A Probabilistic Method for Aligning and Merging Range Images with Anisotropic Error Distribution

Ryusuke Sagawa   Nanaho Osawa   Yasushi Yagi
The Institute of Scientific and Industrial Research, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, JAPAN
{sagawa,n-oosawa,yagi}@am.sanken.osaka-u.ac.jp

Abstract

This paper describes a probabilistic method of aligning and merging range images. We formulate these issues as problems of estimating the maximum likelihood. By examining the error distribution of a range finder, we model it as a normal distribution along the line of sight. To align range images, our method estimates the parameters based on the Expectation Maximization (EM) approach. By assuming the error model, the algorithm is implemented as an extension of the Iterative Closest Point (ICP) method. For merging range images, our method computes the signed distances by finding the distances of maximum likelihood. Since our proposed method uses multiple correspondences for each vertex of the range images, errors after aligning and merging range images are less than those of earlier methods that use one-to-one correspondences. Finally, we tested and validated the efficiency of our method by simulation and on real range images .

1. Introduction

Many researchers have studied modeling shapes of real world objects by scanning them using three dimensional digitizers such as laser range finders [5, 11, 3] and structured-light range finders [22]. For example, a major subject of research is the modeling of cultural heritage objects [12, 9] because they are seen to be candidates having the worthiest shapes and appearances for modeling.

3D modeling of the shape of an object is accomplished by following three steps:

1. Acquiring the range images (scanning).

2. Aligning the acquired range images from different viewpoints (aligning).

3. Reconstructing a unified 3D mesh model (merging).

In the first step, a target object is observed from various viewpoints. In the second step, multiple range images are aligned into a common coordinate system using registration algorithms that establish point correspondences and minimize the total distance between those points; using, for example, feature-based [24, 10], or Iterative Closest Point (ICP)-based [1, 29, 2, 14, 17, 18] methods. The third step is to merge multiple pre-aligned range images. Several approaches have been proposed; for example, mesh-based [25, 23] and volume-based [8, 4, 26, 19] methods.

When aligning and merging multiple range images it is necessary to find corresponding points between them. There is some variation among the methods used. WithICP-based methods, the corresponding points are defined by the closest Euclidean distance [1, 29], and the projection of the source point onto the destination mesh [2, 14, 17, 15]. If an Euclidean distance is used, it is equivalent to assuming that the error distribution of a range image is isotropic. However, the error distribution of an actual range finder does not seem isotropic. If the distance is computed by projecting the source point onto the destination mesh, the error is distributed on the line of projection. It seems plausible because the error distribution of a range finder is along the line of sight, especially for a laser range finder. However, as these methods find one-to-one correspondences, it is difficult to find a correct correspondence when the measurement error of a range finder is large, even if the initial position of range images is a sufficiently good guess.

The methods for merging range images also find corresponding points by using the closest Euclidean distance [25, 8, 26, 19], and projecting a point onto a mesh [4]. For the same reason as in the case of aligning range images, it is difficult to find a correct correspondence when the measurement error of a range finder is large, and the assumption of the error distribution is far from the actual distribution.

Some approaches that model the error distribution of the range measurement have been proposed. Okatani and Deguchi[16] proposed a method to align range images with an error model of range images and modeled it as a normal distribution along the line of sight. However, they deterministically computed the one-to-one corresponding points. Williams and Bennamoun [28] proposed a proba-

bilistic method to align range images. Though they modeled the error distribution by a covariance matrix, they also used deterministic correspondences. Sagawa et al. [20] proposed a method to remove the noise in range images along the line of sight, but they did not simultaneously consider the error distribution of multiple range images.

If we assume that the error distribution of a range finder is along the line of sight, the noise of each pixel can be reduced by simple temporal averaging by obtaining multiple range images from the same viewpoint. However, it is not applicable in the case that the sensor or object moves when acquiring those range images. Our proposed method can be applied even if the sensor is moving, and only one range image is obtained for each view direction.

In this paper, we propose probabilistic methods to align and merge range images. We regard aligning range images as problems of finding optimal parameters of maximum likelihood. Our merging method finds a surface of maximum likelihood during the conversion from range images to a volumetric representation. Because we analyze the error distribution of a laser range finder, we simplify the model for the measurement error and using an anisotropic error model we propose algorithms for aligning and merging range images. In Section 2, we formulate the problems as maximum likelihood estimates. Then, we analyze the error distribution and propose practical algorithms in Section 3. We evaluate our method in Section 4 and finally summarize this paper in Section 5.

2. Probabilistic Method for Aligning and Merging Range Images

This section describes a method to align and merge range images based on a probabilistic framework. First, we define a probabilistic model of the error distribution of a range image that is obtained by a range finder. Next, we define the problems with the model of error distribution as follows:

- Aligning range images has the problem that estimates the parameters that have maximum likelihood.

- Merging range images has the problem that finds a surface that has a maximum likelihood.

