Resolution of a Texture Image Mapped on a Building Surface of a 3D City Model

Masato Ishikawa [1], Tomoaki Inazawa [1], Yoshihiko Nakanishi [1], Futa Kawamata [1], Masahito Takada [1], Takuya Danjo [1], Ryuji Matsuoka [1]

[1] Kokusai Kogyo Co., Ltd., 2-24-1 Harumi-cho, Fuchu-shi, Tokyo 183-0057, Japan - {masato_ishikawa, tomoaki_inazawa, yoshihiko_nakanishi, futa_kawamata, masahito_takada, takuya_danjo, ryuji_matsuoka}@kk-grp.jp

Keywords: 3D City Model, Texture image, Vertical surface, Aerial nadir camera, Aerial oblique camera.

Abstract

It is desirable that a 3D city model at level of detail 2 (LOD2) has texture images mapped on building surfaces. Texture images mapped on building surfaces are utilized not only for providing a realistic appearance but also for enabling some analyses, such as investigation of stories of a building and detection of doors and windows on a building surface. We investigated on the resolution of texture images produced from aerial nadir and oblique images. First, we set up the required resolutions for the investigation of stories and the detection of doors and windows by using simulated images of various resolutions. Next, we obtained the general formula to calculate the resolution of a texture image mapped on a vertical surface. Then we conducted a numerical analysis on a texture image mapped on a vertical surface produced from a nadir image and an oblique image collected by aerial cameras popularly utilized in Japan. The results of the analysis indicate that a texture image mapped on a vertical surface produced from a nadir image collected in the usual procedure of aerial photogrammetry would be unable to satisfy the required resolution for the detection of doors and windows. On the contrary, the results indicate that a texture image mapped on a vertical surface produced from an oblique image can be utilized for not only investigation of stories but also detection of doors and windows. However, there remain some problems to solve in the utilization of an oblique image for texture mapping.

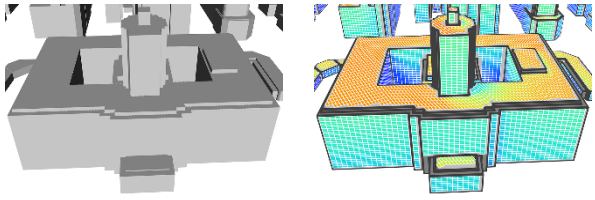
1. Introduction

3D city model data in more than 250 Japanese cities is currently provided by Project PLATEAU, which is a project started by the Ministry of Land, Infrastructure, Transport and Tourism of Japan (MLIT) in 2020. The project aims to construct 3D city models as a platform for urban activities, provide 3D city model data as open access data, and promote the utilization of 3D city models (MLIT, 2025a).

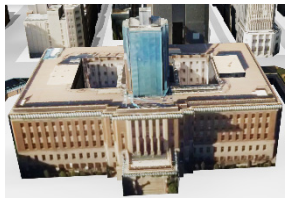
Project PLATEAU recommends that a 3D city model at level of detail 2 (LOD2) should have texture images mapped on building surfaces to make the appearance of buildings of the 3D city model realistic (MLIT, 2025b, MLIT, 2025c). At present, the production of a 3D city model with texture in Japan utilizes mostly aerial nadir images, which provide the geometry of a 3D city model from the aspect of the production cost of a 3D city model.

The photorealistic 3D city models are utilized in various geospatial applications related to virtual city tourism, 3D GIS, urban planning, and real estate management. However, texture images mapped on building surfaces of a 3D city model are utilized not only for providing a realistic appearance but also for enabling some analyses, such as investigating stories of a building and detecting doors and windows on a building surface. Lippoldt proposed a method to detect windows on building facades by using aerial texture images of a 3D city model. Although Lippoldt did not mention utilization of detection results in the paper, we expect that analysis results of texture images attached to a 3D city model would be used as follows:

The photovoltaic (PV) potential estimation system developed by Matsuoka et al. (Matsuoka et al., 2024) utilizes 3D city model data at LOD2. The system handles an arrangement of solar panels on a building surface to estimate PV potential flexibly. Although the system currently does not utilize texture images, the system will be able to estimate more realistic PV potential if the system uses some information derived from texture images, as Figure 1 shows. For instance, it should be possible to establish the condition that no solar panels should be placed on a façade of the ground floor of a building and the condition that any solar panels should be placed on neither a door nor a window.



Current PV potential estimation considering the arrangement of solar panels using a 3D city model without texture



Information on doors and windows derived from a texture image would make PV potential estimation more realistic.

Future PV potential estimation using a 3D city model

Figure 1. Possible use of a texture image in PV potential estimation using a 3D city model

Although there are some studies on producing a texture image of a 3D city model (Buyukdemircioglu, Oude Elberink, 2024; He et al., 2022), there are no studies on the evaluation of a texture image of a 3D city model. Therefore, we decided to investigate on the evaluation of a texture image. Although high-quality texture images can be obtained by using terrestrial images and UAV images, we decided to investigate the quality of a texture

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W6-2025-89-2025

image produced from an aerial image from the aspect of production cost to produce a 3D city model with texture.

A texture image can be evaluated by image qualities on grayscale information such as gradation, contrast, and noise, and image qualities on spatial information such as resolution and sharpness. An image quality, except for the resolution of a texture image, is much influenced by that of a source aerial image. Hence, the investigation was focused on the resolution of a texture image mapped on a building surface. Some image processing techniques at texture mapping, such as cubic convolution using a single image (Meijering, 2002) and super resolution using multiple images (Yang, Huang, 2017), provide higher resolution of a texture image, clearly suitable to make the appearance of buildings of the 3D city model realistic. Since the effect of such image processing techniques on some analyses by a computer using texture images of a 3D city model, such as investigation of stories of a building and detection of doors and windows on a building surface, is unknown, the study took no account of the utilization of such image processing techniques.

The required resolution of a texture image depends on the purpose for using it. A texture image produced from an oblique image would satisfy the resolution required for some purposes that a texture image produced from a nadir image cannot satisfy. Accordingly, we investigated texture images produced from aerial nadir images, which are popularly utilized in the production of a 3D city model in Japan, and aerial oblique images, which are not popularly utilized in the production of a 3D city model in Japan.

