



## Article

# True<sup>2</sup> Orthoimage Map Generation

Guoqing Zhou <sup>1,2,3,\*</sup> , Qingyang Wang <sup>1,2</sup> , Yongsheng Huang <sup>4</sup>, Jin Tian <sup>2,3</sup>, Haoran Li <sup>2,3</sup> and Yuefeng Wang <sup>2,3</sup><sup>1</sup> College of Earth Sciences, Guilin University of Technology, Guilin 541004, China<sup>2</sup> Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin University of Technology, Guilin 541004, China<sup>3</sup> College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, China<sup>4</sup> Guangxi Zhuang Autonomous Region Natural Resources Remote Sensing Institute, Nanning 530219, China

\* Correspondence: gzhou@glut.edu.cn; Tel.: +86-0773-589-6097

**Abstract:** Digital/true orthoimage maps (D/TOMs) are one of the most important forms of national spatial data infrastructure (NSDI). The traditional generation of D/TOM is to orthorectify an aerial image into its upright and correct position by deleting displacements on and distortions of imagery. This results in the generated D/TOM having no building façade texture when the D/TOM superimposes on the digital building model (DBM). This phenomenon is no longer tolerated for certain applications, such as micro-climate investigation. For this reason, this paper presents the generation of a true<sup>2</sup> orthoimage map (T<sup>2</sup>OM), which is radically different from the traditional D/TOM. The basic idea for the T<sup>2</sup>OM generation of a single building is to orthorectify the DBM-based building roof from up to down, the building façade from front to back, from back to front, from left side to right side, and from right side to left side, as well as complete a digital terrain model (DTM)-based T<sup>2</sup>OM, of which a superpixel is proposed to store building ID, texture ID, the elevation of each pixel, and gray information. Two study areas are applied to verify the methods. The experimental results demonstrate that the T<sup>2</sup>OM not only maintains the traditional characteristics of D/TOM, but also displays building façade texture and three-dimensional (3D) coordinates (XYZ) measurable at any point, and the accuracy of 3D measurement on a T<sup>2</sup>OM can achieve 0.025 m (0.3 pixel).

**Keywords:** true<sup>2</sup> orthoimage map; true orthoimage map; superpixel; three-dimensional measurement; orthorectification



**Citation:** Zhou, G.; Wang, Q.; Huang, Y.; Tian, J.; Li, H.; Wang, Y. True<sup>2</sup> Orthoimage Map Generation. *Remote Sens.* **2022**, *14*, 4396. <https://doi.org/10.3390/rs14174396>

Academic Editor: Devrim Akca

Received: 6 August 2022

Accepted: 1 September 2022

Published: 4 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

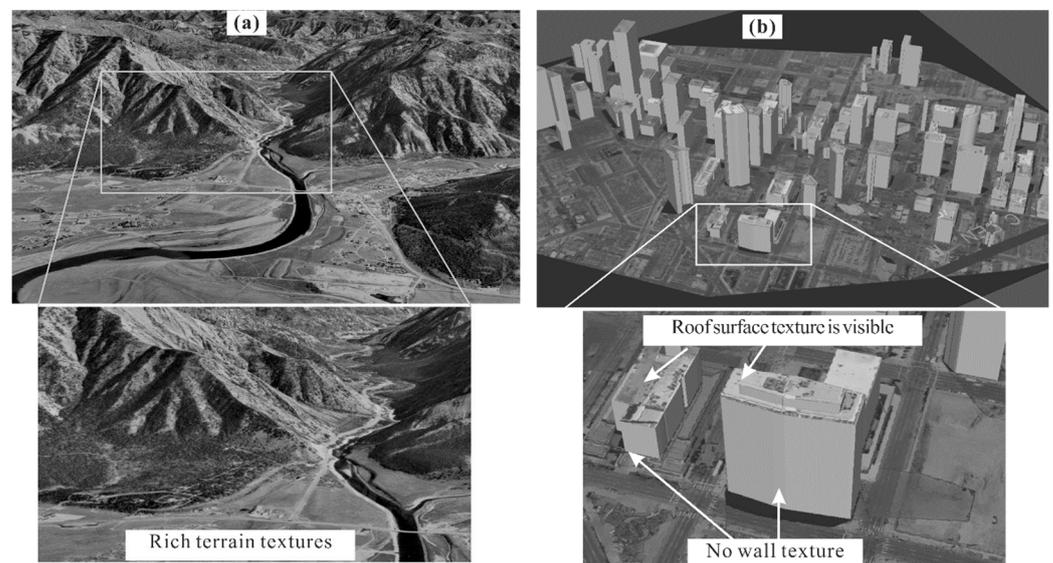
## 1. Introduction

Digital orthophotomaps (DOMs) are a critical component of national spatial data infrastructure (NSDI) [1–4]. DOMs (1) serve as a geospatial foundation upon which to add detail and attach attribute information; (2) provide a base on which to accurately register and compile other themes of data; and (3) orient and link the results of an application to the landscape [5,6]. Especially, a highly detailed DOM is capable of serving as a source for locating the features to be mapped and measured [7].

Many investigations have demonstrated that the generation of high-resolution urban DOMs using the existing procedures and algorithms, proposed in 1990 by the USA National Digital Orthophoto Program (NDOP), has encountered many problems, such as incomplete orthorectification, occlusion, ghost image, shadow, etc. A comprehensive discussion regarding these problems can be found in [8,9]. Thus, the generation of so-called true orthophoto maps (TOMs) has become obligatory, and has been researched by many studies at the end of the 20th century and the beginning of the 21st century, such as [10–39].

However, when a TOM superimposes onto a corresponding digital terrain model (DTM) in a flat and/or hilly area, the terrestrial texture is clearly visible (Figure 1a), but when a TOM superimposes onto a digital building model (DBM), the detailed façade textures of a building are not visible (Figure 1b). The impacts of these problems have significantly influenced the usefulness of TOMs in industries such as micro-climate monitoring,

micro-environment analysis, and cellphone transmission station distribution, since incomplete building façade information cannot be tolerated for these applications. Therefore, this paper proposes a true<sup>2</sup> orthoimage map (T<sup>2</sup>OM) generation method which can provide three-dimensional (3D) and detailed textures of a building's roof and façades.



**Figure 1.** (a) Terrestrial textures visible in flat and hilly area when TOM is superimposed on DTM (<http://www.pcvr.com.cn/html/software/softwarei.html> (accessed on 5 August 2022)), and (b) building façade texture visible when DOM is superimposed on DBM.

## 2. Related Works

The study of TOM began in the 1980s and has continued for decades. A successful and complete TOM generation method consists of an orthorectification algorithm, occlusion detection and compensation, shadow detection and recovery, a seamless mosaic, etc. The review of related works is, therefore, classified accordingly as follows.

**Occlusion detection and compensation:** Occlusion detection is one of the most major components of TOM generation. Amhar et al. [10] and Schickler and Thorpe [11] considered the hidden effects introduced by abrupt changes of surface height (e.g., buildings and bridges). Jauregui et al. [20] presented a procedure for orthorectifying aerial photographs to produce and update terrain surface maps. Vassilopoulou et al. [40] used IKONOS images to generate orthoimages for monitoring volcanic hazards on Nisyros Island, Greece, and Siachalou [21] used IKONOS images to generate the urban orthoimage. Cameron et al. [41] analyzed orthorectified aerial photographs to measure changes in the native pinewood of Scotland, and Passini and Jacobsen [42] analyzed the accuracy of orthoimages from very-high-resolution imagery. Biason et al. [22] further explored the automatic generation of true orthoimages. Piatti and Lerma [43] address the problem of image orthorectification through photogrammetric simulation and its generation based on digital elevation/building/surface models, as well as internal and external orientation parameters of image sensors (i.e., digital cameras). This method appears to be able to create the high-resolution 3D models needed for accurate orthophotos. Zhou et al. [29] proposed a new urban orthophoto occlusion detection method, which first establishes a model describing the relationship between each ghost image and the corresponding building occlusion boundary, and then applies an algorithm that uses building displacement to identify the occlusion region in the ghost image. The method can effectively avoid pseudo-occlusion detection and the drawback of simultaneous occlusion detection and orthophoto generation, providing a key technique for DBM-based T<sup>2</sup>OM. Yoo and Lee [44] proposed a facet-based method for generating realistic orthophotos of building surface facets. The method identifies occluded areas based on the unit surface of the building and uses multiple images and

high-detail digital building model data (i.e., DBM) to recover the occluded areas from each other. Oliveira et al. [45] proposed a new occlusion detection method for true radiographic image mosaic generation. The proposed method uses irregularly spaced point clouds to identify occluded regions, avoiding the interpolation process as an initial step in occlusion detection and thus avoiding the insertion of additional errors in the surface representation. Zhou and Sha [46] proposed a method to simultaneously detect building roof and ground shadows using the DBM as an overlay model. The method determines the solar zenith angle and solar altitude angle by selecting the geographic information of the corner points on the shadow boundary in the aerial image, and then displays the actual shadow area determined in the DBM on the ghost image. The method is independent of ground reflectivity and illumination conditions, and provides technical support for producing high-quality true projection images. Marsetič [47] proposed a method to automatically generate true orthophotos from optical super-resolution satellite images. The automatic workflow consists of five modules, starting with the extraction of ground control points, then the geometric processing of image patches, occlusion detection, orthorectification, and finally generating a real orthophoto. The quality of the true radiographic images produced by this method depends on the accuracy of the geometric correction and the number of images. The occlusion compensation for TOM generation, Skarlatos [13] and Greenfeld [15] demonstrated that building occlusions significantly influenced not only image quality but also the accuracy of the orthoimages. Rau et al. [18] treated enhancements in image radiometry, demonstrating a suitable enhancement technique to restore information within building shadow areas. Sheng et al. [48] used a model-based method to reconstruct a true orthophoto model (CSM) to replace the DEM to generate a true photogram of the forest scene, which mainly focused on the efforts of occlusion and distortion caused by trees in the forest area. Zhou et al. [9] compensated by conjugating blocks of orthoimages, i.e., by refilling the masked area adjacent to the orthoimages. Zhou et al. [29] used adjacent overlapping “slave” orthophotos to fill the occluded region using the filling method proposed by Zhou et al. [9] to compensate for the occluded region in the “master” orthophoto. With such occlusion compensation, a complete true orthophoto can be created for the study area.

For shadow detection and recovery, many efforts have been made. For example, Leone and Distanto [49] performed image shadow detection by improving the classification, segmentation and localization of detected objects, which improved the effect of shadow detection. Makarau et al. [50] proposed an alternative robust method for shadow detection. The method is to adaptively calculate the parameters of a specific scene and allow one to use many different sensors and images obtained under different lighting conditions, improving the accuracy of shadow detection. Tiwari et al. [51] propose an improved algorithm to obtain rough shadows by changing the ratio of intensity to hue and then performing shadow compensation using local thresholding. The experimental results show that the method is more suitable for the shadow detection of low-intensity and medium-intensity images, and the shadow compensation algorithm is suitable for all test images.

For the mosaic of multiple TOMs. Many researchers have made many efforts for the generation of high-quality TOM through the improvement of TOM mosaicking, such as [11,17,26,36–38,52,53]. Their studies resulted in clearer features in the shadows and more continuous and natural grayscale of the filled areas and surrounding images. For example, Pan and Wang [53] adopted a multi-scale processing strategy which can automatically locate the specific positions of the splicing lines and transition areas and improve the image quality after mosaicking. Gharibi and Habib [36] proposed a weighted averaging method to mitigate seam line effects and spectral differences that may occur in true orthophoto mosaics.

Despite many previous efforts, traditional TOM only provides 2D ( $XY$ ) coordinates and building roofs' texture information, while the 3D attributes and the textures of the building facades cannot be provided at all. For this reason, the generation of  $T^2OM$  is presented in this paper. The organization of this paper is arranged as follows: The principle of  $T^2OM$  is presented in Section 3, and Section 4 presents the experimental results and analysis. The conclusions are drawn in Section 5.

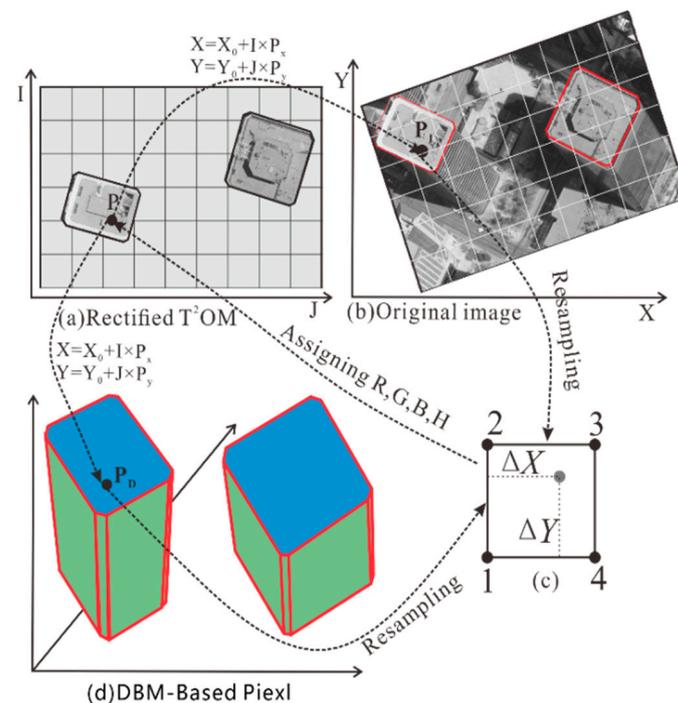
### 3. Principles of True<sup>2</sup> Orthoimage Map (T<sup>2</sup>OM) Generation

T<sup>2</sup>OM is defined as a DOM that can provide measurable 3D (XYZ) coordinates and textures for both the roof and the facade of a building. This means that the T<sup>2</sup>OM not only has traditional TOM characteristics, but also provides the 3D geometric information (X,Y,Z) and the textures for the facade of a building. The method of generating a T<sup>2</sup>OM includes the four basic steps below:

- (1) DBM-based single-building T<sup>2</sup>OM generation, which consists of orthorectifying both the building roof and building facades: a concept, named “superpixel” is proposed for the storage of building texture, building ID, etc. information.
- (2) DBM-based multiple-building T<sup>2</sup>OM generation: merging the DBM-based single-building T<sup>2</sup>OM, including organization of the building ID, building façade, building corner coordinates, etc.
- (3) DTM-based T<sup>2</sup>OM generation for the orthorectification of gentle and continuously elevated hilly areas.
- (4) DTM- and DBM-based T<sup>2</sup>OM merging, which is for merging DTM- and DBM-based T<sup>2</sup>OM for the creation of an entire T<sup>2</sup>OM.