2.1. Probabilistic Model of Error Distribution in Range Measurement

A range image is a set of 3D points acquired by a range finder and which are connected if their positions are close to each other [25]. Therefore, a range image is represented as a mesh model, consisting of vertices and patches. Because of the noise in range measurement, a 3D point in a range image has an error. Figure 1 shows an example where there are a true surface S and a observed range image A. x is a point of A. If x is a measurement of a point s of S, the probability is represented as a posterior probability $P(x|s)$. If a general model of error distribution is assumed, x can
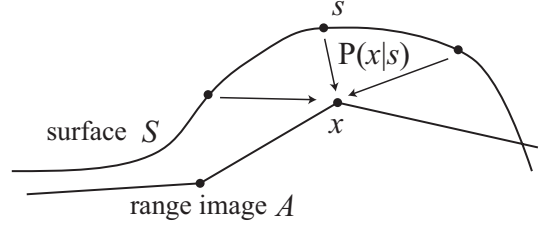
Figure 1. A range image A is observed by scanning a true surface S. A point x of A is a measurement of a point s of S. The posterior probability is $P(x|s)$.
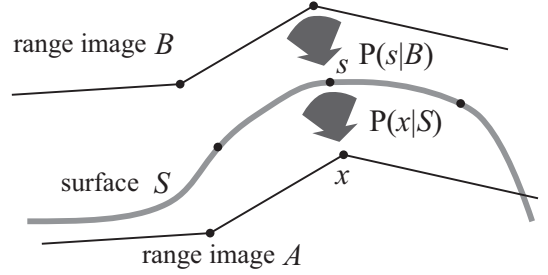
Figure 2. Two range images A and B are two samples of observing a true surface S. If B is observed, the probability of S becomes $P(S|B) = \prod_s P(s|B)$, and the probability that A is observed is $P(A|S) = \prod_x P(x|S)$ if S exists.

be caused by all parts of S. Thus, we denote the posterior probability as $P(x|S)$ and the probability of a whole range image A becomes

$$P(A|S) = \prod_x P(x|S). \qquad (1)$$

2.2. Aligning Range Images by Maximum Likelihood Estimation

Figure 2 shows a situation where there are two range images A, B and a true surface S. Now, we consider the alignment of the range image A to B by changing the parameters of rotation R and by translation t of the range image A. In a probabilistic framework, the parameter is computed by maximizing the probability $P(A|B; \theta)$, where θ is a vector that represents the parameters R and t. $P(A|B; \theta)$ is computed as follows:

$$P(A|B; \theta) = \int_{S \in \Omega} P(A|S; \theta) P(S|B) dS, \qquad (2)$$

where Ω represents the scope of the shape of the true surface S.

Using the Expectation Maximization (EM) approach, the parameter θ that maximizes $P(A|B; \theta)$ is computed by estimating the following conditional expectation

$$E[\log P(S|B)|A] = \int_{S \in \Omega} P(S|A; \theta) \log P(S|B) dS. \quad (3)$$

If S consists of n points $\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n$, (3) becomes

$$\int_{\boldsymbol{s}_1} \cdots \int_{\boldsymbol{s}_n} \prod_j P(\boldsymbol{s}_j|A;\theta) \log \prod_i P(\boldsymbol{s}_i|B) d\boldsymbol{s}_1 \cdots d\boldsymbol{s}_n$$

$$= \int_{\boldsymbol{s}_1} \cdots \int_{\boldsymbol{s}_n} \prod_j P(\boldsymbol{s}_j|A;\theta) \sum_i L(\boldsymbol{s}_i|B) d\boldsymbol{s}_1 \cdots d\boldsymbol{s}_n \quad (4)$$

where $L(\boldsymbol{s}_i|B)$ is the log likelihood $\log P(\boldsymbol{s}_i|S)$. Because $\int_{\boldsymbol{s}_i} P(\boldsymbol{s}_i|A;\theta) = 1$ for all $i = 1, \cdots, n$, it is simplified as

$$E[\log P(S|B)|A] = \sum_i \int_{\boldsymbol{s}_i} P(\boldsymbol{s}_i|A;\theta) L(\boldsymbol{s}_i|B) d\boldsymbol{s}_i. \quad (5)$$

Since an actual range finder has a characteristic error distribution, we propose an algorithm to compute and maximize (5) based on the error distribution in Section 3.2.

2.3. Merging Range Images by Finding Surface of Maximum Likelihood

To merge range images, we compute a signed distance field (SDF) as an intermediate representation. Since it is an implicit representation, the merged surface becomes a set of points $\boldsymbol{x}$ that satisfies $f(\boldsymbol{x}) = 0$, where $f(\boldsymbol{x})$ is a signed distance at $\boldsymbol{x}$. The implicit surface represented by a SDF is converted to a mesh model by the Marching Cubes algorithm [13]. In this section, we propose a method to merge range images by computing a signed distance $f(\boldsymbol{x})$ of maximum likelihood from multiple range images.

Figure 3 shows an example where there are two range images A, B, a true surface S and a point $\boldsymbol{x}$. $D_i(\boldsymbol{x})$ is the distance from $\boldsymbol{x}$ to a point $\boldsymbol{s}_i$ of S. The probability function for $D_i(\boldsymbol{x}) \geq d$ is computed by

$$P(D_i(\boldsymbol{x}) \geq d) = \int_{\|\boldsymbol{x}-\boldsymbol{s}_i\| \geq d} P(\boldsymbol{s}_i|A, B;\theta) d\boldsymbol{s}_i, \quad (6)$$

where θ is the parameter of alignment. Thus, $P(D(\boldsymbol{x}) = d)$ is computed by taking average of all neighbor points as follows:

$$P(D(\boldsymbol{x}) = d) = \lim_{\delta \to 0} \frac{\sum_i^N P(D_i(\boldsymbol{x}) \geq d-\delta) - P(D_i(\boldsymbol{x}) \geq d+\delta)}{2\delta N}, \quad (7)$$

where N is the number of neighbor points.

