Image and Vision Computing 28 (2010) 124–137

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

A new method for the registration of three-dimensional point-sets: The Gaussian Fields framework

Faysal Boughorbel, Muharrem Mercimek, Andreas Koschan *, Mongi Abidi

Imaging, Robotics and Intelligent Systems Laboratory, University of Tennessee, 334 Ferris Hall, Knoxville, TN 37996, USA

ARTICLE INFO

Article history:
Received 18 October 2006
Received in revised form 22 May 2008
Accepted 7 May 2009

Keywords:
Rigid registration
Gaussian Fields
Moment invariants
Fast Gauss Transform
Optimization

ABSTRACT

In this paper, we present a 3D automatic registration method based on Gaussian Fields and energy minimization. A continuously differentiable energy function is defined, which is convex in a large neighborhood of the alignment parameters. We show that the size of the region of convergence can be significantly extended reducing the need for close initialization and overcoming local convergence problems of the standard Iterative Closest Point (ICP) algorithms. Moreover, the Gaussian criterion can be applied with linear computational complexity using Fast Gauss Transform methods. Experimental evaluation of the technique using synthetic and real datasets demonstrates the usefulness as well as the limits of the approach.

Published by Elsevier B.V.

1. Introduction

One of the crucial tasks in computer vision and graphics is the reconstruction of 3D models. The spatial alignment, or registration, of 3D imagery is a critical step in modeling pipelines. Given the narrow field of view of cameras and the complex topologies of real-world objects, we need to view the scenes and objects of interest from different positions in order to build useful descriptions. The primary emphasis of this work is on the registration of multiple rigid shapes for object modeling. We are employing the most basic representation of shapes assuming their description as a cloud of points. Therefore, our techniques will belong to the same class as that of the ubiquitous Iterative Closest Point (ICP) algorithm. It is the limitations of ICP that we are attempting to overcome. The latter method, while increasingly popular since its discovery in 1992 [2], has several shortcomings. Our goal is to extend the range of convergence of point-based registration, which is one of the major problems with ICP. ICP techniques can ensure an accurate registration only when closely initialized. We depart from the ICP criterion by designing a new energy function that quantifies registration.

From a Boolean matching criterion, which is consistent with a rigorous definition of the registration task, we obtain a continuously differentiable energy function by applying mollification, which is simply a smoothing by convolution with a Gaussian kernel. The basic idea is the same with the Kernel Density Estimation, to cluster the information about the sources and target points, and to replace the smooth Gaussians sums of the distances or the feature similarity differences between all pairs of the model and data points for the smaller summations for a given level of precision. We will interpret this criterion physically in terms of Gaussian force fields that are exerted by one of the point-sets on the other. The strength of this field will depend on the distance between the points and on the similarity of their shape and other attributes. Our formulation is straightforward to implement, especially given the possibility of using standard gradient-based optimization techniques instead of the specialized heuristic employed by ICP. Moreover, this technique allows for the reduction of computational complexity from square to linear when applying the Fast Gauss Transform [12,37]. By combining these elements and by studying the properties of the Gaussian criterion we obtain a practical framework for 3D registration. The original idea of using Gaussian Fields for registration was introduced in [3]. A more efficient version of the approach applying the Fast Gauss Transform was presented in [4] and a preliminary comparison with ICP was shown in [5]. In this paper, we present a detailed description of the algorithm and an in depth evaluation of the framework in comparison with ICP results. We study the effect of noise and overlap on the registration results using synthetic data with ground truth. Furthermore, we evaluate the results obtained when changing the parameterization for several real data sets (acquired with laser range finders and confocal microscope). In addition, the basins of convergence are studied in detail for the proposed framework and the ICP method.

* Corresponding author. Tel.: +1 865 974 0229; fax: +1 865 974 5459.
E-mail address: akoschan@utk.edu (A. Koschan).

0262-8856/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.imavis.2009.05.003

The rest of the paper is laid out as follows: Section 2 discusses related work in 3D registration, and in Section 3, Gaussian Fields are introduced for rigid shape registration. The underlying optimization technique is described in Sections 4 and 5 shows its efficient computation using the Fast Gauss Transform. Experimental results for synthetic and real images are presented in Sections 6 and 7. These results are also compared to those obtained when applying ICP. Finally, the paper concludes with Section 8.

2. Background and related work

Due to their limited field of view and to occlusions most 3D imaging systems will provide partial scans of a scene. In order to build a complete description of scene geometry, several of these partial views need to be merged together. Since these datasets are originally represented in the local sensor coordinates frame, registration is a fundamental step in most 3D modeling pipelines. In the applications most relevant to our work, the views are related by rigid transformations (R,t): 3D rotations and translations. Surface shape registration has been studied during the last several years for recovering the rigid transformations using a set given of 3D correspondences [13,18].

In the literature, a common distinction is found between fine and coarse registration methods [6]. Coarse or preliminary registration can be obtained by several techniques, including interactive initialization of corresponding points as in most commercial 3D modeling packages. Fully automatic 3D matching using exhaustive constrained search techniques were attempted [9]. However, these methods are computationally expensive and sensitive to noise. Invariant features are more frequently employed to reduce the search space. In order to generate a complete 3D model, we need to refine the data sets after coarse registration. One of the most popular refinery techniques is the Iterative Closest Point (ICP) algorithm and its very numerous extensions, which starts with pre-estimated overlapping regions. Namely, ICP requires a priori knowledge about an approximate estimation of the transformations, so starts with pre-estimated overlaps. Otherwise ICP tends to converge monotonically to the nearest local minimum of a mean square distance metric. This method is very accurate when used after an initial coarse registration.

Several feature descriptors were used to represent free-form surfaces and point-sets. In the class of global descriptors spherical representations such as the Spherical Attribute Image (SAI), which mapped surface curvature values into a tessellated sphere, were employed for 3D registration [17], Park and Subbarao [24] employed the invariant Stable Tangent Plan (STP) for crude registration. Many local representations were also proposed to encode surface shape information. Stein and Medioni [31] used the variation of the surface normals at local patches for matching, defining a local feature map called Splash. Thirion [34] extracted critical points and contours using Gaussian curvature extrema. Spin Images are another popular invariant local representation, proposed by Johnson et al. [20]. The image encodes the coordinates of points on the surface of an object with respect to the local basis; it is a local description of the global shape of the object and is invariant to rigid transformations. Most of these features rely on local normal or curvature information. Taati et al. [33] proposed a new set of variable dimensional local shape descriptors and performed registration on both angular and free-form objects. We are mostly focused on the registration of unorganized noisy point-sets, where surfaces were not defined yet; differential surface attributes will not be used from the start. Nevertheless, using a voting process, we are able to infer local surface information in noisy and sparse datasets. $S = \{s_1, \ldots, s_{N_s}\}$ of N_s 3D scene points with a geometric model $M = \{m_1, \ldots, m_{N_m}\}$ of N_m points by minimizing the sum of the squared distances between the scene points and the model. The ICP algorithm was shown to converge monotonically to a local minimum. Therefore, the initial estimate of the transformation should be sufficiently close to the correct registration. Independently, Chen and Medioni [8] developed an algorithm similar to ICP that allowed the incorporation of local shape information, as well as the handling of partially overlapping datasets.

Active research on the surface registration problem gave a rise to a proliferation of improved ICP algorithms. Zhang [38] proposed a method based on robust statistics, while Masuda and Yokoya [21] used a Least Mean Squares (LMS) error measure, and Dorai et al. [11] proposed a Minimum Variance Estimate (MVE) of the registration error. Trucco et al. [35] employed Least Median Squares (LMedS), in the so-called Robust ICP version (RICP), to register noisy point-sets with missing data, while Zinßer et al. [39] designed the 'Picky ICP' which combines the strengths of different currents variants to achieve both robustness and increased speed. For points sets-to-surface registration a recent work by Pottmann et al. [25] devised a technique based on instantaneous kinematics, achieving substantial gains in speed over ICP methods. Furthermore, several researchers extended ICP for the registration of multiple (more than 2 views) 3D datasets [1,19,30,32]. For recent surveys and comparisons of several common variants of the ICP algorithm the following references [10,27,28] can be consulted.

Several of the limitations of the ICP framework stem from the non-differentiability of its cost function which imposed the use of a specialized heuristic for optimization. Addressing the registration in the context of gradient-based optimization has attracted some interest recently. We mainly refer here to the work of Fitzgibbon [14], who showed that a Levenberg–Marquardt approach to the point-set registration problem offered several advantages over current ICP algorithms. This work is one of the most closely related to our efforts in 3D registration, since we also aim at designing a criterion that can be optimized through general purpose nonlinear techniques. The main disadvantage of Fitzgibbon's technique is that it is limited to datasets on a grid, where the Chamfer distance transforms and discrete derivatives are easily evaluated. For sparse unorganized point-sets this method cannot be directly applied. Our work is also in the same class of techniques as the one by Charpiat et al. [7], which approximates the Hausdorff distance with a differentiable metric on shape space.

By introducing a method that uses a straightforward differentiable cost function, directly and explicitly expressed in terms of the point coordinates and the registration parameters, we are able to overcome several problems with the ICP-based methods described above. The smooth behavior of our registration criterion, as well as other characteristics which will be discussed in this paper, combined with the use of a standard optimization scheme extends the range of convergence. Given the generality of our framework we can apply the Gaussian Fields method to single and multi-sensor image registration tasks.

3. Gaussian Fields for surface registration

In the formulation of our point-sets registration criterion our aim is to accomplish the following five objectives. (1) The criterion to be maximized should be differentiable and preferably convex in the neighborhood of the registered position. (2) The method should not need any explicit set of point correspondences. (3) The method will incorporate, in addition to point coordinates, as much available information as possible including local shape descriptors or associated intensity values. (4) The method will allow for as large of a region of convergence as possible and reduce dependence on close initialization. (5) The resulting algorithm must be computationally efficient.

3.1. A Boolean combinatorial criterion

Registration can be regarded as a special sub-problem of pattern matching, where the purpose is to locate a 'model' in the 'data'. In the registration task, the goal is the recovery of both the correspondence and spatial transformations that ensure the best match. The assumption is that a significant overlap between the model and the data exists. A good intuitive definition can state that the registered position is the one resulting in the maximum point-to-point overlap of model and data (considering the noiseless case). Such definition uses a minimum amount of information about the datasets, just the position of points, and could be augmented by requiring local shape similarity between the points. We will show the enhancing effect of using this local information on the quality of the registration.

