Robust Point Set Registration Using EM-ICP with Information-Theoretically Optimal Outlier Handling

Jeroen Hermans, Dirk Smeets, Dirk Vandermeulen, Paul Suetens
K.U.Leuven, ESAT-PSI,
Medical Imaging Research Center, University Hospitals Gasthuisberg,
Herestraat 49 - bus 7003, B-3000 Leuven, Belgium
jeroen.hermans@uz.kuleuven.ac.be

Abstract

In this paper the problem of pairwise model-to-scene point set registration is considered. Three contributions are made. Firstly, the relations between correspondence-based and some information-theoretic point cloud registration algorithms are formalized. Starting from the observation that the outlier handling of existing methods relies on heuristically determined models, a second contribution is made exploiting aforementioned relations to derive a new robust point set registration algorithm. Representing model and scene point clouds by mixtures of Gaussians, the method minimizes their Kullback-Leibler divergence both w.r.t. the registration transformation parameters and w.r.t. the scene's mixture coefficients. This results in an Expectation-Maximization Iterative Closest Point (EM-ICP) approach with a parameter-free outlier model that is optimal in information-theoretical sense. While the current (CUDA) implementation is limited to the rigid registration case, the underlying theory applies to both rigid and non-rigid point set registration. As a by-product of the registration algorithm's theory, a third contribution is made by suggesting a new point cloud Kernel Density Estimation approach which relies on maximizing the resulting distribution's entropy w.r.t. the kernel weights.

The rigid registration algorithm is applied to align different patches of the publicly available Stanford Dragon and Stanford Happy Budha range data. The results show good performance regarding accuracy, robustness and convergence range.

1. Introduction

Pairwise scene-to-model registration of point sets is a fundamental computer vision problem which involves the determination of a geometrical transformation and/or a (dense) correspondence field relating a model to a scene observed. Both rigid and non-rigid point set registration approaches are used in several applications including range data fusion, object tracking and face recognition.

Confining the discussion to iterative approaches, surface registration algorithms incrementally estimate the unknown transformation parameters by minimizing a distance measure between the point sets considered. Regardless of the transformation parametrization used, these algorithms can be subdivided according to whether or not point correspondences are required to evaluate the point set distance measure. In the sequel of this paper, these approaches are referred to as correspondence-based and information-theoretic point set registration methods. Both categories are briefly discussed in the following paragraphs.

Correspondence-based registration approaches typically alternate between establishing point correspondences given an estimate of the transformation parameters and updating the transformation parameters as to reduce the distance between the point correspondences established. Point correspondences can be estimated either deterministically or stochastically resulting in the Iterative Closest Point (ICP) [1] or Expectation Maximization Iterative Closest Point (EM-ICP) [5] algorithm. Although both methods handle point correspondences differently, ICP can be regarded as a limiting case of EM-ICP. Over the past decades many efforts have been undertaken to improve both ICP (e.g. [2, 15]) and EM-ICP (e.g. [3, 9]) regarding their convergence behavior and robustness against outliers.

Information-theoretic point set registration algorithms do not rely on point correspondences. Instead, they represent the point sets by probability density functions. Generally Gaussian mixture models are used which are either obtained by clustering the point sets or by defining a mixture component in each input point. In the latter case, the mixture model is also referred to as Kernel Density Estimate (KDE) [11]. After construction of the probability density functions, registration reduces to the com-

putation of transformation parameters minimizing the distance between these density functions. Thereto, several information-theoretic measures are used such as Kullback-Leibler (KL) divergence [14], Kernel Correlation [11], the $L2$-norm [7, 10], Jensen-Shannon Divergence [13] or Jensen-Renyi Divergence [12]. Due to their independence to point correspondences, some of these methods are inherently robust against outliers.

Although both point set registration categories are related to each other, their links have not been exploited yet. In this paper, the relation between EM-ICP and information-theoretic point set registration approaches is formalized in the context of pairwise model-to-scene registration. Starting from the observation that the outlier handling of existing methods relies on heuristically determined models, the relations between both categories are exploited to develop a new robust point set registration algorithm. Representing model and scene point clouds by mixtures of Gaussians, the method minimizes their Kullback-Leibler divergence both w.r.t. the registration transformation parameters and w.r.t. the scene's mixture coefficients. This results in an Expectation-Maximization Iterative Closest Point (EM-ICP) approach with a parameter-free outlier model that is optimal in information-theoretical sense. While the theory is applicable to both rigid and non-rigid point cloud registration, the current CUDA implementation is restricted to the rigid case. As a by-product of the registration algorithm's theory, a new Kernel Density Estimation approach is suggested which relies on maximizing the resulting distribution's entropy w.r.t. the kernel weights.

The remainder of this paper is organized as follows. In Sect. 2.1 the relation between correspondence-based and information-theoretic point set registration approaches is formalized. In Sect. 3 the new EM-ICP point set registration method with optimal outlier handling is described. In Sect. 4 the method is validated in the context of range image fusion. Finally, Sect. 5 draws some conclusions.