Since the magnitude of a signed distance $f(\boldsymbol{x})$ is computed as the distance from $\boldsymbol{x}$ to the nearest neighbor point of the surface [19], then if $\boldsymbol{s}_i$ is the nearest neighbor point, the other points, for example $\boldsymbol{s}_k$ in Figure 3, are farther than $\boldsymbol{s}_i$ from $\boldsymbol{x}$. Since the probability $P_1(D(\boldsymbol{x}) = d)$ that only one point is within $[d-\delta, d+\delta)$ is

$$\lim_{\delta \to 0} \frac{\prod_i P(D_i(\boldsymbol{x}) \geq d-\delta) - \prod_i P(D_i(\boldsymbol{x}) \geq d+\delta)}{2\delta}, \quad (8)$$

the probability density function for $|f(\boldsymbol{x})| = d$ is computed by

$$P(|f(\boldsymbol{x})| = d) = P(D(\boldsymbol{x}) = d) \cdot P_1(D(\boldsymbol{x}) = d) \quad (9)$$

By finding the maximum likelihood of $P(|f(\boldsymbol{x})| = d)$, we determine the magnitude of the signed distance.
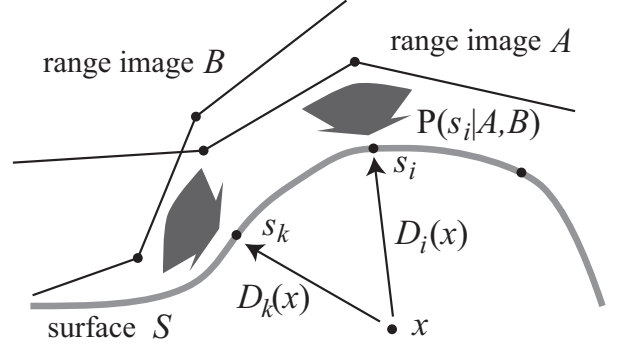
Figure 3. An example of merging two range images. $D_i(\boldsymbol{x})$ is the distance from $\boldsymbol{x}$ to a point $\boldsymbol{s}_i$ of a true surface S.

The remaining issue is to determine the sign of $f(\boldsymbol{x})$. In [19], the sign is determined by considering the angle between $\boldsymbol{p} - \boldsymbol{x}$ and the normal vector $\boldsymbol{n}$ at $\boldsymbol{p}$, if the nearest point is $\boldsymbol{p}$. Namely, the sign of $f(\boldsymbol{x})$ becomes positive if $(\boldsymbol{p} - \boldsymbol{x}) \cdot \boldsymbol{n} < 0$, and negative otherwise.

In our method, if the magnitude of $f(\boldsymbol{x})$ is d, we determine the sign by computing a weighted sum of the probability:

$$W(\boldsymbol{x}, d) = \int_{\boldsymbol{p}} \int_{\|\boldsymbol{x}-\boldsymbol{s}_i\|=d} P(\boldsymbol{s}|\boldsymbol{p}) \text{sgn}((\boldsymbol{p} - \boldsymbol{x}) \cdot \boldsymbol{n}) d\boldsymbol{p} d\boldsymbol{s}, \quad (10)$$

where $\boldsymbol{p}$ is a point of the range images and $\text{sgn}(a)$ is 1 if $a < 0$ and negative otherwise. Therefore, the sign of $f(\boldsymbol{x})$ becomes $\text{sgn}(W(\boldsymbol{x}, d))$.

In this section, we explain our method for merging just two range images; however, it can easily be extended to cases of three or more range images.

3. Algorithm with Anisotropic Error Model

This section describes a practical algorithm to align and merge range images. Since it is difficult to compute the probabilities described in Section 2 for general probabilistic models, we first analyze the error distribution of an actual laser range finder and simplify the model of error distribution. Next, we propose algorithms based on the model.

3.1. Simplified Model of Anisotropic Error Distribution

Here, we used a Canesta DP200 [3] as the laser range finder to analyze the error distribution. As a preliminary experiment, we evaluated the error distribution of the sensor by observing a small object. Since it has a light source and a camera, it obtains a 2D range image by observing the reflected light. The position and size of the small object is adjusted so that it is projected within a pixel of the image.