We start by introducing a simple combinatorial Boolean criterion satisfying the maximum (point-to-point) overlap of two point-sets $M = \{P_i\}_{i=1\ldots N_M}$ and $D = \{Q_j\}_{j=1\ldots N_D}$, that are registered by a transformation Tr^*. We assume at this point the noiseless case. For the problem to be well-posed we need also to assume that M and D have a maximum point-to-point overlap at the aligned position. Even though it is easy to search for the only global maximum rather than controlling its validity, in some occasions the global maximum does not correspond to the correct alignment. ICP algorithm was implicitly proposed based on the same assumption. In the real-world case, registration requires that we have sufficient shape complexity. Then the following measure (1) will have a global maximum at Tr^*:

$$E(Tr) = \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \delta(d(Tr(P_i), Q_j)) \quad \text{with } \delta(t) = \begin{cases} 1 & \text{for } t = 0 \\ 0 & \text{otherwise} \end{cases}, \tag{1}$$

where $d(P,Q)$ is the distance (in our case Euclidean) between points. Incorporating local shape similarity in this criterion is straightforward and requires just using a higher dimensional representation of the datasets where points are defined by both position and a vector of shape attributes: $M = \{(P_i, S(P_i))\}_{i=1\ldots N_M}$ and $D = \{(Q_j, S(Q_j))\}_{j=1\ldots N_D}$.

3.2. Mollification and the Gaussian criterion

The resulting discrete criterion is not continuous with respect to the alignment transformations and it will be difficult to find the global maximum of discrete combinatorial functions. Our approach is to find a smooth approximation of the combinatorial criterion using an analytical method known as *Mollification*. This approach was used as a tool to regularize ill-posed problem with respect to differentiability [23].

Given the Gaussian kernel $\rho_\sigma(t) = \exp\left(\frac{-t^2}{\sigma^2}\right)$ and an arbitrary non-differentiable function $f(t)$ defined on $\Omega \subset \Re^d$, a 'mollified' function $f_\sigma(t)$ can be obtained by convolution:

$$f_\sigma(t) = (\rho_\sigma * f)(t) = \int_\Omega \exp\left(\frac{-(t-s)^2}{\sigma^2}\right) f(s)ds. \tag{2}$$

The resulting function will be an approximation of the original one such that $\lim_{\sigma\to 0} f_\sigma(t) = f(t)$. Furthermore, we will have $f_\sigma \in C^\infty(\Omega)$. This operation is also known as the Gauss Transform. When applying discrete mollification to our combinatorial registration criterion (1), we obtain

$$E_\sigma(Tr) = \int \exp\left(-\frac{(d(Tr(P_i), Q_j) - s)^2}{\sigma^2}\right) \left\{ \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \delta(d(Tr(P_i), Q_j)) \right\} ds,$$

for simplicity we can rewrite $d(Tr(P_i),Q_j)$ as d_{ij},

$$\begin{aligned} E_\sigma(Tr) &= \int \exp\left(-\frac{(d_{ij}-s)^2}{\sigma^2}\right) \left\{ \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \delta(d_{ij}) \right\} ds \\ &= \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \int \exp\left(-\frac{(d_{ij}-s)^2}{\sigma^2}\right) \delta(d_{ij}) ds \\ &= \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \int \exp\left(-\frac{s^2}{\sigma^2}\right) \delta(d_{ij}-s) ds \end{aligned}$$

given that $\delta(d_{ij} - s) \neq 0$ only for $s = d_{ij}$, the integral will become

$$\begin{aligned} &\int \exp\left(-\frac{s^2}{\sigma^2}\right) \delta(d_{ij}-s) ds = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \\ &\sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \int \exp\left(-\frac{s^2}{\sigma^2}\right) \delta(d_{ij}-s) ds = \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \\ &\quad = \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \exp\left(-\frac{d^2(Tr(P_i), Q_j)}{\sigma^2}\right) \end{aligned} \tag{3}$$

The mollified criterion is a straightforward sum of Gaussians of distances between all pairs of model and data points. Expression (3) can be re-interpreted physically as the integration of a potential field whose sources are located at points in one of the datasets dataset and targets in the other one. In the noisy case the Gaussian criterion can account for the noise affecting the position of points by relaxing the parameter σ to values near that of noise variance. Fig. 1 illustrates the functioning of the discrete combinatorial criterion and the mollified version.

Having met the first of our objectives, which is differentiability, we now examine the possibility of extending the basin of convergence of our criterion. Here we are focusing on the case of rigid registration with $Tr(Q_j) = RQ_j + t$. Being the sum of closely packed Gaussian functions, the profile of the criterion with respect to the transformation parameters will generally have the appearance of a Gaussian, with local convexity in the neighborhood of the registered position. Extending the width of the basin of convergence is easily done by increasing the parameter σ. However, this relaxation will come at the price of decreasing the localization accuracy of the criterion. The tradeoff between registration accuracy and size of the region of convergence (ROC) is mainly due to the effect of outliers (i.e. the areas that are outside the intersection of model and data). Assuming that at the registered position the model point-set $M = \{P_i\}_{i=1\ldots N_M}$ is completely included in the data point-set $D = \{Q_j\}_{j=1\ldots N_D}$, and that the points of D with corresponding matches in M are labeled from 1 to N_M, the criterion can be broken into two components representing both inliers–inliers and outliers–inliers interaction:

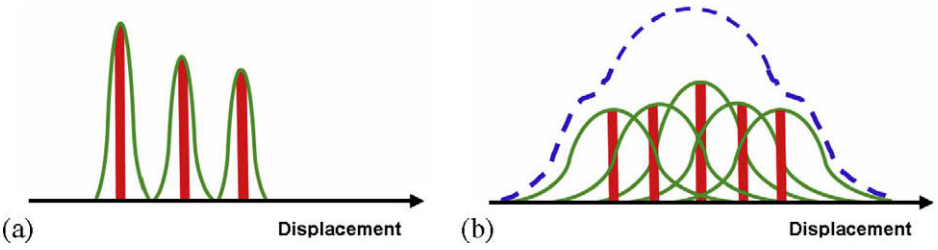

Fig. 1. Mollification converts the Boolean criterion into a smooth sum of Gaussians (a). For σ relaxed we will have an overlap between the different Gaussians. The mixture of these will be our registration criterion, having a dominant peak around the registered position (b).

$$E_\sigma(Tr) = \sum_{\substack{i=1\dots N_M \\ j=1\dots N_M}} \exp\left(-\frac{d^2(Tr(P_i),Q_j)}{\sigma^2}\right) + \sum_{\substack{i=1\dots N_M \\ j=N_M+1\dots N_D}} \exp\left(-\frac{d^2(Tr(P_i),Q_j)}{\sigma^2}\right)$$

$$E_\sigma(Tr) = E_\sigma^{in-in}(Tr) + E_\sigma^{out-in}(Tr). \tag{4}$$

For small values of the decay parameter, the second term will have no effect on the global maximum of the function at the registered position, since $\lim_{\sigma\to 0} E_\sigma^{out-in}(R^*,t^*) = 0$ (again assuming the uniqueness of the aligned position). To reduce the effect of the outliers and ensure good localization error for the Gaussian registration criterion, while at the same time increasing the area of convergence, it is suitable to associate as much information as possible to the points. The inclusion of this additional information is achieved by extending the distance measure between points in the criterion as follows:

$$E_{\sigma,\Sigma}(Tr) = \sum_{\substack{i=1\dots N_M \\ j=1\dots N_D}} \exp\left(-\frac{\|Tr(P_i)-Q_j\|^2}{\sigma^2} - (S(Tr(P_i)) - S(Q_j))^T\Sigma^{-1}(S(Tr(P_i)) - S(Q_j))\right), \tag{5}$$

where $\|\dots\|$ is the Euclidean distance in 3D and the Σ associated with the attribute vector is a diagonal matrix with positive components generalizing the mollification to higher dimensions. With $\omega_\Sigma^{ij}(Tr) = \exp(-(S(Tr(P_i)) - S(Q_j))^T\Sigma^{-1}(S(Tr(P_i)) - S(Q_j)))$ the Gaussian criterion will be

$$E_{\sigma,\Sigma}(Tr) = \sum_{\substack{i=1\dots N_M \\ j=1\dots N_M}} \omega_\Sigma^{ij}(Tr)\exp\left(-\frac{\|Tr(P_i)-Q_j\|^2}{\sigma^2}\right) + \sum_{\substack{i=1\dots N_M \\ j=N_M+1\dots N_D}} \omega_\Sigma^{ij}(Tr)\exp\left(-\frac{\|Tr(P_i)-Q_j\|^2}{\sigma^2}\right) = E_{\sigma,\Sigma}^{in-in}(Tr) + E_{\sigma,\Sigma}^{out-in}(Tr). \tag{6}$$

In the case where the attributes are invariant to the aligning transformations, ω_Σ^{ij} will not depend on Tr. If Σ is chosen with as small components as possible (just the noise level), the term ω_Σ^{ij} can be expected to reduce the value of $E_{\sigma,\Sigma}^{out-in}(Tr)$ more than that of $E_{\sigma,\Sigma}^{in-in}(Tr)$. This will be helpful when the spatial decay parameter σ is relaxed to extend the region of convergence by allowing good localization of the registered position and reducing the need for close initialization. Moreover, our framework accounts for two important aspects in surface registration, namely the quantity and quality of information available in the datasets, and the amount of overlap between the two datasets. In the context of Gaussian Fields the effect of the outliers will be generally important if the overlap region is small relative to the size of the two datasets. The increase in information content results in the decrease of the weights ω_Σ^{ij} and reduces $E_{\sigma,\Sigma}^{out-in}(Tr)$ further than $E_{\sigma,\Sigma}^{in-in}(Tr)$, since the latter term accounts for the corresponding sets.

