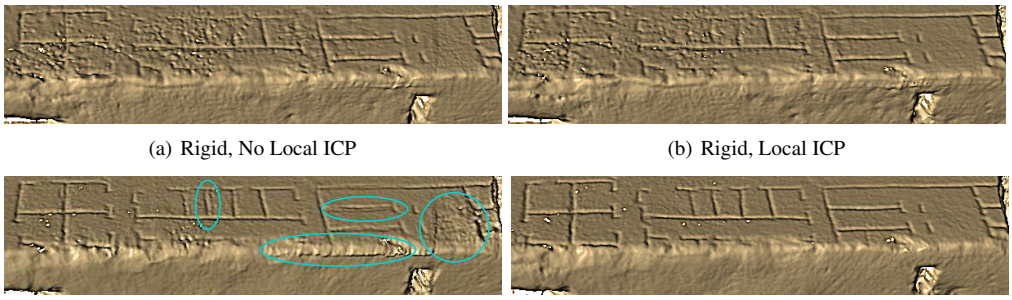
Global Non-Rigid Alignment of 3-D Scans

Benedict J. Brown*
Princeton University

Szymon Rusinkiewicz†
Princeton University



(a) Rigid, No Local ICP (b) Rigid, Local ICP

(c) Non-Rigid, No Local ICP (d) Non-Rigid, Local ICP

Figure 1: Using non-rigid global alignment techniques, we are able to preserve more surface detail with fewer merging artifacts, as seen in closeups of the upper-left edge of this *Forma Urbis Romae* fragment. *(c)* Using non-rigid alignment (section 3.3) improves alignment enough to eliminate most surface noise. *(d)* Incorporating locally weighted ICP-based feature correspondences (section 3.2) further improves detail alignment, notably eliminating double edges.

Abstract

A key challenge in reconstructing high-quality 3D scans is registering data from different viewpoints. Existing global (multiview) alignment algorithms are restricted to rigid-body transformations, and cannot adequately handle non-rigid warps frequently present in real-world datasets. Moreover, algorithms that can compensate for such warps between pairs of scans do not easily generalize to the multiview case. We present an algorithm for obtaining a globally optimal alignment of multiple overlapping datasets in the presence of low-frequency non-rigid deformations, such as those caused by device nonlinearities or calibration error. The process first obtains sparse correspondences between views using a locally weighted, stability-guaranteeing variant of iterative closest points (ICP). Global positions for feature points are found using a relaxation method, and the scans are warped to their final positions using thin-plate splines. Our framework efficiently handles large datasets — thousands of scans comprising hundreds of millions of samples — for both rigid and non-rigid alignment, with the non-rigid case requiring little overhead beyond rigid-body alignment. We demonstrate that, relative to rigid-body registration, it improves the quality of alignment and better preserves detail in 3D datasets from a variety of scanners exhibiting non-rigid distortion.

1 Introduction

Range scanning is a common method for acquiring three-dimensional models of real objects. Because a range scanner sees only a single viewpoint at a time, a complete model requires merging several scans. With small objects, it is easy to obtain calibrated motion of the scanner or object — by placing the object on a motorized turntable, for example — so that the relative alignments between range scans are known. However, this is not possible for large objects such as Michelangelo's statues, nor when both the top and bottom of an object need to be scanned as well as the sides. In these cases, the range scans must be aligned after acquisition. Because they must be merged and retriangulated to obtain a final mesh, even small misalignment leads to smoothing of high-frequency details in the merged mesh. Larger misalignments may result in corrupt reconstructions.

Although existing global (multiview) registration techniques assume that the scans can be aligned using only rigid-body transformations, there are several practical situations in which this assumption is invalid. In the case of data acquired by the Digital Michelangelo Project [Levoy et al. 2000], the desired precision of approximately 0.25 mm combined with viewing volumes over 5 m made perfect scanner calibration unattainable, leading to warping on the order of several millimeters. Even much smaller scanners, which acquire similar-resolution geometry in a much smaller working volume, exhibit warping when their resolution is pushed to the maximum (Figure 2). Over time, we believe that calibration will get harder, not easier, because sensor resolution is increasing faster than lens and motor precision.

We propose a method for simultaneously computing non-rigid warps to a large collection of 3D meshes, in order to minimize scan-to-scan distances in *all* overlapping regions. The alignment minimizes the warp applied to each scan, and distributes error evenly across all meshes. By correctly aligning surface details, the method ensures that low-frequency errors do not corrupt high frequencies of the reconstruction: the merged models exhibit less smoothing and fewer artifacts than their rigidly-aligned counterparts (Figure 1).

Our registration goals are high alignment accuracy and efficiency on large datasets. Specifically, we preserve the overall shape of the scans, which is generally correct, while improving the alignment of surface detail so that the final model is sharper. Since the warp is low frequency, we use a relatively small number of control points, which improves performance and helps prevent high-frequency distortions. We also avoid performing matrix operations such as singular value decomposition on the pairwise constraint matrix so that we can scale to arbitrarily large data sets. Our main contributions are the following:

1. *Feature correspondences*: We introduce a novel variant of iter-

*e-mail: bjbrown@cs.princeton.edu
†e-mail: smr@cs.princeton.edu

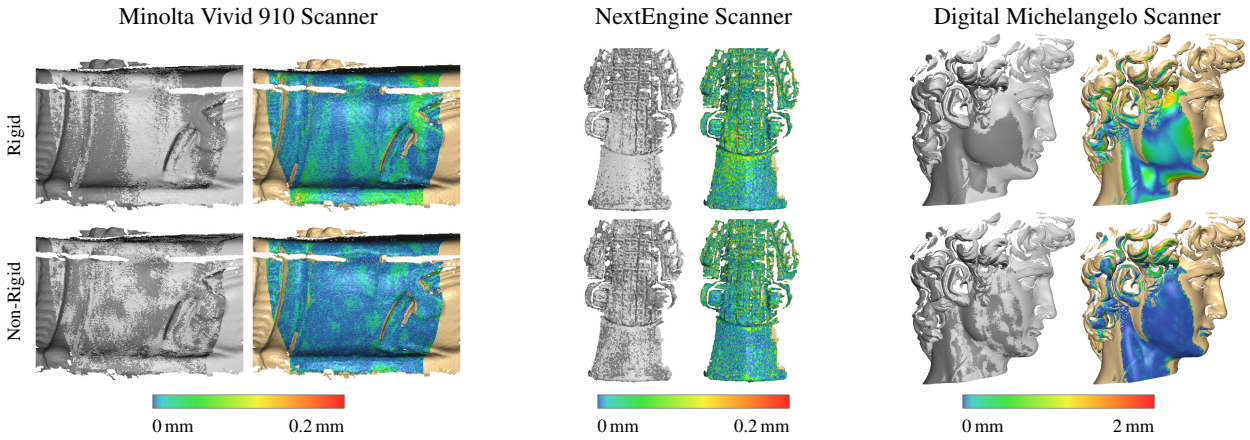


Figure 2: Rigid vs. non-rigid alignments of range scans from different 3-D scanners. The gray figures show two range scans (light and dark) superimposed, while the color figures show the alignment error between the two scans. Low-frequency warp in the initial range scans leads to biased alignment error, which in turn causes one scan to be in front of the other over large contiguous regions. Since the non-rigid alignments compensate for this warp, bias is removed, and neither scan appears consistently in front of the other. Non-rigid registration also reduces alignment error across the scan, as shown in the color figures.

ative closest points (ICP), locally weighted ICP, for efficiently finding correspondences between deformed ranged scans.

