

# A Closest Point Proposal for MCMC-based Probabilistic Surface Registration

Dennis Madsen\*, Andreas Morel-Forster\*, Patrick Kahr, Dana Rahbani,  
Thomas Vetter, and Marcel Lüthi

Department of Mathematics and Computer Science, University of Basel,  
Basel, Switzerland

{dennis.madsen, andreas.forster, patrick.kahr, dana.rahbani,  
thomas.vetter, marcel.luethi}@unibas.ch

**Abstract.** We propose to view non-rigid surface registration as a probabilistic inference problem. Given a target surface, we estimate the posterior distribution of surface registrations. We demonstrate how the posterior distribution can be used to build shape models that generalize better and show how to visualize the uncertainty in the established correspondence. Furthermore, in a reconstruction task, we show how to estimate the posterior distribution of missing data without assuming a fixed point-to-point correspondence.

We introduce the closest-point proposal for the Metropolis-Hastings algorithm. Our proposal overcomes the limitation of slow convergence compared to a random-walk strategy. As the algorithm decouples inference from modeling the posterior using a propose-and-verify scheme, we show how to choose different distance measures for the likelihood model.

All presented results are fully reproducible using publicly available data and our open-source implementation of the registration framework.

**Keywords:** Probabilistic registration, Gaussian Process Morphable Model, Metropolis-Hastings Proposal, Point Distribution Model

## 1 Introduction

The ability to quantify the uncertainty of a surface registration is important in many areas of shape analysis. It is especially useful for the reconstruction of partial data, or for the analysis of data where the exact correspondence is unclear, such as smooth surfaces. Within the medical area, the uncertainty of a partial data reconstruction is needed to make informed surgical decisions [27]. Uncertainty estimates can also be used to build better generalizing Point Distribution Models (PDMs) by assigning an uncertainty measure to each landmark [20, 12].

In this paper, we propose an efficient, fully probabilistic method for surface registration based on Metropolis-Hastings (MH) sampling. Efficiency is gained

---

\* Joint first authors

Code available at <https://github.com/unibas-gravis/icp-proposal>

by introducing a specialized proposal based on finding the closest points between a template and a target mesh. With this, we benefit from the geometry-aware proposal, while at the same time obtaining an uncertainty estimate for the registration result. We formulate the non-rigid surface registration problem as an approximation of the posterior distribution over all possible instances of point-to-point correspondences, given a target surface. With this approach, the registration uncertainty is the remaining variance in the posterior distribution. We use the MH algorithm to sample surface registrations from the posterior distribution. Our method can escape local optima and aims to capture the full posterior distribution of registrations.

Our method improves on previous works in the literature in different ways. In [28, 17], the MH algorithm is used to estimate the uncertainty in non-rigid registration. These papers are working on the image domain, and are not transferable to the surface registration setting. MH has also been used in [21] to fit an Active Shape Model to images and in [30] to fit a Morphable Face Model to an image, both of which only make use of the framework to avoid local optima and to easily integrate different likelihood terms. The main problem with the MH algorithm is the commonly used random-walk approach, which suffers from very long convergence times when working in high-dimensional parameter spaces. To overcome this problem, informed proposal distributions can be designed to perform directed sample updates. In [16], a Bayesian Neural Network is learned to produce informed samples for the MH framework in the case of 3D face reconstruction from a 2D image. This, however, requires training a neural network for each class of shapes to be registered. In [14], local random-walk is combined with an image-dependent global proposal distribution. This image dependent distribution is, however, not directly transferable to the problem of surface registration. We perform registration by warping a single template mesh to a target surface. No training data is required with our method, whereas most state-of-the-art neural-network-based non-rigid registration methods require thousands of meshes in correspondence for training [6, 10]. These methods work well for newer non-rigid registration challenges such as the FAUST dataset [4], where the focus is to learn shape articulation from a training dataset. This is not something we would advocate using our method for. Our method targets settings with no, or limited, available training data and adds to existing methods by providing an uncertainty measure, which is especially important within the medical domain.

Even though newer registration methods exist, Iterative Closest Point (ICP) and Coherent Point Drift (CPD) are still most commonly used for surface registration scenarios without available training data.<sup>2</sup> The ICP algorithm was originally developed for rigid alignment of point sets [5, 3]. ICP iteratively estimates the point-to-point correspondences between two surfaces and then computes a transformation based on the established correspondence. Finally, this transformation is applied to the reference surface. The algorithm has later been modified

<sup>2</sup> In this paper, we focus on non-rigid registration. We, therefore, refer to their non-rigid versions whenever ICP or CPD is mentioned.

for non-rigid surface registration [7]. In general, standard ICP is very efficient and produces good results. However, a bad initialization may lead the algorithm to local optima from which it is unable to recover. Moreover, it is impossible to estimate the correspondence uncertainty of the registration result. Multiple extensions have been proposed to make the algorithm more robust in scenarios such as missing data or large articulation deformation differences [1, 13, 25]. In [18], the ICP method is improved using Simulated Annealing and MCMC. Their method is a robust version of the ICP algorithm, which can find the global optimal rigid registration of point clouds. They do not, however, measure the registration uncertainty, nor are they able to perform non-rigid registrations. An extensive review of different ICP methods can be found in [26]. The CPD method [23, 22] is a probabilistic alternative to ICP. Yet, it does not provide an uncertainty estimate for the registration result.

As it is common in non-rigid registration, we do not infer the rigid alignment. Fortunately, MH allows for easy integration of proposal distributions of parameters other than the shape. The proposal distribution can therefore easily be extended to include translation, rotation, scaling as shown in [21], or texture, illumination, and camera position as in [30].

