Industrial Application of Accented Visualization Based on Augmented Reality

Anton Ivaschenko, Michael Milutkin
Samara State Technical University
Samara, Russia
anton.ivashenko@gmail.com

Pavel Sitnikov
ITMO University
Saint-Petersburg, Russia
sitnikov@o-code.ru

Abstract—**This paper presents an overview and analysis of IT solutions based on implementation of Augmented Reality in industrial applications. Conclusions are based on an authors' experience of AR implementation in the last years. A new concept of accented visualization is provided to solve typical problems. There are presented the results of AR technologies implementation in practice for manual operation control. Main attention is given to the usability of interactive user interfaces and knowledge-based data processing to improve the quality of intelligent technologies efficient use in digital industry applications. The paper develops the ideas of accentuated visualization based on adaptive construction and virtual consideration of the content of the current real scene in the field of view of a person, as well as the viewer's experience that contains perceptions, points of view and expected behavior. The proposed solution is used to identify gaps and failures of operator in real time, predict possible operating mistakes and suggest better procedures based on comparing the sequence of actions to an experience of highly qualified operators.**

I. Introduction

Augmented Reality (AR) is one of the most promising technologies nowadays that considerably improve the capabilities of interactive user interfaces. Modern AR devices become comparatively cheap and easy to use in most areas of human-computer interaction. Starting from various 3D games that provide computer generated virtual content additive to the real environment AR solutions become natural to multiple areas that include service, manufacturing, and transport industries, medicine, tourism and education.

Despite rich functionality and high usability of AR devices, their implementation in practice remains challenging so far. The problems are concerned with both usability and perception. Most users undergo difficulties in using head mounted devices while executing routine manipulations. People treat with inequality the digital intervention into their private reality, which makes it difficult to develop universal solutions for decision-making support.

Based on the experience of solving of these types of issues there was developed an original approach of accented visualization. This concept considers the main feature of AR based solutions that bring digital objects into a person's perception of the real world. Consequently, virtual and real objects are integrated together generating immersive sensations that are perceived as natural parts of an environment. Accented visualization provides formalization and control over this process taking into account the individual user behavior and adapting the user interface accordingly.

II. State of the Art

Most modern AR technologies [1], [2] provide powerful and adaptive user interfaces. Various AR devices (goggles, head mounted displays or widely spread tablets) provide overlaid information additive or masking the real environment that can be informative for the users. AR technology allows developing interactive and context-dependent user interfaces that provide the possibilities of computer vision and object recognition in real time.

Modern AR technologies are tailored with adaptive user interfaces (UI) [3]. AR devices need to present the adequate information required at a certain stages of user story with no redundant data and maximum attraction to important items. Therefore, AR based UI differs from the usual UI of personal computers. Additional data and control elements can be placed and presented in all accessible space, in 3D with no limitations of screens or panel size. This data is overlapping the real objects and can hide some of them and use the others for docking.

In case of presenting textual information it should appear in the focus of the user to make it readable with no disturb. These problems remain open for most of practical applications of AR and are solved by adapting the UI to concrete user cases.

The possibilities of AR as a ubiquitous user interface to the real world are discussed in [4], [5], [6]. User Interfaces are usually designed with the focus on maximizing usability and user experience. The goal of user interface design is to make the user's interaction as simple and efficient as possible, in terms of accomplishing user goals.

To provide adequate and stable identification of real objects and items with a complex shape there can be implemented an algorithm based on a combination of standard technologies of image analysis and neural network-based pattern recognition [7]. Methods of identification of components that have a simple form or consist of elements with a simple shape were taken from free sources and modified for the required quality and performance.

ISSN 2305-7254

Contextual data visualization [8], [9] is provided to combine several data sets to analyze multiple layers of a biological system at once. The system should interlink all related data sets (e.g., images, text, measured values, scans) and offer visual analytics to support experts.

This approach is widely used for medical data processing but can be easily disseminated for a cyber-physical system. This approach supports the idea of maximum effective visualization of complex data for professionals instead of automatic decision-making. In addition to this it is proposed to involve the decision maker into the process of data processing and visualization by means of continuous interacting with the system, which helps optimizing the learning behavior of both humans and algorithms.