#### 3.1. Generation of a DBM-Based Single-Building T<sup>2</sup>OM

In order to clearly describe the process for the generation of T<sup>2</sup>OM for buildings, a single building is first taken as an example (see Figure 2), presuming that the DBM for the single building and the exterior/interior orientational parameters (EOPs/IOPs) for an image are known. The steps for the generation of a T<sup>2</sup>OM for a single building are below.



**Figure 2.** The procedure of DBM-based building roof orthorectification. (a) Rectified T<sup>2</sup>OM; (b) Original image; (c) Resampling; (d) DBM-Based pixel.

#### Step 1: DBM-based building roof orthorectification, which consists of:

- (1) *Determining the size of the T<sup>2</sup>OM*: The resulting DBM-based single-building T<sup>2</sup>OM is expressed as a raster image with pixels arranged in rows and columns. Since the resulting orthoimage is orthorectified from raster image input (called original image) using the DBM data, the size of the output image is defined [9] as

$$\begin{aligned} X_0 &= \max\{\min\{X_D\}, \min\{X_I\}\}, Y_0 = \max\{\min\{Y_D\}, \min\{Y_I\}\} \\ X_1 &= \min\{\max\{X_D\}, \max\{X_I\}\}, Y_1 = \min\{\max\{Y_D\}, \max\{Y_I\}\} \end{aligned} \tag{1}$$

where  $X_0$  and  $Y_0$  are the coordinates of the lower-left corner of the output image;  $X_1$  and  $Y_1$  are the coordinates of the upper-right corner of the output image;  $X_D$  and  $Y_D$  are the X and Y coordinates of the DBM;  $X_I$  and  $Y_I$  are the X and Y coordinates of the original image; and max and min denote maximum and minimum of the elements in the blanket. All of the coordinates here refer to the geodetic coordinate system required in the resulting T<sup>2</sup>OM.

(2) *Computing the X, Y coordinate of each pixel:* In Figure 2,  $P(I, J)$  is a given point pixel on the roof of the T<sup>2</sup>OM building roof, and their raster rows and columns can be transformed to the coordinates of the output T<sup>2</sup>OM, i.e.,

$$\begin{aligned} X &= X_0 + I \times P_x \\ Y &= Y_0 + J \times P_y \end{aligned} \tag{2}$$

where  $X$  and  $Y$  represent the coordinates of pixels;  $X_0$  and  $Y_0$  are the coordinates of the lower-left corner of the output roof T<sup>2</sup>OM;  $P_x$  and  $P_y$  are the sizes of the pixels in the X and Y directions, respectively; and  $I$  and  $J$  are the rows and columns of  $P$  points, respectively.

(3) *Computing the Z coordinate of P(I, J):* In order to perform orthorectification, we also need to know the Z coordinates of the pixel  $P(I, J)$  in the output roof T<sup>2</sup>OM and this is obtained from DBM. However, DBM data only have vector coordinates at corner points. Therefore, it is necessary to interpolate an elevation to the roof pixel of the building. As shown in Figure 3, the elevation (height) is obtained only for pixels with corner points (blue pixels in Figure 3), while the other pixels (orange in Figure 3) are calculated by:

$$h = a_0j - a_1i - a_2 \tag{3}$$

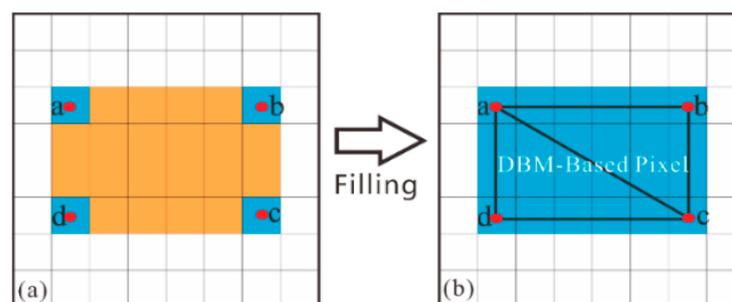
where  $h$  is the raster height value;  $i, j$  is the row number of the raster;  $a_0, a_1$  and  $a_2$  are the equation weight values defined by:

$$Aj + Bi + Ch + D = 0 (C \neq 0) \tag{4}$$

$$h = -\frac{A}{C}j - \frac{B}{C}i - \frac{D}{C} \tag{5}$$

$$a_0 = -\frac{A}{C}, a_1 = -\frac{B}{C}, a_2 = -\frac{D}{C} \tag{6}$$

where each triangular surface weight can be calculated from Equation (2) in the case of three known vertices of the triangular surface, and  $A, B$  and  $C$  are the equation weight values.



**Figure 3.** Assigning an elevation value to the building’s roof. (a) Before the elevation value is filled; (b) After the elevation value is filled, where a, b, c and d represent corner of DBM.

When the pixels on the roof of the building have an elevation, we have to convert the pixel's geodetic coordinates back into the column and row by:

$$R_D = (X - X_{0D})/P_D, C_D = (Y - Y_{0D})/P_D \tag{7}$$

where  $P_D$  denotes the pixel size of the DBM image.  $X_{0D}$  and  $Y_{0D}$  are the lower left corner coordinates of the DBM image.  $K_D$  and  $L_D$  will not generally be exact integers. Thus, an interpolation must be performed to determine  $Z$ . Usually, a bilinear interpolation method of the following form is employed (see Figure 2c):

$$Z = \{[Z_1\Delta X + (1 - \Delta X)Z_4] + [Z_2\Delta X + (1 - \Delta X)Z_3] + [Z_1\Delta Y + (1 - \Delta Y)Z_4] + [Z_2\Delta Y + (1 - \Delta Y)Z_3]\}/4 \tag{8}$$

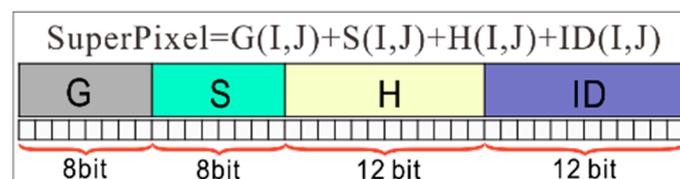
where  $\Delta X = R_D - R_m$  and  $\Delta Y = C_D - C_m$ , in which  $R_m$  is the  $R_D$  rounded to its maximal integer and  $C_m$  is  $C_D$  rounded to its maximal integer. After this estimation, we then know the coordinate  $(X, Y, Z)$  of the pixel.

- (4) *Computing the corresponding coordinate in the original image:* In order to orthorectify the source image, the corresponding coordinate of the source image pixel in the output image is calculated by:

$$\begin{aligned} x_I &= x_{0I} - f \frac{a_1(X-X_s)+b_1(Y-Y_s)+c_1(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)} \\ y_I &= y_{0I} - f \frac{a_2(X-X_s)+b_2(Y-Y_s)+c_2(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)} \end{aligned} \tag{9}$$

where  $x_I$  and  $y_I$  are the corresponding coordinates of the pixel  $P(X,Y)$  in the source image;  $X_s, Y_s$  and  $Z_s$  are the exposure stations;  $f$  is the focal length;  $ai = \{a_1, a_2, a_3\}$ ,  $bi = \{b_1, b_2, b_3\}$  and  $ci = \{c_1, c_2, c_3\}$  are the elements of the rotation matrix, and are the functions of the three exterior orientation angles  $(\phi, \omega, \kappa)$ . These elements have to be computed using at least three ground control points (GCP).

- (5) *Assigning the gray value to pixels:* Since the grid of pixels in the source image rarely matches the grid of the output orthoimage, a re-sampling of the pixels has to be performed in order to assign gray value to the pixels in the output image. The nearest neighbor is employed because it directly transfers the original data values without averaging them. The computational procedure is illustrated in Figure 2.
- (6) *Storing the data for DBM-based building roof T<sup>2</sup>OM:* As can be seen from the above, T<sup>2</sup>OM needs to store more information than the traditional TOM does, such as building roof texture, facade texture, and façade Z coordinates. For this reason, "superpixel" is presented and has the following characteristics (see Figure 4): (1) it inherits the original image gray information; (2) the gray value, elevation, building ID, and facade texture index ID are stored; (3) each pixel coordinate is directly interconnected with the building ID and façade texture ID.



**Figure 4.** Superpixel occupies a total of 40 bits in the computer, where  $I, J$  represent the row and column,  $G(I, J)$  denotes gray,  $S(I, J)$  denotes corner coordinate subdivision grid identification value,  $H(I, J)$  denotes elevation, and  $ID(I, J)$  represents the identification of a building.

The detailed descriptions of  $G, S, H,$  and  $ID$  in the superpixel are as follows:

- (1)  $G(I, J)$  stands for storage of the gray at  $i$ -th row and  $j$ -th column in the image coordinate system whose value is 0–255.

- (2)  $S(I, J)$  stands for storage of the building corner coordinate subdivision grid identification value, which occupies 8 bits. That is, by dividing a single pixel into 256 subdivision sequences, the accuracy of vector to grid data conversion is improved. For a given point  $P(X_p, Y_p)$ , this can be expressed as  $(i_p, j_p)$  after the conversion of the vector to the grid. The lost information is  $(X_p - X_0) - i_p \cdot \Delta X$ . With this information,  $S_x$  can be calculated through Equation (8).

$$S_x = \left\lfloor \frac{(X_p - X_0 - \Delta x \cdot i_p) \cdot 16}{\Delta x} \right\rfloor \tag{10}$$

where  $\Delta X$  is the image resolution and  $\lfloor \cdot \rfloor$  is the function of rounding down.  $(X_0, Y_0)$  is the top-left point of the image.

The calculation method used for  $S_y$  is the same as that used for  $S_x$ . At this time,  $S$  is expressed as two values. This makes the storage of  $S$  very difficult. Therefore,  $(S_x, S_y)$  is converted into a one-dimensional form by means of the Morton code [54].

- (3)  $H(I, J)$  stands for the storage of the building height or DTM height with a floating format.
- (4)  $ID$  stands for the storage of building  $ID$ . An  $ID$  can be used to call for the facade texture. A large city may have hundreds of buildings; therefore, 12 bits are designed to store 0 to 4095 buildings.

The T<sup>2</sup>OM data for the roof of the building generated by the above steps are stored in the superpixel and shown in Table 1.

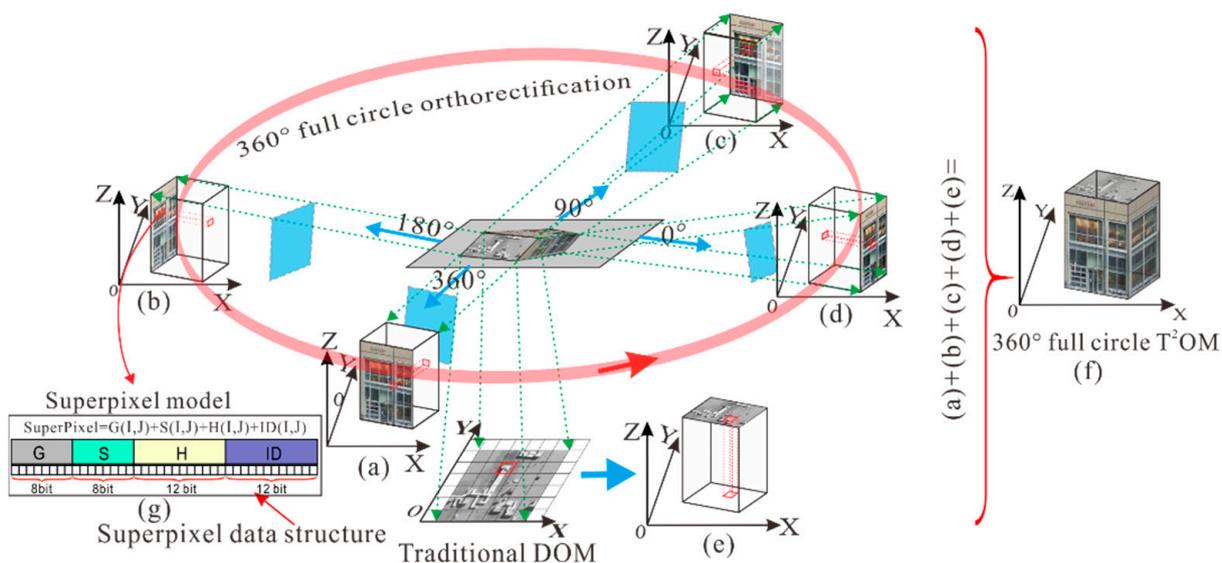
**Table 1.** The superpixel for the roof texture data-generated DBM-based building.

DBM-Based Pixel	Row	Column	Gray Value	Sub Ordinal	Height	BuildingID
BP1	I(BP1)	J(BP1)	G1	$id_{(BP1)}(r,c)$	$h_{(BP1)}$	$B_1$
BP2	I(BP2)	J(BP2)	G2	0	$h_{(BP2)}$	$B_1$
...	...	...	...	...	...	...
BPi	I(BPi)	J(BPi)	Gi	$id_{(BPi)}(r,c)$	$h_{(BPi)}$	$B_1$
BPi + 1	I(BPi + 1)	J(BPi + 1)	Gi + 1	0	$h_{(BPi+1)}$	$B_1$
...	...	...	...	...	...	...

