Projecting Product-Aware Cues as Assembly Intentions for Human-Robot Collaboration

Joe David[1,2(✉)], Eric Coatanéa[1], and Andrei Lobov[2]

[1] Tampere University, 33720 Tampere, Finland
{Joe.David,Eric.Coatanea}@tuni.fi
[2] Norwegian University of Science and Technology, Verkstedteknisk, 213,
Gløshaugen, Norway
joesd@stud.ntnu.no, andrei.lobov@ntnu.no

Abstract. Collaborative environments between humans and robots are often characterized by simultaneous tasks carried out in close proximity. Recognizing robot intent in such circumstances can be crucial for operator safety and cannot be determined from robot motion alone. Projecting robot intentions on the product or the part the operator is collaborating on has the advantage that it is in the operator's field of view and has the operator's undivided attention. However, intention projection methods in literature use manual techniques for this purpose which can be prohibitively time consuming and unscalable to different part geometries. This problem is only more relevant in today's manufacturing scenario that is characterized by part variety and volume. To this end, this study proposes (oriented) bounding boxes as a generalizable information construct for projecting assembly intentions that is capable of coping with different part geometries. The approach makes use of a digital thread framework for on-demand, run-time computation and retrieval of these bounding boxes from product CAD models and does so automatically without human intervention. A case-study with a real diesel engine assembly informs appreciable results and preliminary observations are discussed before presenting future directions for research.

Keywords: Intention · Human-robot collaboration · Multi-agent systems · Digital thread · Product-aware · Knowledge-based engineering · CAD

1 Introduction

The transition of manufacturing from the fourth to fifth industrial revolution places the well-being of the human workforce at its core and leverages the synergy between them and autonomous machines [17]. In a human-robot collaborative environment, this means that humans will work alongside fence-less robots that exchange intentions and desires in a seamless and safe fashion between them. This would enable flourishing a trusted autonomy between interacting agents that would contribute towards an overall efficient manufacturing process [17].

When working in close proximity with a robot as in the case of a collaborative assembly, the operator must be aware of robot intentions as they directly translate to the operator's safety. One way to do this is for the robot to express its intentions with the part it is interacting with and a popular approach for the same has been by augmenting the operator's reality by projecting intentions [3,22,26]. Using such augmented reality cues has demonstrated benefits in real manufacturing scenarios [8].

However, augmenting reality with head worn displays has proven not to be suitable for industrial environments due to bad ergonomics among others [12,14]. Using only a projector to do so has the advantage that it requires no equipment that the operator needs to wear, supports easy operator switching and supports simultaneous usage by multiple operators [25]. Further, spatially augmenting these projections on the product is said to reduce ambiguities and miscommunication as the operator is not required to divide attention between the task and an externally projected display [3].

To this end, this paper contributes with an approach to project spatially augmented product-aware assembly intentions. The novelty in the approach is twofold. First, bounding boxes are introduced as a novel information construct that approximates effectively and efficiently regions of interactions and are purposed to convey intentions associated with assembly and sub-assembly parts. Second, the approach taken to obtain them entails using the ubiquitous assembly design software that uses a digital thread framework for on-demand, online and dynamic computation of data that defines the bounding boxes and spatially augments the operator's reality. This is in contrast with most works in literature that manually extracts required information (e.g. wireframes [3] or reference geometries [27]) from the product's CAD model for the said purpose, sometimes using specialized software. Our approach, once completed, automates its extraction without human intervention and albeit simple, scales well with product subassembly parts of different sizes at any position within the assembly without the need for any reprogramming. Thus, it realizes a scalable approach that robots (or human operators) can use to project intentions about product assembly.

Such flexible approaches were deemed necessary in a recent study that reported interviews with automation and shop-floor operators [11]. Specifically, they expressed the need for intelligent robots that are updated automatically and aware of the product type it should work with in the context of human-robot collaboration. According to them, switching smoothly between products would better the efficiency of manufacturing processes. The work presented herein addresses these in the context of intent communication in human-robot collaborative product assembly.

The next section reviews existing works in literature related to projecting product-aware intentions. This alludes to research objectives of this study and a description of the research setting and scope in Sect. 3. Section 4 presents theoretical background for the approach described in Sect. 5. Section 6 presents results of the study and preliminary observations are discussed. The paper is summarized in a conclusion section in Sect. 7 that also presents future directions of this research.