In this paper, we show how closest-point information can be incorporated into the MH algorithm to take geometric information into account. We introduce a novel closest-point-proposal (CP-proposal) to use within the MH algorithm. Our proposal can make informed updates while maintaining the theoretical convergence properties of the MH algorithm. To propose probabilistic surface deformations, we make use of a Gaussian Process Morphable Model (GPMM) as our prior model. In [8], GPMMs are applied to face surface registration. However, the authors formulate the registration problem as a parametric registration problem, which does not provide an uncertainty measure for the established correspondence. Several alternatives already exist to the random-walk proposal, such as MALA [9], Hamiltonian [24] or NUTS [11]. While the mentioned proposals work well in lower-dimensional spaces and for smoother posteriors, we experienced that they get computationally demanding in high dimensional spaces and have problems when the posterior is far from smooth.

In our experiments, we register femur bones, where the biggest challenge is establishing correspondence along the long smooth surface of the femur shaft. In another experiment, we use our method to reconstruct missing data. To that end, we compute the posterior distribution of registrations for faces where the nose has been removed. Unlike ICP or CPD, we can give an uncertainty estimate for each point in the surface reconstruction. Furthermore, we show how the standard non-rigid ICP algorithm can end up in local optima, while our method consistently provides good registrations and can also quantify the correspondence uncertainty. The three main contributions of this paper are:

- We introduce a theoretically sound, informed closest-point-proposal for the MH framework, with a well-defined transition ratio, section 3.2.
- We show the usefulness of the posterior distribution of registrations for completing partial shapes and for building better generalizing PDMs, section 4.2.

- We demonstrate that the MH algorithm with our proposal leads to better and more robust registration results than the standard non-rigid ICP and CPD algorithms, section 4.3.

## 2 Background

In this section, we formally introduce the GPMM as presented in [19], and we show how the analytic posterior, which we use in our CP-proposal, is computed.

### 2.1 Gaussian Process Morphable Model (GPMM)

GPMMs are a generalization of the classical point distribution models (PDMs) [19]. The idea is to model the deformations, which relate a given reference shape to the other shapes within a given shape family, using a Gaussian process. More formally, let  $\Gamma_R \subset \mathbb{R}^3$  be a reference surface. We obtain a probabilistic model of possible target surfaces  $\Gamma$  by setting

$$\Gamma = \{x + u(x) | x \in \Gamma_R\} \quad (1)$$

where the deformation field  $u$  is distributed according to a Gaussian process  $u \sim GP(\mu, k)$  with mean function  $\mu$  and covariance function  $k : \Gamma_R \times \Gamma_R \rightarrow \mathbb{R}^{3 \times 3}$ . Let  $\Gamma_1, \dots, \Gamma_n$  be a set of surfaces for which correspondence to the reference surface  $\Gamma_R$  is known. From this, it follows that we can express any surface  $\Gamma_i$  as:

$$\Gamma_i = \{x + u_i(x) | x \in \Gamma_R\}. \quad (2)$$

Classical PDMs define the mean and covariance function as:

$$\begin{aligned} \mu_{\text{PDM}}(x) &= \frac{1}{n} \sum_{i=1}^n u_i(x) \\ k_{\text{PDM}}(x, x') &= \frac{1}{n-1} \sum_{i=1}^n (u_i(x) - \mu_{\text{PDM}}(x))(u_i(x') - \mu_{\text{PDM}}(x'))^T. \end{aligned} \quad (3)$$

However, GPMMs also allow us to define the mean and covariance function analytically. Different choices of covariance functions lead to different well-known deformation models, such as radial basis functions, b-splines, or thin-plate-splines. To model smooth deformations, we choose a zero-mean Gaussian process with the following covariance function:

$$k(x, x') = g(x, x') * I_3, \quad (4)$$

$$g(x, x') = s \cdot \exp\left(\frac{-\|x - x'\|^2}{\sigma^2}\right), \quad (5)$$

where  $I_3$  is the identity matrix and  $g(x, x')$  is a Gaussian kernel.

The model, as stated above, is a possibly infinite-dimensional non-parametric model. In [19], they propose to use the truncated Karhunen-Loève expansion to

obtain a low-rank approximation of the Gaussian process. In this representation, the Gaussian process  $GP(\mu, k)$  is approximated as:

$$u[\boldsymbol{\alpha}](x) = \mu(x) + \sum_{i=1}^r \alpha_i \sqrt{\lambda_i} \phi_i(x), \quad \alpha_i \sim \mathcal{N}(0, 1) \quad (6)$$

where  $r$  is the number of basis functions used in the approximation and  $\lambda_i, \phi_i$  are the  $i$ -th eigenvalue and eigenfunction of the covariance operator associated with the covariance function  $k$ . Consequently, any deformation  $u$  is uniquely determined by a coefficient vector  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)$ . Hence, we can easily express any surface  $\Gamma$  as:

$$\Gamma[\boldsymbol{\alpha}] = \{x + \mu(x) + \sum_{i=1}^r \alpha_i \sqrt{\lambda_i} \phi_i(x) | x \in \Gamma_R\} \quad (7)$$

with associated probability

$$p(\Gamma[\boldsymbol{\alpha}]) = p(\boldsymbol{\alpha}) = (2\pi)^{-\frac{r}{2}} \exp(-\|\boldsymbol{\alpha}\|^2). \quad (8)$$

## 2.2 Analytical Posterior Model

GPMMs make it simple and efficient to constrain a model to match known correspondences, such as user annotations or the estimated correspondence from taking the closest point. Indeed, the corresponding *posterior model* is again a Gaussian process, whose parameters are known in closed form.

