Range Image Registration Using a Photometric Metric under Unknown Lighting

Diego Thomas, *Member, IEEE*, and Akihiro Sugimoto, *Member, IEEE*

Abstract—Based on the spherical harmonics representation of image formation, we derive a new photometric metric for evaluating the correctness of a given rigid transformation aligning two overlapping range images captured under unknown, distant, and general illumination. We estimate the surrounding illumination and albedo values of points of the two range images from the point correspondences induced by the input transformation. We then synthesize the color of both range images using albedo values transferred using the point correspondences to compute the photometric reprojection error. This way allows us to accurately register two range images by finding the transformation that minimizes the photometric reprojection error. We also propose a practical method using the proposed photometric metric to register pairs of range images devoid of salient geometric features, captured under unknown lighting. Our method uses a hypothesize-and-test strategy to search for the transformation that minimizes our photometric metric. Transformation candidates are efficiently generated by employing the spherical representation of each range image. Experimental results using both synthetic and real data demonstrate the usefulness of the proposed metric.

Index Terms—Range image, registration, photometry, spherical harmonics, photometric reprojection

◆

1 INTRODUCTION

A**UTOMATING** the 3D modeling process of real objects is a topic of major importance. Three-dimensional models of real objects are widely used in the industry for applications ranging from digitalization of cultural heritage to medical imagery, but also for postproduction verification or situational awareness, to cite a few.

Recent laser scanning techniques allow the acquisition of high-accuracy range images with color attached to each point. Since only a part of the object is visible from one viewpoint, a wide class of 3D modeling techniques starts by acquiring multiple range images from different viewpoints or for different object poses to ensure that full coverage of the object's surface is captured. All range images then have to be aligned together before being merged and integrated into a common 3D model. This is because each range image is represented in the local coordinate system defined by the sensor position and orientation. The process of aligning pairs of range images is called pairwise range image registration.

Let us here clarify our situation of interest. Recently, with the development of RGB-D video cameras, a new trend of work has emerged on registering and fusing dense sequences of depth images [9], [11]. Our situation, however, assumes using a high-accuracy laser scanner for constructing a high-accuracy 3D model of a single object. Because of the time of acquisition and heavy experimental manipulations for such a laser scanner, it is in general

preferred to capture as few range images as possible. This is why we focus on pairwise registration only. Since we are considering only two range images, fusion techniques such as the volumetric method proposed in [11], which are used in the frame-to-global-model registration strategy, are not applicable in our case. Moreover, note that the output data obtained with a laser scanner and an RGB-D video camera are different. The first difference is the accuracy in the depth measurements. The second difference is that a laser scanner can be precisely tuned to focus on a desired area of the scene. This allows discarding points that are away from the object of interest. It is even possible to directly acquire nicely segmented out objects with a laser scanner. RGB-D sensors such as the Microsoft Kinect camera do not have this property and provide, in general, cluttered scenes containing not only the object of interest but also most of the surrounding points. In this paper, we focus on registering range images acquired with a laser scanner and thus we do not consider either cluttered scenes or background.

A concrete scenario may be, for example, the 3D modeling of archeological objects. For study and/or analysis of archeological objects, it is important to obtain high-accuracy 3D models that fit well to the real shape of the object. As a consequence, laser scanners (which are still the most accurate depth sensors) are in general used to capture data of high quality. By putting a target object on a black table and tuning the focus of the laser scanner, it then becomes possible to capture range images of a single, cleanly segmented out object that can be directly input for registration.

Many studies have been proposed that allow accurate pairwise registration of range images ([2], [3], [12], [13]). However, they highly rely on the geometry and/or texture of the object to be modeled and inevitably fail in some situations. In particular, a challenge arises when the input

---

● *The authors are with the National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan.*
E-mail: {diego_thomas, sugimoto}@nii.ac.jp.

range images, captured under unknown lighting, lack discriminative geometric features. Such a situation is likely to occur when modeling man-made objects outside of the laboratory. A concrete scenario may be the on-site digitalization of cultural heritage such as ancient pottery (which are likely to exhibit strong symmetries in their shape).

When an object's surface lacks discriminative geometric features, textural features (derived from the object's textured surface) are often used to guide the registration ([2], [17]). However, under general illumination conditions, the assumptions required for using textural features (such as color consistency or color normalization) do not hold true, and using illumination-invariant features such as reflectance properties (albedo for the diffuse reflectance) becomes preferable [23]. Without known correspondences between two range images or without known illumination, however, computing such features is not possible. Therefore, the feature-based 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 [4], [18] without using explicit point correspondences. In general, the cost function is based on reprojection error. Defining such functions for 2D images has been widely studied and famous methods such as the cross-correlation method are available. However, defining such functions for range images remains an open challenge.

Some works exist in which the cost function is derived using a combination of robust geometric metrics and robust textural metrics (those textural metrics are, in general, direct extensions of the 2D textural metrics) [17]. If the combination is done independently, such as a weighted sum of geometric entities and textural ones, however, the results are not completely satisfactory in that it requires fine tuning and still presents several limitation cases. Therefore, there is a need to define a metric that accounts for both entities concurrently and that is supported by the theory. Photometry, which is the theory that relates geometry, reflectance properties, and incident illumination, is well suited for our purpose.

We investigate the use of photometry for 3D registration and propose: 1) a novel photometric metric for evaluating the correctness of a given transformation and 2) a practical registration method. We consider the situation where the object pose changes during acquisition while the viewpoint and illumination stay fixed. We assume a Lambertian reflection with no interreflections nor any cast shadows. We note that when the object's pose changes while the illumination stays fixed, its appearance in both shape and color changes. By contrast with state-of-the-art methods that strive to cancel lighting to independently combine geometric and textural entities, our proposed method makes use of lighting by concurrently taking into account geometric and textural entities to gain discriminative power that facilitates the search for the best alignment.

The spherical harmonics give a compact yet accurate representation of image formation. For a given transformation and its induced point 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 images and the captured color images (we will refer to it as the photometric reprojection error) defines our photometric metric without using any a priori information on the incident lighting. This function evaluates the consistency of the relationship between geometry, texture, and illumination. Note that not only the color of corresponding points but also normal at points are used concurrently to produce the synthesized images. In the extreme case, this allows us to work even when the surface of the object is untextured.

We use 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 point correspondences but by generating transformations and evaluating them to find the best one. Directly evaluating each possible transformation is computationally unrealistic, even using RANSAC-like methods [5]. To efficiently generate transformations, we use the spherical representation of each range image. 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 twofold: 1) We derive a photometric function 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. We remark that a part of this work has been reported in [24].

2 RELATED WORK

In general, pairwise registration methods can be divided into two categories: 1) the one that searches for the best correspondences between points of the two range images to find the best transformation, 2) the one that searches for the best transformation by minimizing a cost function. While the methods that fall into the first category are fast in general, methods falling into the second category are, in general, more robust and accurate.

2.1 Registration Using Point Correspondences

In general, methods searching for the best correspondences 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) [3] 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 the ICP-like approach 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 textural features. In early work, Johnson and Kang [14] or Okatani and Sugimoto [17]

proposed using color or chromaticity to match points. However, the color and chromaticity of a point are not distinctive by themselves and good quality of correspondences cannot be achieved.

The SIFT and SURF descriptors ([2], [16]) are popular textural features which are extensively used for aligning 2D images. Seo et al. [21] 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 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 the pose-illumination relationship.

Thomas and Sugimoto [23] proposed using albedo (which is invariant to pose-illumination relationship) to register pairs of range images devoid of salient geometric features. However, this approach assumes simple illumination (a single point light source), which is rarely the case in a practical situation.

In a broader sense, photometry states the relationship between geometry, reflectance properties, and incident 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 feature-based approach, where features are computed for two range images independently, fails in some situations when illumination is not known.

2.2 Registration by Minimizing a Cost Function

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 the optimization strategy such as Gauss-Newton ([4], [8], [12], [18]), and some use the hypothesis-and-test strategy such as RANSAC or brute-force search ([7]). In [8], for example, the similarity measure is defined as the cross correlation of the spherical representations of surfaces, and it is customized according to the surface-intrinsic attributes while the spherical harmonics speed up the optimization process. The optimization method is fast but sensitive to the initial alignment, while the hypothesis-and-test strategy does not depend on initialization, even though it may be slower.

Several geometric cost functions as well as 2D textural cost functions have already been explored. However, to the best of our knowledge, there is not yet work done on defining a 3D photometric metric for aligning pairs of range images, and as far as we know, no photometric metric under unknown lighting has been reported.

3 The Photometric Metric

We introduce our photometric metric under unknown lighting that does not compare features but computes reprojection error. By doing so, we simultaneously take into account geometry, reflectance properties, and illumination to derive a metric that makes full use of photometry. Fig. 1 illustrates the derivation of our proposed metric.
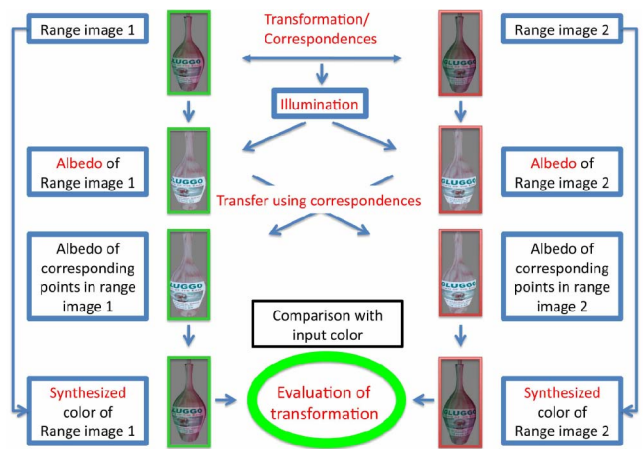

Fig. 1. Procedural evaluation of a given transformation.

3.1 Review of Spherical Harmonics Representation

The spherical harmonics have been shown to be a powerful tool to model image formation [19], 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 percent accuracy [1]. We briefly recall the principles of the spherical harmonics representation and refer to [1] for more details.

We consider a convex, Lambertian object illuminated by distant isotropic lights. The intensity of the reflected light is known to be a function of the normal and albedo. Namely, according to the Lambert's law and for an incident light ray of intensity δ, the intensity of the reflected ray at a point $\boldsymbol{x}$, is $\delta \max(\cos(\theta), 0)$, where θ is the angle between the incident light ray and the normal at the surface at point $\boldsymbol{x}$. Then, the irradiance $E(\boldsymbol{x})$ at a point $\boldsymbol{x}$ for a distant global illumination L and the diffuse reflection kernel[1] R is given by an integral over the sphere:

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

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

The irradiance at a point $\boldsymbol{x}$ is then scaled by the albedo $\rho(\boldsymbol{x})$ to have the color of the point: $I(\boldsymbol{x}) = \rho(\boldsymbol{x})E(\boldsymbol{x})$. The spherical harmonics allow a compact representation of the image formation. Namely, the color $I(\boldsymbol{x})$ of a point $\boldsymbol{x}$ on the Lambertian surface is approximated as

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

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

---

1. $R(\theta) = \max(\cos(\theta), 0)$.

3.2 Evaluation Metric for a Transformation

We use the spherical harmonics representation of image formation to derive our 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; neither requires any a priori information on the incident illumination.

