IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 29 May 2023, accepted 27 June 2023, date of publication 3 July 2023, date of current version 3 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3291461

RESEARCH ARTICLE

An Innovative Approach to PMI Analysis and Enhancing Information Flow Efficiency

HU QIAO, SI-BO HU, LI ZHANG, JI-HANG LI, AND LIANG YOU
School of Mechatronic Engineering, Xi'an Technological University, Xi'an 710021, China

Corresponding author: Hu Qiao (qiaohu@xatu.edu.cn)

This work was supported in part by the Shaanxi Science and Technology Resources Open Sharing Platform under Project 2021PT-006; in part by the Shaanxi Innovative Talent Promotion Plan-Science and Technology Innovation Team under Project 2018TD-036; in part by the Shaanxi International Science and Technology Cooperation Base under Project 2022GHJD-14; and in part by the Advanced Manufacturing Institute, School of Mechatronic Engineering, Xi'an Technological University, Shaanxi Provincial Key Science.

ABSTRACT Model Based Definition (MBD) technology has evolved over the last two decades to become the core of digital product development, containing all Product Manufacturing Information (PMI) as a source of data for the future product lifecycle, and offers unparalleled advantages in terms of information storage and intuitive labelling compared to paper-based engineering drawings. However, as digitisation increases, the number of product types and the correspondingly more complex manufacturing processes, neutral data formats are often distorted or even lost when used across systems, leading to design and manufacturing errors. Therefore, there is a need to build a bridge between information interactions so that models can be accurately and efficiently transferred from one computer-aided design system to another. This paper proposes a method for unifying PMI structures by analysing STEP files to establish geometric dimensions and tolerances, annotating and decomposing several types of generic geometric elements into basic dimensional elements and positional information, establishing intuitive model-PMI relationships, and extracting the elements required for the exchange process and storing them in a STEP parser for PMI product information construction between different systems. Through experimental analysis, the method proposed in this paper has been shown to be highly accurate in exchanging information between three different CAD systems, achieving clarity, certainty and consistency at different stages of the product lifecycle. This STEP file approach enables structured analysis of data and significant savings in time costs and human resources.

INDEX TERMS Feature extraction, PMI, semantic representation, STEP, tolerance semantics.

I. INTRODUCTION

In today's machine engineering industry, computer-aided design (CAD) and manufacturing software is used throughout the entire production and maintenance chain and shows a trend towards increased complexity in technology and operational processes [1]. Digital technology as a process for assessing the degree of digitisation and exchange of product data has become key to the development of global manufacturing. The development model for the integration of product design, manufacturing processing, inspection, operation and maintenance throughout the product life cycle has evolved from 2D engineering drawings to 2D engineering drawings+3D models and then to MBD full 3D digital models [2]. With increasing specialisation, the dichotomy between design and manufacturing used to transfer information from different areas in the development phase [3] often leads to redundancy and errors in document information, while often lacking readability for the user. To address this increasingly prominent problem, data representation standards are even more important for digital manufacturing when exchanging products between CAD and manufacturing software applications [4], [5].

In terms of research in the direction of MBD data capture and visualisation, the visual and clear representation of geometric design intent to the user is essential to enhance the product model and its widespread use [6]. Therefore an intuitive display of geometric model information is an important

The associate editor coordinating the review of this manuscript and approving it for publication was Chaitanya U Kshirsagar.

requirement for the future planning and manufacturing of MBD model processes. Currently 3D geometric models are considered as carriers for information storage and a framework for information reuse is developed [7]. This approach, which is aimed at 3D geometric models as the object of study, is not fully applicable under the more complex MBD models of product model integration, especially as dimensional integrity checks are becoming increasingly important in product design. To address this issue, a number of verification methods based on computer-aided design and manufacturing (CAD/CAM) technology have emerged in recent years. One of the more popular methods is the intelligent reasoning method for dimensional integrity checking of mechanical parts with geometric element constraint states [8].

Furthermore, in terms of document readability, the PDF/E format has become the standard for the exchange and archiving of engineering documents in recent years [5], [9], yet user readability is still an issue in terms of the effectiveness of achieving this. STEP AP242, the standard protocol for MBD, states that MBD needs to include geometry data information, manufacturing and inspection information, etc [10]. In fact, this data cannot be used effectively. Among other things, although the syntax, structure and annotation of PMI in 3D models are consistent under the definition of this standard, in reality CAD manufacturers do not fully comply with this standard. Also different engineering and manufacturing applications represent the same PMI data differently [11]. STEP File Analyzer and Viewer, a software designed by the National Institute of Standards and Technology (NIST), is used for the syntactic, semantic correctness of the implementation of CAD exported STEP files in PMI [12]. Since the STEP AP242 protocol is based on the EXPRESS language, the PMI of the product data model can only be presented in the correct syntax, but not its semantics [13]. All these difficulties make product information interoperability still as an obstacle to be overcome [14].

Based on the above analysis, and in response to the increasing demand for information interaction, this paper proposes a construction method based on CAX-IF recommended practices to improve interoperability between systems and reduce information redundancy to increase document reading speed. The paper is then organised as follows: Section II introduces the MBD 3D semantic association and related research. Section III elaborates and explains the details of the proposed approach. Sections IV and V shows the experimental results and analyses the comparisons. Finally conclusions and future work are drawn.

II. RELATED WORKS

A. DATASETS UNDER MODEL DEFINITION

The manufacturing industry has developed to date and digital design has assumed an increasingly important role in it. Among the products in the design and manufacturing chain, such as machinery, vehicles, ships, electronics and many other complex areas of the manufacturing industry, data collection, data visualisation and data representation have been explored in depth and MBD methodology has been developed [15]. Among them, MBD integrates the design information, geometric information, manufacturing information and other information of the model with the 3D model in the form of a data set. And in 2003 in the American Society of Mechanical Engineers (ASME), the definition formed the digital product definition data practice ASME Y14.41-2003 [16], which was revised nine years later to ASME Y14.41-2012. the International Organization for Standardization (ISO) referred to ASME Y14.41 in 2006 to develop a technical product document "Digital Product Defining Data Practices ISO 16792-2006" [17]. These protocols provide a complete definition of the basic content of an MBD dataset, including engineering drawings, models, revision history, and provide a complete definition of parts lists, materials, annotations, and analysis applications.

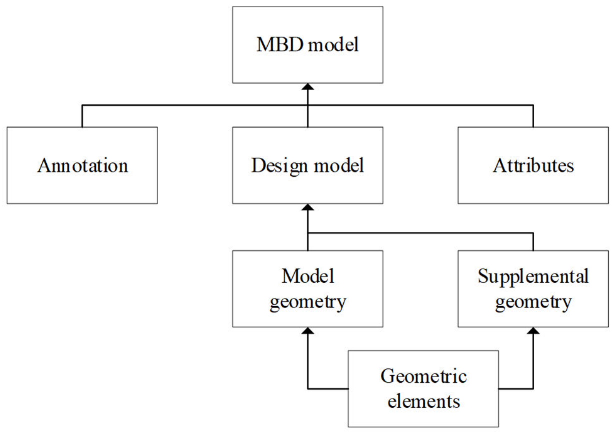


FIGURE 1. The content of MBD model specified in ASME Y14.41.

Unlike the 3D geometry model, the geometry model of the MBD dataset is integrated into the geometry model by integrating geometric dimensions, roughness and tolerance degrees from technical information [18], which is shown in Fig 1.

Where design model refers to the geometric definition and auxiliary geometry of the product. Notes are the dimensions, tolerances, roughness and various textual descriptions displayed on the model. Attributes are the dimensions, tolerances, notes, text or symbols required to fully define a product or function and can be obtained by querying the model [19].

