Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

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

This is a post print of the following article

Original Citation:
Converting maintenance actions into standard symbols for Augmented Reality applications in Industry 4.0 / Scurati, G.W., Gattullo, M., Fiorentino, M., Ferrise, F., Bordegoni, M., Uva, A.E.. - In: COMPUTERS IN INDUSTRY. - ISSN 0166-3615. - STAMPA. - 98:(2018), pp. 68-79. [10.1016/j.compind.2018.02.001]

Availability:
This version is available at http://hdl.handle.net/11589/149634 since: 2022-06-08

Published version
DOI:10.1016/j.compind.2018.02.001

Publisher:

Terms of use:

(Article begins on next page)

19 August 2026

Link to publisher version with DOI

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

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

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.

Keywords:

Augmented Reality, Industry 4.0, 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, 3, 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 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, 21, 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, 24, 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 depending on the user's knowledge, which would require different amount of information and different visual features. The stronger guidance level includes 3D models and arrows, the softer one only highlighting the object with a circle, so the former uses visual features to provide instructions, while the latter only supports the object's localisation. In the system presented by Feiner et al. [24] the augmentations change as the user approaches the object, focusing on localisation when he/she is far and on descriptions and instructions when he/she is close.

The presented works confirm the need for a new approach to the management of technical information using AR. This approach is based on presenting only relevant pieces of information, and is suitable to the design principles of Industry 4.0 [32]. However, there is not a proved methodology that defines how to display information in AR interfaces, taking also into account the process of authoring. In fact, the first AR test cases made large use of 3D models that could seem the most intuitive way to communicate instructions. However, they present many drawbacks about authoring effort and user perception. Text labels occlude user sight of the real scene, and the novel trends push in the direction of text reduction. Then, based on these considerations, we propose a strategy based on the use of 2D symbols to convey instructions, reducing the amount of text and requiring less authoring effort. While arrows are commonly used, there is not much work in literature that specifically focuses on symbols.

A graphical symbol is a visually perceptible figure with a particular meaning used to transmit information independently of language (definition from

ISO 17724:2003) [33]. In [34] are defined the policy guidance and the procedure for the preparation and the acceptance of all new graphical symbols to use in European Standards (EN). Standardized symbols are intended to be widely understood, unambiguous, language neutral and unique. The proposed graphical symbols must be applied to ISO Technical Committee (TC) ISO/TC 145 (”Graphical symbols”) or ISO/TC 10 (”Technical product documentation”). When the graphical symbol proposed has been agreed it receives a registration number and is published. The choice of graphic symbols to assign to action verbs is not trivial. In fact, in the literature, we can find symbols used in AR to explain the action to carry out on the object. However, because there are not common standards, the same symbol could express different meanings: for example, arrows are used to indicate movements in a specific direction or forces to exert [31, 35], but are also used as pointers to focus the user’s attention on an object or area [18, 35]. In such cases, ambiguity could be easily avoided by conferring different appearances to the arrows, according to their meaning: ISO 80416 [36] assigns precise visual attributes to each kind of arrow, but most works in literature do not consider ISO regulations. Figure 1 shows a particularly clear example, as the Belgian head arrow is used as a command to install a component, while according to ISO 80416, it represents the movement of people in public signs.

For the reasons we explained in the beginning of this section, we started analysing ISO symbols standard [39].

ISO symbols have a recognisable style: they are commonly represented using black contours or shapes on a white background, unless they have a specific meaning (e.g. a yellow background stands for warning, green for

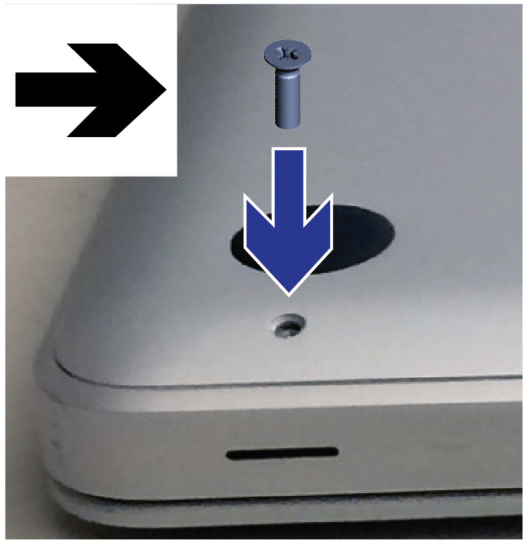


Figure 1: An example of use of arrows in AR applications (adapted from [37] and [38]). A Belgian head arrow, which according to ISO 80416 represents the movement of people in public signs, is used to indicate how to install a component.

safety, blue for instruction). In a previous work, the authors [40] integrated in an AR application, a set of 2D symbols for frequent maintenance/assembly actions: unscrew, screw, location indication, warning, disassemble and assemble. However, these symbols were not taken from a standard vocabulary so that they could be unquestionably understood by anyone. The creation of a set of standard symbols is a key factor for their use in real industrial AR applications. In this work, we propose a methodology for the creation of this vocabulary, starting from the analysis of existing technical documentation to design symbols representing maintenance actions.

3. Design and methods

As described by the authors in [40] and shown in Figure 2, the methodology we propose is the following:

- Analysis of technical documentation: we select and analyse technical documentation with the aim of identifying the most important maintenance actions. We analyse text instructions since the text is the first and most widespread mean to provide instructions in the industrial field. All the verbs used to describe the actions should be approved by the STE vocabulary. Therefore, in case they are not, they need to be translated and then ranked according to their frequency.

- Analysis of ISO symbols: we analyse existing symbols used in the industry to verify if symbols corresponding to the actions already exist, and to obtain design guidelines based on ISO symbols visual attributes. As our aim is creating a standard vocabulary of symbols, we refer to the most widespread standard already in use, ISO graphical symbols for use on equipment [39]. Although each company may have its own standards, our objective is the one of allowing different companies worldwide, and operators with various cultural and technical backgrounds, to communicate.

- Symbols design and test: symbols are designed according to the guidelines obtained in the previous step, and tested using an elicitation test to evaluate the level of consensus among participants. Elicitation studies are an established practice for participatory design of interfaces' commands which can improve usability [41].

- Definition of symbols vocabulary: we create a series of symbols vocabularies considering the level of consensus among participants, and defining additional criteria and constraints. Those result from considerations which emerged after analysing the test results.

- Integration of vocabulary with an AR application: the vocabulary that we consider to globally satisfy all those requirements is finally integrated into an existing AR collaborative system.