2. *Global optimization*: Our global optimization framework is highly parallelizable and memory efficient, making it practical for arbitrarily large datasets. It is agnostic to the final transformation, so that it is suitable for rigid as well as non-rigid alignment.
3. *Thin-plate spline warp*: We show how to use non-rigid thin-plate spline warps to consistently warp many scans simultaneously. Their use in range scan alignment has previously been limited to pairs of range scans.

2 Previous Work

Pairwise Rigid-Body Alignment: The most common algorithm for pairwise range-scan alignment is Iterative Closest Points (ICP), originally introduced by Besl and McKay [1992] and Chen and Medioni [1992]. ICP begins with an initial estimate of the alignment, refines it by repeatedly selecting points on one or both models, finding closest points in the other mesh, and computing the rigid body transform that minimizes the least squares error between the two point sets. Gelfand et al. [2003] improve alignment performance and quality by selecting feature points that will constrain all degrees of freedom in the rigid-body transform. Because rigid-body transforms have only six parameters, they are efficient to compute and apply, and do not overfit the (possibly) poor correspondences in early iterations. On the other hand, the algorithm relies on these properties, and does not extend easily to more complex transforms.

Global Rigid-Body Alignment: Several algorithms have been proposed for simultaneously registering a collection of scans, using rigid-body alignments. Chen and Medioni [1992] perform global alignment by incrementally aligning new range scans to all previous ones. Since error accumulates with successive scans (Figure 4), more recent algorithms focus on distributing the error more evenly across scans. Bergevin et al. [1996] iteratively align each scan to all others, until the errors stabilize and are well-distributed. Pulli [1999] improves the practicality of this approach for large collections of scans by finding corresponding points between pairs of meshes, then repeatedly aligning single scans to all others using these correspondences until the system converges. Neugebauer [1997], Benjemaa and Schmitt [1998], Williams and Bennamoun [2000], Li and Guskov [2005], and Krishnan et al. [2005] describe closed-form methods for determining the alignment of all scans simultaneously, while Wen et al. [2005] simultaneously solve for both rigid pose and target feature positions, assuming correspondences are known. Huang et al. [2006] incorporate a non-intersection constraint during global registration. Our method is most similar to Pulli's: we scale to large datasets by dividing registration into two phases, a computationally expensive yet parallelizable pairwise alignment for all pairs of overlapping scans followed by a global optimization based only on point pairs computed during the first phase.

Non-Rigid Alignment Techniques: Ikemoto et al. [2003] address the rigid-body constraint of ICP by dicing each range scan, and performing global rigid-body alignment on the overlapping pieces. This can accommodate low-frequency warps, but causes smoothing. Also, its running time is quadratic in the amount of dicing performed. The Florentine Pietà Project [Bernardini et al. 2002] uses ICP alignment but also incorporates image-based alignment and conformance smoothing of overlapping scans along scanner lines-of-sight to improve the registration further. Hänel et al. [2003] present an extension to ICP that allows deformable objects to be aligned by computing rigid-body warps on different parts of a jointed skeleton. Allen et al. [2003] 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. In contrast, we assume a low-frequency deformation which allows us to compute a warp more efficiently using many fewer parameters. Shum and Szeliski [2000] perform nonrigid alignment of images using feature correspondences, which is similar to our work in many ways. However, the global optimization is performed only on the *rigid* part of the transformation, with the nonrigid warp based on an average of correspondence locations. Due to the irregular nature of 3-D scan overlaps, such an approach leads to unacceptable artifacts in our case (Figure 7).

Thin-Plate Splines: The thin-plate spline that we use to represent our warp has been extensively used in the pairwise context in medical imaging applications, where its use was first proposed by Bookstein [1989]. Chui and Rangarajan [2003] compute point correspondence iteratively using a softassign/deterministic annealing framework to 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). Jian and Vemuri [2005] match Gaussian mixture models as opposed to the data points themselves. These algorithms assume that high frequency information is unreliable (as in the case of scans of two different brains). This

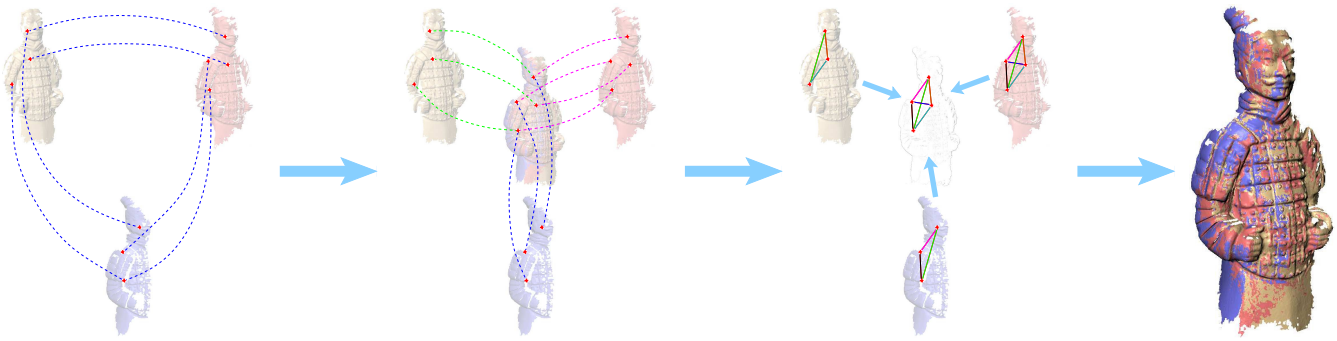


Figure 3: The global alignment framework. We compute pairwise correspondences between features selected on each range scan and all other range scans. From these we obtain a set of global features whose positions on each range scan are known. Next we optimize the locations of these features using relative distances between them as constraints. Finally, we warp and align each range scan using the features as control points of a thin-plate spline.

makes them inappropriate for range scan alignment, where accurately aligning high frequency detail is a primary goal.

The thin-plate spline has also been used for *pairwise* registration of large range scans [Brown and Rusinkiewicz 2004]. Feature correspondences are obtained by dicing one range scan hierarchically, and aligning each piece separately to the other range scan using ICP. The method provides improved pairwise alignments, but includes no mechanism for computing any globally consistent warping among many range scans. Also, the discontinuities between piecewise alignments in the feature correspondence computation are apparent in the structure of the spline warp.

3 Global Alignment Framework

Global alignment aims to find alignments for all scans which minimize overall alignment error while distributing it evenly among scans. Significant error generally results in smoothing and/or artifacts in the final mesh; since low-frequency error in individual range scans produces substantial rigid-body misalignment, it effectively causes error at *all* frequencies in the merged result. In the non-rigid case, the undesirable high frequency misalignment can be traded for a low frequency warp in the range data, restricting the final error to the frequencies where error already existed. On the other hand, it is possible to produce a badly warped global alignment which nevertheless has low alignment error, so it is important to minimize and evenly distribute the warp.

