Registration of Range Images with Different Scanning Resolutions

Hongbin Zha †, Makoto Ikuta, Tsutomu Hasegawa
Department of Intelligent Systems, Kyushu University
6-10-1 Hakozaki, Higashiku, Fukuoka 812-8581 JAPAN
E-mail: zha@is.kyushu-u.ac.jp
(† also affiliated with Center for Information Science, Peking University, CHINA)

Abstract

The paper presents a novel method of registering two range images that are taken at different 3-D viewpoints. Distinct from classic methods, it allows an accurate estimation of a scaling parameter of the images in addition to the general motion parameters. It is made possible owing to the following efforts: 1) Extended signature images (ESIs) are utilized to establish precise correspondence between image objects even if they are different in scale. 2) Fine image alignment is achieved by incorporating scale estimates into a modified iterative closest point (ICP) algorithm. Results of experiments demonstrate the method is of practical use even for images acquired by real digitizers.

1 Introduction

A fast growing domain in computer science is 3-D object modeling which aims at creating a digital duplicate of any real object by using 3-D data. As more and more 3-D sensors become available, the technique has found a range of applications in entertainment, industry and scientific research. At the same time, however, we also encounter many difficulties, such as data acquisition, registration, and organization, in order to make the technology more robust [8, 9].

In the paper, we propose a new method of registering range images that are taken by a digitizer at arbitrary viewpoints. The images usually have diverse scanning resolutions determined by the distances from the sensor to the object surfaces. This requires that the registration algorithms be able to estimate the scaling parameter as well as the general motion parameters between images. To address the issue, our method attempts to achieve an accurate registration by carrying out iteratively the following steps: 1) Get an estimation of the scaling parameter between images by identifying and matching scale-invariant signature images (SI) normalized by surface curvatures. 2) Compute the remaining motion parameters on the basis of correspondences between some control points. The correspondence is established by closeness in both location and orientation of the control points.

The new aspect of the work is that all parameters involved in a general registration problem can be derived in a unified framework. As well-known in the vision community, registration in 3-D space is a highly nonlinear optimization problem, and to avoid troubles in its implementation, the scaling factor has been ignored until now [1, 9]. The problem is solved here mainly owing to the following two strategies: 1) Each loop is a global, rough matching step followed by a local, fine alignment step. 2) Scale-invariant curvatures are put into good use in defining the global features required in the first step.

2 Statements of Problem

The objective of range image registration is to match multi-view sensor images with proper relative positions [4, 10]. In principle, it is an optimization problem identifying parameters required for describing the transformation between images. There are three groups of parameters to be taken into account: translational, rotational, and scaling parameters.

While the first two groups were treated as main subjects in almost all existing methods, the scaling problem has attracted little attention [1, 2]. In reality, however, it arises frequently when we use a sensor that can be placed freely in 3-D space. For example, an object will be scanned with different resolutions if the images are taken from viewpoints at varying distances, resulting in projected objects scaled by different factors in a fixed image frame. Fig.1 shows two images for a bunny that is scanned at remote viewpoints. In this case, if the sampling interval on the real surfaces can not be calibrated precisely, the relative image scales have to be estimated.

In the following, we state how the parameter estimation problem is formalized in our research. Given two sensor images, we pick out one arbitrarily as a
base image \( I_b \). Hence, the scale \( s \) of the other image \( I \) can be defined with respect to \( I_b \). Moreover, assume that the motion between the images are described by a translation vector \( t \) and a rotation matrix \( R \). For any point \( p_i \), we can thus transform it into the world coordinate system attached to \( I_b \) as

\[
p'_i = sRp_i + t.
\]

(1)

If the corresponding point on \( I_b \) is known as \( p_{bi} \), the parameters are estimated by minimizing

\[
E = \frac{1}{N} \sum \|p'_i - p_{bi}\|,
\]

(2)

where \( N \) is the total number of the control points in each image, and \( \|x - y\| \) the Euclidean distance between \( x \) and \( y \). In our method, the optimization problem is solved by a recursive algorithm including the following steps in each loop.

1. Estimate the scale parameter by matching corresponding point pairs that are extracted from extended signature images (ESIs).

2. Derive initial values for the other motion parameters using the obtained image correspondence.

3. Better the estimates by a modified iterative closest point (ICP) registration algorithm.

These processes will be explained in detail in the following sections.

3 Extend Signature Images (ESIs)

As stated above, our method is relied on a pointwise shape description called extended signature images (ESIs), which is explained in the section.

3.1 Definition of ESI

Signature images (SI) are proposed originally to generate a global object representation for any point on the object surface [6, 9]. It is a 2-D image mapping the lowest three degrees of differential quantities of the other points onto the image array. It is computed mainly by the following steps. For a given point \( p \) and for any other point \( q_i \), 1) compute the distance \( d_i \) between them and the angle \( \alpha_i \) between the normal \( n_p \) and the line \( pq_i \), as illustrated in Fig.2; 2) index \( q_i \) in the image array by quantizing \( d_i \) and \( \alpha_i \); 3) set the pixel value by a normalized curvature at \( q_i \). The steps are carried out for each of the other point, and the final image is referred to as the SI for \( p \). An example of SI is given in Fig.3 for a real image.