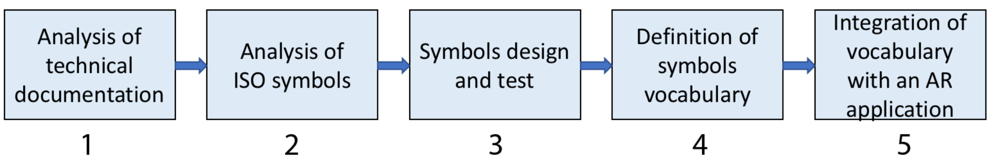


Figure 2: The methodology to transform traditional manuals into symbolic manuals suitable for AR applications

3.1. Analysis of technical documentation

The first step of our methodology is the selection of an adequate technical documentation to be analysed. We decided to analyse maintenance manuals provided by iFixit [8] because we could easily access to a large database of case studies and procedures. IFixit is a global DIY (Do It Yourself) community of people helping each other repairing things, ranging from automotive to game consoles. More than 25000 manuals can be consulted on iFixit. As it is a wiki-based site, anyone can create and share a repair manual for a device, and anyone can also edit the existing set of manuals to improve them. iFixit provides some guidelines to create manuals, requiring sharing on topic information, use simple sentences and correct terms.

However, as maintenance and repair instructions on iFixit are not produced by technical writers, the manuals do not use a controlled language. The controlled language which we refer to in this work is ASD Simplified Technical English (STE), which provides writing rules and a vocabulary of terms approved and alternative suggestions for not approved terms (e.g. "place" is translated into "put", "rotate" into "turn").

We selected 110 maintenance and repair manuals regarding a wide range of procedures. Industrial maintenance includes a variety of processes performed on several components and on both mechanical and electrical parts. Since there are not industrial maintenance cases reported on iFixit, we selected a number and typology of procedures we considered to be the most comparable to the industrial ones. We scanned the manuals' webpages through a Python algorithm to identify all the verbs corresponding to maintenance commands, and we counted them. The algorithm analyses the parts of speech to identify verbs. To select action verbs, we considered verbs in the imperative tense, and verbs in the progressive tense. In fact, it is common using this tense to express the way an action should be accomplished (e.g., remove the component by pulling it). The verbs are then saved into a text file, reporting their frequency. After obtaining all the text files, they were merged using another algorithm in Matlab, to sum up all the equal verbs, ranking them according to the number of counts. The data we obtained were then cleaned, excluding all the verbs related to webpages (e.g., comment, edit, buy) from the list. Phrasal verbs were also discarded, as it is not possible to automatically identify their meaning. Other verbs not expressing an action (e.g., start, see, follow) were excluded as well. The remaining ones were then compared to the STE dictionary.

All the verbs that were not STE approved, were converted into the approved forms, as suggested by the STE manual. In this way, a unique symbol can be created for verbs with the same meaning, and it is associated with the approved verb. Among the 7178 verbs analysed, only 57.6% of them were approved as STE verbs. Once all the verbs were converted, it was possible to sum up official and unofficial synonyms. Some verbs were synonyms even among the approved terms. For instance, "assemble" and "attach" verbs have similar definitions. In this case we checked their reversibility: while "disassemble" is approved as well, the verb "detach" is not. Hence, we preferred the term "assemble" to the term "attach".

At the end of this analysis, we obtained a list of 21 verbs. We clustered those verbs in high-level and low-level ones. Low-level verbs describe actions which require one step or simple movements. Conversely, high level verbs are those actions requiring more than one step or complex movements. "Turn", "lift", "push/pull" are low-level verbs, expressing simple gestures. "Tighten/loosen", "open/close", "cut", "fill", "drain" and "clean" are low level verbs as well, since their outcome and procedures to follow are clear. "Install/remove", "assemble/disassemble", "replace", "move" and "hold" are high-level verbs. Low-level verbs can express how to accomplish a high-level action (e.g. pull the component to remove it). This means the user would have more than a choice to provide instructions. If the action to perform on the object is intuitive, he/she will be able to use a high-level verb. Conversely, if understanding how to obtain a certain change of state in the object is complicated he/she will prefer using a low-level verb. This may occur when an operator has to deal with conventions, which are different from the ones he/she knows. For instance, generally, it would not be necessary to explain how to remove a screw: the verb "remove" would be an effective command. However, if the operator is used to ISO screws and he/she needs to remove an American one, it would not be as immediate. In this case, it would be more appropriate to instruct the maintainer with "turn clockwise".

3.2. Graphical symbol design

By examining a sample of ISO symbols [39], we could recognise some common traits regarding the use of thicknesses and contours or black shapes:

- When possible, contours are preferred to full black shapes and a single type of thickness is used;
- When needed (e.g. to better distinguish two different objects), a second type of thickness is used, which is double the first one in order to be clearly distinguished;
- Fully black shapes are used when they make an object more clearly recognisable than contours (e.g., the figure has small details). In fact, even the lighter one is quite bold, as ISO symbols should be recognised by a certain distance and in small dimensions.

We considered those points as guidelines to design symbols representing maintenance tasks. Symbols were designed by a group of four experts in the fields of technical drawing, industrial design, and maintenance. In the designing phase we took into account the familiarity of users with the proposed symbols. When possible, we included visual features already in use in other symbols (e.g., a wrench, a hand), so that the users could be already familiar with them, and because their perceptual quality had already been proved. Many symbols were inspired by existing ISO symbols. In fact, in most of the cases there was no symbol representing a specific action in general, but there was more than one for the action accomplished on specific objects. For example, while there is no symbol for the task "remove", there is one for "remove the battery", and another one for "remove used mastered container".

20 symbols were designed for each action to represent. Those symbols differed from each other considering both the visual features present in the symbols, and the way those elements were represented. By visual features we mean elements that may compose a symbol, which may vary in the following ways:

- Type of object in the symbol: simple, i.e., defined with less than three primitives as line, triangle, square, circle; complex, i.e. defined with three or more primitives; real object; no object.
- Presence of an anthropomorphic element (e.g., a hand) in the symbol: yes or not.

The way those elements are represented may vary according to other parameters:

- Direction of the arrow/s in the symbol: vertical, horizontal, rotational, multidirectional, oblique, no arrow.
- Contour of the arrow: thin, thick, outline, filled, no arrow.
- Contour of the object: thin, thick, outline, filled, mixed, no object.

- Three-dimensionality of the object: 3D object, 2D object, no object.
- Three-dimensionality of the arrow: 3D arrow, 2D arrow, no arrow.

An example of alternative symbols using 2D or 3D objects is represented in Figures 3. In Figure 4 we can see the presence of a real object rather than anthropometric elements and abstract objects. The arrows also differ, using a thin outline in the first case, and thin lines in the second. Figure 5 shows an example of different directions and object orientations. However, we designed the symbols without following categories based on these characteristics, rather than using the ones that felt more natural for the action to be represented. For instance, not for all the actions the presence of an anthropomorphic element makes sense, and not for all of them different orientations are possible.

