Non-Rigid Range-Scan Alignment Using Thin-Plate Splines

Benedict J. Brown Princeton University bjbrown@cs.princeton.edu

Abstract

We present a non-rigid alignment algorithm for aligning high-resolution range data in the presence of lowfrequency deformations, such as those caused by scanner calibration error. Traditional iterative closest points (ICP) algorithms, which rely on rigid-body alignment, fail in these cases because the error appears as a non-rigid warp in the data. Our algorithm combines the robustness and efficiency of ICP with the expressiveness of thin-plate splines to align high-resolution scanned data accurately, such as scans from the Digital Michelangelo Project [14]. This application is distinguished from previous uses of the thin-plate spline by the fact that the resolution and size of warping are several orders of magnitude smaller than the extent of the mesh, thus requiring especially precise feature correspondence.

1. Introduction

Range scanners are useful tools for acquiring three dimensional models of real objects. Because a scanner can see only a single viewpoint at a time, multiple scans must be assembled to acquire such a model. For small objects, it is feasible to perform only calibrated motions (using a turntable, for example) to obtain the necessary scans. But for large objects, such as Michelangelo's statues, this is simply not possible. Some method must therefore be used to recover the alignment of the different scans. Even small errors in scanner calibration lead to further complication, as they results in a low-frequency, non-rigid warp in the acquired data.

We are interested in aligning range scans from the Digital Michelangelo Project [14]. These scans have approximately quarter-millimeter precision, for very large statues—5 meters high in the case of the David—and were acquired using a laser range scanner mounted on a custom gantry. In the case of the David, the gantry had to extend up to 7.5m high, since the David is mounted on a pedestal, and also had to be highly reconfigurable so that the scanner would be able to see all parts of the statue without Szymon Rusinkiewicz Princeton University smr@cs.princeton.edu

ever touching it. These extreme size and configurability requirements make accurate calibration of the gantry difficult if not impossible to achieve, and the range scans consequently exhibit enough low-frequency warping to prevent even pairwise rigid-body alignment in some cases. Even if the scanner calibration were improved, rescanning these statues is a practical impossibility. Since these engineering challenges will not go away, we expect that similar calibration problems will arise in other very large-scale scanning projects, as well as in very low-cost ones. An algorithm that can recover and correct for these warps is consequently necessary.



Figure 1. In this closeup of David's hair, the original source (green) and target (blue) meshes have been aligned using ICP, but still contain substantial error.

After alignment, range scans are merged to produce a full model. This can be accomplished using a variety of algorithms, including VRIP [8] and ball pivoting [2]. These algorithms merge the two meshes and retriangulate them, and must smooth out noise in the process. The resolution of the output model which can be produced is therefore dependent not only on the range scan resolution, but also on the alignment quality—range scans must be aligned to within the desired precision of the output model. Thus, on Michelangelo's David, for instance, even alignment errors on the order of .5mm can prevent accurate merging of the range scans.

We present a non-rigid alignment algorithm that can handle the non-rigid warps present in range scans from the Digital Michelangelo Project, while retaining the robustness offered by a rigid-body transformation with few degrees of freedom. We use a thin-plate spline to represent the warp, based on feature correspondences computed using a hierarchical iterative closest points (ICP) method. We have found that this reduces overall alignment error. Since the thin-plate spline parameters are computed by solving a linear system of equations, the algorithm is efficient, and is dominated by the ICP computations.

In the following sections, we review existing work in rigid and non-rigid alignment, and discuss how it relates to the problem of range-scan alignment in the presence of non-rigid warps. We then describe the thin-plate spline and show how to compute it. Next we describe our feature correspondence algorithm. Finally, we present and discuss the results of our system, and areas of future work.

