Range Data Registration Using Photometric Features

Joon Kyu Seo*, Gregory C. Sharp**, and Sang Wook Lee*
*Dept. of Media Technology, Sogang University, Seoul, Korea
**Dept. of Radiation Oncology, Massachusetts General Hospital, MA, USA
samoosa@sogang.ac.kr, gsharp@partners.org, slee@sogang.ac.kr

Abstract

This paper investigates the use of photometric features for the pair-wise registration of range images. Many artificial and natural objects exhibit abundant surface texture that may not be revealed in range data, and most structured light and laser range sensors are capable of capturing either grayscale or color photometric intensity in addition to range data. Nevertheless, the use of photometric features has not been widely investigated for range data registration, despite widespread research into local feature descriptors for object recognition in 2D photometric images. This paper addresses some of the problems that arise in using photometric features for range data registration, and presents a systematic method for their use. Potentially useful photometric features are detected on planar regions in 3D, and then reprojected to 2D to remove the perspective distortion. Then, a well-established 2D rotation- and brightness-invariant image feature descriptor is used for matching. Range data alignment is performed using a RANSAC algorithm, with verification performed in 3D. Experimental results demonstrate the effectiveness of this method.

1. Introduction

Range sensors provide a fast and accurate way of measuring three-dimensional (3D) surfaces, but the because of limitation in field of view, multiple scans are required for most objects or scenes. When the relative position and orientation of each scan is unknown, various registration algorithms may be used to align the scans. Most registration methods try to optimize the distances between overlapping surfaces, and some methods employ additional 3D features derived from the range points. However, few methods use photometric features that can be extracted from the 2D photometric images, despite the fact that most commercial or homemade range sensors are capable of capturing registered 2D photometric and 3D range data. This paper discusses a strategy whereby photometric features can be used to simply and quickly register range data in 3D.

Traditional pair-wise registration methods construct the correspondence sets by extracting salient features from both views, and perform a search procedure to match the features. When reliable features can be extracted and correspondences correctly established, the motion that best aligns the two images can be solved directly. For low-profile 2.5D range images, however, there may not be sufficiently distinctive local shape features to establish correct correspondences. In other words, when a scene is rather plain, most features look the same. The distinctiveness further degrades when these features are processed to achieve rotational or affine invariance.

The iterative closest point (ICP) algorithm and its variants are a popular alternative to the feature-based approach for range image registration [1]. It is an iterative descent procedure that minimizes the sum of the squared distances between the points in the two views. The correspondence problem is solved by assuming that the scene is approximately aligned with the model and therefore that each scene point corresponds with its closest model point. However, the ICP method has the disadvantage that a good initial alignment is required to achieve a satisfactory final registration. Without a good initial alignment, many initial conditions must be tried, and convergence to a local minimum is unlikely. One solution to this problem is return to shape features as a guide to the registration process [2-7]. But shape features will still fail on scenes that are dominated by planar or smoothly varying surfaces.

The work presented in this paper is based on the notion that features detected in the intensity channel of the range data can provide a higher level of distinctiveness than the shape features. This is an intuitively simple idea, yet few methods are well known. One example is Johnson and Kang [8] who
describe an ICP algorithm that matches points using their color and range. However, this approach still suffers from correspondence errors, because by themselves, color values are not distinctive. Roth used “points of interest” extracted from the intensity images by the Harris corner detector [9] [10]. Nonetheless, the method by Roth does not utilize the local image information for feature matching, but relies only on the 3D range points associated with the detected feature points. On the other hand, Bendels et al. employed a well-established image feature descriptor [13] for the registration of bas-relief-like archaeological objects, but without any compensation for 3D rotation or projective distortion. [11]

We introduce an approach that uses distinctive photometric features for the alignment of range images. The chief focus of our work is on the modification of the 2D image features using the 3D range information for greater invariance to projective distortion than can be achieved by existing 2D image feature descriptors. A fast, randomized matching algorithm is used for direct global registration without initial alignment.

