Published in
Computer Vision - ACCV'98 (Asian Conference on Computer Vision), 490-497, 1998
which should be used for any reference to this work

Semi-Automatic 3D Object Digitizing System Using Range Images

C. Schütz, T. Jost and H. Hügli

Institute of Microtechnology, University of Neuchatel,
rue Breguet 2, CH-2000 Neuchatel, Switzerland

Abstract. Manual object digitizing is a tedious task and can be replaced by 3D scanners which provide an accurate and fast way to digitize solid objects. Since only one view of an object can be captured at once, several views have to be combined in order to obtain a description of the complete surface. In this paper a digitizing system is proposed which captures and triangulates views of a real world 3D object and semi-automatically registers and integrates them into a virtual model. This process is divided into three steps. First, an object is placed at different poses and its surfaces are sensed by a range scanner. Then, the different surfaces are aligned automatically starting from a pose estimate entered interactively. Finally, the overlapping triangle meshes of the registered surfaces are fused in order to obtain one unique mesh for the entire object.

1 Introduction

The increasing use of virtual object representations for various applications creates a need for fast and simple object digitizing systems. Therefore, 3D surface digitizers get used more and more since the model construction with a standard modeler is a quite tedious task especially for objects of arbitrary shape.

Range scanners give direct access to the 3D geometric information of object surfaces. They allow an accurate digitizing of an object surface at low cost and high speed. However, since most objects self occlude, one acquisition captures only a subpart of the entire object surface. Therefore, there exists a need to combine several range scanner views into one unique object representation.

The view combination is straightforward if the object is moved in a well known coordinate system like a rotation table: the relative transformation of two acquisitions is known. However, this implies a sophisticated mechanical system used to orientate the object or the scanner and to measure its pose. To avoid such complex pose systems, we propose to work with views from unknown object pose. The idea is to combine views based on the sole features of the geometric measurements.

We present a digitizing system which captures and triangulates views of a real world 3D object and finally registers and integrates them into a virtual model. The following steps have to be performed to combine the acquired surfaces.

An object to be digitized is placed in different poses on the acquisition field. The surface points measured by a range scanner are triangulated in 2.5D. The different measurements have to be transformed into a common frame. An interactive interface allows the operator to roughly align the acquired surfaces in 3D space. The precise surface registration is calculated with an automatic registration algorithm which matches the surfaces precisely by minimizing the distance between the common surface parts. In a last step called mesh fusion, the aligned triangle meshes are fused into an unique mesh for the entire object. Here, an erosion process eliminates the redundant surface parts and the remaining triangle meshes are joined with triangles by a gap filling algorithm.

A novel aspect of the presented system is the fact that the view registration and the integration modules are linked together and working completely in 3D space. Table 1 compares our work to systems proposed by other authors.

surface based registration	3D mesh fusion	2.5D mesh fusion
[TUR], [BES]	[RUT], [PIT], [HIL]	[TUR], [SOU]
presented system		

Table 1. Comparison of digitizing systems

The next section presents the global architecture of the 3D digitizing system. Its two main modules are detailed in the following sections and its effectiveness is shown in a further section devoted to experimental results on real objects.

2 System Architecture

The digitizing system consists of two blocks: view digitizing and view integration. The view digitizing block generates a virtual view of the observed object surface. The view integration block iteratively integrates each new virtual view in the virtual model under construction. This allows an incremental construction of the virtual model.

The view digitizing block measures the points of the visible object surface, filters the range data and triangulates the surface points. The resulting output is a triangle mesh representing the virtual view of the real object. Section 3 explains the detailed implementation of the modules used for the view digitizing.

The view integration block combines the virtual views and builds one virtual object becoming an entire model of the real object. The view integration block is composed of the view registration block and the mesh fusion module. The view registration block aligns the different views using interactive pose estimation and automatic registration. The mesh fusion module combines the individual triangle meshes into a new global mesh covering the union surface of the single meshes. Section 4 presents the different methods used for these modules. Fig. 1 gives an overview of the modules and the data flow during the digitizing process.

