A research about the automatic construction of Web/3D model house from 2D digital pictures

Etsuji Kitagawa, Graduate School Informatics, Kansai University, Osaka, Japan
(kitagawa@kansai-labo.co.jp)

Shigenori Tanaka, Faculty of Informatics, Kansai University, Osaka, Japan
(tanaka@res.kutc.kansai-u.ac.jp)

Hitoshi Furuta, Faculty of Informatics, Kansai University, Osaka, Japan
(furuta@res.kutc.kansai-u.ac.jp)

Toshiyuki Sugimachi, Graduate School Informatics, Kansai University, Osaka, Japan
(sugimachi@kansai-labo.co.jp)

Summary

In recent years, the survey is performed for repair, such as a bridge and a building built in past, spending great expense. And it is anxious for the survey technique that doesn't need cost and time more. Then, we made an idea of the technique of precise 3D model creation by 2D pictures. However, the technique of performing the improvement in accuracy of convergent photographing and automatic acquisition of corresponding points was not established.

Therefore, in this research, we try to obtain a semi-automation of corresponding points acquisition from initial corresponding points and the improvement in accuracy of convergent photographing. Moreover, we applied the research to the used house of Japanese real estate, and the applicable field was selected as the high needs of the residence of 3D model. And we developed the system that everyone could create Web / 3D model house by VRML easily without requiring expensive apparatuses or expertise.

1 Introduction

In recent years, there is a tendency that the various information are digitized on Internet and those are opened to public. Especially housing information is closely connected with the life of people. And it requires a immediacy and diversity. At present, a buyer or a lease person is interested in the used building rather than the new building by influence of depression. However, only presenting of the 2D information or the photographic information on a sketch is opened in the public presentation of used housing information. Therefore, if neither a purchaser nor a lease person visits a used house, they can not understand the present condition of it.

On the other hand, existing 3D modeling technique of the housing requires a cost. In addition, since there are many subjects, 3D modeling is difficult technically. This technique in present research makes it possible to create a 3-dimensional model house with high accuracy by only taking the digital photograph. Then, in this research, we aim at the development of a system that everyone can create 3D model virtual house on Web.

2 Outline of system

2.1 Input data

Firstly, in this system, we use 2D digital images to input to two or more sheets. Moreover we put a photographic subject to the simple room without furniture. Secondly, as input data, users select three or more corresponding points on each plane. If there are no suitable datum points of

three or more points among the selected points, users have to select the datum points of three or more points by using a mouse.

2.2 Details of system

In the previous research, we used the Genetic Algorithm (GA) suitable for the nonlinear combination problem. And we established the technique for determining the sub-optimal path points. As the result, we developed a system that is available to create 3D model spaces easily without expensive equipments or expertise. However, there were two subjects in the research. The 1st is that 3D model cannot be created at high speed because corresponding points are selected manually. The 2nd is that the accuracy of convergent photographing is low compared with parallel photographing. Then, in this research, we realize the improvement in accuracy of convergent photographing and the semi-automatic acquisition of corresponding points.

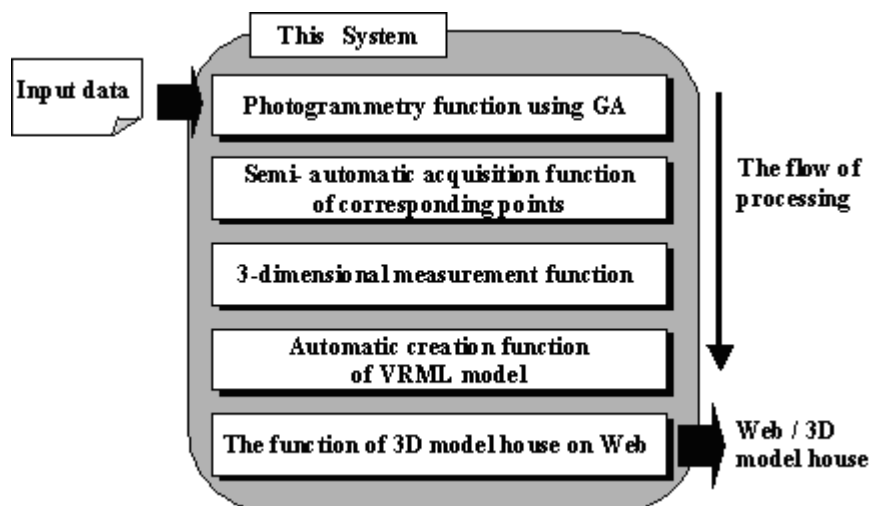This system consists of five functions, as shown in Figure 1.

