Sensors and Materials, Vol. 37, No. 6 (2025) 2569–2587
MYU Tokyo

S & M 4075

Examining the Development of a Geographic Information System Georeferencing Algorithm for Building Informaiton Modeling

Young Sun Kim,[1] Seon Cheol Yu,[2] Yoon Seok Jwa,[3] and Jong Wook Ahn[2*]

[1]Smart Urban Space Institute, Anyang University,
22, Smadeok-ro 37beon-gil, Manan-gu, Anayang-si, Gyeonggi-do 14028, Korea
[2]Department of Smart City Engineering, Anyang University,
22, Smadeok-ro 37beon-gil, Manan-gu, Anayang-si, Gyeonggi-do 14028, Korea
[3]Platform Development Division, Soosung Engineering,
13, Jeongui-ro 8-gil, Songpa-gu, Seoul 05836, Korea

(Received September 26, 2024; accepted April 14, 2025)

Keywords: BIM, GIS, georeferencing, BIM to GIS, outermost polygon of building

Digital twins are rapidly expanding in the market as an important foundation for industrial revitalization. Among these, building information modeling (BIM) plays a critical role in the digital twin field, and its integration with the Geographic Information System (GIS) enables applications in various areas, including urban planning and digital twin development. However, BIM and GIS were developed for different purposes, resulting in distinct data structures and characteristics, which makes integration challenging. In particular, the integration process requires positional correction owing to differences in coordinate systems, but in practice, location information is not prioritized during BIM model creation, making coordinate matching difficult. To address the discrepancies in location information between BIM and GIS, in this study, we propose an automatic algorithm for coordinate transformation to align Local Coordinate System (LCS)-based data with Reference Coordinate System (RCS)-based GIS digital maps. The algorithm goes through a five-step process: measuring location information levels, extracting outermost polygon vectors, polygon-based model matching, extracting coordinate transformation coefficients, and BIM model georeferencing. Owing to the limited availability of BIM data, the proposed method was only applied experimentally in this study. However, future research should involve extensive experiments and discussions using various BIM models.

1. Introduction

1.1 Research background and objectives

The Fourth Industrial Revolution is driving the convergence of technologies and data across industries through “hyper-connectivity”, “super-intelligence”, and “convergence”. In particular, this revolution is having a significant impact on the Geographic Information System (GIS) field

*Corresponding author: e-mail: ajw0603@anyang.ac.kr
https://doi.org/10.18494/SAM5381

by integrating various technologies and connecting heterogeneous data. One notable area of growth is the digital twin, which integrates data from urban planning, architecture, and communication fields based on GIS. The digital twin market is rapidly expanding as a foundation for industrial revitalization, facilitating the sharing of 3D data and services.

Within the digital twin domain, building information modeling (BIM) is considered one of the most important data sources. BIM is fundamentally a model that manages building information in a 3D virtual environment, and when integrated with GIS, it enables applications in various fields, such as efficient facility management and urban planning. According to Song *et al.*,[1] while BIM offers advantages in managing building life cycles and detailed information, GIS excels in visualizing large areas of land and modeling specific locations to aid decision-making. Therefore, the integration of BIM and GIS has emerged as a significant and rapidly growing trend in both practical and academic fields over the past decade.

However, because GIS and BIM were developed for different purposes, they have distinct data structures and characteristics, making integration challenging. BIM is designed to provide detailed and accurate representations of individual structures, making it suitable for design and specific visualizations. However, it does not account for terrain or spatial constraints, which limits its ability to analyze and process data at an urban scale. In contrast, GIS is focused on spatial analysis utilizing data, which makes it less ideal for specific data representation compared with BIM. However, GIS contains terrain and attribute data, making it suitable for statistical analysis and data processing related to other objects.

Furthermore, the integration process requires position correction owing to the use of different coordinate systems. GIS data typically uses a Reference Coordinate System (RCS), while BIM data relies on an arbitrary Local Coordinate System (LCS). Positional accuracy between the two systems is a critical factor in determining the applicability of specific tasks in 3D spatial information services.[2] To utilize GIS and BIM data on the same platform, the coordinates must be aligned, which involves converting BIM's local coordinates into a geographic coordinate system centered on Earth. When integrating with GIS, coordinate mismatches often require significant manual work, and the accuracy of the position alignment can depend heavily on the worker's expertise. Moreover, in practice, location information is not always prioritized when creating BIM models, making coordinate matching a complex task.

Therefore, the purpose of this study is to propose an automatic coordinate transformation algorithm that positions objects corresponding to GIS maps (digital maps) based on RCS from LCS. This approach aims to resolve the discrepancies in location information that arise owing to differences between BIM, which focuses on the detailed and accurate representation of individual objects, and GIS, which emphasizes spatial analysis. This will be achieved by reviewing previous studies on the assignment of location information in BIM and by understanding the level of location information contained in the Industry Foundation Class (IFC) format, the standard data format for BIM.

1.2 Research scope and methods

The scope of this study is limited to proposing a method for automatically matching two models, extracting matching points, calculating coordinate transformation coefficients, and assigning the calculated absolute coordinate information to location-related entities in the IFC model, the standard data model for BIM. In this study, we assume the availability of a polygon model extracted from a digital map corresponding to the BIM model. The spatial scope focuses on the headquarters building of LX Korea Land and Geospatial Informatix Corporation (located in Jeonju). To parse the IFC model schema, the IfcOpenShell open-source library was used to extract and utilize LCS-based location-related entities and attribute information. The overall research method can be summarized in Fig. 1.

First, through a theoretical examination, we explored the meaning of the term “georeferencing” as used in the spatial information field and reviewed the level of location information contained in IFC files, the standard format for BIM data. This step was essential for evaluating the location information level of IFC to propose an algorithm for the automatic assignment of location information to BIM models.

Second, we reviewed previous studies addressing the limitations of georeferencing in IFC and the process of assigning location information through the georeferencing of BIM data. From this review, we derived implications and highlighted the unique aspects of this study.

Third, we proposed a shape-based matching algorithm for BIM and GIS models and conducted experiments to derive georeferencing results.

Finally, on the basis of the outcomes of the proposed technique, we discussed the limitations of the study and provided directions for future research.