3.3. Attributes

In the case of 3D surface registration, 3D moment invariants can be used as point attributes. The three moment invariants [29] were also employed in registration algorithms such as in the extension of ICP by Sharp et al. [30]. The moments J_1, J_2, and J_3 are defined for a local neighborhood N around a point $P(X_P, Y_P, Z_P)$ by

$$\begin{aligned} &J_1 = \mu_{200} + \mu_{020} + \mu_{002}, \\ &J_2 = \mu_{200}\mu_{020} + \mu_{200}\mu_{002} + \mu_{020}\mu_{002} - \mu_{110}^2 - \mu_{101}^2 - \mu_{011}^2, \\ &\text{and} \quad J_3 = \mu_{200}\mu_{020}\mu_{002} + 2\mu_{110}\mu_{101}\mu_{011} - \mu_{002}\mu_{110}^2 - \mu_{020}\mu_{101}^2 \\ &\qquad - \mu_{200}\mu_{011}^2 \end{aligned} \tag{7}$$

with

$$\mu_{pqr} = \sum_{(X,Y,Z)\in N} (X - X_P)^p (Y - Y_P)^q (Z - Z_P)^r. \tag{8}$$

In the case of noisy point-sets, where surfaces and curves are not well extracted, we devised a local measure of visual saliency based on the framework of Tensor Voting, introduced by Medioni et al. [22]. We are considering non-oriented point-sets as our primary input and we just use the first pass of the tensor-voting scheme. To evaluate the saliency at a site P_i, we collect votes from neighboring sites P_j which cast the so-called stick tensor at P_i in the case of 2D voting, and the plate tensor for 3D. This tensor encodes the unit vector $t_{ij} = \frac{P_i - P_j}{\|P_i - P_j\|}$ lying on the line joining the two points using its covariance matrix for which the expression in 2D is given by

$$\begin{aligned} T_{ij} &= \begin{bmatrix} \left(t_{ij}^x\right)^2 & t_{ij}^x t_{ij}^y \\ t_{ij}^x t_{ij}^y & \left(t_{ij}^y\right)^2 \end{bmatrix} \\ &= \frac{1}{\left(P_i^x - P_j^x\right)^2 + \left(P_i^y - P_j^y\right)^2} \begin{bmatrix} \left(P_i^x - P_j^x\right)^2 & \left(P_i^x - P_j^x\right)\left(P_i^y - P_j^y\right) \\ \left(P_i^x - P_j^x\right)\left(P_i^y - P_j^y\right) & \left(P_i^y - P_j^y\right)^2 \end{bmatrix} \end{aligned} \tag{9}$$

These tensors are collected from sites in a small neighborhood around P_i using summation

$$T_i = \sum_{j\neq i} T_{ij}. \tag{10}$$

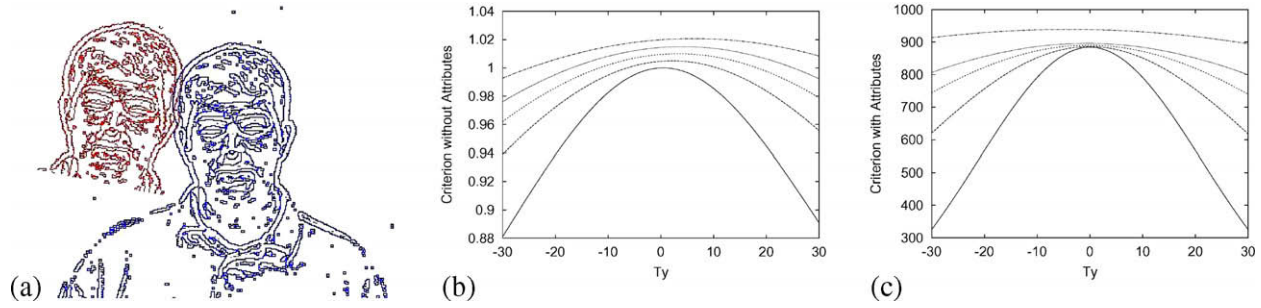

Fig. 2. Profiles of the Gaussian energy function for a displacement around the registered position of the datasets are shown in (a). In (b) the profiles are plotted for σ = 30, 50, 70, 90, 150 obtained without attributes (from narrowest to widest). Plots with moment invariants as attributes for the same values of σ are shown in (c). The scale of the datasets is about 200 × 200 pixels.

To obtain a scalar measure of point saliency, we use the determinant of T_i which is the product of the principle axes of its eigen-system ellipsoid. This feature is invariant to rigid transformations and accounts for degenerate tensors in flat surfaces were its value is zero. The previous local descriptors will be summed in the expression of the energy function (2). Here we use the same approach as in [30] where the covariance matrix of each moment is estimated in a nearly planar region of the datasets, then used to scale the vector of raw features into a set of uncorrelated invariant features with unit variance. For computational efficiency and for robustness to noise we will employ in our later experimental analysis only two descriptors J_1 and the point saliency measure S.

The tradeoff between the region of convergence and localization accuracy can be illustrated with the behavior of the matching criteria with and without attributes as shown in Fig. 2. The profile of the criterion was plotted for a relative displacement of the two point-sets of Fig. 2(a). Several plots are shown with increasing σ. For the non-attributed case, where Euclidean distance between point locations is employed (Fig. 2(b)), we notice that as σ increases the width of the Gaussian bell increases too. However, the maximum will also drift away from the correct location. When we use the Gaussian criterion augmented with moment invariants, as attributes associated with the points, the maximum is more stable (Fig. 2(c)), with nearly no drift for the range of values of σ shown. In the analysis section we will study the localization error as a function of the force range parameter σ.

4. Optimization

For rigid transformation our criterion is always continuously differentiable. For sufficiently dense point-sets this mode will itself have a shape close to a Gaussian given that it is a mixture of Gaussians closely located in parameter space. We can safely assume a smooth convex behavior around the registered position. For a small value of σ and small rigid displacements near the registered position (i.e. a ball of radius ε around the rotation angle and translation vector (φ, t)) the Gaussian criterion (3) can be approximated using the output of the rigid transformation $Tr(P_i)$ and the Taylor series of the functions as follows

$$\begin{aligned}
&\sum_{\substack{i=1\dots N_M\\ j=1\dots N_D}} \exp\left(-\frac{d^2(Tr(P_i),Q_j)}{\sigma^2}\right)\\
&= \sum_{\substack{i=1\dots N_M\\ j=1\dots N_D}} \exp\left(-\frac{\left(\cos\varphi P_i^x - \sin\varphi P_i^y + t_x - Q_j^x\right)^2 + \left(\sin\varphi P_i^x + \cos\varphi P_i^y + t_y - Q_j^y\right)^2}{\sigma^2}\right)\\
&\approx \sum_{\substack{i=1\dots N_M\\ j=1\dots N_D}} 1 - \frac{\left(P_i^x - \varphi P_i^y + t_x - Q_j^x\right)^2 + \left(\varphi P_i^x + P_i^y + t_y - Q_j^y\right)^2}{\sigma^2}. \qquad (11)
\end{aligned}$$

Using the two approximations for a small rotation $\cos\varphi \approx 1$ and $\sin\varphi \approx \varphi$ in addition to the first order approximation resulting from the small displacement compared with σ we obtain

$$\exp\left(-\frac{d^2(Tr(P_i),Q_j)}{\sigma^2}\right) \approx 1 - \frac{d^2(Tr(P_i),Q_j)}{\sigma^2}.$$

Expression (11) is quadratic in the rigid parameters demonstrating the convexity of the criterion. In practice even in the relaxed case we can assume safely convexity in the neighborhood of the registered position. To optimize the Gaussian Fields criterion, we employed the quasi-Newton algorithm [26]. The gradient of the criterion (5) with respect to a transformation parameter α is given by

$$\begin{aligned}
\frac{\partial E_{\sigma,\Sigma}(Tr)}{\partial\alpha} = &\sum_{\substack{i=1\dots N_M\\ j=1\dots N_D}} \frac{-2\omega_\Sigma^{ij}}{\sigma^2}\frac{\partial Tr(P_i)}{\partial\alpha}\\
&(Tr(P_i) - Q_j)\exp\left(-\frac{\|Tr(P_i) - Q_j\|^2}{\sigma^2}\right). \qquad (12)
\end{aligned}$$

The quasi-Newton scheme uses the analytic expression for the gradient along with an approximation of the Hessian to update descent directions. In each descent direction a line search routine is used to find the optimum.

The compromise here is between an accurate localization with a small value of σ and larger region of convergence for a larger σ at the expense of registration accuracy. To strike a balance between these two constraints, we devised a simple scheme based on adapting the values of σ during the optimization process. The scheme consists of two or more runs of the quasi-Newton routine with values of σ decreasing according to suitable schedule (Fig. 3). Using such approach we can start far from the registration parameter while having a good chance of ending with an accurate registration. The main issue here will be a choice of a good reduction schedule for σ. The constraints to be considered are: (1) the lower bound on σ is the noise level, and (2) avoiding being trapped at local minima.

We can try to avoid a local maximum by studying the rate at which the global maximum is drifting with change of the force range parameter. We need to ensure that this drift is not resulting in the next run starting from outside the dominant (usually convex) mode. By studying for several datasets the rate at which the maximum of the criterion drifts with respect to σ and the width of the dominant mode at half the maximum for different values of the force range parameter we can determine the value $\eta \in (0,1)$ which multiplies σ at each run of the global scheme such that we avoid the local maxima. Experimental results show that in practice and with no prior knowledge about initialization we will only need two steps.

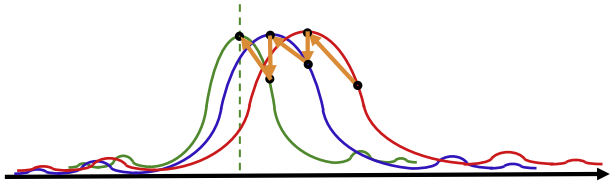

Fig. 3. The global optimization strategy that strikes a balance between the width of the region of convergence and the accuracy of registration is based on adapting the parameter σ, starting from a large value and reducing σ until convergence. We will show experimentally that only two steps are required.