A texture image mapped on a horizontal surface of a building produced from a nadir image, which is utilized to provide the geometry of a 3D city model, would be able to satisfy requirements for most purposes. Hence, the paper discusses texture images mapped on vertical surfaces of buildings of a 3D city model.

The paper focuses on the analytical and numerical analysis of the resolution of texture images mapped on vertical surfaces of buildings of a 3D city model produced in Japan. The aims of the paper are to present the general formula to calculate the resolution of a texture image and to indicate the comparison results between using nadir images and oblique images for texture mapping.

In the paper, first we discuss the required resolutions for the investigation of stories and the detection of doors and windows by using simulated images of various resolutions. Next we show a process to derive the general formula to calculate the resolution of a texture image mapped on a vertical surface. Then we present a numerical analysis of a texture image mapped on a vertical surface by using the obtained general formula. The analysis deals with a texture image produced from a nadir image and an oblique image collected by aerial cameras popularly utilized in Japan. Finally, we summarize the conclusion of the study.

2. Related works

P. Meixner and F. Leberl compared vertical and oblique aerial images from the point of view of their usefulness in the analysis, such as counting floors and counting windows of a building using façade images (Meixner, Leberl, 2010a, 2010b). The resolution on a façade, which they called the façade sampling distance (FSD), was utilized as one of the means of the comparison in the comparison.

As for a nadir image, they presented the following formula to calculate FSD:

$$\mathrm{FSD} = \frac{\mathrm{GSD}}{\tan(\alpha)} \tag{1}$$

where GSD is the ground sampling distance and α is the look angle off-nadir. GSD is calculated by the following equation:

$$\mathrm{GSD} = \frac{ph}{f} \tag{2}$$

where p is the pixel size in the image plane, h is the flying height above ground, and f is the focal length of the camera.

As for an oblique image, they showed that FSD is defined by two angles, α and β. α is the angle between the optical axis and the actual imaging ray, and β is the orientation of the optical axis off-nadir. The off-nadir angle β produces different GSD values in two perpendicular directions. They named the GSD in the direction of the inclination of the optical axis GSDr, with r being the range direction or direction between nadir and the optical axis. They named the GSD in the perpendicular direction, which was called azimuth direction, GSDa as well. They presented the following formulas to calculate GSDr and GSDa:

$$\mathrm{GSDr} = \frac{ph\cos(\alpha)}{f\cos^2(\beta)} \tag{3}$$

$$\mathrm{GSDa} = \frac{ph\cos(\alpha)}{f\cos(\beta)} \tag{4}$$

They presented the following formula to calculate the sampling distance on a vertical façade, which was denoted FSDv:

$$\mathrm{FSDv} = \frac{ph\cos(\alpha)}{f/\sin(\beta)} \tag{5}$$

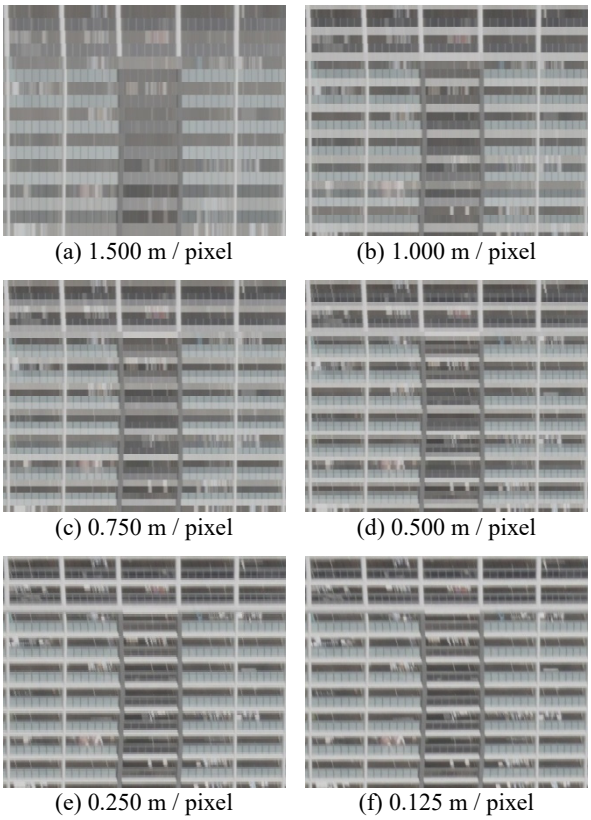
We judged that Equation (1) to calculate FSD of a nadir image is right. On the contrary, we judged that Equations (3), (4), and (5) as to an oblique image are wrong. The reason for our judgment is that substituting 0 for β, which corresponds to a nadir image, yields results that are inconsistent with GSD defined in Equation (2) and FSD outlined in Equation (1) for a nadir image. Since their papers did not show the process to derive the formula to calculate GSDr, GSDa, and FSD, we were unable to find the error source. Accordingly, we decided to derive the general formula to calculate FSDv.

3. Required resolutions of an image mapped on a vertical surface of a 3D city model of LOD2

The resolution required for texture images depends on the purpose of utilizing texture images. It is difficult to specify the required resolution for providing a realistic appearance of buildings in a 3D city model numerically. Hence, we decided to discuss the required resolution for the investigation of stories of a building and the detection of doors and windows on a building surface.

We produced simulated texture images of various resolutions of various buildings from aerial images to specify a required resolution. Figure 2 shows some of the simulated images.

The essential resolutions in the column direction of the simulated images produced from the same aerial image, that is to say, the horizontal direction in the object space, are kept identical, while those in the row direction of the simulated images, that is to say, the vertical direction in the object space, vary with the projection resolution, as shown in Figure 2.



(a) 1.500 m / pixel (b) 1.000 m / pixel

(c) 0.750 m / pixel (d) 0.500 m / pixel

(e) 0.250 m / pixel (f) 0.125 m / pixel

Figure 2. Simulated texture images of a vertical surface

We specified the required resolutions by human inspection of the simulated images, considering the cost of image acquisition. Table 1 shows the specified required resolutions for the investigation of stories of a building and the detection of doors and windows on a building surface, respectively.