Figure.1 The Outline of System

3 Photogrammetry function by using GA

3.1 Calculation of a camera parameter

The parameter of a camera is required to compute a photographic subject's 3-dimensional information by photogrammetry. Then, the parameter of a camera, used by Semi-automatic acquisition function of corresponding points and 3-dimensional measurement function, is calculated by using GA. The parameters of a camera are the position (Bx1, By1, Bz1) (Bx2, By2, Bz2) of lens centers and the inclination ($\omega1$, $\kappa1$, $\varphi1$) ($\omega2$, $\kappa2$, $\varphi2$) when the photograph of right and left side is taken. The parameters of a camera are the position of a lens center (Bx1, By1, Bz1) (Bx2, By2, Bz2) and the inclination ($\omega1$, $\kappa1$, $\varphi1$) ($\omega2$, $\kappa2$, $\varphi2$). The determination of the path points is the most important to calculate the parameter of a camera. However, an advanced specialist's knowledge is needed for the determination of path points. Then, in the previous research, we established the technique of determining the combination of the sub-optimal path points by using GA.

3.2 Correspondence to convergent photographing

In the previous research, we established the technique of precise 3-dimensional model creation that does not need expertise. However, there was the subject that compared with the parallel photographing whose average errors were less than about 2%, the accuracy of convergent photographing was low. The cause of this subject is that the combination of the path points that make errors less than about 2% by Photogrammetry function by using GA in convergent photographing is incomputable. The improvement in the accuracy of convergent photographing is applicable to various uses. For example, it is applicable to the correspondence to the case where all photographic subjects are not reflected in parallel photographing, or the correspondence to free hand photographing. Then, we devise the technique of computing the combination of the sub-optimal path points like parallel photographing in convergent photographing.

In the previous research, we calculated the 3-dimensional coordinate data of corresponding points by the method of connecting center of projection using the coplanarity condition. In this technique, since the parallel positions of two cameras are in the initial state of convergent calculation, the initial values of κ_1, φ_1, ω_2, κ_2, and φ_2 are set to 0. Therefore, compared with the parallel photography near an initial state, the convergent photographing spends a long time for the convergence calculation of them. And the possibility of lapsing into a limited part solution becomes high. This reason is that the camera leans greatly from the initial state. Therefore, the accuracy of the convergent photographing was low compared with parallel photographing. Then, we consider that this problem is solvable by extending the search domain more than the previous research. In evaluation of GA, the evaluation value is calculated in the case where an index coordinate system is both a right-hand coordinate system and a left-hand coordinate system. And we utilize the one of the higher evaluation.

3.3 Proof experiment of convergent photographing

In this experiment, we use a photograph (Figure 2) that is 7.5m away from the starting point and is taken with about 55-degree convergent photographing. These photographs have been used in the previous research.

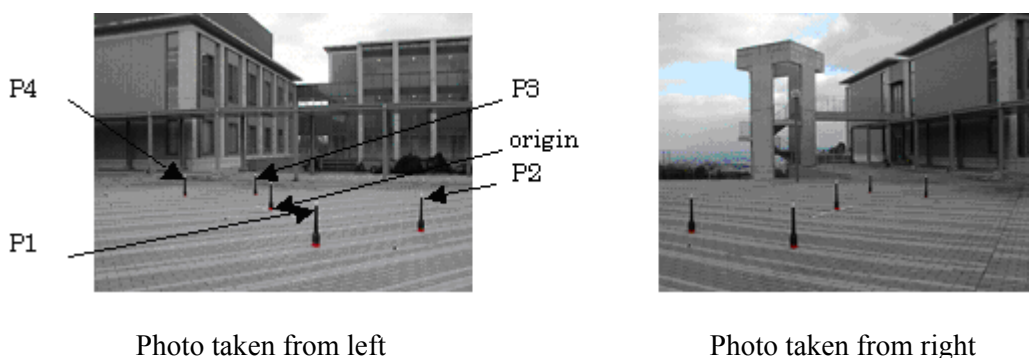Photo taken from left   Photo taken from right

Figure.2 Convergent Photographing

We did not perform the correct of lens, and fixed the height and inclination of a camera by using the tripod. The camera is about 2 million pixels. The camera is the focal distance as 6.5mm, the pixel size as 0.0039mm and the number of pixels as 1600*1200. And we use the same datum point and path point as the previous research. We evaluate measurement accuracy by using four points of datum points for inspection other than path points. A result in this experiment is shown in Table 1.

