COMBINING INTENSITY AND RANGE IMAGES FOR 3D MODELLING

Paulo Dias[(1)], Vítor Sequeira[(2)], Francisco Vaz[(1)], João G.M. Gonçalves [(2)]
[(1)] University of Aveiro/IEETA, Portugal
[(2)] European Commission, Joint Research Centre, Italy

{paulo.dias, fvaz}@ieeta.pt, {vitor.sequeira, joao.goncalves}@jrc.it

ABSTRACT

This paper presents a process combining range and intensity based techniques, in order to get better 3D models than those obtained using these techniques separately. The procedure needs an initial estimation for internal and external camera parameters for two or more intensity images. The technique uses passive triangulation to refine initial camera calibrations and ensure a good registration of range and video data sets. Afterwards, corresponding points from the intensity images are triangulated and introduced in the original range cloud of points. The objective is to complete the models in areas where data is missing or to increase the resolution in areas of high interest and 3D contents.

1. INTRODUCTION

3D reconstruction techniques are typically divided into two families: Range Based techniques (using lasers or structured lights systems) measure directly the distance between the sensor and points in the real world [1,2,3]; Intensity image Based techniques (such as stereo or photogrammetry) triangulate the 3D position of points from 2D Images [4,5]. Each one of these techniques has qualities and limitations. Range images are slow to acquire and have a limited spatial resolution, but provide precise and accurate 3D points location. Intensity images are easy to acquire and have a high resolution, but need more processing to calibrate cameras and compute 3D models precisely. Combining both techniques allows the compensation of each sensor limitation with the qualities of the other one providing a final result of better quality and building a bridge between two worlds traditionally independent.

2. PASSIVE TRIANGULATION

2.1. Fundamental matrix estimation

The main difficulty to overcome when working with intensity images in 3D reconstruction is the matching problem: how to get accurately corresponding points over different images? The technique we adopted uses the fundamental matrix to link intensity images through epipolar geometry, according to the following equation:

$$\boldsymbol{x}'^{T}_{i} F \boldsymbol{x}_{i} = 0 \qquad (1)$$

where $\boldsymbol{x} = (x, y, \zeta)$ is a homogeneous co-ordinate in the first image, $\boldsymbol{x}'$ is the corresponding co-ordinate in the second image and F is the 3×3 Fundamental matrix [4,6].
Our technique involves the computation of an initial estimation for the camera internal and external parameters (Tsai Camera model [7]). This estimation is based on the automatic registration of reflectance and intensity images. More details on the algorithm can be found in [8]. These initial camera calibrations are used to guide a cross correlation-matching algorithm over the intensity images. The correspondences feed a Ransac 8-point algorithm that computes the fundamental matrix between images [6,9]. This matrix is used to find corresponding epipolar lines and build the rectified images where rows correspond to the same epipolar lines. Using rectified images to compute correspondences between images reduces the matching problem from two to one dimension. Figure 1 shows the rectified images of two intensity images acquired at JRC Campus.

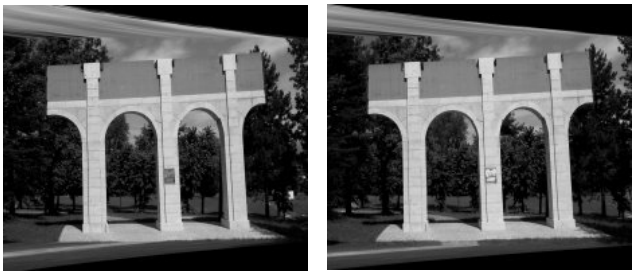

Figure 1: Two rectified images from the arches at JRC campus.

2.2. Passive triangulation

The rectified images computed in 2.1 are used to find matching points. A canny edge detector is applied in one of the intensity images to detect points of interest. The output of the edge detector is then sub-sampled to reduce the number of matching candidates. For each selected point of interest, cross correlation is computed along the epipolar line to find the best matching candidate. Once more, initial calibrations are used to estimate the probable position of the matching point and reduce the search area.
The matching points in the images are used to compute the camera projective rays by inverting the perspective projection equations from Tsai model. The result is the parametric equation of the 2 rays:

$$\vec{L}_0(s) = \vec{B}_0 + s\vec{M}_0 \quad (2)$$

$$\vec{L}_1(t) = \vec{B}_1 + t\vec{M}_1 \quad (3)$$

The squared-distance function is then:

$$Q(s,t) = \left|\vec{L}_0(s) - \vec{L}_1(t)\right|^2 \quad (4)$$

$$Q(s,t) = as^2 + 2bst + ct^2 + 2ds + 2et + f \quad (5)$$

where, $d = \vec{M}_0 \bullet (\vec{B}_0 - \vec{B}_1), e = -\vec{M}_1 \bullet (\vec{B}_0 - \vec{B}_1)$