2. Previous Work

Most range-scan alignment is currently done using variants of ICP [4] [6], which assumes a rigid body transformation (i.e. no measurement error). Although the lack of degrees of freedom allows ICP to effectively handle high frequency noise, low-frequency warps result in poor alignments, with high measured error. The Digital Michelangelo Project and the Florentine Pietà Project [3] have both used variants of ICP for alignment. Several improvements to ICP have been developed in the context of the Digital Michelangelo Project and accompanying Forma Urbis Romae Project, in order to improve the mesh alignment. For example, [10] and [12] introduce improved techniques for sampling the range scans and evaluating the stability of these point sets for ICP alignment. These techniques are of general applicability in ICP, including the hierarchical ICP alignment phase of our algorithm. However, they do not address the situation in which a rigid-body transformation cannot represent the alignment transformation adequately.

The Florentine Pietà Project also incorporated imagebased alignment and conformance smoothing of overlapping scans along scanner lines-of-sight to improve the registration further. Image-based alignment is most useful if the image resolution is higher than the geometric resolution, as in the case with the Pietà project's structured light scanning system. We expect that conformance smoothing, while of general applicability in reducing measurement noise, will not help in the case of calibration error, since the error will not be along the lines of sight.

Ikemoto, *et al.* [12] have also looked at the alignment problem in the presence of calibration error. They use a hierarchical ICP algorithm to align range scans from both the Digital Michelangelo Project, and especially the accompanying Forma Urbis Romae project. They therefore directly address the same alignment problems as ours. Instead of using a non-rigid warp, the pair of range scans to be aligned are diced into a number of overlapping pieces, which are then globally aligned to each other. When the mesh is diced, the overlapping area must be large enough (and contain enough high frequency features) to make the estimate of pairwise alignment well-defined. On the other hand, it must not be too large, or any difference at all will result in high alignment error. The overall size issue is addressed by a user-controlled parameter. The stability of the ICP (*i.e.* the presence of features) is verified by computing the eigenvalues of the covariance matrix between the pair of scans. If the ratio of the largest and smallest eigenvalues is close to one, the covariance matrix is well-conditioned, and ICP is stable; if not, the matrix is ill-conditioned, and ICP will be unstable. (For instance, two flat planes can slide in two directions relative to each other without affecting alignment error.)

The primary drawback of this method is its speed. Because a global alignment phase is required, it takes minutes or hours to align mesh pairs, where our method takes only seconds. Since David's head alone consists of 98 different range scans, this is a substantial issue. Moreover, since the range scan is diced and each piece is aligned separately, the transformation is neither smooth nor even continuous, and relies on the merging stage to compensate. Since there are now many more overlapping pieces that must be merged, the precision constraints are even higher, and the process is slower.

Hänel *et al.* [11] present an extension to ICP that allows deformable objects to be aligned. It does this by attaching links to individual point pairs, which allow translation and rotation, then use a skeletonization of the model to hierarchically optimize an energy functional combining both rigid-body transformation and link deformations. While this approach is suited to range scan data, the deformations it is designed to handle are large-scale model deformations (such as arms moving), not subtle warps.

Allen *et al.* [1] use an affine transformation at each vertex of the source mesh to allow non-rigid registrations of full-body scans to a high-resolution template. There are therefore twelve degrees of freedom per vertex. In order to guarantee smoothness, a term is added to the energy functional which penalizes differences between the transformations applied to neighboring vertices. Initial registration is performed using markers placed on the bodies during scanning to help prevent convergence to a local minimum. Because we know that our warps will be slowly varying, and our range scans often have millions (rather than hundreds of thousands) of vertices, it makes sense to represent our warp using a more computationally efficient and compact representation which nevertheless guarantees a smooth transformation.

The thin-plate spline which we use to represent our warp has been used extensively in medical imaging applications, where its use was first proposed in [5]. Chui *et al.* [7] use a softassign/deterministic annealing framework to iteratively compute point correspondences and align both medical and non-medical data in 2-D and 3-D using thin-plate splines. Rather than assigning features based on closest points as in ICP, each pair of points is assigned a probability of corresponding based on a Gaussian function of their distance from each other (softassign). A thin-plate spline is computed based on these weighted probabilities, and a new correspondence is calculated with a narrower Gaussian (deterministic annealing), until the system converges to exact correspondences. This framework is well-suited to situations where two point sets can be almost completely paired, but not to the partial overlap case we address. Rohr et al. [17] [16] extend the thin-plate spline computation to take covariance matrices and vertex normals into account, and are discussed more fully below. These constraints can substantially improve the quality of a large warp, but have very little practical effect for us. The reason is that our alignment constraints are so stringent that interpolation quality effectively dominates any other consideration.

