FUSION OF INTENSITY AND RANGE DATA FOR IMPROVED 3D MODELS

Paulo Dias[(1)], Vítor Sequeira[(1)], João G.M. Gonçalves[(1)], Fransisco Vaz[(2)]
[(1)] European Commission, Joint Research Centre, Italy
[(2)] University of Aveiro/IEETA, Portugal
{paulo.dias, vitor.sequeira, joao.goncalves}@jrc.it, fvaz@ieeta.pt

ABSTRACT

In this paper, we present techniques combining intensity and range data to achieve more realistic three-dimensional models from real world scenes. The principal novelty is the use of the intensity information not only for texturing the final 3D model but also to improve the segmentation of the range data. The technique registers video and range, i.e., the camera parameters are computed in order to get both data in the same reference frame. Edge extraction techniques are then applied to select corresponding edges and a line-straightening algorithm is used to correct the 3D model. Results are presented for both indoor and outdoor environments.

1. INTRODUCTION

Computing Three Dimensional models from data acquired from real scenes is a challenging task due to the details and discontinuities inherent to real world. A commonly approach to reach this goal is to use laser range sensors combined with texture colour images [1,2,3,4]. But some problems remain since the laser beam used in these scanning devices is not an ideal point but more like a disc. This effect appears clearly in discontinuities and is known as "the mixed point problem"(Figure 1a). In this particular case, the measured distance does not reflect the real distance, but is rather a combination of the distances to both parts of the footprint. As a result, straight lines may appear like a sawtooth. The techniques presented use edge information from digital photographs (that have a spatial resolution far better than range images) to straighten the edges in the model.

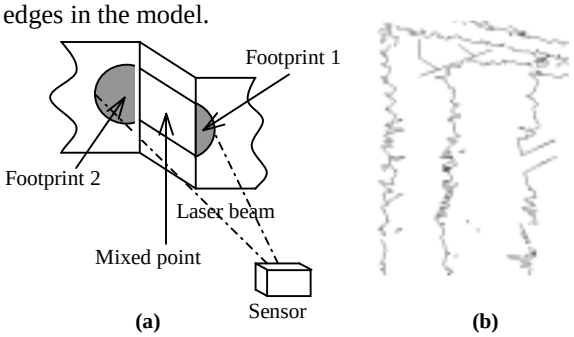


Figure 1 - The mixed point problem (a) and its effect (b).

2. CALIBRATION

The first step of our procedure is the registration of the video images in order to get both the range and intensity information in the same coordinate reference. The calibration is done using a semi-automatic procedure in three steps [5]. First, video and reflectance (an infrared image acquired by most of the current laser scanners) images are resized based on edge detection and then matching techniques are used to seek for correspondences. Finally, a Tsai camera calibration technique [6] is used to find out the 11 parameters defining the intrinsic and extrinsic parameters of the camera.

Figure 2 presents the different steps and results of the calibration for a laboratory image. The final image can be directly used as a texture map for the model and allow the evaluation of the quality of the registration.

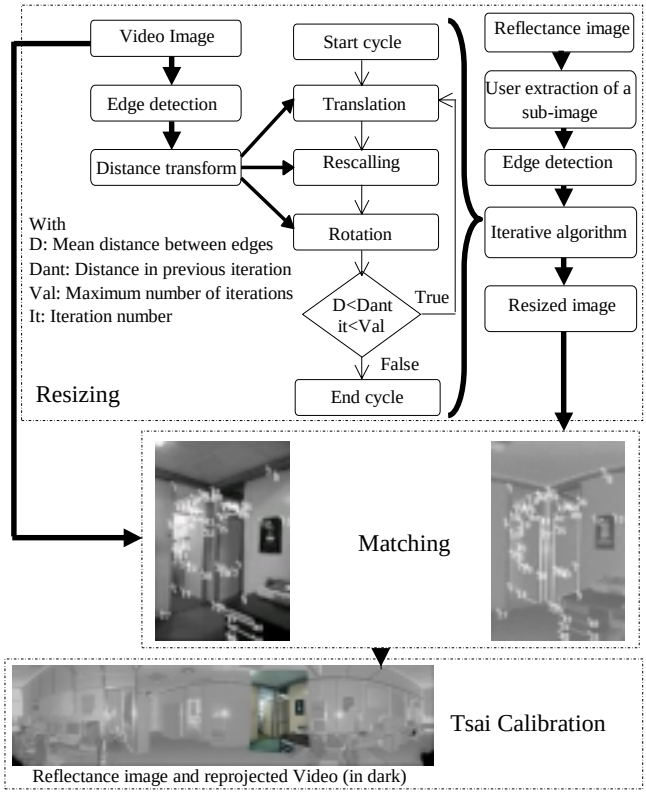


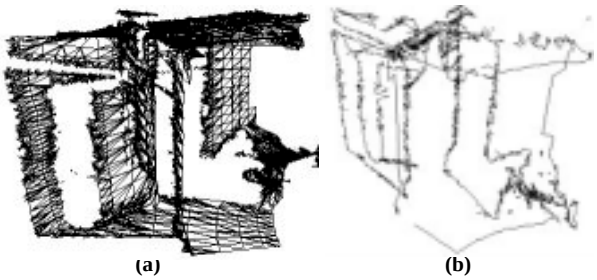
Figure 2: The calibration algorithm.