Table.1 A Result in This Experiment

Point Name		Calculated Value	Measured Value	Error (m)	Error Rate (%)
P1	x	0.90592	0.91	0.00408	0.44839
	y	2.83820	2.85	0.01180	0.41396
	z	0.58621	0.60	0.01379	2.29906
P2	x	-0.90336	-0.92	-0.01664	1.80901
	y	2.93529	2.88	-0.05529	-1.91983
	z	0.58296	0.60	0.01704	2.83939
P3	x	-1.04641	-1.05	-0.00359	0.34175
	y	-2.93448	-2.90	0.03448	-1.18884
	z	0.59534	0.60	0.00466	0.77667
P4	x	0.73571	0.72	-0.01571	-2.18245
	y	-3.52322	-3.55	-0.02678	0.75428
	z	0.58799	0.60	0.01201	2.00114

With convergent photographing, 3-dimensional coordinate values could be calculated with an error within about 2% by Photogrammetry function by using GA. Therefore, we confirm that this system is an effective technique in the various photographing methods.

4   Semi-automatic acquisition function of corresponding points

4.1   Flow of semi-automatic acquisition of corresponding points

Many residences are constituted by plane combination. This function acquires corresponding points automatically to all planes. First selects two points arbitrary from corresponding points on the same plane. Second calculates the 3-dimensional coordinate of corresponding points from photographic coordinate with an arbitrary interval on a picture until it reaches the boundary lines on plane. Finally, processing of first and second is repeated with the arbitrary numbers of times. Flow of semi-automatic acquisition of corresponding points is shown in Figure 3.
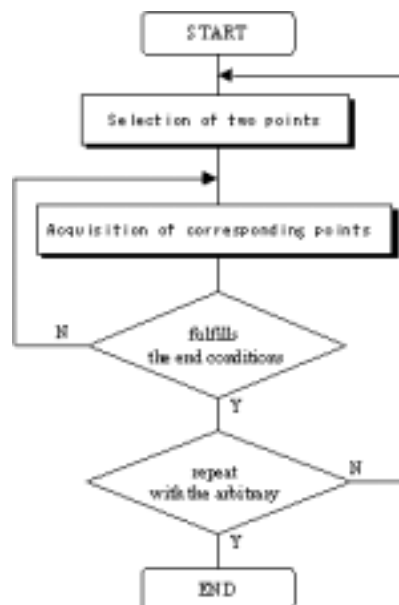

Figure.3 Flow of semi-automatic acquisition of corresponding points

4.2　Selection of two points

In selection of two points, the datum points for automatic acquisition of corresponding points are selected. In the first time, two points are selected from the datum points. From the second time, two points are arbitrarily selected from corresponding points acquired. However, if the combination of two points has been already selected, two points is selected once again.

4.3　Acquisition of corresponding points

In Photogrammetry function by using GA, 3-dimensional coordinate is calculated from corresponding points. Firstly, 1) the equation of straight line that passes along the two selected points is calculated from the 3-dimensional coordinate. Secondly, 2) we calculate 3-dimensional coordinate that is moved with arbitrary values from a point on the computed straight line. And it is transformed by reverse the affine transformation and the reverse transformation. We acquire the calculated coordinate as a corresponding point. The reverse affine transformation is to transform pixel coordinate system from photograph coordinate system. Moreover, the reverse transformation is to transform a 3-dimensional coordinate system into a photograph coordinate system. The equation used for the reverse affine transformation is shown in the following equation (1).

