Camera 3D positioning mixed reality-based interface to improve worker safety, ergonomics and productivity

Adolfo Muñoz, Ana Martí, Xavier Mahiques, Luis Gracia,
J. Ernesto Solanes*, Josep Tornero

*Instituto de Diseño y Fabricación, Universitat Politècnica de València, Camino de Vera s/n, 46022 València, Spain (*Corresponding author e-mail: esolanes@idf.upv.es)*

Abstract

This research develops a new mixed reality-based worker interface for industrial camera 3D positioning, which is intuitive and easy to manage, in order to enhance the worker safety, ergonomics and productivity. An experimental prototype to be used in the car body quality control is developed in the paper. The benefits and drawbacks of the proposed interface are discussed along the paper and sustained through several usability tests conducted with users familiar and not-familiar with mixed reality devices. Furthermore, the feasibility of the proposed approach is demonstrated by tests made in an industrial environment with skilled workers from Alfatec Sistemas company.

Keywords: Camera 3D positioning, camera calibration, mixed reality interface, assisted maintenance, quality control.

---

1. Introduction

1.1. Motivation

Industrial systems based on robotics [1] and artificial vision have allowed the automation of complex industrial processes [2, 3]. Technological advances have made it possible to introduce artificial vision sensor networks, powerful image processing and artificial intelligence algorithms [4] that have improved the productivity [5] and both worker safety and ergonomics [6] in many industrial processes.

Camera calibration is the process of estimating the parameters of a camera [7] and it is one of the most important processes in computer vision since the success or failure of vision applications depend to a great extent on it

Preprint submitted to CIRPJ of Manufacturing Science & Tech. March 16, 2020

[8, 9]. Camera parameters include intrinsic, extrinsic, and distortion coefficients. Classical calibration methods are based on mapping 3D points, which are given by a known calibration pattern, in order to obtain internal parameters of a camera [10, 11]. Although this method is possibly the most used, readers are referred to [12, 13] in order to learn about other approaches.

While the camera calibration using mappings of 3D feature points simplifies calibration process for intrinsic and extrinsic parameters of each camera of a network, estimating camera 3D poses with respect to a fixed global coordinate system is still required and it is one of the most time consuming tasks of the calibration process in industry.

For most of complex industrial applications using camera networks, e.g., 3D metrology, quality control, etc., two main methods are used for setting and calibrating the cameras: 1) the resolution and framerate of each camera is fixed and, once the system has been developed according to the specifications of the client and the restrictions of the industrial plant, the cameras are placed and the lenses are chosen so that the product can be inspected; 2) a virtual design of the system is carried out using a specific simulation software such as Blender or Siemens NX, among others. Thus, taking into account the customer specifications and the industrial plant restrictions, the optimal values for the cameras' extrinsic parameters are obtained: location, resolution, focal length, etc.

The first method mentioned above usually requires less knowledge and time to configure the camera network. However, engineers tend to oversize the camera system (number of cameras, camera resolution, focal lenses, etc.) due to the lack of knowledge about the network optimal configuration. In contrast, the second method mentioned above requires a higher level of knowledge in order for the engineers to use the simulation software and, usually, a fine-tuning process is also required once the camera network is placed in the industrial plant. Without loss of generality, the approach proposed in this work can be included within the framework of the second method mentioned above.

Once the extrinsic and intrinsic camera parameters are obtained from the simulated scenario, a virtual image of each camera is taken as reference image. Expert workers position each camera in the real workspace accordingly to the extrinsic parameters obtained in simulation. Since there are always errors and discrepancies, the reference images are used by workers to perform the fine tuning of the cameras using a laptop or digital tablet in order to compare the real camera view with simulated camera view, and reducing thus the

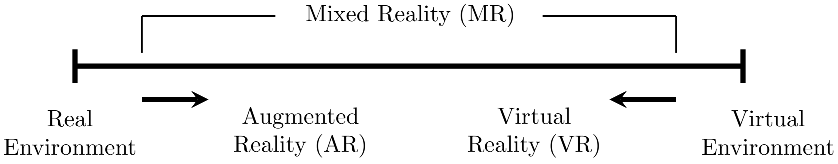


Fig. 1. Illustration of the Reality-Virtuality Continuum presented in [14].

errors. When the camera has to be positioned in accessible locations, e.g., at floor level, the worker holds the device whilst performing the task. This process usually lasts a mean of 20 minutes depending on the skill of the worker. When the camera is located in areas with difficult access, e.g. at high heights, the worker cannot hold the device while setting the camera for safety reasons. In this case, two workers are necessary: one from the ground holds the device, watches the streaming of the camera and gives indications about the corrections to be made, whilst the other is setting the camera according to this indications. Note that this method is not productive neither ergonomic.

1.2. Related work

The development of new approaches and techniques of mixed reality (MR), augmented reality (AR) and virtual reality (VR) [15] allow to enhance many industrial applications from an ergonomic and economic point of view. Fig. 1 illustrates the Reality-Virtuality Continuum described in [14], where MR is the gap between virtuality and reality. The MR technology has been recently introduced to enhance maintenance and production systems [16, 17, 18, 19] and in other industrial processes [20, 21, 22, 23, 24, 25]. As detailed in [25], 14.52% of the industrial applications using these devices are concerned with assembly, maintenance and design operations, while only 1.71% of the applications are related with quality control. In this regard, this is the first research presenting an MR-based solution for industrial camera 3D positioning tasks.

1.3. Objective

The aim of this paper is to enhance ergonomics, safety and productivity of operators working in industrial camera 3D positioning in general throughout a new MR-based interface. The proposed interface replaces current 2D devices (e.g. laptops or digital tablets) and provides a more safety and easy way of 3D positioning industrial cameras. The proposed interface is designed according to the specifications given by the expert workers of Alfatec Sistemas, company that has developed the QEyeTunnel systems in Volkswagen Pamplona and Mercedes-Benz Vitoria plants, located in Spain. In this work, several usability test are carried out with participants with and without knowledge about MR devices, comparing the commanding of the proposed interface using gestures and voice commands. Furthermore, the proposed MR-based worker interface is validated in an industrial environment with expert workers using as MR device the Microsoft® *HoloLens glasses*.

The paper is organized as follows: next section introduces the preliminaries and problem statement, while Section 3 describes in detail the proposed MR-based interface for industrial camera 3D positioning. The feasibility and benefits of the proposed method are shown in Section 4 by experimental tests. Finally, some conclusions are given.

2. Preliminaries and problem statement

This work is focused, without loss of generality, on the camera 3D positioning in automatic vision based defect detection systems on specular surfaces [26]. These systems are basically composed of two parts (see Fig. 2): a lighting system, which can physically move, as described in [27, 28] or stay static, as described in [29, 30, 31], to project light on the surface to be inspected [32, 26]; and a network of cameras positioned around the workspace so that the area of the surface to be inspected is maximized.

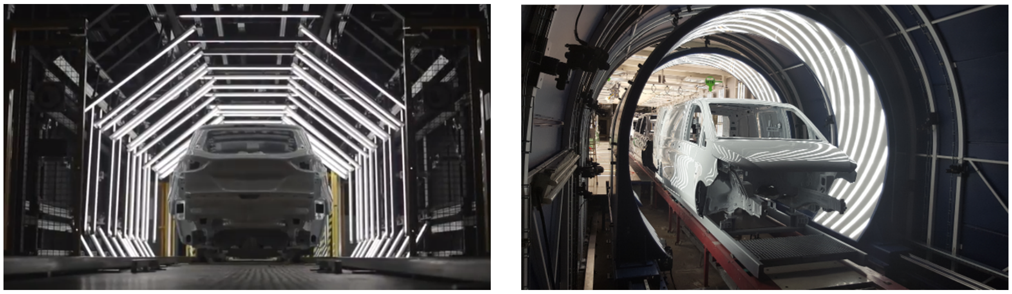
In the design stage of these systems, simulated environments are used (e.g., Blender or Siemens NX), which have render capabilities, in order to be able to see the effect of introducing a new camera in a specific 3D location. In this manner, designers can check a priori the amount of surface and minimum defect size detected by each camera accordingly to the light and resolution[1].

Fig. 3 shows an example of simulated scenario for the QEyeTunnel inspection system (see [28, 33]). Fig. 3(a) shows the entire modeled workspace, whilst Fig. 3(b) and Fig.3(c) show two examples of camera views.

Once the design meets the client specifications, the setup obtained has to be transferred to the real physical system. Note that errors or differences between the simulated camera views and the real camera views will affect

---

[1] Note that the resolution determines the minimum size of defect that can be detected [27].