ASME Y14.41 and ISO 16792 primarily specify the content of MBD datasets and product geometric dimensions and tolerances (GD&T) [20] for the symbolic language of engineering drawings and models, and later ASME Y14.9 and ISO 1101 further specify GD&T for the display of permissible errors in the geometry of defined features in three dimensions. Based on the establishment of such standards, the problem of user-oriented GD&T was basically solved, making the promotion of MBD a more solid basis. In order for MBD to support the implementation of 3D digital design and manufacturing, MBD data sets need to be represented as information that can be recognised and processed by computers.

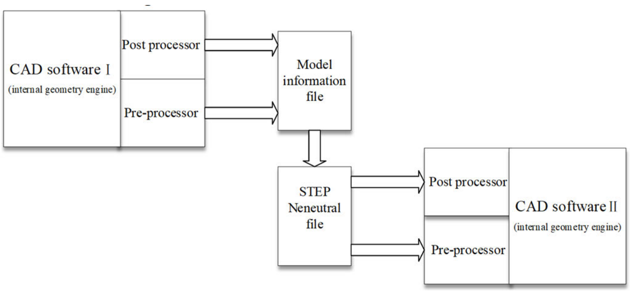


FIGURE 2. Data exchange between two CAD systems by neutral file.

Therefore, the existing Application Protocol (AP) and STEP (Standard for the Exchange of Product Model Data) parts 203 and 214 were built upon. The STEP AP242 neutral file, also known as an AP242 file [21], is the most commonly used implementation of the STEP standard today, enabling access to the MBD dataset of the system in which it is located and the interaction of data between different CAD systems via a STEP neutral file. This implementation defines a set of fixed text formats and fixed entity data representations that can be converted to neutral files between systems to enable the exchange of data between systems [22].

B. EXCHANGE OF INFORMATION BETWEEN SYSTEM

Data exchange and translation in CAD systems is usually divided into:

(a) direct translation between two CAD systems

(b) Conversion of model information into a neutral file format for information exchange

Since the direct interaction of data between different CAD systems is limited by the format, neutral formats are widely used for data exchange [23]. The neutral file exchange information is shown in Fig 2. Both CAD systems need to be processed by a pre-processor and a post-processor in order to interact with other systems for product data. The software outputs the model as a file of a certain format, which can also be read by the post-processor, usually in STEP, IGES, STL, etc. As the core component of the CAD system, the collection engine, the data exchange between CAD systems is in fact an internal collection engine that interchanges data.

AP242 gives important support to the application of MBD methodology in the case of STEP as a neutral format, offering great convenience to engineers in terms of information interaction. However, AP242 also has certain limitations: on the one hand, when the model file in which it is located is too large, the model tends to cause problems with the model file being too slow to read and taking too long to load when it is read. This can result in a lot of unnecessary time being spent on reading the file, so there is a need to further improve the read speed of neutral files. Furthermore, although AP242 provides a detailed representation of the MBD, ambiguities in the product manufacturing information arise due to defects in the CAD software itself, or due to the designers themselves not following the appropriate design rules in the model design. This problem is not avoided by converting the data to a neutral file in these and other situations that may lead to defects in the product. The first problem is to improve the read rate of STEP neutral files. After studying the structure of neutral files and the framework of the read process, a solution is proposed.

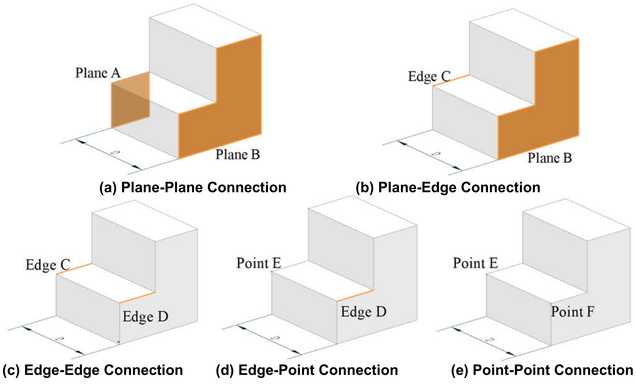


FIGURE 3. Example of a dimension-associative object.

C. MBD MODEL SEMANTIC ASSOCIATION

In this paper, tests in UG NX 12.0, CATIA V5 and Creo 9.0 revealed that there are various forms of association between 3D annotations and geometric features in 3D CAD software. This will have an impact on the subsequent understanding of the STEP file, which will not be able to achieve the correspondence between the geometry and PMI of the neutral file; it will be inconvenient for the subsequent design of the information interaction between the various CAD systems and the speed of the file reading, so it is necessary to study the semantic association of 3D annotation features.

PMI is used as supplementary data to product geometry to transfer product definition information in 3D CAD systems for product manufacturing, maintenance and inspection. PMI includes product dimensions, tolerances, material specifications, weld symbols, surface finishes, user defined property data and 3D annotations. As dimensional and geometric tolerance (GD&T) data is the most important information in PMI, in order to assess the ability of CAD systems to implement the AP 242 protocol, the STEP file analyser software developed by NIST performed verification and validation tests on the output files of each CAD system and found that inconsistencies still existed. Its information is mainly combined into model geometry through boundary representations (B-rep), which provide a rich and detailed description of the product. The B-rep in the definition of AP 224 works well with the machining feature library, but does not establish any relevant semantic information [24]. The AP 242 information standard is able to capture its information, however this combined representation does not provide accurate information on the key features or geometry of the part to be manufactured, making it only understandable to designers and process planners with a high level of work experience, which will greatly limit future rollouts. Therefore, before the data analyser is built, the model is grouped with different calibration measurement features according to the different calibration information.

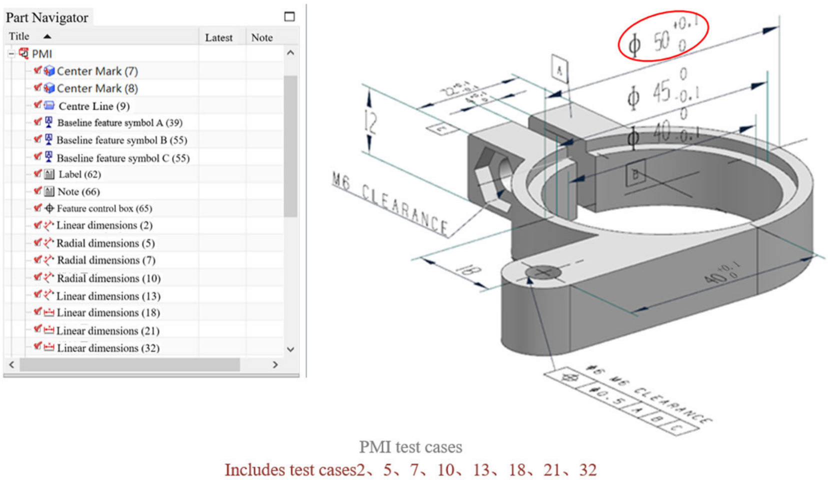


FIGURE 4. **PMI test case drawing.**

The tests were carried out in UG NX 12.0 software, by using the PMI annotation module of this software. The test artefacts are annotated in Fig 3. The measurement dimensions D are associated to different objects, 1) when face and face are associated objects (Fig 3(a)); 2) when face and line are associated objects (Fig 3 (b)); 3) when line and line are associated objects (Fig 3(c)); 4) when line and point are associated objects (Fig 3(d)); and 5) when point and point are associated objects (Fig 3(e)). When the above five are selected for dimension D, only a correctly expresses the designer's design intent, which is meant to indicate the distance between face A and face B. and not to indicate the distance between face B and line C, nor is it intended to indicate the distance between edges; therefore the way b, c, d and e are labelled, creates ambiguity and error.

If they can be expressed in a computer-interpretable language and the STEP file parser can be extended to read and process the language, then it can be the framework software used in any neutral file.

