

Article

# Orthorectification of Skin Nevi Images by Means of 3D Model of the Human Body

Piotr M. Szczypiński \*  and Katarzyna Sprawka 

Institute of Electronics, Lodz University of Technology, Al. Politechniki 10, 93-590 Łódź, Poland;  
katarzyna.sprawka@dokt.p.lodz.pl

\* Correspondence: piotr.szczypinski@p.lodz.pl; Tel.: +48-42-631-2642

**Abstract:** Melanoma is the most lethal form of skin cancer, and develops from mutation of pigment-producing cells. As it becomes malignant, it usually grows in size, changes proportions, and develops an irregular border. We introduce a system for early detection of such changes, which enables whole-body screening, especially useful in patients with atypical mole syndrome. The paper proposes a procedure to build a 3D model of the patient, relate the high-resolution skin images with the model, and orthorectify these images to enable detection of size and shape changes in nevi. The novelty is in the application of image encoding indices and barycentric coordinates of the mesh triangles. The proposed procedure was validated with a set of markers of a specified geometry. The markers were attached to the body of a volunteer and analyzed by the system. The results of quantitative comparison of original and corrected images confirm that the orthorectification allows for more accurate estimation of size and proportions of skin nevi.

**Keywords:** visual sensing; skin cancer; image reconstruction; whole-body model



**Citation:** Szczypiński, P.M.; Sprawka, K. Orthorectification of Skin Nevi Images by Means of 3D Model of the Human Body. *Sensors* **2021**, *21*, 8367. <https://doi.org/10.3390/s21248367>

Academic Editor: Ondrej Krejcar

Received: 1 November 2021

Accepted: 12 December 2021

Published: 15 December 2021

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



**Copyright:** © 2021 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/>).

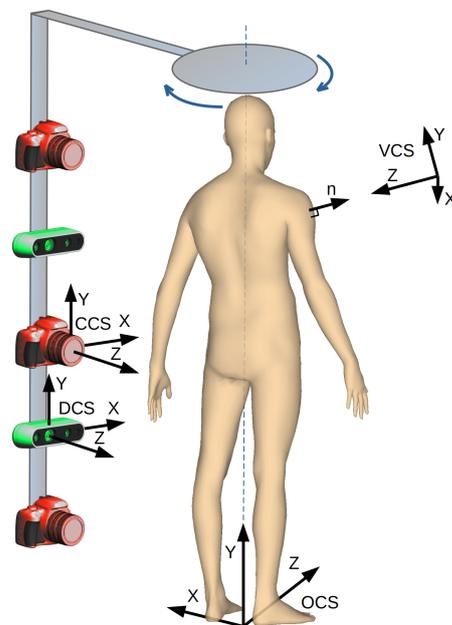
## 1. Introduction

The incidence of skin cancers among fair-skinned populations has increased over the past decades, becoming a growing public health problem [1,2]. According to [3], reports of malignant skin changes have increased by over 600% between the 1940s and 2010 worldwide. In 2020 alone, skin cancer affected over 1,500,000 new individuals, while more than 120,000 died from this condition globally [4]. Some of the skin cancer forms include squamous cell carcinoma, basal cell carcinoma, and melanoma [1]. Although melanomas constitute only around 2% of all diagnosed cases of skin cancers, they are the most deadly of all types [1,5,6]. Melanoma develops as a flat lesion from a mutation of the pigment-producing cells melanocytes. Initially, the skin spots are small, but they grow in size quite quickly and infiltrate deeper layers of the skin to finally spread to the lymph nodes, bones, and lungs if not cured. The 5-year survival rate of melanoma can reach 99% if it is diagnosed and treated at an early stage. This percentage drops sharply with the successive stages of the disease [6,7].

The primary screening method for melanoma is the total body skin examination. The dermatologist searches the skin of the patient for individual nevi. Then, one by one, nevi are examined with a dermatoscope [8]. The device is placed on the skin surface, on top of the nevi. It enables magnified visualization of the skin fragment, measurement of the nevi size, and to make a judgment on its benign or malignant form. The attributes of melanoma-evaluated nevi are known as ABCD, which stands for asymmetry, border irregularity, color, and diameter [9]. Moreover, the growth of the nevi in size, compared with the former examination result, is also used as an indicator of possible malignancy. This kind of examination is cumbersome, time-consuming, and prone to human error, especially for patients with atypical mole syndrome, who may have hundreds of nevi to examine. Automation of this procedure would facilitate monitoring of the suspected skin lesions and make the screening procedure more accessible for patients.

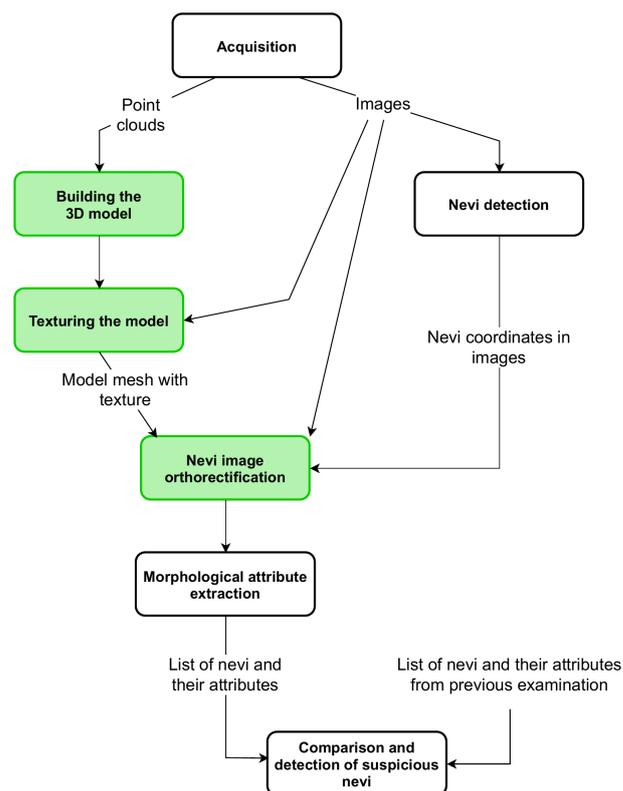
The solution would be to capture skin images covering the whole body of the patient. Each image would present a vast fragment of the skin, showing not a single one, but a group of nevi at the same time. Then, with the computer vision algorithms, the individual nevi would be identified, outlined, characterized in terms of their color and morphological attributes, and finally classified. However, as the image is acquired, distances between the camera lens and particular nevi vary, and the nevi are visible at different angles. As a result, the objects which are at the distance from the camera appear smaller, and the surfaces visible at the angle have distorted proportions. Therefore, such images alone do not carry sufficient information to measure the diameter or asymmetry or to enable size control in time.

