

A Multi-View Extension of the ICP algorithm

Pooja A.
Independent Consultant
Bengaluru, Karnataka
apooja.june@gmail.com

Venu Madhav Govindu^{*}
Department of Electrical Engineering
Indian Institute of Science
Bengaluru, Karnataka 560012
venu@ee.iisc.ernet.in

ABSTRACT

Although the Iterative Closest Point (ICP) algorithm has been an extremely popular method for 3D points or surface registration, it can only be applied to two point sets at a time. By only registering two scans at a time, ICP fails to exploit the redundant information available in multiple scans that have overlapping regions. In this paper, we present a multi-view extension of the ICP algorithm by a method that simultaneously averages the redundant information available in the scans with overlapping regions. Variants of this method that carry out such simultaneous registration in a causal manner and that utilise the transitive property of point correspondences are also presented. The improved accuracy of this motion averaged approach in comparison with ICP and some multi-view methods is established through multiple tests. We also present results of our method applied to some well-known real datasets.

1. INTRODUCTION

Recent advances in scanning techniques and algorithms have resulted in an improved ability to build 3D computer models that can be rendered, manipulated and analysed in a variety of applications like cataloging and display of cultural artifacts, modelling for engineering applications, virtual reality, prosthetic design etc. Since an object cannot be observed in its entirety from a single viewpoint, such 3D modelling entails scanning from multiple viewpoints so as to cover the entire object surface. A key processing step is to *register* these scans in a single co-ordinate system resulting in a merged model that represents the entire object. Since its introduction in [3, 7], the Iterative Closest Point (ICP) algorithm in a variety of modified forms has been the workhorse for 3D registration. While the ICP method yields satisfactory results with an appropriate initialisation, it uses only two scans at a time, implying that the registration of multiple scans has to be carried out in a sequential manner. This is a significant limitation of ICP since we do not exploit all the information available in the set of scans to solve the registration problem. For example, if we consider a series of scans of an object rotated on a turntable, sequential registration using ICP ignores the

^{*}Corresponding Author. This work was carried out when the first author was a student at the Indian Institute of Science.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICVGIP '10, December 12-15, 2010, Chennai, India
Copyright 2010 ACM 978-1-4503-0060-5/10/12 ...\$10.00.

fact that the last scan and the first one have overlapping areas that can yield an additional constraint for registration. Many such additional constraints are available in the full set of scans and can be efficiently utilised. In this paper, we develop a *multi-view* extension of the ICP method that simultaneously solves for the registration of all scans. By simultaneously utilising all available scan-pair relationships, our method yields superior results since registration errors are reduced by a process of averaging and are also distributed evenly across the scans.

2. ICP AND ITS VARIANTS

In this Section we briefly outline the ICP algorithm and discuss some variants of interest in our context. If correspondences for two point sets are known, we can solve for the rotation and translation (i.e. rigid 3D motion) required to register them in a common co-ordinate system. However, in many scenarios such correspondence information is unavailable and needs to be estimated from the points themselves. First introduced in [3, 7], ICP is an iterative method that simultaneously solves for the correspondences between two point sets and registers them. Assuming an initial guess for the rigid 3D motion between point sets, we compute a correspondence map between points in the two sets based on a measure of closeness (**correspondence step**). Using this correspondence map, we can compute the required 3D motion (**motion step**). These two steps are repeated till convergence. ICP can be shown to converge to a local minima and performs well with a good initialisation of the 3D motion. When the point sets lack structure, the common way of assigning correspondence for a point in the first set is to pick its closest point in the second set [3]. Throughout this paper we use this method to determine point correspondences. For structured points forming a mesh or scan surface, [7] proposed to compute the correspondences based on the distance of a point to tangent planes of the second surface. For a discussion on the many variants and issues of efficiencies involved in ICP computations, see [10, 11]. Although the majority of techniques apply to a pair of scans at a time, there are other approaches that address the multi-view registration problem [1, 2, 5, 8, 9, 12]. However, to allow for a comparison with the approach of this paper, the discussion of these multi-view methods is given in Sec. 4.1 and a comparison of their performance and our method is given in Sec. 6.2.

3. MOTION AVERAGING

As noted above, although there are multi-view methods based on ICP, they do not fully utilise all the motion information available in a set of scans. In [4], all available relative motions were averaged to solve for the motion of a camera sequence. Our approach

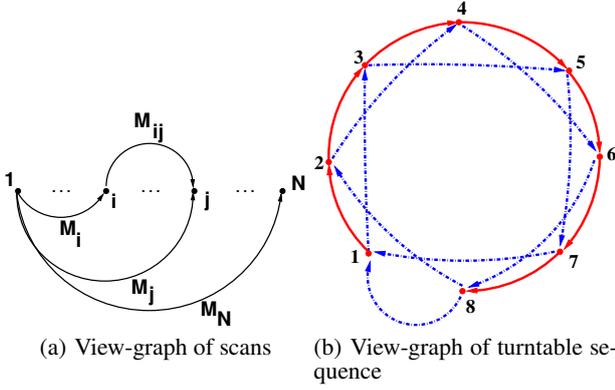


Figure 1: (a) View-graph representing a set of scans to be registered. M_{ij} represents the relative motion estimated from the point sets i and j . $M_{global} = \{I, \dots, M_i, \dots, M_j, \dots, M_N\}$ represents the global camera motion to be estimated. **(b) View-graph of a typical turntable sequence.** The edges in solid red represent the relative motions used for sequential registration with ICP. The edges in dashed blue represent the additional relationships available for motion averaging.