(a) AIS system developed in Ford factories https://www.youtube.com/watch?v=HroEU8XsaTU.

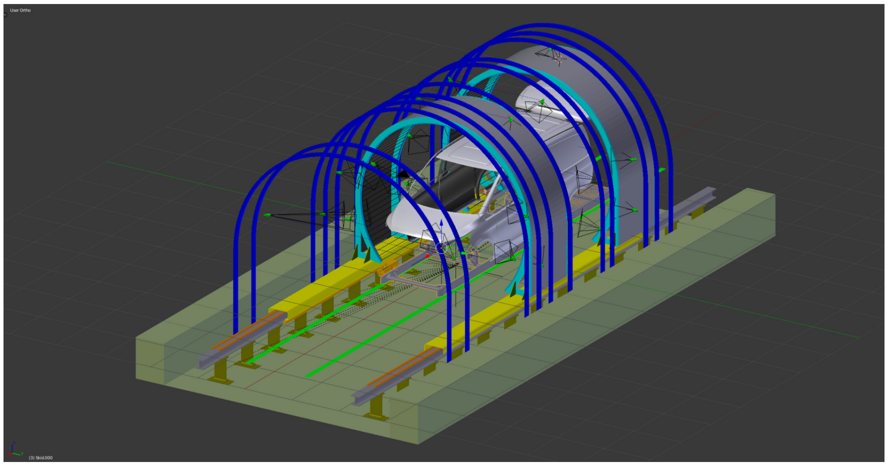
(b) QEyeTunnel system developed in Mercedes factories https://www.youtube.com/watch?v=jN8vazudEXc.

(c) Expert performing the camera 3D positioning task in the QEyeTunnel system.

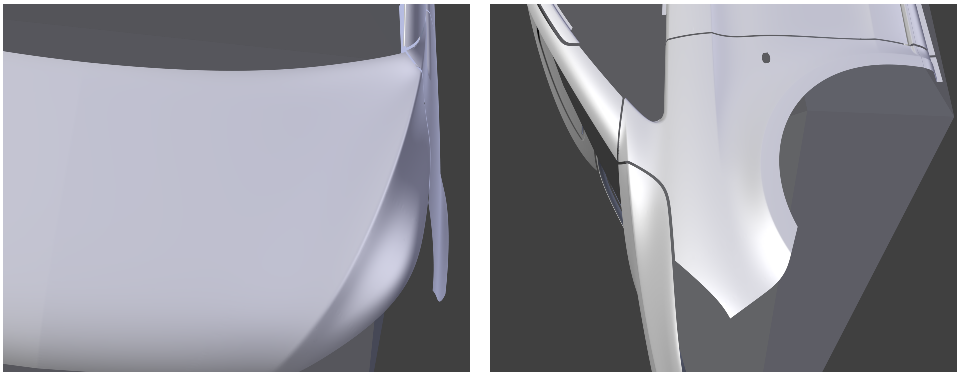
Fig. 2. Automatic inspection systems: Ford (courtesy of Ford España S.A.U., Almussafes plant) and Mercedes (courtesy of Mercedes-Benz España S.A.U., Vitoria plant).

the system performance, i.e., reduction of the total area inspected or the minimum defect size that the system will be able to detect. In order to place the cameras in the actual workspace, two stages are carried out by the workers:

- First, there is a "rough camera 3D positioning", where the worker places the camera in X-, Y- and Z-axes and introduces the orientation values given in simulation. Since errors when fixing the camera in the actual workspace occur, there will be discrepancies between the real camera view and the simulated camera view.
- Second, there is a "fine camera 3D positioning" in order to match to



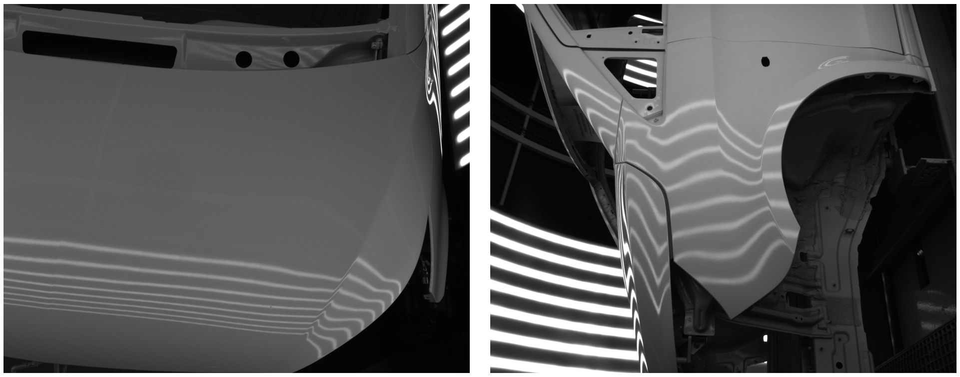
(a) 3D view of the system.



(b) Camera 1 view.

(c) Camera 5 view.

Fig. 3. Simulated scenario of the QEyeTunnel in [28, 33].



(a) Camera 1 view. (b) Camera 5 view.

Fig. 4. Real camera views of the QEyeTunnel in [28, 33].

a certain extent the real view and the simulated view. In this stage, the worker receives feedback from the actual camera and, using references such as fixed parts of the system structure, is able to match both images.

Examples of camera's view from the QEyeTunnel are depicted in Fig. 4. Note that, after the camera is set, the camera views in Fig. 4(a) and Fig. 4(b) match the simulated camera views in Fig. 3(b) and Fig. 3(c), respectively.

Several issues arise from this method affecting directly the worker safety, ergonomics and productivity. Firstly, to compare the simulated camera view and the real camera view in order to reduce the error between them, the worker needs to obtain real-time feedback from the real camera. Currently, workers use a laptop or similar devices to obtain the camera feedback. Several problems arise from the use of this kind of devices as discussed below. With respect to the worker ergonomics, the current weight of laptops or similar devices make it very difficult to hold them for a long time (e.g., more than 10 minutes) with the arm, which can cause muscle problems. In addition, having to look at the laptop screen while changing the position/orientation of the camera produces tension in the neck and cervical vertebrae, leading the worker to long-term cervical diseases. With regard to safety, the worker performs most of the camera position tasks working at height using ladders or safety harnesses. In this situation, the workers are not able to carry heavy devices (e.g., laptops) or, if they are, they put themselves into an unnecessary risk, reducing their stability and increasing the danger of falls.

To avoid the above problems, at present, two workers are needed to perform the task of positioning cameras at heights: one worker holds the laptop and indicates corrections to be made to a second worker, who modifies the camera position according to the instructions made by the first worker. In addition to increasing costs, this method usually requires more time to properly position the camera since the visual feedback is not received directly by the worker who modifies the position of the camera.

3. Camera 3D positioning MR-based worker interface

Next, the proposed MR-based interface for industrial camera 3D positioning tasks is fully described. Without loss of generality, this work assumes the existence of a local server connected to both the inspection system (i.e., video streaming from cameras) and the factory database such as in [33, 28]. It is also assumed that wireless connections in the factory are allowed.

Table 1 shows the main functionalities of the MR-based interface proposed in this work. Three main menus were designed: a first menu, where the worker introduces his credentials and, according to this, the worker will have complete or restricted access to the interface functionalities; a second menu, where the worker with full access can modify the device/interface settings such as connection settings or application settings; and a main panel, in which the camera view is always in foreground and that includes a navigation panel in order to manage image aspects such as zoom in/out or image displacement, among others, and also a state indicator in order to know if the interface has received the worker command correctly. It is worthy to mention that the specifications and main functionalities were determined by means of interviews with expert workers in industrial camera 3D positioning tasks of Alfatec Sistemas company.

Fig. 5 shows the block diagram of the camera 3D positioning process with the introduction of the proposed MR approach. A server UDP/IP runs within the local server workstation with the following tasks: allowing the connection with MR devices via UDP/IP; accessing the factory databases in order to load the correspondent reference features and other necessary information to perform the task; applying worker commands to the acquired video stream, i.e., image zoom in/out, image displacement, among others; and saving the data introduced by the worker to factory databases. The MR device is in charge of running a client application with the following tasks: introducing into the real workspace a virtual interface which meets the specifications

Table 1. MR-based interface functionalities.

