Wearable Visual Intelligence to Assist Humans in Workplaces

Giovanni Maria Farinella[1,*], Antonino Furnari[1] and Francesco Ragusa[1]

[1]Next Vision s.r.l. - Spin-off of the University of Catania, Italy

Abstract

Wearable devices equipped with a camera and a display allow to develop human-centric applications providing specific services able to improve worker's productivity and increase safety in industrial environments. Despite there are different wearable devices in the market, current solutions are mainly focused only on passive augmented reality and real-time remote assistance. This contribution presents Artificial Intelligence technologies for wearable devices provided by NEXT VISION s.r.l. - Spin-off of the University of Catania. In particular, we present three technologies developed by NEXT VISION able to 1) localize humans in unfamiliar environments and support them to navigate the space to reach a specific destination, 2) understanding human-object interactions to provide support to the workers in industrial workplaces during complex procedures of maintenance and 3) interact with humans thanks to a conversational intelligent agent exploiting natural language and the artificial vision to answer questions regarding the surrounding environment.

Keywords

Artificial Intelligence, Egocentric Vision, Wearable devices

1. Introduction

Artificial Intelligence technologies allow to support humans in industrial environments improving workers' safety and their productivity. Several vision systems have been adopted with the aim to support humans visually using fixed cameras (e.g., surveillance cameras) to observe the surrounding environment from third person point of view.

More recently, different wearable devices equipped with a camera and a computational unit have been introduced in both consumers (e.g., Microsoft Hololens 2[1], Nreal Light[2], Vuzix Blade [3] and industrial (e.g., Realwear[4], Vuzix M400 [5]) market.

Differently from fixed cameras, these devices are useful to look at the scene from the user point of view (first person vision) to assist workers in industrial workplaces. These devices can be used without engaging the hands, and can display real-time information thanks to the wearable screen which allow to see holograms placed in the real world through Augmented Reality. Also these devices can provide artificial intelligence services analyzing images and videos acquired from the user's point of view.

Despite these devices are present in the market and their adoption in industrial environments is growing, current solutions are focused mainly on augmented reality technologies and real-time remote assistance without exploiting Artificial Intelligence and Computer Vision capabilities.

This contribution presents Artificial Intelligence technologies for wearable devices provided by NEXT VISION s.r.l. - Spin-off of the University of Catania. The technologies discussed here are focused on industrial use cases with the aim to improve workers' safety and their productivity.

In the following sections, we present three technologies: *NAIROBI*, an Artificial Intelligence assistant able to localize users in unknown environments and provide routes to reach specific objects or points of interest; *NAOMI*, an intelligent assistant able to detect and recognize interactions between humans and objects present in an industrial environments, and to guide them during the execution of complex procedures (e.g., installation or maintenance); *HERO*, a conversational agent able to interact with a worker exploiting natural language artificial vision.

The technologies presented in this contribution are covered by two patents [6, 7]. Please, refer to our web page https://www.nextvisionlab.it/ for further information about the technologies developed by NEXT VISION.

2. Indoor Visual Navigation

The ability to localize workers in indoor environments is important to support humans in performing complex tasks and to improve safety in industrial environments (e.g., for rescue applications). For example, a worker could need to move from side to side in a big factory avoiding dangerous areas with suspended loads as well as areas reserved for the transit of vehicles. If a system

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy

*Corresponding author.

gmfarinella@nextvisionlab.it (G. M. Farinella); afurnari@nextvisionlab.it (A. Furnari); fragusa@nextvisionlab.it (F. Ragusa)

Figure 1: A route to reach the chosen destination shown in Augmented Reality by NAIROBI.

is able to localize humans, then it can advice workers informing them in case are on dangerous areas. In case of emergencies (e.g., fires), the system can localize and guide workers to the closest fire extinguisher or the closest emergency exit helping to get to safety. When an operator is working in an unfamiliar environment, his ability to navigate it is initially reduced, indeed, this can have negative affect to his productivity and decrease his safety. A system able to localize and suggest paths to the user can help to navigate the new unfamiliar environments.

To support workers in these complex scenarios, NEXT VISION developed NAIROBI (Navigating Autonomously Indoor Routes by Observing Building Information), an Artificial Intelligence assistant able to guide the user to reach specific areas (e.g., a specific workbench), or the closest points of interest (eg., fire extinguisher or emergency exit). NAIROBI is able to localize the user in the industrial environment through computational vision allowing the implementation of rescue procedures in dangerous situations for the operator. The intelligent assistant uses Computer Vision and Artificial Intelligence algorithms to localize the worker inside the building from images, compute the best route and to guide the user with a path shown in the display of the wearable device using the Augmented Reality. An example of navigation provided by NAIROBI in an indoor environment is shown in Figure 1. NAIROBI is also able to provide a map (i.e., "You are here" map shown through holograms on the display of the wearable device) of the indoor environment showing the real-time user's position to help users navigate a building (Figure 2).