Purpose	Permissible	Preferable
Investigation of stories	1.00 m / pixel	0.75 m / pixel
Detection of doors / windows	0.50 m / pixel	0.25 m / pixel

Table 1. Required resolutions for the investigation of stories and the detection of doors and windows

4. General formula to calculate the resolution of the image mapped on a vertical surface

4.1 Assumptions to obtain a general formula

We made the following assumptions to obtain a general formula to calculate the resolution of the image mapped on a vertical surface.

The image coordinate system is considered a planar rectangular coordinate system (x, y), while the object space coordinate system is regarded as a spatial rectangular coordinate system (X, Y, Z), as illustrated in Figure 3.

Consider that an aerial image to be utilized to be mapped on a surface of a 3D city model is taken by a camera with focal length f, field of view θ, and pixel size r. Let $\mathrm{O}(X_0, Y_0, Z_0) = (0,0,H)$ and $\mathrm{R}(\omega, \phi, \kappa)=(0, \phi_0, 0)$ be the position and the attitude of the camera in the object space coordinate system (X, Y, Z), respectively. Here $0 < X_\mathrm{P}$, $0 \le Z_\mathrm{P} < H$, and $0 \le \phi_0 < 90°$. On the assumption we are able to consider the vertical resolution of an image mapped on a vertical surface by using the 2D coordinate system (X, Z) and the 1D coordinate system (x) as Figure 4 shows.

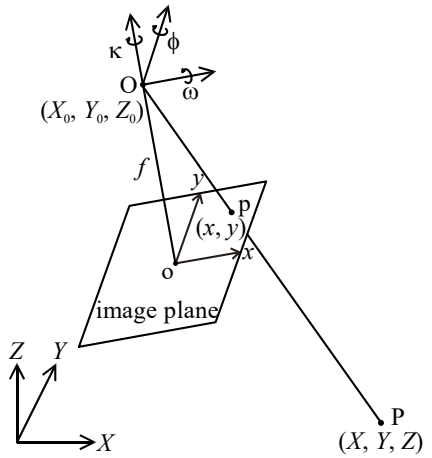


Figure 3. Image coordinate system (x, y) and object space coordinate system (X, Y, Z)

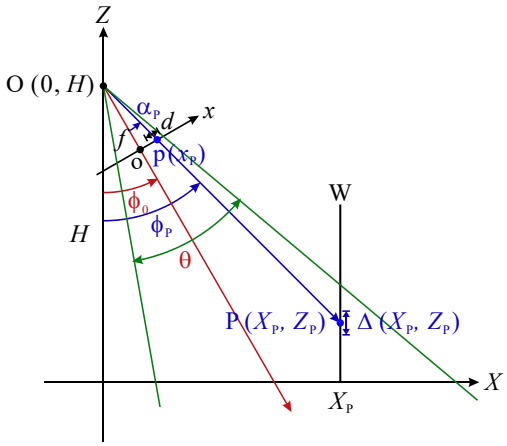


Figure 4. 2D coordinate system (X, Z) and 1D coordinate system (x) in the study

4.2 Formula from an image point $\mathrm{p}(x_\mathrm{P})$ on an aerial image to a point $\mathrm{P}(X_\mathrm{P}, Z_\mathrm{P})$ on a vertical surface

Now let a point $\mathrm{P}(X_\mathrm{P}, Z_\mathrm{P})$ be on the vertical surface $\mathrm{W}: X = X_\mathrm{P}$, and let the point $\mathrm{P}(X_\mathrm{P}, Z_\mathrm{P})$ be projected to the image point $\mathrm{p}(x_\mathrm{P})$ on an aerial image. Here the existing region of $\mathrm{P}(X_\mathrm{P}, Z_\mathrm{P})$ is as follows:

$$\tan\left(\phi_0 - \frac{\theta}{2}\right) \le \frac{X_\mathrm{P}}{H - Z_\mathrm{P}} \le \tan\left(\phi_0 + \frac{\theta}{2}\right) \tag{6}$$

A view angle α_P of the image point $\mathrm{p}(x_\mathrm{P})$ measured from the optical axis of the camera is expressed by the following equation:

$$\tan \alpha_\mathrm{P} = \frac{x_\mathrm{P}}{f} \tag{7}$$

Similarly, a view angle ϕ_P of the point $\mathrm{P}(X_\mathrm{P}, Z_\mathrm{P})$ measured from the nadir is expressed by the following equation:

$$\tan \phi_\mathrm{P} = \frac{X_\mathrm{P}}{H - Z_\mathrm{P}} \tag{8}$$

Equation (8) gives the following equation:

$$Z_\mathrm{P} = H - \frac{X_\mathrm{P}}{\tan \phi_\mathrm{P}} \tag{9}$$

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W6-2025-89-2025 | © Author(s) 2025. CC BY 4.0 License.

Here the viewing angle ϕ_P of the point $P(X_P, Z_P)$ is represented by the following equation:

$$\phi_P = \alpha_P + \phi_0 \tag{10}$$

Equation (10) brings the following equation:

$$\frac{1}{\tan\phi_P} = \frac{1}{\tan(\alpha_P+\phi_0)} = \frac{1-\tan\alpha_P \tan\phi_0}{\tan\alpha_P+\tan\phi_0} \tag{11}$$

Substituting Equation (8) for Equation (11) and rearranging gives the following equation:

$$\frac{1}{\tan\phi_P} = \frac{1-\frac{x_P}{f}\tan\phi_0}{\frac{x_P}{f}+\tan\phi_0} = \frac{-x_P\tan\phi_0+f}{x_P+f\tan\phi} \tag{12}$$

Substituting Equation (12) for Equation (9) and rearranging gives the following equation:

$$Z_P = H - X_P\frac{-x_P\tan\phi_0+f}{x_P+f\tan\phi_0} = H + X_P\frac{x_P\tan\phi_0-f}{x_P+f\tan\phi_0} \tag{13}$$

Equation (13) expresses that the image coordinate x_P of the image point p on an aerial image gives the object space coordinate Z_P of the point P on the vertical surface W.

