Extended Coherent Point Drift Algorithm with Correspondence Priors and Optimal Subsampling

Vladislav Golyanik* Bertram Taetz* Gerd Reis Didier Stricker
German Research Center for Artificial Intelligence,
Trippstadter Str. 122, 67663 Kaiserslautern, Germany
{Vladislav.Golyanik, Bertram.Taetz, Gerd.Reis, Didier.Stricker}@dfki.de

Abstract

The problem of dense point set registration, given a sparse set of prior correspondences, often arises in computer vision tasks. Unlike in the rigid case, integrating prior knowledge into a registration algorithm is especially demanding in the non-rigid case due to the high variability of motion and deformation. In this paper we present the Extended Coherent Point Drift registration algorithm. It enables, on the one hand, to couple correspondence priors into the dense registration procedure in a closed form and, on the other hand, to process large point sets in reasonable time through adopting an optimal coarse-to-fine strategy. Combined with a suitable keypoint extractor during the preprocessing step, our method allows for non-rigid registrations with increased accuracy for point sets with structured outliers. We demonstrate advantages of our approach against other non-rigid point set registration methods in synthetic and real-world scenarios.

1. Introduction

Point set registration is a key component in many computer vision tasks such as 3D reconstruction, medical image registration, computer graphics and shape recognition. The problem consists in finding correspondences and transformations between a *reference* and a *template point set* that might show a deformed, partially overlapped and noisy part of the reference point set. Thus, the objective is to register two point sets into a common reference frame. We refer to *rigid point set registration*, if the transformation between the template point set and the reference point set can be entirely determined by parameters of the rigid body motion. In case of *non-rigid point set registration*, transformations can be arbitrary for each point and non-rigid deformations are possible. We consider a special case of the non-rigid registration problem, when a sparse set of *correspondence priors*, possibly with outliers, is known in advance.

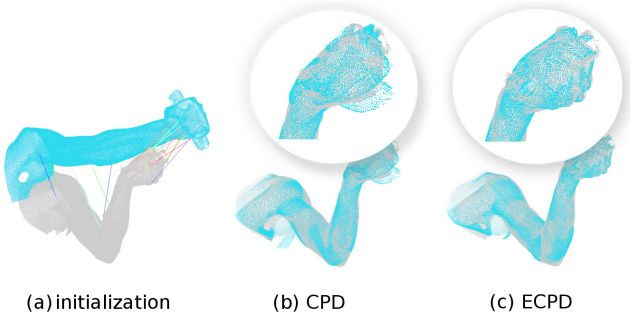

Figure 1: Non-rigid registration of two 3D point sets from [2] representing arms in different poses. (a): initial alignment with correspondences obtained through a comparison of the Persistent Feature Histograms at ISS 3D keypoints; (b): result of the non-rigid registration with CPD; (c): result of the non-rigid registration with ECPD using the correspondence priors. Complex non-rigid deformations (combination of supination, flection and abduction of the arm) in the area of the hand were resolved with higher precision, compared to (b). See Fig. 3 for detailed comparison.

Non-rigid point set registration is an ill-posed problem. To obtain a unique solution, additional constraints on the solution space are required. Constraints originate from assumptions about point sets as well as the type of displacements that the point sets may undergo. They can be formulated as prior knowledge (priors) and included into the registration procedure in different ways (e.g. through regularization). An overview of registration methods can be found in [26] and in Section 2. In this paper we present the Extended Coherent Point Drift (ECPD) algorithm allowing to include prior information in form of point correspondences into the non-rigid registration process to influence it in a favourable way (see Fig. 1). Our method builds upon the Coherent Point Drift (CPD) [20] and thus broadens its scope. It allows to combine application specific correspondence search algorithms operating on different information sources with the robust non-rigid point set registration algorithm. Furthermore, we adopt a coarse-to-fine registration strategy which maintains the impact of the correspondence priors and accelerates the registration process linearly.

* the authors contributed equally to this work

2. Related Work

Early works on non-rigid point set registration used probabilistic Gaussian Mixture Models (GMM) positioned along contours. The contours were modelled by splines allowing non-rigid transformations, but the method was restricted to contour-like registrations [14], [23]. Several extensions of the Iterative Closest Point (ICP) for the non-rigid case were proposed [11], [3]. They evince the same drawbacks as their rigid counterpart, namely high sensitivity to outliers. One of the widely used non-rigid point set registration methods is based on modelling the transformation with thin plate splines (TPS) [6] followed by robust point matching (RPM) and is known as the TPS-RPM [9]. It uses deterministic annealing and alternates between updates of the soft assignment and estimation of transformation parameters. The authors showed how the expectation-maximization (EM) algorithm can be embedded into a deterministic annealing scheme [8]. An optimized implementation of the TPR-RPM on GPU for point sets with thousands of elements was addressed in [19]. A correlation based approach was proposed in [27] and later improved in [16]. It tries to align two distributions whereby each of the point sets represents GMM centroids. CPD method was introduced in [21] and further improved in [20]. It employs an EM algorithm to optimize the GMM and the regularization originating in the motion coherence theory [29], [30]. Compared to TPS-RPM, CPD offers superior accuracy and stability with respect to non-rigid deformations in presence of outliers. An additional parameter for outlier modelling of CPD was introduced for a hybrid optimization with the Nelder-Mead simplex method in [28]. CPD was also modified by imposing the Local Linear Embedding topological constraint to cope with highly articulated non-rigid deformations [13]. However, this extension performs poorly on data with inhomogeneous density and is more sensitive to noise than CPD. Recently, a non-rigid registration method based on Student's Mixture Model (SMM) showed to be even more robust and accurate on noisy data than CPD method [33].

Several methods were proposed for improving registration quality in challenging cases through embedding prior knowledge and constraining the solution space. Some are based on shape priors and address registration of a human template to human scans [15]. A recent approach for human shape registration [5] not only registers the geometric shape, but also the appearance of shapes and the optical flow between shapes obtained from textured 3D scans. Some spectral differential geometry methods, constituting rather a separate class of algorithms, are able to regard prior correspondences [17]. Compared to point set registration methods, they are more noise sensitive and typically operate on meshes (i.e. need surfaces and normals), whereas real-world scans are usually noisy point clouds. In the context of the related problem of medical image registration, several methods utilize correspondences between SIFT keypoints or hybrid detectors and couple them into the registration procedure [18], [31]. No further assumptions about the content of the image are made, which allows to decouple correspondence search and image registration.