Let  $u \sim GP(\mu, k)$  be a GPMM and  $\epsilon \sim \mathcal{N}(0, \Sigma)$ ,  $\Sigma = \sigma_{noise} I_3$  be the certainty of each known landmark. Every landmark  $l_R$  on the reference surface can then be matched with its corresponding landmark  $l_T$  on the target. The set  $L$  consists of the  $n$  reference landmarks and its expected deformation to match the target

$$L = \{(l_R^1, l_T^1 - l_R^1), \dots, (l_R^n, l_T^n - l_R^n)\} = \{(l_R^1, \hat{u}^1), \dots, (l_R^n, \hat{u}^n)\}, \quad (9)$$

with  $\hat{u}$  being subject to Gaussian noise  $\epsilon$ . Using Gaussian process regression, we obtain the posterior model  $u_p \sim GP(\mu_p, k_p)$ , which models the possible surface deformations that are consistent with the given landmarks. Its mean and covariance function are given by:

$$\begin{aligned} \mu_p(x) &= \mu(x) + K_X(x)^T (K_{XX} + \epsilon)^{-1} \hat{U} \\ k_p(x, x') &= k(x, x') + K_X(x)^T (K_{XX} + \epsilon)^{-1} K_X(x'). \end{aligned} \quad (10)$$

Here we defined  $K_X(x) = (k(l_R^i, x))_{i=1, \dots, n}$ , a vector of the target deformation as  $\hat{U} = (l_T^i - l_R^i)_{i=1, \dots, n}$  and the kernel matrix  $K_{XX} = (k(l_R^i, l_R^j))_{i, j=1, \dots, n}$ .

## 3 Method

We formulate the registration problem as Bayesian inference, where we obtain the posterior distribution of parameters  $\boldsymbol{\alpha}$  given the target surface as:

$$P(\boldsymbol{\alpha} | \Gamma_T) = \frac{P(\Gamma_T | \boldsymbol{\alpha}) P(\boldsymbol{\alpha})}{\int P(\Gamma_T | \boldsymbol{\alpha}) P(\boldsymbol{\alpha}) d\boldsymbol{\alpha}}. \quad (11)$$

The prior probability, computed with eq. (8), pushes the solution towards a more likely shape given the GPMM space, by penalizing unlikely shape deformations. The likelihood term can easily be customized with different distance measures and probability functions, depending on the application goal at hand. We are usually interested in modeling the L2 distance between two surfaces ( $d_{l^2}$ ), for which we can use the independent point evaluator likelihood:

$$P(\Gamma_T|\boldsymbol{\alpha}) = \prod_{i=1}^n \mathcal{N}(d_{l^2}(\Gamma_T^i, \Gamma[\boldsymbol{\alpha}]^i); 0, \sigma_{l^2}^2), \quad (12)$$

as also used in [21]. The L2 distance between the  $i$ -th point  $\Gamma[\boldsymbol{\alpha}]^i \in R^3$  and its closest point on the surface  $\Gamma_T$  is rated using a zero-mean normal distribution with the expected standard deviation for a good registration. The variance  $\sigma_{l^2}^2$  is the observation noise of the points of our target surface. We can register for a better Hausdorff distance [2] by changing the likelihood to:

$$P(\Gamma_T|\boldsymbol{\alpha}) = \text{Exp}(d_H(\Gamma_T, \Gamma[\boldsymbol{\alpha}]); \lambda_H) \quad (13)$$

with  $d_H$  being the Hausdorff distance between the two meshes and  $\text{Exp}$  being the exponential distribution with pdf  $p(d) = \lambda_H e^{-\lambda_H d}$ .

### 3.1 Approximating the Posterior Distribution

The posterior distribution defined in eq. (11) can unfortunately not be obtained analytically. Yet, we can compute the unnormalized density value for any shape described by  $\boldsymbol{\alpha}$ . This allows us to use the Metropolis-Hastings algorithm [29] to generate samples from the posterior distribution in the form of a Markov-chain. The MH algorithm is summarized in algorithm 1. A general way to explore the parameter space is to use a random-walk proposal, i.e. a Gaussian update distribution in the parameter space

$$Q(\boldsymbol{\alpha}'|\boldsymbol{\alpha}) \sim \mathcal{N}(\boldsymbol{\alpha}, \sigma_l). \quad (14)$$

We usually combine differently scaled distributions, each with a specified  $\sigma_l$ , to allow for both local and global exploration of the posterior distribution. For each proposal, one distribution is chosen at random.

---

#### Algorithm 1 Metropolis-Hastings sampling

---

- 1:  $\boldsymbol{\alpha}_0 \leftarrow$  arbitrary initialization
  - 2: **for**  $i = 0$  to  $S$  **do**
  - 3:    $\boldsymbol{\alpha}' \leftarrow$  sample from  $Q(\boldsymbol{\alpha}'|\boldsymbol{\alpha}_i)$
  - 4:    $t \leftarrow \frac{q(\boldsymbol{\alpha}_i|\boldsymbol{\alpha}')p(\Gamma_T|\boldsymbol{\alpha}')p(\boldsymbol{\alpha}')}{q(\boldsymbol{\alpha}'|\boldsymbol{\alpha}_i)p(\Gamma_T|\boldsymbol{\alpha}_i)p(\boldsymbol{\alpha}_i)} \cdot \{\text{acceptance threshold}\}$
  - 5:    $r \leftarrow$  sample from  $\mathcal{U}(0, 1)$
  - 6:   **if**  $t > r$  **then**
  - 7:      $\boldsymbol{\alpha}_{i+1} \leftarrow \boldsymbol{\alpha}'$
  - 8:   **else**
  - 9:      $\boldsymbol{\alpha}_{i+1} \leftarrow \boldsymbol{\alpha}_i$
-

