

Reflection Removal for Large-Scale 3D Point Clouds

Jae-Seong Yun and Jae-Young Sim

School of Electrical and Computer Engineering

Ulsan National Institute of Science and Technology, Ulsan, Korea

jssyun@unist.ac.kr and jysim@unist.ac.kr

Abstract

Large-scale 3D point clouds (LS3DPCs) captured by terrestrial LiDAR scanners often exhibit reflection artifacts by glasses, which degrade the performance of related computer vision techniques. In this paper, we propose an efficient reflection removal algorithm for LS3DPCs. We first partition the unit sphere into local surface patches which are then classified into the ordinary patches and the glass patches according to the number of echo pulses from emitted laser pulses. Then we estimate the glass region of dominant reflection artifacts by measuring the reliability. We also detect and remove the virtual points using the conditions of the reflection symmetry and the geometric similarity. We test the performance of the proposed algorithm on LS3DPCs capturing real-world outdoor scenes, and show that the proposed algorithm estimates valid glass regions faithfully and removes the virtual points caused by reflection artifacts successfully.

1. Introduction

We often take glass images, for example, taking photos of goods through show windows and taking photos of buildings with glass curtain walls. In such cases, resulting glass images include transmitted scenes captured through the glass as well as undesired scenes reflected on the glass, such as sky, buildings and trees as shown in Figure 1(a). Reflection removal has been studied which automatically decomposes glass images into transmission scenes and reflected scenes. A single glass image based technique exploits the characteristics of reflection images such as gradient sparsity [18], relative smoothness [20], ghosting cue [25] and Laplacian fidelity [3]. Reflection removal techniques based on multiple glass images employ polarization [17, 24] or flash [1] to extract different characteristics from multiple images. Moreover, multiple glass images taken from different capturing positions are also used for reflection removal, where the gradient sparsity [19], motion fields [27] and low-rank matrix representation are investigated [13].

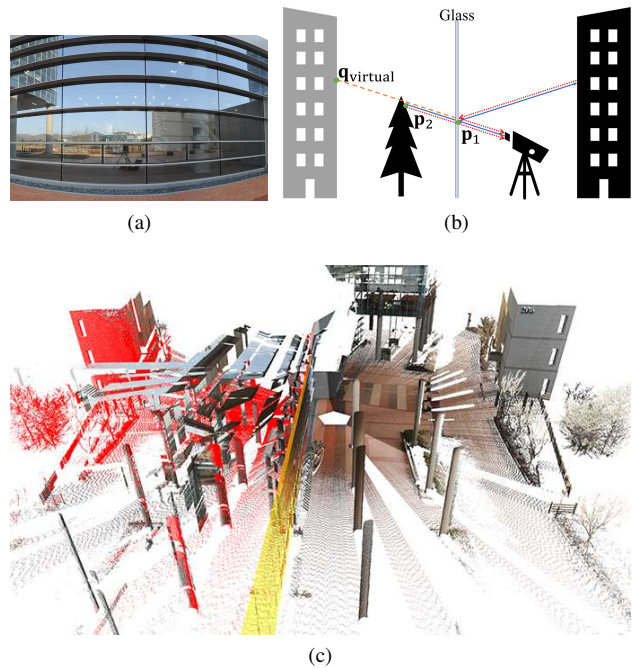


Figure 1: Reflection by glass. (a) Reflection artifact on 2D image. (b) The principle of reflection in a LiDAR laser scanner. The black building and tree are real-world objects, while the gray building denotes a virtual object generated by reflection. (c) A LS3DPC model where the virtual points are shown in red and the glass plane is shown in yellow, respectively.

With the advent of high-performance Light Detection And Ranging (LiDAR) scanners, lots of effort has been made to develop efficient techniques for processing 3D point clouds. A real-time LiDAR scanner and stereo cameras were used to generate a test dataset [12] for the applications of autonomous vehicles, which is used to study object detection [8, 7] and road detection [6, 7]. Moreover, high-resolution terrestrial LiDAR scanners provide Large-Scale 3D Point Clouds (LS3DPCs) capturing 360° environment

of 3D real world scene, which are recently used for the research issues of data compression [2] and saliency detection [26, 28].

LS3DPCs captured by using terrestrial LiDAR scanners also suffer from the reflection artifacts since many outdoor real-world structures include glasses, e.g., vehicles and buildings. A LiDAR scanner measures the distances of target scene by emitting laser pulses and receiving their echo pulses. Figure 1(b) shows the principle of reflection caused by LiDAR scanners. It calculates the distance from the scanner to an object by measuring the time it takes to emit a laser pulse and receive the echo pulse. A single laser pulse emitted from the scanner first hits the glass and its echo pulse is come back to the scanner creating a 3D point \mathbf{p}_1 on the glass. Also, penetration and reflection of light occur simultaneously on the glass. The penetrated laser pulse hits the tree, a real-world object behind the glass, and its echo pulse is received at the scanner creating another 3D point \mathbf{p}_2 . On the other hand, the reflected laser pulse hits the building, a real-world object in front of the glass, and the scanner receives its echo pulse to create a virtual 3D point $\mathbf{q}_{\text{virtual}}$ at the opposite side of the glass plane. Consequently, from a single emitted pulse, multiple echo pulses are generated which generate three different 3D points. Among the points, \mathbf{p}_1 and \mathbf{p}_2 are valid points sampled on real-world objects, but $\mathbf{q}_{\text{virtual}}$ locates on a wrong position in 3D space, i.e., gray building. Such reflection artifacts occur since the scanner regards a received pulse is reflected on a real-world object only once. Therefore, the resulting LS3DPC includes virtual scene which may degrade the performance of the related processing techniques of LS3DPCs. Figure 1(c) shows the obtained LS3DPC capturing a real-world outdoor scene including the scene in Figure 1(a), where the virtual points of building and tree are shown in red and the plane of glass is depicted in yellow, respectively.

To the best of our knowledge, no research has been conducted to remove the reflection artifacts from LS3DPCs. In this paper, we propose a first reflection removal algorithm for LS3DPCs. We investigate the capturing mechanism of terrestrial LiDAR scanner and estimate the glass regions by modeling the distribution of the number of received echo pulses. Then we detect a point as a virtual point when it has a corresponding real point with similar geometric feature at the opposite side about the glass plane. We perform the experiments on LS3DPC models by capturing real-world outdoor scenes and show that the proposed algorithm removes the reflection artifacts faithfully.