CREDENTIALS MENU	EXPERT	FULL ACCESS
	WORKER	RESTRICTED ACCESS
CONFIGURATION MENU	CONNECTION SETTINGS	IP ADDRESS CONNECT TO SERVER END COMMUNICATION RESET COMMUNICATION
	APPLICATION SETTINGS	CAMERA ID INTERACTION MODE (GESTURE OR VOICE)
MAIN PANEL	CAMERA WINDOW (GESTURE OR VOICE)	ADJUSTABLE SIZE ZOOM GRID
	NAVIGATION PANEL (GESTURE OR VOICE)	LEFT/RIGHT/UP/DOWN OPTION BACK/RESET OPTION
	STATUS INDICATOR	COMMUNICATION STATUS COMMAND ACK/NACK CAMERA STATUS

shown in Table 1; receiving both gesture and voice commands from worker and send them via UDP/IP to the server; and displaying the modified video streaming from the local server workstation. The worker closes the loop by modifying the camera position/orientation according to the visual feedback received and also giving commands in order to zoom in/out the image or move through the canvas.

Fig. 6 shows the block diagram of the overall process. As can be seen, the MR device acts as a client and another device (e.g., laptop, Workstation or High Performance Computing platforms) acts as a server.

The client receives the video streaming provided by the server. This video streaming contains the reference image and the real image overlapped. The client projects the received video streaming in the form of a hologram, which is used by the worker as visual feedback. Based on this information, the worker modifies the position/orientation of the camera to reduce the matching error between the reference image and the real image. In addition, the worker can interact by voice or gesture with the MR device using all the programmed functionalities, see Table 1.

The server receives the camera number to be viewed by the client and accesses to its video streaming. At the same time, the server loads from a

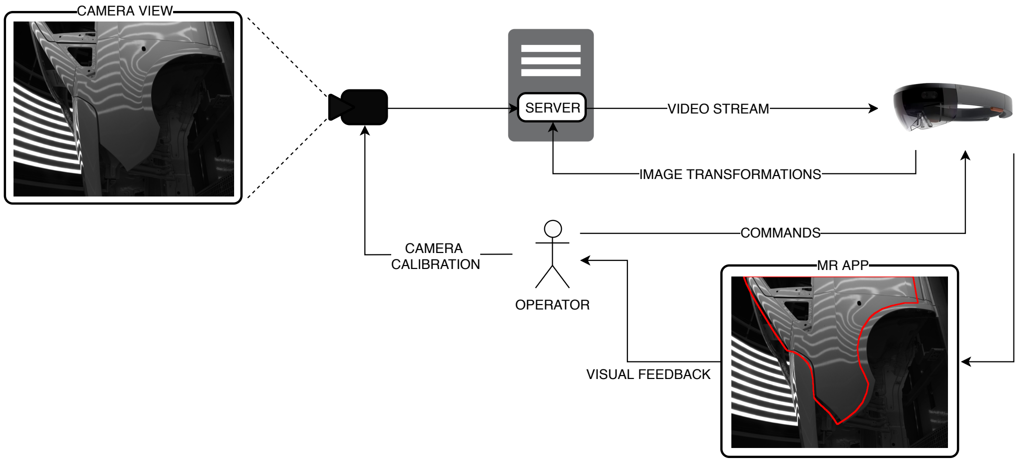


Fig. 5. Graphical representation of the proposed approach for the quality control of car body surfaces: red contour in MR APP block refers to the reference for the worker to be used to perform the matching task.

database the reference image obtained in simulation, overlapping both the received video streaming and the reference image. Depending of the client demands (i.e., zoom and scroll), the server applies the correspondent transformation to the overlapped video streaming and sends the result to the client.

4. Experimental Results

To show the viability of the proposal approach, the MR-based worker interface for industrial camera 3D positioning has been implemented in the Microsoft® *HoloLens glasses* [34], although other MR devices could also be used.

4.1. Usability analysis results

Similarly to [35, 36, 37], several methods such as usability tests of applications, which are traditionally used to validate hardware and software, and interviews were conducted to show the advantages of the proposed approach.

The set-up used for the usability tests is shown in Fig. 8. It is composed of an industrial camera, which is ceiling mounted using a 3 degrees of freedom camera support *Manfrotto 405 Pro Geared Tripod Head*, a training hood and two testing car body parts: a hood and a door. The camera is located

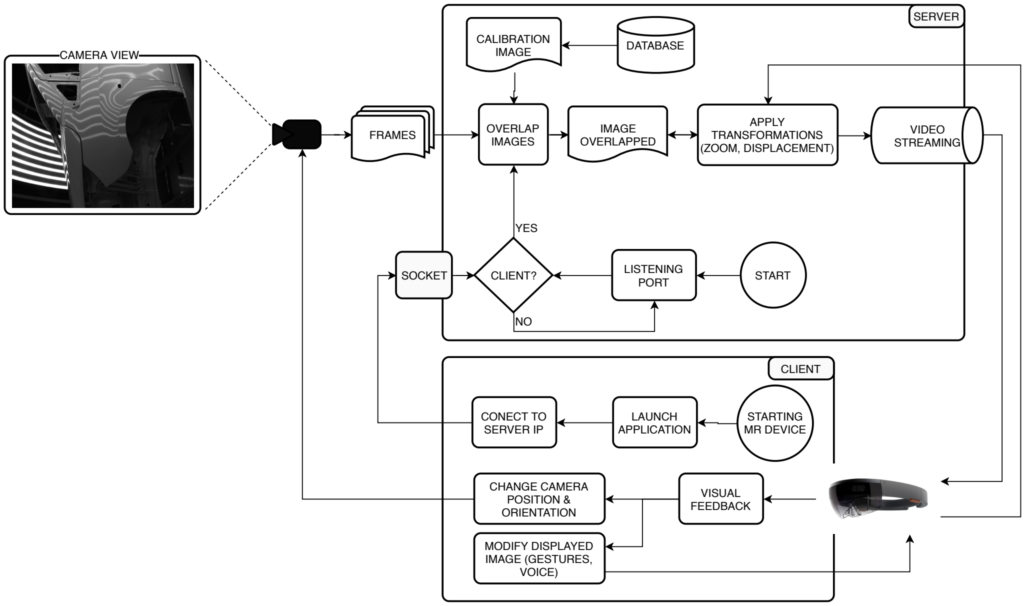


Fig. 6. Client/Server application flow chart.

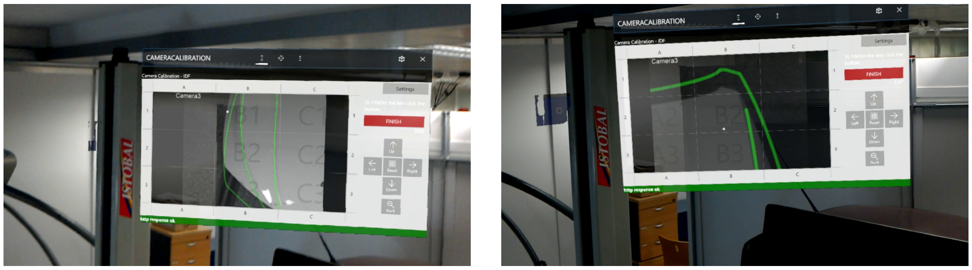
two meters high so participants have to climb a ladder to manipulate it, simulating in this way stress situations that occur in the industry.

Twenty (20) participants, with and without experience using MR devices, where selected for the usability tests. The main information about these participants is shown in Fig. 9. It is worth to mention that none of these participants had ever been involved in camera 3D positioning tasks and the 70% of them had never used MR devices.

An explanation of the task was given to each participant followed by a short training (15 min approx.) in order to get them used to the MR interface and the camera 3D positioning task. Hereafter, each participant performed two tests: first, a camera 3D positioning task commanding the interface by gestures (a video demonstrative can be played at `https://media.upv.es/player/?id=33be4390-a4af-11e9-abe1-f718df9623c1`); and second, a camera 3D positioning task commanding the interface by voice (a video demonstrative can be played at `https://media.upv.es/player/?id=096fab50-18d7-11ea-a59e-3f45266cd80b`).

Fig. 10 shows the results of the sum square error (SSE) in degrees committed by the participants and the time that took them to finish the task. It is remarkable the fact that both error and time have been reduced con-

(a) Overall view.



(b) Example of zoom x3

(c) Example of zoom x9.

Fig. 7. Proposed MR-based worker interface for industrial camera 3D positioning tasks: green features in the video panel represent the reference features to be matched.

siderably for the second task, despite it is more difficult. This is due in part to the gain of experience by the participants as they perform new camera 3D positioning tasks and, also to the fact that the interaction with the proposed MR-based worker interface is more intuitive by voice command than by gesture command.