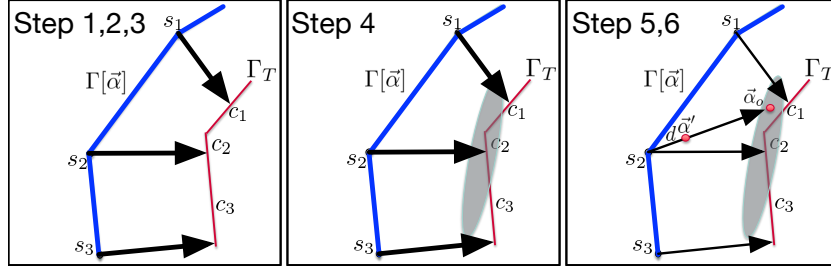


Fig. 1: Visualization of the CP-Proposal with the current instance (blue) from the model, and the target surface (red). The grey ellipse in the centre window shows the landmark noise for  $s_2$ . The right window shows how an update for  $s_2$  is generated based on the posterior sample  $\alpha_o$  and the step size  $d$ .

### 3.2 CP-proposal

A random-walk in the parameter space of the prior GPMM model  $\mathcal{M}$  is time-consuming as it usually is high-dimensional. Instead, we propose to accelerate convergence by using an informed proposal. For the proposal to reach a unique stationary distribution, we have to be able to compute the transition probability, which requires the proposal to be stochastic. We propose the CP-proposal, which takes the geometry of the model and the target into account to guide the proposed change. Internally, we use a posterior model  $\mathcal{M}_\alpha$  based on estimated correspondences to propose randomized informed samples. From a current state  $\alpha$ , we propose an update  $\alpha'$  by executing the following steps (visualized in fig. 1):

1. Sample  $m$  points  $\{s_i\}$  on the current model instance  $\Gamma[\alpha]$ .
2. For every point  $s_i$ ,  $i \in [0 \dots m]$  find the closest point  $c_i$  on the target  $\Gamma_T$ .
3. Construct the set of observations  $L$  based on corresponding landmark pairs  $(s_i, c_i)$  according to eq. (9) and define the noise  $\epsilon_i \sim \mathcal{N}(0, \Sigma_{s_i})$  using eq. (16).
4. Compute the analytic posterior  $\mathcal{M}_\alpha$  (eq. (10)) with  $L$  and  $\{\Sigma_{s_i}\}$ .
5. Get  $\alpha_o$  by first drawing a random shape<sup>3</sup> from the posterior model  $\mathcal{M}_\alpha$  and then projecting it into the prior model  $\mathcal{M}$ .
6. We generate

$$\alpha' = \alpha + d(\alpha_o - \alpha) \quad (15)$$

with  $d \in [0.0 \dots 1.0]$  being a step-length.

The noise  $\epsilon$  in step 3 is modeled with low variance along the normal direction and high variance along the surface. The variance at each point  $s_i$  in  $\Gamma[\alpha]$  is computed by:

$$\Sigma_{s_i} = [\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} \sigma_n^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_v^2 \end{bmatrix} [\mathbf{n}, \mathbf{v}_1, \mathbf{v}_2]^T, \quad (16)$$

<sup>3</sup> Sampling all  $\alpha_i$  independently from  $\mathcal{N}(0, 1)$  and constructing the shape with eq. (7).

where  $\mathbf{n}$  is the surface normal at the position  $s_i$  in the mesh and  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are perpendicular vectors to the normal. The variances along the vectors are set to  $\sigma_n^2 = 3.0 \text{ mm}^2$  and  $\sigma_v^2 = 100.0 \text{ mm}^2$ . This noise term ensures that the posterior model from step 4 takes the uncertain correspondence along the surface into account, which is not well defined in flat regions.

If a small step-length is chosen in step 6, the current proposal is only adjusted slightly in the direction of the target surface, resulting in a locally restricted step. With a step size of 1.0, the proposed sample is an independent sample from the posterior in eq. (10).

In practice, closest-point based updates often find wrong correspondences if the sizes of  $\Gamma$  and  $\Gamma_T$  greatly differ. This is especially problematic in the case of elongated thin structures. It is, therefore, useful also to establish the correspondence from  $\Gamma_T$  to  $\Gamma$  from time to time.

**Computing the Transition Probability** For each new proposal  $\alpha'$  from the CP-proposal distribution, we need to compute the transition probability as part of the acceptance threshold (see algorithm 1 step 4). The transition probability  $q(\alpha'|\alpha)$  is equal to the probability of sampling the shape corresponding to  $\alpha_o$  from the posterior model  $\mathcal{M}_\alpha$ , computed in step 4 of the CP-proposal. For  $q(\alpha|\alpha')$ , the transition probability is computed in the same way. We solve eq. (15) for  $\alpha'_o$  after swapping  $\alpha$  and  $\alpha'$  and evaluate the corresponding shape likelihood under the posterior distribution  $\mathcal{M}_{\alpha'}$ .

## 4 Experiments

In the following, we perform registration experiments on surfaces of femur bones as well as a reconstruction experiment of missing data on face scans. For the face experiment, the face template and the face GPMM from [8] are used together with 10 available face scans. For the femur experiments, we use 50 healthy femur meshes extracted from computed tomography (CT) images<sup>4</sup>. Each surface is complete, i.e. no holes or artifacts. This setting is optimal for the standard ICP algorithm and therefore serves as a fair comparison to the CPD algorithm and our probabilistic implementation. The CPD experiments are performed with the MATLAB code from [22] and all other experiments with the Scalismo<sup>5</sup> library.

### 4.1 Convergence Comparison

We compare the convergence properties of our CP-proposal and random-walk. The CP-proposal configuration is mentioned together with its definition in section 3.2. For the random-walk, we use a mixture of the proposals defined in eq. (14), with  $\sigma_l$  being set to six different levels, from 1.0 mm to 0.01  $\mu\text{m}$ , and all six proposal distributions equally likely to be sampled from. In fig. 2, the con-

<sup>4</sup> Available via the SICAS Medical Image Repository [15].

