WORLD CONFERENCE ON TIMBER ENGINEERING 2025
BRISBANE, AUSTRALIA
Advancing Timber for the Future Built Environment

RESEARCH ON CONVERTING SFM 3D POINT CLOUD TO 3D COORDINATE SYSTEM USING STLP POINT CLOUD CONNECTION DEVICE IN THE "JODO-SHU SAINEN-JI HONDOU"

Zhixing Wang[1] , Hito Sato[2] , Hisamitsu Kajikawa[3]

ABSTRACT: We used SfM(Structure from Motion) technology to generate several 3D point clouds through photography of the "SAINEN-JI HONDOU". These point clouds include the interior, exterior, and roof. To combine the three point clouds, we proposed the STLP(Stereo Tags-Platform Leveling) point cloud connection device.With this device we can adjusted the scale and leveling of the 3D point cloud and combine they into a 3D model.Furthermore, we convert the 3D model into CAD drawings.

KEYWORDS: Temple, SFM Technology,STLP Point Cloud Connection Device,3D Coordinate System,Automation

1 – INTRODUCTION

1.1 BACKGROUND AND PREVIOUS RESEARCH

SfM[1] is a technology that generates 3D point clouds from photos taken from multiple angles.Automated driving technology uses SfM(Structure from Motion) technology to generate a 3D point cloud of the surrounding environment using images and then determines the distance of surrounding objects from the point cloud while evaluating whether the object is a person or not[2].Compared to traditional laser scanning methods for generating point clouds, SFM (Structure from Motion) point clouds offer lower costs and greater flexibility. For example, by utilizing drones, it is possible to generate point clouds of building rooftops.

However there are few research on SFM in architecture due to the complexity of building structures. Generating 3D point clouds of buildings requires a large number of photographs. Additionally, factors such as lighting conditions and photographic equipment often result in the generation of multiple 3D point clouds.

1.2 RESEARCH OBJECTIVES AND METHOD OF PROCEEDING

In previous studies, we took photographs of various parts of the building. Since there were no common areas between the interior and exterior, multiple separate point clouds were generated from the photos — specifically, interior, exterior, and roof point clouds[3,4].This research seeks to clarify the impact of the STLP (Stereo Tags-Platform Leveling) point cloud connection device on the generated 3D point clouds.According to previous studies, the use of artificial tags can effectively improve the quality of point cloud generation in Structure from Motion (SfM), especially for challenging surfaces such as monochromatic walls where point clouds are difficult to generate[5,6].Additionally, research aims to verify the method of creating architectural drawings from the generated 3D point clouds and to assess the accuracy of these drawings.

First we set the STLP point cloud connection device in the main hall of temple which is captured from both inside and outside,with the device's top surface set to level.After photographing the interior, exterior, and roof, we use the SfM technology to generate the 3D point clouds form these photograph,the STLP device automatically linked the three since of 3D point clouds to generate a complete

3D point cloud of the main hall of the temple.Then we use the stereo tags to level the 3D point cloud,adjust the scale of it and generated a 3D model.We get the coordinates of pillars and earthen walls from the 3D model,then we determined the dimensions and positions,use the program displayed structural elements, dimensions, and reference lines in the 3D coordinate system.

[1] Zhixing, Architecture. Meiji University , Kanagawa, Japan, station2498@gmail.com

[2] Hito, Architecture. Meiji University , Kanagawa, Japan, station2498@gmail.com

[3] Hisamitsu , Architecture. Meiji University , Kanagawa, Japan, kajihisa@meiji.ac.jp

https://doi.org/10.52202/080513-0368

2 – EXPERIMENTAL RESEARCH FOR GENERATING 3D POINT CLOUDS WITH THE STLP POINT CLOUD CONNECTION DEVICE

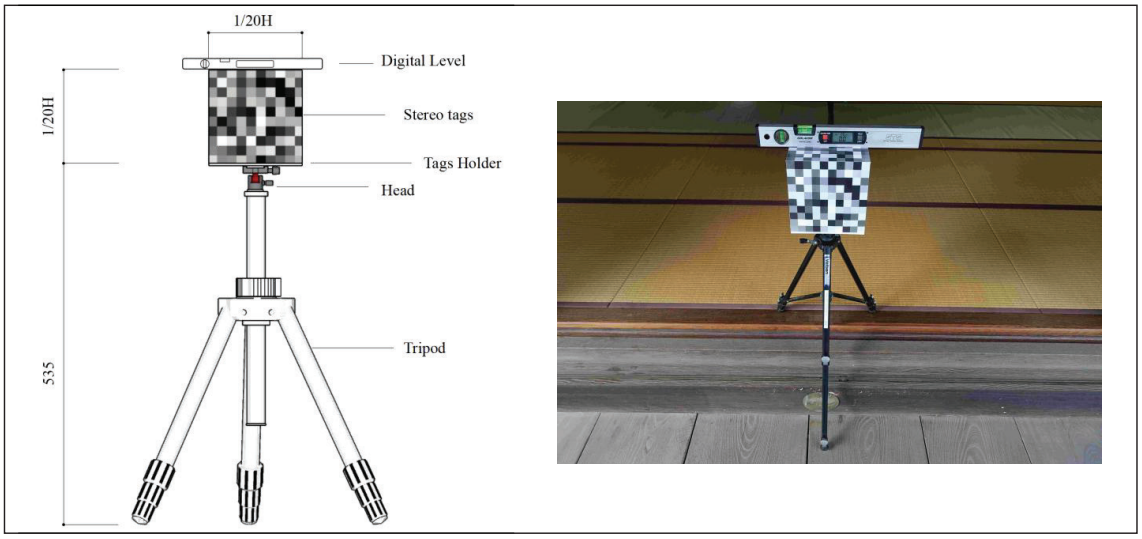

Figure 1.STLP point cloud connection device

2.1 OVERVIEW OF THE STLP POINT CLOUD CONNECTION DEVICE

The STLP point cloud connection device consists of a tripod, leveling device, marker fixation device, stereo marker, and level. The height of the device, as well as the dimensions, material, and design of the stereo marker, are adjustable. Figure 1 shows a schematic of the STLP point cloud connection device.

