Occlusion understanding and recovery

Castellani, U.; Livatino, Salvatore

Published in:
On-line Compendium of Computer Vision

Publication date:
2002

Document Version
Også kaldet Forlagets PDF

Occlusion Understanding and Recovery

U.Castellani and S.Livatino

1 Introduction

Range images are used in a wide range of applications. So far they have been used extensively in object recognition [10, 13], reverse engineering [2], and other applications, nearly all focusing on small and rather complex objects and scenes. While extending the use of range images to a whole environment rather than well-delimited objects (an important example of these applications is the CAMERA EU project [4]) new issues arose.

Occlusion is a major cause of information loss: even in moderately complicated scenes it is either impossible or impractical to obtain complete range scans [12]. On the other hand, an exhaustive description of the observed objects or environment is needed for some applications, like construction of a 3D model [4] and environment object recognition.

The problem to solve is the reconstruction of partially occluded simple-shaped areas, like parts of a wall hidden behind furniture pieces, the corner area of a cupboard hidden by an open door, a collection of objects on a floor hiding each other, etc. Figure 1 shows the problem, i.e. a typical occlusion in range images.

Our solution is a procedure to fill in the gaps without performing extra scans. This procedure is termed reconstruction and is able to automatically infer the shape of the occluded areas by exploiting information from the surroundings.

There have been few attempts in the literature to reconstruct occluded surfaces [1], [6], [9]. They are mostly related to simplified occlusion cases. In fact, occlusion reconstruction remains a new and little explored research field. Despite this research many unsolved cases still remain (e.g. reconstruction of object back sides) and the current research state is still far away from a general solution to recover all occlusions.

2 Handling Occlusions

There are different types of occlusions which may arise when a scene is scanned with a laser beam or structured light sensor. Lots of these occlusions are not resolvable without domain specific knowledge, or models, helping to derive structural interpretation of the images. For this work, we only consider indoor scenes containing mostly man-made objects, for which both general architectural and scene specific knowledge might be available.

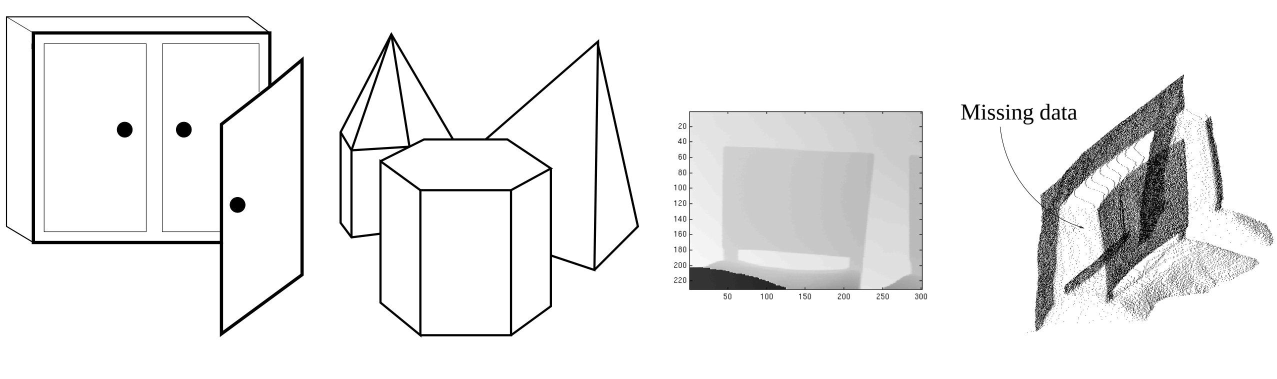

Figure 1. The Occlusion Problem

Previous research on occlusion reconstruction focused on the reconstruction of a single large area occluded by one object. In that context two cases were considered: Occlusions Preserving Surfaces, Occlusion Breaking Surfaces.

- **Occlusions Preserving Surfaces** In this case different image regions correspond to different objects and the occluding region is entirely surrounded by the occluded one. For example, a book lying in the middle of a table scanned by a range sensor from above. Figure 2.a shows an example of these occlusions: the surface B occludes surface A.

