ITERATIVE CLOSEST POINT ALGORITHM

Arundhati Banerjee
14CS30043

September 3, 2017

1 Introduction

The ICP (Iterative Closest Point) algorithm was introduced originally in [1] as the basis for a general purpose, representation-independent method for the accurate and computationally efficient registration of 3-D shapes, including free-form curves and surfaces. It can handle the full six degrees of freedom, requiring only a procedure to find the closest point on a geometric entity to a given point. ICP has therefore become the dominant method for aligning 3-D models based purely on the geometry, and sometimes color, of the meshes. The algorithm is widely used for registering the outputs of 3D scanners, which typically only scan an object from one direction at a time.

Given 3-D data in a sensor coordinate system, which describes a data shape that may correspond to a model shape, and given a model shape in a model coordinate system in a different geometric shape representation, the idea is to estimate the optimal rotation and translation that aligns, or registers, the model shape and the data shape minimizing the distance between the shapes and thereby allowing the determination of the equivalence of the shapes via a mean-square distance metric. The proposed shape registration algorithm in [1], can be used with the following representations of geometric data:

1. Point sets
2. Line segment sets (polylines)
3. Implicit curves:
$$\vec{g}(x, y, z) = 0$$
4. Parametric curves: (x(u), y(u), z(u))
5. Triangle sets (faceted surfaces)
6. Implicit surfaces: g(x,y,z) = 0
7. Parametric surfaces: (x(u,v), y(u,v), z(u,v))

Other representations can also be handled by providing a procedure for evaluation of the closest point on the given shape to a given digitized point.

2 The ICP algorithm

Let $P = \vec{p}_i$ be a measured data point set to be aligned with the model point set $X = \vec{x}_i$, where $N_x = N_p$ and each $\vec{p}_i$ corresponds to each $\vec{x}_i$ with the same index.

Let $\mathbf{R}$ be the 3x3 rotation matrix generated by the unit quaternion $\vec{q}_R = [q_0 q_1 q_2 q_3]^t$ and $\vec{q}_T = [q_4 q_5 q_6]^t$ be a translation vector. The complete registration state vector is $\vec{q} = [\vec{q}_R | \vec{q}_T]^t$. The mean square objective function to be minimized can be now defined as

$$f(\vec{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\vec{x}_i - \mathbf{R}(\vec{q}_R)\vec{p}_i - \vec{q}_T\|^2. \tag{1}$$

The distance metric d between an individual data point $\vec{p}$ and the model shape X is denoted as

$$d(\vec{p}, X) = \min_{\vec{x} \in X} \|\vec{x} - \vec{p}\|. \tag{2}$$

The closest point in X at the minimum distance is denoted as $\vec{y}$ such that $d(\vec{p}, \vec{y}) = d(\vec{p}, X)$ where $\vec{y} \in X$. Let Y be the set of closest points in X for all points in P and C be the closest point operator

$$Y = C(P, X). \tag{3}$$

Using this corresponding set Y, the least squares registration is computed as described above:

$$(\vec{q}, d) = Q(P, Y). \tag{4}$$

The positions of the points in the data shape can then be updated as $P = \vec{q}(P)$.

The ICP algorithm is stated as follows:

- The point set P from the data shape and the model shape X are given. The data shape and model shape can be represented in any one of the allowable forms discussed. For our purpose however, we consider the data shape to be decomposed into a point set with N_p points $\vec{p}_i$, and X having N_x supporting geometric primitives - points, lines or triangles.
- The iteration is initialized by setting $P_0 = P$, $\vec{q_0} = [1, 0, 0, 0, 0, 0, 0]^t$ and $k = 0$. Until convergence within a tolerance τ, the following steps are repeated:
  - Compute $Y_k = C(P_k, X)$
  - Compute the registration state vector: $(\vec{q}_k, d_k) = Q(P_k, Y_k)$
  - Apply $P_{k+1} = \vec{q}_k(P_0)$, assuming the registration vectors are defined relative to P_0
  - Terminate if desired precision is reached: $d_k - d_{k+1} < \tau, \tau > 0$

