A 3D ACQUISITION AND MODELLING SYSTEM

S. Mahadevan, H. Pandzo, M. Bennamoun , J.A. Williams
Space Centre for Satellite Navigation
Queensland University of Technology, Brisbane, Australia
Email: m.bennamoun@qut.edu.au

ABSTRACT

This paper presents our implementation of an integrated 3D acquisition and modelling system. The proposed system consists of an acquisition module, a registration algorithm proposed by the last two authors and a reconstruction module. Techniques for addressing the implementation of the modules are first briefly followed by a more detailed discussion of the techniques implemented in the system is given. Results are also presented to demonstrate the system's operation.

1 INTRODUCTION

The applications of 3D vision systems are increasing rapidly. One area in particular that is receiving more interest is 3D object modeling. In this paper we present the implementation of a system that can generate 3D surface models of objects or parts thereof.

Applications of such a system are numerous. One application is in reverse engineering of manufacturing parts. A typical scenario is if there are no construction plans for a certain part. In this case, a 3D model of the object can be generated and exported to a CAD program where it can be used to make the construction plans for future use. Another application is the generation of 3D human models. These models can be used in a wide variety of ways such as for ordering custom-made clothing or playing 3D games. The systems can also be used for general 3D content creation and animation. By utilising this technology, fast 3D models can be generated and then animated by graphics produces for use in movies.

Our proposed system is composed of 3 main components, namely 1) view acquisition, 2) view registration and 3) the surface reconstruction modules. The basic operation of the system is as follows. Firstly, the entire surface of the object must be captured. This requires multiple views from different viewpoints to be acquired as the acquisition system can only generate data along the camera line-of-sight. The resulting view data is relative to the sensor-centred coordinate system, which is different for each view. The views must therefore be transformed into an object-centred coordinate system before the surface can be extracted. This is achieved by view registration. Surface reconstruction is then performed to merge the views into a single surface and to extract the surface. Figure 1 shows the high level acquisition and modeling process.

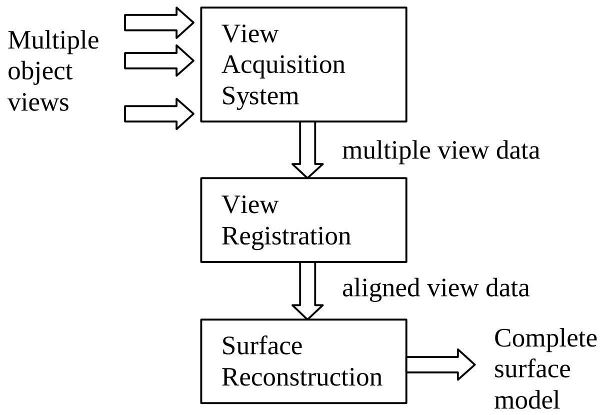

Figure 1. Block diagram of the Acquisition and Modelling System

The acquisition, registration and reconstruction algorithms have been integrated into a stand-alone Windows based program that runs on the host PC. The PC also provides the interface to the CCD camera, through a frame grabber, and the structured light projector required for acquisition. The camera and projector have been mounted on a fully adjustable stand (Figure 2).

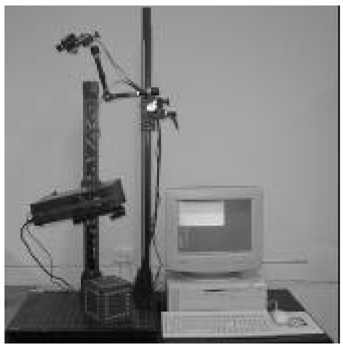

Figure 2. The System Hardware

The significance of this work is the integration of the acquisition with our previously proposed registration [7,8] and a reconstruction algorithm into a stand-alone system. Additionally, the system has been developed from off-the-shelf hardware components and is thus a cost-effective approach to 3D data acquisition and modelling.

The structure of the paper is as follows. Section 2 presents a brief review of 3D acquisition, registration and reconstruction techniques. Sections 3, 4 and 5 then describe the acquisition, registration and reconstruction components of this system in more detail. Results from the integrated system are presented in section 6 and finally conclusions are given in section 7.

