

Simultaneous Covariance Driven Correspondence (CDC) and Transformation Estimation in the Expectation Maximization Framework

Michal Sofka, Gehua Yang, and Charles V. Stewart

Department of Computer Science

Rensselaer Polytechnic Institute

Troy, New York 12180-3590 *

{sofka,yang2,stewart}@cs.rpi.edu

<http://www.vision.cs.rpi.edu/CDC/>

Abstract

This paper proposes a new registration algorithm, Covariance Driven Correspondences (CDC), that depends fundamentally on the estimation of uncertainty in point correspondences. This uncertainty is derived from the covariance matrices of the individual point locations and from the covariance matrix of the estimated transformation parameters. Based on this uncertainty, CDC uses a robust objective function and an EM-like algorithm to simultaneously estimate the transformation parameters, their covariance matrix, and the likely correspondences. Unlike the Robust Point Matching (RPM) algorithm, CDC requires neither an annealing schedule nor an explicit outlier process. Experiments on synthetic and real images using a polynomial transformation models in 2D and in 3D show that CDC has a broader domain of convergence than the well-known Iterative Closest Point (ICP) algorithm and is more robust to missing or extraneous structures in the data than RPM.

1. Introduction

The circular dependence between point correspondences and transformation parameters is at the heart of many problems in registration. A common approach to addressing this problem is to start from an initial estimate of the transformation parameters, use this to generate an initial set of correspondences, re-estimate the parameters, and iterate. The iterative closest point (ICP) algorithm [2, 5], along with its variants [9, 10, 14, 17, 19], is a widely used incarnation of this idea. An important limitation of such algorithms is their narrow domain of convergence. Figure 1 (a) shows an example of this using synthetic points sampled from lines in the shape of an 'H'. From the configuration shown, ICP is unable to align the point sets correctly. The cause of the problem is two-fold. First, for moving points along the hor-

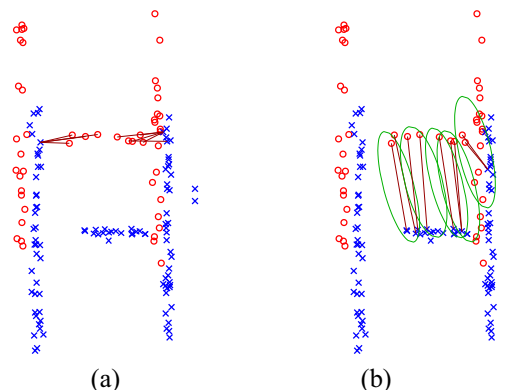


Figure 1. Difficulties in correspondence estimation and alignment of two 'H' shapes. Points on the moving 'H' are shown as red circles, while points on the fixed 'H' are shown blue 'x's. In (a), ICP correspondences are shown for points on the cross-bar of the 'H' as red line segments. These mismatches prevent ICP from properly aligning the shapes. When alignment uncertainty is properly accounted for (b) these same points preferentially establish correspondence with points on the opposing cross-bar.

izontal line of the 'H', the closest fixed points are along one of the two vertical lines instead of the horizontal line. This prevents formation of constraints that "pull" the transformation of the moving 'H' in the desired direction, down on top of the fixed 'H'. Second, noise in the locations and surface normals of points along the sides of the 'H' produces constraints that resist changes in the transformation, again preventing the desired vertical movement of the mapping.

An alternative approach embodied by the Robust Point Matching Algorithm (RPM) [6] starts by initially considering all possible matches, and then, using an annealing schedule, gradually moves toward unique correspondences as the transformation is refined. By virtue of this, RPM effectively aligns many cases on which ICP fails. It works poorly, however, when there are missing or extraneous structures (Figure 2(a)). In the early stages of its computation RPM tends toward aligning the centers of mass of the points, producing a bias in the estimate when there are extraneous structures that can not be overcome in later stages

*This article was supported by the DOD and the Medical University of South Carolina under DOD Grant No. W81XWH-05-1-0378. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Department of Defense or the Medical University of South Carolina.

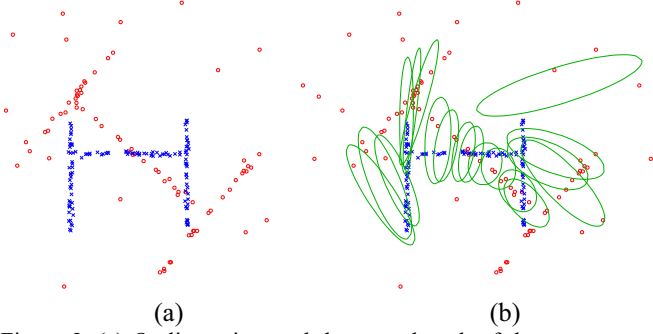


Figure 2. (a) Outlier points and the extra length of the top arms of the moving ‘H’ shape (red circles) contribute constraints that prevent RPM [6] from properly aligning these two shapes. The CDC correspondence uncertainties — a subset of the uncertainty ellipses is shown in (b) — initially include constraints from these unmatched points, but as the algorithm moves toward the correct estimate, these uncertainty ellipses will shrink, causing all correspondences for these points to be treated as outliers (Fig. 5(f)).

of the computation.