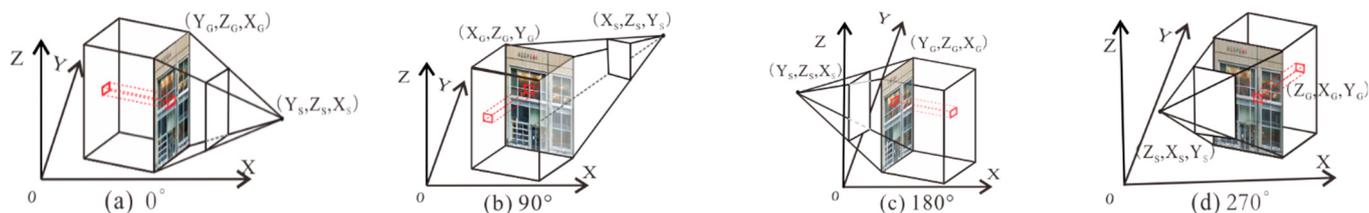
**Step 2: DBM-based orthorectification for a single-building facade**

In order to obtain the façade texture and the 3D coordinates of a building, the orthorectification for building facade in four directions is also performed in T<sup>2</sup>OM, and the four directions are determined according to the minimum bounding box (a detailed description has been given by [55]). The basic idea for the orthorectification of a building façade is: four directions at 0° for the front façade (Figure 5a), 90° for the left façade (Figure 5b), 180° for the back façade (Figure 5c), and 270° for the right façade (Figure 5d) are orthorectified, respectively. For example, the co-linear equation for the 0° directional facade texture is adopted (see Figure 6a), i.e.,

$$\begin{aligned} y_g &= -f \frac{a_1(Y_G - Y_S) + b_1(Z_G - Z_S) + c_1(X_G - X_S)}{a_3(Y_G - Y_S) + b_3(Z_G - Z_S) + c_3(X_G - X_S)} + y_0 \\ z_g &= -f \frac{a_2(Y_G - Y_S) + b_2(Z_G - Z_S) + c_2(X_G - X_S)}{a_3(Y_G - Y_S) + b_3(Z_G - Z_S) + c_3(X_G - X_S)} + z_0 \end{aligned} \tag{11}$$



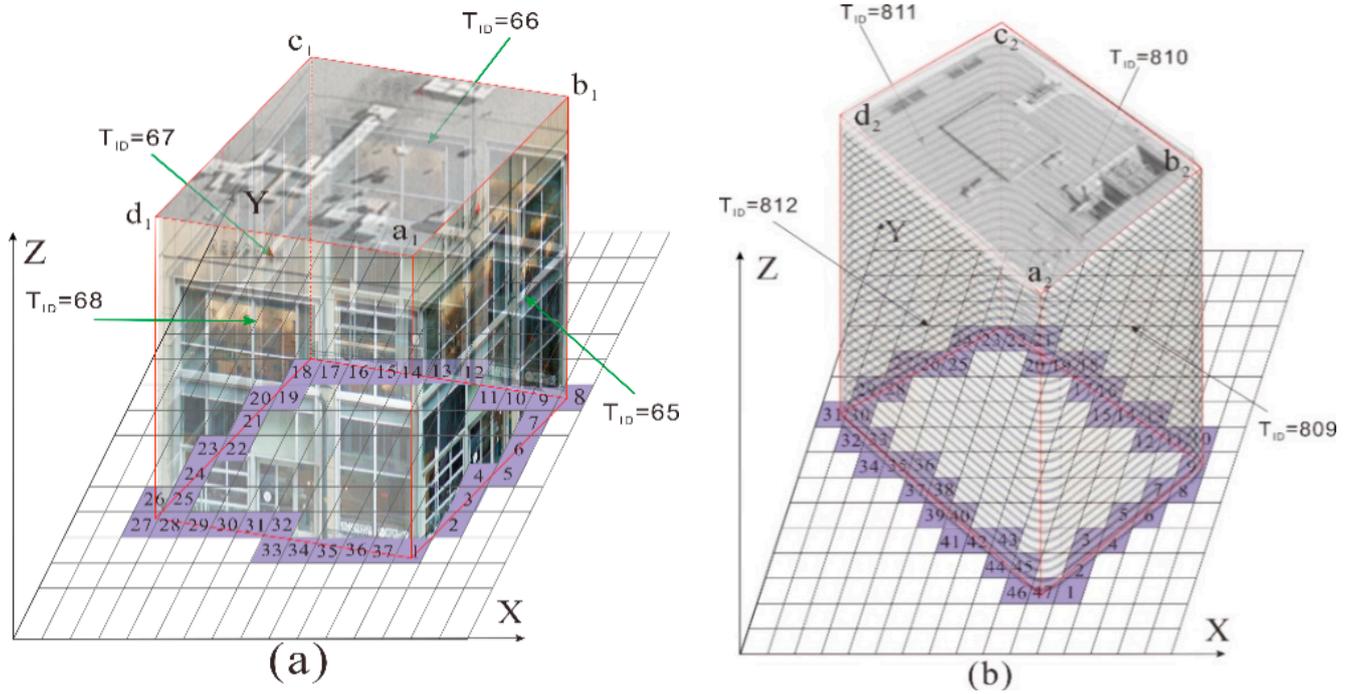
**Figure 5.** True three-dimensional (360°) full-circle T<sup>2</sup>OM generation. (a) Orthorectification for a building facade in 360° direction; (b) orthorectification for a building facade in 180° direction; (c) orthorectification for a building facade in 90° direction; (d) orthorectification for a building facade in 0° direction; (e) orthorectification for the building’s roof; (f) the generation of T<sup>2</sup>OM for a building after a full circle (360°) orthorectification; (g) explanation of a superpixel data structure.



**Figure 6.** Orthorectification method for the building facade through 4 different directions. (a) Orthorectification for a building facade in 0° direction; (b) orthorectification for a building facade in 90° direction; (c) orthorectification for a building facade in 180° direction; (d) orthorectification for a building facade in 270° direction.

The facade texture in 0° direction is orthorectified into the ZOY plane (see Figure 6a). Similarly, the collinear equation for the orthorectification of the other three directions at 90°, 180°, 270° can be orthorectified into the ZOY (see Figure 6b), YOZ (see Figure 6c), and YOX (see Figure 6d) planes, respectively.

The details of the orthorectification of a single building can be further described as follows: As shown in Figure 7, the buildings in Figure 7a,b are 25.8 m and 46.4 m, respectively, with an elevation resolution of 0.2 m. There are four planes ( $a_1b_1$ ,  $b_1c_1$ ,  $c_1d_1$  and  $d_1a_1$ ) for the building facade in Figure 7a, and the corresponding texture index data are 65, 66, 67 and 68, respectively. There are four planes ( $a_2b_2$ ,  $b_2c_2$ ,  $c_2d_2$  and  $d_2a_2$ ) for the building facade in Figure 7b, and the corresponding texture index data are 809, 810, 811 and 812, respectively (see Tables 2 and 3).



**Figure 7.** Building 3D spaghetti data structure, where TID represents the wall texture ID (e.g.,  $T = 809$ );  $a_1, b_1, c_1, d_1, a_2, b_2, c_2,$  and  $d_2$  represent building’s roof corners. (a) 3D data structure of a building with a height of 25.8 m; (b) 3D data structure of a building with a height of 46.4 m.

**Table 2.** The building wall texture superpixel data in Figure 7a.

Pixel	Gray (8 bit)	Hight ( $H$ , 12 bit)	TextureID (TID, 12 bit)	Notes
1	11000110	000100000010	000001000001	Wall 1 $H = 25.8$ TID = 65
2	11010011	000100000010	000001000001	
3	11010000	000100000010	000001000001	
...	...	...	...	
8	11010011	000100000010	000001000001	Wall 2 $H = 25.8$ TID = 66
9	11001000	000100000010	000001000010	
10	11101000	000100000010	000001000010	
...	...	...	...	Wall 3 $H = 25.8$ TID = 67
18	11011101	000100000010	000001000010	
19	11011101	000100000010	000001000011	
20	11011000	000100000010	000001000011	
...	...	...	...	Wall 4 $H = 25.8$ TID = 68
27	11011000	000100000010	000001000011	
28	11010001	000100000010	000001000100	
29	11011010	000100000010	000001000100	
...	...	...	...	
37	11011010	000100000010	000001000100	

**Table 3.** The building façade texture superpixel data in Figure 7b.

Pixel	Gray (8 bit)	Hight ( $H$ , 12 bit)	TextureID (TID, 12 bit)	Notes
1	11010101	000111010000	001100101001	Wall 1 $H = 46.4$ TID = 809
2	11111101	000111010000	001100101001	
3	11110001	000111010000	001100101001	
...	...	...	...	
9	11010100	000111010000	001100101001	Wall 2 $H = 46.4$ TID = 810
10	11010111	000111010000	001100101010	
11	11010101	000111010000	001100101010	
...	...	...	...	
22	11010101	000111010000	001100101010	Wall 3 $H = 46.4$ TID = 811
23	11011111	000111010000	001100101011	
24	11010111	000111010000	001100101011	
...	...	...	...	
30	11010111	000111010000	001100101011	Wall 4 $H = 46.4$ TID = 812
31	11010101	000111010000	001100101100	
32	11010111	000111010000	001100101100	
...	...	...	...	
47	11010101	000111010000	001100101100	

The buildings in Figure 7a,b have 37 and 47 façade superpixels, respectively, and the superpixels are stored for each pixel as shown in Tables 2 and 3. The corresponding pixels, textures, and elevations of each facade of the building in Figure 7 can be obtained, and the detailed storage contents are shown in Tables 4 and 5.

**Table 4.** The building wall texture index in Figure 7a.

WallID	Wall Pixel Index	Texture Index	Hight
$a_1b_1$	1, 2, 3, 4, 5, 6, 7, 8	65	25.8
$b_1c_1$	9, 10, 11, 12, 13, 14, 15, 16, 17, 18	66	25.8
$c_1d_1$	19, 20, 21, 22, 23, 24, 25, 26, 27	67	25.8
$d_1a_1$	28, 29, 30, 31, 32, 33, 34, 35, 36, 37	68	25.8

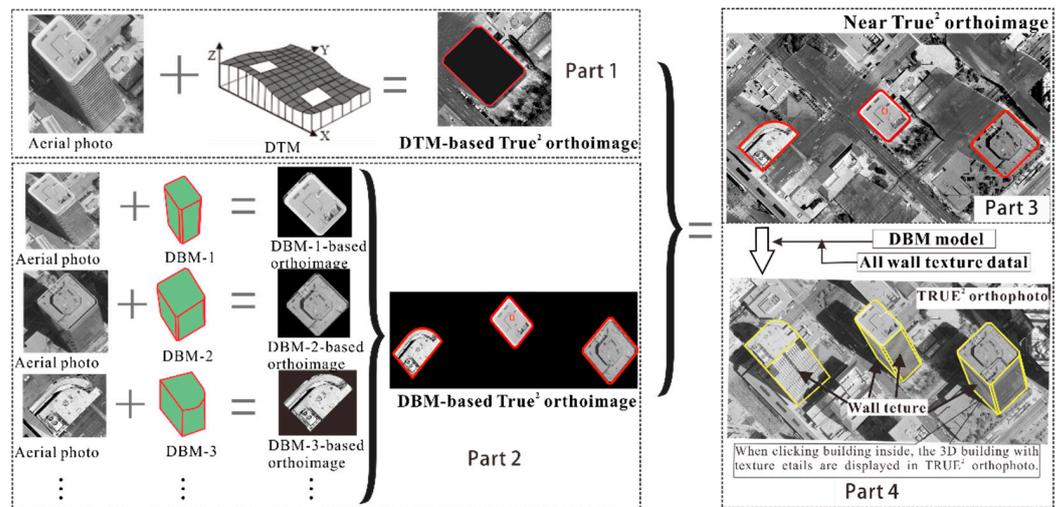
**Table 5.** The building wall texture index in Figure 7b.

WallID	Wall Pixel Index	Texture Index	Hight
$a_2b_2$	1, 2, 3, 4, 5, 6, 7, 8, 9	809	46.4
$b_2c_2$	10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22	810	46.4
$c_2d_2$	23, 24, 25, 26, 27, 28, 29, 30	811	46.4
$d_2a_2$	31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47	812	46.4

### 3.2. Generation for DBM-Based Multiple-Building T<sup>2</sup>OMs

There are usually many buildings in a city. Therefore, the next step is to generate the multiple-building T<sup>2</sup>OM on the basis of the generation of a single-building T<sup>2</sup>OM. Part 2 in Figure 8 shows the process of generation for multiple-building T<sup>2</sup>OMs, in which each building is assigned a unique identification (BuildingID) and the building ID is used to control the display and hiding of each building (see Table 6). In addition, the information of a single-building model is divided into top surface, elevation and bottom surface. Additionally, each roof and facade are assigned a separate identity (RoofID, WallID) (see Tables 7 and 8). The face IDs are associated with the building top surface table and building wall table, and each face point ID (Points) is recorded in the roof table and wall table, and the point IDs are associated with the building corner point information table (see Table 9). The corner point information is expressed using multiple horizontal projection polygons, and each corner point is provided with 3D vector information in the corner point information table. Building textures are divided into top surface textures and wall textures

in the table. In the building, the top surface table does not need to identify the top surface texture identity; its information can be associated with the 2D T<sup>2</sup>OM multiple-building data table in having the building identity (BuildingID) from which the top surface texture value and 3D coordinate information value of the building is obtained, and the real building top surface texture is obtained by rendering a single pixel at a single vector point in turn. For the wall texture of the building, the texture ID of the wall is associated with the texture data in the database. The texture name (TextureName), the address of the file uploaded in the computer (FileAddress), the data of the texture saved in binary (Binary), the format of the texture (Format), the size of the texture (Size) recorded in bytes (Byte) in the database, and the date when the texture was saved (Date) are recorded (see Table 10). Parts 1, 3 and 4 in Figure 8 are described in detail in Sections 3.3 and 3.4.



**Figure 8.** DBM-based generation for multiple-building T<sup>2</sup>OM and merged DTM-based and DBM-based T<sup>2</sup>OM.

**Table 6.** Building relationships.

BuildingID	Type	RoofID	WallID	Properties	Others
B <sub>1</sub>	Volume	R <sub>11</sub>	W <sub>11</sub> , W <sub>12</sub> , W <sub>13</sub> ...	Brick structure	...
B <sub>2</sub>	Volume	R <sub>12</sub>	W <sub>21</sub> , W <sub>22</sub> , W <sub>23</sub> ...	reinforced concrete structure	...
B <sub>3</sub>	Volume	R <sub>13</sub>	W <sub>31</sub> , W <sub>32</sub> , W <sub>33</sub> ...	steel structure	...
...	...	...	...	...	...