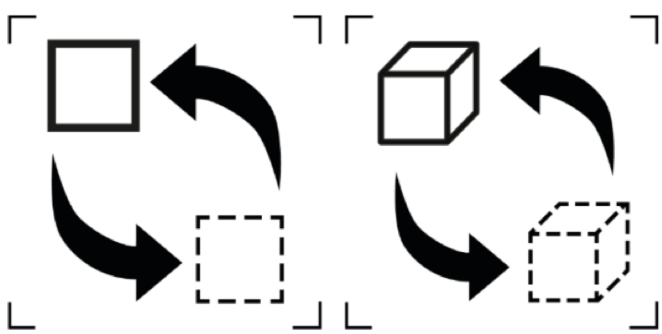


Figure 3: In this example the same symbol can contain two or three-dimensional geometries

3.3. Symbol elicitation

To test the designed symbols we prepared a questionnaire on Google survey. Since we refer to ISO symbols for use on equipment [39], we followed

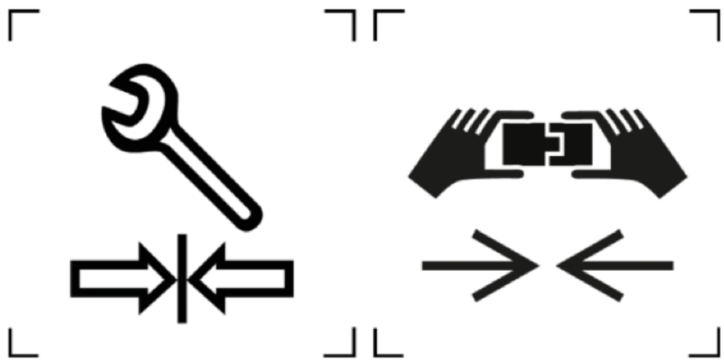


Figure 4: In this example the symbol can contain anthropometric visual elements or simple tools

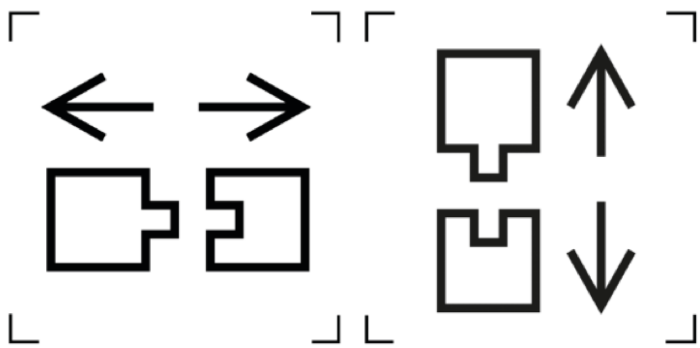


Figure 5: An example of the same symbol presented in two different directions

the relative rules listed in ISO 9186-3, Method for testing symbol referent association [42]. The symbols should be presented as black shapes on a white background, unless the colour is an integrant part of the coding information (e.g., safety symbols). The image size of a symbol should not be less than 28×28 mm, in case of screen presentation and assuming a distance between 400 and 700 mm. As the test does not take place in a controlled environment, we first tested the legibility of text and images on a user screen.

After an initial page of anagraphic questions, the questionnaire begins showing the question for the first action. For each action we presented its definition according to the STE vocabulary, in order to avoid ambiguity and prevent the users to provide random answers. The test consists of selecting the symbol that the user considered the most appropriate for expressing the action among the 20 different symbols presented. This was repeated for all the 21 maintenance actions selected. For each user, the sequence of the actions displayed was the same, while the order of the symbols was randomised. For each action and at the end of the questionnaire users had the possibility to share comments and suggestions.

The test was submitted to generic people, trying to select English speakers, which is the reason why they were mainly students. We did not consider necessary to submit the test to professional maintainers, as suggested by ISO 9186-3 [42]. This normative is for testing symbol referent association and it is used to test symbols to be represented in specific contexts, as the industrial one. Since our test is performed by generic people rather than by field's experts, the participants need to familiarise with the symbols' referents before answering. We did that by providing the STE definitions.

192 valid data were gathered out of 193 persons that started the survey. One of them reported the symbols to be not clearly legible on the screen (so this result was not considered).

147 (76,2%) of the participants were male, and 46 (23%) were female. 79,3% of them were students, and 20,7% were workers. Most of them had an engineering background, or other technical background. Additional information regarding participants' education and English level can be found in Tables 1 and 2.

	High School Diploma	**Bachelor Degree**	**Master Degree**	**PhD**
Education	31,6%	39,9%	23,3%	5,2%
	Italian	**Other European countries**	**Asian countries**	**Other countries**
Nationality	75,5%	10,0%	11,9%	2,6%

Table 1: Demographic information of people who participated to the online survey

	1	**2**	**3**	**4**	**5**	**6**	**7**
English level	1,6%	3,6%	11,4%	24,9%	39,4%	18,7%	0,5%

Table 2: English level of people who participated to the online survey

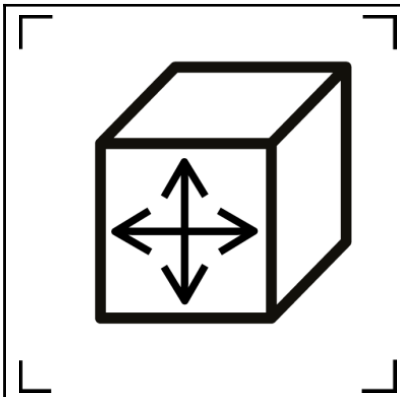
3.4. Criteria for vocabulary definition

The creation of vocabularies results from choosing a symbol to represent each referent (i.e., in our case, the actions). As argued by Wobbrock et al. [41], in today's computing systems, high guessability is essential, particularly for symbolic input. They defined guessability as that quality of symbols that allows a user to access intended referents via those symbols despite a lack of knowledge about them. They introduced a measure of guessability for symbolic input. The guessability G of the resultant symbol set (i.e., vocabulary) S for the captured set of proposed symbols P is:

$$G = \frac{\sum_{s \in S} |P_s|}{|P|} \cdot 100\% \tag{1}$$

In Equation 1, P is the set of proposed symbols for all referents, and Ps is the set of proposed symbols using symbol s, which is a member of the resultant symbol set S. In our test we have 21 referents and 192 participants, so $|P| = 21 \times 192 = 4032$ is the number of proposed symbols for all referents.