The motivating idea of this paper is that these problems may be overcome with effective modeling of uncertainty in the registration process, including uncertainty in the point locations, in the transformation parameter estimate, and ultimately in the correspondences themselves. Figure 1(b) illustrates this: The uncertainty in the transformation estimate should produce uncertainty in the mapped point positions that is predominantly in the vertical direction. This should allow correspondences between points on the cross-bars of the two H’s to be preferentially established, which should then cause the estimation process to “pull” the transformation in the vertical direction. Moreover, since the correspondences between points along the two sides of the ‘H’ should indicate a large amount of uncertainty in the vertical direction, these correspondences should not work against the desired changes in the transformation.

This paper presents the Covariance Driven Correspondence (CDC) algorithm, built from this motivation. After Sec. 2 provides further discussion of the literature, the correspondence alignment error covariance matrix is derived in Sec. 3. This is formed into an objective function taken over the correspondences, the parameter estimate, and the parameter estimate covariance matrix in Sec. 4. Sec. 5 presents an EM-like optimization algorithm, briefly describes a second algorithm that uses an approximation to the CDC objective function, and outlines implementation details. Experimental results presented in Sec. 6 compare the performance of the algorithms to both ICP and RPM on synthetic data, on range scans, and on intensity images.

2. Background

Uncertainty has been used in a number of ways in existing image matching and registration algorithms. The RPM

algorithm [6, 27, 15] and related work on EM versions of ICP [8, 11] model the uncertainty in correspondences using an isotropic covariance matrix. This matrix starts large and then is gradually reduced through an annealing schedule. By combining annealing with constraints from a doubly-stochastic correspondence matrix, these algorithms eventually move toward unique correspondences. Outliers — points that have no analog in the other image — are handled through the formation of an explicit outlier process. Some of these algorithms include point location covariances as well.

Many other algorithms use point location covariance matrices during parameter estimation without incorporating uncertainty in the resulting estimate in a feed-back loop. This includes registration algorithms as well as related techniques in fundamental matrix estimation, stereo and structure-from-motion [7, 12, 26].

The final common use of uncertainty in image registration occurs following estimation — for model selection and region growth [20], for generating additional constraints in multi-image alignment [24], or for evaluating the stability of the estimate [10]. All of these algorithms compute covariance using a fixed set of unique correspondences, which, as illustrated in Fig. 3, leads to significantly underestimated uncertainties [22].

3. Covariance of the Alignment Error

We start by deriving the covariance matrix for the alignment error between two points given (a) the covariance matrix of the point locations, (b) a vector of estimated transformation parameters, and (c) the covariance matrix of these parameters. Let $\mathbf{p}_i$ be a moving image point and let $\mathbf{q}_j$ be a fixed image point. Let $\hat{\boldsymbol{\theta}}$ be the vector of estimated transformation parameters. Finally, let $\mathbf{T}$ be the mapping function. Then the alignment error vector between $\mathbf{q}_j$ and $\mathbf{p}_i$ is

$$\mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}}) = \mathbf{q}_j - \mathbf{T}(\mathbf{p}_i; \hat{\boldsymbol{\theta}}). \quad (1)$$

All three quantities that $\mathbf{e}$ depends on are random variables — $\mathbf{p}_i$ and $\mathbf{q}_j$ because they are measurements (or computed from measurements) and $\hat{\boldsymbol{\theta}}$ because it is estimated. Therefore, in order to derive the covariance matrix of $\mathbf{e}$, we consider the uncertainty in all three quantities:

$$\mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}}) = \mathbf{e}(\tilde{\mathbf{p}}_i + \Delta\mathbf{p}_i, \tilde{\mathbf{q}}_j + \Delta\mathbf{q}_j; \tilde{\boldsymbol{\theta}} + \Delta\hat{\boldsymbol{\theta}}), \quad (2)$$

where $\tilde{\mathbf{p}}_i$, $\tilde{\mathbf{q}}_j$, and $\tilde{\boldsymbol{\theta}}$ denote the error-free values. The error in the alignment distance is then

$$\Delta\mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}}) = \mathbf{e}(\tilde{\mathbf{p}}_i + \Delta\mathbf{p}_i, \tilde{\mathbf{q}}_j + \Delta\mathbf{q}_j; \tilde{\boldsymbol{\theta}} + \Delta\hat{\boldsymbol{\theta}}) - \mathbf{e}(\tilde{\mathbf{p}}_i, \tilde{\mathbf{q}}_j; \tilde{\boldsymbol{\theta}}). \quad (3)$$

Taking a Taylor expansion with respect to all three random variables and dropping quadratic and higher terms yields:

$$\Delta\mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}}) \approx \Delta\mathbf{q}_j + \mathbf{J}_{\mathbf{p}_i} \Delta\mathbf{p}_i + \mathbf{J}_{\boldsymbol{\theta}} \Delta\hat{\boldsymbol{\theta}} \quad (4)$$

where $\mathbf{J}_{\mathbf{p}_i} = \partial \mathbf{T} / \partial \mathbf{p}_i$ is the Jacobian of the transformation with respect to the moving point and $\mathbf{J}_{\boldsymbol{\theta}} = \partial \mathbf{T} / \partial \hat{\boldsymbol{\theta}}$ is the Jacobian of the transformation with respect to the current parameter estimate. Both Jacobians are evaluated at the current moving point and parameter estimates, $\mathbf{p}_i$ and $\hat{\boldsymbol{\theta}}$.