  The detection and the reconstruction of the occlusions is based on searching for regions entirely contained within the boundaries of another region, so that they can then be extended across the occluding area [5].

- **Occlusions Breaking Surfaces** In this case the occluding part breaks the occluded surface into two regions which correspond to a single surface. The wall part obscured by the chair in figure 2.b is an example of this case. The surface B occludes the surface A which is split into regions A' and A".

  The detection and the reconstruction is based on identifying compatible regions which can then be merged in order to reconstruct the missing occluded part [1, 6].

In both the above cases the occluding object does not obscure the boundaries of the occluded one. The work of [9] explored the case when the occluding part partially obscures boundaries of an occluded object. This case is called Occlusion Breaking Boundaries. The corner area of a cupboard occluded by an open door, as in figure 1, is an example of this case. Figure 2.c shows an example of these occlusions.

- **Occlusions Breaking Boundaries** The proposed case is harder to solve than the previous ones since a reconstruction of unseen boundaries is also needed. Furthermore, it can happen that more objects are obscured under a single occlusion. The proposed approach is different from the previous ones since the focus is on the detection of the occluded part of an object.

  The detection and reconstruction is based on establishing a foreground - background relation between regions during the analysis of region boundaries. In particular, the background region is occluded by the foreground region and reconstruction starts from its boundary.

In order to reconstruct occlusions we need to exploit available information that constrains a scene. In particular:

1) good surface continuation. That is, the occluded surface keeps the same shape of its visible part [8, 11].

2) good boundary continuation. That is, the occluded boundary keeps the same slope of its visible part.

3) architectural constraints. That is, the occluded surface is bounded by an architectural constraint. This can be a wall, a floor, a door, a window etc. [3, 7].

The reader should note that in order to solve cases of Occlusion Preserving Surfaces and Occlusion Breaking Surfaces, the surface good continuation constraint suffices. In the proposed case of Occlusion Breaking Boundaries it is also needed to apply the constraint of good boundary continuation and the architectural constraints.

3 Occlusion Understanding

The goal is to understand the kind of occlusions that arose by the scanning. Our input is a range data set that provides 3D spatial information related to the sensor position. The data contains depth information that we exploit for detecting occlusions.

In order to detect occlusion is more convenient to switch from true 3D point-sets to a 2.5D representation, i.e. a range image. In this way, we observe the scene from the sensor viewpoint, that is, we look at occlusions the same way the sensor does.

3.1 Method for Occlusions Preserving Surfaces and Occlusions Breaking Surfaces [1, 6]

Image Segmentation

The image range can be segmented by extracting depth discontinuities either across each coordinate axis [1], or between neighbouring fitted 3D surfaces [6]. The case of depth discontinuities is represented in figure 3, while the case of 3D surfaces is represented in figure 4.