utilises this idea of motion averaging to solve the ICP **motion step** for all scans simultaneously. Here we briefly develop the idea of motion averaging by considering a graph representation of the relationships between scans (see Fig. 1(a)). We denote the 3D rotation or Euclidean motion of a scan with respect to a frame of reference as M . Without loss of generality, we attach the frame of reference to the first scan, i.e. $M_1 = I$. The motions required for full registration of all scans is denoted as $M_{global} = \{I, M_2, \dots, M_N\}$ where N is the number of scans to be registered. Let the graph shown in Fig. 1(a) be $G = (V, E)$. Here, each vertex in V represents the motion associated with a given scan and each edge in E implies that we have sufficient overlap between the two scans (vertices) to be able to register them. The motion required to register scan j onto scan i is given by M_{ij} implying that it can be estimated from scans i and j . Each relative motion M_{ij} can also be written in terms of the global motion model since

$$M_j M_i^{-1} = M_{ij} \quad (1)$$

$$\Rightarrow \underbrace{\left[\dots \quad M_{ij} \quad \dots \quad -I \quad \dots \right]}_{\text{Relative Motions : } \mathbf{A}_{ij}} \underbrace{\begin{bmatrix} \vdots \\ M_i \\ \vdots \\ M_j \\ \vdots \end{bmatrix}}_{\mathbf{M}_{global}} = \mathbf{0}$$

Each relative motion between a pair of scans gives us a single equation of the form in Eqn. 1 and is the information represented by a single edge in E . By using all equations arising from the relative motions (i.e. from all edges in graph G), we can estimate the global motion using $\mathbf{A} \mathbf{M}_{global} = \mathbf{0}$ where \mathbf{A} is composed of the \mathbf{A}_{ij} 's in Eqn. 1 stacked as rows. It can be seen from Fig. 1(a) that a global motion can always be estimated if E contains a spanning tree since in such a case, each vertex is reachable from the frame of reference, i.e. vertex 1. More specifically, for N scans we need only $N - 1$ relative motions available as long as they form a span-

ning tree. In contrast, from a set of N scans we can have as many as $\frac{N(N-1)}{2}$ relative motions (i.e. the case represented by a fully connected graph). These relative motions form a highly redundant system of equations since to specify M_{global} we need only $N - 1$ motions with the requirement that they form a spanning tree.

The rigid motions required to register scans are either 3D rotations (\mathbf{R}) or Euclidean motions (\mathbf{M}) which are elements of the Special Orthogonal and Special Euclidean respectively. The rigid 3D Euclidean motion has the form

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

where $\mathbf{M} \in SE(3)$, $\mathbf{R} \in SO(3)$ and $\mathbf{t} \in \mathbb{R}^3$ are the 3D rotation and translation respectively. Both the groups $SO(3)$ and $SE(3)$ are also Lie groups and have a smooth differentiable structure. Consequently the averaging of many relative motions M_{ij} can be efficiently carried out using the corresponding Lie-algebras [4]. The algorithm proposed in [4] for motion averaging is summarised in Appendix A.

In most scenarios we do not expect to have all the $\frac{N(N-1)}{2}$ relative motions available since many scan pairs would not have any overlapping regions between them. Nevertheless, we do have a significant amount of redundant information available in a set of scans that can be utilised to our advantage. For pedagogical purposes, consider an object on a turntable that moves by 45° between consecutive scans. The resulting scan relationships are shown in Fig. 1(b). The ICP approach would be to register adjacent scans and use the estimated motions to solve for the full registration of all scans. The relative motions between adjacent scans used in this approach are shown in solid red. So if we consider scan 8, its registration with scan 1 will be defined by the composition of all the relative motions along the path from scan 1 to scan 8. In particular, this will result in the individual errors along the path being added up.

However, in most typical scenarios we can expect a substantial overlap between scans that are two steps from each other, say between scans 1 and 3, 2 and 4 etc. As a result we can also solve for the relative motions between such scan pairs (shown in dashed blue). Thus, while the ICP-based global motion estimate uses 7 relative motions between adjacent scans, we now have an additional 8 constraints arising from the dashed blue edges. In fact, one could have more constraints if other scan pairs have sufficient overlap. All of these relative motions, i.e. both solid red and dashed blue ones, can be averaged to simultaneously solve for the global motion. Note that the additional motions provided by the dashed blue edges will act as constraints on the global solution and will prevent the accumulation of error. In particular, notice that we can also add a dashed blue edge directly between scan 8 and scan 1. The relative motion computed by this relationship will act as an ‘anchor’ and distribute the motion errors over the entire global motion and prevent a drift of the motion for scan 8. This idea of motion averaging can be applied to any situation where we can compute some additional relative motions beyond the minimal number defined by a spanning tree. Note that motion averaging is not specific to an ordered sequence of scans and can be applied to any set of scans for which multiple relative motions can be computed.

4. MOTION AVERAGED ICP

We can now introduce our approach to solve for the global reg-

istration of scans by combining ICP and Motion Averaging. Let $\mathbf{M}_E = \{\mathbf{M}_{i_1j_1}, \mathbf{M}_{i_2j_2}, \dots, \mathbf{M}_{i_Sj_S}\}$ denote the set of relative motions represented by the edges in E and $S = |E|$. Here the subscript indices i_1j_1, i_2j_2 etc. denote different pairs of vertices (i, j) defined by the edges in E . Using this representation we can define the equivalent **correspondence step** and **motion step** for each iteration in our multi-view extension.

Correspondence Step : This step is a straightforward extension of the correspondence step in the two-scan ICP. Let us consider a single edge in E that represents a scan pair (i, j) . For each point in scan i , we compute its closest point in scan j and thereby define a pair of point correspondences between the two scans i and j . Here, following the ideas summarised in [10] we use a distance threshold, rigidity constraints and also avoid point correspondences that lie on the boundaries of a scan. We also use a k - d tree to speed up the search for the nearest point on scan j . This correspondence step is carried out for all S scan pairs (i, j) represented by the edges in E . At the end of the correspondence step we have correspondences determined for every scan pair and we can use this information to carry out the motion step.

Motion Step : In the ICP algorithm, this step consists of computing the optimal rigid 3D motion between the two sets of corresponding points and registering the scans using this motion estimate. Furthermore, for the optimal solution, the estimation of rotation and translation can be delinked. For estimating rotations in this step, we use the method based on the SVD decomposition described in [13]. In our multi-view extension, for every scan pair (i, j) we use the same estimation procedure as in the case of ICP. As a result, we have S estimates of the relative motions \mathbf{M}_{ij} which can be averaged. While our methods of computing the rotations and translations are the same as that for ICP, in the context of motion averaging in our multi-view extension we utilise the $\{\mathbf{R}, \mathbf{t}\}$ estimates in a slightly different manner.

