An Extension of the ICP Algorithm for Modeling Nonrigid Objects with Mobile Robots

Dirk Hähnel[†‡], Sebastian Thrun[†], Wolfram Burgard[‡†]

[†]Carnegie Mellon University, School of Computer Science, PA, USA,
[‡]University of Freiburg, Department of Computer Science, Germany

Abstract

The iterative closest point (ICP) algorithm [2] is a popular method for modeling 3D objects from range data. The classical ICP algorithm rests on a rigid surface assumption. Building on recent work on nonrigid object models [5; 16; 9], this paper presents an ICP algorithm capable of modeling nonrigid objects, where individual scans may be subject to local deformations. We describe an integrated mathematical framework for simultaneously registering scans and recovering the surface configuration. To tackle the resulting high-dimensional optimization problems, we introduce a hierarchical method that first matches a coarse skeleton of scan points, then adapts local scan patches. The approach is implemented for a mobile robot capable of acquiring 3D models of objects.

1 Introduction

In recent years, there has been a flurry of work on acquiring 3D models from range data. The classical setting involves a range sensor (e.g., a 3D range camera or a stereo vision system) used to acquire range images of the target object from multiple vantage points. The problem of integrating multiple range scans into a 3D model is commonly known as *scan registration*. Most state-of-the-art implementations are based on the popular *iterative closest point* algorithm [2]. The topic has received significant attention in fields as diverse as computer vision [12; 9] and medical imaging [7], large-scale urban modeling [14], and mobile robotics [10; 8; 15].

ICP aligns range scans by alternating a step in which closest points are identified, and a step by which the optimal translation and rotation of scans relative to each other is computed. In doing so, most of them are making a rigid object assumption: range scans, if aligned correctly, must be spatially consistent with each other.

Many objects are deformable. For example, people change shape, as do trees, pillows, and so on. A natural research goal is therefore to extend ICP to accommodate local object transformations. Following recent work primarily found in the medical imaging literature [5; 16; 6; 3], we proposes an approach suited for scan registration and 3D modeling of nonrigid objects that can efficiently deal with hundreds of thousands of variables. To accommodate local deformations, our approach transforms scans into Markov random fields, where nearby measurements are linked by a (nonlinear) potential function. All links are soft. They can be bent, but bending them incurs a penalty. Figure 1 illustrates this transformation: rigid links between the measurement coordinates and the robot sensor are replaced by soft links between adjacent measurement points. The resulting problem of scan registration under these soft constraints becomes a high-dimensional optimization problem with orders of magnitude more variables involved than in regular ICP. We show how to solve this problem via Taylor series expansion (linearization), and then propose a coarse-to-fine hierarchical optimization technique for carrying out the optimization efficiently.

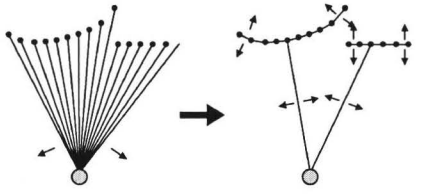

Figure 1: The essential idea: Rather assuming that the relation of measurement coordinates are fixed relative to the location from which the measurement was taken, the approach proposed here constraints the relation of measurement coordinates in a soft way. The exact configuration of a scan is calculated while scans are registered to each other.

Our new algorithm is applied to the problem of learning 3D models of non-stationary objects with a mobile robot. We describe an implemented robot system that utilizes a model differencing technique similar to the one described in [1] to segment scans. By acquiring views of the target objects from multiple sides, our approach enables a robot to acquire a 3D model of a non-stationary object.

2 Scan Registration

This section describes a variant of the well-known *iterative closest point* algorithm (ICP) [2] for *rigid* objects. The algorithm alternates two phases, one in which nearest points is identified, and one in which the distance between all pairs of nearest points is minimized.

2.1 Scans

The input to the ICP algorithm is a set of 3D scans denoted

$$\mathbf{d} = \mathbf{d}_1, \mathbf{d}_2, \ldots \quad (1)$$

(a) mobile robot with scanner (b) scene acquired by scanner (c) result of model differencing