The covariance of the alignment error may be approximated by taking the expected value of the outer product of $\Delta \mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}})$ with itself and assuming all three random variables are pairwise independent. The independence of $\mathbf{p}_i$ and $\mathbf{q}_j$ is a safe assumption. The independence of $\hat{\boldsymbol{\theta}}$ is more problematic because in an iterative refinement framework the points are used to estimate the parameters. Still, if there are many correspondence pairs, then the influence of any one pair is small and the independence assumption is reasonable. Therefore,

$$\begin{aligned} E[\Delta \mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}}) \Delta \mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}})^\top] \\ = E[\Delta \mathbf{q}_j \Delta \mathbf{q}_j^\top] + \mathbf{J}_{\mathbf{p}_i} E[\Delta \mathbf{p}_i \Delta \mathbf{p}_i^\top] \mathbf{J}_{\mathbf{p}_i}^\top + \mathbf{J}_{\boldsymbol{\theta}} E[\Delta \hat{\boldsymbol{\theta}} \Delta \hat{\boldsymbol{\theta}}^\top] \mathbf{J}_{\boldsymbol{\theta}}^\top \\ = \mathbf{S}_{\mathbf{q}_j} + \mathbf{J}_{\mathbf{p}_i} \mathbf{S}_{\mathbf{p}_i} \mathbf{J}_{\mathbf{p}_i}^\top + \mathbf{J}_{\boldsymbol{\theta}} \mathbf{S}_{\boldsymbol{\theta}} \mathbf{J}_{\boldsymbol{\theta}}^\top = \mathbf{S}_{ij}. \end{aligned} \quad (5)$$

Thus, the uncertainty in the error of the correspondence, which we write simply as $\mathbf{S}_{ij}$, depends on (a) the covariance of the fixed point location $\mathbf{S}_{\mathbf{q}_j}$, (b) the covariance of the moving point location $\mathbf{S}_{\mathbf{p}_i}$, mapped using the Jacobian $\mathbf{J}_{\mathbf{p}_i}$, and (c) the transformation parameter estimate $\hat{\boldsymbol{\theta}}$ and its covariance $\mathbf{S}_{\boldsymbol{\theta}}$, mapped using the Jacobian $\mathbf{J}_{\boldsymbol{\theta}}$. $\mathbf{S}_{\mathbf{q}_j}$ and $\mathbf{S}_{\mathbf{p}_i}$ are pre-computed from the data and are independent of the transformation [1] (Sec. 5). The third covariance, $\mathbf{S}_{\boldsymbol{\theta}}$, must be estimated during registration. The term it forms for each correspondence, $\mathbf{J}_{\boldsymbol{\theta}} \mathbf{S}_{\boldsymbol{\theta}} \mathbf{J}_{\boldsymbol{\theta}}^\top$, is often called the “transfer error covariance” [13, Ch. 4]. As will be shown below (Fig. 4), during the initial stages of registration, the transfer error covariance dominates $\mathbf{S}_{ij}$, causing elongation in the direction of the error in the transformation, but as the estimate converges, the transfer error covariance shrinks dramatically and the location covariances come to play an equal or greater role in $\mathbf{S}_{ij}$.

4. Objective Function

Having derived the covariance matrix for each correspondence match, we now formulate an objective function in terms of the correspondences, the parameters $\boldsymbol{\theta}$, and the covariance $\mathbf{S}_{\boldsymbol{\theta}}$. We model the alignment error for a correct correspondence as a normal distribution with zero mean and covariance $\mathbf{S}_{ij}$. Writing the negative log-likelihood over a set $\mathcal{C}_T$ of correct correspondences $\mathbf{p}_i, \mathbf{q}_j$ and dropping constant terms, we obtain

$$\mathcal{L}_T(\boldsymbol{\theta}, \mathbf{S}_{\boldsymbol{\theta}}) = \sum_{\mathbf{p}_i, \mathbf{q}_j \in \mathcal{C}_T} \frac{1}{2} (\mathbf{e}_{ij}^\top \mathbf{S}_{ij}^{-1} \mathbf{e}_{ij} + \ln |\mathbf{S}_{ij}|) \quad (6)$$

where $\mathbf{e}_{ij} = \mathbf{e}(\mathbf{p}_i, \mathbf{q}_j; \hat{\boldsymbol{\theta}})$.

Now, since the set of true correspondences is unknown, we expand this function to include all possible correspondences. In doing so, we need to account for incorrect matches and for points that have no analog in the other image. We do this in two parts, first by replacing the quadratic loss function with a robust loss function and second by factoring in a competitive weight between different correspondences. This results in the function

$$\mathcal{F}(\boldsymbol{\theta}, \mathbf{S}_{\boldsymbol{\theta}}) = \sum_{i \in \mathcal{P}} \sum_{j \in \mathcal{Q}} w_{ij} \left(\rho \left(\sqrt{\mathbf{e}_{ij}^\top \mathbf{S}_{ij}^{-1} \mathbf{e}_{ij}} \right) + \ln |\mathbf{S}_{ij}| \right). \quad (7)$$

Robust loss function $\rho(\cdot)$ is the standard Beaton-Tukey [21] scaled by a multiplier $k = 2.9872$ found by minimizing the L_2 -norm difference between $\rho(u)$ and u^2 over the inlier interval $[-3..3]$.

The competitive weight term is formed to enforce competition between matches:

$$w_{ij} = \frac{w_{ij}^r}{\sum_{\mathbf{p}_i \in \mathcal{P}} w_{ij}^r} \frac{w_{ij}^r}{\sum_{\mathbf{q}_j \in \mathcal{Q}} w_{ij}^r}. \quad (8)$$

Here, w_{ij}^r is the usual M-estimator weight used in reweighted least-squares and formed from the ρ function [21]. The two factors in w_{ij} enforce competition between correspondence $(\mathbf{p}_i, \mathbf{q}_j)$ and other correspondences for $\mathbf{p}_i$ in the fixed image and other correspondences for $\mathbf{q}_j$ in the moving image.