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix}^{-1} \left\{ \begin{pmatrix} x' \\ y' \end{pmatrix} - \begin{pmatrix} c \\ f \end{pmatrix} \right\}$$

(1)

a, b, c, d, e, f : Affine coefficient

(x, y) : Pixel coordinate

(x', y') : Photographic coordinate

The equation used for the reverse coordinate transformation of a photo taken from left is shown in the following equation (2).

$$A = \begin{pmatrix} \cos\varphi_1 & 0 & \sin\varphi_1 \\ 0 & 1 & 0 \\ -\sin\varphi_1 & 0 & \cos\varphi_1 \end{pmatrix} \qquad B = \begin{pmatrix} \cos\varkappa_1 & -\sin\varkappa_1 & 0 \\ \sin\varkappa_1 & \cos\varkappa_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ y_1 \\ -c \end{pmatrix} = B^{-1} * A^{-1} \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} Bx_1 \\ By_1 \\ Bz_1 \end{pmatrix} \right\}$$

(2)

c : Focal distance

(x_1, y_1) : Photographic coordinate

(X, Y, Z) : 3-dimensional coordinate

(Bx_1, By_1, Bz_1) : Coordinate of lens center

The equation used for the reverse coordinate transformation of a photo taken from right is shown in the following equation (3).

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_2 & -\sin\omega_2 \\ 0 & \sin\omega_2 & \cos\omega_2 \end{pmatrix} \qquad B = \begin{pmatrix} \cos\varphi_2 & 0 & \sin\varphi_2 \\ 0 & 1 & 0 \\ -\sin\varphi_2 & 0 & \cos\varphi_2 \end{pmatrix}$$

$$C = \begin{pmatrix} \cos\kappa_2 & -\sin\kappa_2 & 0 \\ \sin\kappa_2 & \cos\kappa_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} x_2 \\ y_2 \\ -c \end{pmatrix} = C^{-1} * B^{-1} * A^{-1} \left\{ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} Bx_2 \\ By_2 \\ Bz_2 \end{pmatrix} \right\} \qquad (3)$$

c : Focal distance

(x_2, y_2) : Photographic coordinate

(X, Y, Z) : 3-dimensional coordinate

(Bx_2, By_2, Bz_2) : Coordinate of lens center

Finally, 3) the processing of 2) is repeated until it fulfills end condition. The end condition is that the position, which acquires a corresponding point, arrives at the boundary lines on plane in the direction of the starting point and the terminal point. However, if the straight line equation calculated in processing of 1) has been already calculated once, acquisition of a corresponding point is not performed. Moreover, we perform the edge detection by using the Laplace operator in order to extract a vivid boundary lines on plane. Processing of from 1) to 3) is repeated. The arbitrary values and the repeat times are set properly according to a photographic subject. Acquisition of corresponding points is shown in Figure 4.
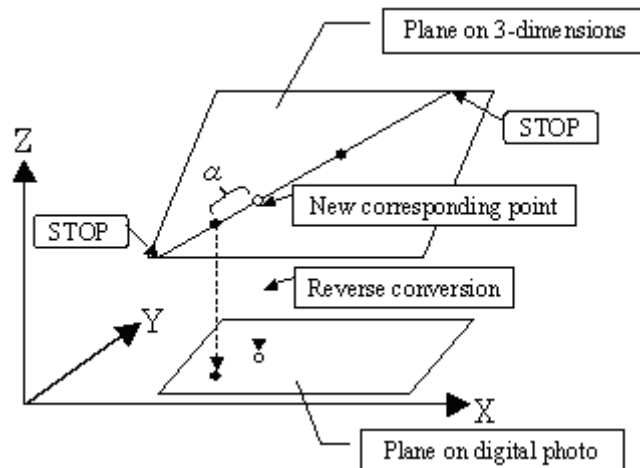

Figure.4 Acquisition of corresponding points

4.4 Inspection of function

As inspection of function, we acquired corresponding points automatically on conditions; free hand photographing with the 3 million pixels camera generally marketed, the focal distance (7.1mm), the pixel size (0.00345mm), and the number of pixel (2048*1536), and the glass board of the shape of 38mm/mass. We calculated the parameter of a camera by acquiring the datum points on the glass board. As shown in Figure.5, early corresponding points were set as the