We clarify these differences by first considering the motion estimation in the basic ICP algorithm. Let the rotation and translation estimated at iteration k of the basic ICP algorithm be denoted as $\Delta\mathbf{R}(k)$ and $\Delta\mathbf{t}(k)$ respectively. The usual procedure would be to apply these transformations to the second scan and register it to the first one. In other words, at each iteration we put both scans in the same frame of reference and try to adjust the positioning of the second scan. However, if we consider the second scan in its *original* frame of reference the net effective motion applied to it at the k -th iteration is nothing but the composition of all the previous motion estimates, i.e.

$$\begin{aligned} \mathbf{M}(k) &= \prod_{i=1}^k \Delta\mathbf{M}(i) = \Delta\mathbf{M}(k) \mathbf{M}(k-1) \\ &= \Delta\mathbf{M}(k) \Delta\mathbf{M}(k-1) \dots \Delta\mathbf{M}(2) \Delta\mathbf{M}(1) \end{aligned} \quad (3)$$

where $\Delta\mathbf{M}(i)$ is the Euclidean motion computed at iteration i by composing the estimated rotation $\Delta\mathbf{R}(i)$ and translation $\Delta\mathbf{t}(i)$ according to Eqn. 2 and the product of matrices is carried out in *pre*-multiplication order. Therefore, at the end of the k -th iteration, the effective transformation applied to the second scan is $\mathbf{M}(k)$.

The basic ICP iterations can now be carried out in two analogous ways : a) we repeatedly apply $\Delta\mathbf{M}$ to the second scan so as to maintain both scans in the same frame of reference, or b) we apply $\mathbf{M}(k)$ given in Eqn. 3 to the second scan in its *original* frame

of reference. As far as ICP is concerned, both these approaches are equivalent and the results would be identical. However, in our method, since we are interested in carrying out a motion averaging on all the relative motions at the k -th iteration, we use the latter approach, i.e. at the k -th iteration, we use $\mathbf{M}(k)$ in the *original* frame of reference instead of $\Delta\mathbf{M}(k)$. Given this set of relative motions i.e. $\mathbf{M}_E(k) = \{\mathbf{M}_{i_1j_1}(k), \mathbf{M}_{i_2j_2}(k), \dots, \mathbf{M}_{i_Sj_S}(k)\}$ computed at the k -th iteration, we average them using the Lie-algebraic averaging scheme described in [4]. This results in $\mathbf{M}_{global}(k) = \{\mathbf{I}, \mathbf{M}_2(k), \dots, \mathbf{M}_N(k)\}$ which is our estimate for the global motion at the k -th iteration. The crucial point to be noted here is that since we are interested in the motion averaged estimates to drive the registration process, once we estimate the global motion, all previously computed relative motions are discarded, i.e. $\mathbf{M}_{global}(k)$ replaces the information contained in $\mathbf{M}_E(k)$. Consequently, for the next iteration, the relative motion $\mathbf{M}_{ij}(k)$ is replaced by its motion averaged version, i.e. $\mathbf{M}_j(k)\mathbf{M}_i^{-1}(k)$. We can now state the complete **motion step** at the k -th iteration.

1. Compute all incremental relative motions $\Delta\mathbf{M}_E(k) = \{\Delta\mathbf{M}_{i_1j_1}(k), \dots, \Delta\mathbf{M}_{i_Sj_S}(k)\}$
2. Update all the relative motions in $\mathbf{M}_E(k)$ $\mathbf{M}_{i_sj_s}(k) = \Delta\mathbf{M}_{i_sj_s}(k)\mathbf{M}_{i_sj_s}(k-1) \forall s = \{1, \dots, S\}$
3. Compute $\mathbf{M}_{global}(k) = \{\mathbf{I}, \mathbf{M}_2(k), \dots, \mathbf{M}_N(k)\}$ by applying the Motion Averaging algorithm of Appendix A to all the relative motions in $\mathbf{M}_E(k)$ from step 2 above
4. Update all relative motions using the global motion estimate, i.e. $\forall s = \{1, \dots, S\} \mathbf{M}_{i_sj_s}(k) \leftarrow \mathbf{M}_{j_s}(k)\mathbf{M}_{i_s}^{-1}(k)$ where i_s and j_s represent the indices of the pair of scans represented by the s -th edge in E .

The multi-view ICP algorithm can now be stated as a repeated iteration using the **correspondence** and **motion** steps till some convergence criteria is satisfied.

4.1 Comparison with multi-view methods

In this Subsection we briefly compare some existing multi-view methods with our algorithm. An early method that carried out multi-view registration is by Bergevin *et. al.* [2] which was refined by [1]. Here, motion estimation in each iteration was a single round of updates of the motion estimate of each scan, i.e. \mathbf{M}_i in our notation. For the update of the i -th motion \mathbf{M}_i , every other motion in \mathbf{M}_{global} is held fixed, i.e. in Eqn. 1 we solve for \mathbf{M}_i by holding \mathbf{M}_j fixed $\forall j \neq i$. Given that this approach does not simultaneously solve for all motions, it results in a suboptimal solution. In [9], a set of relative motion estimates are obtained and then the scans are added one at a time to the set of registered scans by using the relevant correspondences and estimated relative motions and adjusting to distribute the errors. In [6], multi-view registration is carried out via a set of key points. Points from each scan are registered to these key points and in turn the key point locations are updated using the registered points. In [8], a robust multi-view cost function is minimised using conjugate gradient search. In [5] a registration error metric is minimised on the manifold of 3D rotations. However, this method is only applicable when the point matches are known and not usable in the scenario where the point matches need to be updated in each iteration. Finally, a graph representation of relative motions is the basis for the registration approach of [12]. This method makes strict, limiting assumptions on the graph in which each edge (i.e. relative motion \mathbf{M}_{ij}) is present in at least one graph cycle. Using such cycles, the errors are distributed across

the vertices of the cycle. Such an approach is not truly multi-view and does not distribute the errors in an optimal fashion. Instead it sub-optimally solves the registration problem over cycles and averages these results. Thus, although there are some multi-view approaches, our method is differentiated from them on two counts. In the first instance, the other multi-view methods do not make full use of all the relative motion information to simultaneously register the scans by updating correspondences and motion estimates. Secondly, none of these methods make use of the Lie-group structure of the motion group to carry out the averaging process in an optimal fashion as is done in [4] (summarised in Appendix A) and used in our method.