Our goal is to develop a device and algorithms for fast examination of skin nevi all over the human body [10]. The analysis is to be carried out automatically. The results should be stored as a reference for the follow-up examination to be carried out after several months. The proposed solution includes a whole-human-body scanning platform (Figure 1). It consists of a rotary arm with high-resolution cameras and depth scanners mounted on it. The cameras are pointed toward the patient, and the arm rotates around the vertical axis to acquire images of, possibly, the whole body surface.



**Figure 1.** The data acquisition system with a rotating arm, cameras, and depth scanners.

The operation of the device is controlled by a computer program, which also includes the acquisition of point clouds from depth scanners and the acquisition of high-resolution images from digital cameras. Point clouds of different sides of the patient are combined to build a 3D model of the skin surface. Then, the model is covered with high-resolution images—the model is textured. During this process, the relations between the coordinates in the 3D model and the coordinates of the image fragments used for texturing are stored. The next procedure identifies the nevi in the images and generates their list with the information in which image and in what coordinates of the image the nevus was found. For each element from the list, the appropriate image fragment is geometrically corrected—orthorectified. Corrected images are used for morphological attribute extraction, such as surface area, asymmetry, or diameter of each nevus. If the result of a former examination of the patient is available, then the values of the morphological attributes of nevi are compared. If the values of the surface area or diameter increased from the previous examination, the program informs the operator about this case. This article focuses exclusively on the algorithms for building a three-dimensional model and for orthorectifying high-resolution images (Figure 2).



**Figure 2.** A block diagram of complete algorithm for nevi growth detection. The green background indicates algorithms discussed in this article.

As already mentioned, due to the perspective effect, images of the nevi which are closer to the camera look bigger than the images of objects seen from a distance. Moreover, if a skin fragment is oriented at an angle with respect to the optical axis of the camera, the resulting image has distorted proportions. This makes it difficult to judge the actual size and shape. To reconstruct the real shape and size of the nevi, we need to know the properties of image projection onto the image sensor of the camera. In particular, we need to know the distance and the orientation angle of the skin surface in relation to the camera.

Such image reconstruction is a subject of orthophotography [11]. For decades, this technique has been applied for aerial and satellite image acquisition [12,13]. The original images are orthorectified, which means they are projected onto the topographic surface and reconstructed in the way that they were seen in frontal view with respect to this surface. The corrected image, an orthophoto, is scale uniform and can be used to measure true distances or object sizes and proportions [14,15]. This procedure is rarely applied in the correction of biomedical images. However, the need for such image correction manifests in problems of intestine unfolding [16] or in face mapping [17]. Our goal is to apply orthorectification to properly measure the sizes and characterize the shapes of skin nevi.

To retrieve the information required for the orthorectification, we use a whole-body model of the person under screening. To create the model, we use depth scanners mounted on the rotating arm similarly to the cameras. The devices have a built-in structured light source and stereo vision camera system. This arrangement enables the acquisition of depth maps, i.e., measuring the distance from the scanner to any point of the skin surface. The set of such measurements constitutes a so-called point cloud. As a result of rotation, the scanners acquire point clouds of the skin surface from various points of view. All the point clouds are, in sequence, registered and merged. The next steps involve the processing of the combined point clouds to form a surface defined by the mesh of triangles.

By having the triangle mesh model and knowing the coordinates of the camera related to the model, it is possible to establish the distances between every point of the surface and the camera. In addition, by finding the normal vector to the model's surface and knowing

the direction of the camera's optical axis, we can find the orientation angle of the skin surface in relation to the camera. This enables the image orthorectification.

In this paper, we present methods for building a three-dimensional model of the whole human body, we describe the method of texture mapping to overlay camera images onto the surface of the triangular mesh, and we describe the algorithm for reconstruction of nevi images in their actual sizes and proportions.

## 2. Materials and Methods

### 2.1. Building a Model

The scanning procedure is as follows. The person under screening, naked or wearing underwear, stands on a platform in a designated location. The arm rotates around the patient, making a full angle ( $360^\circ$ ). During this process, the arm stops for a moment at every  $45^\circ$  to acquire high-resolution images and point clouds. Therefore, the collected data present different sides of the body and cover it entirely. The angle step of  $45^\circ$  was chosen to enable acquisition at eight various orientations. The orientations are dense enough to enable visualization of the whole surface of the skin, including areas under the armpits and between the legs. At the same time, the number of point clouds is small enough to build the 3D model and analyze the images within a few minutes. This way, the person under screening obtains test results in an acceptably short time.

The point clouds are sets of densely spaced three-dimensional coordinates, which are samples of the scanned surfaces. The coordinate system of each point cloud is related to the temporary location of the depth scanner whilst it was acquiring the data. The points in the cloud represent not only the surface of the human body but also surfaces of other objects within the scanner's field of view. Therefore, the first step in point cloud processing is cropping, to keep only the points related to the human object. It is fairly simple since the space which the scanned person occupies is known and limited. The second step is to align and combine the point clouds. This requires a transformation of all the points to the common coordinate system related to the scanned object. The origin of the object coordinate system (OCS) is located at the level of the floor. The Y-axis is the rotation axis of the arm and is perpendicular to the transverse plane of the body. The Z-axis is oriented toward the front of the body and perpendicular to the coronal plane.

All the cameras and depth scanners point toward the rotation axis; each of them is mounted at the specified height  $h_D$  above the floor, and their distance  $r_D$  to the rotation axis is known. To transform (1) a point cloud from its own depth sensor coordinate system (DCS), the point coordinates are translated by the distance  $r$  along the Z-axis and then rotated around the Y-axis by the rotation angle of the arm at which the data were collected.

$$\begin{bmatrix} x_{OCS} \\ y_{OCS} \\ z_{OCS} \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \left( \begin{bmatrix} x_{DCS} \\ y_{DCS} \\ z_{DCS} \end{bmatrix} + \begin{bmatrix} 0 \\ h_D \\ -r_D \end{bmatrix} \right) \quad (1)$$

After the transformation of points to OCS, their number is downsampled using a voxelized grid approach. It works by applying the 3D grid on the point cloud, bucketing the points into voxels. Next, every voxel is approximated by the spatial average of all contained points. The carefully chosen size of the voxel enables obtaining a uniform output point cloud with decimated points while preserving its fidelity. The lowered number of points results in reduced computation time and memory usage during the following processing stages. The next step is the estimation of normal vectors at every point. The applied function determines the principal axis of an arbitrary chosen number of nearest neighboring points using the covariance analysis. The resulting vector should point outside, toward the original camera location.