mass of three corners on a glass board. The parameter value was set as $38\sqrt{2}$ mm in the oblique direction and as 38mm in the vertical and horizontal direction. We acquired five new corresponding points between early corresponding points. And we acquired fifteen points.
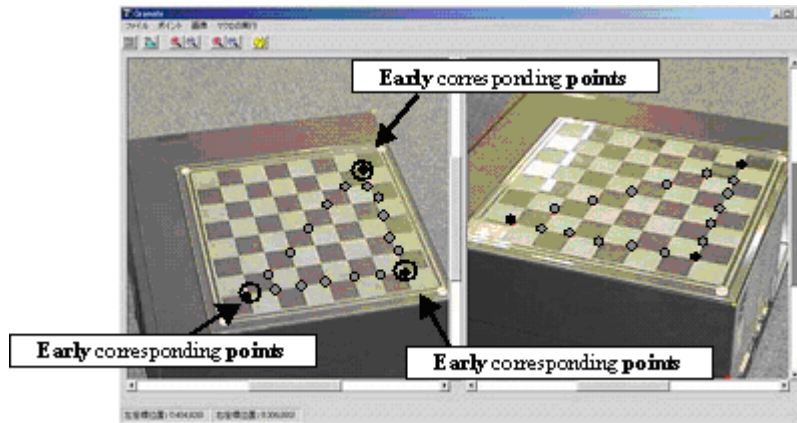

Figure.5 Result of Automatic acquisition of corresponding points

We calculated the sum of errors between x coordinates and y coordinates of fifteen corresponding points and the intersection of the mass on a glass board. And we calculated the average error, the maximum error, and the minimum error. The result is shown in Table 2. The average error was 1.8 pixels. And the maximum error was also 4 pixels. Therefore, we confirmed that this function had sufficient accuracy.

Table.2 The error table of corresponding points

Total of Points	Average Error (pixel)	Maximum Error	Minimum Error
18	1.8	4.0	0.0

5   3-dimensional measurement function

This function calculates the 3-dimensional coordinate data of corresponding points, like Photogrammetry function by using GA, by the method of connecting center of projection using coplanarity condition. Firstly, the affine transformation of the arbitrary coordinate system used for acquisition of pixel coordinates is transformed into photographic coordinate system. The equation of the affine transformation used in this function is shown in the following equation (4).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix}$$

(4)

a, b, c, d, e, f : Affine coefficient
(x, y): Pixel coordinate
(x', y'): Photographic coordinate

Secondly, we calculate 3-dimensional coordinates data by transforming model coordinate system from photographic coordinates transformed by the affine transformation. The equation used for transformation of the coordinates into the model coordinate system of a photo taken from left is shown in the following equation (5). And the equation used in a photo taken from right, is shown in the following equation (6).

$$\begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} = A * B \begin{pmatrix} x_1 \\ y_1 \\ -c \end{pmatrix} + \begin{pmatrix} Bx_1 \\ By_1 \\ Bz_1 \end{pmatrix} \tag{5}$$

c : Focal distance
(X_1, Y_1, Z_1) : Model coordinate
(x_1, y_1) : Photographic coordinate
(Bx_1, By_1, Bz_1) : Coordinate of lens center

$$\begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = A * B * C \begin{pmatrix} x_2 \\ y_2 \\ -c \end{pmatrix} + \begin{pmatrix} Bx_2 \\ By_2 \\ Bz_2 \end{pmatrix} \tag{6}$$

c : Focal distance
(X_2, Y_2, Z_2) : Model coordinate
(x_2, y_2) : Photographic coordinate
(Bx_2, By_2, Bz_2) : Coordinate of lens center

6   Automatic creation function of VRML model

6.1   Automatic acquisition function of plane information

In order to perform 3-dimensional modeling from 3-dimensional coordinate calculated by 3-dimensional measurement function, the plane information of an object is needed. In the previous research, we developed the interface that anyone is available to acquire the plane information easily. However, in the case of the big photographic subject or the complicated object that had free curved surfaces, there was the subject that needed the remarkable time and labor. Therefore, in this function, we acquire the information of planes automatically by the Delaunay triangulation. If there is the dented plane on an object, decision of the boundary for creation of planes is very difficult in the 3-dimensional Delaunay triangulation. Therefore, we perform the 2-dimensional Delaunay triangulation by transforming 3-dimensional information into 2-dimensional information.

Firstly, three arbitrary points on the same plane are selected. By setting the one of three points as datum point, the related information between it and other two points is calculable. There are two kinds of this related information; 1) Distance between the 1st point and the 2nd point, 2) Distance between the 3rd point and the straight line that passes along these two points. These points are used for calculation of the parameter of coordinate transformation into X-Y plane by absolute orientation. This related information is to acquire other two points by setting the 1st point as the origin. This acquisition method is shown in Figure 6. Secondly, 3-dimensional coordinates of corresponding points are transformed into the X-Y plane by absolute orientation.