As a result of this design, robustness and competition are achieved without an explicit outlier process and without an annealing schedule. Unique correspondences are never explicitly enforced and, indeed, because of different sampling in the two images we do not want to enforce them. Moving points, $\mathbf{p}_i$ having all bad correspondences in the fixed image will have $w_{ij} = 0, \forall j$ (since w_{ij}^r is 0 for large outliers when using the Beaton-Tukey ρ function), and will not influence the estimate.

5. Algorithm

Minimizing the objective function (7) in order to compute the transformation parameter estimate $\hat{\boldsymbol{\theta}}$ requires simultaneously estimating the covariance matrix $\mathbf{S}_{\boldsymbol{\theta}}$ and weights w_{ij} . It is important to note that $\mathbf{S}_{\boldsymbol{\theta}}$ is being estimated at the same time as $\hat{\boldsymbol{\theta}}$ instead of afterwards. Since $\hat{\boldsymbol{\theta}}$ is used to weight correspondences during registration but is itself uncertain, its uncertainty should be incorporated into the registration process. The estimation of $\mathbf{S}_{\boldsymbol{\theta}}$ occurs naturally in an EM framework, but this has not been used previously in registration.

The minimization problem is solved using a multi-step process starting from an initial transformation parameter vector $\hat{\boldsymbol{\theta}}_0$ as well as the point sets $\mathcal{P}$ and $\mathcal{Q}$ and the point covariances. Here is an outline of the algorithm, with details provided subsequently:

1. Compute an initial estimate of transformation parameter covariance matrix $\mathbf{S}_\theta$ (Section 5.1).
2. Iterate until convergence
 - (a) Recompute the weights $\{w_{ij}\}$ for all correspondences keeping $\hat{\theta}$ and $\mathbf{S}_\theta$ fixed (Equation 8).
 - (b) Update the parameter estimate $\hat{\theta}$ holding $\{w_{ij}\}$ and $\mathbf{S}_\theta$ fixed (Section 5.2).
 - (c) Recompute the weights $\{w_{ij}\}$ as in Step 2(a).
 - (d) Update the covariance matrix estimate $\mathbf{S}_\theta$ holding $\{w_{ij}\}$ and $\hat{\theta}$ fixed (Section 5.2).

5.1. Covariances and Initialization

The point covariance matrices may be viewed as a local summary of shape, with more elliptic matrices along edges and planes (in 3d) and more rounded matrices near corners. When the points are given in a discrete set, as with range data, we compute the covariance matrix at each point ($\mathbf{p}_i$ or $\mathbf{q}_j$) by forming a scatter matrix from the points in a small, surrounding neighborhood [1]. For an intensity image, points are placed at local maxima of the intensity gradient, as in edge detection, and the covariance matrices are formed from the summed outer product of the surrounding gradient vectors [4].

The initial estimate of $\mathbf{S}_\theta$ — Step 1 of the algorithm — is found using a two-part process. First, using the initial parameter estimate $\hat{\theta}_0$, the moving points are mapped into the fixed image, the closest fixed point to each is found, and a robustly-weighted, least-squares objective function is formed. Just as in the covariance computation for normal distance ICP [10, 22], $\mathbf{S}_\theta$ is assigned to be the inverse of the Hessian of this objective function. Second, this starting estimate of $\mathbf{S}_\theta$ is refined by running optimization Steps 2(c) and 2(d) of the algorithm, with the parameter estimate, $\hat{\theta}_0$, fixed. Five iterations prove sufficient in practice. Fig. 3 illustrates the effect of this on the correspondence covariance matrices $\mathbf{S}_{ij}$ and their uncertainty ellipses.

We ensure that $\mathbf{S}_\theta$ remains positive definite throughout the computation by using the Cholesky decomposition: $\mathbf{S}_\theta = \mathbf{U}^\top \mathbf{U}$, where matrix $\mathbf{U}$ is upper triangular. The parameters of $\mathbf{U}$ are estimated by the optimizer. These parameters describe the covariance of the overall transformation for similarity (in 2d) and affine transformations (in 2d and in 3d). When computing the 3d rigid transformation they represent the covariance of the incremental transformation estimated using a small angle approximation.

5.2. Optimization

The iterations of updating the weights and then the parameters are reminiscent of the EM algorithm. Given fixed parameters and covariance matrix, the weights are updated

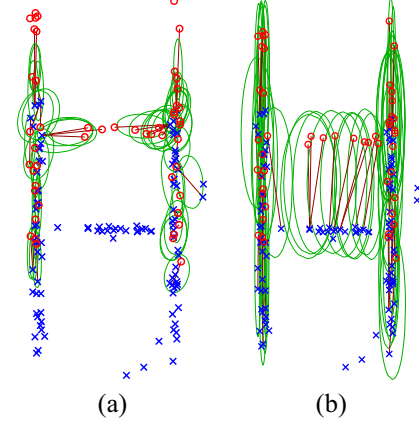


Figure 3. Correspondence uncertainty ellipses formed from covariance matrices $\mathbf{S}_{ij}$. In (a) the $\mathbf{S}_{ij}$ were formed using the starting parameter covariance matrix $\mathbf{S}_\theta$ obtained from the ICP-like objective function, while in (b) they were formed from $\mathbf{S}_\theta$ after five refinement iterations. Only the ellipses for the highest-weighted matches with their covariance matrices are shown. The ellipses in (b) more effectively capture the uncertainty in the alignment.

as defined in (8). In this and in all other computations, the covariance matrix $\mathbf{S}_\theta$ enters indirectly through the computation of the individual correspondence covariance matrices $\mathbf{S}_{ij}$ (5), which also depend on the point covariances and two Jacobians. The former are fixed throughout the computation, while the latter are easily computed at each point given the transformation parameters.