5. Fast Gauss Transform for efficient computation

The registration criterion is essentially a mixture of N_D Gaussians evaluated at N_M points then summed together. The cost of direct evaluation will be $O(N_M \times N_D)$, which for large datasets is computationally expensive. Recently, the Fast Gauss Transform was employed [12] in order to reduce the computational complexity of Gaussian Mixture evaluation, which is a way to find kernel density estimation, to $O(N_M + N_D)$ considerably speed up kernel density estimation. The method was first applied to potential field computations [15,16]. The basic idea is that all calculations are required to be only up to certain accuracy. To evaluate sums of the form $S(t_i) = \sum_{j=1}^{N} f_j \exp\left(-\left(\frac{s_j - t_i}{\sigma}\right)^2\right), \quad i = 1, \ldots, M$, where $\{s_j\}_{j=1,\ldots,N}$ are the centers of the Gaussians known as sources and $\{t_i\}_{i=1,\ldots,M}$ the targets, the following shifting identity and expansion in terms of Hermite series are used:

$$\begin{aligned}
&\exp\left(\frac{-(t-s)^2}{\sigma^2}\right) \\
&= \exp\left(\frac{-(t - s_0 - (s - s_0))^2}{\sigma^2}\right) \\
&= \exp\left(\frac{-(t-s_0)^2}{\sigma^2}\right) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{s - s_0}{\sigma}\right)^n H_n\left(\frac{t - s_0}{\sigma}\right), \qquad (13)
\end{aligned}$$

where H_n are the Hermite polynomials. Given that these series converge rapidly, and that only few terms are needed for a given precision, this expression can be used to replace several sources by s_0 with a linear cost at the desired precision, these clustered sources can then be evaluated at the targets. For a large number of targets the Taylor series (14) can similarly be used to group targets together at a cluster center t_0, further reducing the number of computations. Hence,

$$\begin{aligned}
\exp\left(-\left(\frac{t-s}{\sigma}\right)^2\right) &= \exp\left(\frac{-(t - t_0 - (s - t_0))^2}{\sigma^2}\right) \\
&\approx \sum_{n=0}^{p} \frac{1}{n!} h_n\left(\frac{s - t_0}{\sigma}\right)\left(\frac{t - t_0}{\sigma}\right)^n, \qquad (14)
\end{aligned}$$

where the Hermite functions $h_n(t)$ are defined by $h_n(t) = e^{-t^2} H_n(t)$. The method was shown to converge asymptotically to a linear behavior as the number of sources and targets increases. Implementation details and analysis can be found in [12,16]. The main problem with the original Fast Gauss Transform is the exponential increase of the complexity with the number of dimensions. To address this limitation, a new variant of the method Improved Fast Gauss Transform (IFGT) was proposed by Yang et al. [37] where a data-clustering scheme along with a more intelligent multivariate Taylor expansion allowed for substantial further computational gains even in high dimensions. These gains increase dramatically with the size of the datasets as illustrated in the plots of Fig. 4, which are based on 64 clusters. A general procedure for determining the number of clusters was given in Yang et al. [37] consisting

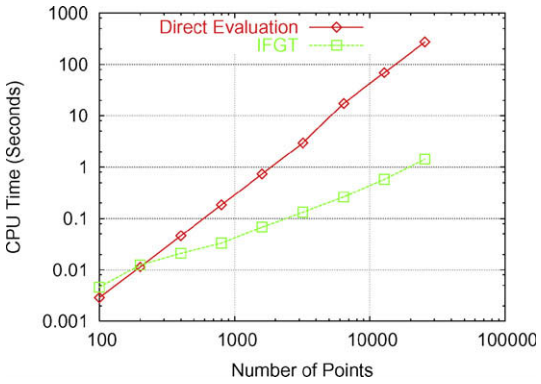

Fig. 4. Plot showing the CPU time (Pentium IV, 2.3 GHz) required for the evaluation of the Gaussian criterion for a given number of data points, direct evaluation versus the Fast Gauss Transform as implemented by Yang et al. [37](called Improved Fast Gauss Transform: IFGT).

of clustering successively for different values of k (number of clusters) until the maximum radius of the clusters does not decrease significantly. Fig. 5 shows an example where we applied the scheme to one of our datasets. We used Euclidean distance over the five-dimensional space embedding 3D points augmented by the two local feature descriptors.

The IFGT approach presents an alternative technique to the standard technique. Given that we are dealing with at least three-dimensional space (coordinates of the points), to a dimensionality of 5 when using moment invariants J in addition to the measure of S, the IFGT is more suitable. The application to our task is straightforward and does not require any other modification, except including the de-correlation step before using the generalized distance in both 3D coordinate and attribute space. One other noteworthy point is the case of the gradient of the Gaussian criterion, which boils down to the computation of a weighted version of the Gaussian function which is similar to the one in (13).

6. Experimental results for synthetic data

One of the main advantages of the Gaussian Fields registration method is the relatively limited number of free parameters used. The only parameter that can change during the registration process is σ to which a large part of the experimental analysis will be devoted. Other parameters on which the method depends are generally computed only once and in most cases are derived in a process similar to sensor characterization. For instance the main purpose of the de-correlation matrix Σ is to create the orthogonal features necessary for effective fusion. Moreover, a confidence parameter

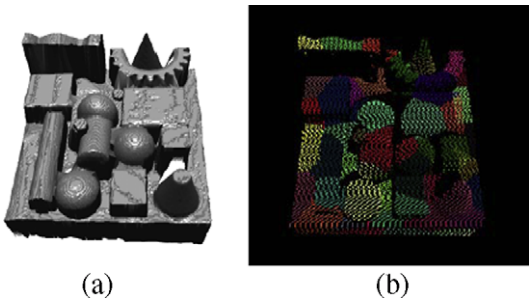

Fig. 5. Result of applying the farthest point clustering algorithm to the 3D model (a). The obtained point clusters (the points of each cluster are labeled with a different color) shown in (b) correspond to regions that are roughly homogeneous with respect to our generalized distance measure (including position and attributes).

C_a is employed to control the effects of noise on the features used in the Gaussian criterion. In addition, we will investigate the effect of the amount of overlap between the datasets to be registered to studying the effects of noise on the algorithm. This first set of experiments is conducted on synthetic datasets to isolate and focus on the factors being studied.

6.1. Effects of noise

In the Gaussian criterion framework noise will influence both the position of the point-sets and consequently the descriptors that are computed from them. When considering very high levels of noise, local shape descriptors can become so corrupted by noise that they are practically useless. It is for this reason that we add the confidence factor C_a to our criterion which will balance the contribution of the descriptors with respect to the coordinates,

$$E_{\sigma,\Sigma}(Tr) = \sum_{\substack{i=1\ldots N_M \\ j=1\ldots N_D}} \exp\left(-\frac{\|Tr(P_i)-Q_j\|^2}{\sigma^2} - \frac{(S(Tr(P_i))-S(Q_j))^T\Sigma^{-1}(S(Tr(P_i))-S(Q_j)))}{C_a}\right), \quad (15)$$

where Σ is the covariance matrix of the feature descriptors used in the algorithm. We will focus our experimental analysis on uniform noise. We do not believe that other forms of noise would affect the performance of the algorithm adversely. Dealing with uniform noise constitutes studying a bad enough case scenario given that in practical applications noise is more biased in one direction (usually the radial component with respect to the camera's coordinate frame is dominant). In this experiment, we use a 3D model of a head which we divide into two partially overlapping sections (Fig. 6(a)). To each of these datasets we add uniform noise of amplitude going up to ±12% of the length of the head. Each of the head sections has about 3500 points. The main purpose of the experiment is to study the drift in the maximum of the criterion under the effect of noise. This is achieved by initializing the algorithm close to the ground truth registered position and starting the optimization scheme. The plots shown in Fig. 6(b) give the resulting rotation error in degrees and translation error as a fraction of the length of the model. They show the rate of increase in the error for two confidence values $C_a = 0.001$ and $C_a = 1.0$.

Rotation and translation errors are calculated as,

$$e_r = \|r - \hat{r}\| \quad \text{and} \quad e_t = \|t - \hat{t}\|/L \cdot 100(\%) \quad (16)$$

where e_r and e_t are rotation and translation errors, r and t are the real rotation and translation vectors, $\hat{r}$ and $\hat{t}$ are and the approximated rotation and translation vectors, and L is the biggest dimension of the bounding box of the 3D dataset.

6.2. Overlap and outliers

One of the fundamental issues in registration, and pattern matching in general, is the effect of outliers, by which we mean areas of the datasets that are not shared by them. Outliers will result in a drift of the maximum of the Gaussian criterion away from the correct position. This drift will only be moderated by a decrease in the value of σ (which itself results in a smaller region of convergence). The error can also be moderated by high complexity in the datasets. To study the effect of overlap on our registration algorithm we used a surface reconstructed from a digital elevation map (DEM). Several partially overlapping surfaces are generated from this DEM. Fig. 7(a) shows four surfaces in the registered position with an amount of overlap ranging from approximately 25% to 70%. In this analysis we keep the same common area and increase the outliers. The drift of the criterion's maximum caused by the outliers is studied for several pairs and for four values of σ (20%, 40%, 60%, and 80%). The plots in Fig. 7(b) and (c) show that up to about 50% of overlap, the criterion maintains a very low localization error. The error starts to increase rapidly for values of overlap less than 30%. As expected the slowest drift is the one corresponding to the lowest value of σ. The other curves evolve closely for most of the range, with the ones corresponding to the two higher values showing a more oscillatory behavior. This experiment gives an idea about the setting of force range parameter to minimize the effect of outliers. The analysis also shows that for practical applica-

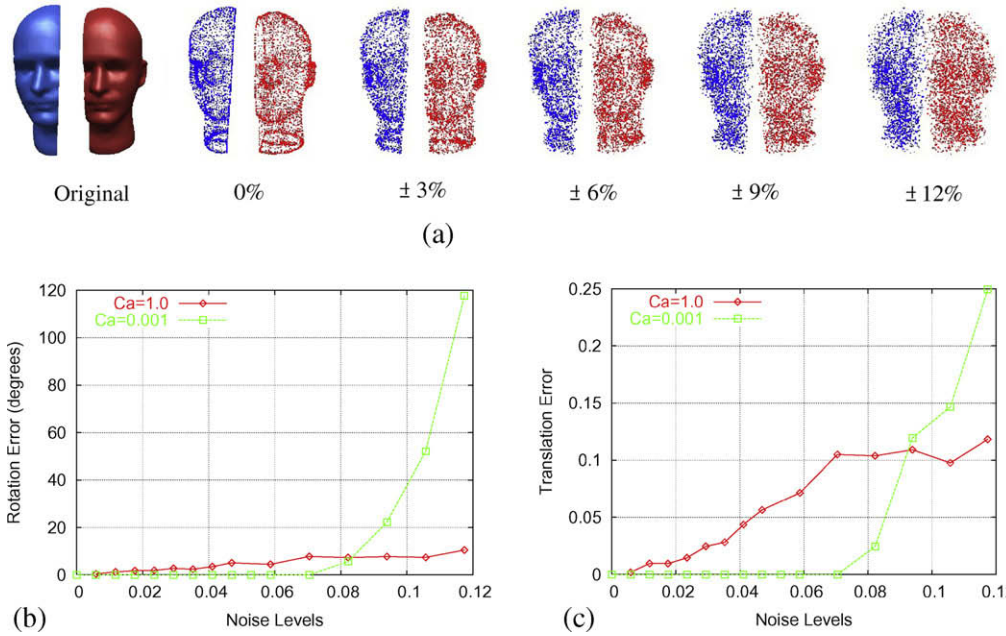

Fig. 6. Datasets used for the noise analysis. (a) Two overlapping sections of a head model are used to study the effects of uniform noise. The point-sets corrupted with increasing levels of uniform noise are shown with the noise value expressed as a fraction of the model's length. Registration error versus uniform noise: (b) rotation error in degrees, (c) translation error as a fraction of the length the head model. We show plots for two values of the confidence parameter.