Figure 4 shows the results of the measurements. Figure 4(a) is the histogram of the number of measurements for each pixel along the horizontal axis of the image. Since most of the measurements are projected onto only one pixel,

(a)

(b)
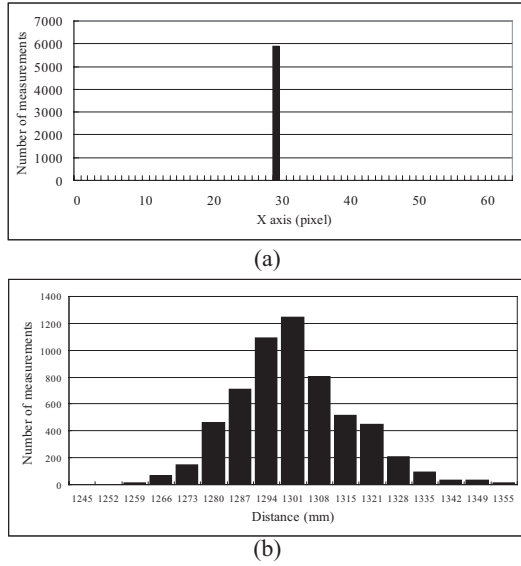
Figure 4. Histograms of measuring a small object: (a) along the horizontal axis of a range image, (b) along the line of sight.

it can be assumed that a pixel is not affected by other pixels, which is reasonable if an object is in focus. It is equally reasonable to assume that the error exists only along the view direction. Figure 4(b) is the histogram of the number of measurements for a pixel along the line of sight. Since the distribution is similar to a normal distribution, we assume that it is a normal distribution $N(d, \sigma^2)$, where d is the distance to the true surface and σ is the standard deviation. Since [21] analyzed for the standard deviation of the error, we use the result. Consequently, since we assume that the error only exists along the view direction, the model of error distribution is simplified to a 1D normal distribution. Though this error model is validated only for the Canesta DP200, it can be applied to other range finders, especially laser range finders.

3.2. Aligning Range Images

Figure 5 shows a situation where there are two range images A, B to be aligned and an assumed true surface S. If we assume the model of error distribution described in Section 3.1, a vertex x_i of the range image A is a measurement of s_i of the surface S, which is on the same line of sight of the range image A. The probability of x_i becomes

$$P(\boldsymbol{x}_i|\boldsymbol{s}_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp(\frac{- \parallel \boldsymbol{x}_i - \boldsymbol{s}_i \parallel^2}{2\sigma^2}). \quad (11)$$

Meanwhile, the measurement of s_i in the range image B is $\boldsymbol{y}_i$, which is on the same line of sight of B. The probability of $\boldsymbol{y}_i$ is similar to (11). Since $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ only depend on $\boldsymbol{s}_i$,
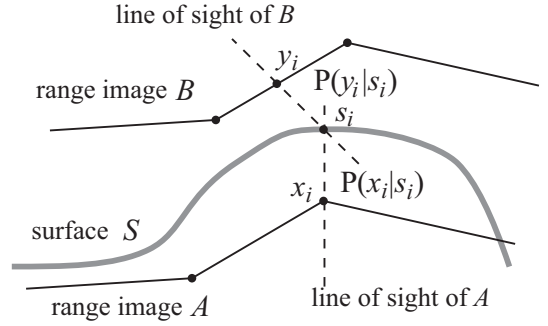
Figure 5. Correspondence points between range images A, B and an assumed true surface S. $\boldsymbol{s}_i$ is on the lines of sight of both $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$.

(5) becomes

$$E[\log P(S|B)|A] = \sum_i \int_{\boldsymbol{S}_i} P(\boldsymbol{s}_i|\boldsymbol{x}_i; \theta) L(\boldsymbol{s}_i|\boldsymbol{y}_i) d\boldsymbol{s}_i. \quad (12)$$

In the maximization step of the EM algorithm, the estimated parameter $\hat{\theta}$ becomes by Bayes's rule

$$\begin{aligned} \hat{\theta} &= \arg_\theta \max \sum_i \int_{\boldsymbol{S}_i} P(\boldsymbol{s}_i|\boldsymbol{x}_i; \theta) L(\boldsymbol{s}_i|\boldsymbol{y}_i) d\boldsymbol{s}_i \\ &= \arg_\theta \max \sum_i \int_{\boldsymbol{S}_i} P(\boldsymbol{x}_i|\boldsymbol{s}_i; \theta) L(\boldsymbol{y}_i|\boldsymbol{s}_i) d\boldsymbol{s}_i, (13) \end{aligned}$$

where we drop $P(\boldsymbol{s}_i), P(\boldsymbol{x}_i)$ and $P(\boldsymbol{y}_i)$ because they are constant with respect to the parameters if we assume that there is no constraint on the shape of S.

To resolve (13), we compute the weighted sum of several samples of $\boldsymbol{s}_i$ instead of computing the integral of $\boldsymbol{s}_i$ as shown in Figure 6. Since the error distribution is a normal distribution, it becomes

$$\hat{\theta} = \arg_\theta \min \\ \sum_i \sum_j P(\boldsymbol{s}_{i,j}|\boldsymbol{x}_i) \parallel \boldsymbol{y}_{i,j} - \boldsymbol{s}_{i,j} \parallel^2, \quad (14)$$

where $\boldsymbol{s}_{i,j}$ is a sample of $\boldsymbol{s}_i$ and $\boldsymbol{y}_{i,j}$ is the corresponding point of $\boldsymbol{s}_{i,j}$. The distribution of $\boldsymbol{s}_{i,j}$ is given by user. $P(\boldsymbol{s}_{i,j}|\boldsymbol{x}_i)$ is computed by

$$\frac{1}{2}|\text{erf}(\frac{t_j + t_{j+1}}{2\sqrt{2}\sigma}) - \text{erf}(\frac{t_j + t_{j-1}}{2\sqrt{2}\sigma})|, \quad (15)$$

where $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$, and $\boldsymbol{s}_j = \boldsymbol{x} + t_j \boldsymbol{v}$ ($\boldsymbol{v}$ is the unit vector along the line of sight). Therefore, the EM algorithm iteratively minimizes the distance of the corresponding points. This is similar to the ICP-based methods, especially [14, 15], which find corresponding points along the view direction. However, since they do not consider the error distribution, they use only the correspondence between $\boldsymbol{x}_i$ and $\boldsymbol{y}_{i3}$ in the case of Figure 6. Thus, our algorithm is the same as [15] for minimizing an energy function, but it uses a different energy function (14).

