

Creation of an Experimental Engineering Toolbox for the Digital Transformation of Manual Jet Engine Assembly

Konstantinos Spanoudakis¹ and Fabian Jonen² and Christian Borck³

B-TU Cottbus-Senftenberg, 03046 Cottbus, Germany

¹konstantinos.spanoudakis@b-tu.de

²fabian.jonen@b-tu.de

³christian.borck@b-tu.de

Abstract. Modern aircraft engines require appropriate care and understanding of design and manufacturing. This is even more important, as the production of aerospace engines remains a manual process in many cases with limited data sources. Quality control has to consider verification of manufacturing and assembly steps through specific checks and controls whilst implementing additional data sources.

In this article, a review of the challenges with regard to controls, automation and process, and technical understanding for aerospace engine production and repair is provided. As this requires the collaboration of many teams and partners, an improvement and step change towards deeper understanding and process efficiency is required. As many operations remain manual, innovations in the way humans interact with technology and collaborate with industrial environments are needed.

This article demonstrates the creation and usage of the proposed solutions for collaboration, troubleshooting and error correction in jet engine assembly.

Keywords: Industry 4.0, Autonomous Robot Programming, Collaborative Cognitive Robots, Augmented Reality, Virtual Reality.

1 Introduction

If transportation is a blood-vascular system of our world, then all motions elements of this are red blood cells. The systematic prophylactic is needed.

The area of the jet aircraft is a high frontier of the new technologies. It is jet aviation that allows people to travel the fastest traveling. Unlike high-speed trains, the network of air routes covers the whole world.

It has been proven that technological processes and quality control of aircraft engine maintenance are key factors for advancing the process of continuing air-worthiness support. [7]

Unlike many other modern high-tech areas like car and combustion engine assembly lines and production lines of electric motors, the assembly and maintenance of jet engines is quite conservative. In production and in service many manual labor operations

are used. This is a factor of an additional risk in spite of numerous checks and controls. Automatic assembly (e.g. with the use of industrial robots) is not applicable.

This makes cooperation difficult. But maintenance and repair of jet engines requires the coordinated work of a large number of people and factories.

To ensure the reliability of aircraft engines, it is of great importance to develop methods for complex multi-parameter diagnostics of technical condition. [15]

In the research projects Cockpit 4.0 and Insight Industrial Analytics - running until mid-2020 - Rolls-Royce Germany and the chair of automation of the Brandenburg technical university Cottbus-Senftenberg is already developing the first digital technologies in the field of assembly, logistics and the linking of development and production / assembly.

Important is not only the use of several manual operations during the maintenance and repair but the use of old communications and data recovery methods during a lot of these operations. It seems first of all about the use of “paper documentation” in the working area and the problem of its actualization.

To solve these problems, considering the restrictions of the production environment and the strict regulations in the aircraft industry, the following tools were chosen within the research project:

1. Object Detection
2. Collaborative Cognitive Robot (CCR)
3. Smart Assembly Model (SAM)
4. Smart Human Machine Interaction

This is already the first step forward in product understanding and increased operational efficiency.

However, the partners expect a significant quantum leap through the far-reaching combination of collaborative cognitive robot, smart human-machine interaction, operational machine-related data, logistics information and the exploitation of the so-called digital thread from concept, development, production / assembly, testing to service.

The approach can be generally applied to all maintenance fields. In this research, the before mentioned tools are described and the way they improve manual maintenance is examined.

2 Object Detection

2.1 The Detector

For the robotic object detector in the jet engine assembly, the faster R-CNN detector [17] is used in combination with the large Resnet101 network, which has a high average precision (more on that in the next section). Detectors are divided into region based and single shot detectors, where single shot detectors (SSD) have a much higher frame rate enabling to perform real time object detection at the cost of lower accuracy. Region based detectors like the mentioned faster R-CNN detector demonstrate an accuracy advantage if real-time speed is not needed.

According to the elaboration of Hui from 2018, compared to other detectors, faster R-CNN is not the fastest but one of the most accurate detectors [10].