2 LITERATURE REVIEW

The following sections discuss the general techniques for the modules of the proposed modeling system. The aim is to introduce the reader to the advantages and disadvantages of the approaches and as such in depth details of these methods is not provided.

2.1 3D Acquisition Systems

3D acquisition techniques can be categorised into two main classes - active and passive. In active systems, an external source projects energy onto the scene of interest in order to measure surface data. Structured light and laser finders are two examples of active systems. The advantage of active systems is the dense 3D data they generate and the ease of implementation. The main disadvantage stems from the use of the external energy source. This makes active systems sensitive to the ambient light conditions.

In passive systems, no external energy is imposed onto the scene of interest. The most common passive technique is stereo. Stereo has the advantage over active methods in that it is less sensitive to ambient light levels and is lightweight (no energy source required). This makes stereo systems more suitable for mobile vision platforms. The major drawback with stereo is the correspondence problem. The process of matching between stereo pairs is complex, as it must consider problems such as repetitive patterns, occlusions and bland regions to name just a few [1].

Given this, we decided to implement an acquisition system based on a structured light approach. The main reason for this choice was the dense data produced which will allow more accurate models to be reconstructed. The issue of ambient light is not significant, as the system will be operated in a controlled environment.

2.2 View Registration

The approaches to view registration can also be categorised into 2 classes, namely 1) pair-wise and 2) global.

In pair-wise techniques, the views are registered two at a time, with the resulting view being subjected to a further registration with another view, until all views are registered. The advantage of pair-wise approaches is their simplicity. However, the pair-wise registrations performed attempt to minimise the errors for the pairs and not the global mesh. A consequence of this is that the global mesh constructed may be far different to the true surface. In addition to this, error estimates made are on a pair basis and have no real significance to the complete surface mesh.

In contrast to this, global techniques register all related view data simultaneously to achieve an optimal surface mesh. This is at the expense of algorithm complexity. A disadvantage of global techniques until recently was the inability to determine a global error estimate. This has now been addressed by Williams et al. [7,8].

We decided to use our global registration technique briefly described in [8] and in more detail in [7]. The use of this global technique ensures that an optimal overall surface mesh will be generated. The error estimation capability of this method also allows for the acquisition system error to be estimated and propagated into the registration module. Estimation of the acquisition error has not been addressed yet.

2.3 Surface Reconstruction

Techniques for surface reconstruction fall into two main categories – volumetric and geometric.

In volumetric approaches each surface view is represented by a 3D function which is then merged to form the complete object. The surface can then be extracted. The advantage of volumetric methods is their simplicity and the guarantee that a watertight surface will result.

In geometric approaches, the original point and mesh geometry is used to merge the views and extract the surface. In contrast to volumetric approaches, geometric approaches produce an interpolation not an approximation. This results in a more accurate surface but also more complex algorithms.

A volumetric reconstruction technique was chosen because of the advantages mentioned above. The accuracy of volumetric techniques is sufficient not to warrant the use of a geometric method.

3 3D VIEW ACQUISITION

The basic idea behind the structured light approach is as follows. A series a binary light stripes/planes are projected onto the object of interest and images of each projection are taken. The images are then combined using a simple weighting scheme to produce an image defining a number of unique planes in the scene. This image is commonly called a stripe image. The pixels in the stripe image are then interpreted as rays and each pixel has a corresponding stripe id that defines the plane the pixel lies in. Thus if we know the parameters of the particular ray and plane, we can determine the 3D data point at which they intersect. Commonly, a series of 8 stripe patterns are projected onto the scene thus defining 256 (2^8) distinct planes.

Our acquisition system has 2 modes of operation, these being calibration and view reconstruction. In the calibration phase the aim is to find the transformation between the world coordinate system and the camera and projector coordinate systems respectively (ie.

parameterise the rays and planes). To do this, a calibration cube, which has a number of fiducial marks on it, is used as the world coordinate reference. The stripe image of the cube is constructed and the image locations of the centroids of the fiducial marks are found. Thus, for each fiducial mark we know its image location, stripe/plane id and the corresponding world coordinate. This information is then used to estimate the necessary transformations.