2. Theoretical and Literature Review

In the theoretical review, we examined the meaning of georeferencing as used in the spatial information field and the level of location information contained in IFC files, the standard format for BIM data. Georeferencing can be interpreted differently across various fields. Without a clear definition, confusion may arise during the coordinate transformation process owing to the term’s multiple interpretations. Therefore, to apply the correct technical approach, it is essential to have a precise understanding of the meaning of georeferencing. Regarding the

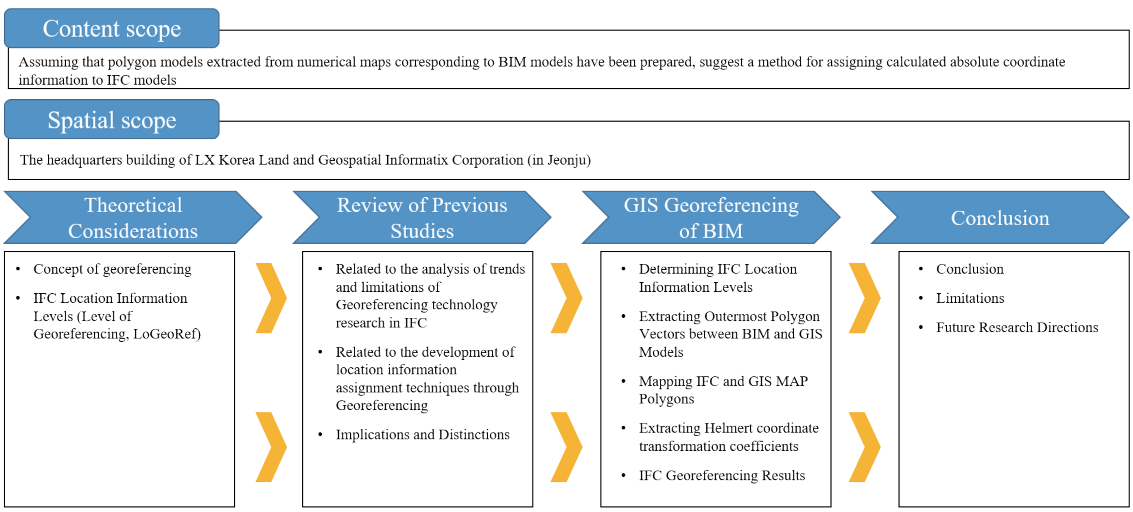


Fig. 1. (Color online) Research scope and methodology.

level of location information contained in IFC files, it is important to note that BIM and GIS use different coordinate systems. An accurate conversion requires a detailed examination of the level of location information included in BIM's IFC files. Additionally, to address the issue of mismatched location information, an understanding of the accuracy and details of location information used in BIM is necessary.

The review of previous research was divided into two main parts: (1) an analysis of trends and limitations in georeferencing technology research in the IFC format, the standard format for BIM data, and (2) a review of previous studies on the development of technology for assigning location information through georeferencing. On the basis of this review, we presented key implications and highlighted the differentiating factors of this study.

2.1 Concept of georeferencing

As georeferencing has various interpretations in different fields, we examined the concept of the term as used in the spatial information field. We explored the concept of georeferencing, broadly dividing it into spatial-information-related institutions, existing literature, and spatial-information-related software.

First, in the OGC GeoTIFF standard,[3] georeferencing is defined as positioning an object using a correspondence model derived from points where both ground and image coordinates are known. Here, the correspondence model refers to the mathematical relationship between ground and image coordinates.

The United States Geological Survey (USGS)[4] defines georeferencing as linking the internal coordinate system of a digital map or an aerial photograph to a ground system of geographic coordinates. Georeferenced digital maps or images are connected to a known Earth coordinate system, allowing users to determine the location of all points on maps and aerial photographs relative to Earth.

Next are the concepts of georeferencing as defined in existing literature. Hackeloeer *et al.*[5] defined georeferencing as a comprehensive term for technologies related to the unique identification of geographical objects, stating that geographical objects include items such as points of interest (POIs), buildings, and infrastructure, which are associated with specific geographical locations. Yao[6] defined georeferencing as the process of assigning locations to geographical objects within a geographic reference frame. He categorized georeferencing methods into two types: those that are directly linked to a GIS database containing referenced spatial features and those that generate georeferenced locations based on attribute data, such as the name and index of the target object. Uggla and Horemuz[7] defined georeferencing as the process of assigning geodetic coordinates related to real-world locations. They argued that in the infrastructure design phase, georeferencing is performed under the assumption that the plane of the Cartesian coordinate system aligns with the map projection plane, making the vertical directions of the LCS in BIM and the RCS in GIS parallel. On the basis of this geometric concept, they calculated the coordinate transformation parameters required for georeferencing. Noardo *et al.*[8] defined georeferencing as the method used to determine the location of a point on Earth's surface. Similarly, Jaud *et al.*[9] interpreted it as the process of positioning assets on

Earth. Lastly, Im(10) defined georeferencing as the process of projecting data defined in an image or object coordinate system onto a real-world coordinate system, thereby assigning actual coordinates.

Next are the concepts of georeferencing as presented in spatial information software. In Esri's ArcGIS GIS,(11) georeferencing is described as a method for adjusting the location of raster files, such as orthoimages (e.g., aerial and drone images), whose coordinates are not defined. Similarly, QGIS,(12) an open-source desktop GIS application, defines georeferencing as the process of assigning real-world coordinates to each pixel of raster data.

Table 1 shows the concepts of the term georeferencing used in the spatial information field. On the basis of the following content, we define IFC georeferencing as "the process of assigning RCS-based location information to BIM objects in LCS."

2.2 IFC level of georeferencing (LoGeoRef)

Christian and Hendrik(13) proposed a method for classifying the level of location information in IFC into five distinct levels through the concept of LoGeoRef (level of georeferencing). This classification enhances the understanding of coordinate information when georeferencing between the BIM and GIS domains. The classification ranges from LoGeoRef10, the simplest level that stores postal addresses, to LoGeoRef50, which can store attribute information for coordinate transformation using IfcMapConversion, a newly introduced entity in the IFC4

Table 1
georeferencing concepts.

	Figure	
Related organizations	GeoTIFF standard (OGC)(3)	Positioning objects using a correspondence model derived from point clouds where both ground and image coordinates are known
	United States Geological Survey (USGS)(4)	Linking the internal coordinate system of a digital map or an aerial photograph to a ground system of geographic coordinates
Existing literature	Hackeloeer *et al.*(5)	Linking spatial data such as maps and images to a geographic or projected coordinate system (RCS) through coordinate transformation
	Yao(6)	Assigning locations to geographical objects within a geographic reference frame
	Uggla and Horemuz(7)	Assigning geodetic coordinates related to the real world
	Noardo *et al.*(8)	Method used to define the position of a point on Earth's surface
	Jaud *et al.*(9)	Positioning assets on Earth
	Im(10)	Task of projecting data defined in image or object coordinate systems onto real-world coordinate systems to assign actual coordinates
Software	ArcGIS(11)	Method of adjusting the positions of raster files (such as aerial images, drone images, and other orthoimages) that do not have defined coordinates
	QGIS(12)	Assigning real-world coordinates to each pixel of raster data

version. This system classifies the level of location information based on the attribute information held by location-related entities within IFC. Although higher LoGeoRef levels correspond to higher-quality georeferencing, these levels do not necessarily include the conditions of lower levels; rather, each level has independent characteristics. Table 2 shows the conditions for each LoGeoRef level.

2.3 Review of previous studies

Table 3 shows the review of previous studies, categorized into two main areas: the analysis of trends and limitations in georeferencing technology research in IFC, the standard format for BIM data, and research on the development of technology for assigning location information through georeferencing.