Finally, we perform the Delaunay triangulation to the transformed points, and generate the triangular polygons. We acquire the information of each plane automatically by acquiring the information of points that constitutes this triangular polygon.
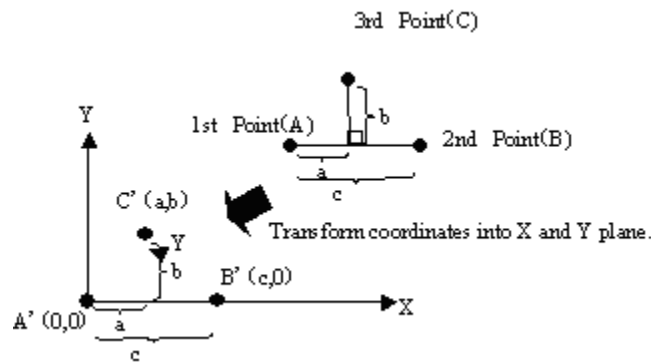

Figure.6 The acquisition method of points by the datum point

6.2   Creation function of VRML model

One of the uses of the 3-dimensional coordinate data calculated by 3-dimensional measurement function is the visualization by CG. We use VRML (ISO/IEC 14772-1) as the technique of visualization. In this function, by using the 3-dimensional coordinates data calculated by 3-dimensional measurement function and the plane information by Automatic acquisition function of plane information, we output the data of VRML as a text automatically. And the file of a VRML model is automatically created from the text data.

6.3   Automatic attachment function of texture

In this function, we stick automatically the textures on the triangular polygons created by Automatic acquisition function of plane information. Transforming texture coordinate system from photographic coordinate system is needed for the above-mentioned function. Then we select one of photos taken from right and left. And we transform each point (x, y) into texture coordinate (X, Y) according to the scale of a photo by using information of constitution points of triangular polygons. An equation of relations to transform each photographic coordinate into texture coordinate is the following equation (7).

$$X = x / width$$
$$Y = 1 - y / Height$$

(7)

Witdh : A pixel number of photo in horizontal direction

Height : A pixel number of photo in vertical l direction

6.4   Inspection of function

Firstly, we create triangular polygons from a group of correspondence points calculated. And we create VRML data by a VRML model creation function automatically Secondly, we stick the textures on VRML data in automatically. The result that pixel coordinate of correspondence points was transformed into texture coordinate by using an equation (7) is shown in Table.3. VRML model stuck textures on automatically is shown in Figure 7. We were able to confirm the effectiveness of this function by this inspection result.
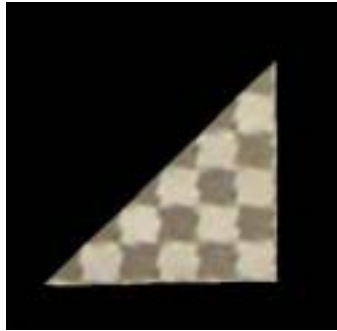
Figure.7 An execution result (texture)

Table.3 Pixel coordinate and texture coordinate of corresponding points

Number of Point	Pixel Coordinate		Texture Coordinate	
1	1301	874	0.635	0.431
2	992	920	0.484	0.401
3	1216	670	0.594	0.564
4	1251	881	0.611	0.426
5	1199	889	0.585	0.421
6	1147	896	0.560	0.417
7	1097	904	0.536	0.411
8	1045	912	0.510	0.406
9	1285	837	0.627	0.455
10	1271	801	0.621	0.479
11	1257	767	0.614	0.501
12	1243	733	0.607	0.523
13	1229	701	0.600	0.544
14	1181	708	0.577	0.539
15	1148	746	0.561	0.514
16	1110	787	0.542	0.488
17	1073	830	0.524	0.460
18	1033	873	0.504	0.432

7 The function of 3D model house on Web

7.1 Outline of function

Because of VRML, this research makes it possible that everyone serves 3D model house on Web. The function of 3D model house on Web is constituted of the following two functions.

1) Correlation function of VRML file realizes that the correlative 3D model house is indicated on Web by selecting a room in a sketch. The VRML files are correlated with a whole room in a sketch by simple mouse operation. And, for realization of this function, we unite beforehand the VRML files a whole room in sketch.

2) Automatic creation function of HTML codes realizes automatic creation of necessary HTML codes for the indication on browser. The necessary codes for the layout on browser and the integration of VRML data united by the Correlation function of VRML file are created automatically.