The highest guessability is reached if, for every referent, we selected the symbol with the highest preference among the users. However, there are also



Direction of the arrow/s	Multidirectional
Three-dimensionality of the arrow	No
Contour of the arrow	Thin
Type of the object	Simple
Three-dimensionality of the object	Yes
Contour of the object	Thin
Presence of an anthropomorphic element	No

Table 3: An example of tagging a symbol

other criteria and constraints to consider in the composition of a vocabulary. Those are the "homogeneity", the "reversibility", and the "conflict set" criteria [41, 43, 44].

A vocabulary is homogeneous if the included symbols have the same attributes. In our study, these attributes could not be defined a priori, because we were not able to identify a category of symbols that was more effective than the others (e.g., symbols containing anthropomorphic elements, symbols containing 2D/3D elements) as the choices are cross to the categories. In the design of symbols, we mixed these attributes. Then a vocabulary can contain symbols with different attributes at different scales. The attributes considered are the following:

- Direction of the arrow/s in the symbol: vertical, horizontal, rotational, multidirectional, oblique, no arrow.
- Three-dimensionality of the arrow: 3D arrow, 2D arrow, no arrow.

- Contour of the arrow: thin, thick, outline, filled, no arrow.
- Type of the object in the symbol: simple, i.e., defined with less than three primitives as line, triangle, square, circle; complex, i.e., defined with three or more primitives; real object; no object.
- Three-dimensionality of the object: 3D object, 2D object, no object.
- Contour of the object: thin, thick, outline, filled, mixed, no object.
- Presence of an anthropomorphic element (e.g., a hand) in the symbol: yes or not.

We tagged each symbol based on the previous attributes (Table 3).

For each attribute we computed a homogeneity index of a vocabulary H:

$$H(a) = \frac{\sum_{P_i \subseteq P} \frac{1}{2}|P_i|(|P_i| - 1)}{\frac{1}{2}|P|(|P| - 1)} \cdot 100\% \tag{2}$$

Where P is the set of all proposals for the attribute a, $|P|$ is the size of the set, and Pi is the subsets of identical proposals for the attribute a (e.g. all the symbols with vertical arrow/s). This definition was deduced from that of the Agreement Rate in Wobbrock et al. [41]. In fact, we can define homogeneity of a vocabulary according to an attribute a, as the number of pairs of symbols in agreement (same proposal) with each other divided by the total number of pairs of symbols that could be in agreement [45]. In this way, we got a homogeneity index $H(a)$ for each attribute a. Therefore, to obtain a unique homogeneity index H for a vocabulary, we computed the mean of these seven indices.

Another constraint to respect is that of reversibility: for each symbol representing an action, the opposite action should be represented by the opposite symbol. In fact, if each opposite symbol differed from the others, the vocabulary would have an increased number of visual features to recognise, which might affect the usability of the system, as recognising all the objects would require more time.

Finally, the conflict set constraint must be checked since one symbol could be assigned just to one referent. If two or more referents have the same symbol, it should be assigned to the referent with the highest number of proposals. We extracted three vocabularies, following these principles:

1. Maximising the guessability, ignoring the homogeneity;
2. Maximising the homogeneity, ignoring the guessability;
3. Hybrid approach to have both good guessability and homogeneity.

4. Results

Vocabulary 1 is obtained assigning the symbol with the highest preference to each referent. Guessability of vocabulary 1 is $G = 32.9\%$ and homogeneity is $H = 42.2\%$. This vocabulary respects the conflict set constraint, but not the reversibility constraint for the pairs "tighten/loosen", "install/remove", and "push/pull". For "install/remove" we decided to maintain the symbol type of the referent "install" in the vocabulary because there is not an opposite symbol of the symbol type of the referent "remove". For the other two pairs, we examined the four vocabularies obtained combining the two symbol types for each referent. Among these four vocabularies, we decided to maintain that with the highest guessability that is vocabulary 1.1 ($G = 31.8\%$, $H = 42.2\%$).

Vocabulary 2 is obtained putting together all the symbols that shared the highest possible number of attributes. Homogeneity of vocabulary 2 is $H = 75.8\%$, while its guessability is $G = 5.5\%$. The conflict set constraint is respected, while the reversibility constraint is violated by the pair "install/remove". However, there is not an opposite symbol for the symbol type of the referent 'install', then the resulting vocabulary is just one: vocabulary 2.1 with $G = 6.1\%$ and $H = 74.1\%$.

Vocabulary 3 is obtained considering for each referent the five proposals with the highest preferences. Then we created a Matlab [46] code to find the vocabulary with the highest homogeneity among those obtained combining the symbols. However, the possible combinations of symbols were 5^{21}, which made it impossible to analyse them all. Therefore, we divided the referent in three classes:

- High level verbs: move, disassemble, assemble, install, remove, replace, hold;
- Low level verbs with agreement rate [47] > 0.2: examine, fill, clean, cut;
- Low level verbs with agreement rate < 0.2: tighten, loosen, lift, push, pull, turn clockwise, drain, turn counter clockwise, close, open.

In this way, the possible combinations reduced to $5^7 + 5^{10} + 5^4 = 9844375$, and it took approximately 20 minutes to find the optimal vocabulary. Homogeneity of vocabulary 3 is $H = 45.3\%$, while its guessability is $G = 24.3\%$

. Conflict set constraint is respected, while reversibility constraint is violated for the pairs 'install/remove' and 'push/pull'. We examined the four vocabularies obtained by combining the two symbol types for each referent. Among these four vocabularies, we decided to maintain that one with the highest guessability, which is vocabulary 3.1 ($G = 23.9\%$, $H = 45.3\%$).

The three vocabularies extracted at the end of this user evaluation are displayed in Table 4, with the relative guessability and homogeneity. The confidence value for each referent of the three vocabularies can be found in Tables 5, 6 and 7. Confidence value is defined as the percentage of participants who proposed that symbol, and is computed by using the definition given by Vatavu [47]:

$$C_r = \max_{P_i \subseteq P_r} \frac{|P_i|}{|P_r|} \cdot 100\% \tag{3}$$

where Pi is the set of participants that proposed the same symbol I, while Pr is the set of all the participants. According to Vatavu [47], confidence value is important to describe the goodness of each command. Confidence value is high if C is higher than 50%, medium if C is between 25% and 50%, low if C is lower than 25%. Vocabulary 1.1 has 5 symbols with low confidence value and 2 with high confidence value. Vocabulary 2.1 has 20 symbols with low confidence value and none with high value. Vocabulary 3.1 has 10 symbols with low confidence value and 2 with high confidence value. Given these results, we decided to use the vocabulary 1.1 for the case study.

5. Case Study

The vocabulary we selected (vocabulary 1.1) was integrated into an existing AR application to support collaborative maintenance. The first version of this application was presented in [40] and improved by the authors in [40]. The application was developed using the Unity 3D 4.3.4 environment, integrated with the Qualcomm Vuforia 2.8.9 tool used for tracking the real scene, and allowing to superimpose digital information onto the real environment.