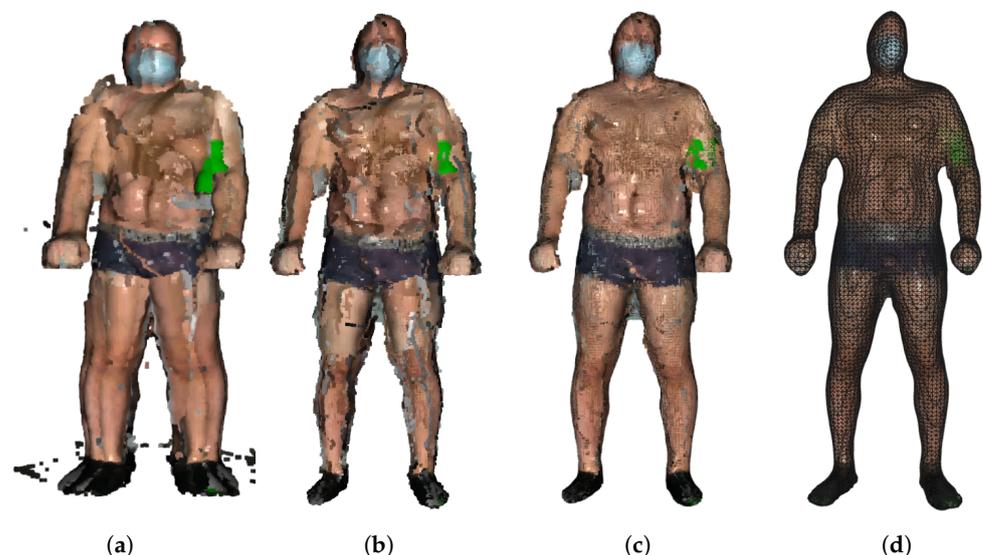
It must be noted that point clouds are transformed from their DCSs to OCS using a rigid transformation based on the assumption that the exact orientation of the cameras is known. This assumption is not completely true, since the devices are mounted on an arm with some inaccuracy, and the arm itself is not perfectly rigid. Moreover, the human

subject moves during the acquisitions, even though we use hand supports to stabilize his or her posture. Therefore, matching the point clouds with each other and building the model is a challenging task. To solve it, we make use of state-of-the-art algorithms and their implementations in Open3D [18] and PCL [19] data processing libraries.

Since the coordinates may be inaccurate, the resulting point clouds are usually misaligned. Therefore, we apply coregistration by the iterative closest point (ICP) registration algorithm which refines the properties of the rigid transformations by minimizing the sum of point-to-point distances [20]. The ICP is a pairwise registration run in a loop for sequentially ordered pairs of point clouds and thus is prone to error. The error between the subsequent point cloud pairs propagates and may accumulate, resulting in the loop-closure problem. Therefore, an additional global optimization for loop closure was applied. Finally, the transformed point clouds are merged to form a common point cloud representing a human posture.

The next step is to remove the sparse outliers often contained in the raw data collected from depth sensors resulting from the measurement errors. They are determined by performing statistical analysis of the points with reference to their local neighborhood. A point is considered an outlier if its mean distance to its neighbors is larger than a given threshold. The threshold which was used was the standard deviation of the average distance from the neighbors. The applied value was chosen empirically and enables removing “loose points” without losing information at the edges of point clouds, which could distort the ICP procedure decreasing overlapping areas.

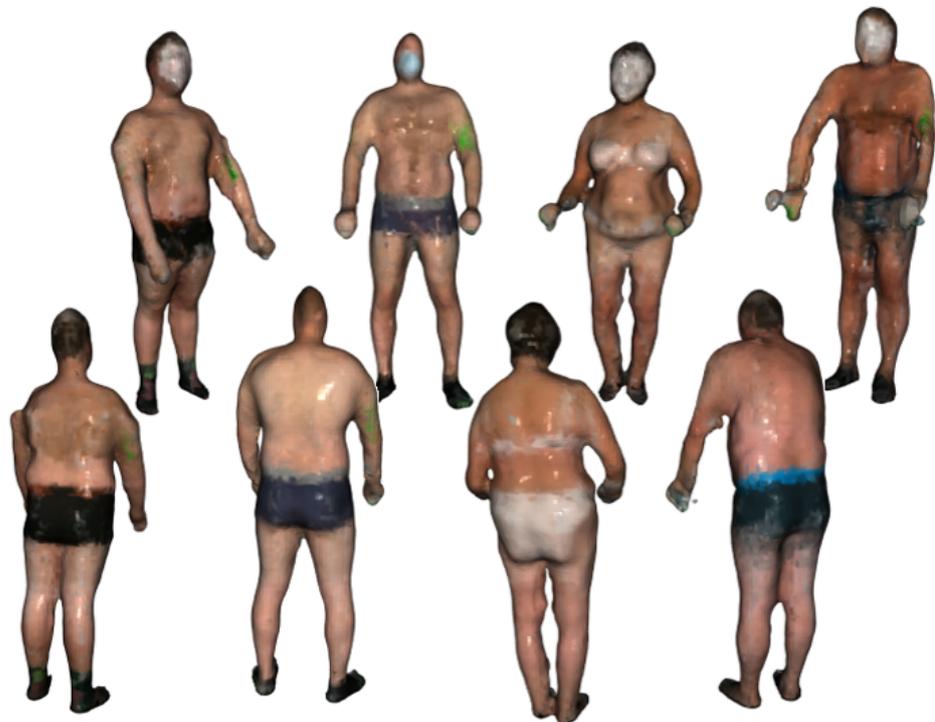
Still, registering may produce double wall artifacts—multiple sets of points forming close and parallel surface structures. Furthermore, the point cloud is likely to contain irregularities caused by measurement errors of the depth sensor, which appears as a wavy surface. These artifacts can be corrected and smoothed by the moving least squares (MLS) surface reconstruction algorithm, which works by locally approximating the surface with polynomials, minimizing the least squares error [21]. The smoothing effect of the MLS algorithm is presented in Figure 3.