8 A proof experiment

We created the 3D model of a classroom in Kansai University as a proof experiment of this system on same conditions of inspection of Semi-automatic acquisition function of corresponding points. And we set the interval of acquisition of corresponding points as 300mm.

As an initial value, we selected the three points of corresponding points on the six planes by manual operation. The six planes used for acquisition of corresponding points and an input photo taken from left is shown in Figure 8.
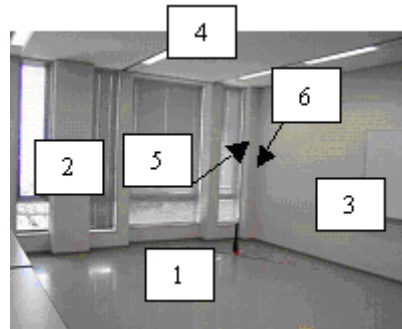

Figure.8 The input photo taken from left and the six planes

The confusion of an edge by the light to take the distance between the three points acquired by manual operation broadly can be solved because an edge between two points selected is not an end condition. The number of corresponding points acquired every a plane in this experiment is shown in Table 4. However, the points that distance with a plane exceeded 300mm were eliminated with the system side automatically. 3D model created in this experiment is shown in Figure 9, 10. As shown in Figure.10, we could create automatically 3D model that unevenness of the plane No.5 and No.6 in Figure 8 was expressed precisely. It is clear that a result has high precision by the precision of the datum points in Table.5 and the indoor height in Table 6.
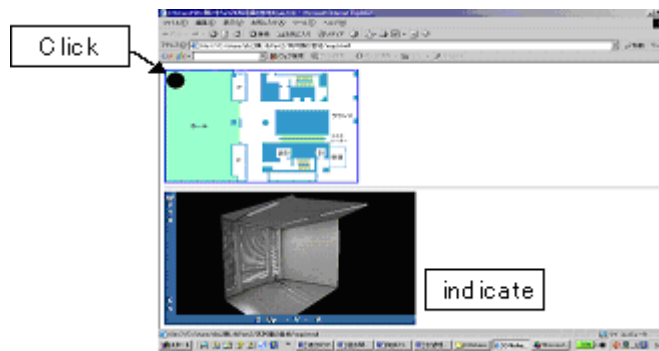

Figure.9 An execution result

Figure.10 3D model (the upper view)

table.4 The number of corresponding points acquired every planes

Number of Plane	Number of Corresponding Points Acquired
1	50
2	28
3	35
4	43
5	12
6	10

Table.5 The precision of datum points

Number of Point	Calculated Value			Measured Value		
	x	y	z	x	y	z
1	0	0	0	5.97	36.24	37.16
2	1000	0	0	1033.23	1.10	-1.55
3	0	1000	0	1.10	961.18	-1.35
4	0	0	600	-40.31	1.47	585.73

Table.6 The precision inspection of indoor height

Measured Value	Average Calculated Value (mm)	Error (mm)	Error Rate (%)
2,530	2,661	131.0	5.178

9   Conclusions

We realized the semi-automatic acquisition of the corresponding points by the five functions. We also realized the high accuracy in convergent photographing. And we established the technique that everyone is available to create the high accurate 3D model easily. However, in this research, we inputted the points on same plane as an initial value manually. Therefore, we tried to realize automatic acquisition of the points on same plane. Because we limited the photographic subject to a room without furniture, we realized the automatic acquisition of corresponding points in a free curved surface.

10   References

Muraki, H. Tanaka, S. Furuta, H and Kitagawa, E. A Study for Producing 3D Model Space from 2D Digital Picture using Photogrammetry Technology. (2000). Journal of Civil Engineering Information Processing System. Japan Society of Civil Engineers. Vol.9,pp.67-74

Japanese Society for Photogrammetry in Japan. (1997). Analytical Photogrammetry (Revised edition).

Panton, D.J. (1975). A Flexible Approach to Digital Stereo Mapping. Photogrammetric Engineering and Remote Sensing. Vol.44, No.12, pp.1499-1512.

Remote Sensing Society of Japan. (1986). Processing and Analysis of Images. Kyoritsu Shuppan Co. Ltd.

Takagi, M. and Shimoda, H. (1998). Handbook of Image Analysis. Publishing Circle. Tokyo University.