2   Related Work

Projecting and consequently communicating intentions associated with a product between humans and robots is not a new idea in itself with many originating circa 1993 at least [16]. Notable works include that of Terashima and Sakane [25] that prototypes a 'digital desk' that allows two levels of interaction via a virtual operational panel (VOP) and an interactive image panel (IIP). While the VOP is used to communicate task dependent operations, the IIP streams the robot's workspace by a separate vision system with which the operator is able to convey target object intentions by touching with his/her hands. Thus, it did not capture the physical aspects of a collocated setup as collaborative relationships with robots were not necessarily the goal back then and such systems where used for "guiding and teaching robot tasks" as opposed to true intent communication and also worked in a single direction from the operator to the robot. Later around the same period, Sato and Sakane [21] added a third subsystem in addition to the VOP and IIP called the 'interactive hand pointer (IHP)' that allowed the operator to point directly at an object in the robot's workspace to convey his/her intentions thereby removing the need for separate a workspace or display. However, this too worked in the direction of the operator to the robot. Needless to say, the ability for the robot to convey intentions is crucial for operator safety.

More recent works include that of Schwerdtfeger et al. [22] that explores the use of a mobile head-mounted laser projector on a helmet in an attempt to do away with the discomfort of conventional head-mounted displays to display product-aware intentions. The device projects simple 3D aligned augmentations for welding points on the surface of the part the operator interacts with (a car door) while instructions are provided on a standard stationary computer monitor. However, the position of the weld points were defined off-line by a tracked pointer. Further, the device was later reported as "too heavy and big" for use as a head-mounted device [23]. A subsequent developed hybrid solution entails a tripod mounted projector that benefits from partially mobile - partially stationary degrees of freedom but requires careful pose estimation each time the tripod is moved [23]. Sand et al. [20] present '*smARt.assembly*' which is a projection-based AR system to guide the operator to pick parts from an assembly rack during manufacturing assembly. The projector projects nested rectangles as an animation on the label of the part the operator is supposed to pick. The projection also entails a 2D image of the digital 3D model of the corresponding step that is presented on a panel on one side of the assembly station which is separated from the assembly workspace.

Uva et al. [27] present a spatial augmented reality solution that bears a close resemblance to the work reported in this paper. As the system was built with the goal of projecting technical guidance instructions, they use the reference geometries in the CAD model of only the base (fixed) part as an occlusion model and not the assembled part. This, as they say, was to reduce effort in the authoring phase which is evidence to the difficulties involved in extracting required geometry data for projection by manual means. Further, to use reference geometries of only the base part to project intent in a collaborative human-robot

scenario would be a problem as the assembled part can take any form factor that is not accounted for while projecting intent and can be dangerous, say for example when a robot is placing a component that spans to an area where the operator is simultaneously working. Also, the solution makes use of a third party software, Unity, that requires specialized expertise for development not commonly found in run-of-the-mill manufacturing enterprises. Andersen et al. [3] present a product aware intention projection solution that tracks the object in real-time and projects wireframes of the object. However, the approach involves generating "a large" number of edge maps offline from the CAD model of the object. Further, it is not clear how the they manage to illuminate the parts of the door the robot works with. As noted previously, manual approaches can be prohibitively time consuming and difficult to scale to different object parts at poses not determined beforehand.

3   Research Objectives, Setting and Scope

Existing approaches reviewed in the previous section either require pre-processing of the CAD model of the product, uses manual techniques, requires special developer expertise or are unsuitable for use in collaborative environments. Thus, we identify a gap for a simple, scalable solution in the automation of intent projection methods for use in collaborative environments between humans and robots. To this end, this research sets the following as its objectives:

1. Realize a generalizable intent information construct that can be used to project product-aware intentions for product assembly.
2. Use it to do so in a manner with minimal human intervention, preferably with *in-situ* software.

The research is carried in a laboratory environment shown in Fig. 1a. It consists of a DLP projector (1920×1080) and a Kinect Camera (RGB-D) mounted atop a height adjustable table that acts as a collaborative working space between a table-mounted UR5 collaborative robot and a human operator. The experiments are conducted with respect to an assembly of a real diesel engine.