For acceleration purposes, subsampling of point sets can be applied. In the rigid case the recovered transformation refers simultaneously to all points. Hence, the transformation, recovered for a subsampled point set, can be directly generalized to the initial one [24]. In opposite, in the case of non-rigid registration such transformation does not generalize to the initial dense point set.

Our contributions are stated as follows. We derive ECPD method by embedding the correspondence priors into CPD in a closed form and thus extend its scope. To the best of our knowledge, utilizing the correspondence priors for point set registration in a closed-form was not shown in the literature so far, at least for the probabilistic case. We suggest how subsampling can be adopted in the context of non-rigid registration resulting in linear speedup as a function of the subsampling factor and point sets sizes. We also show an efficient implementation of ECPD in a heterogeneous environment with a GPU. Finally, we demonstrate several application specific approaches for finding correspondence priors and show that their appropriate utilization in combination with the proposed ECPD can address diverse issues, especially 1) resolving complex non-rigid deformations, not addressed by the original coherency constraint; 2) obtaining accurate registrations in the regions of interest (ROI) for point sets with structured outliers; 3) obtaining registration results linearly faster with respect to CPD.

3. Coherent Point Drift

CPD method [20] considers alignment of two D-dimensional point sets $\mathbf{X}_{N\times D} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)^T$ and $\mathbf{Y}_{M\times D} = (\mathbf{y}_1, \ldots, \mathbf{y}_M)^T$ as a probability density estimation problem where one point set represents the GMM centroids ($\mathbf{Y}_{M\times D}$) and the other one represents the data points ($\mathbf{X}_{N\times D}$). According to different deformation models between the two point sets, an appropriate rigid or non-rigid transformation can be selected. At the optimum two point sets become aligned and the correspondences are obtained using the maximum of the GMM posterior probability for a given data point. Core to CPD method is to force the GMM centroids to move coherently as a group to preserve the topological structure of the point sets. The GMM probability density function of CPD method can be written as

$$p(\mathbf{x}) = \sum_{m=1}^{M+1} P(m)p(\mathbf{x}|m). \tag{1}$$

To explicitly model outliers, the density function is split in the following form

$$p(\mathbf{x}) = w\frac{1}{N} + (1-w)\sum_{m=1}^{M} P(m)p(\mathbf{x}|m), \qquad (2)$$

with $p(\mathbf{x}|m) = \frac{1}{(2\pi\sigma^2)^{D/2}}\exp(-\frac{\|\mathbf{x}-\mathbf{y}_m\|^2}{2\sigma^2})$ and $0 \leq w \leq 1$ — a weight parameter that reflects the assumption about the amount of outliers in the reference point set. The GMM centroids are adjusted by a set of transformation parameters θ that can be estimated by minimizing the negative log-likelihood function

$$E(\theta, \sigma^2) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M+1} P(m)p(\mathbf{x}_n|m). \qquad (3)$$

An EM algorithm is used to find θ and σ^2. The E-step constructs a guess of the parameter values based on the previous ("old") values and then uses the Bayes' theorem to compute a posteriori probability distribution $P^{old}(m|\mathbf{x}_n)$. The M-step updates the parameters by minimizing an upper bound of the negative log-likelihood function in Eq. (3). Leaving out the terms constant w.r.t. σ^2 and θ the objective function can be written as

$$\begin{aligned} Q(\theta, \sigma^2) &= \frac{1}{2\sigma^2}\sum_{n=1}^{N}\sum_{m=1}^{M} P^{old}(m|\mathbf{x}_n)\|\mathbf{x}_n - T(\mathbf{y}_m, \theta)\|^2 \\ &+ \frac{N_P D}{2}\log\sigma^2 \end{aligned} \qquad (4)$$

where $T(\mathbf{y}_m, \theta)$ is a transformation applied to $\mathbf{Y}$ and $N_p = \sum_{n=1}^{N}\sum_{m=1}^{M} P^{old}(m|\mathbf{x}_n)$ and

$$P^{old}(m|\mathbf{x}_n) = \frac{\exp(-\frac{1}{2}\|\frac{\mathbf{x}_n - T(\mathbf{y}_m, \theta^{old})}{\sigma^{old}}\|^2)}{\sum_{k=1}^{M}\exp(-\frac{1}{2}\|\frac{\mathbf{x}_n - T(\mathbf{y}_k, \theta^{old})}{\sigma^{old}}\|^2) + c} \qquad (5)$$

with $c = (2\pi\sigma^2)^{D/2}\frac{w}{1-w}\frac{M}{N}$. Based on the formulas above the transformation T can be specified for rigid, affine and non-rigid point set registration.

4. Extended CPD

Let $(\mathbf{y}_j, \mathbf{x}_k)$ be a set of correspondence priors with indices $(j,k) \in N_c \subset \mathbb{N}^2$. We model correspondence priors by a product of particular independent density functions

$$P_c(N_c) = \prod_{(j,k)\in N_c} p_c(\mathbf{x}_j, \mathbf{y}_k) \qquad (6)$$

with

$$p_c(\mathbf{x}_j, \mathbf{y}_k) = \frac{1}{(2\pi\alpha)^{D/2}}\exp(-\frac{\|\mathbf{x}_j - T(\mathbf{y}_k, \theta)\|^2}{2\alpha^2}). \qquad (7)$$

Due to the Gaussian form of the distribution the parameter $\alpha > 0$ reflects the priors' degree of reliability. We incorporate correspondence priors into CPD method by including the prior probability $P_c(N_c)$ into the GMM in Eq. (1) and obtain a modified GMM with the density function

$$\tilde{p}(\mathbf{x}) = P_c(N_c)p(x). \qquad (8)$$

The modified energy function can be derived by considering the negative logarithm of the combined modified GMM

$$\begin{aligned} \tilde{E}(\theta, \sigma^2) &= -\log\left(P_c(N_c)\prod_{i=1}^{N} p(\mathbf{x}_n)\right) \\ &= E(\theta, \sigma^2) - \sum_{(j,k)\in N_c}\log(p_c(\mathbf{x}_j, \mathbf{y}_k)). \end{aligned} \qquad (9)$$