The application works on different devices and operating systems and its architecture is of a client-server type. The application allows the user to decide whether to be server or client, that means choosing to start the session as skilled or unskilled operator, there can be more clients. The skilled operator can give instructions through a personal computer, equipped with the traditional input and output devices (keyboard, mouse and monitor), or he/she can use a tablet or a mobile device equipped with the Android operating system or iOS. The unskilled operator can use a tablet (which is also equipped with the Android operating system or iOS) through which he/she can see a series of information and commands projected directly on the objects of which he/she is doing maintenance.

The unskilled operator/operators (client) use the application to capture pictures of the real environment and send them to the skilled one (server). These pictures are used on the client side as a marker to visualise the instructions on the real environment, and on the server side to point out the maintenance task to perform. The skilled operator uses sketches and symbols to point out respectively where to act, and what to do. Furthermore, he/she can add a text message to send further information to the client. The sketch, the symbols and the text will be displayed in Augmented Reality modality to the client.

The authors in [40] listed different requirements to improve the previous version of this application, among which the use of a universal language for operators of different languages and cultures, and a maintenance recorder module to register the operation as a potential manual in case the same failure occurs on another machine, or for legal issues.

The version of the application we propose, which is based on conveying technical information through 2D symbols, attempts to satisfy both these requirements. The maintenance instruction is provided with a sketching tool and the 2D symbols selectable from a list. The sketching tool is used to identify the area or object on which maintenance should be performed (Figure 6), while the symbol communicates the action to perform on the live image taken from the client camera (Figure 7). In a first version of the application the symbol was displayed near the sketch made in correspondence of the real part. However, we realised that there was a low added value respect to its placement fixed on the GUI, at the expense of an occlusion of the scene, even if small.

A recorder module has been implemented as well, saving screenshots of the application each time a piece of information is provided from the skilled operator. The operators communicate through vocal messages, which are also saved onto the same folder reporting date and time of the maintenance action.

The system is made as to meet the needs of maintainers with different skills. First because, as explained in section 3.1, the vocabulary includes low-level and high-level verbs, adapting to actions with different levels of difficulty. The same can be said about users' skills: unskilled workers will probably require the use of low-level verbs more frequently. Secondly, when an action implies other ones, the skilled operator can choose if guiding the maintainer by providing instructions only about the main action, or instead about each step to complete. Furthermore, if additional communication is needed, the microphone and message tools can be used.

Figure 6: Application on client. The sketch drawn by the expert operator is displayed, indicating the component on which the action should be accomplished. The action is illustrated by the symbol presented on the left.

6. Discussion and conclusion

In this paper, we present a methodology to convert maintenance actions into 2D symbols to convey instructions in AR. According to the authors, the advantages of this approach are the following:

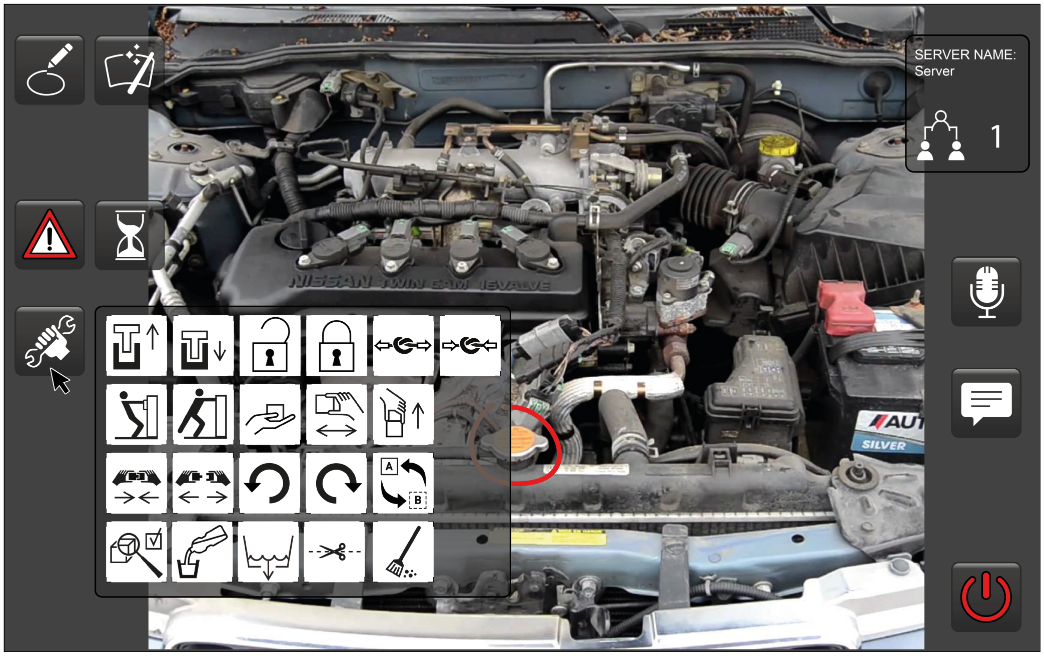


Figure 7: Application on server. After sketching to indicate the component on which the action should be accomplished, the expert operator can select the symbol corresponding to the action from the vocabulary panel. The symbol and the sketch will be sent to the maintainer.

- simpler authoring: technical writers should just select the desired symbol from a vocabulary;
- lower mental load for the users: no 3D perception skills are needed;
- working without tracking: 2D symbols can be displayed on the GUI, and the real component they refer to can be highlighted with a graphic sign.

Through this method we aimed to create a vocabulary of symbols to cover most of the maintenance actions. Our strategy is not intended to replace 3D models with 2D symbols, but to use the latter for the most frequent general purpose tasks and the former for very complex customized operations. We think that the implementation of an extensive vocabulary is necessary for an effective integration of 2D symbols in AR applications. Furthermore, this work may be considered as a starting point of a process to determine standard guidelines for the communication of instructions in AR. We believe that those guidelines are required for AR manuals to be accepted and come in use in the industrial world. In fact, the design principles of Industry 4.0 require that technical documentation should be: digital, connected to the internet, updatable in real-time, standardized, easy to translate, organized as a service, modular, easy to access. The use of AR manuals based on STE and standardized symbols pursues these design principles.

To define the symbols vocabulary, we had to solve preliminary questions, about the definition of the maintenance actions to convert in symbols, and the design of symbols. In the first part of this work, we analysed technical documentation to identify maintenance actions. The aim was to create a complete vocabulary of actions to enable expert maintainers to provide maintenance commands. We noted that there are verbs with the same meaning, for example "install" and "mount". In these cases, we defined a unique symbol, referring to the STE approved verb. Even after this association, there were still lots of action verbs, so we decided to limit the list to the most frequent. With the defined list, we covered about 84% of the task actions found in the analysed manuals, corresponding to 21 verbs.