Although, we consider only pick and place tasks in this study, the presented concept may be used for other tasks that require the representation of part geometries in the manner proposed herein (e.g. screwing). Also, the scope of this paper does not extend beyond the identification of the information construct and a reflection on its use in the case of diesel engine assembly. Consequently, details of the digital thread framework that accesses and processes the CAD model have been intentionally left out after a brief overview. However, the reader is provided references to our previous work [6,7] that has them as the focus of the study. Further, we do not deal with the pose estimation problem of the part and assume that its pose is known. Pose estimation based on the CAD model has been the subject of several focused studies [10,18,19] and the approach presented herein is expected to be built upon any of them.
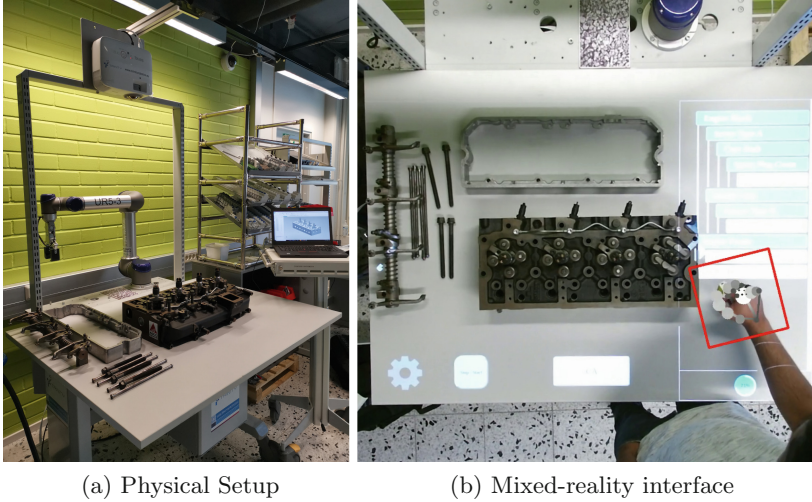
(a) Physical Setup   (b) Mixed-reality interface

Fig. 1. Human-robot collaboration setting

4   Theoretical Background

4.1   Oriented Minimum Bounding Boxes

In three dimensional euclidean space ($\mathbb{R}^3$), a minimum[1] bounding box around
a 3D object is the minimum or smallest cuboid (by volume) that completely
encloses it. If the edges of the bounding box are parallel to a coordinate axes of
a Cartesian coordinate system, it is an axis-aligned bounding box (AABB) with
respect to that coordinate system. On the other hand, if its edges are inclined
at an angle to the coordinate axes of a Cartesian coordinate system, then they
are oriented-bounding boxes (OBB) with respect to that coordinate system.

In this paper, we use bounding boxes to compute the minimum enclosing
cuboid of a sub-assembly part geometry. Depending on the geometry of the sub-
assembly part and how it is aligned within the entire assembly, the AABB may
or may not be the smallest enclosing cuboid, but an OBB will always be. For
this reason, in this study, we only refer to OBBs. AABBs will be the smallest
cuboid and same as the OBB when the part is positioned such that its OBB is
aligned with the coordinate system of the assembly part. Bounding boxes are
further discussed in the Sect. 5 with examples (Fig. 3).

4.2   Camera and Projector Model

A camera can be modelled using the pin-hole camera model [24] that describes
how a point in the 3D world is mapped onto its image plane. The projector too

---

[1] In this paper, we deal with only 'minimum' bounding boxes and we omit the word
'minimum' henceforth for brevity.

can be considered as an inverse camera where the rays of light are reversed, i.e. the light is projected instead of being captured [9]. Hence, the ideas underlying the calibration techniques that determine its intrinsic parameters used for a camera such as Zhang's [28], can be used for a projector as well [9].

The homogeneous transformation for a point X in the world coordinate system $\{\mathbb{W}\}$ ($\mathbb{R}^3$), to a point x in the pixel coordinate system $\{I_K\}$ ($\mathbb{R}^2$) of the image plane whose coordinate system origin is located at X_O is given by equation:

$$\mathrm{x}_{\{I_K\}} = \mathrm{P} \mathrm{X}_{\{\mathbb{W}\}} \tag{1}$$

where P is the direct linear transform (DLT)

$$\mathrm{P} = \mathrm{K} R[I_3| - X_O] \tag{2}$$

where K is a 3×3 matrix and defines five intrinsic parameters obtained through calibration and $R[I_3| - X_o]$ defines the 6 (3 translational + 3 rotational) extrinsic parameters or the rigid body transformation in a 3×4 matrix.

$$R[I_3| - X_o] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \tag{3}$$

K and Eq. 3 can be multiplied together to realize the transform as a 3×4 matrix substituted in Eq. 1 as

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{4}$$

where x, y is the pixel coordinate on the image plane of the projector of a point in the real world with coordinates X, Y, Z, both expressed in homogeneous coordinates and defined up to a scale factor.