An objective function $\tilde{Q}$ can now be derived utilizing the same derivation as in Eq. (4). Rewriting the last term of Eq. (9) and leaving out the constants the modified objective function reads

$$\begin{aligned} \tilde{Q} &= Q + \frac{1}{2\alpha^2}\sum_{(j,k)\in N_c}\|\mathbf{x}_j - T(\mathbf{y}_k, \theta)\|^2 \\ &= Q + \frac{1}{2\alpha^2}\sum_{n=1}^{N}\sum_{m=1}^{M}\tilde{P}_{m,n}\|\mathbf{x}_n - T(\mathbf{y}_m, \theta)\|^2 \end{aligned} \qquad (10)$$

with the $M \times N$ matrix $\tilde{\mathbf{P}}$ of entries

$$\tilde{p}_{j,k} = \begin{cases} 1 & \text{for } (j,k) \in N_c \\ 0 & \text{else} \end{cases}. \qquad (11)$$

This matrix can be precomputed once. The EM algorithm of ECPD with correspondence priors now reads: 1) in the E-step compute the probability matrix (Eq. (5)); 2) in the M-step, the modified objective function (Eq. (10)) has to be minimized w.r.t. (θ, σ^2).

4.1. Non-rigid Point Set Registration

In the proposed approach the case of general non-rigid alignment can be addressed by defining the displacement field for $\mathbf{Y}$ as

$$T(\mathbf{Y}, v) = \mathbf{Y} + v(\mathbf{Y}). \qquad (12)$$

We build upon the regularizing prior on the displacement field of CPD [20]. In the Bayesian framework it can be formulated as $p(v) = \exp(-\frac{\lambda}{2}\phi(v))$ with a weighting parameter $\lambda \in \mathbb{R}$ and a regularization function $\phi(v)$. By multiplying the density function (1) with $p(v)$ or equivalently adding the exponent to the negative likelihood function (3) we integrate the regularizing prior into the GMM and obtain

$$f(v, \sigma^2) = E(v, \sigma^2) + \frac{\lambda}{2}\phi(v). \qquad (13)$$

In this work we extend this framework by including the additional correspondence priors (6) into Eq. (9) and obtain the following energy function

$$\tilde{f}(v, \sigma^2) = \tilde{E}(v, \sigma^2) + \frac{\lambda}{2}\phi(v). \tag{14}$$

To solve the M-step, the energy function (14) needs to be minimized w.r.t. v and σ^2. Following the same principles as in Eq. (4) we derive an upper bound for the energy function (14) that reads

$$\begin{aligned}\tilde{Q}(v, \sigma^2) = \\ \frac{1}{2\sigma^2}\sum_{n=1}^{N}\sum_{m=1}^{M} P^{old}(m|\mathbf{x}_n)\|\mathbf{x}_n - T(\mathbf{y}_m, v, \sigma^2)\|^2 \\ + \frac{1}{2\alpha^2}\sum_{n=1}^{N}\sum_{m=1}^{M} \tilde{P}_{m,n}\|\mathbf{x}_n - T(\mathbf{y}_m, v, \sigma^2)\|^2 \\ + \frac{N_P D}{2}\log\sigma^2 + \frac{\lambda}{2}\phi(v).\end{aligned} \tag{15}$$

The regularization function of CPD can be written in the Reproduction Kernel Hilbert Space as

$$\phi(v) = \int_{\mathbb{R}^D} \frac{|\tilde{v}(\mathbf{s})|^2}{\tilde{G}(\mathbf{s})}\, d\mathbf{s}. \tag{16}$$

Here, $\tilde{G}$ is the Fourier transform of a kernel function G which in turn describes a positive function that approaches zero as $\|\mathbf{s}\| \to \infty$. As in [20] we use a Gaussian kernel function $G(\mathbf{s}) = \exp(-\|\frac{\mathbf{s}}{\beta}\|^2)$. Furthermore, $\tilde{v}$ is the Fourier transform of the displacement field v in Eq. (12) and $\mathbf{s}$ is a frequency domain variable. This regularization can be interpreted as low-pass filtering with respect to the displacement field v. The filtered frequencies can be adjusted via the parameter β. To obtain v, we minimize the objective function (15) with respect to v by treating σ^2 as a constant. Note that due to our choice to integrate the correspondence priors (Eq. (6)) the minimizing function for the corresponding energy (14) still has the form of a radial basis function as in [21, 20], i.e.

$$v(\mathbf{z}) = \sum_{m=1}^{M} \mathbf{w}_m G(\mathbf{z} - \mathbf{y}_m) \tag{17}$$

(*we provide a proof as well as clarification of the symbol* $\mathbf{w}_m$ *in supplementary material*). To compute the displacement field $v(\mathbf{Y}) = \mathbf{GW}$ minimizing the energy (14) (where in our case $\mathbf{G}$ is a symmetric Gram matrix with entries $G_{i,j} = \exp(-\|\frac{\mathbf{y}_i - \mathbf{y}_j}{\beta}\|^2)$), we substitute v into Eq. (15) and leave out all terms independent of v. This yields

$$\begin{aligned}\tilde{Q}(\mathbf{W}) &= \frac{1}{2\sigma^2}\sum_{n=1}^{N}\sum_{m=1}^{M} P^{old}(m|\mathbf{x}_n)\|\mathbf{x}_n - T(\mathbf{y}_m, \mathbf{W})\|^2 \\ &+ \frac{1}{2\alpha^2}\sum_{n=1}^{N}\sum_{m=1}^{M} \tilde{P}_{m,n}\|\mathbf{x}_n - T(\mathbf{y}_m, \mathbf{W})\|^2 \\ &+ \frac{\lambda}{2}\mathbf{W}^T\mathbf{G}\mathbf{W}.\end{aligned} \tag{18}$$

Minimizing $\tilde{Q}$ with respect to $\mathbf{W}$ will minimize the energy function (14) with respect to v. Setting the derivative of $\tilde{Q}$ to zero in matrix form yields

$$\begin{aligned}\frac{\partial\tilde{Q}}{\partial\mathbf{W}} = \mathbf{G}\left(\frac{1}{\sigma^2}[d(\mathbf{P1})(\mathbf{Y} + \mathbf{GW}) - \mathbf{PX}]\right. \\ \left. + \frac{1}{\alpha^2}[d(\tilde{\mathbf{P}}\mathbf{1})(\mathbf{Y} + \mathbf{GW}) - \tilde{\mathbf{P}}\mathbf{X}]\right) + \lambda\mathbf{GW} = 0.\end{aligned} \tag{19}$$