It can be proved that the iterative closest point algorithm always converges monotonically to a local minimum with respect to the mean-square distance objective function. Given $P_k = \vec{p}_{ik} = \vec{q}_k(P_0)$ and X, $Y_k = \vec{y}_{ik}$ can be computed as described before. The correspondence error can then be defined as

$$e_k = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\vec{y}_{ik} - \vec{p}_{ik}\|^2 \tag{5}$$

and the alignment error as

$$d_k = \frac{1}{N_p}\sum_{i=1}^{N_p}\|\vec{y}_{ik} - \mathbf{R}(\vec{q}_{kR})\vec{p}_{i0} - \vec{q}_{kT}\|^2. \tag{6}$$

It is always true that $d_k \leq e_k$, otherwise the identity transformation on the point set would give a smaller mean square error than the least squares registration, which cannot be true. Now, applying the registration $\vec{q}_k$ to get P_{k+1} from P_0, if the previous correspondence Y_k was still maintained, then the mean square error would still be d_k, that is:

$$d_k = \frac{1}{N_p}\sum_{i=1}^{N_p}\|\vec{y}_{ik} - \vec{p}_{i,k+1}\|^2. \tag{7}$$

However, since the new corresponding point set Y_{k+1} is obtained by the closest point operator C, so,

$$\|\vec{y}_{i,k+1} - \vec{p}_{i,k+1}\| \leq \|\vec{y}_{i,k} - \vec{p}_{i,k+1}\| \tag{8}$$

for all $i = 1(1)N_p$. Hence, $e_{k+1} \leq d_k$. So, we get the inequality:

$$0 \leq d_{k+1} \leq e_{k+1} \leq d_k \leq e_k \text{ for all k.} \tag{9}$$

Since the MSE sequence is non-increasing and bounded below, so the ICP algorithm must converge monotonically to a minimum value. Each iteration includes the 3 main steps - finding the closest points, which is $O(N_pN_x)$ in the worst case, and $O(N_p \log N_x)$ on the average; computing the registration with cost $O(N_p)$; and applying the registration with cost $O(N_p)$. The iterative in ICP comes from the fact that the correspondences are reconsidered as the solution comes closer to the error local minimum. As any gradient descent method, the ICP is applicable when we have a relatively good starting point in advance since it might otherwise be trapped into the first local minimum.

The key benefit of the ICP is that the convergence is fast and monotonic. No expensive closest point evaluations are spent on registration vectors that have worse MSEs than the current state. Because of the ICP convergence theorem, it is not needed to search around blindly in the multidimensional space to determine the direction in which to move.

3 Variants of the basic ICP algorithm

Several variants of the basic ICP algorithm have been proposed based on the following stages:

- Selecting source points
- Matching to points in the other mesh
- Weighting the correspondences
- Rejecting certain (outlier) point pairs
- Assigning an error metric to the current transform

- Minimizing the error metric with respect to the transformation

Their respective performances are assessed based on factors like speed, stability, tolerance to noise and/or outliers, and maximum initial misalignment.

3.1 Selection of Points

Some of the strategies proposed for selection of point pairs include always using all available points, uniform subsampling of the available points, random sampling having a different sample of points at each iteration, feature based sampling, as well as normal space sampling. Each of these schemes may select points from only one mesh, or from both meshes. In normal space sampling, the points are chosen such that the distribution of normals among selected points is as large as possible. The motivation for this strategy is the observation that for certain kinds of scenes (such as the "incised plane" data set) small features of the model are vital to determining the correct alignment. A strategy such as random sampling will often select only a few samples in these features, which leads to an inability to determine certain components of the correct rigid-body transformation. Thus, one way to improve the chances that enough constraints are present to determine all the components of the transformation is to bucket the points according to the position of the normals in angular space, then sample as uniformly as possible across the buckets. Normal-space sampling is therefore a very simple example of using surface features for alignment; it has lower computational cost, but lower robustness, than traditional feature-based methods. An extension to this technique is the point selection strategy called stable sampling which improves the geometric stability of the ICP algorithm. It is aimed at sampling those features of the input meshes which provide the best convergence of the algorithm to the correct pose. The sampling strategy is therefore based on estimating the transformations that can cause unstable sliding in the ICP algorithm and picking points that best constrain this sliding.