2.2 Network Architecture

In the robotic vision application, the faster R-CNN detector sits on top of a neural network which was trained by the synthetic training data (mentioned in the next section). There are two methods to train a neural network – one where the network architecture will be assembled manually layer by layer with convolutional layers, pooling layers, fully-connected layers, hidden layers etc. It takes a lot of computing power and time to train the layers and their millions of neurons to detect low level contours, colors, shapes etc. in order to recognize objects. The other method is to use a pre-trained model which can already detect various objects like animals, furniture, humans etc. and only overwrite the classifier layers with the objects that need to be detected. However, in this method, the base layers (feature extraction layers) -where the neurons already learned how to identify contours, colors, shapes etc.- remain pre-trained. This is giving the advantage of a short training time and layers already trained by experts for state-of-the-art feature extraction. Within the research project the last approach was used. As a pre-trained model, the Resnet101 with the already trained coco image dataset [6] was used. Everything that needed to be done was the training of the top classifier layers with training data of objects that need to be recognized.

2.3 Synthetic Training Data Generation

To generate enough training data and increase the diversity of object appearance, we create multiple sets of rendered images. In these images, we differentiate the position, orientation, scale and material of the 3d object as well as the intensity of the light source. Two sets of 3d objects are sampled, root objects and non-root objects. For root objects, we create an animation sequence of 10.000 frames. The object is animated to rotate in Z and X axes during the frame range. This process provides images of the object at every angle with a 2 degrees threshold. Furthermore, is taken into consideration the pose of the object in the real world and no image is created where the object will not be accessible with the camera, thus, avoiding feeding the CNN with unnecessary data. For non-root objects the same process is followed, and the images created correspond to every angle of the referred object with a threshold of 4 degrees. In order to improve the perception of our CNN we use material diffuse colors as close as possible to real ones. To achieve that we use the actual camera that will be later used, to take photos of the object in different light conditions and different angles to receive realistic range of texture color. Subsequently we pick a range of RGB colors from the photos and use it as diffuse color in our 3d material. Throughout the animation mentioned above, the diffuse color of the object changes smoothly between all received colors. During scale and position transformations, we keep at least 25 by 25 pixels of the object's image inside the safe frame of the camera.

The parts to be inspected will not only be free in space but also parts of an assembly, thus the need for our CNN to be able to recognize parts fused with each other. To

generate appropriate data, we integrate the assembly steps into the 3d environment. Starting with the full assembly, we dedicate one render channel per object class and setting its name to the class representative. With all but the first layer of objects, according to fitting instructions, hidden and with regard in transparent background image format and alpha channel enabled we render the entire animation sequence. In each next step of the rendering process we unhide the next visible object class in the assembly and render the animation sequence.

After completion, if the assembly consists of 10 different object classes, thus assembled in 10 steps, we have 100.000 training images.

Furthermore, with the multi-channel rendering, each object which is visible in an assembly step, also has rendered images. The 550000 images of the multi-channel renderings are used only to generate annotation data in a VOC-Pascal format for each image.

The next section describes the approach of how to automatically generate annotation data for our training images.

Creating annotation data. With OpenCV the bounding boxes can be obtained for each object in the assembly sequence. Therefore, multi-channel images are created from one assembly scene. The main image of a specific assembly step has a different channel for each object. These multi-channel images only display the (partial) objects, which are not concealed by other objects. From these separated objects, OpenCV can detect bounding boxes of pixel groups. The coordinates of the bounding boxes will then be written into the annotation file. All the bounding boxes of all the objects from one image (main image) of an assembly step will be gathered in one annotation file. Both the main image and the annotation file will then be fed as training data into the neural network. This method allows to automatically generate training data from 3d models in any partial assembled state, allowing the classification layers of the neural network to recognize partial concealed or assembled objects. Combining the large number of synthetic training images, with a small number of real images with ground truth human annotations, will form a training image set with high performance [8].