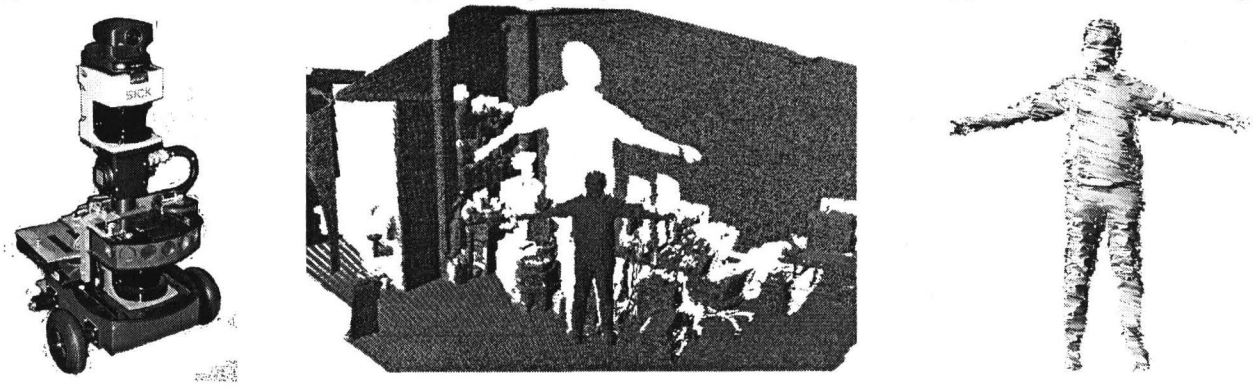

Figure 2: (a) 3D range scan acquired by mobile robot shown in (b). (c) Scene from which the scan in (a) is extracted through background differencing.

Each such scan $\mathbf{d}_k$ consists of a collection of 1D range measurements, arranged as a 3D "range image":

$$\mathbf{d}_k = d_{1k}, d_{2k}, \ldots \tag{2}$$

Figure 2c shows a typical scan acquired by a robot. Further below, we will denote the horizontal angle of the k-th range measurement by α_k and the vertical one by β_k. In our experimental setup, scans are obtained by a SICK laser range finder mounted on a tilt unit shown in Figure 2a.

The problem of scan registration can be formulated as the problem of recovering the vantage points from which the scans were taken. In our approach, scans are taken from a mobile robot; hence each vantage point is described by three variables: its x-y location in Cartesian coordinates and its orientation γ:

$$\mathbf{x}_k = (x_k, y_k, \gamma_k)^T \tag{3}$$

Here $\mathbf{x}_k$ denotes the vantage point from which scan $\mathbf{d}_k$ was acquired. The set of all vantage points will be denoted $\mathbf{x}$. Recovering the vantage points $\mathbf{x}$ is equivalent to registering the scans if one vantage point is (arbitrarily) defined to be $\mathbf{x}_1 = (0,0,0)^T$.

2.2 Measurement Model

Scans are registered in world coordinates. To do so, measurements d_{ik} must be mapped into 3D world coordinates. This is achieved by a projective function π, which takes as an argument a range measurement and a vantage point and returns as its output the corresponding coordinate in 3D world coordinates:

$$\pi(d_{ik}, \mathbf{x}_k) = \begin{pmatrix} x_k + d_{ik}\cos(\gamma_k + \alpha_i)\sin\beta_i \\ y_k + d_{ik}\sin(\gamma_k + \alpha_i)\sin\beta_i \\ z + d_{ik}\cos\beta_i \end{pmatrix} \tag{4}$$

Here α_i and β_i are the orientation of the range measurement d_{ik} relative to the sensor. The variable z is the generic height of the sensor which in our robot system is fixed, but is easily generalized to variable heights. For brevity, we will sometimes write π_{ik} instead of $\pi(d_{ik}, \mathbf{x}_k)$.

We define the quality of a pairwise scan registration by the probability representing the likelihood $p(d_{ik})$ of a range measurement under a fixed (hypothetical) registration:

$$p(d_{ik} \mid \mathbf{d}_l, \mathbf{x}_k, \mathbf{x}_l) = \tag{5}$$
$$\begin{cases} \max\limits_j \; |2\pi\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\pi_{ik}-\pi_{jl})^T\Sigma^{-1}(\pi_{ik}-\pi_{jl})\right\} & \text{if } \pi_{ik} \in \mathcal{F}_l \\ \text{const.} \quad \text{if } \pi_{ik} \notin \mathcal{F}_l \end{cases}$$

This likelihood distinguishes two cases. The first case models the noise in range perception by a Gaussian with a sensor-specific covariance Σ (usually a diagonal matrix). The measurement error under this Gaussian is given by the distance between the point π_{ik} under consideration, and the point j in scan $\mathbf{d}_l$ that maximizes this Gaussian. As is easily to be seen, this point π_{jl} minimizes the Mahalanobis distance to π_{ik}. Thus, π_{jl} is simply the point "nearest" to π_{ik} in the scan $\mathbf{d}_l$ under Σ^{-1}.

However, finding the nearest point only makes sense if π_{ik} falls within the perceptual range of scan $\mathbf{d}_l$. If π_{ik} is occluded, it might be perfectly well-explained by an object not detectable from the vantage point $\mathbf{x}_l$. This is captured by the second case in (5), which applies when π_{ik} lies outside the free space of scan l. The free space of scan l is denoted $\mathcal{F}_l$. It is defined as the region between the robot and the detected objects. If π_{ik} lies outside this region, the measurement probability is assumed to be uniform. The value of this uniform depends on the range of occluded space, but it plays no role in the optimization to come; hence we leave it unspecified.