4.3 Formula to calculate the resolution $\Delta(X_P, Z_P)$ of the image mapped on a vertical surface

Now we are going to obtain a formula to express the resolution of an image mapped on a vertical surface by using Equation (13) and the following Equation (14):

$$dZ_P = \frac{\partial Z_P}{\partial x_P}dx_P \tag{14}$$

where the resolution $\Delta(X_P, Z_P)$ of the mapped image at the point $P(X_P, Z_P)$ and the pixel size r of the aerial image are given by the following equations, respectively:

$$\Delta(X_P, Z_P) = dZ_P \tag{15}$$
$$r = dx_P \tag{16}$$

Substituting Equations (15) and (16) for Equation (14) and rearranging gives the following equation:

$$\Delta(X_P, Z_P) = r\frac{\partial Z_P}{\partial x_P} \tag{17}$$

Here, Equation (13) gives the partial differential $\frac{\partial Z_P}{\partial x_P}$ as follows:

$$\frac{\partial Z_P}{\partial x_P} = X_P\frac{\partial}{\partial x_P}\left(\frac{x_P\tan\theta_0-f}{x_P+f\tan\phi_0}\right) = X_P\frac{f}{(x_P\cos\phi_0+f\sin\phi_0)^2} \tag{18}$$

Substituting Equation (18) for Equation (17) and rearranging gives the following equation:

$$\Delta(X_P, Z_P) = X_P\left(\frac{r}{f}\right)\frac{1}{\left\{\left(\frac{x_P}{f}\right)\cos\phi_0+\sin\phi_0\right\}^2} \tag{19}$$

Equations (7), (8), and (10) give the following equation:

$$\frac{x_P}{f} = \tan(\theta_P-\phi_0) = \frac{X_P-(H-Z_P)\tan\phi_0}{X_P\tan\phi_0+(H-Z_P)} \tag{20}$$

Finally, substituting Equation (20) for Equation (19) and rearranging gives the following equation:

$$\Delta(X_P, Z_P) = \left(\frac{r}{f}\right)\frac{\{X_P\sin\phi_0+(H-Z_P)\cos\phi_0\}^2}{X_P} \tag{21}$$

Equation (21) is a general formula to calculate the resolution $\Delta(X_P, Z_P)$ of the image mapped on a vertical surface.

4.4 Maximum and minimum values of the image resolution $\Delta(X_P, Z_P)$ on a vertical surface

To obtain the maximum and minimum values of the resolution $\Delta(X_P, Z_P)$ the partial differentials $\frac{\partial\Delta(X_P,Z_P)}{\partial X_P}$ and $\frac{\partial\Delta(X_P,Z_P)}{\partial Z_P}$ are going to be introduced. Equation (22) gives the partial differential $\frac{\partial\Delta(X_P,Z_P)}{\partial X_P}$ as follows:

$$\frac{\partial\Delta(X_P,Z_P)}{\partial X_P} = \left(\frac{r}{f}\right)\left\{\sin^2\phi_0 - \frac{(H-Z_P)^2}{X_P^2}\cos^2\phi_0\right\} \tag{22}$$

X_P expressed by the following equation brings $\frac{\partial\Delta(X_P,Z_P)}{\partial X_P} = 0$ and makes $\Delta(X_P, Z_P)$ the smallest:

$$X_P = \frac{H-Z_P}{\tan\phi_0} \tag{23}$$

On the other hand, Equation (21) gives the partial differential $\frac{\partial\Delta(X_P,Z_P)}{\partial Z_P}$ as follows:

$$\frac{\partial\Delta(X_P,Z_P)}{\partial Z_P} = -2\cos\theta_0\left(\frac{r}{f}\right)\left\{\sin\phi_0 + \frac{(H-Z_P)}{X_P}\cos\phi_0\right\} \tag{24}$$

The assumptions $0 < X_P$, $0 \le Z_P < H$, and $0 \le \phi_0 < 90°$ bring $\frac{\partial\Delta(X_P,Z_P)}{\partial Z_P} < 0$, and $Z_P = 0$ makes $\Delta(X_P, Z_P)$ the largest. That is to say, the image resolution $\Delta(X_P, Z_P)$ on a vertical surface is the lowest at the point P on the ground, where $Z = 0$, and becomes better as Z_P increases.

5. Numerical analysis on the resolution of a texture image produced from an aerial image

We conducted a numerical analysis on the resolution of a texture image mapped on a vertical surface from an aerial image collected by a nadir camera and an oblique camera. We examined whether the resolution of a texture image was able to satisfy the requirement for the investigation of stores and for the detection of a door and a window.

As for using aerial images for texture mapping, the lower flight height H brings the higher resolution $\Delta(X_P, Z_P)$ of an image mapped on a vertical surface of a building. To discuss practically in the analysis of a nadir camera, we assume that a texture image will be produced from a nadir image utilized to produce 3D city models. On the other hand, in the analysis of an oblique camera, we assume that a texture image will be produced from an oblique image collected simultaneously with the collection of a nadir image utilized to produce 3D city models.

Project PLATEAU requires that 3D city models of LOD2.2 and LOD3.x be produced by using aerial nadir images of 0.08 m GSD (ground sampling distance) (MLIT, 2025d). Here we assume that aerial nadir images of 0.08 m GSD will be used in the production of 3D city models.

5.1 Using images collected by a nadir aerial camera

As the off-nadir angle ϕ_0 of a nadir camera is 0, Equation (21) is expressed by the following equation:

$$\Delta(X_P, Z_P) = \left(\frac{r}{f}\right)\frac{(H-Z_P)^2}{X_P} \tag{25}$$

At the point $P(X_P, 0)$ on the ground, where $Z = 0$, Equation (25) becomes the following equation:

$$\Delta(X_P, 0) = \left(\frac{r}{f}\right)\frac{H^2}{X_P} \tag{26}$$

Rearranging Equation (26) brings the following formula:

$$X_P = \left(\frac{r}{f}\right)\frac{H^2}{\Delta(X_P,0)} \tag{27}$$

Here we assume that aerial nadir images of 0.08 m GSD collected by the Vexcel Imaging UltraCam Eagle Mark 3 (UCE M3) will be used in the production of 3D city models. The UCE M3 is one of the popular aerial cameras utilized to produce a 3D city model in Japan. The specifications of the assumed camera UCE M3 are shown in Table 2 (Vexcel Imaging, 2025a). The footprint of an image collected by the assumed camera is 2117 m wide (across track) by 1360 m long (along track).

