

Integration of Range Images in a Multi-View Stereo System

Xiangyang Ju Tim Boyling J Paul Siebert
Dept. of Com. Sci., Univ. of Glasgow, UK

Nigel McFarlane Jiahua Wu Robin Tillet
Sensing Group, Silsoe Research Institute, UK

Abstract

A novel method for integrating multiple range images in a multi-view stereo imaging system is presented here. Due to self-occlusion an individual range image provides only a partial model of an object surface. Therefore multiple range images from differing viewpoints must be captured and merged to extend the surface area that can be captured. In our approach range images are decomposed into subset patches and then evaluated in a "confidence competition". Redundant patches are removed whilst winning patches are merged to complete a single plausible mesh that represents the acquired object surface.

1. Introduction

Although many viable 3-D surface measurement techniques are now available, 3-D stereo photogrammetry is the only technique which is capable of capturing a full field high resolution (>4 Mpixel) range image of an object in less than 1ms [7]. This technique has the advantage that with a high degree of fidelity it can capture the shapes of moving and unstable (e.g. deformable) objects. Having captured a stereo pair of images of an object, a range image of the object can be computed by means of image matching and stereo photogrammetry (i.e. triangulation). Because of self-occlusion, an individual range image provides only a partial model of the object's surface, and thus inevitably the integration of several range images is required to recover the complete description of the object surface.

The goal of the multiple range image integration process is to merge partial object surfaces captured from different view angles. Clearly, the range images have to be registered to a common coordinate framework before they can be merged. Here we focus only on merging range images, assuming that accurate range image registration has already been achieved by camera calibration (we capture multiple views from multiple stereo-pairs of cameras simultaneously). Many researchers have investigated a variety of merging approaches to computing an integrated surface model from multiple range images. Mostly these are volumetric approaches employing implicit surfaces [1,3,4,5,9,12] and approaches involving direct surface merging [10,13,14,15,17,18]. Dyer [2] presented a review of reconstruction methods based on multiple views. There are also approaches [8,11,13,16] which combine range images and

2-D silhouette information [6] to assist the integration process.

Most of the integration approaches reported was designed to operate on 3D data obtained from active 3-D scanners. Because of self occlusion, some parts of an imaged object surface are hidden and the scanner therefore produces a depth discontinuity. This depth discontinuity can be detected by calculating the angle between the normal to the acquired object surface and the optical axis of the camera (or projector, in the case of active illumination systems). Local surface patches whose surface normal-optical axis angles are larger than a certain threshold are treated as regions corresponding to depth discontinuities [14]. Alternative measures based on the same concept have also been reported [4,5,9,17]. Because of the low level of noise in range data recovered by active scanners during surface measurement, the above approach is usually adequate to label all of the data corrupted by depth discontinuities. In the case of a stereo photogrammetry based 3D vision system, the problem of detecting depth discontinuities becomes more complicated. Firstly, in an area-based image matching system, the search process can incorrectly select correspondences biased towards depth discontinuities that produce high correlation scores, but unfortunately incorrect matches. Secondly, the noise level of the recovered range data is higher than that acquired by active scanners. Hence, the surface normals calculated from range images obtained via stereo photogrammetry are also correspondingly noisier. Such compromised range data will tend to cause classical volumetric integration approaches to fail. The practical effect of failed integration is the corruption of merged object surfaces. If direct range image merging is utilised as an alternative integration approach, the surface noise generates ambiguities when attempting to select valid surface patches. Narayanan and Kanade [9] addressed the problem and modified the volumetric approach of Curless and Levoy [1] to handle erroneous surfaces. Model based approaches [11,16] may be capable of solving the problem but require prior knowledge of the object.

Here we proposed a new approach for merging range images computed from multi-view stereo pairs, making use of stereo and silhouette information. Our approach is based on the work of Pito [10] and, Soucy and Laurendeau [15]. Soucy and Laurendeau [15] decomposed multiple range images into canonical subset views that correspond to areas common to a unique subset of range images. The subset views were parameterized and then merged into a single mesh using a weighted

average. In this method the attendant difficulties of estimating the normals of surfaces adjacent to depth discontinuities must also be addressed. Pito [10] triangulated the range images and then defined a confidence measurement for each triangle based on the position and orientation of the range scanner. In the case of multiple triangles (from differing views) sampled from the same location of the object surface patch, the triangle with the highest confidence value is retained, and the others discarded. We designed a confidence competition to remove patches from range images instead of removing individual triangles, which is sufficiently robust to allow us to use the noisy confidence data obtained from stereo matching when computing confidence for each patch.