4
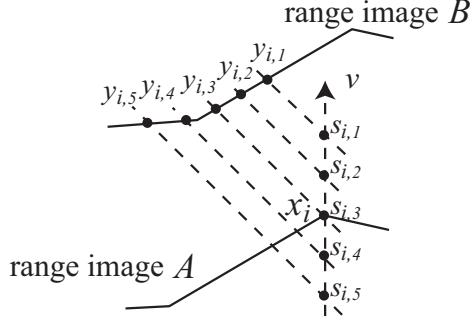
Figure 6. Computing the weight sums of samples of $s_{i,j}$ instead of computing the integral of s_i.

3.3. Merging Range Images

The next step is to merge range images. To find the maximum of the probability of $P(|f(\boldsymbol{x})| = d)$, we create a histogram instead of computing the limit of the probability expressed in (9). Our algorithm to compute the signed distance $f(\boldsymbol{x})$ is as follows:

1. Find several neighbor points of range images whose distances from $\boldsymbol{x}$ are smaller than $d_{\min} + 3\sigma$, where $d_{\min}$ is the distance to the nearest neighbor point and σ is the standard deviation of measurement.

2. Compute $P(D_i(\boldsymbol{x}) \geq d_k)$ for $k = 1, 2, ...n$ for each found neighbor point, where the bin of a histogram is defined between d_k and d_{k+1}.

3. Compute the sum and product of the probability function for all neighbor points:

$$P(D(\boldsymbol{x}) = d_k) = \sum_i \{P(D(\boldsymbol{x}) \geq d_{k+1}) - P(D(\boldsymbol{x}) \geq d_k)\}$$
$$P_1(D(\boldsymbol{x}) = d_k) = \prod_i P(D_i(\boldsymbol{x}) \geq d_{k+1}) - \prod_i P(D_i(\boldsymbol{x}) \geq d_k)$$

4. Find the bin b_k that has the maximum value of $P_{b_k} = P(D(\boldsymbol{x}) = d_k) \cdot P_1(D(\boldsymbol{x}) = d_k)$.

5. Interpolate the distance $|f(\boldsymbol{x})|$ of the maximum probability by fitting a quadratic function to $P_{b_{k-1}}, P_{b_k}$ and $P_{b_{k+1}}$.

6. The sign is determined by computing (10) for the bin b_k.

To find the neighbor points, we use a k-d tree [7] to reduce the cost with a small modification to find points whose distances are within $d_{\min} + 3\sigma$. Since other points of the range images farther than $d_{\min} + 3\sigma$ do not contribute to the probability $P(D(\boldsymbol{x}) = d_k)$, we omit them from the computation. The width of a bin is defined by the user because it is a trade-off between cost and accuracy. Since our algorithm is based on [19], the position of $\boldsymbol{x}$ is determined by an octree; for details, please refer to [19]. Figure 7 shows an example
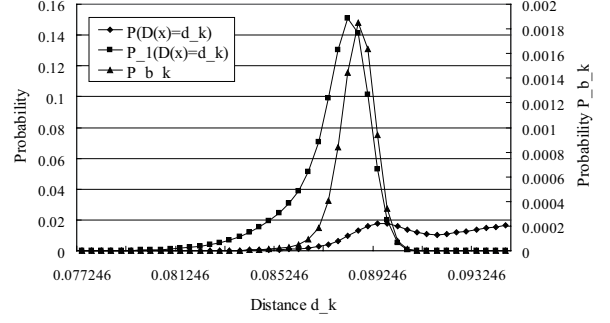
Figure 7. An example of the probabilities $P(D(\boldsymbol{x}) = d_k)$, $P_1(D(\boldsymbol{x}) = d_k)$ and P_{b_k} according to the distance d_k.
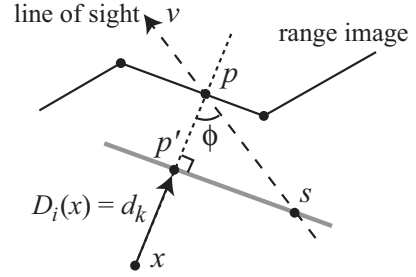
Figure 8. Our method finds the nearest point $\boldsymbol{p}$ for each patch. The standard deviation of d_k is approximated by $\sigma' = \sigma \cos\phi$.

of the probabilities $P(D(\boldsymbol{x}) = d_k)$, $P_1(D(\boldsymbol{x}) = d_k)$ and P_{b_k} according to the distance d_k.