2.3 Registration as Likelihood Maximization

The goal of scan registration is to determine the vantage points $\mathbf{x}$ that maximize the joint likelihood of the scans $\mathbf{d}$ and $\mathbf{x}$:

$$p(\mathbf{d}, \mathbf{x}) = p(\mathbf{d} \mid \mathbf{x})\, p(\mathbf{x}) \tag{6}$$

The first term of this product is obtained by calculating the product over all individual measurement likelihoods, assuming noise independence in each measurement:

$$p(\mathbf{d} \mid \mathbf{x}) = \prod_k \prod_i \prod_{l \neq k} p(d_{ik} \mid \mathbf{d}_l, \mathbf{x}_k, \mathbf{x}_l) \tag{7}$$

The term $p(\mathbf{x})$ in (6) is the *prior* on the individual vantage points (and hence on the registration). The prior of the point $\mathbf{x}_k$ is expressed as a Gaussian with mean $\bar{\mathbf{x}}_k$ and covariance matrix Ψ:

$$\begin{aligned} p(\mathbf{x}) &= \prod_k p(\mathbf{x}_k) \qquad (8) \\ &= \prod_k |2\pi\Psi|^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\Psi^{-1}(\mathbf{x}_k - \bar{\mathbf{x}}_k)\right\} \end{aligned}$$

In our robot system, this prior is obtained from the robot's 2D localization routines, supplied by a public domain software package [11].

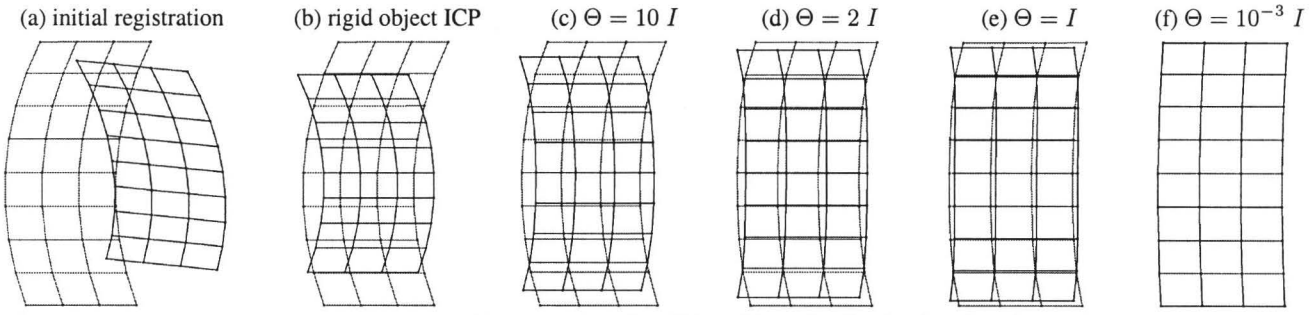

Figure 3: Illustration in 2D: (a) initial configurations of two scans in red and blue, (b) result of optimal registration, (c-f) result of recovering the object configuration by our new adjustable shape technique.

The negative logarithm of the joint likelihood (6) is given by the following sum:

$$\begin{aligned} G &= -\log p(\mathbf{d}, \mathbf{x}) \\ &= \text{const.} + \tfrac{1}{2} \sum_k \left[(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \Psi^{-1} (\mathbf{x}_k - \bar{\mathbf{x}}_k) \right. \\ &+ \left. \sum_i \sum_{l \neq k} \min_j \ (\pi_{ik} - \pi_{jl})^T \Sigma^{-1} (\pi_{ik} - \pi_{jl}) \right] \end{aligned} \tag{9}$$

Scans are registered by minimizing the sum, which is equivalent to maximizing the likelihood function.

Unfortunately, minimizing G is not possible *in closed form* for the following three reasons: (1) the exact terms in the sum depend on the occlusion constraint which are a function of the vantage points $\mathbf{x}$; (2) the result of the minimization over j involves a discrete search for a nearest neighbor, whose outcome is again a function of $\mathbf{x}$; (3) the projection functions π are nonlinear in the points $\mathbf{x}$, and the resulting nonlinear optimization problem is hard.

2.4 Optimization Procedure

ICP-style algorithms minimize the negative log-likelihood by calculating a sequence of vantage points (registrations):

$$\mathbf{x}_k^{[0]}, \mathbf{x}_k^{[1]}, \mathbf{x}_k^{[2]}, \ldots \tag{10}$$

The first set of points are obtained from the prior (e.g., the robot odometry): $\mathbf{x}_k^{[0]} = \bar{\mathbf{x}}_k$. The $(n+1)$-th vantage points are computed from the n-th ones by the following five step algorithm:

Step 1. The set of occluded points is determined for the n-th points $\mathbf{x}_k^{[n]}$. This step involves the calculation of a relative orientation of a point π_{ik} to scan $\mathbf{d}_l$ via $\mathbf{x}_l$. It furthermore involves a range comparison with the corresponding measurement in scan $\mathbf{d}_l$, to determine if π_{ik} is occluded relative to $\mathbf{d}_l$.