2. Overview of Proposed Algorithm

We use a 3D terrestrial laser scanner, RIEGL VZ-400 [22], to acquire LS3DPCs for real-world outdoor scenes including glasses. In general, glass is highly specular,

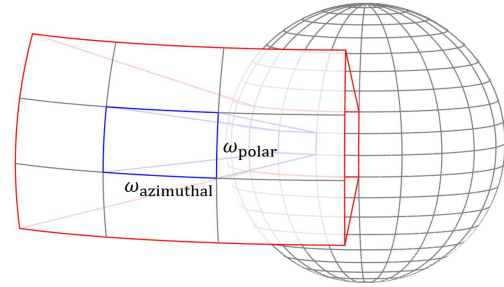


Figure 2: Partitioning of the unit sphere into local surface patches.

and therefore valid 3D points are sampled over a relatively small area on a glass plane where the directions of emitted lasers are close to the normal direction of the glass plane. However, many buildings have coated glass curtain walls and windows which produce sampled points over relatively larger areas than uncoated ones. Also, when capturing a typical real-world scene by LiDAR scanners, while a single glass plane is located close to the scanner associated with dominant reflection artifact, the other glass regions yield small numbers of points with negligible reflection artifacts. Therefore, in this work, we first estimate a glass plane of dominant reflection artifact in a captured scene. Then we detect and remove the virtual points by comparing the features between a pair of symmetric points about the glass plane.

3. Glass Region Estimation

In the research of reflection removal for 2D images, we usually consider glass images are captured such that the reflection occurs over the entire image area. However, LiDAR scanners capture 360° environment of real-world scene, and therefore, the virtual points associated with glasses are distributed at local regions in a single LS3DPC model. We first estimate the glass plane where the dominant reflection occurs using the characteristics of LiDAR scanning.

3.1. Patch Classification by Point Projection

In general, a single 3D point is created for each laser pulse, since the light is reflected on a real-world object only once in most cases. But, as shown in Figure 1(b), a laser pulse hitting the glass plane may produce more than two points: one from the penetrated laser pulse and the others from the reflected pulses on the glass plane. Each laser pulse is emitted periodically with predefined azimuthal and polar angular resolutions, $\omega_{\text{azimuthal}}$ and ω_{polar} , respectively. By using the measured distance and the associated azimuthal and polar angles of an emitted laser pulse, the coordinates of 3D points are computed in the spherical coordinate system. As shown in Figure 2, we consider the unit

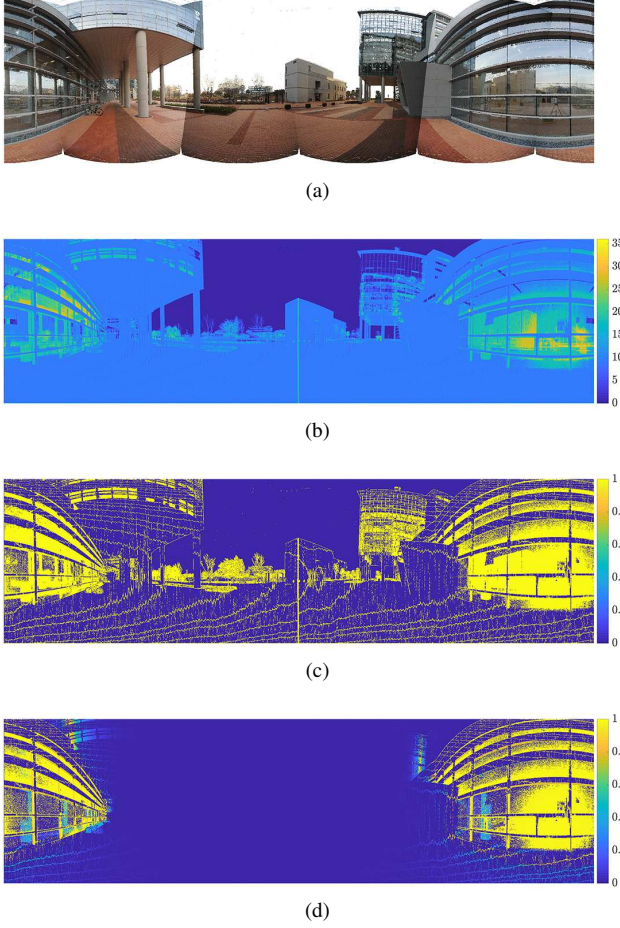


Figure 3: Glass region estimation. (a) The panorama image for a target scene. (b) The number of projected points n_i . (c) The posterior probability of $p_1(n_i)$ of belonging to glass regions. (d) The reliability map of ρ_i where dominant glass regions are highlighted.

sphere with the origin at the scanner location, and partition the sphere surface into local patches where the blue rectangular area becomes a primitive surface patch which covers the angular range of $\omega_{\text{azimuthal}} \times \omega_{\text{polar}}$. Then we count a number of points corresponding to each patch by projecting the 3D points onto the surface of the unit sphere, and we classify the patches into the ordinary patches including only a single 3D point and the glass patches where two or more 3D points are projected.

However, a laser scanner acquires the points based on whisk broom scanning which samples one point at a certain time instance by mechanically rotating the sensor. Hence, to reduce the sampling error of points, we employ a larger surface patch covering a wider angular range of $m\omega_{\text{azimuthal}} \times m\omega_{\text{polar}}$ where m is a positive integer. We set $m = 3$ empirically as depicted by the red rectangular area in Figure 2,

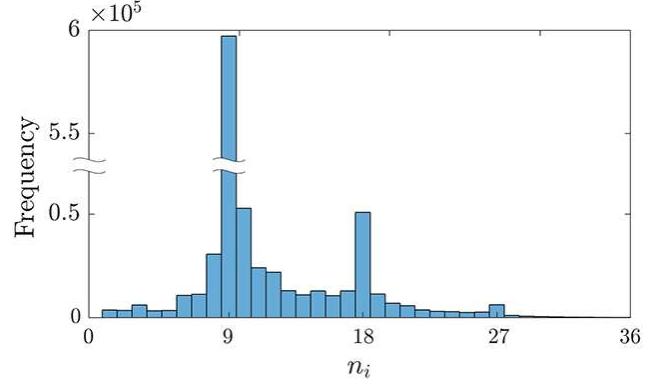


Figure 4: Histogram of n_i 's for the patches with $n_i > 0$ in 'Engineering Building' model.

since it is the smallest integer detecting the glass regions of dominant reflection reliably. We see that each patch is associated with approximately m^2 laser pulses.