3. RANGE AND INTENSITY EDGE DETECTION

Following the video and range registration, the algorithm detects edges both in the 3D model and the video image. The 3D geometry is used to select edges and discontinuities in the model. The lines in the intensity image are extracted using a standard feature detection technique (e.g. Sobel operator). The final corrections to apply to the model are computed from the comparison of the 2D (intensity) and 3D (model) edges based on the camera calibration.

3.1. 3D model

The edges are computed from the 3D triangulated mesh. All triangle segments belonging to a single triangle are considered as edges and add to a list. A tracking algorithm orders the vertexes inside each one of the edges creating a list of segments. Figure 3b shows a snapshot of the detected edges for a laboratory scene.



(a) (b)

Figure 3: the triangulated model (a) and the results of the edge detection (b) in a laboratory Scene.

3.2. Intensity image

Line detection is computed in the registered video images to extract the straight edges. The points with a response higher than a given threshold to the Sobel operator are thinned (through morphological operations). The points are then split into straight segments according to the following conditions: minimum number of points per segment and the distance from the points to the segments.

To avoid redundant lines, a filter based on the orientation and distance between segments is applied. Two user defined thresholds set the minimum distance and orientation between segments (see Figure 4a). A segment S is valid only if for all segment i:

Dist (i,S) > th1

OR Dist (i,S) < th1 and Angle (i,S) > th2

OR Dist (i,S) <th1 and Angle(i,S) <th2 and N(S)>N(i)

where Dist(i,S) is the orthogonal distance between segments i and S, Angle(i,S) is the angle between segment i and S, and N(i) is the number of points in segment i.

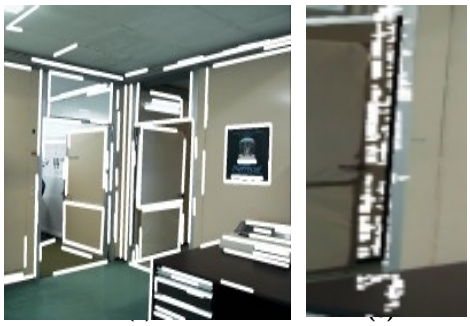


Figure 4: (a) Segment detection results. (b) Original video segment (black) and the reprojected 3D points (white).

4. 3D EDGE CORRECTION

4.1. Selection of a first 3D line approximation

For each detected segment in the video image, the algorithm selects the 3D edge points that are projected in the segment's neighbourhood using the registration computed in section 2. Figure 4b, presents a segment in the video (black) and the reprojection of the selected 3D points (white). A RANSAC technique [7] is applied to pairs of points to determine the best 3D line **L**.

4.2. Selection of the edges to correct

In the second step, the algorithm follows the range discontinuities (according to the ordered list defined in section 3.1) to select the 3D edges that have high probability to correspond to the 2D segment.

Two main conditions have to be fulfilled by an edge to be selected:

- All the points of the edge have to be spatially close to the line **L** in three dimensions. This is performed in 3D using the orthogonal distance; given A and B, the orthogonal distance **D** of a point C to line AB is defined as:

$$\mathbf{D} = \frac{\sqrt{AC^2 AB^2 - (AC \bullet AB)^2}}{AB} \qquad (1)$$

- The linear regression of the 3D edge points reprojected in the video image must have an orientation similar to the original segment. Given θ_s the orientation of initial segment and θ_l the orientation of the median line define by the reprojected 3D points, we have:

$$\theta = \text{abs}\,(\theta_l - \theta_s) < \text{threshold} \qquad (2)$$

4.3. Correction of the 3D coordinates

4.3.1. Refining of the 3D line parameters

At this point, the algorithm has already selected all the points to be corrected according to the 3D line **L**. It is then possible to refine the 3D line parameters by computing a linear fitting based on orthogonal regression [8]. This allows finding the 3D line **L'** that best fits all the points. The two vectors (Origin (**O**) and Direction (**Dir**)) that define the optimal line between all points to correct $\mathbf{X}_i(x_i, y_i, z_i)$ are:

$$\mathbf{O} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{X}_i ; \qquad (3)$$

given $\mathbf{O}(a,b,c)$, the direction vector (**Dir**) is defined as a unit length eigenvector corresponding to the smallest eigenvalue of the matrix:

$$\delta\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} - \begin{bmatrix}\sum_{i=1}^{m}(x_i-a)^2 & \sum_{i=1}^{m}(x_i-a)(y_i-b) & \sum_{i=1}^{m}(x_i-a)(z_i-c)\\ \sum_{i=1}^{m}(x_i-a)(y_i-b) & \sum_{i=1}^{m}(y_i-b)^2 & \sum_{i=1}^{m}(y_i-b)(z_i-c)\\ \sum_{i=1}^{m}(x_i-a)(z_i-c) & \sum_{i=1}^{m}(y_i-b)(z_i-c) & \sum_{i=1}^{m}(z_i-c)^2\end{bmatrix}$$

$$\text{with } \delta = \sum_{i=1}^{m}(x_i-a)^2 + (y_i-b)^2 + (z_i-c)^2 \qquad (4)$$

4.3.2.Correction of single point coordinates

After the optimal Line **L'** is found, the three dimensional coordinates of the orthogonal projection of each point to correct are computed using the following formula: given the original point $X(x,y,z)$, the projected point on the line $X'(x',y',z')$, $O(O_x,O_y,O_z)$ and $D(D_x,D_y,D_z)$ the origin and direction vectors of **L'**, we have:

$$\begin{aligned} X' &= O_x + D_x * t \\ Y' &= O_y + D_y * t \\ Z' &= O_z + D_z * t \end{aligned} \qquad (5)$$

with

$$t = \frac{XD_x + YD_y + ZD_z - O_xD_x - O_yD_y - O_zD_z}{D_x^2 + D_y^2 + D_z^2} \qquad (6)$$

4.4. Iteration with adaptive thresholds

To make the algorithm more effective and less dependent on the user defined thresholds (depending on the images, the results after one iteration can be fairly good or not) the procedure is ran iteratively, each time with new thresholds. The algorithm stops when no correction is computed or when a user-defined number of iteration is reached. The modified thresholds from iteration to iteration are:

- The distance **D** between the points and the 3D line **L**. This allows a larger search space around the line in successive iterations.
- The difference θ in the orientation between the segment in the video image and the reprojected edge.
- The minimum number of points in an edge. This ensures that only relatively long edges are corrected as the thresholds increase.

5. EXPERIMENTAL RESULTS

The main test scene is composed of a room with two doors and several windows. It was chosen because it is rich in discontinuities making it an ideal scene for testing the improvement our technique can bring. In effect, the acquisition of reliable range data is particularly difficult in such a scene since range measurements are prone to errors in discontinuities. At the same time, the scene is rich in geometric contents that can be recovered from the video.

Figure 5, presents a snapshot of the original edges (a) and the results of the correction (b) after three iterations of the algorithm. The whole procedure takes approximately one minute in a PII, 400Mhz processor.

Figure 6, shows in the same snapshots the original and the corrected edges from two different viewpoints. These snapshots show clearly how the mixed point effects are minimised.

(a) **(b)**

Figure 5: Original (a) and Corrected (b) edges.

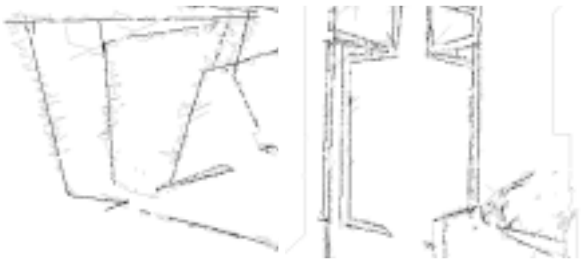


Figure 6: Snapshots of the original (light grey) and the corrected edges (black).

Figure 7 presents a snapshot of the indoor model before any correction. The corrected model, after three iterations of the algorithm, appears in Figure 8.

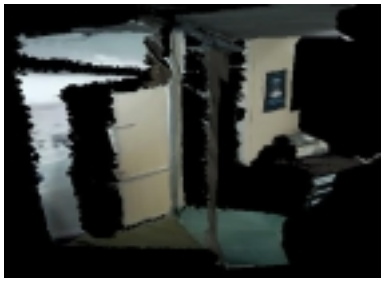


Figure 7: Snapshot of the original indoor model.

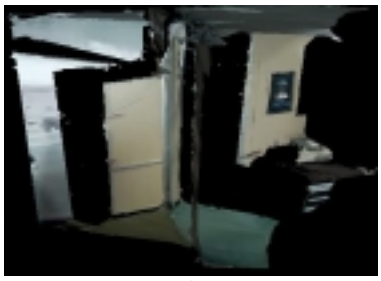


Figure 8: Snapshot of the corrected model.