We then designed a set of 20 symbols for each verb. They were designed following guidelines that we defined observing ISO symbols and some of these symbols were already present in ISO norms. Then, we made an elicitation study to identify which symbols user preferred for each referent. From the results of this test, we answered to our research question. We created twelve vocabularies applying these criteria: guessability, including the most preferred symbols; homogeneity, including symbols with similar visual features; conflict set, including different symbols for each verb; reversibility, assigning opposite symbols to opposite actions. We found that the vocabulary with the highest homogeneity $(H = 74.1\%)$ presented low guessability $(G = 6.1\%)$, i.e., the symbols included in this vocabulary had few preferences among the users. On the other side, the vocabulary with the highest guessability $(G = 31.8\%)$ had an average homogeneity $(H = 42.2\%)$. We discarded the vocabulary with the highest homogeneity because it had very low confidence values and some symbols were not chosen by any of the users. We then tried to create a new vocabulary, following a hybrid approach to have both a good guessability and homogeneity. However, homogeneity was 7% increased $(H = 45.3\%)$, but guessability was 25% reduced $(G = 23.9\%)$. Comparing this vocabulary with that with the highest guessability we had more symbols (10 against 5) with a low confidence value. For this reason, we decided to integrate in the AR application, used as case study, the vocabulary with the highest guessability (vocabulary 1.1) among the ones respecting the reversibility constraint.

The results of this study lead to additional considerations. For example, considering the distinction between low-level and high-level actions. In our previous work [40], we made this distinction expecting that users could associate a precise mental image to low-level actions, since they express simple movements and clear procedures. We then expect that low-level verbs would report higher confidence values than high-level ones. Our result confirms this hypothesis since, for the vocabulary chosen, three out of seven high level actions had low confidence values. Furthermore, only the verbs "pull" and "loosen" had low confidence values among low-level verbs, but just because we replaced the symbol to respect the reversibility constraint. As a result, in our future work we plan to validate the intuitiveness of the symbols in real maintenance procedures, exploring the advisability to convert all the verbs into symbols or just low-level verbs.

The results we present integrate the research works concerning the ways to convey information in Augmented Reality for remote collaboration applications. In particular, Nuernberger et al. found that users tend to draw arrows, circles, and outlines more than other types of drawings, and that arrows are used more with action tasks [48]. This result could seem to not be in accordance with our strategy based on the use of 2D symbols. However, they let the users free to draw whatever they wished to complete the task. It is then reasonable that participants preferred to draw simpler objects like arrows, lines and circles rather than complex sketches. But these simple features are not enough to convey the precise action to be performed. In a remote collaboration application this could be done through oral communication between workers, but there could be misunderstandings and translation issues. The use of pre-designed standard 2D symbols overcome this issues, by integrating the information provided with sketches of simple features such as arrows, circles and outlines, which can be sketched freely by the user.

The issue of optimising the authoring process of technical documentation for Augmented Reality began to be addressed in the last years. Various strategies to convert existing manuals are being proposed, mostly making use of CAD models. For example, Mohr et al. proposed a system capable of automatically transferring traditional printed documentation to AR [49]. Their system recognises an annotated explosion diagram on the paper manual and generates 3D annotations and animations displayed in AR. Engelke et al. proposed a concept that allows technically skilled persons to author and transfer maintenance task descriptions from traditional instructions to AR where useful and required without creating the descriptions from scratch [50]. However, both these methods convert pieces of documentation that are already in a visual form as explosion drawings. They do not suggest methods to convert text information. Our research, that instead is focused on the replacement of text task instructions, could be easily combined with what found in the previously cited works.

Bibliography

[1] T. D. Oesterreich, F. Teuteberg, Understanding the implications of digitisation and automation in the context of industry 4.0: A triangulation approach and elements of a research agenda for the construction industry, Computers in Industry 83 (Supplement C) (2016) 121 – 139. doi:https://doi.org/10.1016/j.compind.2016.09.006. URL http://www.sciencedirect.com/science/article/pii/S0166361516301944

[2] S. M. Harrison, A comparison of still, animated, or nonillustrated on-line help with written or spoken instructions in a graphical user interface, in:

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '95, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1995, pp. 82–89. doi:10.1145/223904.223915.
URL http://dx.doi.org/10.1145/223904.223915

[3] J. Heiser, D. Phan, M. Agrawala, B. Tversky, P. Hanrahan, Identification and validation of cognitive design principles for automated generation of assembly instructions, in: Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '04, ACM, New York, NY, USA, 2004, pp. 311–319. doi:10.1145/989863.989917.
URL http://doi.acm.org/10.1145/989863.989917

[4] R. E. Mayer, W. Bove, A. Bryman, R. Mars, L. Tapangco, When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons., Journal of educational psychology 88 (1) (1996) 64.

[5] P. Thompson, The digital natives as learners: Technology use patterns and approaches to learning, Computers & Education 65 (Supplement C) (2013) 12 – 33. doi:https://doi.org/10.1016/j.compedu.2012.12.022.
URL http://www.sciencedirect.com/science/article/pii/S0360131513000225

[6] M. Prensky, Digital natives, digital immigrants part 1, On the horizon 9 (5) (2001) 1–6.

[7] S. Bennett, K. Maton, L. Kervin, The 'digital natives' debate: A critical

review of the evidence, British journal of educational technology 39 (5) (2008) 775–786.

[8] https://www.ifixit.com/, Last checked 18 January 2018.

[9] http://www.instructables.com, Last checked 18 January 2018.

[10] S. Stokes, Visual literacy in teaching and learning: A literature perspective, 2002.

[11] D. Andrulis, C. Brach, Integrating literacy, culture, and language to improve health care quality for diverse populations 31 Suppl 1 (2007) S122–33.

[12] ASD, Rue Montoyer 101000, Brussels, Belgium, Simplified Technical English, ASD-STE100, issue 6 Edition (January 2013).

[13] M. Hermann, T. Pentek, B. Otto, Design principles for industrie 4.0 scenarios, in: System Sciences (HICSS), 2016 49th Hawaii International Conference on, IEEE, 2016, pp. 3928–3937.

[14] D. Reiners, D. Stricker, G. Klinker, S. Müller, Augmented reality for construction tasks: Doorlock assembly, in: IN PROC. IEEE AND ACM IWAR'98 (1. INTERNATIONAL WORKSHOP ON AUGMENTED REALITY, AK Peters, 1998, pp. 31–46.