<sup>5</sup> <https://scalismo.org>



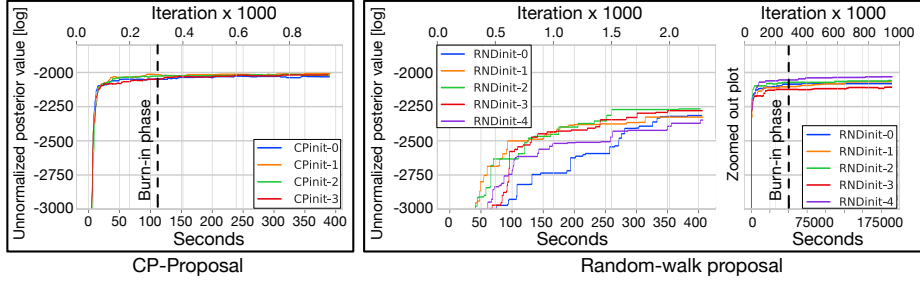


Fig. 2: Convergence plots for the femur GPM registrations with 50 components. Our CP-proposal is shown to the left and the random-walk (including a zoomed out plot) to the right. The CP-proposal needs 300 iterations, while the random-walk needs more than 200k samples for the burn-in phase without reaching the same registration quality even after 1M iterations. Run-time in seconds is shown on the lower x-axis and number of MH iterations on the upper.

vergence time of the standard random-walk and the CP-proposal is shown. The experiment is performed with a GPM with a low-rank approximation of 50 components, see section 2.1. We randomly initialize the model parameters and start 5 registrations in parallel. As expected, our proposal leads to much faster convergence. In fig. 3 we see a posterior plot comparison of the two proposals. Notice how less likely samples are often accepted, which makes it different from deterministic methods such as ICP and CPD.

#### 4.2 Posterior Estimation of Missing Data

We use a face GPM to estimate the posterior distribution of noses from partial face data, where the nose has been removed. In fig. 4, we see that there is no perfect overlap between the face model and the scan. Therefore, we need to change our likelihood function to adjust for a possible changing number of correspondence points during the registration. To obtain a close fit on average, while also penalizing far away points, we use the collective average likelihood introduced in [30] and extend it with the Hausdorff likelihood,

$$P(\Gamma_T|\alpha) \propto \mathcal{N}(d_{CL}(\Gamma_T, \Gamma[\alpha]); 0, \sigma_{CL}^2) \cdot \text{Exp}(d_H(\Gamma_T, \Gamma[\alpha]); \lambda_H), \quad (17)$$

where

$$d_{CL} = \frac{1}{N} \sum_{i=1}^N \|\Gamma_R^i - \Gamma_T^i\|^2. \quad (18)$$

Here the closest point of  $\Gamma_R^i$  on the target is  $\Gamma_T^i$  and  $N$  is the number of landmarks in the reference mesh  $\Gamma_R$ .

On a technical note, using ICP to predict correspondence in a missing data scenario maps all points from the reference, which are not observed in the target to the closest border. To counter this effect, we filtered away all predicted

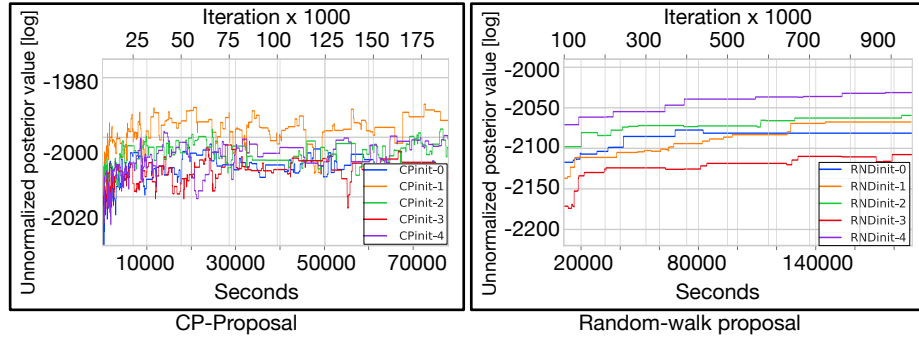


Fig. 3: Posterior plot comparison of the CP-proposal and random-walk. Even with a very small update step for the random-walk, it has difficulties to explore the posterior in the high-dimensional setting. The CP-proposal, on the other hand, can more efficiently explore the high-dimensional space.

correspondences, where the predicted target point is part of the target surface’s boundary. This is also done in e.g. [1].

In fig. 4, we show the correspondence uncertainty from the posterior registration distribution. Our method infers a larger correspondence uncertainty in the outer region. However, as the surface is observed, the uncertainty is low in the direction of the face surface normals but high within the surface. This is because there is no anatomical distinctive feature to predict the correspondence more precisely. High uncertainty is also inferred on the nose, where data is missing. In contrast to the outer region, the uncertainty in the direction of the normal of the reconstructed nose surface is large. This shows that uncertainty visualization can be used to detect missing areas or to build better PDMs by incorporating the uncertainty of the registration process as demonstrated next.

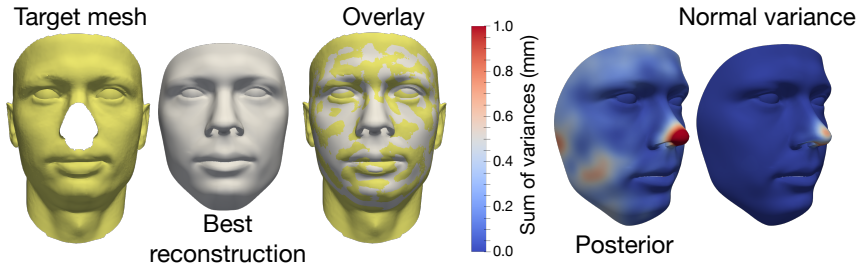


Fig. 4: Nose reconstruction with the face GPMM. No uncertainty in the outer region of the face along the normal indicates that only the correspondence is uncertain compared to the nose.