Given fixed weights, the optimization of the objective function (7) with respect to $\hat{\theta}$ or $\mathbf{S}_\theta$ is accomplished with a quasi-Newton method known as the BFGS algorithm [18]. This requires computation of the gradient of the objective function at each iteration (an exercise in calculating matrix derivatives and applying the chain rule). We allow only a few iterations in each direction prior to stopping and recomputing weights, similar to the Generalized EM algorithm [3, Ch. 9]. This, together with the separation of updates to $\hat{\theta}$ and $\mathbf{S}_\theta$, produces more reliable convergence.

A simplified version of the algorithm, with a quadratic objection function in place of $\rho(\cdot)$ and with competitive weighting only based on alternative correspondences for moving points, matches the form of the Generalized EM algorithm [3, Ch. 9], and therefore is provably convergent. As yet we do not have a convergence proof for the full form of the algorithm. In practice, convergence is always obtained.

5.3. An Approximation Algorithm

For a k -dimensional parameter vector, CDC estimates $k(k+1)/2$ covariance matrix parameters. While this is not as inefficient as it seems — experiments show that approximately equal time is spent on weight computation, transformation parameter estimation, and covariance parameter

estimation — more efficient approximations are possible. We describe one here.

CDC uses the estimated S_{θ} to compute the transfer error covariance matrix $J_{\theta} S_{\theta} J_{\theta}^{\top}$ for each moving point separately, using this in turn to compute the alignment error covariance S_{ij} . The approximation, CDC', instead estimates a global approximation for the transfer error covariance matrix: $S_T \approx J_{\theta} S_{\theta} J_{\theta}^{\top}$. Using this,

$$S_{ij} = S_{q_j} + J_{p_i} S_{p_i} J_{p_i}^{\top} + S_T,$$

and Step 2(d) of the algorithm estimates S_T . This reduces the number of covariance unknowns to $d(d+1)/2$, where $d = 2$ for intensity images and $d = 3$ for range data. This is somewhat akin to the use of a global correspondence covariance matrix in RPM. The difference is that our covariance matrix is non-isotropic and is estimated instead of annealed. Preliminary experiments using CDC' are presented below.

5.4. Implementation Details

During weight computation, the bounded robust objective function limits the range over which correspondences for each moving point can have non-zero weight. A k-d tree storing the points is then searched to find the fixed points in this range. Only these points are considered in the weight computation, reducing the quadratic cost of computing (7) much closer to linear.

The points are organized into a multiresolution representation, but the actual use of this representation is novel. The average size of the transfer error uncertainty ellipse for the moving points is used to adaptively switch resolutions. Larger ellipses and therefore broader search ranges cause a switch to a coarser resolution, while smaller ellipses cause a switch to a finer resolution. The switch ensures that each moving point considers at most 4-5 fixed points with non-zero weight.

When estimating parameters of transformations that have an associated scaling (similarity and affine), we incorporate a symmetry component to the estimation, with additional correspondences (and weights and covariances) computed by reversing the roles of the fixed and moving point sets. This leads to some redundancy, but stabilizes the scaling parameters.

6. Experimental Results

We present experimental evaluation of the proposed algorithms on a variety of 2D and 3D data sets, focusing primarily on CDC but presenting a brief evaluation of the approximation, CDC'. We use both types of data to compare CDC with robust ICP. We use synthetic data to compare the abilities of CDC and RPM to handle missing or extraneous structures. Finally, we evaluate using CDC in place of ICP

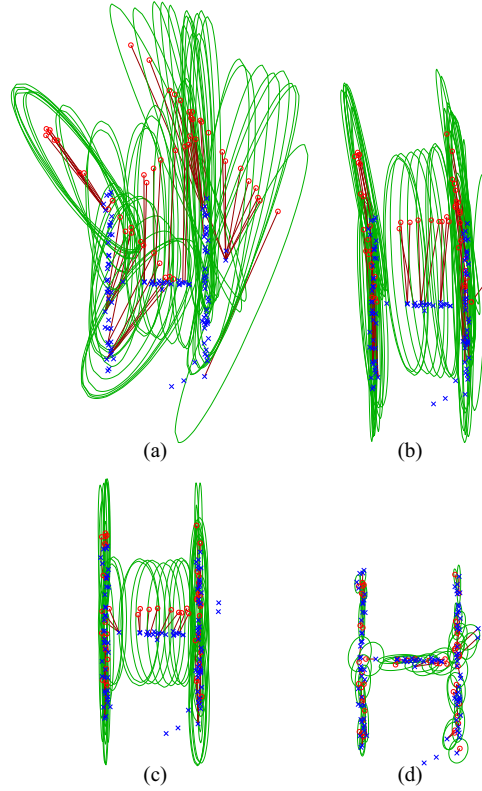


Figure 4. Example of aligning two H shapes using a similarity transformation. Correspondence covariance matrices S_{ij} are shown as oriented uncertainty ellipses. The initial alignment and CDC covariances are shown in (a). Frames (b) through (d) illustrate intermediate results, ending in convergence.

as the core alignment engine within the Dual-Bootstrap algorithm [20, 25].