The transformations are defined by a 3x4 projective transform matrix (PTM's) mapping between the camera and world coordinates and a 2x4 PTM mapping between stripe projector and world coordinates. A full derivation of the PTM's is given in [4].

In view reconstruction we aim to reconstruct the 3D surface of an object along the line of sight of the camera. This begins by constructing the stripe image of the object. Background segmentation is then performed to ensure that only valid 3D data points belonging to the object's surface are generated. This is an important phase, as spurious data obtained from the background will result in failure of the registration component. Once the valid stripe image pixels have been determined, they are transformed into the corresponding 3D data points using triangulation. The transformation applied is determined by the PTM's found in calibration, the pixel location and the stripe id at the pixel. The derivation of the transformation is given in [4].

A problem that affects the accuracy of structured light systems is the finite thickness of the projected light stripes. This results in bands of pixels having the same stripe id (lying in the same plane), which is not ideal. Valkenburg et al. proposed a sub-stripe technique to overcome this effect and we have implemented it in our system.

4 VIEW REGISTRATION

The views acquired are initially expressed in a sensor-centred coordinate system and need to be registered into an object-centred system before the surface can be extracted.

The registration process is as follows. Firstly, a set of view pairs is formed from adjacent views that contain the overlapping information. The initial correspondence information for each view pair is then entered by interactively selecting 3 points in each view. This correspondence information is used to provide an initial estimate of the best alignment of views. The model of the overall object thus consists of the several pairs of views, each of which has an approximate initial correspondence information. The model is then subjected to a series of iterative registrations to determine the best fit. This uses a variant of Besl and McKay's Iterative Closest Point (ICP) algorithm [2].

The original ICP algorithm has a few shortcomings. Firstly, it was designed to register a data set *Y* to a model *X* (ie. *Y* is a subset of *X*). This causes a problem in view registration because only a subset of points in each view corresponds to each other but ICP will incorrectly try to generate correspondences for every point in the two data sets. This problem is solved by applying several heuristic constraints including 1) maximum edge distance 2) normal vector compatibility 3) boundary match rejection and 4) top *p* percent [5].

The second shortcoming of the original ICP algorithm is that it is pair-wise. The last two authors already solved this problem by applying the algorithm simultaneously to all views representing the model [7,8].

Our modified ICP algorithm can be stated briefly as follows:

```
- Assume all views are approximately in
  their correct positions

- For each pair in model
    For each point in view 1
    (repeat for view 2)
       - Match with closest point in
         other view.
       - Apply constraints to limit the
         point pair set.
    End
  End

- Register using our algorithm in [7,8]
  to compute rotations and translations
  which minimise the distance between
  all point pairs matched before.
```

The resulting rotations and translations, when applied to each view, aligns them as they are in real life.

5 SURFACE RECONSTRUCTION

Once the views have been aligned in a common coordinate system they must be merged and a single surface extracted to represent the object. The method used here is a volumetric integration technique proposed by Hilton et al. [3].

In this technique, an implicit 3D function, $f_i(\mathbf{x})$, is defined for each input mesh, *i*. The functions represent the signed Euclidean distance from any point $\mathbf{x}$ to the surface, with negative values indicating the point is inside the surface, positive values indicating the point is outside the surface and zero indicating that the point is on the surface. A binary function $b_i(\mathbf{x})$ is also defined for each view, *i*, which returns the value of 1 if the nearest point on the mesh to $\mathbf{x}$ is on the boundary of the mesh, or zero otherwise.

These functions are then merged into a global distance function $f(\mathbf{x})$ using a simple weighting scheme. Finally the Marching Cubes algorithm is used to extract the zero set of the global function, which approximates the object's surface.

6 RESULTS

The system has been extensively tested. In this section we present results from two different tests on the acquisition and modelling system. The first test verifies the operation of the system using a plastic moulded shoe. The second test shows the effectiveness of the sub-stripe technique mentioned in section 3.

Figure 3 shows four different views of the plastic shoe captured by the acquisition system.