The remaining issue is how to compute $P(D_i(\boldsymbol{x}) \geq d_k)$. As shown in Figure 8, we find the nearest point $\boldsymbol{p}$ for each patch of range images from $\boldsymbol{x}$. A patch is usually a triangle that connects adjacent vertices. Though the point $\boldsymbol{s}$ of the true surface exists along the line of sight, since the nearest point $\boldsymbol{p}'$ is interpolated inside of the patch, the distribution of the nearest point becomes smaller. In this paper, we simplify the computation of $\boldsymbol{p}'$ by approximating the standard deviation of d_k by $\sigma' = \sigma \cos\phi$. Therefore, $P(D_i(\boldsymbol{x}) \geq d_k)$ is computed by

$$P(D_i(\boldsymbol{x}) \geq d_k) = \frac{1}{2}(1 - \text{erf}(\frac{t}{\sqrt{2}\sigma'})) \qquad (16)$$

where $\boldsymbol{p}' = (1 + t)(\boldsymbol{p} - \boldsymbol{x})$.

4. Experiments

We evaluate our method using synthetic data and real range images. First, we compare our aligning and merging methods with previous methods by using synthetic range images, to which noises are added along the line of sight. Second, we show that our method successfully creates a model of an object from noisy range images.
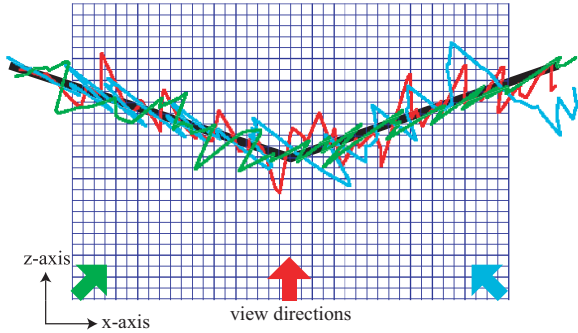
Figure 9. Three examples of slices of synthetic range images in the xz-plane. Their shape and view directions are indicated by colors. The original mesh model is a thick black line. The interval of grids is 0.02.

4.1. Comparison by using Synthetic Range Images

We created one hundred synthetic range images by adding noise of normal distribution to a mesh model from ten different viewpoints. Figure 9 shows an original mesh model and three examples of slices of synthetic range images in the xz-plane. The original model consists of two planes, which are perpendicular to the xz-plane. In this case, the standard deviation of noise is 0.05 while the width of a original mesh model is 1.0 and the interval between the vertices is 0.025. The maximum difference between viewing angles is 90 degrees.

4.1.1 Aligning Range Images

We estimate the robustness of the alignment with respect to the initial parameters by comparing our method with one of the previous methods [15], which finds corresponding points along the align of sight of one of range images. Since the original mesh model consists of two planes, it has ambiguity if we align the models. Thus, we estimate the error of the alignment results by comparing the difference of translation along the x- and z-axes and the view direction from the ground truth. We align two range images starting with various initial positions. The offsets from the ground truth are uniformly distributed within $[-0.25, 0.25]$ along each axis. Table 1 shows the robustness of the estimated parameters. X-std. and Z-std. are the standard deviations of translation along the x- and z-axes. RMS is the root-mean-square distance from the ground truth. Dir. is the difference of the viewing directions from the ground truth in degrees. Since our method has smaller standard deviation and errors than [15], our method can estimate the parameters robustly with respect to the variation of the initial parameters.

4.1.2 Merging Range Images

To test the proposed merging algorithm, we merge two range images that are acquired from different viewpoints. The positions of the input range images are at the ground truth. Figure 10 shows two examples of the results of merging by the proposed method (red line) and a previous

Table 1. Robustness of the estimated parameters.

|	Our method	Previous method [15]
X-std.	0.051	0.080
Z-std.	0.0025	0.0063
RMS	0.121	0.146
Dir.	5.81	7.29
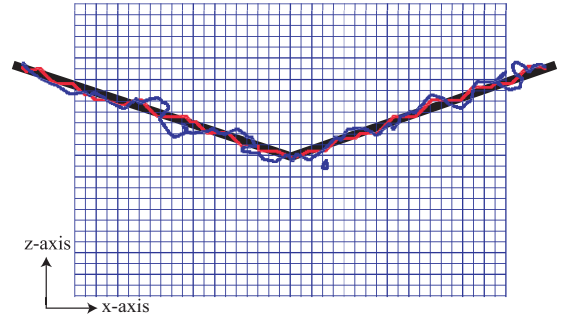
Figure 10. The result of merging 10 range images. The red line indicates the results of the proposed method and the blue line indicates the results of the previous method [19].

method [19] (blue line), which uses the average Euclidean distance to the nearest points. To estimate the error of the merged mesh model, we compute the offset from the original model along the z-axis of the vertices. The RMS error of the input range images is 0.0402. We test ten times for each algorithm; the RMS error of the result of our method becomes 0.0268, while that of the previous method is 0.0293. Our method obtains a better result than the previous method and reduced the noise to 66% of the input range images.

4.2. Aligning and Merging Real Range Images

Next, we align and merge real range images captured by the Canesta DP200, which obtains range images at 30 frames/sec, but is less accurate than other slower laser range finders. The standard deviation is 8mm along the line of sight if an object is at about 1m from the sensor. We sequentially captured 72 range images of an object shown in Figure 11 from various directions by rotating it on a turntable. Figure 12 shows one of the range images. From the top view, we can see that the range image contains large noise compared to the object size, which is about 420mm wide.