3. Thin-plate splines

Thin-plate splines are a class of non-rigid spline mapping functions with several desirable properties for our application. They are globally smooth, easily computable, separable into affine and non-affine components, and contain the least possible non-affine warping component to achieve the mapping. By the last statement, we mean that the sum of squares of all second order partial derivatives is minimized. So, if $f : \mathbb{R}^n \to \mathbb{R}$ is an *n*-dimensional thin-plate spline, the *bending energy*,

$$J = \int \left(\sum_{i,j} f_{x_i x_j}^2\right) \mathrm{d}x_1 \dots \mathrm{d}x_n \tag{1}$$

is minimal. (This definition is often extended to minimize the sum squares of all d'th order derivatives, subject to the constraint that 2d - n > 0). Note that since affine transformations are linear, they contribute no error under this metric.

Duchon [9] proves that, for two corresponding point sets $X = \{x_1, ..., x_m\}$ and $Y = \{y_1, ..., y_m\}$, there is a unique function f such that $f(x_i) = y_i$ whose bending energy is minimal. Furthermore, this function takes the form xd + Kw, where x is a point written in homogeneous coordinates, d is an affine transformation, w is a fixed mdimensional column vector of non-affine warping parameters constrained to $X^tw = 0$, and K is an m-dimensional row vector where K_i is the Green's function $U(|x - x_i|)$. In our case (minimizing second order partials in \mathbb{R}^3), this is simply $|x - x_i|$; the constant factor is implicitly folded into w.

Thin-plate splines need not be interpolating. Instead, they can minimize the energy functional [9] [18]

$$E_{\lambda} = \frac{1}{m} \sum |y_i - f(x_i)| + \lambda J$$
(2)

The spline will not be interpolating in this case, but for any fixed λ , there will still be a unique minimum, of the form described above. Rohr *et al.* [17] [16] give extensions to allow a confidence σ_i to be attached to each x_i , or to assign a covariance matrix to each x_i , or even to constrain the normals at each point (although the latter changes the form of the spline slightly). In all cases, the energy functional changes, but the overall manner of computation does not. Chui *et al.* [7] add a constraint on the affine transformation, which can help prevent flips in early stages of feature correspondence. This is not a problem in our case, and complicates the parameter computation somewhat.

For the interpolating case, the thin-plate spline specification provides a linear system of equations, which [5] solves directly:

$$\left(\frac{w}{d}\right) = \left(\frac{K \mid X}{X^t \mid 0}\right)^{-1} \left(\frac{Y}{0}\right) \tag{3}$$

where $K_{ij} = U(|x_i - x_j|)$. For the approximating case [18, eqs. 2.4.23 and 2.4.24] derives a similar system of equations by rewriting equation 2 in matrix form, performing a QR decomposition on *X*, and simplifying:

$$Xd + (K + m\lambda I)w = Y$$

$$X^{t}w = 0$$
 (4)

An analogous derivation yields a similar equation when confidence values are used: *I* is replaced by diag(σ_i) [17]. When covariance matrices are used, *I* is replaced by diag(Σ_i) where Σ_i is the 4 × 4 covariance matrix associated with x_i . In this case K_{ij} and X_{ij} are replaced by $K_{ij}I_4$ and $X_{ij}I_4$, expanding these matrices by a factor of 4 in each direction. *w*, *d*, and *Y* are rewritten as column vectors, reading across the rows and down the columns [16]. The derivation is slightly more complex in this case, but follows the same pattern.