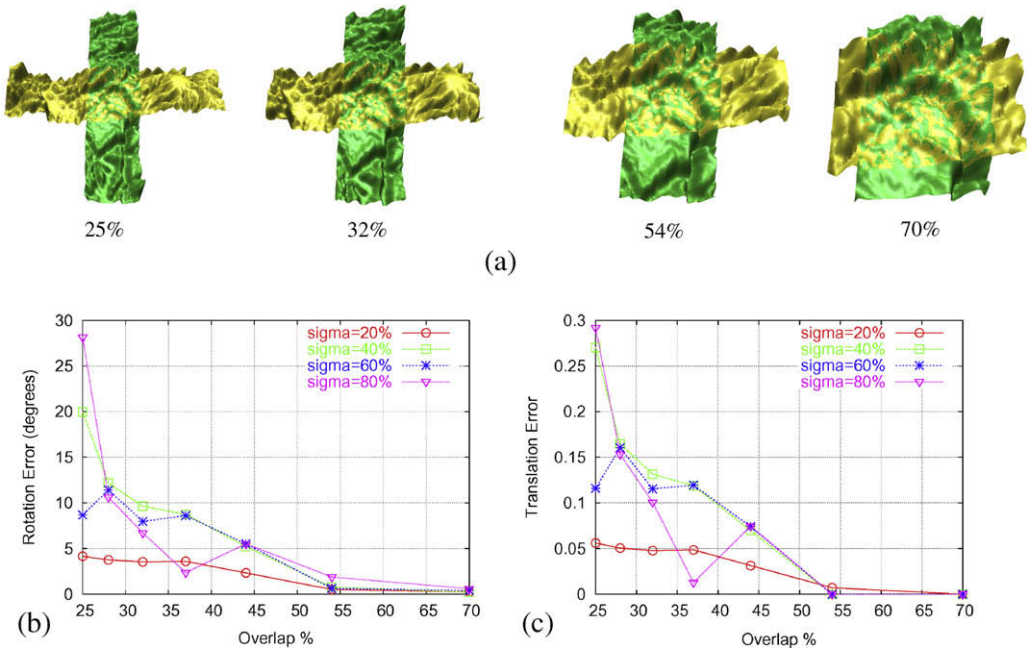

Fig. 7. The amount of overlap between two surfaces and its effect on the registration accuracy: (a) four of the DEM pairs used in this experiment; plots of the (b) rotation and (c) translation errors versus the percentage of overlap.

tions it is suitable to have at least around 40–50% overlap. Nevertheless, the algorithm is able to handle lower values, at the cost of paying more attention to the way values of σ are adapted during the optimization scheme.

6.3. Effect of resolution

For the experiments on real-world datasets, it is assumed that the sampling densities of the points are the same for model and data pairs. In case of having the datasets acquired at two significantly different resolution levels, some of the details available in the more densely sampled one cannot be matched in the other one. From a sophisticated point of view, we need to be able to classify and weight any sampled data in terms of the sampling resolution. We aim to mimic registration task for the pairs obtained at different resolution levels (Fig. 8), as if we obtained them with two different sensors of the same modality (3D range data in our case). The effect is studied for the synthetic head pairs through changing the resolution of the data uniformly to a limited set of values and keeping the resolution of the reference model the same. The response of the criterion to the resolution change is investigated. The analysis for this data shows, that the algorithm is able to handle matching point clouds up to the resolution value of 0.6.

7. Experimental results for real 3D datasets

Several 3D datasets acquired by various sensors, which will be described in the following, are used to study the performance of the algorithm in one of its primary areas of applications, namely in the case of scene reconstruction from multiple 3D point-sets obtained from range images. The main focus in this second set of experiments is on the parameter σ and on designing an optimal scheme for accurate registration without close initialization. This required studying the registration error and basins of convergence of the algorithm for several datasets. A comparison of the robustness of algorithm with the standard Iterative Closest Point (ICP) algorithm, with respect to initialization, was also undertaken.

In our second set of experiments we employed several 3D range sensors operating at largely different scales to show the scope of the algorithm:

- The first sensor employed in this analysis, the LMS Z-210 from RIEGL Laser Measurement Systems GmbH, is a laser-based system that uses the time of flight principle for range measurement. The operating range of the LMS Z-210 goes from about 2 to 350 m, with an estimated noise level of about 20 mm and a range resolution of 25 mm.
- The second scanner is the IVP Ranger 2200 laser scanner from SICK IVP Integrated Vision Products and it is based on the principle of triangulation. This scanner has a camera with a two-dimensional CMOS sensor that captures a laser sheet of light which illuminates the imaged objects. Each profile contains 512 samples. A controlled relative motion between the camera and the scanned objects allows for the acquisition of a 2D range map with a depth resolution of about 2 mm.
- For smaller objects we acquired 3D imagery with a Leica SP2 LSCM confocal microscope that resolves objects as small as 1 μm. The confocal microscope uses depth from focus to build several layers of the scanned objects. A full three-dimensional model is reconstructed by assembling these layers.
- In addition, two commonly used datasets from the Stanford 3D Scanning Repository [36], Bunny and Happy Buddha, are used to demonstrate our technique on standard databases.

The first six datasets have been acquired using the sensors mentioned above, two per scanner. The datasets consist of a pair of 3D scans of various objects corresponding to the typical operating ranges of the cameras. The two scans are acquired from significantly different viewing points. Ground truths for the standard datasets are already given. To build a ground truth registration for our datasets, the surface is reconstructed from the original high resolution scans, corresponding points are picked by hand and the registering transformation is computed using the classic absolute orientation SVD-based technique [18]. A refinement step is performed using an extended ICP version that takes into account nor-

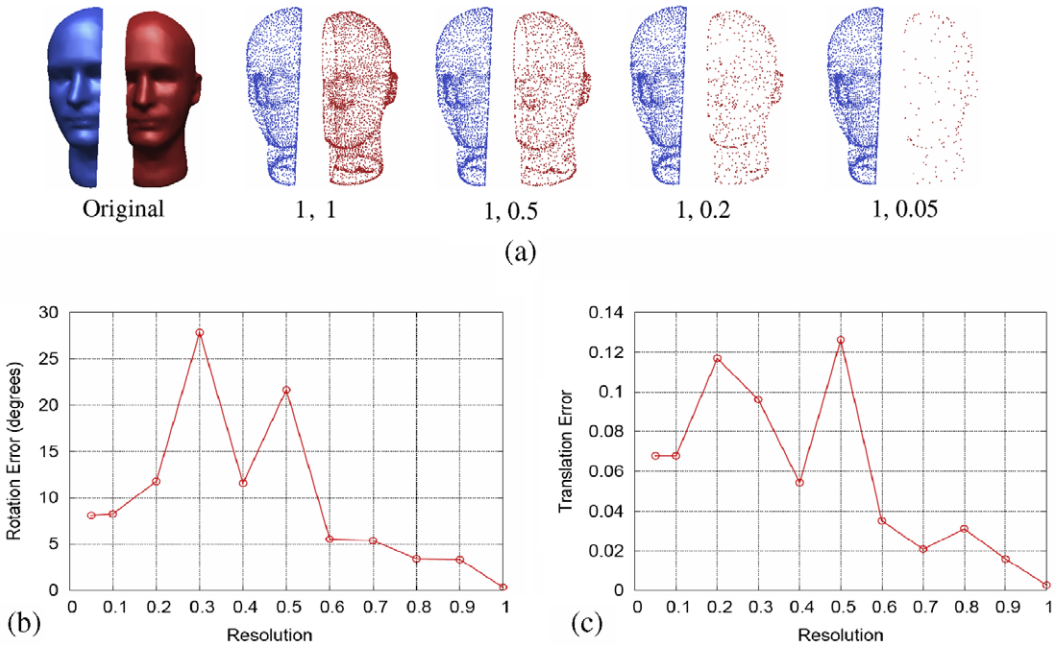

Fig. 8. Different dataset resolutions and its effect on the registration accuracy: (a) the model is kept at the same resolution and the data is stored at different resolutions. The plots of the (b) rotation and (c) translation errors versus the resolution of the data.

mal information. In our actual analysis we use lower resolution sub-sampled models (less then 10,000 points), and all the results are obtained in the context of point-set registration. Table 1 gives general information about the dimensions and the sensor which they were acquired with. Fig. 9 shows the different datasets used in the experiments.

7.1. Profiles

We start by plotting the profile of the Gaussian criterion around the ground truth registration for all the datasets. The behavior is quite similar for all the six pairs of views, with the usual dominant mode resembling the one of the simple 2D experiment shown in Fig. 2. We show in Fig. 10 sample plots obtained using the 'Parts' dataset. The plots are generated for various values of the Gaussian force range parameter σ ranging from 5% to 100% of the length of the datasets. The original plots for one translation parameter along with the plots for one rotation parameter are shown in Fig. 10(a). The increase in σ results in an increase of the amplitude (dimensionless) of the criterion. It also leads to a drift in the position of the maximum away from the correct position, a behavior that we explained in the theory. To emphasize the increase in the width of the dominant mode of the Gaussian criterion we show the same plots scaled by the maximum in Fig. 10(b). In these latter plots we can see an increase that is almost proportional to the increase in the values of σ for the lower ranges of this parameter but which is slower for the higher ones. This behavior is discussed in the next sections.