Step 2. The minimization in (9) is carried out by determining the closest measurement j in each scan relative to each point π_{ik}:

$$l_{ikl}^{[n]} = \min_j \ (\pi_{ik} - \pi_{jl})^T \Sigma^{-1} (\pi_{ik} - \pi_{jl}) \tag{11}$$

For $\Sigma = \text{const.} \cdot I$, this calculation is equivalent to finding the closest points (hence the name of the algorithm).

Step 3. The projection function π is linearized to obtain a quadratic objective function. This is achieved by the Taylor series expansion:

$$\pi(d_{ik}, \mathbf{x}_k) \approx \pi_{ik}^{[n]} + \mathbf{J}_{ik}^{[n]} (\mathbf{x}_k - \mathbf{x}_k^{[n]}) \tag{12}$$

Here $\mathbf{J}_{ik}^{[n]}$ is the Jacobian (gradient) of π at $\mathbf{x}_k^{[n]}$, which is directly obtained from the definition of π in (4):

$$\mathbf{J}_{ik}^{[n]} = \begin{pmatrix} 1 & 0 & -d_{ik} \sin(\gamma_k + \alpha_i) \sin \beta_i \\ 0 & 1 & d_{ik} \cos(\gamma_k + \alpha_i) \sin \beta_i \\ 0 & 0 & 0 \end{pmatrix} \tag{13}$$

With this approximation, $G^{[n]}$ (the negative log-likelihood function G in the n-th iteration) becomes quadratic in $\mathbf{x}_k$ and is of the form

$$\begin{aligned} G^{[n]} &= \text{const.} + \tfrac{1}{2} (\mathbf{x} - \bar{\mathbf{x}})^T \Psi^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \\ &+ \tfrac{1}{2} (\mathbf{A}^{[n]} \mathbf{x} - \mathbf{c}^{[n]})^T \tilde{\Gamma}^{-1} (\mathbf{A}^{[n]} \mathbf{x} - \mathbf{c}^{[n]}) \end{aligned} \tag{14}$$

Here $\mathbf{A}^{[n]}$ is a matrix, $\mathbf{c}^{[n]}$ a vector, and $\mathbf{x}$ the vector of all vantage points (the unknowns). The various sums in (9) are simply integrated into $\mathbf{A}^{[n]}$ and $\mathbf{c}^{[n]}$, and $\tilde{\Sigma}$ generalizes Σ to the full space of all points $\mathbf{x}$:

$$\tilde{\Sigma}^{-1} = \begin{pmatrix} \Sigma^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma^{-1} \end{pmatrix}^{-1} \tag{15}$$

Step 4. The x-y-values of the vantage points $\mathbf{x}^{[n+1]}$ are now obtained through a closed-form solution of minimizing $G^{[n]}$. This solution is obtained by setting the first derivative of $G^{[n]}$ with respect to $\mathbf{x}$ to zero (derivation omitted, see [13]):

$$\begin{aligned} &\mathbf{x}^{[n+1]} \\ &= (\Psi^{-1} + \mathbf{A}^{[n]T} \Sigma^{-1} \mathbf{A}^{[n]})^{-1} \left(\Psi^{-1} \bar{\mathbf{x}} + \mathbf{A}^{[n]T} \Sigma^{-1} \mathbf{c}^{[n]} \right) \end{aligned} \tag{16}$$

In the original ICP algorithm [2], this step is implemented by calculating the the "center of mass" of each scan and shifting the scans accordingly (which is computationally simpler); our approach also accommodates a prior $\bar{\mathbf{x}}$ on $\mathbf{x}$.

Step 5. Because of the linearization, only Cartesian coordinates are actually updated in Step 4; the orientation coordinates remain unmodified. For this reason, the original ICP literature introduced a separate step for calculating orientations. This is achieved by a singular value decomposition (SVD) step, in which the correlation between residual errors in the scan is determined, and each scan is rotated so as to minimize these error correlations. Details on this step can be found in the ICP literature [2].

All steps are iterated until a convergence criterion is reached. Satisfactory registrations are usually obtained within the first three to four iterations. If the total number of scans is small (e.g., less than 10), the computationally most expensive

step is the determination of the closest points in Step 2. This step is usually implemented efficiently by representing scans through kd-trees [4].

Figure 3a-b illustrate the result of scan registration in 2D. The initial configuration in Figure 3a is transformed into the one shown in Figure 3b, which is the one that minimizes the squared distance (maximizes the likelihood). Clearly, both scans are incompatible in shape. Pure registration techniques are unable to handle such shape deformations, but the technique presented in the next section is.