2. The 3-D Stereo Image System

Our stereo imaging system acquires stereo image pairs from calibrated stereo cameras (left and right cameras). Silhouettes of objects acquired are extracted to create masks of the object that constrain stereo matching to areas related to the object. Applying stereo matching to stereo pair images, we obtain horizontal and vertical disparities, and a confidence image which reflects the local correlation scores of corresponding patches in the stereo pair images (Figure 1). The value of each element in the confidence image is the correlation $c(d_H, d_V)$ between the local corresponding patches $I_L(i, j)$ on the left image and $I_R(i-d_H, j-d_V)$ on the right image, where d_H and d_V are the horizontal and vertical disparities respectively, and

$$c(d_H, d_V) = \arg \max \{ I_L(i, j) \circ I_R(i-d_H, j-d_V) \} \quad (1)$$

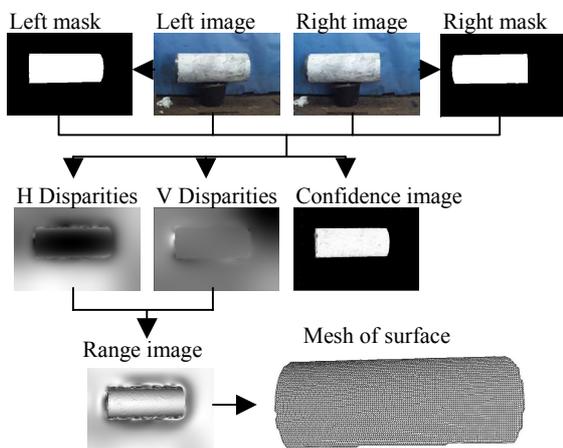


Figure 1. A pair of stereo images acquired and processed through our 3-D image system

A range image can be calculated from the disparities through the camera calibration parameters. In our approach, range surface registration is achieved by determining, through calibration, the relative orientation of each recovered range image. Each range image captured from each stereo pair can then be transformed into the same system of world coordinates. Surface merging is then achieved through our novel integration

algorithm. Our range image is defined in a generalized form,

$$r = r(i, j) \quad (2)$$

where (i, j) is the coordinates of a range image element, r the distance from the optical centre of the left camera to the point $P(x, y, z)$ on the object surface (Figure 2). P is calculated from the disparities at (i, j) .

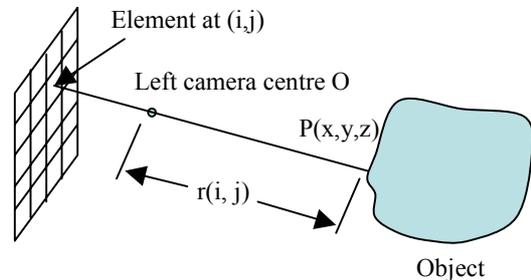


Figure 2. Range image geometry

3. The Integration algorithm

Our integration algorithm addresses three major issues:

- decomposing range images to subset patches
- removing overlapping redundant patches through confidence competition
- merging meshes from *winning* patches into a single mesh

3.1. Range Image Decomposition

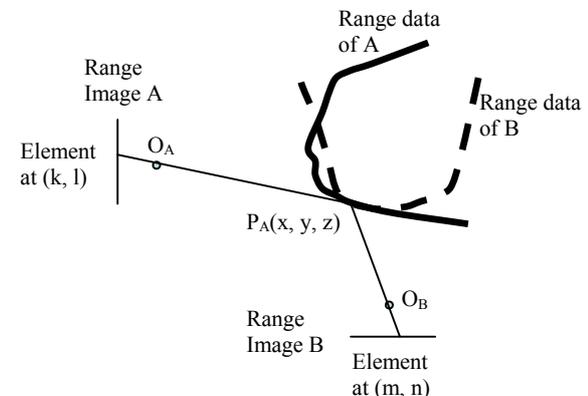


Figure 3. Comparison of range images

Our range image decomposition process makes use of range images and their masks: data in the bright area of the mask is processed and data in dark area ignored. A range image can be decomposed into *visible*, *invisible*, *overlapping* and *unprocessed* patches, when compared with a second range image, for example (Figure 3), range image $r_A(i, j)$ compared with $r_B(i, j)$.

We can label an element $p_A(k, l)$ in the range image r_A as *visible*, *invisible*, *overlapping* or *unprocessed* when comparing the range image r_A with the range image r_B . A 3-D point $P_A(x, y, z)$ can be derived from the range data

$r_A(k, l)$ at the element $p_A(k, l)$. At the same time, we can find the corresponding element of $P_A(x, y, z)$ in range image r_B , $p_B(m, n)$ with a range value $r_B(m, n)$. Now we can label the element $p_A(k, l)$ as,

visible if $r_B(m, n) > O_B P_A + \epsilon$

invisible if $r_B(m, n) < O_B P_A - \epsilon$

overlapping if $|r_B(m, n) - O_B P_A| \leq \epsilon$

unprocessed if $p_B(m, n)$ is outside the image B

where ϵ is related to the measurement uncertainty.