2. Methods

Consider two D-dimensional point sets $\mathcal{S} = \{\boldsymbol{s}_i\}$ $(i = 1, \ldots, N_{\mathcal{S}})$ and $\mathcal{M}(\boldsymbol{\theta}) = \{\boldsymbol{m}_j(\boldsymbol{\theta})\}$ $(j = 1, \ldots, N_{\mathcal{M}})$ with $\boldsymbol{s}_i \in \mathbb{R}^D$ and $\boldsymbol{m}_j(\boldsymbol{\theta}) \in \mathbb{R}^D$ their constituting points. While the first point set is regarded as the scene, the second one represents a model that has to be registered to the scene. Hence, the model's geometric configuration is parameterized by a vector $\boldsymbol{\theta}$. Both point sets $\mathcal{S}$ and $\mathcal{M}(\boldsymbol{\theta})$ are assumed to be constituted of i.i.d. samples drawn from the scene's and the model's probability density function $p_{\mathcal{S}}(\boldsymbol{r})$ and $p_{\mathcal{M}}(\boldsymbol{r}|\boldsymbol{\theta})$ (with $\boldsymbol{r} \in \mathbb{R}^D$).

Within this context, the following section formalizes relations between correspondence-based and information theoretic point set registration approaches.

2.1. Relating Correspondence-Based and Information-Theoretic Point Set Registration Methods

In order to determine the transformation parameters $\boldsymbol{\Theta}$ registering the model with the scene, EM-ICP algorithms maximize the log-likelihood

$$\boldsymbol{\Theta} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N_{\mathcal{S}}} \log\left(p_{\mathcal{M}}(\boldsymbol{s_i}|\boldsymbol{\theta})\right). \quad (1)$$

Mostly, the model's probability density function is given by a Gaussian mixture model

$$p_{\mathcal{M}}(\boldsymbol{r}|\boldsymbol{\theta}) = \sum_{j=1}^{N_{\mathcal{M}}} \gamma_j f_\sigma(\boldsymbol{r}|\boldsymbol{m}_j(\boldsymbol{\theta})) \quad (2)$$

with $\sum_j \gamma_j = 1$ and $\gamma_j \geq 0$ $(j = 1, \ldots, N_{\mathcal{M}})$, while

$$f_\sigma(\boldsymbol{r}|\boldsymbol{\mu}) = \frac{1}{(2\pi\sigma^2)^{\frac{D}{2}}} \exp\left(-\frac{\|\boldsymbol{r}-\boldsymbol{\mu}\|^2}{2\sigma^2}\right). \quad (3)$$

In many practical cases, the scene point set is corrupted with *outliers*. Outliers are defined as data points that cannot be declared by the model. Hence, even if the model is perfectly aligned with the scene, outliers will have low likelihood values in the model's probability density function. As such they have a large effect on the data log-likelihood of Eq. 1 and tend to disturb the registration process. Therefore, outlier models are introduced to reduce their influence on the registration process. There are two main approaches of outlier handling in EM-ICP procedures.

The first approach adds at least one additional outlier mixture component to the model's probability density function of Eq. 2. Outlier data should have a sufficiently large likelihood w.r.t. this component as to reduce their effect on the log-likelihood of Eq. 1. Generally, this outlier class is modeled with a broad Gaussian or a uniform (pseudo-) distribution. However, there is no reason to assume that these outlier models are optimal in any sense.

Instead of adapting the model's probability density function, the second outlier handling approach manipulates the scene's probability density function. Following Latecki *et al.* [8], this approach is motivated by the observation that in the absence of outliers, the log-likelihood of Eq. 1 corresponds to a (negated) Monte Carlo integral approximation of the Kullback Leibler (KL) divergence between the model's and the scene's probability density function

$$KL(p_{\mathcal{S}}\|p_{\mathcal{M}}) = \int_{\boldsymbol{r}} p_{\mathcal{S}}(\boldsymbol{r}) \log\left(\frac{p_{\mathcal{S}}(\boldsymbol{r})}{p_{\mathcal{M}}(\boldsymbol{r}|\boldsymbol{\theta})}\right) d\boldsymbol{r}. \quad (4)$$

Hence, EM-ICP can also be considered as an information-theoretic registration approach minimizing the KL divergence between a model and a scene. However, in practical

cases, a part of the scene points is generated by an outlier process. Hence, the scene's probability density function is assumed to be composed of an inlier and an outlier component. As EM-ICP only aims at minimizing the KL divergence between the model's probability density function and the inlier part of the scene's probability density function, a weighted log-likelihood function

这里主要介绍了两种处理outlier的方法，一种是引入一个额外的噪声模型；另一种是对inlier进行加权

$$L_w(\boldsymbol{\theta}) = \sum_{i=1}^{N_S} w_i(\boldsymbol{s_i}) \log\left(p_{\mathcal{M}}(\boldsymbol{s_i}|\boldsymbol{\theta})\right) \tag{5}$$

proves to be a better approximation of the intended KL divergence. The weights $w_i(\boldsymbol{s_i})$ are referred to as *inlier beliefs* and increase with increasing likelihood of $\boldsymbol{s_i}$ being an inlier. This approach is strongly related to the theory of robust M-estimators [6] providing a substantial amount of possible weighting functions $w_i(\boldsymbol{r})$. However, there is no indication that any of these weighting functions is optimal in any sense. In addition, using the Gaussian mixture model of Eq. 2 and the weighting function $w_i(\boldsymbol{r}) \propto p_{\mathcal{M}}(\boldsymbol{r}|\tilde{\boldsymbol{\theta}})$ with $\tilde{\boldsymbol{\theta}}$ a current registration transformation estimate, it can be shown that the resulting EM-ICP approach reduces to Kernel Correlation. The latter is closely related to other information-theoretic point set registration measures such as the $L2$-norm and the Renyi entropy. However, there is no indication that these methods have optimal outlier handling in any sense.

