Illumination-Free Photometric Metric for Range Image Registration

Diego Thomas
National Institute of Informatics
Chiyoda, Tokyo, Japan
diego_thomas@nii.ac.jp

Akihiro Sugimoto
National Institute of Informatics
Chiyoda, Tokyo, Japan
sugimoto@nii.ac.jp

Abstract

This paper presents an illumination-free photometric metric for evaluating the goodness of a rigid transformation aligning two overlapping range images, under the assumption of Lambertian surface. Our metric is based on photometric re-projection error but not on feature detection and matching. We synthesize the color of one image using albedo of the other image to compute the photometric re-projection error. The unknown illumination and albedo are estimated from the correspondences induced by the input transformation using the spherical harmonics representation of image formation. This way allows us to derive an illumination-free photometric metric for range image alignment. We use a hypothesize-and-test method to search for the transformation that minimizes our illumination-free photometric function. Transformation candidates are efficiently generated by employing the spherical representation of each image. Experimental results using synthetic and real data show the usefulness of the proposed metric.

1. Introduction

A wide class of 3D object modeling techniques starts with acquiring multiple scans from different poses or viewpoints. The acquired scans are then aligned with each other using rigid transformations to obtain the complete 3D model. Aligning different scans is called registration, and over the years, many studies have been proposed that allow accurate registration ([2, 6, 8, 10]). However, they rely highly on geometry and/or texture of the object to be modeled and inevitably fail in some situations. In particular, registering 3D scans of an object using photometry is still an unresolved problem.

In this paper, we exploit photometry to successfully register two overlapping range images with color of an object surface devoid of salient geometric features under general and distant, unknown lighting. This expands applications such as on-site modeling of man-made objects (whose surfaces often exhibit symmetries). High precision range sensors are heavy and difficult to manipulate. Therefore, when capturing 3D scans of an object's surface, it is in general easier to move the object in front of the sensor rather than moving the sensor around the object. In this work, we thus consider the situation where the object pose changes during the acquisition, while the viewpoint and illumination stay fixed. We assume a Lambertian reflection with no inter-reflection, nor any cast shadows. We notice that the relationship between normals and incident illumination changes with the object's pose, and so does the object's appearance. By using a modern laser range scanner, we can capture both the 3D shape and a color image of an object from a fixed viewpoint.

When an object surface lacks discriminative geometric features, photometric features are used to guide the registration ([6, 15, 19, 20]). In particular, illumination-invariant photometric features are preferable. Without known correspondences between two range images or without known illumination, however, computing such features is not possible. Therefore, the ICP-like approach where features are used to find correspondences does not work under unknown lighting.

Another approach to registration can be found where alignment is achieved by minimizing a cost function ([3, 16]) without using explicit point correspondences. In general, the cost function is based on re-projection error. Defining such functions for 2D images have been widely studied and methods such as the cross-correlation method are available. However defining a photometric metric for 3D images remains an open challenge.

We propose a new illumination-free photometric metric for registering two range images of an object lacking in geometric features and show its usefulness with a practical registration method. The spherical harmonics give a compact yet accurate representation of image formation. For a given transformation and its induced correspondences, this representation allows the estimation of illumination and albedo, both of which are used with one range image to synthesize the color image of the other range image. The difference between the synthesized color image and the cap-

tured color image (we will refer to it as the photometric re-projection error) is independent of illumination and defines our illumination-free photometric metric. This function evaluates the consistency of the relationship between geometry, texture and illumination.

We employ a hypothesis-and-test registration method to demonstrate the usefulness of our proposed photometric metric. Our method carries out registration not by estimating transformations from correspondences but by generating transformations and evaluating them to find the best one. A transformation is defined by a triplet of correspondences, and with about 10^4 points in each range image, we have potentially around 10^{12} possibilities. Directly evaluating each possibility is thus computationally unrealistic, even using RANSAC-like methods. To efficiently generate transformations, we use the spherical representation of each range image. The sphere is homeomorphic to a close-zero genus surface, and a natural representation for a 3D surface. Over the sphere, we generate rigid transformations and evaluate them to reach the best one for the final result.

The contribution of our work is two-fold: (1) we derive an illumination-free photometric metric to evaluate given transformations under general and distant, unknown lighting; and (2) we demonstrate the effectiveness of our photometric metric by implementing a practical registration method using the hypothesis-and-test search strategy. To the best of our knowledge, this is the first work that registers range images devoid of salient geometric features under general and distant, unknown lighting using photometry.