2.2 PARAMETERS OF THE STLP POINT CLOUD CONNECTION DEVICE

In this research, we use two kinds of stereo tags. We designed a 20 cm cubic box covered with a 2 cm grid pattern on its surface. One type features randomly distributed grayscale colors, while the other consists of black and white alternating patterns in figure2. The box is made of 3D-printed PVC material, and the surface is covered with standard printing paper.

2.3 SETTING OF THE STLP POINT CLOUD CONNECTION DEVICE

We installed one STLP device at the junction of each of the three corridors and the interior. These three devices form a sufficiently large isosceles triangle to facilitate subsequent coordinate system adjustments. Using an electronic level, we adjusted the top surfaces of the three devices to be perfectly horizontal. Then, with a laser level, we aligned the top surfaces of the three cubic boxes to the same height. Through these adjustments, we ensured that the top surfaces of the three boxes formed a large, flat horizontal plane.the device place show in the figure3.

3 – GENERATION OF 3D POINT CLOUDS ABOUT THE "JODO-SHU SAINEN-JI HONDOU"

3.1 CAMERA SETTINGS FOR PHOTOGRAPHY

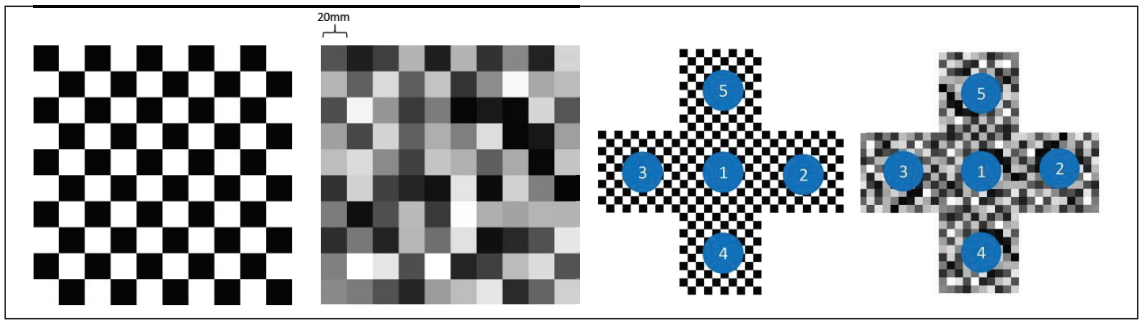

Figure 2.stereo tags

https://doi.org/10.52202/080513-0368

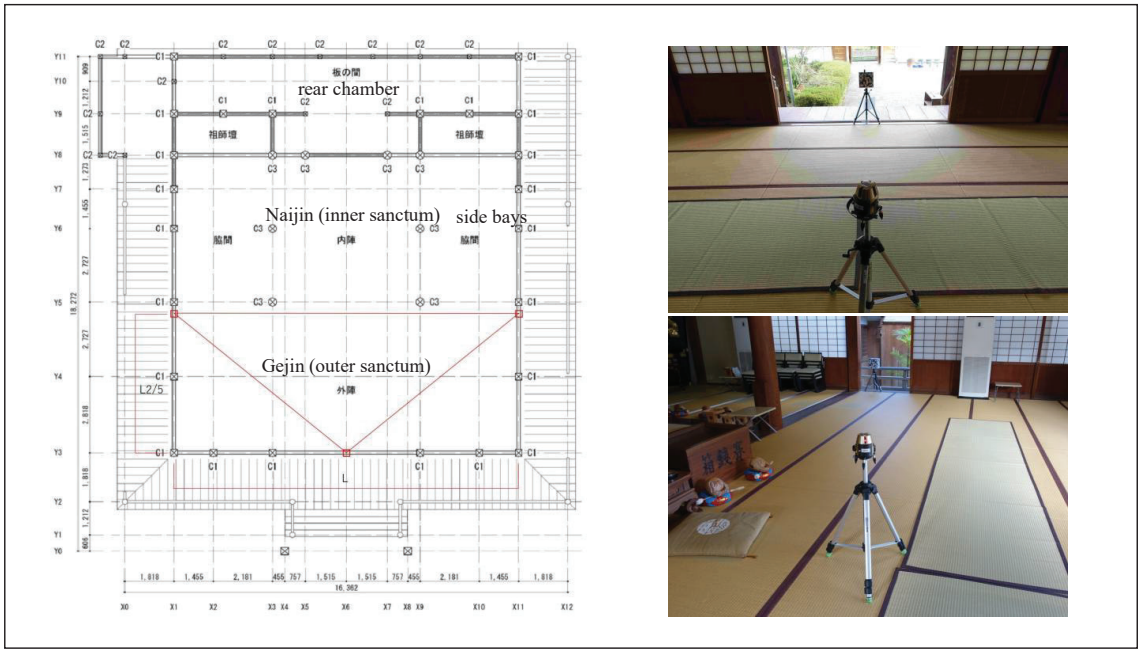

Figure 3.setting of the STLP point cloud connection device

In this research a single-lens camera and a drone were used for photo capturing. The shooting points were determined using a measuring tape on the ground. Then, a laser beneath the tripod was used to align the tripod's position with the ground markers. The specific parameters are shown in the table2.