For feature matching, a 2D invariant to rotation is employed in our work to overcome the ambiguity that remains even after the compensation for projective distortion. There has been a considerable amount of research performed for generating local features with: rotation invariance [12], scale invariance [13], and affine invariance [14-15]. Our work benefits much from the recent advances in invariant feature generation research for object recognition.

The rest of this paper is organized as follows. Section 2 addresses the problems in using the 2D photometric features for range data registration and presents systematic solutions. Section 3 discusses the registration and alignment using photometric feature correspondences, and Section 4 presents experimental results. Discussions and summary are given in Sections 5 and 6, respectively.

2. Photometric feature matching

We propose a method that uses photometric features to register 3D range images. While photometric features are widely used in matching 2D images, corrections are needed for geometric and illumination variations when used for 3D images.

2.1 The problems with photometric features

Photometric features are a valuable source of information for image registration. However, two challenges must be addressed when computing feature values on three dimensional range images. The first challenge is geometric distortion due to the change in camera pose, and the second challenge is photometric distortion due to change in illumination pose.

We will address those problems by concentrating on the photometric features for planar or near-planar regions, where image plane homography and intensity normalization are sufficient. When the camera pose changes between views, features become distorted in the image plane. For example, rotating the camera about the target changes the projective warping as illustrated in Figure 1. Our approach to this problem is based on the feature rectification in 3D space using plane homography.

![Figure 1. Photometric feature matching](image)

Features can also be distorted by changes in illumination pose. For example, when the light source is rotated about the target, the illumination change is not generally uniform. For diffusely reflecting surfaces, the intensity $I(x,y)$ is related to the albedo $\rho$, the illumination $L$ and the surface normal $N(x,y)$ according to:

$$I(x,y) = \rho L \cdot N(x,y).$$

Discounting the effect of illumination change on a surface patch where $N(x,y)$ varies substantially is not trivial since the accurate computation $N(x,y)$ on a highly curved surface is difficult. If $N(x,y)$ is constant, on the other hand, the intensity changes uniformly with changes in $L$, and normalization of feature with respect to overall intensity is sufficient. We enforce this condition by limiting our attention to features that lie in nearly planar surface patches.

2.2 Feature rectification

Feature rectification is preprocessing step that allows us to accurately compute photometric features from 3D range images. It is designed to address the distortions in image geometry and intensity caused by changes in camera pose and illumination pose.

The geometric rectification algorithm comprises the following three stages:

1. **Feature candidate detection:** The first stage is to find feature candidates in the photometric images. In our implementation, we use the Harris corner detector [9] over 5×5 or 7×7 windows. All locations that have a
detector response over a predefined threshold are selected as features.

(2) **Planarity check and surface orientation estimation**: For each detected corner, the algorithm next checks the planarity of region near the feature. If the feature is nearly coplanar, we will use the feature for matching. Otherwise, the feature is rejected. These checks are performed over a 15×15 region using the singular value decomposition (SVD). Given \( n \) points in 3D where \( \vec{X}_i = [X_i \quad Y_i \quad Z_i]^T \), we solve the SVD as:

\[
UDV^T = \sum_{i=1}^{n} (\vec{X}_i - \vec{X}_m)(\vec{X}_i - \vec{X}_m)^T ,
\]

where \( \vec{X}_m = \sum_{i=1}^{n} \vec{X}_i \) is the mean position of the points.

To check for coplanarity, we inspect the ratio of the second largest singular value with the smallest singular value. If this ratio is greater than a predefined threshold, the feature is used for matching. For these features, we estimate the surface normal \( \vec{N} \) as the column of \( \vec{V} \) that corresponds to the smallest singular value.

\[
\text{Figure 2. Surface and projection geometry}
\]

(3) **Image warping**: Based on the surface orientation, a window containing each photometric image feature is warped to face the image plane.