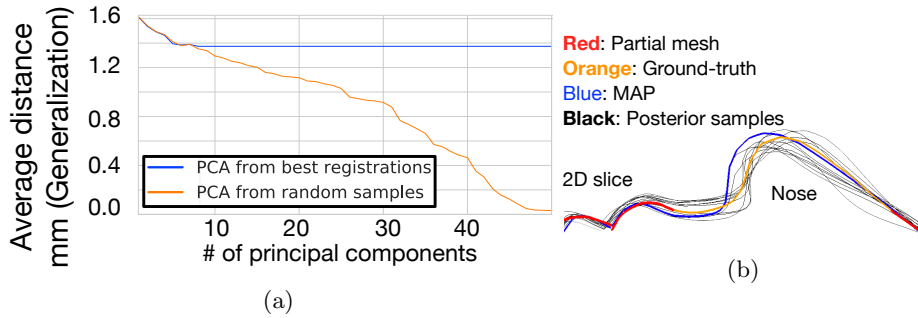


Fig. 5: (a) shows the better generalization of a PDM using 100 random samples from the posterior distributions of each of the 10 targets, compared to only using the most likely samples. The blue line flattens out as no more than 9 principal components are computed with 10 meshes. (b) 2D slice view of posterior face samples. The MAP solution does not explain the ground-truth shape, whereas the random samples cover the ground-truth shape.

**PDMs from Posterior Distributions** In this experiment, we show how to benefit from probabilistic registration when building new PDMs from a small dataset with partially missing data, which is common in medical imaging. If much more data is available, the influence of using the posterior is reduced as the variability gained from the probabilistic framework can be gained by including a lot more data. In section 4.2, we registered the 10 available face scans by computing the posterior distribution of reconstructions for each target. In fig. 5b we show samples from the posterior distribution of nose reconstructions.

In fig. 5a, we compare two different PDMs’ generalization abilities, i.e. the capability to represent unseen data [31]. One PDM is built following the classical approach of only using the most likely registrations. The other PDM is built from 100 random samples from each of the 10 posterior distributions. We compute the average generalization of all 10 PDMs built using a leave-one-out scheme. The plot shows that the PDM built from the posteriors generalizes better.

### 4.3 Registration Accuracy - ICP vs CPD vs CP-proposal

In this experiment, we compare the best sample, also known as Maximum a posteriori (MAP) from our probabilistic method, with the ICP and CPD methods. We use a femur GPMM approximated with 200 basis functions. From the model, we sample 100 random meshes as starting points for the registration. We, therefore, end up with 100 registrations for each target. In fig. 7 we show the summary of all 5000 registrations (All) and some individual representative target meshes. The naming scheme combines the registration method (ICP, CPD, or MH with CP) with the likelihood function which was used (L2). The box plot shows the variation of the average L2 surface distances from all 100 registrations of each

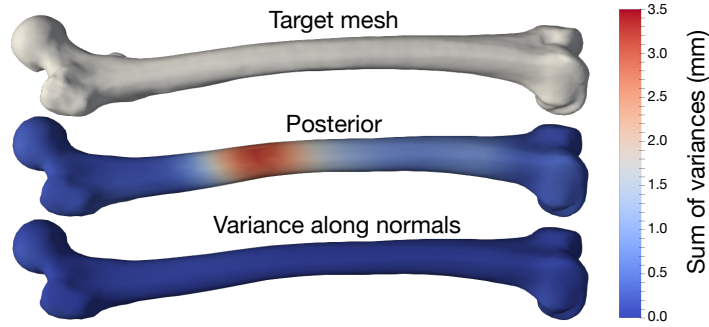


Fig. 6: Femur registration with uncertainty. The registration uncertainty is visualized with the point-wise sum of variances. As expected, the higher uncertainty of the established correspondence coincides with the shaft region with the least characteristic shape. No variance is observed along the normals, so the uncertainty is only in the correspondence along the surface.

target. As the ICP and CPD methods converge at maximum 100 iterations, we also restrict our sampling method to 100 samples. The chain usually converges within 100-300 samples as shown in fig. 2.

The CP-proposal is consistently better than the ICP and CPD registrations, and at the same time provides much less fluctuation in the quality. The few outliers for our method are cases where the chain has not converged. In fig. 6, we show the uncertainty of the established correspondence of a registration from 1000 samples (300 samples for the burn-in phase). Depicted is the uncertainty of individual points from the posterior distribution from a single registration. Note the high uncertainty values along the shaft, which indicate that the established

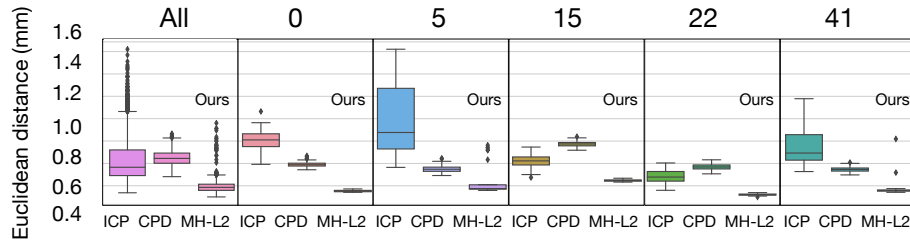


Fig. 7: Distances between the final registrations and their target meshes. For each target mesh, we randomly initialize 100 registrations. The MAP sample from our CP-proposal is superior to ICP and CPD. The femur target id (0 to 49) is shown on top of each plot.