Table 1:the setting of the camera

Item	Specification
Camera	SONY α7R (Full-frame mirrorless single-lens camera)
Image Sensor	35mm Full-frame (35.9×24.0mm), "Exmor" CMOS sensor
Lens	SONY Vario-Tessar, focal length 16–35mm F4 (wide-angle zoom lens)
Shutter Speed	1/60
F-number	Auto
ISO	Auto

Table 2:Shooting Rules

Item	Value
Object Height (Ceiling Height)	H
Shooting Distance	0.7H
Shooting Height	0.45H
Shooting Angles (Inner Viewpoints)	Front 0°, Left 30°, Right 30°, Up 30°, Down 30°
Shooting Angles (Outer Viewpoints)	Front 0°, Left 30°, Right 30°, Up 30°, Upper Left 30°, Upper Right 30°, Down 30°, Lower Left 30°, Lower Right 30°
Distance Between Shooting Locations	800 mm

TAKING THE PHOTOGRAPH

After the shooting points was set on the ground,we take the photograph of each elevations,some of then are the wall of "HONDOU",and some of the elevations just have column and roof beam.We divided the photographic subjects into elevations along the X-axis and Y-axis using a floor plan. A total of X1, X3, X9, X11, Y3, Y5, Y6, and Y8 elevations were photographed. Except for Y8, all elevations were captured from two directions to generate a complete point cloud model. The detailed shooting locations for elevation Y3 are illustrated in Figure 4 below.

sed a camera to photograph most of the interior and exterior facades of the temple. For the roof, we used a drone to capture video footage and then extracted consecutive frames from the video. This method effectively generates a large number of overlapping images, which can be used to efficiently create a point cloud.

3.2 RESULTS OF THE PHOTOGRAPH

Because the shooting positions at X1 East, Y3 North, and X11 West were close to the axes, the camera angle was rotated by 60° during photography. Consequently, nine photographs were taken at each of these points. Additionally, there were several planned shooting positions where photography could not be carried out due to unexpected obstacles.Tables 3 show the number of photographs taken using random pixel and black-and-white pixel markers,the number of photographs in which markers appeared,and the proportion of photographs containing markers

https://doi.org/10.52202/080513-0368

Table 3:number of the photograph for each axis

Y03	South	70
Y03	North	171
Y05	South	70
Y05	North	52
Y06	South	55
Y06	North	60
Y08	North	46
X01	East	153
X01	West	50
X03	East	55
X03	West	30
X09	East	25
X09	West	55
X11	East	60
X11	West	180
	Total	**1,132**

3.3 GENERATION OF THE 3D POINT CLOUD OF THE HONDOU FROM PHOTOGRAPHS

Considering that we set up two different markers in total, we captured two sets of photos, amounting to 3056and2564 images altogether. We imported both sets of photos into Metashape software to generate point clouds. The results are shown in the figure5 below.

From the results, the stereo markers with randomly arranged pixels generated a larger number of point clouds. Moreover, the interior, exterior, and roof sections of the point clouds were automatically combined into one complete point cloud. However, the stereo tags with alternating black-and-white pixels resulted in missing points on the external wall along the x3 axis. Photos taken from outside the x3 axis were not automatically combined with the internal photos via the stereo marker into a single, complete point cloud; instead, they formed a separate point cloud.