In order to derive a robust point set registration approach with optimal outlier handling, the relation between EM-ICP and the information-theoretical Kullback-Leibler divergence is exploited.

3. EM-ICP with Optimal Outlier Handling

In this paper, surface registration is, following Eq. 4, performed by the minimization of the Kullback-Leibler divergence between the model's and the scene's probability density function. However, instead of using the Monte Carlo integral approximation, a non-parametric kernel density estimate is considered as the scene's probability density function. Hence, similar to the model's probability density function, the scene's probability density function is given by

$$p_{\mathcal{S}}(\boldsymbol{r}) = \sum_{i=1}^{N_S} \omega_i f_\sigma(\boldsymbol{r}|\boldsymbol{s}_i) \tag{6}$$

with $\sum_i \omega_i = 1$ and $\omega_i \geq 0$ $(i = 1, \ldots, N_S)$. Note that for both model and scene the same KDE bandwidth σ is chosen.

Moreover, opposed to existing methods, the Kullback-Leibler divergence is not only minimized w.r.t. the registration transformation parameters, but also w.r.t. the scene's KDE mixture coefficients ω_i. As such an optimal outlier rejection scheme is obtained in information-theoretical sense. Indeed, according to Eq. 4, the KL divergence between scene and model corresponds to

$$KL(p_{\mathcal{S}}\|p_{\mathcal{M}}) = H(p_{\mathcal{S}}, p_{\mathcal{M}}) - H(p_{\mathcal{S}}) \tag{7}$$

with $H(p_{\mathcal{S}})$ the scene distribution's Shannon entropy and $H(p_{\mathcal{S}}, p_{\mathcal{M}})$ the Shannon cross entropy between scene and model. Hence, minimizing the KL divergence corresponds to the simultaneous maximization of the scene's Shannon entropy while minimizing the Shannon cross entropy between scene and model.

Maximizing the scene's Shannon entropy w.r.t. its mixture coefficients, results in a scene probability density function that is as flat and as extended as possible. As such, it invokes a tendency to generate as much non-zero scene mixture coefficients as possible. On the other hand, minimization of the Shannon cross entropy between scene and model enforces the suppression of all scene mixture coefficients belonging to points that cannot be declared by the model. Combining both opposite forces results in a point set registration approach that aligns the model with the largest part of the scene that can be declared by the model. Besides the fact that outlier rejection is performed in an information-theoretically optimal manner, this approach does not require the heuristical selection of a parametric outlier class model or a log-likelihood weighting function.

这句话的含义就是rejection的含义，即减少对不会出现在model中的点的系数；

一个是最大化熵，一个是最小化熵。

The use of the scene probability density function's entropy as a regularizer for the outlier model, suggests that the generally used KDE representations with equal mixture coefficients are not optimal. In order to improve the KDE representation, the mixture coefficients of a KDE should be optimized as to maximize its entropy. This is in accordance with the maximum entropy principle stating that the optimal probability distribution obeying certain constraints is the one with maximal entropy.

As there does not exist a closed-form expression for the Kullback-Leibler divergence between Gaussian mixture models, it will be minimized through an upper bound

$$KL(p_{\mathcal{S}}\|p_{\mathcal{M}}) \quad \leq \quad KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi}) \tag{8}$$

As will be shown below, the latter is constructed in a working point defined by an estimate $\tilde{\Psi}$ of the parameter set $\Psi = \{\boldsymbol{\theta}, \gamma_j, \sigma\}$ containing the transformation parameters, the model's mixture coefficients and the KDE bandwidth. Hence, the Kullback Leibler divergence is minimized through iterative construction and minimization of its upper bound. This results in the following Space-Alternating generalized EM [4] algorithm:

while not converged **do**

E-Step: Construct $KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi})$ with current estimate of $\tilde{\Psi}$ of the parameter set Ψ.

M-Step 1: Reduce $KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi})$ by updating the registration transformation parameters $\boldsymbol{\theta}$. This results in a new estimate $\tilde{\Psi}$ of the parameter set Ψ.

这个所谓的"generalized EM" 有点像ECM，就是当有多个参数的时候，分布求解。

但是否会引入更大的计算量呢?

E-Step: Construct $KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi})$ with current estimate of $\tilde{\Psi}$ of the parameter set Ψ.

M-Step 2: Reduce $KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi})$ by updating the scene's mixture coefficients ω_i.

end while

Although the model weights γ_j and the KDE bandwidth σ can also be re-estimated, this is not included in the current implementation. In the following paragraphs the algorithm's E-Step and its M-Steps are discussed in more detail.