**Table 7.** The relationship of building roof textures.

RoofID	Type	TextureID	PointID	Others
R <sub>11</sub>	Polygon	T <sup>R</sup> <sub>11</sub>	P <sup>R</sup> <sub>11</sub> , P <sup>R</sup> <sub>12</sub> , P <sup>R</sup> <sub>13</sub> ...	...
R <sub>12</sub>	Polygon	T <sup>R</sup> <sub>12</sub>	P <sup>R</sup> <sub>21</sub> , P <sup>R</sup> <sub>22</sub> , P <sup>R</sup> <sub>23</sub> ...	...
R <sub>13</sub>	Polygon	T <sup>R</sup> <sub>13</sub>	P <sup>R</sup> <sub>31</sub> , P <sup>R</sup> <sub>32</sub> , P <sup>R</sup> <sub>33</sub> ...	...
...	...	...	...	...

**Table 8.** The relationship of building wall textures.

WallID	Type	TextureID	PointID	Others
W <sub>11</sub>	Polygon	T <sup>W</sup> <sub>11</sub>	P <sup>W</sup> <sub>11</sub> , P <sup>W</sup> <sub>12</sub> , P <sup>W</sup> <sub>13</sub> ...	...
W <sub>12</sub>	Polygon	T <sup>W</sup> <sub>12</sub>	P <sup>W</sup> <sub>21</sub> , P <sup>W</sup> <sub>22</sub> , P <sup>W</sup> <sub>23</sub> ...	...
W <sub>13</sub>	Polygon	T <sup>W</sup> <sub>13</sub>	P <sup>W</sup> <sub>31</sub> , P <sup>W</sup> <sub>32</sub> , P <sup>W</sup> <sub>33</sub> ...	...
...	...	...	...	...

**Table 9.** Building vertex relationships.

PointID	X, Y, Z Coord.			Pixel Coord.		Others
P <sup>W</sup> <sub>11</sub>	X <sup>W</sup> <sub>11</sub>	Y <sup>W</sup> <sub>11</sub>	Z <sup>W</sup> <sub>11</sub>	I <sup>W</sup> <sub>11</sub>	J <sup>W</sup> <sub>11</sub>	...
P <sup>W</sup> <sub>12</sub>	X <sup>W</sup> <sub>12</sub>	Y <sup>W</sup> <sub>12</sub>	Z <sup>W</sup> <sub>12</sub>	I <sup>W</sup> <sub>12</sub>	J <sup>W</sup> <sub>12</sub>	...
P <sup>W</sup> <sub>13</sub>	X <sup>W</sup> <sub>13</sub>	Y <sup>W</sup> <sub>13</sub>	Z <sup>W</sup> <sub>13</sub>	I <sup>W</sup> <sub>13</sub>	J <sup>W</sup> <sub>13</sub>	...
...	...	...	...	...	...	...
P <sup>R</sup> <sub>11</sub>	X <sup>R</sup> <sub>11</sub>	Y <sup>R</sup> <sub>11</sub>	Z <sup>R</sup> <sub>11</sub>	I <sup>R</sup> <sub>11</sub>	J <sup>R</sup> <sub>11</sub>	...
P <sup>R</sup> <sub>12</sub>	X <sup>R</sup> <sub>12</sub>	Y <sup>R</sup> <sub>12</sub>	Z <sup>R</sup> <sub>12</sub>	I <sup>R</sup> <sub>12</sub>	J <sup>R</sup> <sub>12</sub>	...
P <sup>R</sup> <sub>13</sub>	X <sup>R</sup> <sub>13</sub>	Y <sup>R</sup> <sub>13</sub>	Z <sup>R</sup> <sub>13</sub>	I <sup>R</sup> <sub>13</sub>	J <sup>R</sup> <sub>13</sub>	...
...	...	...	...	...	...	...

**Table 10.** Data texture table.

TextureD	TextureName	FileAddress	Date	Fomat	Others
T <sup>W</sup> <sub>11</sub>	WTN11	WFA11	WD11	WF11	...
T <sup>W</sup> <sub>12</sub>	WTN12	WFA12	WD12	WF12	...
T <sup>W</sup> <sub>13</sub>	WTN13	WFA13	WD13	WF13	...
...	...	...	...	...	...
T <sup>R</sup> <sub>11</sub>	PTN11	PFA11	PD11	PF11	...
T <sup>R</sup> <sub>12</sub>	PTN12	PFA12	PD12	PF12	...
T <sup>R</sup> <sub>13</sub>	PTN13	PFA13	PD13	PF13	...
...	...	...	...	...	...

3.3. Generation of DTM-Based T<sup>2</sup>OM

Part 1 in Figure 8 is DTM-based T<sup>2</sup>OM generation, which orthorectifies the displacement caused by terrestrial elevation, i.e., orthorectifies the terrains into an upright position in a given map coordination. Therefore, the digital differential orthorectification method is applied for this purpose. The details of this method can be found in [8,37]. Similarly, the given DTM data, data structure and data storage for each pixel in the DTM-based T<sup>2</sup>OM are similar to that of the DBM-based T<sup>2</sup>OM, but the building ID is assigned as “none”.

3.4. Merging DTM- and DBM-Based T<sup>2</sup>OMs

In view of the different structures of the DBM- and DTM-based T<sup>2</sup>OMs, an entire T<sup>2</sup>OM generation needs merging algorithms (see Figure 8). To do this, the logic operation <or> is performed with the superpixel ID of 0 or non-zero. In order to eliminate possible boundary confusion in the merging process, the following judgment conditions are executed: with the DTM-based T<sup>2</sup>OM as the base map (see Part 1 in Figure 8), when the same grid number appears in the DTM- and DBM-based T<sup>2</sup>OMs at the same time, only the DBM-based T<sup>2</sup>OM is retained (see Part 2 in Figure 8). This is because the building area determined by the horizontal projection polygon is not a regular rectangle or divided along the grid direction. Therefore, there are actually more grid elements located at the building boundary than at the real building boundary. Retaining the grid of the DBM-based T<sup>2</sup>OM will ensure the accuracy of the building location to the greatest extent. This is helpful for the 3D T<sup>2</sup>OM display. Merge the DTM-based and DBM-based true-squared orthophotos to obtain a near-true-squared image (see Part 3 in Figure 8). Linking the DBM model and wall textures to the building data enables the ability to display building façade textures and three-dimensional (3D) coordinates (XYZ) measurable at any point (see Part 4 in Figure 8).

4. Experiments and Analysis

Figure 9 is a flowchart for T<sup>2</sup>OM generation, divided into five parts. Part 1 shows two experimental datasets (high-resolution images, control points, orientation parameters, DBM and DTM) from Denver, CO, USA, and Nanning, China, which are used as input data. Part 2 shows DBM-based T<sup>2</sup>OM generation, which consists of orthorectifying both building roof and building facades, with which “superpixel” is used for the storage of

building texture, building ID, etc. Part 3 shows DTM-based T<sup>2</sup>OM generation, which is for orthorectification of gentle and continuously elevated hilly areas. Part 4 shows the merging of DBM-based and DTM-based T<sup>2</sup>OMs as the output data, i.e., T<sup>2</sup>OM. Part 5 shows the accuracy evaluation for the generated T<sup>2</sup>OM using ground control points.

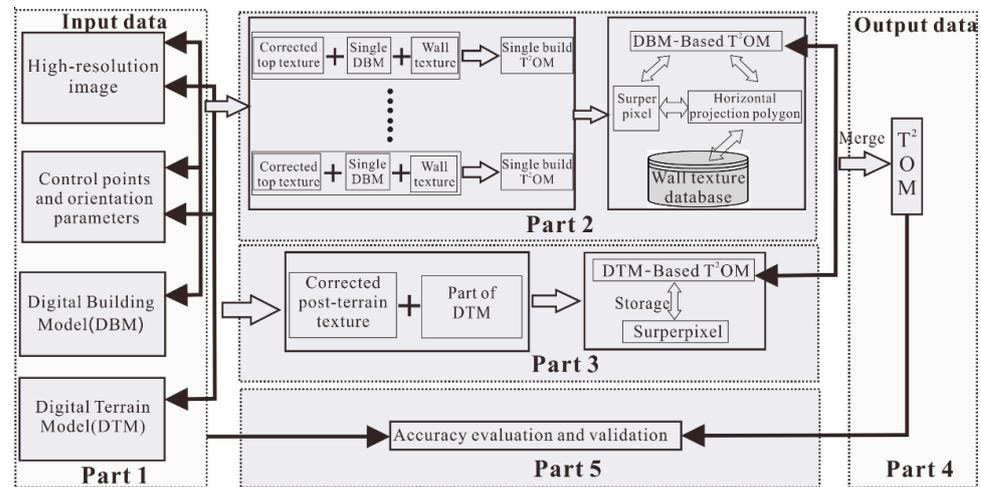


Figure 9. The proposed flowchart for T<sup>2</sup>OM generation.

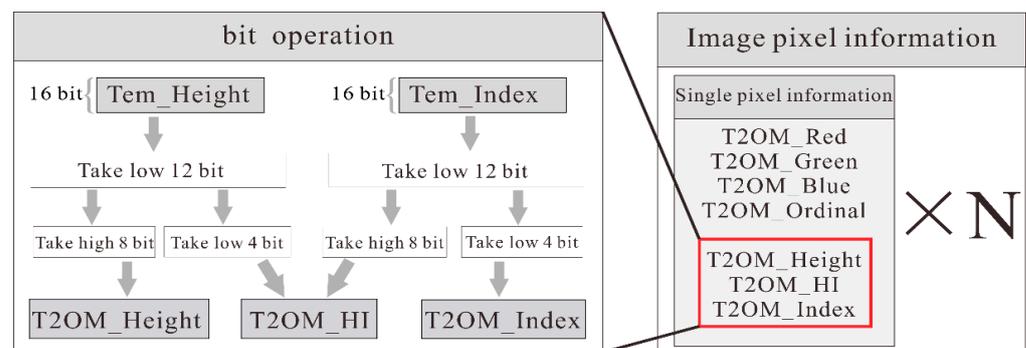
#### 4.1. Metadata of T<sup>2</sup>OM

This experiment was implemented using the programming language C++. To store the T<sup>2</sup>OM in binary form, a file format *fus* needed to be designed to store superpixels due to the restriction of traditional bit-storing in computers. Table 11 shows the entire file format, which consists of a file flag block, image header file information, and image pixel information. These three parts were written to the *fus* file in binary form.

Table 11. The file format of the T<sup>2</sup>OM *fus*.

File Section	Properties	Description
File flag block	m_FileProperty	Identifier “fus” (char type)
	m_Version	Version number (int type)
Image header information	m_UpleftCoordinateX	Image coordinate lower right X value (double type, units: meters)
	m_UpleftCoordinateY	Image coordinate lower right Y value (double type, units: meters)
	m_TMaxZ	The highest point in the DTM file (double type, in meters)
	m_TMinZ	The lowest point in the DTM file (double type, in meters)
	m_BMaxZ	Maximum building height in DBM (double type in meters)
	m_BMinZ	Minimum building height in DBM (double type in meters)
	m_IntervalX	Unit interval in X-axis direction (double type in meters)
	m_IntervalY	Unit interval in Y-axis direction (double type, units: meters)
	m_FileHigh	Image height (int type)
	m_FileWidth	Image width (int type)
Image Pixels Information	Z_Resolution	Topographic data unit elevation level (double type, units: meters)
	Z_Bresolution	Building data unit elevation level (type double, in meters)
	Build_Num	Number of building objects elements (type int)
	T <sup>2</sup> OM_Grey	Pixel grey component (unsigned char type)
	T <sup>2</sup> OM_Ordinal	Subdivision grid order (unsigned char type)
	T <sup>2</sup> OM_Height	Elevation level high 8 bits (unsigned char type)
	T <sup>2</sup> OM_HI	Elevation level low 4 bits, logo high 4 bits (unsigned char type)
	T <sup>2</sup> OM_Index	Marker data low 8 bits (unsigned char type)

In the superpixel data structure, information about the elevation level  $H$  and building identification  $ID$  need 12 bits of memory. However, there is no data type of this size in the computer. Therefore, 12 bits of information are saved by the bit operation. As shown in Figure 10, this method first opens up two unsigned short types of data (16 bit), TEM\_Height and TEM\_Index, and uses them to record  $H$  and  $ID$ , respectively. Then, the data are put into variables ( $T^2OM\_height$ ,  $T^2OM\_hi$  and  $T^2OM\_index$ ) through the bit operation. Finally, the information in the superpixel is combined, as shown on the right side of the figure.



**Figure 10.** Saving pixel information by bit manipulation.

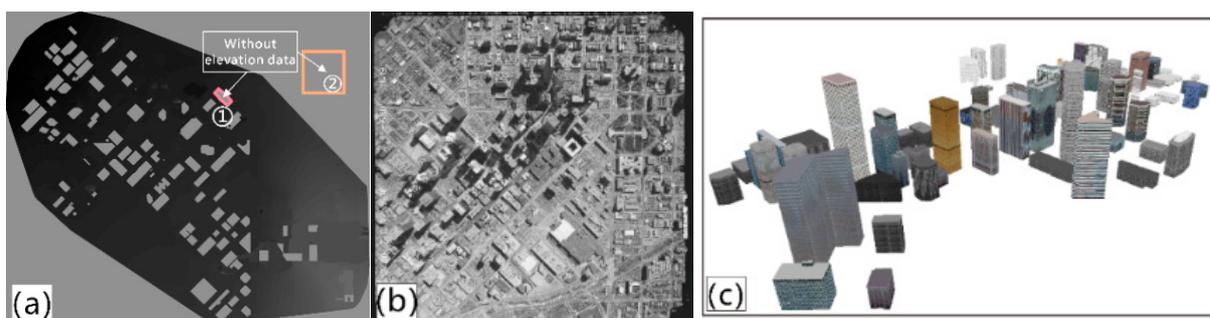
#### 4.2. $T^2OM$ Generation

In this section, this paper uses Dataset 1 and Dataset 2 to describe the generation process of generating  $T^2OM$  and to verify the feasibility of the method in this paper.

##### 4.2.1. Experimental Result with Dataset 1

The experimental Dataset 1 includes digital surface model data, aerial imagery data, and digital building model data. A brief description is given as follows:

- (1) *DTM data*: Figure 11a shows DTM data from Denver, CO, USA, which is represented as a height–depth map, where the darker the color is, the lower the height, and vice versa, because the topography of the city is relatively flat. Thus, the elevations shown on the ground are relatively similar (the colors shown are similar). The accuracy of plane surface coordinates and vertical coordinates are about 0.1 m and 0.2 m, respectively. The horizontal datum is GRS 1980, and the vertical datum is NAD83.
- (2) *Aerial Image data*: Figure 11b shows the original aerial image acquired using the RC30 aerial camera lens in Denver. The flight altitude in Denver is 1650 m higher than the mean ground elevation of the imaged area. Aerial photographs were initially recorded on film and then scanned into digital format at a pixel resolution of 25  $\mu\text{m}$ .
- (3) *DBM data*: Figure 11c shows Denver DBM data, and these buildings with a ground resolution of about 25.4 cm per pixel were identified. Each building model contains building corner point information and elevation texture information.