An AR system must address three overlapping concepts: tracking, registration and calibration [10]. To display virtual objects registered to real objects in 3D space, we must know at least the relative pose – that is the position and orientation of the AR display relative to the real objects. Because AR operates in real-time, pose measurements must be continuously updated, i.e. tracked over time.

In the field of AR, tracking of three-dimensional position of real entities is opposed to the idea of tracking two-dimensional objects in image space, which is common in traditional computer vision. Calibration is the process of comparing measurements made with two different devices – a reference device and a device to be calibrated. The reference device can be replaced with a known reference value or with a known coordinate system. The objective is to determine parameters for using the device to be calibrated to deliver measurements on the known scale. For AR we need to calibrate the components of AR system – especially the devices used for tracking.

While tracking means performing measurements continuously, calibration is usually carried out only at discrete times. Registration in AR refers to the alignment of coordinate systems between virtual and real objects [11]. Computer graphics elements should align with real-world objects. This requires tracking of the user's head or of the camera providing a video background, or both. Obtaining static registration, when the user or the camera is not moving, requires calibration of the tracking system so as to establish a common coordinate system between virtual and real object. Obtaining dynamic registration, when the user or the camera is moving, requires tracking.

Eye tracking may have the broadest and most profound impact of all human facing signals given that eye tracking measures human attention. Rendering processes are redesigned only to show user a portion of what they are looking at in full detail. Those techniques act on the scientifically proven fact that our eyes have a narrow field of view at full resolution and blurring occurs outside the fovea [12]. This is what accented visualization is based on.

First results of focused or accented visualization were presented in [13], [14], [15]. Based on an experience of their implementation in practice there was developed a concept of focus and context combination in AR based solution to improve decision-making support in practical applications. As an example, the proposed approach was implemented in a specialized intelligent system for industrial manual operation control. Such a system implements the ideas of Industry 4.0 for smart manufacturing by introduction of cyber-physical systems [16], [17].

III. AR CASE STORIES

In addition to the theoretical overview in this section there is provided a description of authors' practical experience of several projects targeting implementation of AR solutions in real cases.

The first project was dedicated to development of a customized AR based interactive guide for a robot design kit. The project was successfully performed for for Abilix robot creative brick series, using tablets or Epson Moverio goggles as AR devices. Resulting AR software solution can be deployed on a tablet or AR goggles and used by children to make the process of finding the small pieces of robot easy and comfortable. First experiments and practical use of AR devices at schools have proven high potential of their application and interest demonstrated by students.

Modern robot design kits provide powerful capabilities of developing various constructions with programmable behaviour starting from elementary to advanced levels of complexity. These design kit has high perspective to be used for education purpose to train a child's creative, analytic, and practical abilities. Due to this fact they are widely implemented to educational process at middle schools.

Technical challenges of implementation of AR use as a part of AR guide for robot design kit are concerned with low performance and quality of image recognition algorithms capable of functioning in real time. The main problem of AR implementation was concerned with stable and reliable recognition and identification of the objects in view.

For example, one of the main challenges of AR user guides is a necessity to identify the object that is required at the current moment, attract the user's attention and give complete and comprehensive annotation. To provide such features the technology should be capable of processing substantial number of images in real time.

AR based solution was dedicated to support the process of finding small pieces (details) easy and comfortable. There was developed special software (see Fig. 1 – 3) that can identify the components using a combination of existing image recognition techniques.

Intelligent systems for manual operation control are close in use cases and functionality to interactive user guides that also implement AR for decision making support at certain stages of operating processes. Intelligent software provides image recognition of the objects and their matching with a corresponding description in a knowledge base. Video panels or AR goggles are used to present the corresponding contextual information to an operator.

