Gaussian Energy Functions for Registration without Correspondences

Faysal Boughorbel, Andreas Koschan, Besma Abidi, and Mongi Abidi
Imaging Robotics and Intelligent Systems Laboratory
The University of Tennessee, Knoxville, TN, 37996
fboughor@utk.edu

Abstract

A new criterion based on Gaussian fields is introduced and applied to the task of automatic rigid registration of point-sets. The method defines a simple energy function, which is always differentiable and convex in a large neighborhood of the alignment parameters; allowing for the use of powerful standard optimization techniques. We show that the size of the region of convergence can be extended so that no close initialization is needed, thus overcoming local convergence problems of Iterative Closest Point algorithms. Furthermore, the Gaussian energy function can be evaluated with linear complexity using the Fast Gauss Transform, which permits efficient implementation of the registration algorithm. Analysis through several experimental results on real world datasets shows the practicality and points out the limits of the approach.

1. Introduction

The task of automatically registering 3D free-form surfaces is critical in many computer vision applications, such as scene modeling from multiple range views and object recognition. In the case of volumetric datasets, several similarity measures were employed to align multi-modal and single-modality datasets, which are mostly based on correlating intensity values. In this paper we focus on the problem of rigid registration at the point level. In real applications, surface samples are provided by a variety of imaging sensors, such as laser scanning devices or stereovision systems.

The registration of 3D datasets consists of the recovery of transformations that align the partial views. In the case of rigid registration, the rotation and translation parameters could be computed in a closed form solution if point correspondences are available [7]. Recovering those matches automatically is the first task in many modeling systems. This step, however, results usually in a rough registration that needs to be refined using Iterative Closest Point (ICP) algorithms [1][8].

To establish point correspondences in 3D datasets, several invariant feature extraction techniques were proposed. Methods surveyed by Campbell and Flynn [2] are mostly surface-based and use differential properties to build their representations. Most invariant feature methods assume that the partial surfaces were accurately reconstructed from the range maps. In real applications, however these surfaces are affected by noise, reducing the accuracy of invariant features registration, hence the need for further refinement using point-based techniques. The ICP algorithm is a locally convergent scheme that requires parameter initialization close to the aligned position. It operates at the point level, minimizing the mean squared distance between the datasets. Despite its popularity, ICP has several shortcomings, including the need for sufficient overlap between the datasets and local convergence.

The main contribution of this work is the design of a point-sets registration criterion which is differentiable and convex in a large neighborhood of the aligned position. The underlying motivation is to overcome the problems of standard registration techniques, and in particular ICP. One of the main reasons behind these limitations is the non-differentiable cost function associated with ICP that impose local convergence, so that in real applications the preliminary point-feature extraction and matching step is necessary before proceeding to what is considered a refinement step. Some work was done recently on the design of approximations to non-differentiable matching and similarity measures including the work by Charpiat et al [3] on approximating Hausdorff distances by a differentiable metric on shape space. In our case we will use a straightforward sum of Gaussian distances that is defined for point-sets with associated attributes. These attributes are local moments computed from the datasets. We show that this criterion can be used for accurate registration, while at the same time extending the region of convergence so that we do not need any close initialization.

Our energy function is convex in the neighborhood of the solution and always differentiable allowing for the use of a wide range of well proven optimization techniques. More importantly, the criterion can be evaluated, with linear complexity, using the recent numerical technique known as the Fast Gauss Transform [4], making it computationally less expensive than current registration algorithms. In the following sections we first present the Gaussian energy function, describe the local attributes used, present an overview of the fast evaluation method,

0-7695-2128-2/04 $20.00 (C) 2004 IEEE

and finally show an analysis of our approach based on several experimental results.

2. Gaussian fields and the energy function

Our basic idea is to use a Gaussian field to measure both the spatial proximity and visual similarity of two points belonging to the two datasets. Consider first the two point-sets $M = \{(P_i, S(P_i))\}_{i=1...N_M}$ and $D = \{(Q_j, S(Q_j))\}_{j=1...N_D}$, with their associated attribute vectors. Those vectors can include curvature for smooth surfaces and curves, invariant descriptors, as well as color attributes when available. The Gaussian measure of proximity and similarity between two points is given by:

$$F(P_i, Q_j) = \exp\left(-\frac{d^2(P_i, Q_j)}{\sigma^2} - (S(P_i) - S(Q_j))^T \Sigma^{-1} (S(P_i) - S(Q_j))\right) \quad (1)$$