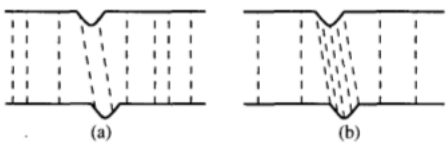

Figure 1: Corresponding point pairs selected by the (a) random sampling and (b) normal space sampling strategies for an incised mesh. Using random sampling, the sparse features may be overwhelmed due to the presence of noise or distortion, causing the ICP algorithm to not converge to a correct alignment.

3.2 Matching Points

This has the greatest effect on the convergence and speed of the ICP algorithm. The following are some of the strategies that have been proposed as variants for the correspondence finding stage.

- Finding the closest point in the other mesh. It is generally stable, but slow and requires pre-processing. However, the computation may be accelerated using a k-d tree and/or closest point caching.

- Normal shooting, i.e.finding the intersection of the ray originating at the source point in the direction of the source point's normal with the destination surface. It performs slightly better than the closest point finding method for smooth structures, but worse for noisy or complex structures.

- Reverse calibration, i.e. projecting the sample point onto the destination mesh, from the point of view of the destination meshś range camera. Although the projection algorithm does not offer the best convergence per iteration, each iteration is faster than an iteration of closest point finding or normal shooting because it is performed in constant time, rather than involving a closest-point search (which, even when accelerated by a k-d tree, takes $O(\log n)$ time). As a result, the projection-based algorithm has a significantly faster rate of convergence vs. time.

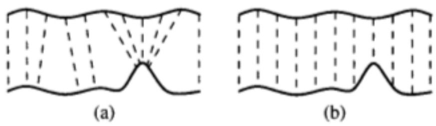

Figure 2: (a) In the presence of noise and outliers, the closest-point matching algorithm potentially generates large numbers of incorrect pairings when the meshes are still relatively far from each other, slowing the rate of convergence. (b) The projection matching strategy, in comparison, is less sensitive to the presence of noise.

- Any of the above methods, restricted to only matching points compatible with the source point according to a given metric. Compatibility metrics based on color and angle between normals have been explored.

3.3 Weighting of Point Pairs

The corresponding point pairs determined can be assigned different weights according to different algorithms.

- Constant weight assigned to all point pairs.
- Assigning lower weights to pairs with greater point-to-point distances.
- Weighting based on compatibility of normals, or even compatibility of colours.
- Weighting based on the expected effect of scanner noise on the uncertainty in the error metric.

It is observed that even with the addition of extra noise, all of the weighting strategies have similar performance, with the uncertainty and compatibility of normals options having better performance than the others.

Selection vs. Weighting : The same effect could be achieved with weighting as with selection of points. However, it is often hard to guarantee that enough samples of important features will be present, except at high sampling rates. Besides, weighting strategies turn out to be dependent on the data. The preprocessing/run-time cost tradeoff should also be considered in determining how to find the correct weights.

3.4 Rejecting Point Pairs

The purpose of this stage is usually to eliminate outliers, which may have a large effect when performing least-squares minimization. Some rejection strategies include rejecting corresponding points more than a given (user-specified) distance apart, rejecting pairs whose point-to-point distance is larger than some multiple of the standard deviation of distances, and even rejection of pairs containing points on mesh boundaries. The latter strategy, of excluding pairs that include points on mesh boundaries, is especially useful for avoiding erroneous pairings (that cause a systematic bias in the estimated transform) in cases when the overlap between scans is not complete. The Trimmed ICP (TrICP) algorithm involves sorting all correspondences with respect to their error metric and deleting the worst $t\%$, where t is to be estimated with respect to the overlap. TrICP is fast, applicable to overlaps under 50%, robust to erroneous measurements and shape defects, and has easy-to-set parameters. ICP is a special case of TrICP when the overlap parameter is 100%.

Figure 3: (a) When two meshes to be aligned do not overlap completely, allowing correspondences involving points on mesh boundaries can introduce a systematic bias into the alignment. (b) Disallowing such pairs eliminates many of these incorrect correspondences.

In general, it is observed that even though outlier rejection may have effects on the accuracy and stability with which the correct alignment is determined, it generally does not improve the speed of convergence.