5. VARIANTS OF MOTION AVERAGED ICP

In this section we consider two variants of the motion averaging method proposed in Sec. 4. The first one is a causal version that only uses the relationship of a scan with previous ones in a sequence and the second method speeds up the computation of the motion averaged ICP by exploiting the transitive relationship for correspondence mappings.

5.1 Causal Motion Averaging

Although the method of motion averaging described in Sec. 4 takes advantage of all the redundant information available in a sequence of scans, it assumes that all scans are available before we carry out the registration process using this method. However, in certain situations like navigation using localisation and mapping (SLAM) or 3D modelling in real-time, we desire to build 3D representations in a causal manner. If the registration process has to respect the causality principle, we cannot take advantage of all the relative motion computations possible. Instead we are restricted to use the information available with respect to previous scans alone. Thus, for the k -th scan, we can only compute the relative motions \mathbf{M}_{ik} where $i < k$. However, as we shall describe now, even in such a restricted scenario it is possible to take advantage of the motion averaging principle while respecting causality.

Let us assume that we want to register the k -th scan when all previous scans have been registered. In the standard ICP scenario, we would compute the relative motion between scan k and $k - 1$, i.e. $\mathbf{M}_{k-1 k}$ and use it to register scan k . However, we note that even under the causality constraint, redundant information can be utilised. As shown in Fig. 2(a), if we assume that scan k shares overlapping areas with scans $k - 1$ and $k - 2$, then the registration of scan k can be carried out by simultaneously considering its relative motions with respect to the two previous views (shown by solid red edges). If we rewrite the motion averaging in Eqn. 1 as $\mathbf{M}_{ij}\mathbf{M}_i = \mathbf{M}_j$, then in the scenario of Fig. 2(a) we have two constraints available for the estimate of \mathbf{M}_k , i.e. $\mathbf{M}_k = \mathbf{M}_{k-2 k}\mathbf{M}_{k-2}$ and $\mathbf{M}_k = \mathbf{M}_{k-1 k}\mathbf{M}_{k-1}$. We can estimate \mathbf{M}_k by averaging these representations using the method for averaging on the $SE(3)$ group described in [4] (summarised in Appendix A). This estimation process can be incorporated into the **motion step** for \mathbf{M}_k . In this causal motion averaging approach, the estimate of \mathbf{M}_k is obtained by simultaneously registering scan k with the two previous scans. This results in a reduction in the accumulation of error as the scan k is ‘tethered’ to both the previously registered scans, thereby reducing drift in the estimate. In this approach using a greater number of previous scans for registration gives better results by the process of averaging of redundant information. It will also be noted that when scan k has overlap with the previous scan alone, our causal motion

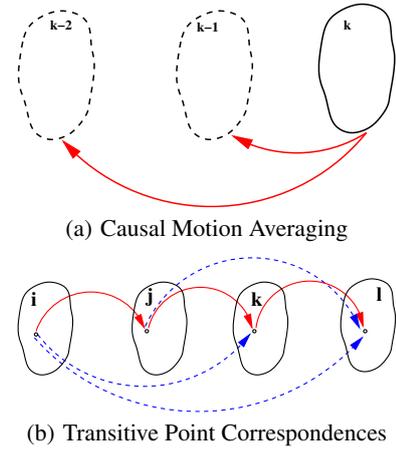


Figure 2: (a) In causal motion averaging, each scan is simultaneously registered to previous scans in the sequence via motion averaging. (b) For a single point in scan i , the arcs in solid red denote its correspondences in adjacent scans in a sequence. The arcs in dashed blue represent other correspondence mappings for the same point. The transitivity property of the correspondence mappings imply that the correspondences represented by the dashed blue arcs can be constructed out of the solid red ones resulting in reduced computational needs. See text for details.

averaging method boils down to the conventional ICP method. In other words, our approach results in a motion averaging registration method that takes full advantage of all information available in a causal manner and is a generalisation of the conventional sequential registration procedure that uses ICP.

5.2 Using Transitivity for Correspondences

The additional computation load due to the motion averaging step is insignificant compared to the correspondence step that is known to be the most significant computational load for ICP. As discussed earlier, motion averaging is useful since it allows us to average many relative motions to achieve a solution that is both accurate and also distributes the errors uniformly across all the scans. This advantage accrues to us when $S > (N - 1)$. In this sense, since in our correspondence step we compute all S correspondence mappings, it would seem that the multi-view motion averaged ICP algorithm stated above would involve a substantially increased computational load due to the additional sets of correspondences needed. However, we can simplify this computational load by using the following observations.

In general, for an iterative approach like ICP the exact motion estimates during each iteration do not matter as long as their cumulative effect is to lead the algorithm to the local minima. In our context, using this observation we wish to reduce the computational load due to the additional steps of solving for correspondences. The reduction in computational load is achieved by using the transitivity of the correspondence mappings as illustrated in Fig. 2(b). Consider a set of 4 ordered scans denoted as $\{i, j, k, l\}$. We consider a single point \mathbf{p} that has correspondences in all scans and denote this point in scan i as \mathbf{p}_i . Let the correspondence map from scans i to j be denoted as π_j^i , i.e. for a point in scan i , \mathbf{p}_i , the corresponding point in scan j is given as $\mathbf{p}_j = \pi_j^i(\mathbf{p}_i)$. To register the scans using ICP, we would solve for the correspondences between adjacent