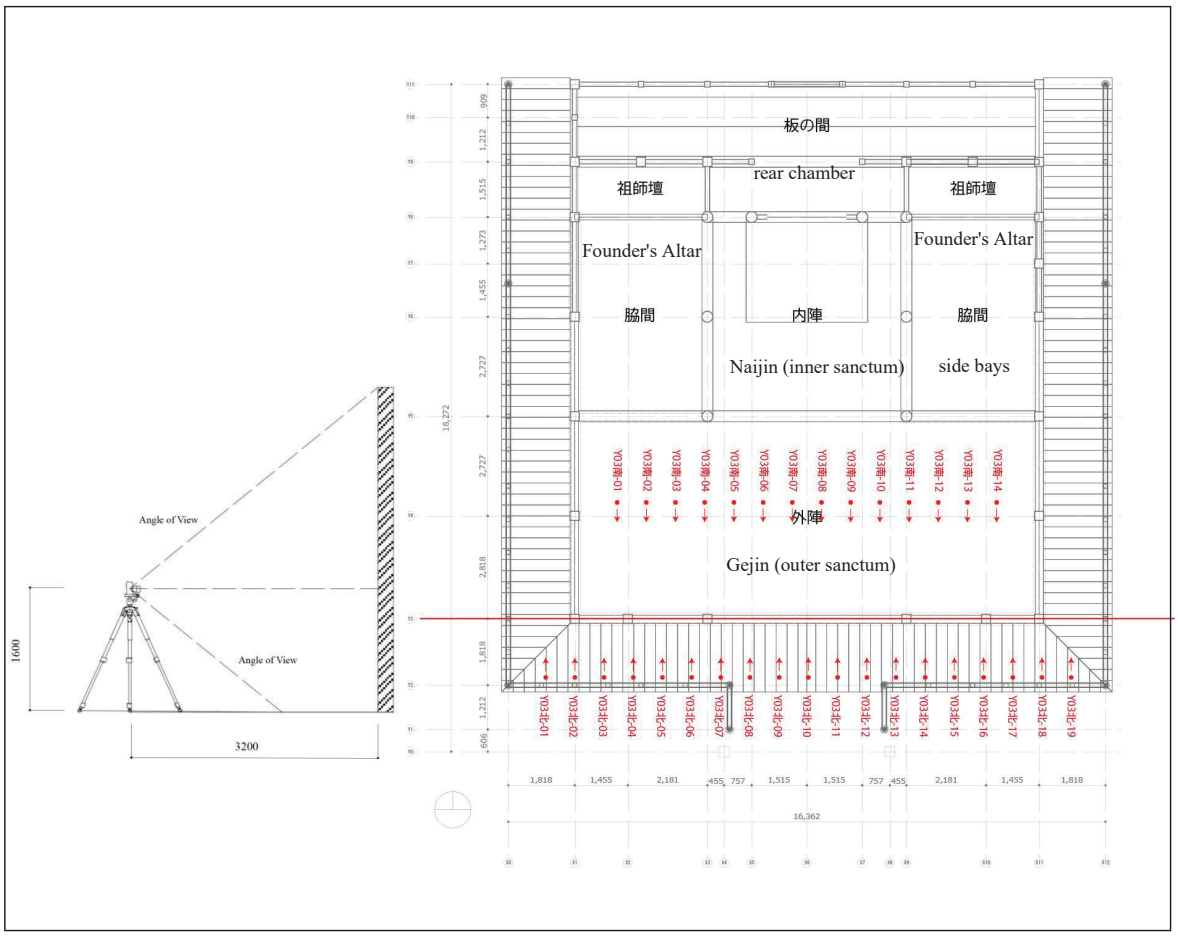

Figure 4.Relationship Between Camera Position and Object Height H and shooting locations for elevation Y3

Future tasks include addressing the issue that drones do not have fixed calibration positions like ground-based DSLR cameras, resulting in significant variations in drone flight paths between two different marker setups. In future research, it will be necessary to consider predefined drone trajectories to generate stable rooftop point clouds. Additionally, some photographs exhibited overexposure. Due to the automatic aperture and shutter settings used in this study, certain photos taken in the evening were overexposed, causing marker regions to appear entirely white and negatively impacting the point cloud alignment

represented as A (Ax, Ay, Az), B(Bx, By, Bz), and C(Cx, Cy, Cz). The direction of vector AB defines the positive X-axis, The positive Y-axis direction is defined by the vector originating from point C and perpendicular to line AB.and plane ABC is the horizontal plane. The origin of the XY-plane coordinate system is defined by the XY coordinates of point C.The GL plane is determined by the center of the stone stairs located in front of the temple.

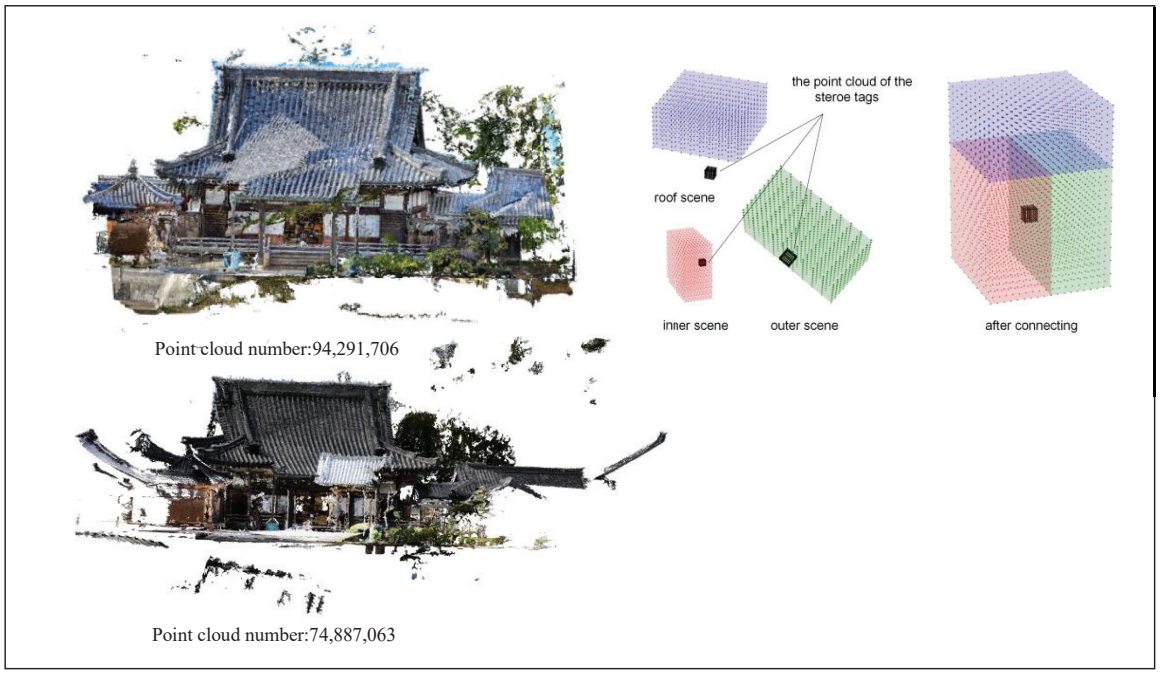

Figure 5.point cloud of the HONDOU and the principle of the STLP point cloud connection device