The simplest approach to aligning multiple scans is to sequentially align each scan to either the single previous scan, or to all previous scans. While this frequently works for small models, it does not scale well. Small alignment errors accumulate with each successive scan so that the last scans to be incorporated have far higher error than the first ones. This is easily seen in examples such as 4, where the scans form a closed ring. Although these scans from the edge of a *Forma Urbis Romae* fragment in fact form a closed loop, the first and last scans in the sequence do not align because of the accumulated error. Successive non-rigid alignment fares even worse, because warp can accumulate as well. A more successful alternative for rigid registration is to iteratively align each scan either to all scans or to all previously aligned scans. Unfortunately, the additional freedom of the non-rigid warp causes serious problems for this class of algorithms. Figure 5 shows the result of a range scan being pulled in two directions at once by simultaneous alignment to two slightly misaligned scans: distortion and high frequency artifacts then propagate through the alignment process. The long running times for convergence and the requirement that all scans be in memory simultaneously make these iterative technique unsuitable for large datasets in any case.

Figure 4: Edge scans from *Forma Urbis Romae* fragment 033 aligned sequentially, using non-rigid alignment. Error accumulates with each scan, such that the first and last scans do not align at all. In this case, the scans started correctly aligned; all error is due to the alignment process.

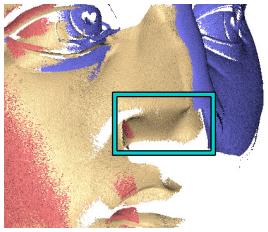


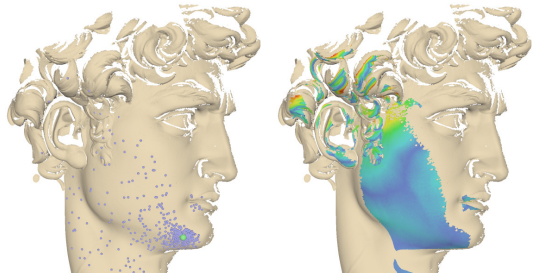
Figure 5: The peach scan has been simultaneously aligned to the blue and red scans using non-rigid alignment. The result is that it has been pulled in two directions at once, leading to stretching and high-frequency artifacts such as the pinched region on the nose.

We solve the problem of minimizing and distributing warp (Figure 3) by first finding the locations of a common set of features on all range scans, using the locally weighted ICP method described in Section 3.2. Conceptually, we might find the locations of the warrior's mouth, elbow, shoulder and collar bone in every range scan. We next find positions in space for each of these feature points, such that each range scan can be mapped to these locations with minimal warp. Finally, using these features as control points, we compute a thin-plate spline warp for each range scan that brings it into alignment with the global point positions (and hence with all other range scans). Note that, as in most previous work on global registration, we assume that an approximate initial alignment is available.

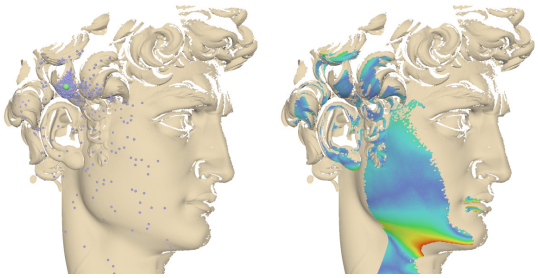
3.1 Feature Point Selection

Feature points must cover the entire model and yield stable correspondence computations between range scans. While it may be intuitive to think of these features as derived from the object (e.g. the tip of David's nose), we do not select them this way because we have only the initial range scans available at this stage. Instead, we select features on each range scan, then find their correspondences on every other range scan. We select half the features using uniform random sampling (to ensure coverage in smooth areas of the model), and half using the stable sampling technique described in section 3.2 (to ensure we can find good feature correspondences wherever possible).

Since the feature points will eventually define a spline, the number of points required depends in principle on the degree of warp in the scans. In practice, however, we have found the algorithm to be relatively insensitive to the number of points that are selected. Instead, we sample a fixed percentage of points from each range scan, typically 1% for small models, 0.1% for medium-size models, and less for very large models with a high degree of overlap, such as the Digital Michelangelo scans.



(a) Selected samples and alignment for locally weighted ICP of a feature point on the chin.



(b) Selected samples and alignment for a point in the hair.

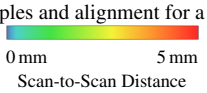


Figure 6: Locally weighted ICP focuses the alignment on a small region, producing a better result in that area and a worse one elsewhere. We use locally weighted ICP centered at feature points to compute accurate correspondences.

3.2 Locally Weighted ICP for Correspondences

The next stage of our pipeline is to find the correspondence for each feature point $\mathbf{f}_i$ on each scan that overlaps it. In order to find correspondences efficiently for thousands of points distributed across many scans, we use a novel variant of the ICP algorithm. Although ICP is usually used for aligning entire scans, it has also been used on "diced" sections of scans to provide more local alignment [Ikemoto et al. 2003; Brown and Rusinkiewicz 2004]. We extend this idea with "locally weighted ICP," which gives higher weight in the registration to the portion of the mesh immediately surrounding a feature $\mathbf{f}_i$. After alignment, the closest point on the other mesh is used as the correspondence.

Our ICP variant is based on the "baseline" algorithm described by Rusinkiewicz [2001], with a point-to-plane error metric, rejection of matches to edge points, closest-point computations accelerated with a k-D tree, and matching based on compatibility of normals. For additional stability, we also use compatibility of mean curvatures to improve matching.

Our major change is to the point-selection stage: source points are sampled using a probability distribution that makes it more likely that points near $\mathbf{f}_i$ are used. The major component of the probability is a decreasing function of distance from the feature point $\mathbf{f}_i$ under consideration:

$$p_{feature}(\mathbf{x}) = \frac{1}{\varepsilon + \|\mathbf{x} - \mathbf{f}_i\|^2}. \tag{1}$$

If, however, the weight were based entirely on this, there would be a danger of selecting only points that lie on a region of the mesh without sufficient geometric variation to constrain all six degrees of freedom of a rigid-body transformation. Therefore, we augment the probability function with a term that assigns high probability to locations on the mesh that constraint the alignment:

$$p_{stability}(\mathbf{x}) = \begin{pmatrix} \mathbf{x} \times \mathbf{n}_\mathbf{x} & \mathbf{n}_\mathbf{x} \end{pmatrix} C^{-1} \begin{pmatrix} \mathbf{x} \times \mathbf{n}_\mathbf{x} \\ \mathbf{n}_\mathbf{x} \end{pmatrix}, \tag{2}$$

where C is the 6×6 point-to-plane ICP covariance matrix computed over the entire region of overlap of the scans (see [Gelfand et al. 2003] for a derivation of the expression for ICP stability). Intuitively, the covariance matrix has small eigenvalues corresponding to transformations that are not well-constrained, hence using its inverse will assign higher probability to points that are important for these under-constrained degrees of freedom.

Note that Gelfand et al. [2003] samples for stability in a slightly different fashion. Whereas we compute a probability distribution, that work computes the stability of every point with respect to each of the six degrees of freedom, sorts these six lists and selects points from the top of each. For the *Forma Urbis Romae* data considered in that paper, the set of constraining points is of nearly measure zero, and sorting is essential for a stable alignment. For less pathological data sets, however, sorting leads to "clumping" of the samples and is less effective than sampling the probability distribution.