**Figure 3.** Stages of model building: (a) original point clouds, (b) registered point clouds after removal of outliers, (c) application of MLS algorithm, and (d) final model after PSR.

When applying the MLS smoothing, the size of the neighborhood used for fitting has to be chosen carefully. The larger value produces a smoother surface, yet it deforms the true geometry. On the other hand, the low value can result in a weak smoothing effect. We use the neighborhood radius of 5 cm. Again, we chose this value empirically, so that the skin appears as a smooth surface, but the original geometry of the input point cloud is preserved.

The final step is to create a triangulated approximation (triangle mesh) of the human body surface, which is formed on the point cloud refined by the previous processing steps. We use the Poisson surface reconstruction (PSR) algorithm developed by Kazhdan et al. [22] to reconstruct the surface from so-called oriented points (points with estimated normal vectors). It produces a watertight and smooth surface. PSR applies the marching cubes algorithm for surface extraction and stores the result in an octree of a given depth. An octree of depth  $D$  produces a three-dimensional mesh of maximal resolution  $2^D \times 2^D \times 2^D$ . Thus, as the octree depth increases, mesh resolution also increases. The higher depth value results in a detailed surface; however, it also emphasizes the remaining noise and artifacts. Choosing the lower value results in greater smoothing but lower details. We chose the depth of 7 as a trade-off between smoothness and resolution. The example final models of volunteers are presented in Figure 4.



**Figure 4.** The final body models of four example volunteers.

### 2.2. Texture Mapping

In principle, the models in Figure 4 define the shape of the surface. For a more complete visualization, additional information about the color is superimposed on the models. It comes from the cameras built into the depth scanner. However, this color data is of low resolution and does not allow for the identification and analysis of nevi. For identification and analysis of nevi images, high-resolution cameras are used. However, as stated before, the individual nevi visible in the images are geometrically distorted and need to be orthorectified. This process requires information from the 3D model. Therefore, correspondences between the vertices of the model mesh and the coordinates of the corresponding locations in the high-resolution images must be found. By knowing these correspondences or relations and knowing coordinates on the surface of the model, one can find the corresponding fragment of the high-resolution image. On the other hand, by selecting a point in a high-resolution image, it is possible to find a corresponding location on the surface of the model.

In computer graphics, the process of finding the relations between coordinates in raster images and coordinates of vertices in the 3D mesh is called texture mapping. The high-resolution image, the texture, is superimposed on the triangle mesh to visualize the model with appropriate detail. In the case of our approach, a known fragment of the high-

resolution image, which may present an individual nevus, is matched to the coordinates on the surface of the model. Since the coordinates of the camera at the time of taking the picture are known, the distance between the camera and the location of the nevus on the model can be calculated. Moreover, the model enables determination of the normal vector to the skin surface and, thus, the orientation of this surface in relation to the optical axis of the camera. These two pieces of information are sufficient to locally orthorectify the image at the point where the nevus is visible.

In the following, we formally explain how the texture mapping is carried out, and how the relations between high-resolution images and the 3D model are used to locally orthorectify the images.

Texture mapping is a process of finding relations between the vertices of the triangle mesh  $[x_{OCS} y_{OCS} z_{OCS}]^T$  and coordinates  $[x_t y_t]^T$  in the texture image. It is accomplished by transforming the coordinates of every vertex from OCS to CCS. The transformation (Equation (2)) is described by the matrix defining relative rotation of the object with respect to the arm by the angle  $\alpha$ , and by the translation vector including camera elevation  $h_C$  above the floor level and distance  $r_C$  from the rotation axis. These make Equation (2) similar to Equation (1), which defines relations between DCS and OCS. However, the camera orientation may not be exactly as desired due to assembly imperfections. It can be mounted slightly askew or shifted. To account for these discrepancies, the calibration matrix  $\mathbf{Q}_C$  and the calibration vector  $\mathbf{T}_C$  are introduced into the equation. The matrix and the vector should be determined individually for each camera as a result of the calibration procedure consisting of exact alignment of the texture images with the model.

$$\begin{bmatrix} x_{CCS} \\ y_{CCS} \\ z_{CCS} \end{bmatrix} = \mathbf{Q}_C \left( \begin{bmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \begin{bmatrix} x_{OCS} \\ y_{OCS} \\ z_{OCS} \end{bmatrix} - \begin{bmatrix} 0 \\ h_C \\ -r_C \end{bmatrix} \right) + \mathbf{T}_C \quad (2)$$

After transforming the coordinates of the mesh vertices to the CCS, using the model of a pinhole camera, the vertex coordinates are projected onto the space of a two-dimensional texture image (Equation (3)). The  $d_C$  parameter reflects the properties of the optical system of the camera and the resolution of the image sensor. As a result, we obtain coordinates in the texture image related to the vertices of the mesh.

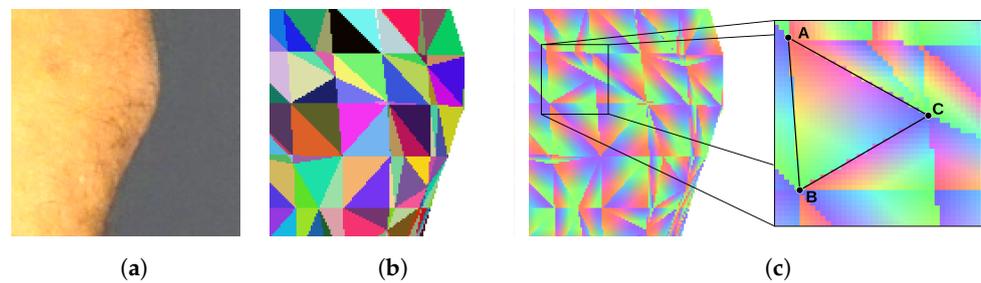
$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \frac{1}{d_t z_{CCS}} \begin{bmatrix} x_{CCS} \\ y_{CCS} \end{bmatrix} \quad (3)$$

It must be noted that there are many images of the human subject taken at various  $\alpha$  angles and by various cameras mounted at different heights  $h_C$  above the floor level. Therefore, the computation defined by Equations (2) and (3) has to be repeated for all available texture images individually.