NAIROBI is multiplatform and not device specific. It can be integrated into new wearable device equipped with a camera. Currently, it has been developed on Microsoft Hololens 2 device and on iOS and Android smartphones. For both settings, the user is able to receive augmented reality information to navigate the unknown environment.

3. Human-Object Interaction Understanding

Human-object interactions algorithms are fundamental to understand how humans relate with the surrounding environments while achieving their goals. In particular, in industrial places, the usage of specific objects and machines need an initial training phase, often with the cooperation of another worker which has more experience, to transfer specific skills to the new workers. For example, a testing procedure of a machine, may require the usage of an electric panel composed of several buttons that must be pressed following a specific sequence. These complex operations often require people which have an expertise on the field, indeed a company has significant costs in terms of time and money for providing specialized training courses. Considering a complex procedure, even an expert operator could consult a manual during the execution of the procedure, which can slow down the task or it can lead to making mistakes.

To support workers during the execution of complex procedures, NEXT VISION proposes NAOMI (Next Active Object for Monitoring Interactions) [8, 9], an Artificial Intelligence assistant able to understand and monitor interactions between humans and the objects present in the

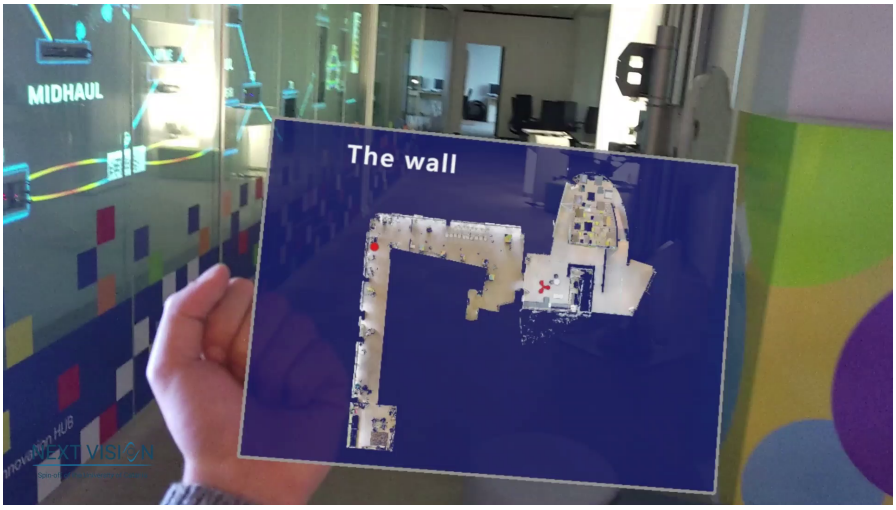

Figure 2: "You are here" map which shows where the user is in the building.

Figure 3: On the left, a procedure of maintenance on an electric panel where NAOMI suggest how interact with the objects observing the surrounding environment from the first person point of view. On the right, an example of human-object interaction recognized by NAOMI in a industrial workbench.

surrounding environment. NAOMI is able to recognize the objects and the interactions with them and to guide the user during the execution of complex procedures suggesting the different steps to perform and understanding if each step has been performed correctly. Moreover, it is able to notify alarms if an interaction is dangerous (e.g., electrick risk) avoiding mistakes which could cause a machine breakdown or sending information to IoT systems present in the environments (e.g., to immediately turn off the electric power for safety). All the information about the human-object interactions, predicted by the artificial intelligence and computer vision algorithms composing NAOMI, can be aggregated and processed to provide statistics about the interacted objects (e.g., usage time, maintenance date, electric consumption, temperature). These information are useful to understand how the worker interacts with the surrounding environment, for example to check the wear of a specific industrial tool in order to carry out preventive maintenance, or if a specific machine has been used correctly and safely. Furthermore, the obtained information can be used to check the quality of procedures.

Object recognition and interaction monitoring are addressed thanks to the artificial intelligence algorithms developed by Next Vision. The communication with the users through the Augmented Reality is done thanks to the Microsoft Hololens 2 capabilities (Figure 3).