A given transformation between two range images (range images 1 and 2) induces point correspondences across the two images. We use the spherical harmonics representation of image formation to derive a linear system from the point correspondences with the illumination as unknown. The estimated illumination then allows us to compute albedo values at points of the range images, which are transferred to their corresponding points. The transferred albedo values are used together with the estimated illumination and geometry to synthesize colors of the two range images. The synthesized colors are then compared with the captured colors of the two range images to evaluate the photometric consistency (i.e., photometric reprojection error) of the alignment induced by the given rigid transformation.

Let T denote a given transformation and

$$(\boldsymbol{x}_i, \Gamma(T(\boldsymbol{x}_i)))_{i \in [0, n-1]} \text{ and } (\Gamma(T^{-1}(\boldsymbol{y}_i)), \boldsymbol{y}_i)_{i \in [0, m-1]}$$

denote the induced point correspondences, where $\boldsymbol{x}_i \in \boldsymbol{R}^3$ belongs to range image 1 (denoted as I^1), $\boldsymbol{y}_i \in \boldsymbol{R}^3$ belongs to range image 2 (denoted as I^2), and Γ denotes the point correspondences identification operator (see Section 3.3). 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 $\rho(\boldsymbol{x}_i) = \rho(\Gamma(T(\boldsymbol{x}_i)))$ (similarly, $\rho(\boldsymbol{y}_i) = \rho(\Gamma(T^{-1}(\boldsymbol{y}_i)))$).

R is known and depends on only the surface normals. Therefore, using (2), we can derive a linear system $LM = 0$ with L as unknown, where $L = [L_{0,0}, L_{1,-1}, L_{1,0}, L_{1,1}, L_{2,-2}, L_{2,-1}, L_{2,0}, L_{2,1}, L_{2,2}]$ is a row vector in 9D and $M = [M_i]_{i \in [0, n+m-1]}$ is a $9 \times (n+m)$ matrix, where n and m are the number of corresponding points from range images 1 and 2, respectively, with

$$\begin{cases} M_i = [I^1(\mathbf{x}_i)R_{0,0}(\Gamma(T(\mathbf{x}_i))) - I^2(\Gamma(T(\mathbf{x}_i)))R_{0,0}(\mathbf{x}_i), \\ \qquad \ldots, I^1(\mathbf{x}_i)R_{2,2}(\Gamma(T(\mathbf{x}_i))) - I^2(\Gamma(T(\mathbf{x}_i))) \\ \qquad R_{2,2}(\mathbf{x}_i)]^\top \qquad\qquad (\text{if } i < n), \\ M_i = [I^2(\mathbf{y}_i)R_{0,0}(\Gamma(T^{-1}(\mathbf{y}_i))) - I^1(\Gamma(T^{-1}(\mathbf{y}_i))) \\ \qquad R_{0,0}(\mathbf{y}_i), \ldots, I^2(\mathbf{y}_i)R_{2,2}(\Gamma(T^{-1}(\mathbf{y}_i))) \\ \qquad - I^1(\Gamma(T^{-1}(\mathbf{y}_i)))R_{2,2}(\mathbf{y}_i)]^\top \quad (\text{if } n \le i < n+m). \end{cases}$$

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

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

We need to carefully choose an attribute for our evaluation. For example, comparing the estimated albedo of corresponding points is not effective. This is because the

photometric solution for a given transformation has scale ambiguity, and regardless of the relationship between geometry, illumination, and albedo, a solution with a small-scaled albedo always gives better results. Namely, the reprojection error $\|\tilde{L}(T)M\|$ or the residual error in albedo $\|\rho(\boldsymbol{x}_i) - \rho(\Gamma(T(\boldsymbol{x}_i)))\|$ 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 if T is accurate.

Corresponding points $(\boldsymbol{x}_i, \Gamma(T(\boldsymbol{x}_i)))$ should have the same albedo if T is accurate. We thus synthesize the color of $\boldsymbol{x}_i$ in range image 1 as follows:

$$\tilde{I}_T^1(\boldsymbol{x}_i) = \rho(\Gamma(T(\boldsymbol{x}_i))) \sum_{l=0}^{2} \sum_{m=-l}^{l} \tilde{L}_{l,m}(T) R_{l,m}(\boldsymbol{x}_i). \qquad (4)$$

Similarly, we synthesize the colors $\tilde{I}_T^2(\boldsymbol{y}_i)$ of points $\boldsymbol{y}_i$ in range image 2.

We now define our photometric reprojection error of T:

$$Eval(T) = \frac{\sum_{i=0}^{n-1} \|I^1(\boldsymbol{x}_i) - \tilde{I}_T^1(\boldsymbol{x}_i)\| + \sum_{i=0}^{m-1} \|I^2(\boldsymbol{x}_i) - \tilde{I}_T^2(\boldsymbol{x}_i)\|}{n+m}. \qquad (5)$$

We remark that the unknown scaling factor λ that arises when estimating albedo in (3) is no longer present in (5).

We notice that the shape of $Eval$ depends on T. Namely, $\tilde{I}_T^1$ and $\tilde{I}_T^2$ change even for the same point, depending on T. Thus, the derivation of $Eval$ is procedural and we do not have an analytical formula for the function. In addition, $\tilde{I}_T^1$ and $\tilde{I}_T^2$ are only piecewise continuous with sufficiently similar transformations. This is because the distribution of albedo over the surface is only piecewise continuous. As a consequence, the values of the entries of the matrix M in the linear system defined above vary piecewise continuously with sufficiently similar transformations and so does the estimated photometric properties as well as the synthesized images.

3.3 Point Correspondences Identification

Though it is a simple task, identifying the point correspondences from the given transformation T is the most time-consuming one for our evaluation metric. It is thus of major importance to perform it as fast as possible. We use *projective data association* [20] to realize fast point correspondences estimation.

For two range images I^1 and I^2, their corresponding depth maps D_1 and D_2 with the associated intrinsic matrix K, and the given transformation T aligning I^1 to I^2, the corresponding point $\Gamma(\boldsymbol{x}) \in I^2$ of a point $\boldsymbol{x} \in I^1$ is identified as follows: 1) $\boldsymbol{x}$ is transformed into the camera coordinate system of I^2 ($\boldsymbol{y} = T\boldsymbol{x}$), 2) the point $\boldsymbol{y}$ is perspective projected into image coordinates ($(i, j, 1) = K\boldsymbol{y}$), 3) $\Gamma(\boldsymbol{x})$, the closest point in I^2 of $\boldsymbol{x}$, is then identified as the point associated to the pixel (i, j) in D_2. Searching for the closest points from I^2 to I^1 is carried out similarly.
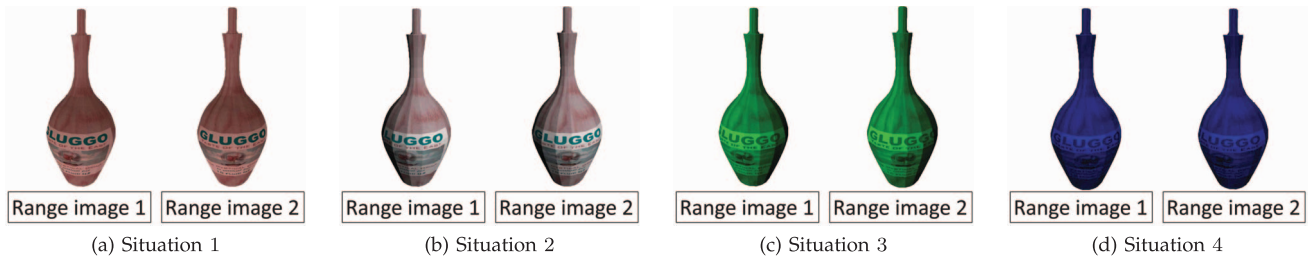
Fig. 2. The four different illumination conditions.

(a) Situation 1 (b) Situation 2 (c) Situation 3 (d) Situation 4

3.4 Stable Points Identification

Points in the overlapping area do not always correspond exactly. This is due to different digitization of the overlapping area depending on the object pose. As a consequence, even for the best transformation aligning the range images, there may be some point correspondences $(\boldsymbol{x}, \Gamma(T(\boldsymbol{x})))$ that do not satisfy the statement $\rho(\boldsymbol{x}) = \rho(\Gamma(T(\boldsymbol{x})))$. This is because the distribution of albedo at an object's surface is not continuous. In such a case, the quality of the estimated photometric properties (illumination and albedo) may be significantly degraded, which would reduce the reliability of our photometric metric. To overcome this problem, we first identify a stable point, i.e., a point whose albedo and normal values 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. We extract stable points independently from two range images as a preprocessing step.

It is well known that in a small vicinity and for diffuse reflection, the difference in chromaticity approximates the difference in albedo well. Accordingly, we define a stable point using both difference of chromaticity and difference of normals in a small vicinity. Namely, a point $\boldsymbol{x}$ is identified to be stable if

$$\forall \boldsymbol{y} \text{ such that } \|\boldsymbol{y} - \boldsymbol{x}\| < \epsilon_s,$$
$$\|c(\boldsymbol{x}) - c(\boldsymbol{y})\| < \epsilon_c \text{ and } \|\boldsymbol{n}(\boldsymbol{x}) - \boldsymbol{n}(\boldsymbol{y})\| < \epsilon_n, \quad (6)$$

where $\boldsymbol{y}$ is a point in the range image concerned, c is chromaticity, $\boldsymbol{n}$ represents the surface normals, and ϵ_s, ϵ_c, and ϵ_n are three thresholds.

4 ANALYSIS OF THE PHOTOMETRIC METRIC

We analyze the behavior of our photometric metric under various parameters. Starting from the ground-truth transformation that perfectly aligns two range images, we generate several transformations by randomly perturbing the parameters of the ground-truth transformation and plot the photometric reprojection error as the function of the registration error (7) for each generated transformation. We notice that when the function is not defined (insufficient number of corresponding points), the photometric metric returns $+\infty$. We then clamp the photometric reprojection error to $[0; 255]$ for better visualization.[2] The random perturbation was obtained by perturbing the rotation angles inside the range $[-0.3; 0.3]$ radians and the translation

___

2. All metric values greater than 255 are set to 255.

values inside the range $[-0.7; 0.7]$ mm. First, a uniform noise was applied with range $[-0.3; 0.3]$ radians for perturbations in the angles and $[-0.7; 0.7]$ mm for the perturbations in the translation. To increase the density of transformations generated close to the ground truth, we then successively applied a uniform noise with ranges $[-0.15; 0.15]$, $[-0.07; 0.07]$, and $[-0.03; 0.03]$ radians for perturbations in the angles and $[-0.3; 0.3]$, $[-0.1; 0.1]$, and $[-0.01; 0.01]$ mm for the perturbations in the translation.

We performed this procedure for different illumination conditions and different values of ϵ_s. We chose to test our photometric metric against different values of ϵ_s because it is the parameter that reflects the size of the neighborhood used in identifying the stable points.