2. Related work

In general, pair-wise registration methods consist of feature detection and description, followed by feature matching (the transformation aligning two range images is then estimated using the obtained correspondences). The iterative closest point (ICP) [2] and its variants are popular methods where points are matched to their closest point. The list of correspondences is iteratively updated until the estimated transformation converges to a stable solution. The major drawback of ICP is that it may get trapped in local minima and thus depends highly on the initial alignment.

Many interest point detectors and feature descriptors have been proposed in the last decade to establish point correspondences using color-like features. In early work, Johnson and Kang [11] or Okatani *et al.* [15] proposed to use color or chromaticity to match points. However, color and chromaticity of a point are not distinctive by themselves and good quality of correspondences cannot be achieved.

The SIFT and SURF descriptors ([6, 13]) are the most popular textural features, which are extensively used for aligning 2D images. However, they are 2D features and thus suffer from projective distortion when used to align range images. Seo *et al.* [9] extended the SIFT feature for range image alignment by accounting for the projective distortion. However, SIFT-like methods do not account for changes in pose-illumination relationship, but rely on the color normalization assumption (i.e. intensity changes uniformly along with changes in illumination, and normalization of features with respect to the overall intensity is sufficient). The above mentioned methods thus all suffer from changes in pose-illumination relationship.

More recently, Thomas *et al.* [19, 20] explored the use of albedo. They proposed a robust feature descriptor based on adaptive regions using albedo. However, this approach assumes known illumination or a few unknown point light sources. In practical situations, however, illumination is more complex and difficult to estimate; this technique is therefore not applicable.

In a broader sense, photometry states for the relationship between texture, geometry and illumination. As a consequence, from a single range image it is not possible to estimate one of the three without knowing the others. Therefore, the ICP-like approach, where features are computed for two range images independently, fails in some situations when illumination is not known.

On the other hand, matching features is not the only way we can take to align range images. Other transformation search methods can be found in the literature where a cost function is minimized over a parameter space. Some use optimization strategy such as Gauss Newton ([3, 5, 8, 16]), and some use the hypothesis-and-test strategy such as RANSAC or brute-force search ([14]). The former is efficient but sensitive to the initial alignment while the latter does not depend on initialization even though it may be less efficient.

Several geometric and 2D textural cost functions have already been explored. However, less work has been done on defining a 3D photometric metric for aligning pairs of range images, and as far as we know no photometric metric independent of illumination has been reported.

3. Illumination-free photometric metric

We introduce our illumination-free photometric metric that does not compare features but compute re-projection error. By doing so, we simultaneously take into account geometry, texture and illumination. Figure 1 illustrates the derivation of our proposed metric.

3.1. Review of spherical harmonics representation

The spherical harmonics have been shown to be a powerful tool to model image formation [17], and, for the Lambertian reflectance, up to the second-order spherical harmonic expansion is known to be sufficient to approximate the image formation with more than 98% accuracy [1].

The irradiance $E(\mathbf{x})$ at a point $\mathbf{x}$ for a distant illumination L and the diffuse reflection kernel R is given by an

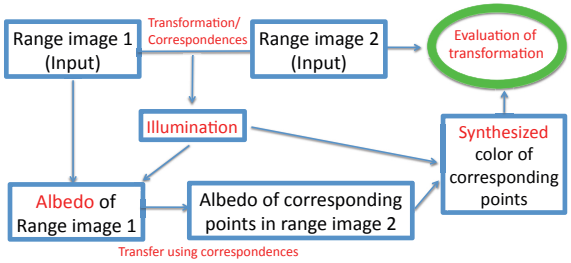


Figure 1. Procedural evaluation of a given transformation.

integral over the sphere.

$$E(\mathbf{x}) = \int_0^{2\pi} \int_0^{\pi} L(\theta,\phi) R(G_{\mathbf{x}}(\theta,\phi)) \sin(\theta) \mathrm{d}\theta \mathrm{d}\phi, \quad (1)$$

where (θ, ϕ) are the incident angles in the global coordinate system and $G_{\mathbf{x}}$ is the transformation from the global coordinate system to the local coordinate system centered around the normal of the point $\mathbf{x}$.

The irradiance at a point $\mathbf{x}$ is scaled by the albedo $\rho(\mathbf{x})$ to have the color of the point: $I(\mathbf{x}) = \rho(\mathbf{x})E(\mathbf{x})$.

The color $I(\mathbf{x})$ of a point $\mathbf{x}$ on the Lambertian surface is approximated as

$$I(\mathbf{x}) \approx \rho(\mathbf{x}) \sum_{l=0}^{2} \sum_{m=-l}^{l} L_{l,m} R_{l,m}(\mathbf{x}), \quad (2)$$

where $L_{l,m}$ and $R_{l,m}(\mathbf{x})$ represent the spherical harmonic coefficients of L and $R \circ G_{\mathbf{x}}$, respectively. We notice that the spherical harmonic coefficients $R_{l,m}(\mathbf{x})$ of the Lambertian reflection kernel $R \circ G_{\mathbf{x}}$ are known as functions of the normal at point $\mathbf{x}$ [1].