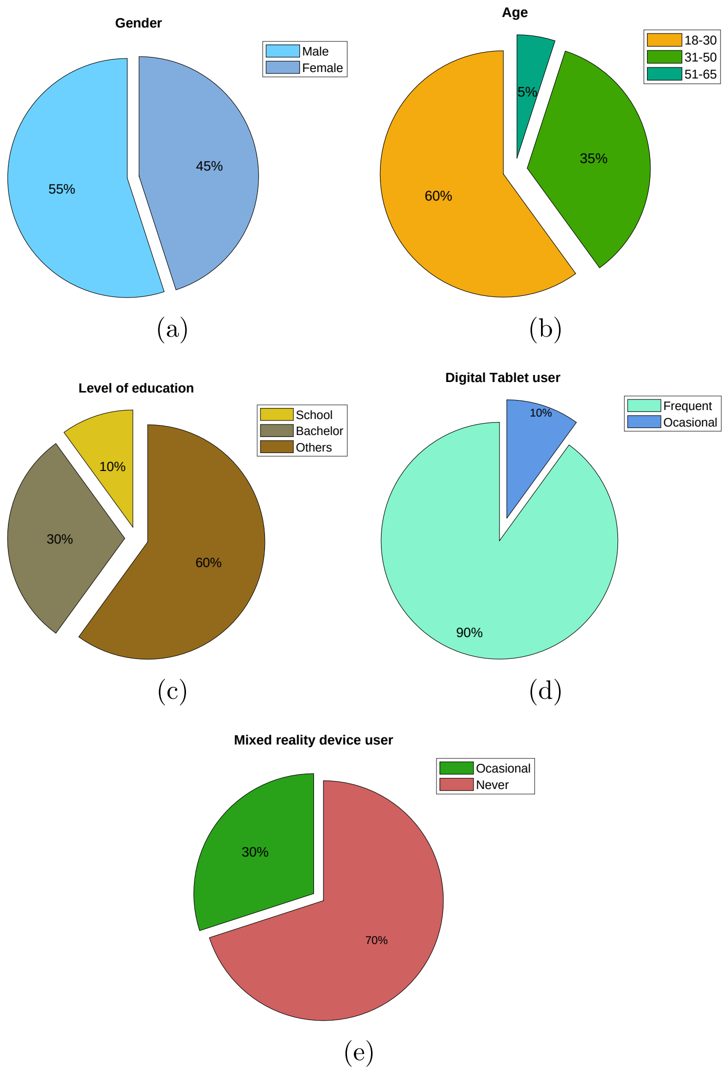
After the test, the participants were asked to complete two standard questionnaires: the NASA Task Load index (NASA-TLX) [38] and the System Usability Scale (SUS) [39]. The NASA-TLX questionnaire was chosen because it is widely used to evaluate physical and digital experiences in working environments, although other questionnaires could also have been used, e.g., the Subjective Workload Assessment Technique (SWAT) [40]. The SUS questionnaire was used to test the usability of the proposed interface because it is short, concise and widely used. However, other similar test could also be used, e.g., the Unified Theory on Acceptance and Use of Technology (UTAUT) [41, 42], the Questionnaire for User Interaction Satisfaction (QUIS) [43] or the USE Questionnaire [44].

On the one hand, the NASA-TLX questionnaire was conducted in order to

Fig. 8. Experimental set-up for the usability test.

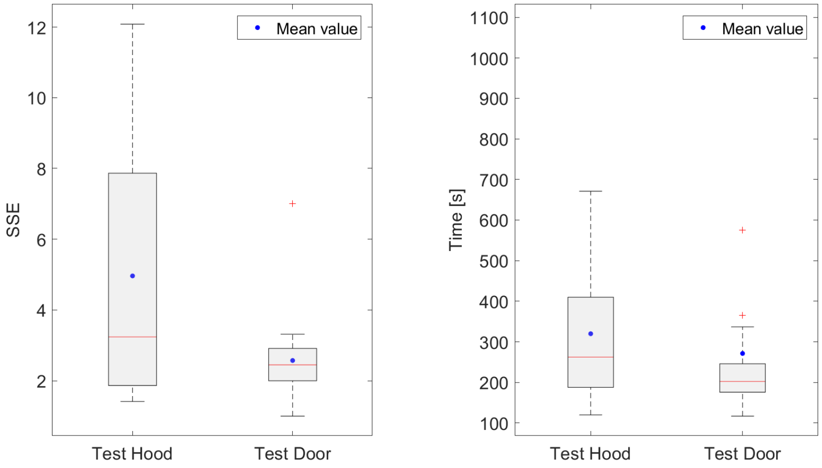
perform subjective workload assessments on the participants. NASA-TLX scores were calculated using the official NASA-TLX application `https://software.nasa.gov/software/ARC-15150-1A`. Accordingly to it, workload is defined as the weighted score of six different categories: mental demand, physical demand, temporal demand, effort, performance and frustration level. The participants rated each category on a scale from 0 (low) to 100 (high). On the other hand, the SUS questionnaire was conducted to evaluate whether the participants were comfortable using the proposed interface for the camera positioning task.

Fig. 11 shows the scores of the NASA-TLX questionnaire. Note that the mental and physical demand scores are 48% and 42%, respectively, which are



(a) (b) (c) (d) (e)

Fig. 9. Main information of the participants involved in the usability tests.



(a) Axis error (degrees). (b) Time (seconds).

Fig. 10. Performance of the usability test participants.

typical scores for these kind of tasks. Note also that the frustration score is around 35%, which is lower than expected, since the participants were not familiar with the task. This result is also consistent with the performance score, which indicates that around 81% of the participants were satisfied about their performance.

Regarding the SUS questionnaire, the overall perceived usability was 81,5 out of 100 (min: 42,5 max:100; SD: 14.38), which means that the proposed MR-based interface has reached a high level of usability. In addition, Fig. 12 shows the results obtained for each question of the SUS questionnaire, which are detailed in Table 2. Note that most of the participants would use this interface frequently and found the interface easy to use. The participants also indicated that all the interface functionalities were well integrated and that the proposed interface was consistent. Moreover, the participants felt confident with the interface and indicated that it could be used by workers of all educational levels.

In addition to both standard questionnaires mentioned above, a third questionnaire was conducted to specifically evaluate the usability of some aspects of the proposed mixed reality-based interface. This questionnaire had two objectives: the first one was to determine the strengths and weaknesses

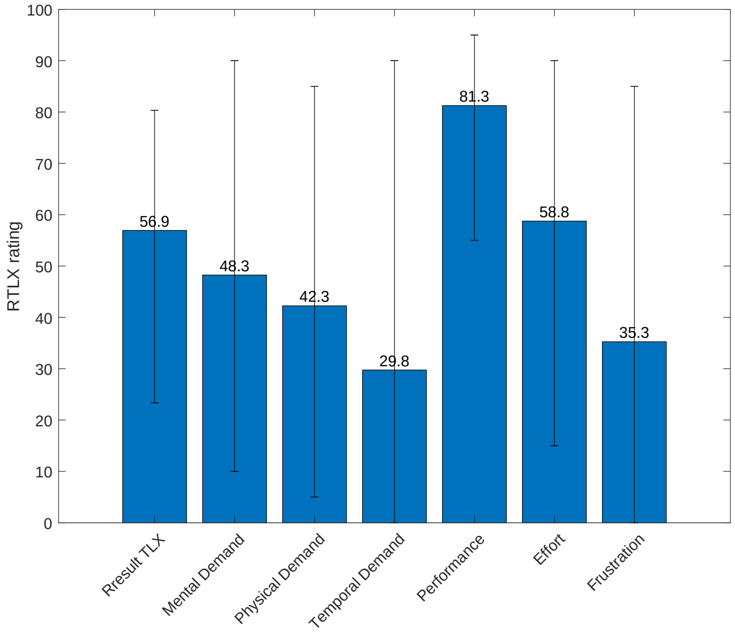


Fig. 11. Results of the NASA-TLX questionnaire.

Table 2. Questions of the SUS questionnaire [39].