E-Step In the E-step an upper bound for the Kullback Leibler divergence is estimated given the current estimates $\tilde{\Psi}$ of the parameter set $\Psi = \{\boldsymbol{\theta}, \gamma_j, \sigma\}$ containing the transformation parameters, the model's mixture coefficients and the KDE bandwidth. This upper bound is composed of two terms according to

KL散度不是一个balanced的测量方式，那么可否引入一些比较"balanced"的测量方式呢?

$$KL_{UB}(p_{\mathcal{S}}\|p_{\mathcal{M}}|\tilde{\Psi}) = H_{UB}(p_{\mathcal{S}}, p_{\mathcal{M}}|\tilde{\Psi}) - H_{LB}(p_{\mathcal{S}}) \quad (9)$$

with $H_{LB}(p_{\mathcal{S}}) \leq H(p_{\mathcal{S}})$ a lower bound for the Shannon entropy of a Gaussian mixture model and $H_{UB}(p_{\mathcal{S}}, p_{\mathcal{M}}) \geq H(p_{\mathcal{S}}, p_{\mathcal{M}}|\tilde{\Psi})$ an upper bound for the Shannon cross entropy between Gaussian mixture models. Both terms are formalized in the following sections.

Shannon Cross Entropy Upper Bound The Shannon cross entropy between model and scene is given by

$$H(p_{\mathcal{S}}, p_{\mathcal{M}}) \quad (10)$$
$$= -\sum_{i=1}^{N_{\mathcal{S}}} \omega_i \int_{\boldsymbol{r}} f_\sigma(\boldsymbol{r}|\boldsymbol{s}_i) \log\left(\sum_{j=1}^{N_{\mathcal{M}}} \gamma_j f_\sigma(\boldsymbol{r}|\boldsymbol{m}_j(\boldsymbol{\theta}))\right) d\boldsymbol{r}$$

In order to construct an upper bound for Eq. 10, following the soft assign weights are introduced

$$\varphi_{i,j} = \frac{\tilde{\gamma}_j f_{\tilde{\sigma}}(\boldsymbol{s}_i|\boldsymbol{m}_j(\tilde{\boldsymbol{\theta}}))}{\sum_{\ell=1}^{N_{\mathcal{M}}} \tilde{\gamma}_\ell f_{\tilde{\sigma}}(\boldsymbol{s}_i|\boldsymbol{m}_\ell(\tilde{\boldsymbol{\theta}}))} \quad (11)$$

with $\tilde{\boldsymbol{\theta}}$, $\tilde{\gamma}_j$ and $\tilde{\sigma}$ the current estimates of the registration transformation parameters $\boldsymbol{\theta}$, the model's mixture coefficients γ_j and the KDE bandwidth σ. Inserting these soft assign weights into the logarithmic terms of Eq. 10 and applying Jensen's inequality results in

添加了一项的目的是什么?

$$\log\left(\sum_{j=1}^{N_{\mathcal{M}}} \gamma_j f_\sigma(\boldsymbol{r}|\boldsymbol{m}_j(\boldsymbol{\theta}))\right)$$
$$= \log\left(\sum_{j=1}^{N_{\mathcal{M}}} \gamma_j \frac{\varphi_{i,j}}{\varphi_{i,j}} f_\sigma(\boldsymbol{r}|\boldsymbol{m}_j(\boldsymbol{\theta}))\right)$$
$$\geq \sum_{j=1}^{N_{\mathcal{M}}} \varphi_{i,j} \log\left(\frac{\gamma_j}{\varphi_{i,j}} f_\sigma(\boldsymbol{r}|\boldsymbol{m}_j(\boldsymbol{\theta}))\right) \quad (12)$$

Substituting this into Eq. 10 results in the following upper bound

$$H(p_{\mathcal{S}}, p_{\mathcal{M}})$$
$$\leq \sum_{i=1}^{N_{\mathcal{S}}} \omega_i \sum_{j=1}^{N_{\mathcal{M}}} \varphi_{i,j} \left(\frac{\|\boldsymbol{m}_j(\boldsymbol{\theta}) - \boldsymbol{s}_i\|^2}{2\sigma^2} - \log\left(\frac{\gamma_j}{\varphi_{i,j}}\right)\right)$$
$$+\frac{D}{2} + \frac{D}{2}\log(2\pi\sigma^2)$$
$$\equiv H_{UB}(p_{\mathcal{S}}, p_{\mathcal{M}}|\tilde{\Psi}) \quad (13)$$

Shannon Entropy Lower Bound The Shannon Entropy of the scene's distribution is given by

$$H(p_{\mathcal{S}}) \quad (14)$$
$$= -\sum_{i=1}^{N_{\mathcal{S}}} \omega_i \int_{\boldsymbol{r}} f_\sigma(\boldsymbol{r}|\boldsymbol{s}_i) \log\left(\sum_{\ell=1}^{N_{\mathcal{S}}} \omega_\ell f_\sigma(\boldsymbol{r}|\boldsymbol{s}_\ell)\right) d\boldsymbol{r}$$