5   Methods

5.1   Estimating Intrinsic and Extrinsic Parameters

We used a manual approach to establish correspondences between the projector pixels and the calibration landmarks (a printed planar checkerboard pattern) for calibration using the OpenCV library [4] to estimate the projector intrinsic matrix, K. To determine projector extrinsics, we used the Perspective-n-Point (PnP) pose computation method using similar correspondences, again using the OpenCV library.
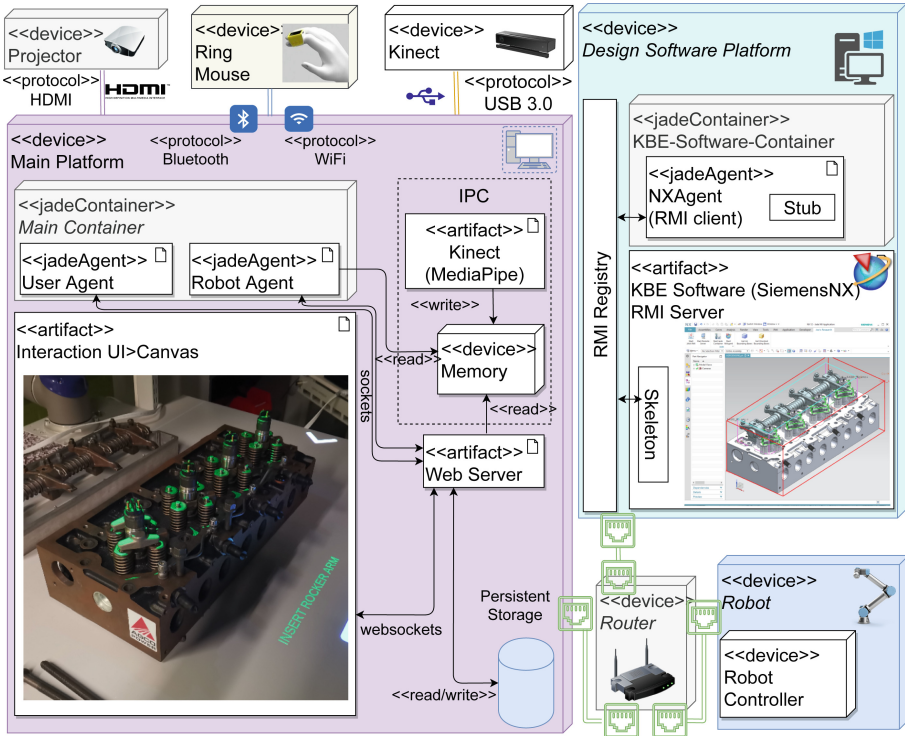
Fig. 2. Deployment diagram representing the architecture of the system

5.2   Digital Thread Framework

Figure 2 shows a condensed deployment diagram of the architecture of the digital thread framework employed in the HRC environment. The digital thread framework uses an agent-based framework (JADE) that integrates the assembly design environment (`Design Software Platform`) that exposes the product model via an API that provides with the data needed to project intentions from the product CAD model. It also consists of a purpose-built web application (`Interaction UI`) that is projected onto the shared work table that acts as a real-world canvas to project intentions and which the operator interacts with (Fig. 1b) to facilitate bi-directional communication with the robot. The operator's hand is tracked via an open-source hand recognition framework, MediaPipe (`Kinect`), while input is received from a ring mouse (`Ring Mouse`) worn by the operator. Further details pertaining to the digital thread framework, associated components and the interaction model can be found in our earlier works focused on the framework [7] and the web-based interaction model [6].

5.3   Product Design Environment

The digital thread framework maintains an online connection with the product design environment, Siemens NX. As a software also built for knowledge-based engineering (KBE), Siemens NX has a rich set of API, that allows to interact with the product geometry via the NXOpen API [2] and permits building digital thread applications [5] that help integrate product lifecycle information. It is with this API that the core functionality of our approach is realized.