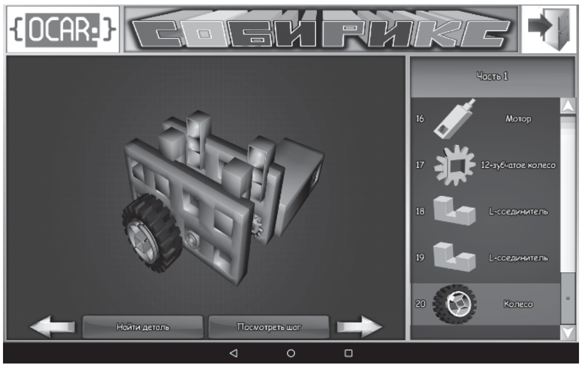

Fig. 1. Robot design navigator

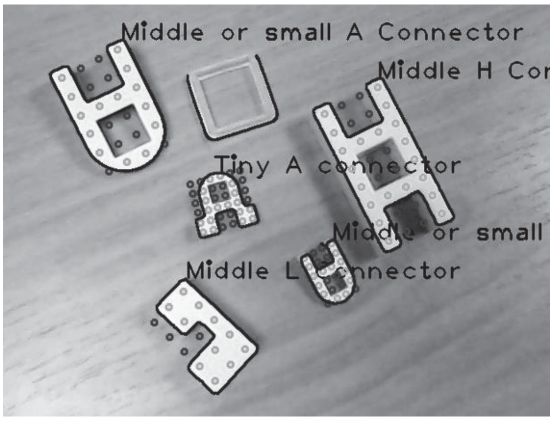

Fig. 2. Design kit components identification

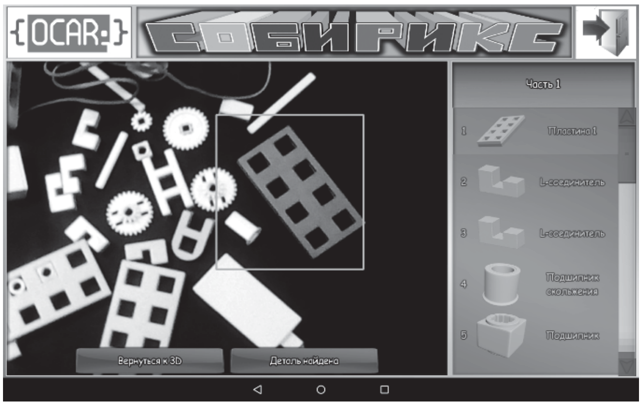

Fig. 3. Identification algorithm result

One particular truck part – radial compressor (turbomachinery) was taken as an object to recognize. Objects identification is illustrated by Fig. 4 – 7. The parts of required object are contextually highlighted according to the production process.

The main condition of the experiment was the use of an ordinary cheap web-camera for the task of recognition and tracking. The requirements for solution were: it should visually (in AR mode) step by step demonstrate the process of assembling the layout in accordance with the procedure described in the technological process for this product; it would display an error if the product is not assembled in accordance with the technological process; it would display an error if during assembly of the product typical errors are listed that are listed in the process.