4.1 Our Metric against Different Illumination Conditions

Fig. 3 illustrates our photometric metric with respect to the different illumination conditions illustrated in Fig. 2. We remark that many points are accumulated on the line $error = 255$, and that the more the registration error increases, the more points are accumulated on this line. This is due to clamping our metric to $[0; 255]$. In all situations, we used the same range images, with the same initial positions. Situation 1 (Fig. 2a) is obtained by illuminating the two range images with a light probe using the spherical harmonics representation of image formation. Situations 2, 3, and 4 (Figs. 2b, 2c, and 2d) are obtained by illuminating the range images with an ambient light and a point light source in different positions and with different colors. In Situations 2, 3, and 4, the classic Lambertian reflection model was used to simulate image formation. For all situations, the stable points were identified using the parameters $\epsilon_s = 0.03$, $\epsilon_c = 0.01$, and $\epsilon_n = 0.2$.

First, we observe that while the variation in the photometric reprojection error is large for transformations with large registration error (most of the metric values are then above 255), it becomes small for transformations with small registration errors (most of the metric values are then below 255). This is naturally explained by the fact that the transformations with larger registration errors are much more different from each other than those with smaller registration errors (and so for the point correspondences).

Second, we observe that the behavior of our photometric metric is similar in all situations, even though the illumination conditions are completely different (extended light sources in situation 1, point light source with different positions and different colors in the other three situations). This is because no a priori information on the illumination is used in deriving the photometric cost function.

(a) Situation 1   (b) Situation 2   (c) Situation 3   (d) Situation 4
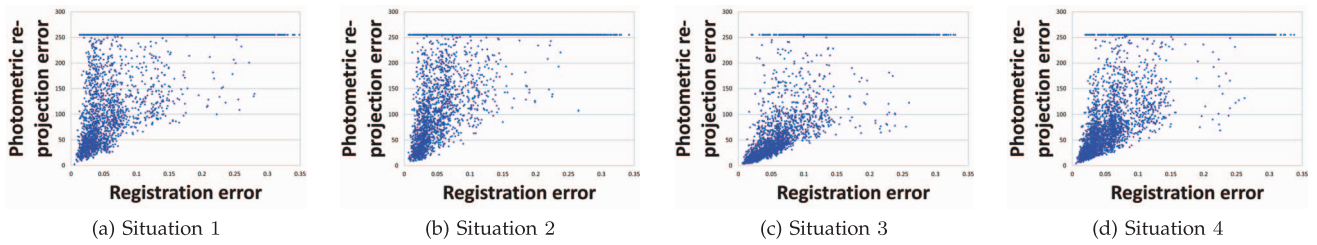
Fig. 3. Photometric reprojection error in function of the registration error for the different situations.

Third, we observe that for the same registration error, our photometric metric can have different values. Two factors explain this phenomena. 1) While the registration error is the same, the transformations are different. For example, from the ground-truth position, sliding one range image to the left or to the right with the same translation magnitude gives the same registration error. However, the point correspondences are different, which gives different photometric reprojection errors (Fig. 4). 2) The texture at the surface of an object is not continuous. As a consequence, while in some directions the normals and albedo values vary smoothly, in some other directions the albedo values may change brutally (Fig. 5). This explains why in some cases, even for a small change in the registration error, the photometric reprojection error changes drastically and why in some cases, even for significant change in the registration error, it changes slowly.

Finally, we observe that, as expected, the transformations with the minimum photometric reprojection errors are also the ones with the minimum registration errors. We notice, however, that the extreme case exists where two different transformations give the minimum photometric reprojection error. Namely, this situation occurs if albedo values and normals exhibit the same symmetries. This extreme case rarely occurs in real situations and we may ignore such a case from the practical point of view.

From these observations, we can conclude that 1) a transformation with a small photometric reprojection error is close to the ground-truth transformation aligning the range images, 2) the minimal solution is almost insensitive to changes in illumination, and 3) our proposed photometric metric will be difficult to minimize using traditional optimization methods.

4.2 Our Metric against Different Values of ϵ_s

Fig. 6 illustrates our photometric metric with respect to different values for ϵ_s. We remark that ϵ_s reflects the minimum size of the neighborhood containing points with similar albedo and normal values required for a point to be identified as stable.

On one hand, we observe that when ϵ_s increases, the values taken by our metric become more concentrated for small registration errors and more sparse for large registration errors. The first effect can be explained by the fact that for a large ϵ_s, the chromaticity and normals around the stable points are homogeneous inside a large area. Therefore, for slightly different transformations, the albedo and normal values of the corresponding points change slightly and then the matrix M defined in Section 3.2 also does so. Thus, the photometric reprojection error varies more slowly with respect to the transformation when ϵ_s increases. The second effect can be explained by the fact that when ϵ_s increases, the number of stable points decreases, and then the number of situations where the number of correspondences is insufficient for estimating $Eval$ increases.

On the other hand, we observe that when ϵ_s increases, our photometric metric becomes less discriminative for evaluating the quality of the registration. This is because the albedo and normal values of the stable points become less discriminative (similar chromaticity and normals for correspondences obtained for slightly different transformations).

In conclusion, a large value for ϵ_s brings more robustness to our metric, while a small ϵ_s brings more discriminative power to our metric (and thus increases its accuracy).

5 REGISTRATION

Given two overlapping range images, we seek a rigid transformation that minimizes the photometric reprojection error defined in (5). When minimizing the photometric reprojection error, we have to decide the strategy we use. As we discussed in Section 4, our proposed photometric metric is hard to minimize using traditional optimization methods. Therefore, we choose the hypothesis-and-test search. Fig. 7 illustrates the flowchart of our proposed registration method.
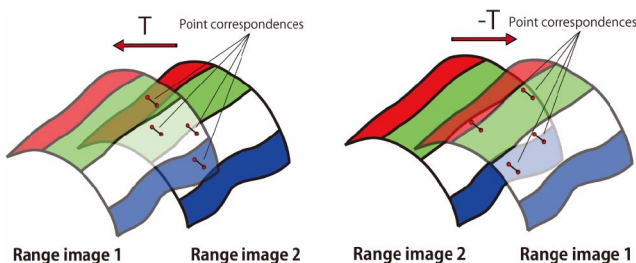

Fig. 4. Two different transformations having the same registration errors give different photometric reprojection errors.
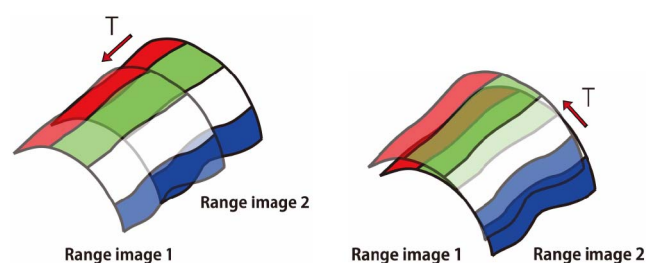

Fig. 5. The photometric reprojection error varies differently for different directions of the transformation.
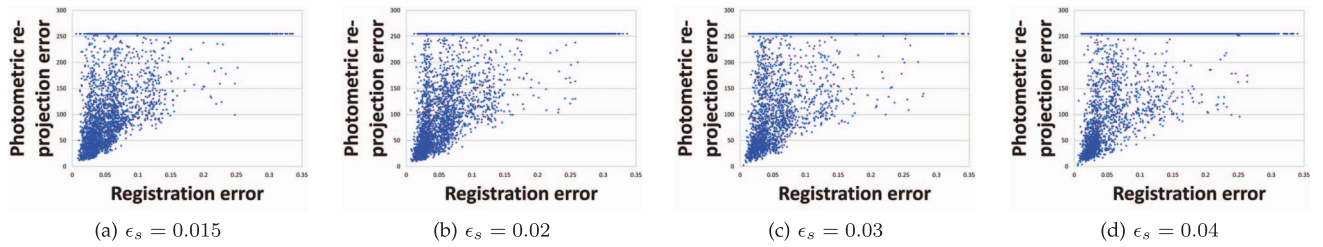
Fig. 6. Photometric reprojection error in the function of the registration error for different values of ϵ_s.

5.1 Transformation Search

The hypothesis-and-test search is performed by testing the correctness 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,[3] the spherical representation is pose invariant [26]. 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 belong to $SO(3)$. We thus employ the spherical representation for range images. This representation reduces the transformation parameter space from $SO(3) \times \boldsymbol{R}^3$ 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 tackle 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 using the method proposed by Horn [10] 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 [10] to obtain the corresponding rigid transformation in $SO(3) \times \boldsymbol{R}^3$.

We remark that though we can use the RANSAC method 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. In the experiments, we set $step = 2^i$ for the ith iteration, with i ranging from 0 to max_{iter}. Note that we uniformly sampled the current searching space at every iteration. In the experiments, at every iteration, we sampled 11 points in each dimension.

5.2 Spherical Representation and Refinement

Our spherical representation method is inspired by the method proposed in [26] 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.

5.2.1 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 hull. The vertices are used to generate Delaunay triangulations. The set of 3D
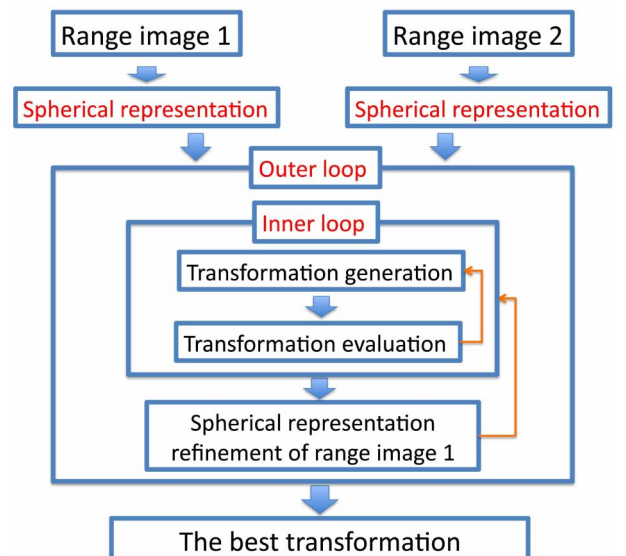
---

3. A closed surface is defined as a surface that is compact and without a boundary; a nonclosed surface is a compact surface with a boundary.

Fig. 7. Flowchart of transformation search.

points in the range image corresponding to the vertices of the convex hull 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 [15], 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 illustrated in Fig. 8.

Algorithm 1. Progressive spherical representation

Require: range image RI

Ensure: spherical representation S of RI that preserves the local structure of RI

$Mesh_0 \leftarrow$ triangulated 2D convex hull of RI, centered around its center of mass

$PtMesh \leftarrow$ 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**

    $\Psi \leftarrow$ a point of $List$

    $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) \leftarrow$ three points of the enclosing triangle of Ψ in $Mesh_i$

    $Mesh_{i+1} \leftarrow$ Delaunay triangulation of $Mesh_i$ where Ψ has been added

    $P \leftarrow$ polygon in $Mesh_{i+1}$ containing the points $(\Psi, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$

    $U \leftarrow$ flatten polygon obtained by conformal mapping of P [15]

    $(\alpha, \beta, \gamma) \leftarrow$ barycentric coordinates of Ψ in U for $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$

    $\Psi' \leftarrow \alpha \boldsymbol{a}' + \beta \boldsymbol{b}' + \gamma \boldsymbol{c}'$

    $S_{i+1} \leftarrow S_i + \{\Psi'\}$, with the same connectivity as in $Mesh_{i+1}$

    $List \leftarrow List - \{\Psi\}$

end for

return S_n

5.2.2 Spherical Representation Refinement

Without loss of generality, we consider the problem of aligning range image 1 to range image 2. After each iteration, we refine the spherical representation of range image 1 with respect to range image 2.