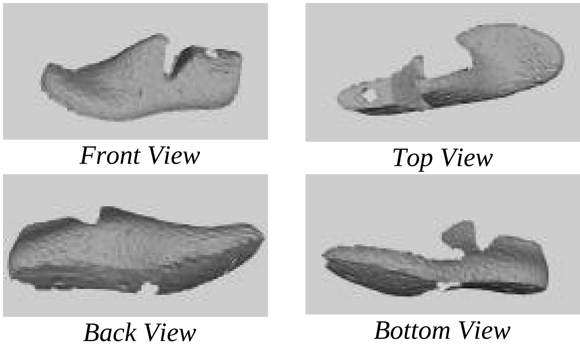

Figure 3. Four acquired views of a shoe

The result following the registration of these views is shown in Figure 4. The different shades indicate data from different views.

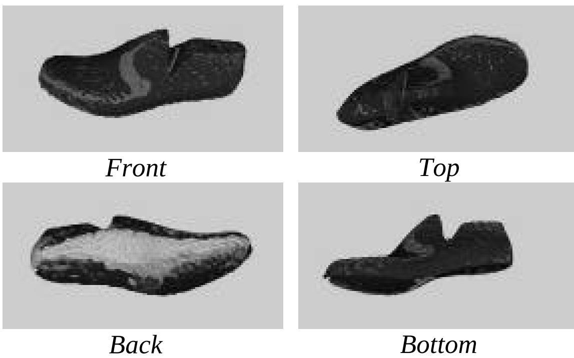

Figure 4. Registered views from different viewpoints.

The final surface mesh extracted by surface reconstruction is shown in Figure 5.

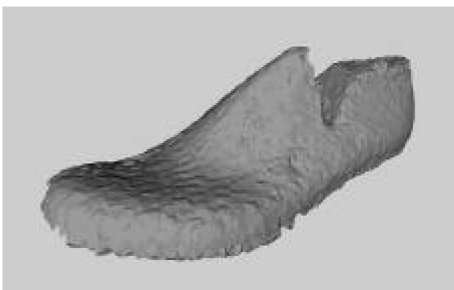

Figure 5. Reconstructed model

To test the effectiveness of the sub-stripe technique, a single view from a plane was acquired, with and without the application of sub-stripe estimation. A plane was then fitted to the two data sets and the RMS error was calculated. In the case where sub-stripe estimation was applied the RMS error was 0.68mm. This compares favourably against the RMS error without sub-stripe estimation, which was 1.2mm.

7 CONCLUSIONS

We have presented our implementation of a 3D acquisition and modelling system. Results from the reconstruction of an object have been included to verify its operation. The application of a sub-stripe technique in the view acquisition has also been tested and shown to have a significant effect on the results.

8 REFERENCES

[1] J. Banks, M. Bennamoun, and P. Corke, "Fast and Robust Stereo Matching Algorithms for Mining Automation", Digital Signal Processing (DSP), 9(3):137-148, July 1999.

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

[3] A. Hilton, A.J. Stoddart, J.Illingworth, T. Windeatt, "Reliable surface reconstruction from multiple range images", Proceedings 4th European Conference on Computer Vision, 117-126, 1996.

[4] A.M. M^{c}Ivor, R.J. Valkenburg, "Calibrating a Structured Light System", Industrial Research Limited Report 362, Auckland, New Zealand, 1995.

[5] K. Pulli, "Multiview Registration for Large Data Sets", Proceedings 2nd International Conference on 3D Digital Imaging and Modeling (3DIM'99), 160-168, 1999.

[6] R.J. Valkenburg, A.M. M^{c}Ivor, "Accurate 3D measurement using a Structured Light System", Industrial Research Limited Report 576, Auckland, New Zealand, 1996.

[7] J.A. Williams, M. Bennamoun, "Simultaneous Registration of Multiple Corresponding Point Sets", Computer Vision and Image Understanding (to appear in January 2001).

[8] J. Williams, and M. Bennamoun, "Simultaneous Registration of Multiple Point Sets Using Orthogonal Matrices", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'2000), Istambul, Turkey, June 5-9, 2000, pp.2049-2052.