Here, $d(\cdot)$ is a diagonal matrix. Multiplying the whole equation with $\mathbf{G}^{-1}\sigma^2$ (which always exists for a Gaussian kernel [21]) and rearranging it, we obtain

$$\begin{aligned}\left(d(\mathbf{P1})\mathbf{G} + \frac{\sigma^2}{\alpha^2}d(\tilde{\mathbf{P}}\mathbf{1})\mathbf{G} + \lambda\sigma^2\mathbf{I}\right)\mathbf{W} = \\ \mathbf{PX} - d(\mathbf{P1})\mathbf{Y} + \frac{\sigma^2}{\alpha^2}(\tilde{\mathbf{P}}\mathbf{X} - d(\tilde{\mathbf{P}}\mathbf{1})\mathbf{Y}).\end{aligned} \tag{20}$$

The transformed positions of $\mathbf{y}_m$ are found according to Eq. (12) as $T(\mathbf{Y}, \mathbf{W}) = \mathbf{Y} + \mathbf{GW}$. Thereafter, we obtain σ^2 by setting the corresponding derivative of Eq. (15) to zero. This yields the same result as in [20], i.e.

$$\begin{aligned}\sigma^2 = \frac{1}{N_P D}\sum_{n=1}^{N}\sum_{m=1}^{M}\|\mathbf{x}_n - T(\mathbf{y}_m, \mathbf{W})\|^2 \\ = \frac{1}{N_P D}(tr(\mathbf{X}^T d(\mathbf{P}^T\mathbf{1})\mathbf{X}) - 2tr((\mathbf{PX})^T\mathbf{T}) \\ + tr(\mathbf{T}^T d(\mathbf{P1})\mathbf{T})).\end{aligned} \tag{21}$$

ECPD is summarized in Algorithm 1. Convergence criteria are defined in terms of the maximum number of iterations and the smallest feasible step size of the EM algorithm.

5. Implementation

For fast registration of large point sets several approaches were introduced in [20] for CPD. First, the *Fast Gauss Transform* (FGT) allows to efficiently compute sums of exponentials in $\mathcal{O}(M + N)$ time. It can be applied in exactly the same way for ECPD. The second method is the *low-rank matrix approximation* of the Gram matrix $\hat{\mathbf{G}} = \mathbf{Q}\Lambda\mathbf{Q}^T$, where $\Lambda_{K\times K}$ denotes the diagonal matrix of the largest eigenvalues and $\mathbf{Q}_{M\times K}$ the corresponding

Algorithm 1 ECPD with Correspondence Priors

Input: Point sets $\mathbf{X}, \mathbf{Y}$ and index set for correspondence priors N_c
Output: Aligned point set $T(\mathbf{Y})$

- **Initialization:** $\mathbf{W} = 0\,, \sigma^2 = \frac{1}{DMN}\sum_{m,n=1}^{M,N}\|\mathbf{x}_n - \mathbf{y}_m\|^2$, $0 \leq w \leq 1, \beta > 0, \lambda > 0$.
  Construct $\mathbf{G}$: $G_{i,j} = \exp^{(-\frac{1}{2\beta^2}\|\mathbf{y}_i - \mathbf{y}_j\|^2)}$,
  Precompute $\tilde{\mathbf{P}}$, as in Eq. (11).
- **EM optimization**, repeat until convergence:
  - E-Step: compute $\mathbf{P}$, with $p_{m,n} = P^{old}(m|\mathbf{x}_n)$ as in Eq. (5).
  - M-Step: Solve for T, σ^2.
  - Solve Eq. (20) with respect to $\mathbf{W}$.
  - $N_P = \mathbf{1}^T\mathbf{P}\mathbf{1}, \mathbf{T} = \mathbf{Y} + \mathbf{G}\mathbf{W}$.
  - $\sigma^2 = \frac{1}{N_P D}[tr(\mathbf{X}^T d(\mathbf{P}^T\mathbf{1})\mathbf{X}) - 2tr((\mathbf{P}\mathbf{X})^T\mathbf{T}) + tr(\mathbf{T}^T d(\mathbf{P}\mathbf{1})\mathbf{T})]$.
- The aligned point set is $T(\mathbf{Y}, \mathbf{W}) = \mathbf{Y} + \mathbf{G}\mathbf{W}$.
- The probability for correspondences is given by $\mathbf{P}$.

eigenvectors in matrix form. Also the Woodbury matrix identity can be applied to efficiently solve the linear system for the transformation $\mathbf{W}$. In our case we need to solve the modified system (20) and rewrite it as follows

$$\underbrace{([d(\mathbf{P}\mathbf{1}) + \frac{\sigma^2}{\alpha^2}d(\tilde{\mathbf{P}}\mathbf{1})]}_{d(\hat{\mathbf{P}}\mathbf{1})}\mathbf{G} + \lambda\sigma^2\mathbf{I})\mathbf{W} =$$

$$\underbrace{\mathbf{P}\mathbf{X} - d(\mathbf{P}\mathbf{1})\mathbf{Y} + \frac{\sigma^2}{\alpha^2}(\tilde{\mathbf{P}}\mathbf{X} - d(\tilde{\mathbf{P}}\mathbf{1})\mathbf{Y})}_{F}$$

$$\iff (\mathbf{G} + \lambda\sigma^2 d(\hat{\mathbf{P}}\mathbf{1})^{-1})\mathbf{W} = d(\hat{\mathbf{P}}\mathbf{1})^{-1}F. \quad (22)$$

Therefore, the Woodbury matrix identity for solving the modified system (20) with the low-rank matrix approximation reads

$$(\mathbf{Q}\Lambda\mathbf{Q}^T + \lambda\sigma^2 d(\hat{\mathbf{P}}\mathbf{1})^{-1})^{-1} =$$
$$\frac{1}{\lambda\sigma^2}\left[F - d(\hat{\mathbf{P}}\mathbf{1})\mathbf{Q}\left(\lambda\sigma^2\Lambda^{-1} + \mathbf{Q}^T d(\hat{\mathbf{P}}\mathbf{1})\mathbf{Q}\right)^{-1}\mathbf{Q}^T F\right).$$