4 – CONVERT THE 3D POINT CLOUD OF "JODO-SHU SAINEN-JI HONDOU" INTO THE CAD COORDINATE SYSTEM

4.1 OVERVIEW OF THE BASIC CAD COORDINATE SYSTEM

Since each point cloud has its own local coordinate system, we need to convert these point clouds into a standard building coordinate system through rotation, translation, and scaling. In this experiment, by extracting the point clouds of three stereo markers, the central coordinates of the top square sections of the markers were calculated. These three points formed a large triangular plane. Through rotation and translation, the local coordinate system was transformed into the CAD base coordinate system. Additionally, the length of the triangle's longer axis was measured on-site, allowing the point cloud to be scaled to match the actual building dimensions. The coordinates of the three points are

4.2 CONVERT THE 3D POINT CLOUD INTO THE CAD COORDINATE SYSTEM

To transform the local coordinate system of the generated 3D point cloud, it is necessary to find the center coordinates of the top surface of the three cubic connection devices in the STLP point cloud. Therefore, we need to classify the point cloud to extract the point cloud corresponding to the cubic marker. Here, the DBSCAN method[7] is used for point cloud classification, and its principle is as follows.

Table 4: the parameter of DBCSAN

Parameter	Description
ε(Epsilon,neighborhood radius)	Determines the search range of the neighborhood.
MinPts(Minimum number of points)	Determines the criterion for identifying a core point.
p (Target point)	The point being examined to determine if it is a core point.

https://doi.org/10.52202/080513-0368

q (Neighbor point)	A point that may belong to the same cluster.
D (Dataset)	The entire point cloud or data collection.
dist(p, q) (Distance function)	Measures the distance between two points, usually using Euclidean distance.
Nε(p) (ε-Neighborhood)	The set of all points within a radius ε of the target point p.

$$N\varepsilon(p) = q \in D | dist(p,q) \leq \varepsilon \quad (1)$$

$$|N\varepsilon(p)| \geq MinPts \quad (2)$$

Point cloud classification was performed using different neighborhood radii of 0.5, 1, and 2, while setting MinPts to 10. As shown in the figure below, when the neighborhood radius was set to 1, the point cloud corresponding to the 3D marker was well three STLP components, and the point clouds of these parts were extracted. The results are shown in the figure below.Next, the extracted cubic point cloud was classified based on normal vectors, where vectors with an angle less than 15° were grouped into the same category, allowing the five faces of the cube to be classified separately. The top face was then extracted, and the RANSAC method was used to fit a plane using the point cloud from this surface. Subsequently, the entire cubic point cloud was projected onto this plane, resulting in a 2D point cloud with a square shape. Finally, RANSAC was applied to fit three edges of the square, and the center coordinates of the square were calculated.

Table 5:Top central coordinates of three 3D tags

Random		x	y	z
DBSCAN	A	-25.256	-6.885	-15.532
	B	47.235	-6.042	-61.544
	C	-85.873	-2.176	-75.629
White and Black				
DBSCAN	A	-94.383	-22.482	-24.858
	B	121.762	-47.943	-270.606
	C	-367.497	-7.519	-181.363

After computing the coordinates of points A, B, and C, the normal vector of the plane formed by these three points is first determined. Then, the angle between this normal vector and the Z-axis of the CAD coordinate system is calculated. A rotation matrix is then computed to align the ABC plane with the XY plane, and the rotated coordinates are represented as A', B', and C'.Next, the point cloud is translated so that the XY coordinates of point A become (0,0), with the new coordinates represented as A". Subsequently, the point cloud is scaled based on the measured distance between points B and C, obtained from an on-site laser rangefinder, and the new coordinates are denoted as A'".Finally, the point cloud is translated in the Z-direction to ensure that the height of the GL plane is set to 0.The final point cloud coordinate system is shown in Figure 7

Table 6:Top central coordinates of three 3D tags after translation

Random		x	y	z
DBSCAN	A	-6.667	5959.94	2361.878
	B	-6335.702	920.854	2245.907
	C	6330.912	925.538	2255.659
White and Black				
DBSCAN	A	6.189	6071.91	2254.709
	B	6523.615	11141.7	2245.579
	C	-6267.875	11228.4	1727.661

5 –CREATING A FLOOR PLAN USING A CAD 3D POINT CLOUD

5.1 REMOVAL OF NOISE POINTS

Before generating a floor plan from the CAD 3D point cloud, it is first necessary to process the point cloud by removing noise points. This includes filtering out sparse