There are three practical problems with texturing. The first one is the occlusion of some body fragments by the other fragments which are closer to the camera, for example, armpit areas obscured by arms. This means that the fragment of the mesh related to the occluded fragment should not be textured with the particular texture image, as the texture does not picture this fragment. Secondly, the texture should not be mapped on parts of the model that are on the opposite side—not facing the camera lens. Finally, the third difficulty is that the same regions of the skin may be visible in several images taken by different cameras or at different angles. In this case, there is a need to decide which of the images presents the fragment with the greatest accuracy and therefore which of the images is the most appropriate for texturing the related fragment of the model.

To determine if the particular triangle of the mesh should be textured by the particular piece of the image, we render the triangle mesh. The model is rendered as if it was visible from the point of view of the camera that has taken the texture image. Every triangle of the mesh is assigned a unique color for identification. Thus, the rendering result is an image consisting of a mosaic of colored triangles (Figure 5b). The rendered surface should exactly align with the corresponding body fragment in the texture image, and the two images

should be of the same size. Therefore, the pixel of the texture image which displays a particular point of the skin can be related to the particular triangle of the mesh by checking the color of the pixel at the same location in the rendered image.



**Figure 5.** Images used to establish texture–model relations: (a) image of a texture, (b) mosaic of mesh triangles identified by unique colors, (c) barycentric coordinates coded by the three color components.

The rendered image of a triangle mosaic can be quickly produced by means of the contemporary graphics processing units (GPUs). The GPUs run the back-face culling and the occlusion culling procedures to determine which areas of the rendered surface are not visible. Therefore, the resulting image does not show triangles, which are obscured or facing away from the camera lens. As the vertices of a triangle are being projected onto the texture image, with Equations (2) and (3), we verify if the particular triangle is visible in the mosaic. If the color of the triangle in the mosaic does not match the projected one, then the particular texture is not assigned to the triangle. This way, comparison of the mosaic image with the texture image enables determination of obscured skin fragments and solves the first two difficulties.

In a multicamera system, the same triangle may be textured with information from several images obtained from various cameras or at different angles. In our approach for texturing the triangle, we select the image which presents the particular skin fragment with the highest resolution. Therefore, after projecting the triangle vertices onto the texture image, we establish the area of the projected triangle. Then, we select the image for which the area is the largest. This solves the third difficulty.

The model with superimposed images of the skin surface serves multiple purposes. It enables visualization of the human body, localization of a skin nevus on its surface, and retrieval of the information on the distance and orientation of the skin fragment. When viewing, the user can rotate or change the size of the model, and indicate a specific place on its surface, for example, a mole that he or she wants to analyze. The user should be able to immediately see the original image or orthorectified visualization of the indicated location. The other scenario is image segmentation and automatic nevi detection in texture images. After detecting the nevus in the image, it is also necessary to retrieve the information from the model needed for orthorectification.

Identification of the place indicated by the user takes place in two stages. We assume that the user observes the model at some chosen orientation and scale. This means that the textured model has been rendered in this orientation and is presented on the screen. By means of the mouse or a touch, the user indicates a particular point on the rendered image. Then, the algorithm renders the triangle mosaic image in the same scale and orientation as the model visible to the user. The algorithm checks the color of the mosaic image at the same location as indicated by the user. The color enables the identification of the corresponding triangle on the model's mesh. This technique is known in computer graphics as mouse picking [23–25].

In the second scenario, the input for the procedure are coordinates of the nevus identified by image processing. Again, the mosaic of triangles is rendered from the same point of view as the texture image was taken. The color of the mosaic taken from the same coordinates enables identification of the model's triangle.

In both scenarios, identification of the triangle is not enough. Thus, in the next step, we render an image in which each triangle is linearly shaded with RGB color components (Figure 5c). Each of the three vertices of the triangle is assigned one of the three color components. As the color over the surface of the triangle is interpolated, every point of the triangle has a unique color shade. It can be noticed that the proportions of color components at a specific location correspond with the barycentric coordinates of this location. As a result, the coordinates  $\mathbf{V}$  of the indicated location are obtained as a sum of coordinates of the triangle vertices  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  weighted by the color component proportions  $r$ ,  $g$ , and  $b$  (Equation (4)).

$$\mathbf{V} = \frac{r\mathbf{A} + g\mathbf{B} + b\mathbf{C}}{r + g + b} \quad (4)$$

The method explained in this section finds a link between a point in the texture image and the corresponding point on the model surface. This in turn provides data for further computation of the distance between the point and the camera and the local orientation of the surface. The method is time-efficient as it utilizes graphics processing units available on almost every computer, tablet, or smartphone, which are able to render images in a fraction of a second.

### 2.3. Image Reconstruction

Another application of the model is to obtain information about the distance and orientation of the skin surface in relation to the optical axis of the camera—the camera by which a high-resolution texture image was taken. This information is necessary to orthorectify the image from the camera. As explained in the previous section, the point indicated on the high-resolution texture image can be linked with the corresponding point and the corresponding triangle in the model's triangle mesh.

We know the  $[x_t \ y_t]^T$  coordinates of the center of a nevus. In the mosaic image, we read the color and identify the indicated triangle. From the image of the shaded triangles, we read the color components, convert them to barycentric coordinates, and calculate the coordinates of the center of the nevus in the OCS (Equation (4)). Next, we calculate the normal vector  $\mathbf{n}$  to the surface of the triangle (Equation (5)), where  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  are vectors of coordinates of the triangle vertices.

$$\mathbf{n} = \frac{(\mathbf{B} - \mathbf{A}) \times (\mathbf{C} - \mathbf{A})}{|\mathbf{B} - \mathbf{A}| |\mathbf{C} - \mathbf{A}|} \quad (5)$$

Let us consider an imaginary orthogonal camera which serves as a means to orthorectify the image of the nevi. The virtual camera coordinate system (VCS) is created in such a way (Figure 1) that the camera is pointed at the  $\mathbf{V}$  point. Its optical axis  $\mathbf{Z}_{VSC}$  is perpendicular to the surface of the model at this point and antiparallel to the normal vector  $\mathbf{n}$  (Equation (6)). We also assume that the  $\mathbf{X}_{VSC}$  is parallel to the XZ plane of the OCS and perpendicular to the  $\mathbf{Z}_{VSC}$  (Equation (7)). As the  $\mathbf{Y}_{VSC}$  is perpendicular to the  $\mathbf{X}_{VSC}$  and  $\mathbf{Z}_{VSC}$ , we calculate its axis as a vector product of the two axes (Equation (8)). Finally, we obtain Equation (9) to transform the coordinates from the OCS to the VCS.