Fig. 4. Intelligent manual operations' control. Working place

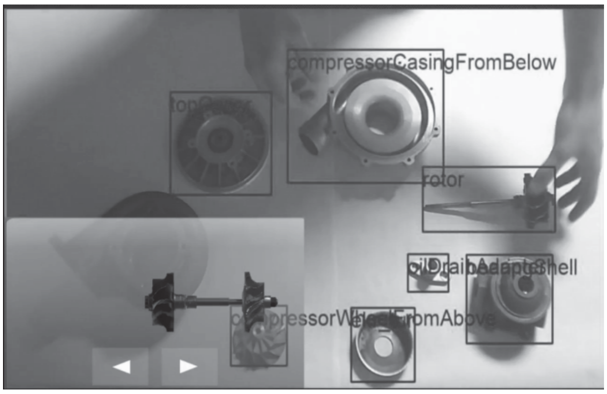

Fig. 5. Intelligent manual operations' control. Objects identification

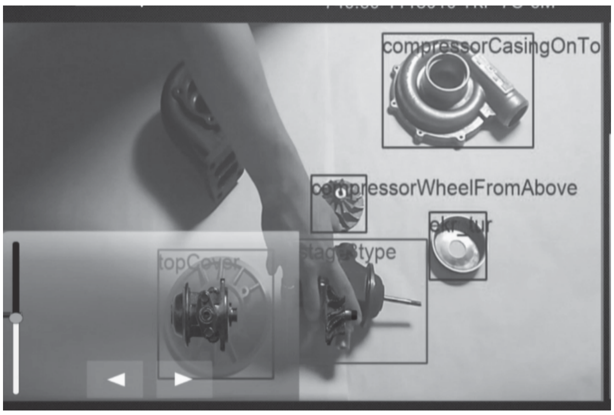

Fig. 6. Intelligent manual operations' control. Relevant object identified and highlighted

Fig. 7. Intelligent manual operations' control. Head mounted device

Software controls the correctness of the product assembly as much as possible. The initial data were the assembly process with typical errors, the digital model and the assembly video. The main parts of the product are following: compressor wheel, cover, compressor housing, bearing housing, turbine housing, turbine screen, oil-drain adapter, rotor, bearing housing with a rotor, redial compressor assembled.

However, there were revealed the problems of using AR goggles in practice due to a necessity to personalize them and provide continuous usability and comfort.

In this respect there were explore other possible ways of the system deployment using cheap and widespread equipment. A number of video cameras are used to track operations according to production process and identify the objects to be operated in real scene.

The overall solution is used to identify gaps and failures of operator in real time, predict possible operating mistakes and suggest better procedures based on comparing the sequence of actions to an experience of highly qualified operators.

IV. Accented Visualization Model

In this section, there is proposed a combined formal model for contextual visualization of overlay data using an AR device. This model was originally developed for maintenance manuals of technical equipment and based on the experience of AR and VR projects but can be extended and used in industrial applications. The model described below is proposed for an example of manufacturing decision making support using AR.

The accentuated notion of visualization is based on adaptive construction and virtual consideration of the content of not only the current real scene in the field of view of a person, but also the viewer's experience that contains perceptions, points of view and expected behavior. This is especially true for the actual support of work and decision-making, when a correct solution to the situation is required. This system is allowed to see the solution of the situation performed by different performers with different education, knowledge and skills in a unified way.

Let us consider the scene s_j, where $j = 1..N^s$ is the scene number that contains a number of real objects $w_{i,j}$, $i = 1..N_j^w$ and corresponding virtual entities $q_{i,j,l}$, $l = 1..N_{i,j}^q$:

$$s_j = \left\{ w_{i,j}, q_{i,j,l} \right\}. \tag{1}$$

Objects are considered independent in view of various features of visualization in various scenes. Objects can have different displays in different scenes, but refer to the only enterprise in the real world.

These objects are considered equal. It is possible to describe this equality using the compliance matrix for all scenes:

$$C\left(w_{i_1,j_1}, w_{i_2,j_2}\right) = \begin{cases} 1, \; w_{i_1,j_1} \equiv w_{i_2,j_2}, \\ 0, \; elswise. \end{cases} \tag{2}$$

Attention to the viewer at a certain point in time is given to many objects (one or more). Each center of an object can be described by an event represented by a Boolean variable:

$$v_{i,j,k} = v\left(w_{i,j}, t_{i,j,k}\right) \in \left\{0,1\right\}. \tag{3}$$

This focus may end or require action. In this case, the sequence of actions can be considered as a scenario. For example, the service process for a particular device can be described by a scenario of actions organized by the sequence.

The work process is formulated in a standard scenario:

$$c_{j,m} = \left\{ e_{i,j,m,n} \right\}, \tag{4}$$

where $e_{i,j,m,n} = e\left(w_{i,j}, d_{i,j,m,n}, t_{i,j,m,n}, \Delta t_{i,j,m,n}\right) \in \left\{0,1\right\}$.

Events $e_{i,j,m,n}$ represent the facts that actions $d_{i,j,m,n}$ that refer to the objects $w_{i,j}$ have to be done at lap time $t_{i,j,m,n} \pm t_{i,j,m,n}/2$.