Applying Jensen's inequality to Eq. 14, a lower bound of the scene distribution's Shannon Entropy is obtained:

$$H(p_{\mathcal{S}})$$
$$\geq -\sum_{i=1}^{N_{\mathcal{S}}} \omega_i \log\left(\int_{\boldsymbol{r}} f_\sigma(\boldsymbol{r}|\boldsymbol{s}_i) \sum_{\ell=1}^{N_{\mathcal{S}}} \omega_\ell f_\sigma(\boldsymbol{r}|\boldsymbol{s}_\ell) d\boldsymbol{r}\right)$$
$$= -\sum_{i=1}^{N_{\mathcal{S}}} \omega_i \log\left(\sum_{\ell=1}^{N_{\mathcal{S}}} \omega_\ell f_{\sqrt{2}\sigma}(\boldsymbol{s}_\ell|\boldsymbol{s}_i)\right)$$
$$\equiv H_{LB}(p_{\mathcal{S}}). \quad (15)$$

After construction of an upper bound for the Kullback-Leibler divergence, it can be minimized w.r.t. the registration transformation parameters and the scene distribution's mixture coefficients. This is discussed in the following sections.

M-step 1 In the first M-step the registration transformation parameters are updated as to reduce the value of the upper bound of the Kullback-Leibler divergence. According to Eq. 9, Eq. 13 and Eq. 15, this is achieved by reducing the upper bound $H_{UB}(p_{\mathcal{S}}, p_{\mathcal{M}}|\tilde{\Psi})$ for the Shannon cross entropy between model and scene. This is done using an approximate Newton update step

$$\boldsymbol{\theta}_{new} = \tilde{\boldsymbol{\theta}} - Hess(H_{UB}(\tilde{\boldsymbol{\theta}}))^{-1} \nabla(H_{UB}(\tilde{\boldsymbol{\theta}})) \quad (16)$$

with $Hess(H_{UB}(\tilde{\boldsymbol{\theta}}))$ the diagonal of the upper bound's Hessian and $\nabla(H_{UB}(\tilde{\boldsymbol{\theta}}))$ its gradient for the current estimate $\tilde{\boldsymbol{\theta}}$ of the registration transformation parameters.

Eq. 13 is proportional to the lower bound of Eq. 5 traditionally used by EM-ICP algorithms. Hence, in the first M-Step the parameters are updated as to increase the weighted log-likelihood of the scene. However, instead of computing the inlier beliefs of Eq. 5 using some heuristically chosen loss function, they are co-estimated during the registration process in an information-theoretically optimal sense.

M-step 2 In the second M-step the inlier beliefs or scene mixture coefficients are updated as to reduce the upper bound of the Kullback-Leibler divergence between model and scene. In order to enforce positivity of the scene's mixture coefficients as well as their partition of unity, they are parameterized as

$$\omega_i = \frac{\exp(\rho_i)}{\sum_{\ell=1}^{N_\mathcal{S}} \exp(\rho_\ell)} \tag{17}$$

Using a gradient descent approach with step size ξ, the scene's mixture coefficients are update according to

$$\omega_{i,new} \propto \omega_i \exp\left(-\xi \frac{\delta KL_{UB}(p_\mathcal{S} \| p_\mathcal{M}|\tilde{\Psi})}{\delta \rho_i}\right) \tag{18}$$

with

$$\frac{\delta KL_{UB}(p_\mathcal{S} \| p_\mathcal{M}|\tilde{\Psi})}{\delta \rho_i} = \frac{H_{UB}(p_\mathcal{S}, p_\mathcal{M}|\tilde{\Psi})}{\delta \rho_i} - \frac{H_{LB}(p_\mathcal{S})}{\delta \rho_i} \tag{19}$$

while

$$\begin{aligned}\frac{H_{LB}(p_\mathcal{S})}{\delta \rho_i} = \\ \omega_i \left(1 + \sum_{\ell=1}^{N_\mathcal{S}} \omega_\ell \log\left(\frac{\sum_{j=1}^{N_\mathcal{S}} \omega_j f_{\sqrt{2}\sigma}(\boldsymbol{s}_\ell|\boldsymbol{s}_j)}{\sum_{j=1}^{N_\mathcal{S}} \omega_j f_{\sqrt{2}\sigma}(\boldsymbol{s}_i|\boldsymbol{s}_j)}\right)\right) - \\ \omega_i \sum_{\ell=1}^{N_\mathcal{S}} \frac{\omega_\ell f_{\sqrt{2}\sigma}(\boldsymbol{s}_i|\boldsymbol{s}_\ell)}{\sum_{j=1}^{N_\mathcal{S}} \omega_j f_{\sqrt{2}\sigma}(\boldsymbol{s}_j|\boldsymbol{s}_\ell)}\end{aligned} \tag{20}$$