3 Recovering the Surface Configuration of Nonrigid Objects

The key idea for extending ICP to nonrigid objects was already discussed in the introduction to this paper, and is highlighted in Figure 1. Technically, it involves two modifications: First, the static relationship between points π_{ik} and the corresponding vantage points $\mathbf{x}_k$ is replaced by nonrigid links between adjacent points. These links can be bent (but at a probabilistic penalty), to accommodate nonrigid surfaces. Second, and as consequence of this, the optimization now involves the determination of the location of all points π_{ik}, in addition to the robot poses $\mathbf{x}_k$. This optimization problem is much higher dimensional, and we will discuss a hierarchical optimization technique for tackling it efficiently. A key characteristic of the approach proposed here is that it fits neatly into the ICP methodology above: Again, under appropriate linearization the target function is quadratic, and estimates are obtained just as in (16).

3.1 Links

The definition of links between pairs of adjacent points makes it necessary to augment the notion of a measurement point. In particular, our approach associates an (imaginary) coordinate system with each node. The origin of each coordinate system is the familiar coordinate π_{ik}, and its orientation is specified by three Euler angles (an alternative formulation may use quaternions):

$$\mathbf{r}_{ik} = (\phi_{ik}, \theta_{ik}, \psi_{ik})^T \tag{17}$$

The orientation is initialized arbitrarily; e.g. $\mathbf{r}_{ik} = (0,0,0)^T$. (The result of the optimization is invariant with respect to this initialization.) A link is now given by the affine coordinate transformations among the coordinate systems of adjacent measurements. Each link possesses six parameters, three for rotation (denoted $\Delta\mathbf{r}_{i\to j,k}$) and three for translation (denoted $\Delta\pi_{i\to j,k}$). They are calculated as follows:

$$\begin{aligned} \Delta\mathbf{r}_{i\to j,k} &= \mathbf{r}_{jk} - \mathbf{r}_{ik} \\ \Delta\pi_{i\to j,k} &= R_z(-\psi_{ik}) \cdot R_y(-\theta_{ik}) \cdot R_x(-\phi_{ik}) \begin{pmatrix} x_{jk} - x_{ik} \\ y_{jk} - y_{ik} \\ z_{jk} - z_{ik} \end{pmatrix} \end{aligned} \tag{18}$$

Here the R's are the rotation matrices around the three coordinate axes. Links enable us to recover a node's coordinates from any of its neighbors:

$$\begin{aligned} \pi_{jk} &= \underbrace{\pi_{ik} + R_x(\phi_{ik}) \cdot R_y(\theta_{ik}) \cdot R_z(\psi_{ik}) \Delta\pi_{i\to j,k}}_{=:\ \hat{\pi}_{i\to j,k}} \\ \mathbf{r}_{jk} &= \underbrace{\mathbf{r}_{ik} + \Delta\mathbf{r}_{i\to j,k}}_{=:\ \hat{\mathbf{r}}_{i\to j,k}} \end{aligned} \tag{19}$$

To model nonrigid surfaces, our approach allows for violation of these link constraints. This is obtained by introducing the following Gaussian potentials for each link

$$\begin{aligned} h_{i\to j,k} &=; |2\pi\Theta|^{-\frac{1}{2}} \\ &\exp\left\{ -\tfrac{1}{2} \begin{pmatrix} \pi_{jk} - \hat{\pi}_{i\to j,k} \\ \mathbf{r}_{jk} - \hat{\mathbf{r}}_{i\to j,k} \end{pmatrix}^T \Theta^{-1} \begin{pmatrix} \pi_{jk} - \hat{\pi}_{i\to j,k} \\ \mathbf{r}_{jk} - \hat{\mathbf{r}}_{i\to j,k} \end{pmatrix} \right\} \end{aligned}$$

Here Θ defines the strength of the link (the resulting structure is a Markov random field [17]).

3.2 Target Function

The negative logarithm of these potentials, summed over all links, is given by the following function H (constant omitted):

$$H = \tfrac{1}{2} \sum_{i\to j,k} \begin{pmatrix} \pi_{jk} - \hat{\pi}_{i\to j,k} \\ \mathbf{r}_{jk} - \hat{\mathbf{r}}_{i\to j,k} \end{pmatrix}^T \Theta^{-1} \begin{pmatrix} \pi_{jk} - \hat{\pi}_{i\to j,k} \\ \mathbf{r}_{jk} - \hat{\mathbf{r}}_{i\to j,k} \end{pmatrix} \tag{20}$$

As in the scan registration problem, all these terms are nonlinear in the node coordinates π_{ik} and the orientations $\mathbf{r}_{ik}$. To obtain a closed-form solution for the resulting equation system, the link function is linearized via a Taylor series expansion:

$$\begin{pmatrix} \pi_{jk} \\ \mathbf{r}_{jk} \end{pmatrix} \approx \begin{pmatrix} \pi_{ik}^{[n]} \\ \mathbf{r}_{ik}^{[n]} \end{pmatrix} + K_{ik,ik}^{[n]} \begin{pmatrix} \pi_{ik} - \pi_{ik}^{[n]} \\ \mathbf{r}_{ik} - \mathbf{r}_{ik}^{[n]} \end{pmatrix} \tag{21}$$

where $\pi_{ik}^{[n]}$ and $\mathbf{r}_{ik}^{[n]}$ specify the coordinate system for the node ik in the n-th iteration of the optimization. The matrix $K_{ik,ik}^{[n]}$ is a Jacobian matrix of dimension six by six, which is obtained as the derivative of the functions (19).