	UCE M3
Image size	26 460 × 17 004 pixels
Physical pixel size r	4.0 μm
Lens system focal length f	100 mm
Field of view θ across track × along track	55.8°× 37.6°
Flying height H for 0.08 m GSD	2000 m

Table 2. Specifications of the assumed camera UCE M3 (Vexcel Imaging, 2025a)

Substituting the fields of view θ of across track and along track shown in Table 2 for Equation (6), the range of X_P, that is to say, the coverage by the assumed camera UCE M3, can be obtained as follows:

$$\text{[across track]}\ \tan\left(0° - \frac{55.8°}{2}\right) \le \frac{X_P}{2000} \le \tan\left(0° + \frac{55.8°}{2}\right) \tag{28}$$

$$\text{[along track]}\ \tan\left(0° - \frac{37.6°}{2}\right) \le \frac{X_P}{2000} \le \tan\left(0° + \frac{37.6°}{2}\right) \tag{29}$$

Rearranging Equations (28) and (29) yields the following range of X_P where $X_P > 0$:

$$\text{[across track]}\ 0 < X_P \le 1058 \text{ m} \tag{30}$$

$$\text{[along track]}\ 0 < X_P \le 680 \text{ m} \tag{31}$$

Substituting the specifications of the assumed camera shown in Table 2 for Equations (26) and (27) gives the following formulas:

$$\Delta(X_P, 0) = \left(\frac{0.004}{100}\right)\frac{2000^2}{X_P} = \frac{160}{X_P} \tag{32}$$

$$X_P = \left(\frac{0.004}{100}\right)\frac{2000^2}{\Delta(X_P,0)} = \frac{160}{\Delta(X_P,0)} \tag{33}$$

Substituting Equations (30) and (31) for Equation (32) brings the following range of $\Delta(X_P, 0)$:

$$\text{[across track]}\ \Delta(X_P, 0) \ge 0.151 \text{ m} \tag{34}$$

$$\text{[along track]}\ \Delta(X_P, 0) \ge 0.235 \text{ m} \tag{35}$$

Figure 5 shows the resolution $\Delta(X_P, 0)$ of an image mapped on a vertical surface from a nadir image collected by the assumed camera UCE M3. Table 3 shows the range of X_P satisfying the requirement of the resolution of an image mapped on a vertical surface as for the assumed camera UCE M3.

To obtain the distribution of $\Delta(X_P, Z_P)$ at $X_P = 200$ m, substituting the specifications of the assumed camera UCE M3, as detailed in Table 2, into Equation (25) yields the following formula:

$$\Delta(200, Z_P) = \left(\frac{0.004}{100}\right)\frac{(2000 - Z_P)^2}{200} \tag{36}$$

Figure 6 shows the resolution $\Delta(200, Z_P)$ of an image mapped on a vertical surface from a nadir image collected by the assumed camera UCE M3. Considering the heights of taller buildings in Japan, we show the resolution $\Delta(200, Z_P)$ up to $Z_P = 300$ m in Figure 6. The resolution $\Delta(200, Z_P)$ becomes better almost linearly as Z_P increases.

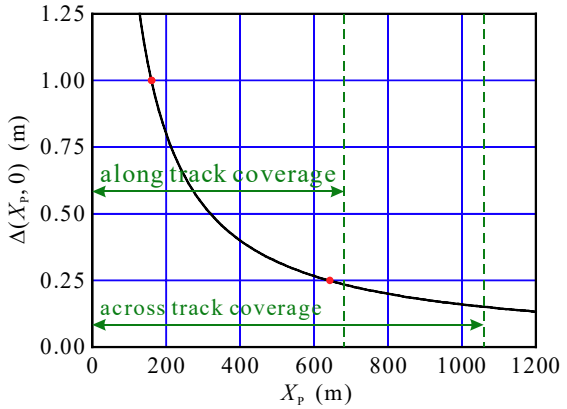


Figure 5. Resolution $\Delta(X_P, 0)$ of an image mapped on a vertical surface from a nadir image collected by the UCE M3

Resolution $\Delta(X_P, 0)$	Range of X_P
$\Delta(X_P, 0) \le 1.00$ m	$X_P \ge 160$ m
$\Delta(X_P, 0) \le 0.75$ m	$X_P \ge 213$ m
$\Delta(X_P, 0) \le 0.50$ m	$X_P \ge 320$ m
$\Delta(X_P, 0) \le 0.25$ m	$X_P \ge 640$ m

Table 3. Range of X_P satisfying the required resolution $\Delta(X_P, 0)$ as for the UCE M3

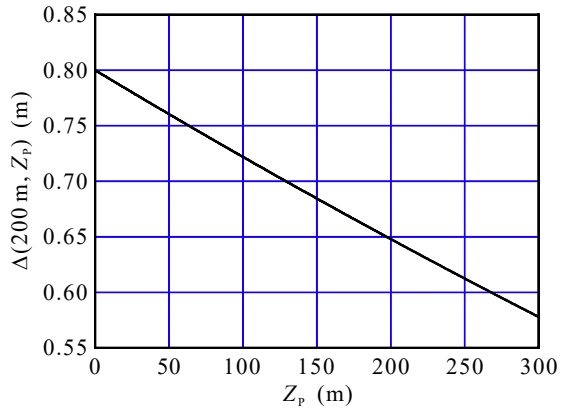


Figure 6. Resolution $\Delta(200, Z_P)$ of an image mapped on a vertical surface from a nadir image collected by the UCE M3

Now the production of a 3D city model is assumed to be conducted using aerial nadir images collected by a series of flight lines, similarly to ordinary aerial photogrammetry. Table 4 shows the minimum overlapping ratio satisfying the required resolution $\Delta(X_P, 0)$ for the assumed camera UCE M3, where the footprint of an image collected by the assumed camera is 2117 m wide (across track) by 1360 m long (along track).