First, in the review, the authors examine previous studies related to the analysis of trends and limitations in georeferencing technology research in IFC. Typically, the georeferencing of BIM is not prioritized during the object design phase, which is why this topic is primarily addressed within the GIS field. This indicates that proper georeferencing work within GIS systems is necessary to link IFC models with surrounding terrain and environmental elements. According to Noardo *et al.*,[14] there has been limited research on the capability of commercial software to optimize georeferencing using location-related attribute information from IFC standards when georeferencing BIM models. Additionally, Irizarry *et al.*[15] noted that in IFC2×3, an early version of IFC, it was not possible to store multiple georeferenced object models on a server or edit their attributes within a single IFC file. To address this limitation, they used CityGML, a data exchange format, to complement the shortcomings of IFC2×3.

On the basis of Christian and Hendrik's[13] LoGeoRef criteria, Noardo *et al.*[16] reviewed 57 IFC models and found that most contained incorrectly stored georeferencing information. Specifically, errors in the storage of planar coordinates were identified in 42% of the models, likely due to the application of default location options in BIM software used to generate the IFC models. Additionally, 12% of the models included very approximate georeferencing information, often based on randomly generated location data from within the country where the object was located. It remains unclear whether these inaccuracies originated from the BIM software itself or from the inclusion of default meta-information. Only 16% of the models provided sufficient

Table 2
Characteristics of LoGeoRef level.

LoGeoRef Level	Description
LoGeoRef 10	There must be attribute information values for IfcPostalAddress, and this entity must be referenced by either IfcSite or IfcBuilding.
LoGeoRef 20	IfcSite must have latitude, longitude, and elevation information values.
LoGeoRef 30	The IfcCartesianPoint value referenced by IfcAxis 2P placement 3D must be greater than (0, 0, 0). → This is the stage where projected coordinate system values are entered.
LoGeoRef 40	The coordinate value of IfcCartesianPoint referenced by IfcGeometricRepresentationContext must be greater than (0, 0, 0).
LoGeoRef 50	This level can only be verified in the current IFC4 version. The type of CRS information among the attribute information of IfcMapConversion must be IfcGeometricRepresentationContext.

Table 3
Review of previous research related to georeferencing.

Previous Research		Content
Related to the Analysis of Trends and Limitations in georeferencing Technology Research	Noardo *et al.*[14]	It is mentioned that there is insufficient research on the capability to optimize georeferencing work using the location-related attribute information of IFC in commercial software for georeferencing BIM models. It is argued that the methods each software uses to store georeferencing information do not always follow LoGeoRef. Even within the same software, when georeferencing different files, information may be stored in different entities according to different criteria.
	Irizarry *et al.*[15]	It is argued that in IFC2×3, it was impossible to store multiple georeferenced object models on a server and edit their attributes in a single IFC file, so CityGML, a data exchange format, was used to complement this.
	Noardo *et al.*[16]	A total of 57 IFC models were reviewed, and it is claimed that most of them had incorrectly stored georeferencing information. None of the 57 models included the attribute information of georeferencing-related entities newly added in the IFC4 version.
Related to Assigning Location Information through georeferencing	Kim *et al.*[17]	Research was conducted on generating a 3D visualization model through CAD-GIS polygon-based georeferencing. After editing the numerical map to fit the CAD drawing, georeferencing was performed by extracting the reference points of the floor polygon of the shape building.
	Kim and Hong[2]	An absolute coordinate transformation algorithm was developed to generate a 3D texturing model based on absolute coordinates through IFC coordinate information-based georeferencing and to integrate it into BIM/GIS platforms. However, this algorithm alone does not place the model in the exact location, so detailed adjustment is needed by mapping the initially transformed IFC data with texturing data.
	Diakite and Zlatanova[18]	It deals with methods to automate IFC-GIS polygon-based georeferencing. A method was developed to extract the outer polygons of the IFC and 3D models in GIS, find the most similar matching points, and extract the transformation matrix. If the shape of the building is perfectly symmetrical, it is difficult to distinguish the correct direction, so an additional information input is required for BIM data and map.

georeferencing information that could be properly integrated with GIS, and none of the 57 models included georeferencing-related entity attributes introduced into the IFC4 version. Furthermore, Noardo *et al.*[14] categorized the results of georeferencing IFC models using four commercial software programs, according to the five LoGeoRef criteria established by Christian and Hendrik.[13] The methods by which each software stored georeferencing information did not always align with LoGeoRef standards, and in some instances, even within the same software, georeferencing information for different files was stored in different entities based on varying criteria. This inconsistency highlights the diversity of georeferencing options within the IFC standard and underscores that clear and accurate georeferencing information is not consistently stored.[14] In conclusion, this emphasizes the importance of exercising caution when performing georeferencing tasks, as there is currently no established method, rule, or standard for storing georeferencing information within the IFC standard.

Next, several studies have focused on assigning location information through the georeferencing of BIM data. Kim *et al.*[17] conducted a study on generating a 3D visualization model after georeferencing in GIS using CAD data. This process involved extracting wall and floor lines by selecting the relevant layers from CAD drawings and creating polygons on the basis of these extracted lines. georeferencing was performed by referencing the floor polygon of the building shape, which was generated using data from GPS surveys on numerical maps and building roofs. The georeferencing process was executed using Affine and Similarity coordinate transformations. When the numerical map and CAD drawing shapes did not align, georeferencing became challenging. To overcome this limitation, the numerical map was edited to match the CAD drawing before georeferencing. Kim and Hong[2] studied a method for the position correction of 3D models for BIM and GIS data fusion using the IFC coordinate-information-based georeferencing technology. This technology can be described as a georeferencing method that updates coordinate information on the basis of the level of detail (LOD), reference points, and volume information of the IFC model to fit the BIM/GIS data schema structure. It calculates the absolute coordinates of a building using database index information, LOD, and reference position values within IFC. However, manual corrections are still required to achieve a precise placement, as the building is not automatically positioned in the exact location. Diakite and Zlatanova[18] studied an IFC-GIS polygon-based georeferencing method that can be applied even when the two datasets do not have the same number of points or when corresponding points are not known in advance. This technology extracts the outer polygons of the IFC and 3D models in GIS, identifies the most similar matching points through polygon comparison, and extracts the transformation matrix. georeferencing is then performed by adjusting the scale and rotation of the two datasets, whose approximate positions are aligned using the transformation matrix. Notably, while most georeferencing-related studies involve primarily manual processes, this study is significant for its focus on an automated georeferencing method. Typically, to match the positions of two datasets located in different places, they must have the same number of points, or corresponding points between them must be identified in advance, which complicates the automation of georeferencing. However, Diakite and Zlatanova[18] assumed that the shapes of the map and BIM polygons are similar. The greater the dissimilarity between these shapes, the larger the error that occurs. Additionally, if the building's shape is perfectly symmetrical, it may be difficult to determine the correct orientation, which could lead to the misplacement of elements such as the main entrance. This limitation requires additional information input for both the BIM data and the map.