and

$$\begin{aligned}\frac{H_{UB}(p_\mathcal{S}, p_\mathcal{M}|\tilde{\Psi})}{\delta \rho_i} = \\ -\omega_i \left(\sum_{\ell=1}^{N_\mathcal{S}} \omega_\ell \sum_{j=1}^{N_\mathcal{M}} \varphi_{\ell,j} \frac{\|\boldsymbol{m}_j(\boldsymbol{\theta}) - \boldsymbol{s}_\ell\|^2}{2\sigma^2}\right) \\ +\omega_i \left(\sum_{j=1}^{N_\mathcal{M}} \varphi_{i,j} \frac{\|\boldsymbol{m}_j(\boldsymbol{\theta}) - \boldsymbol{s}_i\|^2}{2\sigma^2}\right) \\ +\omega_i \left(\sum_{\ell=1}^{N_\mathcal{S}} \omega_\ell \sum_{j=1}^{N_\mathcal{M}} \varphi_{\ell,j} \log\left(\frac{\gamma_j}{\varphi_{\ell,j}}\right)\right) \\ -\omega_i \left(\sum_{j=1}^{N_\mathcal{M}} \varphi_{i,j} \log\left(\frac{\gamma_j}{\varphi_{i,j}}\right)\right).\end{aligned} \tag{21}$$

Analyzing these gradients gives a more intuitive insight into the outlier scheme.

The gradient of the Shannon Entropy's lower bound of Eq. 20 drives the scene's mixture coefficients to a configuration for which all scene points have equal likelihoods under a bandwidth shifted version of the scene's probability density function $p_\mathcal{S}(\boldsymbol{r})$. Apart from the bandwidth shift with a factor $\sqrt{2}$, this is an intuitively justifiable result as neglecting the model to be aligned with the scene, all scene points should be considered equally likely. The counter-intuitive introduction of the bandwidth shift is caused by the particular Shannon entropy's upper bound used. Further research will focus on alternative approaches to approximate the Shannon entropy without introducing this bandwidth shift. However, it must be emphasized that this will not influence the general idea of this work.

On the other hand, the gradient of the Shannon Cross Entropy's upper bound of Eq. 21 drives the scene's mixture coefficients to configurations for which all inlier scene points have approximately the same distance to the model's point set. Hence, instead of imposing some kind of heuristical threshold or heuristically weighting on this scene-to-model distance, the method proposed compares the distances of all scene points to the model in order to suppress outliers.

3.1. Implementation Details

Although the registration framework applies to both rigid and non-rigid registration the current implementation considers rigid point set registration only. In order to speed up the computations, the algorithm is implemented on the GPU using CUDA.

4. Results

In order to validate the rigid point set registration approach in a realistic setting, it is applied to range data of the standing *Happy Budha* and the standing *Dragon*. These data sets are publicly available from The Stanford 3D Scanning Repository of the Stanford University Computer Graphics Laboratory [1]. Both data sets are acquired with a Cyberware 3030 MS scanner and contain fifteen range data acquired from directions with 24^o intervals. In the upper left part of Fig. 1 two patches of the Dragon are shown, while the upper right part displays two patches of the Happy Budha.

The registration algorithm proposed is applied to each pair of Dragon patches and each pair of Happy Budha patches that are acquired from directions with angular difference less than 90^o. As range data sets acquired from directions with a larger angular difference only share a very small portion of points, their registration is nearly impossible. Therefore, these data are neglected for validation purposes. As the registration algorithm is not symmetrical each pair of retained range data is aligned in both directions separately. This results in 105 pairwise registration cases for each model (including the trivial pairs of identical patches).