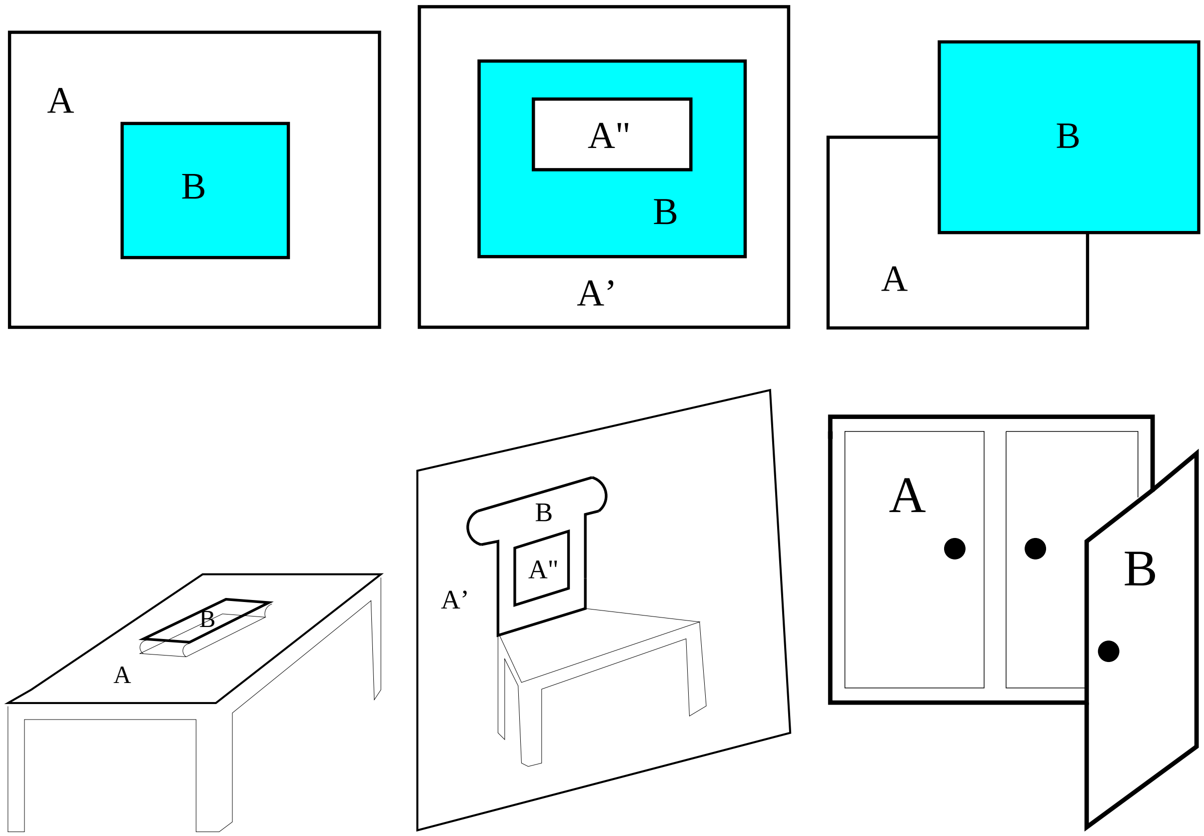

Figure 2. Type of occlusions

Figure 3. A Range image (left) and corresponding extracted regions after labelling (right).

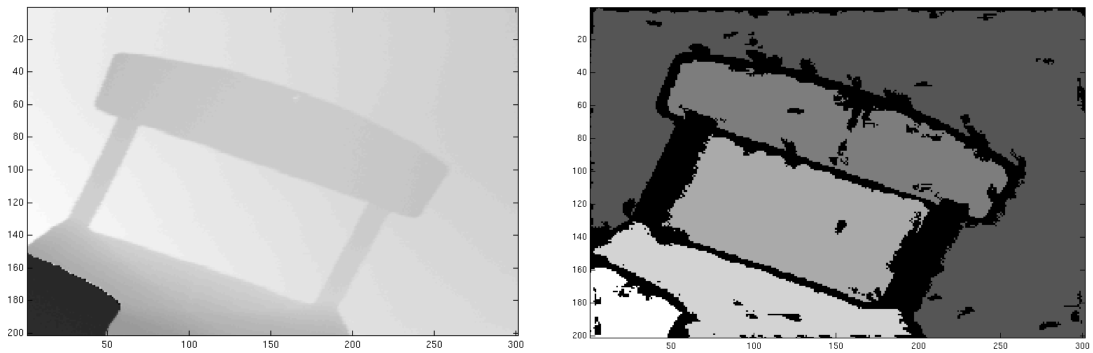

Figure 4. The first segmentation.