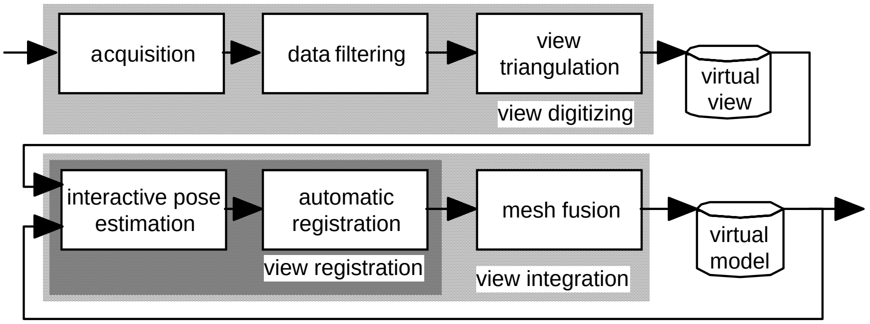

Fig. 1. Digitizing system architecture

3 View Digitizing

Range scanners give access to accurate and fast scanning of an object shape without any need of contact or visual marks. In our laboratory, the geometry of the object surface is acquired by a range scanner working on the principle of space coding with projected stripe pattern and triangulation. The coordinates measurements are arranged in a two dimensional array corresponding to the CCD camera image of the range scanner and can be visualized as a range image where the pixel intensity represents the camera to object distance as shown for a duck toy in Fig. 2. The virtual object can be easily textured since both images are represented in the same coordinate frame and the intensity and range information do correspond for every pixel.

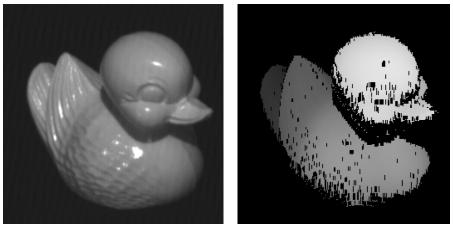

Fig. 2. Intensity and range images of a duck toy

The black pixels in the range image represent non-valid data (holes). They are found where the measurement confidence is low and are often caused by shadow regions or failed stripe coding. The filter module uses morphological operations such as erosion and closing to remove speckles and to fill small holes of non-valid data. Furthermore, the measurement noise present in the range data is filtered with a Gaussian filter.

Since the measurements are ordered in a regular grid namely the range image the triangulation of the surface becomes straightforward as proposed by [RUT]. Adjacent valid pixels are connected by triangles which results in a triangle mesh for the visible object surface. Checking the validity of the range points is not sufficient to avoid bad triangles. Other authors [TUR] [RUT] showed that additional checks are necessary to avoid the connection of range points separated by a discontinuity step in the range image. We have to ensure that occluded parts are not covered by triangles. Only triangles with small edges and an angle below 80° between the triangle normal vector and the sensor view axis are valid.

4 View Integration

Adding a new view to the virtual model under construction requires first the view to be registered to the model and then their meshes to be fused together.

4.1 View Registration

As stated in the introduction, the scanning process is kept as simple as possible and allows an operator to place the real world object in any stable pose on the acquisition field. Therefore, the transformation between the reference frames of a new acquisition and the virtual model is not known a priori and has to be determined in a first step. Since no external measurements of the object pose are available, we have to rely on the object surface characteristics in order to register it with the virtual model. This assumes that the virtual model and the new object view have at least some surface area in common which allows to establish correspondences between them.

View registration is performed in a two step process: interactive rough pose estimation and automatic matching.

Interactive Rough Pose Estimation.

Since human perception easily identifies corresponding surface parts for any object type and shape, we use an interactive graphic interface that permits an operator to enter a pose estimate for the two objects to be aligned. Both the virtual model and the new view are rendered in 3D and can be manipulated in all six degree of freedoms using a space mouse as input device. However, even a sophisticated object rendering and pose manipulation hardware is not sufficient to align the objects precisely. In fact, there is no measure apart from the visual feedback indicating the quality of the surface matching. Therefore, the interactive interface only provides a rough pose estimate to be used as a starting pose for the automatic precise registration. Fig. 3 shows an example of two roughly aligned surfaces used as starting configuration for the automatic registration.

Fig. 3. Roughly aligned views of a cat toy

Automatic Registration.

As stated before the precise alignment of two 3D surfaces is a tedious task if it has to be done manually. Besl proposed an automatic surface registration algorithm called ICP which avoids this problem [BES]. This algorithm registers two surfaces starting from an initial pose estimate. The algorithm proceeds iteratively. First, it pairs every point of one surface called P with the closest point of an other surface called X. These pairs of closest points are used to calculate the rigid transformation ($\mathbf{R}$, $\mathbf{t}$) which minimizes the mean square coupling distance or error. The surface P is then translated and rotated by the resulting transformation and the algorithm starts again with the closest point coupling.

This algorithm has been shown to converge fast but not necessarily towards the optimal solution. A good starting configuration for the two surfaces P and X is preliminary to a successful convergence. However, the range of successful starting configurations is quite large (see [HUG] and Fig. 3) which does not impose difficult constraints to the operator when entering a pose estimate for P and X.