III. MODEL SPECIFICATION AND SEMANTIC CONSTRUCTION

A. CONSTRUCTION OF PMI

PMI constructs consist of one or more annotated entities whose constructs contain dimensional dimensions, datum features, feature control boxes (parallelism, position, perpendicularity, flatness, cylindricity, coaxiality, etc.). The test geometry model used in this paper contains 2, 5, 7, 10, 13, 18, 21, and 32 of the constructs. In the CAD example shown in Fig 4, the information marked as a red circle with a radial dimension of 50, an upper deviation of +0.1 and a lower deviation of 0 is sorted as 13 in the PMI construction and belongs to the equal bilateral tolerance dimension. The documentation of this radial dimension in the CAx-IF recommendation is shown in subsection 3.2.2 [25]. In the block diagram of this example, the rectangular box represents the entity name and the attribute is shown next to the connecting line between the entity names. The end of the connecting line connecting the rectangular box with a small circle connects the range of the attribute, while the side without the small circle belongs to the area of the attribute. For example, the dimensional_characteristic_representation as an entity has two attributes (dimension and representation). dimensional_location and shape_definition_representation represent the range of these two attributes respectively. representation represent the range of these two properties respectively.

Fig 4. shows a block diagram of the instantiation of geometric tolerances. Note that some properties of the entity classes are not labelled in the diagram in order to simplify the instantiation diagram.

B. SEMANTIC REPRESENTATION CONSTRUCTION FOR PATTERN SPECIFICATION

CAx-IF is a forum co-sponsored by ProSTEP iViP and PDES, Inc. to promote the development of CAx system conversions and to meet the needs of users for interoperability. models against which to compare their detection ranges. In order to achieve PMI interactivity in STEP files, this paper builds PMI constructs based on PMI representation in recommended practice. The constructed PMI can be used as a data source for general use with other CAx software to achieve PMI interactivity and also as a framework for PMI information consistency checking for STEP neutral files.

The scope of PMI as a non-geometric attribute for conveying CAD systems can be broadly divided into GD&T information, 3D text annotations, 3D symbols and custom attributes [26]. The construction of an axiomatic set of PMI domain terms as recommended by CAx-IF in practice describes the encompassing relationships between

TABLE 1. The Construct, syntactic and semantics representations associated with concepts in SROIQ(D).

Construct	Syntactic	Semantics
Concept	C	$C^I \subseteq \Delta^I$
Negation	$\neg C$	$(\neg C)^I = \Delta^I \setminus C^I$
Role	R	$R^I \subseteq \Delta^I \times \Delta^I$
Disjunction	$C \vee D$	$(C \vee D)^I = C^I \wedge D^I$
Conjunction	$C \wedge D$	$(C \wedge D)^I = C^I \wedge D^I$
Value restriction	$\forall R.C$	$(\forall R.C)^I = \left\{x \mid \forall y, \langle x, y\rangle \in R^I \rightarrow y \in C^I\right\}$
Existential restriction	$\exists R.C$	$(\exists R.C)^I = \left\{x \mid \exists y, \langle x, y\rangle \in R^I \wedge y \in C^I\right\}$
Cardinality	$\leq nR.C$	$(\leq nR.C)^I = \left\{x \mid \#\left\{y, \langle x, y\rangle \in R^I \wedge y \in C^I\right\} \leq n\right\}$
Restriction	$\geq nR.C$	$(\geq nR.C)^I = \left\{x \mid \#\left\{y, \langle x, y\rangle \in R^I \wedge y \in C^I\right\} \geq n\right\}$

all terms and SROIQ(D) based terms. For example, we can define the follo-wing domain terms:shape_aspect, geometric_representation_item, geometric_item_specific_usage,dimensional_location, product_definition_shape,plus_minus_tolerance, dimensi-onal_ characteristic_representation, shape_dimension_representation,tolerance_value, measure_with_unit. unit. The specific meanings of these terms are as follows:

1. shape_aspect is used to represent the concept of GD&T characteristics as a composite concept (which may be defined by another concept) and is defined as follows:
$shape\_aspect \equiv \exists ID.decimal \wedge \exists name.string \wedge \exists description.string \wedge \exists of\_shape.product\_definition\_shape \wedge \exists product\_definitional.boolean$
where ID, description, of_shape and product_definition are attributes of shape_aspect. decimal, string, boolean, product_definition_shape are ranges of the corresponding attributes and are constructors of SROIQ(D). To simplify the description, only the _shape attribute is labelled in Fig 8.
2. geometric_item_specific_usage is used as a concept for an intermediate entity connecting the shape_aspect entity to the geometric_representation_item entity and is defined in a simplified way as follows:
$geometric\_item\_specific\_usage \equiv \exists definition.shape\_aspect \wedge \exists identified\_item.geometric\_representation\_item$
3. dimensional_ location as a concept for representing dimensional locations with the following simplified definition:
$dimensional\_characteristic\_representation \equiv \exists dimension .dimensional\_characteristic \wedge \exists representation .shape\_dimension\_representation$
4. dimensionl_ character_ representation as a concept representing a dimensional feature with the following simplified definition:
$dimensional\_characteristic\_representation \equiv \exists dimension .dimensional\_characteristic \wedge \exists representation. shape\_dimension\_representation$
5. plus_ minus_ tolerance as representing the concept of having upper and lower tolerances, with the following simplified definition:
$plus\_\min us\_tolerance \equiv \exists range.tolerance\_value \wedge \exists toleranced\_\dim ension.\dim ensional\_location$
6. shape_ dimension_ representation indicates a dimensional tolerance with the following simplified definition:
$shape\_dimension\_representation \equiv \exists items.measure\_with\_unit$
7. tolerance_ value indicates a tolerance value, which is defined in simplified form as follows:
$tolerance\_value \equiv lower\_bound.measure\_with\_unit \wedge upper\_bound.measure\_with\_unit$
8. measure_ with_ unit indicates the type and value of the parameter being measured, which is defined as follows:
product_definition_shapeisusedtorepresenttheentityofthe product
9. dimensional_ characteristic indicates a dimensional or position-related characteristic with the following simplified definition:

$$dimensional\_characteristic \equiv dimensional\_location \vee dimensional\_size$$

The above nine expressions belong to the concept definition metric in SRIOQ(D). In the practice recommended by CAx-IF, another concept metric can be used to define the hierarchical relationships between classes. In the following expressions, angle_location is a subclass of dimensional_location and angle_size is a subclass of dimensional_size:

$$angle\_location \subseteq dimensional\_location$$
$$angle\_size \subseteq dimensional\_size$$

We can provide an interpretable semantics for computers by defining the expression of the above concepts. In $I = (\Delta^I, \bullet^I)$,where Δ^I is a non-empty universe of domain values,

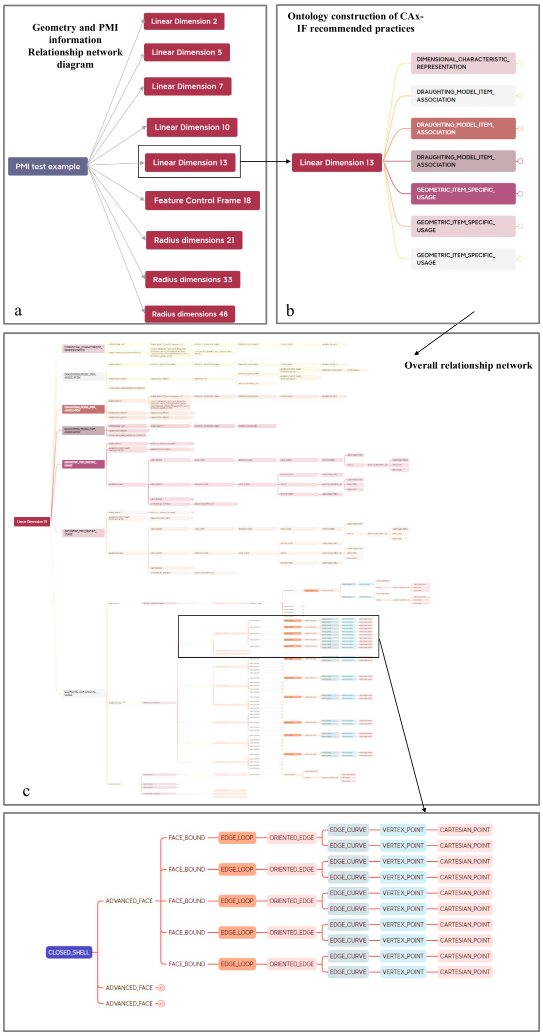


FIGURE 5. Class view based on the test model.