3.2. Evaluation metric for a transformation

We use the spherical harmonics representation of image formation to derive our illumination-free photometric evaluation metric for a given rigid transformation. We remark that our photometric metric does not suffer from scale ambiguity that arises when estimating photometric features.

A given transformation between two range images (range images 1 and 2) induces point correspondences across the two images. We estimate illumination from these point correspondences. The estimated illumination (obtained with respect to the given transformation) allows us to compute albedo at points of range image 2, which is transferred to their corresponding points in range image 1. The transferred albedo are used together with the estimated illumination and geometry to synthesize colors of range image 1. The synthesized colors are then compared with the captured colors of range image 1 to evaluate the photometric consistency (i.e. photometric re-projection error) of the alignment induced by the given rigid transformation.

Let T and $(\mathbf{x}_i, T(\mathbf{x}_i))_{i\in[0,n-1]}$ respectively denote a given transformation and the induced point correspondences, where $\mathbf{x}_i \in \mathbf{R}^3$. If T accurately aligns the two range images, then two corresponding points represent the same point of the surface viewed in different poses, and their albedo is the same $\rho(\mathbf{x}_i) = \rho(T(\mathbf{x}_i))$.

R is known and depends on only the surface normals. Therefore, using (2), we can derive a linear system $LM = \mathbf{0}$ with $L_{l,m}$ as unknowns, where $L = [L_{0,0}, L_{1.-1}, L_{1,0}, L_{1,1}, L_{2.-2}, L_{2,-1}, L_{2,0}, L_{2,1}, L_{2,2}]$ and $M = [M_i]_{i\in[0,n-1]}$ with $M_i = [I(\mathbf{x}_i)R_{0,0}(T(\mathbf{x}_i)) - I(T(\mathbf{x}_i))R_{0,0}(\mathbf{x}_i), \ \ldots \ , I(\mathbf{x}_i)\ R_{2,2}(T(\mathbf{x}_i)) - I(T(\mathbf{x}_i))R_{2,2}(\mathbf{x}_i)]^\top$. Note that, L is a row vector in 9D and M is a $9 \times n$ matrix, where n is the number of corresponding points.

The matrix M is known, and we can estimate $\tilde{L}(T)$ using the SVD, up to an unknown scaling factor λ ($\lambda \neq 0$). We can then estimate albedo of each point.

$$\rho(\mathbf{x}) = \frac{1}{\lambda} \left(\frac{I(\mathbf{x})}{\sum_{l=0}^{2} \sum_{m=-l}^{l} \tilde{L}_{l,m}(T) R_{l,m}(\mathbf{x})} \right). \quad (3)$$

A given transformation is not likely correct in most cases. Therefore, contrary to the case where correspondences are assumed to be perfect ([18]), we need to carefully choose an attribute for our evaluation. For example, comparing the estimated albedo of corresponding points is not effective because $\|\rho(\mathbf{x}_i) - \rho(T(\mathbf{x}_i))\|$ is different for $\tilde{L}$ and $\lambda\tilde{L}$, with $\lambda \neq 1$ while $\tilde{L}$ and $\lambda\tilde{L}$ correspond to the equivalent photometric solution. We thus use the captured color images as the ground truth to evaluate the transformation T. This is justified by the fact that the estimated photometric properties should be coherent with the correspondences and the captured images.

Corresponding points $(\mathbf{x}_i, T(\mathbf{x}_i))$ should have the same albedo if T is accurate. We thus synthesize the color of $T(\mathbf{x}_i)$ as follows:

$$\tilde{I}_T(T(\mathbf{x}_i)) = \rho(\mathbf{x}_i) \sum_{l=0}^{2} \sum_{m=-l}^{l} \tilde{L}_{l,m}(T) R_{l,m}(T(\mathbf{x}_i)). \quad (4)$$

We now define our photometric re-projection error of T.

$$Eval(T) = \frac{\sum_{i=0}^{n-1} \|I(T(\mathbf{x}_i)) - \tilde{I}_T(T(\mathbf{x}_i))\|}{n}. \quad (5)$$

We remark that λ in (3) is no longer present in (5).

4. Registration

Given two overlapping range images, we seek a rigid transformation that minimizes the photometric re-projection error (5). When minimizing the photometric re-projection error, we have to decide the strategy we use.