The overlapping areas between the 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 $\boldsymbol{x}$ in the overlapping area O_1 of range image 1, $B(\boldsymbol{x}) = closest(\boldsymbol{x})$ if $\boldsymbol{x} = closest(closest(\boldsymbol{x}))$; $B(\boldsymbol{x})$ 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
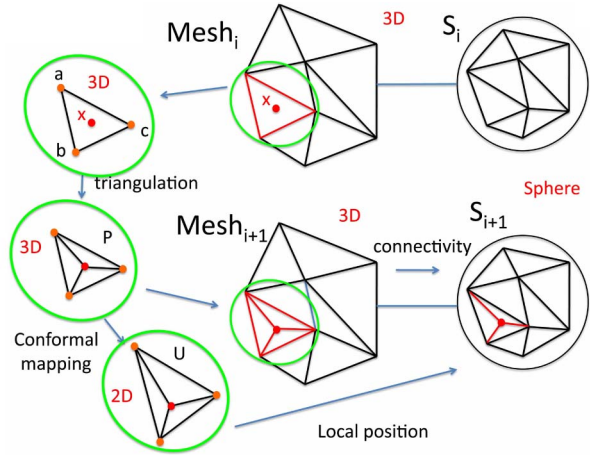

Fig. 8. Illustration of the loop "for" of Algorithm 1.

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.

5.3 Time Complexity Analysis

Our proposed algorithm has its input of two range images with n and m points, respectively, and outputs the transformation T that best aligns the two input range images. Here, we briefly analyze the computational complexity of our proposed algorithm. We refer to Fig. 7 for the different steps of our method and give the computational complexity for each of these steps.

At each inner iteration, our method performs two steps: 1) 3D rigid transformation candidate generation, 2) evaluation of the candidate transformation. During step 1, we identify point correspondences in the spherical domain and then estimate the rigid transformation for the correspondences in the original domain. The closest point identification in the spherical domain is done in $O(n)$ computations (by projective assignment). Estimating the rigid transformation for the point correspondences is done in $O(n)$ computations and evaluating the transformation takes $O(n + m)$ computations. Therefore, for l inner iterations, the inner loop takes $O(l \times (n + m))$ computations.
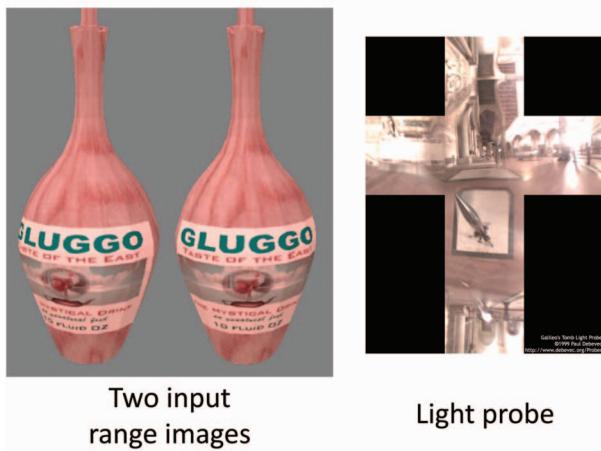
Identifying the current overlapping areas takes $O(n + m)$ computations. Therefore, refining the spherical representation of range image 1 is done in $O(n + m)$ computations. Finally, for l inner iterations and h outer iterations, our registration method takes $O(h \times l \times (n + m))$ computations. We notice that all inner iterations are completely independent. Therefore, using parallel implementation would drastically reduce the computational time.

6 EXPERIMENTS

To demonstrate the usefulness of our proposed method, we evaluate our algorithm in several challenging situations using synthetic and real data. For the comparison, we used seven methods. For methods that search for best matches, we used the method proposed in [23] using albedo with a given directional light source (Method 2), the method proposed in [23] using chromaticity instead of albedo (Method 3), the SIFT algorithm [16] (we used the available

TABLE 1
Description of All Methods Used for Comparison

	Metric	Minimization
Proposed Method	Proposed photometric metric	Proposed exhaustive search
Method 1	NCC: $\frac{1}{3} \sum_{c \in [\![1,3]\!]} \frac{\sum_{x \in S} chrom_c(x) chrom_c(closest(x))}{\sqrt{\sum_{x \in S} chrom_c(x)^2 \sum_{x \in S} chrom_c(closest(x))^2}}$	Proposed exhaustive search
Method 2	Differences between matched adaptive regions using estimated albedo	ICP
Method 3	Differences between matched adaptive regions using chromaticity	ICP
SIFT-RANSAC	SIFT descriptor	RANSAC
RGBD	SIFT descriptor and L_2 distances between closest points	RANSAC and ICP
Stochastic	$\sum_{x \in S} -\frac{1}{1 + \alpha ||x - closest(x)||^2 + \mu ||RGB(x) - RGB(closest(x))||^2}$	Stochastic minimization
Second Order	$\sum_{x \in S} \sum_{y \in M} \left(\frac{\exp \frac{-1}{2\sigma^2}(||x-y||_2^2 + \nu ||RGB(x) - RGB(y)||_2^2)}{\sum_{z \in M} \exp \frac{-1}{2\sigma^2}(||x-z||_2^2 + \mu ||RGB(x) - RGB(z)||_2^2)} \right) ||x - y||_2^2$	Second order optimization

(a) The first set-up.

(b) The second set-up.

Fig. 9. Initial state for the two different setups.

Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	**0.005**	0.008	0.415	0.415
Error (in res)	**0.5**	0.8	41.5	41.5
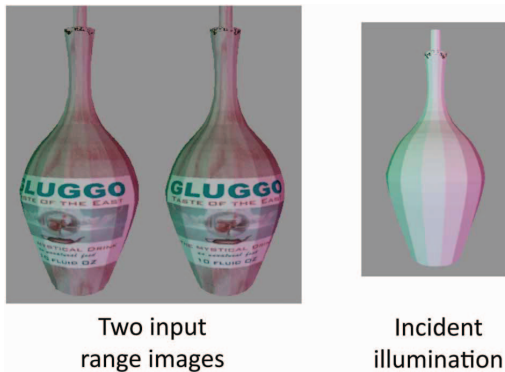
Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	0.0764	0.0840	0.252	0.235
Error (in res)	7.64	8.40	25.2	23.5

Fig. 10. Results obtained with the eight methods for the first setup.

TABLE 2
Description of the Data *Vase*

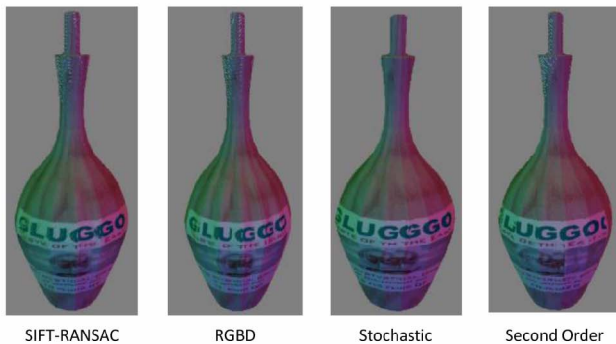| Nb_Points	Resolution	Expected_rot (angle; axis)	
30650	0.01 mm	$(20.0; -0.01, 0.93, 0.34)$	
Expected_translation		$(0.19, 0.01, -0.03)$	
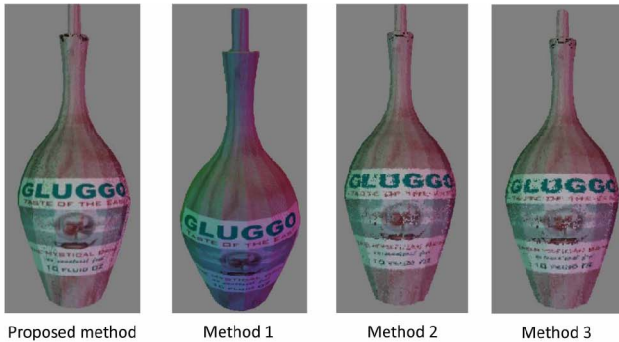
Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	**0.005**	0.08	0.415	0.415
Error (in res)	**0.5**	0.8	41.5	41.5

Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	0.0478	0.0502	0.227	0.229
Error (in res)	4.78	5.02	22.7	22.9

Fig. 11. Results obtained with the eight methods for the second setup.

	(Situation 1)	(Situation 2)	(Situation 3)	(Situation 4)
Error	0.007	0.007	0.011	0.011

(a) Different illumination

ε	(a) $\varepsilon = 0.015$	(b) $\varepsilon = 0.02$	(c) $\varepsilon = 0.03$	(d) $\varepsilon = 0.04$
Error	0.028	0.005	0.007	0.007

(b) Different ϵ_c

Fig. 12. Registration results obtained with our method for the various situations presented in Section 4. The transformations evaluated during the registration process are plotted in red and superimposed onto the graphs shown in Section 4.

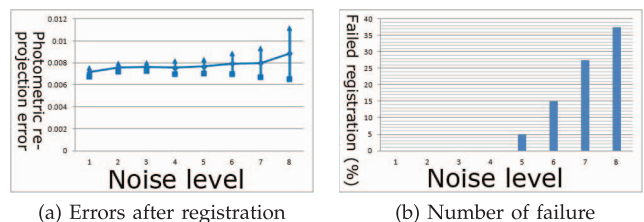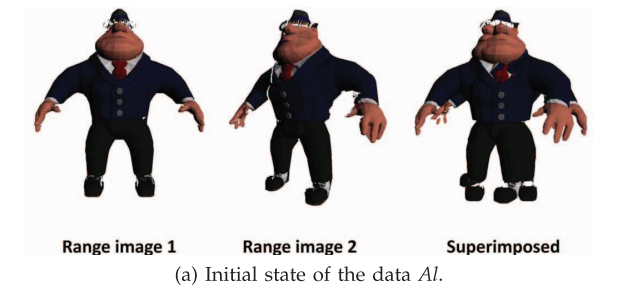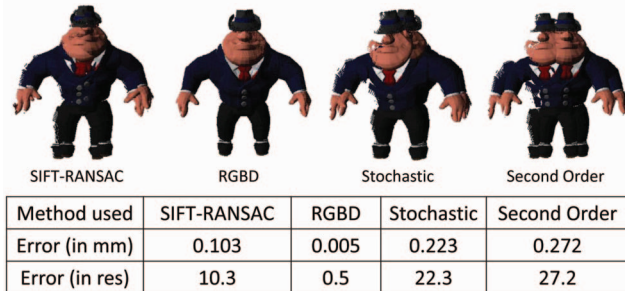(a) Errors after registration   (b) Number of failure

Fig. 13. Experiments with different levels of noise added to the color.

code provided by Vedaldi [25]) with a RANSAC outlier removal postprocess (SIFT-RANSAC), and the registration method proposed in [9] that combines SIFT-RANSAC and ICP algorithms (RGBD). For methods that minimize cost functions, we extended the stochastic minimization method [18] (Stochastic) and the second order minimization method using quadratic approximation [4] (Second Order) (both initially account for geometry only) so that they can also consider color information.[4] Namely, for both methods, we combined the L_2 color distance term with the euclidean distance term to derive the cost functions. We also used our proposed method using chromaticity instead of our photometric evaluation function (Method 1). The employed metrics and searching strategies of all the methods used for comparison are summarized in Table 1. In this table, S denotes the source image (range image 1, for example), M denotes the model image (range image 2, for example), $chrom()$ denotes the chromaticity vector of a point ($chrom_1()$, $chrom_2()$, and $chrom_3()$ denote the red, green,

and blue channels of the chromaticity vector, respectively), $RGB()$ denotes the color vector of a point, and $closest()$ denotes the closest point in the model image of a point in the source image.