Let n_i be the number of projected points to the i -th surface patch \mathcal{S}_i . Figure 3(b) visualizes the distribution of n_i 's associated with the target scene shown in Figure 3(a). We see that while the ordinary patches usually exhibit $n_i = m^2$, the glass patches tend to have $n_i > m^2$ due to multiple echo pulses by reflection. Also, Figure 4 shows the histogram of n_i 's counted over all the patches having valid projected points, i.e., $n_i > 0$ on a LS3DPC model in Figure 3. We observe that there are two strong peaks at $n_i = 9$ and $n_i = 18$. The first peak at $n_i = 9$ implies that m^2 laser pulses produce the same number of points which is highly probable to describe the ordinary patches. On the other hand, the second peak at $n_i = 18$ is associated with $2m^2$ pulses yielded mainly by the glass regions where each pulse generates two points on average, one from the penetrated laser pulse and the other from the reflected laser pulse. We also see a weak peak at $n_i = 27$ associated with $3m^2$ pulses on average.

3.2. Reliability for Glass Patches

We classify the patches into two categories of ordinary patches and glass patches by modeling the distribution of n_i using the mixture of K Gaussian distributions [5]. The density of Gaussian mixture model is given by

$$f(n_i) = \sum_{k=0}^{K-1} \lambda_k \mathcal{N}(n_i | \mu_k, \sigma_k^2), \quad (1)$$

where $\mathcal{N}(n_i | \mu_k, \sigma_k^2)$ is the k -th Gaussian density with the mean μ_k and the variance σ_k^2 , respectively, and λ_k is the mixing coefficient. We set the number of Gaussians to $K = 2$, one for the ordinary patches and the other for the glass patches. We introduce a two dimensional binary random vector $\mathbf{z} = [z_0, z_1]^T$ where $z_k \in \{0, 1\}$ and $\sum_{k=0}^{K-1} z_k = 1$. Without loss of generality, we assume that $\mu_0 \leq \mu_1$, then

$\mathbf{z} = [1, 0]^T$ at ordinary patches and $\mathbf{z} = [0, 1]^T$ at the glass patches, respectively.

To estimate the parameters μ_k , σ_k and λ_k , we use the Expectation Maximization (EM) algorithm [9]. For given randomly initialized parameters μ_k , σ_k and λ_k , the EM algorithm evaluates the posterior probability $p_k(n_i)$ as

$$p_k(n_i) = \frac{\lambda_k \mathcal{N}(n_i | \mu_k, \sigma_k^2)}{\sum_{j=0}^{K-1} \lambda_j \mathcal{N}(n_i | \mu_j, \sigma_j^2)}. \quad (2)$$

Then the parameters of μ_k , σ_k and λ_k are updated by using $p_k(n_i)$. This process is iteratively applied to yield optimal parameters. Figure 3(c) shows the resulting probability distribution of $p_1(n_i)$, where we see that the patches corresponding to the glass regions are assigned relatively high probabilities. However, the patches corresponding to the complex scene structures such as trees as shown in far background also have high probabilities as well, since a laser pulse may produce multiple echo pulses due to arbitrary diffuse reflection. In addition, some small glass regions located far from the scanner are also assigned high probabilities, but their reflection artifacts are negligible.

Therefore, in order to select only the region on a glass plane yielding dominant reflection artifact, we compute a reliability for each patch. Let us define a set of points $\mathcal{C} = \{\mathbf{c}_i\}$, where \mathbf{c}_i is the closest point to the scanner among all the projected points to \mathcal{S}_i . Then, we define the set $\mathcal{C}_{\text{candidates}} \subseteq \mathcal{C}$ as

$$\mathcal{C}_{\text{candidates}} = \{\mathbf{c}_i | p_1(n_i) > p_0(n_i), \mathbf{c}_i \in \mathcal{C}\}. \quad (3)$$

Since the points sampled on the dominant glass plane should have smaller distances from the scanner than that of the transmission and virtual points, we assume that $\mathcal{C}_{\text{candidates}}$ consists of the points belonging to the glass patches and some complex objects. By applying RANSAC [10] to fit a plane to $\mathcal{C}_{\text{candidates}}$, we estimate the glass plane Π . Then we define a reliability ρ_i for each patch \mathcal{S}_i by weighting the probability $p_1(n_i)$ as

$$\rho_i = e^{-d_i} p_1(n_i), \quad (4)$$

where d_i is the Euclidean distance between Π and \mathbf{c}_i . If \mathbf{c}_i is close to Π , we assign a high reliability to \mathcal{S}_i . In contrary, if \mathbf{c}_i deviates from Π too much, we assign a low reliability to \mathcal{S}_i . Figure 3(d) shows the resulting reliability map, where only the dominant and closest glass plane is highlighted while the high probabilities of $p_1(n_i)$ associated with the far and small glass regions, trees and ground are suppressed.