2.4 Implications and differentiation

As reviewed above, previous studies on assigning location information for linking BIM data in GIS systems have been conducted from various perspectives. However, it has been found that, to date, no standardized method or rule exists for storing georeferencing information in IFC, leading to inconsistent storage practices. In some cases, incorrect information is stored in IFC models because designers omit or fail to enter the necessary data during the modeling process. Additionally, current georeferencing-related research is often tailored to specific projects or lacks detailed descriptions of the methods used. Furthermore, it has been observed that the georeferencing methods currently employed to position BIM models in GIS are largely manual or semi-automated, with no fully automated approaches in widespread use. This is because, to align the positions of two datasets from different locations, they often require an equal number of points or that corresponding points must be known in advance. The need to verify these constraints has made the automation of georeferencing challenging. Diakite and Zlatanova(18) proposed an automated georeferencing method that can be applied even when the two datasets do not contain the same number of points or when corresponding points are not known in advance. However, their method assumes that the shapes of the map and BIM polygons are similar. The more dissimilar the shapes, the greater the resulting error. Additionally, if the building's shape is perfectly symmetrical, it may be difficult to determine the correct orientation, potentially leading to misalignment, such as placing the main entrance incorrectly. In such cases, additional information is required for both the BIM data and the map.

In this study, we address these limitations by proposing a shape-based automatic georeferencing technology that aligns objects corresponding to RCS-based GIS from an LCS basis. The aim is to develop an algorithm that automatically transforms coordinates for models at or below the LoGeoRef10 level, which do not contain coordinate system information.

3. Development of GIS georeferencing Algorithm for BIM Information Model

3.1 Algorithm development overview

In the context of linking BIM data with GIS data, in this study, we propose a georeferencing algorithm to align objects on a GIS map (digital map) based on an RCS with LCS-based data, by evaluating the level of location information in IFC.

The target site for this study is the headquarters building of LX Korea Land and Geospatial Informatix Corporation (located in Jeonju), using both the IFC FC2×3 and IFC4 versions. The corresponding map model in the GIS model is a 1:5000 scale digital map of the building, which is easily accessible for practical applications.

Figure 2 illustrates the five-step process for the georeferencing of the BIM model.

3.2 LoGeoRef measurement

To measure the level of location information in the IFC file, we utilized the IfcOpenShell library, an open-source tool that implements the mapping functionality for IFC files. The version of the IFC file is stored in the FILE_SCHEMA section of the IFC data, and we extracted this value using the schema function of the IfcOpenShell library. For IFC2×3 versions, we checked levels up to LoGeoRef10-40, and for IFC4 versions, we extended the check to the LoGeoRef50 level.

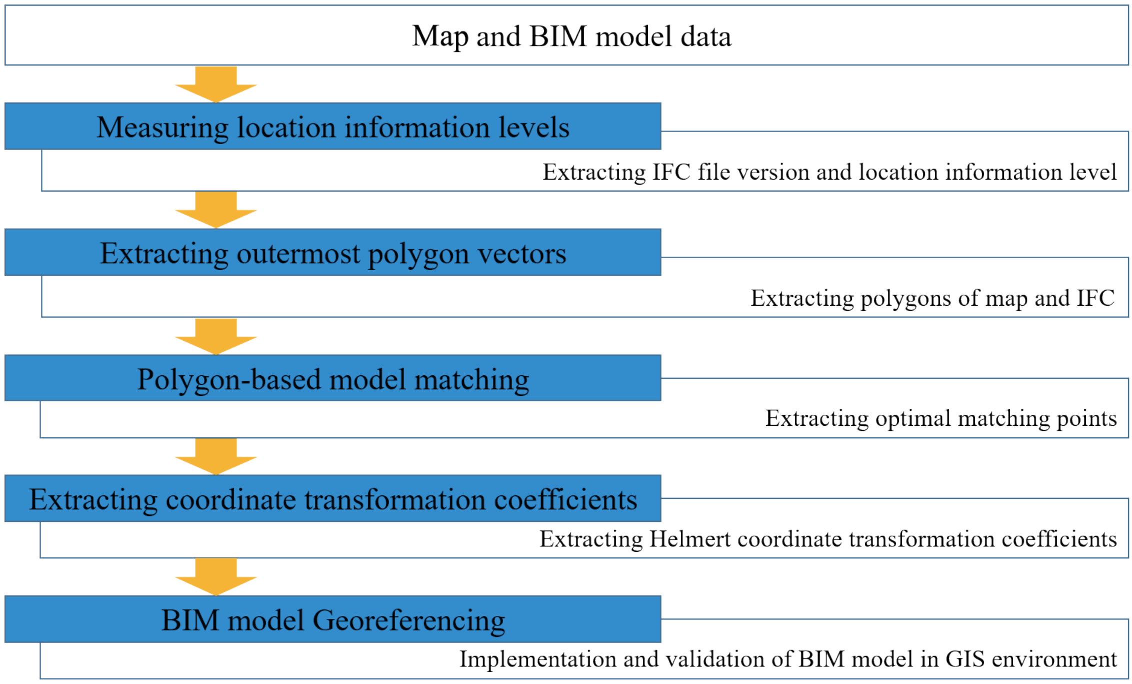


Fig. 2. (Color online) Five-step process for georeferencing of BIM model.

Once the version was confirmed, we verified the property information corresponding to each LoGeoRef level and returned the value indicating the level of location information for the respective IFC file.

Among the LoGeoRef levels, the stage that includes coordinate system information starts at LoGeoRef20. Therefore, for models at LoGeoRef10 or lower, we assigned actual location information values, assuming that accurate information was provided in the IFC model's entities. For IFC files below LoGeoRef20, georeferencing was performed, followed by the extraction of the outermost polygon.

3.3 Extraction of outermost polygon vectors between BIM and GIS models

This stage involves extracting polygons from each model to compare the IFC model, a 2D projection of the BIM, with the map model. The method connects the outermost points of the IFC model to generate the outer polygon. As the points comprising the IFC model consist of corner points of IFC objects, we selected key components of the IFC model such as IfcWall, IfcFooting, IfcColumn, IfcDoor, IfcWindow, and IfcSlab to extract the vertices of these corresponding objects.

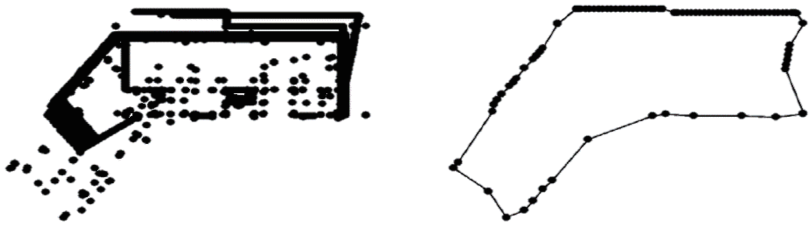
Each object contains information regarding its deviation in the form of (x, y, z) coordinates, representing the deviation from the coordinate value of the IfcLocalPlacement entity, which holds the reference coordinate information of the IFC model. By obtaining the reference coordinate value from the IfcLocalPlacement entity and adding the deviation value for each object, we were able to extract the corner points forming the object.

From the extracted point cloud, we applied an improved convex hull algorithm to identify the outermost points,(19) which were used to generate the outermost polygon vector for the BIM model. The map model, corresponding to the IFC, is a 1:5000 scale digital building model, commonly used in practical applications. We extracted the 2D outermost polygon of the target object based on the RCS.

Figure 3 illustrates the processes of extracting corner points from the IFC model and generating a 2D-based outermost polygon vector, while Fig. 4 shows the extracted polygon vector of the map model.