We sequentially aligned the range images by pairwise registration. The images are aligned automatically by setting the initial parameter of rotation and translation to be the same as the aligned result of the previous frame. Figure 13 shows the estimated sensor position and viewing direction. Because the result of our method (red line) becomes an accurate circle, our method successfully aligned the range images while the result of the previous method [15] (blue line) collapsed during registration.
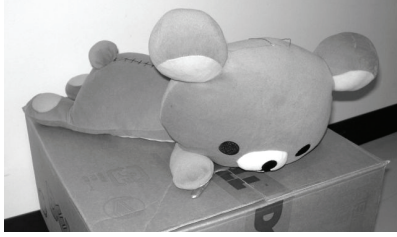
**Figure 11.** A stuffed bear on a box.
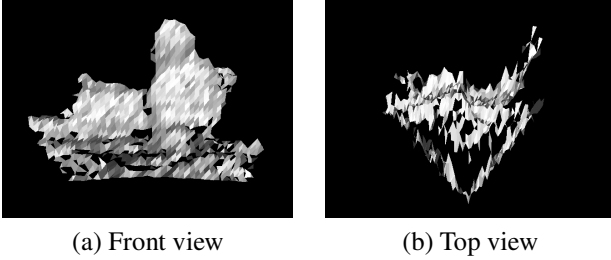
| (a) Front view | (b) Top view |

Figure 12. A range image of the stuffed bear contains large noise compared to the object size.

Figure 15(a) shows the results of alignment, and Figure 14(a) is a horizontal slice of the aligned range images. Though the shape of the object can be recognized from them, it is represented as a cloud of points or meshes. By applying our merging method to the range images, we construct a SDF. In Figure 14(b) the red lines are a slice of the mesh model extracted from the SDF. By comparing the merged model by the previous method [19] (blue lines), our method reduces the noise in the merged model. Figure 15(b) is the result of merging range images. Since our method does not assume any prior model of the object, the mesh model is not smooth. After smoothing the SDF by a mean filter, a smooth mesh model is extracted as shown in Figure 15(c).

5. Conclusion

This paper proposed a probabilistic method of aligning and merging range images. We formulated the issues involved as maximum likelihood estimates. By examining the error distribution of a range finder, we modeled it as a normal distribution along the line of sight. For aligning range images our method estimates the parameters based on the EM approach. By assuming an error model, the algorithm is implemented as an extension of an ICP-based method. For merging range images, our method computes the signed distances by finding the distances of maximum likelihood, and successfully reduced the sensor noise. Since our proposed method uses multiple correspondences for each vertex of range images, the errors after aligning and merging range images are less than those of earlier methods that use one-to-one correspondences. In this paper, we do not assume any prior model of the shape of a true surface. Thus, the resulting surface of merging may not be smooth, as was
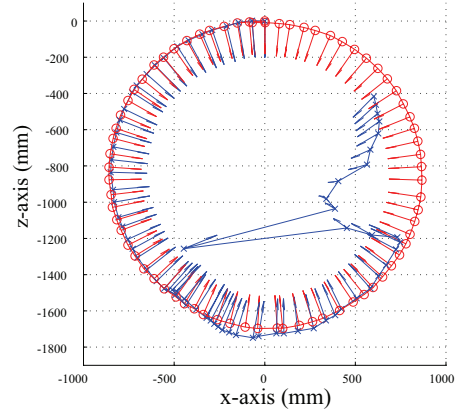
Figure 13. The estimated sensor position and viewing direction by our proposed method (red) and the previous method [15] (blue).
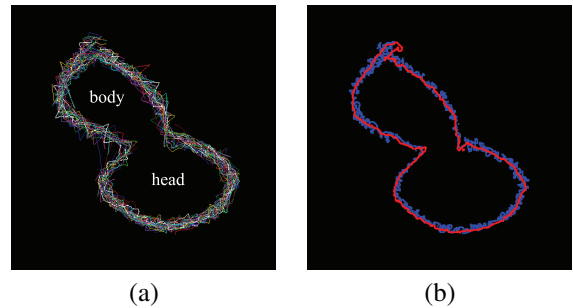

| (a) | (b) |

Figure 14. Horizontal slices of models: (a) aligned range images, (b) merged models by the previous method [19] (blue) and the proposed method(red).

shown in the results. Earlier models have also been introduced during the reconstruction of 3D models [27, 6]. In future work we intend to incorporate a prior model of the surface into our framework to generate a smooth surface.

References

[1] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Trans. Patt. Anal. Machine Intell.*, 14(2):239–256, Feb 1992.

[2] G. Blais and M. Levine. Registering multiview range data to create 3d computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, 1995.

[3] Canesta, Inc. CanestaVision EP Development Kit. http://www.canesta.com/devkit.htm.

[4] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proc. SIG-GRAPH'96*, pages 303–312. ACM, 1996.

[5] Cyra Technologies, Inc. Cyrax 2500. http://www.cyra.com.