Q1	I think that I would like to use this system frequently
Q2	I found the system unnecessarily complex
Q3	I thought the system was easy to use
Q4	I think that I would need the support of a technical person to be able to use this system
Q5	I found the various functions in this system were well integrated
Q6	I thought there was too much inconsistency in this system
Q7	I would imagine that most people would learn to use this system very quickly
Q8	I found the system very cumbersome to use
Q9	I felt very confident using the system
Q10	I needed to learn a lot of things before I could get going with this system

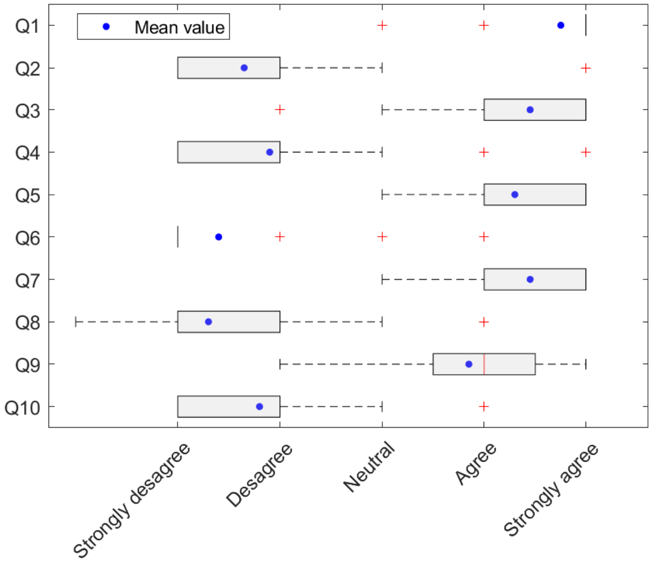


Fig. 12. Results of the SUS questionnaire.

of the proposed interface, both hardware and software; and the second was to determine whether the participants preferred to command the interface using gestures or voice.

Regarding the first objective mentioned above, the YES/NO questions shown in Table 3 were asked to the participants in order to evaluate the following indicators: ergonomics, security, easiness, comfort, handling and weight. Fig. 13(a) shows the results for these questions, where it can be seen that: 70% of the participants agreed that the proposed interface was ergonomic despite that only 55% of them thought that the device weight was not a problem; 85% of the participants agreed that the proposed interface helped to improve the operator safety; 75% of the participants indicated that it improved the worker comfort; and almost all the participants agreed that the proposed interface was easy to manage, use and handle.

Regarding the second objective mentioned above, the questions shown in Table 4 were asked to the participants to know their commanding preference according to the following indicators: easiness, effectiveness, intuitive and satisfaction. Fig. 13(b) shows the results of these questions, where it can be seen that: 75% of the participants indicated that was easier to command the interface by voice than by gestures; 70% of the participants indicated

Table 3. Specific usability questions in the third questionnaire.

Ergonomics	I think that the mixed reality glasses are ergonomics for this task (YES/NO)
Safety	I did not notice any type of risk when interacting with the interface (dizziness, distraction, etc.) (YES/NO)
Easiness	I found the interface was easy to use (YES/NO)
Comfort	I felt comfortable during the task (YES/NO)
Handling	I found the interface easy to manage (YES/NO)
Weight	I think that the mixed reality device weighted too much (YES/NO)

Table 4. Specific questions in the third questionnaire to compare gesture and voice commands.

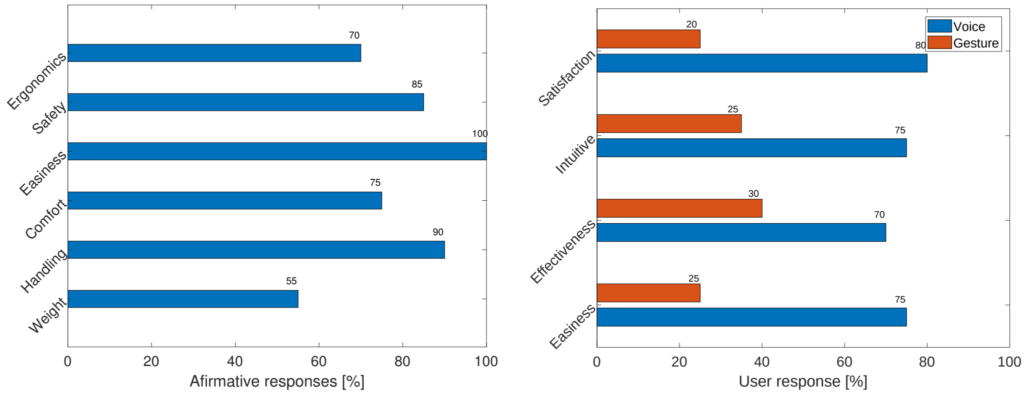
Easiness	I found easier to command the interface by (gestures or voice?)
Effectiveness	I found more effective commanding the interface by (gestures or voice?)
Intuitive	I found more intuitive commanding the interface by (gestures or voice?)
Satisfaction	I felt very satisfy commanding the interface by (gestures or voice?)

that voice commands were more effective (i.e., command recognition by the application) than gesture commands; 75% of the participants indicated that commanding by voice was more intuitive than gestures; and 80% of the participants were more satisfied using voice commands than gesture commands.

Finally, the participants were allowed to add comments about their experience with the proposed procedure, see the discussion in Section 5.

4.2. Industrial tests

Recently, automatic inspection systems for detecting car body surface defects have been developed in specialized well known companies [29, 30, 31, 28]. Fig. 14 shows the inspection system developed by Proemisa and Alfatec companies in Valencia (Spain) for detecting defects on car body surfaces. This system comprises an external fixed structure with a certain number of cameras that are optimally placed to inspect the body surface, as well as an



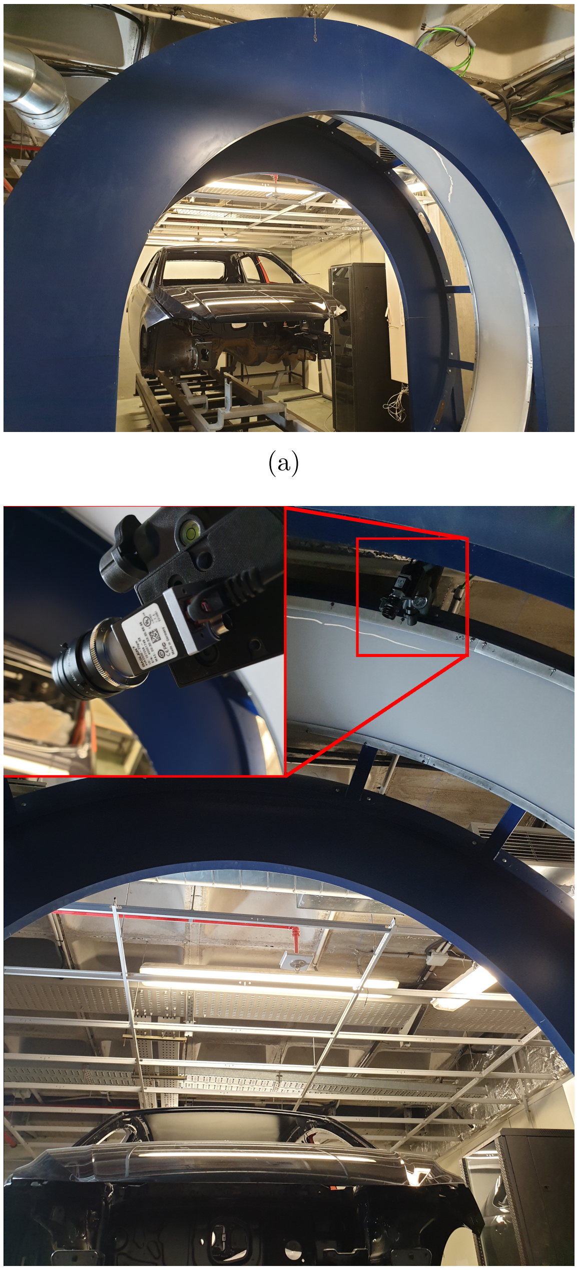
(a) Usability. (b) Gesture and voice comparison.

Fig. 13. Results of the third questionnaire.

internal structure that houses a curved screen that projects the illumination on the body surface. The number of required cameras and the size of the curved screen depend on the car body to be analyzed by the system. In this particular case, the system consists of 30 monochrome cameras of 5MP working at 30 fps, and a curved screen of 3000 mm in diameter and a resolution of 1152x64 pixels. The total size of the system, including the cover, is 3250x3250x1200 mm.