3.4. Polygon-based model matching of BIM and GIS

This stage involved comparing the polygons of the IFC and map models extracted in the previous step, identifying the optimal combination, and extracting the matching points of that combination. Moreover, geometric matching was performed, accounting for differences in coordinate systems and orientation, while assuming minimal differences in scale, as both the digital map and the BIM model represent real-world objects based on metric units. The polygon-based model matching process proposed in this study is illustrated in Fig. 5.



(a) (b)

Fig. 3. (a) Corner points of extracted IFC data and (b) extracted IFC outermost polygon.

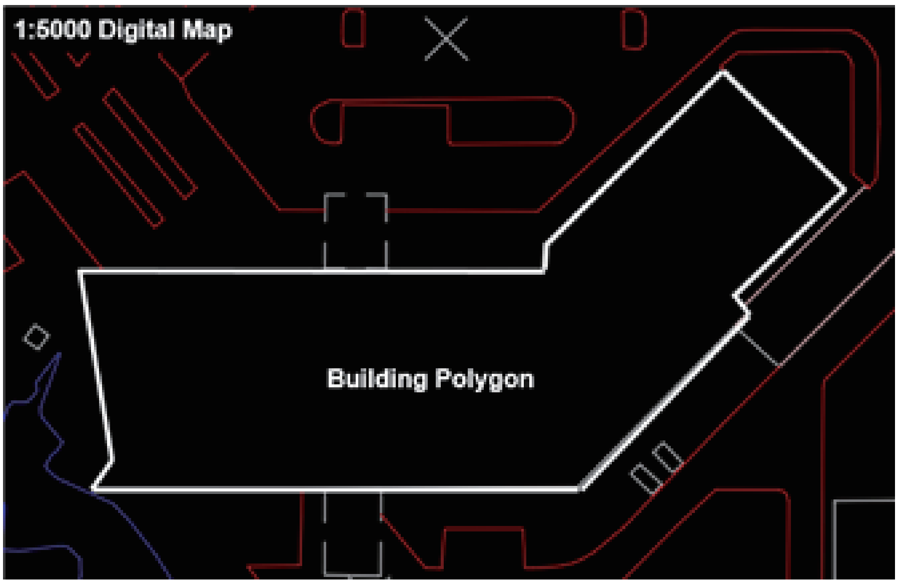


Fig. 4. (Color online) Polygon of map model.

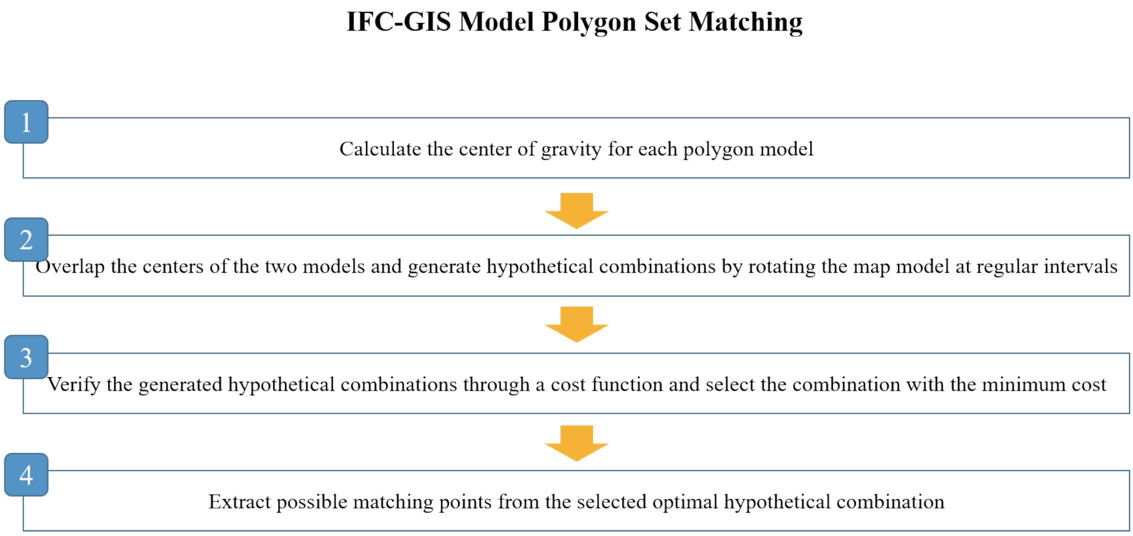


Fig. 5. (Color online) Matching algorithm of BIM-GIS model polygon set.

Before initiating the polygon-based model matching process, the IFC polygon model must be simplified using the widely applied Douglas–Peucker algorithm for feature point extraction.[20] Simplifying the IFC polygon model is necessary because comparing a simplified version with the map polygon reduces processing time. The Douglas–Peucker algorithm, commonly used in GIS and data compression, approximates a curve or line segment into a simpler form.

The process begins by selecting the start and end points with the greatest distance from the given set of points, followed by identifying the point furthest from the line segment connecting the start and end points. This distance serves as the criterion for dividing the points into two subsets. The same process is then applied recursively to these subsets, following a recursive approach. The algorithm continues until the maximum distance is less than a predefined threshold. The original curve or line segment is then approximated as a straight line connecting the start and end points.

Figure 6 shows the approximation of the IFC model's polygon through the Douglas–Peucker algorithm.

To compare the approximated IFC polygon with the map polygon by overlapping them, we calculated their centers of gravity. As shown in Fig. 7, we generated hypothetical combinations by rotating the map polygon in eight directions at regular intervals on the overlapped models.

Calculating the proximity between the points that form the map and IFC polygons becomes time-consuming and complex as the models increase in size. Therefore, as shown in Fig. 8(a), we divided the outer polygon of the IFC model into sections based on corner points. Then, as illustrated in Fig. 8(b), we calculated the proximity by measuring the distances between the points of the map model and each section of the IFC model.

To quantitatively verify the hypothetical combinations generated on the basis of the proximity of the IFC and map polygons, rotated at regular angles, we constructed a cost function using the concept of model closeness (MC). MC is represented in Eq. (1), indicating the goodness-of-fit between a given model M and data D.

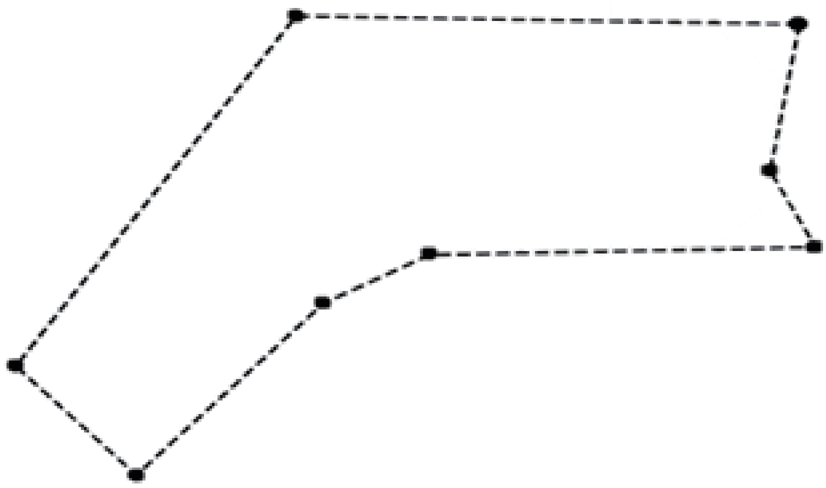


Fig. 6. Approximated polygon of IFC model.