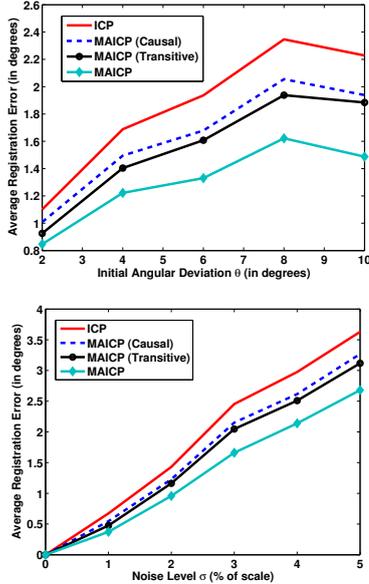


Figure 3: Average Registration Errors (in degrees) with respect to different initial rotation values (θ) and noise levels (σ). See Sec. 6.1 for explanation of the experimental setup and the definition of θ and σ .

views, i.e. between pairs $\{i, j\}$, $\{j, k\}$ and $\{k, l\}$. These mappings are shown in solid red in Fig. 2(b). For our case of motion averaging using all relative motions available, we would need to consider the correspondences between additional pairs of scans in the sequence, i.e. $\{i, k\}$, $\{i, l\}$ and $\{j, l\}$ which are denoted by dashed blue arcs in Fig. 2(b). As a result, it would seem that the computational load for establishing correspondences is doubled. However, if the correspondence mappings are correct, we make the crucial observation that such correspondence mappings are transitive. For instance, to establish the match in scan k for point \mathbf{p}_i we could use the path between \mathbf{p}_i and \mathbf{p}_k through \mathbf{p}_j , i.e.

$$\mathbf{p}_k = \pi_i^k(\mathbf{p}_i) = \pi_j^k(\mathbf{p}_j) = \pi_j^k(\pi_i^j(\mathbf{p}_i))$$

implying that $\pi_i^k = \pi_j^k \circ \pi_i^j$. Notice that both the mappings π_i^j and π_j^k are already computed for registering using ICP and therefore computing the correspondence map π_i^k does not incur any additional computational load. This observation can now be generalised to the following approach.

To globally register all scans in a single frame of reference we need $N - 1$ correspondence maps that form a spanning tree (ST) on the graph G (equivalent to the solid red arcs in Fig. 2(b)). All other correspondence maps that we require (i.e. the dashed blue ones) can be established by following the path between the pair of vertices on ST. Thus, given the ST, we establish all correspondence maps between adjacent vertices. All additional correspondence maps belong to paths that are more than one edge long on ST. Such correspondence maps are established by using the transitivity relationship along the path on ST that connects the two vertices. In addition, we note that a point correspondence can be ‘propagated’ along ST only to the extent that it has correspondences along the edges of ST. As a result, the total computational load for establishing correspondences is the same as in the case of basic ICP, i.e. $N - 1$ mappings and the additional load incurred in creating the

composition maps using the transitive relationship is small.

In the discussion above, we have assumed that the correspondence maps are transitive, i.e. $\pi_i^k = \pi_j^k \circ \pi_i^j$. However, this is only true if the correspondences maps are correct. In the general scenario of ICP, this is not true especially if the scans are noisy or if they are significantly far apart from each other. In such a case, $\pi_i^k \neq \pi_j^k \circ \pi_i^j$. However, we note that even in the case of the basic ICP algorithm, the correspondence maps are only approximate and are expected to improve after each iteration. In the basic ICP algorithm, while the correspondences are approximate we can expect the estimated correspondence to be within a small distance from the true corresponding point. Translating this to our scenario, when we take a series of finite transitive steps along the spanning tree, the effective error introduced is also bounded. An additional consideration can be illustrated by revisiting Fig. 2(b). If we consider the direct correspondence (in dashed blue) between \mathbf{p}_i and \mathbf{p}_l , the error could be substantial since the two views i and l are far apart. In contrast, each of the single steps (shown in solid red) from i to l , i.e. $i \rightarrow j$, $j \rightarrow k$ and $k \rightarrow l$ are between adjacent scans that are closer to each other. As a result each of these correspondence maps are expected to be reasonably accurate and their composition $\pi_l^i \circ \pi_j^k \circ \pi_i^j$ can be expected to be close to π_l^i , i.e. $\pi_l^i \circ \pi_j^k \circ \pi_i^j \approx \pi_l^i$.

6. RESULTS

In this Section we present results that compare the performance of the different motion averaged registration methods proposed in this paper and the basic ICP approach. We characterise the performance of our approach by using synthetic data. In addition, we present the results of applying our method to some well-known 3D scan datasets and also compare its performance with some other multi-view registration approaches.

6.1 Results on Synthetic Data

In order to characterise the different methods we need to study their comparative performance in relation to a) the amount of noise and b) the (initial) amount of rotation prior to registration. These variables are parametrised by σ and θ respectively. To enable us to focus on these factors, we use synthetic data for these experiments. We randomly sample 5000 points from the first scan of the Stanford Bunny dataset. Using this set of 5000 points we generate 4 additional point sets by applying a rotation to the points and adding Gaussian noise to the 3D co-ordinates of these points. This results in a total of 5 rotated and noisy point sets that we register using the various methods.

To simulate for common scanning scenarios we use two different types of rotations which we call ‘turntable’ and ‘random’ rotations. For ‘turntable’ rotations, the axis of rotation is fixed along the original axis of rotation for the scans. For a given choice of θ , the rotation between adjacent views is randomly chosen from a uniform distribution over $(0, \theta]$. In the case of ‘random’ rotations, the axis of rotation is also chosen randomly. In effect, for both types of rotation, each subsequent scan is increasingly further from the first one. To determine the impact of noise in the point locations we use the following approach. We denote the radius of the circumsphere of the object as r and the noise level σ is defined as a given percentage of r . In our experiments we vary θ as $\{2, 4, 6, 8, 10\}$ degrees and the noise-level σ as $\{1, 2, 3, 4, 5\}\%$ of r . For each combination of σ and θ , we run 25 trials for each type of motion and average the registration errors. Since the comparative performance for both the ‘turntable’ and ‘random’ rotations are very similar, we present re-