Effective manufacturing process requires correspondence of $e_{i,j,m,n}$ and $v_{i,j,k}$:

$$\begin{aligned} M\left(e_{i_1,j,m,n} \cdot v_{i_2,j,k}\right) &= e_{i_1,j,m,n} \cdot v_{i_2,j,k} \cdot \left[w_{i_1,j_1} \equiv w_{i_2,j_2} \right] \cdot \\ &\cdot \left[w_{i_1,j_1} \equiv w_{i_2,j_2} \right] \cdot \left[t_{i,j,k} = t_{i_1,j,m,n} \pm \Delta t_{i_1,j,m,n} \right]. \end{aligned} \tag{5}$$

AR implementation in this case should support maximum correspondence of $v_{i,j,k}$ to $e_{i,j,m,n}$ and result in prompt data contextual visualization according to the current viewer's focus. It can be formalized as generating the set of focus attractors in textual, mark or highlight form:

$$\mathrm{Q_{i,m}} = \left\{ q_{i,j,l} \right\}, q_{i,j,l} = q\left(w_{i,j}, t_{i,j,l}\right) \in \left\{0,1\right\}. \tag{6}$$

to meet the following objective:

$$\begin{aligned} I\left(s_j, c_{j,m}\right) &= \sum_{n=1}^{N_{j,m}^c} e_{i,j,m,n} \cdot \\ &\cdot \left(1 - \left[\sum_{k=1}^{N_j^v} \sum_{l=1}^{N_j^q} \sum_{i=1}^{N_j^w} \left(M\left(e_{i,j,m,n} \cdot v_{i,j,k}\right) \vee q_{i,j,l} \right) > 0 \right] \right) = 0 \end{aligned} \tag{7}$$

This means that, in order to provide a standard AR scenario, an interactive guide should produce virtual enterprises with a context that leads the data in moments when the viewer's attention should be drawn to specific objects.

V. Solution Approach

Based on the methods of contextual data visualization and interactive computer-human interaction there was developed a concept of accented visualization. System architecture is presented in Fig. 8.

The concept is based on formalization of <focus, context, and overlay context> for each actor in Operator Ontology. Context is a set of concepts that describe the current situation and background that determines the decision. Focus is a concrete object processed at a certain moment. Overlay context includes virtual entities (textual items, marks or highlights) that attract user's attention to the required scene objects when needed. Such fragmentation allows introducing a control loop, where the correct focus is stimulated according to the context in real time.

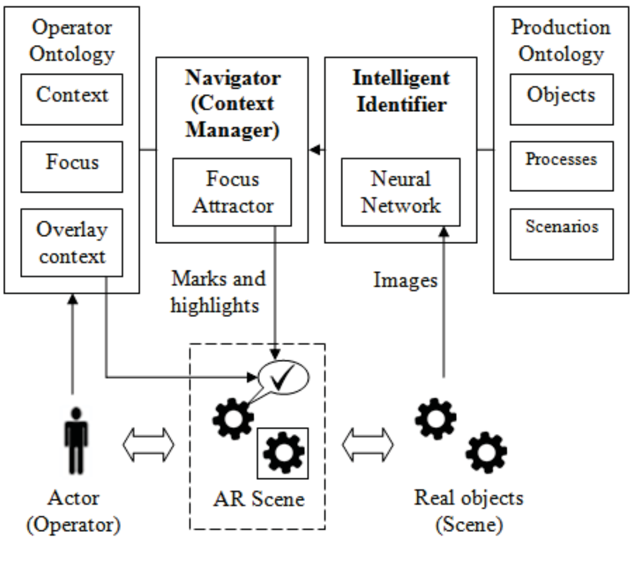

Fig. 8. Accented visualization solution for interactive user guides

Navigator module (context manager) implements user's focus coordination is based on intelligent analysis of production processes. User's focus is captured in the form of event chains and compared with typical scenarios. The system tracks user attention and adapts additional data of overlay context introduced to virtual scene according to the current need.

Identifier module implements intelligent pattern recognition using e.g. neural networks. It is a basic component of the whole solution and provides identification of all the objects in view and Navigator data support.