Fig. 15 shows the components of the hardware architecture and the communications between them for the automatic quality control system prototype. The main elements of this architecture are detailed below:

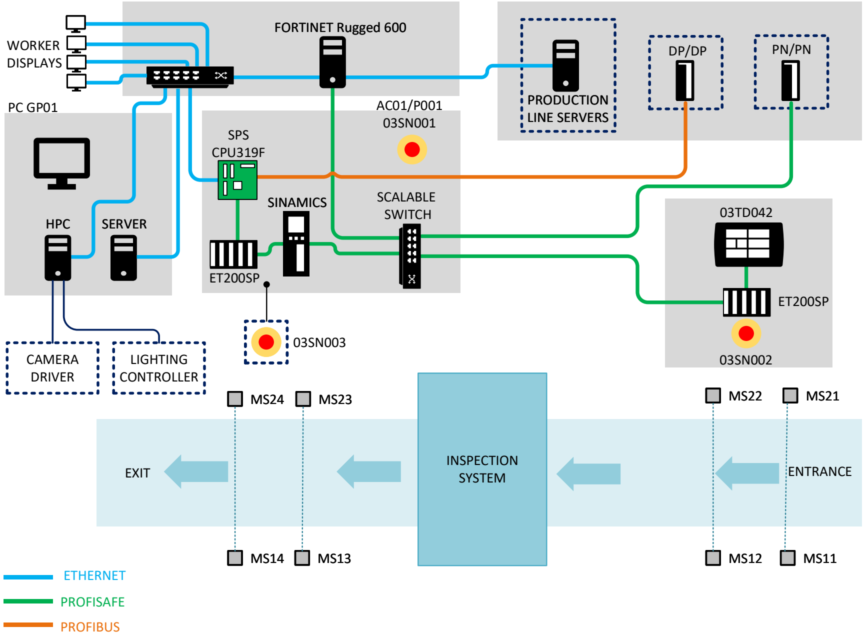
- *High Performance Computing Platform (HPC):* this element controls the program flow. It is an industrial vision system from *Matrox*® named *Supersight*, which is an entry-level configurable single-node high-performance computing (*HPC*) platform supporting two multi-core *Intel*® *Xeon* processors, third-party *GPU*s and *Matrox*® *FPGA* boards for demanding industrial imaging applications. It is equipped with 8 *GigE Card PCIe AdLink*® *GIE64+* with Power over Ethernet technology (*PoE*), and 5 ports for supporting 30 cameras. Moreover, it is equipped with 2 *Gigabyte*® *GeForce GTX 1080 8GB GDDR5X Dual Link DVI-D HDMI 3X DisplayPort PCI-E* graphic cards, which were used to implement the algorithm for defect detection and also to run the program that generates the patterns to be projected. This system



(a)

(b)

Fig. 14. Inspection system developed by Proemisa and Alfatec companies in Valencia, Spain.



(a)

Fig. 15. Hardware architecture and communications for the automatic quality control system prototype (see [33] for more details).

communicates with the *SERVER* using the *TCP-IP* communication protocol.

- *System High-level Controller (SERVER):* this element provides the inspection results to the workers of the production line, as well as acts as an interface between line workers and the *HPC* on maintenance issues. It is an industrial *PC*-based system that communicates with the *Worker Displays*, *Production Line Servers* and the *HPC* systems using the *TPC-IP* communication protocol.
- *Worker displays:* several screens controlled by small *PC*s are placed throughout the production line to display the results of the inspection to the workers, who use this information to locate the defects and act accordingly (i.e., fixing them whenever possible or marking them for

later repair).

- *Production Line Servers:* the results of each inspection and the system backups are saved on these servers. The communication with the *SERVER* is through *TCP-IP* but using a *FORTINET* protocol. This element is not only a data security system but also performs bigdata analysis to identify problems in the painting process, yielding significant money savings and a better quality of the final product.

In order to evaluate the MR-based interface proposed in this work, it has been compared with the PC-based interface currently used by expert workers. The evaluation was carried out with 7 experts from Alfatec Sistemas company. Since these workers were already familiar with the camera 3D positioning tasks, only a short training was made in order to introduce them to the Microsoft® *HoloLens glasses* and the developed application. Fig. 16 shows two expert workers setting a camera, whilst Fig. 17 shows several instants of the camera 3D positioning process: Fig. 17(a) and Fig. 17(b) correspond to the rough 3D positioning; Fig. 17(c) to Fig. 17(e) correspond to the fine 3D positioning; and Fig. 17(f) shows the final result of the camera positioning task.

Three indicators were evaluated in this experiment: productivity, ergonomics and safety.

To assess productivity, as it was done in the previous section, it was measured the time that took each worker to perform the task of positioning a camera at height, i.e., the workers had to use a ladder to position the camera. The average time to complete the camera positioning task using the PC-based interface was around 15 minutes, while that using the proposed MR-based interface was only around 8 minutes. Hence, the time to complete the task was reduced by around 50% using the proposed approach, which is a significant improvement in the productivity of the expert workers.

To assess ergonomics and safety, two questionnaires were considered: the SUS questionnaire and another one specifically designed to evaluate the proposed application. Table 5 shows the six questions included in this questionnaire: Q1 to Q3 are related to the safety indicator, while Q4 to Q6 are related to the ergonomics indicator. In addition, the expert workers were allowed, as before, to add comments about their experience with the proposed procedure, see the discussion in Section 5.

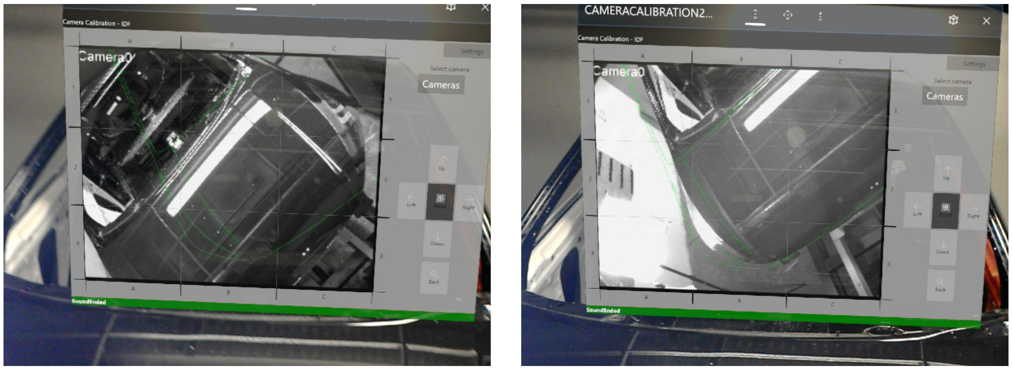
Regarding the SUS questionnaire, the overall perceived usability of the PC-based interface was 33,5 out of 100 (min 31, max 36, SD 1.87), while that


(a) (b)

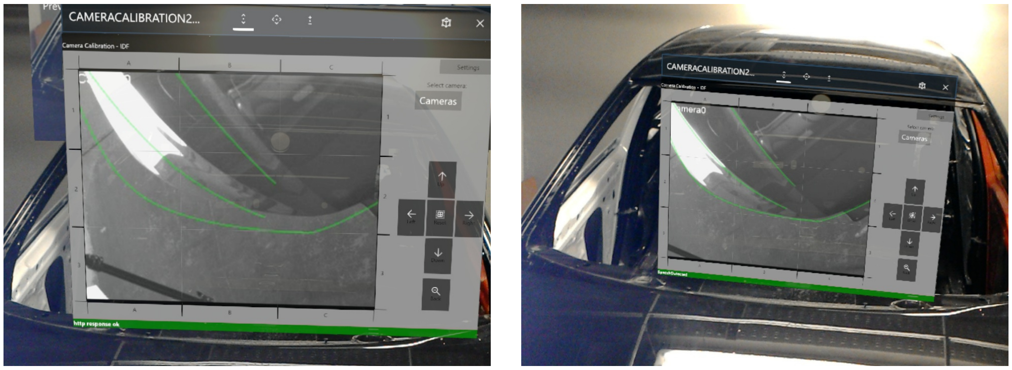
Fig. 16. Experts working in camera 3D positioning tasks using the interface proposed in this work.

of the MR-based interface was 80,25 out of 100 (min: 79.9 max: 80.7; SD: 0.32). Hence, it is concluded from these results that the proposed interface clearly provides a higher level of usability.

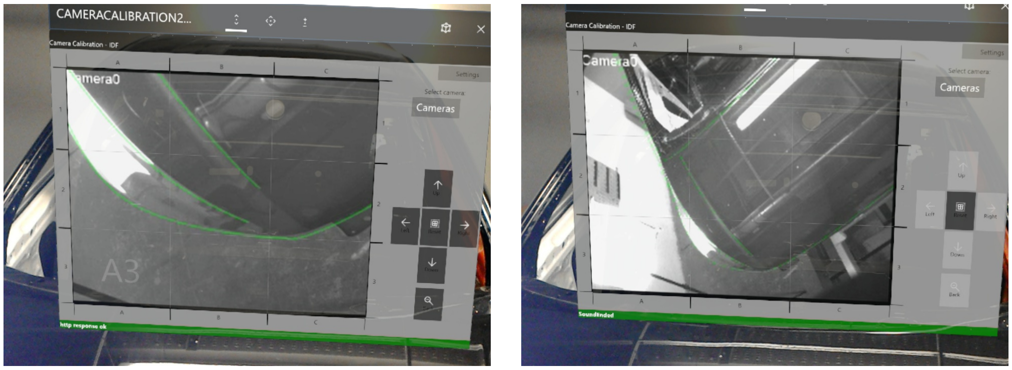
Regarding the questionnaire shown in Table 5, Fig. 18(a) and Fig. 18(b) show the results obtained when using the PC-based and MR-based interfaces, respectively. In particular, the results of the six questions indicate: Q1 - the expert workers felt more risk when using the PC-based interface that when using the MR-based interface; Q2 - the expert workers had almost the same feeling of distraction or dizziness using both interfaces; Q3 - the expert workers had to take more risks when the task was performed using the PC-



(a) Rough rotation camera placement 1. (b) Rough translation camera placement 2.