When all elements in the masked range image A have been labelled, adjacent elements of the same labels are grouped as *visible*, *invisible*, *overlapping* and *unprocessed* patches. In order to resolve the inherent ambiguity in selecting *visible* or *invisible* patches, we run a confidence competition (explained in section 3.2) with masked confidence images of A and B. The losers of the competition are removed by assigning their area in the mask A from bright into dark. The above decomposition and patch deletion steps are repeated for those remaining range images.

3.2. Confidence competition

For an individual (*visible* or *invisible*) patch \mathcal{S} , we run a confidence competition to decide whether we retain or remove the patch. For example, a correlation score c_A at element $p_A(k, l)$ can be found and the score c_B of the corresponding element (in the second range image) $p_B(m, n)$, so that

$$w(k, l) = \begin{cases} 1 & \text{if } c_A \geq c_B \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

and the winning rate of the patch \mathcal{S} is defined as

$$w = \sum_{(k, l) \in \mathcal{S}} w(k, l) / N \quad (4)$$

where N is the number of elements in the patch \mathcal{S} . If $w > 0.5$, the patch wins the competition and is kept, otherwise it is removed.

3.3. Mesh Merging

After decomposition and patch deletion has been completed for all the range images, triangular meshes are generated from the processed range images. The boundaries of the meshes are eroded until they no longer overlap. Any gaps that appear between the meshes are filled by growing new triangles to connect all the meshes, the procedure is similar to that in [10, 17].

4. Results

We have succeeded in capturing the surface shape of an object by developing a high-resolution stereo imaging system. Our image acquisition system comprises six commercial high-resolution colour digital cameras. These high-resolution digital cameras were configured as three

stereo pairs in order to capture the stereo pair images of the object from three perpendicular directions (namely X, Y and Z directions). The system is calibrated before and after each imaging experiment. Stereo pair images of the object are captured simultaneously and then transferred to a PC for 3-D shape recovery. Through application of the integration algorithm, a single plausible mesh of the object surface is obtained.

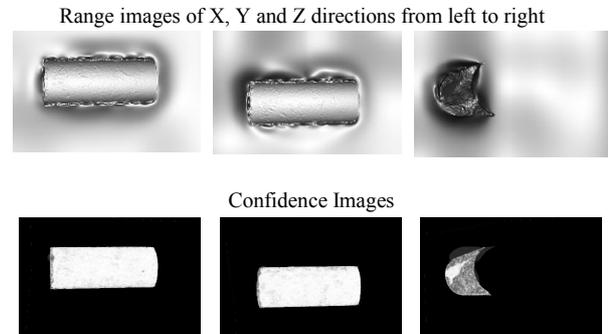


Figure 4. Range images and their confidence images of a cylinder

Figures 4 and 5 show the results of recovering the shape of a cylinder. After stereo matching, the range images and their confidence images were obtained (Figure 4). The range image in the Z viewing direction was distorted and only a fraction of this image contained valid cylinder surface. A plausible mesh of the cylinder (Figure 5) from the three range images was created by selecting the winning patches from the range images and removing the incorrect range data.

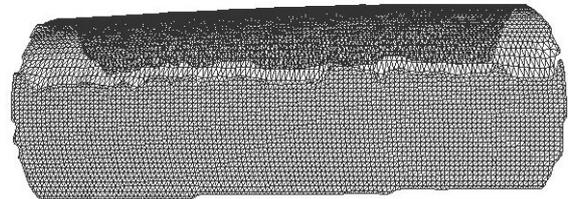


Figure 5. Integrated mesh of the cylinder

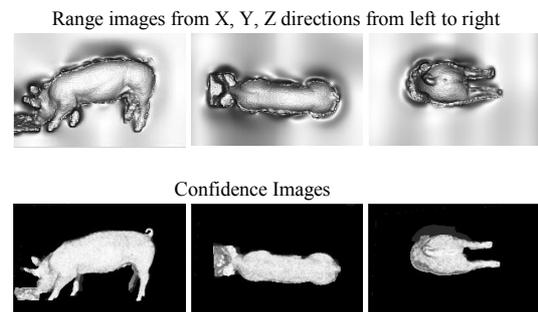


Figure 6. Range images and their confidence images of a pig

We also captured stereo pair images of 32 pigs each week over a duration of 14 weeks in order to monitor the growth of these animals. The pig's surfaces were recovered successfully by applying our integration approach.

Figure 6 shows the recovered range images (above) and confidence images (below) for each of the three views of a pig. A single mesh of the pig, integrated from these range images, is shown in Figure 7.



Figure 7. Direct mesh integration

In order to illustrate the difference between the results of a naïve volumetric approach and that of our integration approach (Figure 7), the results using a marching cubes approach (that incorrectly carves away part of the valid surface of the pig) is also shown in Figure 8.

