Robust Registration of Coarse Binary Volume Models

Jochen Wingbermühle
Institut für Theoretische Nachrichtentechnik
und Informationsverarbeitung, Universität Hannover
Appelstr. 9a, 30167 Hannover, Germany
wingbermuehle@tnt.uni-hannover.de

Abstract

In this paper we propose a robust approach to register two coarse binary volume models. The presented method is based on a combination of maximising the spatial cross-correlation and an extended version of the iterative closest point algorithm (ICP). During an iterative optimisation the cross-correlation. i.e. the section volume of both models, is maximised and the mean squared distance between surface points of the first model and their respective closest surface points on the second model is minimised simultaneously by varying the six parameters of an arbitrary spatial movement. The proposed method has been applied to merge different, independently created binary volume models of a rigid object. Each model is created by a shape from silhouette approach using a set of views captured by a single camera that observes the object rotating on a turntable. Since rotation around a single axis doesn't allow the camera to examine the entire object surface, different models, captured with varying orientations of the object on the turntable, have to be integrated to obtain a complete 3D model. Thus, merging two or more sequentially created coarse volume models is an essential technique to obtain high quality 3D models of real objects with a simple turntable setup.

1. Introduction

Several approaches for the automated reconstruction (scanning) of rigid 3D objects have been developed within the last decade. Today, 3D scanning systems that consist of a calibrated fixed sensor observing the respective object on a turntable are frequently in use. Such a setup is cheap and easy to build, but the relative motion between object and sensor is restricted to a rotation around a single axis, i.e. each scanning run can be understood as a sensor cycling on a fixed planar orbit around the resting object. Thus, the object surface can not be completely scanned, in particular the respective top and bottom part of the object can be reconstructed with reduced accuracy only. To overcome this limitation, it is desirable to allow the user to change the object's orientation relative to the turntable and scan it again, yielding two or more complementary models of the same object. Since the object motion between such subsequent scans is unknown the resulting locally erroneous models have to be registered in order to merge them into one improved model. In case of the shape from silhouette approach we use [8], each model is represented as a discrete binary volume model, i.e. a set of voxels describing the object volume. Whereas in computer vision applications registration is usually applied to merge partially overlapping sets of surface data of rigid objects (e.g. stereo- or laserscan data) [5, 6], in medical applications registration of surface and volumetric (e.g. tomography) data of non-rigid objects is subject to recent research activities [7]. In both fields the iterative closest point (ICP) algorithm of Besl and McKay [1] is frequently used and several extensions have been proposed [2, 3, 4]. The basic idea of the ICP algorithm is to iteratively minimise the mean distance between surface points from one data set and the respective closest points from the second set. Aiming at the registration of binary volume models, a straight forward approach would be to extract all voxels on the model's surfaces and apply the ICP as is. However, during our work it turned out that the additional use of spatial cross-correlation as a matching criterion significantly enhances the method's robustness against local shape errors.

After a brief description of the scanning system is given in the next section, the third section outlines the determination of suitable start-values. Some computational aspects are discussed in section 4 and the iterative optimisation is explained in section 5. Results are given in section 6 and we end up with some concluding remarks and future directions in section 7.

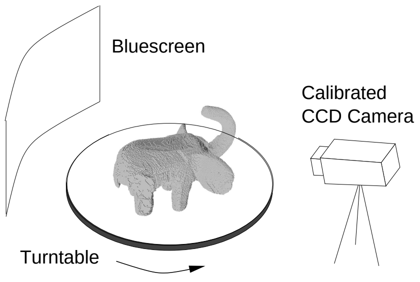

Figure 1. Turntable setup for data acquisition.

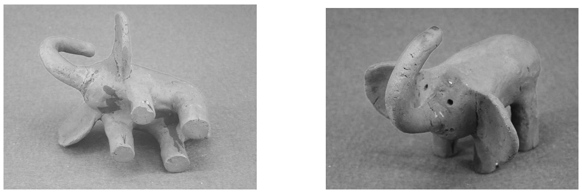

Figure 2. Examples of input images.

2. Data Acquisition

Our goal is to create high quality 3D models of a real object with a homogenous distribution of errors over the model surface. Following the approach of Niem [8], the respective object is placed on the turntable and rotated in front of a calibrated CCD camera in steps of 10 to 30 degrees (Fig.1). Between two subsequent steps the camera takes an image of the object in front of a blue screen (Fig.2). A chroma keying technique is then applied to extract the object's silhouettes from each image and the obtained silhouettes are back projected into space, each yielding a pyramidal volume. The spatial section of all of these volumes encloses the real object and is used as an approximation of the object shape. It is stored as a discrete binary volume model. After the first scanning run, the object is removed from the turntable, arbitrarily rotated (typically about 90 degrees), placed on the turntable again and a second model is acquired in the same way.

Fig.3 shows two models of an elephant sculpture. The model in the top row was acquired with the sculpture standing in an upright position, the second model while it was lying on it's side. Since only silhouette information is used to reconstruct the shape, both models show significant errors. Whereas the first model looks good from the side, it looks quite poor from the top. The second model shows a complementary error distribution. Thus, it is desirable to integrate both models in a way that the errors are mutually eliminated.