Table 4 suggests that a texture image mapped on vertical surfaces produced from a nadir image collected in the usual procedure of

aerial photogrammetry by using the assumed camera UCE M3 would be able to be utilized for investigation of stories of a building. On the contrary, Table 4 indicates that it should be somewhat difficult to satisfy $\Delta(X_P, 0) \leq 0.50$ m in practical operation, and furthermore $\Delta(X_P, 0) \leq 0.25$ m should be practically impossible. That is to say, a texture image mapped on vertical surfaces produced from a nadir image collected in the usual procedure of aerial photogrammetry by using the assumed camera UCE M3 would be unable to satisfy the required resolution for the detection of doors and windows.

Resolution $\Delta(X_P, 0)$	Minimum overlap ratio	
	across track	along track
$\Delta(X_P, 0) \leq 1.00$ m	57.6%	61.8%
$\Delta(X_P, 0) \leq 0.75$ m	60.1%	65.7%
$\Delta(X_P, 0) \leq 0.50$ m	65.1%	73.5%
$\Delta(X_P, 0) \leq 0.25$ m	80.2%	97.0%

Table 4. Minimum overlapping ratio satisfying the required resolution $\Delta(X_P, 0)$ as for the UCE M3

5.2 Using images collected by an oblique aerial camera

In Japan, the Vexcel Imaging UltraCam Osprey 4.1 and the Leica CityMapper-2 are popular aerial camera systems with the function of collecting oblique images. Both camera systems can collect photogrammetry-grade nadir images and oblique images of off-nadir angles of 45° in four directions, such as forward, backward, leftward, and rightward.

Hence, we assumed the off-nadir angle ϕ_0 of an oblique aerial camera will be 45°. Then Equation (21) can be expressed by the following equation:

$$\Delta(X_P, Z_P) = \left(\frac{r}{f}\right)\frac{\{X_P + (H - Z_P)\}^2}{2X_P} \quad (37)$$

At the point $P(X_P, 0)$ on the ground, where $Z = 0$, Equation (37) can be expressed as follows:

$$\Delta(X_P, 0) = \left(\frac{r}{f}\right)\frac{(X_P + H)^2}{2X_P} \quad (38)$$

Rearranging Equation (38) gives us the following quadratic equation as to X_P:

$$\left(\frac{r}{f}\right)X_P^2 + 2\left\{\left(\frac{r}{f}\right)H - \Delta(X_P, 0)\right\}X_P + \left(\frac{r}{f}\right)H^2 = 0 \quad (39)$$

Solving Equation (39) brings the following formulas:

$$X_P = \frac{-\left\{\left(\frac{r}{f}\right)H - \Delta(X_P, 0)\right\} \pm \sqrt{D}}{\left(\frac{r}{f}\right)} \quad (40)$$

$$D = -\left\{2\left(\frac{r}{f}\right)H - \Delta(X_P, 0)\right\}\Delta(X_P, 0) \quad (41)$$

The condition $D \geq 0$ gives the following condition:

$$\Delta(X_P, 0) \geq 2\left(\frac{r}{f}\right)H \quad (42)$$

Substituting $\phi_0 = 45°$ for Equation (23), X_P that makes the resolution $\Delta(X_P, Z_P)$ the smallest is expressed by the following equation:

$$X_P = H - Z_P \quad (43)$$

Here we assume that aerial nadir images of 0.08 m GSD collected by the nadir camera of the UltraCam Osprey 4.1 (UCO 4.1) will be used in the production of 3D city models, and images collected simultaneously by four oblique cameras of UCO 4.1 will be used to be mapped on vertical surfaces of the 3D city model. The specifications of the assumed camera system UCO 4.1 are shown in Table 5 (Vexcel Imaging, 2025b). The footprint of an image collected by the nadir camera of the assumed camera system UCO 4.1 is 1644 m wide (across track) by 1121 m long (along track). Figure 7 shows the ground footprints of a nadir camera and four oblique cameras of the assumed camera system UCO 4.1 (Vexcel Imaging, 2025b).

	UCO 4.1 Nadir
Image size	20 544 × 14 016 pixels
Physical pixel size r	3.76 µm
Lens system focal length f	80 mm
Field of view θ across track × along track	51.8°× 36.6°
Flying height H for 0.08 m GSD	1702 m
	UCO 4.1 Oblique
Image size	14 144 × 10 560 pixels
Physical pixel size r	3.76 µm
Lens system focal length f	123 mm
Field of view θ image column × image row	24.3°× 18.4°

Table 5. Specifications of the assumed camera system UCO 4.1 (Vexcel Imaging, 2025b)

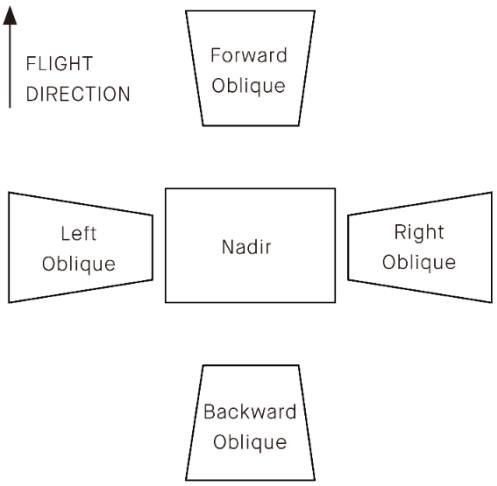


Figure 7. Ground footprint of the UCO 4.1 (Vexcel Imaging, 2025b)

When we discuss the resolution of an image mapped on a vertical surface in utilizing an oblique image collected by the assumed camera system UCO 4.1, the field of view θ of the short side of an oblique camera, that is to say, the image row direction, should be used.

Substituting the field of view θ of the short side of an oblique camera within the assumed camera system UCO 4.1 into Equation (7), the range of X_P —specifically, the coverage provided by the oblique camera—can be determined as follows:

$$\tan\left(45° - \frac{18.4°}{2}\right) \leq \frac{X_P}{1702} \leq \tan\left(45° + \frac{18.4°}{2}\right) \quad (44)$$

Rearranging Equation (43) brings the following range of X_P.

$$1229 \text{ m} \leq X_P \leq 2357 \text{ m} \quad (45)$$

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W6-2025-89-2025 | © Author(s) 2025. CC BY 4.0 License.