Computationally, CDC averages about 2.5 seconds per iteration using 1700 points on a Pentium 4 3.2GHz PC with 2GB memory computer, requiring 63 iterations, on average, to converge for 3d experiments. Time tends to be split somewhat equally between weight, transformation and covariance estimation steps.

6.1. 2D Registration of Synthetic Point Sets

An example of aligning two synthetic H shapes using a similarity transformation is presented in Fig. 4 to illustrate the behavior of CDC. In generating this data, points were uniformly sampled on two H shapes 200 points wide and 400 points tall, and isotropic Gaussian noise with std. dev. 4.0 was added to their locations. Correspondence covariance matrices S_{ij} are shown as uncertainty ellipses whose axis half-lengths are the square roots of the eigenvalues of the S_{ij} . One ellipse is plotted for each moving point using its highest weight correspondence. As the algorithm converges, the covariance matrices change from long and thin

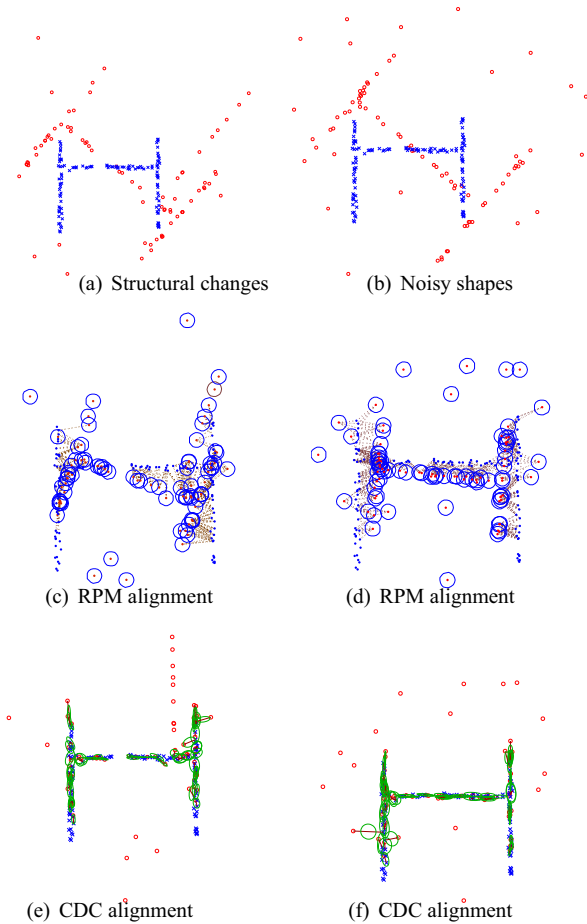


Figure 5. Example of aligning two H shapes with structural changes **(a)** and noise **(b)**. Robust Point Matching (RPM) fails because of the correspondences generated with extraneous structures **(c),(d)**, while the proposed CDC algorithm locks onto the correct correspondences and correctly aligns the shapes **(e),(f)**.

to small and round, reflecting a decrease in the transformation uncertainty both overall and relative to point location uncertainty.

We also study the domain of convergence of CDC and CDC' as compared to that of ICP and of RPM. We used our own implementation of robust ICP, but tested RPM [6] using code posted on-line by the authors. The moving shape was initialized at positions sampled on a circle whose radius is half of the fixed shape width. Initial rotations of $\pm 40^\circ$ in 20° increments were tested. Results and an example initialization are shown in Fig. 7. Clearly, ICP is significantly worse than RPM, CDC and CDC'. This shows the limitations using single matches and Euclidean closest-point distance measures. RPM and CDC have the same number of failures, 14 each, while CDC' has a few more.

The next step in comparing RPM [6] with CDC is to consider the effect of outliers and extraneous structures. This is important for registration when there are illumination or

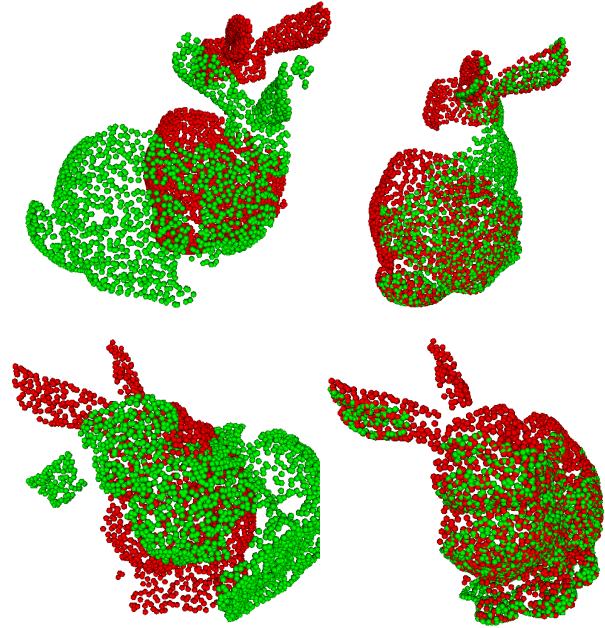


Figure 6. Example initial transformations **(left column)** and final CDC alignments **(right column)** on the Stanford bunny [23] dataset (best seen in color). In both cases robust ICP using normal distance constraints failed to converge.

physical changes or low overlap between images. Two comparison examples are shown in Fig. 5. In the top example, an extra line of points has been added to the moving shape, while in the bottom example the top arms of the 'H' have been extended in the moving shape. RPM failed to align the two shapes, effectively producing a compromise between the correct correspondences and the correspondences with the extra structures that causes both rotation and scale errors. CDC, as shown in the figure, and CDC' are both able to align these data sets correctly. This difference in performance persists for other examples as well.

