STEP-based Digital Twin Model Construction for Assembly

Yazui Liu¹,a, Haodong Shen²,a, Xishuang Jing³,a, Gang Zhao⁴,b

¹Research Institute of Aero-Engine, Beihang University, Beijing,
²MIIT Key Laboratory of Aeronautics Intelligent Manufacturing, Beihang University, Beijing,
³Beihang University, liuyazui5684845@buaa.edu.cn
⁴Beihang University, zhaog@buaa.edu.cn

Abstract. Assembly is a crucial phase in the product manufacturing process, which influences ultimately product quality. Digital twin (DT) technology can mirror the physical product assembly in the digital space, enabling precise and quantitative analysis of assembly parameters. The acquired data and simulations can be implemented for monitoring and guiding the actual assembly process. Despite its usefulness, the large volume of heterogeneous data generated by various systems in the assembly process poses a significant challenge. This paper proposes a STEP-based digital modeling method that can be employed for assembly. Firstly, this paper presents the general framework of the assembly digital twin model. Secondly, based on the characteristics of the assembly process, the paper divides the assembly elements into the assembly object, auxiliary assembly equipment, and inspection devices. The assembly features are systematically extracted and classified, and the assembly process is expressed by portraying the relationships between these features. The paper constructs a data structure that adheres to the assembly process by referencing, extending, or redefining entities through the existing STEP standard. This enables information exchange between diverse systems in the assembly digital twin framework. Finally, the study employs a specific case. Results show that this method can describe all kinds of information in the process of product assembly well. From the results, it is concluded that this method can effectively solve the problem of heterogeneous data interaction in the process of digital twin assembly.

Keywords: Digital twin, STEP, Assembly model, Data fusion, Data exchange

DOI: https://doi.org/10.14733/cadaps.2024.1031-1044
1 INTRODUCTION

The assembly process is a critical factor affecting product quality and is also an engineering challenge, especially for precision products [16]. Complex product assembly is gradually moving towards digitization and intelligence, and digital twin technology, as one of the key technologies for achieving the integration of information and physical systems, has a wide range of application prospects in the assembly field. Essentially, digital twin technology achieves a true mapping of physical entities in virtual space through parameterized modeling of these entities [7]. By employing virtual and physical interaction feedback, data fusion analysis, and iterative decision-making optimization, it monitors and enhances the performance of physical entities.

With the in-depth study of digital twin technology in the assembly field and the development of high-performance digital equipment, intelligent assembly becomes possible. Currently, there is an urgent need for a standardized expression method to model the assembly process. At the same time, due to the complexity of assembly processes, a large amount of heterogeneous data generated by different systems leads to difficulties in information exchange. Structured semantic expression provides a reliable method for modeling the assembly process, demonstrating significant advantages when describing hierarchical relationships. The STEP standard represents a structured semantic representation based on its rich entities and flexible extensibility. This is the focus of the research in this paper.

Digital twin technology has undergone rapid development in recent years, primarily focusing on the construction of macro frameworks. To address the shortage of digital twin technology applications in industry and manufacturing, Tao et al. [10] first put forward the framework of digital twin shop-floor (DTS) and discussed the operation mechanism of DTS. In response to the high complexity and the large number of physical objects in the assembly floor environment, Zhuang et al. constructed a digital twin framework for intelligent production management and control on the shop floor, Zhuang et al. [18] construct the shop-floor intelligent production management and control digital twin framework. For the study of product lifecycle data is mainly focused on the physical space and ignores the processing of data in the virtual space, resulting the problem of data fragmentation, Tao et al. [8] proposed a method and framework for the application of product design, manufacturing and service driven by digital twin technology. Intelligent manufacturing is a highly automated and self-optimizing approach, Zhou et al. [17] have introduced an innovation framework that leverages the power of digital twin technology, dynamic knowledge bases, and knowledge-based intelligent skills to create the knowledge driven digital twin manufacturing cell (KDTMC). In the application of digital twin technology, it is necessary to collect and merge various types of data. Aiming at the problem of difficult selection of related technologies and tools, Qi et al. [4] put forward a five-dimensional digital twin model, which provides a reference for understanding and implementing digital twin. Yi [13] divided the digital twin model of intelligent assembly design into three layer, including physical space layer, interaction layer and virtual space layer, and applied them to the design of simplified satellite assembly process solutions. In assembly implementation process, Bao [1] researched an ontology-based assembly process semantic modeling method for the problem of difficult in formulating non-standard assembly processes.