When the robot is interacting with any of the sub-assembly parts, it requests the design software for the information it needs pertaining to the interacting part to project its intent. For a task such as that for an incoming placing operation which requires, conveying as intention, the relative position of a part in a base part (the part it is assembled into) but not the whole part itself, the coordinates for only the lower face of the bounding box is requested. For tasks that require, conveying as intention, the geometry of the whole part (as guidance instruction for example), the coordinates that defines the entire bounding box is requested. Thus, necessary information of any sub-assembly part can be obtained from the CAD file of the assembly and communicated to agents that require it dynamically at system run-time on request.
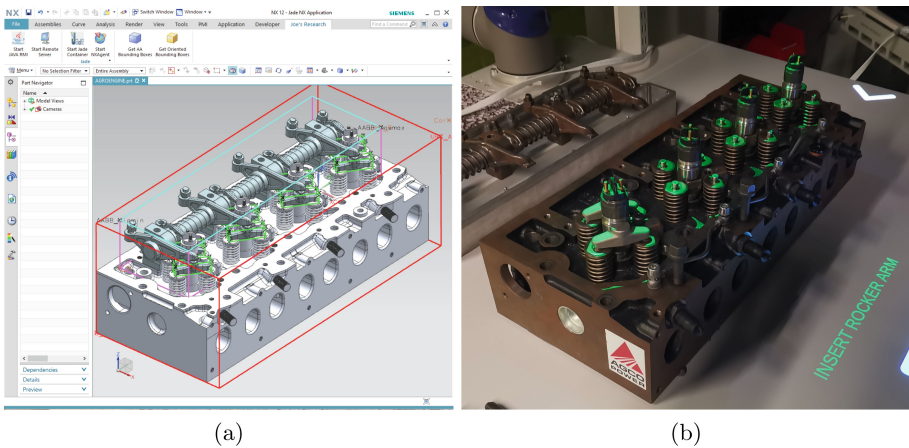

(a)   (b)

Fig. 3. (a) Computed bounding boxes viewed in the assembly design environment (b) Intent projection for the rocker arm

Figure 3a shows the CAD model of a real diesel engine loaded in the design software, NX. The diesel engine consists of many other parts but here only two sub-assembly parts (besides the fixed base part, engine block), a rocker arm shaft and rocker arm (8 nos), are shown to keep the demonstration concise and clear. However, the method scales well for parts of different sizes at any sub-assembly pose with the base part. Bounding boxes that the robot estimates are superimposed on the respective parts in NX and shown in Fig. 3a. Note the

orientation of the absolute coordinate system $\{\mathbb{P}\}$ (ACS) located at the bottom left of the viewport. The bounding boxes of the engine block (in red) and the rocker arm shaft (in pink) have their edges aligned along the axes of the ACS. Hence both their bounding boxes are axis aligned. However, the rocker arms are positioned such that they are inclined with respect to the ACS. Hence the computed bounding boxes are oriented bounding boxes (with respect to ACS).

5.4   Intent Projection

The robot agent uses the data it receives from the assembly design software to project intentions through a web-based mixed reality user interface [6] projected onto the shared environment (`Interaction UI` in Fig. 2 and Mixed Reality Interface in Fig. 4). Specifically, the robot agent uses the HTML5 Canvas API [1] to draw shapes and to write text to reveal its intent. As earlier mentioned, the current iteration of the development works on known poses of parts with respect to the real world external coordinate system $\{\mathbb{W}\}$. The coordinates that define the bounding box are computed from the CAD model and is transformed to $\{\mathbb{W}\}$ using the known pose and subsequently to the projector's image plane using the DLT (Eq. 2 & Eq. 4). Once the coordinates are mapped to the projector plane, a convex hull algorithm [13] calculates the smallest convex polygon that contains these points and fills it with colour using the HTML Canvas API. The overall steps taken for intent projection are summarized in Fig. 4:
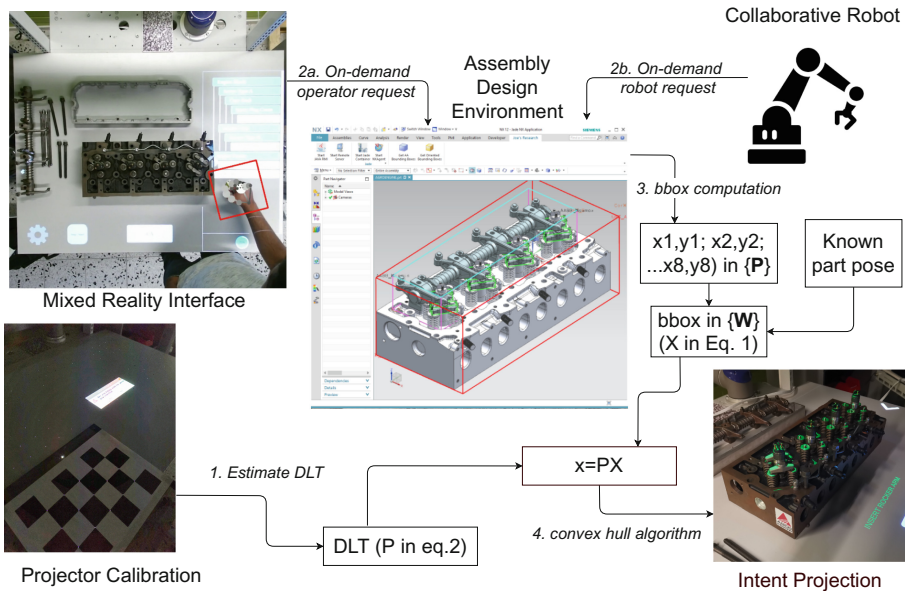