with $d(P_i, Q_j)$ being the Euclidean distance between the points. The expression can be seen as a force field whose sources are located at one point and are decaying with distance in Euclidean and attribute space. The parameter σ controls the decay with distance while Σ, a diagonal matrix with small components, punishes the difference in attributes. We can now define an energy function that measures the registration of M and D as:

$$E(Tr) = \sum_{\substack{i=1...N_M \\ j=1...N_D}} \exp\left(-\frac{d^2(P_i, Tr(Q_j))}{\sigma^2} - (S(P_i) - S(Tr(Q_j)))^T \Sigma^{-1} (S(P_i) - S(Tr(Q_j)))\right) \quad (2)$$

where Tr is the transformation that registers the two point-sets. In this paper we will focus mainly on the rigid case: $Tr(Q_j) = RQ_j + t$. More general cases can be handled as well. We also use as visual attributes moments which are invariant to rigid transformations [9] which are de-correlated and combined using the same approach adopted by Sharp et al [10]. If we choose the decay parameters to be very small the energy function E will just 'count' the number of points that overlap at a given pose. This is due to the exponential being very small except for $P_i = (RQ_j + t)$ and $S(P_i) = S(Q_j)$. In particular, if M is a subset of D we will have at the registered position (R^*, t^*): $\lim_{\substack{\sigma \to 0 \\ \Sigma \to 0}} E(R^*, t^*) = N_M$. Thus, for this case we meet a rigorous definition of registration as maximization of both overlap and local shape similarity between the datasets.

The Gaussian energy function is convex around the registered position and is always differentiable, allowing for the design of efficient registration algorithms. Several powerful optimization techniques can be used for this task such as the quasi-Newton and conjugate gradient algorithms. The parameter σ controls the size of the convex safe region of convergence. Therefore we are interested in increasing σ as much as possible. This will be mostly limited by the decrease in the localization accuracy of our criterion, hence the tradeoff between large region of convergence and precision. Studying the cost function, we found that the region of convergence can be extended considerably in the case of complex datasets and when using as many independent local descriptors as possible. This can be illustrated with the behavior of the matching criteria with and without attributes as illustrated in Fig. 1. The profile of the criterion was plotted for relative displacement of the two point sets of Fig. 1(a). Several plots are shown with increasing σ. For the non-attributed case (Fig. 1(b)) we notice that as σ increases, the width of the Gaussian bell increases too, but the maximum starts to drift away from the correct position. When we use the Gaussian criterion with moment invariants, as attributes associated with the points, the maximum is stable for the same values of σ (Fig. 1(c)). In the analysis section we use additional real datasets to study the localization error as a function of σ.

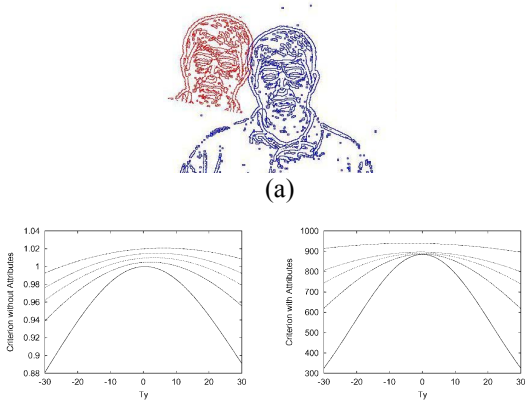

Figure 1. Profiles of the Gaussian energy function for a displacement around the registered position of the datasets shown in (a). In (b) the profiles are plotted in the case without attributes for $\sigma = 30,50,70,90,150$ (from narrowest to widest). Plots with moment invariants as attributes for the same values of σ are shown in (c). The scale of the datasets is about 200×200. (For (b) magnitudes were rescaled for comparison).

3. The Fast Gauss Transform

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. Similar limitations are encountered in other

computer vision tasks, especially for Gaussian kernel density estimation. A new numerical method, called the Fast Gauss Transform, was recently employed in color modeling and tracking applications [4] in order to reduce the computational complexity of Gaussian Mixture evaluation to $O(N_M + N_D)$. The method, which belongs to a new class of fast evaluation algorithms known as "fast multipole" methods, was first introduced by Greengard and Strain [5][6] and applied to potential fields computations. The basic idea is to exploit the fact that all calculations are required only up to a certain accuracy. In this framework, the sources and targets of potential fields were clustered using suitable data structures, and the sums were replaced by smaller summations that are equivalent to a given level of precision.