4. Virtual Point Detection

We detect and remove the virtual points associated with the glass patches estimated in Section 3. Figure 5 illustrates

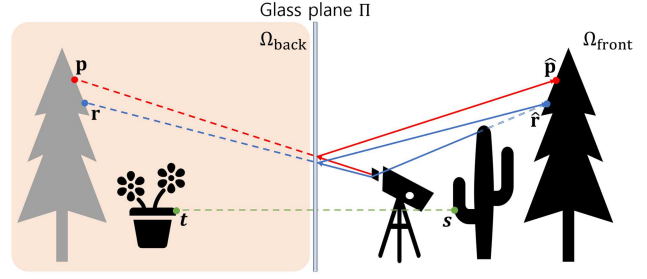


Figure 5: Symmetry relation of reflection between a pair of real point and virtual point.

the situation when reflection occurs on a glass plane, where the virtual points \mathbf{p} and \mathbf{r} correspond to the real points $\hat{\mathbf{p}}$ and $\hat{\mathbf{r}}$. Note that the virtual points are always created at the opposite side to the corresponding real points with respect to the glass plane, i.e., located behind the glass from the scanner. We divide the space of a target scene into Ω_{front} and Ω_{back} by taking the glass plane Π as the boundary, where Ω_{front} contains the scanner location. Then we detect the virtual points only within the half space Ω_{back} , since Ω_{front} contains the real points only. A point $\mathbf{p} \in \Omega_{\text{back}}$ is highly probable to be a virtual point when 1) it is projected to a patch with a high reliability and 2) there is a corresponding real point $\hat{\mathbf{p}} \in \Omega_{\text{front}}$ which has symmetry relation to \mathbf{p} about the glass plane Π , and yields a similar geometric feature to \mathbf{p} .

4.1. Reflection Symmetry

For a given point $\mathbf{p} \in \Omega_{\text{back}}$, we first evaluate a symmetry score $\gamma_{\text{symmetry}}(\mathbf{p})$ which measures how an actually acquired point $\hat{\mathbf{p}} \in \Omega_{\text{front}}$ is close to the symmetric position of $\mathbf{p} \in \Omega_{\text{front}}$. We find the symmetric position of \mathbf{p} about the glass plane Π using the Householder matrix [15], which describes a linear transformation of reflection about a plane. The Householder matrix \mathbf{A} for a given plane is defined as $\mathbf{A} = \mathbf{I} - 2\mathbf{n}\mathbf{n}^T$, where \mathbf{I} is the identity matrix and \mathbf{n} is the unit vector of plane normal. Hence, with the plane equation $ax + by + cz + d = 0$, it is given by

$$\mathbf{A} = \begin{bmatrix} 1 - 2a^2 & -2ab & -2ac & -2ad \\ -2ab & 1 - 2b^2 & -2bc & -2bd \\ -2ac & -2bc & 1 - 2c^2 & -2cd \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

Note that the Householder matrix is orthogonal.

$$\begin{aligned} \mathbf{A}^T \mathbf{A} &= (\mathbf{I} - 2\mathbf{n}\mathbf{n}^T)^T (\mathbf{I} - 2\mathbf{n}\mathbf{n}^T) \\ &= \mathbf{I} - 4\mathbf{n}\mathbf{n}^T + 4\mathbf{n}\mathbf{n}^T \mathbf{n}\mathbf{n}^T = \mathbf{I}, \end{aligned} \quad (6)$$

since $\mathbf{n}\mathbf{n}^T$ is a symmetric matrix and $\mathbf{n}^T \mathbf{n} = 1$. Similarly, $\mathbf{A}\mathbf{A}^T = \mathbf{I}$.

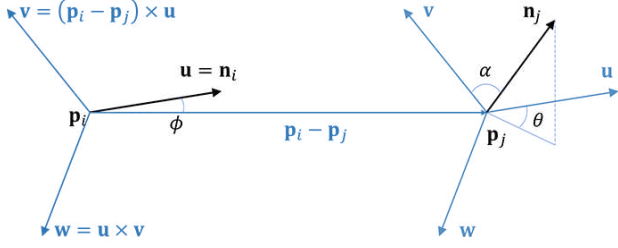


Figure 6: Fast Point Feature Histogram which describes the three angular variations (α, ϕ, θ) associated with the two points \mathbf{p}_i and \mathbf{p}_j .

Then we have the relation $\hat{\mathbf{p}} = \mathbf{A}_\Pi \mathbf{p}$ where \mathbf{A}_Π is the Householder matrix of glass plane Π . The homogeneous coordinates are used to represent the translation of plane. We use a k -d tree to find the closest point $\mathbf{q} \in \Omega_{\text{front}}$ to $\hat{\mathbf{p}}$, and compute

$$\gamma_{\text{symmetry}}(\mathbf{p}) = e^{-\frac{\|\hat{\mathbf{p}} - \mathbf{q}\|}{\beta_1}}, \quad (7)$$

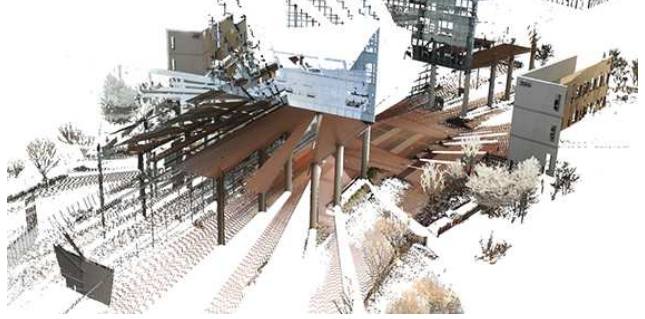
where $\|\hat{\mathbf{p}} - \mathbf{q}\|$ is the Euclidean distance between $\hat{\mathbf{p}}$ and \mathbf{q} .

4.2. Geometric Similarity

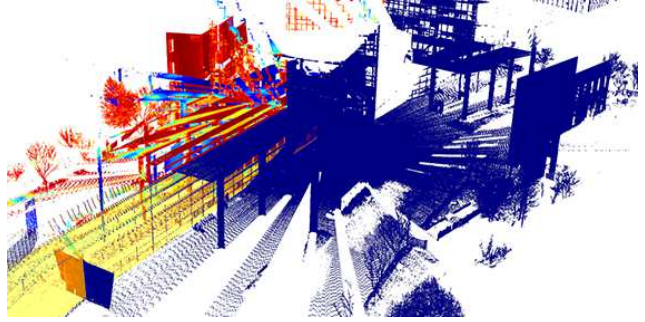
The symmetry score $\gamma_{\text{symmetry}}(\mathbf{p})$ may not be sufficient to detect the virtual points faithfully. For example, as shown in Figure 5, a real point $\mathbf{t} \in \Omega_{\text{back}}$ has a relatively high score of $\gamma_{\text{symmetry}}(\mathbf{t})$, since we can find another real point $\mathbf{s} \in \Omega_{\text{front}}$ which yields a small distance $d(\hat{\mathbf{t}}, \mathbf{s})$ where $\hat{\mathbf{t}} = \mathbf{A}_\Pi \mathbf{t}$. Therefore, we also evaluate a geometric similarity $\gamma_{\text{similarity}}(\mathbf{p})$ between \mathbf{p} and \mathbf{q} using a 3D feature descriptor. Note that the employed 3D feature descriptor should be reflection invariant, since the two points \mathbf{p} and \mathbf{q} are assumed to have the relation of reflection symmetry.