6.2. 3D Rigid Registration

Our experiments with rigid registration in 3D use the well-known Stanford bunny dataset [23], which has ten scans. In these experiments, we initialize the transformation to the identity matrix and run registration on all scan pairs having any overlap. We compare only normal-distance ICP to our proposed CDC algorithm because code for a 3D version of RPM was unavailable. It is important to observe that we are attempting to align scans taken from different viewpoints rather than prerotating a set of data points to register against itself. Our test is harder because the difference in scanner viewpoints causes a decrease in the fraction of common surfaces and therefore an increase in the number of missing or extraneous data points. Combined with the result of the previous experiment, this suggests (without

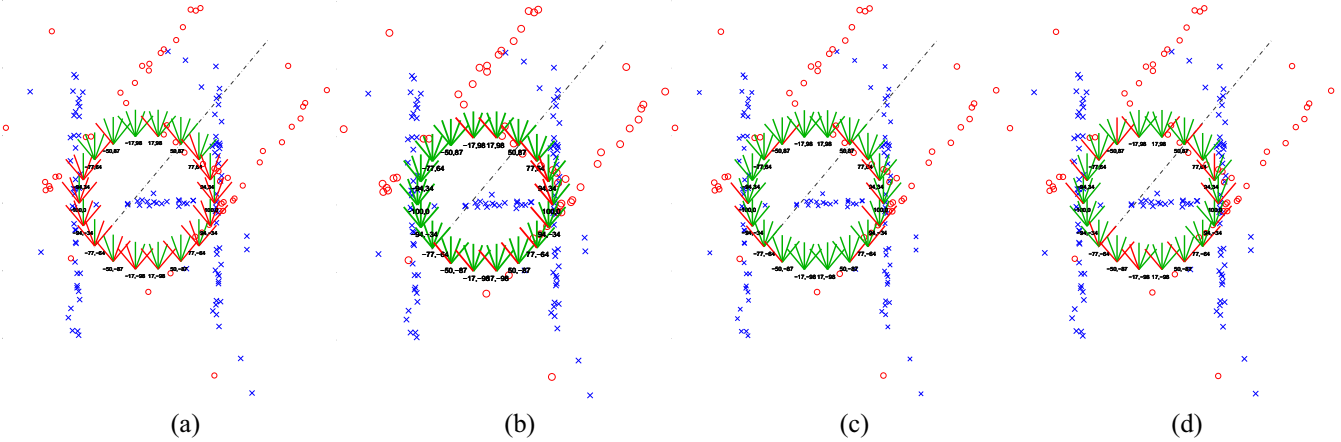


Figure 7. Convergence analysis for two H shapes using ICP with normal distances (a), RPM (b), the CDC algorithm proposed here (c), and the approximation, CDC' (d). The moving shape (red circles) was initialized at the marked locations around a circle with five different orientations, shown by the orientations of the green/red line segments, at each location. A green line segment indicates a location and orientation from which the algorithm converged correctly, while a red segment indicates a failure.

proof) that RPM would work poorly on this data.

ICP successfully aligned nine scan pairs, but CDC correctly aligned 16 including the nine ICP aligned. CDC failed on only two pairs whose inter-scan rotation angle was less than 80 degrees — the failures involved 56 and 59 degree rotations — whereas ICP failed on most pairs rotated by more than 45 degrees. Examples of initial and final alignments of pairs for which CDC succeeded and ICP failed are shown in Fig. 6. Notice the large differences between the point sets in these scans. The overall RMS alignment accuracy is close to the noise and sample spacing of the scanner. Occasionally it is slightly higher for low-overlap scans, especially when different parts of the ear are sampled in the two scans. Such alignments could be further refined by incorporating information about surface normals during weight computation [19].

6.3. CDC in the Dual-Bootstrap Algorithm

Our final set of experiments explores the use of CDC in place of ICP in the Dual-Bootstrap algorithm [20, 25]. The Dual-Bootstrap is really a system of algorithms consisting of (a) generation of initial transformations, each roughly accurate over only a small image region, (b) a core estimation, refinement and region growth procedure, applied to each initial transformation separately, and (c) a sophisticated set of decision criteria. Step (b) is currently based on a robust form of ICP. Experiments have shown that 81% of the initial estimates that are roughly accurate in only the small region are successfully refined to an accurate final transformation. This allows the Dual-Bootstrap to correctly align 19 of 22 image pairs from a challenging test suite.

Testing CDC in place of ICP within the Dual-Bootstrap resulted in the correct alignment of 36% of the roughly-accurate initializations that failed when using ICP. While



Figure 8. Image pair from the Oxford data set [16] that the Dual-Bootstrap algorithm aligned using CDC but not using ICP as its core alignment engine.

this is only a modest improvement, it occurs in a context of an overall algorithm — particularly the region-growing step — specifically designed to limit the effects of ICP's short-comings. Although use of CDC did not result in a correct alignment of the three pairs from the original test suite on which the Dual-Bootstrap failed, experiments with the data set in [16] did result in one increase in the correctly-aligned pairs (Fig. 8). This indicates the significance of CDC in helping the Dual-Bootstrap align extremely challenging pairs.