(c) Fine translation camera placement (second level of zoom). (d) Fine rotation camera placement 1 (first level of zoom).



(e) Fine translation camera placement (second level of zoom). (f) Result of the calibration.

Fig. 17. Expert working in camera 3D positioning tasks using the proposed MR-based worker interface (worker's view).

Table 5. Questions for the experts to evaluate the traditional PC-based and proposed MR-based interfaces.

Q1	I think that the task has little accident risk
Q2	I noticed some type of risk when interacting with the interface (dizziness, distraction, etc.)
Q3	I had to take some risks during the task
Q4	I felt aches caused by bad body postures
Q5	I think that the task requires low physical effort
Q6	I think that the task requires low visual effort

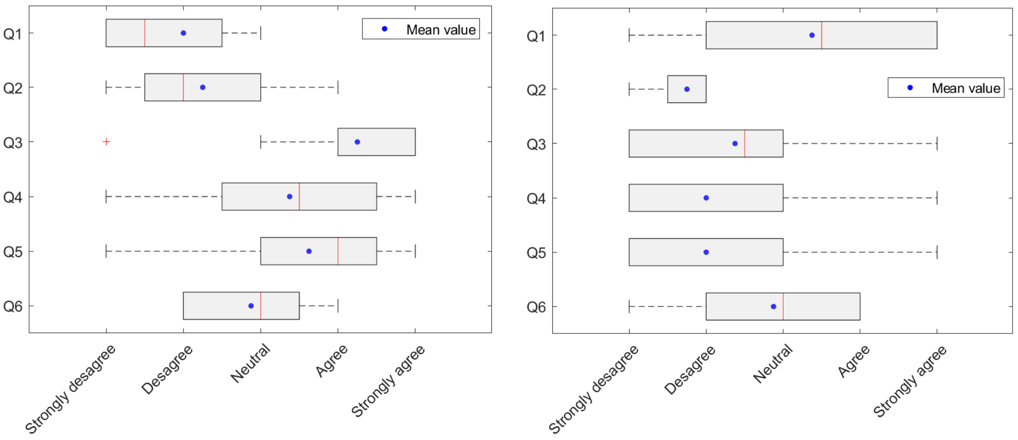
based interface; Q4 - the expert workers felt more aches produced by bad body postures when using the PC-based interface; Q5 - the task requires more physical effort when the expert workers used the PC-based interface; and Q6 - both interfaces require the same visual effort.

Therefore, from the data analyzed above it is concluded that the proposed interface improves the productivity, ergonomics and safety of the worker during the accomplishment of the camera positioning task.

5. Discussion

The results of previous section show a significant improvement in terms of the worker's productivity, ergonomics and safety when performing the camera 3D positioning task using the proposed MR-based interface with respect to the traditional PC-based interface. However, some comments made by the expert workers should be taken into account when developing future industrial versions of the proposed interface, as discussed below.

The expert workers highlighted the great robustness of the application working in an industrial environment and pointed out that a second expert was not required to place the cameras at heights, which means an improvement of the expert self-sufficiency. The experts also pointed out that the device and virtual objects did not affect their view, which means an improvement of the worker safety. Moreover, they highlighted the easiness to place the virtual interface anywhere in the environment and to modifying its size, reducing the physical stress of the expert while performing the task,



(a) Results using the PC-based interface. (b) Results using the MR-based interface.

Fig. 18. Comparative results of the experts for the questionnaire shown in Table 5.

which means an improvement of the worker ergonomics.

Regarding the last comment, one expert worker suggested to introduce the capability of moving the virtual interface in unison with the head (i.e., the MR glasses) so that the worker does not have to place it manually in a specific place and, thus, the interface is always seen. The authors consider that this is a good suggestion but requires further studies and considerations before its implementation, e.g., the interaction with other real objects should be taken into account.

Arguably, the main complaint of both non-expert participants and expert workers was that the field of view (FoV) of the MR device was too small. In particular, the FoV of the Microsoft® *HoloLens glasses* used in the tests is 34-degrees diagonal. Hence, the participants could see digital objects interacting with the real world while looking straight ahead, but, if they turned their head a little, digital objects disappeared or got cut off. In order to alleviate this issue, according to Microsoft's Alex Kipman, the FoV has been doubled from *HoloLens* to *HoloLens 2* [45], i.e., the FoV of *HoloLens 2* is 52-degrees diagonal.

The expert workers also expressed their worries about the fatigue produced by the weight of the Microsoft® *HoloLens glasses*. Moreover, they pointed out the difficulty of wearing them together with safety helmets. To solve this issue, Microsoft® has developed the *Trimble XR10* with the *HoloLens 2* [46], which is the first certified safety helmet that integrates the

MR device.

6. Conclusions

A mixed reality-based worker interface for industrial camera 3D positioning tasks has been developed in this work to improve the worker ergonomics, safety and productivity. The main functionalities and characteristics of the proposed interface have been fully described.

Without loss of generality, the proposed mixed reality-based interface was implemented using the Microsoft® *HoloLens glasses* (first generation) and was tested for the particular case of camera 3D positioning in inspection systems of car body surfaces.

In order to evaluate the usability and intuitiveness of the proposed interface, a first test was conducted with non-expert users. The results showed that almost all the participants found the interface easy to use, consistent, easy to learn and comfortable. In addition, around 80% of the participants indicated that commanding the interface by voice was more intuitive than by gestures.

In order to prove the feasibility and advantages of the proposed mixed reality-based interface, several expert workers from Alfatec Sistemas company compared the traditional PC-based interface and the proposed mixed reality-based interface in an industrial environment. The results showed that the proposed interface enhance around 50% the productivity of the expert worker with respect to the PC-based interface. The results also indicated that the expert workers felt more safety using the proposed mixed reality-based interface when working at heights. Finally, the results also showed a significant improvement of the expert workers ergonomics when performing the 3D positioning task with the proposed approach.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the Spanish Government under the Project DPI2017-87656-C2-1-R.

References

[1] G. Michalos, J. Spiliotopoulos, S. Makris, G. Chryssolouris, A method for planning human robot shared tasks, CIRP Journal of Manufacturing Science and Technology 22 (2018) 76–90.

[2] Y. Ogren, P. Tóth, A. Garami, A. Sepman, H. Wiinikka, Development of a vision-based soft sensor for estimating equivalence ratio and major species concentration in entrained flow biomass gasification reactors, Applied Energy 226 (2018) 450 – 460. doi:https://doi.org/10.1016/j.apenergy.2018.06.007.

[3] J. Arents, R. Cacurs, M. Greitans, Integration of computer vision and artificial intelligence subsystems with robot operating system based motion planning for industrial robots, Automatic Control and Computer Sciences 52 (5) (2018) 392–401. doi:10.3103/S0146411618050024.

[4] A. Belitzki, C. Stadter, M. Zaeh, Distortion minimization of laser beam welded components by the use of finite element simulation and artificial intelligence, CIRP Journal of Manufacturing Science and Technology 27 (2019) 11–20.

[5] M. Colledani, T. Tolio, A. Yemane, Production quality improvement during manufacturing systems ramp-up, CIRP Journal of Manufacturing Science and Technology 23 (2018) 197–206.

[6] M. Dalle Mura, G. Dini, Optimizing ergonomics in assembly lines: A multi objective genetic algorithm, CIRP Journal of Manufacturing Science and Technology 27 (2019) 31–45.

[7] Z. Zhang, Camera Parameters (Intrinsic, Extrinsic), Springer US, Boston, MA, 2014, pp. 81–85. doi:10.1007/978-0-387-31439-6_152.

[8] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, New York, NY, USA, 2003.

[9] R. Usamentiaga, D. Garcia, Multi-camera calibration for accurate geometric measurements in industrial environments, Measurement 134 (2019) 345 – 358. doi:https://doi.org/10.1016/j.measurement.2018.10.087.

[10] O. D. Faugeras, G. Toscani, Camera calibration for 3d computer vision, 1987.