Fig. 4. Summary of steps involved in intent projection

6   Experimental Results and Observations

The approach described in the previous section is used to project intentions in a real diesel engine assembly in a laboratory environment. The results are shown for the two sub-assembly parts, namely the rocker arms and the rocker arm shaft in Fig. 3b and Fig. 5a respectively. While a comprehensive user study is not in the objectives of this paper, this section documents some preliminary observations made during the development and experimental process along with a general discussion.

6.1   Occlusions

Inherent to the single projector setup, the projections suffer from occlusions both from the operator and the robot. This can be seen in Fig. 5a where the shadow of the robot arm is cast onto the middle portion of the rocker arm shaft. Further protruding parts too cause occlusion. In Fig. 3b, it can be seen that the ignition coil occludes portions of the green bounding boxes projected on the rocker arm. While for small assemblies, such as the one presented here, this can be solved by skewing the projector or placing it vertically on top, larger assemblies are bound to suffer from occlusions from part geometry. However, considering the objectives of the study, this is not a limitation of the presented approach but that of the hardware setup. Depending on part geometry, multi-projector systems or a mobile projector setup [15] are two ways literature have minimized occlusions and with such additional hardware occlusions may be minimized.
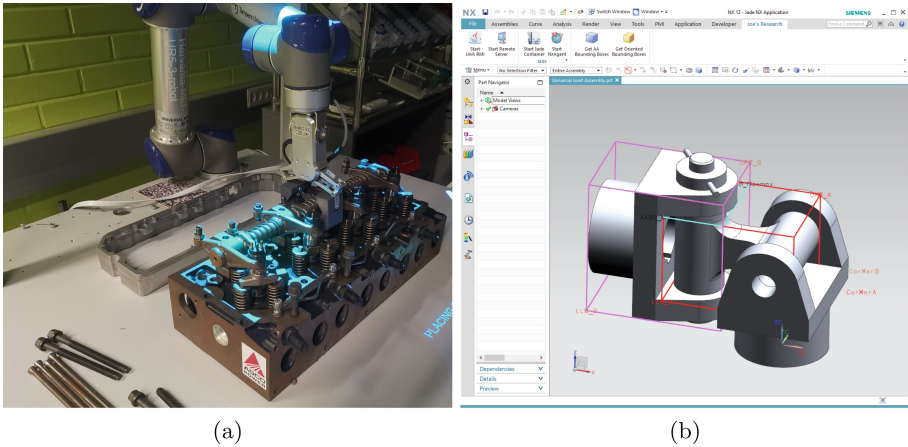

(a)   (b)

Fig. 5. (a) Intentions projected for the rocker arm shaft (b) Axis-aligned bounding boxes for a universal joint assembly and their intersection (cyan)

6.2   Bounding Boxes for Intent Projection

Bounding boxes generalize the problem of projecting intentions by approximating its shape quite well and do so efficiently. A 3D bounding box can be defined completely by six floating point numbers (minX, minY, minZ; maxX, maxY, maxZ) rather than, wireframes for example, that are defined by a series of points that define the boundary. Computing their intersection is a common and efficient method of collision detection that most if not all 3D software come with built-in functions for its computation. If not, it is possible to iterate through the geometry to compute them manually using an algorithm. Further, intersection of bounding boxes of assembly components can be used to locate the mating positions between sub-assembly components within an assembly when there is no third part involved (e.g. welded joints). Figure 5b shows the intersection of two axis-aligned bounding boxes of a universal joint in cyan to illustrate this. Note that the same code that was used in the diesel engine assembly case presented earlier was used, which demonstrates the scalability of the approach. In our work, while the robot uses bounding boxes to project its intentions to the operator, in a similar way the operator can use the mixed-reality interface (Fig. 1b) to request similar projection cues of sub-assembly parts. However, these projection cues projected at the behest of the operator are not used by the robot to perceive the operator's intentions. Rather, it is to reduce the cognitive load on the operator and to assist with the assembly process in general. To the best of our knowledge, such flexible functionality have not been implemented for intention projection purposes.