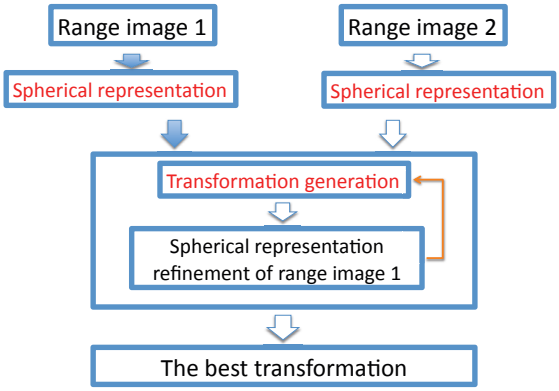


Figure 2. Flowchart of transformation search.

The complex behavior of our photometric re-projection error makes optimization methods difficult to use. First, our proposed metric may not even be continuous over the space of all rigid transformations because of the discontinuities in the albedo values of the object's surface. Second we do not rely on an initial guess for the alignment, and falling into local minima may be inevitable. Therefore, we choose the hypothesis-and-test search. Figure 2 illustrates the flowchart of our proposed registration method.

4.1. Transformation search

The hypothesis-and-test search is performed by testing the quality of the registration for a set of rigid transformation candidates. The search ends when a rigid transformation accomplishing accurate registration is found or when all candidates are tested (the rigid transformation minimizing the cost function is then selected).

The most famous hypothesis-and-test search is the RANSAC method where candidates are generated from random triplets of correspondences. Straightforwardly using RANSAC is computationally unrealistic. This is because we have potentially around 10^{12} possibilities for range images with 10^4 points. Therefore, how to efficiently search the best rigid transformation aligning range images becomes a critical issue.

On one hand, rigid transformations aligning two range images can be equivalently represented by sets of rigid point correspondences induced by the transformations. Therefore, if we represent the range image in another domain while keeping the rigidity of point correspondences, we can discuss the problem of searching the best transformation aligning the range images in this new domain.

On the other hand, the unit sphere is a convenient representation of a close-zero genus 3D surface. For closed surfaces[1], the spherical representation is pose invariant [22]. Therefore, the local structure in the spherical domain does not change and the rigidity of point correspondences is kept.

In addition, the rigid transformations aligning two spheres (with the same radius and the same center) belong to $SO(3)$. We thus employ the spherical representation for range images. This representation reduces the transformation parameter space from $SO(3) \times \mathbf{R}^3$ to $SO(3)$.

The spherical representation of range images is, unfortunately, not pose invariant because surfaces in a range image are not closed. As a consequence, the local structure in the spherical domain may change in the original domain. This means that the rigidity of point correspondences in the spherical domain may not be kept in the original domain. To overcome this problem, we introduce refinement of the spherical representations throughout the registration process to reduce changes of the local structure in the spherical domain as much as possible. Due to the possibility of violating rigidity of point correspondences in the original domain, we also have to generate the rigid transformations in the original domain from the point correspondences obtained in the spherical domain. For this, we use the method proposed by Horn [7] as follows. A 3D rotation in $SO(3)$ gives us point correspondences in the spherical domain. In the original domain, we use the same point correspondences as the input of [7] to obtain the corresponding rigid transformation in $SO(3) \times \mathbf{R}^3$.

We remark that though we can use RANSAC to generate transformation candidates from $SO(3)$ we prefer to use an exhaustive search to ensure convergence to the optimal solution. To reduce the computational time, we reduce $[0 : 360]^3$ to $[0 : \frac{20}{step}]^3$, where $step$ increases during the iteration[2]. In the experiments, we set $step = 1$ at the first iteration and then $step = step + 3$.

4.2. Spherical representation and refinement

Our spherical representation method is inspired by the method proposed in [22] for closed surfaces. The input is an unorganized point set represented in the global coordinate system, and the output is a structured mesh with corresponding coordinates on the unit sphere that preserves the local structure. We note that the spherical representation of each range image is computed independently. We also note that because we refine the representation throughout the registration process, the quality of the initial spherical representation does not affect the quality of the final registration. This is because our illumination-free re-projection error is always evaluated in the original domain.

Spherical representation. We first orthogonally project all the 3D points of a range image along the z axis (viewing direction) to a plane. We then compute the convex hull of the projected points and identify the vertices of the convex

[1] A closed surface is defined as a surface that is compact and without a boundary; a non-closed surface is a compact surface with a boundary.

[2] The computation can be done in parallel for a fast implementation.