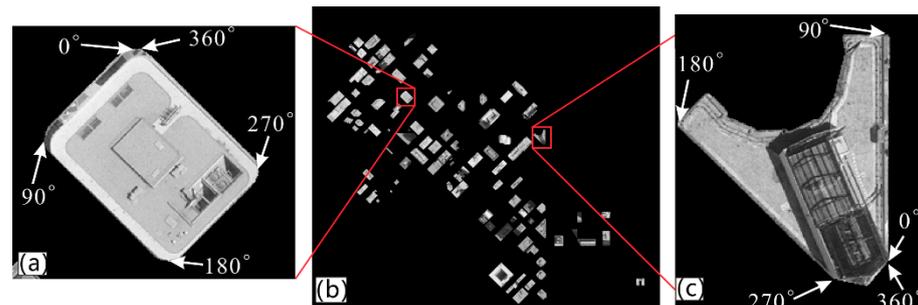


**Figure 11.** The experimental Dataset 1. (a) DTM data, where area ① and ② represent areas without elevation data; (b) Aerial image data; (c) DBM data.

### Step 1. Generation of DBM-based T<sup>2</sup>OM

#### (1) DBM-based building roof orthorectification

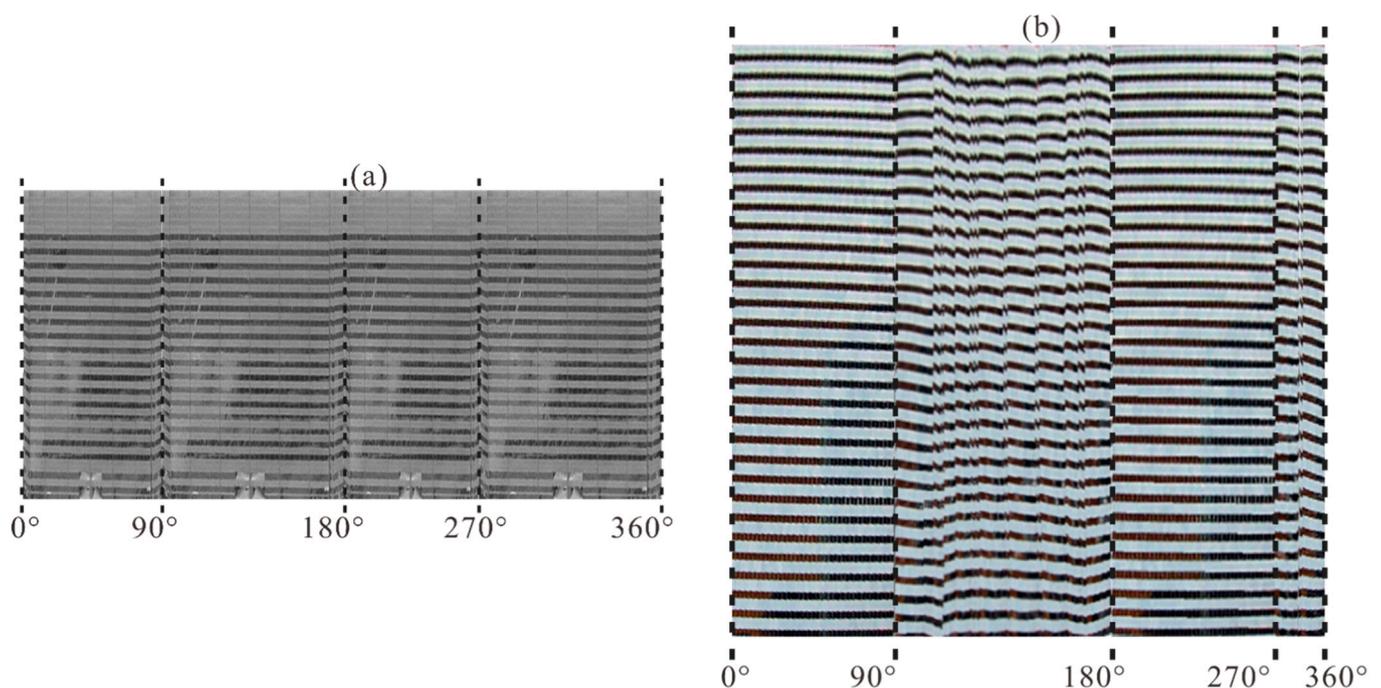
DBM-based roof orthorectification corrects only the displacements caused by the buildings and does not take into account the displacements caused by the terrain. Therefore, the generated DBM-based T<sup>2</sup>OM (see Figure 12) only corrects the building texture and not the texture of the terrain area, so the texture at the terrain is black (background value). Where (a) and (c) are two buildings in the T<sup>2</sup>OM, it can be seen that the roof textures of the buildings are obtained accurately.



**Figure 12.** Generation results of DBM-based T<sup>2</sup>OM on dataset 1. (b) DBM-based T<sup>2</sup>OM; (a,c) are enlarged windows of the two regions.

#### (2) DBM-based building façade orthorectification

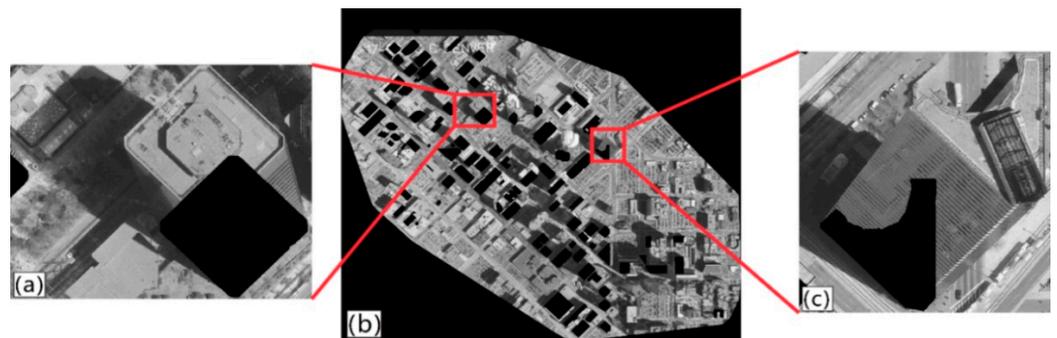
In order to obtain the façade textures and 3D coordinates of the buildings, the orthorectification of the building façade textures was also performed in T<sup>2</sup>OM. The building façade textures for Dataset 1 were selected from the existing texture library and the same textures were used in all four directions, and the results of façade texture correction are shown in Figure 13. Figure 13a,b show the results of the façade texture correction for (a) and (c) in Figure 12, respectively, and the angles marked in Figure 13 are consistent with those in Figure 12.



**Figure 13.** Façade texture orthorectification results. (a,b) are the orthorectification results of the building Façade texture in a and c in Figure 12.

### Step 2. Generation of DTM-based T<sup>2</sup>OM

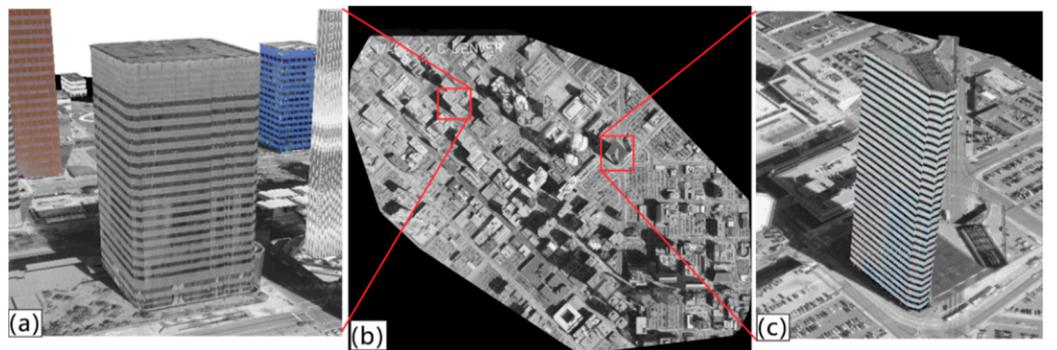
The correction of buildings needs to be followed by the orthorectification of non-building areas. The DTM-based differential correction of non-buildings is performed to obtain the corrected texture of the image. At the same time, superpixels are generated by overlaying each data information, and finally the DTM-based T<sup>2</sup>OM is obtained (see Figure 14). Because this part only corrects the terrain texture and not the building area texture, the texture at the building is black (background value). The DTM-based T<sup>2</sup>OM also has ghosts and shadows in the texture area, because Dataset 1 lacks complementary images, so occlusion detection, compensation for textures and shadow detection and compensation operations are not performed in this step.



**Figure 14.** Generation results of DTM-based T<sup>2</sup>OM on dataset 1. (b) DTM-based T<sup>2</sup>OM; (a,c) are enlarged windows of the two regions.

### Step 3. Merging DTM- and DBM-based T<sup>2</sup>OMs

Finally, the merging of the DBM with the DTM's T<sup>2</sup>OM yields the result shown in Figure 15b, which fills the superpixels at the roof texture exactly to the area where the original background value is black. Figure 15a,c show enlarged views of two of the buildings.

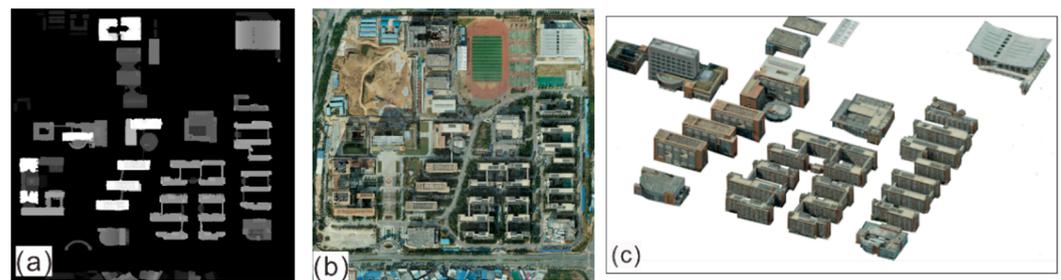


**Figure 15.** Merging DTM- and DBM-based T<sup>2</sup>OM on dataset 1. (b) T<sup>2</sup>OM; (a,c) are enlarged windows of the two regions.

#### 4.2.2. Experimental Result with Dataset 2

The experimental Dataset 2 includes digital surface model data, aerial imagery data, and digital building model data. A brief description is given as follows.

- (1) *DTM data*: Figure 16a shows DTM data from Nanning, China, represented as a height–depth map, where the darker the color is, the lower the height, and vice versa, because the topography of the city is relatively flat. Thus, the elevations shown on the ground are relatively similar (the colors shown are similar). The accuracy of the plane surface coordinates and vertical coordinates are about 0.1 m and 0.2 m, respectively. The horizontal datum is GRS 1980, and the vertical datum is NAD83.



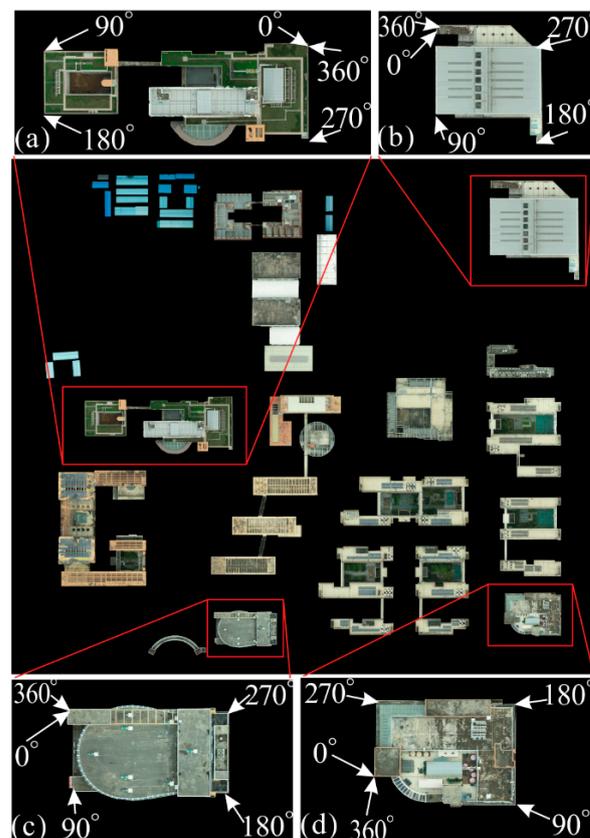
**Figure 16.** The experimental Dataset 2. (a) DTM data; (b) Aerial image data; (c) DBM data.

- (2) *Aerial Image data:* Figure 16b shows the original aerial image acquired using the CMOS lens in Nanning. The flight altitude in Nanning is 200 m higher than the average ground elevation of the imaging area.
- (3) *DBM data:* Figure 16c shows Nanning DBM data, and buildings with a ground resolution of about 25.4 cm per pixel were identified. Each building model contains building corner point information and elevation texture information.

