Computers in Industry 98 (2018) 68–79

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Converting maintenance actions into standard symbols for Augmented Reality applications in Industry 4.0

Giulia Wally Scurati[a], Michele Gattullo[b], Michele Fiorentino[b], Francesco Ferrise[a,*], Monica Bordegoni[a], Antonio Emmanuele Uva[b]

[a] Politecnico di Milano, Department of Mechanical Engineering, Via La Masa 1, 20156 Milano, Italy
[b] Politecnico di Bari, Department of Mechanics, Mathematics and Management, Viale Japigia 182, 70126 Bari, Italy

ARTICLE INFO

Article history:
Received 30 November 2017
Received in revised form 22 January 2018
Accepted 22 February 2018
Available online xxx

Keywords:
Augmented Reality
Industry 4.0
Maintenance

ABSTRACT

The evolution of technical documentation in the age of Industry 4.0 is going towards the use of visual manuals, in particular exploiting Augmented Reality (AR) technology. Traditional manuals are rich of text instructions that in AR applications are not advisable. In fact text occludes the real scene behind and it is an issue for the translation. For this reason, we propose to create and adopt a controlled and exhaustive vocabulary of graphical symbols, to be used in AR to represent maintenance instructions. In particular, in this work we identified the most frequent maintenance actions used in manuals, and converted them into graphical symbols. Then, we made an elicitation of the symbols designed and created different candidate vocabularies of symbols basing on the criteria found in literature of *guessability* and homogeneity. Moreover, the vocabularies had to respect two constraints: *conflict set* and *reversibility*. Finally, we identified the best of symbols and integrated this one in a real AR application for remote maintenance.

1. Introduction

Industry 4.0 is an emerging paradigm for the increasing digitisation and automation of the manufacturing environment as well as the creation of a digital value chain to enable the communication between products and their environment and business partners [1]. This digitisation process regards the entire production system, and in particular the maintenance of systems and machines with all the relative technical documentation. The Industry 4.0 paradigms are becoming the driver to the development of a new generation of digital technical instructions, mainly based on the new display technologies such as Augmented and Virtual Reality that exploit more graphical and visual elements, whose role becomes primary.

Studies have shown that visual instructions are cognitively favourable by people as they are easier to comprehend and remember than text information [2–4]. This trend is also confirmed by studies on digital natives that show how children who grow up immersed in digital media think and learn differently from those who grew up with printed text [5]. In particular, they prefer graphics to text [6]. These studies involved the education system where we are assisting to a dramatic shift from text-based to multimedia educational resources [7].

In traditional manuals, usually available in paper and digital versions, text is prevalent and occasionally supported by images. However, there are other emerging typology of cloud-based manuals in which this relationship is inverted: pictures are aided by graphic contents and minor text descriptions. It is the case of instructional websites (e.g. IFixit [8], Instructables [9]). In a similar way, the technical documentation in Augmented Reality (AR) shows the real environment augmented with geo-located graphic contents and eventually, concise text descriptions. In AR applications, text reduction is not simply preferable, it is mandatory because text boxes, as well as large images, may occlude user's sight.

Text reduction is also demanded by a globalised industrial world and economy: in this context, minimising text is a way to overcome language and cultural barriers, as it happens in other fields [10,11]. However, text instructions have a long time and established tradition in the industry, and they have been regulated over the years through the development and use of controlled languages (e.g. STE: Simplified Technical English [12]), which define detailed and precise standards regarding vocabulary, grammar and syntax. While in text based instructions the standardisation process is ongoing, the Augmented Reality technical documentation is at an early stage and no guidelines

* Corresponding author.
E-mail address: francesco.ferrise@polimi.it (F. Ferrise).

https://doi.org/10.1016/j.compind.2018.02.001

are available. However, the design principles of Industry 4.0 [13] orientate the developers in the definition of these guidelines.

The first step in defining those guidelines is selecting the type of augmentation which the manual should display. 3D models, which can be superimposed on the real model and integrated with the environment, represent one of the most popular methods to visualise actions in AR. They can also be animated, showing how to accomplish a change of state of the objects, which may be especially useful when complex movements are required. However, they cause occlusion of the real world [14,15] require a strong authoring effort [16,17], and need time to be understood by users [18,19]. Conversely, simple 2D elements as graphical symbols could be a better alternative. Graphical symbols are easier to recognise and comprehend than other complex 2D elements such as technical drawings, which may be quite elaborated and detailed. Moreover, symbols cause a minor occlusion and do not require a precise superimposition since they can be displayed on a fixed place on the Graphical User Interface (GUI). Finally, a documentation mainly based on symbols would be less dependent on the product, limiting the problems related to product updating and customisation. This adaptability can significantly diminish authoring time and effort.