Because we need alignment error of less than .5mm over a mesh which spans roughly a half meter, the spline must be heavily weighted toward interpolation, and any effects of covariance matrices or more exotic restrictions are minimal. For this reason, we use no covariance constraints, and rely on very accurate correspondences to produce a good alignment. We set λ to a very small value because Equation 4 becomes unstable when $\lambda = 0$ and many point pairs are used.

The alignment algorithm therefore works as follows. We randomly select a subset of vertices on one mesh, and find the corresponding vertices on the other mesh, using the technique described below. Correspondence pairs in which either vertex lies on the boundary of its mesh, or whose maximum distance exceeds a threshold are rejected. This guarantees that the sample vertices are all in the overlapping portion of the two meshes. A thin-plate spline is then computed to map the sample source vertices to their corresponding targets. This spline defines a global warping of space, and is therefore used to warp the entire source mesh onto the target mesh. Since the warp is smoothly extrapolated from the overlapping area to all of space, no discontinuities are introduced. Furthermore, the minimal warping property of the thin-plate spline guarantees that the extrapolation will be reasonable.

4. Feature Correspondence

In order to align two meshes, and measure the precision of the alignment, a correspondence must be found between them. This correspondence may be computed separately from the alignment [16], or in alternation with it [4] [6] [7]. In the latter case, a rough feature correspondence gives rise to a rough alignment, which in turn provides a better correspondence, and so on. The iterated closest point algorithm [4] [6], which is predominantly used for mesh alignment, is an example of this. A subset of points on mesh X is selected, and the "closest" points on mesh Y are chosen as the corresponding ones. Chen et al. [6] define closest using a point-to-plane metric ($|\vec{n}_{y}.(x-y)|$), allowing flat areas of the mesh to slide relative to each other, while constraining areas of high curvature. The best rigid-body transformation is then found to map the points on X to the points on Y, and the process is repeated. Assuming the initial alignment is fairly close, this will converge to a good rigid-body alignment. The initial alignment is usually performed either manually [14], or using a technique such as spin images [13].

ICP cannot be directly extended to use a thin-plate spline mapping rather than a rigid-body mapping because it can interpolate the point set regardless of the correspondence precision. As a result, an incorrect correspondence simply warps *X* incorrectly, causing equally bad—or worse—correspondences on the next iteration. Even imposing a smoothness constraint does not help in our case, because the warp is too small compared to the scale of the data points.

Chui *et al.* [7] do successfully use an iterative correspondence algorithm with the thin-plate spline, but makes a key assumption which does not hold in our case. Specifically, it assumes that there is an almost complete matching between the two data sets (although some outliers may exist). To be in that condition, we would need to use *every* point in the overlapping area of the the two meshes—often on the order of a million points for the David. That is simply not feasible. One possible workaround to this is to select closest points on the target mesh at each iteration, similar to ICP. But since the thin-plate spline will always interpolate its points, the poor initial correspondences will incorrectly warp the source mesh, leading to an incorrect local minimum. Furthermore, we have found in practice that at later iterations of the deterministic annealing, most points get rejected as outliers even in this scenario, leading to an unstable system.

To solve these problems, we use a hierarchical ICP approach to find good feature correspondences, then compute a single thin-plate spline to smoothly warp the mesh. We start with a global ICP phase, then divide the source mesh through the middle of the longest axis of its bounding box, and realign each half separately. We repeat this process a fixed number of times, or until all pieces are unstable under ICP [10]. (The latter case occurs when too few correspondences can be found, such as when the given piece of the source mesh does not overlap the target mesh at all.) A fixed error threshold could be used in place of a fixed number of iterations. In practice, we are trying to recover a very low-frequency warp, so large sections of the meshes can be aligned well with ICP. As a result, only a small number of decompositions is necessary to achieve good piecewise alignment.



Figure 2. Hierarchical alignment [12] was used to dice a source mesh, and align the resulting pieces to a target mesh (not shown). Closeups of the circled areas show that the different pieces can overlap or drift apart substantially from each other. These discontinuities are smoothed out by the thin-plate spline computed by our algorithm.