$a = \vec{M}_0 \bullet \vec{M}_0, b = -\vec{M}_0 \bullet \vec{M}_1, c = \vec{M}_1 \bullet \vec{M}_1$

$f = (\vec{B}_0 - \vec{B}_1) \bullet (\vec{B}_0 - \vec{B}_1)$

The minimum distance occurs at a point where the gradient of the function is null:

$$\vec{\nabla} Q = 2(as + bt + d, bs + ct + e) = (0,0) \quad (6)$$

A user threshold defines the maximum distance between two rays above which triangulated points are not considered. Otherwise, the triangulated point is computed as the centre of the segment between the two closest points in the rays [10].
Figure 2 illustrates the triangulation process. It shows the range cloud of points, the position of the camera for the two considered images and the projective rays for a few points.

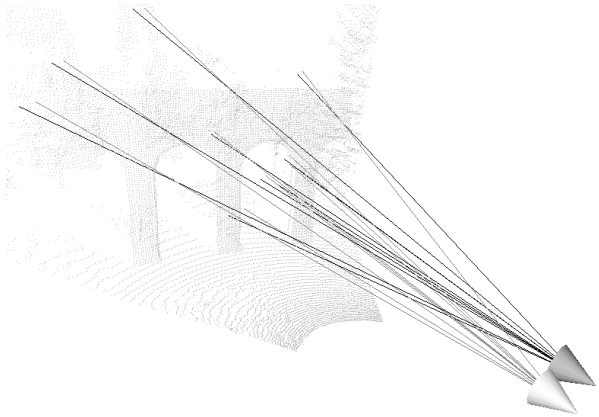

Figure 2: The passive triangulation process.

3. REFINING THE CALIBRATIONS

3.1. Association triangulated/range 3D points

The process presented in 2.2 gives, for each point of interest, the 3D position of the triangulated point. It is possible to use this information to measure the quality of the camera calibration, by measuring the distance from the triangulated points to the 3D cloud acquired with the laser. To optimise this computation, the range cloud of points is bucketed into referenced small cubes. This permits a fast navigation inside the cloud of points and a fast computation of the closest point. For each triangulated point, the closest discontinuity point in the range image is found. Figure 3 shows, in the same image, the cloud of points from the range image and the 3D triangulated points obtained from the initial calibrations.

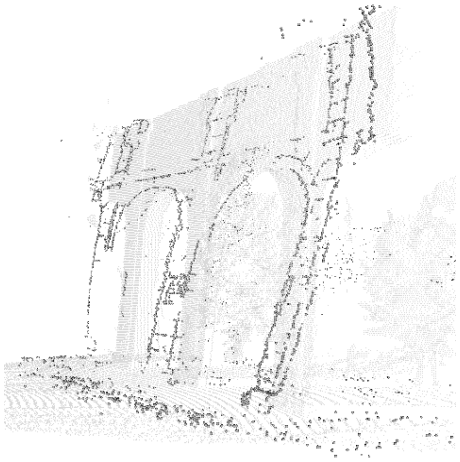

Figure 3: Range points (light grey) and intensity triangulated points (black).

3.2. Iterative calibration process

To improve the calibration and force triangulated points to converge into the 3D cloud, the closest point in the range image is used with the matching points to perform a new Tsai camera calibration for each image. The process is iterative and points are triangulated and compared with the range points in the 3D cloud, in each loop of the cycle.
This process continues as long as the average distance between triangulated and range points decreases. At this stage an additional optimisation is introduced that tries also to minimize 2D matching errors and allow for a fine tune of the camera parameters. In these new cycles of the process, pixel positions of the correspondences are updated using the current camera model. The closest 3D range points (used to calibrate the cameras) are re-projected into the images, and the new 2D match coordinates in every intensity image are computed as the centre of the segment defined by the re-projected point and the original position of the matching. Figure 4

presents the results of the triangulation at the end of the process using the optimised camera parameters. Figure 5 shows the evolution of the average distance between the range points and the intensity based triangulated points. In this example, the additional optimisation in the matching begins at the fifth iteration.