[1] http://www-graphics.stanford.edu/data/3Dscanrep/

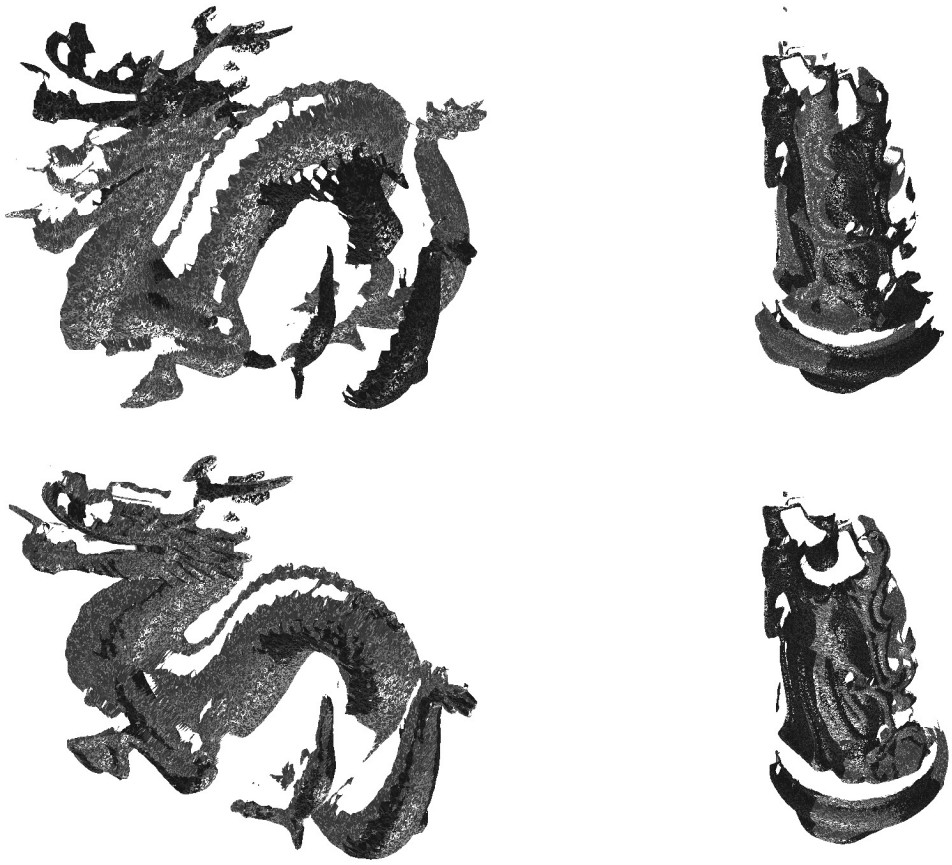

Figure 1. Left: two range data sets of the Stanford Dragon before (upper) and after (lower) robust point set registration. Right: two range data sets of the Stanford Happy Budha before (upper) and after (lower) robust point set registration. In both cases the registration was successful.

In order to limit the execution time of the registration algorithm, only 1500 randomly selected points were used on both scene and model point set. In order to increase the method's capture range, a deterministic annealing scheme is applied on the KDE bandwidth used to represent scene and model. Specifically, an initial bandwidth of 15mm is chosen with an annealing rate of 0.9. The minimal KDE bandwidth was set to 1mm while at most 500 iterations are performed on each annealing level. At the start of each annealing stage, the model's mixture coefficients are optimized as to maximize the entropy of the resulting distribution. Averaged over all registration cases considered, the algorithm finished after 5659 iterations and 152.9 seconds. Neglecting the computational overhead from e.g. model entropy maximization, the computational speed obtained is 27 milliseconds per iteration. It must be noted that the stop criteria used were relatively stringent resulting in a large number of iterations performed in each annealing stage. Although it is expected that the number of iterations can be reduced without decreasing the registration performance, this fine tuning is considered future work.

In Fig. 2 a cumulative histogram of the number of registrations performed and the number of successful registrations is plotted w.r.t. the initial angular misalignment. A registration is considered successful if the final orientational error is less than 4^o. Both for the Dragon and the Happy Budha the likelihood of successful registration decreases with increasing initialization error. For initial orientational errors up to 48^o success rates of 83.6% or more are observed. Even after discarding the trivial registration cases, the success rates remain 78.2% or more. However, within the range from 48^o to 90^o initial rotational misalignment, the success rate drops to 40% for the Dragon and 50% for the Happy Budha. This performance drop is caused by two factors.

First, as an iterative registration algorithm is considered, it's capture range is limited to a certain range of initialization errors. Second, the less overlap between the scene and the model the higher the likelihood of convergence towards an erroneous solution. The latter is visualized in Fig. 3

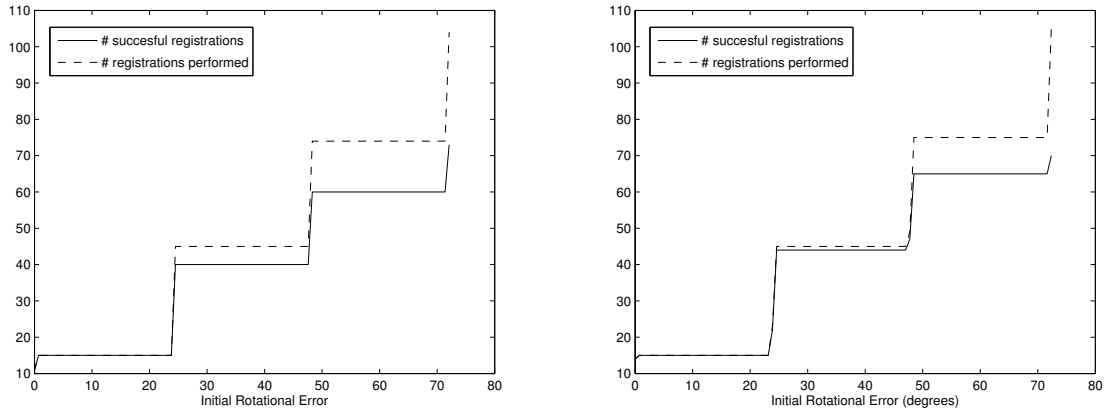

Figure 2. Cumulative histograms of the amount of registrations performed (dashed) and the number of successful registrations w.r.t. initial rotational misalignment for the Stanford Dragon (left) and the Stanford Happy Budha (right).

showing a cumulative histogram of the number of registrations performed and the number of successful registrations w.r.t. the outlier fraction in the scene. From these figures it is readily seen that the success rate starts decreasing when the outlier fraction is within a range of $40\% - 50\%$ and beyond. It must be noted that the outlier configurations encountered are not generated by some artificial random outlier process. Instead, the outliers originate from the fact that the scanned parts of the model differ between different measurements. Moreover, there is a collaboration between aforementioned factors that cause the performance drop of the registration algorithm. The higher the orientational initialization error, the higher the angular difference between the acquisition directions of the range images considered and the less overlap between them.