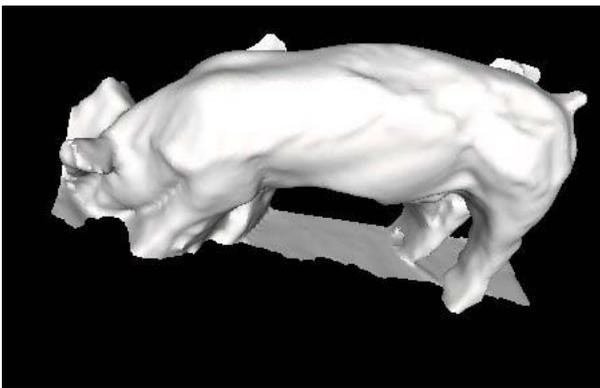


Figure 8. Mesh integration using Marching Cubes

5. Conclusions

A novel approach to integrating range images captured by a 3-D stereo photogrammetry system has been presented. Our contributions comprise: range image decomposition, redundant data removal by means of a confidence competition and a demonstration of the success of this approach. We conclude that satisfactory multi-view integration can be achieved when occlusions have been explicitly detected and removed through our confidence competition.

6. Acknowledgements

The authors would like to acknowledge the support of the BBSRC and also the UK Imaging Faraday Partnership.

References

1. B. Curless and M. Levoy (1996) "A Volumetric Method for Building Complex Models from Range Images". *SIGGRAPH'96*, pp303-312.
2. C. Dyer (2001) "Volumetric Scene Reconstruction From Multiple Views", Chapter 16 of *Foundations of Image Understanding*, L. S. Davis, ed., pp. 469-489.
3. H. Hoppe, T. DeRose and T. Duchamp (1992) "Surface Reconstruction from Unorganized Points", *Computer Graphics*, 26, 2, July, pp71-78.
4. A. Hilton, A. J. Stoddart, J. Illingworth and T. Winder (1996) "Reliable Surface Reconstruction from Multiple Range Images", *ECCV'96*, 4th European Conference on Computer Vision, Cambridge, UK, April 15-18.
5. A. Hilton, A. J. Stoddart, J. Illingworth and T. Winder (1998) "Implicit Surface-Based Geometric Fusion", *Computer Vision and Image Understanding*, March 1998, pp 273-291.
6. J. Isidoro and S. Sclaroff (2003) "Stochastic Mesh-Based Multiview Reconstruction". Boston University Computer Science Tech. Report No. 2003-016, July 1.
7. X. Ju, T. Boyling and P. Siebert (2003) "A High Resolution Stereo Imaging System". *3D Modelling 2003*, 23-24 April 2003, Paris, France.
8. M. Lhuillier and L. Quan (2003) "Surface Reconstruction by Integrating 3-D and 2D Data of Multiple Views". *ICCV2003*, pp 1313-1320.
9. P. J. Narayanan and T. Kanade (1998) "Virtual Worlds using Computer Vision". *Proceedings of the 1998 IEEE and ATR Workshop on Computer Vision for Virtual Reality Based Human Communications*, January, 1998, pp. 2 - 13.
10. R. Pito (1996) "Mesh Integration Based on Co-Measurements". *IEEE Int. Conf on Image Proc. Lausanne, Switzerland*, vol2, pp 397-400.
11. R. Plankers and P. Fua (2003) "Articulated soft objects for multiview shape and motion capture". *PAMI* vol 25(9), pp 1182- 1187.
12. K. Pulli, T. Duchamp, H. Hoppe, J. McDonald, L. Shapiro and W. Stuetzle (1997) "Robust Meshes from Multiple Range Maps", *Int. Conf. On Recent Advances in 3-D Digital Imaging and Modelling*, May.
13. Y. Kawai, T. Ueshiba, T. Yoshimi and M. Oshima (1992) "Reconstruction of 3-D Objects by Integration of Multiple Range Data", *ICPR'92*, pp154-157.
14. M. Rutishauser, M. Stricker and M. Trobin (1994) "Merging Range Images from Arbitrarily Shaped Objects", *IEEE Conference of Computer Vision and Pattern Recognition*, pp 573-580.
15. M. Soucy and D. Laurendeau (1995) "A General Surface Approach to the Integration of a Set of Range Views". *PAMI* 17(4), pp 344-358.
16. J. Starck and A. Hilton (2003) "Model-Based Multiple View Reconstruction of People". *ICCV'2003*, pp915-922.
17. G. Turk and M. Levoy (1994) "Zippered Polygon Meshes from Range Images", *SIGGRAPH'94*, Orlando, Florida, July 24-29.
18. J. Wang and M. Oliveira (2002) "Improved Scene Reconstruction from range images", *EUROGRAPHICS 2002*, 21(3).