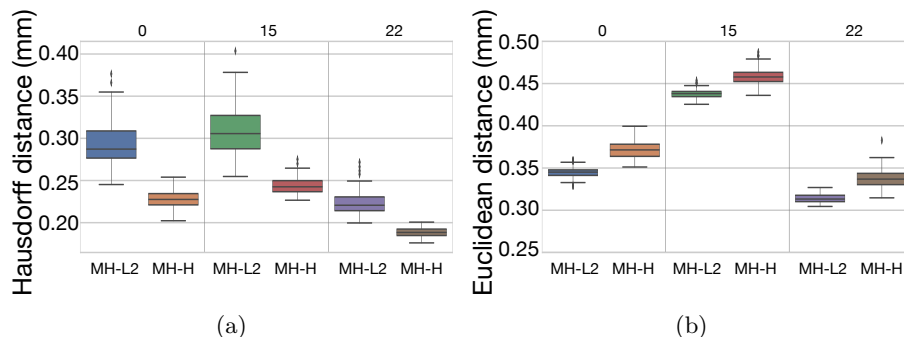


Fig. 8: Registration result comparison using either the Hausdorff or the Euclidean distance likelihood. (a) Euclidean distance and (b) Hausdorff distance between the MAP samples and the target meshes. The plots show that the average L2 surface distance is only slightly worse when the Hausdorff ( $H$ ) likelihood is used.

correspondence is less reliable in that region. No variance along the normals indicates that the uncertainty is purely correspondence shift within the surface.

**Alternative Hausdorff Distance Likelihood** In fig. 8a, we compare the registration results based on their Hausdorff distance, and we compare results from sampling using the L2 likelihood, eq. (12), and the Hausdorff likelihood, eq. (13). As expected, we can focus the registration to avoid large Hausdorff distances. The equivalent L2 likelihood experiment is shown in fig. 8b and shows that while optimizing for the Hausdorff distance, the average L2 surface distance is increasing only slightly. This demonstrates the capability to change the likelihood in our framework based on the application’s needs.

**Drawbacks of Deterministic Methods** The main problem with ICP and CPD is that they cannot recover from local optima. If the algorithm finds the closest point on the wrong part of the target, we end up with a bad registration. In fig. 9, we show a registration result of the 3 registration methods. The ICP method can get the overall length of the bone correct but ends up with a registration, where the structure is folding around itself. The CPD approach is more robust than ICP as it preserves the topological structure of the point sets. Our method additionally provides an uncertainty map.

**Run-time Comparison** The number of components in the low-rank approximation can be seen as regularization of the deformations. More complex local deformations can be obtained using more components. The algorithm run-time scales linearly in the number of components, with the run-time being 2.5 times slower for each time the model rank doubles. For models with rank 50, 100 and

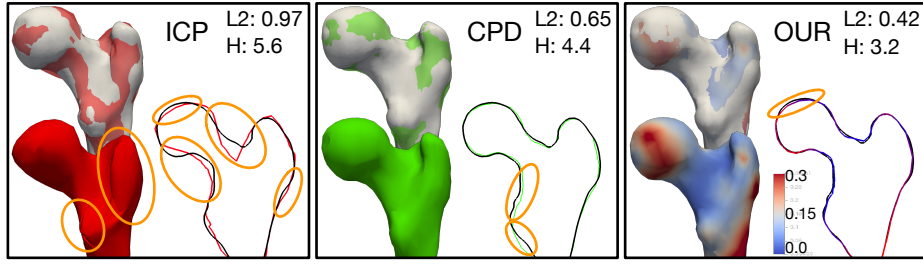


Fig. 9: The 3 registration methods (ICP, CPD, OURS) are shown in separate windows. The registration accuracy for the same target is noted in the form of Euclidean- (L2) and Hausdorff- (H) distances (mm). The orange ellipses highlight problematic areas of the registration for each method. For each method, we show the target with the registration overlaid, the 3D-registration, and a 2D slice of the registration (colored) and the target (black). Notice how our method (in comparison to ICP and CPD) shows the correspondence uncertainty (summed point variances ( $\text{mm}^2$ ) for each landmark).

200, the CP-proposal takes: 46 s, 110 s and 275 s. In comparison, the ICP implementation takes 30 s, 69 s and 155 s. The CPD implementation does not make use of the GPMM model and takes 75 s for 100 samples with our setup. While our method is slower than the two others, we still get reasonable run-times while inferring more accurate results and estimate the full posterior instead of a single estimate.

## 5 Conclusion

In this paper, we presented a probabilistic registration framework. Our main contribution is the informed proposal for the MH algorithm, which makes it possible to work in a high-dimensional model space that would be difficult to explore with pure random-walk. Our informed proposal integrates geometry awareness in the update step, which results in faster convergence. In the case of missing data, our method provides an estimate of the posterior over possible reconstructions. Thus our framework can provide uncertainty measures for critical tasks such as surface reconstruction in the medical domain, as required for surgical decision making. Using our framework, different likelihood terms can be combined and used, while the choice is restricted to the L2 norm in standard ICP and CPD. Finally, we showed how to build PDMs that generalize better using the posterior distribution of registrations.