Production Ontology is a knowledge base that contains the description of production processes, technologies and necessary equipment as well as critical features or characteristics of each object. Rules of their processing are taken from production processes and coordinated with other scene objects and steps of operating scenarios. Eye tracking software and hardware is introduced to identify possible gaps in viewer's perception, if no required attention is given to certain scene objects at necessary times.

Despite the variety of shapes of the details most of them can be identified using the same approach. Some components look similar; some differ only in size or color. In addition to this AR device should be capable of identification of overlapped details with the lack of required visual data. To overcome this difficulty and increase the efficiency of AR interactive user guide there was developed a new approach base on accented visualization considering the current context.

Implementation of accented visualization approach that considers that combination of focus and context can be presented in the form of two layers: Identifier directly for intelligent pattern recognition and Navigator acting as a wrapper for Identifier. Navigator can be used for pre-processing of input data reducing obvious mistakes and interference. It can either be used for post-processing of identification results to improve their interpretation as a part of decision-making support.

Considering the context at the second layer can provide better targeting the intelligent algorithm or choose the most appropriate algorithm from a certain list. From this perspective it can be implemented using the formal logic predefined by an expert (like in the given above example). There can be proposed a special visual tool to describe this logic in the form of decision-making tree or semantic network as a part of ontology. This tool can get special user interface for several experts and be later used for the system configuration.

Navigator can also implement its own neural network as an alternative that is trained in a different way than a first layer and therefore becomes capable producing additional logic of data pro-processing and interpreting. Input of this layer in this case will take the distribution of objects identification probabilities from the first layer and determine the context in the form of i.e. stage of the process.

Possible issues of implementation of the proposed approach contain high dependency on the first layer that remains the main source of information and basic identification tool and complexity of decision-making trees or semantic networks in the second layer in case of processing multiple objects, stages, cases and processes. Nevertheless, this approach remains effective for use of intelligent technologies in practice and allows increasing their adequacy and sufficiency for industrial applications.

VI. Implementation

Implementation was divided into three stages. The task was to detect objects on the video stream, regardless of their location (objects could be on the floor, on the table, near foreign objects, next to other objects). Also, the object could not be in the frame at all.

At first stage, the task was to automatically determine the coordinates of parts of a product and assembled product on video images of a test sample. To provide adequate and stable identification of objects with a complex shape we developed an algorithm based on a combination of standard computer vision methods and neural network-based pattern recognition. Methods of identification of components that have a simple geometric shape and/or consist of elements with a simple geometric shape were taken from free sources and modified for the required quality and performance. The proposed identification algorithm includes the following steps:

- outer contour is identified for each object;
- the contour is drawn around the rectangle with the smallest possible area;
- the ratio of the length and width of the rectangle is calculated, which must correspond to the analogous value at the contour of the expected part;
- a grid of dots is applied to the rectangle;
- a map is drawn, describing which points are inside the contour (belong to a figure), and which are not;
- the compiled map is compared with a pre-prepared map;
- in case of coincidence, the parts are identified.

The task of the second stage was to realize the loading of multiple video files — data sets — after which the result would be output in the form of a csv-file containing a table with the coordinates of recognized product elements. Each video file was processed independently of the other video files in the sample, i.e. no additional information about other video files was used during processing. A log of the solution of the problem in real time was also kept. The average processing time for a single video file did not exceed the playback time of the video file itself.

At the third stage, a real-time solution was launched with a video stream received from a webcam mounted above the workspace. Software functionality on third stage included the following features:

- scene object identification based on image analysis;
- contextual description of the object in view;
- search and highlighting of the object required;
- operating scenario processing, tracking, and control;
- complex devices analysis including components identification by partial view and assemble tips generating;
- user attention identification and contextual add-ons generating according to the principles of accented visualization.

In order to provide the required performance in addition to the described algorithm there was implemented an approach of accented visualization. User's focus coordination was introduced and based on analysis of the design process. The system tracks user attention and adapts additional data according to the current context and need.