[6] H. Dinh, G. Slabaugh, and G. Turk. Reconstructing surfaces using anisotropic basis functions. In *Proc. International Conference on Computer Vision*, pages 606–613, Vancouver, Canada, July 2001.

| (a) Aligned range images | (b) Merged mesh model | (c) Smoothed mesh model |
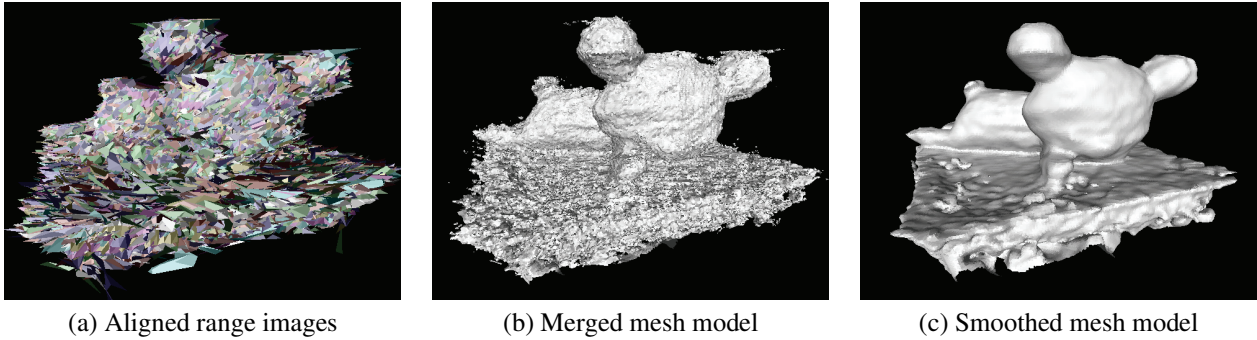
Figure 15. The results of aligning and merging range images. (b) is extracted from the SDF and (c) is the result after smoothing the SDF.

[7] J. Friedman, J. Bentley, and R. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 3(3):209–226, 1977.

[8] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. In *Proc. SIGGRAPH'92*, pages 71–78. ACM, 1992.

[9] K. Ikeuchi. Digitally archiving cultural heritage. In *Proc. 10th International Conference on Virtual Systems and Multimedia (VSMM2004)*, November 2004.

[10] A. Johnson and M. Hebert. Surface registration by matching oriented points. In *Proc. Int. Conf. On Recent Advances in 3-D Digital Imaging and Modeling*, pages 121–128, May 1997.

[11] Konica Minolta Photo Imaging U.S.A., Inc. Vivid 9i non-contact digitizer. http://www.minoltausa.com/vivid/.

[12] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The digital michelangelo project: 3D scanning of large statues. In *Proc. SIGGRAPH 2000*, pages 131–144, 2000.

[13] W. Lorensen and H. Cline. Marching cubes: a high resolution 3d surface construction algorithm. In *Proc. SIGGRAPH'87*, pages 163–170. ACM, 1987.

[14] P. Neugebauer. Geometrical cloning of 3d objects via simultaneous registration of multiple range images. In *Proc. Int. Conf. on Shape Modeling and Application*, pages 130–139, Mar 1997.

[15] T. Oishi, R. Sagawa, A. Nakazawa, R. Kurazume, and K. Ikeuchi. Parallel alignment of a large number of range images. In *Proc. 3DIM 2003*, pages 195–202, 2003.

[16] I. Okatani and K. Deguchi. A method for fine registration of multiple view range images considering the measurement error properties. *Computer Vision and Image Understanding*, 87:66–77, 2002.

[17] K. Pulli. Multiview registration for large data sets. In *Proc. Second International Conference on 3D Digital Imaging and Modeling*, pages 160–168, 1999.

[18] S. Rusinkiewicz and M. Levoy. Efficent variant of the ICP algorithm. In *Proceedings of the 3rd International Conference on 3-D Digital Imaging and Modeling*, pages 145–152, 2001.

[19] R. Sagawa, K. Nishino, and K. Ikeuchi. Adaptively merging large-scale range data with reflectance properties. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):392–405, March 2005.

[20] R. Sagawa, T. Oishi, A. Nakazawa, R. Kurazume, and K. Ikeuchi. Iterative refinement of range images with anisotropic error distribution. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 79–85, October 2002.

[21] R. Sagawa, N. Osawa, T. Echigo, and Y. Yagi. Real time 3d environment modeling for a mobile robot by aligning range image sequences. In *Proc. British Machine Vision Conference 2005*, volume 1, pages 330–339, Oxford, UK, September 2005.

[22] K. Sato and S. Inokuchi. Range-imaging system utilizing nematic liquid crystal mask. In *Proc. International Conference on Computer Vision*, pages 657–661, 1987.

[23] M. Soucy and D. Laurendeau. A general surface approach to the integration of a set of range views. *IEEE Trans. Patt. Anal. Machine Intell.*, 17(4):344–358, April 1995.

[24] F. Stein and G. Medioni. Structural indexing: efficient 3-d object recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(2):125–145, 1992.

[25] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *Proc. SIGGRAPH'94*, pages 311–318, Jul 1994.

[26] M. D. Wheeler and K. Ikeuchi. Sensor modeling, probabilistic hypothesis generation, and robust localization for object recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(3):252–265, March 1995.

[27] R. Whitaker. A level-set approach to 3d reconstruction from range data. *International Journal of Computer Vision*, 29(3):203–231, October 1998.

[28] J. Williams and M. Bennamoun. Multiple view surface registration with error modeling and analysis. In *Proc. IEEE Int. Conf. on Image Processing (ICIP '00)*, volume I, pages 545–548, Sep 2000.

[29] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.