	ICP	Other Multi-view Methods		Multi-view Motion Averaged ICP			
		Benjemaa <i>et al</i> [1]	Sharp <i>et al</i> [12]	Closure	Causal	MAICP	Transitive
Happy Buddha	5.78	1.96	1.53	2.62	3.67	0.85	0.91
Stanford Bunny	0.92	0.61	0.61	0.78	0.72	0.59	0.55
Stanford Dragon	7.69	3.75	5.54	5.51	5.72	2.74	3.39
Ohio Bunny	1.55	0.78	0.69	0.27	1.18	0.28	0.47
Ohio Pooh	13.33	1.55	1.00	0.68	4.52	0.5	0.36

Table 1: Average Registration Errors for real datasets (in degrees). Our motion-averaged registration methods significantly outperform ICP and the other multiview registration methods listed. See text for details.

sults that are averaged over both these types of rotation. In Fig. 3 we show this average rotation error for all scans with respect to the first one which is the reference scan. The rotation error is calculated as the angular difference between the true rotation and the estimated value. As shown in the plots in Fig. 3, for each σ or θ these errors are averaged over the other variable. We denote the sequential ICP-based registration method as **ICP**, the multi-view motion averaging method of Sec. 4 as **MAICP**, the causal variant of Sec. 5.1 as **MAICP (Causal)** and the averaging method using transitive correspondences described in Sec. 5.2 as **MAICP (Transitive)**. In all experiments in this paper, **MAICP (Causal)** results are obtained by initialising with **ICP**. In turn, **MAICP** is initialised with the results of **MAICP (Causal)**. As is easily seen from Fig. 3, our method **MAICP** significantly outperforms the **ICP** algorithm. It can also be seen that the two variants of motion averaging also perform better than **ICP** and their error values are approximately in between that of **MAICP** and **ICP**.

6.2 Registration Accuracy for Real Datasets

In this Subsection we describe the results of our registration methods applied to some well-known datasets. Specifically, we demonstrate results on three datasets from the Stanford repository i.e. the Stanford Bunny, Happy Buddha (standing) and Dragon which have 10, 15 and 15 scans respectively.¹ While the Bunny scan set consists of varied viewpoints to cover the surface, the Happy Buddha and Dragon set are turntable sequences with consecutive scans taken 24° apart. For all the datasets in the Stanford repository, the ground truth motion is provided. In addition, we also show results for the registration of two sets from the Ohio State repository, i.e. the Ohio Bunny and Pooh which consist of 18 turntable scans that are 20° apart.² However, for these two datasets, the axis of rotation is not provided. For measuring the registration error, for the Stanford repository datasets we compute the angular difference between the ground truth rotation and the estimated rotation matrix. In the case of the Ohio State repository datasets, since the ground truth axis of rotation is not known, the angular error is simply the difference between the ground truth angle and the angle of the estimated rotation matrix (i.e. independent of the axes of rotation).

For the experiments on the three Stanford datasets, we average the results over 25 trials using a protocol as follows. To evaluate the performance of ICP and our multi-view methods we use

¹<http://graphics.stanford.edu/data/3Dscanrep>.

²<http://sampl.ece.ohio-state.edu/data/3DDB/RID/index.htm>.

the registered set of scans and perturb them by rotations about the turntable axis. For each trial, the rotation angle is randomly drawn from a uniform distribution over $(0, 5^\circ]$. For estimating the motion between two scans we take 1000 randomly selected points from the first scan and find their correspondences on the second scan. In addition to the different methods of registration detailed above, we consider an additional averaging scheme that demonstrates the power of our method. As described in Sec. 3, in a turntable sequence, the last scan often has a substantial area of overlap with the first one and hence their relative motion can be estimated (see Fig. 1(b)). Typically, due to sequential registration using ICP there is a significant accumulation of error in the registration as more scans are added. In our experiments, we test the simplest possible motion averaging multi-view registration scheme, i.e. by ‘closing the loop’. Thus, along with the edges between adjacent scans (solid red ones in Fig. 1(b)) we also consider the additional dashed blue edge between the last scan and the first one which closes the loop along the sequence. Thus, instead of $(N - 1)$ edges of the spanning tree, we have N edges in this case. This method is denoted as **MAICP (Closure)**. In the case of the two datasets from the Ohio repository, since we do not have the ground truth motions, we simply register the scans with identity matrices as our initial motion estimates and then carry out the other multi-view registration methods. Thus, the error rates are for a single experiment and not averaged over multiple trials as in the case of the Stanford datasets.

In Table. 1, we give the comparative performance of all of our methods on these datasets. We also compare our results with that obtained with the multi-view methods due to Benjemaa *et al* [1] and Sharp *et al* [12]. We notice that **MAICP (Causal)** performs poorly compared to the other multi-view methods since it does not have all the multi-view constraints available for averaging. This implies that it is crucial to utilise all the available motion constraints to solve for the global motions required for registration. This view is borne out by the superior performance of all other **MAICP** variants. As can be seen, even the addition of a single motion constraint in **MAICP (Closure)** reduces the error by a large amount since this constraint forces the errors to be distributed uniformly over the sequence thereby preventing drift in registration that is common with the sequential registration using ICP. The basic method of **MAICP** does better compared to the other multiview approaches considered since it utilises all available information in the scan sequence. It is also of interest to note that **MAICP (Transitive)** which uses only the correspondence maps obtained by the spanning tree of the ICP method is very close in performance to the solution of **MAICP**. This implies that the assertion about the transitivity of correspon-

dences in Sec. 5.2 is correct and we can obtain the simultaneous multi-view registration solution by motion averaging with only a marginal increase in the computational load when compared with the basic ICP approach. Interestingly, for some of the datasets, the method of generating correspondences using transitivity does better than that of **MAICP** which is probably due to the accuracy of the transitive mappings over adjacent scans. While the **MAICP (Transitive)** method is as good as **MAICP**, it will be recalled that the transitive method does not perform as well in the experiments of Sec. 6.1 (See Fig. 3). This is because in the experiments of Sec. 6.1 we use 5000 points throughout and do not use the full or subsampled scans to determine correspondences. As a result, the individual correspondence maps are inaccurate resulting in poorer compositions. If we use the full or subsampled scans to search for correspondences, the **MAICP (Transitive)** method is seen to be almost as accurate as **MAICP** as evidenced by the experiments in this Section.