VII. Analysis and Tests

The problem of AR implementation was concerned with stable and reliable recognition and identification of the objects in view. For example, one of the main challenges of AR user guides is a necessity to identify the object that is required at the current moment, attract the user's attention and give complete and comprehensive annotation. To provide such features the technology should be capable of processing substantial number of images in real time.

One can see that no extra limitations for the working place and lighting (like e.g. green background) are required. Simple web cameras can be used to track and capture the objects: the quality of video is good enough for the most op-to-date models. Still it is recommended to introduce minimum two cameras to reduce the defects of lighting caused by occultation and blur. The quality of lighting turns out to be not so critical, which makes it beneficial being deployed at real enterprises. Considerable lighting changes require additional calibration.

Intelligent image recognition and objects identification in practice is usually performed using standard neural networks. There were considered several alternative libraries, including Tensorflow (which is the most fast), Keras, Theano (no longer supported by the developers) and Deeplearning4j (supports Java). Thesorflow was chosen for implementation, in addition to high productivity it is distributed under the open license Apache 2.0, provides access from Python, C++, Java, Haskell, Go, Swift API, supports Linux, Windows, macOS, iOS, Android, supports Google and cloud computing and gains high popularity among the developers.

Among the possible neural network topologies (AlexNet, VGG16, GoogLeNet/Inception, etc.) there was chosen ResNet-50, a Residual Network powered by Microsoft. ResNet-50 is a convolution neural network that is trained on more than a million images from the ImageNet database. It contains is 50 layers deep with direct links between the neurons located by one level.

Despite the variety of shapes of the objects in practice most of industrial components look similar. In addition to this the solution should be capable of identification of overlapped details with the lack of required visual data. For example, the assembly unit of two details is combined of a bearing case and adapter (see Fig 6). Adapter does not add much identity (adapter produces 10 – 15 % of the overall picture space), so the two pictures of bearing case and assembly unit look close to each other. Standard intelligent identification algorithm gets them mixed up.

The problems like this can be solved using the proposed concept of accented visualization. According to this approach intelligent identification algorithm is wrapped by knowledge driven logic that coordinates its target on the basis of correlation between the user's focus and context.

To provide the context driven decision-making support there was built a table of objects with possible distribution and probability of occurrence is view. Based on this table there is developed a decision-making tree that determines the sequence of stages and requirements for a details' list to be presented in view for each stage. Therefore the current stage determines the priority of objects identification.

For the given above example there can be specified a rule that in case for a collision of bearing case/assembly there is identified an adapter in view separate to an analyzed object with a probability of 0,7 then the object is identified as a bearing case. This logic corresponds to 5S methodology of workplace organization that requires no foreign objects exterior objects to appear in scene.

Implementation results for a real case of radial compressor assembly and identification of its details considering the requirements of production process are given in Fig. 9. The functions represent the dependency of a number of faults of identification step.

The faults were calculated in the following way. The video of identification procedure was recorded (duration time 9,13 min) and analyzed by an expert. Every 10th frame was taken for analysis and the object identified by the system was compared to the object required for identification according to the production process. The total amount of steps was 1271 and it is presented in the figure. Each step corresponds to a certain frame scene.

Two types of identification are investigated: considering the context and without it. The case without context, where neural network implemented only for intelligent identification of objects in view is illustrated by the upper function. Total amount of faults is 802, which takes the range of 63.1%. The second case considers the context and results with 86 faults which takes 6.8% of all steps.

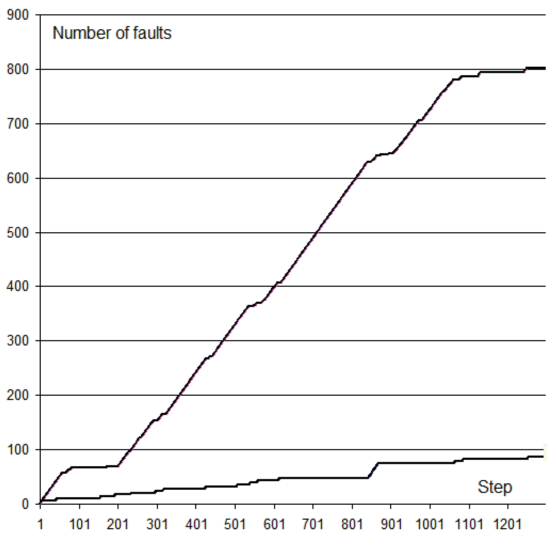

Fig. 9. Accumulation of identification faults (errors): considering the context (lower graph) and no context (upper graph)

This example shows that the recognition rate of objects successful identification using accented visualization if 93,2% comparing to standard technology that presents 36,9%.