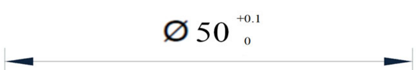


FIGURE 6. Example Nominal Value with Plus/Minus Bounds.

and $\bullet^I$ is a mapping function. In Table 1 the constructs, syntax and semantics associated with the concepts in SROIQ(D) are listed. This is then explained by the definition of the first expression in this subsection. Its concrete interpretation can be seen in Fig 5.

$$(shape\_aspect)^I \equiv (\exists ID.decimal \wedge \exists name.string \wedge \exists description.string \wedge \exists of\_shape.product\_definition\_shape \wedge \exists product\_definitional.boolean)^I$$

$$= (\exists ID.decimal)^I \wedge (\exists name.string)^I \wedge (\exists description.string)^I \wedge (\exists of\_shape.product\_definitional\_shape)^I \wedge (\exists product\_definitional.boolean)^I$$

$$= \left\{x|\exists y, \langle x, y\rangle \in ID^I \wedge y \in decimal^I\right\} \wedge \left\{x|\exists z, \langle x, z\rangle \in name^I \wedge z \in string^I\right\} \wedge \left\{x|\exists h, \langle x, h\rangle \in description^I \wedge h \in string^I\right\}$$

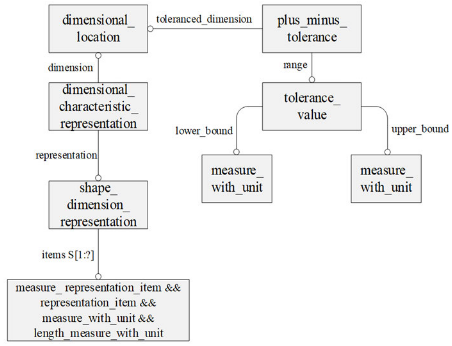


FIGURE 7. Instantiated block diagram of Nominal Value with Plus/Minus Bounds.

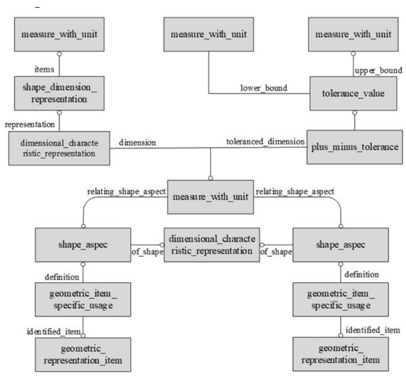


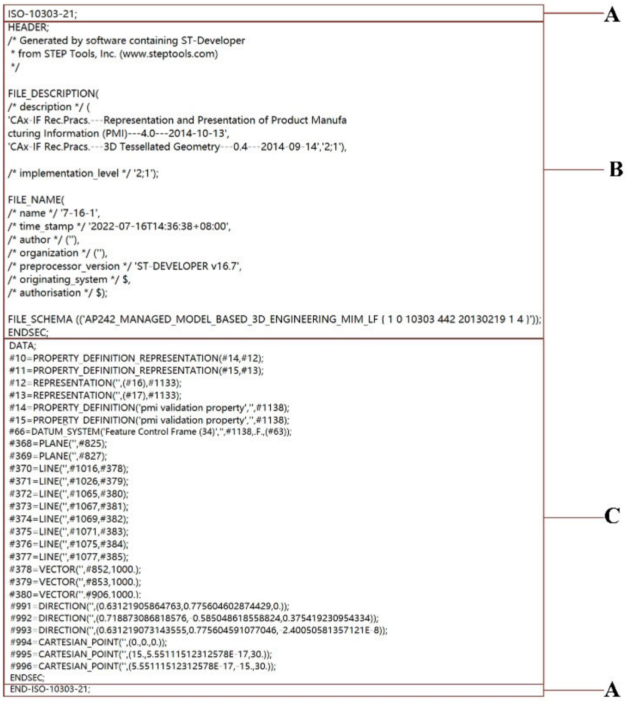
FIGURE 8. Instantiated block diagram of the geometric tolerance.

C. CONSTRUCTION OF SEMANTIC ASSOCIATION OF PMI MODEL

The CAD model in Fig 4, is examined in order to provide a more visual representation of the PMI information in the CAx-IF recommended practice for construction methods, and subsequently in order to test the validity and accuracy of the proposed method. The corresponding product information intent is collated according to the geometric product specification and relevant design standards. As an example, the model was collated using the product design information Linear size(13). In the PMI Test example, there are nine annotated entities, as shown in Fig 5(a). Among them, Linear size (13) consists of DIMENSIONAL CHARACTERISTIC_REPRESENTATION, DRAUGHTING_MODEL_ITEM_ASSOCIATION, GEOMETRIC_ITEM_SPECIFIC_USAGE shown in Fig 5(b), the build consists of 7 items. These product information will eventually be composed as CARTESIAN_POINT information with VERTEX_POINT as shown in Fig 5(c).

D. STRUCTURAL KNOWLEDGE OF PATTERN SPECIFICATION SEMANTIC REPRESENTATION

Taking the example of the large Radius dimension (21), shown in Fig 6, which is represented as the nominal value of a tolerance and carries a set of positive and negative deviations




```
ISO-10303-21;
HEADER;
/* Generated by software containing ST-Developer
 * from STEP Tools, Inc. (www.steptools.com)
 */

FILE_DESCRIPTION(
/* description */ (
'CAx-IF Rec.Pracs.---Representation and Presentation of Product Manufa
cturing Information (PMI)---4.0---2014-10-13',
'CAx-IF Rec.Pracs.---3D Tessellated Geometry---0.4---2014-09-14','2;1'),

/* implementation_level */ '2;1');

FILE_NAME(
/* name */ '7-16-1',
/* time_stamp */ '2022-07-16T14:36:38+08:00',
/* author */ (''),
/* organization */ (''),
/* preprocessor_version */ 'ST-DEVELOPER v16.7',
/* originating_system */ $,
/* authorisation */ $);

FILE_SCHEMA (('AP242_MANAGED_MODEL_BASED_3D_ENGINEERING_MIM_LF { 1 0 10303 442 20130219 1 4 }'));
ENDSEC;
DATA;
#10=PROPERTY_DEFINITION_REPRESENTATION(#14,#12);
#11=PROPERTY_DEFINITION_REPRESENTATION(#15,#13);
#12=REPRESENTATION('',(#16),#1133);
#13=REPRESENTATION('',(#17),#1133);
#14=PROPERTY_DEFINITION('pmi validation property','',#1138);
#15=PROPERTY_DEFINITION('pmi validation property','',#1138);
#66=DATUM_SYSTEM('Feature Control Frame (34)','',#1138,.F.,(#63));
#368=PLANE('',#825);
#369=PLANE('',#827);
#370=LINE('',#1016,#378);
#371=LINE('',#1026,#379);
#372=LINE('',#1065,#380);
#373=LINE('',#1067,#381);
#374=LINE('',#1069,#382);
#375=LINE('',#1071,#383);
#376=LINE('',#1075,#384);
#377=LINE('',#1077,#385);
#378=VECTOR('',#852,1000.);
#379=VECTOR('',#853,1000.);
#380=VECTOR('',#906,1000.);
#991=DIRECTION('',(0.63121905864763,0.775604602874429,0.));
#992=DIRECTION('',(0.718873086818576,-0.585048618558824,0.375419230954334));
#993=DIRECTION('',(0.631219073143555,0.775604591077046,-2.40050581357121E-8));
#994=CARTESIAN_POINT('',(0.,0.,0.));
#995=CARTESIAN_POINT('',(15.,5.55111512312578E-17,30.));
#996=CARTESIAN_POINT('',(5.55111512312578E-17,-15.,30.));
ENDSEC;
END-ISO-10303-21;
```

FIGURE 9. STEP AP242 in textual format.

or bounds from that tolerance, then the entity shown would be used to instantiate a STEP file. An example of this type of dimension is shown in Fig 7 below.