**Acknowledgements.** This research is sponsored by the Gebert R f Foundation under the project GRS-029/17.

## References

1. Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid icp algorithms for surface registration. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. IEEE (2007)
2. Aspert, N., Santa-Cruz, D., Ebrahimi, T.: Mesh: Measuring errors between surfaces using the hausdorff distance. In: Proceedings. IEEE International Conference on Multimedia and Expo. vol. 1, pp. 705–708. IEEE (2002)
3. Besl, P.J., McKay, N.D.: Method for registration of 3-D shapes. In: Sensor Fusion IV: Control Paradigms and Data Structures. vol. 1611, pp. 586–607. International Society for Optics and Photonics (1992)
4. Bogo, F., Romero, J., Loper, M., Black, M.J.: Faust: Dataset and evaluation for 3d mesh registration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3794–3801 (2014)
5. Chen, Y., Medioni, G.: Object modelling by registration of multiple range images. *Image and vision computing* **10**(3), 145–155 (1992)
6. Deprelle, T., Groueix, T., Fisher, M., Kim, V., Russell, B., Aubry, M.: Learning elementary structures for 3d shape generation and matching. In: Advances in Neural Information Processing Systems. pp. 7433–7443 (2019)
7. Feldmar, J., Ayache, N.: Rigid, affine and locally affine registration of free-form surfaces. *International journal of computer vision* **18**(2), 99–119 (1996)
8. Gerig, T., Morel-Forster, A., Blumer, C., Egger, B., Luthi, M., Schönborn, S., Vetter, T.: Morphable face models-an open framework. In: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). pp. 75–82. IEEE (2018)
9. Grenander, U., Miller, M.I.: Representations of knowledge in complex systems. *Journal of the Royal Statistical Society: Series B (Methodological)* **56**(4), 549–581 (1994)
10. Groueix, T., Fisher, M., Kim, V.G., Russell, B.C., Aubry, M.: 3d-coded: 3d correspondences by deep deformation. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 230–246 (2018)
11. Hoffman, M.D., Gelman, A.: The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research* **15**(1), 1593–1623 (2014)
12. Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H.: Generation of a statistical shape model with probabilistic point correspondences and the expectation maximization-iterative closest point algorithm. *International journal of computer assisted radiology and surgery* **2**(5), 265–273 (2008)
13. Hufnagel, H., Pennec, X., Ehrhardt, J., Handels, H., Ayache, N.: Point-based statistical shape models with probabilistic correspondences and affine em-icp. In: *Bildverarbeitung für die Medizin 2007*, pp. 434–438. Springer (2007)
14. Jampani, V., Nowozin, S., Loper, M., Gehler, P.V.: The informed sampler: A discriminative approach to bayesian inference in generative computer vision models. *Computer Vision and Image Understanding* **136**, 32–44 (2015)
15. Kistler, M., Bonaretti, S., Pfahrer, M., Niklaus, R., Büchler, P.: The virtual skeleton database: an open access repository for biomedical research and collaboration. *Journal of medical Internet research* **15**(11), e245 (2013)
16. Kortylewski, A., Wieser, M., Morel-Forster, A., Wiczeorek, A., Parbhoo, S., Roth, V., Vetter, T.: Informed mcmc with bayesian neural networks for facial image analysis. *arXiv preprint arXiv:1811.07969* (2018)

17. Le Folgoc, L., Delingette, H., Criminisi, A., Ayache, N.: Quantifying registration uncertainty with sparse bayesian modelling. *IEEE transactions on medical imaging* **36**(2), 607–617 (2016)
18. Liu, H., Liu, T., Li, Y., Xi, M., Li, T., Wang, Y.: Point cloud registration based on mcmc-sa icp algorithm. *IEEE Access* (2019)
19. Lüthi, M., Gerig, T., Jud, C., Vetter, T.: Gaussian Process Morphable Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* pp. 1–1 (2018)
20. Ma, J., Lin, F., Honsdorf, J., Lentzen, K., Wesarg, S., Erdt, M.: Weighted robust pca for statistical shape modeling. In: *International Conference on Medical Imaging and Augmented Reality*. pp. 343–353. Springer (2016)
21. Morel-Forster, A., Gerig, T., Lüthi, M., Vetter, T.: Probabilistic fitting of active shape models. In: *International Workshop on Shape in Medical Imaging*. pp. 137–146. Springer (2018)
22. Myronenko, A., Song, X.: Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence* **32**(12), 2262–2275 (2010)
23. Myronenko, A., Song, X., Carreira-Perpinán, M.A.: Non-rigid point set registration: Coherent point drift. In: *Advances in neural information processing systems*. pp. 1009–1016 (2007)
24. Neal, R.M., et al.: Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo* **2**(11), 2 (2011)
25. Pan, Y., Yang, B., Liang, F., Dong, Z.: Iterative global similarity points: A robust coarse-to-fine integration solution for pairwise 3d point cloud registration. In: *2018 International Conference on 3D Vision (3DV)*. pp. 180–189. IEEE (2018)
26. Pomerleau, F., Colas, F., Siegwart, R., et al.: A review of point cloud registration algorithms for mobile robotics. *Foundations and Trends® in Robotics* **4**(1), 1–104 (2015)
27. Risholm, P., Fedorov, A., Pursley, J., Tuncali, K., Cormack, R., Wells, W.M.: Probabilistic non-rigid registration of prostate images: modeling and quantifying uncertainty. In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. pp. 553–556. IEEE (2011)
28. Risholm, P., Pieper, S., Samset, E., Wells, W.M.: Summarizing and visualizing uncertainty in non-rigid registration. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 554–561. Springer (2010)
29. Robert, C., Casella, G.: *Monte Carlo statistical methods*. Springer Science & Business Media (2013)
30. Schönborn, S., Egger, B., Morel-Forster, A., Vetter, T.: Markov chain monte carlo for automated face image analysis. *International Journal of Computer Vision* **123**(2), 160–183 (2017)
31. Styner, M.A., Rajamani, K.T., Nolte, L.P., Zsemlye, G., Székely, G., Taylor, C.J., Davies, R.H.: Evaluation of 3d correspondence methods for model building. In: *Biennial International Conference on Information Processing in Medical Imaging*. pp. 63–75. Springer (2003)