Given that the viewing direction is known, and that we can recover the surface normal direction for a given point, we can solve for the distortion of viewing direction and scaling. The relationship between the captured and the warped image features is plane homography.

When a range sensor is calibrated, the 3×4 projection matrix \( \vec{P} \) in the following equation is known:

\[
\vec{x} = \vec{P} \vec{X} = \vec{K} [\vec{I} \ | \ 0] \vec{X} ,
\]

where \( \vec{x} = [x \quad y \quad 1]^T \) is the homogeneous coordinates in 2D image, \( \vec{X} = [X \quad Y \quad Z \quad 1]^T \) is the homogeneous coordinates of 3D range points in each view, and \( \vec{K} \), \( \vec{R} \) and \( \vec{T} \) are the intrinsic, rotation and translation components of the projection matrix \( \vec{P} \), respectively. Figure 2 illustrates the surface and camera projection geometry. The first step is to rotate the points so that the fitted plane is aligned with the original image plane. This requires finding a rotation \( \vec{R}_s \) that aligns the surface orientation \( \vec{N} \) with the direction of the principle axis of the camera \( \vec{e}_Z \), i.e., the unit vector in the Z direction. This is an underdetermined problem, and can be solved by adding an additional constraint. Let \( \vec{r} \) be the unit vector in the direction \( \vec{N} \times \vec{e}_Z \). Then, we have the 3×3 rotation matrix as:

\[
\vec{R}_s = \begin{bmatrix} \vec{e}_Z \quad \vec{r} \times \vec{r} \end{bmatrix} \begin{bmatrix} \vec{N} \quad \vec{r} \end{bmatrix} ,
\]

We may take \( x \) as the projection from a reference space as:

\[
\vec{x} = \vec{K} [\vec{I} \ | \ 0] \vec{X} .
\]

Then, given \( \vec{R} \), the image coordinates for the rotated surface is given as:

\[
\vec{x}' = \vec{K} [\vec{R}_s \ | \ 0] \vec{X} ,
\]

and the plane homography for image warping between \( \vec{x} \) and \( \vec{x}' \) can be easily obtained from:

\[
\vec{x}' = \vec{H} \vec{x} , \text{ where } \vec{H} = \vec{K} \vec{R}_s \vec{K}^{-1} .
\]

Finally, it remains to correct for the scale of the feature. If the original focal length is \( f \), and the distance to the fitted plane is \( \rho \), we have:

\[
\vec{x}' = \frac{\rho}{f} \vec{H} \vec{x} , \text{ where } \rho = -N^T \vec{X}_m .
\]

\[
\text{Figure 3. Rectification using surface orientation information}
\]

After rectification, the corresponding features should appear identical in the image plane up to rotation about the surface normal as illustrated in Figure 3. For the ambiguity due to this rotation on the image, a 2D feature invariant to rotation can be employed for matching.

**2.3 2D feature detection and matching**

After the rectified images are generated, we perform 2D local feature matching using the “Scale Invariant Feature Transform (SIFT)” algorithm, proposed by Lowe [13]. The SIFT algorithm was chosen because it incorporates most of the recent advances in invariant
feature generation for recognition and it is one of the most effective, fast and reliable feature descriptors in our experience. It has several important properties:

1) **2D-Rotation invariance**: By aligning the direction of the maximum intensity gradient with the x-axis, we can achieve invariance to the unknown 2D-rotation of the rectified image.

2) **Robustness to illumination changes**: Because the lighting pose changes for each view, the intensity varies within the planar region. The effect of intensity change is discounted in SIFT in the process of extracting the edge of the image.

3) **Affine compensation**: By using the surface orientation to correct perspective distortions, the rectified images are sensitive to noise of range data. Affine compensation reduces the effect of imprecise surface normal computations.

