S.A.M.I.R.: Supporting Tele-Maintenance with Integrated Interaction Using Natural Language and Augmented Reality

Fabio de Felice, Anna Rosa Cannito, Daniele Monte, Felice Vitulano

S.A.M.I.R.: Supporting tele-maintenance with integrated interaction using Natural Language and Augmented Reality

Fabio De Felice, Anna Rosa Cannito, Daniele Monte, Felice Vitulano

Exprivia S.p.A. Molfetta, Italy
Fabio.DeFelice@exprivia.com

Abstract. Remote maintenance is becoming a crucial aspect in the industrial life cycle, especially in context where different countries, languages, and time zones are involved. Supplying on-site operators with smart tools to rapidly and easily request support can boost maintenance procedure execution times and solving of unexpected problems, can reduce the number of interventions and can speed up novice training, resulting in cost reduction and equipment use maximization. Authoring tools are needed to properly create new multimedia content and to process, organize and index legacy company knowledge, regardless of type. In this paper the SAMIR solution is presented, it defines a SaaS platform integrating Augmented Reality and Natural Language Processing to supply an effective support during maintenance intervention together with authoring tools to create multimedia content and procedures.

Keywords: Augmented reality, tele-maintenance, Architecture for AR services, Middleware, Natural Language Processing, Natural Language Interaction

1 Introduction

Today, the interaction between production and maintenance is gaining importance in the "Industry 4.0" context. Proper maintenance management helps maximize utilization of equipment, to maintain a constant level of productivity, to reduce the number and timing of interventions, leading to costs reduction [1]. The tele-presence approach for remote maintenance, also because of the current pandemic emergency, is gaining the attention of companies in the manufacturing industry. In this context, the use of Virtual Reality/Augmented Reality (AR/VR) combined with multimodal Human-Machine Interaction, such as using gestures or natural language with text or voice [2][3], allows effective reduction of execution times. Stress and weariness can be reduced enhancing security [4] and human error [5] and intervention costs can be minimized. Knowledge sharing of specialized skills is also enhanced facilitating continuous learning of operators.

In this paper we present SAMIR (Smart Assistance for Maintenance with Intelligent Research), developed for the Italian project FINDUSTRY4.0 – Future Internet for Industry 4.0 (PON FESR 2014-2020 project code F/080002/01-04/X35, from MISE– Ministero dello Sviluppo Economico), a SaaS platform aimed at supplying integrated tools to support on-site operators during maintenance, remotely and in real time.

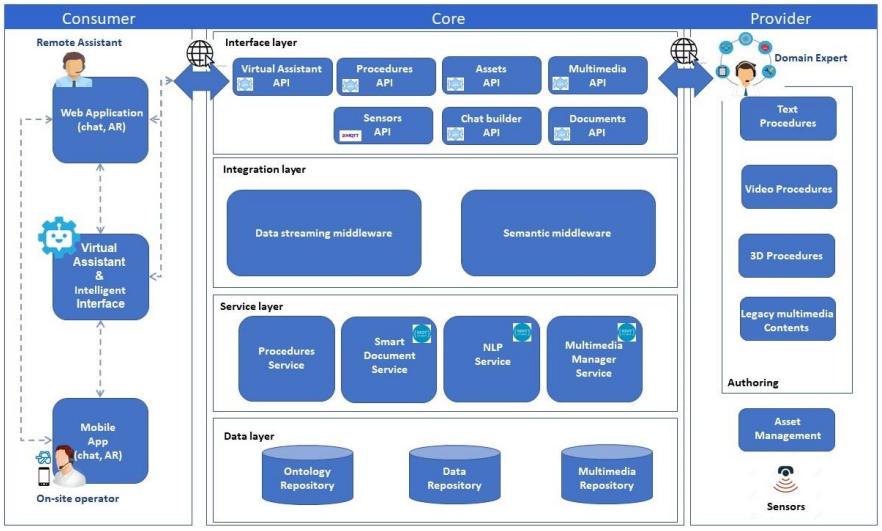

Fig. 1. The SAMIR architecture

SAMIR also defines authoring tools to easily create multimedia contents (video, text, image, 3D models) and uploads legacy company content.