![Figure 2. \( d_i \) and \( \alpha_i \) for the points \( p \) and \( q_i \).](image)

![Figure 3. SI for the image point \( p \).](image)

SI is a unique representation for any point if the object surface does not possess a strong symmetry characterized typically by a surface of revolution[7]. Hence, it has been widely used for identifying corresponding point pairs in images related by a 3-D motion. The identification is usually implemented by a template matching operation. Assume pixels in two SIs are denoted by \( g(i, j) \) and \( t(i, j) \) respectively. Then, the similarity of the two image points are evaluated by

\[
E^2 = \frac{1}{N_d^2} \sum_{(i,j) \in D} \|g(i,j) - t(i,j)\|^2, \quad O = \frac{N_d}{N_d},
\]

(3)

where \( D \) is the overall region of the SIs containing \( N_d \)
points, and $N_{dx}$ the number of points overlapping in the SIs. $O$ is used here for coping with occlusions in the images.

Unfortunately, the SI can not be used straightforwardly in our cases because the uniqueness of SI does not hold true if the scale factor is taken into consideration. To clarify the point, we make a simple investigation into the relationship between the image scale and surface curvatures. Assume that we are given two images $I_b$ and $I$, in which the same object is projected with different scales. Let the scale parameter of $I$ be $s$. It is not hard to see that the mean curvatures $H_b$ and $H$ of the two images are related by

$$H = \frac{1}{s} H_b$$  \hspace{1cm} (4)

for the same point on the real surface. This means that the change in the image scale will make the SI values vary in inverse proportion to $s$.

By the same way, we find that the scale has also a direct effect on the indices of the SIs, as shown in Fig.4. In order to eliminate these effects, we normalize the SI for $I$ by: 1) scaling the $d_i$ index by $s$; 2) scaling the pixel values by $1/s$, where $s$ is the ratio of the mean curvatures of $I$ to $I_b$. We refer to such a SI as the extended signature image (ESI) to distinguish it from the original SI.

![Figure 4. Effects of scale factors on SIs.](image)

3.2 Scale Estimation Based on ESIs

The ESIs defined above maintain the uniqueness regardless of scale changes. Therefore, we can improve the reliability in establishing correspondence between two images by matching the ESIs of some control points. To save the computational cost, we reduce the number of the control points by the following steps: 1) Select image points of relatively large curvatures as candidates of the control points; 2) Match the candidates to determine correct pairings. The candidates that fail to find adequate counterparts are abandoned.

Once the false pairings are removed partially in this manner, we obtain an estimate of the scale by calculating the average ratio of curvatures of the point pairs. The accuracy of the estimation can be enhanced further by a simple clustering in the histograms of curvature ratios.

4 Fine Alignments of Images

In actuality, the above estimate of the scale parameter is merely a loose approximation due to errors in the surface features, in particular, in the surface curvatures. To remedy the drawback, we developed a fine alignment algorithm that is based on the so-called iterative closest point (ICP) algorithm [1, 3].

![Figure 5. Modified ICP algorithm.](image)

A sketch of the algorithm is given in Fig.5. The inner loop is almost the same as the original ICP algorithm except for the way of finding the closest points. In the original method, the closeness is measured strictly by the 3-D Euclidean distances [11]. For a point $p_b$ in an image $I_b$, the closest point $p'$ in another image $I$ is searched out completely on the basis of

$$p' = \min ||p_i - p|| \quad \text{for all} \ p_i \in I.$$  \hspace{1cm} (5)

In our method, the decision on the closest point is a little more complicated. At first, the closest point $q$ in terms of distances is found as did in the original method. Next, a neighborhood $B_q = \{p_{q_i,t_0} = ...$
of \( q \) is extracted. Then, the closest point \( q' \) is determined by

\[
q' = \min | \cos^{-1} p_q^T p_{q_i} | \tag{6}
\]
in \( B_q \). Here, we take into account not only the closesness in distance but also the similarity in local shapes. As results, the fine alignment converges more quickly than the original algorithm.

Once the inner loop converges, the correspondence between the control points can be modified according to the estimated motion. The procedure proceeds as follows: 1) Transform the image \( I \) into the world coordinate system by using the resulting affine parameters. 2) Compute the distances between the transformed points and the base image points. 3) Annul the correspondence if the distance is large as compared with the mean of the distances. This procedure benefits greatly the fine alignment of the next loop, as demonstrated in Fig. 6. In Fig. 6(a), the corresponding point pairs include a number of false ones. If these pairs are delivered directly to the next loop, the resulting registration is not as good as expected. After our correspondence modification, however, the result gains remarkably, as shown in Fig. 6(b).