Averaged over all successful registrations, the RMS translation error is 0.43mm for the Dragon and 0.45mm for the Happy Budha. RMS angular registration errors after successful registration are 0.45^o and 1.5^o for the Dragon and the Budha respectively. The circular shape of the Budha's pedestal is assumed to be the cause of the reduced rotational accuracy.

Overall the registration algorithm shows good convergence behavior, both in terms of robustness and accuracy.

5. Conclusions

This paper considered the problem of pairwise model-to-scene point set registration. Firstly, the relations between correspondence-based and some information theoretic point cloud registration algorithms were formalized. Starting from the observation that the outlier handling of existing methods relies on heuristically determined models, aforementioned relations were exploited to obtain a new robust point set registration algorithm. Representing model and scene point cloud by mixtures of Gaussians, the method minimizes their Kullback-Leibler divergence both w.r.t. the registration transformation parameters and w.r.t. the scene's mixture coefficients. This resulted in an Expectation-Maximization Iterative Closest Point (EM-ICP) approach with a parameter-free outlier model that is optimal in information-theoretical sense. Although the current (CUDA) implementation is restricted to rigid registration, the method's underlying theoretical framework applies to both rigid and non-rigid registration. As a by-product of the registration algorithm's theory, it was suggested that the traditional representation of a point cloud by a Kernel Density Estimate with equal kernel weights is suboptimal. Instead, the kernel weights should be optimized as to maximize the resulting distribution's entropy.

The rigid registration algorithm was applied to align different patches of the publicly available Stanford Dragon and Stanford Happy Budha range data. The results showed good performance regarding accuracy, robustness and convergence range.

References

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

[2] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *Proceedings of IEEE International Conference on Robotics and Automation*, page 27242729, 1991.

[3] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.

[4] J. A. Fessler and A. O. Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, 42(10):2664–2677, 1994.

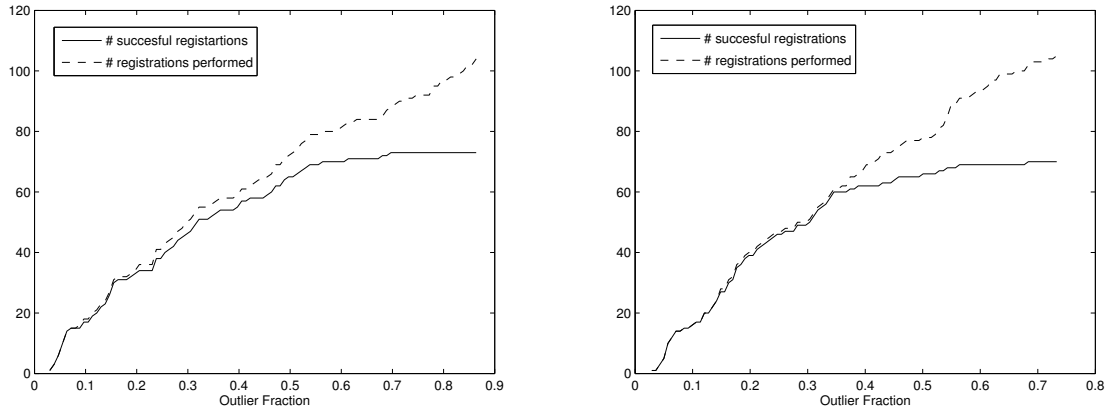

Figure 3. Cumulative histograms of the amount of registrations performed (dashed) and the number of successful registrations w.r.t. scene's outlier fraction for the Stanford Dragon (left) and the Stanford Happy Budha (right).

[5] S. Granger and X. Pennec. Multi-scale em-icp: A fast and robust approach for surface registration. In *Proceedings of European Conference on Computer Vision*, pages 418–432, 2002.

[6] P. J. Huber. *Robust Statistics*. Wiley series in probability and mathematical statistics. 1981.

[7] B. Jian and B. Vemuri. A robust algorithm for point set registration using mixture of gaussians. In *Proceedings of the IEEE International Conference on Computer Vision*, page 12461251, 2005.

[8] L. J. Latecki, M. Sobel, and R. Lakaemper. New em derived from kullback-leibler divergence. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 267–276, 2006.

[9] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262 – 2275, 2010.

[10] A. S. Roy, A. Gopinath, and A. Rangarajan. Deformable density matching for 3d non-rigid registration of shapes. In *Proceedings of the 10th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part I*, page 942949, 2007.

[11] Y. Tsin and T. Kanade. A correlation-based approach to robust point set registration. In *Proceedings of the 8th European Conference on Computer Vision*, pages 558–569, 2004.

[12] F. Wang, T. Syeda-Mahmood, B. C. Vemuri, D. Beymer, and A. Rangarajan. Closed-form jensen-renyi divergence for mixture of gaussians and applications to group-wise shape registration. In *Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part I*, pages 648–655, 2009.

[13] F. Wang, B. C. Vemuri, and A. Rangarajan. Groupwise point pattern registration using a novel cdf-based jensen-shannon divergence. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1283–1288, 2006.

[14] Y. Wang, K. Woods, and M. McClain. Information-theoretic matching of two point sets. *IEEE Transactions on Image Processing*, 11(8):868–872, 2002.

[15] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.