Among many 3D feature descriptors [16, 11, 21], we use Fast Point Feature Histogram (FPFH) [23]. For a given query point \mathbf{p}_i , FPFH computes the three angular variations of (α, ϕ, θ) for all the 50 nearest neighboring points \mathbf{p}_j to \mathbf{p}_i in terms of the Euclidean distance, as shown in Figure 6. The FPFH vector $\Phi(\mathbf{p})$ at \mathbf{p} is defined as the histogram of the three angular variations. We prove the reflection invariance of FPFH. Let $\hat{\mathbf{p}}_i = \mathbf{A}\mathbf{p}_i$ and $\hat{\mathbf{p}}_j = \mathbf{A}\mathbf{p}_j$ be the reflected points of \mathbf{p}_i and \mathbf{p}_j , respectively. Then, the axes satisfy $\hat{\mathbf{v}} = \mathbf{A}\mathbf{v}$, $\hat{\mathbf{w}} = \mathbf{A}\mathbf{w}$, and $\hat{\mathbf{u}} = \mathbf{A}\mathbf{u}$. Let us denote the normals of \mathbf{p}_i and \mathbf{p}_j as \mathbf{n}_i and \mathbf{n}_j , respectively, then we also have $\hat{\mathbf{n}}_i = \mathbf{A}\mathbf{n}_i$ and $\hat{\mathbf{n}}_j = \mathbf{A}\mathbf{n}_j$. The angular variation of α is given by $\alpha = \langle \mathbf{v}, \mathbf{n}_j \rangle$. The angular variation $\hat{\alpha}$ associated with the two reflected points $\hat{\mathbf{p}}_i$ and $\hat{\mathbf{p}}_j$ can be derived as

$$\begin{aligned} \hat{\alpha} &= \langle \hat{\mathbf{v}}, \hat{\mathbf{n}}_j \rangle = \langle \mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{n}_j \rangle \\ &= (\mathbf{A}\mathbf{v})^T \mathbf{A}\mathbf{n}_j = \mathbf{v}^T \mathbf{A}^T \mathbf{A} \mathbf{n}_j \\ &= \mathbf{v}^T \mathbf{n}_j = \langle \mathbf{v}, \mathbf{n}_j \rangle = \alpha, \end{aligned} \quad (8)$$



(a)



(b)

Figure 7: Scores for virtual points. (a) An input LS3DPC model. (b) The distribution of score $\gamma(\mathbf{p})$ where the red and blue colors depict high and low scores, respectively.

since $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ by (6). Similarly, the other angular variations are given by $\phi = \langle \mathbf{u}, ((\mathbf{p}_i - \mathbf{p}_j) / \|\mathbf{p}_i - \mathbf{p}_j\|) \rangle$, and $\theta = \arctan(\langle \mathbf{w}, \mathbf{n}_j \rangle, \langle \mathbf{u}, \mathbf{n}_j \rangle)$, and we can prove $\hat{\phi} = \phi$ and $\hat{\theta} = \theta$ similarly. Consequently, FPFH implies the reflection invariance.

We compute the similarity score $\gamma_{\text{similarity}}(\mathbf{p})$ as

$$\gamma_{\text{similarity}}(\mathbf{p}) = e^{-\frac{H(\Phi(\mathbf{p}), \Phi(\mathbf{q}))}{\beta_2}}, \quad (9)$$

where $\Phi(\mathbf{p})$ is the FPFH vector at \mathbf{p} . $H(\Phi(\mathbf{p}), \Phi(\mathbf{q}))$ is the Hellinger distance[14] defined by

$$H(\Phi(\mathbf{p}), \Phi(\mathbf{q})) = \sqrt{\frac{\sum_i (\sqrt{\Phi_i(\mathbf{p})} - \sqrt{\Phi_i(\mathbf{q})})^2}{2}}, \quad (10)$$

where $\Phi_i(\mathbf{p})$ denotes the i -th element of $\Phi(\mathbf{p})$.

4.3. Detection of Virtual Points

We combine the symmetry score and the similarity score together to compute a final score $\gamma(\mathbf{p})$ given by

$$\gamma(\mathbf{p}) = \gamma_{\text{symmetry}}(\mathbf{p}) \gamma_{\text{similarity}}(\mathbf{p}). \quad (11)$$

Figure 7 shows the resulting scores, where we set $\gamma(\mathbf{p}) = 0$ for $\mathbf{p} \in \Omega_{\text{front}}$. We see that a point $\mathbf{p} \in \Omega_{\text{back}}$ yields a high

score $\gamma(\mathbf{p})$ when it has a corresponding point $\mathbf{q} \in \Omega_{\text{front}}$ which yields a similar geometric feature to \mathbf{p} and is close to $\mathbf{A}_{\Pi}\mathbf{p}$. We basically separate the virtual points from a LS3DPC model using the resulting scores by assigning a binary label l_i to each point \mathbf{p}_i such that $l_i = 1$ when \mathbf{p}_i is virtual and $l_i = 0$ when \mathbf{p}_i is real.