Despite those advantages, the use of graphical symbols in technical documentation is not so widespread because it introduces a new research question that is addressed in this work: how to define a vocabulary of 2D symbols to convey technical instructions?

Most research works reported in literature do not adopt an approach specifically based on graphical symbols. Rather symbols are mixed with 3D models and text instructions, so that not all the actions to accomplish are conveyed through symbols. Hence, to our knowledge, no work in literature is focused on determining an extensive vocabulary of maintenance actions to be converted into symbols. In the following sections, we present a methodology to convert maintenance actions into graphical symbols. Despite it would be impossible to entirely cover technical documentation, it is possible to convert most of it. Using a limited set of symbols would also allow users to learn and remember them. The first step of this methodology, described in Section 3.1 was the analysis of a wide set of instructions to select the actions that could be converted into symbols. Then, we designed a set of 2D symbols referring to ISO symbols, as described in Section 3.2. Finally, we performed an elicitation of the symbols, described in Section 3.3, and from the analysis of the elicitation results, showed in Section 4, we defined a potential vocabulary that was integrated in a real AR application for maintenance, used as case study, as described in Section 5.

2. Related works

The main advantage of using Augmented Reality for maintenance and assembly instructions is related to the intuitiveness associated to this innovative displaying technique, because information could be displayed directly on the object it refers to [20–22]. However, this implies a novel approach in the authoring of technical documentation. The way information is presented in AR, has a crucial impact on user experience. The two key issues related to the presentation of technical information in AR, are the modality to convey instructions (text labels, CAD models, 2D symbols) and the management of the information displayed in this novel way.

Regarding the first issue, since from the first works related to AR for maintenance and assembly, in the early nineties, there has never been a single approach prevailing on the ways instructions are conveyed.

Especially in the first AR applications [23–25] text labels were the primary element used to explain maintenance actions. One of the major problems related to the use of labels is that they may overlap each other and also hide relevant objects present in the scene, as reported by Ong et al. [21]. There was also a huge amount of 3D models displayed, most of which were useless because they just covered the real objects. The problem of occlusion is not limited to text labels but also involves 3D models.

However, many research works were focused on the problem of occlusion [14,15]. In particular, Bell et al. [26] studied the problem of occlusion in the context of authoring of technical information. In fact, they considered both the constraints of visibility (if and in which cases an object can be occluded) and priority (which objects and constraints are more important).

The main concern about the use of 3D models is the authoring time and effort required to place them in an AR scene. A solution could be that of automating the placement of CAD models in the real scene [16,17]. However, the tracking accuracy is not still enough in some cases to make this process reliable and scalable.

A different approach to the one of showing a 3D model to indicate an object change of state (e.g., a component to be removed), is that one of communicating to the user the action to perform in order to obtain that change of state. A possible way to do it is showing hands 3D models performing gestures on the object, as proposed by Yin et al. [27] or 3D models of the tools used to perform the operations [27,28]. Even if the authoring effort is reduced since they do not require a precise placement in the scene and are independent from the product, we can reasonably expect similar issues related to occlusion. Furthermore, the gesture alone does not provide information about the outcome. For instance, a screwdriver turning does not communicate if the aim is unscrewing or loosening. Additional text instructions would be required.

Another method is the one of conveying maintenance actions using symbols. Symbols are simple 2D visual features and can be displayed in relatively small dimensions on the GUI. Their use would have the following advantages compared to the other kinds of augmentations described so far: minor view management related issues (e.g. occlusion and cluttering), lower authoring efforts, and lower mental load. The first two advantages can be easily understood from the discussion regarding the previous cited works. As to the reduction of mental load, 2D elements require less time to be understood, diminishing the time required to complete the task. Conversely, the user needs more time to perceive and process the 3D data: the distorted perception of distances, positions and angles of 3D interfaces may cause uncertainty, delaying the user's action [18]. Phatomaree et al. [19] conducted a user study on an assembly task using 2D and 3D puzzles. They noticed that the same task to be accomplished using a 3D or 2D frame, requires a significantly shorter completion time for the latter.

The second issue we mentioned is related to the management of the information displayed in an AR application for technical documentation. Information management is important to build effective interfaces: Hollerer et al. [29] suggest to make user interfaces as clear and obvious as possible. Zarraonadia et al. also suggest that the number of symbols depicted at the same time should be reduced to optimise user performances [30]. Following this strategy implies that the use of virtual elements should be limited to only those needed. Although the effectiveness of presenting only relevant information is undeniable, it may be more complicated to define which instructions are both necessary and sufficient to provide the user, and how this affects not only the quantity but also the form of the augmentations displayed. Possible factors for information filtering are task difficulty, user skills, and the operator's position in the work environment. Radkowski et al. [18] make the hypothesis that visual features should differ basing on tasks degree of difficulty for assembly operations. Webel et al. [31] speculate different levels of guidance