In Figs. 4, 5 and 6, we present a comparison of the results of different methods applied to the Stanford Happy Buddha, Stanford Dragon and the Ohio State Pooh scans respectively. The results of a single trial are shown in the form of cross-sections and the corresponding regions on the 3D models are indicated. As is evident, our method accurately registers the different scans for all the three datasets. Due to its sequential nature, the ICP method does not work well in any of the cases and results in significant misalignments. In the Happy Buddha dataset in Fig. 4, it will be noted that the method of [1] also does poorly as it is unable to distribute the errors well enough. Similarly in the case of the Dragon dataset, the method of [12] (Fig. 5(e)) shown a significant error in registering some of the scans. However, in both the cases of the Happy Buddha and Dragon datasets, our motion averaging scheme works well since it uses all the relative motion constraints in a single motion averaging step thereby forcing the errors to be evenly distributed over the entire scan sequence. In the case of the Pooh dataset, all the multi-view methods work well and their quality of registration is virtually indistinguishable in Fig. 5. However, as can be clearly seen in the row in Table. 1 corresponding to this dataset, our algorithm performs substantially better than the other methods.

7. CONCLUSION

In this paper we have introduced a multi-view extension of the ICP algorithm that uses the redundant information in a set of scans via motion averaging. We have also introduced two significant variants of this approach and demonstrated the utility of our multi-view formulation. Our approach performs significantly better than existing multi-view registration methods. Future work will involve computational refinements that exploit the possibilities available in an algorithmic design space using the transitivity property of correspondences and the redundancy available in the motion averaging framework.

8. REFERENCES

- [1] R. Benjema and F. Schmitt. Fast global registration of 3d sampled surfaces using a multi-z-buffer technique. *International Conference on 3D Digital Imaging and Modeling*, pages 113–120, 1997.
- [2] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau. Towards a general multi-view registration technique. *PAMI*, 18(5):540–547, 5 1996.
- [3] P. Besl and N. McKay. A method for registration of 3-d shapes. *PAMI*, 14(2):239–256, February 1992.

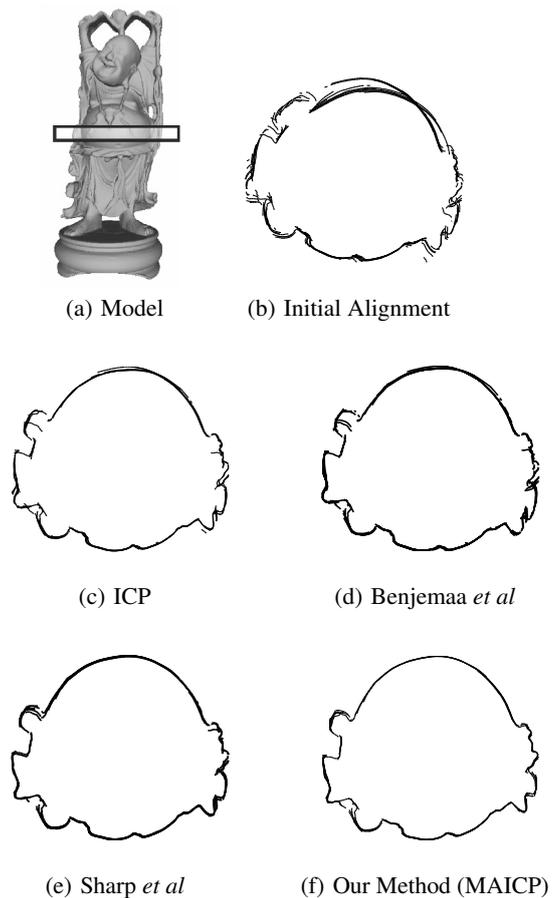


Figure 4: Cross-section of the registered Happy Buddha scans for the different methods. See text for details.

- [4] V. M. Govindu. Lie-algebraic averaging for globally consistent motion estimation. In *CVPR04*, pages I: 684–691, 2004.
- [5] S. Krishnan, P. Lee, J. Moore, and S. Venkatasubramanian. Global registration of multiple 3d point sets via optimization-on-a-manifold. In *Eurographics Symposium on Geometry Processing*, 2005.
- [6] T. Masuda. Generation of geometric model by registration and integration of multiple range images. In *3DIM*, pages 254–261, 2001.
- [7] G. Medioni and Y. Chen. Object modeling by registration of multiple range images. In *CRA91*, pages 2724–2729, 1991.
- [8] K. Nishino and K. Ikeuchi. Robust simultaneous registration of multiple range images. In *ACCV02*, 2002.
- [9] K. Pulli. Multiview registration for large data sets. In *3DIM99*, pages 160–168, 1999.
- [10] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *3DIM01*, pages 145–152, 2001.
- [11] J. Salvi, C. Matabosch, D. Fofi, and J. Forest. A review of recent range image registration methods with accuracy evaluation. *IVC*, 25:578–596, 2007.
- [12] G. Sharp, S. Lee, and D. Wehe. Multiview registration of 3d scenes minimizing error between coordinate frames. *PAMI*, 26(8):1037–1050, 8 2004.
- [13] S. Umeyama. Least-squares estimation of transformation

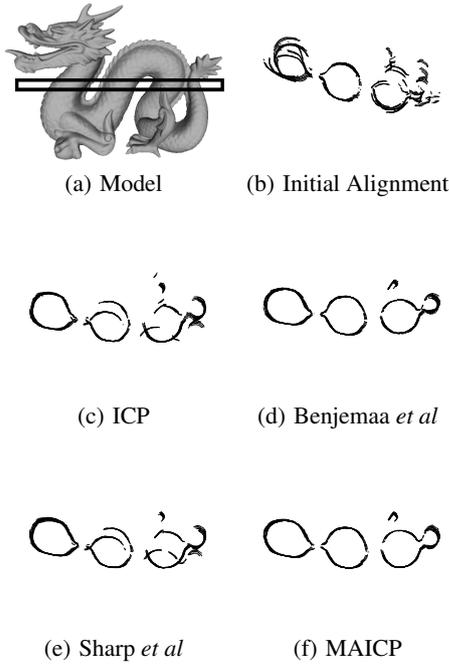


Figure 5: Cross-section of the registered Dragon scans for the different methods. See text for details.

parameters between two point patterns. *PAMI*, 13(4):376–380, 1991.