2.4 Overfit-Resistant Image Synthesis

We synthesize thousands of images for each 3d model. While the realistic effect considered to be better, images with simple lighting conditions and materials prevent deep CNNs from picking up unreliable patterns. We simplified the creation of clutteriness embedding all important randomizing elements together. For image rendering, we explore parameters of rotation, scale, move and light intensity. For the position, scale and orientation condition, we sample values according to the original size of each object as for the deformed object to occupy enough space in the final image thus, to be exploited by the CNN. For the lighting condition, intensity of the light source is sampled. Images rendered as above have a fully transparent background, and the object boundaries are highly contrasted. To prevent classifiers from overfitting such unrealistic boundary patterns, we synthesize the background by a simple and scalable approach [19].

For each rendered image, we sample multiple backgrounds of solid colors, photos of industrial environments, nature and of same-textured-parts assemblies. We use alpha-

composition to blend a rendered image as foreground and a scene image as background [19]. Images cropped by camera's safe frame and images with cropped parts by other objects from the multi-layer renders teach the CNN to recognize occluded or truncated images. Finally, we use 75% of our created images as a training data, while the rest 25% of it is used as validation data.

3 Collaborative Cognitive Robot (CCR)

3.1 Cognitive Vision

With Chapter 2 we created a visual feedback, where a robot can identify objects in his workspace and calculate rough positions in his base coordinate system. The rough estimation of an object's position in the 3d workspace of the robot can be achieved by using the knowledge of the size of the detected object, the constraint that the object always stands on a known height (floor) relative to the robot and the xyz-coordinates and angles of the robotic tool center point (camera position). Since the accuracy of the calculated pose will be far from good and only represent a rough estimation, a second iteration is necessary for the robot to get the exact position of the detected object – getting into close contact with the haptic feedback.

3.2 Haptic Feedback

Haptic calibration. For a robot it is necessary to know the exact object position or damage might occur. Making a robot calibrate itself is great for versatile workspaces without fixtures, granting possible human collaborators freedom in their own ways of working. The CCR achieves this self-adjustment with the help of camera-based object detection and a force-torque sensor. Both can be attached to any off-the-shelf collaborative robot. In the experimental study a force-torque sensor (FT-sensor) from Robotiq (FT300) and a camera module with the google coral edge TPU dev board was used. The process of self-calibration involves calculating a rough position of the object with the camera-based object detector and finally the haptic calibration. Within the haptic calibration, at first the robot slowly moves close to the object's estimated position with a TCP speed slower than 15cm per second. Then the robot will be close to the object performing a slower tactile move (5cm per second) towards the object's surface. When the surface is reached the force applied to the sensor should exceed 2 N or 0.12 Nm, which is gentle enough to not move the object if its weight is above 2 kg. The force threshold for recognizing the touched object should not be lower, since the FT-sensor has a certain noise level. When touching the first contact point of the object, two or more additional points are required for calculating the middle point and angle of the object in the base coordinate system of the robot. After the self-calibration with haptic feedback, the object's position will be accurate within ± 5 mm, depending on how the touching tip of the TCP is designed. A very small touch tip would grant the best accuracy, however, since the application is collaborative according to EN ISO/TS 15066, a surface on the robot must never be smaller than 1mm² resulting in a larger tip and a

lower accuracy. The resulting position error of $\pm 5\text{mm}$ can be challenging for follow up assembly or pick and place operations. To compensate the error, again the FT-sensor can be used for force-controlled operations.

Force-controlled operations. If the position of an object is not programmed or calibrated correctly, damage can be caused to involved components. To become more flexible and versatile a robot can be equipped with a force-torque sensor. With robotic skills like “spiral search” or “linear search” the robot can use haptic feedback to identify object positions or assembly positions simply by monitoring a force drop or rise on contact or successful insertion. Sometimes, the misalignment of the case makes the first insertion attempt fail. With the spiral search algorithm, the bushing is able to be inserted in a flexible way. Each time a bushing is inserted, the correct assembly position is used to correct misalignment or calibration errors. This way the robot learns from his assembly attempts to identify the exact position of the object.

Combining visual with haptic feedback clears the path for flexible and versatile collaboration between robots and humans while using self-calibration and force-controlled operations.

4 Smart Assembly Model (SAM) and Autonomous Robot Programming

With the usage of skills, workspace data and process data, robot programs can be generated and kept updated automatically. These key-components are gathered and managed in a Smart Assembly Model (SAM). SAM is a high- and low-level assembly resource planning system which integrates elements for robotics, humans and processes. It also serves as a data collector and sends out commands to robotic devices, machine-human interaction systems and other technology in order to control and guide assembly.