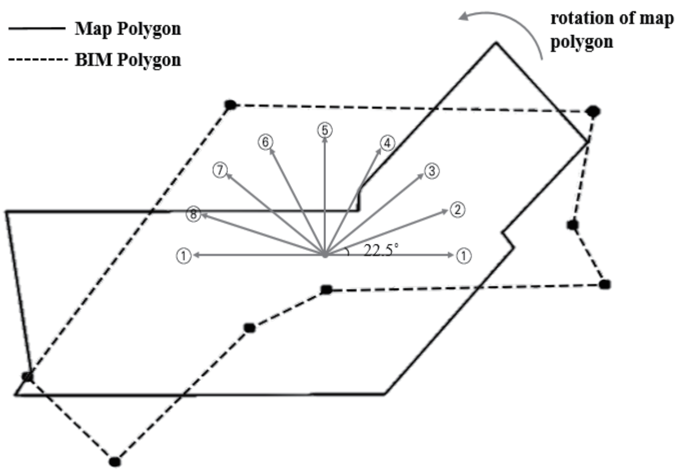
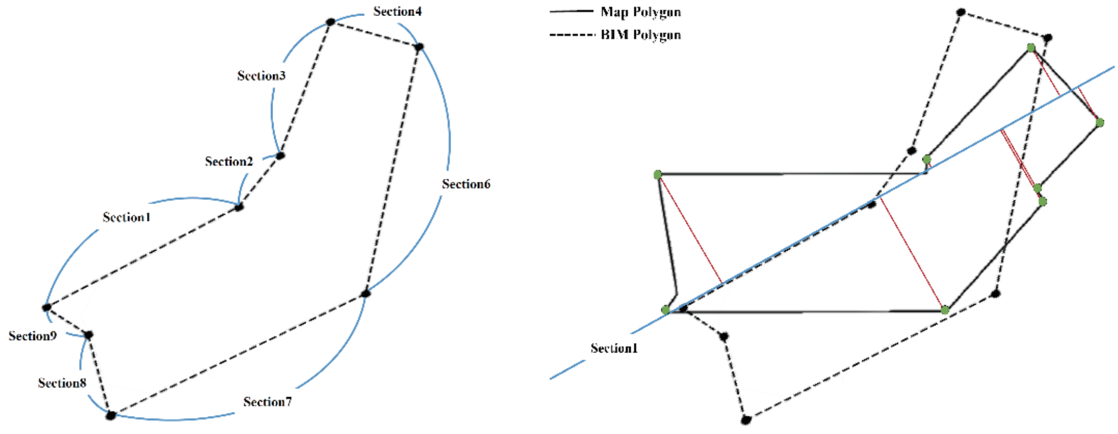


Fig. 7. Rotation hypothesis combinations of IFC and map polygons.

$$MC(D \mid M) = \frac{\Omega}{2 \ln 2} \tag{1}$$

Here, Ω denotes the weighted sum of the squared residuals between *M* and *D*, $[D-M]^T P[D-M]$ in the matrix form.

On the basis of this approach, we assigned the map and IFC polygons to M and D, respectively, and quantified the degree of difference between the two models in each hypothetical combination as residuals, assuming that the residuals follow a Gaussian distribution. Equation (2) represents the cost function used to select the optimal model M^* from the generated hypothetical combinations. The hypothetical combination with the minimum cost is then determined as the optimal overlapping configuration between the IFC and map models.



(a) (b)

Fig. 8. (Color online) (a) IFC polygon divided into sections and (b) relationship between map points and IFC sections.

$$M^{*} = \arg\min_{M \in \Phi} \{MC(D \mid M)\} \tag{2}$$

Figure 9 presents the results of calculating MC for each of the eight hypothetical combinations using the IFC and map polygon-based matching algorithms. Among these, #8 was selected as the optimal configuration, as it exhibited the minimum cost.

The matching points extracted from this optimal hypothetical combination are subsequently used in the next step to estimate the Helmert coordinate transformation coefficients.

3.5. Extraction of Helmert coordinate transformation coefficients

buildingSMART[21] presented an example of utilizing the property information of the IfcMapConversion entity to extract and apply Helmert coordinate transformation coefficients using corresponding points to georeferencing between the LCS and RCS models. IfcMapConversion consists of six property attributes, namely, Eastings (E), Northings (N), OrthogonalHeight (O), XAxisAbscissa (XAA), XAxisOrdinate (XAO), and Scale (S), which are linked to the Helmert coordinate transformation function as shown below.

$$\begin{bmatrix} X^R \\ Y^R \\ Z^R \end{bmatrix} = \begin{bmatrix} S\cos\theta & -S\sin\theta & 0 \\ S\sin\theta & S\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X^L \\ Y^L \\ Z^L \end{bmatrix} + \begin{bmatrix} E \\ N \\ O \end{bmatrix} \tag{3}$$

Here, $[X^R\ Y^R\ Z^R]$ and $[X^L\ Y^L\ Z^L]$ denote transformed coordinates with RCS and initial coordinates with LCS, respectively; $[E\ N\ O]$ represents translation in X, Y, and Z; θ is an orientation derived from atan2(XAO, XAA); and S represents a scale.

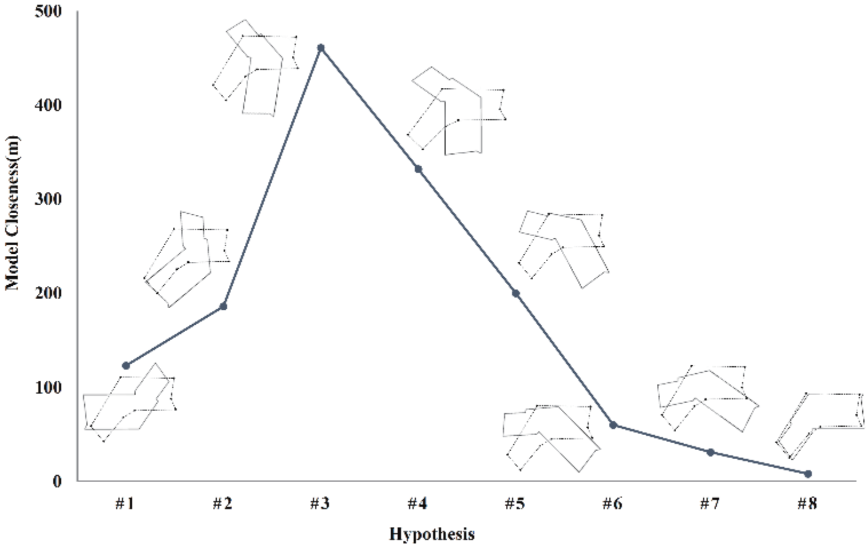


Fig. 9. MC relationship diagram by combination.