### Step 1. Generation of DBM-based T<sup>2</sup>OM

#### (1) DBM-based building roof orthorectification

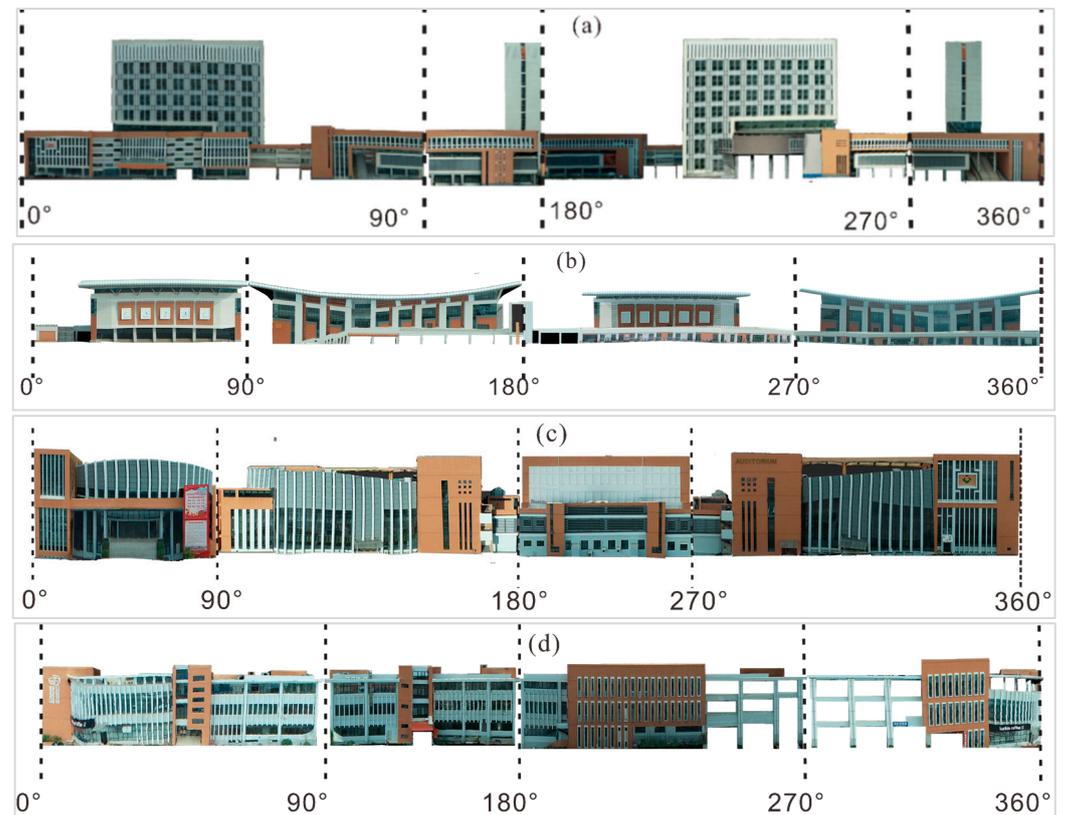
The DBM-based roof orthorectification corrects only the displacements caused by the buildings and does not take into account the displacements caused by the terrain. Therefore, the generated DBM-based T<sup>2</sup>OM (see Figure 17) only corrects the building texture and not the texture of the terrain area, so the texture of the terrain is black (background value). Four buildings of the T<sup>2</sup>OM are denoted as (a–d), and the roof textures of the buildings are obtained accurately.



**Figure 17.** Generation results of DBM-based T<sup>2</sup>OM on dataset 2. (a–d) are enlarged windows of the four regions.

## (2) DBM-based building façade orthorectification

In order to obtain the building facade texture and 3D coordinates, the building facade texture is also orthorectified in T<sup>2</sup>OM. The orthorectification results are shown in Figure 18. Figure 18a–d show the results of the façade texture correction for (a–d) in Figure 17, respectively, and the angles marked in Figure 18 are consistent with those in Figure 17.



**Figure 18.** Façade texture orthorectification results. (a–d) are the orthorectification results of the building Façade texture in (a–d) in Figure 17.

### Step 2. Generation of DTM-based T<sup>2</sup>OM

The correction of buildings needs to be followed by the orthorectification of non-building areas. The DTM-based differential correction of non-buildings is performed to obtain the corrected texture of the image. At the same time, superpixels are generated by overlaying each data information, and finally the DTM-based T<sup>2</sup>OM is obtained (see Figure 19). Because this part only corrects the terrain texture and not the building area texture, the texture of the building is black (background value). The DTM-based T<sup>2</sup>OM also has ghosting and shadows in the texture region, and the occlusion detection and compensation of textures and shadow detection and compensation operations in this paper are adopted from [9].

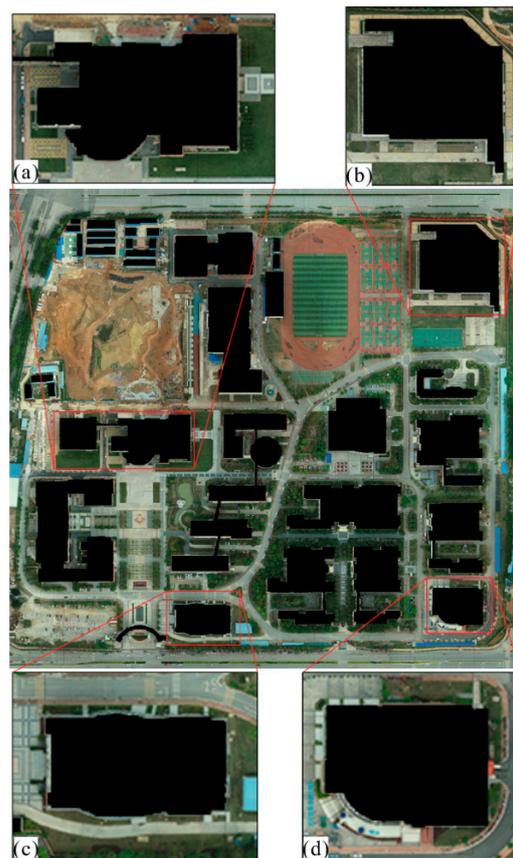


Figure 19. Generation results of DTM-based T<sup>2</sup>OM; (a–d) are enlarged windows of the four regions.

### Step 3. Merging DTM- and DBM-based T<sup>2</sup>OMs

Finally, the merging of the DBM with the DTM's T<sup>2</sup>OM yields the result shown in Figure 20, which fills the superpixels at the roof texture exactly to the area where the original background value is black. Figure 20a–d show enlarged views of four of the buildings.

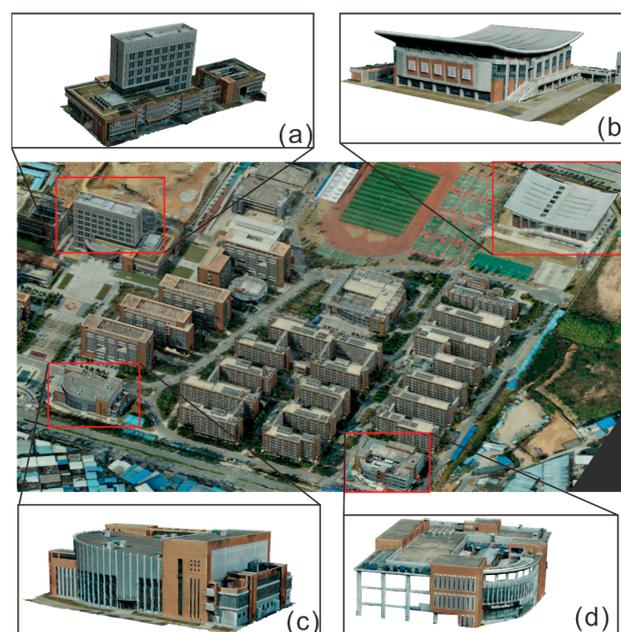


Figure 20. Merging DTM- and DBM-based T<sup>2</sup>OMs on dataset 2; (a–d) are an enlarged windows of the four regions.

### 4.2.3. T<sup>2</sup>OM 3D Measurement

With the 3D measurement function, the elevation of any point in the scene can be obtained. Figure 21 shows elevation information from a point on the selected façade. Figure 22 shows the color information, true 3D coordinates, and attribute information of each pixel in the acquired point. “Selective Hide” can also be used to display individual buildings, as shown in Figure 23. In addition, based on the fact that the horizontal projection polygon of a building can completely record building information, we can obtain complete information about corner points and façade through the “3D Building Information Display” function, as shown in Figure 24. The “Building Distance Measurement” function calculates the minimum and maximum distances between two buildings by using the building corner information, as shown in Figure 25.

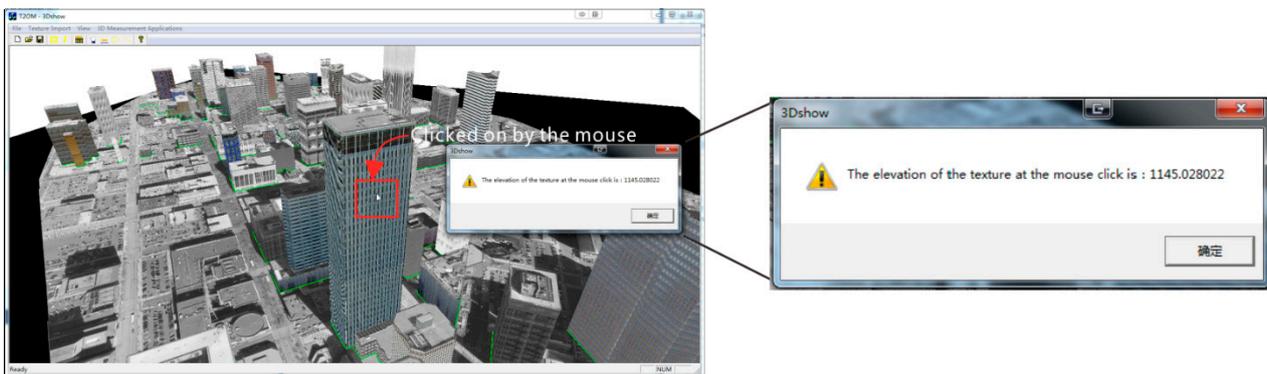


Figure 21. Measuring the elevation of any point on the façade.

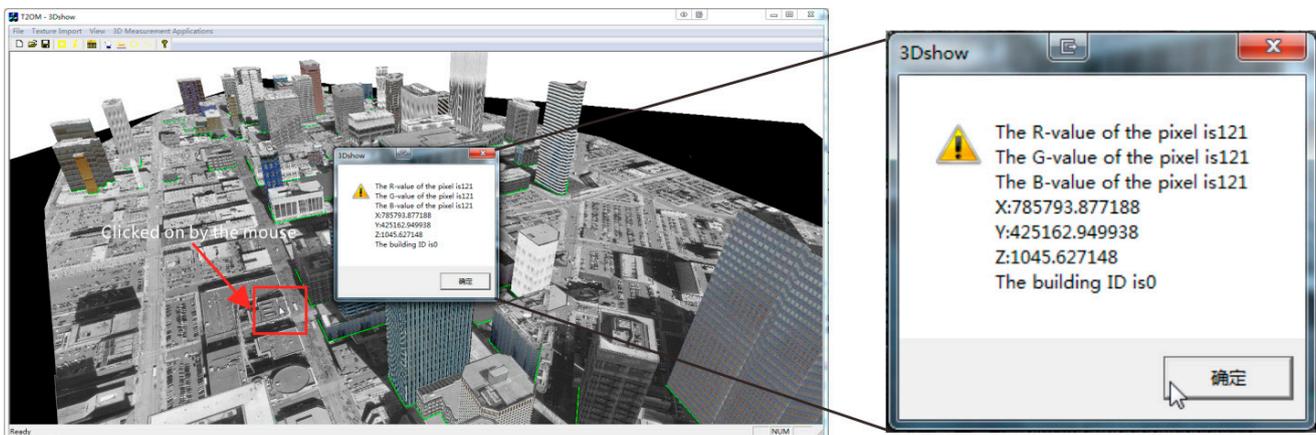


Figure 22. View of the superpixel information.

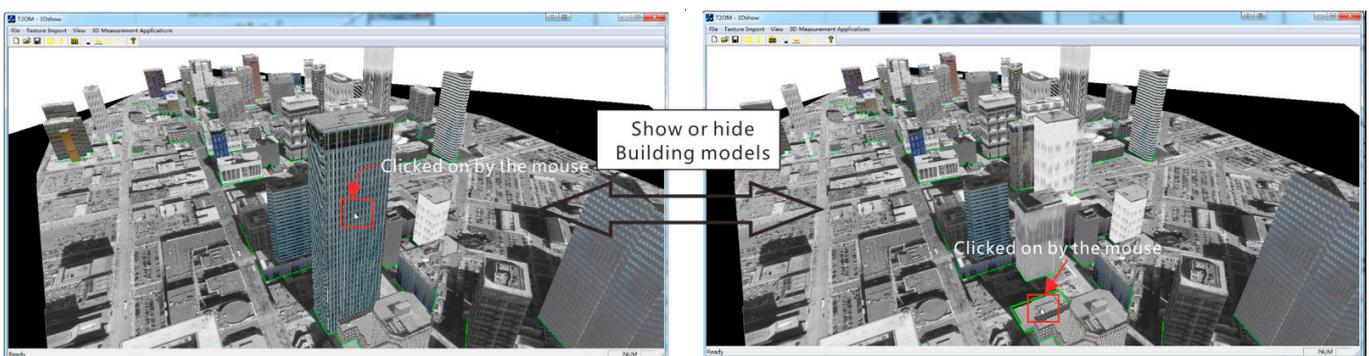


Figure 23. Selectively showing or hiding information.

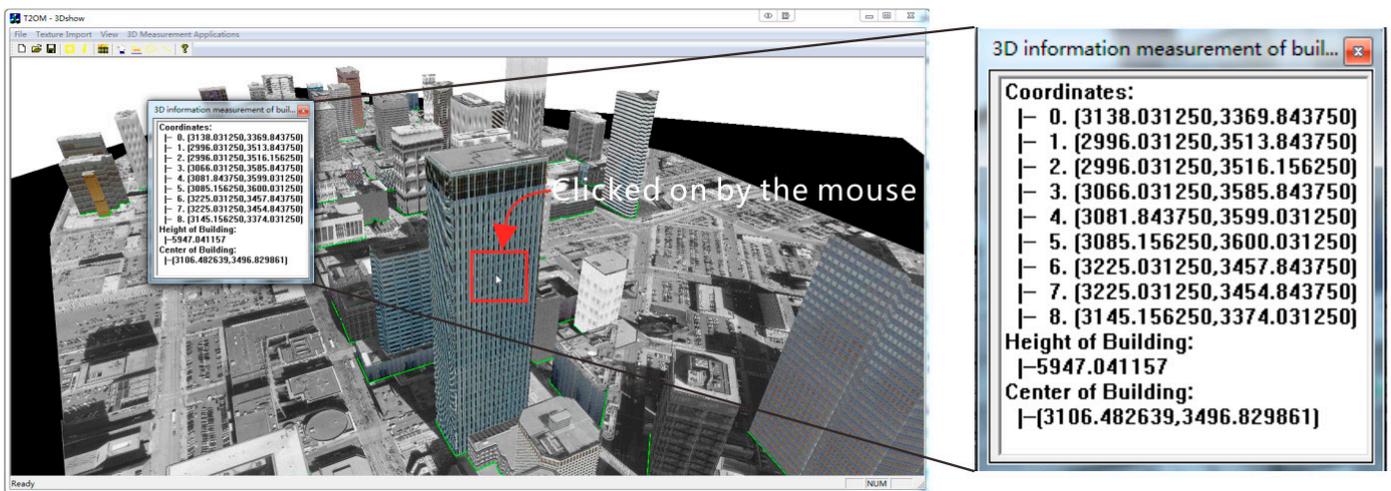


Figure 24. View information about the building.