We aim at the non-rigid registration of large point sets ($> 10^5$ points) and thus opt for an optimized implementation of ECPD on a system with a multicore CPU and a GPU. The FGT was implemented as a heterogeneous function applying acceleration techniques similar to those described in [10]. Here, the tangible impact on the performance yields both the enabling of high occupancy and extensive use of faster shared and register memory for reusable data. For computation of the eigenvalue decomposition, which the low-rank approximation of the Gram matrix G is based on, we rely on Implicitly Restarted Lanczos Method [7] (provided by the library ARPACK [1]) in combination with the FGT. In the final iterations of the EM optimization — when the width of Gaussians becomes small — the algorithm switches to the truncated Gaussian approximation mode (truncated mode) and the complexity increases to $\mathcal{O}(MN)$. The acceleration of this mode is achieved by taking advantage of the parallel portable shared memory programming model offered by OpenMP [22]. Even though the implementation of the truncated mode is scalable with the number of multiprocessors, the improvement on a multicore CPU is still not sufficient to compensate for the increase in the computation time caused by the $\mathcal{O}(MN)$ complexity.

5.1. Coarse-To-Fine Strategy

To reduce the registration time further, we introduce the procedure of *correspondence preserving subsampling* (CPS) and adopt a coarse-to-fine strategy. CPS is a joint subsampling procedure of two (generally k) point sets which reduces the number of points in one or several point sets according to some subsampling rule (e.g. uniformly) and guarantees to preserve established correspondence priors between those sets. Suppose n_c to be the number of correspondence priors between the point sets and t the subsampling factor. Then, the subsampled template contains at most $\lfloor\frac{N}{t}\rfloor + n_c$ points.
ECPD in subsampled mode operates as follows: 1) perform the CPS on the template point set; 2) register the subsampled template point set with the reference point set; 3) register the registered subsampled template (result of the second step) as a reference with the initial dense template. Thereby amount of points in the result remains unchanged, independently of the subsampling factor. We observe an important property of the CPS. Since correspondences between the subsampled template and the initial template are known, they can be taken as strong correspondence priors on the final registration step. In the following we give a rough comparison of a number of operations required to register a given dataset with and without CPS.
For the number of operations T_{plain} required by the plain non-rigid registration it holds

$$T_{plain} = c_1 MN, \quad (23)$$

and for the number of operations $T_{subs.}$ required by ECPD with subsampling it holds

$$T_{subs.} = c_2\Big(M \cdot \frac{N}{t} + \frac{N}{t} \cdot N\Big) = c_2\frac{N(M+N)}{t}, \quad (24)$$

where c_1 and c_2 are constants. Thus, the speedup of the algorithm with subsampling can be estimated as

$$s(m,n,t) = \frac{T_{plain}}{T_{subs.}} = \frac{c_1 MN}{\frac{c_2 N(M+N)}{t}} = c_s\frac{Mt}{M+N} \quad (25)$$

To achieve a speedup s, according to Eq. (25), the subsampling factor t should be chosen as $\lceil\frac{s(M+N)}{M}\rceil$. Note that the

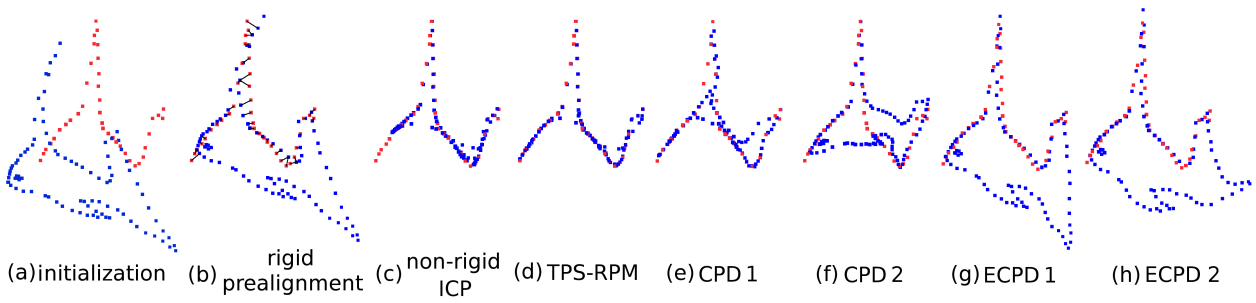

Figure 2: Non-rigid registration of a 2D "Fish" dataset. The reference is shown in red, the template is shown in blue. (a): initialization of the point sets; (b): rigid prealignment with established correspondence priors; note that rigid CPD recovers prior correspondences by far not correct; results of the registrations with non-rigid ICP [11] (c); TPS-RPM [9] (d); CPD [20] with two different parameter sets: $\beta = 1.0$, $\lambda = 1.0$ (e) and $\beta = 2.0$, $\lambda = 2.0$ (f); proposed approach ECPD with two different parameter sets: $\beta = 14.0$, $\lambda = 14.0$, $\alpha = 10^{-8}$ (g) and $\beta = 8.0$, $\lambda = 0.8$, $\alpha = 10^{-8}$ (h).

given estimate does not consider the constant time spent for the template subsampling and approximation of the Gram matrix G for the template.

6. Evaluation

We run ECPD on a system with 32 GB RAM, Intel Xeon processor and NVIDIA GTX 660 Ti graphics card and evaluate it on synthetic and real data. The experiments are designed to verify the main advantages of ECPD — improvement of registration quality — both in the presence of structured outliers and without those, when correspondence priors are available. Where appropriate we compare our approach against CPD (which we consider the baseline) and other non-rigid registration methods.