In Eq. (3), there are four unknown parameters (E, N, θ, and S). To estimate these parameters, at least two matching pairs of points are needed between models that have both coordinates before transformation and actual coordinates. In this study, on the basis of the model M^* composed of the optimal hypothesis combination, adjacent points between IFC and map polygons were matched by setting the allowable error of the map polygon model as a threshold. Figure 10 shows the extraction of a total of six vertex matching pairs from the optimal hypothesis combination, and these redundant points were used to calculate the Helmert coordinate transformation parameters.

3.6. BIM georeferencing results

This stage involved georeferencing the IFC model, which was previously at a LoGeoRef10 level or below, using the extracted Helmert coordinate transformation coefficients. By estimating these coefficients, all the extracted points from the existing IFC model were transformed. On the basis of these transformed points, the IfcLocalPlacement coordinate values, which serve as the reference coordinates for the IFC model, were calculated. This process enabled the IFC model to possess transformed coordinate data. Figure 11 presents an overlay image of the IFC model polygon, composed of the transformed points, with the map polygon.

To provide the IFC file with the transformed location information, we utilized IfcPatch, an open-source library that offers functionality for manipulating and processing IFC file formats. The IfcPatch library includes functions for modifying IFC file information. Among these functions, we used the OffsetObjectPlacements function to adjust the coordinate values of the IfcLocalPlacement entity. When the IfcLocalPlacement value obtained from the transformed IFC points is entered as a function parameter, it returns an IFC file containing those updated coordinate values. Figure 12 illustrates the resulting IFC file overlaid on a digital map.

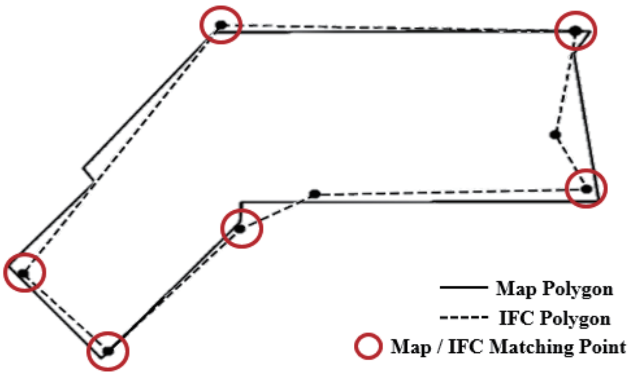


Fig. 10. (Color online) Matching points between IFC and map polygons.

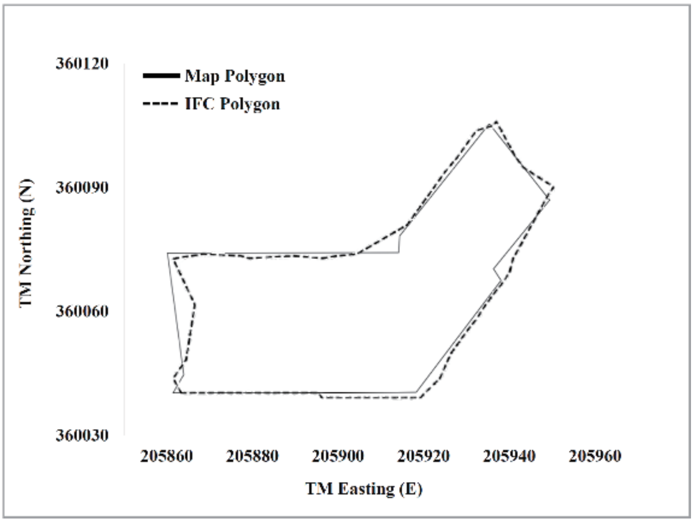


Fig. 11. Transformed IFC and map polygons.

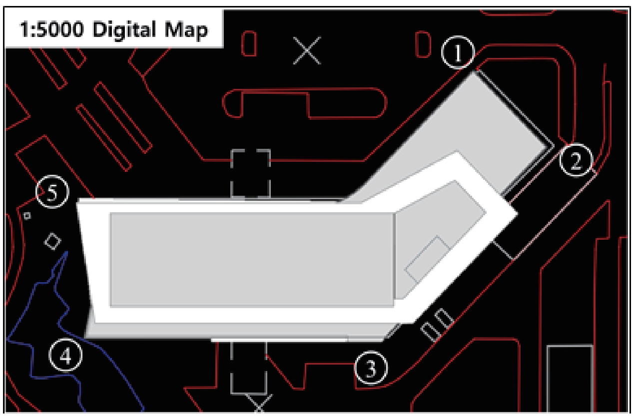


Fig. 12. (Color online) View of IFC file overlaid on a digital map.

As shown in Fig. 12, when the IFC and map models were overlaid, position errors were calculated for the five reference points. The results are presented in Table 4. The average error was 0.94 m, with minimum and maximum values of 0.66 and 1.20 m, respectively. According to Article 56 of the Aerial Photogrammetry Work Regulations issued by the National Geographic Information Institute of Korea, the standard deviation for a 1:5000 scale map is 0.72 m and the maximum allowable value is 1.44 m. On the basis of these standards, it can be confirmed that the georeferencing of the IFC model was performed using the 1:5000 digital map within an acceptable error range.

On the basis of the georeferencing results, 3D building model visualization was applied to the test platform. A projection of the IFC model onto OpenStreetMap (OSM) is shown in Fig. 13.

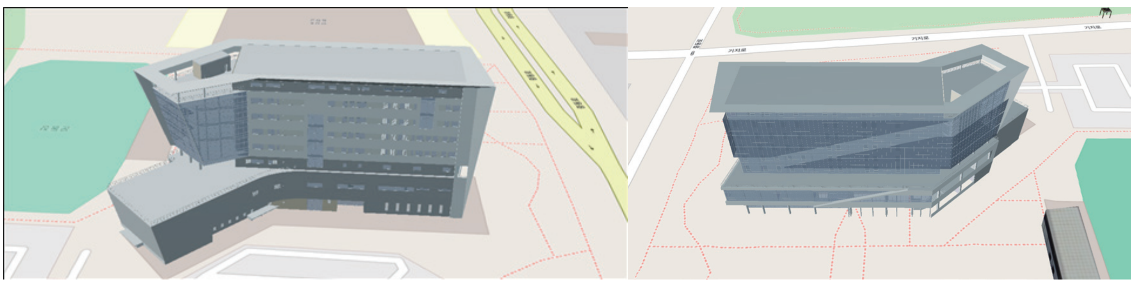
4. Conclusion

In this study, we examined the theoretical aspects of georeferencing with respect to integrating BIM and GIS data, reviewed previous research on the limitations of georeferencing in IFC (BIM standard), and explored studies on assigning location information through the georeferencing of BIM data. Through this, we identified limitations in providing absolute coordinates in IFC models that can be linked and utilized with GIS systems. To address these challenges, we proposed a pilot-automated georeferencing algorithm to position LCS-based IFC models on RCS-based maps.

The proposed technique offers the following advantages by applying a polygon-based geometric matching method instead of points or lines: First, the direction and corresponding

Table 4
Error in matching points between the map and IFC models.