APPENDIX

A. LIE-ALGEBRAIC AVERAGING

In this Appendix we briefly summarise the motion averaging approach of [4] that is used by our method. The analysis presented applies to all matrix Lie-groups. Given a set of rotation matrices $\{\mathbf{R}_1, \dots, \mathbf{R}_N\}$, their arithmetic mean is not necessarily a valid rotation. However, if we consider these rotations as Lie-group elements, a suitable definition of the mean is given by $\bar{\mathbf{R}}$ which minimises the variance $\sum_{i=1}^N d^2(\mathbf{R}_i, \bar{\mathbf{R}})$, where $d(\cdot, \cdot)$ is the Riemannian distance on the $SO(3)$ group. For such a minimisation, an efficient solution is obtained by successively approximating the Riemannian distance on the tangent space of the Lie-group, i.e. its corresponding Lie-algebra. Thus, the average of matrix Lie-group elements can be obtained using :

A1 : Algorithm for Intrinsic Average

Input : $\{\mathbf{M}_1, \dots, \mathbf{M}_N\} \in G$ (Matrix Group)

Output : $\mu \in G$ (Intrinsic Average)

Initialise : $\mu = \mathbf{I}$ (Identity)

Do

$$\Delta \mathbf{M}_i = \mu^{-1} \mathbf{M}_i$$

$$\Delta \mathbf{m}_i = \log(\Delta \mathbf{M}_i)$$

$$\Delta \mu = \exp\left(\frac{1}{N} \sum_{i=1}^N \Delta \mathbf{m}_i\right)$$

$$\mu = \mu \Delta \mu$$

Repeat till $\|\Delta \mu\| < \epsilon$

where $\log(\cdot)$ and $\exp(\cdot)$ are matrix operations and \mathfrak{m} is the Lie-algebra of \mathbf{M} . Using similar arguments, we can successively approximate the relative motion relationship, using the relationships

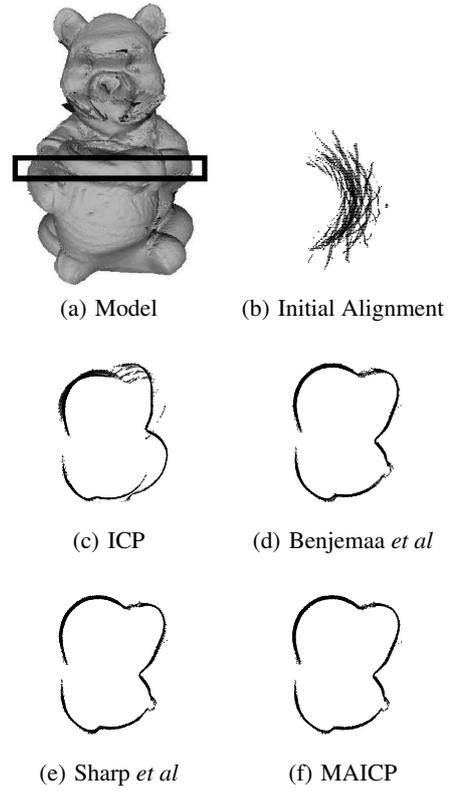


Figure 6: Cross-section of the registered Pooh scans for the different methods. See text for details.

$$\begin{aligned} \mathbf{M}_{ij} &= \mathbf{M}_j \mathbf{M}_i^{-1} \Rightarrow \mathbf{m}_{ij} = \mathbf{m}_j - \mathbf{m}_i \\ \Rightarrow \mathbf{v}_{ij} &= \mathbf{v}_j - \mathbf{v}_i \Rightarrow \mathbf{v}_{ij} = \underbrace{[\dots - \mathbf{I} \dots \mathbf{I} \dots]}_{\mathbf{D}_{ij}} \mathfrak{V} \end{aligned}$$

where \mathbf{v} is a vector extracted from \mathfrak{m} representing the independent variables that define the Lie-algebra \mathfrak{m} . \mathfrak{V} contains all the variables \mathbf{v}_i stacked together into a single vector. By collecting all the relative motion constraints columnwise, we have $\mathbb{V}_{ij} = [\mathbf{v}_{ij1}; \mathbf{v}_{ij2}; \dots]$ and $\mathbf{D} = [\mathbf{D}_{ij1}; \mathbf{D}_{ij2}; \dots]$ which leads to the solution to the motion averaging in the tangent space as $\mathfrak{V} = \mathbf{D}^\dagger \mathbb{V}_{ij}$, where \mathbf{D}^\dagger is the pseudo-inverse of \mathbf{D} . This leads to the following Lie-algebraic averaging algorithm :

A2 : Algorithm for Motion Averaging

Input : $\{\mathbf{M}_{ij1}, \mathbf{M}_{ij2}, \dots, \mathbf{M}_{ijn}\}$ (n relative motions)

Output : $\mathbf{M}_{global} : \{\mathbf{I}, \mathbf{M}_2, \dots, \mathbf{M}_N\}$ (N image global motion)

Set \mathbf{M}_{global} to an initial guess

Do

$$\Delta \mathbf{M}_{ij} = \mathbf{M}_j^{-1} \mathbf{M}_{ij} \mathbf{M}_i$$

$$\Delta \mathbf{m}_{ij} = \log(\Delta \mathbf{M}_{ij})$$

$$\Delta \mathbf{v}_{ij} = \text{vec}(\mathbf{m}_{ij})$$

$$\Delta \mathfrak{V} = \mathbf{D}^\dagger \Delta \mathbb{V}_{ij}$$

$$\forall k \in [2, N], \mathbf{M}_k = \mathbf{M}_k \exp(\Delta \mathbf{m}_k)$$

Repeat till $\|\Delta \mathfrak{V}\| < \epsilon$

where $\text{vec}(\cdot)$ extracts the parameters \mathbf{v} from \mathfrak{m} . The reader should consult [4] for further details.