It is evident that digital twin is a key technology for realizing intelligent manufacturing. However, current researches remain focuses on the definition of the framework, lacking standardized data definition solutions. In view of the numerous systems and poor compatibility involved in the design and manufacturing processes of complex products, an open and universal data framework is needed for data exchange and information sharing. STEP has been widely studied in recent years as a content-rich standard for product information exchange. X.F. [14] developed a prototype system consisting of a CAD system, a product modeling system, and an assembly planning and evaluation system based on the STEP approach for integrated design and planning of assemblies, thereby exploring the application of STEP in the field of assemblies by modeling the hierarchical relationships of assemblies. Mehmet et al. [5] addresses the difficulty of interpreting data semantically due to the lack of intuitive semantic information contained in the STEP modeling language (express), provides a intuitive semantic interpretation of geometric dimensions and tolerances (GD&T) in STEP based on OntoSTEP to
address the requirements of different product life cycle. Economen et al. [6] extended the STEP standard in the field of one-dimensional tolerance analysis and enable the exchange of data between different systems based on this model. Zha et al. [15] integrate the EXPRESS/XML schema-based model and NIST object-oriented UML-based open assembly model for assembly evaluation, which combines the STEP product definition and the fuzzy hierarchical assembly evaluation process with flexibility. Kwon et al. [3] proposed an intelligent product quality assessment system with decision potential in ensuring product quality by fusing design data represented by STEP and inspection data represented by QIF in the digital thread based on ontology and knowledge graphs.

Although STEP is currently the most widely defined standard data architecture, theoretically serving as a carrier for data interaction between different systems, existing research primarily focuses on using specific methods to address particular problems. There is a lack of comprehensive and detailed definitions to describe the elements involved in the assembly process. To address this limitation, this paper introduces the assembly digital twin model centered on assembly features and delineates the method of data interaction within the model by employing the STEP standard.

The rest of this paper is organized as follows. In section 2, the framework of the assembly digital twin model is constructed according to the actual assembly process. In section 3, a generic product assembly model is built in the STEP standard language EXPRESS/EXPRESS-G to represent the assembly data, including assembly model, tooling fixture model and inspection model. In section 4, the method is applied to a case study of a plunger pump assembly to demonstrate its applicability. And a conclusion is given in section 5.

2 ASSEMBLY DIGITAL TWIN MODEL FRAMEWORK

The actual product assembly process is characterized by an array of parts, equipment and processes. In order to achieve a true mapping of the assembly process between physical space and virtual space, a high-fidelity, full-scene virtual model needs to be constructed. In this section, a versatile assembly digital twin framework is proposed, as shown in Fig.1. The framework comprises four integral elements: the physical assembly space, the virtual assembly space, the data space, and the service space [12].

2.1 Physical Assembly Space

The physical assembly space comprises physical entities such as parts, tools, and measuring equipment utilized in the manufacturing process. These components come together according to the process documentation and undergo operations within the space. During these operations, data on the features of parts and the assembly process are collected. The acquired data is then processed in the data space and subsequently transferred to the virtual space, where high-fidelity modeling can be carried out or the service space for iterative computations.

To enable intelligent applications of the assembly digital twin, the assembly-related elements are divided into three categories: assembly objects, auxiliary assembly objects, and data perception objects. Assembly objects consist of product components. Auxiliary assembly objects include equipment that facilitates product assembly such as tooling systems and fixtures. Data perception objects are made up of various high-precision sensors and measuring equipment, which enable real-time or non-real-time data collection on assembly processes, features, and performance. The timeliness of data depends on its intended purpose. The data collected may include, but are not limited to: geometric accuracy data of parts, surface condition data of assembly features, interference check, process parameter test data during assembly, pre-tightening force measurement data, and stage performance data after the completion of each operation.