Table 1
Information about the real-world datasets used in the analysis.

Name of the dataset	Bounding box dimensions	Sensor
Van, Fig. 9(a)	5.23 × 4.42 × 2.53 m	RIEGL, time of flight laser scanner
Building, Fig. 9(b)	33.59 × 21.87 × 12.82 m	RIEGL, time of flight laser scanner
Parts, Fig. 9(c)	237 × 235 × 127 mm	Ranger 2200, laser triangulation system
Boat, Fig. 9(d)	462 × 273 × 140 mm	Ranger 2200, laser triangulation system
Gear, Fig. 9(e)	1385 × 1462 × 125 μm	Leica, confocal microscope
Microchip, Fig. 9(f)	149 × 149 × 8 μm	Leica, confocal microscope
Bunny, Fig. 9(g)		Cyberware 3030 MS [36]
Happy Buddha, Fig. 9(h)		Cyberware 3030 MS [36]

7.2. Effects of the parameter σ

From the graphs generated for all six datasets, with the different values of σ we can measure the drift of the dominant mode in the different dimensions of the registration parameter space and plot their evolution. This is shown in Fig. 11 which renders the evolution of localization error with σ. The overall behavior is similar for the six datasets, in the sense that it starts with an almost linear increase in the drift as a function of σ. For larger values this drift is much slower and tends toward an asymptotic limit. This can be explained by the fact that, as shown earlier, the force range parameter controls the influence of outliers, hence the relatively rapid increase in the lower range in particular for a dataset with low amount of overlap such as the 'Microchip'.

The asymptotic stabilization is explained by the fact that as σ exceeds the average distance between points in the datasets the exponential can be approximated by its first order development

$$\exp\left(-\frac{d^2(Tr(P_i),Q_j)}{\sigma^2}\right) \approx 1 - \frac{d^2(Tr(P_i),Q_j)}{\sigma^2}. \tag{17}$$

The actual optimization problem will not depend on σ anymore. In the case of a large force range parameter the problem of maximizing the Gaussian criterion is equivalent to minimizing the sum of average distances from the points of one dataset to the other dataset:

$$\min_{Tr} \sum_{i=1\ldots N_M} \bar{d}^2(Tr(P_i),D) \quad \text{with} \quad \bar{d}^2(Tr(P_i),D) = \frac{1}{N_D} \sum_{j=1\ldots N_D} d^2(Tr(P_i),Q_j). \tag{18}$$

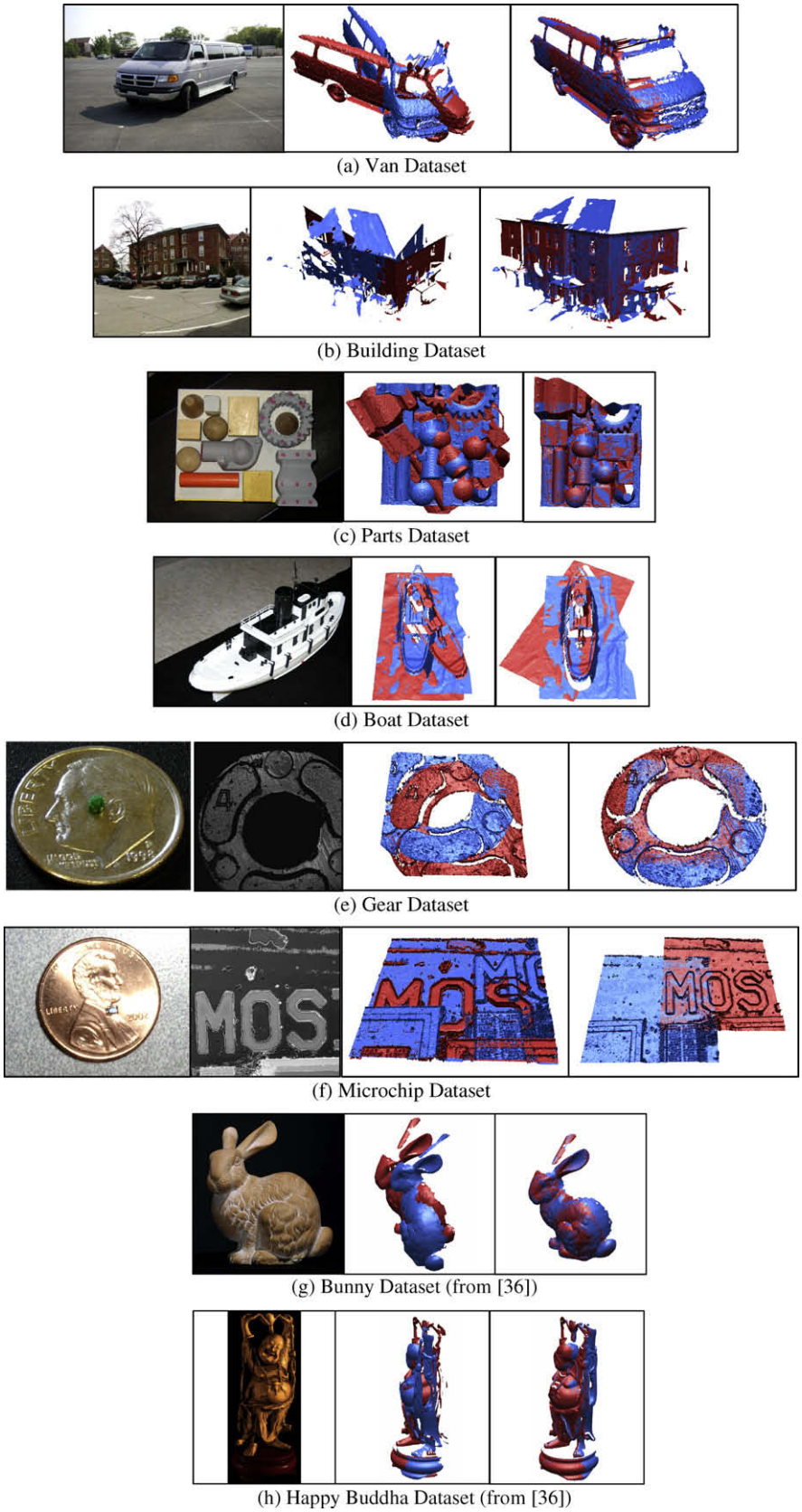

Fig. 9. Real-world datasets used in the analysis. For each dataset a color image is shown to the left, then the 3D views in unregistered position in the middle, followed by the registered views to the right.

This problem is independent of the value of σ, whence the asymptotic behavior. This fact insures that the registration error is bounded from above, for the plots shown the maximum error for most datasets is between 7% and 10% for translation and between 2° and 20° for the rotation, excluding the 'Microchip' dataset which has an overlap area that we consider too small and which is very

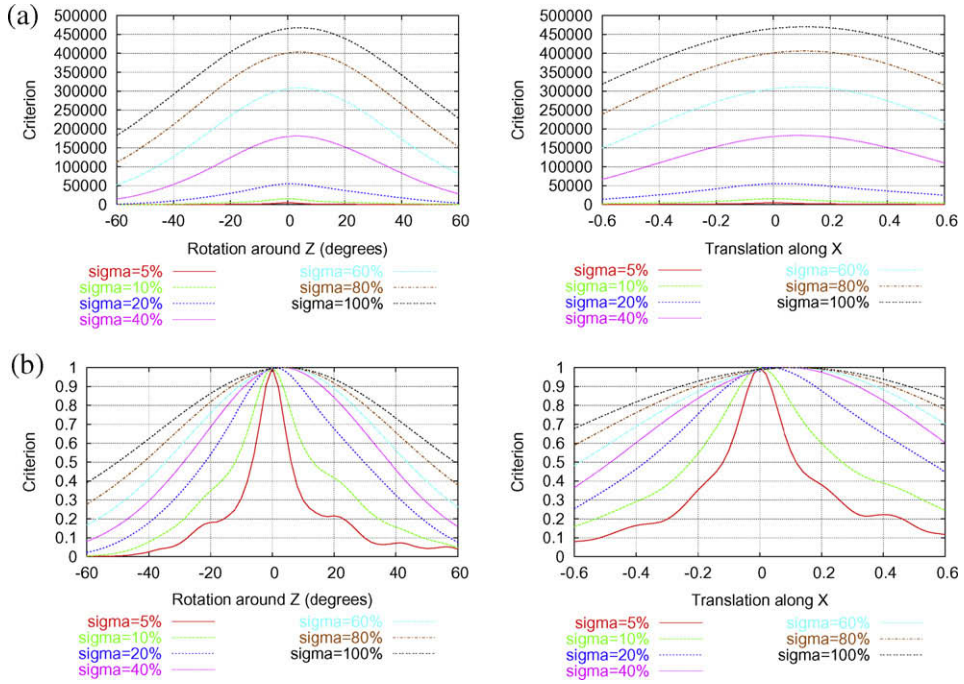

Fig. 10. (a) The plots of the Gaussian criterion versus the rotation around the z-axis (perpendicular to the plan of view of the datasets in Fig. 9), and the translation along the x-axis (the horizontal axis in Fig. 9) for seven values of σ (as a % of the length). (b) The plots are the scaled versions (of unit maximum) of those in (a) emphasizing the relative increase in the width of the dominant mode.

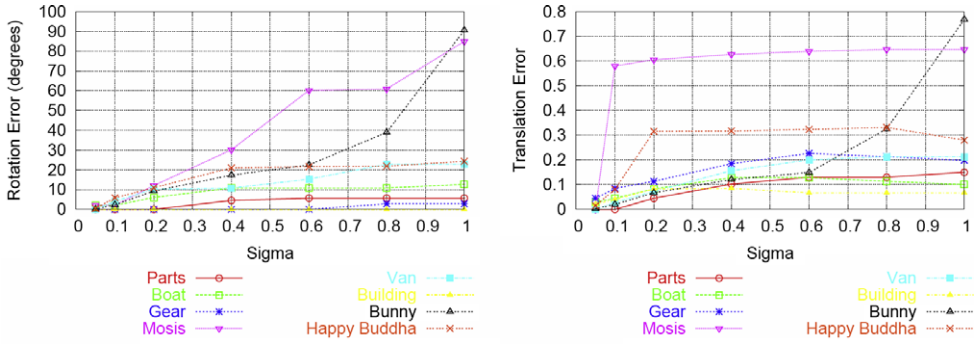

Fig. 11. Plots showing the evolution of the rotation and translation localization error, represented as a function of the parameter σ (as a fraction of the length).

flat. Such dataset presents great difficulty for registration algorithms.