Our final probability function consists of the product of $p_{feature}$ and $p_{stability}$. The function is normalized, integrated into a cumulative distribution function, then numerically inverted to transform a uniform random variable into samples to be selected (i.e., the "inversion method" for importance sampling). A separate locally weighted ICP is run for each feature point $\mathbf{f}_i$, and the nearest point on the target range scan is selected as the correspondence to $\mathbf{f}_i$. Figure 6 shows the results of aligning a pair of scans for two different feature points. Note that the points used for ICP are selected in the vicinity of the feature points, leading to lower alignment errors in those regions, and therefore more accurate correspondences.

Our feature correspondence algorithm adds little to the total computation time, relative to the time that would be required for rigid-body registration. Because an initial overall alignment of each pair of range scans is performed first, it takes only a few iterations of ICP (taking a few milliseconds) to perform each weighted alignment. The most expensive parts of the ICP, building k-D trees and computing the overlapping region between each pair of meshes, are computed only once per pair. Per-vertex normals and curvatures are generated and smoothed (to remove noise) as a preprocess, so this computation is performed only once per range scan.

3.3 Global Point Positioning

For a given feature i, we denote the location of its correspondence on mesh m by $\mathbf{f}_i^m$. In order to compute a consistent alignment, we need to find a "global position" $\mathbf{g}_i$ for each feature point, based on the known correspondences $\mathbf{f}_i^m$. Moreover, we must find global positions that do not induce unnecessary warp, as in Figures 5 and 7. Intuitively, for any mesh m, the mutual arrangement of all feature locations $\mathbf{f}_i^m$ should differ as little as possible from the spatial arrangement of the corresponding $\mathbf{g}_i$. If all correspondences are correct, and there exists a rigid-body alignment of all scans, then these spatial arrangements should be identical; in the presence of a low frequency warp this will be almost true, with nearby features maintaining their spatial arrangement better than distant ones.

Two simple (albeit inadequate) heuristics for computing the position of each $\mathbf{g}_i$, illustrate the relationship between each $\mathbf{g}_i$ and the corresponding $\mathbf{f}_i^m$. The first places each $\mathbf{g}_i$ at the location where the feature was originally selected. By this, we mean that if feature i was initially chosen among the vertices of mesh m, then $\mathbf{g}_i = \mathbf{f}_i^m$. Unless the range scans were perfectly aligned to start with, this will result in mutually inconsistent positions for the $\mathbf{g}_i$, and we will have a situation akin to Figure 5. Alternatively, we could set the position of $\mathbf{g}_i$ to be the average over all $\mathbf{f}_i^m$. This too will fail, since even nearby features have correspondences on different sets of range scans (Figure 7), again resulting in mutually global positions.

Instead, guided by the intuition that global point positions should be affected by their geometric configuration relative to neighboring features, we optimize for the $\mathbf{g}_i$ by attempting to preserve their relative distances, as computed within each mesh. This is in some ways

similar to the approach of Laplacian mesh coordinates [Sorkine et al. 2004], in which each vertex is represented relative to its neighbors. Conceptually, we consider placing a "spring" with *non-zero rest length* between all pairs of features i and j. Assuming feature i was originally selected on mesh m and feature j on mesh n, we set this rest length to be $\frac{1}{2}(|\mathbf{f}_i^m - \mathbf{f}_j^m| + |\mathbf{f}_i^n - \mathbf{f}_j^n|)$ (Figure 8).

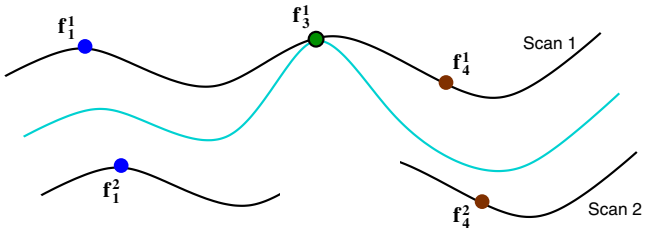


Figure 7: Mutually consistent global feature positions must be selected to prevent unnecessary warp from being introduced. Here the average position for each feature was used; these are not mutually consistent because the green feature has no correspondence on the bottom scan.

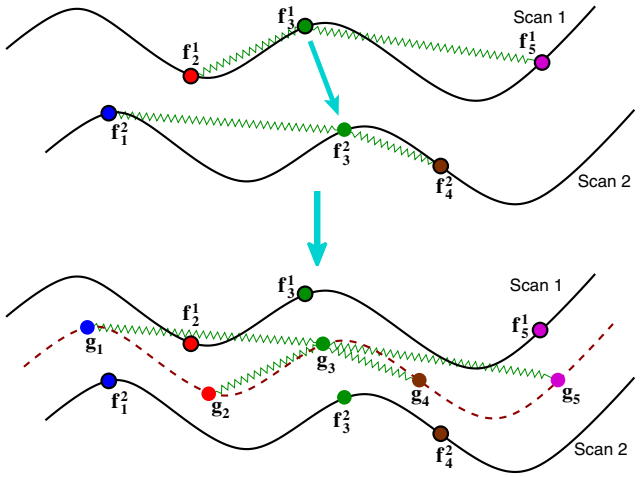


Figure 8: *Above:* Global feature points are attached to each other by springs whose desired length is based on distances along range scans. Since the red ($\mathbf{f}_2$), green ($\mathbf{f}_3$) and magenta ($\mathbf{f}_5$) features are all associated with the top scan, the rest length of the springs attaching them is the actual distance between the features. However the rest length of the springs attaching $\mathbf{f}_3$ to the blue ($\mathbf{f}_1$) and brown ($\mathbf{f}_4$) ones is based on the distance from the green feature's *correspondence* on the bottom scan. *Below:* Point positions are adjusted to minimize the total spring energy, ensuring that the aligned meshes will have minimal warp.

Mathematically, we form an error metric for each point pair:

$$\mathscr{E}_{ij} = \sum_m w_{ij} \left(|\mathbf{g}_i - \mathbf{g}_j| - |\mathbf{f}_i^m - \mathbf{f}_j^m| \right)^2, \tag{3}$$

where the weighted sum is taken over all meshes m such that either:

- Both features i and j were originally selected on mesh m, or
- Feature i was selected on m and feature j has a valid correspondence $\mathbf{f}_j^m$ on m.

The overall energy function is the sum of all the pairwise functions. To minimize it, we initialize each $\mathbf{g}_i$ to be the average of all valid $\mathbf{f}_i^m$, then move one point at a time using gradient descent. Despite the simplicity of this approach, we have found it to be efficient and robust (in particular, it is more stable than Newton's method), though of course it is not guaranteed to converge to a global minimum.

3.4 Outlier Rejection

Our global point positioning approach shares the drawback of all least-squares methods: it is sensitive to outliers. Incorrect correspondences, which occur most commonly in flat areas, lead to incorrect springs and poor feature positioning. Therefore, we implement aggressive pruning of outlier correspondences, based on the error and stability information we obtain from ICP, both in isolation and in relation to nearby features.

The first stage of outlier rejection is based on absolute error thresholds. If the unweighted ICP we perform between a pair of scans before computing correspondences fails (due to lack of stability or insufficient overlap), we do not record any correspondences between that pair of meshes. We also reject any correspondence (1) whose RMS ICP error (point-to-plane distance) is above a threshold (30 mm), (2) which is more than 50 mm from the average feature location, or (3) whose stability (as reported by locally-weighted ICP — Equation 2) is lower than 0.001. These thresholds are only designed to eliminate egregious correspondences, and are therefore set to be large relative to the mesh sample spacing.