Where the values of positive and negative tolerances can be found in the measure_with_unit entities which are referred to by the "upper" and "lower" attributes of the tolerance_value entity. Even if the tolerance_value is given as an upper or lower limit, the value should still be considered as an offset from the nominal value. For example, the lower limit should normally be a negative number and the upper limit a positive number. The same can be said for two positive or two negative markers [27].

E. STEP FILE STRUCTURE

With a strong descriptive language such as EXPRESS [28], the definition and specification of the data model and the related constraint information are described in detail. Thus STEP neutral files have a strict standard specification and are designed to enable fast data exchange and sharing of products, without relying on a specific exchange system. The format structure is non-dualistic and unrelated, the text code contains geometric, topological and management information, and the entire file structure forms the basic structure of STEP [29], [30]. It is easy to process by computer software.

It is important to note that what is displayed in the STEP file depends on the export settings defined by the respective CAD system. As shown in Fig 9, each entity instance has a unique definition and this ID can be used as a reference to other entities as well as referencing other entities. In subsequent readings of the STEP file, only the relevant information

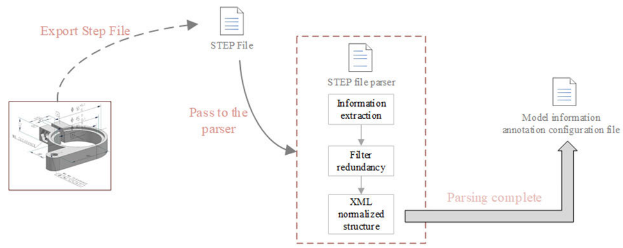


FIGURE 10. The overall architecture of the STEP parser.

can be extracted and stored. References to tolerances and datum features can also be extracted when using the relationships of the individual entities in the STEP file when reading GD&T information. The three parts of STEP (A to C) are summarised as follows:

Sec A: Neutral documents start with "ISO-10303-21;" at the head and end with "END-ISO-10303-21;" at the end. This indicates that the format conforms to the rules required by part 21 of ISO-10303 (STEP standard) and that the head and tail sections appear only once in the document.

Sec B: The HEARED section immediately follows to indicate the file start flag, followed by the file creation date, file name, computer-aided design software system version and licensing information, and the file mode indicates the name of the STEP application protocol that provides the file [31].

Sec C: Each line in the data segment section is represented as a specific example data information, which contains model geometry information, topology information, etc. The data structure is represented as follows: # Instance Identifier = Instance Type (Attribute 1, Attribute 2, ……). It is important to note that the instance identifier is randomly assigned by the system to uniquely identify the instance data; the instance type is used as a specific entity type defined in the STEP standard to describe the instance type to which the instance belongs. The content in brackets immediately following the instance type is used to correspond to the displayed attributes of the entity type, which may be one or more, with commas separating the attribute values; the attribute values may consist of specific numbers, strings, three-dimensional coordinates ((0, 0, 0)), references to identifiers of other attributes (#14), logical values (.F.), etc.

IV. RESULTS

A. STEP FILE PARSER OVERALL BUILD PROCESS

In sections II and III the MBD model and PMI information are defined and represented. In this section the following issues are addressed: how to build a STEP file parser based on STEP neutral files, to achieve neutral files to reduce file redundancy, in different applications and to speed up their process. Quickly build associations with entities to support further applications.

In this paper, C# is used as the language for the STEP parser and the parsing process is based on the representation structure presented in section III-C. The overall construction

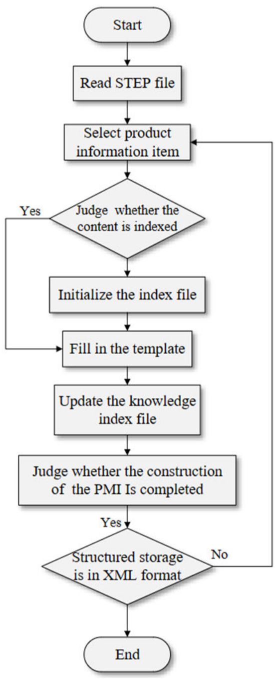


FIGURE 11. PMI information extraction optimization algorithm.

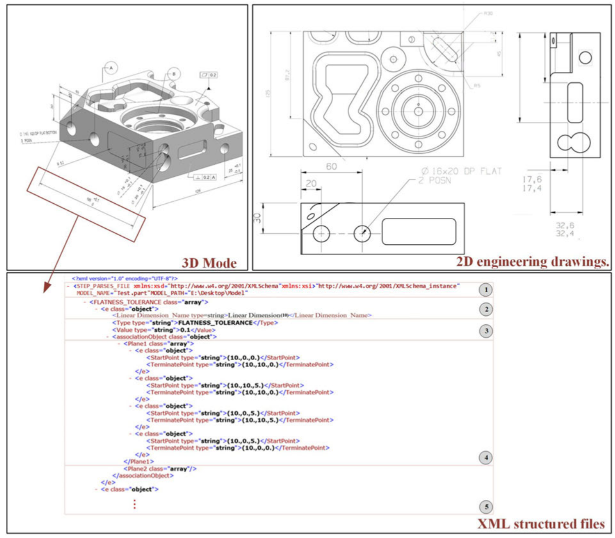


FIGURE 12. PMI index structure.

process for parsing a STEP file of PMI information is shown in Fig 10. First, a file describing the product information is generated in CAD software A by constructing a 3D model with PMI information; the AP242 file with the product geometry information model and the technical information model exported by the CAD software is used. In the STEP file parsing, the relevant technical and geometric information is extracted and its associated groups are constructed. The product information is then associated with the geometric information to form the STEP file parser. The extraction of product information and the association of geometric information are the key steps in the construction of the above parsing network. The network construction process is described below.

Before writing to the AP242 output file, the system performs a basic check on the model. It checks the following.

- Is the entity a feature or a sizing feature? For size features, relate the tolerance to the shape_aspect or derived_shape_aspect entity.
- Avoid redundant data, for example by reusing existing data rather than redefining it.
- Ensure that dimensions, sizes and shape tolerances have no datum associated with them.

B. DESIGN INFORMATION EXTRACTION PROCESS

At this stage, CAD software mostly uses the AP242 edition to store 3D annotation information in STEP neutral files, whose mosaic processing method fragments the geometry set and leads to slow information processing, and does not preserve the semantics of technical information. This paper proposes an information extraction algorithm based on STEP242 neutral files to extract the corresponding geometric and technical information. The algorithm reads the neutral files sequentially, and the annotated information forms a geometric information relationship network and a geometric information annotation network in turn. Finally, the neutral file reading network is constructed by forming a file normalisation structure through XML, and the flow is shown in Fig 11.