On-site operators interact with the system by using natural language during dialogue with remote human experts, by means of chat or video call, and with an AI Virtual Assistant by means of chat or Voice/Vocal User Interface (VUI). The adoption of VUI in conjunction with normal chats allows keeping the hands and eyes focused on the task at hand, speeding up the intervention, as the speech recognition process is at least three times faster than keyboard input on smartphones [6]. Remote human experts can enhance their explanation by the use of 3D symbols, predefined or drawn in freehand, visualized by the operator in AR mode, moreover, AR can be used to represent entire maintenance procedures. The closer the interaction between human and real object, as in the maintenance case, the more virtuality should be non-invasive and constructive by enriching the real world with virtual contents, for this reason the AR approach is preferable to VR [7][8].

The rest of the paper is organized as follows: in section 2 an overview of the main architecture components is given with focus on interaction implementation, in section 3 the interaction design is described, while in section 4 and 5 a real use case and conclusions are reported respectively.

2 SAMIR Architecture

In this section the SAMIR architecture, depicted in figure 1, is described highlighting how contents are created and consumed. The architecture is organized in three main areas, the Provider, including tools used by the domain expert to create and organize

contents, the Consumer, with tools to request and use contents, and the Core implementing the back-end, which integrates processing services and stores data.

The Provider tier contains applications to create maintenance procedures in different formats (text, video, 3D) by means of authoring tools and to upload legacy multimedia contents to populate the solution company knowledge. Also, sensors monitoring the equipment fall in this area as they provide live data that can be visualized by the operator during intervention or can be used to trigger condition-based maintenance procedures.

On the Consumer side there are three main components, the mobile app used by the on-site operator, running on smartphone or tablet, the web application used by the remote expert, and the Virtual Assistant which provides an intelligent interface to interact directly with the search engine of the system. The Core tier is based on a three-tiers architecture, extended with an integration layer to decouple interface and service layers in order to promote extensibility and modularity. In the Service Layer the Multimedia Manager Service (MMS) is responsible for the semantic segmentation and indexing of uploaded videos and for successively retrieving them as results of user queries. The Smart Document Service (SDS) handles the textual contents, both generic and procedures, and retrieves them when queried, while Procedures Service handles both video and 3D procedures. The NLP Service is responsible for processing user requests in natural language both syntactically and semantically to create a formal query to be used in input to the MMS and SDS.

3 Interaction design

Two types of interactions take place, Provider interaction, involving authoring activities, and Consumer interaction, involving how operators interact with each other and with the system. In Provider interaction, to create video or text procedures, intuitive interfaces are available; videos can be subdivided into tasks and steps, analogously a given text can be annotated to create an index and to mark section of interests. This preprocessing step allows direct access to the portion of interests of a given procedure. Generic contents can also be uploaded, after tagging them, for free research.

In the Consumer interaction the on-site operator is first guided by the Virtual Assistant to the required information, if this support does not satisfy the user needs the dialogue can switch to a human operator. The Virtual Assistant tries to understand user requests by grouping them into three categories: Planned Procedures, Live Data and Contents Research. The user asks for Planned Procedures during scheduled maintenance to find a particular procedure about particular equipment, the assistant guides the user by showing the procedures available for that equipment and the format available for a given procedure (text, video or AR), at the end of this dialogue the procedure is shown to the user in the chosen modality. In an analogue dialog, Live Data requests show sensor data and other quantitative measurements to the user. Finally, Contents Research allows the user to submit a generic request using natural language, this is the case when NLP comes in by extracting the key words needed to make formal queries to the MMS and SDS to retrieve contextual information based on videos and text.