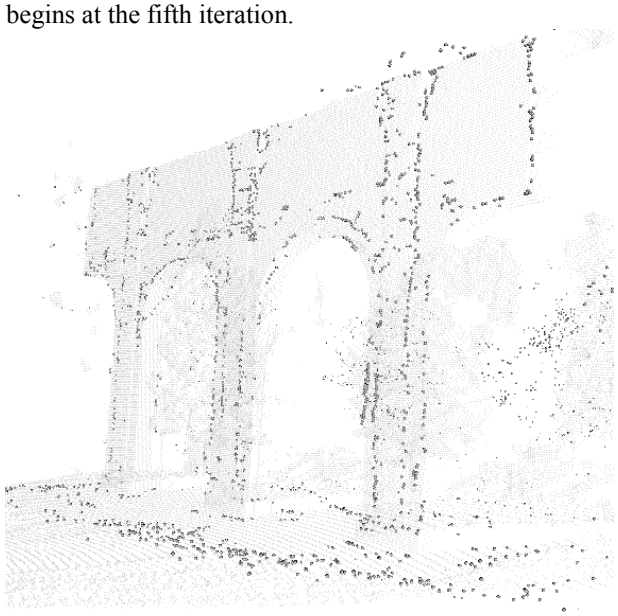

Figure 4: The range points (light grey) and the intensity triangulated points (black) after the optimisation process.

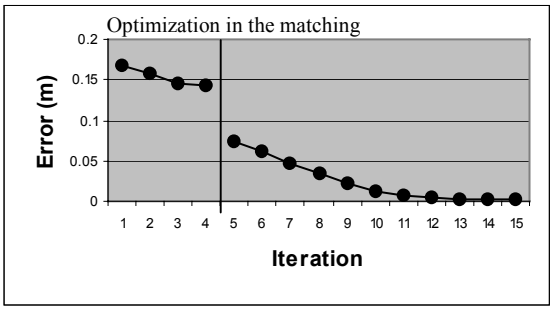

Figure 5: Evolution of the error (m) along the optimisation process.

The example presented here uses two intensity images, but the passive triangulation can be extended to as many images as desired: the process presented in 2.2 is applied to each pair of images and the final triangulated point is the centre of gravity between all the triangulated points.
Regarding processing times, it takes approximately 2 minutes on a PIV 2GHz to perform the whole optimisation loop (15 iterations with 16415 triangulated points).

4. ADDING POINTS TO THE 3D RANGE CLOUD

Thanks to the optimisation process, we ensure that the cameras are well registered with the range data, making possible the use of intensity images as an additional source of 3D information. This can be particularly useful in areas where information is missing in the range data (non reflective areas, occlusions, etc…) or also in parts of the models highly textured and rich in 3D content. Intensity images can be a valuable source of data since it is possible to acquire them easily, fast and with a very high resolution.
The process here is approximately the same as the one presented in section 2 but applied to as many points as possible. This corresponds to the dense depth-mapping step in stereo/photogrammetry techniques. Practically, all points where the variation of the gradient is significant are triangulated. Only uniform areas where the matching is not reliable are not considered (e.g. large areas of same colour such as white walls in the Barza data set, see Figure 7). The triangulated points can then be introduced into the range cloud of points before entering a 3D reconstruction process, but with data coming from both range and intensity sensors.
In Figure 6 we present some results obtained on the "São Vicente" arches. Figure 7 shows similar results with the model of a church in Barza (VA), Italy. In this case, the intensity images cover only a reduced part of the model in order to demonstrate how photographs of high resolution can be used to increase 3D point density in the range cloud.

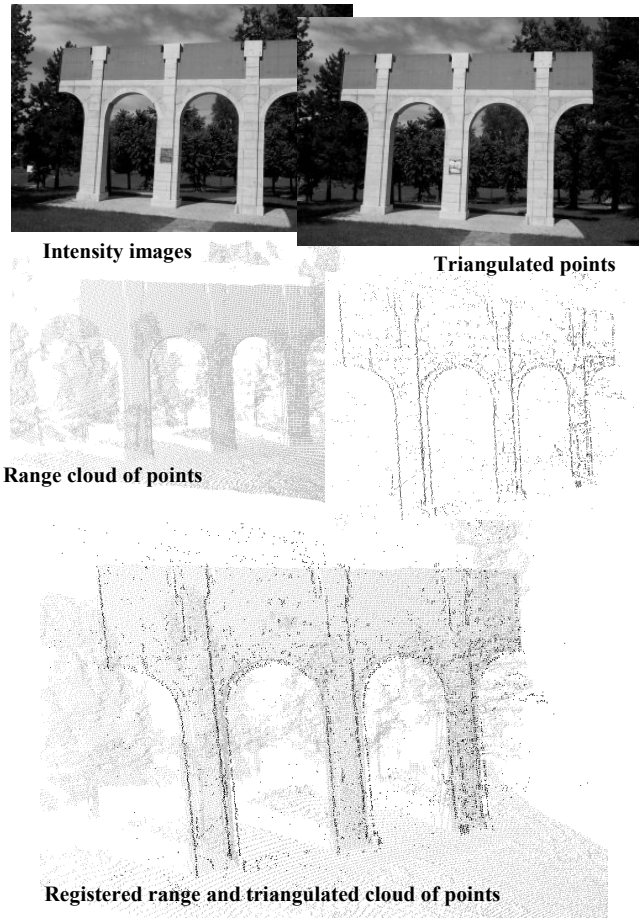

Figure 6: Addition of points in the arch example.