By the same logic by which G can be approximated by a quadratic function in the scan registration problem, substituting our approximation back into the definition of H gives us a quadratic function in all variables π_{ik} and $\mathbf{r}_{ik}$. This function can be written in the form

$$H = \tfrac{1}{2} \left[\mathbf{B}^{[n]} \begin{pmatrix} \pi \\ \mathbf{r} \end{pmatrix} - \mathbf{f}^{[n]} \right]^T \tilde{\Theta}^{-1} \left[\mathbf{B}^{[n]} \begin{pmatrix} \pi \\ \mathbf{r} \end{pmatrix} - \mathbf{f}^{[n]} \right] \tag{22}$$

where π is the vector of all coordinate system origins, $\mathbf{r}$ the vector of all Euler angles, and $B^{[n]}$ is a matrix and $\mathbf{f}^{[n]}$ a vector. Calculating $B^{[n]}$ and $\mathbf{f}^{[n]}$ is involved but mathematically straightforward.

3.3 Optimization Procedure

Our new version of ICP now minimizes the combined target function $G + H$, which is again quadratic in all parameters ($\mathbf{x}$, π, and $\mathbf{r}$). By doing so, it simultaneously recovers the scan registration and the surface configuration of the object. The solutions for $\mathbf{x}$ and π are completely analogous to the one in (16):

$$\begin{pmatrix} \pi^{[n+1]} \\ \mathbf{r}^{[n+1]} \end{pmatrix} = (\mathbf{B}^{[n]T}\mathbf{B}^{[n]})^{-1}\, \mathbf{B}^{[n]T}\mathbf{f}^{[n]} \tag{23}$$

The global orientation is still optimized by a single global SVD as above. Our new augmented optimization leads to relative adjustments between measurement points, in which the links play the role of soft constraints.

This is illustrated in Figure 3c-f, for different values for Θ. As the various diagrams illustrate, scans are deformed to improve their match. The degree of the deformation depends on the value of Θ, which defines the rigidity of the surface. Figure 3c-f illustrates that our approach succeeds in locally rotating and even rescaling the model.

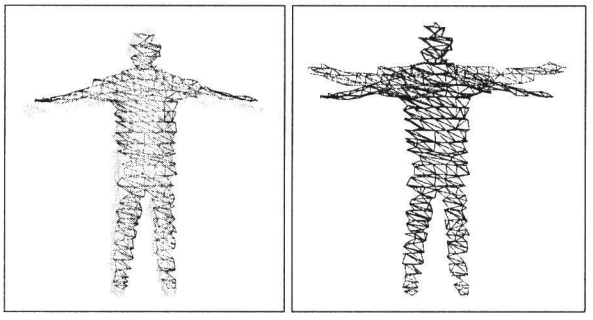

Figure 4: Example of a thinned graph superimposed to the original scan left) and before and after adjustment (right). Thinning is necessary to perform the optimization efficiently.

3.4 Efficient Variable Resolution Optimization

The main problem with the approach so far is its enormous complexity. The number of variables involved in the optimization is orders of magnitude larger than in scan registration. This is because the target function H is a function of all measurement points π and orientations $\mathbf{r}$, whereas G has only the vantage points $\mathbf{x}$ as its arguments. The matrix $\mathbf{B}^{[n]}$ in (23) is, thus, a (sparse) matrix with many thousand dimensions.

To tackle such problems efficiently the optimization is reduced to a sequence of nested optimization problems. In a first step, scans are analyzed for connected components (regions without large disparities); links exist only between connected components in each scan; hence H factors naturally into different subproblems for different connected components. Next, the resulting scan patches are thinned. Thinning proceeds by identifying a small number of representative landmark measurements that are approximately equally spaced. This computation is performed by stipulating a grid over the scan (in workspace coordinates), and selecting measurements closest to the center points of each grid cell. An all point shortest path search then associates remaining measurements with landmark measurements. The optimization is first performed for the thinned scan; after the landmark scans are localized (and the corresponding coordinate transformation are computed), the remaining measurements are optimized locally, in groups corresponding to individual landmark measurements. Smoothness is attained by using multiple landmark measurements as boundary conditions in this optimization. Figure 4 shows an examples of a thinned graph, for which the optimization can be carried out in seconds.