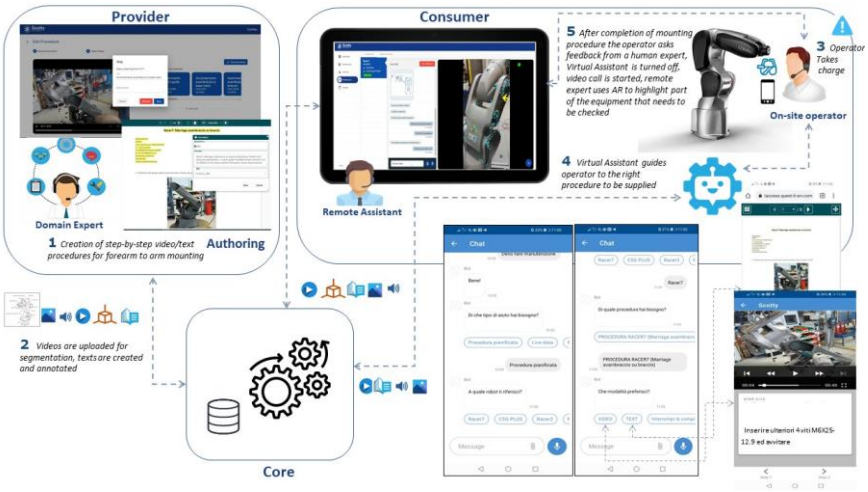

Fig. 2. Forearm to arm mounting scenario

4 Use Case

The ongoing development of the solution is constantly tested within a real use case. The application has been deployed in the industrial automation domain with the collaboration of the project partner Comau [9], a leading company in the field. Maintenance procedures about robotic equipment have been created and some user experiences have been conducted as depicted in figure 2. In this image the case of the Racer7 robot is depicted concerning the procedure of robot forearm to arm mounting.

The implementation integrates two existent solutions, while the integration back-end is developed from scratch in Exprivia [10]. The Scotty application from Wideverse [11] is used for AR interaction and authoring for video and 3D procedures, while Algho Virtual Assistant from Quest-it [12] is used for VUI, dialogues and NLP, with a proprietary algorithm [13].

5 Conclusions

In this paper the SAMIR solution has been presented with focus on integrated interaction combining NLP and AR supporting on-site operators with effective contextual information research during maintenance operations. Multimodal procedures can be supplied in real time created with appropriate authoring tools used by domain experts. Also, live data from sensors and free contents research based on requests submitted in natural language are possible. The development is in progress and constantly tested in real use cases in collaboration with a leading company in the domain of industrial automation. The interest from a real stakeholder proves how such approach is a real need

from the industry field, and allows better understanding of final user requirements in order to realize a quality product.

References

1. Mourtzis, D., Zogopoulos, V., Vlachou, E.: Augmented Reality Application to Support Remote Maintenance as a Service in the Robotics Industry. Procedia CIRP (63), 46-51 (2017).
2. Flatt, H., Koch, N., Röcker, C., Günter, A., Jasperneite, J.: A context-aware assistance system for maintenance applications in smart factories based on augmented reality and indoor localization. In: 20th Conference on Emerging Technologies & Factory Automation (ETFA), IEEE, Luxemburg (2015).
3. Gorecky, D., Schmitt, M., Loskyll, M., & Zühlke, D.: Human-machine-interaction in the industry 4.0. In: 12th IEEE International Conference on Industrial Informatics (INDIN). (2014), pp. 289-294, IEEE, Porto Allegre, Brazil (2014).
4. Kaasinen, E, Schmalfuß, F., Özturk, C., Aromaa, S., Boubekeur, M., Heilala, J., Heikkilä, P., Kuula, T., Liinasuo, M., Mach, S., Mehta, R., Petäjä, E., Walter, T.: Empowering and engaging industrial workers with Operator 4.0 solutions. Computers & Industrial Engineering. Elsevier, (2020).
5. Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., Tecchia, F.: Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. Interactive Learning Environments 23(6), 778–798, (2015)
6. Ruan, S., Wobbrock, J. O., Liou, K., Ng, A., Landay, J. A.: Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 1, no. 4, pp.° 1–23. (2018).
7. Krupitzer, C., Müller, S., Lesch, V., Züfle, M., Edinger, J., Lemken, A., Schäfer, D., Kounev, S. & Becker, C. A Survey on Human Machine Interaction in Industry 4.0. Human Computer Interaction (1), 1-45 (2020).
8. Eschen, H., Kötter, T., Rodeck, R., Harnisch, M., Schüppstuhl, T.: Augmented and Virtual Reality for Inspection and Maintenance Processes in the Aviation Industry. Procedia Manufacturing (19), 156-163 (2018).
9. Comau homepage, https://www.comau.com/it last accessed 2021/04/20
10. Exprivia homepage, https://www.exprivia.it/it/ last accessed 2021/04/20
11. Wideverse homepage, https://www.wideverse.com/it/home-it/ last accessed 2021/04/20
12. Algho homepage, https://www.alghoncloud.com/ last accessed 2021/04/15.
13. Melacci, S, Globo, A., Rigutini, L: Enhancing Modern Supervised Word Sense Disambiguation Models by Semantic Lexical Resources. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (2018).