After optimizing feature point locations, we do additional pruning of points based on aggregate error. The first step is to thin out feature points, so that no two features are closer than some threshold to each other. Starting from each remaining feature point, we find all other feature points within this distance, and discard all but the one with lowest spring energy. This helps prevent high-frequency artifacts due to positioning errors, but more importantly it reduces the running time of computing thin-plate splines. The appropriate threshold to use depends on the sampling density and size of the dataset. For Digital Michelangelo data we use a minimum distance of 8 mm, while for small datasets such as the penguin we allow points to be closer — about 2.5 mm.

Next we discard features which move too far relative to their nearby features on the same mesh. For each mesh m, we consider the list of features with correspondences on that mesh. For each of these features $\mathbf{f}_i^m$, we find the 8 nearest features (again with correspondences on the mesh), and calculate the median distance d between the feature locations on mesh m and their global positions. If $|\mathbf{g}_i - \mathbf{f}_i^m| > 4d$, we discard feature i. After all of these pruning steps we re-optimize the locations of the remaining feature points.

3.5 Warping with Thin-Plate Splines

Once the global features are positioned, all range scans must be warped to align to them. That is, we compute the warp that maps all the $\mathbf{f}_i^m$ on scan m to their global positions $\mathbf{g}_i$, and apply that warp to the entire scan. (Since these warps are computed independently for each scan, we drop m from our notation in this section, referring instead to $\mathbf{f}_i$.) The nature of the warp can be complex, difficult to model, and vary between data sets. For instance, calibration error, lens distortion, combining input from different sensors, and object movement all cause different kinds of warp. Each of these is difficult to model on its own, and we would like to handle all of them seamlessly. As in a number of other non-rigid alignment systems, such as [Brown and Rusinkiewicz 2004] and [Chui and Rangarajan 2003], we turn to the thin-plate spline. However, we extend the use of the thin-plate spline to global registration by mapping each feature point onto its *global* position.

Here we briefly summarize the properties of the thin-plate spline, originally introduced by Duchon [1977]. It is a non-rigid, globally smooth function that is easily computable, separable into affine and non-affine components, and contains 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 $\mathscr{S} : \mathbf{R}^n \to \mathbf{R}$ is an n-dimensional thin-plate spline, the *bending energy*,

$$J = \int \left(\sum_{i,j} \mathscr{S}_{\mathbf{f}_i \mathbf{f}_j}^2 \right) \mathrm{d}\mathbf{f}_1 \ldots \mathrm{d}\mathbf{f}_n \tag{4}$$

is minimal. Note that since affine transformations are linear, they contribute no error under this metric.

Duchon proves that, for two corresponding point sets $F = (\mathbf{f}_1 \dots \mathbf{f}_n)^t$ and $G = (\mathbf{g}_1 \dots \mathbf{g}_n)^t$, there is a unique function $\mathscr{S}$ such that $\mathscr{S}(\mathbf{f}_i) = \mathbf{g}_i$ and whose bending energy is minimal. Furthermore, this function takes the form $A\mathbf{x} + WK(\mathbf{x})$, where $\mathbf{x}$ is a point written in homogeneous coordinates, A is an affine transformation, W is a fixed n-dimensional row vector of non-affine warping parameters constrained to $WF^t = 0$, and $K(\mathbf{x})$ is an n-dimensional column vector where $K(\mathbf{x})_i$ is the Green's function $U(|\mathbf{x} - \mathbf{f}_i|)$. In our case (minimizing second order partials in $\mathbf{R}^3$), U is simply $|\mathbf{x} - \mathbf{f}_i|$; the constant factor is implicitly folded into W.

Thin-plate splines need not be interpolating. Instead, they can minimize the energy functional [Duchon 1977; Wahba 1990]

$$E_\lambda = \frac{1}{n} \sum |\mathbf{g}_i - \mathscr{S}(\mathbf{f}_i)|^2 + \lambda J \tag{5}$$

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.

For the interpolating case, the thin-plate spline specification provides a linear system of equations. For the approximating case, [Wahba 1990, eqs. 2.4.23 and 2.4.24] derives a similar system of equations by rewriting Equation 5 in matrix form, performing a QR decomposition on $\mathscr{S}$, and simplifying:

$$\begin{aligned} AF + W(K + n\lambda I) &= G \\ WF^t &= 0 \end{aligned} \tag{6}$$

where K is $n \times n$ and $K_{ij} = U(|\mathbf{f}_i - \mathbf{f}_j|)$.

Because we would like very precise alignments (since multiple measurements of rigid objects should align precisely), the spline must be heavily weighted toward interpolation. For this reason we must rely on accurate correspondences to produce a good alignment. Nevertheless, we set λ to a very small value (10^{-10}), because doing so improves the numerical stability of Equation 6 when many point pairs are used.

4 Results

We present alignment results from a variety of sources. The Digital Michelangelo scanner acquires range scans up to 50 cm long, with a warp of several millimeters. The commercially available NextEngine and Minolta Vivid scanners have much lower warp for a similar resolution, but acquire commensurately smaller viewing volumes. We also show applications of our methods to a photometric stereo dataset, which exhibits a different and much larger warp.

4.1 Digital Michelangelo Scanner

Michelangelo's David: Figure 10 shows closeups of the head of Michelangelo's David, created with global rigid-body alignment, and with our non-rigid alignment. Figure 9 shows cross sections of the range scan positions for each range scan at the right edge of the eye after ICP alignment and after our non-rigid alignment, as well as histograms of alignment error under rigid and non-rigid alignments. The original data set for David's head is 1406 range scans, and the final model contains 28 million vertices when merged at 0.25 mm precision with VRIP [Curless and Levoy 1996].

The most obvious difference between the rigid and non-rigid models is the reduction in merging artifacts and smoothing due to misalignment. This is especially apparent in areas of high detail, such as the closeup of David's hair shown in Figures 10(e) and 10(f). Here, some of the chisel marks between locks of hair are badly smoothed in the rigid-body version, while others are nearly obliterated. The non-rigid version is sharper and clearer. More subtly (and counter-intuitively), rigid-body alignment can distort shapes, as it has in the pupil of David's right eye (Figure 10(a)). This occurs because the merging process follows different misaligned range scans in different areas. With nonrigid alignment this artifact is removed, and the local shape is better preserved.

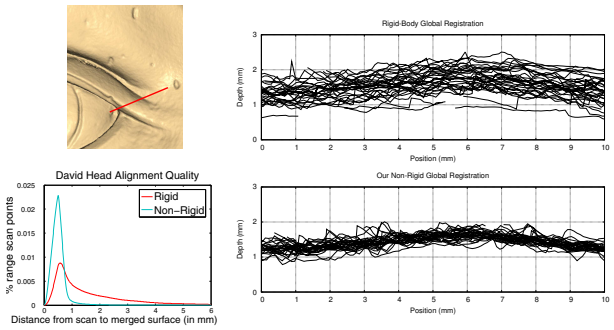


Figure 9: Alignment error plots for David's head. The line plots at right show the depths of the aligned range scans along the marked line on David's eye. The larger spread on the rigid scans *(above)* compared to the non-rigid ones *(below)* causes the smoothing and artifacts seen in figure. At the lower left are histograms of alignment error under rigid-body and non-rigid registration. The rigid-body alignment histogram reveals a long tail of misalignments by several millimeters, which corresponds to the greater spread of range scans seen in the depth plots.