2.2 Virtual assembly Space

The virtual assembly space is a digital representation of the physical assembly space, integrating geometric, physical, and rule-based characteristics [9]. It consists of assembly resource modeling, data assimilation, and
Figure 1: Digital twin assembly model

Assembly resource modeling, which entails creating a proportionate entity in the virtual space aligned with the physical space, ensuring that postures, states, and other physical information in the virtual space match those in the physical space, is imperative in constructing the assembly digital twin model[11]. With data synchronization, data-driven methods accurately reflect service space data analysis and processing into the virtual and physical spaces. The assembly process in the virtual space ought to be the same as in the physical space, albeit the dynamic nature of the process in the physical environment calls for dynamic updating of the corresponding assembly twin in the virtual space. The model updating relies on synchronization results.
2.3 Data Space

The data space serves as an extension of the virtual assembly space, primarily responsible for data storage, processing, and transmission. Meanwhile, the service space focuses on data analysis and computation. The data space contains operation measurement data from the physical space, design data from the virtual space, and calculation result data from the service space, connecting the other three spaces simultaneously. As the data distribution center for the assembly digital twin model, the data space can manage all data during the assembly process in a unified manner while providing targeted information for different systems and reducing the handling of redundant data. Initially, the data processing uses the perception dataset from the physical assembly space, which has large volumes, heterogeneity, high dimensionality, and high noise characteristics. Data preprocessing techniques like noise point elimination, feature extraction, and structuring are utilized. After preprocessing, the data participates in the interaction of various processes. Therefore, constructing a unified data format can significantly simplify the processing time. For more details, refer to the next section.

![Figure 2: Digital twin assembly model](image)

2.4 Service Space

The service space is another extension of the virtual assembly space that focuses on data analysis and computation. It comprises various data processing modules that use data from the data space for assembly process quality monitoring, assembly process design, assembly quality prediction, and assembly quality optimization. The process quality monitoring module ensures that each assembly step is within the expected range, and if a problem is detected, it can be stopped in time for troubleshooting. The assembly process design module improves the information collaboration capability among product designers, process designers, and assembly operators by packing lightweight 3D models containing assembly instructions for output. The assembly quality prediction module is used in conjunction with the monitoring and optimization modules. Using a high-fidelity digital twin model, algorithms like Jacobian-torsor[2] and finite element analysis are employed for tolerance and force analysis, respectively, to predict assembly quality and performance. The assembly quality optimization module calculates the impact of assembly parameters on assembly quality based on data in the physical assembly space through sensitivity analysis after quality prediction or measurement, as shown in Figure 2. The calculated results can be used to evaluate the feasibility and effectiveness of compensation or adjustment. Real-time feedback can be provided to the assembly execution stage based on the compensation of the assembly process to complete the assembly parameter optimization. Therefore, the service space is a critical component in achieving closed-loop control of the assembly digital twin model.

The assembly digital twin model framework has been completed, enabling monitoring and perception of the assembly process through collaboration between the four different spaces. However, it is essential to note that there is a vast amount of heterogeneous data between different systems and modules in the assembly digital twin model framework. Therefore, it is necessary to establish a standardized data exchange format. In the upcoming section, a data structure for expressing assembly resources, based on assembly features and
under the STEP standard framework, will be defined. This definition will further promote the development of assembly digital twin technology by utilizing the structural characteristics of the assembly digital twin model.

3  STEP SCHEME FOR ASSEMBLY DIGITAL TWIN MODEL

STEP (ISO 10303) is an important means to achieve product information management and sharing. This standard addresses explicit representation and exchange mechanisms for machine-readable product information throughout the product lifecycle, providing the foundation for data exchange[19]. However, although STEP contains a large number of entity definitions, the standard has not yet been detailed in the assembly domain, resulting in the inability to directly apply the standardized model. Fortunately, STEP was designed with flexible extensibility from the outset. In this study, based on existing standard entity definitions, necessary extensions are made to construct a data model suitable for the information exchange of assembly digital twin model, ultimately achieving seamless data exchange and unified information integration in assembly digital twin.