No	Map model		IFC model		Error max distance (m)
	X(*E*)	*Y*(*N*)	*X*(*E*)	*Y*(*N*)	
1	205940.677	360099.078	205939.733	360099.002	0.947
2	205956.065	360083.951	205954.865	360083.870	1.203
3	205922.055	360045.845	205920.908	360045.529	1.190
4	205859.986	360045.679	205860.000	360046.342	0.663
5	205858.576	360073.48	205858.24	360072.867	0.699



(a) (b)

Fig. 13. (Color online) View of IFC model overlaid on OpenStreetMap.

vertices between BIM and map polygons do not need to be predetermined, nor is it necessary for the polygons to have the same number of vertices. Second, by extracting a polygon model that appropriately preserves the outermost shape of the BIM model and fixing the translation element to the model's center while generating hypothetical models using only the rotation element for matching with the GIS model, the number of hypothetical combinations is significantly reduced. This approach allows the method to handle BIM models with high geometric complexity to a certain extent.

The limitations identified in this study are as follows. First, uncertainty arises in identifying corresponding vertices when overlapping two models owing to errors between the calculated object center positions in BIM and map polygon models that represent the same object but have different geometric shapes. Second, uncertainty regarding orientation (front/back) occurs when overlapping BIM and map polygons in models with generalized symmetrical structures (e.g., rectangles and circles). Third, owing to constraints in available BIM data, we applied the proposed method on a pilot basis in this study. Hence, it will be necessary to conduct extensive experiments and discussions using various types of BIM models in future research.

Future research should not only address these limitations but also verify the level of georeferencing information provided by BIM models. On the basis of this, by exploring corresponding multidimensional GIS object models, further studies should aim to establish a series of automated pipelines that can be used to replace IFC models.

Acknowledgments

This work was supported by the Digital Land Information Technology Development Project (Project No. RS-2022-00143804) of the Ministry of Land, Infrastructure, and Transport and the Korea Agency for Infrastructure Technology Advancement.

References

1 Y. Song, X. Wang, Y. Tan, P. Wu, M. Sutrisna, J. C. P. Cheng, and K. Hampson: ISPRS Int. J. Geo-Inf. **6** (2017) 12. https://doi.org/10.3390/ijgi6120397
2 J. E. Kim and C. H. Hong: J. Korea Acad.-Ind. Coop. Soc. **18** (2017) 3 (in Korean). https://doi.org/10.5762/KAIS.2017.18.3.56
3 OGC GeoTIFF standard: https://docs.ogc.org/is/19-008r4/19-008r4.html#_georeferencing (accessed September 2024).
4 USGS: https://www.usgs.gov/faqs/what-does-georeferenced-mean (accessed September 2024).
5 A. Hackeloeer, K. Klasing, J. M. Krisp, and L. Meng: Ann. GIS **20** (2014) 1. https://doi.org/10.1080/19475683.2013.868826
6 X. Yao: Int. Encyclopedia of Human Geography (2nd ed.) (2020) 111–117. https://doi.org/10.1016/B978-0-08-102295-5.10548-7
7 G. Uggla and M. Horemuz: Autom. Constr. **96** (2018) 554. https://doi.org/10.1016/j.autcon.2018.10.014
8 F. Noardo, T. Wu, K. Ohori, T. Krijnen, and J. Stoter: arXiv 2020 arXiv:2011.03117v1 1–39. https://doi.org/10.48550/arXiv.2011.03117
9 S. Jaud, A. Donaubauer, O. Heunecke, and A. Borrmann: Autom. Constr. **118** (2020) 103211. https://doi.org/10.1016/j.autcon.2020.103211
10 R. H. Im: Korea Res. Inst. Hum. Settlements Working Paper (KRIHS, Sejong, Korea, 2020) (in Korean).
11 Esri: https://pro.arcgis.com/en/pro-app/latest/help/data/imagery/overview-of-georeferencing.htm (accessed September 2024).
12 Qgis: https://www.qgis.org/ (accessed September 2024).

13 C. Christian and G. Hendrik: Level of Georeferencing (LoGeoRef) Using IFC for BIM (2019).
14 F. Noardo, L. Harrie, K. Arroyo Ohori, F. Biljecki, C. Ellul, T. Krijnen, H. Eriksson, D. Guler, D. Hintz, M. A. Jadidi, and M. Pla: ISPRS Int. J. Geo-Inf. **9** (2020) 502. https://doi.org/10.3390/ijgi9090502
15 J. Irizarry, E. P. Karan, and F. Jalaei: Autom. Constr. **31** (2013) 241. https://doi.org/10.1016/j.autcon.2012.12.005
16 F. Noardo, K. Arroyo Ohori, T. Krijnen, and J. Stoter: Appl. Sci. **11** (2021) 2232. https://doi.org/10.3390/app11052232
17 J. S. Kim, J. H. Yom, and D. C. Lee: J. Korean Soc. Surv. Geod. Photogramm. Cartogr. **25** (2007) 117.
18 A. Diakite and S. Zlatanova: Comput. Environ. Urban Syst. **80** (2020) 101453. https://doi.org/10.1016/j.compenvurbsys.2019.101453
19 E. Dey, M. Awrangjeb, F. Tarsha Kurdi, and B. Stantic: Proc. 2021 Digital Image Computing: Techniques and Applications (DICTA, Gold Coast, Australia) (2021) 1–8.
20 U. Ramer: Comput. Graph. Image Process. **1** (1972) 224. https://doi.org/10.1016/S0146-664X(72)80017-0
21 buildingSMART: User Guide for Geo-Referencing in IFC: https://www.buildingsmart.org/wp-content/uploads/2020/02/User-Guide-for-Geo-referencing-in-IFC-v2.0.pdf (accessed September 2024).

About the Authors

Young Sun Kim received her B.S. and M.S. degrees from Anyang University, Korea, in 2022 and 2024, respectively. Since 2022, she has been a research engineer at the Smart Urban Space Institute at Anyang University. Her research interests include urban planning, smart cities, and spatial information. (kysun4592@gmail.com)

Seon Cheol Yu received his B.S., M.S., and Ph.D. degrees from Anyang University, Korea, in 2003, 2005, and 2011, respectively. From 2011 to 2014, he was an associate research fellow at the Korea Research Institute for Human Settlement. Since 2021, he has been a professor at Anyang University. His research interests include urban planning, big data, smart cities, and urban regeneration. (scyu@anyang.ac.kr)

Yoon Seok Jwa received his B.S. degree from Inha University, Korea, in 2001, his M.S. degree from Ohio State University, USA, in 2006, and his Ph.D. degree from York University, Canada, in 2014. From 2016 to 2019, he was an R&D manager at Industrial SkyWorks, Canada. Since 2019, he has been working as a technical director in the platform development division at Soosung Engineering. His research interests are in drones, spatial big data, digital twin, and BIM. (ysjwa@soosungeng.com)

Jong Wook Ahn received his B.S., M.S., and Ph.D. degrees from Anyang University, Korea, in 2000, 2002, and 2007, respectively. From 2007 to 2011, he was an associate research fellow at the Korea Research Institute for Human Settlement, Korea. Since 2011, he has been a professor at Anyang University. His research interests include urban planning, smart cities, big data, and spatial information. (ajw0603@anyang.ac.kr)