$$\mathbf{Z}_{VSC} = [g_{zx} \ g_{zy} \ g_{zz}]^T = -\mathbf{n} \quad (6)$$

$$\mathbf{X}_{VSC} = [g_{xx} \ g_{xy} \ g_{xz}]^T = \frac{[g_{xx} \ 0 \ -g_{zx}]^T}{\sqrt{g_{zz}^2 + g_{zx}^2}} \quad (7)$$

$$\mathbf{Y}_{VSC} = [g_{yx} \ g_{yy} \ g_{yz}]^T = \mathbf{X}_{VSC} \times \mathbf{Z}_{VSC} \quad (8)$$

$$\begin{bmatrix} x_{VCS} \\ y_{VCS} \\ z_{VCS} \end{bmatrix} = \mathbf{G} \left( \begin{bmatrix} x_{OCS} \\ y_{OCS} \\ z_{OCS} \end{bmatrix} - \mathbf{V} \right) = \begin{bmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{bmatrix} \left( \begin{bmatrix} x_{OCS} \\ y_{OCS} \\ z_{OCS} \end{bmatrix} - \mathbf{V} \right) \quad (9)$$

By having the coordinates in the VCS, we can project them onto the surface of the sensor of the virtual camera. The equation defining orthogonal projection of the  $[x_{\text{OCS}} y_{\text{OCS}} z_{\text{OCS}}]^T$  onto the surface of the orthorectified image takes the form of Equation (10), where  $s$  is an arbitrary scaling factor.

$$\begin{bmatrix} x_g \\ y_g \end{bmatrix} = s \begin{bmatrix} x_{\text{VCS}} \\ y_{\text{VCS}} \end{bmatrix} \quad (10)$$

Equations (8) and (9) enable computation of the coordinates of a point in the corrected, orthorectified image on the basis of coordinates of this point in the OCS. However, when generating the corrected image, we usually have to solve the inverse problem. While knowing the coordinates  $[x_g y_g]^T$ , we need to find coordinates on the surface of the model in the OCS and then coordinates  $[x_t y_t]^T$  in the original texture image. To solve this problem, let us note that the matrix  $\mathbf{G}$  in Equation (8) is orthonormal. Therefore, the inverse matrix can be simply obtained by its transposition  $\mathbf{G}^{-1} = \mathbf{G}^T$ . In addition, the center of the corrected image ( $[x_g y_g]^T = [0 0]^T$ ) corresponds to the point  $\mathbf{V}$  on the model's surface. As a result of these observations, we obtain the relationship of Equation (11).

$$\begin{bmatrix} x_{\text{OCS}} \\ y_{\text{OCS}} \\ z_{\text{OCS}} \end{bmatrix} = \frac{1}{s} \begin{bmatrix} g_{xx} & g_{yx} \\ g_{xy} & g_{yy} \\ g_{xz} & g_{yz} \end{bmatrix} \begin{bmatrix} x_g \\ y_g \end{bmatrix} + \mathbf{V} \quad (11)$$

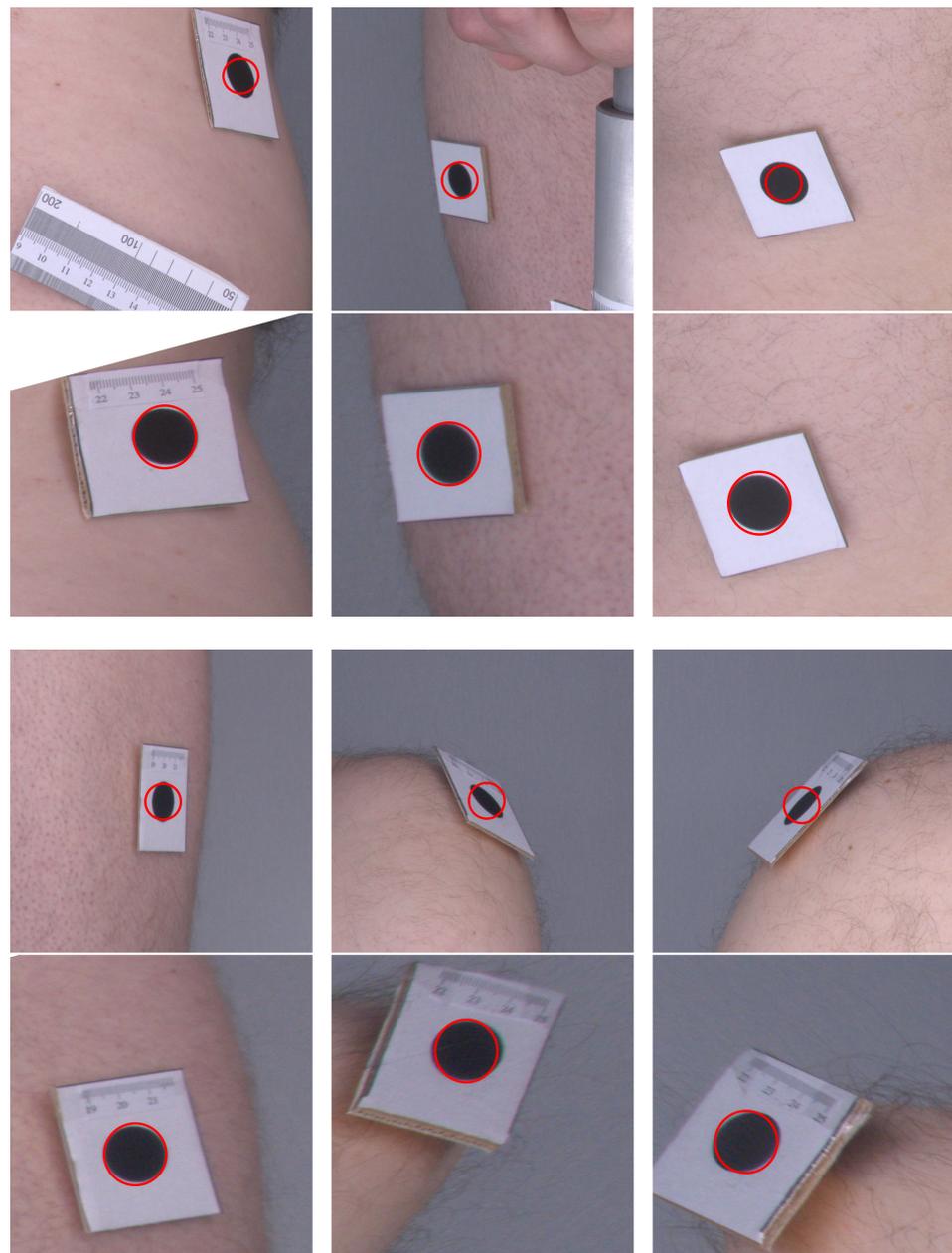
The image reconstruction procedure is based on the fact that for each point of the reconstructed image, we find its counterpart in the texture image. For this, we use Equation (10), and then Equations (2) and (3). When coordinates are found, the color components at the point  $[x_t y_t]^T$  are interpolated. It should be noted that changing the scaling factors enables adjustment of the resolution of the resulting image in pixels per unit length. When knowing the resolution of the corrected image, the size of the nevi in the length units can be computed.