3.5 Error metric and minimization algorithm

- **Sum of squared distances between corresponding points:** For an error metric of this form, there exist closed-form solutions for determining the rigid-body transformation that minimizes the error. The solution-methods proposed are based on Singular Value Decomposition, quaternions, orthonormal matrices and dual quaternions.

- **Point to plane metric** takes into account the sum of squared distances from each source point to the plane containing the destination point and oriented perpendicular to the destination normal at that point. The error metric can be defined as:

$$E := \sum_{i=1}^{N_p} ((\mathbf{R}\vec{p}_i + \vec{t} - \vec{x}_i) \cdot \vec{n}_i)^2 \tag{10}$$

where R is the rotation matrix and $\vec{t}$ is the translation vector, $\vec{n}_i$ is the unit normal vector to model shape X at $\vec{x}_i$ which is the corresponding point to $\vec{p}_i$ of the data point set. In this point-to-plane case, no closed-form solutions are available. The least-squares equations may be solved using a generic non-linear method (e.g. Levenberg-Marquardt), or by simply linearizing the problem (i.e., assuming incremental rotations are small, so $\sin\theta \sim \theta$ and $\cos\theta \sim 1$). Using point-to-plane distance instead of point-to-point lets flat regions slide along each other, and therefore leads to better performance.

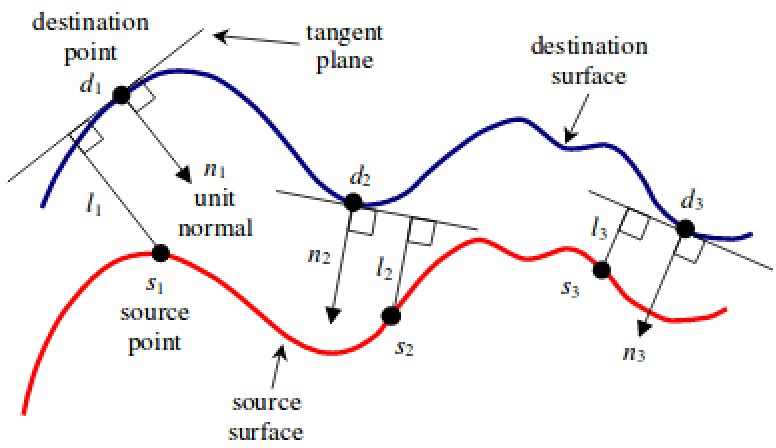

Figure 4: Illustration of the point-to-plane metric computation.

4 Advantage and Disadvantage

The primary advantages of the basic ICP algorithm are as follows:

- It handles the full six degrees of freedom and is independent of shape registration.
- It does not require pre-processing of 3-D point data, such as smoothing, as long as the number of statistical outliers is near zero.
- It does not require any derivative estimation or any local feature extraction.
- Global matching is achieved at a predictable cost based on shape complexity. Local matching is also achieved at predictable cost based on shape complexity and the percentage of allowable occlusion.

- It can handle a reasonable amount of normally distributed vector noise, with standard deviations of upto 10% of object size.
- The method generalizes to n dimensions and is relatively insensitive to minor data segmentation errors.

The following are some disadvantages of the basic ICP algorithm:

- It is susceptible to gross statistical outliers unless a statistically robust method is employed either in preprocessing or regisration computation for outlier detection.
- The cost of local matching can get quite large for small allowable occlusions, e.g., 10% or less.
- The generalization to matching deformable models with high order deformations is not straightforward without, e.g., enumerating a dense set of possible deformations.
- For any given fixed initial set of rotations, the global shape matching capability can be defeated even without sensor noise by constructing sea urchin or planetoid shaped objects based on the set of rotations such that the correct registration cannot be guaranteed.

5 Analysis

The github repository includes an implementation of the ICP algorithm using SVD-based optimisation of the MSE error metric to obtain the rotation and translation matrices. It is tested for the cases of point set matching and curve matching. In figure 5, the data point set having 8 points is to be matched with the model point set containing 11 points. In this case, no initial rotation or translation state was specified.