Substituting the specifications of the oblique camera from the assumed camera system UCO 4.1, as shown in Table 5, into Equations (38) and (42) yields the following formulas:

$$\Delta(X_\mathrm{P}, 0) = 3.0569 \times 10^{-5} \frac{(X_\mathrm{P}+1702)^2}{X_\mathrm{P}} \quad (46)$$

$$\Delta(X_\mathrm{P}, 0) \geq 2\left(\frac{0.00376}{123}\right) 1702 = 0.104 \text{ m} \quad (47)$$

Equation (43) indicates $X_\mathrm{P} = H = 1702$ m provides the minimum resolution $\Delta(H, 0) = 0.104$ m. By using Equations (45) and (46), the resolutions at both ends of the coverage of an oblique camera of the assumed camera system UCO 4.1 become $\Delta(X_\mathrm{P}, 1229\text{m}) = \Delta(X_\mathrm{P}, 2357 \text{ m}) = 0.107$ m. Figure 8 shows the resolution $\Delta(X_\mathrm{P}, 0)$ of an image mapped on a vertical surface from an oblique image collected by the assumed camera system UCO 4.1.

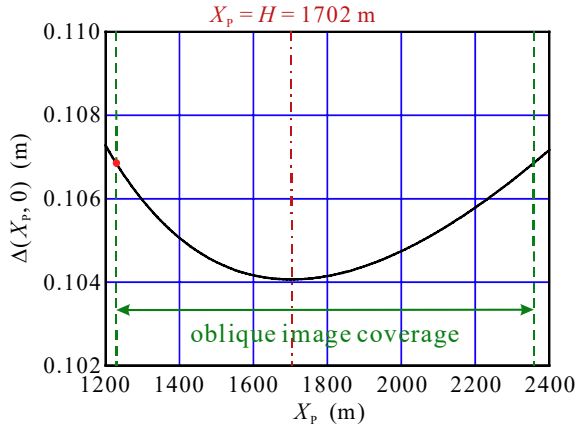


Figure 8. Resolution $\Delta(X_\mathrm{P}, 0)$ of an image mapped on a vertical surface from an oblique image collected by the UCO 4.1

Figure 8 indicates that all the mapped images satisfy the harshest condition $\Delta(X_\mathrm{P}, 0) \leq 0.25$ m in the whole coverage of an oblique camera of the assumed camera system UCO 4.1 expressed by Equation (45). Figure 8 shows that the variation of the resolution $\Delta(X_\mathrm{P}, 0)$ of an oblique image is extremely small in the whole coverage in contrast to that of a nadir image shown in Figure 5 as well. Furthermore, Equation (47) represents that $\Delta(X_\mathrm{P}, 0) < 0.104$ m should be impossible on the assumption in the paper.

To obtain the distribution of $\Delta(X_\mathrm{P}, Z_\mathrm{P})$ at $X_\mathrm{P} = H = 1702$ m, substituting the specifications of the camera shown in Table 2 for Equation (37) gives the following formula:

$$\Delta(H, Z_\mathrm{P}) = \left(\frac{0.00376}{123}\right) \frac{(2\times1702-Z_\mathrm{P})^2}{2\times1702} \quad (48)$$

Figure 9 shows the resolution $\Delta(H, Z_\mathrm{P})$ of an image mapped on a vertical surface from an oblique image collected by the assumed camera system UCO 4.1. Considering the heights of taller buildings in Japan, we show the resolution $\Delta(H, Z_\mathrm{P})$ up to $Z_\mathrm{P} = 300$ m in Figure 6. The resolution $\Delta(H, Z_\mathrm{P})$ becomes better almost linearly as Z_P increases.

Now the production of a 3D city model is assumed to be conducted using aerial nadir images collected by a series of flight lines. The most common image acquisition plan for topographic mapping using aerial nadir images is that the overlapping ratio of successive images along a flight line is 60 percent, and that of adjacent flight lines is 30 percent. Since the footprint of a nadir image collected by the assumed camera system is 1644 m wide (across track) by 1121 m long (along track), according to the image collection plan, the distance of adjacent flight lines is 1150 m and the distance between the ground nadir points of successive images along a flight line is 449 m. Table 6 shows which oblique camera images will be utilized for texture mapping of a vertical surface of a building when a flight for image collection is assumed from south to north.

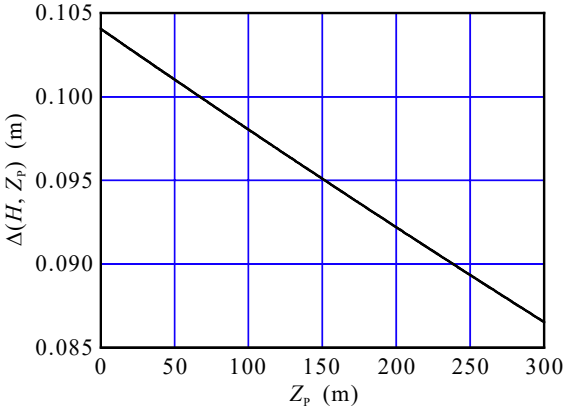


Figure 9. Resolution $\Delta(H, Z_\mathrm{P})$ of an image mapped on a vertical surface from an oblique image collected by the UCO 4.1

Vertical surface	Oblique camera
South-facing	Forward camera
North-facing	Backward camera
East-facing	Leftward camera
West-facing	Rightward camera

Table 6. Oblique camera utilized for texture mapping of a vertical surface

The coverage of an oblique image expressed by Equation (45) shows that there is no gap between the coverages of oblique images by the image collection plan. Accordingly, a texture image mapped on vertical surfaces produced from an oblique image collected in the usual procedure of aerial photogrammetry by using the assumed camera system UCO 4.1 would be able to be utilized for investigation of stories of a building and detection of doors and windows.

However, texture mapping with an oblique image necessitates collecting images over a larger area compared to producing a 3D city model using nadir images. Equation (45) suggests that two-by-two more flight lines and two-by-three images along a flight line are necessary for only texture mapping.

6. Conclusion

We investigated an evaluation of a texture image of a 3D city model produced from an aerial image. The investigation was focused on the resolution of a texture image mapped on a building surface. The required resolution of a texture image depends on the purpose of utilizing it. We decided to discuss the required resolutions for the investigation of stories and the detection of doors and windows by using simulated images of various resolutions.