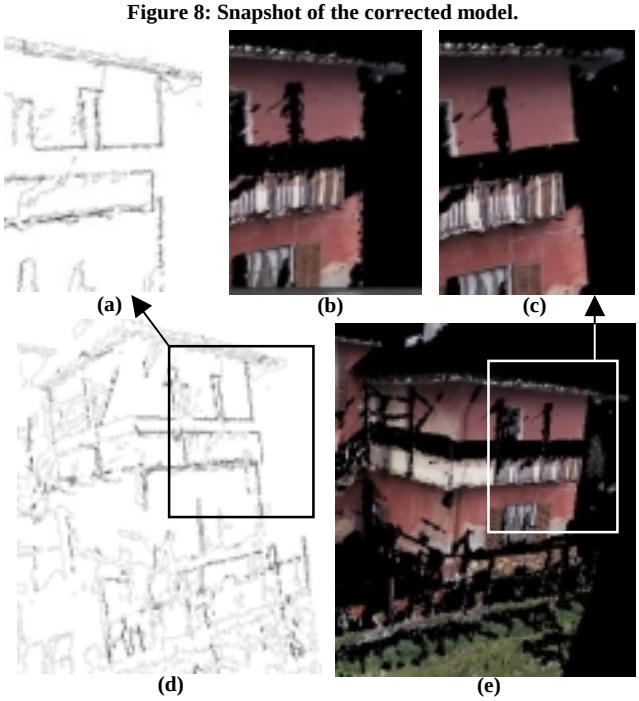


Figure 9: Results of the algorithm with the farmhouse scene

The algorithm (with the same parameters and number of iterations) was run in a model of an outdoor scene giving similar results. Figure 9**,** presents these results: (a)- Details of the edges (light grey: original edges, black: corrections); (b)-Details of the original model; (c)- Details of the corrected model; (d)-Global view of the edges; (e)- Global view of the corrected model.

6. DISCUSSION

The technique introduced in this paper improves the 3D segmentation by fusing range and intensity data. This is particularly of interest in man made environments rich in geometric information.

The algorithm is independent of the scanning devices and can be used with any range scanner or camera.

This technique is a starting point to get a description of the model by lines. A possible extension is to directly create CAD models of buildings from real data.

On going work aims at improving the edge detection in the 3D model in order to correct all the type of range edges including roof and crease edges.

A further development is the extension of the technique to edges with any 3D geometry, i.e., straight, curved and planar, non-planar.

7. ACKNOWLEDGMENTS

This research was partly funded by the EC CAMERA TMR (Training and Mobility of Researchers) network, contract number ERBFMRXCT970127 and by the Portuguese Foundation for Science and Technology through the Ph.D. grant PRAXIS XXI/BD/19555/99. A word of gratitude is necessary to Mauro ruggeri for his data structure to interface with VRML formats.

8. REFERENCES

[1] V. Sequeira, *Active Range Sensing for Three-Dimensional Environment Reconstruction*, Ph.D. Thesis, Dept. of Electrical and Computer Engineering, IST-Technical University of Lisbon, Portugal, 1996.

[2] V. Sequeira, K. Ng, E. Wolfart, J.G.M. Gonçalves, D. Hogg, "Automated Reconstruction of 3D Models from Real Environments", *ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier)*, vol. 54, pp. 1-22, 1999.

[3] S. F. El-Hakim, C. Brenner, G. Roth, "A multi-sensor approach to creating accurate virtual environments", *ISPRS Journal for Photogrammetry and Remote Sensing*, Vol. 53(6), pp. 379-391, December 1998.

[4] I. Stamos, P. K. Allen, "3-D Model Construction Using Range and Image Data", IEEE International Conference on Computer Vision and Pattern Recognition, pages 531-536 (volume I), South Carolina, June 2000.

[5] P. Dias, V. Sequeira, J.G.M Gonçalves, F. Vaz – "Automatic Registration of Laser Reflectance and Colour Intensity Images for 3D Reconstruction", SIRS'2000, 8th International Symposium on Intelligent Robotic Systems, Univ. Reading-England, pp. 191-197, July 18-20 2000.

[6] R.Y. Tsai, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and automation*, Vol. RA-3, No. 4, pp. 323-344, August 1987.

[7] M.A. Fischler, R.C. Bolles, "Random Sample Consensus: a paradigm for model fitting with application to image analysis and automated cartography", *Communications of the ACM 24 (6)*, pp. 381-395, 1981.

[8] Eberly D.H., *3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics*, Morgan Kaufmann Publishers, September 2000.