However, it can be argued that using bounding boxes can cause to loose the shape of the part geometry and thus loose the ability to identify the part from the box projection alone. In our work, we compensate for this by presenting textual descriptions that are automatically loaded from the design software with matching colors as that of the bounding box. Another issue with projecting bounding boxes is that while projecting intentions for a large part, the entire base part is illuminated. For example, the engine frame that lies beside the engine block in Fig. 5a is largely hollow in the center but spans along the edges of the engine block in its assembled position (not shown). As such, its bounding box would illuminate the entire engine block which can be difficult for the operator to understand given the box projection alone. In such cases, textual descriptions are important to prevent any confusions for the operator. On the contrary, the large bounding box encompasses all the areas that require to be clear of any activity to guarantee safety.

6.3   Digital Thread Framework

The digital thread framework provides important information necessary for correct projection of the bounding boxes from a (type of) software that is commonly used in manufacturing enterprises, i.e. the product design software. Thus, the requirement of a third party software that requires specialized development

expertise is avoided, C# for Unity as an example. Rather, modern KBE software vendors, expose their CAD kernel with a rich set of API that can be then used to drive applications that use the product in the manufacturing processes. NX, in particular exposes that CAD kernel via a Common Object Model that has bindings in four general purpose languages, Java, C++, Python and .NET which considerably reduces the barrier for such application developers. However exposing the entire CAD kernel with APIs means an overwhelming amount of programming constructs and our experience is that finding the right constructs to perform simple operations can sometimes be time consuming. However, as we got acquainted with the API, we experienced this less. Lastly, since such software is already well integrated with a traditional manufacturing enterprise, interfacing it with related systems is expected to be easy as was in our case and such solutions could be expected to be well received by the involved stakeholders.

7   Conclusions and Future Work

Collaborative tasks between humans and robots are becoming commonplace and recognizing intentions of agents that behave autonomously is pivotal in guaranteeing operator safety. The work presented in this paper presents a generalizable information construct in the form of oriented bounding boxes that is expected to foster greater situational awareness between agents engaged in collaborative assembly. The approach uses only the ubiquitous assembly design software and exploits the flexibility of a KBE software API to realize a scalable solution for on-demand, online computation of the required information dynamically at system run-time.

As future work, we aim to develop an information model or vocabulary that semantically grounds agent interactions. The work presented herein is expected to support the notion of agent intentions during these interactions. Another possible direction of future research includes pose estimation of the parts. We would like to investigate if we could, in a similar manner, automate the extraction of sufficient information that could train models for run-time identification and pose detection of assembly parts.

References

1. Canvas API - Web APIs: MDN. https://developer.mozilla.org/en-US/docs/Web/API/Canvas_API
2. Siemens Documentation: NX Open Programmer's Guide. https://docs.plm.automation.siemens.com/tdoc/nx/12/nx_api#uid:xid1162445:index_nxopen_prog_guide
3. Andersen, R.S., Madsen, O., Moeslund, T.B., Amor, H.B.: Projecting robot intentions into human environments. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 294–301 (2016). https://doi.org/10.1109/ROMAN.2016.7745145
4. Bradski, G.: The OpenCV library. Dr. Dobb's J. Softw. Tools **25**(11), 120–123 (2000)

5. David, J., Järvenpää, E., Lobov, A.: Digital threads via knowledge-based engineering systems. In: 2021 30th Conference of Open Innovations Association FRUCT, pp. 42–51 (2021). https://doi.org/10.23919/FRUCT53335.2021.9599986

6. David, J., Järvenpää, E., Lobov, A.: A web-based mixed reality interface facilitating explicit agent-oriented interactions for human-robot collaboration. In: 2022 8th International Conference on Mechatronics and Robotics Engineering (ICMRE), pp. 174–181 (2022). https://doi.org/10.1109/ICMRE54455.2022.9734094

7. David, J., Lobov, A., Järvenpää, E., Lanz, M.: Enabling the digital thread for product aware human and robot collaboration - an agent-oriented system architecture. In: 2021 20th International Conference on Advanced Robotics (ICAR), pp. 1011–1016 (2021). https://doi.org/10.1109/ICAR53236.2021.9659352

8. Doshi, A., Smith, R.T., Thomas, B.H., Bouras, C.: Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing. Int. J. Adv. Manuf. Technol. **89**(5), 1279–1293 (2017). https://doi.org/10.1007/s00170-016-9164-5

9. Falcao, G., Hurtos, N., Massich, J.: Plane-based calibration of a projector-camera system. VIBOT Master **9**, 1–12 (2008)