To evaluate sums of the form:

$$S(t_i) = \sum_{j=1}^{N} f_j \exp(-(\frac{s_j - t_i}{\sigma})^2), \; i = 1,..., M \tag{3}$$

where $\{s_j\}_{j=1,...,N}$ are the centers of the Gaussians known as sources and $\{t_i\}_{i=1,...,M}$ the targets. The following shifting identity and expansion in terms of Hermite series are used:

$$\exp(\frac{-(t-s)^2}{\sigma^2}) = \exp(\frac{-(t-s_0-(s-s_0))^2}{\sigma^2}) \tag{4}$$
$$= \exp(\frac{-(t-s_0)^2}{\sigma^2}) \sum_{n=0}^{\infty} \frac{1}{n!} (\frac{s-s_0}{\sigma})^n H_n(\frac{t-s_0}{\sigma})$$

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 (5) can similarly be used to group targets together at a cluster center t_0, further reducing the number of computations.

$$\exp(-(\frac{t-s}{\sigma})^2) = \exp(\frac{-(t-t_0-(s-t_0))^2}{\sigma^2}) \tag{5}$$
$$\approx \sum_{n=0}^{p} \frac{1}{n!} h_n(\frac{s-t_0}{\sigma})(\frac{t-t_0}{\sigma})^n$$

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 [4] [6].

4. Analysis

Two critical issues for registration techniques are parameter estimation accuracy, and initialization for iterative methods. The complexity of the datasets is another important factor. In our case the parameter that controls the range of convergence is σ. Any two datasets that we want to register are supposed to have a certain amount of overlap. Hence, the two point-sets can be initialized at least in an arbitrarily overlapped position. The parameter σ determines the reach of the Gaussian field for each source point. If σ can be increased to ensure interaction between all the points without losing too much accuracy, that will allow fully automatic registration that is not sensitive to initialization.

For our experiments we use the two real datasets shown in Fig. 2. For both test datasets we have ground truth registration obtained by accurate position and orientation instruments. In our implementation of the Gaussian criterion, we used the three 3D moment invariants as attributes in addition to curvature, when surfaces were extracted. For optimization, a quasi-Newton algorithm was employed and convergence was obtained in few iterations. The speed of the algorithm was enhanced by employing the Fast Gauss Transform method. On these datasets, we studied the effect of increasing σ on the registration accuracy. The results of this experiment, performed on both the 'Van' and 'Parts' datasets, are shown in Fig. 2(c). An interesting fact that emerges from these plots is that registration accuracy first degrades rapidly up to a value of σ roughly equal to the length of the objects then deteriorates at a much slower rate. For σ between one and two times the objects length the average translational registration error is around 12% for the 'Parts' dataset and 17% for the 'Van'. The rotation error, in the same region, is 7º and 13º for the 'Parts' and 'Van' datasets respectively. For smaller values of σ (less than one length) the registration error becomes very small. Based on this behavior we can design an algorithm that starts from unknown initial pose with a large σ then decreases the value of this parameter until accurate transformations are obtained.

Using the same datasets, we studied the effect of uniform sub-sampling on the accuracy of the algorithm. The plots of Fig. 2(d) show that the Gaussian Energy function is robust to sub-sampling. The registration error remains very low until we reach a number of points of less than 1% of the original one (we have around 70000 points for the 'Parts' dataset and 35000 points for the 'Van' dataset). This also, in addition to the Fast Gauss

Transform, permits very fast implementations that improve on the current combination of local feature matching and ICP-based algorithms.

5. Conclusion

In this paper we show that a simple new criterion, based on the sum of Gaussians of distances, leads to a fully automatic registration framework. The method overcomes the current dichotomy between (1) initialization using feature extraction and matching techniques and (2) refinement using ICP based algorithms. Using a single energy function, we can start from arbitrary initial positions and converge accurately to the pose parameters. Furthermore, the method can be implemented with linear complexity using the recent numerical technique known as the Fast Gauss Transform, which clearly improves on current algorithms. Our experiments, performed on real noisy datasets, illustrate the behavior of the method and show its applicability to real world data. While in this paper we focused mainly on 3D rigid registration, the Gaussian energy approach can be used for non-rigid registration as well. We are currently investigating this task.