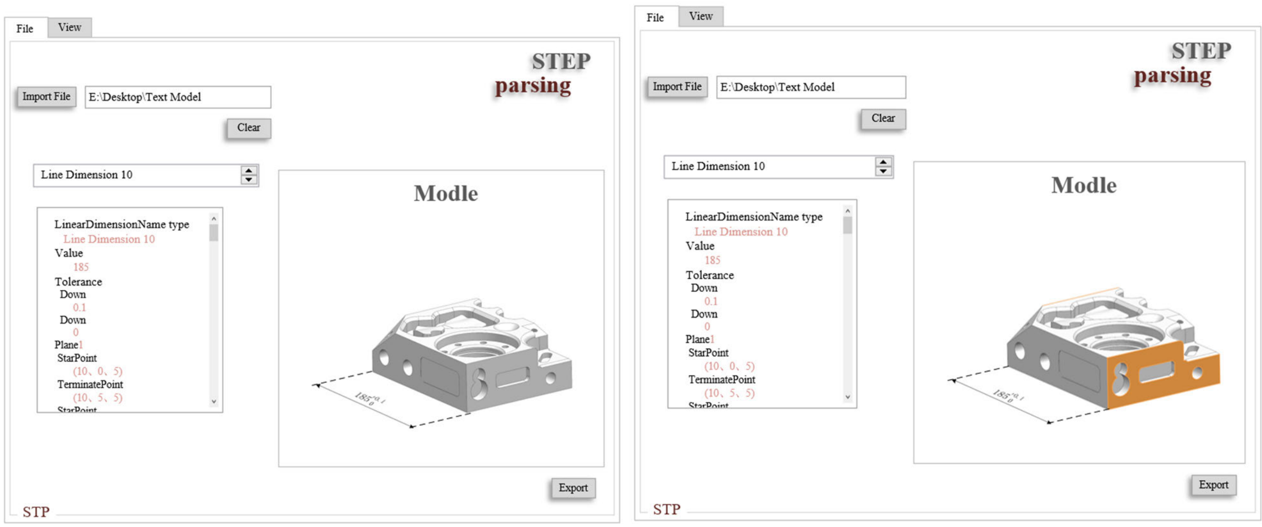
Due to the wide range of PMI information extraction, the general process of information extraction is the same, and the block diagram of the proposed information extraction optimisation algorithm takes Linear Dimension as an example to briefly describe the information extraction process.

Step 1: Extract the PMI annotation information Linear Dimension Read the STEP file and iterate through the files in the file, when the index information is the same, extract the key information to get the relevant information in the STEP file Linear Dimension (3) Linear Dimension (4), etc., establish the relationship between Linear Dimension() and extraction, merge the same categories, and the merged ones as child nodes. The merged ones are used as child nodes to form a network of information associations for the next step.

Step 2: Extract relationships between geometric and technical information. GEOMETRIC_ITEM_SPECIFIC_USAGE to associate SHAPE_ASPECT and GEOMETRIC_MODEL_ITEM (ADVA integrates such relationships into the geometric information relationship network through traversal to form a geometric and technical information network.

Step 3: The parsed PMI annotations and the coordinates of the geometric objects to which they are attached are structured and stored as .xml files, which can be called model information annotation profiles.

In the parsing of STEP files, the information extraction and structuring module is called from the toolbar. This module reads and processes the PMI part of the STEP file to



(a) Import the part model in STEP neutral format and analyze.

(b) Map the parts to which the PMI is attached.

FIGURE 13. GUI of the STEP parser.

form an Extensible Markup Language (XML) index file. This index file makes it possible both to establish the relationship between products and product information and to optimise the content of the information. Take ''Linear Dimension (10)'' as an example, the three-dimensional dimen-sions marked by ''PMI'', which can be automatically mapped to two-dimensional engineering drawings, are organised in XML as shown in the Fig 12. The structured information of the Dimension (10) is associated with the corresponding plane as follows:

1. the root node STEP_PARSES_FILE contains the basic attributes such as the name of the product model, the path to the file where the product is located, etc. Each STEP neutral file index file corresponds to the corresponding product model.
2. The sub-node FlATNESS_TOLERANCE_CLASS represents the class used to store the model PMI, while the Linear Dimension (10) represents the corresponding model information.
3. FlATNESS_TOLERANCE represents the attributes of FlATNESS_TOLERANCE_CLASS. Value type indicates the tolerance dimension of the Linear Dimension (10).
4. Plane1 class is an identifier in the product model, containing StartPoint and TerminalPoint information. It is used as a unique identifier to solve the problem of mapping information between PMI and the model fusion process in AP242 file extraction.
5. Another model identifier for Linear Dimension (10).

C. APPLICATIONS

The Fig 13 shows the GUI of the STEP parser. The first tab 'File' shown in the interface is divided into two import modules. After selecting the STEP file to be parsed, the boundary representation model will be visualised in the 'Model' box in the GUI and the STEP file will be structured by the parser to provide the initial PMI information in the CAD model, in order of PMI information. When the information to be queried is selected, the block diagram shows the specified name and tolerance type, the corresponding tolerance values, and the reference information and dependent geometry. When the analysis is complete, the analysed STEP file is exported by clicking on the tab 'Export'. Fig 13(a), for example, shows the Linear Dimension (10) of the model, where the PMI information has been formalised in a row for easier understanding.

The second tab 'View', shown in Fig 13(b), allows you to change the feature information you need to view by using the drop down, and the block diagram on the left side clearly shows the Cartesian coordinate point information for the associated plane. When the tab 'Show' is clicked, the Model box on the right shows the position of the two associated planes of PMI information.

V. DISCUSSION

In this paper, the proposed method for parsing PMI in MBD models is constructed to include the extraction of structured information and the presentation of semantic relationships in MBD models. Next, the effectiveness of the proposed method is illustrated in the following two aspects.

A. RELIABILITY VERIFICATION

The broad applicability of ontologies, especially in the manufacturing domain, makes the interaction of information between different systems a daunting task. The diversity of concepts and the extensive semantic mismatch make seamless interpretation still a challenge [32].

With STEP File Analyzer, the PMI is structured in the STEP file and can be used for a specific representation of

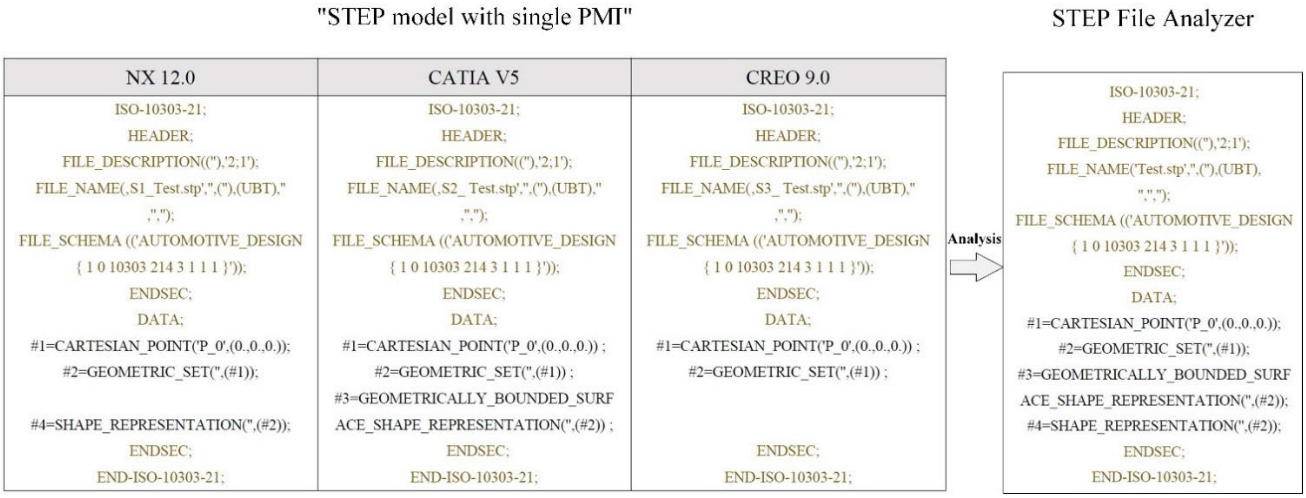


FIGURE 14. Single feature model from three different CAD systems converted to STEP File Analyzer.