We sampled 0.0075% of the range scan vertices to use as features, resulting in 16499 feature points, of which 5832 were retained for global registration. Much of the sample rejection comes during the point thinning stage, so that increasing the sampling rate would not improve alignment quality. Correspondence required 78 hours of CPU time, distributed across 60 nodes of a cluster (1.5 hours of wall clock time), while global point positioning required an additional 30 minutes on a single processor.

Forma Urbis Romae: Figure 1 shows piece 033 from the *Forma Urbis Romae* data set [Levoy et al. 2000]. Because the only features available on the top of these scans are $^1/_2$ mm-deep incisions, alignment is particularly tricky. Furthermore, scans taken at an angle so as to capture both the top and sides of the pieces cannot generally capture the incision due to self-occlusion; instead these scans contain holes with noisy boundaries. Because no full model of this piece previously existed, we have generated four different versions using our algorithms, ranging from a classic rigid-body alignment to the full non-rigid treatment. For Figures 1(a) and 1(c) we disabled locally weighted ICP; a single unweighted ICP alignment is computed for each pair of range scans, but no additional locally weighted ICPs are performed before reading off correspondences. This is similar to the pairwise alignments/correspondences computed in other multiview rigid-body algorithms. Similarly, for Figures 1(a) and 1(b), we did not compute thin-plate spline transforms. Instead, we perform our usual global point positioning, but compute only the best rigid-body alignment of each scan.

As is evident in Figures 1(a) and 1(b), rigid-body alignment leads to substantial misalignment of the top, resulting in surface roughness. Harder to see are the phantom incisions due to misalignment, which show up in both rigid-body alignments and when locally weighted ICP is removed. Some of these are circled in 1(c). We have found that the top is consistently better with weighted ICP correspondences and thin-plate spline alignments. The top edges are also cleaner due to better alignment of the sides with the top.

The parameters for aligning this model are only slightly different than those used for David's head. Because there are only a tenth as many scans, we used ten times the sampling rate to achieve a similar feature density (0.075%). Also, as discussed in Section 3.2, the top surface of these fragments required taking only the highest-stability points for ICP, while the non-pathological edges and bottom used the (more robust) stability sampling. Correspondences required 48 CPU hours, while global registration took 27 minutes.

Awakening: This model (Figure 11) consists of 1836 range scans. While this is less than 50% more scans than David's head, the model is much larger, with a lower degree of overlap. As with

the other models, our non-rigid alignment improves the sharpness while reducing artifacts in areas of high detail such as the face and thigh. Correspondences required 51.5 CPU hours, and global alignment required 1 hour.

4.2 Other 3-D Scanners

While the Digital Michelangelo scanner achieves high resolution for its working volume, and consequently has especially large deformations, the problem is not limited to this scanner. As shown in figure 2, the commercially-available NextEngine Desktop 3D Scanner and Minolta Vivid 910, both based on laser triangulation, contain small yet measurable deformation, and we expect that other commercial scanners do as well. Once again, the non-rigid alignments reduce the alignment error (Figure 12). While these deformations are tenths of millimeters rather than millimeters, the working volume is also an order of magnitude smaller.

We have also tested our methods on datasets produced by scanners not based on the principles of triangulation. In particular, the penguin dataset in Figure 13 was captured using photometric stereo [Woodham 1980], which estimates the surface *normals* based on observations of a surface under multiple lighting conditions. These normals are integrated to obtain surface geometry, and any small bias in the normals, such as that due to interreflection, accumulates to produce warped geometry. Here too, non-rigid alignment can help. Figure 13 shows the penguin model scanned using photometric stereo and aligned using rigid and non-rigid alignment. While Nehab et al. [2005] have demonstrated the ability to correct for this bias by combining photometric normals with triangulation data, our methods can work with the biased normal scans directly.

The penguin required 25 minutes for correspondence computation and one minute for global alignment, while the warrior required 1.5 hours for correspondences and 2 minutes for global registration. Because these models are so much smaller, the feature sampling rate was increased to 0.1%, the minimum distance between features was reduced to 3 mm, and the minimum stability threshold for correspondences was raised to 0.01.

5 Discussion and Future Work

We have presented an algorithm for global non-rigid alignment of three dimensional range scans. Our algorithm is robust not only to noise, but also to non-linear warp caused by scanner calibration error. We recover accurate feature correspondences using a locally weighted, stable ICP computation, then obtain a consistent set of range scans using thin-plate spline warps. The resulting merged models contain more detail than rigidly-aligned versions when scans contain warp. The algorithm scales well with data size, requiring running time proportional to the number of overlapping pairs of scans. Because of its scalability and efficiency, we have found it to be a practical method for rigid as well as nonrigid alignment of models of all sizes.

It is important to be clear about the ways in which our results are "better" than those produced by rigid-body alignment. Indeed, we can not claim to produce results whose overall shape is necessarily closer to the original, since there is just as much uncertainty in the overall position of the scans. However, by avoiding the loss of high-frequency detail caused by low-frequency warps, our final meshes are more *precise* than those produced using rigid-body alignment. Moreover, because of the deformation-minimizing properties of thin-plate splines, our results reflect the minimum non-rigid warp necessary to compensate for deformation in the data. As imagers and 3D scanners become more ubiquitous and inexpensive, and as they are pushed to their resolution and working-volume limits, we believe that in many situations calibration will become the ultimate limit on achievable quality. Our algorithm produces results that maximize the high-frequency fidelity of the result while introducing minimum deviation from the original data.

There are some areas of the algorithm which will benefit from further refinement. In certain areas, such as the top of David's head, the acquired range scans are simply of poor quality. Large, flat areas on the sides and bottoms of *Forma Urbis* fragments are fundamentally unstable. By incorporating existing confidence data into our framework, detecting outlier scan regions during the merging process, and dynamically selecting whether to perform locally weighted ICP in unstable regions, we can improve the final mesh quality in these areas. Doing so would likely reduce high frequency error further at the expense of simplicity.

There is also room for improvement in the outlier rejection itself. In particular, we believe it would be better to reject *correspondences* which large spring energy relative to their neighbors or to other correspondences of the same feature rather than rejecting feature points themselves. While this is simple to do in our framework, we have not yet implemented it.

Acknowledgments

Numerous people have provided helpful suggestions and input on this project, as well as code and data. We would particularly like to thank Natasha Gelfand, Marshall Brown, Misha Kazhdan, Diego Nehab, and all the members of the Princeton graphics group, as well as the Digital Michelangelo Project at Stanford University and the Aim@Shape repository. We would also like to acknowledge the support of the Sloan Foundation and NSF grant CCF-0347427.

References

ALLEN, B., CURLESS, B., AND POPOVIĆ, Z. 2003. The space of human body shapes: Reconstruction and parameterization from range scans. *ACM Trans. Graphics (Proc. SIGGRAPH) 22*, 3, 587–594.

BENJEMAA, R., AND SCHMITT, F. 1998. A solution for the registration of multiple 3d point sets using unit quaternions. In *Proc. ECCV*.

BERGEVIN, R., SOUCY, M., GAGNON, H., AND LAURENDEAU, D. 1996. Towards a general multi-view registration technique. *IEEE Trans. PAMI 18*, 5, 540–547.