According to the assembly objects and the auxiliary equipment involved in the assembly process, we establish three data models. The assembly project model, which represents assembly objects; the tooling fixture model, which directly participates in the assembly process; and the inspection model, contributing indirectly to the assembly process. These three models are defined using the EXPRESS/EXPRESS-G language, which inherently integrates object-oriented structures into the novel data structure. The primary STEP standard definition model examined in this research includes ISO 10303-203, ISO 10303-21, ISO 10303-42, ISO 10303-47, and ISO 14649-10. To accurately describe the characteristics of the assembly process product data, the research defines new entities by inheriting, combining, and extending existing standard entities. Based on this discourse, the following subsections describes the details of the definition for the STEP-compliant model for the assembly digital twin.

3.1  Assembly Project Model

During actual assembly operations, the complete assembly project typically comprises multiple distinct assembly procedures. As the smallest constituent unit of the assembly project, each assembly procedure necessitates a thorough analysis of objectives to be portrayed and develop an apt assembly model.

The assembly project is a structural representation of the entire assembly process, as shown in Fig.3. The construction of the assembly project entity unit begins with defining the geometry shape of the product, known as the nominal shape. The product name is then specified, and its attributes are defined as name. The technical requirement lists the necessary requirements for assembling the product. It is necessary to define a global coordinate system (placement) for assembly, which is used to describe the positioning of assembly features. The performance_index is used in this study to describe the performance of the product, including assembly deviation and running stability, with theoretical ranges and actual values given, respectively. Finally, the assembly procedure is defined.

As shown on Fig.3, in order to comprehensively describe the assembly process, the assembly procedure entity contains numerous attributes, such as its ID, operation_time, assembly_process, assembly_feature (primary and secondary), final_fit_state, assembly_environment and inspection. The core attributes of the assembly procedure entity are the assembly_feature, final_fit_state, and assembly_process, describing assembly objects, assembly targets and assembly operations respectively. In this study, it is assumed that pre-processing work, such as cleaning and grinding, has already been completed before the parts are assembled. Each procedure describes the process of fitting two assembly features of two parts. To distinguish between the two parts involved in the process, the assembly feature of the part where the assembly reference is located is called the primary assembly feature, while the other is the secondary assembly feature. It should be noted that the assembly features are attached to the parts; thus, it is critical to associate these features with the part to ensure accurate and precise assembly. The definition of part entity is shown in Fig.4.
The assembly process usually involves assembly features appearing in pairs. These features are classified into three categories based on their respective functions: single feature, replicate feature, and compound feature. The single feature consist of four basic types, while the replicate feature is a repeated combination of a single feature distributed in a specific pattern. The compound feature is combinations of two or more single feature. Four basic single features are assembly planar face, assembly shaft hole, assembly spherical, and assembly thread. The assembly planar face is used to depict plane fit features, assembly shaft hole represents the shaft-hole fit feature, assembly spherical outlines the spherical features, and assembly thread is used to characterize thread features in assembly. The attributes contained in each single feature the definition specifics are illustrated in Fig.5(a).

Once the assembly features have been defined, the next step is to model the assembly targets and the assembly operations. As shown in Fig.6, several constraint forms of assembly target state are defined according to the relationship between feature datum. For assembly operation, the first prerequisite for this operation is the definition of the location operation, which involves constraining the datum of two assembly features in accordance with specific rules such as coincidence, parallel and vertical. The datum is the fundamental unit, which can be either a point, line, or plane, that describes the assembly feature, and this varies depending on the feature in question. After completing the location operation, the mating or fastening operation is implemented.

During the multi-operation entity definition, the assembly path of this process is documented via the path_list. Based on this item, subsequent assembly path planning can be considered. Concurrently, the technology is used to describe the theoretical reference range and actual measured values of forces, moments,
and geometric requirements pertinent to the assembly operation.

At this point, all elements in the assembly object are defined.

3.2 Tooling Fixture Model And Inspection Model

The assembly digital twin model requires a comprehensive mapping of all elements involved in the assembly process from the physical space to the virtual space. Consequently, in addition to the assembly object, the data models should include fixtures, tooling and measurement-related equipment.