#### 2.4. Experiment

To assess the feasibility of image orthorectification with the presented method, we performed the following experiment. A volunteer with markers stuck to his skin was scanned. Every marker shows a black circle of 2 cm diameter. The volunteer was scanned using a prototype device consisting of the rotating arm with four Sony DSC RX100 cameras and three Intel RealSense D435 depth scanners. As the arm was turning, the images and point clouds were acquired at each subsequent 45 degree angle. The whole-body model was built, and the marker centers were indicated on the model. Next, the procedure for image orthorectification was applied.

Figure 6 presents a comparison of the original images as they were acquired by the cameras and their corrected counterparts. One can see that, in the original images, the markers are elliptical in shape, and their sizes vary depending on their distance from the camera. In the corrected images, the markers are even in size and appear more similar to circles. The angles at the corners of the marker images also move closer to the right angle after applying the orthorectification procedure. Therefore, the visual inspection and comparison of the original and corrected images confirm that the orthorectification meets the expectations, and the resulting images allow for a more accurate assessment of the size and proportions.

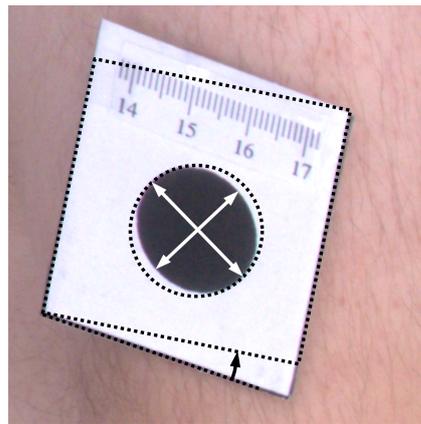
To quantitatively assess results of orthorectification (Figure 7), the ellipse was fitted to every marker, and proportions of semi-major to semi-minor axes were calculated. For such correctly rectified images, the proportion should be equal to one. Moreover, the area of the markers was computed, which for the circle of 2 cm diameter should be equal to  $2\pi \text{ cm}^2$ . Furthermore, the shape of the tag that contains the marker is rectangular. Therefore, measuring the angle between the two edges of the tag and computing the difference between the angle and the right angle also facilitates the assessment of the rectification accuracy. The mean values of such measurements are presented in Table 1.



**Figure 6.** Example images before and after orthorectification. Original images are in the top rows and corrected images are in the bottom rows. Red circles indicate the expected shape of markers.

The resulted values of the morphological attributes are closer to the desired ones after the images were orthorectified. The shear angle was reduced more than two times. The ratio of the semi-major to semi-minor axes is closer to one, with the error of the ratio reduced 2.7 times. The standard deviations of the area computations decreased almost two times, and the estimates of the area match the expected value.

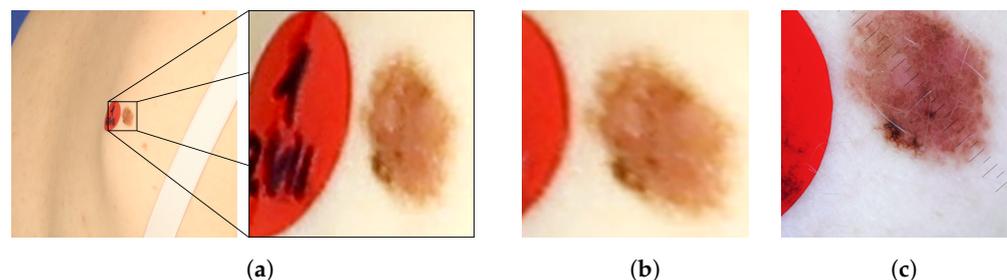
Figure 8 presents images of an exemplary noncancerous nevus. In the original image, the nevus's proportions are distorted. Moreover, it is impossible to judge its actual size and the proportion of length to width. In the orthorectified image, the assessment of these morphological features becomes possible. For comparison, a reference image obtained with a dermatoscope is also presented. The shape and size of the corrected and reference images are similar.



**Figure 7.** Illustration of morphological attributes measured for the assessment of the orthorectification procedure.

**Table 1.** Comparison of the resulting morphological attributes of marker images before and after orthorectification. The values in the brackets represent relative standard deviations.

	Original	Orthorectified
Angle	13.14 (0.53)	5.65 (0.21)
Ratio	1.43 (0.16)	1.16 (0.03)
Area	10.74 (0.44)	6.10 (0.23)



**Figure 8.** Example of a nevus: (a) as it appears in the original image, (b) the same image after orthorectification, and (c) the reference image acquired by means of dermatoscope.

It should be noted that the results were obtained with an imperfect 3D model of the human body. The surface of the model is not as smooth as in a real human body. There are apparent errors in the orientations of the model triangle faces, which contribute to the errors in orthorectified images. Building a model that reproduces the shape of the skin more accurately would, even more, improve the outcomes of the orthorectification procedure.

### 3. Conclusions

We have presented a complete procedure for building a whole-body model to enable the correction of nevi geometries. An appropriate choice of the model reconstruction parameters enables obtaining a relatively smooth model with preserved true structure and shape of the body.