TABLE 2. Compatibility testing between Linear Dimension and various systems under a model.

		Import in								
		NX 12.0			CATIA V5			CREO 9.0		
		I	D	N	I	D	N	I	D	N
Single feature model export	NX 12.0	+	+	+	+	-	-	-	-	-
	CATIA V5	+	+	+	+	+	+	+	+	+
	CREO 9.0	+	+	+	+	-	-	+	+	+
STEP File Analyzer		+	+	+	+	+	+	+	+	+

I: Import, D: Display, N: Naming, +: transferred sufficiently, -: transferred insufficiently

TABLE 3. Compatibility testing between Radial dimensions and various systems under a model.

		Import in								
		NX 12.0			CATIA V5			CREO 9.0		
		I	D	N	I	D	N	I	D	N
Single feature model export	NX 12.0	+	+	+	+	-	-	-	-	-
	CATIA V5	+	+	+	+	+	+	+	+	+
	CREO 9.0	+	+	+	+	-	-	+	+	+
STEP File Analyzer		+	+	+	+	+	+	+	+	+

I: Import, D: Display, N: Naming, +: transferred sufficiently, -: transferred insufficiently

the same product component in different systems. Here it is necessary to find out the basic annotation information for each CAD software, create the same product component model in each software, annotate the component with PMI and generate and export it programmatically as a STEP file. Next, the elements that are being displayed but are not required are removed and the remaining files are the entries necessary for the components represented by each software. As the components described by the three different CAD software have the same entries as necessary and only the same PMI information needs to be displayed. By this method, four comparative models can be created.

In this paper three different CAD systems (NX 12.0 (Siemens), CATIA V5 (Dassault Systèmes), CREO 9.0 (PTC)) are now selected. Single feature information Linear Dimension() and Radial dimensions() were created for them respectively, which are displayed in the same way, but with different additional entries displayed in the STEP file under different CAD software, as shown in Fig 14 through STEP File Analyzer.

Based on the above single piece of information, to verify STEP File Analyzer compatibility and also to compare the impact of information interaction between the CAD systems. The structured files were imported into the three

TABLE 4. Example of a step-neutral file containing PMI.

File name	Abbreviations	Size of the file
Fan-cooled Magnetic Particle Clutch	MPC	2.37M
Clutch lubricator	CL	9.22M
Through-hole chamfering machine	TCH	37.32M
Pulse heat presses	PHP	89.31M
Double clutch valve body testing equipment	DCTE	105.94M
Double clutch valve body airtightness testing equipment	DCATE	338.07M

TABLE 5. Example of a step-neutral file containing PMI.

File name	Size of the file	Step file parser Execution time (s)	Open CAS-CADE Execution time (s)
MPC	2.37M	3.243	2.878
CL	9.22M	8.987	8.754
TCH	37.32M	28.474	28.442
PHP	89.31M	67.221	68.447
DCTE	105.94M	89.452	93.751
DCATE	338.07M	288.337	314.854

systems for information interaction testing, as shown in Table 2-3.

The single feature information models generated by each system and the information processed by STEP File Analyzer were imported into the three systems. In both examples, the quality of the information interaction was significantly improved after processing by STEP File Analyzer. Only 75% (Linear Dimension) and 70% (Radial dimensions) of the test features could be transferred in the single feature interaction tests, with CATIA V5 performing best in Table 2 and CREO 9.0 in Table 3. In order to be able to transfer different feature information models at the same time, models normalised by STEP File Analyzer are the best choice.

B. PARSING SPEED COMPARISONS

For CAD systems, the data analysis module must be used to exchange product data with other systems. The STEP File Analyzer in the previous section has been used successfully to read STEP neutral files. However, when the STEP neutral file is too large, the file loading speed is slow and the waiting time for the file to be read is too long.

In order to test the superiority of PMI information extraction, a large number of document samples were analysed and extracted in this study, and six representative document sample results were selected for experimental analysis. The experimental tests were conducted on a PC configured with Windows 10 operating system and a PC configured with a quad-core Intel i7 processor (2.30GHz) by selecting file samples of different sizes [33]. It is important to note that the experimental results kind of vary depending on the host configuration, but the experimental conclusions are consistent.

The selected 3D solid models are presented in Table 4 and the selected file names are abbreviated for ease of subsequent description. The models were exported as AP242 files using a unified CAD software to facilitate the subsequent experimental comparison. In this paper, Open CASCADE was selected as the experimental control group. Among the existing CAD geometry engines, the Open CASCADE kernel is open source [34] and the STEP file reading function can be used separately. The proposed file reader is therefore used to compare the read results of the six representative file samples described above with the Open CASCADE read function.

Comparison of the time taken by the two methods to read the file is shown in Table 5. When reading smaller files, the difference in time between the two methods to read the file is not significant, with the method proposed in this paper taking longer; as the file grows larger, the reading method in this paper performs better. In contrast to the algorithm in this paper, Open CASCADE implements the read task with C++ class objects defined for all entity objects of the STEP standard. During the read, a series of processes are performed to transform the data into C++ class objects, which are then converted into data structure objects for the geometry engine. The method in this paper eliminates the intermediate steps and converts the data in the STEP file directly into data structure objects for the geometry engine. This shows that the PMI build in the STEP parser not only accurately extracts relevant product information, but also effectively filters redundant information from the STEP file.

VI. CONCLUSION

PMI is used as an important component of digital smart manufacturing to deal with the exchange of information that exists between products at different stages and between different computer-aided design software. This paper shows the semantics and syntax of PMI as defined by the ASME standard and analyses the implementation specification for PMI under the CAx-IF recommendations. With the help of this specification, the STEP neutral document analyser used in this paper establishes links to identify the most important components of geometric dimensions and tolerances in product manufacturing information, and then represents the identified features with the corresponding information. Storing the identified feature information in a structured form enhances the compatibility of the information and avoids redundancy in the documentation. Identification and extraction of feature PMIs in AP242 format to improve consistency of information in collaborative manufacturing processes. The feasibility of the approach proposed in this paper is also verified through practical examples, saving computer arithmetic power and proposing a new solution for future data exchange of product manufacturing information.

The next step is intended to extend the implementation for product manufacturing information in neutral documents, which requires extending the types of recognition of neutral documents as well as product information by the STEP parser. A number of case studies are in progress to test this model.

REFERENCES

[1] K. Efthymiou, A. Pagoropoulos, N. Papakostas, D. Mourtzis, and G. Chryssolouris, "Manufacturing systems complexity: An assessment of manufacturing performance indicators unpredictability," *CIRP J. Manuf. Sci. Technol.*, vol. 7, no. 4, pp. 324–334, 2014.

[2] S. Ramnath, P. Haghighi, A. Venkiteswaran, and J. J. Shah, "Interoperability of CAD geometry and product manufacturing information for computer integrated manufacturing," *Int. J. Comput. Integr. Manuf.*, vol. 33, no. 2, pp. 116–132, Feb. 2020.

[3] R. S. Srinivasan, K. L. Wood, and D. A. McAdams, "Functional tolerancing: A design for manufacturing methodology," *Res. Eng. Design*, vol. 8, no. 2, pp. 99–115, Jun. 1996.

[4] J. Geng, X. Tian, M. Bai, X. Jia, and X. Liu, "A design method for three-dimensional maintenance, repair and overhaul job card of complex products," *Comput. Ind.*, vol. 65, no. 1, pp. 200–209, Jan. 2014.