In the original ICP algorithm the surface P is a subpart of X which is not the case in our application where both surfaces contain data not present in the other. The ICP algorithm needs therefore to be modified as proposed by Turk [TUR]. Closest points which are too far apart are not considered to be corresponding points and marked as invalid so they have no influence during the error minimization. The modified ICP algorithm is defined as follows:

- input: Two 3D surfaces P and X containing respectively N_p and N_x vertices.
- output: Transformation ($\mathbf{R}$, $\mathbf{t}$) which registers P and X
- iteration:

1. Build the set of closest point pairs (**p**, **x**):

$$\forall \mathbf{p} \in P \text{ find } \mathbf{x} \in X \text{ with } d_k = \min\left(\left\|\mathbf{p} - \mathbf{x}_j\right\|^2\right),\ j \in \left[1, \dots \mathrm{N_x}\right] \tag{1}$$

2. Weight every closest point pair (**p**, **x**) by applying the following distance threshold:

$$w_k = \begin{cases} 1 & d_k < (\mathrm{c} \cdot \mathrm{s} \cdot \mathrm{r})^2 \\ 0 & \text{else} \end{cases} \qquad \text{with } d_k = \left\|\mathbf{p}_k - \mathbf{x}_k\right\|^2 \text{ and } k \in [1, \dots, \mathrm{N_p}] \tag{2}$$

3. Find the rigid transformation (**R**, **t**) that minimizes the mean square error

$$e(\mathbf{R}, \mathbf{t}) = \frac{1}{\mathrm{W}} \sum_{\mathrm{N_p}} w_k \left\|\mathbf{R}\mathbf{p}_k + \mathbf{t} - \mathbf{x}_k\right\|^2 \quad \forall \text{pairs of } (\mathbf{p}_k, \mathbf{x}_k) \text{ and } \mathrm{W} = \sum_{\mathrm{N_p}} w_k \tag{3}$$

4. Apply the transformation (**R**, **t**) to P

The decision threshold for a valid coupling distance is set to the product $(\mathrm{c} \cdot \mathrm{s} \cdot \mathrm{r})^2$ where s equals the range scanner sampling distance and r equals the range image subsampling factor. The constant c allows to control the convergence of the automatic matching. It is set to a relatively large value at the beginning when the two surfaces are far apart. When the surfaces are superimposed the value of c can be lowered so only similar points are coupled which results in higher matching precision.

In order to verify the registration quality and to stop the iteration, a pertinent matching quality measure is needed. The minimization error e(**R**, **t**) corresponding to the mean μ of the square distances is a measure generally used to qualify the matching. Another statistical measure which has been used successfully to qualify matched surfaces in object recognition [SCH] is the deviation σ of the square distances indicating the matching regularity. The matching is stopped if the sum of μ and σ does not change any more.

$$\mu = \frac{1}{\mathrm{N_p}} \sum_{\mathrm{N_p}} d_k \qquad \sigma = \sqrt{\frac{1}{\mathrm{N_p} - 1} \sum_{\mathrm{N_p}} \left(d_k - \mu\right)^2} \tag{4}$$

In order to detect cases where only very few points are coupled, matchings with a high number of coupled points on the surface P are selected, as proposed by Krebs [KRE] and expressed by a high value of the coupling measure ε.

$$\varepsilon = \frac{\mathrm{W}}{\mathrm{N_p}} 100 \quad \text{with } \mathrm{W} = \sum_{\mathrm{N_p}} w_k \tag{5}$$

As mentioned before, the two surfaces should have enough common data points for successful matching. 30 to 50 % of common surface has been observed to be a good amount.

4.2 Mesh Fusion

There exist several methods to integrate registered surfaces acquired from different views [TUR] [RUT] [PIT] [SOU] [HIL]. They differ mainly in how they treat the redundant overlapping zone of the two registered surfaces and can be separated into two groups: partial erosion and complete retriangulation of the surface points.

Methods using a partial erosion approach [PIT] [TUR] erode the overlapping surfaces until the overlap disappears. The two triangle meshes are then linked at their frontiers in order to have one unique mesh for the union of the two surfaces. Other authors [HIL] [SOU] [RUT] discard the mesh information from the triangulated views if calculated at all and retriangulate the overlapping zone or even the complete point set.

Since the object views can be easily triangulated using the range image structure, we opt for the partial erosion approach which keeps intact as much as possible of the triangle mesh structure. We propose a new mesh fusion algorithm that benefits from the closest point relationships established during the geometric matching. There is no need to run an extra routine to erode overlapping surfaces and to detect the surface frontiers as done in other work [PIT] [TUR].