As a data collector, it stores CAD-files of assemblies and parts, 3d-models of assembly sites, digitized fitting instructions, point clouds, assembly data etc.

Through its management system it is able to perform a particular set of actions that include path planning, Image processing, point cloud processing, digital fitting instruction interpretation, animations, digital twin processing etc.

These days robot programming is a manual task. The programmer needs knowledge about the assembly routine and the workspace. Adapting robot programs in a changing environment is time consuming and increases non-value-added tasks.

With skill-based programming an autonomous programming system can be implemented which reads digitized fitting instructions (FI) and assembly control records (ACR) with semantic product data (assembly routine) and overlays instructions with 3d factory model (workspace).

Subsequently, the data is processed with a neural network in order to create a robot program consisting of pre-programmed skills and path plans. This automatically generated robot program can be reviewed by a robotic specialist and used as a framework to do fine-tuning.

5 Technology Enhanced Qualification (TEQ)

Technology enhanced learning enables the worker to learn new processes and work steps independently and immersively, without having to be instructed by a supervisor. It enables the worker to access work instructions, training and additional information on various work steps. The aim is to make the knowledge about the process accessible to the employee. The field of Technology Enhanced Learning for Automation serves on the one hand as a database for knowledge about process flows and on the other hand as a platform for training courses which can be demanded independently by the worker with the right technological means (VR glasses, CAVE, tablet, etc.). Corresponding learning scenarios are automatically created from digital fitting instructions and simulated installation paths (called keyframes). Embedded in a virtual shop floor and using reduced 3D models derived from the product data, immersive learning units can be automatically generated from existing digital fitting instructions.

After the worker has internalized a process in technology enhanced learning, he can train it alone or in groups. The training scenarios are generated from the "Smart Assembly Model". The results of the training can be analyzed and compared with reference values. The deviations can be used to determine which assembly steps are mastered and to what extent (good, sufficient, insufficient). In this way, recommendations can be made for training courses.

6 Smart Human Machine Interaction

We use a smart system for the interaction between robot and human that employs Augmented, voice/gesture control etc. Using that system, the operator can interact efficiently with the robot and its digital twin. Through that interaction data flow in both directions, from the user to the robot and vice versa. With the proper processing the data is converted into information that both the interacting parts can use. That information processing and exchange leads to gradual continuous development of knowledge. Through the abovementioned interaction with the robot and its digital twin the operator can see visualized trajectories and next steps, future poses and workspace, status, errors etc. Furthermore, collaborative skills like hand over can be controlled using the Augmented reality (AR) interface. AR can be used in many fields of everyday life and the economy [3]. Reduction of error search and correction can be achieved by using AR for assistance tasks in the areas of training [5] [2], assembly, repair [9] or automation [13]. AR technologies have great potential in commissioning and maintenance. For example, when servicing equipment or complex machines, it is important for the technician to have all relevant information available. This includes, for example, information about what activities are to be performed on this system and what has already been done in the past [18].

The research on possible applications of AR has been done for many years [11] [16] [5] because the AR and AR applications are used in many fields, such as project and factory planning. Areas of mechanical engineering and manufacturing automation are interesting for AR. Primary AR is used for the visualization of assembly tasks, machine

operation, or reparation processes. With a camera (e.g., integrated in the tablet or smartphone) the system automatically recognizes the component to which need to be repaired and allows all relevant information as well as the exact assembly instructions to flow directly into the real image. E.g., the automatic fade-in of the circuit diagram while working on electrical system components is conceivable. It is much easier for the technician to record very complex system and to check the relevant measured values [18].

Despite numerous research and development projects in the field of AR technologies, it has not yet been widely used in mechanical engineering. The main reason is because of the hardware. Since the first experiments from the year 2003 up to the increasing interest since the year 2008, the hardware available for AR applications was too heavy and / or too big or too expensive.

Now the actual stage of the hardware and software make a developing and a using of an AR-application easy and possible for everybody [4]. There is a huge amount of free software that can be used easily for complex applications. Thus, it is possible to program easy applications for complex mechanical problems [4].