Experiments on Synthetic Data. In our first experiment we compare the performances of ECPD, CPD as well as two widely used non-rigid registration methods (non-rigid ICP [11] and TPS-RPM [9]) on a synthetic 2D "Fish" dataset (Fig. 2, (a)). Implementations of the latter ones are taken from [9] and parameters of the algorithms are not altered. The reference point set represents a non-rigidly transformed section of the template point set. Accordingly, not all points of the template have valid correspondences in the reference point set. Points without valid correspondences are outliers which are not uniformly distributed, but are rather structured. We are especially interested in preserving the topology of these points. In the preprocessing step the point sets were prealigned using rigid CPD [20]. The rigid registration produces a rough estimation of correspondences (Fig. 2, (b)) and its outcome serves as an input for the non-rigid registrations. Fig. 2, (c), (d) shows registration results of non-rigid ICP and TPS-RPM algorithms respectively. Both algorithms minimize energy functions without further constraints on the non-rigid transformations. This circumstance leads to flattening and breaking of the topology of the template point set. The experiments show that these algorithms perform poor in the presence of structured outliers. Segmentation of the point sets could improve performance in this case, but would assume outlier detection beforehand. In opposite, the coherency constraint valid both in CPD and ECPD, preserves the outliers from breaking up the topology. Though in the case of CPD, either the correspondences between inliers are assigned precisely while the topology of the outliers is broken (Fig. 2, (e)); or the topology of the outliers is preserved while a significant part of correspondences between the inliers still is not assigned precisely (Fig. 2, (f)). Through the synergetic effect of its constraints on the displacement field, ECPD is able to recover the non-rigid transformation between the inliers reliably (Fig. 2, (g), (h)). The synergetic effect originates in simultaneous impact of appropriately weighted coherency constraint and prior correspondences. Crucial is that structured outliers are modelled as a part of the point sets. Prior correspondences enforce the registration to restrict itself to predefined areas, decreasing the influence of structured outliers. Concurrently, the coherency constraint insures point topology to be preserved in all regions.

The above experiment was designed to illustrate the advantage of ECPD against other non-rigid registration methods in presence of structured outliers. The parameters were chosen for emphasizing particular effects peculiar to the algorithms in general. The correspondences obtained through rigid prealignment were assumed to be reliable. This assumption of course may not always be valid in practice. Prior knowledge about the underlying point sets (e.g. shape prior) often allows to extract a sparse set of correspondences which is eventually used to prealign point sets rigidly. However, robust correspondence priors establishment is application specific. Most registration algorithms handle correspondences established through rigid prealignment as correspondence priors implicitly. For instance, non-rigid ICP assigns higher weights to those points from the beginning on. In opposite, ECPD can additionally incorporate an arbitrary set of correspondences explicitly.

Experiments on Real Data. In the second experiment we show how embedding the correspondence priors improves

the accuracy of non-rigid registrations in real world applications. We register 3D scans of arms in different poses. The scans were taken from [2]. First, initial correspondences between the scans are precomputed (Fig. 1, (a)). For this purpose we use an approach similar to those described in [25]. We extract 3D keypoints on the point clouds with an Intrinsic Shape Signature descriptor [32], determine Persistent Feature Histograms (PFH) [25] at those keypoints and establish correspondences by comparing the PFH's. In the next step we register the scans with CPD and ECPD. In both cases the established correspondences are used to rigidly prealign the point clouds, whereas in the case of ECPD the correspondences are additionally applied as correspondence priors. In the case of CPD we show the result with $\beta = 8.0$, $\lambda = 8.0$ and in case of ECPD with $\beta = 1.0$, $\lambda = 1.0$, $\alpha = 10^{-2}$ and $t = 20$ (Fig. 1, (b) and (c) respectively). Note that for CPD this is the best achieved result (with the smallest cloud-to-cloud mean distance) and parameters β and λ differ from those of ECPD. Again, in the case of CPD only the coherency constraint is applied and complex non-rigid deformations (combination of supination, flection and abduction of the arm) in the area of the hand can not be resolved. Employing the correspondence priors improves registration accuracy significantly. Moreover, performance of ECPD is less sensitive to the selection of algorithm's parameters in the latter case as was ascertained in the course of the experiment. Fig. 3 displays the comparisons of the arm registrations in the area of the hands in more detail.

In another experiment we show an application of ECPD in the scenario of scan-template registration of human heads recovered with [12] from multiple views under real-world conditions. Point sets representing scans of human heads exhibit a high variety compared to other body parts. Often they contain areas with structured outliers (parts of the clothes, hair). We are especially interested in high registration quality in the facial area. The first dataset we analyse is the "woman with a scarf" shown in Fig. 4. The scan and the template contain $2.9 \cdot 10^5$ and $9 \cdot 10^4$ points respectively. After prealignment with rigid CPD we run CPD and ECPD with correspondence priors. We obtain correspondence priors between the scan and the template in the following way. Positions of the 3D keypoints on the template are known and the template is kept the same for all registrations (shown in Fig. 4, (c)). For detecting 3D keypoints on the scan we start with extracting 2D facial keypoints with the Chehra face tracker [4] on one of the frontal views (Fig. 4, (a)). We obtain normalized texture coordinates by dividing 2D keypoint coordinates by extent of the image in respective dimensions and project the frontal view onto the mesh (we know the projection matrices from the multiview setting). Next, we determine a vertex with texture coordinate, closest to the keypoint's one, and retrieve the corresponding 3D keypoint on the scan (Fig. 4, (b)). Finally, the correspondences are established, since the order of extracted keypoints both on the scan and on the template is known. The result of the registration with CPD can be observed in the Fig. 4, (d) and (e). As expected, the method is not able to register the facial area accurately, since the outliers are not uniformly distributed. Eventually we registered the point clouds with ECPD with facial correspondence priors. The results can be observed in Fig. 4, (f). We notice significantly enhanced alignment precision in the ROI.

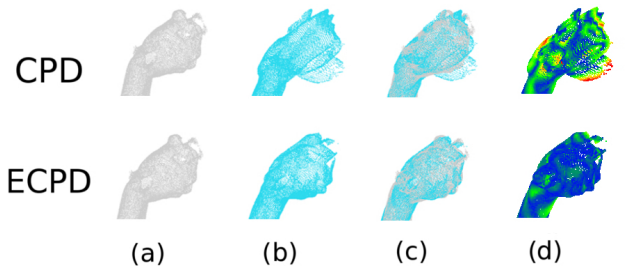

Figure 3: Comparison of the registration results of the arms in the area of the hands with CPD (top row) and ECPD (bottom row) methods. (a): reference; (b): registered template; (b): overlapped view ((a) + (b)); (d): cloud-to-cloud distance in Blue<Green<Yellow<Red scale. Red corresponds to 0.0171 distance units. The mean distance and standard deviation amounts to (0.0029; 0.003) distance units for CPD and (0.0024; 0.0015) distance units for ECPD.