The following features characterize the proposed mesh fusion algorithm. We refer to the same surfaces P and X as introduced for the geometric matching, where the vertices on P are coupled with points on the triangles of X.

1) **overlap detection:** The registration algorithm calculates for the vertices on surface P the closest points on the triangles of the surface X. Closest points with an Euclidean distance below a defined threshold are coupled. During the geometric matching iterations, the coupled points on P converge to the overlapping area of the two surfaces P and X.
2) **overlap remove:** The redundant part of the surface P is deleted by removing the triangles with one or more coupled vertex. The remaining meshes are separated by a gap defined by a frontier on P and X.
3) **frontier detection:** Triangles where the geometric matching coupled only one vertex are connected to the frontier on P. Actually, the two non-coupled vertices build an edge of the frontier on P. The frontier on X is detected by a closest point search.
4) **gap filling:** The gap enclosed by the two frontiers is filled iteratively with triangles. Vertices on the two frontiers are used as candidates to build a filling triangle. The triangulation does not need projection into tangential planes which allows a correct triangulation of sharp edges. Triangles with a maximal opening angle are constructed in order to have an optimal approximation.

The implementation details are discussed below and illustrated by examples shown in the figure Fig. 4.

Overlap detection

The closest point routine of the automatic registration module marks the vertices on surface P which overlap surface X. This results in a set $\mathrm{C_V}$ of coupled vertices with $\mathrm{C_V} = \left\{ \mathbf{p}_k \in \mathrm{P} \middle| w_k = 1 \right\}$ and therefore $\mathrm{C_V} \subseteq \mathrm{P}$.

Overlap remove

The remove process eliminates all the vertices member of $\mathrm{C_V}$ from the surface P. The clipped surface $\mathrm{P_C} = \left\{ \mathbf{p}_k \in \mathrm{P} \middle| \mathbf{p}_k \notin \mathrm{C_V} \right\}$ is separated from the surface X by a gap since $\mathrm{C_V}$ is slightly larger than the actual overlap area due to a coupling threshold which is not zero.

Frontier detection

A frontier is defined by the frontier list F which contains the ordered vertices of the surface $\mathrm{P_C}$ which limit the gap created by the above overlap remove process. The list F is build as follows: During the automatic matching the list

$$\mathrm{T_F} = \left\{ \mathbf{t}_l = \left\{ \mathbf{v}_{l,0}, \mathbf{v}_{l,1}, \mathbf{v}_{l,2} \right\} \middle| \mathbf{v}_{l,i} \in \mathrm{P} \text{ and } \sum_{i=0}^{2} w_{l,i} = 1 \right\}$$

containing the triangles with only one coupled vertex is established. For every triangle in $\mathrm{T_F}$, the vertices which are not coupled are inserted to

$$\mathrm{F} = \left\{ \mathbf{v}_{l,i} \in P \middle| \mathbf{v}_{l,i} \in \mathbf{t}_l \text{ with } \mathbf{t}_l \in \mathrm{T_F} \text{ and } w_{l,i} = 0 \right\}.$$

Such a frontier list is established for every frontier on P_C.

Gap filling

The gap between the two surfaces X and P_C is filled with triangles in order to join the two meshes. The different frontiers on P_C delimiting these gaps are processed sequentially. The filling process is initialized for a frontier on P_C with the search of the first vertex $\mathbf{x}_N$ on the frontier of X. To do so, the first two vertices $\mathbf{f}_0$ and $\mathbf{f}_1$ of the frontier list F are selected and the nearest point $\mathbf{x}_N$ to $\mathbf{f}_0$ and $\mathbf{f}_1$ on the frontier of X is calculated. Then, the first bridge triangle joining the two frontiers is constructed with the vertices $\mathbf{x}_N$, $\mathbf{f}_0$ and $\mathbf{f}_1$ as shown in Fig. 4. The frontier list F is updated by setting its first vertex $\mathbf{f}_0$ equal to $\mathbf{x}_N$.

The following algorithm fills the gap iteratively starting with the above initialization. Two candidate vertices are selected to build the next bridge triangle. One is $\mathbf{f}_2$, the third vertex in the frontier list F and the other one is $\mathbf{x}_C$, the next vertex on the frontier of X. These two candidates form together with the vertices $\mathbf{f}_0$ and $\mathbf{f}_1$ of F the next potential bride triangles as shown in Fig. 4. The candidate which encloses the maximal angle is selected in order to obtain a regular triangulation. The frontier list F is updated with the new vertices as follows: $\mathbf{f}_0$ is set equal to $\mathbf{x}_C$ if $\mathbf{x}_C$ is chosen or $\mathbf{f}_1$ is removed from F if the candidate $\mathbf{f}_2$ is selected. The candidate selection starts again with the modified frontier list and the above procedure is applied until F contains only two vertices.