BERNARDINI, F., RUSHMEIER, H., MARTIN, I. M., MITTLEMAN, J., AND TAUBIN, G. 2002. Building a digital model of Michelangelo's Florentine Pietà. *IEEE Computer Graphics and Applications 22*, 1, 59–67.

BESL, P. J., AND MCKAY, N. D. 1992. A method for registration of 3-D shapes. *IEEE Trans. PAMI 14*, 2, 239–256.

BOOKSTEIN, F. L. 1989. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. PAMI 11*, 6 (June), 567 – 585.

BROWN, B., AND RUSINKIEWICZ, S. 2004. Non-rigid range-scan alignment using thin-plate splines. In *Proc. 3DPVT*.

CHEN, Y., AND MEDIONI, G. 1992. Object modelling by registration of multiple range images. *Image and Vision Computing 10*, 3, 145–155.

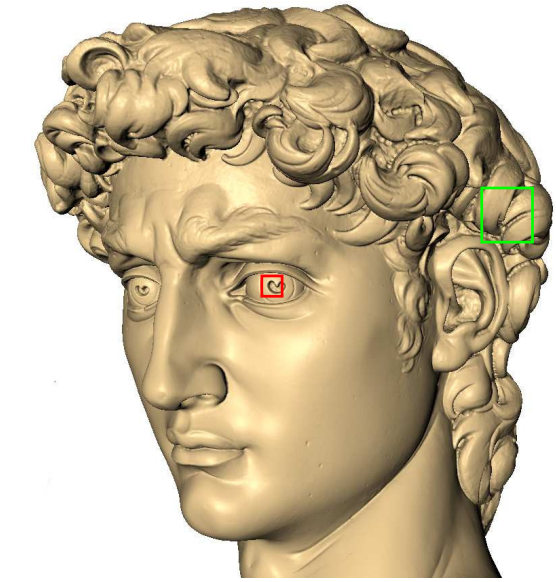
CHUI, H., AND RANGARAJAN, A. 2003. A new point matching algorithm for non-rigid registration. *CVIU 89*, 2-3 (February-March), 114–141.

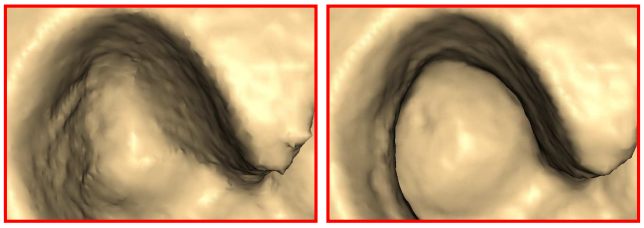
CURLESS, B., AND LEVOY, M. 1996. A volumetric method for building complex models from range images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press, 303–312.

DUCHON, J. 1977. Splines minimizing rotation-invariant seminorms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables*, Springer-Verlag, Berlin, 85–100.

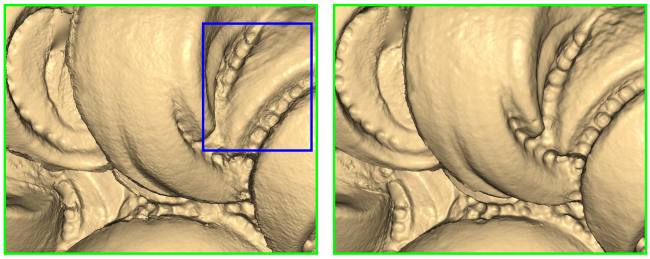
GELFAND, N., IKEMOTO, L., RUSINKIEWICZ, S., AND LEVOY, M. 2003. Geometrically stable sampling for the ICP algorithm. In *Proc. 3DIM*.

HÄHNEL, D., THRUN, S., AND BURGARD, W. 2003. An extension of the ICP algorithm for modeling nonrigid objects with mobile robots. In *Proc. IJCAI*, IJCAI.

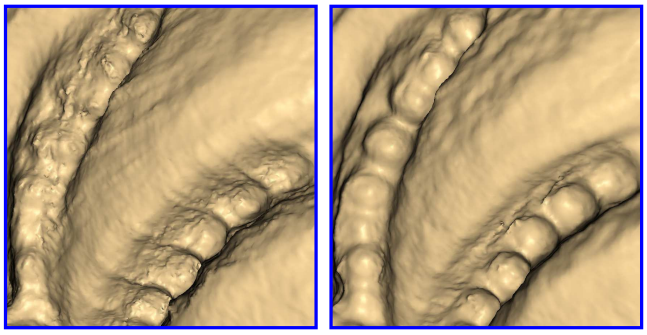




(a) Rigid Pupil (b) Non-Rigid

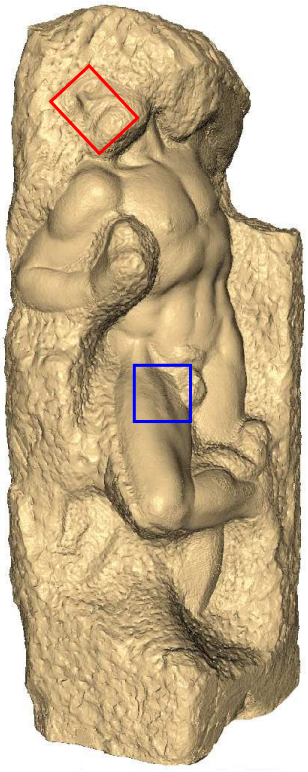


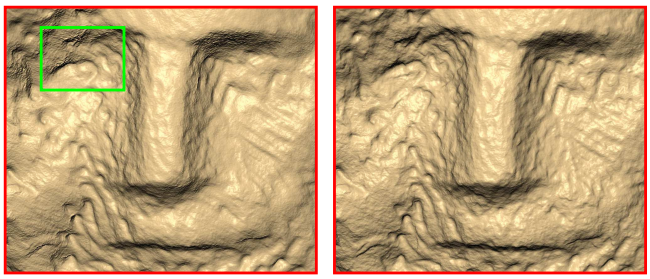
(c) Rigid Hair (d) Non-Rigid



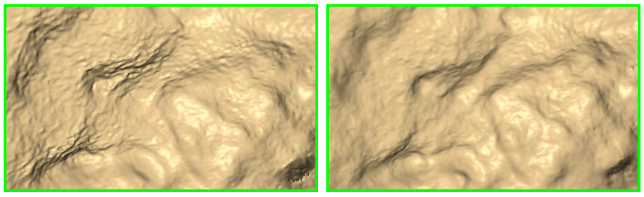
(e) Rigid Hair Detail (f) Non-Rigid

Figure 10: Comparisons of the rigid-body and non-rigid alignments of David's head. The non-rigid alignments are consistently sharper, with fewer merging artifacts. Note in particular the distorted shape and edges of the pupil (Figure (a)) and artifacts in the chisel marks in the hair (Figure (e)).

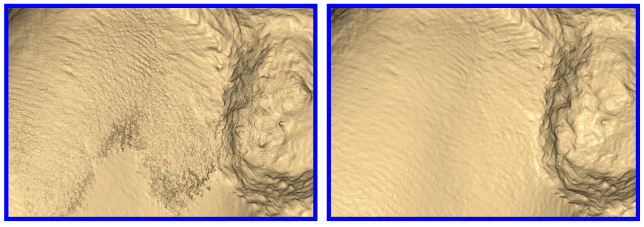




(a) Rigid Face (b) Non-Rigid



(c) Rigid Eye (d) Non-Rigid



(e) Rigid Leg (f) Non-Rigid

Figure 11: Michelangelo's Awakening. Non-rigid alignment yields sharper detail, especially on the face, while reducing misalignment artifacts such as on the leg.