[5] S. Abdulrahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi, and M. Guizani, "A survey on federated learning: The journey from centralized to distributed on-site learning and beyond," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5476–5497, Apr. 2021.

[6] J. Camba, M. Contero, M. Johnson, and P. Company, "Extended 3D annotations as a new mechanism to explicitly communicate geometric design intent and increase CAD model reusability," *Comput.-Aided Design*, vol. 57, pp. 61–73, Dec. 2014.

[7] M. Lundin, S. Sandberg, and M. Näsström, "Knowledge retention and reuse: Using CAD models as carriers of knowledge in product development," in *Proc. 30th Comput. Inf. Eng. Conf.*, Jan. 2010, pp. 1173–1182.

[8] H. Liu, A. Zhang, and L. Wang, "An intelligent approach for dimensioning completeness inspection in 3D based on transient geometric elements," *Robot. Comput.-Integr. Manuf.*, vol. 55, pp. 191–203, 2019.

[9] *Document Management—Engineering Document Format Using PDF—Part 1: Use of PDF 1.6 (PDF/E-1)*, Standard ISO 24517-1:2008, ISO/TC 171/SC 2, 2008.

[10] *Industrial Automation Systems and Integration-Product Data Representation and Exchange-Part 242: Application Protocol: Managed Model-Based 3D Engineering*, Standard ISO 10303-242:2014, Geneva, Switzerland, 2014.

[11] R. Lipman and J. Lubell, "Conformance checking of PMI representation in CAD model STEP data exchange files," *Comput.-Aided Design*, vol. 66, pp. 14–23, Sep. 2015.

[12] D. Cheney and B. Fischer, "Measuring the PMI modeling capability in CAD systems: Report 3-fully-toleranced test case verification," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NISTGCR:15-999, 2015.

[13] W. Lu, Y. Qin, X. Liu, M. Huang, L. Zhou, and X. Jiang, "Enriching the semantics of variational geometric constraint data with ontology," *Comput.-Aided Design*, vol. 63, pp. 72–85, Jun. 2015.

[14] T. Hedberg, J. Lubell, L. Fischer, L. Maggiano, and A. B. Feeney, "Testing the digital thread in support of model-based manufacturing and inspection," *J. Comput. Inf. Sci. Eng.*, vol. 16, no. 2, Jun. 2016, Art. no. 021001.

[15] V. Quintana, L. Rivest, R. Pellerin, F. Venne, and F. Kheddouci, "Will model-based definition replace engineering drawings throughout the product lifecycle? A global perspective from aerospace industry," *Comput. Ind.*, vol. 61, no. 5, pp. 497–508, Jun. 2010.

[16] *ASME Y14.41 Digital Product Definition Data Practices*, Amer. Soc. Mech. Eng., New York, NY, USA, 2003.

[17] *Technical Product Documentation-Digital Product Definition Data Practices*, International Organization for Standardization: Geneva, Switzerland, Standard ISO 16792, 2006.

[18] R. Huang, S. Zhang, X. Bai, and C. Xu, "Multi-level structuralized model-based definition model based on machining features for manufacturing reuse of mechanical parts," *Int. J. Adv. Manuf. Technol.*, vol. 75, nos. 5–8, pp. 1035–1048, Nov. 2014.

[19] S. Ramnath, P. Haghighi, A. Chitale, J. K. Davidson, and J. J. Shah, "Comparative study of tolerance analysis methods applied to a complex assembly," in *Proc. 15th CIRP Conf. Comput. Aided Tolerancing CIRP CAT*, Milan, Italy, 2018, pp. 1–6.

[20] *ASME Y14.5 Dimensioning and Tolerancing*, Amer. Soc. Mech. Eng., New York, NY, USA, 2003.

[21] M. Hallmann, S. Goetz, and B. Schleich, "Mapping of GD&T information and PMI between 3D product models in the STEP and STL format," *Comput.-Aided Design*, vol. 115, pp. 293–306, Oct. 2019.

[22] W. Xiang, Z. Wansheng, and W. Li, "Three-dimensional geometric representation," *Aerosp. Manuf. Technol.*, no. 4, pp. 40–43, 2002.

[23] J. Yang, S. Han, and S. Park, "A method for verification of computer-aided design model errors," *J. Eng. Design*, vol. 16, no. 3, pp. 337–352, Jun. 2005.

[24] F. Ahmed and S. Han, "Interoperability of product and manufacturing information using ontology," *Concurrent Eng.*, vol. 23, no. 3, pp. 265–278, Sep. 2015.

[25] *Test Suite for the CAx Implementor Forum Round 35J*. Accessed: Oct. 2022. [Online]. Available: https://www.caxif.org/documents/test_suite_35j_v1.pdf

[26] J. J. Shah, Y. Yan, and B.-C. Zhang, "Dimension and tolerance modeling and transformations in feature based design and manufacturing," *J. Intell. Manuf.*, vol. 9, no. 5, pp. 475–488, 1998.

[27] *Industrial Automation Systems and Integration-Product Data Representation and Exchange-Part 11: Description Methods: The Express Language Reference Manual*, International Organization for Standardization, Standard ISO 10303-11: 1994, ISO T (1994) 184/SC 4, 1994.

[28] *Industry Automation Systems and Integration Product Data Representation and Exchange-Part 11, Description Methods: The Express language Reference Manual*, Standard ISO 10303-11, 1994.

[29] *Industrial Automation Systems and Integration-Product Data Representation and Exchange*, Standard ISO IS 0.10303, 1994.

[30] *Application Protocol for the Exchange of Configuration Controlled 3D Product Design Data*, Standard ISO 10303-203, 1991.

[31] D. Baisheng, "Research on STEP-based manufacturing feature recognition and 3D reconstruction technology," M.S. thesis, Northeast Agricult. Univ., Harbin, China, 2010.

[32] F. Ahmed and S.-H. Han, "Semantic mismatches for interoperability of product and manufacturing information," *Int. J. CAD/CAM*, vol. 13, no. 2, pp. 97–108, 2013.

[33] *Traceparts Parts Library*. Accessed: Mar. 22, 2023. [Online]. Available: https://www.traceparts.cn/zh

[34] *Open CASCADE*. Accessed: Oct. 2022. [Online]. Available: https://dev.opencascade.org/release

HU QIAO received the Ph.D. degree from Northwestern Polytechnical University, Shaanxi, China, in 2015. He is currently a Lecturer with the School of Mechatronic Engineering, Xi'an Technological University, Shaanxi. His research interests include digital design and manufacture, 3Dmodel retrieval and reuse, and assembly automation.

SI-BO HU was born in Shaanxi, China, in 1998. He received the bachelor's degree in mechanical engineering from Henan Agricultural University, in 2021. He is currently pursuing the master's degree in mechanical engineering with the School of Mechatronic Engineering, Xi'an Technological University, Shaanxi. His research interests include digital design and fabrication and in-machine measurements.

LI ZHANG was born in Shaanxi, China, in 1996. He is currently pursuing the master's degree in mechatronic engineering with Xi'an Technological University. His research interests include digital design and manufacturing.

JI-HANG LI was born in Shaanxi, China, in 1998. He received the bachelor's degree in mechanical engineering from Hebei University, in 2021. He is currently pursuing the master's degree in mechanical engineering with the School of Mechatronic Engineering, Xi'an Technological University. His research interests include the application of on-machine measurement technology in the field of mechanical manufacturing.

LIANG YOU received the bachelor's degree from the North China Institute of Science and Technology, in 2018. He is currently pursuing the master's degree with the School of Mechatronic Engineering, Xi'an Technological University, Shaanxi, China. His research interests include 3D model retrieval and reuse, discharges, discharges in liquids, spectroscopic diagnostics, plasma propulsion, and innovation plasma.