Algorithm 1 Progressive spherical representation

Require: Range image RI represented using the barycentric coordinate system

Ensure: spherical representation S of RI

$Mesh_0 \leftarrow$ triangulated 2D convex hull of RI
$PtMesh \leftarrow$ set of vertices belonging to $Mesh_0$
$S_0 \leftarrow$ projection of $Mesh_0$ to the sphere
$List \leftarrow$ list of points in RI but not in $PtMesh$
$nb \leftarrow$ size of $List$
for $i = 0$ to $nb - 1$ **do**
  $\mathbf{\Psi} \leftarrow$ a point of $List$
  $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \leftarrow$ three points of the enclosing triangle of $\mathbf{\Psi}$ in $Mesh_i$
  $Mesh_{i+1} \leftarrow$ Delaunay triangulation of $Mesh_i$ where $\mathbf{\Psi}$ has been added
  $P \leftarrow$ polygon in $Mesh_{i+1}$ containing the points $(\mathbf{\Psi}, \mathbf{a}, \mathbf{b}, \mathbf{c})$
  $U \leftarrow$ flatten polygon obtained by conformal mapping of P [12]
  $(\alpha, \beta, \gamma) \leftarrow$ barycentric coordinates of $\mathbf{\Psi}$ in U for $(\mathbf{a}, \mathbf{b}, \mathbf{c})$
  $\mathbf{\Psi}' \leftarrow \alpha\mathbf{a}' + \beta\mathbf{b}' + \gamma\mathbf{c}'$
  $S_{i+1} \leftarrow S_i + \{\mathbf{\Psi}'\}$, with the same connectivity as in $Mesh_{i+1}$
  $List \leftarrow List - \{\mathbf{\Psi}\}$
end for
return S_n

hull. The vertices are used to generate Delaunay triangulations. The set of 3D points in the range image corresponding to the vertices is then projected to the unit sphere by normalizing the coordinates of each point. Next, we select a point (in the range image) that is not included in the vertices of the convex hull and carry out the following process: we progressively construct triangulations by adding the point and compute the local position of the point with respect to the new triangulation. The local position is computed in the flattened vicinity of the point, obtained using conformal mapping [12], to accurately represent the local structure. The point is then positioned on the sphere using this local position. This series of processes is carried out until all the points in the range image are involved. The concrete procedure is described in Algorithm 1, and the progressive spherical representation is illustrated in Fig.3.

Spherical representation refinement. The overlapping areas between two range images from the current best transformation are first identified. The bijection B between points of the two overlapping areas is then computed. Namely, for a point $\mathbf{x}$ in the overlapping area O_1 of range image RI_1, $B(\mathbf{x}) = closest(\mathbf{x})$ if $\mathbf{x} = closest(closest(\mathbf{x}))$, $B(\mathbf{x})$ is undefined otherwise. Here, $closest$ stands for the closest point (in the sense of the Euclidean distance) in the overlapping area of the other range image. Then, for each point of O_1, its coordinates on the sphere are set to those of its closest point. The remaining points of O_1 that do not have an image for B are placed on the sphere using the local positions as we did in the spherical representation above.

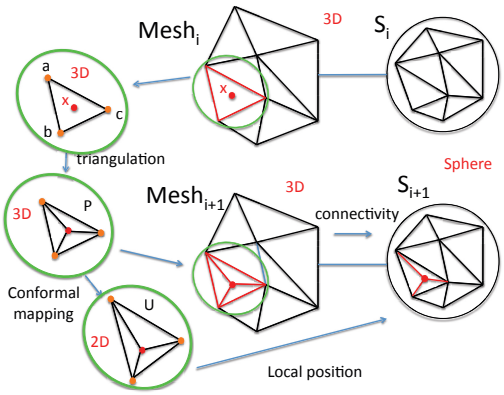


Figure 3. Illustration of the loop "for" of Algorithm 1.

5. Experiments

We evaluate our algorithm in several challenging situations using both synthetic and real data. For the comparison, we used three methods: the proposed method using chromaticity instead of our evaluation function (Method 1); the method in [20] using albedo with a given directional light source (Method 2), and the method [20] using chromaticity instead of albedo (Method 3). Whereas comparing our method with Method 1 shows the advantage of using photometry, comparing it with Methods 2 and 3 shows the advantage of our search strategy. We also compared our method with the alignment obtained by matching using SIFT (we used the available code provided by Andrea Vedaldi [21]). We notice that all data are devoid of salient geometric features, and thus, using geometric feature-based registration methods does not work in these cases.