7.3. Basins of convergence

The effect of σ on the region of convergence (ROC) was theoretically discussed when we derived the criterion by mollification and relaxation from the Boolean function. Here we analyze the basins of convergence of the algorithm for four of our datasets and two standard datasets (see Fig. 12). The plots show the relationship between the initial value of the transformation parameters provided to the algorithm and the residual registration error at the end of the process. These so-called basins of convergence were obtained for several values of σ. What these plots confirm is the tradeoff between a large basin of convergence for a large value of σ associated with a large residual error as well, and a smaller basin of convergence for lower values of σ that come with a better registration accuracy. This fact argues again for the two-steps scheme discussed before. We note that the width of the basins will first grow fast but will not increase much after a certain value of the force range parameter which was already deduced from the profiles of the criterion. Also the width of these basins is significantly larger than the value of σ (generally around 10 times for values less then 5%).

These basins are compared with those of the point-based ICP algorithm and we notice that they are wider for all datasets even for low values of σ. This is to be expected, since ICP is a close-range locally convergent scheme. On the other hand, ICP has a smaller residual error except when compared with the algorithm tuned for close-range Gaussian Fields. A balance between residual error and ROC size may be achieved by the adaptive optimization strategy.

7.4. Mean squared error comparison

In this section we will show: (1) that our method has an optimum parameter σ which balances the accuracy of registration and the region of convergence with respect to the initial starting parameters of the Gaussian Fields algorithm and (2) that for a large range of σ values our algorithm outperforms the ICP algorithm with respect to the Mean Squared Error (MSE) criterion. We start by decomposing the MSE into Bias and Variance terms. The MSE

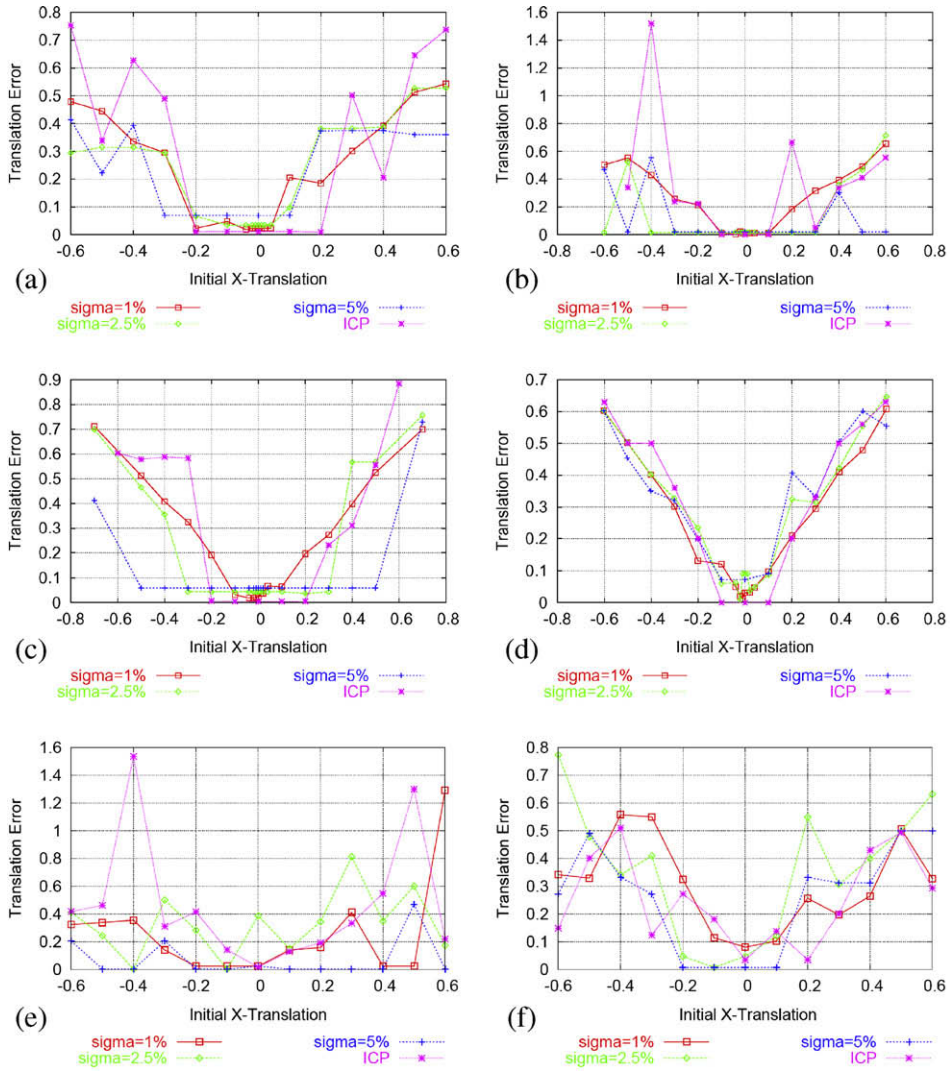

Fig. 12. Basins of convergence of our method for several values of σ and of the ICP method. The plots show the residual registration error as a function of initialization for the translation along the x-axis for (a) Building, (b) Boat, (c) Gear, (d) Microchip (e) Bunny, and (f) Happy Buddha datasets. The region of convergence can be directly measured in the graphs by evaluating the width of the bottom of the basin. When done over several datasets, as is the case in our experiments, the graphs give a good idea about the improvements brought by our algorithm as compared to ICP. An example of ROC measurement from (c) is, $ROC_{Gaussian}(Gear) = 1.0$ and $ROC_{ICP}(Gear) = 0.4$.

criterion offers a natural benchmark for quantifying the quality of registration across a range of perturbations. Since we are primarily interested in the sensitivity to initialization (which is the main weakness of current algorithms) we will compute the MSE with respect to a distribution of initial starting points of the algorithm. Furthermore, the MSE will give us a way of directly comparing the performance of our method with the ICP algorithm. Let the actual registration parameters for a pair of datasets be Tr^* (the ground truth in our case), and let $Tr^\sigma(x)$ the result of the registration for a given initial guess x and for a given σ. The MSE with respect to the distribution of initial parameters is given by

$$\begin{aligned} MSE &= E_x((Tr^\sigma(x) - Tr^*)^2) \\ &= E_x((Tr^\sigma(x) - E_x(Tr^\sigma(x)))^2) + E_x(Tr^\sigma(x) - Tr^*)^2. \end{aligned} \tag{19}$$

Here the subscript x of E_x denotes that the MSE values are calculated especially using two translation vectors, first the translation vector outputs of the criterion for the initial translations along x axis and second the ground truth translation vector. The term $E_x((Tr^\sigma(x) - Tr^*))^2 = B_\sigma^2$ represents the square of the Bias of the registration, while $E_x((Tr^\sigma(x) - E_x(Tr^\sigma(x)))^2) = V_\sigma$ is the Variance of the registration parameters with respect to the initial relative positions of the datasets. Hence, $MSE = V_\sigma + B_\sigma^2$.

We have shown that if σ is large, the criterion will become independent of this parameter. Therefore $\lim_{\sigma\to\infty} V_\sigma = 0$ and $\lim_{\sigma\to\infty} B_\sigma = B$, a constant value. For the lower values of σ the bias will decrease given that the maximum of the criterion will be closer to the correct position as discussed earlier but the narrowing of the width of the dominant mode and the appearance of local maxima will result in the increase of the variance of the algorithm with respect to initialization. This behavior is illustrated by the plots of Fig. 13. In the case of ICP, we know that the method is precise locally, hence characterized by a local low bias. Thus, we can expect a high variance over a large range of initializations.

For a quantitative comparison of the performance of the Gaussian Fields method and ICP we use the same four real-world datasets. We illustrate the difference between the two methods by using a uniform distribution of initial translations (along the x-axis), in the same way that we obtained the basins of convergence. For each dataset we compute $MSE = V_\sigma + B_\sigma^2$ over the different initializations. The results are obtained for several values of σ set as a

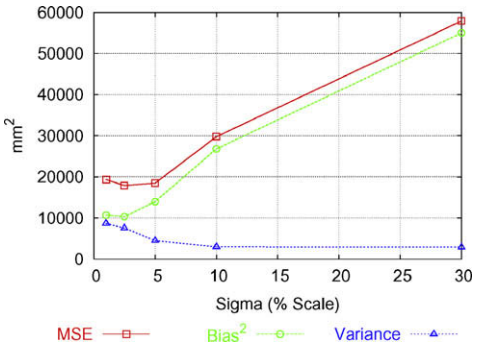

Fig. 13. Evolution of the Bias and Variance components of the Mean Squared Error computed with respect to a uniform distribution of initial translations for the Gaussian Fields method and for increasing values of σ ('Boat' dataset). The Bias will increase as σ increases while the Variance becomes smaller.

fraction of the size of the datasets (ranging from 0.5% to 30%). The plots of Fig. 14 show the variation of MSE criterion w.r.t. σ as compared to the MSE of the ICP algorithm. For all datasets there is a point at which the Gaussian method will outperform the ICP algorithm. A minimum of the MSE is obtained w.r.t. σ corresponding to an optimal behavior balancing the Bias and Variance constraints. The threshold below which the Gaussian method is better than ICP as well as the optimal σ are inevitably dependant on the datasets, as can be seen from the different plots.

8. Conclusion

This paper describes an effort at further automating the registration task, which is one of the main bottlenecks in modeling and fusion pipelines. The ICP algorithm is the basis for close-range accurate registration, a task for which it is well suited. In this paper, we have stressed the fundamental limits of this method pointing to the non-differentiable nature of the criterion that required a specialized optimization heuristic. We addressed the question of how to design a new criterion which balances the two conflicting goals of registration accuracy and large range of convergence. We have shown through theoretical and experimental analysis and comparison that the Gaussian Fields method developed in this research offers a good answer to this question. The criterion stems from a clear and rigorous formulation of the registration task as a search for the maximum overlap between the datasets. We extend the overlap to the multi-dimensional space of both position and local attributes.