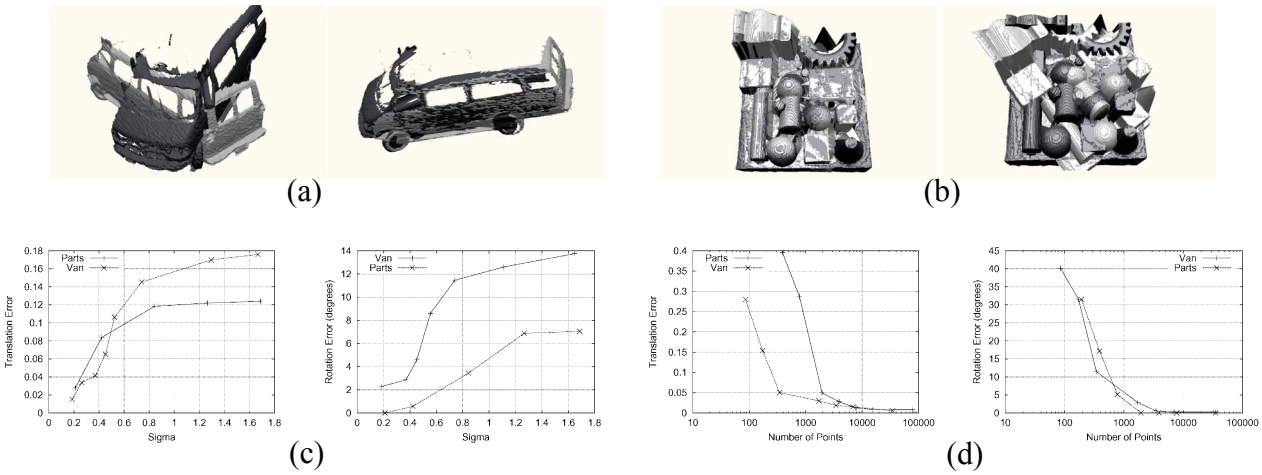

Figure 2. Two 3D views of a 14-passenger Van (a) and multiple parts and objects (b), shown in un-registered and registered positions. The registration error as a function of σ is shown in (c). Also the effect of sampling on the alignment using the Gaussian function is illustrated in (d).

Acknowledgments

This work is supported by the University Research Program in Robotics under grant DOE-DE-FG02-86NE37968, by the DOD/RDECOM/NAC/ARC Program, R01-1344-18, by the US Army under grant Army-W56HC2V-04-C-0044, and by FAA/NSSA Program, R01-1344-48/49.

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*, vol. 14, no. 2, pp. 239-256, 1992.

[2] R. Campbell and P. Flynn, "A Survey of Free-form Object Representation and Recognition Techniques", *CVIU*, vol. 81, no. 2, pp. 166-210, 2001.

[3] G. Charpiat, O. Faugeras, and R. Keriven, "Shape Metrics, Warping and Statistics," *in Proc. International Conference on Image Processing*, vol. 2, pp. 627-630, Barcelona, 2003.

[4] A. Elgammal, R. Duraiswami, and 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*, vol. 25, no. 11, pp. 1499-1504, 2003.

[5] L. Greengard, *The Rapid Evaluation of Potential Fields in Particle Systems*, Cambridge, Massachusetts, MIT Press, 1988.

[6] L. Greengard and J. Strain, "The Fast Gauss Transform," *SIAM J. Scientific Computing*, vol. 12, pp. 79-94, 1991.

[7] B. Horn, "Closed-form Solution of Absolute Orientation using Unit Quaternions," *Journal of the Optical Society of America A*, vol. 4, no. 4, pp. 629-642, 1987.

[8] S. Rusinkiewics, and M. Levoy, "Efficient Variants of the ICP Algorithm," *in Proc. of 3D Digital Imaging and Modeling, IEEE Computer Society Press*, pp. 145-152, 2001.

[9] F. A. Sedjadi and E. L. Hall, "Three-Dimensional Moment Invariants," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2. no. 2, pp. 127-136, 1980.

[10] G. C. Sharp, S. W. Lee, and D. K. Wehe, "ICP Registration using Invariant Features," *IEEE Transactions on Pattern Analysis Machine Intelligence,* vol. 24, no. 1, pp. 90-102, 2002