Along with [20], we evaluated the registration results using the distance between the estimated transformation and the ground truth transformation obtained manually:

$$err = \frac{\Theta d + \|T_g - T_e\|}{res}, \tag{6}$$

where Θ is the angle between the normalized ground truth rotation quaternion and the estimated normalized rotation quaternion, d is the depth of the 3D scans, T_g and T_e are the ground truth translation and the estimated translation respectively and res stands for the resolution of the 3D scans. The unit of this measure is the resolution of the 3D scans.

Since range data are acquired by changing the poses of an object, even points in the overlapping area may not cor-

respond exactly. This is due to different digitization of the overlapping area depending on the pose. We thus first identify a stable point, i.e., a point whose photometric features are sufficiently similar to those of its corresponding point even though the correspondence may not be exact. We then use only stable points to evaluate our photometric function.

Stable points identification. We extract stable points independently for two range images as a pre-processing step. In a small vicinity and for diffuse reflection, the difference in chromaticity approximates well the difference in albedo. Accordingly, we define a stable point using both difference of chromaticity and difference of normals in a small vicinity. Namely, a point $\mathbf{x}$ is identified to be stable if

$$\begin{aligned} \forall \mathbf{y} \text{ such that } \|\mathbf{y}-\mathbf{x}\| < \epsilon_s, \\ \|c(\mathbf{x})-c(\mathbf{y})\| < \epsilon_c \text{ and } \|n(\mathbf{x})-n(\mathbf{y})\| < \epsilon_n, \end{aligned} \quad (7)$$

where $\mathbf{y}$ is a point in the range image, c is chromaticity, n represents the surface normals, and ϵ_s, ϵ_c, and ϵ_n are three thresholds. ϵ_s, ϵ_c, and ϵ_n were set to $3 * res$, 0.02, and 0.1 respectively for all experiments with synthetic data and to $2 * res$, 0.05, and 0.2 for all experiments with real data.

5.1. Synthetic data

The synthetic data, called *vase* were obtained using a 3D modeler software (3D Studio Max) (see Table 1). The exact albedo is known and we simulated lighting under different illuminations. This data set is challenging for registration in that the shape is rotationally symmetric, the texture of the objects presents several repetitive patterns, and no exact correspondences exist between the two range images.

The first set-up is illustrated in Fig. 4. The range images were illuminated by the light probe *galileo* from the Debevec database [4], and the color was synthesized using spherical harmonics with a Lambertian reflection kernel. Figure 5 shows the results obtained with the five methods mentioned above. In this situation, the illumination is dominated by ambient light, and the changes in appearance between the two range images are small. This situation is thus well adapted to using chromaticity. Nevertheless, our evaluation function worked well compared with Method 1. Moreover, because of repetitive textures and distant initial positions, Methods 2 and 3, i.e., methods using an ICP-like framework, fell into a local minimum and failed pitifully in aligning the range images, even when using chromaticity. Our proposed search strategy, on the contrary, efficiently found the global minimum. The estimated transformation T_e by our proposed method was a rotation of $(-19.90; 0.001, 0.936, 0.350)$ and a translation of $(0.19, -0.01, 0.03)$, and $Eval(T_e)$ was[3] 1.2. It was a rotation of $(-19.89; 0.000, 0.936, 0.350)$ and a translation of $(0.19, -0.01, 0.03)$ when using chromaticity instead of our evaluation function.

Table 1. Description of the data *vase*.

Nb_Points	Resolution	Expected_rot (angle; axis)
30650	0.01mm	$(-20.00; 0.006, 0.933, 0.339)$
Expected_translation		$(0.19, -0.01, 0.03)$

Figure 4. The first set-up.

Method used	Proposed method	Method 1	Method 2	Method 3	SIFT
Error	0.56	0.61	41.51	41.53	15.62

Figure 5. Results obtained with the five methods for the data *vase*.

The second set-up is illustrated in Fig. 6. The illumination was composed of three directional light sources of different intensities as well as an ambient light source. The images were rendered using the standard Lambertian model. Figure 7 shows the results obtained with the five methods. In this situation, the illumination induces significant changes in the object appearance (e.g. the color of several points changed from reddish to white), and the use of chromaticity is no longer effective for both our search strategy and ICP-like methods. Method 2 failed in aligning the range images because it assumes that the object is illuminated by only a single light source. Our method is the only one that achieved accurate alignment (similar to the one obtained in the previous set-up). The estimated transformation T_e by our proposed method was a rotation of $(-20.12; 0.001, 0.936, 0.349)$ and a translation of $(0.19, -0.01, 0.03)$, and $Eval(T_e)$ was 3.4.

[3]The color was coded in RGB with values between 0 and 255. Since the color of a point is approximated with 98% accuracy in Eq. (2), the re-projection error of around 1.2 means that the optimal solution is found provided that the distribution of RGB values is uniform over the range.