This formulation allows for the easy incorporation of intensity and shape information in the registration framework. Our formulation derives from the use of simple Boolean matching principles, along with mollification and relaxation techniques. The Gaussian criterion gives a straightforward understanding of the effect of outliers as well and provides a mechanism for their control. An intuitive physical interpretation of the method can be provided by analogy to particle physics where points are subject to some exponentially decaying force fields. The criterion has the nice properties of continuous differentiability and can be also shown to be locally convex in the neighborhood of the registered position. Such properties allow for the use of the well developed gradient-based optimization techniques. In addition to easy optimization the criterion allows for the increase of the range of convergence by tuning the Gaussian smoothing parameter to have a two-stage global scheme. Hence we have with the same framework both initialization and refinement without need for explicit point-feature matching. Finally by using the powerful Fast Gauss Transform numerical technique, we save our method from what was going to be its main drawback which is the nearly $O(N^2)$ computational complexity. By employing clever analytic 'tricks' it clusters several data points substituting them with a small number of field sources and targets, thus achieving a remarkable linear $O(N)$ complexity. Most current implementations of ICP are of the order $O(N \cdot \log(N))$. We think that several additional multi-dimensional matching tasks can benefit from the Gaussian Fields framework, and we are currently extend-

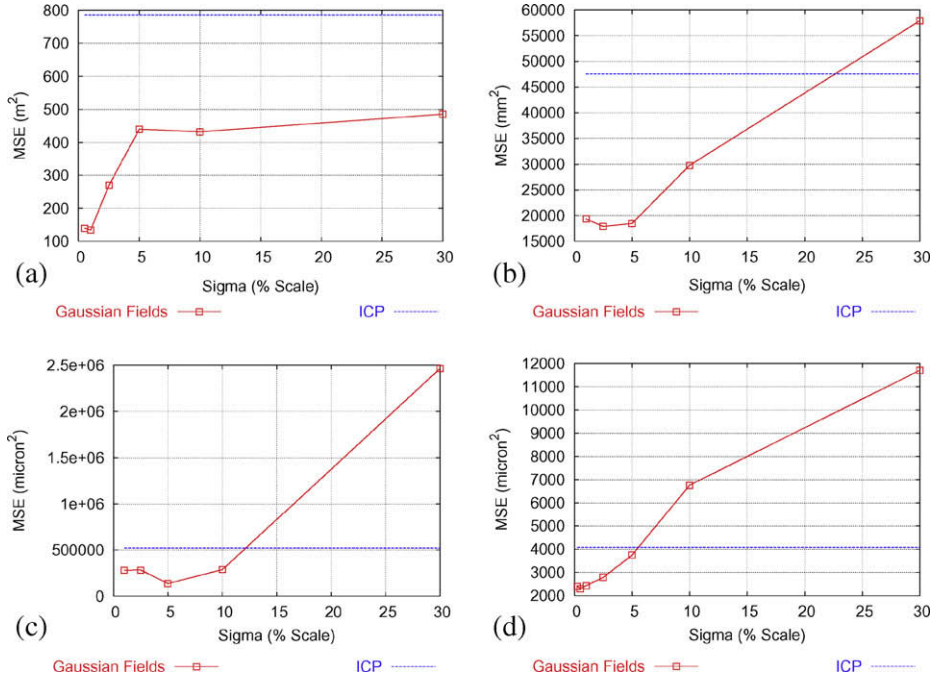

Fig. 14. Comparison of the MSE of the Gaussian Fields method with the MSE of ICP (MSE computed with respect to initialization) for (a) Building, (b) Boat, (c) Gear, and (d) Microchip dataset. We notice in general: (1) there is a threshold for σ below which the Gaussian Fields method will have a lower MSE than the ICP algorithm, and (2) there exists an optimum σ that minimizes the MSE uncertainty criterion.

ing the method to handle non-rigid alignment and video tracking using color and shape.

Acknowledgements

This work was supported in part by the DOE University Research Program in Robotics under Grant DOE-DE-FG02-86NE37968 and by the DOD/TACOM/NAC/ARC Program, R01-1344-18.

References

[1] R. Bergevin, M. Soucy, H. Gagnon, D. Laurendeau, Towards a general multi-view registration technique, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (5) (1996) 540–547.
[2] P.J. Besl, N.D. McKay, A method for registration of 3-d shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (2) (1992) 239–256.
[3] F. Boughorbel, A. Koschan, B. Abidi, M. Abidi, Gaussian fields: a new criterion for 3d rigid registration, Pattern Recognition 37 (7) (2004) 1567–1571.
[4] F. Boughorbel, A. Koschan, B. Abidi, M. Abidi, Gaussian energy functions for registration without correspondences, in: Proc. Int'l Conf. Pattern Recognition, vol. 3, IEEE, Cambridge, UK, 2004, pp. 24–27.
[5] F. Boughorbel, A. Koschan, M. Abidi, Automatic registration of 3d datasets using Gaussian fields, in: Proc. Int'l Conf. Image Processing, vol. III, IEEE, Genoa, Italy, 2005, pp. 804–807.
[6] R. Campbell, P. Flynn, A survey of free-form object representation and recognition techniques, Computer Vision and Image Understanding 81 (2) (2001) 166–210.
[7] G. Charpiat, O. Faugeras, R. Keriven, Shape metrics, warping and statistics, in: Proc. Int'l Conf. Image Processing, vol. 2, IEEE, Barcelona, Spain, 2003, pp. 627–630.
[8] Y. Chen, G. Medioni, Object modeling by registration of multiple range images, Image and Vision Computing 10 (3) (1992) 145–155.
[9] C. Chen, Y. Hung, J. Cheng, RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (11) (1999) 1229–1234.
[10] G. Dalley, P. Flynn, Pair-wise range image registration: a study in outlier classification, Computer Vision and Image Understanding 87 (1–3) (2002) 104–115.
[11] C. Dorai, J. Weng, A.K. Jain, Optimal registration of object views using range data, IEEE Transactions on Pattern Analysis Machine Intelligence 19 (10) (1997) 1131–1138.
[12] A. Elgammal, R. Duraiswami, L. Davis, Efficient kernel density estimation using the Fast Gauss Transform with applications to color modeling and tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (11) (2003) 1499–1504.
[13] O. Faugeras, M. Hebert, The representation recognition and locating of 3-d objects, International Journal on Robotic Research 5 (3) (1986) 27–52.
[14] A.W. Fitzgibbon, Robust registration of 2d and 3d point sets, Image and Vision Computing 21 (12–13) (2003) 1145–1153.
[15] L. Greengard, The Rapid Evaluation of Potential Fields in Particle Systems, MIT Press, Cambridge, Massachusetts, 1988.
[16] L. Greengard, J. Strain, The fast Gauss transform, SIAM Journal on Scientific Computing 12 (1) (1991) 79–94.
[17] M. Hebert, K. Ikeuchi, H. Delingette, A spherical representation for recognition of free-form surfaces, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (7) (1995) 681–690.
[18] B.K.P. Horn, Closed-form solution of absolute orientation using unit quaternions, Journal of Optical Society of America A 4 (4) (1987) 629–642.
[19] D. Huber, M. Hebert, Fully automatic registration of multiple 3D data sets, Image and Vision Computing 21 (7) (2003) 637–650.
[20] A.E. Johnson, M. Hebert, Using spin images for efficient multiple model recognition in cluttered 3-d scenes, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (5) (1999) 433–449.
[21] T. Masuda, N. Yokoya, A robust method for registration and segmentation of multiple range images, Computer Vision and Image Understanding 61 (3) (1995) 295–307.
[22] G. Medioni, M.S. Lee, C.K. Tang, A Computational Framework for Segmentation and Grouping, Elsevier, Amsterdam, 2000.
[23] D.A. Murio, The Mollification Method and the Numerical Solution of Ill-Posed Problems, Wiley, New York, 1993.
[24] S.Y. Park, M. Subbarao, Automatic 3d reconstruction based on novel pose estimation and integration techniques, Image and Vision Computing 22 (8) (2004) 623–635.
[25] H. Pottmann, S. Leopoldseder, M. Hofer, Registration without ICP, Computer Vision and Image Understanding 95 (1) (2004) 54–71.
[26] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, New York, NY, 1986.
[27] M. Rodrigues, R. Fisher, Y. Liu, Special issue on registration and fusion of range images, Computer Vision and Image Understanding 87 (1–3) (2002) 1–7.
[28] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, in: Proc. 3D Digital Imaging and Modeling, IEEE, Québec City, Canada, 2001, pp. 145–152.
[29] F.A. Sadjadi, E.L. Hall, Three-dimensional moment invariants, IEEE Transactions on Pattern Analysis and Machine Intelligence 2 (2) (1980) 127–136.
[30] G.C. Sharp, S.W. Lee, D.K. Wehe, ICP registration using invariant features, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (1) (2002) 90–102.
[31] F. Stein, G. Medioni, Structural indexing: efficient 3-d object recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (2) (1992) 125–145.
[32] A.J. Stoddart, A. Hilton, Registration of multiple point sets, in: Proc. 13th Int'l Conf. on Pattern Recognition, IEEE, Vienna, Austria, 1996, pp. 40–44.
[33] B. Taati, M. Bondy, P. Jasiobedzki, M. Greenspan, Automatic registration for model building using variable dimensional local shape descriptors, in: Proc. 3DIM 2007, Montreal, Quebec, Canada, 2007, pp. 265–272.
[34] J.-P. Thirion, New feature points based on geometric invariants for 3d image registration, International Journal of Computer Vision 18 (2) (1996) 121–137.
[35] E. Trucco, A. Fusiello, V. Roberto, Robust motion and correspondence of noisy 3-d point sets with missing data, Pattern Recognition Letters 20 (9) (1999) 889–898.
[36] G. Turk, M. Levoy, Zippered polygon meshes from range images, in: Proc. ACM SIGGRAPH, 1994.
[37] C. Yang, R. Duraiswami, N.A. Gumerov, L. Davis, Improved fast Gauss transform and efficient kernel density estimation, in: Proc. Ninth Int'l Conf. Computer Vision, IEEE, Nice, France, 2003, pp. 464–471.
[38] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, International Journal of Computer Vision 13 (2) (1994) 119–152.
[39] T. Zinßer, J. Schmidt, H. Niemann, A refined ICP algorithm for robust 3-D correspondence estimation, in: Proc. Int'l Conf. Image Processing, vol. 2, IEEE, Barcelona, Spain, 2003, pp. 695–698.