The ICP stage alone provides good feature correspondences, at the cost of substantial discontinuities in the source mesh. These are eliminated in the thin-plate spline mapping, since it is a globally smooth mapping of the entire mesh. This is in contrast to [12], which requires an ICPstable overlap between each piece for global alignment. In that case, the mapping is not smooth, and the overlap size determines a trade-off between smoothness and warping flexibility.

5. Results

We have run our thin-plate spline alignment algorithm on all overlapping pairs of range scans from David's head. The data set we used contains 98 range scans, with data points spaced at approximately .5mm. There were 2996 overlapping pairs of range scans in total. The source mesh was diced 15 times, and 1000 point pairs were used for alignment. We computed the root-mean-square error of the initial ICP alignment, and the thin-plate spline alignment using 10,000 randomly selected point pairs in the overlapping area. Under ICP alignment, the mean error was 0.299mm, with a standard deviation of 0.354mm. For the thin-plate spline alignment, the mean error was reduced by a factor of 3 to 0.101mm, and the standard deviation was 0.132mm.

In order to visualize the alignment quality and warp, we show the alignment of two range scans of Michelangelo's David which cannot be adequately aligned using ICP alone. The source mesh is face4_e, which we warp to the target mesh, right_shoulder_chest_a. The second-highest resolution is used for each mesh: 906.791 vertices for face4_e and 780.469 vertices for right_shoulder_chest_a. We diced the mesh 5 times, and used 250 control points for the thin-plate spline. On a Pentium IV 2GHz with 1GB RAM, the warping took 49.45 seconds to compute. Of this, 24.55 seconds were spent in hierarchical ICP, 4.75 seconds in computing the thin-plate spline, and 20.15 seconds in performing the warp. The warping time is directly proportional to the number of control points, and is only 8.4 seconds with 100 control points (at only a nominal loss in quality).

Figures 3(a) and 3(b) show the target and source meshes. Figure 3(c) shows the alignment quality of the ICP and non-rigid alignments. The green channel of each vertex shows the distance to the nearest vertex on right_shoulder_chest_a. Bright green means the two vertices coincide; a green value of zero indicates the nearest vertex on right_shoulder_chest_a is more than 5mm away. Similarly, the red channel encodes the quality of the non-rigid alignment. Therefore, yellow areas indicate where both alignments are good, green areas have worse non-rigid than ICP alignment, and red areas have better non-rigid alignment. The large red areas in the hair and under the chin show that the non-rigid alignment is substantially better than the ICP alignment.

Figures 3(d), 3(e), and 3(f) show how the non-rigid warp is affecting face4_e. Black vertices did not move at all, white vertices moved 5mm (note that the scale is non-linear due to gamma correction). Figure 3(d) shows only the affine component of the warp, figure 3(e) shows only the non-affine component, and figure 3(f) shows the full warp. The affine component is very small—a slight

scaling—which shows that the initial ICP alignment was nearly optimal for a rigid-body alignment. The non-rigid component is largest in areas near the eye and in the hair, where the original alignment was worst. It is also fairly substantial in areas of face4_e that do not overlap right_shoulder_chest_a. This may seem counterintuitive, but we place no constraints on the warp here, and allowing the mesh to move in these locations produces a globally smoother warp.

Finally, figure 3(g) shows the two meshes merged using VRIP after non-rigid alignment. Figures 3(h) and 3(i) show a closeup near the ear of VRIP merges after ICP and after non-rigid alignment. Note in particular the improvements at the top of the ear and the large lock of hair to the left of the ear.

6. Discussion and Future Work

Our non-rigid alignment algorithm can accurately and efficiently handle low-frequency warps present in scanned range data. By relying in ICP for feature correspondences, it is both robust an efficient, while the thin-plate spline allows it to flexibly and compactly represent the warp. The principal drawback we have noticed is that very high resolution data is required for accurate warps. This is because closest point computations are made only to vertices in the target mesh; if those vertices are spaced too far apart, the feature correspondence will be skewed. In our experience, the resulting warp still provides a good alignment, but is less smooth and warps non-overlapping areas of the source mesh more than necessary.