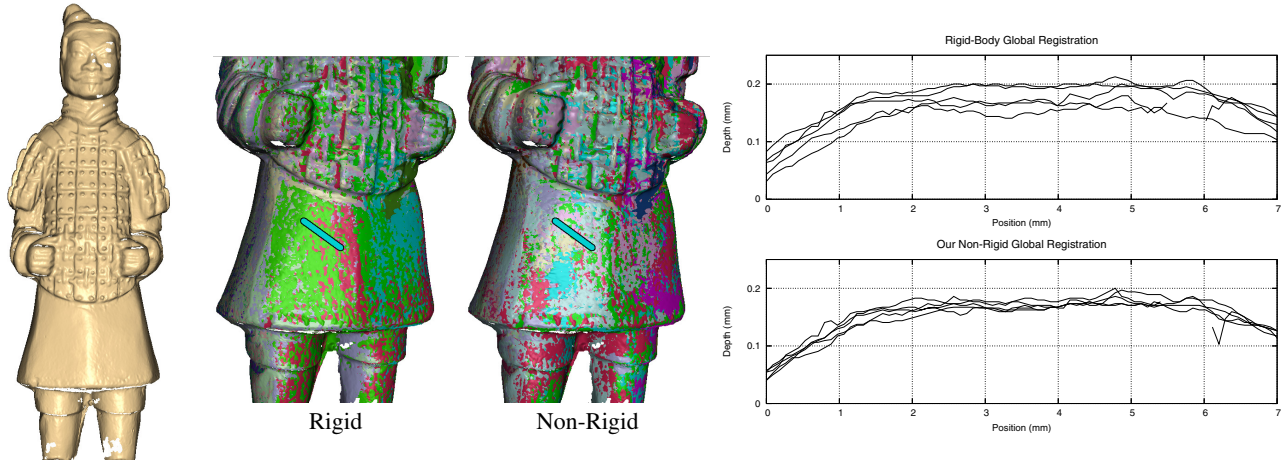


Figure 12: The warrior was scanned using the commercially-available NextEngine 3D Desktop Scanner — a laser triangulation scanner. The central images show each range scan in a different color under rigid and non-rigid alignment. The large fields of a single color in the rigid alignment indicate biased misalignment due to warp; the more mottled color of the non-rigid alignment shows that this bias has been removed. The plots at the right show the depths of each range scan along the marked line.

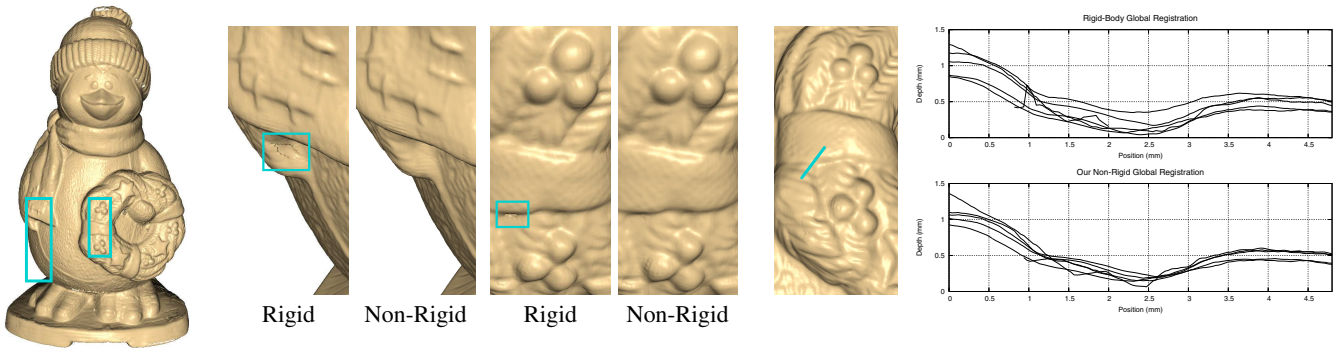


Figure 13: This penguin was acquired using photometric stereo, which acquires and integrates a normal field rather than capturing geometry directly. Small errors and bias in the normal field (due to interreflection, for instance) compound to produce large warps in the geometry. Under rigid-body alignment, these lead not only to increased blurring, but also to "tears," two of which are shown here. At the right, we show a depth plot of all range scans under rigid and non-rigid alignment along the marked line.

HUANG, Q.-X., FLÖRY, S., GELFAND, N., HOFER, M., AND POTTMANN, H. 2006. Reassembling fractured objects by geometric matching. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, ACM Press, New York, NY, USA, 569–578.

IKEMOTO, L., GELFAND, N., AND LEVOY, M. 2003. A hierarchical method for aligning warped meshes. In *Proc. 3DIM*.

JIAN, B., AND VEMURI, B. 2005. A robust algorithm for point set registration using mixture of gaussians. In *Proc. ICCV*.

KRISHNAN, S., LEE, P., MOORE, J., AND VENKATASUBRAMANIAN, S. 2005. Simultaneous registration of multiple 3d point sets via optimization on a manifold. In *Proc. Symposium on Geometry Processing*.

LEVOY, M., PULLI, K., CURLESS, B., RUSINKIEWICZ, S., KOLLER, D., PEREIRA, L., GINZTON, M., ANDERSON, S., DAVIS, J., GINSBERG, J., SHADE, J., AND FULK, D. 2000. The Digital Michelangelo Project: 3-D scanning of large statues. In *Proc. SIGGRAPH*.

LI, X., AND GUSKOV, I. 2005. Multiscale features for approximate alignment of point-based surfaces. In *Symposium on Geometry Processing*, 217–226.

NEHAB, D., RUSINKIEWICZ, S., DAVIS, J., AND RAMAMOORTHI, R. 2005. Efficiently combining positions and normals for precise 3D geometry. *ACM Transactions on Graphics (Proc. SIGGRAPH) 24*, 3 (Aug.).

NEUGEBAUER, P. J. 1997. Geometrical cloning of 3D objects via simultaneous registration of multiple range images. In *Proc. SMA*, IEEE Computer Society, 130.

PULLI, K. 1999. Multiview registration for large data sets. In *Proc. 3DIM*.

RUSINKIEWICZ, S. 2001. Efficient variants of the ICP algorithm. In *Proc. 3DIM*.

SHUM, H.-Y., AND SZELISKI, R. 2000. Construction of panoramic mosaics with global and local alignment. *International Journal of Computer Vision 36*, 2, 101–130.

SORKINE, O., LIPMAN, Y., COHEN-OR, D., ALEXA, M., RÖSSL, C., AND SEIDEL, H.-P. 2004. Laplacian surface editing. In *Proceedings of the Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, ACM Press, 179–188.

WAHBA, G. 1990. *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia, PA, ch. 2.4.

WEN, G., ZHU, D., XIA, S., AND WANG, Z. 2005. Total least squares fitting of point sets in m-d. In *Computer Graphics International*.

WILLIAMS, J., AND BENNAMOUN, M. 2000. A multiple view 3D registration algorithm with statistical error modeling. *IEICE Transactions on Information and Systems E83-D(8)* (August), 1662–1670.

WOODHAM, R. J. 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering 19*, 1, 139–144.