[15] J. Gimeno, P. Morillo, J. M. Orduña, M. Fernández, A new ar authoring tool using depth maps for industrial procedures, Comput. Ind. 64 (9) (2013) 1263–1271. doi:10.1016/j.compind.2013.06.012.
URL http://dx.doi.org/10.1016/j.compind.2013.06.012

[16] I. Stock, M. Weber, E. Steinmeier, Metadata based authoring for technical documentation, in: Proceedings of the 23rd Annual International Conference on Design of Communication: Documenting & Designing for Pervasive Information, SIGDOC '05, ACM, New York, NY, USA, 2005, pp. 60–67. doi:10.1145/1085313.1085330.
URL http://doi.acm.org/10.1145/1085313.1085330

[17] C. Knopfle, J. Weidenhausen, L. Chauvigne, I. Stock, Template based authoring for ar based service scenarios, in: IEEE Proceedings. VR 2005. Virtual Reality, 2005., 2005, pp. 237–240. doi:10.1109/VR.2005.1492779.

[18] R. Radkowski, J. Herrema, J. Oliver, Augmented reality-based manual assembly support with visual features for different degrees of difficulty, International Journal of Human–Computer Interaction 31 (5) (2015) 337–349. arXiv:http://dx.doi.org/10.1080/10447318.2014.994194, doi:10.1080/10447318.2014.994194.
URL http://dx.doi.org/10.1080/10447318.2014.994194

[19] N. Pathomaree, S. Charoenseang, Augmented reality for skill transfer in assembly task, in: Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on, IEEE, 2005, pp. 500–504.

[20] S. J. Henderson, S. K. Feiner, Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret, in: ISMAR, 2009.

[21] S. K. Ong, M. L. Yuan, A. Y. C. Nee, Augmented reality applications in manufacturing: a survey, International Journal of Production Research 46 (10) (2008) 2707–2742. arXiv:http://dx.doi.org/10.1080/00207540601064773, doi:10.1080/00207540601064773.
URL http://dx.doi.org/10.1080/00207540601064773

[22] W. Dangelmaier, M. Fischer, J. Gausemeier, M. Grafe, C. Matysczok, B. Mueck, Virtual and augmented reality support for discrete manufacturing system simulation, Computers in Industry 56 (4) (2005) 371 – 383, the Digital Factory: An Instrument of the Present and the Future. doi:https://doi.org/10.1016/j.compind.2005.01.007.
URL http://www.sciencedirect.com/science/article/pii/S0166361505000266

[23] D. Mizell, Boeing's wire bundle assembly project, in: W. Barfield, T. Caudell (Eds.), Fundamentals of wearable computers and augmented reality, Lawrence Erlbaum Associates New York, NY, USA, 2001, pp. 447–467.

[24] S. Feiner, B. Macintyre, D. Seligmann, Knowledge-based augmented reality, Commun. ACM 36 (7) (1993) 53–62. doi:10.1145/159544.159587.
URL http://doi.acm.org/10.1145/159544.159587

[25] U. Neumann, A. Majoros, Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance, in: Proceedings of the Virtual Reality Annual International Symposium, VRAIS '98, IEEE Computer Society, Washington, DC, USA, 1998, pp.

4–.
URL http://dl.acm.org/citation.cfm?id=522258.836112

[26] B. Bell, S. Feiner, T. Höllerer, View management for virtual and augmented reality, in: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology, UIST '01, ACM, New York, NY, USA, 2001, pp. 101–110. doi:10.1145/502348.502363.
URL http://doi.acm.org/10.1145/502348.502363

[27] X. Yin, Y. Gu, S. Qiu, X. Fan, Vr & ar combined manual operation instruction system on industry products: A case study, in: 2014 International Conference on Virtual Reality and Visualization, 2014, pp. 65–72. doi:10.1109/ICVRV.2014.55.

[28] H. Regenbrecht, G. Baratoff, W. Wilke, Augmented reality projects in the automotive and aerospace industries, IEEE Comput. Graph. Appl. 25 (6) (2005) 48–56. doi:10.1109/MCG.2005.124.
URL http://dx.doi.org/10.1109/MCG.2005.124

[29] T. Hollerer, S. Feiner, D. Hallaway, B. Bell, M. Lanzagorta, D. Brown, S. Julier, Y. Baillot, L. Rosenblum, User interface management techniques for collaborative mobile augmented reality, Computers & Graphics 25 (5) (2001) 799 – 810, mixed realities - beyond conventions. doi:https://doi.org/10.1016/S0097-8493(01)00122-4.
URL http://www.sciencedirect.com/science/article/pii/S0097849301001224

[30] T. Zarraonandia, I. Aedo, P. Diaz, A. M. Montes, Augmented pre-

sentations: Supporting the communication in presentations by means of augmented reality, International Journal of Human-Computer Interaction 30 (10) (2014) 829–838. arXiv:http://dx.doi.org/10.1080/10447318.2014.927283, doi:10.1080/10447318.2014.927283. URL http://dx.doi.org/10.1080/10447318.2014.927283

[31] S. Webel, U. Bockholt, T. Engelke, N. Gavish, M. Olbrich, C. Preusche, An augmented reality training platform for assembly and maintenance skills, Robotics and Autonomous Systems 61 (4) (2013) 398 – 403, models and Technologies for Multi-modal Skill Training. doi:http://doi.org/10.1016/j.robot.2012.09.013. URL http://www.sciencedirect.com/science/article/pii/S0921889012001674

[32] M. Hermann, T. Pentek, B. Otto, Design principles for industrie 4.0 scenarios, in: 2016 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 3928–3937. doi:10.1109/HICSS.2016.488.

[33] International Organization for Standardization, Switzerland, ISO 17724:2003 Graphical symbols – Vocabulary, 1st Edition (August 2003).

[34] https://boss.cen.eu/reference material/guidancedoc/pages/graphsymb.aspx, Last checked 18 January 2018.

[35] V. Havard, D. Baudry, X. Savatier, B. Jeanne, A. Louis, B. Mazari, Augmented Industrial Maintenance (AIM): A Case Study for Evaluating and Comparing with Paper and Video Media Supports, Springer International Publishing, Cham, 2016, pp. 302–320. doi:10.1007/

978-3-319-40621-3_22.
URL https://doi.org/10.1007/978-3-319-40621-3_22

[36] International Organization for Standardization, Switzerland, ISO 80416-2 Basic principles for graphical symbols foruse on equipment —Part 2:Form and use of arrows, 1st Edition (July 2001).

[37] Augmented reality for maintenance and repair (armar) - http://graphics.cs.columbia.edu/projects/armar/index.htm, Last checked 18 January 2018.

[38] S. Henderson, S. Feiner, Exploring the benefits of augmented reality documentation for maintenance and repair, IEEE Transactions on Visualization and Computer Graphics 17 (10) (2011) 1355–1368. doi:10.1109/TVCG.2010.245.