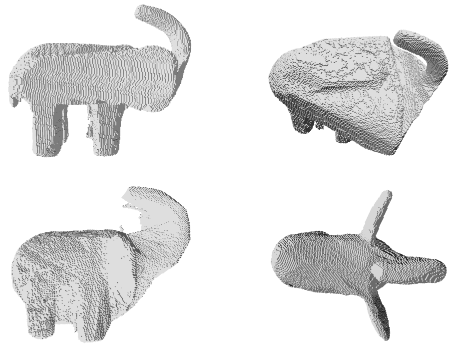

Figure 3. Side and top views of two captured volume models of an elephant sculpture.

3. Determination of Start-Values

In order to integrate both models they have to be registered, i.e. the six parameters of a spatial transformation needed to align both models as good as possible have to be determined. This is achieved during an iterative optimisation procedure which requires suitable start-values to prevent it from converging to local minima. As outlined by Besl and McKay [1], suitable start-values can usually be provided by aligning the centres of gravity and principle axes of both models. However, shape errors do affect the principle axes directions and therefore it is advisable to verify the resulting start-parameters. This is done by evaluating the ICP error criterion, i.e. calculating the mean squared distance between surface points of first model and the respective closest surface points of the second model after applying the obtained initial transformation. If the resulting value exceeds a predefined reasonable threshold, the start-values are rejected and a better set of parameters is heuristically searched by rotating the transformed second model in steps of 15 degrees around all three principle axes. The combination of angles which yields the smallest value of the ICP error criterion is then used as start-value for the subsequent optimisation.

4. Computational Aspects

The calculation of the ICP error criterion requires for each evaluated surface point the determination of its respective closest surface point of the second data set. The volume models we use have a typical resolution of 256 x 256 x 256 voxels (or higher) and an accordingly huge number of surface points. Instead of evaluating all these surface points we found the mutual evaluation of $N = 10000$ randomly selected surface points of both models to be sufficient. The ICP error criterion, E_{ICP}, can then be written as:

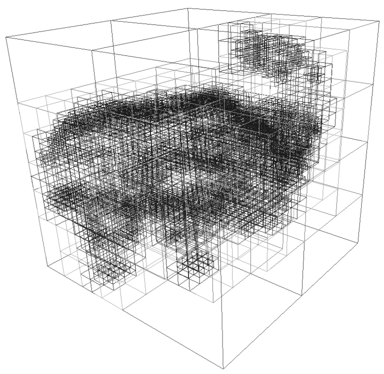

Figure 4. Octree based distance map.

$$E_{ICP} = \frac{D_{12} + D_{21}}{2N}$$

The terms D_{12} and D_{21} are the respective contributions of both models. $\mathbf{H}$ is a homogeneous 4x4 matrix which represents the transformation that has to be applied to model 2 in order to align it with model 1.

$$D_{12} = \underbrace{\sum_{i=1}^{N} \min\{\underbrace{|\mathbf{H}^{-1}\vec{p}_{1,i} - \vec{p}_{2,1}|^2, \ldots, |\mathbf{H}^{-1}\vec{p}_{1,i} - \vec{p}_{2,J}|^2}_{\text{all } J \text{ surface points of data set 2}}\}}_{\text{N surface points of data set 1}}$$

$$D_{21} = \sum_{j=1}^{N} \min\{|\mathbf{H}\vec{p}_{2,j} - \vec{p}_{1,1}|^2, \ldots, |\mathbf{H}\vec{p}_{2,j} - \vec{p}_{1,I}|^2\}$$

Even if D_{12} and D_{21} are restricted to N surface points of each model, they still require the evaluation of all (I or J) surface points of the respective other model. To avoid this time consuming calculation we pre-calculate octree distance maps for each model. Similar distance maps have been proposed by Szeliski and Lavallée [7]. Therefor the surface points of both models are extracted and used to construct corresponding octrees as follows: Starting with a root cube, e.g. the model's bounding box, surface points are subsequently added to this cube until a predefined point capacity is reached. Adding another point causes the cube to split in 8 equally sized sub-cubes and to distribute all his points, with respect to their position, among the sub-cubes. This procedure is recursively repeated until all surface points have been added. Finally, empty leaf cubes, i.e. cubes without sub-cubes are removed from the tree. Fig. 4 shows the resulting octree for one of the initial models of the elephant sculpture. It can be seen that the octree's leafs are getting smaller if we approach the model surface. Thus, we can use the leaf size, $LfS(\vec{p})$, i.e the edge length of the respective cube, at an arbitrary position, $\vec{p}$, as an unsigned approximation for the distance of this point to the model surface. For all points outside the initial cube the distance to the model surface is assumed to be equal to the edge length of the initial cube. This allows the following approximations for D_{12} and D_{21}:

$$\tilde{D}_{12} = \sum_{i=1}^{N} LfS_2^2(\mathbf{H}^{-1}\vec{p}_{1,i}) \ , \quad \tilde{D}_{21} = \sum_{j=1}^{N} LfS_1^2(\mathbf{H}\vec{p}_{2,j})$$

Searching the respective closest point for each evaluated point is hereby replaced with simply inserting the point in an octree and looking up the size of the obtained leaf, which gains a significant speed up.