4 Setup and Experimental Results

Our approach was implemented using a mobile robot, in an attempt to acquire 3D models of non-stationary objects. In a technique adopted from [11], we first acquired a map of the environment (see Figure 5). Non-stationary objects were detected through differencing of scans, using the robot's localization routines to get a rough estimate of this pose. Figure 2b illustrates the segmentation process. Red scans are retained while the black scans are assumed to correspond to the background and are henceforth discarded.

	start	loop #1	loop #2	loop #3
moving arms				
scan 1	2.3266	0.8993	0.7986	-
scan 2	2.5320	0.8704	0.8001	-
stretched body				
scan 1	1.9369	1.2915	1.2008	1.1975
scan 2	2.5087	1.2964	1.2220	1.2120

Table 1: Average distance to the closest points to the matched model *after* scan registration. The decrease of this distance measures the improvement of the model through local surface deformations.

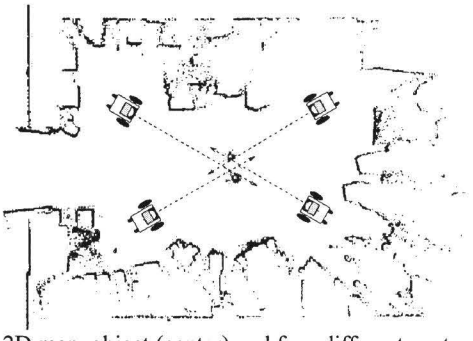

Figure 5: 2D map, object (center) and four different vantage points.

Figure 6 illustrates one iteration of the algorithm in all essential steps, using data acquired by the robot. Results for matching three scans with different postures are shown in Figure 7(a-b). While the standard registration procedure leads to a model with six arms, our approach correctly deforms the scan to arrive at an improved model, with two arms. A similar result is shown in Figure 7(c-g), which shows three raw scans on the left, followed by the result of (rigid) scan registration and the result of our approach. Another example is the chair in Figure 8(a-d) scanned in different heights. The standard registration will lead to multiple feet, our approach correctly aligns them. Table 1 shows the cumulative distance between points in the nearest neighbor calculation. The value marked as "start" is the result of an initial registration phase, reflecting the remaining distances under the rigid body assumption. All other columns correspond to further iterations of our algorithm as it adjusts the shape of the scans. This result illustrates numerically the integrity of the result is indeed improved by iterate the process.

5 Discussion

This paper proposed a technique for simultaneous scan registration and scan deformation for modeling nonrigid objects. The deformation was made possible through the definition of (soft) links between neighboring scan points, whose configuration was calculated during registration. To tackle the resulting optimization problem efficiently, we described a hierarchical optimization techniques that operated on thinned graphs. Experimental results obtained using a mobile robot illustrated the viability of this approach.

There are many problems in object modeling that this paper does not address, but whose inclusion shall be the subject of future research. For example, the present segmentation approach is somewhat simplistic: It will fail if more than one non-stationary object appears in the scene. The approach requires deformations to be small, and the target object may not move very far during acquisition. Objects are not subsegmented. This will cause difficulties when components of objects are adjacent to different other components, or missing entirely, which can happen if components are combined and

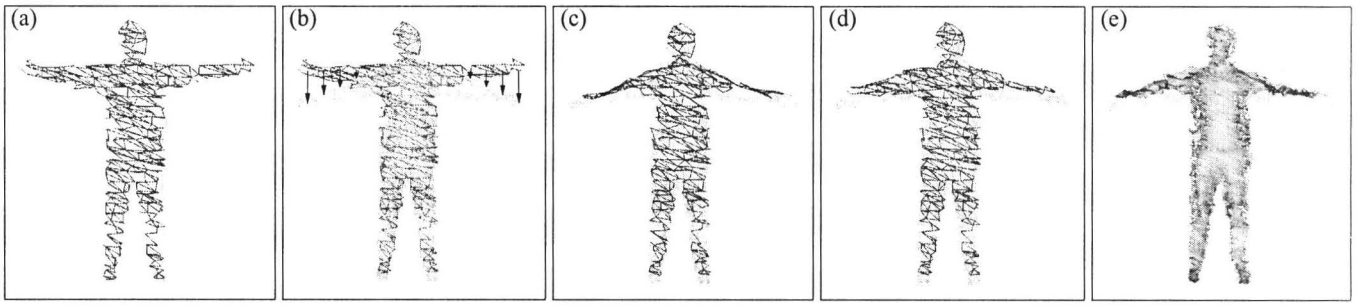

Figure 6: (a) Thinning, (b) nearest neighbor search, (c) optimization (first iteration), (d) optimization (second iteration), (e) optimization of remaining nodes.