Figure 6. The second set-up.

Proposed method	Method 1	Method 2	Method 3	SIFT
Method used	Proposed method	Method 1	Method 2	Method 3	SIFT
Error	0.51	41.02	41.48	41.51	12.03

Figure 7. Results obtained with the five methods for the data *vase*.

Results obtained using SIFT key-point detector and descriptor for estimating point correspondences are also illustrated in both Fig. 5 and Fig. 7. Due to the projective deformations, changes in intensity and repetitive patterns, the SIFT-based method did not work in our situation.

5.2. Real data

We employed a Konica Minolta Vivid 910 range scanner, which captures the 3D shape and the texture of an object. The ground truth transformation was obtained manually. We note that there is a gamma correction factor in the obtained color images which should preferably be cancelled. In our experiments, however, we did not know this factor, and thus the gamma correction was not cancelled.

We obtained two range images of a rotationally symmetric can that is approximately 10cm high and has a diameter of about 5cm (Fig. 8). Details on the data called *can* are given in Table 2, and the results are shown in Fig. 9. These data are challenging in that the quality of the image is low, and there is an unknown gamma correction factor. Moreover, this data exhibits several repetitive patterns such as similar letters while the texture is either red or white with large uniform areas. Nevertheless, our proposed method accurately registered the two range images. The obtained accuracy was under the resolution of the range sensor, and our method worked extremely well compared to the other methods. The estimated transformation T_e obtained with our proposed method was a rotation of $(19.66; 0.015, 0.941, 0.338)$ and a translation of $(8.93, 0.37, -1.19)$, and $Eval(T_e)$ was 2.56. We obtained a rotation of $(19, 23; 0.028, 0.952, 0.303)$ and a translation of $(9.06, -0.18, -1.11)$ by Method 1.

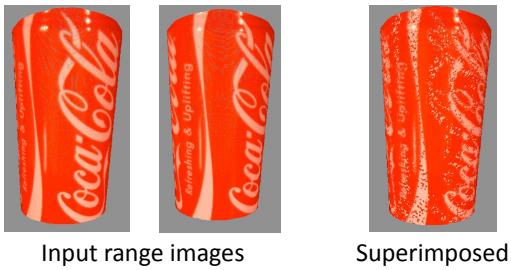
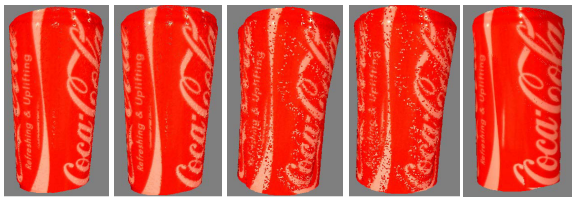


Figure 8. Input range images and initial positions for the data *can*.



Proposed method	Method 1	Method 2	Method 3	SIFT
Method used	Proposed method	Method 1	Method 2	Method 3	SIFT
Error	0.94	1.58	31.85	30.41	15.97

Figure 9. Results obtained with the five methods for the data *can*.

Another data item called *hand* is presented in Fig. 10 and in Table 3. Registration results are shown in Fig. 11. For this data, the intensity of a point in two range images changed drastically (e.g. points at the middle of the images). Therefore, the use of chromaticity to evaluate the goodness of transformations becomes unreliable. Our method is the only one that achieved accurate registration of the two range images. The gap in accuracy between our proposed method and Method 1 became larger than that for the data *can*. This is because drastic changes in intensity degrade reliability of chromaticity while our method uses an illumination-free metric. Method 2 and Method 3 still failed pitifully.

Results obtained using SIFT key-point detector and descriptor for estimating point correspondences are also illus-

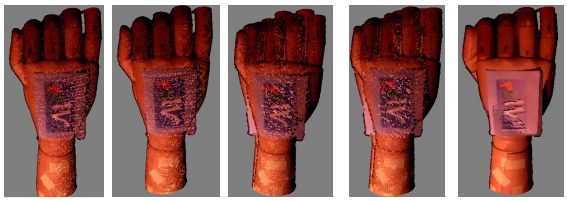
Table 2. Description of the data *can*.

Nb_Points	Resolution	Expected_rot (angle; axis)
28000	0.55mm	$(20.00; 0.010, 0.930, 0.340)$
Expected_translation		$(9.00, 0.10, -1.00)$

Table 3. Description of the data *hand*.

Nb_Points	Resolution	Expected_rot (angle; axis)
50000	0.55mm	$(20.00; 0.010, 0.930, 0.340)$
Expected_translation		$(6.60, -2.3, -0.40)$

Figure 10. Input range images and initial positions for data *hand*.