However, as shown in Figure 5, a virtual point \mathbf{r} can be generated by reflection, but its corresponding real point $\hat{\mathbf{r}}$ is not actually acquired by scanner due to the occlusion by the cactus. In such a case, \mathbf{r} may not be detected as a virtual point due to a low score of $\gamma(\mathbf{r})$. To overcome this issue, we formulate an energy function given by

$$E(\mathcal{L}) = \sum_i D_i + \tau \sum_i \sum_{\mathbf{p}_j \in N_i} V_{ij}, \quad (12)$$

where $\mathcal{L} = \{l_i\}$ is the set of all labels. The data cost D_i is defined by

$$D_i = \begin{cases} -\rho(\mathbf{p}_i)\gamma(\mathbf{p}_i), & l_i = 1 \\ -(1 - \rho(\mathbf{p}_i)\gamma(\mathbf{p}_i)), & l_i = 0 \end{cases} \quad (13)$$

where $\rho(\mathbf{p}_i)$ is the reliability of the patch where \mathbf{p}_i is projected. By multiplying the reliability, we effectively detect the points associated with the glass patches of dominant reflection. V_{ij} is the smoothness cost that forces the neighboring points to have same labels, and N_i is the set of neighboring points to \mathbf{p}_i . We find 48 nearest neighboring points using a k -d tree, but omit the points \mathbf{p}_j 's having Euclidean distances $d(\mathbf{p}_i, \mathbf{p}_j)$ larger than 0.1% of the bounding box's diagonal distance for a given LS3DPC model. We select the parameters empirically by testing the performance with variable numbers of neighboring points from 16 to 96 and variable threshold values for $d(\mathbf{p}_i, \mathbf{p}_j)$ from 0.01% to 1%, respectively. V_{ij} is computed by

$$V_{ij} = \begin{cases} e^{-\frac{d(\mathbf{p}_i, \mathbf{p}_j)}{\beta_1}} e^{-\frac{H(\Phi(\mathbf{p}_i), \Phi(\mathbf{p}_j))}{\beta_2}}, & l_i \neq l_j \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

We use the Iterated Conditional Modes (ICM) [4] to obtain an optimal solution of (12). Finally we remove the detected virtual points from an input LS3DPC model.

5. Experimental Result

We acquire LS3DPCs by capturing real-world outdoor scenes including glasses using a 3D terrestrial LiDAR scanner, RIEGL VZ-400 [22] with the angular resolutions $\omega_{\text{azimuthal}} = 0.06^\circ$ and $\omega_{\text{polar}} = 0.06^\circ$. We evaluate the performance of the proposed algorithm on six LS3DPC models shown in Figure 8: 'Architecture Building', 'Engineering Building', 'Natural Science Building', 'Botanical Garden', 'Gymnasium' and 'Terrace', where a single LS3DPC model has approximately 5~6 millions of points

in general. β_1 and β_2 in (7) and (9) are empirically chosen as 0.5 and 0.5 for 'Botanical Garden' model and 1.5 and 1.5 for the other models. τ in (12) are set to 1.

We first show the estimated glass regions in Figure 8, where we see that the glass patches on dominant glass planes are detected successfully in most cases. Note that no points are sampled over a large area of the dominant glass plane in Figure 8(a), since the LiDAR scanner is located too close to the glass. When outdoor scenes are captured, the sky is often reflected on the glasses. In such cases, however, no valid real points are sampled on the sky due to the limit of capturing distance, and therefore, no virtual points are generated associated with the sky region.

Figure 9 shows the results of reflection removal in LS3DPCs. In 'Architecture Building' model, some trees are not removed since they are reflected on the glass region where no sampled points are obtained as shown in Figure 8(a). However, most of the reflected building and trees including a reflected person are detected and removed successfully despite a massive absence of sampled points on the glass plane. As shown in Figure 8(b), most of the reflection artifacts are well removed including building, trees and ground. However, some part of the reflected building shown in the second column still remains since the corresponding glass patches are assigned relatively low reliability values due to the lack of sampled points. In 'Natural Science Building' model, a linear region of real trees that intersect with the extended glass plane is also classified as virtual points, since the trees have multiple echo pulses due to complex scene structures and also yield short distances to the glass plane. However, the removal of these points is inconspicuous in the reconstructed 3D model. Also note that some virtual points of reflected building are not removed due to the lack of corresponding real points by occlusion. 'Botanical Garden' model exhibits a relatively complex scene that similar trees appear on both sides of the glass plane. In this model, the building and trees reflected on the glass are removed, while the real trees in the garden survive successfully.

In addition, we measure the processing time of the proposed algorithm on Intel i7-4790k Processor (4.38GHz), and provide the results in Table 1. Note that the processing times of glass region estimation and virtual point detection are not linearly proportional to the number of points. The descriptor computation to find normal and FPFH consumes more than a half of the total processing time in most cases. However, the pre-computed descriptors can be also used for further processing of point clouds in various applications.

6. Conclusion

In this paper, we proposed a reflection removal algorithm for LS3DPCs captured by a terrestrial LiDAR scanner. We investigated the characteristics of received echo pulses asso-