[11] W.-S. L. Sheng-Wen Shih, Yi-Ping Hung, Efficient and accurate camera calibration technique for 3-d computer vision (1992). doi:10.1117/12.57975.

[12] L. Song, W. Wu, J. Guo, X. Li, Survey on camera calibration technique, in: 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Vol. 2, 2013, pp. 389–392. doi:10.1109/IHMSC.2013.240.

[13] S. M. Ayaz, M. Y. Kim, J. Park, Survey on zoom-lens calibration methods and techniques, Machine Vision and Applications 28 (8) (2017) 803–818. doi:10.1007/s00138-017-0863-1.

[14] P. Milgram, H. Takemura, A. Utsumi, F. Kishino, Augmented reality: A class of displays on the reality-virtuality continuum, 1994, pp. 282–292.

[15] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, J. Gain, A survey of augmented, virtual, and mixed reality for cultural heritage, J. Comput. Cult. Herit. 11 (2) (2018) 7:1–7:36. doi:10.1145/3145534.

[16] G. Vignali, M. Bertolini, E. Bottani, L. D. Donato, A. Ferraro, F. Longo, Design and testing of an augmented reality solution to enhance operator safety in the food industry, International Journal of Food Engineering 14 (2017) 1 – 16. doi:https://doi.org/10.1515/ijfe-2017-0122.

[17] R. Masoni, F. Ferrise, M. Bordegoni, M. Gattullo, A. E. Uva, M. Fiorentino, E. Carrabba, M. D. Donato, Supporting remote maintenance in industry 4.0 through augmented reality, Procedia Manufacturing 11 (2017) 1296 – 1302, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy. doi:https://doi.org/10.1016/j.promfg.2017.07.257.

[18] A. E. Uva, M. Gattullo, V. M. Manghisi, D. Spagnulo, G. L. Cascella, M. Fiorentino, Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations, The International Journal of Advanced Manufacturing Technology 94 (1) (2018) 509–521. doi:10.1007/s00170-017-0846-4.

[19] M. Gattullo, G. W. Scurati, M. Fiorentino, A. E. Uva, F. Ferrise, M. Bordegoni, Towards augmented reality manuals for industry 4.0: A methodology, Robotics and Computer-Integrated Manufacturing 56 (2019) 276 – 286. doi:https://doi.org/10.1016/j.rcim.2018.10.001.

[20] V. Elia, M. G. Gnoni, A. Lanzilotto, Evaluating the application of augmented reality devices in manufacturing from a process point of view: An ahp based model, Expert Systems with Applications 63 (2016) 187 – 197. doi:https://doi.org/10.1016/j.eswa.2016.07.006.

[21] R. Hanson, W. Falkenstrom, M. Miettinen, Augmented reality as a means of conveying picking information in kit preparation for mixed-model assembly, Computers and Industrial Engineering 113 (2017) 570 – 575. doi:https://doi.org/10.1016/j.cie.2017.09.048.

[22] C. Liu, S. Cao, W. Tse, X. Xu, Augmented reality-assisted intelligent window for cyber-physical machine tools, Journal of Manufacturing Systems 44 (2017) 280 – 286, special Issue on Latest advancements in manufacturing systems at NAMRC 45. doi:https://doi.org/10.1016/j.jmsy.2017.04.008.

[23] J. P. Lima, R. Roberto, F. Simoes, M. Almeida, L. Figueiredo, J. M. Teixeira, V. Teichrieb, Markerless tracking system for augmented reality in the automotive industry, Expert Systems with Applications 82 (2017) 100 – 114. doi:https://doi.org/10.1016/j.eswa.2017.03.060.

[24] J. Neves, D. Serrario, J. N. Pires, Application of mixed reality in robot manipulator programming, Industrial Robot: the international journal of robotics research and application 45 (6) (2018) 784–793. doi:10.1108/IR-06-2018-0120.

[25] E. Bottani, G. Vignali, Augmented reality technology in the manufacturing industry: a review of the last decade, IISE Transactions 0 (ja) (2018) 1–64. doi:10.1080/24725854.2018.1493244.

[26] L. Armesto, J. Tornero, A. Herraez, J. Asensio, Inspection system based on artificial vision for paint defects detection on cars bodies, in: IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, 9-13 May 2011, 2011, pp. 1–4. doi:10.1109/ICRA.2011.5980570.

[27] J. Tornero, L. Armesto, M. C. Mora, N. Montés, A. Herráez, J. Asensio, Detección de defectos en carrocerías de vehículos basado en visión artificial: Diseño e implantación, Revista Iberoamericana de Automática e Informática Industrial RIAI 9 (1) (2012) 93–104. doi:http://dx.doi.org/10.1016/j.riai.2011.11.010.

[28] L. Arnal, J. E. Solanes, J. Molina, J. Tornero, Detecting dings and dents on specular car body surfaces based on optical flow, Journal of Manufacturing Systems 45 (2017) 306–321.

[29] Micro-Epsilon, Sistema de inspección de defectos opel españa (2015).

[30] ISRA, Paintscan: Paint inspection with in-line deflectometry (2015).

[31] J. Santolaria, J. Velázquez, D. Samper, J. Aguilar, I. Escursell, Sistema de inspección de defectos opel españa (2016).

[32] S. Hofer, J. Burke, M. Heizmann, Infrared deflectometry for the inspection of diffusely specular surfaces, Advanced Optical Technologies 5 (2016) 377 – 387. doi:doi:10.1515/aot-2016-0051.

[33] J. Molina, J. E. Solanes, L. Arnal, J. Tornero, On the detection of defects on specular car body surfaces, Robotics and Computer-Integrated Manufacturing 48 (2017) 263 – 278.

[34] A. Mikkelsen, S. Honningsøy, T.-M. Grønli, G. Ghinea, Exploring microsoft hololens for interactive visualization of uml diagrams, in: Proceedings of the 9th International Conference on Management of Digital EcoSystems, MEDES ’17, ACM, New York, NY, USA, 2017, pp. 121–127. doi:10.1145/3167020.3167039.

[35] J. Blattgerste, B. Strenge, P. Renner, T. Pfeiffer, K. Essig, Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks, in: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments, ACM, 2017, pp. 75 – 82. doi:10.1145/3056540.3056547.

[36] C. Attig, D. Wessel, T. Franke, Assessing personality differences in human-technology interaction: An overview of key self-report scales to

predict successful interaction, in: C. Stephanidis (Ed.), HCI International 2017 – Posters' Extended Abstracts, Springer International Publishing, Cham, 2017, pp. 19–29.

[37] T. Franke, C. Attig, D. Wessel, A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ati) scale, International Journal of Human-Computer Interaction 0 (0) (2018) 1–12. `doi:10.1080/10447318.2018.1456150`.

[38] S. G. Hart, L. E. Staveland, Development of nasa-tlx (task load index): Results of empirical and theoretical research, in: P. A. Hancock, N. Meshkati (Eds.), Human Mental Workload, Vol. 52 of Advances in Psychology, North-Holland, 1988, pp. 139 – 183. `doi:https://doi.org/10.1016/S0166-4115(08)62386-9`.

[39] J. Brooke, "SUS-A quick and dirty usability scale." Usability evaluation in industry, CRC Press, 1996, iSBN: 9780748404605.

[40] G. B. Reid, T. E. Nygren, The subjective workload assessment technique: A scaling procedure for measuring mental workload, in: P. A. Hancock, N. Meshkati (Eds.), Human Mental Workload, Vol. 52 of Advances in Psychology, North-Holland, 1988, pp. 185 – 218. `doi:https://doi.org/10.1016/S0166-4115(08)62387-0`.
URL `http://www.sciencedirect.com/science/article/pii/S0166411508623870`

[41] V. Venkatesh, M. Morris, G. Davis, F. Davis, User acceptance of information technology: Toward a unified view, MIS Quarterly 27 (3) (2003) 425–478.

[42] B. Simeonova, P. Bogolyubov, Evgeny Y., R. Kharabsheh, Cross-cultural Validation of UTAUT: The Case of University VLEs in Jordan, Russia and the UK, Electronic Journal of Knowledge Management 12 (1) (2014) 25–34.

[43] B. D. Harper, K. L. Norman, Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5, in: Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference, 1993, pp. 224–228.

[44] T. Tullis, W. Albert, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008.

[45] C. Davies, HoloLens 2: The four big improvements, accessed 12/02/2019 (2019).
URL https://www.slashgear.com/microsoft-hololens-2-four-key-improvements-comfort-fov-holograms-25567353/

[46] Microsoft, Trimble XR10 with HoloLens 2, accessed 12/02/2019 (2019).
URL https://news.microsoft.com/microsoft-at-mwc19/photos/trimble-xr10-with-hololens-2/