The qualitative examination shows that the models properly, though not perfectly, represent the shape of the surface of the patient's body. The biggest problem is the patient's movement during the scanning procedure, which results in mutually displaced surfaces that are difficult to fit. In our experiment, inaccuracies in the corrected images of markers may result from inaccuracies in the model. In addition, the source of errors is the inaccurate sticking of the markers to the skin, thus the surface of markers does not align with the surface of the skin. We are continuing work on the scanning system and model-building algorithms to minimize these errors. Moreover, as the marker centers during the experiment were indicated manually, we plan to implement image segmentation and nevi identification

algorithms. This would allow for a fully automatic analysis of the entire human body surface. Following this work, we plan to verify the operation of the device and the implemented algorithms quantitatively. Taking into account the results obtained so far, we find them encouraging in confirming the feasibility to apply the proposed solution in practice.

**Author Contributions:** P.M.S.: conceptualization, investigation, software, supervision, writing—original draft, writing—review and editing; K.S.: investigation, software, visualization, writing—original draft, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the European Regional Fund and National Centre for Research and Development project POIR.04.01.04-00-0125/18-00 “Development of a device to support early stage examination of skin lesions, including melanoma using computer imaging, spatial modeling, comparative analysis and classification methods”. The scanning device was built by Skopia Estetic Clinic Sp. z o.o.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki. Ethical review and approval were waived by the Ethics Committee of Regional Medical Chamber in Kraków for this study, due to full non-invasiveness of the conducted research. (no medical procedures, the only action was taking optical pictures and 3D scans of the participants) and to ensuring the anonymity of the volunteers participating in the study.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are not publicly available because optical full-body images and 3D scans of patients constitute sensitive medical information.

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

## References

- Geller, A.C.; Annas, G.D. Epidemiology of melanoma and nonmelanoma skin cancer. *Semin. Oncol. Nurs.* **2003**, *19*, 2–11. [[CrossRef](#)]
- Apalla, Z.; Lallas, A.; Sotiriou, E.; Lazaridou, E.; Ioannides, D. Epidemiological trends in skin cancer. *Dermatol. Pract. Concept.* **2017**, *7*, 1–6. [[CrossRef](#)] [[PubMed](#)]
- Chang, N.B.; Feng, R.; Gao, Z.; Gao, W. Skin cancer incidence is highly associated with ultraviolet-B radiation history. *Int. J. Hygiene Environ. Health* **2010**, *213*, 359–368. [[CrossRef](#)]
- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
- Okur, E.; Turkan, M. A survey on automated melanoma detection. *Eng. Appl. Artif. Intell.* **2018**, *73*, 50–67. [[CrossRef](#)]
- Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2021. *CA Cancer J. Clin.* **2021**, *71*, 7–33. [[CrossRef](#)] [[PubMed](#)]
- Mumford, S.L.; Towler, B.P.; Pashler, A.L.; Gilleard, O.; Martin, Y.; Newbury, S.F. Circulating microRNA biomarkers in melanoma: Tools and challenges in personalised medicine. *Biomolecules* **2018**, *8*, 21. [[CrossRef](#)]
- Ring, C.; Cox, N.; Lee, J.B. Dermatoscopy. *Clin. Dermatol.* **2021**, *39*, 635–642. [[CrossRef](#)]
- Lee, H.; Kwon, K. A mathematical analysis of the ABCD criteria for diagnosing malignant melanoma. *Phys. Med. Biol.* **2017**, *62*, 1865. [[CrossRef](#)]
- Strzelecki, M.H.; Strąkowska, M.; Kozłowski, M.; Urbańczyk, T.; Wielowieyska-Szybińska, D.; Kociotek, M. Skin Lesion Detection Algorithms in Whole Body Images. *Sensors* **2021**, *21*, 6639. [[CrossRef](#)]
- Lear, A.C. Digital orthophotography: Mapping with pictures. *IEEE Comput. Graph. Appl.* **1997**, *17*, 12–14. [[CrossRef](#)]
- Petrie, G. Orthophotomaps. *Trans. Inst. Br. Geogr.* **1977**, *2*, 49–70. [[CrossRef](#)]
- Robinson, A.H.; Morrison, J.L.; Muehrcke, P.C. Cartography 1950–2000. *Trans. Inst. Br. Geogr.* **1977**, *2*, 3–18. [[CrossRef](#)]
- Nale, D. Digital orthophotography: The foundation of GIS. *Am. City County* **1995**, *110*, 26–31.
- 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](#)]
- Sudarsky, S.; Geiger, B.; Chéfd’hotel, C.; Guendel, L. Colon unfolding via skeletal subspace deformation. *Med. Image Comput. Comput. Assist. Interv.* **2008**, *11*, 205–212.
- Perronnin, F.; Dugelay, J.L.; Rose, K. A probabilistic model of face mapping with local transformations and its application to person recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1157–1171. [[CrossRef](#)]
- Zhou, Q.Y.; Park, J.; Koltun, V. Open3D: A modern library for 3D data processing. *arXiv* **2018**, arXiv:1801.09847.

19. Rusu, R.B.; Cousins, S. 3D is here: Point cloud library (pcl). In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4.
20. Besl, P.J.; McKay, N.D. Method for registration of 3-D shapes. Sensor fusion IV: Control paradigms and data structures. *Int. Soc. Opt. Photonics* **1992**, *1611*, 586–606.
21. Alexa, M.; Behr, J.; Cohen-Or, D.; Fleishman, S.; Levin, D.; Silva, C.T. Computing and rendering point set surfaces. *IEEE Trans. Vis. Comput. Graph.* **2003**, *9*, 3–15. [[CrossRef](#)]
22. Kazhdan, M.; Bolitho, M.; Hoppe, H. Poisson surface reconstruction. In Proceedings of the Fourth Eurographics Symposium on Geometry Processing, Cagliari, Italy, 26–28 June 2006; Volume 7.
23. Zhao, H.; Jin, X.; Shen, J.; Lu, S. Fast and reliable mouse picking using graphics hardware. *Int. J. Comput. Games Technol.* **2009**, *2009*, 730894. [[CrossRef](#)]
24. Wu, S.T.; Abrantes, M.; Tost, D.; Batagelo, H.C. Picking and snapping for 3D input devices. In Proceedings of the 16th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2003), Sao Carlos, Brazil, 12–15 October 2003; pp. 140–147.
25. He, J.Y.; Xu, Q.H.; You, J. A 3-D Picking Method Based on OpenGL. *Comput. Eng. Sci.* **2006**, *1*, 45–47.