The tooling fixture entity has a crucial role in the assembly process, as it provides clamping, tightening and positioning functions. To accurately represent its position and shape within the virtual space, it is necessary to construct a CAD model of the fixture and tooling initially, followed by a detailed description of
Figure 6: EXPRESS-G of the assembly feature

Figure 7: EXPRESS-G of the assembly operation parameters

Figure 8: EXPRESS-G of the tooling fixture
the operation. Furthermore, the precision of this system must be measured and documented. Lastly, various types of tooling fixture can be identified, such as those with clamping and posture adjustment mechanisms or tightening mechanisms. These different types can be classified based on their unique characteristics within the assembly process, as shown in Fig. 8.

The integration of the inspection entity mainly serves to monitor and detect the assembly process. Through this model, not only can issues within the assembly process be promptly identified, but critical aspects of the assembly procedure can also be determined based on data, subsequently enabling a reinforcement of operations within these segments. Within this model, the data structure primarily consists of the measuring equipment and the content of the measurements. Considering the characteristics of the mechanical product assembly process, the measurement content include the geometrical morphology, the mechanical property, and the assembly environment. The explicit content is delineated in Fig. 9.

At this point, the data model for information transmission in assembly digital twin has been comprehensively defined. In the subsequent section, the efficacy of this methodology will be substantiated by illustrating the assembly process of a mechanical part.

4 CASE STUDY

In the laboratory assembly environment, taking a typical plunger pump assembly process as an example, the assembly digital twin model is constructed based on the aforementioned method. Fig. 10 shows the parts to be assembled, which are pump body, sleeve, plunger, stuffing cover, and bolts. By analyzing the plunger pump, the assembly project is broken down into four assembly procedures. As shown in the Fig. 11.

Simultaneously, the assembly features of the parts are extracted, the final fit state of the assembly features are defined, corresponding assembly operations are chosen, and the features and operations are linked to the assembly method. As a result, the feature and operation information of the parts can be combined via the assembly procedure, and the assembly procedures are combined via the assembly project to build a whole assembly process file. For a comprehensive assembly digital twin, the data interaction framework requires theoretical and actual measurement data of the assembly components and digitally represented tools and fixtures involved in the assembly process.

To achieve the digital twin of the plunger pump assembly process, a corresponding information exchange system must be constructed. Firstly, the assembly features of plunger pump parts are extracted from the CAD system and represented using a predefined data format. Secondly, assembly operations are instantiated based on each feature of procedure, involving the assembly guidance plan from the design phase and specific conditions of the actual assembly process. These conditions include the theoretical and actual values of geometric parameters in the technology. Lastly, by integrating the instantiated assembly features, assembly
operations and other attributes, the assembly procedures shall be obtained. Based on the defined entities, the information exchange system for the plunger pump assembly digital twin model can be constructed by instantiating the assembly process components, tools, fixtures, and measurement systems continuously. The specific customizations can be found in the appendix, as illustrated in Fig. 12.

In this case, we have completed the construction of the digital twin model for the plunger pump assembly. The model enables the seamless interaction of both physical and virtual data within the system. Consequently, the data interaction framework proposed in this paper for the STEP-based assembly process is effective and convenient, as demonstrated by the completed plunger pump assembly digital twin model.

5 CONCLUSIONS

To address difficulties caused by data heterogeneity during the construction of assembly digital twin models, we propose a STEP-based assembly information description framework. First, we construct the assembly digital twin model; then, we analyze the characteristics of mechanical assembly processes to extract three types of common assembly features and four types of assembly operations. Based on these results, the assembly project is a combination of several procedures, each with an information description that implements the coordination of two assembly features; the assembly process is connected by these procedures. Furthermore, we establish a data structure compatible with assembly logic based on the current STEP standard by inheriting, extending, and redefining STEP entities.

In conclusion, a complete data exchange method is essential for the proper application of digital twin
technology. In conclusion, a complete data exchange method is essential for the proper application of digital twin technology. Furthermore, we will pursue the development of service layer applications based on this framework to improve the service of assembly processes. Lastly, the goal is to develop a complete digital twin system for assembly.

ACKNOWLEDGEMENTS
The work is supported by National Natural Science Foundation of China (Project Nos. 62102011)

Yazui Liu, http://orcid.org/0000-0002-7480-1635
Haodong Shen, http://orcid.org/0009-0002-4968-7643

REFERENCES
Digital Twin towards Smart Manufacturing and Industry 4.0.