We consider the problem of aligning range image 1 to range image 2 and we assume we are given the ground truth (obtained manually for real data). We evaluate the registration result using the distance between the estimated position of points of range image 1 after registration and their ground-truth position. Namely, given T_g and T_e, the ground-truth transformation and the estimated transformation, respectively, the registration error $err(T_e)$ is computed as follows:

$$err(T_e) = \frac{\sum_{i=1}^{n} \|T_g(\boldsymbol{x}_i) - T_e(\boldsymbol{x}_i)\|_2}{n}, \qquad (7)$$

4. This is because there is no existing photometric cost function minimization technique to compare with.

(a) Initial state of the data *Al*.

Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	**0.007**	0.020	0.191	0.222
Error (in res)	**0.7**	2.0	19.1	22.2

Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	0.103	0.005	0.223	0.272
Error (in res)	10.3	0.5	22.3	27.2

(b) Results obtained with the eight methods with the data *Al*.

Fig. 14. Experiments with the data *Al*.

(a) Initial state of the data *Al 2*.

Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	**0.007**	0.786	0.244	0.246
Error (in res)	**0.7**	78.6	24.4	24.6

Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	0.023	0.005	0.082	0.276
Error (in res)	2.3	0.5	8.2	27.6

(b) Results obtained with the eight methods with the data *Al 2*.

Fig. 15. Experiments with the data *Al 2*.

where n is the number of points in range image 1 and $\{\boldsymbol{x}_i\}_{i \in [1;n]}$ are the points of range image 1.

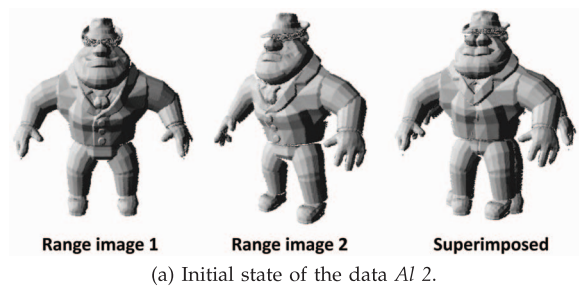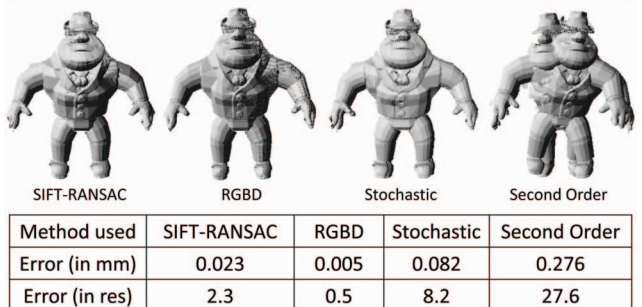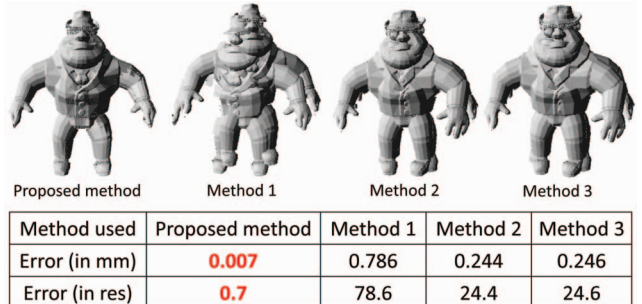ϵ_s, ϵ_c, and ϵ_n were set to $3 * res$, 0.02, and 0.1, respectively for all experiments with synthetic data and to $2 * res$, 0.05, and 0.2 for all experiments with real data. res here means the resolution of the range images (i.e., the average distance between neighboring points).

6.1 Synthetic Data

6.1.1 The Data Vase

The synthetic data, called *vase*, were obtained using a 3D modeler software (3D Studio Max) (see Table 2). The exact albedo is known and we simulated lighting under different illuminations. This dataset is challenging for registration in that the shape is rotationally symmetric, the texture of the object presents several repetitive patterns, and no exact correspondences exist between the two range images.

The first setup is illustrated in Fig. 9a. The range images were illuminated by the light probe *galileo* from the Debevec database [6], and the color was synthesized using spherical harmonics with a Lambertian reflection kernel. Fig. 10 shows the results obtained with the eight methods mentioned above. In this situation, the illumination is dominated by ambient light, and the changes in RGB 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 the

chromaticity cross correlation (Method 1). Moreover, because of repetitive texture patterns (multiple "G" letters for example) and distant initial positions, Method 2, Method 3 as well as SIFT-RANSAC (i.e., methods that match points based on their textural features) did not give satisfactory results. Note that RGBD did not improve the registration result because of the lack of salient geometric features. We observe that Stochastic and Second Order did not perform well on this example, though they do consider color information in addition to geometry. This can be explained by the fact that the color distribution function over the object's surface is noncontinuous and nonsmooth; therefore, the cost function is not differentiable and is highly oscillatory. As a consequence, the minimization strategy using the gradient descend (as done in Second Order) becomes easily trapped into local minima and the stochastic search strategy (as done in Stochastic) fails due to the highly oscillatory behavior of the cost function. Our proposed search strategy, in contrast, efficiently found the global minimum. The estimated transformation T_e by our proposed method was a rotation of $(-19.9; 0.00, 0.94, 0.35)$ and a translation of $(0.19, -0.01, 0.03)$, and $Eval(T_e)$ was[5] 1.2.

---

5. The color was coded in RGB with values between 0 and 255. Since the color of a point is approximated with 98 percent accuracy in (2), the reprojection error of around 1.2 means that the optimal solution is found provided that the distribution of RGB values is uniform over the range.
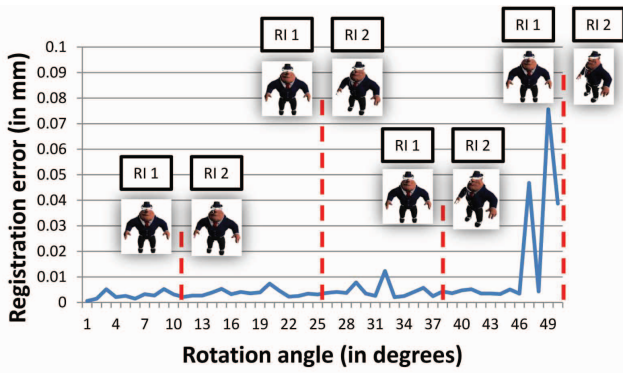
(a) Registration results obtained with our method for the data *Al* for the scenario 1 under various different initial poses ("RI 1" and "RI 2" stand for "range image 1" and "range image 2", respectively).
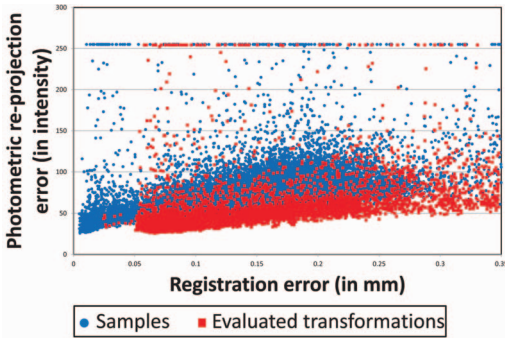
(b) Plot of the registration process for a rotation with an angle of 50.0 degrees.

Fig. 16. Experiments against various initial relative poses.

The second setup is illustrated in Fig. 9b. 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. Fig. 11 shows the results obtained with the eight 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). Results obtained with all the methods were similar to those obtained in the first setup. Our method could achieve accurate alignment (similar to the one obtained in the previous setup). The estimated transformation T_e obtained with our proposed method was a rotation of $(-20.1; 0.00, 0.94, 0.35)$ and a translation of $(0.19, -0.01, 0.03)$, and $Eval(T_e)$ was 3.4.
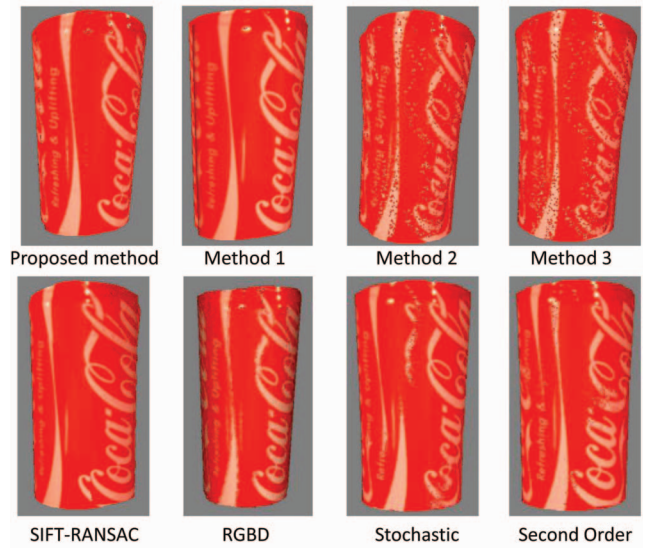
The results obtained with our method for the various situations presented in Section 4 are shown in Fig. 12 along

Fig. 17. The experimental setup.
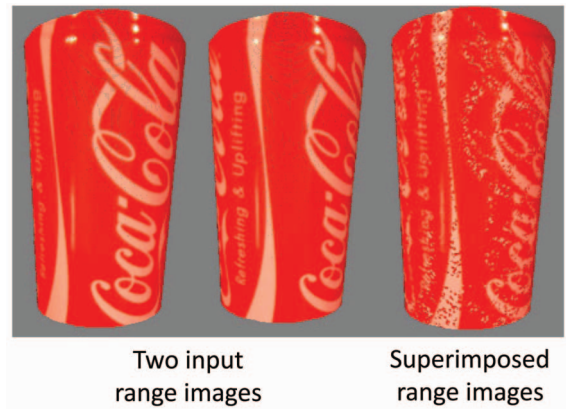
(a) Initial state.

(b) Results obtained with the eight methods.

Fig. 18. Experiments with the data *Can*.

with the graphs of the photometric reprojection error evaluated during the registration process. We can see that during the registration, the lower bound of our photometric metric is always evaluated, which allowed us to converge and validated our search strategy. However, we remark that when ϵ_s is too small, the registration result is degraded. This is because the quality of the point correspondences becomes poor.

Fig. 13 shows the results of intensive experiments when adding noise into the color. Random noise was added to the rendered colors with different levels independently in each RGB channel. For each level of noise, our method was run 40 times. Fig. 13a shows the errors after registration and Fig. 13b shows the number of runs for which our method failed in registering the two range images (i.e., the error was greater than five times the resolution of the range image).