Method used	Proposed method	Method 1	Method 2	Method 3	SIFT
Error	1.01	5.57	30.61	38.72	11.45

Figure 11. Results obtained with the five methods for data *hand*.

trated in both Fig. 9 and Fig. 11. Due to the repetitive patterns for the data *can* and significant changes in intensity for the data *hand*, the SIFT-based method did not work well.

6. Conclusion

We introduced a new illumination-free photometric metric for registering range images of Lambertian surfaces under general and distant, unknown lighting. Our metric evaluates photometric re-projection error by taking into account the relationship between texture, geometry and illumination. We use captured color of range images as the ground truth to eliminate scale ambiguity that arises when estimating photometric features. We also demonstrated the effectiveness of our metric by using the hypothesis-and-test strategy for the registration where a range image is represented over the sphere and its representation is refined throughout the registration process. The experimental results showed that our method significantly outperformed existing, feature matching based methods under distant, general and unknown illumination.

References

[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspace. *IEEE Trans. on PAMI*, 25(2):218–233, 2003. 98, 99

[2] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Trans on PAMI*, 14(2):239–256, 1992. 97, 98

[3] D. Breitenreicher and C. Shnörr. Intrinsic second-order geometric optimization for robust point set registration without correspondence. *In Proc. of EMMCVPR2009*, pages 274–287, 2009. 97, 98

[4] P. Debevec. *Light Probe Image Gallery*, 2004. http://ict.debevec.org/~debevec/Probes/. 102

[5] B. Gutman, Y. Wang, T. Chan, P. M. Thompson, and A. W. Toga. Shape registration with spherical cross correlation. *In 2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy*, pages 56–67, 2008. 98

[6] H.Bay, T.Tuytelaars, and L. van Gool. Surf: Speeded up robust features. *In Proc. of ECCV'06*, pages 404–417, 2006. 97, 98

[7] B. Horn. Closed-form solution of absolute orientation using orthonormal matrices. *J. Opt. Soc. Amer. A*, 5(7):1127–1135, 1987. 100

[8] B. Jian and B. C. Vemuri. A robust algorithm for point set registration using mixture of gaussian. *In Proc. of ICCV'05*, 2:1246–1251, 2005. 97, 98

[9] G. S. J.K. Seo and S. Lee. Range data registration using photometric features. *In Proc. of CVPR'05*, 2:1140–1145, 2005. 98

[10] A. Johnson and M. Hebert. Surface registration by matching oriented points. *In Proc. of 3DIM'97*, pages 121–128, 1997. 97

[11] A. Johnson and S. Kang. Registration and integration of textured 3d data. *Image and vision computing*, 17(2):135–147, 1999. 98

[12] A. Lee, W. Sweldens, P. Shroder, L. Cowsar, and D. Dobkin. Maps: multiresolution adaptive parameterization of surfaces. *In Proc. of SIGGRAPH'98*, pages 343–352, 1998. 101

[13] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60:91–110, 2004. 98

[14] F. J. O. Enqvist and F. Kahl. A brute-force algorithm for reconstructing a scene from two projections. *In Proc. of CVPR'11*, 2011. 98

[15] I. Okatani and A. Sugimoto. Registration of range images that preserves local surface structure and color. *In Proc. of 3DPVT'04*, pages 786–796, 2004. 97, 98

[16] C. Papazov and D. Burschka. Stochastic optimization for rgid point set registratio. *In Proc. of ISVC'09*, pages 1043–1054, 2009. 97, 98

[17] R. Ramamoorthi. Modeling illumination variation with spherical harmonics. *Face Processing: Advanced Modeling Methods*, pages 385–424, 2006. 98

[18] T. Tachikawa, S. Hiura, and K. Sato. Robust estimation of light directions and diffuse reflectance of known shape object. *In Proc. of Vision, Modeling and Visualization Workshop*, pages 37–44, 2009. 99

[19] D. Thomas and A. Sugimoto. Range image registration of specular objects under complex illumination. *In Proc. of 3DPVT'10*, 2010. 97, 98

[20] D. Thomas and A. Sugimoto. Robustly registering range images using local distribution of albedo. *Computer Vision and Image Understanding*, 115:649–667, 2011. 97, 98, 101

[21] A. Vedaldi. *SIFT code for Matlab*, 2006. http://www.vlfeat.org/~vedaldi/code/sift.html. 101

[22] K. Zhou, H. Bao, and J. Shi. 3D surface filtering using spherical harmonics. *Computer Aided Design*, 36:363–375, 2004. 100