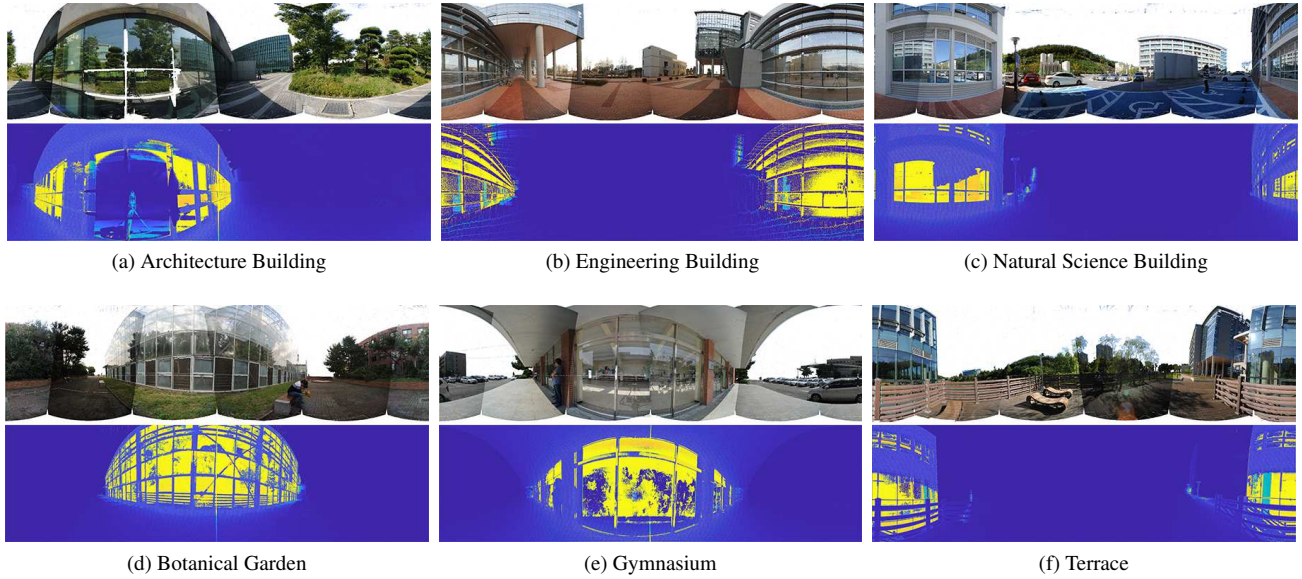


Figure 8: Estimation of glass regions associated with dominant reflection artifacts. In each subfigure, a color panorama image and the resulting reliability distribution is shown in top and bottom, respectively.

Model	Number of points	Processing time (sec.)			
		Glass region estimation	Descriptor comput.	Virtual point detection	Total
(a)	5,562,972	15.9	43.0	34.9	93.7
(b)	9,720,671	10.7	72.5	51.6	134.8
(c)	4,913,710	12.8	37.2	4.7	54.7
(d)	6,140,383	10.7	49.8	11.8	72.3
(e)	5,609,449	10.0	43.3	16.6	69.9
(f)	5,000,902	9.5	38.2	24.7	72.4

Table 1: Processing time of the proposed algorithm. (a) Architecture Building, (b) Engineering Building, (c) Natural Science Building, (d) Botanical Garden, (e) Gymnasium, and (f) Terrace.

ciated with glasses, and computed the reliability to estimate valid glass regions of dominant reflection artifacts. Then we measured a reflection symmetry score and a geometric similarity score for each point, and detected the virtual points reliably which should be removed. Experimental results demonstrated that the proposed algorithm successfully detected and removed the reflection artifacts in LS3DPCs.

7. Future Research Issues

In this work, we assume that only a single dominant glass plane exists in each LS3DPC model. When there are multiple glass planes, the current algorithm selects one of them as a glass plane and removes the virtual points associated with the selected plane only. To overcome this limitation,

we are conducting follow-up research on the reflection removal with multiple glass planes.

Quantitative performance evaluation of the proposed algorithm is quite challenging since it is hard to obtain ground truth 3D point clouds without reflection artifact. We plan to generate reflection-free point clouds by simulating the acquisition process of LiDAR on synthetic 3D point cloud models, which are used to evaluate the quantitative performance of reflection removal algorithms.

The proposed distance based weighting scheme for glass region estimation successfully prevents the vegetation, however some glass regions are also assigned low reliability values. We may apply vegetation detection methods to increase the accuracy of glass region estimation.

In this pioneering work, we intentionally use geometric information only without using the information of 2D images to make the proposed algorithm more widely applicable to LiDAR based applications when no additional devices are required to capture 2D images. As a future research topic, we will investigate the glass region estimation on 2D panoramic image domain, and exploit the geometric information and photometric information together to improve the performance of reflection removal for large-scale colored 3D point clouds.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (No. 2017R1A2B4011970).