TABLE 3
Description of Data *Can*, *Hand*, *Pottery*, *Cylinder* and *Helmet* (Size Denotes the Number of Points)

	Can	*Hand*	*Pottery*	*Cylinder*	*Helmet*
Size	28000	50000	210000	80000	16000
Res	0.55 mm	0.55 mm	0.55 mm	0.55 mm	0.55 mm

TABLE 4
Quantitative Evaluation of Registration Results for the Six Datasets

	Rotation (angle (in degrees); axis)	Translation (in mm)	Error (in mm)	Error (in resolution)
Can				
Ground Truth	$(20.0; 0.01, 0.93, 0.34)$	$(9.00, 0.10, -1.00)$	0.0	0.0
Our Method	$(\mathbf{19.7; 0.01, 0.94, 0.34})$	$(\mathbf{8.93, 0.37, -1.19})$	**0.52**	**0.94**
Method 1	$(20.2; 0.02, 0.94, 0.33)$	$(8.61, 0.52, -1.24)$	0.65	1.18
Method 2	$(1.94; -0.08, 0.58, 0.81)$	$(0.74, 4.26, 1.58)$	11.2	20.4
Method 3	$(8.23; 0.06, 0.92, 0.39)$	$(3.32, 1.77, 0.75)$	9.57	17.4
SIFT-RANSAC	$(20.1; 0.02, 0.95, 0.31)$	$(9.2, 0.37, -1.35)$	0.74	1.35
RGBD	$(19.0; 0.03, 0.94, 0.34)$	$(8.75, 0.35, -1.25)$	0.53	0.97
Stochastic	$(2.73; 0.13, 0.53, 0.83)$	$(-0.09, -0.10, -0.13)$	10.8	19.6
Second Order	$(1.59; -0.2, 0.02, 0.98)$	$(0.32, 1.35, 0.35)$	10.4	18.9
Hand				
Ground Truth	$(20.0; 0.01, 0.93, 0.34)$	$(6.60, -2.3, -0.40)$	0.0	0.0
Our Method	$(\mathbf{19.4; 0.01, 0.94, 0.34})$	$(\mathbf{6.71, -2.71, -0.44})$	**0.55**	**1.00**
Method 1	$(23.5; 0.11, 0.98, -0.13)$	$(12.6, -3.76, -1.46)$	12.2	22.2
Method 2	$(10.3; -0.01, 0.68, 0.73)$	$(-2.81, -1.83, 0.06)$	10.7	19.5
Method 3	$(17.4; 0.04, 0.98, 0.21)$	$(3.90, -2.59, -0.46)$	3.49	6.35
SIFT-RANSAC	$(19.8; -0.03, 0.94, 0.35)$	$(8.58, -0.74, 0.50)$	2.83	5.15
RGBD	$(19.7; 0.01, 0.95, 0.33)$	$(7.23, -1.07, -0.21)$	1.42	2.58
Stochastic	$(25.6; 0.27, -0.92, -0.28)$	$(-1.58, -5.57, 11.6)$	20.6	37.4
Second Order	$(17.52; 0.03, 0.94, 0.33)$	$(6.32, -1.81, -0.67)$	1.09	1.98
Pottery				
Ground Truth	$(20.2; 0.05, 0.91, 0.42)$	$(309, 4.57, -48.1)$	0.0	0.0
Our Method	$(\mathbf{20.1; 0.09, 0.91, 0.41})$	$(\mathbf{307, -10.5, -47.7})$	**0.82**	**1.50**
Method 1	$(15.3; 0.08, 0.92, 0.40)$	$(236, -7.74, -27.6)$	6.45	11.7
Method 2	$(20.0; 0.03, 0.99, -0.12)$	$(307, -14.1, -49.4)$	31.9	58.0
Method 3	$(19.6; 0.07, 0.99, -0.06)$	$(302, -26.5, -47.5)$	32.1	58.4
SIFT-RANSAC	$(0.89; -0.39, 0.89, -0.22)$	$(11.2, 4.69, 1.21)$	25.0	45.5
RGBD	$(11.0; 0.06, 0.91, 0.40)$	$(170, -5.29, -13.19)$	11.6	21.1
Stochastic	$(11.0; 0.06, 0.91, 0.40)$	$(170, -5.29, -13.19)$	12.3	22.4
Second Order	$(7.95; 0.09, 0.89, 0.44)$	$(120, -9.7, -5.79)$	16.1	29.3
Cylinder 1				
Ground Truth	$(32.0; 0.03, 0.96, 0.26)$	$(10.0, -0.50, -4.00)$	0.0	0.0
Our Method	$(\mathbf{31.8; 0.03, 0.97, 0.26})$	$(\mathbf{10.23, -0.27, -3.64})$	**0.52**	**0.93**
Method 1	$(31.6; 0.03, 0.97, 0.25)$	$(10.3, 0.60, -3.41)$	1.33	2.42
Method 2	$(2.22; 0.00, 0.98, 0.21)$	$(-2.40, 0.03, -0.42)$	16.1	30.9
Method 3	$(12.4; 0.03, 0.97, 0.23)$	$(-1.59, -0.70, 0.19)$	18.5	37.0
SIFT-RANSAC	$(31.6; 0.04, 0.96, 0.26)$	$(10.2, -0.07, -3.34)$	0.90	1.63
RGBD	$(32.0; 0.03, 0.97, 0.26)$	$(10.3, -0.03, -3.58)$	0.70	1.28
Stochastic	$(13.3; 0.04, 0.94, 0.33)$	$(1.43, -0.37, -0.70)$	11.1	20.2
Second Order	$(11.7; 0.03, 0.93, 0.37)$	$(0.55, -0.13, -0.89)$	12.0	21.8
Cylinder 2				
Ground Truth	$(32.0; 0.03, 0.96, 0.26)$	$(10.0, -0.50, -4.00)$	0.0	0.0
Our Method	$(\mathbf{33.6; 0.03, 0.96, 0.26})$	$(\mathbf{10.6, -0.09, -3.15})$	**1.20**	**2.18**
Method 1	$(3.24; -0.05, -0.60, 0.75)$	$(-4.7, -1.03, -0.48)$	18.8	34.2
Method 2	$(0.53; -0.17, -0.38, -0.91)$	$(-4.49, -0.30, -0.70)$	18.1	28.2
Method 3	$(3.67; 0.01, 0.84, 0.54)$	$(-2.5, -0.18, -0.45)$	15.9	29.5
SIFT-RANSAC	$(30.75; 0.01, 0.97, 0.22)$	$(14.3, 1.89, -6.93)$	5.72	10.4
RGBD	$(29.0; 0.03, 0.97, 0.26)$	$(12.7, 1.18, -5.65)$	3.65	6.64
Stochastic	$(1.97; -0.73, 0.37, -0.58)$	$(-5.67, 0.00, 0.28)$	18.3	33.3
Second Order	$(1.73; 0.24, 0.90, 0.35)$	$(-4.54, -0.35, -0.91)$	16.7	30.4
Helmet				
Ground Truth	$(20.0; 0.05, 0.96, 0.29)$	$(461, 1.52, -80.8)$	0.0	0.0
Our Method	$(\mathbf{20.1; 0.06, 0.95, 0.3})$	$(\mathbf{460, -2.64, -80.3})$	**0.35**	**0.64**
Method 1	$(21.0; 0.11, 0.96, 0.28)$	$(485, -29.6, -88.7)$	2.93	5.33
Method 2	$(17.4; 0.11, 0.97, 0.21)$	$(410, -33.1, -63.2)$	2.11	3.84
Method 3	$(15.26; 0.04, 0.99, 0.16)$	$(366, -6.35, -50.3)$	2.65	4.82
SIFT-RANSAC	$(17.7; 0.12, 0.92, 0.38)$	$(395, -27.3, -57.8)$	1.42	2.58
RGBD	$(19.7; 0.06, 0.96, 0.3)$	$(457, -5.55, -78.8)$	0.65	1.18
Stochastic	$(19.6; -0.03, 0.97, 0.23)$	$(458, 32.6, -80.6)$	1.01	1.84
Second Order	$(18.1; 0.05, 0.95, 0.29)$	$(417, -1.78, -65.9)$	1.03	1.87

From these results, we can see that our method can achieve accurate results even in the presence of noise in the captured colors. However, when the noise becomes larger, our method happens to fail from time to time. This can be explained by the discretization of the searching space. To reduce the number of failures, it is required to refine the discretization of the searching space (which will increase the computational time).
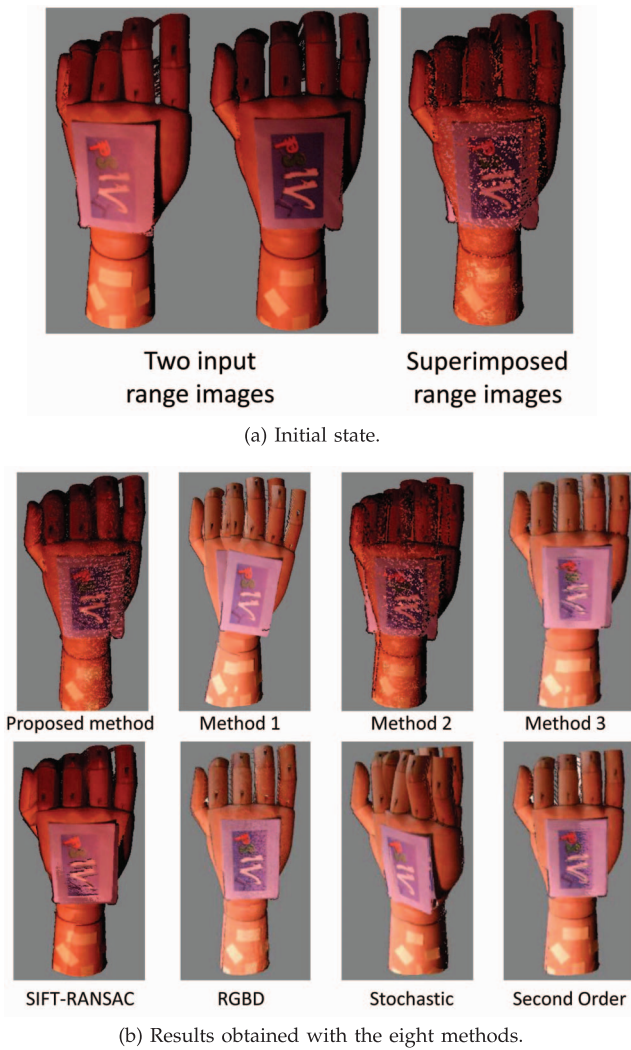
(a) Initial state.

(b) Results obtained with the eight methods.

Fig. 19. Experiments with the data *Hand*.

(a) Data *Pottery*.

(b) Lighting.

(c) Data *Cylinder 1*.

(d) Lighting.

(e) Data *Cylinder 2*.

(f) Lighting.

(g) Data *Helmet*.

(h) Lighting.

Fig. 20. Initial state for data *Base*, *Cylinder 1*, *Cylinder 2*, and *Helmet*.