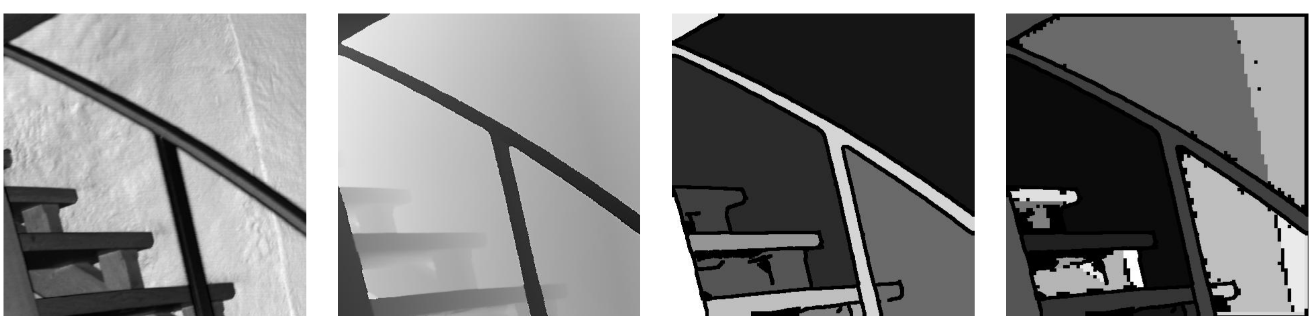

Figure 5. An intensity image (left) and its corresponding range image (center-left). The same image after depth discontinuity processing (center-right) and additional partitioned using fold edges (right).

The extracted surfaces are further segmented by analysing their surface orientation. This process is called fold edge detection. The result of this detection is shown in figure 5. The right-hand figure shows that a fold edge detection leads to a more accurate detection.

Surface Fitting

The extracted surfaces are then fitted by parametric surface shapes in order to obtain good comparative measures. Experiments have shown that planes, cylinders and spheres are usually sufficient to describe most of the simple surfaces in architectural scenes. The following table shows the parameters associated with each surface type.

Surface type	Description
Planar	surface normal $surf\_nl$ displacement $disp$
Cylindrical (circular)	point on axis p unit vector of axis $axis\_uv$ radius r
Spherical	centre c radius r

For two surfaces to be contiguous, they must be first of the same type. The surface parameters introduced above can be used for comparison. The following table gives the parameters and the thresholds used for determining whether two surfaces match or not.

Surface type	Requirements for matching
Plane	$surf\_nl_1 \cdot surf\_nl_2 < 5\deg$ $\lvert disp_1 - disp_2 \rvert < 5cm$
Cylinder	$axis\_uv_1 \cdot axis\_uv_2 < 5\deg$ $axis\_uv_2 \cdot p_1p_2 < 5\deg$ $p_1p_2 \cdot axis\_uv_1 < 5\deg$ $(2 * \lvert r_1 - r_2 \rvert)/(r_1 + r_2) < 5\%$
Sphere	$(2 * dist(c_1, c_2))/(r_1 + r_2) < 10\%$ $(2 * \lvert r_1 - r_2 \rvert)/(r_1 + r_2) < 5\%$

3.2 Method for Occlusions Breaking Boundaries [9].

The detected surfaces are considered in pairs and their boundaries compared. If depth values associated with adjacent surface boundaries are similar, the surfaces are contiguous and so not occluded in the area surrounding those boundaries. The boundaries between the wall and cupboard shown in figure 6 are examples of boundaries between contiguous surfaces. We call them true boundaries.

If instead, depth values associated with adjacent surface boundaries are different, it means that the surface closer to the sensor is in foreground and the other one is in background. In this case the boundary of the background surface is a false boundary. In figure 6 the boundaries between door and cupboard are examples of false boundaries. We have to establish a foreground-background relation for each pair of adjacent regions.

Once false and true boundaries have been identified it is possible to detect the points where the background surface boundary is occluded. We call them *endpoints*. The endpoints are the ends of the background surface true boundaries and they lie just next to the false boundary endpoints. Figure 6 shows the endpoints for the example case. It is important to detect endpoints because they represent the starting points for the reconstruction.

4 Occlusion Recovery

The key to reconstruction is to identify which part of the occluded region can potentially be connected behind the occlusion.

To reconstruct occluded areas is a difficult task since we need to estimate something which is not *visible*. Sometime it is in fact impossible to reconstruct occluded areas, such as the case when the hidden area contains abrupt surface changes which do not follow the concept of continuous constraints. This explains why we propose a conservative approach to the reconstruction. We need to be sure we see the right condition allowing some reconstructions.

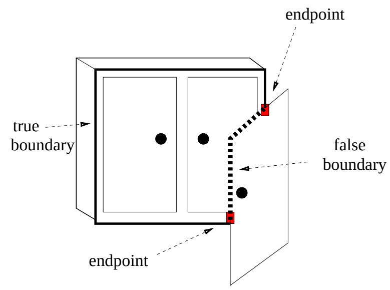

Figure 6. Endpoints and boundaries between adjacent surfaces.