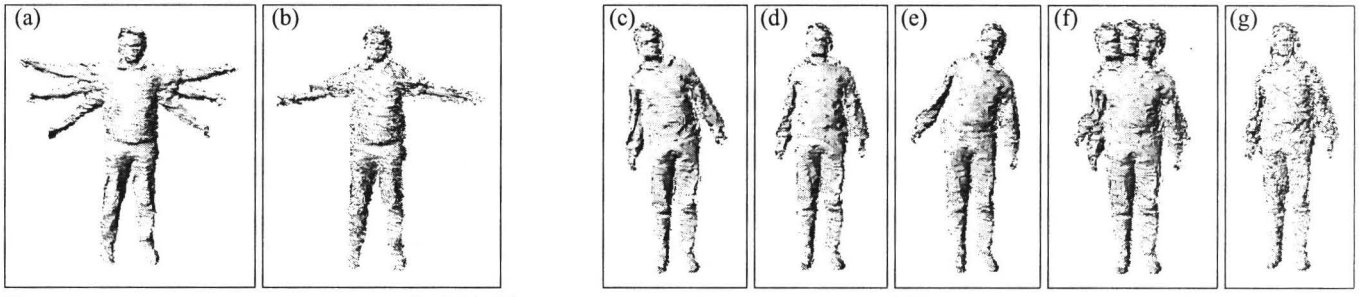

Figure 7: (a) Optimal scan registration with rigid object assumption, (b) corresponding result for our approach. (c-e) three scans, (f) results of optimal scan registration and (g) our approach.

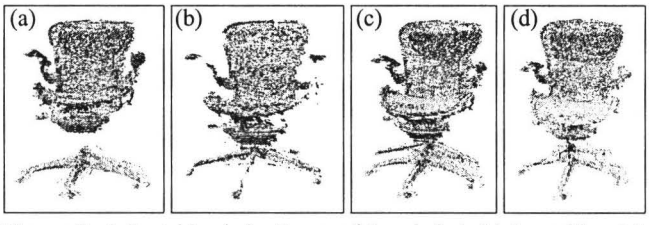

Figure 8: Adjustable chair: Scans of the chair in high position (a), scan from the chair in low position (b), the resulting model of the chair with multiple feet (c), and the resulting model of the chair with the transformed scan (b).

cannot be separated to find the corresponding part. A final direction of future research involves the integration into advanced techniques for between-object data association when modeling multiple objects in non-stationary environments [1].

References

[1] D. Anguelov, R. Biswas, D. Koller, B. Limketkai, S. Sanner, and S. Thrun. Learning hierarchical object maps of non-stationary environments with mobile robots. *UAI-02*.

[2] P. Besl and N. McKay. A method for registration of 3d shapes. *PAMI* 14(2), 1992.

[3] M. Bakircioglu, S. C. Joshi, M. I. Miller. Landmark matching on brain surfaces via large deformation diffeomorphisms on the sphere. *SPIE Medical Imaging*, 1999.

[4] L. Breiman, J. H. Friedman, R. A. Ohlsen, and C. J. Stone. Classification and regression trees. 1984.

[5] H. Chui, A. Rangarajan. A new point matching algorithm for non-rigid registration. *CVPR*, 2000.

[6] M. K. Chung, K. J. Worsley, S. Robbins, P. Paus, J. Taylor, J. N. Giedd, J. L. Rapoport, A. C. Evans. Deformation-Based Surface Morphometry with an Application to Gray Matter Deformation. *NeuroImage*, 18, 2003.

[7] J. Feldmar, N. Ayache, and F. Betting. 3D-2D projective registration of free-form curves and surfaces. *CVIU*, 65(3), 1997.

[8] J.-S. Gutmann and K. Konolige. Incremental mapping of large cyclic environments. *CIRA-00*, 2000.

[9] E. G. Huot, H. M. Yahia, I. Cohen, I. Herlin. Surface Matching with Large Deformations and Arbitrary Topology: A Geodesic Distance Evolution Scheme on a 3-Manifold. *ECCV*, 2000.

[10] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4, 1997.

[11] M. Montemerlo, N. Roy, and S. Thrun. Carnegie mellon robot navigation toolkit, 2002.

[12] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. *3DIM-01*.

[13] G Strang. *Introduction to Linear Algebra*. Wellesley-Cambrigde Press, 1998.

[14] S. Teller, M. Antone, Z. Bodnar, M. Bosse, S. Coorg, M. Jethwa, and N. Master. Calibrated, registered images of an extended urban area. *CVPR-01*.

[15] S. Thrun. A probabilistic online mapping algorithm for teams of mobile robots. *IJRR*, 20(5), 2001.

[16] Y. Wang, L. H. Staib. Shape-Based 3D Surface Correspondence Using Geodesics and Local Geometry. *CVPR*, 2000.

[17] G. Winkler. *Image Analysis, Random Fields, and Dynamic Monte Carlo Methods*. Springer, 1995.