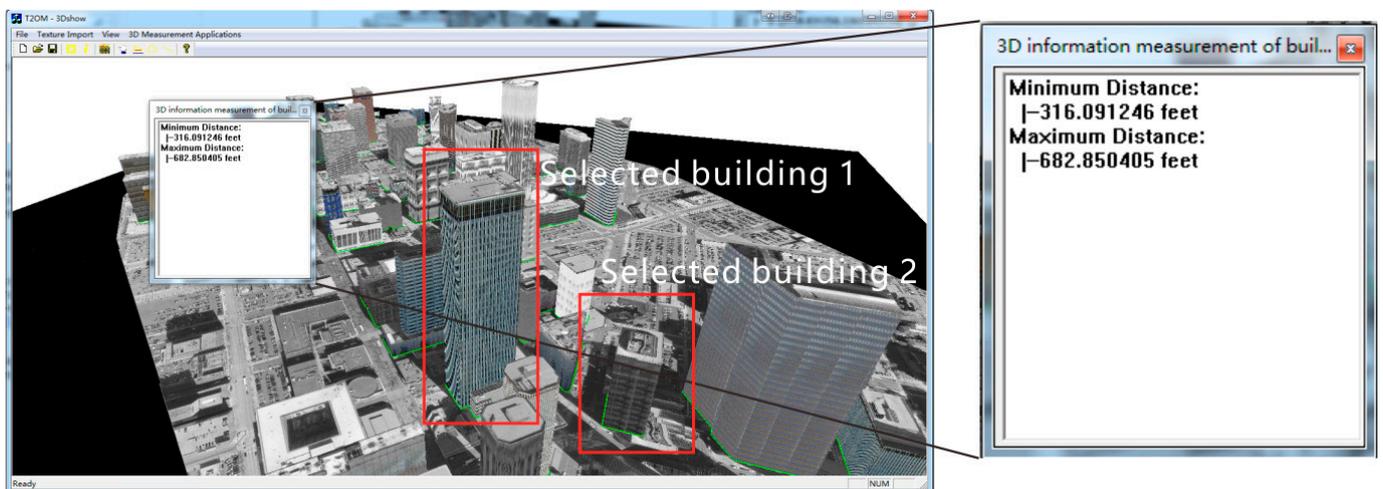
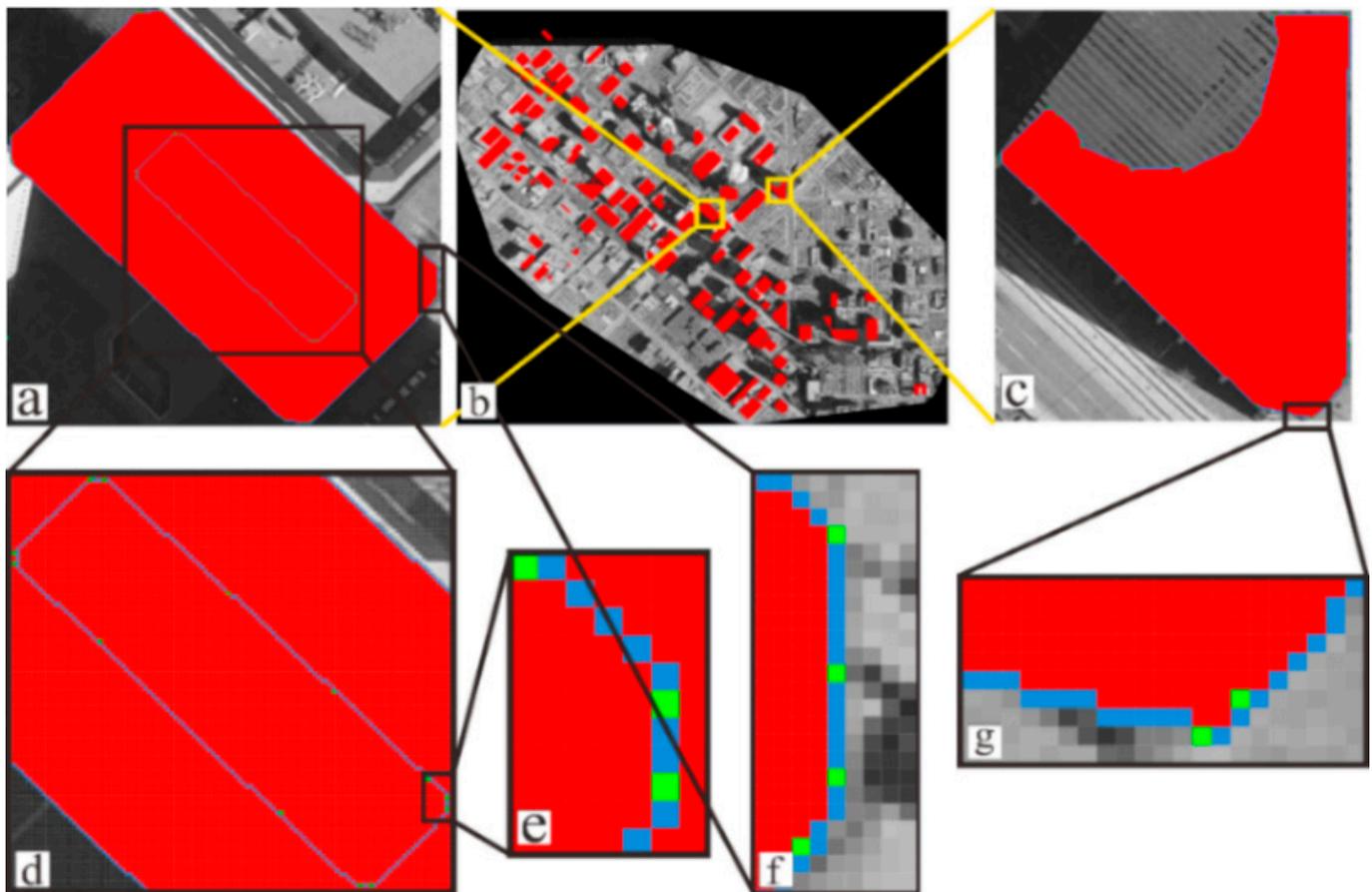


Figure 25. Measuring the maximum and minimum distances between two buildings.

#### 4.3. Accuracy Evaluation and Analysis

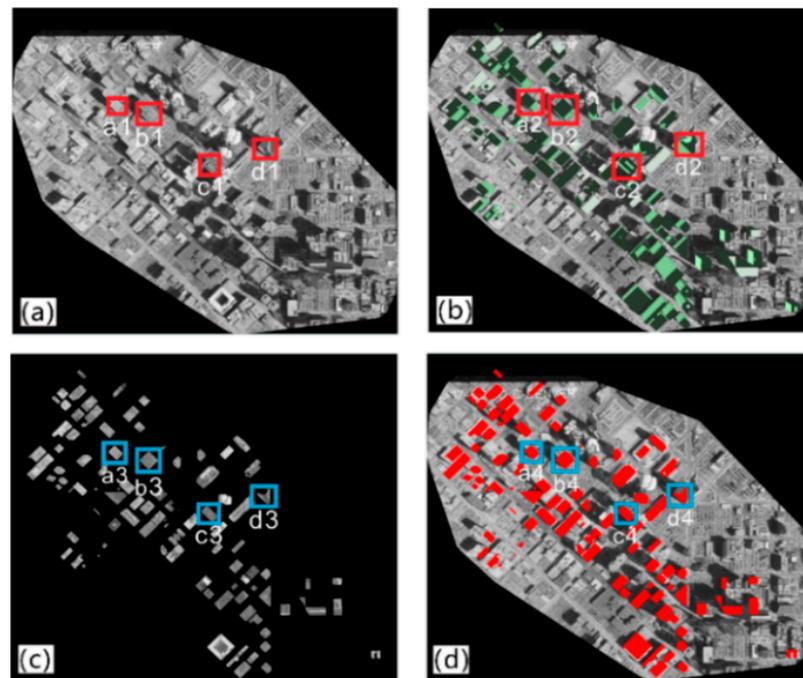
In our method, T<sup>2</sup>OM generation is obtained by merging the DTM- and DBM-based T<sup>2</sup>OMs. There can be errors in multiple steps of the process. In addition, because the superpixel uses an elevation series instead of the double-form true elevation value, errors exist in the recording height of the superpixel. Based on the generated T<sup>2</sup>OM, this section evaluates the matching accuracy between the building horizontal projection polygons.

In order to ensure the accuracy of building location information in T<sup>2</sup>OM, it is necessary to evaluate the construction accuracy of the horizontal projection polygon. In Figure 26, the grid detected as the building edge is represented in blue. Building corners are represented in green. The horizontally projected polygon is displayed in red. As can be seen from the enlarged Figure 26e–g, the extracted blue edge is completely consistent with the green building corner. This proves that the accuracy of the building horizontal projection polygon construction is sufficient.



**Figure 26.** Acquisition of horizontal projection polygon corner points: (a) building model consisting of two voxels; (b) graph of the corner point detection results; (c) building model consisting of a single voxel; (d–f) local enlargement of (a); (g) local enlargement of (c).

Four buildings identified in the T<sup>2</sup>OM data were randomly selected to evaluate the accuracy of information recording in T<sup>2</sup>OM, as shown in Figure 27. Firstly, roof corner points of each building above were extracted from the DBM. Then, the coordinates of the corner points in the DBM were compared with the grid number, subdivision, and elevation series recorded in the superpixels. Table 12 presents the 3D coordinates before and after coding, where  $x_{ori}$ ,  $y_{ori}$ , and  $z_{ori}$  are the 3D coordinates before coding;  $S(r,c)$  and  $H$  are the subdivision and elevation series, respectively; and  $X_{bc}$ ,  $Y_{bc}$ , and  $Z_{bc}$  are the 3D coordinates calculated using superpixels. Through calculations, the average errors of the X, Y and Z coordinate components were determined to be 0.017, 0.025, and 0.09 m, respectively. Compared with the resolution of the original data in plane coordinates of 0.1 m and the elevation of 0.2 m, it can be concluded that the coding method recorded by superpixels can greatly reduce the level of error generated in the process of converting vector data into grid data.



**Figure 27.** Analysis of T<sup>2</sup>OM generation results. (a) output orthophoto, (b) orthophoto superimposed on the DBM, (c) extracted roof texture, (d) red area used to represent the superpixels generated with the DBM. Where a1-a4 are four regions with the same geographical position in (a–d); b1–b4 are four regions with the same geographical position in (a–d); c1–c4 are four regions with the same geographical position in (a–d); d1–d4 are four regions with the same geographical position in (a–d).

**Table 12.** 3D coordinate values before and after encoding.

Building	$X_{ori}$	$Y_{ori}$	$Z_{ori}$	$id(r,c)$	$h$	$X_{bc}$	$Y_{bc}$	$Z_{bc}$
a	1286.850	1306.000	5551.700	14	1973	1286.840	1306.030	5551.600
	1346.850	1245.000	5551.700	14	1973	1346.840	1245.030	5551.600
	1355.000	1237.975	5551.700	241	1973	1355.030	1237.970	5551.600
	1356.000	1237.850	5551.700	209	1973	1356.030	1237.840	5551.600
b	1931.937	1424.000	5533.400	15	1876	1931.910	1424.030	5533.330
	1932.150	1440.000	5533.400	3	1876	1932.160	1440.030	5533.330
	1910.850	1462.000	5533.400	14	1876	1910.840	1462.030	5533.330
	1906.000	1464.150	5533.400	33	1876	1906.030	1464.160	5533.330
c	2990.850	1244.000	5462.500	14	1499	2990.840	1244.030	5462.330
	2912.850	1151.000	5462.500	14	1499	2912.840	1151.030	5462.330
	2911.850	1148.000	5462.500	14	1499	2911.840	1148.030	5462.330
	2964.850	1093.000	5462.500	14	1499	2964.840	1093.030	5462.330
d	3339.000	1882.850	5755.800	209	3057	3339.030	1882.840	5755.760
	3341.096	1883.000	5755.800	2	3057	3341.090	1883.030	5755.760
	3344.850	1881.000	5755.800	14	3057	3344.840	1881.030	5755.760
	3356.054	1864.000	5755.800	1	3057	3356.030	1864.030	5755.760

#### 4.4. Discussion

From the above two sets of experimental results, it is concluded that our proposed T<sup>2</sup>OM can realize the switching display of two-dimensional flat TOMs and three-dimensional buildings, and use superpixels to save three-dimensional information so that the accuracy of three-dimensional measurement can be controlled within 0.0625 m.

In addition, it is feasible to expand the traditional pixel storage method to increase the amount of information expressed by a grid image. In our method, first, a large number

of heterogeneous data are unified to achieve centralized management. This makes the reconstruction and display of a 3D model easier, because superpixels accurately store each location with elevation information. However, the proposed method still has defects. The main problem is that storage of the T<sup>2</sup>OM requires twice as much memory space as that of the TOM. This makes it difficult to store and transfer data, especially across large-scale urban areas. A solution is to compress the bit width of storage components, such as  $S$ ,  $H$  and  $ID$ , by a statistical method. For example, in the compression of  $H$ , the height change curve for the whole study area can be counted, and the elevation series can be reduced from 4096 to 256. In addition, in areas without buildings, the removal of  $ID$  space to reduce memory consumption can be considered. Another defect that cannot be ignored is that the description of the geometric structure of building facades is not refined enough. On facades, there are balconies protruding from walls and windows recessed into walls. The geometric correction and line extraction of these facade structures is still difficult at present [56].

## 5. Conclusions

In light of the problem that traditional DOMs/TOMs only provide building roofs' 2D ( $X, Y$ ) attributes and gray information, and cannot provide 3D information or building facade textures at all, this paper proposes the generation of T<sup>2</sup>OM, which is radically different from the traditional generation of DOMs/TOMs, since the T<sup>2</sup>OM is able to provide three-dimensional (3D) and detailed textures of building roofs and facades.