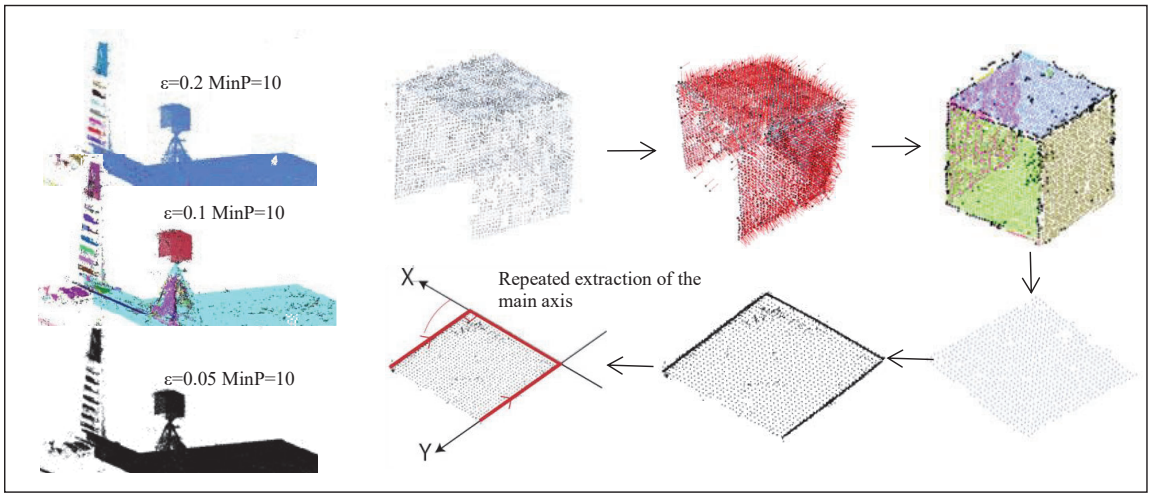

Figure 6.Extraction of the center coordinates of three stereo tags

https://doi.org/10.52202/080513-0368

points and points that are not part of the main building structure, thereby reducing the computational load during the floor plan generation process.In this experiment, point cloud projection and hexagonal voxelization were used to remove some of the noise points. Since the camera captured a large number of high-resolution images, the generated point cloud is generally of high density, including unnecessary elements such as non-building structures, surrounding trees, and the ground. Simply filtering the point cloud based on 3D density makes it difficult to distinguish the building structure from the surrounding noise points.

To address this, the point cloud was projected onto planes from three different directions, ensuring that the 2D projection of the building structure has a higher density. Additionally, instead of using conventional square voxelization, a hexagonal voxelization method was applied, allowing more point cloud data to be retained along the edges of complex building shapes.A hexagonal radius of 50 mm was defined, and regions with point densities below 50 were progressively filtered out. The minimum threshold was increased by 10 in each iteration. After several iterations, the point cloud was reduced to approximately 700 million points, effectively .The point cloud after partial noise removal is shown in Figure 8

5.2 EXTRACTION OF COLUMNS AND WALLS

To generate a standard CAD floor plan from a CAD 3D point cloud, it is necessary to extract and classify the point clouds corresponding to columns and walls from the building structure. From the column point cloud, the width and center coordinates of each column are extracted. Using the center coordinates of all columns, an axis line diagram is constructed. Then, the wall widths are extracted from the wall point cloud, and the walls are drawn along the axis lines to produce a building floor plan.

First, we performed a vertical point cloud density analysis on the CAD 3D point cloud. Two distinct peaks were observed, corresponding to the ceiling and floor point clouds. Therefore, the point clouds representing columns and walls should be extracted within this height range. The results are shown in the figure below.The extracted height of the ceiling is 5735 mm, and the floor height is 1565 mm. Therefore, the calculated height of the columns is 4150 mm in the figure 9.We extracted the point cloud within this range, and then manually selected