4. Human-Machine Interaction

Artificial Intelligence allows to build systems which enable automatic conversation between humans and machines using natural language. These conversational agents could be integrated in more complex systems

with the aim to support humans where they live and work. In particular, user interactions with the surrounding environment and the objects present in an industrial workplace require knowledge of specific information and procedures obtained through the training of operators with specific courses or cooperation with an expert figure. Considering a complex environment, specific information such as "what is the testing procedure of the electric board", "what is the next step of this maintenance procedure", "how to calibrate the oscilloscope", or "how to use this specific item", may not be immediately available and require continuous interactions with laboratory technicians or qualified workers. These information are also more complex to obtain when a visual representation of the surrounding environment is needed to correctly answer to the worker's question avoiding language ambiguity or diversity (for example "what is the object in front of me?").

NEXT VISION developed an artificial intelligence technologies to support operators which need to obtain information during procedures. HERO (Human Expertise Replication from Observation) [10] is an artificial assistant that through the interaction with a conversational agent equipped with artificial vision algorithms, is able to receive questions by workers using natural language, images and videos acquired from the user's point of view through wearable devices. HERO is able to receive the human input, which can be represented by text or audio signals, understand the questions processing the input with the artificial intelligence algorithms as well as the surrounding environments analyzing the images/videos acquired from a camera (e.g., head-mounted camera or smartphone camera) to disambiguate the questions asked by the user. HERO provides the most reasonable answers to the questions performed by users. An example is shown in Figure 4 where the user asked a question to HERO using the chat. The text has been generated using a speech-to-text module. Note that to disambiguate the question "How to use this objects?", HERO asked to the user to send an image of the surrounding environment, understanding that the user is referring to the iron soldering. Then, HERO provided the answer with the information on how to use that specific object.

The conversational assistant is able to cooperate to the other intelligent assistants developed by NEXT VISION and discussed in previous paragraphs. For example, the user can ask "guide me to the closest fire extinguisher" or "start the wizard for the object I'm looking at" to receive support from NAIROBI and NAOMI.

Figure 4: The figure shows a conversation with HERO using a chat where text signal have been retrieved by a speech to text module. Note that to disambiguate the worker's question, an image has been sent to the artificial assistant representing the object for which support has been asked.

5. Conclusions

This manuscript presented technologies developed by NEXT VISION - Spin-off of the University of Catania. The proposed technologies make use of Artificial Intelligence algorithms and wearable devices equipped with a camera to improve workers productivity and to increase their safety in industrial workplaces. In particular, three artificial intelligent assistants have been presented to support workers in different manners. NAIROBI is able to localize the user and provide routes to reach a specific

point of interest, NAOMI understands human-object interactions providing instructions on how to perform a procedure of maintenance, and HERO, a conversational assistant able to interact with workers using the natural language and artificial vision.

References

[1] Microsoft hololens 2, https://www.microsoft.com/en-us/hololens, Last accessed on 2023-04-01.

[2] Nreal light, https://www.nreal.ai/light/, Last accessed on 2023-04-01.

[3] Vuzix blade, https://www.vuzix.com/products/vuzix-blade-2-smart-glasses, Last accessed on 2023-04-01.

[4] Realwear, https://www.realwear.com/, Last accessed on 2023-04-01.

[5] Vuzix m400, https://www.vuzix.com/products/m400-smart-glasses?variant=41517448659110, Last accessed on 2023-04-01.

[6] G. M. Farinella, A. F. F. Ragusa, E. Ragusa, G. Sorbello, A. Lopes, L. Santo, M. Samarotto, B. Scarso, E. Scarso, Metodo di assistenza virtuale relativo dispositivo e sistema, 2022.

[7] G. M. Farinella, G. Signorello, A. Furnari, S. Battiato, E. Scuderi, A. Lopes, L. Santo, M. Samarotto, G. Distefano, D. G. Marano, Metodo integrato con kit indossabile per analisi comportamentale e visione aumentata, 2020.

[8] R. Leonardi, F. Ragusa, A. Furnari, G. M. Farinella, Egocentric human-object interaction detection exploiting synthetic data, in: International Conference on Image Analysis and Processing (ICIAP), 2022. URL: https://iplab.dmi.unict.it/EHOI_SYNTH/.

[9] M. Mazzamuto, F. Ragusa, A. Resta, G. M. Farinella, A. Furnari, A wearable device application for human-object interactions detection., in: International Conference on Computer Vision Theory and Applications (VISAPP), 2023.

[10] C. Bonanno, F. Ragusa, R. Leonardi, A. Furnari, G. M. Farinella, HERO: An artificial conversational assistant to support humans in industrial scenarios, in: International Conference on Signal Processing and Multimedia Applications (SIGMAP), 2022.