The major innovation of this manuscript lies in the new method for the generation of T<sup>2</sup>OM, in which a data structure that can simultaneously store the 2D and 3D information of a building, building roof and building façade is developed. The proposed superpixel data structure takes the grid as the basic unit and successfully integrates a variety of data types by expanding the pixel storage space. The application of subdivision  $S$  and elevation series  $H$  greatly improves the accuracy of the 3D model. The proposed superpixel model is capable of promoting the fusion of multi-source heterogeneous data, so that a single image of data can display both 2D plane information and the 3D real scene. Moreover, the superpixel model can be applied to facade texture images, so that the 3D measurement of any point in a scene can be achieved. These contributions are valuable for large-scale urban DOM generation and applications.

Two sets of experimental results demonstrate that the proposed generation method of T<sup>2</sup>OM can maintain the traditional DOM/TOM characteristics, i.e., provide 2D  $XY$  coordinates and displaced building texture, but also provide the 3D  $XYZ$  coordinates of buildings' roofs and facades. The accuracy of 3D measurement on a T<sup>2</sup>OM can achieve 0.025 m (0.3 pixel).

Nevertheless, the proposed method needs to be improved; for example, when the number of buildings in a city is large, if all of them are loaded into the memory according to the original texture data, the 3D display may occupy a large amount of memory and the refresh speed will be reduced. Therefore, memory loading according to the visible area and the compression of loaded memory are needed to reduce the memory occupancy and improve the refresh speed.

**Author Contributions:** Conceptualization, G.Z. and Q.W.; methodology, G.Z.; software, Y.H.; validation, Q.W., Y.H. and J.T.; formal analysis, H.L.; investigation, Y.W.; resources, Q.W.; data curation, Q.W.; writing—original draft preparation, Q.W.; writing—review and editing, G.Z.; visualization, Y.H.; supervision, Y.W.; project administration, G.Z.; funding acquisition, G.Z. and Q.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science of China (Project No. 41961065), Guangxi Innovative Development Grand Program (Project No. Guike AD19254002, GuikeAA18118038, and GuikeAA18242048); Guangxi Natural Science Foundation for Innovation Research Team (Project No. 2019GXNSFGA245001), the BaGui Scholars program of Guangxi (Guoqing Zhou), Innovation Project of Guangxi Graduate Education (Project No. YCBZ2021061) and Guangxi Key Laboratory of Spatial Information and Geomatics Program (Project No. 19-050-11-14).

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the reviewers for their constructive comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Federal Geographic Data Committee. *Fact Sheet: National Digital Geospatial Data Framework: A Status Report*; Federal Geographic Data Committee: Reston, VA, USA, July 1997; 37p.
2. Liu, Y.; Zheng, X.; Ai, G.; Zhang, Y.; Zuo, Y. Generating a High-Precision True Digital Orthophoto Map Based on UAV Images. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 333. [[CrossRef](#)]
3. Maitra, J.B. *The National Spatial Data Infrastructure in the United States: Standards; Metadata, Clearinghouse, and Data Access*; Federal Geographic Data Committee c/o US Geological Survey: Reston, VA, USA, 1998.
4. Yang, M.; Liu, J.; Zhang, Y.; Li, X. Design and Construction of Massive Digital Orthophoto Map Database in China. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 103–106. [[CrossRef](#)]
5. Zhou, G. *Onboard Processing for Satellite Remote Sensing Images*; CRC Press: Boca Raton, FL, USA, 2022; ISBN 978-10-32-329642.
6. Federal Geographic Data Committee. *Development of a National Digital Geospatial Data Framework*; Federal Geographic Data Committee: Reston, VA, USA, 1995. [[CrossRef](#)]
7. Jamil, A.; Bayram, B. Tree Species Extraction and Land Use/Cover Classification From High-Resolution Digital Orthophoto Maps. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *11*, 89–94. [[CrossRef](#)]
8. Zhou, G.; Schickler, W.; Thorpe, A.; Song, P.; Chen, W.; Song, C. True orthoimage generation in urban areas with very tall buildings. *Int. J. Remote Sens.* **2004**, *25*, 5163–5180. [[CrossRef](#)]
9. Zhou, G.; Chen, W.; Kelmelis, J.A.; Zhang, D. A comprehensive study on urban true orthorectification. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2138–2147. [[CrossRef](#)]
10. Amhar, F.; Jansa, J.; Ries, C. The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM. *Int. Arch. Photogramm. Remote Sens.* **1998**, *32*, 16–22.
11. Schickler, W.; Thorpe, A. Operational procedure for automatic true orthophoto generation. *Int. Arch. Photo-Grammetry Remote Sens.* **1998**, *32*, 527–532.
12. Di, K.; Jia, M.; Xin, X.; Wang, J.; Liu, B.; Li, J.; Xie, J.; Liu, Z.; Peng, M.; Yue, Z.; et al. High-Resolution Large-Area Digital Orthophoto Map Generation Using LROC NAC Images. *Photogramm. Eng. Remote Sens.* **2019**, *85*, 481–491. [[CrossRef](#)]
13. Skarlatos, D. Orthophotograph Production in Urban Areas. *Photogramm. Rec.* **1999**, *16*, 643–650. [[CrossRef](#)]
14. Zhou, G.; Li, H.; Song, R.; Wang, Q.; Xu, J.; Song, B. Orthorectification of Fisheye Image under Equidistant Projection Model. *Remote Sens.* **2022**, *14*, 4175. [[CrossRef](#)]
15. Greenfeld, J. Evaluating the accuracy of digital orthophoto quadrangles (DOQ) in the context of parcel-based GIS. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 199–206.
16. Haggag, M.; Zahran, M.; Salah, M. Towards automated generation of true orthoimages for urban areas. *Am. J. Geogr. Inf. Syst.* **2018**, *7*, 67–74. [[CrossRef](#)]
17. Mayr, W. True orthoimages. *GIM Int.* **2002**, *37*, 37–39.
18. Rau, J.-Y.; Chen, N.-Y.; Chen, L.-C. True orthophoto generation of built-up areas using multi-view images. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 581–588.
19. Shoab, M.; Singh, V.K.; Ravibabu, M.V. High-Precise True Digital Orthoimage Generation and Accuracy Assessment based on UAV Images. *J. Indian Soc. Remote Sens.* **2021**, *50*, 613–622. [[CrossRef](#)]
20. Jauregui, M.; Vélchez, J.; Chacón, L. A procedure for map updating using digital mono-plotting. *Comput. Geosci.* **2002**, *28*, 513–523. [[CrossRef](#)]
21. Siachalou, S. Urban orthoimage analysis generated from IKONOS data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 12–23.
22. Biasion, A.; Dequal, S.; Lingua, A. A new procedure for the automatic production of true orthophotos. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 1682–1777.
23. Shin, Y.H.; Lee, D.-C. True Orthoimage Generation Using Airborne LiDAR Data with Generative Adversarial Network-Based Deep Learning Model. *J. Sensors* **2021**, *2021*, 4304548. [[CrossRef](#)]
24. Yao, J.; Zhang, Z.M. Hierarchical shadow detection for color aerial images. *Comput. Vis. Image Underst.* **2006**, *102*, 60–69. [[CrossRef](#)]
25. Xie, W.; Zhou, G. Experimental realization of urban large-scale true orthoimage generation. In Proceedings of the ISPRS Congress, Beijing, China, 3–11 July 2008; pp. 3–11.
26. Zhou, G.; Jezek, K.C. Satellite photograph mosaics of Greenland from the 1960s era. *Int. J. Remote Sens.* **2002**, *23*, 1143–1159. [[CrossRef](#)]
27. Zhou, G.; Schickler, W. True orthoimage generation in extremely tall building urban areas. *Int. J. Remote Sens.* **2004**, *25*, 5161–5178. [[CrossRef](#)]

28. Zhou, G. Near Real-Time Orthorectification and Mosaic of Small UAV Video Flow for Time-Critical Event Response. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 739–747. [[CrossRef](#)]
29. Zhou, G.; Wang, Y.; Yue, T.; Ye, S.; Wang, W. Building occlusion detection from ghost images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 1074–1084. [[CrossRef](#)]
30. Zhang, R.; Liu, N.; Huang, J.; Zhou, X. On-Board Ortho-Rectification for Images Based on an FPGA. *Remote Sens.* **2017**, *9*, 874. [[CrossRef](#)]
31. Zhou, G.; Zhang, R.; Zhang, D.; Huang, J.; Baysal, O. Real-time ortho-rectification for remote-sensing images. *Int. J. Remote Sens.* **2018**, *40*, 2451–2465. [[CrossRef](#)]
32. Zhou, G.; Bao, X.; Ye, S.; Wang, H.; Yan, H. Selection of Optimal Building Facade Texture Images From UAV-Based Multiple Oblique Image Flows. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1534–1552. [[CrossRef](#)]
33. Jensen, L.B.; Per, S.; Nielsen; Alexander, T.; Mikkelsen, P.S. The Potential of the Technical University of Denmark in the Light of Sustainable Livable Cities. *Des. Civ. Environ. Eng.* **2014**, *90*. [[CrossRef](#)]
34. Huang, X.; Zhang, L. Morphological Building/Shadow Index for Building Extraction From High-Resolution Imagery over Urban Areas. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *5*, 161–172. [[CrossRef](#)]
35. Yu, B.; Wang, L.; Niu, Z. A novel algorithm in buildings/shadow detection based on Harris detector. *Optik* **2014**, *125*, 741–744. [[CrossRef](#)]
36. Gharibi, H.; Habib, A. True Orthophoto Generation from Aerial Frame Images and LiDAR Data: An Update. *Remote Sens.* **2018**, *10*, 581. [[CrossRef](#)]
37. Zhou, G. *Urban High-Resolution Remote Sensing: Algorithms and Modeling*; CRC Press: Boca Raton, FL, USA, 2020. [[CrossRef](#)]
38. Liu, X.; Zhou, G.; Zhang, W.; Luo, S. Study on Local to Global Radiometric Balance for Remotely Sensed Imagery. *Remote Sens.* **2021**, *13*, 2068. [[CrossRef](#)]
39. Wang, Q.; Zhou, G.; Song, R.; Xie, Y.; Luo, M.; Yue, T. Continuous space ant colony algorithm for automatic selection of orthophoto mosaic seamline network. *ISPRS J. Photogramm. Remote Sens.* **2022**, *186*, 201–217. [[CrossRef](#)]
40. Vassilopoulou, S.; Hurni, L.; Dietrich, V.; Baltsavias, E.; Pateraki, M.; Lagios, E.; Parcharidis, I. Orthophoto generation using IKONOS imagery and high-resolution DEM: A case study on volcanic hazard monitoring of Nisyros Island (Greece). *ISPRS J. Photogramm. Remote Sens.* **2002**, *57*, 24–38. [[CrossRef](#)]
41. Cameron, A.; Miller, D.; Ramsay, F.; Nikolaou, I.; Clarke, G. Temporal measurement of the loss of native pinewood in Scotland through the analysis of orthorectified aerial photographs. *J. Environ. Manag.* **2000**, *58*, 33–43. [[CrossRef](#)]
42. Passini, R.; Jacobsen, K. Accuracy analysis of digital orthophotos from very high resolution imagery. International Archives of the Photogrammetry. *Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2004**, *35 Pt B4*, 695–700. [[CrossRef](#)]
43. Piatti, E.J.; Lerma, J.L. Generation of True Ortho-Images Based On Virtual Worlds: Learning Aspects. *Photogramm. Rec.* **2014**, *29*, 49–67. [[CrossRef](#)]
44. Yoo, E.J.; Lee, D.-C. True orthoimage generation by mutual recovery of occlusion areas. *GIScience Remote Sens.* **2015**, *53*, 227–246. [[CrossRef](#)]
45. De Oliveira, H.C.; Dal Poz, A.P.; Galo, M.; Habib, A.F. Surface gradient approach for occlusion detection based on triangulated irregular network for true orthophoto generation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 443–457. [[CrossRef](#)]
46. Zhou, G.; Sha, H. Building Shadow Detection on Ghost Images. *Remote Sens.* **2020**, *12*, 679. [[CrossRef](#)]
47. Marsetič, A. Robust Automatic Generation of True Orthoimages from Very High-Resolution Panchromatic Satellite Imagery Based on Image Incidence Angle for Occlusion Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3733–3749. [[CrossRef](#)]
48. Sheng, Y.; Gong, P.; Biging, G.S. True Orthoimage Production for Forested Areas from Large-Scale Aerial Photographs. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 259–266. [[CrossRef](#)]
49. Leone, A.; Distanto, C. Shadow detection for moving objects based on texture analysis. *Pattern Recognit.* **2007**, *40*, 1222–1233. [[CrossRef](#)]
50. Makarau, A.; Richter, R.; Muller, R.; Reinartz, P. Adaptive Shadow Detection Using a Blackbody Radiator Model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2049–2059. [[CrossRef](#)]
51. Tiwari, S.; Chauhan, K.; Kurmi, Y. Shadow Detection and Compensation in Aerial Images using MATLAB. *Int. J. Comput. Appl.* **2015**, *119*, 5–9. [[CrossRef](#)]
52. Li, D.; Wang, M.; Pan, J. Auto-dodging processing and its application for optical RS images. *Geomat. Inf. Sci. Wuhan Univ.* **2006**, *31*, 753–756.
53. Pan, J.; Wang, M. A Multi-scale Radiometric Re-processing Approach for Color Composite DMC Images. *Geomat. Infor. Sci. Wuhan Univ.* **2007**, *32*, 800–803.
54. Zhou, G.; Pan, Q.; Yue, T.; Wang, Q.; Sha, H.; Huang, S.; Liu, X. Vector and Raster Data Storage based on Morton Code. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLIII-3*, 2523–2526. [[CrossRef](#)]
55. Chan, C.; Tan, S. Determination of the minimum bounding box of an arbitrary solid: An iterative approach. *Comput. Struct.* **2001**, *79*, 1433–1449. [[CrossRef](#)]
56. Fan, H.; Wang, Y.; Gong, J. Layout graph model for semantic façade reconstruction using laser point clouds. *Geo Spatial Inf. Sci.* **2021**, *24*, 403–421. [[CrossRef](#)]