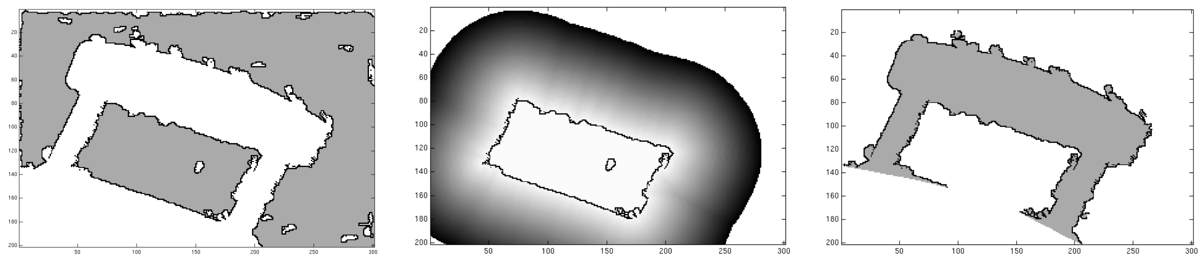

Figure 7. Matching areas. These belong to the same wall lying behind the chair. Areas are gray, perimeter points are shown in black (left). The distance transform for one of the areas shown in the left figure. Perimeter points have been added for clarity. (right) The potentially reconstructible area between the two surfaces shown in the left figure.

4.1 Method for Occlusions Preserving Surfaces and Occlusions Breaking Surfaces [1], [6].

Identifying Allowed Reconstructions

We rely on distance transforms [6]. This method finds perimeter points that are actually facing the connecting area. Then we draw lines between the pixels facing each other with the Bresenham algorithm, as in [1]. The method relies on distance transforms, which encode the distance (in pixels) to the nearest region point. Figure 7 (center) shows the distance transform of one of the matching areas in figure 7 (left). Figure 7 (right) shows the potentially reconstructible area between the two surfaces shown in figure 7 (left). The distance transforms records the distance to the nearest boundary point allowing an easy computation of the hypothesised occluded surface.

Surface Reconstruction

Reconstruction takes place between regions belonging to the same group and satisfying the depth constraint. Given an occluding pixel and an occluded surface, a simple and intuitive way to perform reconstruction is to intersect the ray from the sensor through the occluding point with the occluded surface. As this ray overlaps the optical ray of the laser scanning beam, the reconstructed pixel is placed in a position that could actually have been sensed by the sensor.

For planes there is usually one intersection. For cylinders and spheres we usually find two intersections. If the shape of the surface is concave, the most distant intersection is chosen. Similarly on convex surfaces the closer intersection is chosen.

Due to noise, surfaces extrapolated into the possibly occluded area never perfectly match each other. Consequently, the surfaces are interpolated. A solution is a weighted averaging between all the intersections with the extrapolated surfaces, with weights depending on the distances from the intersection to the closest perimeter point of each surface. The position of the reconstructed pixel can be computed by:

$$\vec{x}_{inter} = \frac{\sum f(d_s)\, \vec{x}_s}{\sum f(d_s)} \tag{1}$$

$$f(d_s) = (d_{max} - d_s)^{\frac{3}{2}} \tag{2}$$

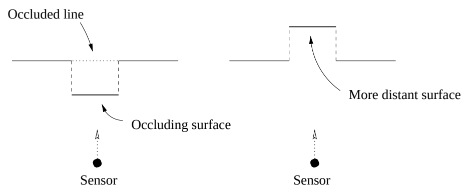

Figure 8. A *real* occlusion and a niche

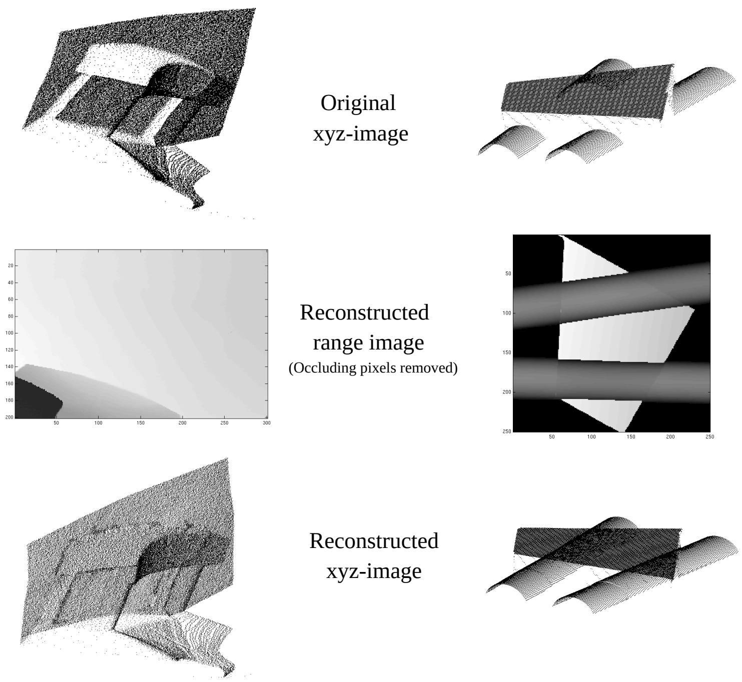

Figure 9. Results for an occluded wall (left) and two occluded cylinders (right)

The summation in equation (1) takes place over all the surfaces involved in the reconstruction. Because points that are closer to a certain surface would be weighted more than points that are further away, the weighting function should be decreasing with the input distance. The weighting function used for the example is shown in equation (2).

Reconstruction Validation

The choice now becomes whether to reconstruct or not. Reconstruction is a severe change to the image, so it is important to be very careful in applying it. If the area between two matching surfaces is further away from the sensor than the area of the surface to be reconstructed, then reconstruction should not happen (see figure 8). This happens on many occasions when niches as door or windows are involved. In cases of Contiguous Occluded Areas or Single Region Occlusions, a proposed solution is to first reconstruct the surface and then compare reconstructed and measured points. Reconstruction is allowed only if reconstructed points lie behind the measured ones. All the pixels in the area between the surfaces are considered. To achieve a more reliable result each pixel vote for reconstruction, and only if enough pixels votes for a reconstruction will it be applied. As any reconstruction requires hypothesising non-observable data, the assumptions lead to a conservative reconstruction having a high likelihood of being correct.

Figure 4.1 shows some reconstructed images. The left image shows the occlusion caused by a chair in front of a wall. The right image is an image acquired by an orthographic scanner. The middle row shows a range image with reconstructed pixels and the bottom row shows the reconstruction in 3D. Because the reconstructed surface appeared unnaturally smooth, the same amount of Gaussian noise in the z-direction as found in the original surfaces was added.

Texture Reconstruction

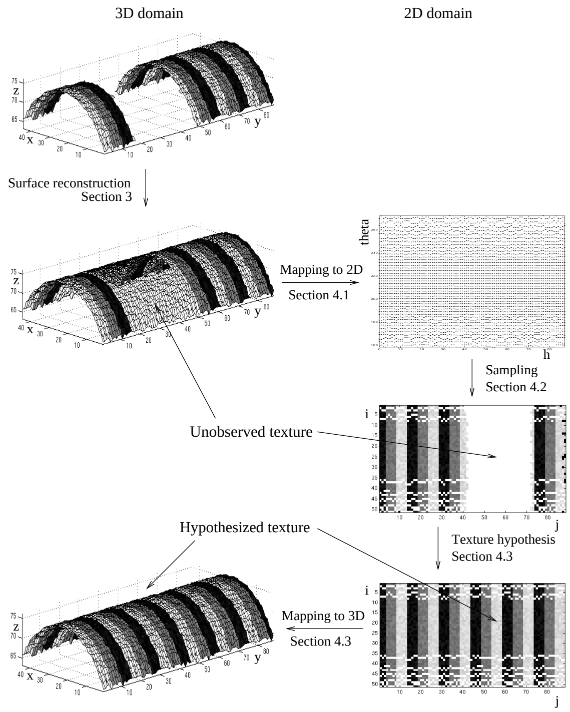

Figure 10. Texture reconstruction

Although reconstructing the occluded surfaces makes the 3D models more realistic (no more chair-shaped holes as in the figure), they still lack intensity texture. The reconstructed surfaces are merely points in 3D space without any information about the light intensity or colour at that point. So, we hypothesise what the intensity on the surface would have been if it was possible to directly sense the surface.