Finally, the scale parameter \( s \) is refined in turn, and the outer loop is repeated. The whole algorithm terminates if the registration error is smaller than a specified threshold.

5 Results of Experiments

To verify the applicability of our method in various situations, we have conducted experiments using simulation data or real images. Results of two of them are reported here.

![Figure 7. Image data for the 1st experiment.](image)

In the first experiment, we use as the base image \( I_a \) a sampling dataset provided by Raindrop Geomagic Studio. It contains 10,968 points and 21,928 triangular facets. Then, we generate three new images which are magnified reproductions of \( I_a \) with scales 1.0, 1.6, and 2.2, respectively. After applied with further motions, the three images \( I_{a1}^{1.0}, I_{a1}^{1.6} \), and \( I_{a1}^{2.2} \) are used as those to be registered with \( I_a \). The images are shown in Fig. 7. The initial values for the scale parameter are set as \( 0.8 < s < 2.4 \).

![Figure 8. Distributions of individual estimates.](image)
Distributions of the estimated scales are illustrated in Fig.8. The different symbols in plots indicate the point pairs for different images. The horizontal axis shows the similarity values ($E_n^2/O$ computed in eq.(3)) for the point pairs, while the vertical axis the value of estimated scale parameters. It is evident that the estimates concentrate, as expected, around the real values indicated by the dot lines.

Next, we show the results of the fine alignments in Fig.9. The upper graph shows the errors, and the below the values of the scale parameters, at the iterations. The values at the 0th iteration amount to those obtained from the initial scales given in Fig.7. Fig.10 shows the final results for the registration of the image $I_a^{2.2}$. It illustrates that the registration is successful despite a large difference in the image scale.

In the second experiment, two images to be registered are both acquired by a real Minolta Vivid700 digitizer. They are taken, respectively, at distances 1.5m and 1.3m to the object, and are denoted by $I_{a,5}$ and $I_{a,3}$.

Since the Vivid700 digitizer is able to recover the scale parameter automatically, we have to operate on the images to simulate the situation we assume in the paper. At first, we make a new image $I_{p,0}$ by magnifying $I_{a,3}$ with a scale of 1.5. Thus, we can regard the real scale between $I_{p,0}$ and $I_{a,3}$ as 1.5. Next, we generate images $I_{p,1}$, $I_{p,2}$, and $I_{p,3}$ that are all parts of $I_{p,0}$, as shown in Fig.11. The purpose of the experiment is to register these images to $I_{p,0}$. The initial values for the scale are $1.0 < s < 2.0$.

Results of the initial estimation of the scale parameters are shown in Fig.12(a)-(d), and the final registration corresponding to the cases is illustrated in Fig.12(e)-(h), respectively. Some data on the registration errors or used CPU times are given in Table 1. We can observe here that our algorithm works well even for an image that is only a part of the base image. For such an image, however, we have to raise the number of control points to guarantee a reasonable initial estimation of motion parameters. Roughly speaking, the control points need to be added as the area the partial image occupies decreases.

6 Conclusions

We have introduced a new range image registration method for a scanner that can be set freely in 3-D space. We envision many applications for the method, including modeling of large sculptures or even whole indoor environments. We will also try to extend the framework into a system that can merge data from different kinds of sensors.
Table 1. Data derived in the 2nd experiment

<table>
<thead>
<tr>
<th></th>
<th>$I_{p0}$</th>
<th>$I_{p1}$</th>
<th>$I_{p2}$</th>
<th>$I_{p3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>num of point</td>
<td>14153</td>
<td>10221</td>
<td>9668</td>
<td>6947</td>
</tr>
<tr>
<td>num of CPs</td>
<td>50</td>
<td>60</td>
<td>88</td>
<td>90</td>
</tr>
<tr>
<td>scale after ESI</td>
<td>1.605</td>
<td>1.502</td>
<td>1.493</td>
<td>1.452</td>
</tr>
<tr>
<td>scale after ICP</td>
<td>1.495</td>
<td>1.495</td>
<td>1.510</td>
<td>1.507</td>
</tr>
<tr>
<td>SI err</td>
<td>10153</td>
<td>10221</td>
<td>9668</td>
<td>6947</td>
</tr>
<tr>
<td>post-ICP err</td>
<td>0.509</td>
<td>1.109</td>
<td>1.160</td>
<td>0.397</td>
</tr>
<tr>
<td>time (ESI, sec)</td>
<td>25.6</td>
<td>27.9</td>
<td>36.12</td>
<td>32.7</td>
</tr>
<tr>
<td>time (ICP, sec)</td>
<td>15.1</td>
<td>11.5</td>
<td>15.2</td>
<td>9.7</td>
</tr>
</tbody>
</table>

References