[39] International Organization for Standardization, Switzerland, ISO 7000:2014 Graphical symbols for use on equipment – Registered symbols, 5th Edition (January 2014).

[40] authors' papers, for blind review.

[41] J. O. Wobbrock, H. H. Aung, B. Rothrock, B. A. Myers, Maximizing the guessability of symbolic input, in: CHI '05 Extended Abstracts on Human Factors in Computing Systems, CHI EA '05, ACM, New York, NY, USA, 2005, pp. 1869–1872. doi:10.1145/1056808.1057043.
URL http://doi.acm.org/10.1145/1056808.1057043

[42] International Organization for Standardization, Switzerland, ISO 9186-3

Graphical Symbols - Test methods - Part3- Method for symbol referent association, 1st Edition (October 2014).

[43] Human interface guidelines v2.0 available at http://download.microsoft.com/download/6/7/6/676611b4-1982-47a4-a42e-4cf84e1095a8/kinecthig.2.0.pdf, Last checked 18 January 2018 (2014).

[44] T. Piumsomboon, A. Clark, M. Billinghurst, A. Cockburn, User-defined gestures for augmented reality, in: CHI '13 Extended Abstracts on Human Factors in Computing Systems, CHI EA '13, ACM, New York, NY, USA, 2013, pp. 955–960. doi:10.1145/2468356.2468527.
URL http://doi.acm.org/10.1145/2468356.2468527

[45] R.-D. Vatavu, J. O. Wobbrock, Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, ACM, New York, NY, USA, 2015, pp. 1325–1334. doi:10.1145/2702123.2702223.
URL http://doi.acm.org/10.1145/2702123.2702223

[46] https://www.mathworks.com, Last checked 18 January 2018.

[47] R.-D. Vatavu, A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments, Journal of Ambient Intelligence and Smart Environments 5 (2) (2013) 187–211.

[48] B. Nuernberger, K. C. Lien, T. Höllerer, M. Turk, Interpreting 2d gesture annotations in 3d augmented reality, in: 2016 IEEE Symposium

on 3D User Interfaces (3DUI), 2016, pp. 149–158. doi:10.1109/3DUI.2016.7460046.

[49] P. Mohr, B. Kerbl, M. Donoser, D. Schmalstieg, D. Kalkofen, Retargeting technical documentation to augmented reality, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, ACM, New York, NY, USA, 2015, pp. 3337–3346. doi:10.1145/2702123.2702490.
URL http://doi.acm.org/10.1145/2702123.2702490

[50] T. Engelke, J. Keil, P. Rojtberg, F. Wientapper, M. Schmitt, U. Bockholt, Content first: A concept for industrial augmented reality maintenance applications using mobile devices, in: Proceedings of the 6th ACM Multimedia Systems Conference, MMSys '15, ACM, New York, NY, USA, 2015, pp. 105–111. doi:10.1145/2713168.2713169.
URL http://doi.acm.org/10.1145/2713168.2713169

	VOC 1	**VOC 1.1**	VOC 1.2	VOC 1.3	VOC 1.4	VOC 2	**VOC 2.1**	VOC 3	**VOC 3.1**	VOC 3.2	VOC 3.3	VOC 3.4
MOVE												
DISASSEMBLE												
ASSEMBLE												
INSTALL												
REMOVE												
REPLACE												
HOLD												
TIGHTEN												
LOOSEN												
EXAMINE												
LIFT												
PUSH												
PULL												
TURN COUNTER CLOCKWISE												
TURN CLOCKWISE												
FILL												
DRAIN												
CLOSE												
OPEN												
CUT												
CLEAN												
GUESSABILITY	32.9%	**31.8%**	31.5%	31.7%	31.3%	5.5%	**6.1%**	24.3%	**23.9%**	23.5%	23.3%	23.0%
HOMOGENEITY	42.2%	**42.2%**	42.2%	43.0%	43.0%	75.8%	**74.1%**	45.3%	**45.3%**	45.3%	45.4%	45.4%

Table 4: Vocabularies 1, 2, and 3 were created considering the guessability and homogeneity criteria. Vocabularies from 1.1 to 1.4, 2.1, and from 3.1 to 3.4 were created considering also the conflict set and reversibility constraints, and new symbols were chosen for some referents.

MOVE C=21.9%	DISASSEMBLE C=34.9%	ASSEMBLE C=30.2%	INSTALL C=19.3%	REMOVE C=11.5%	REPLACE C=35.9%	HOLD C=30.2%
TIGHTEN C=29.7%	LOOSEN C=21.4%	EXAMINE C=39.6%	LIFT C=28.6%	PUSH C=30.2%	PULL C=23.4%	TURN COUNTER CLOCKWISE C=32.8%
TURN CLOCKWISE C=33.9%	FILL C=53.1%	DRAIN C=26.0%	CLOSE C=25.5%	OPEN C=28.6%	CUT C=68.2%	CLEAN C=43.2%

Table 5: Confidence values for vocabulary 1.1 high $\geq$ 50 % (green) , medium 25-50 % (yellow), low $<$ 25 % (red)

MOVE C=7.3%	DISASSEMBLE C=26.0%	ASSEMBLE C=21.9%	INSTALL C=9.9%	REMOVE C=15.1%	REPLACE C=0%	HOLD C=0.5%
TIGHTEN C=1.6%	LOOSEN C=2.1%	EXAMINE C=7.3%	LIFT C=0.5%	PUSH C=1.0%	PULL C=3.6%	TURN COUNTER CLOCKWISE C=9.4%
TURN CLOCKWISE C=8.3%	FILL C=1.6%	DRAIN C=0%	CLOSE C=6.3%	OPEN C=4.2%	CUT C=0%	CLEAN C=2.1%

Table 6: Confidence values for vocabulary 2.1 high $\geq$ 50% (green) , medium 25-50% (yellow), low $<$ 25% (red)

MOVE C=7.3%	DISASSEMBLE C=11.5%	ASSEMBLE C=10.9%	INSTALL C=19.3%	REMOVE C=11.5%	REPLACE C=35.9%	HOLD C=30.2%
TIGHTEN C=7.8%	LOOSEN C=13.0%	EXAMINE C=4.7%	LIFT C=25.5%	PUSH C=30.2%	PULL C=23.4%	TURN COUNTER CLOCKWISE C=32.8%
TURN CLOCKWISE C=33.9%	FILL C=53.1%	DRAIN C=26.0%	CLOSE C=25.5%	OPEN C=28.6%	CUT C=68.2%	CLEAN C=2.6%

Table 7: Confidence values for vocabulary 3.1 high $\geq 50\%$ (green), medium 25-50% (yellow), low $< 25\%$ (red)