### 3. Registration

When $s_i$ is 3D point in scene, and $m_i$ is the corresponding 3D point in model, we seek a rigid transformation $T$ that minimizes the mean squared difference between scene and transformed model:

$$E = \sum |s_i - T(m_i)|^2.$$  \hspace{1cm} (8)

When computing mean squared difference, we have the flexibility of choosing a one to one matching algorithm or a one to many matching algorithm. One to one matching is preferred when most of the correspondences are correct. However, one to many matching can be useful when correspondence errors are common, due to repetitive patterns or a low degree of overlap between scene and model. For such a multiple matching algorithm, we use a random sampling method like RANSAC [16]. First, we compute a candidate transformation from a randomly chosen set of correspondences, and then we verify if the transformation is reasonable or not.

For a rigid 3D transformation, a minimum of three non-collinear point correspondences is required for a unique solution. With more correspondences, the accuracy of the transformation increases, but the probability of mismatched points also increases. For our local photometric features, we found that four features generally suffice.

Verification of the registration was accomplished through a view thresholding technique [17]. After the model is aligned with the scene using the photometric feature correspondences, the 3D points in the model are thresholded into inliers and outliers according to their visibility in the camera viewpoint of the scene. The outliers, points that are not visible in the scene, are discarded. The inliers are used for verification, by computing the mean squared distance from inliers to their matching point in the scene. Transformations can be generated and verified until a sufficiently good match between the surfaces is found.

### 4. Experimental Results

We used a homemade structured light range sensor, which consists of Epson EMP-7700 projector and Sony xc003 3CCD camera. It is calibrated with a calibration object prior to range scanning, and generates registered range and color reflectance images.

Examples of the registration results generated by our method are shown in Figs. 4-6. In Figure 4, the model is rotated approximately 50 degrees relative to the scene, causing considerable distortion of the photometric features due to foreshortening. Two views and feature correspondences generated by the presented algorithm are shown in Figures 4(a). Figure 4(b) shows the matches generated using the Lowe’s reference implementation of SIFT [18]. While SIFT generates a many correct matches on the object to the right, few matches are generated for the object on the left.

Although our algorithm requires that surfaces used for matching be locally planar, mildly curved surfaces can be used as well. Figure 5 shows the registration results for a scene that is composed primarily of curved surfaces. The registration succeeds for these kinds of scenes because our rectification procedure generates nearly identical photometric patches, as shown in 5(d). Figure 6 shows another set of results for a desktop scene.

### 5. Discussions

For many applications, we require a registration algorithm that can align scenes that contain mostly planar or smoothly varying surfaces. For these scenes, algorithms that rely primarily on geometry are faced with the difficult task of resolving the ambiguity in matching planes against planes. The use of photometric features in planar regions is a natural complement to shape features, and we anticipate future algorithms that blend these capabilities.

The computational complexity of our algorithm is low compared with other registration methods. Let each image contain $N$ points, $P$ features be detected, invariant images be size $W$, and let there be $M$ matched points. Then, the complexity of finding feature points is $O(N)$, and the complexity of making the descriptors is $O(FW)$. The registration process is $O(mM)$, where $m$ is the number of random trials used in the RANSAC procedure.
If the probability of a correct match is 50%, we require 35 trials to correctly match 3 pairs with 99% probability. The verification stage is $O(N)$ if correspondences are formed by projection, and average case $O(N \log N)$ if formed using closest point.

6. Summary

In this paper, we propose new method of 3D range data registration using 2D local photometric features. The features are rectified from 3D to 2D using the surface normal, and aligned using randomized matching of 2D affine invariants. Our proposed method is faster than most methods that rely on shape information, and we have demonstrated strong results for scenes with photometric texture.

7. Acknowledgements

This work was supported in part by Grant No. R01-2002-000-004720 from the Basic Research program of Korea Science and Engineering (KOSEF), and in part by the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea.

Figure 4. Plastic bottles: (a) two views and feature correspondences from the presented method, (b) feature detection and correspondences from Lowe’s reference SIFT implementation, (c) aligned range points (white/black), (d) a point-rendered novel view, and (e) rectified feature images. (Note that the displayed image size is 50×50 for viewing clarity.)
8. References