A texture image mapped on a horizontal surface of a building produced from a nadir image, which is utilized to provide the geometry of a 3D city model, would be able to satisfy requirements for most purposes. Hence, the paper discussed texture images mapped on vertical surfaces of buildings of a 3D city model.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W6-2025-89-2025 | © Author(s) 2025. CC BY 4.0 License.

At first, we set up the required resolutions for the investigation of stories and the detection of doors and windows by using simulated images of various resolutions. Next, we derived the general formula to calculate the resolution of a texture image mapped on a vertical surface. We conducted a numerical analysis on the resolution of a texture image produced from an aerial image by using the obtained general formula. The aerial image was assumed to be collected by a nadir camera and an oblique camera popularly utilized in Japan.

The analysis results indicate that it would be possible to produce a texture image with a resolution adequate for investigating the stories of a building from a nadir image. However, the results indicate that it is impossible to produce a texture image with a resolution sufficient for detection of a door and a window of a vertical building surface from a nadir image.

On the other hand, the results indicate that it could produce a texture image with a resolution sufficient not only for investigating the stories of a building but also for detecting a door and a window on a vertical building surface from an oblique image.

As for the utilization of an oblique image, there remain two major problems to solve. One issue is that occlusion caused by buildings on the side of a camera in an oblique image may occur more frequently than in a nadir image. P. Meixner and F. Leberl addressed this issue and concluded that vertical imagery is more suitable for façade analysis (Meixner, Leberl, 2010a, 2010b).

To reduce the effect of occlusion, we should make an image collection plan considering the distribution of taller buildings intensively. We are currently planning to develop software that will use both oblique and nadir images for texture mapping on a single surface, aiming to minimize the impact of occlusion. The software will include a feature that detects occlusions on building surfaces within images by utilizing the geometry of a 3D city model. It will also automatically select appropriate oblique and nadir images for texture mapping.

The other issue is that the multi-directional camera systems with the function of collecting oblique images, such as the Vexcel Imaging UltraCam Osprey 4.1 and the Leica CityMapper-2, might be hard to use in current Japan. The reason is that the acquisition of the geometry of buildings of a 3D city model is conducted by using nadir images in the current Japanese ordinary procedure to produce a 3D city model. The nadir camera of the camera system is not always preferable for producing a 3D city model due to its smaller image size. Moreover, the volume of data collected by the camera system might be huge, and the volume of the data collected by oblique cameras for texture mapping to add value to a 3D city model is more than twice more as that of the data collected by a nadir camera to produce a 3D city model without texture. If oblique images are utilized for acquisition of the geometry of buildings of a 3D city model, the second problem should be smaller.

The study took no account of the utilization of such image processing techniques as cubic convolution using a single image and super resolution using multiple images at texture mapping. Such image processing techniques provide higher resolution of a texture image, clearly suitable to make the appearance of buildings of the 3D city model realistic. We consider that it is necessary to investigate the effect of such image processing techniques on some analyses by a computer using texture images of a 3D city model, such as investigation of stories of a building and detection of doors and windows on a building surface.

References

Buyukdemircioglu, M. and Oude Elberink, S., 2024. Automated texture mapping CityJSON 3D city models from oblique and nadir aerial imagery, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-4/W5-2024, 87–93, https://doi.org/10.5194/isprs-annals-X-4-W5-2024-87-2024.

He, H., Yu, J., Cheng, P., Wang, Y., Zhu, Y., Lin, T., Dai, G., 2022. Automatic, Multiview, Coplanar Extraction for CityGML Building Model Texture Mapping, *Remote Sensing*, 14, 50, https://doi.org/10.3390/rs14010050.

Lippoldt, F., 2019. Window Detection in Facades for Aerial Texture Files of 3D CityGML Models, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 11–19.

Meijering, E., 2002. A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing, *Proceedings of the IEEE*, vol. 90, no. 3, 319–342.

Matsuoka, R., Takemoto, T., Takahashi, G., Inazawa, T., and Sogo, S., 2024. Estimation of Photovoltaic Potential of Urban Buildings Considering a Solar Panel Arrangement Using a 3D City Model, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-4/W4-2024, 123–130, https://doi.org/10.5194/isprs-annals-X-4-W4-2024-123-2024.

Meixner, P., Leberl, F., 2010a. Vertical-or oblique imagery for semantic building interpretation, *Vorträge Dreiländertagung OVG, DGPF und SGPF*, 247–256.

Meixner, P., Leberl, F., 2010b. Interpreting Building Facades from Vertical Aerial Images Using the Third Dimension, *Proceedings of the ISPRS Technical Commission IV and AutoCarto Symposium*.

Ministry of Land, Infrastructure, Transport and Tourism of Japan (MLIT), 2025a. https://www.mlit.go.jp/plateau (April 7, 2025).

Ministry of Land, Infrastructure, Transport and Tourism of Japan (MLIT), 2025b. Standard Data Product Specification for 3D City Model, https://www.mlit.go.jp/plateau/file/libraries/doc/plateau_doc_0001_ver04.pdf.

Ministry of Land, Infrastructure, Transport and Tourism of Japan (MLIT), 2025c. Standard Implementation Procedures for 3D City Model, https://www.mlit.go.jp/plateau/file/libraries/doc/plateau_doc_0002_ver04.pdf.

Ministry of Land, Infrastructure, Transport and Tourism of Japan (MLIT), 2025d. Manual of Survey for 3D City Modeling, https://www.mlit.go.jp/plateau/file/libraries/doc/plateau_doc_0010_ver04.pdf.

Vexcel Imaging, 2025a. ULTRACAM E AGLE MARK 3, https://www.vexcel-imaging.com/brochures/UC_Eagle_M3_en.pdf.

Vexcel Imaging, 2025b. ULTRACAM OSPREY 4.1, https://www.vexcel-imaging.com/brochures/UC_Osprey_4.1_en.pdf.

Yang, J., Huang, T.S., 2017. Image Super-resolution: Historical Overview and Future Challenges, *Super-Resolution* edited by Peyman Milanfar. CRC Press, Boca Raton, 1–23.