7. Discussion and Conclusion

As compared to the widely-used ICP algorithm, CDC has a much broader domain of convergence. In effect CDC can be viewed as estimating and using Mahalanobis distances during the registration process which allows it to generate constraints that ICP's minimum-distance matching misses. As compared to RPM, CDC includes explicit esti-

mation of covariances rather than an annealing of isotropic covariances, and CDC relies on a robust objective function rather than a separate outlier process. These allow CDC to more effectively tolerate extraneous or missing structures in the data.

CDC has two primary weaknesses. First, since it explicitly includes estimation of the covariance matrix during the iterations of registration, the number of unknowns that must be estimated increases quadratically with the number of transformation parameters. While experimentally the cost of this is less than expected, improvements are still needed. Second, CDC has only been demonstrated on low-dimensional transformation models and has not been applied to deformable registration. The approximation used in the CDC' algorithm, where the alignment covariance no longer explicitly depends on the parameter estimate covariance, represents our first step toward solving both of these problems.

In conclusion, we have shown that by explicitly incorporating a model of uncertainty in the registration process, both the robustness of registration to missing structures and the domain of convergence of alignment can be substantially improved.

References

- [1] J. Berkmann and T. Caelli. Computation of surface geometry and segmentation using covariance techniques. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(11):1114–1116, 1994. 3, 4
- [2] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(2):239–256, 1992. 1
- [3] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006. 4
- [4] M. J. Brooks, W. Chojnacki, D. Gawley, and A. van den Hengel. What value covariance information in estimating vision parameters? In *Proc. ICCV*, pages 302–308, 2001. 4
- [5] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *IVC*, 10(3):145–155, 1992. 1
- [6] H. Chui, A. Rangarajan, J. Zhang, and C. M. Leonard. Unsupervised learning of an atlas from unlabeled point-sets. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(2):160–172, 2004. 1, 2, 6
- [7] P. David, D. DeMenthon, R. Duraiswami, and H. Samet. SoftPOSIT: Simultaneous pose and correspondence determination. *Int. J. Comp. Vis.*, 59(3):259–284, Sept. 2004. 2
- [8] R. S. J. Estépar, A. Brun, and C.-F. Westin. Robust generalized total least squares iterative closest point registration. In *Proc. 7th MICCAI*, pages 234–241, Saint-Malo, France, 2004. 2
- [9] A. W. Fitzgibbon. Robust registration of 2d and 3d point sets. *IVC*, 21(13-14):1145–1153, Dec. 2003. 1
- [10] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy. Geometrically stable sampling for the ICP algorithm. In *Proc. 4th Int. Conf. on 3DIM*, pages 260–267, 2003. 1, 2, 4
- [11] S. Granger and X. Pennec. Multi-scale EM-ICP: A fast and robust approach for surface registration. In *Proc. Seventh ECCV*, pages 418–432, 2002. 2
- [12] A. Gruber and Y. Weiss. Multibody factorization with uncertainty and missing data using the EM algorithm. In *Proc. CVPR*, volume 1, pages 707–714, 2004. 2
- [13] R. Hartley and A. Zisserman. *Multiple View Geometry*. Cambridge University Press, 2000. 3
- [14] T. Jost and H. Hugli. A multi-resolution ICP with heuristic closest point search for fast and robust 3d registration of range images. In *Proc. 4th Int. Conf. on 3DIM*, pages 427–433, 2003. 1
- [15] B. Luo and E. R. Hancock. Iterative procrustes alignment with the EM algorithm. *IVC*, 20(5-6):377–396, Apr. 2002. 2
- [16] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *Int. J. Comp. Vis.*, 65(1–2):43–72, 2005. 7
- [17] N. J. Mitra, N. Gelfand, H. Pottmann, and L. Guibas. Registration of point cloud data from a geometric optimization perspective. In *Eurographics Symposium on Geometry Processing*, 2004. 1
- [18] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2000. 4
- [19] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. 3rd Int. Conf. on 3DIM*, pages 224–231, 2001. 1, 7
- [20] C. Stewart, C.-L. Tsai, and B. Roysam. The Dual-Bootstrap Iterative Closest Point algorithm with application to retinal image registration. *IEEE Trans. Med. Imag.*, 22(11):1379–1394, 2003. 2, 5, 7
- [21] C. V. Stewart. Robust parameter estimation in computer vision. *SIAM Rev.*, 41(3):513–537, 1999. 3
- [22] C. V. Stewart. Uncertainty-driven, point-based image registration. In N. Paragios, Y. Chen, and O. Faugeras, editors, *Handbook of Mathematical Models in Computer Vision*, pages 221–235. Springer, 2006. 2, 4
- [23] G. Turk and M. Levoy. Zippered polygon meshes from range images. In *SIGGRAPH*, pages 311–318, Orlando, Florida, USA, July 1994. 6
- [24] G. Yang and C. V. Stewart. Covariance-driven mosaic formation from sparsely-overlapping image sets with application to retinal image mosaicing. In *Proc. CVPR*, pages 804–810, 2004. 2
- [25] G. Yang, C. V. Stewart, M. Sofka, and C.-L. Tsai. The Generalized Dual-Bootstrap ICP algorithm with application to registering challenging image pairs. Technical Report 05-19, Department of Computer Science, Rensselaer Polytechnic Institute, 2005. 5, 7
- [26] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *Int. J. Comp. Vis.*, 27(2):161–195, 1998. 2
- [27] Y. Zheng and D. Doermann. Robust point matching for non-rigid shapes by preserving local neighborhood structures. *IEEE Trans. Pattern Anal. Machine Intell.*, 28(4):643–649, 2006. 2