10. Fan, Z., Zhu, Y., He, Y., Sun, Q., Liu, H., He, J.: Deep learning on monocular object pose detection and tracking: a comprehensive overview. arXiv preprint arXiv:2105.14291 (2021)

11. Gustavsson, P., Syberfeldt, A.: The industry's perspective of suitable tasks for human-robot collaboration in assembly manufacturing. IOP Conf. Ser. Mater. Sci. Eng. **1063**(1), 012010 (2021). https://doi.org/10.1088/1757-899x/1063/1/012010

12. Hietanen, A., Pieters, R., Lanz, M., Latokartano, J., Kämäräinen, J.K.: AR-based interaction for human-robot collaborative manufacturing. Robot. Comput.-Integr. Manuf. **63**, 101891 (2020). https://doi.org/10.1016/j.rcim.2019.101891. https://www.sciencedirect.com/science/article/pii/S0736584519307355

13. Jarvis, R.: On the identification of the convex hull of a finite set of points in the plane. Inf. Process. Lett. **2**(1), 18–21 (1973)

14. van Krevelen, D., Poelman, R.: A survey of augmented reality technologies, applications and limitations. Int. J. Virtual Reality **9**(2), 1–20 (2010). https://doi.org/10.20870/IJVR.2010.9.2.2767. https://ijvr.eu/article/view/2767

15. Leutert, F., Herrmann, C., Schilling, K.: A spatial augmented reality system for intuitive display of robotic data. In: 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 179–180 (2013). https://doi.org/10.1109/HRI.2013.6483560

16. Milgram, P., Zhai, S., Drascic, D., Grodski, J.: Applications of augmented reality for human-robot communication. In: Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 1993), vol. 3, pp. 1467–1472. IEEE (1993)

17. Nahavandi, S.: Industry 5.0-a human-centric solution. Sustainability **11**(16) (2019). https://doi.org/10.3390/su11164371. https://www.mdpi.com/2071-1050/11/16/4371

18. Nguyen, D.D., Ko, J.P., Jeon, J.W.: Determination of 3D object pose in point cloud with CAD model. In: 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), pp. 1–6 (2015). https://doi.org/10.1109/FCV.2015.7103725

19. ten Pas, A., Gualtieri, M., Saenko, K., Platt, R.: Grasp pose detection in point clouds. Int. J. Robot. Res. **36**(13–14), 1455–1473 (2017)

20. Sand, O., Büttner, S., Paelke, V., Röcker, C.: smARt.Assembly – projection-based augmented reality for supporting assembly workers. In: Lackey, S., Shumaker, R. (eds.) VAMR 2016. LNCS, vol. 9740, pp. 643–652. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39907-2_61

21. Sato, S., Sakane, S.: A human-robot interface using an interactive hand pointer that projects a mark in the real work space. In: Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065), vol. 1, pp. 589–595 (2000). https://doi.org/10.1109/ROBOT.2000.844117

22. Schwerdtfeger, B., Klinker, G.: Hybrid information presentation: combining a portable augmented reality laser projector and a conventional computer display. In: Froehlich, B., Blach, R., van Liere, R. (eds.) Eurographics Symposium on Virtual Environments, Short Papers and Posters. The Eurographics Association (2007). https://doi.org/10.2312/PE/VE2007Short/027-032

23. Schwerdtfeger, B., Pustka, D., Hofhauser, A., Klinker, G.: Using laser projectors for augmented reality. In: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, VRST 2008, pp. 134–137. Association for Computing Machinery, New York (2008). https://doi.org/10.1145/1450579.1450608

24. Sturm, P.: Pinhole camera model. In: Ikeuchi, K. (ed.) Computer Vision, pp. 610–613. Springer, Boston (2014). https://doi.org/10.1007/978-0-387-31439-6_472

25. Terashima, M., Sakane, S.: A human-robot interface using an extended digital desk. In: Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C), vol. 4, pp. 2874–2880. IEEE (1999)

26. Uva, A.E., et al.: Design of a projective AR workbench for manual working stations. In: De Paolis, L.T., Mongelli, A. (eds.) AVR 2016. LNCS, vol. 9768, pp. 358–367. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40621-3_25

27. Uva, A.E., Gattullo, M., Manghisi, V.M., Spagnulo, D., Cascella, G.L., Fiorentino, M.: Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. Int. J. Adv. Manuf. Technol. **94**(1), 509–521 (2018). https://doi.org/10.1007/s00170-017-0846-4

28. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000). https://doi.org/10.1109/34.888718