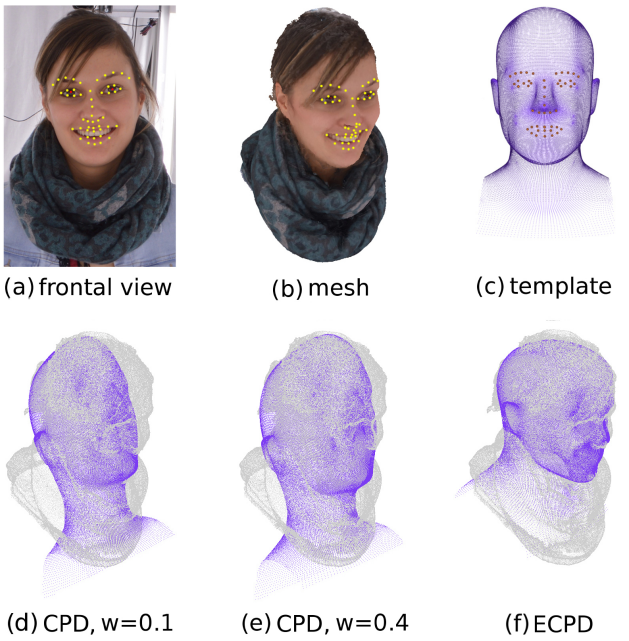

Figure 4: Scan-template non-rigid registration of the "woman with a scarf" dataset. (a): keypoints, extracted on the frontal view; (b): keypoints, transferred to the mesh as 3D keypoints; (c): template with predefined 3D keypoints; (d), (e): results of the registration with CPD, $w = 0.1$ and $w = 0.4$ respectively; (f): result of ECPD with correspondence priors, $\alpha = 10^{-6}$. In this experiment $\beta = 20.5$ and $\lambda = 20.5$ are the same for CPD and ECPD. The mean distance and standard deviation amounts to (0.05; 0.043), (0.048; 0.046) and (0.023; 0.024) distance units for (d), (e) and (f) respectively.

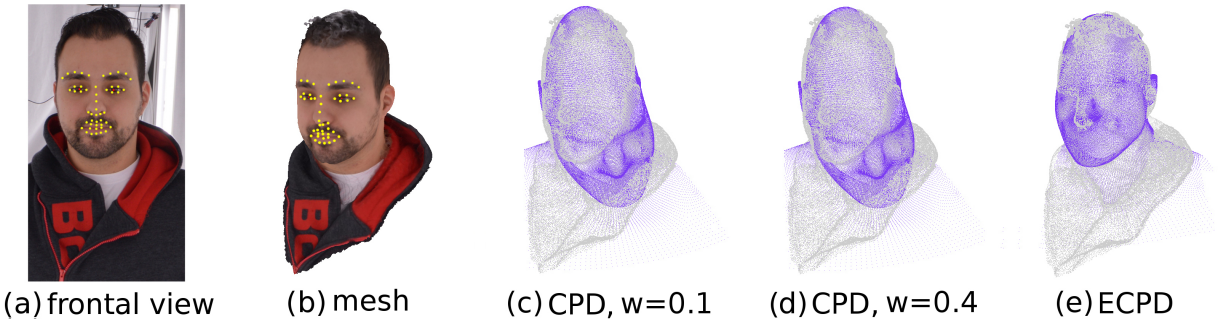

Figure 5: Scan-template non-rigid registration of the "man with a hood" dataset. (a): keypoints, extracted on the frontal view; (b): keypoints, transferred to the mesh as 3D keypoints; (c), (d): results of the registration with CPD, $w = 0.1$ and 0.4 respectively; (e): result of ECPD with correspondence priors, $\alpha = 10^{-5}$. In this expermient $\beta = 8.0$ and $\lambda = 16.0$ are the same for CPD and ECPD. The mean distance and standard deviation amounts to (0.049; 0.047), (0.048; 0.047) and (0.019; 0.018) distance units for (c), (d) and (e) respectively.

We also evaluate the performance of ECPD on varying values of the subsampling factor t. We measure runtimes of the algorithm and compute the speedup and cloud-to-cloud metrics (mean distance error and standard deviation) of the results depending on different t. The speedup is measured as ratio between the runtime of ECPD without subsampling and the runtime of ECPD with a particular value of t.

subs. factor	runtime, sec	speedup		mean distance	std. deviation
		estimated	achieved		
1	4124.0	1.0	1.0	**0.0228**	**0.0240**
4	1360.2	2.99	3.03	**0.0226**	**0.0236**
8	576.4	5.98	7.15	0.0235	0.0243
16	338.78	11.97	12.17	**0.0225**	**0.0240**
32	196.37	23.97	21.00	0.0236	0.0241
64	122.13	47.88	33.77	0.0236	0.0248
128	81.08	95.75	50.86	**0.0293**	**0.0339**

Table 1: Speedup of ECPD as a function of the subsampling factor for the "woman with a scarf" dataset

As expected, a discrepancy between estimated and achieved speedup is observed (see Table 1). For low values of t (4-16) the achieved speedup is larger than estimated, as (25) does not consider convergence criteria hidden in the multiplicative factor. For larger values of t (32-128) the speedup is smaller than estimated. With increasing values of t the time spent for subsampling and computation of the Gram matrix (which the speedup estimation (25) does not consider as well) is also increasing, relative to the algorithm's core computation time. For t varying from 4 to 64 results are qualitatively similar. Indeed, the smallest mean distance and sigma values are obtained with $t = 16$ and not without subsampling ($t = 1$) as one might expect. As the value of t exceeds 128, the amount of points in the subsampled template is not sufficient to capture the variation of the reference point cloud and the mean error/sigma increases. An optimal value of t in this experiment amounted to 64 leading to more than a thirtyfold speedup.
Another challenging dataset in this experiment was the "man with a hood" (the scan contained $7.65 \cdot 10^4$ points). It was processed analogously and we show results in Fig. 5.

Choosing Parameters. Proper parameterization is crucial for desired registration results. The more different the registered point sets, the smaller should be β allowing higher variety in deformations. λ balances the relative impact of the coherency constraint and the energy function with prior correspondences. Empirically, setting $\beta = \lambda$ often works well. We typically choose $\beta \in [1, 20]$ and $\alpha \in [10^{-8}, 1]$. The smaller value of α is most suited for highly reliable prior correspondences, whereas 1 depreciates their influence.

