nee A. Y. C. Augmented reality applications in manufacturing: a survey. International Journal of Production Research. 2008, 46, p. 2707–2742.
[11] Chang C.-C. New Approach for Static Gesture Recognition. J. Inf. Sci. Eng. 2006, 22, p. 1047–1057.
[12] Pires J. N. Robot-by-voice: Experiments on commanding an industrial robot using the human voice. Industrial Robot. 2005, 32, p. 505–511.
[13] Nee A. Y. C., Ong S. K. Virtual and augmented reality applications in manufacturing. IFAC Proceedings Volumes (IFAC-PapersOnline). IFAC Secretariat: 2013.
[14] Murar M., Brad S. Monitoring and controlling of smart equipments using Android compatible devices towards IoT applications and services in manufacturing industry. Proceedings of 2014 IEEE International Conference on Automation, Quality and Testing, Robotics, AQTR 2014. IEEE Computer Society: 2014.
[15] Kühnlenz B., Erhart M., Kainert M., Wang Z. Q., Wilm J., Kühnlenz K. Impact of trajectory profiles on user stress in close human-robot interaction. At-Automatisierungstechnik. 2018, 66, p. 483–491.
[16] Andronas D., Argyrou A., Fourtakas K., Paraskevopoulos P., Makris S. Design of Human Robot Collaboration workstations - Two automotive case studies. Procedia Manufacturing. Elsevier B.V.: 2020.
[17] Brooke J. SUS-A quick and dirty usability scale. 1995