5. Optimisation

Once a suitable start-value has been determined as outlined in section 3, the six registration parameters are iteratively optimised by minimising the following error criterion E:

$$E = \underbrace{\frac{\tilde{D}_{12} + \tilde{D}_{21}}{2N}}_{\text{ICP-term}} + \underbrace{w \cdot \min\{\frac{V_1 - V_s}{V_1}, \frac{V_2 - V_s}{V_2}\}}_{\text{correlation-term}}$$

This criterion is a weighted sum of two terms. The first term is the simplified version of the standard ICP criterion proposed in the previous section, the second term equals one minus the ratio between the section volume, V_s, for the current transformation and the volume, V_1 or V_2, of the respective smaller model. In case of binary volume models the section volume equals the cross-correlation of both models in their respective positions. Thus, minimising the second term can be understood as a maximisation of the cross-correlation of both models, which is, of course, a common matching technique. Whereas the minimisation of the first term still ends up in local minima sometimes, in particular for complex shapes, we found that adding the correlation term significantly enhances the reliability of results even in cases of nearly symmetric objects or very rough initial models. However, if during the optimisation a position is reached where the bigger model includes the smaller model completely, the correlation term vanishes and provides no further information on the registration quality, i.e. the ICP term becomes responsible for the fine adjustment. The fixed weight, w, of the second term is used to balance the influence of both terms, during our experiments we used values from an interval of $w = [10 \ldots 100]$. Although very powerful minimisation strategies for the ICP term exist [9] the introduction of the correlation term forced us to use the less efficient well known Powell method.

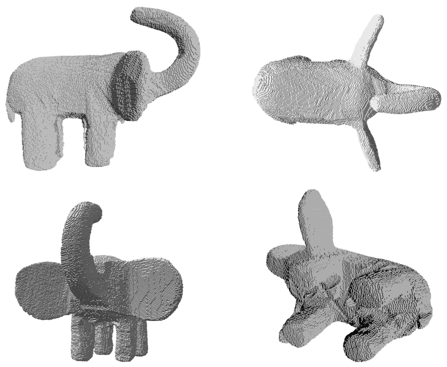

Figure 5. Different views of the merged volume model of an elephant sculpture.

6. Results

Fig.5 shows several views of the final volume model of our elephant. This model was obtained by intersecting the two initial models shown in Fig.3 after applying the transformation described by the resulting registration parameters. Computing these parameters takes about 10 minutes on a standard 400 MHz PII PC. In Fig.6 top and side views of two initial models of a boot and their intersection using the obtained registration parameters are depicted. Due to the limited extend of this article we have to refrain from including a more detailed comparism with other approaches.

7. Conclusions

The presented method for the registration of binary volume models was applied to merge different binary volume models of a rigid object captured in various orientation with a standard turntable setup using a shape from silhouette approach. During extensive tests, the method turned out to be very robust and the obtained registration parameters are sufficiently accurate for the target application. However, the required processing time needs to be reduced. Future work therefore aims at an improved, faster optimisation strategy. Moreover, we plan to extend the method to register deformable, articulated objects.

Figure 6. Rear and side views of two initial models and the resulting merged model of a boot.

References

[1] P.J. Besl and N.D. McKay, *A Method for Registration of 3D Shapes*, IEEE Trans. PAMI, Vol.14 No.2, (1992) 239–256.

[2] S.M. Yamany et al., *Novel Surface Registration Using the Grid Closest Point (GCP) Transform*, Proc. ICIP 1998, (1998) 809–813

[3] J.-P. Tarel and N. Boujemaa, *A Coarse to Fine 3D Registration Method Based on Robust Fuzzy Clustering*, Computer Vision and Image Understanding, Vol.73 No.1, (1999) 14–28.

[4] B. Krebs, P. Sieverding, and B. Korn, *A Fuzzy ICP Algorithm for 3D Free Form Object Recognition*, Proc. ICPR96, volume Track A, (1996) 539–543.

[5] K. Higuchi, M. Hebert, and K. Ikeuchi, *Building 3-D Models from Unregistered Range Images.*, Proc. IEEE Int. Conf. on Robotics and Automation, (San Diego, California, 1994) 2248–2253.

[6] S. Weik, *Registration of 3-D Partial Surface Models Using Luminance and Depth Information*, Proc. Int. Conf. on Recent Advances in 3-D Imaging and Modeling (Ottawa, 1997) 93–100.

[7] R. Szeliski, S. Lavallée, *Matching 3-D Anatomical Surfaces with Non-Rigid Deformations Using Octree-Splines*, Int. Journal of Comp. Vision, 18(2), (1996) 171–186.

[8] W. Niem, *Robust and Fast Modelling of 3D Natural Objects from Multiple Views*, SPIE Proc. "Image and Video Processing II", Vol. 2182, (San Jose 1994) 388–397.

[9] B.K.P. Horn, *Closed-form Solution of Absolute Orientation Using Unit Quaternions*, Journal of the Optical Society of America, Vol.4 No.4, (1987) 629–642.