6.1.2 The Data Al

Results obtained with another synthetic data called *Al* are shown in Fig. 14. This synthetic data was of resolution about 0.01 mm. Two range images were rotated by 40.0 degrees around the vertical axis and illuminated by a fixed directional light source. The observed colors were synthesized using the classic Lambertian reflection model. This situation is challenging in that the change in pose between the two range images induces significant change in intensity. Moreover, there are a large number of occluded points and few points with salient photometric features, which makes the matching problem more difficult to solve.

Our method successfully registered the two range images. In particular, the hypothesis-and-test strategy (our proposed method and Method 1) gave better results when compared with the feature-based methods (Method 2, Method 3, and SIFT-RANSAC). This is because of the advantage of global optimization functions against local functions. We remark that in this case, RGBD worked well thanks to sufficient geometric features and sufficiently good initialization obtained with SIFT-RANSAC, while Stochastic

and Second Order did not work well, probably because the initialization was too far from the ground truth.

To illustrate the advantage of using our proposed photometric metric over state-of-the-art methods, we created the synthetic data *Al 2* shown in Fig. 15a that is the same geometric object as the data *Al*, with the same transformation, initialization, and illumination but without the texture. To be more precise, we set all color channels of all points to 255 and then illuminated the object with a directional gray light using the classical Lambertian model. As we can see from the results shown in Fig. 15b, our method could estimate the transformation that aligns the two input range images. This is because our method takes support from changes in appearance due to illumination, in contrast with the other methods that strive to cancel (or ignore) this effect.

To verify the robustness of our method against the initial relative pose between the input range images, we captured several range images of the synthetic data *Al*, from the same viewpoint but under different poses. We restricted our experiment to the situation where the rotation axis is fixed, while the rotation angle varies. Each pose was obtained by rotating the 3D model around the vertical axis, with

different rotation angles (ranging from 1.0 degree up to 50.0 degrees) and in the clockwise direction.

Fig. 16a shows the results obtained with our method. We can see that for a rotation angle up to 46.0 degrees, we could always obtain accurate registration results (for a rotation angle of 46.0 degrees, the percentage of overlapping area between the two range images was about 80.0 percent). As described in Section 5.1, the rotations of maximum 20 degrees are searched in the spherical domain. In this case, the deformations induced by the projection operator resulted in 20 degrees in the spherical domain covering 46 degrees in the original domain. Therefore, even though in the original domain the rotation angle is about 46 degrees, our method was still working. However, for a rotation angle greater than 46.0 degrees, we observe failures in the registration. This can be explained by the fact that the initial spherical representations of the two input range images (computed independently) became too different. As a consequence, as we can see in Fig. 16b, it prevents us from generating a searching space fine enough for convergence. In fact, the transformation that correctly aligns the two range images (which also minimizes our photometric evaluation function) is not generated.

6.2 Real Data

We employed a Konica Minolta Vivid 910 laser 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 canceled. In our experiments, however, we did not know this factor, and thus the gamma correction was not canceled.

Fig. 17 illustrates the experimental setup used to capture all real data. A target object was put on a black turning table; the laser scanner Vivid 910 was fixed and focused on the target object. We then rotated the table to capture two range images of the object in different poses. All data shown in the following experiments are direct output of the laser scanner; no segmentation algorithm was applied.

6.2.1 The Data Can and Hand

We obtained two range images of a rotationally symmetric can that is approximately 10.0 cm high and has a diameter of about 5.0 cm (Fig. 18a). Details on the data called *Can* are given in Table 3, and the results are shown in Fig. 18b and Table 4. These data are challenging in that the quality of the image is low and there is an unknown gamma correction factor. Moreover, these data exhibit 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 laser scanner, and our method worked extremely well compared to the other methods. For estimated transformation T_e obtained with our proposed method, $Eval(T_e)$ was 2.56.

Another data item called *Hand* is presented in Fig. 19a and Table 3. Registration results are shown in Fig. 19b and Table 4. 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 correctness of transformations becomes unreliable. Our

(a) Data *Pottery*.

(b) Data *Cylinder 1*.

(c) Data *Cylinder 2*.

(d) Data *Helmet*.

Fig. 21. Registration results for data *Pottery*, *Cylinder 1*, *Cylinder 2*, and *Helmet*.

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 a photometric metric. Methods 2 and 3 still failed pitifully.
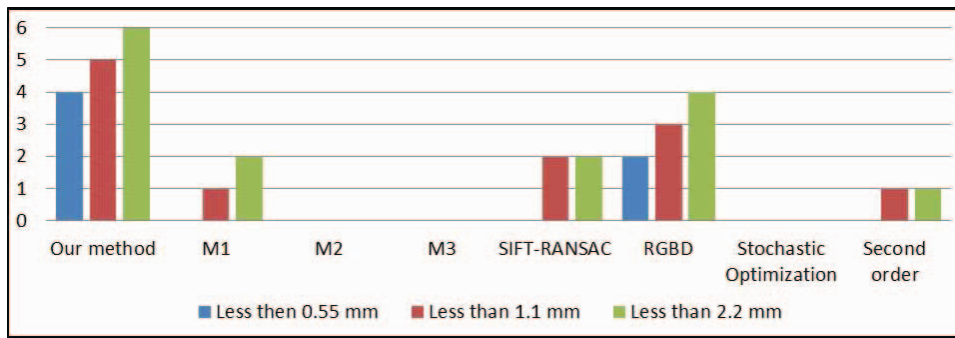
Fig. 22. Number of successful registrations obtained for the six datasets with all the methods.

Results obtained using SIFT-RANSAC, RGBD, Stochastic, and Second Order are also illustrated in both Figs. 18b and 19b, and numerical results are given in Table 4. RGBD gave interesting results for these two data: For the data *Can*, the SIFT-RANSAC obtained good alignment, and RGBD thus obtained good results as well. For the data *Hand*, on the other hand, RGBD obtained less accurate results in spite of more geometric features. This is because SIFT-RANSAC did not provide accurate alignment for RGBD. Stochastic and Second Order did not work for the data *Can* because of poor initialization, discontinuity in color distribution, and symmetry of the geometrical shape. We remark that Second Order gave fair results for the data *Hand* because of the fact that its geometry was sufficient to give good direction for gradient descent minimization. However, Stochastic did not work for the data *Hand*.

6.2.2 Other Data

Fig. 20 illustrates the different objects called *Pottery*, *Cylinder 1*, *Cylinder 2*, and *Helmet*. Table 3 lists the description of all the data. The data *Cylinder 1* and *Cylinder 2* are range images of the same object, with the same changes in pose but under different illumination conditions. Note that the data *Pottery* had a height of about 10 cm and a width of about 20 cm, the data *Cylinder* had a height of about 20 cm and a width of about 8 cm, and the data *Helmet* had a height of about 8 cm and a width of about 5 cm. The results obtained with these objects are shown in Table 4 and illustrated in Fig. 21.

The data *Pottery* is challenging in that its shape is rotationally symmetric while its texture does not exhibit clear key-points with distinctive features that could be used for matching. Because in Methods 2 and 3 we do not use key point detector, the number of outliers in matching becomes larger than that of inliers. This results in failed registration. SIFT-RANSAC uses key point detector and key point descriptor identified in the intensity images for matching. In the intensity images, however, the texture patterns are not distinctive enough, which leads to some mismatches and results in failed registration. Because the shape of the object is rotationally symmetric, using geometry did not help the registration process much, as we can see in the results obtained with RGBD, Stochastic or Second Order. On the other hand, our proposed method uses a global error metric and a hypothesis-and-test search which allowed us to successfully register the two range images. We notice that using chromaticity in this case did

not work. This is because the difference of chromaticity between points of the two range images is not discriminative enough to find the best transformation.

The data *Cylinder 1* and *Cylinder 2* are challenging in that their reflective properties cause several missing points and large noise in the depth values. This is because the accuracy of the laser scanner decreases when the texture at the surface becomes black (the laser beam is then not properly reflected), which is the case for all letters. This effect has a dramatic impact on Methods 2 and 3 because the missing points may prevent the region from growing in one range image while it will continue growing in the other range image. In addition, the noise in the depth values amplifies the distortion between the descriptors of the same point in the two range images. This results in failed registration. SIFT-RANSAC worked fairly well for the data *Cylinder 1* thanks to the small changes in color between the two range images and many textural features, but did not work for the same object under different illumination conditions (data *Cylinder 2*). On the other hand, our proposed method could obtain accurate registration results for both situations.

The data *Helmet* were captured to see the performance of our proposed method to an object with salient geometric features against the state-of-the-art methods. From the results illustrated in Fig. 21d and shown in Table 4, we observe that our proposed method obtained results of the same accuracy as the state-of-the-art methods, though our method is not specifically designed for objects with salient geometric features.

Table 4 summarizes quantitative results obtained for all the objects and with all the methods. It shows the ground-truth transformations, the estimated transformations after registration, as well as the registration errors in mm and in term of resolution. Fig. 22 shows the success rate obtained for our real datasets with all the methods. For each method, we plot the number of successful registrations we obtained over the six datasets. To determine whether a registration is successful, we employed three different criteria (illustrated in blue, red, and green) based on the registration error. Namely, we define 1) a registration is successful if the error is below the data resolution (i.e., 0.55 mm) (in blue), 2) a registration is successful if the error is below twice the data resolution (i.e., 1.1 mm) (in red), and 3) a registration is successful if the error is below four times the data resolution (i.e., 2.2 mm) (in green). We can see that our proposed method obtained the best robustness in terms of a variety of objects. In particular, the advantage of our
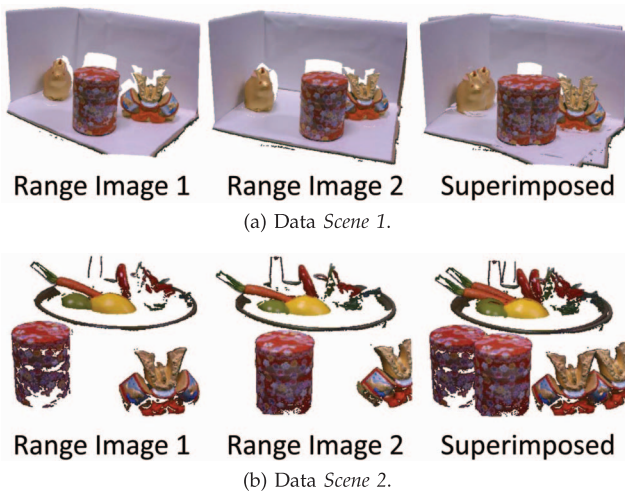
(a) Data *Scene 1*.

(b) Data *Scene 2*.