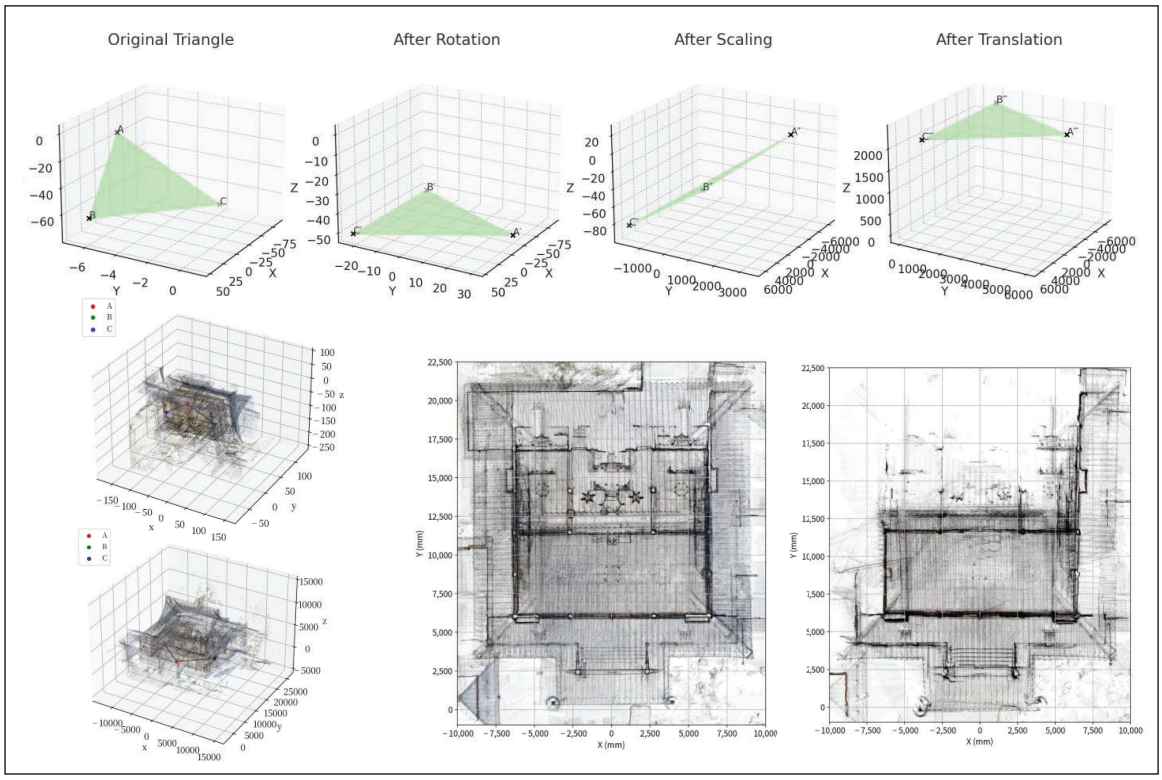

Figure 7.Identification of GL points and transformation of the coordinate system

https://doi.org/10.52202/080513-0368

all the columns along with a portion of the surrounding point cloud on the XY plane. For the square columns, we analyzed the point cloud distribution along the X and Y directions, extracted the four principal axes of each column, and calculated their center coordinates.the point cloud of column show in the figure9.

Table 7:Column width and accuracy

Item	X3Y3 Corner	X9Y3 Corner
SFM 3D Point		
Width w [mm]	235.23	239.95
Depth l [mm]	239.71	236.24
Average [mm]	237.47	238.1
Measured		
Width w [mm]	238.98	241.79
Depth l [mm]	237.86	241.31
Average [mm]	238.42	241.55
SFM3D /		
Width w	0.98	0.99
Depth l	1.01	0.98
Average	1	0.99

generating a CAD floor plan

To establish the structural grid lines (X and Y axes) from the extracted column center coordinates, the following method was used:Each column file (e.g., x3y3.ply) was named according to its structural grid position, where the prefix x3 indicates it is aligned along the third X-axis and y3 denotes alignment along the third Y-axis. Using these identifiers, all columns were grouped based on their respective X and Y axis indices.For each axis group:The average of the X center coordinates was computed for all columns sharing the same X index. This average value was then defined as the position of the corresponding X grid line.Similarly, the average of the Y center coordinates was calculated for all columns with the same Y index, defining the location of the respective Y grid line.The computed axis lines were then visualized and show in the figure9.

6–CONCLUSION

• This experiment verified that the STLP point cloud stitching device can connect different spatial regions lacking common points, successfully generating a unified point cloud.

· It also demonstrated that 3D tags with random colors produce denser point clouds and achieve better spatial connection performance compared to black-and-white 3D tags.

· A method was established for extracting the center point of the top surface of a 3D marker in the STLP point cloud connection system using DBSCAN and RANSAC, from an SFM 3D point cloud without level adjustment, scaling, or coordinate axis settings.

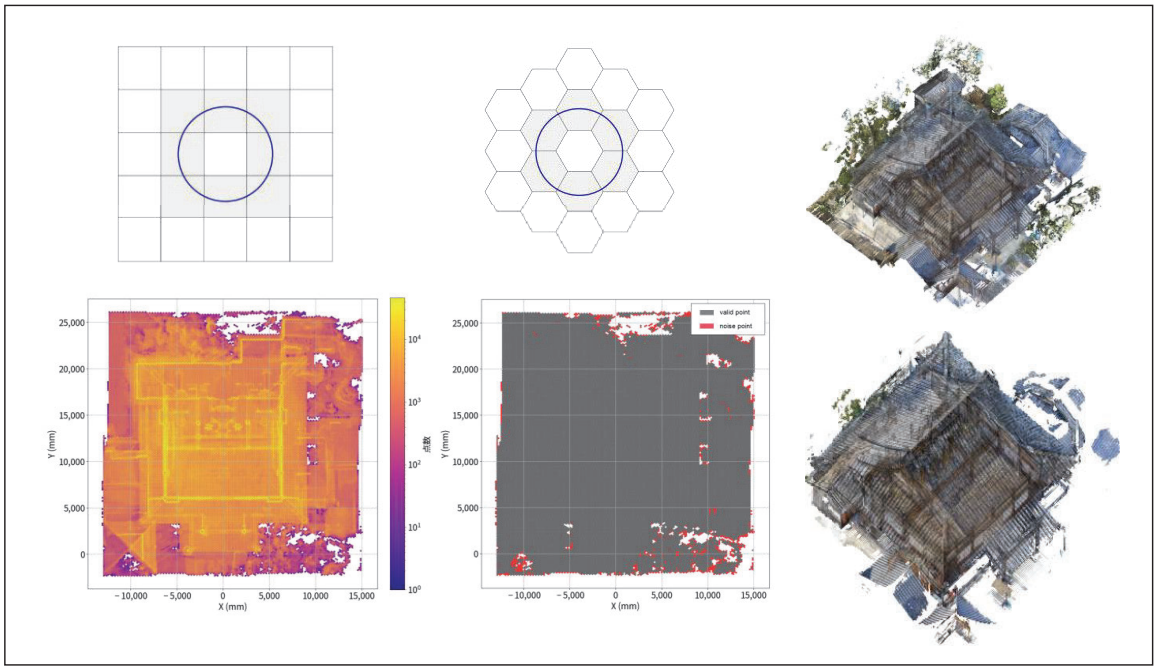

Figure 8.noise removal from the point cloud

https://doi.org/10.52202/080513-0368

· By setting three reference points from the STLP point cloud connection system on the SFM 3D point cloud and combining translation and horizontal movement using Rodrigues' rotation formula, it was possible to establish the coordinate axes in the CAD base coordinate system.