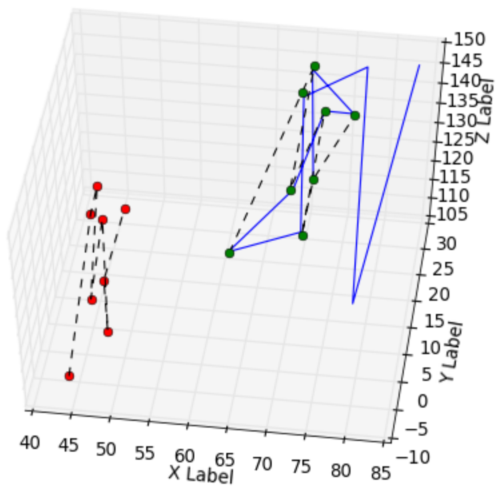

Figure 5: The data point set, in red, is aligned with the model set, in, blue. The data point set is shown in green after the registration is complete.

In order to test for curve matching, the model data points were initially synthesized to generate the 3-D space curve. A copy of it was translated and rotated to be relatively difficult to match. Each x, y, and z component of each point in this set was then corrupted by a zero-mean Gaussian noise with a unit standard deviation.

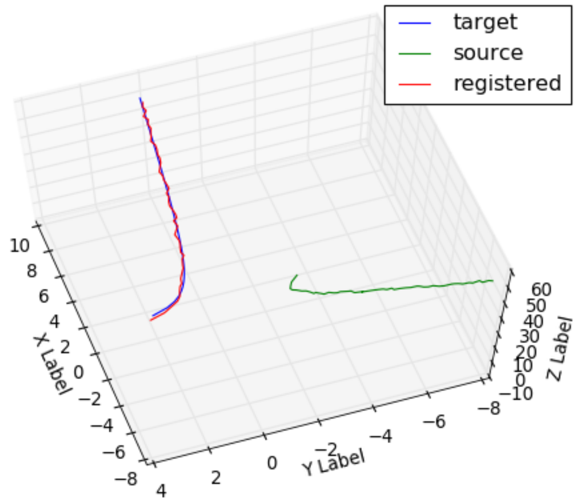

Figure 6: The target and source curves are shown before registration in the 3-D space. The curve in red is after the ICP registration is applied to the source curve.

It is observed that the registration to produce the source curve from the target curve when post-multiplied with the final registration obtained by the ICP algorithm, we obtain a matrix close to the identity matrix except for the effects of noise.

6 Conclusion

The ICP algorithm and its variants find a large number of applications in tasks such as improving spatial, range and intensity estimates of imagery collected during urban terrain mapping, 3-D surface reconstruction, hand tracking as well as pose estimation. The ability to have ICP execute in real time (e.g., at video rates) would permit significant new applications in computer vision and graphics. Several such methods have been proposed in different use case scenarios as mentioned in [4],[5],[6]. Algorithms that switch between variants, depending on the local error landscape and the probable presence of local minima might also provide increased robustness.

7 References

[1] Paul J. Besl and Neil D. McKay. A method for the registration of 3-D shapes. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[2] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *International Conference On 3-D Digital Imaging and Modeling*, 2001.

[3] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *Proc. IEEE Conf. on Robotics and Automation*, 1991.

[4] Tagliasacchi et al. Robust Articulated-ICP for Real-Time Hand Tracking. In *Eurographics Symposium on Geometry Processing*, 2015.

[5] M. Nießner et al. Combining Inertial Navigation and ICP for Real-time 3D Surface Reconstruction. In *EUROGRAPHICS 2014*.

[6] S. Marden and J. Guivant. Improving the Performance of ICP for Real-Time Applications using an Approximate Nearest Neighbour Search. In *Proceedings of Australasian Conference on Robotics and Automation*, 2012.

[7] D. Chetverikov, D. Svirko, D. Stepanov, and Pave1 Krsek. The Trimmed Iterative Closest Point Algorithm. In *Proc. 16th International Conference on Pattern Recognition*, 2002.

[8] K. Pulli. Multiview registration for large datasets. In *International Conference On 3-D Digital Imaging and Modeling*, 1999.

[9] G. Godin, M. Rioux, and R. Baribeau. Three-dimensional registration using range and intensity information. In *Proceedings of SPIE videometric III (Vol. 2350, pp. 279-290)*, 1994.

[10] N. Gelfand et al. Geometrically stable sampling for the ICP algorithm. In *3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. Fourth International Conference on. IEEE*, 2003.