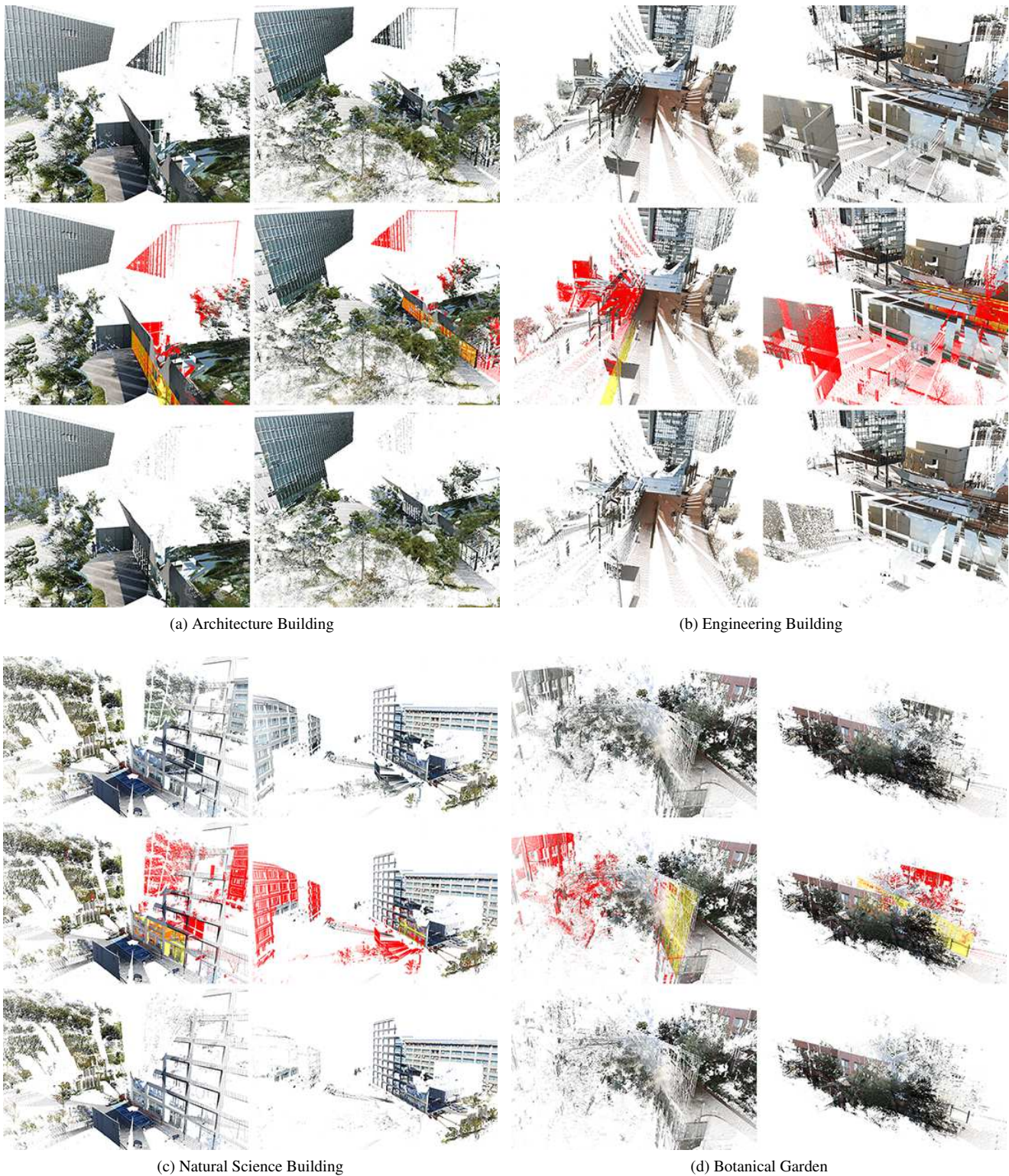


Figure 9: Results of reflection removal for LS3DPCs. In each subfigure, the top row shows an input LS3DPC model, the middle row visualizes the estimated glass regions (yellow) and the detected virtual points (red), and the bottom row shows the resulting LS3DPC model where the virtual points are removed.

References

- [1] A. Agrawal, R. Raskar, S. K. Nayar, and Y. Li. Removing photography artifacts using gradient projection and flash-exposure sampling. *ACM Trans. Graph.*, 24(3):828–835, 2005. [1](#)
- [2] J.-K. Ahn, K.-Y. Lee, J.-Y. Sim, and C.-S. Kim. Large-scale 3D point cloud compression using adaptive radial distance prediction in hybrid coordinate domains. *IEEE J. Sel. Topics Signal Process.*, 9(3):422–434, 2015. [2](#)
- [3] N. Arvanitopoulos, R. Achanta, and S. Susstrunk. Single image reflection suppression. In *Proc. IEEE CVPR*, pages 1752–1760, 2017. [1](#)
- [4] J. Besag. On the statistical analysis of dirty pictures. *J. R. Stat. Soc. Series B (Methodol.)*, 48(3):259–302, 1986. [6](#)
- [5] C. M. Bishop. *Pattern recognition and machine learning*, pages 431–439. springer, 2006. [3](#)
- [6] L. Caltagirone, S. Scheidegger, L. Svensson, and M. Wahde. Fast LiDAR-based road detection using fully convolutional neural networks. In *Proc. IEEE IV*, pages 1019–1024, 2017. [1](#)
- [7] L. Chen, J. Yang, and H. Kong. LiDAR-histogram for fast road and obstacle detection. In *Proc. IEEE ICRA*, pages 1343–1348, 2017. [1](#)
- [8] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia. Multi-view 3D object detection network for autonomous driving. In *Proc. IEEE CVPR*, pages 6526–6534, 2017. [1](#)
- [9] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Series B (Methodol.)*, 39(1):1–38, 1977. [4](#)
- [10] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981. [4](#)
- [11] A. Frome, D. Huber, R. Kolluri, T. Bülow, and J. Malik. Recognizing objects in range data using regional point descriptors. In *Proc. ECCV*, pages 224–237, 2004. [5](#)
- [12] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Proc. IEEE CVPR*, pages 3354–3361, 2012. [1](#)
- [13] B.-J. Han and J.-Y. Sim. Reflection removal using low-rank matrix completion. In *Proc. IEEE CVPR*, pages 5438–5446, 2017. [1](#)
- [14] E. Hellinger. Neue begründung der theorie quadratischer formen von unendlichvielen veränderlichen. *Journal für die reine und angewandte Mathematik*, 136:210–271, 1909. [5](#)
- [15] A. S. Householder. Unitary triangularization of a nonsymmetric matrix. *Journal of the ACM*, 5(4):339–342, 1958. [4](#)
- [16] A. E. Johnson. *Spin-images: a representation for 3-D surface matching*. PhD thesis, Carnegie Mellon University Pittsburgh, PA, 1997. [5](#)
- [17] N. Kong, Y.-W. Tai, and J. S. Shin. A physically-based approach to reflection separation: from physical modeling to constrained optimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 36(2):209–221, 2014. [1](#)
- [18] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(9), 2007. [1](#)
- [19] Y. Li and M. S. Brown. Exploiting reflection change for automatic reflection removal. In *Proc. IEEE ICCV*, pages 2432–2439, 2013. [1](#)
- [20] Y. Li and M. S. Brown. Single image layer separation using relative smoothness. In *Proc. IEEE CVPR*, pages 2752–2759, 2014. [1](#)
- [21] Z.-C. Marton, D. Pangercic, N. Blodow, J. Kleinhellefort, and M. Beetz. General 3D modelling of novel objects from a single view. In *Proc. IEEE/RSJ IROS*, pages 3700–3705, 2010. [5](#)
- [22] RIEGL. RIEGL VZ-400 3D terrestrial laser scanner. [2, 6](#)
- [23] R. B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (FPFH) for 3d registration. In *Proc. IEEE ICRA*, pages 3212–3217, 2009. [5](#)
- [24] Y. Y. Schechner, J. Shamir, and N. Kiryati. Polarization and statistical analysis of scenes containing a semireflector. *J. Opt. Soc. Am. A*, 17(2):276–284, 2000. [1](#)
- [25] Y. Shih, D. Krishnan, F. Durand, and W. T. Freeman. Reflection removal using ghosting cues. In *Proc. IEEE CVPR*, pages 3193–3201, 2015. [1](#)
- [26] E. Shtrom, G. Leifman, and A. Tal. Saliency detection in large point sets. In *Proc. IEEE ICCV*, pages 3591–3598, 2013. [2](#)
- [27] T. Xue, M. Rubinstein, C. Liu, and W. T. Freeman. A computational approach for obstruction-free photography. *ACM Trans. Graph.*, 34(4):79:1–79:11, 2015. [1](#)
- [28] J.-S. Yun and J.-Y. Sim. Supervoxel-based saliency detection for large-scale colored 3D point clouds. In *Proc. IEEE ICIP*, pages 4062–4066, 2016. [2](#)