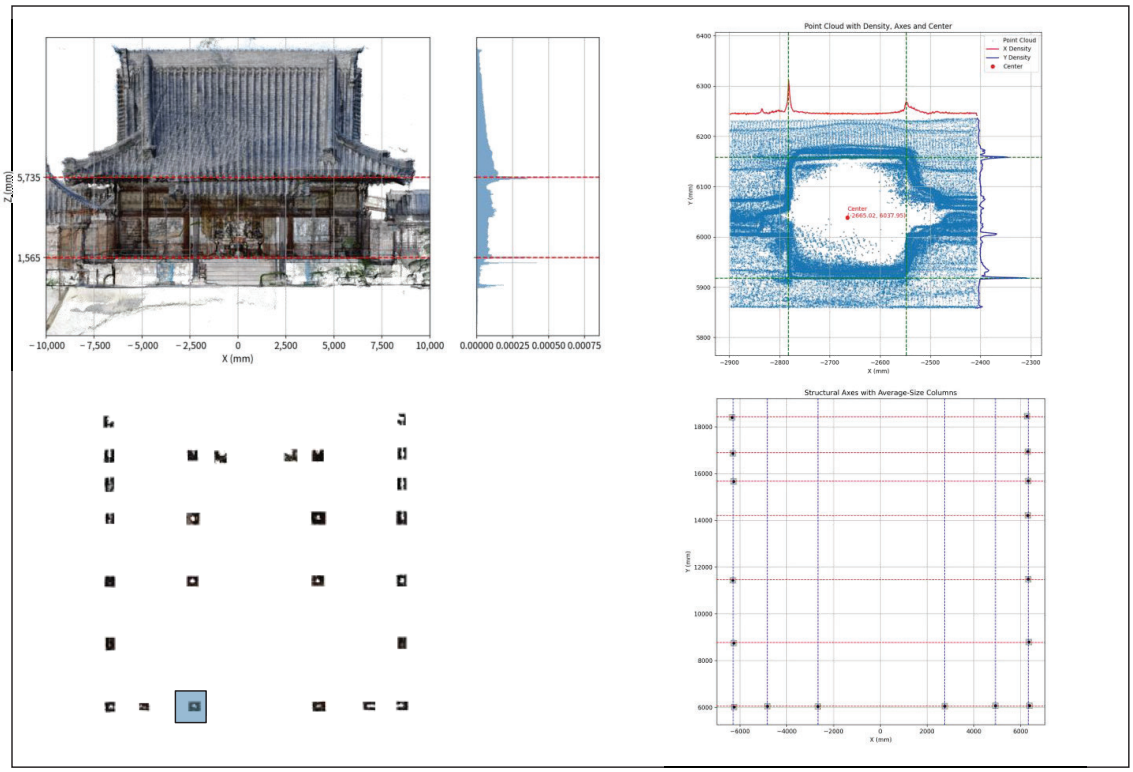

Figure 9.Extraction of column point clouds and generation of floor plans

· The width and center coordinates of each column were extracted from the CAD 3D point cloud, the accuracy of the columns generated using this method was verified, and part of the floor axis diagram was inferred.

7–REFERENCES

[1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge, U.K.: Cambridge University Press, 2004.

[2] Shiyu Song, Manmohan Chandraker, Clark C. Guest：High Accuracy Monocular SFM and Scale Correction for Autonomous Driving ， IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 38, NO. 4, APRIL 2016

[3] Shisei Ou, Hisamitsu Kajikawa, Akito Kikuchi： Verification of Automation of 3D Reconstruction, Drawing and Seismic Diagnosis by SFM Technology Part I: Study on Photography Method and Point Cloud assemble Method , Summaries of technical papers of Annual Meeting, Architectural Institute of Japan(Hokkaido), Sep,2022

[4] Shisei Ou, Hisamitsu Kajikawa, Akito Kikuch, Hito Sato： Research on Automation of 3D Reconstruction, Drawing, and Seismic Diagnosis by SFM Technology Part 2 Research on Photography and 3D Reconstruction of the Jodo-shu Sainen-ji Hondou, Summaries of technical papers of Annual Meeting, Architectural Institute of Japan (kinki) , Sep,2023

[5] [1] J. DeGol, T. Bretl, and D. Hoiem, “Improved Structure from Motion Using Fiducial Marker Matching,” in Computer Vision – ECCV 2018 Workshops, Lecture Notes in Computer Science, vol. 11129, Springer, 2019, pp. 285–300.

[6] [1] K. Iijima and M. Tsuji, "Study on the Utilization of 4D Building Models Created by Photogrammetry, Part 4: Efficient Modeling of Building Walls with Repeating Patterns Using 3D Markers," in Proceedings of the Annual Conference on Information System Technology, vol. 2022, pp. 17–18, Jul. 2022. (in Japanese)

https://doi.org/10.52202/080513-0368

[7] Ryo kashiba, Meguru hiraoka, Eisuke mitsuda,Segmentation and 3D modeling methods for structural analysis by machine learning algorithms on point clouds of traditional wooden buildings,J. Struct. Constr. Eng., AIJ, Vol. 89, No. 824, 1176-1183, Oct., 2024