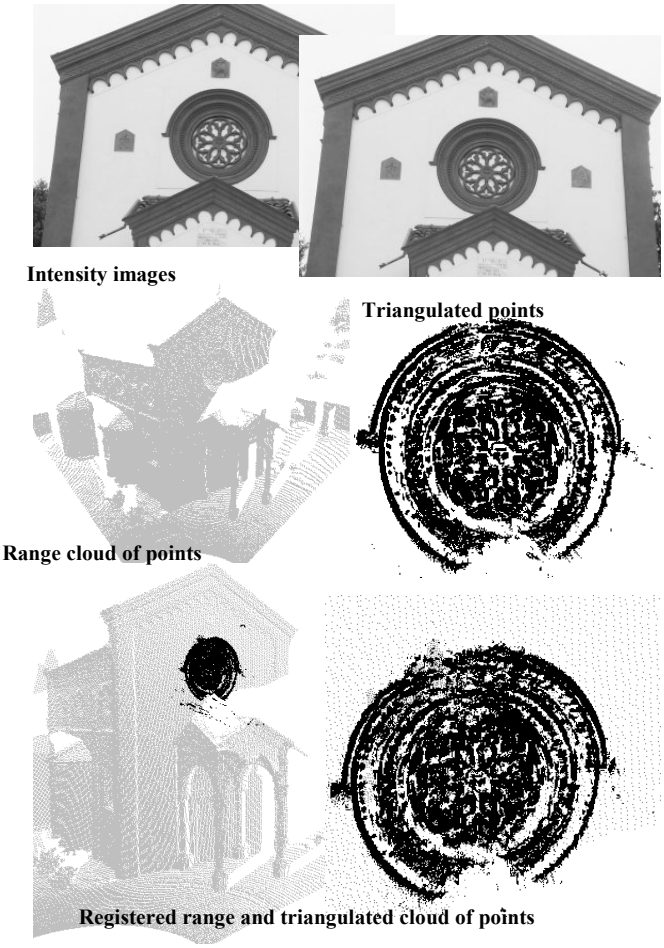

Figure 7: Addition of points in the Barza example.

5. DISCUSSION AND CONCLUSIONS

This paper demonstrates that it is possible to use passive triangulation of intensity images to improve 3D models computed from range data, more specifically in cases when dealing with low-resolution range scanners. Since the data (range and intensity) are fully registered, the 3D information can come from the two sources allowing to select the best data, for a given area of the model or for a given application. The intensity data can also be used to add 3D points in some areas of the model, as shown in the Barza example. In this case, the resolution has been significantly increased in some parts of the church using digital photographs.

The quality of the process depends mainly on two factors. The original parameters of the cameras can influence the process used to refine the camera calibrations since they are used as initial estimation. This means that if images are badly registered initially, the algorithm cannot compensate and the result is of poor quality. The other important factor on the process is the quality of the 3D edge detection in range data, since discontinuities in the cloud of points are used in the calibration refining process to select the closest 3D point for the next calibration. Fortunately new lasers have increased resolutions, making detection of discontinuities easier and much more reliable. The use of such data will probably lead to even better results than the ones presented in this document.

Finally, the acquisition of intensity data is also an important matter. Because of the dense mapping step, it is important to acquire data from close viewpoints and rich in texture information, in order to ensure a reliable matching between features.

6. 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.

7. REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] Faugeras O.D., *Three-Dimensional Computer Vision: A Geometric Viewpoint,* Cambridge, MA: Mit Press, 1993.

[5] M. Pollefeys, *Self-Calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences,* Ph.D. Thesis, Dept. of Electrical, Katholieke Universiteit Leuven, 1999.

[6] P.H.S. Torr, D.W. Murray, "The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix", *International Journal on Computer Vision*, Vol. 24, No. 3, pp. 271-300, September 1997.

[7] 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.

[8] P. Dias, V. Sequeira, J.G.M. Gonçalves, F. Vaz, "Automatic Registration of Laser Reflectance and Colour Intensity Images for 3D Reconstruction", *Robotics and Autonomous System-Elsevier*, Vol. 39 (3-4), pp. 157-168, June 2002.

[9] 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.

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