The described intelligent system for manual operation control with the provided identification quality has won a specialized contest organized by KAMAZ, Skolkovo Foundation and a Foundation for Advanced Studies in a nomination of the best industrial solution in August 2018.

VIII. Conclusion

The results of a presented research help solving technical problems of AR implementation in real applications that is still critical in industry. This solution is proposed to be used in practice used to identify gaps and failures of operator in real time, predict possible operating mistakes and suggest better procedures based on comparing the sequence of actions to experience of highly qualified operators.

References

[1] N. Navab, "Developing killer apps for industrial Augmented Reality". Technical University of Munich, *IEEE Computer Graphics and Applications IEEE Computer Society*, 2004

[2] R. Krevelen, "Augmented Reality: technologies, applications, and limitations". Vrije Universiteit Amsterdam, Department of Computer Science, 2007

[3] S. Julier, M.A. Livingston, J.E. Swan II, Y. Baillot, D. Brown, "Adaptive user interfaces in Augmented Reality". *Proceedings of ISMAR '03*, Tokyo, Japan, 2003

[4] M. Singh, M.P. Singh, "Augmented Reality interfaces". *Natural Web Interfaces IEEE Internet Computing*, 2013. pp. 66-70

[5] K. Lee, "Augmented Reality in education and training". *TechTrends*. Vol 56, No 2, 2012, pp. 13-21

[6] D. Schmalstieg, G. Reitmayr, "The world as a user interface: Augmented Reality for ubiquitous computing". *Central European Multimedia and Virtual Reality Conference*, 2005

[7] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, "Deep learning, vol. 1" MIT press Cambridge, 2016, Vol. 925

[8] A. Holzinger, "Extravaganza tutorial on hot ideas for interactive knowledge discovery and data mining in biomedical informatics". *Lecture Notes in Computer Science*, 8609, 2014, pp. 502-515

[9] A. Holzinger, "Interactive machine learning for health informatics: when do we need the human-in-the-loop?". *Brain Informatics*, Volume 3, Issue 2, 2016, pp. 119-131

[10] Dieter Schmalstieg, Tobias Hollerer "Augmented Reality: Principles and Practice", 2016, Pearson Education

[11] Holloway, R. L. (1997) Registration error analysis for augmented reality. Presence: Teleoperators and Virtual Environments 6, 4, MIT Press, 413– 432.

[12] "The importance of Eye Tracking in AR/VR", 2018, Moor Insights & Strategy

[13] A. Ivaschenko, M. Milutkin, P. Sitnikov, "Accented visualization in maintenance AR guides". *Proceedings of SCIFI-IT 2017 Conference*, Belgium, EUROSIS-ETI. pp. 42-45

[14] A. Ivaschenko, A. Kolsanov, A. Nazaryan, "Focused visualization in surgery training and navigation". *Advances in Intelligent Systems and Computing*, vol 858. Springer, Cham, 2019. pp. 537-547

[15] A. Ivaschenko, P. Sitnikov, M. Milutkin, D. Khasanov, A. Krivosheev, "AR Optimization for Interactive User Guides". *Advances in Intelligent Systems and Computing*, vol 869. Springer, Cham, 2019. pp. 948-956

[16] H. Kagermann,; W. Wahlster, J. Helbig, "Recommendations for implementing the strategic initiative Industrie 4.0". Final report of the Industrie 4.0 Working Group, 2013, 82 p.

[17] H. Lasi, H.-G. Kemper, P. Fettke, T. Feld, M. Hoffmann, "Industry 4.0". *Business & Information Systems Engineering* 4 (6), 2014, pp. 239-242