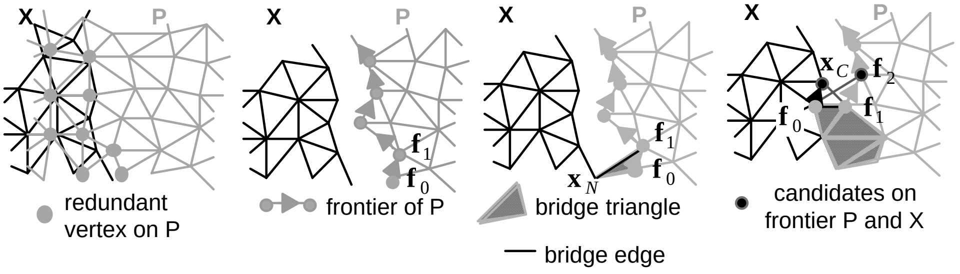

Fig. 4. Mesh fusion showed for an example

5 Results

The described 3D object digitizing system has been implemented and used to create virtual models from several real world objects. Results are presented here for two toy objects. Six views have been used to create the cat toy model whereas ten views have been merged to obtain the rabbit model. The cat object consists of 9700 points and about 19000 triangles whereas the rabbit object contains 15000 points and 30000 triangles. The successful reconstructed objects are shown in Fig. 5.

Fig. 5. Digitized cat and rabbit toy

The range finder assigns an object color to every vertex. If the triangle mesh is fine enough the object texture is maintained, as shown for a rabbit toy.

As discussed above, the object views to be assembled need common surface parts with enough geometric structure in order to allow the automatic matching to converge to a stable solution. For example for a box, a view should contain at least three faces in order to find a stable alignment of two views. The proposed digitizing system is especially suited for objects of complex free-form shape.

6 Conclusions

The presented digitizing system permits to construct models from free-form 3D objects. It integrates object views acquired by a range scanner. In order to provide simplicity and full flexibility during the acquisition, the system does not need any information about the object poses. The various views are integrated by view registration which combines an interactive rough view registration step followed by an automatic precise matching. A new mesh fusion algorithm combines the meshes into a global one.

The particular interest of the system consists in the adequate combination and linking of modules of the digitizing process. For example, the overlap information obtained in the matching module is directly used to remove redundant mesh area.

7 Acknowledgments

This research has been funded by the Swiss national science foundation under project number 2100-43530.

References

[BES] P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 14, no. 2, pp. 239-256, 1992.

[HIL] A. Hilton, A.J. Stoddart, J. Illingworth and T. Windeatt, "Marching triangles: range image fusion for complex object modelling," IEEE International Conference on Image Processing, Lausanne, vol. 2, pp. 381-384, 1996.

[HUG] H. Hugli, Ch. Schutz, "How well performs free-form 3d object recognition from range images?," Intelligent Robots and Computer Vision XV, Algorithms, Techniques, Active Vision and Materials Handling, SPIE, Boston, Vol. 2904, 1996, pp. 66-74.

[KRE] B. Krebs, P. Sieverding and B. Korn, "Correct 3D Matching via a Fuzzy ICP Algorithm for Abitrary Shaped Objects," DAGM96 - Mustererkennung, Heidelberg, 1996.

[PIT] R. Pito, "Mesh Integration Based on Co-Measurements," IEEE International Conference on Image Processing, Lausanne, vol. 2, 1996, pp. 397-400.

[RUT] M. Rutishauser, M. Stricker, M. Trobina, "Merging Range Images of Arbitrarily Shaped Objects," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVRP), Seattle, 1994, pp. 573-580.

[SCH] Ch. Schutz, H. Hugli, "Towards the recognition of 3D free-form objects," Intelligent Robots and Computer Vision XIV, Algorithms, Techniques, Active Vision and Materials Handling, SPIE, Philadelphia, Vol. 2588, pp. 476-484, 1995.

[SOU] M. Soucy, D. Laurendeau, "A General Surface Approach to the Integration of a Set of Range Views," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 4, 1995, pp. 344-358.

[TUR] G. Turk, M. Levoy, "Zippered Polygon Meshes from Range Images," Proceedings ACM Siggraph, Orlando, 1994, pp. 311-318.