As future work, we would like to improve the selection of control points for the warp. Since the time to warp a mesh is directly proportional to the number of control points used, it is desirable to have as few as possible, as well-placed as possible. We would also like to improve our memory efficiency, and explore out-of-core implementations so that we can work with larger data sets.

Another area of future work is to incorporate any knowledge we may have of calibration and measurement error into the thin-plate spline computation. Using this information, to the extent it is available, should improve the accuracy of the resulting model, while maintaining alignment consistency.

Finally, we can currently align meshes only pairwise. We plan to extend our system in the future to support global registration in an efficient way. An obvious possibility is to incrementally add range scans to a global mesh. Unlike rigid-body ICP, we do not expect that alignment error would increase as additional scans are added. However, we do expect that the amount of warp will progressively increase, so the problem is one of minimizing and distributing total warp rather than total error. Combined with the







(a)right_shoulder_chest_a

(b) Unwarped face4_e

(c) Alignment quality.





(d) Affine portion of the warp; brighter is more warp

(e) Non-affine portion of the warp

(f) The full warp



(g) The merged mesh



(h) Ear and hair merged after ICP alignment



(i) Ear and hair merged after non-rigid alignment

many additional degrees of freedom in a thin-plate spline vs. a rigid-body transformation, this makes direct application of existing global registration techniques infeasible.

Acknowledgements

The authors wish the thank Leslie Ikemoto and Natasha Gelfand for providing their Hierarchical ICP implementation. Paul Calamia provided helpful suggestions on the algorithm and paper. We would also like to thank the anonymous reviewers for their many useful comments.

References

- B. Allen, B. Curless, and Z. Popović. The space of human body shapes: reconstruction and parameterization from range scans. *ACM Trans. Graph.*, 22(3):587–594, 2003.
- [2] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin. The ball-pivoting algorithm for surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 5(4):349–359, 1999.
- [3] F. Bernardini, H. Rushmeier, I. M. Martin, J. Mittleman, and G. Taubin. Building a digital model of Michelangelo's Florentine Pietà. *Computer Graphics and Applications, IEEE*, 22(1):59–67, 2002.
- [4] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(2):239– 256, 1992.
- [5] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.
- [6] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image Vision Comput.*, 10(3):145– 155, 1992.
- [7] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, February-March 2003.
- [8] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the* 23rd Annual Conference on Computer Graphics and Interactive Techniques, pages 303–312. ACM Press, 1996.
- [9] J. Duchon. Splines minimizing rotation-invariant seminorms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables*, pages 85–100, Berlin, 1977. Springer-Verlag.
- [10] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy. Geometrically stable sampling for the ICP algorithm. In Fourth International Conference on 3D Digital Imaging and Modeling (3DIM 2003), October 2003.
- [11] D. Hähnel, S. Thrun, and W. Burgard. An extension of the ICP algorithm for modeling nonrigid objects with mobile robots. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI)*, Acapulco, Mexico, 2003. IJCAI.
- [12] L. Ikemoto, N. Gelfand, and M. Levoy. A hierarchical

method for aligning warped meshes. In *Fourth International Conference on 3D Digital Imaging and Modeling (3DIM* 2003), October 2003.

- [13] 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.
- 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: 3-D scanning of large statues. In *Proc. SIGGRAPH*, 2000.
- [15] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Auton. Robots*, 4(4):333–349, 1997.
- [16] K. Rohr, M. Fornefett, and H.S. Stiehl. Spline-based elastic image registration: integration of landmark errors and orientation attributes. *Computer Vision and Image Understanding*, 90:153, May 2003.
- [17] K. Rohr, H.S. Stiehl, R. Sprengel, W. Beil, T.M. Buzug, J. Weese, and M.H. Kuhn. Point-based elastic registration of medical image data using approximating thin-plate splines. In VBC, pages 297–306, 1996.
- [18] G. Wahba. Spline Models for Observational Data, chapter 2.4. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1990.