7. Conclusions

We propose the non-rigid point set registration algorithm Extended Coherent Point Drift. To the best of our knowledge, this is the first non-rigid point set registration algorithm, at least for the probabilistic case, which allows to embed correspondence priors in a closed-form. ECPD employs correspondence preserving subsampling counterbalancing the polynomial complexity by splitting the problem into two subproblems of smaller size and reducing the number of operations by a linear factor. Our experiments demonstrated that in applications where correspondence priors are available, ECPD significantly improves registration quality and enhances robustness of point set registration under complex non-rigid deformations (Fig. 1) and in presence of structured outliers (Fig. 4), compared to the baseline method CPD. In future work we plan to investigate automatic parameter selection as well as combine key point extractors and large-scale non-rigid registration with the dynamic vectorial α into a single framework.

Acknowledgments

Development of ECPD was partially funded by the EU 7th Framework Programme project AlterEgo (600610) and by the BMBF project DYNAMICS (01IW15003).

References

[1] ARPACK - Arnoldi Package. http://www.caam.rice.edu/software/ARPACK/. Accessed: 23.03.2015.

[2] B. Allen, B. Curless, and Z. Popović. Articulated body deformation from range scan data. *ACM SIGGRAPH 2002*, 21(3):612–619, 2002.

[3] B. Amberg, S. Romdhani, and T. Vetter. Optimal step nonrigid icp algorithms for surface registration. In *CVPR*, pages 1–8, June 2007.

[4] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. Incremental face alignment in the wild. In *CVPR 2014*, 2014.

[5] F. Bogo, J. Romero, M. Loper, and M. J. Black. FAUST: Dataset and evaluation for 3D mesh registration. In *CVPR*, 2014.

[6] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *T-PAMI*, 11(6):567–585, June 1989.

[7] D. Calvetti, L. Reichel, and D. Sorensen. An implicitly restarted lanczos method for large symmetric eigenvalue problems. *ETNA*, (2):1–21, 1994.

[8] H. Chui and A. Rangarajan. A feature registration framework using mixture models. In *Mathematical Methods in Biomedical Image Analysis*, pages 190–197. IEEE, 2000.

[9] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.*, 89(2-3):114–141, Feb. 2003.

[10] F. A. Cruz, S. K. Layton, and L. A. Barba. How to obtain efficient gpu kernels: an illustration using FMM & FGT. *Comput. Phys. Commun.*, (10):2084–2098, 2011.

[11] A. W. Fitzgibbon. Robust registration of 2D and 3D point sets. In *BMVC*, pages 662–670, 2001.

[12] Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *T-PAMI*, 32(8):1362–1376, Aug. 2010.

[13] S. Ge, G. Fan, and M. Ding. Non-rigid point set registration with global-local topology preservation. In *CVPR Workshops*, June 2014.

[14] G. E. Hinton, C. K. I. Williams, and M. D. Revow. Adaptive elastic models for hand-printed character recognition. In *NIPS*, pages 512–520. 1992.

[15] D. A. Hirshberg, M. Loper, E. Rachlin, and M. J. Black. Coregistration: Simultaneous alignment and modeling of articulated 3d shape. In *ECCV*, volume 7577 of *Lecture Notes in Computer Science*, pages 242–255. 2012.

[16] B. Jian and B. C. Vemuri. A robust algorithm for point set registration using mixture of gaussians. In *ICCV*, pages 1246–1251, 2005.

[17] A. Kovnatsky, M. M. Bronstein, A. M. Bronstein, K. Glashoff, and R. Kimmel. Coupled quasi-harmonic bases. *CoRR*, abs/1210.0026, 2012.

[18] M. Moradi and P. Abolmaesumi. Medical image registration based on distinctive image features from scale-invariant (sift) key-points. In *Inter'l Congress and Exhib. on Computer Assisted Radiology and Surgery*, volume 1281, 2005.

[19] C. Mourning, S. Nykl, H. Xu, D. M. Chelberg, and J. Liu. GPU acceleration of robust point matching. In *ISVC*, pages 417–426, 2010.

[20] A. Myronenko and X. Song. Point-set registration: Coherent point drift. *T-PAMI*, 2010.

[21] A. Myronenko, X. Song, and M. A. Carreira-Perpiñán. Nonrigid point set registration: Coherent point drift. In *NIPS*, pages 1009–1016. MIT Press, 2006.

[22] OpenMP Architecture Review Board. OpenMP application program interface version 3.0, May 2008.

[23] M. Revow, C. K. I. Williams, and G. E. Hinton. Using generative models for handwritten digit recognition. *T-PAMI*, 18:592–606, 1996.

[24] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *3DIM*, june 2001.

[25] R. B. Rusu, N. Blodow, Z. C. Marton, and M. Beetz. Aligning point cloud views using persistent feature histograms. In *IROS*, 2008.

[26] G. K. L. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X. Sun, and P. L. Rosin. Registration of 3d point clouds and meshes: A survey from rigid to nonrigid. *IEEE Trans. Vis. Comput. Graph.*, 19(7):1199–1217, 2013.

[27] Y. Tsin and T. Kanade. A correlation-based approach to robust point set registration. In *ECCV*, volume 3023 of *Lecture Notes in Computer Science*, pages 558–569. 2004.

[28] P. Wang, P. Wang, Z. Qu, Y. Gao, and Z. Shen. A refined coherent point drift (cpd) algorithm for point set registration. *Science China Information Sciences*, 54(12):2639–2646, 2011.

[29] A. Yuille and N. Grzywacz. Motion coherence theory. In *ICCV*, pages 344–353, Dec 1988.

[30] A. L. Yuille and N. M. Grzywacz. A mathematical analysis of the motion coherence theory. *IJCV*, 3(2):155–175, 1989.

[31] R. Zhang, W. Zhou, Y. Li, S. Yu, and Y. Xie. Nonrigid registration of lung ct images based on tissue features. 2013, 2013.

[32] Y. Zhong. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In *ICCV Workshops*.

[33] Z. Zhou, J. Zheng, Y. Dai, Z. Zhou, and S. Chen. Robust nonrigid point set registration using student's-t mixture model. *PLoS ONE*, 9(3):e91381, 03 2014.