The task of check and control can be simplified with a using of the AR as well. A practical example of using of the AR for this, is a project for Rolls Royce Germany [14].

7 Conclusion and Future Work

The current situation in the maintenance and repair of jet engines is far from optimal. The organizational and manager solutions and methods for technical processes, control of information and other things are not optimal either. These solutions solve the problem of reliability and safety in the field of engine building and engine maintenance. But such decisions require considerable time and material costs.

This is especially evident with the introduction of new engine models. Thus, classical solutions, which imply for example a large number of paper documents, are not flexible enough. This is unacceptable in a constantly changing modern world.

This problem can be formalized as the following points:

- Changing production environments require flexibility and versatility.
- Parts, workspaces and assembly routines can change.
- Adapting to change is time consuming and increases non-value-added tasks.

So, there is not one, but several similar problems. It cannot be enough to use only one new method or approach to solve them all. These problems require a comprehensive solution.

This article proposes and discusses the following methods and technological solutions to overcome the problem:

1. Object Detection
2. Collaborative Cognitive Robot (CCR)

3. Smart Assembly Model (SAM)
4. Smart Human Machine Interaction+AR

The article provides a basic description of these methods and solutions. Showing their main advantages and disadvantages.

The next step of the authors of the article is the practical implementation of several methods in the framework of a small test bench. The main focus will be on the design and use of Smart Human Machine Interaction, machine learning and CCR.

8 Acknowledgement

This publication is funded by the European Union Fund for Regional Development and supported by the Federal State of Brandenburg.

The project is COCKPIT 4.0 (CustOmized Collaborative Knowledge Pilot for Industrial Technology 4.0) with EFRE project number 80170503 (www.b-tu.de/en/automation/research/projects/current-projects#c174887).

References

1. Andulkar, M.: Development of Multimodal Collaborative Robot System using Hybrid Programming Methods. In: Shaker Verlag, Aachen (2018).
2. Bade, C., Hoffmeyer, A.: Industrielle Anwendung von Augmented Reality in der Fertigungsplanung bei der Volkswagen AG (2008).
3. Baitinger, M.: Virtual Factory – 3-D-Visualisierung in der Projektabwicklung von Lackieranlagen. In 11. IFF-Wissenschaftstage: Virtual Reality und Augmented Reality zum Planen, Testen und Betreiben technischer Systeme VR und AR – Automotive, Magdeburg pp.53-60 (2008).
4. Berger, U., Bilous, V., Noack, R., Andulkar, M.: Anwendung von der Augmented Reality für die Mensch-Roboter Interaktion bei der Fehlerbeseitigung und bei der Maschinenbedienung. In: INNTERACT 2016 3DSENSATION, Chemnitz (2016).
5. Buschbacher, J.: Augmented Reality in der Ausbildung – Elektroniker für Geräte und Systeme (2012), <https://buschbacher.wordpress.com/2012/12/12/augmented-reality-in-der-ausbildung-elektroniker-fur-gerate-und-systeme/> last accessed 2020/03/23.
6. Chen, X., Fang, H., Lin, T.-Y., Vedantam, R., Gupta, S., Dollar, P., Zitnick, C.: Microsoft COCO Captions: Data Collection and Evaluation Server. arXiv:1504.00325 (2015).
7. Dmitriev, S., Burlakov, V., Popov, O., Popov, D.: Technological Processes and Quality Control in Aircraft Engine Maintenance. In: Aviation, vol. 19, no. 3, pp. 133–137 (2015).
8. Handa, A., Patraucean, V., Badrinarayanan, V., Stent, S., Cipolla, R. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016).
9. Henkel, E., Reiter, D.-I. R.: Digitale Fabrik – Stand der Anwendung heute und Ausblick auf die Weiterentwicklung, Magdeburg (2008).
10. Hui, J.: Object Detection: Speed and Accuracy Comparison (Faster R-CNN, R-FCN, SSD, FPN, RetinaNet and YOLOv3). In: Medium (2018), https://medium.com/@jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359, last accessed 2020/03/23.