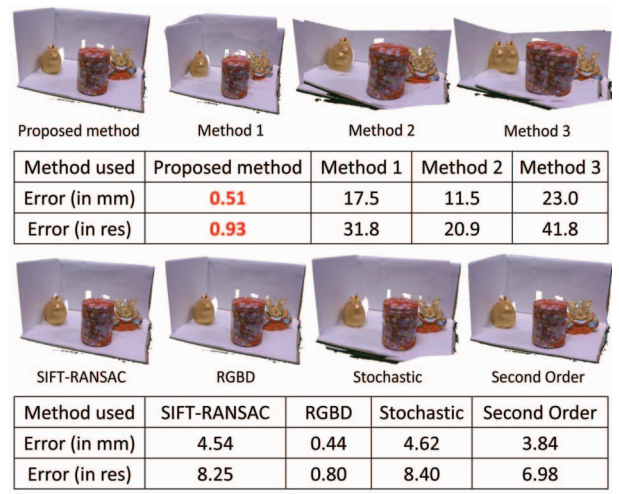Fig. 23. Data *Scenes 1* and *2*.

method is clear under the criterion that the registration error is below the resolution of the range image (in blue).
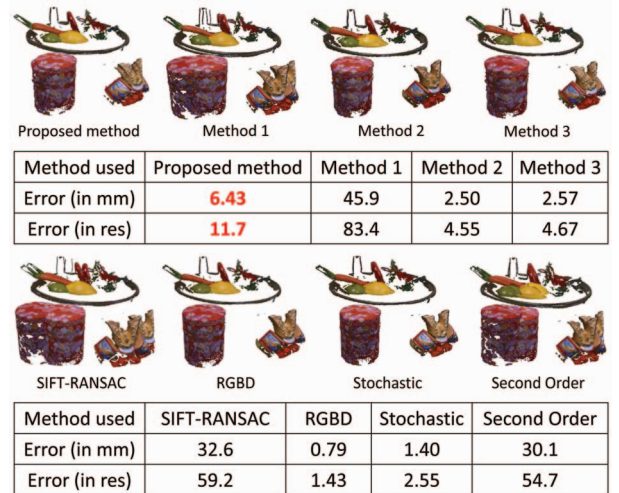
7   SUMMARY AND DISCUSSION

We introduced a new photometric metric for registering range images of Lambertian surfaces under general and distant, unknown lighting. Our metric evaluates photometric reprojection error by taking into account the relationship between reflectance properties, 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. We notice that different approaches to using our photometric metric for registration exist. For example, we can use a coarse registration method as an initialization of a brute-force search in the vicinity of the initial estimate of the transformation. We chose to use the spherical representation of range images to reduce the searching space because it allows us not to depend on the quality of an initial coarse registration, making the overall registration method more stable against the initial conditions.

Though in this paper, we handled only range images of a single segmented out object, with the recent development of depth sensors a new scenario became available. In particular, with an RGB-D video camera, the input data generally consist of a cluttered scene with or without background. Therefore, to see the potential applicability of our proposed method to RGB-D videos, we additionally tested two other data. We captured the data *Scenes 1* and 2 that simulate cluttered scenes with and without background. Fig. 23 shows the input data as well as the initial poses.

From the results shown in Fig. 24, we can see that our proposed method can handle background and multiple objects to some extent. In particular, we observe that with the data *Scene 1*, we obtained accurate results similar to those obtained with RGBD that is specifically designed for

Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	0.51	17.5	11.5	23.0
Error (in res)	0.93	31.8	20.9	41.8

Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	4.54	0.44	4.62	3.84
Error (in res)	8.25	0.80	8.40	6.98

(a) Data *Scene 1*.

Method used	Proposed method	Method 1	Method 2	Method 3
Error (in mm)	6.43	45.9	2.50	2.57
Error (in res)	11.7	83.4	4.55	4.67

Method used	SIFT-RANSAC	RGBD	Stochastic	Second Order
Error (in mm)	32.6	0.79	1.40	30.1
Error (in res)	59.2	1.43	2.55	54.7

(b) Data *Scene 2*.

Fig. 24. Registration results for data *Scenes 1* and *2*.

RGB-D videos. Note that the photometric reprojection error of the transformation estimated with our method was 12.5. The failure of our proposed method with the data *Scene 2* may be due to the large portion of depth measurement errors (as shown in Fig. 25). More precisely, as we can see inside the red box in Fig. 25, depth measurement errors typically arise around the areas where the geometry of the scene presents strong discontinuities. In such cases, the laser scanner often tends to interpolate depth values around discontinuous areas. This induces a large portion of false
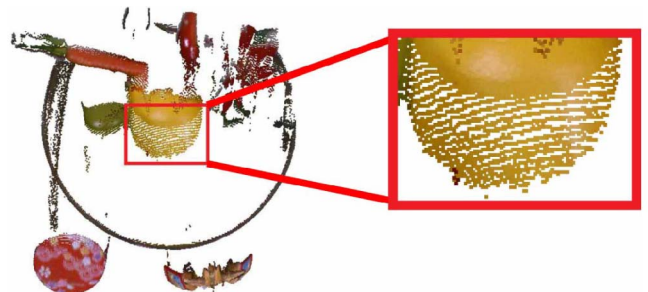

Fig. 25. Top view of range image 1 of *Scene 2*.

normal estimation and, thus, incoherent photometric solutions, even for the best transformation. Indeed, the photometric reprojection error of the transformation estimated with our method for *Scene 2* was 50.4. In cluttered scenes, many geometric discontinuities likely exist due to multiple objects. As a consequence, before directly applying our proposed method to RGB-D videos, we first need to investigate how robust our proposed photometric cost function is against localized large depth measurement errors caused by geometric discontinuities.

ACKNOWLEDGMENTS

REFERENCES

[1] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspace," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 2, pp. 218-233, Feb. 2003.

[2] H. Bay, T. Tuytelaars, and L. van Gool, "SURF: Speeded Up Robust Features," *Proc. European Conf. Computer Vision,* pp. 404-417, 2006.

[3] P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, no. 2, pp. 239-256, Feb. 1992.

[4] D. Breitenreicher and C. Shnörr, "Intrinsic Second-Order Geometric Optimization for Robust Point Set Registration without Correspondence," *Proc. Int'l Conf. Energy Minimization Methods in Computer Vision and Pattern Recognition,* pp. 274-287, 2009.

[5] S. Choi, T. Kim, and Z. Yu, "Performance Evaluation of RANSAC Family," *Proc. British Machine Vision Conf.,* 2009.

[6] P. Debevec, "Light Probe Image Gallery," http://ict.debevec.org/~debevec/Probes/, 2004.

[7] O. Enqvist, F. Jiang, and F. Kahl, "A Brute-Force Algorithm for Reconstructing a Scene from Two Projections," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2011.

[8] B. Gutman, Y. Wang, T. Chan, P.M. Thompson, and A.W. Toga, "Shape Registration with Spherical Cross Correlation," *Proc. MICCAI Workshop Math. Foundations of Computational Anatomy,* pp. 56-67, 2008.

[9] P. Henry and M. Krainin, and E. Herbst, and X. Ren, and D. Fox, "RGB-D Mapping: Using Depth Cameras for Dense 2D Modeling of Indoor Environments," *Proc. Int'l Symp. Experimental Robotics,* 2010.

[10] B.K.P. Horn, "Closed-Form Solution of Absolute Orientation Using Orthonormal Matrices," *J. Optical Soc. Am. A,* vol. 5, no. 7, pp. 1127-1135, 1987.

[11] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "Kinectfusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera," *Proc. ACM Symp. User Interface Software and Technology,* 2011.

[12] B. Jian and B.C. Vemuri, "A Robust Algorithm for Point Set Registration Using Mixture of Gaussian," *Proc. IEEE Int'l Conf. Computer Vision,* vol. 2, pp. 1246-1251, 2005.

[13] A.E. Johnson and M. Hebert, "Surface Registration by Matching Oriented Points," *Proc. IEEE Conf. 3-D Digital Imaging and Modeling,* pp. 121-128, 1997.

[14] A.E. Johnson and S.B. Kang, "Registration and Integration of Textured 3D Data," *Image and Vision Computing,* vol. 17, no. 2, pp. 135-147, 1999.

[15] A. Lee, W. Sweldens, P. Shroder, L. Cowsar, and D. Dobkin, "MAPS: Multiresolution Adaptive Parameterization of Surfaces," *Proc. ACM Siggraph,* pp. 343-352, 1998.

[16] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int'l J. Computer Vision,* vol. 60, pp. 91-110, 2004.

[17] I.S. Okatani and A. Sugimoto, "Registration of Range Images that Preserves Local Surface Structure and Color," *Proc. Int'l Symp. 3D Data Processing, Visualization and Transmission,* pp. 786-796, 2004.

[18] C. Papazov and D. Burschka, "Stochastic Optimization for Rigid Point Set Registration," *Proc. IEEE Int'l Symp. Visual Computing,* pp. 1043-1054, 2009.

[19] R. Ramamoorthi, "Modeling Illumination Variation with Spherical Harmonics," *Face Processing: Advanced Modeling Methods,* pp. 385-424, Academic Press, 2006.

[20] S. Rusinkiewitcz and M. Leroy, "Efficient Variants of the ICP Algorithm," *Proc. Int'l Conf. 3D Digital Imaging and Modeling,* pp. 145-152, 2001.

[21] J.K. Seo, G.C. Sharp, and S.W. Lee, "Range Data Registration Using Photometric Features," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 1140-1145, 2005.

[22] T. Tachikawa, S. Hiura, and K. Sato, "Robust Estimation of Light Directions and Diffuse Reflectance of Known Shape Object," *Proc. Vision, Modeling and Visualization Workshop,* pp. 37-44, 2009.

[23] D. Thomas and A. Sugimoto, "Robustly Registering Range Images Using Local Distribution of Albedo," *Computer Vision and Image Understanding,* vol. 115, pp. 649-667, 2011.

[24] D. Thomas and A. Sugimoto, "Illumination-Free Photometric Metric for Range Image Registration," *Proc. IEEE Workshop Applications of Computer Vision,* pp. 97-104, 2012.

[25] A. Vedaldi, "SIFT Code for Matlab," http://www.vlfeat.org/~vedaldi/code/sift.html, 2006.

[26] K. Zhou, H. Bao, and J. Shi, "3D Surface Filtering Using Spherical Harmonics," *Computer Aided Design,* vol. 36, pp. 363-375, 2004.

Diego Thomas received the master's degree in informatics and applied mathematics from the Ecole Nationale Superieure d'Informatique et de Mathematiques Appliquees de Grenoble, France, in 2008. He received the PhD degree from the Graduate University for Advanced Studies in 2012. His research interests are 3D images registration, 3D reconstruction, and photometric analysis. He is a member of the IEEE.

Akihiro Sugimoto received the BS, MS, and Dr.Eng. degrees in mathematical engineering from the University of Tokyo in 1987, 1989, and 1996, respectively. After working for Hitachi Advanced Research Laboratory, Advanced Telecommunications Research Institute International, and Kyoto University, he joined the National Institute of Informatics, Tokyo, Japan, where he is currently a full professor. From 2006 to 2007, he was a visiting professor at ESIEE, France. He is interested in mathematical methods in engineering. In particular, his current main research interests include discrete mathematics, optimization algorithm, vision geometry, and modeling of human vision. He is the author or coauthor of more than 100 peer-reviewed journal/international conference papers. He is a regular reviewer for international conferences/journals in computer vision. He has also served several international conferences, including PSIVT '09 (general cochair), PSIVT '10 (general cochair), ACCV '12 (general cochair), ACCV '10 (program cochair), ACCV '09 (area chair), ICCV '09 (tutorial cochair), ICCV '11 (industrial liaison cochair), PSIVT '13 (workshop cochair). He received the Paper Award from the Information Processing Society in 2001. He is a member of the IEEE and the ACM.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.