The main idea is illustrated in diagram of figure 10 The reconstructed points and the neighbouring measured ones are mapped onto a planar texture space. This step allows easier and more efficient processing than the unordered points in 3D space. For 3D scene planes, this is a simple translation and rotation. For cylinders we use a h-θ surface parametrization, and for spheres a θ-ϕ parametrization.

An acquired intensity image is mapped on the same plane according to number and position of the mapped surface points and then propogated by a variety of methods. The reconstructed 3D points then acquire a hypothesized intensity from the corresponding point in the surface texture space.

4.2 Method for Occlusions Breaking Boundaries [9].

The proposed method is based on the concept that the occluded area is filled in with the same type of surface which fits the visible area. To achieve this we first need to estimate the boundaries of the occluded region. Consequently, for each endpoint we estimate the direction of its continuation within the occluded area. The method proposed to estimate the occluded boundary is based on the boundary good continuation constraint. In particular, the endpoint prolongation is performed according the Gestal principle of good continuation and proximity in the order: linearity, co-circularity [8], closure [11]. In this way we are able to bound the surface which is going to be reconstructed.

We can distinguish three possible cases based on the relation between visible boundaries lying in the proximity of an intersection with the occluded surface.

case A) coincident boundaries;

case B) convergent boundaries;

case C) divergent or parallel boundaries.

This cases are represented in figure 11.

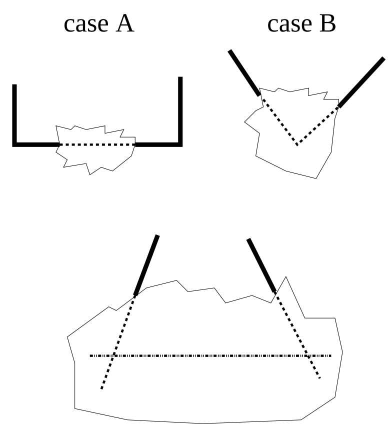

Figure 11. Proposed reconstruction rules

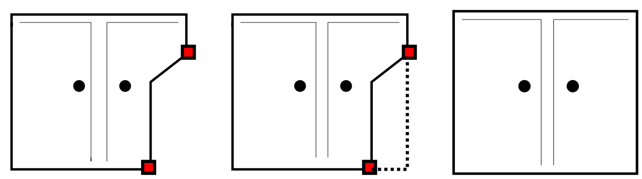

Figure 12. Main steps of the reconstruction process (case of convergent boundaries)

In all the three cases the hidden boundary is the continuation of the visible one. However, in the first case, the boundary extensions across the occluded area are coincident to the line connecting endpoints; in the second case, the boundary extensions over the occluded area intersect within the occluded area (e.g.figure 13); in the third case, the boundary extensions over the occluded area do not intersect (e.g.figure 15). The last case needs an architectural constraint to limit boundary extensions. In the work presented here, we assume that the extension does not pass through walls or the floor. This may overextend some surfaces; the alternative would be more conservative and not reconstruct these cases - awaiting instead additional data.

Hypothetical surfaces can then be created by extending the visible surface regions into the identified bounded area. The method to estimate the occluded surface is based on the surface good continuation constraint. That is, we hypothesise that the surface does not change its shape within the occluded area. The reconstruction is performed in the 3D space in order to achieve a higher accuracy. Given an occluding pixel and an occluded surface we intersect the ray from the sensor through the occluding point with the occluded surface. As this ray overlaps the optical ray of the laser scanning beam, the reconstructed pixel is placed in a position that could actually have been sensed by the sensor. For planes there is usually one intersection. For cylinder and spheres we usually find two intersections. Figure 12 summarizes the recovery method. Figures 13 - 15 show some examples of reconstruction.

Acknowledgements

This work was funded in part by the CAMERA research network (EC Contract No.FMRX-CT970127) under the EU TMR Programme.

References

[1] F. Dell'Acqua and R.B. Fisher. Reconstruction of planar surfaces behind occlusions in range image. *IEEE Trans. Pat. Anal and Mach. Intel*, to appear.

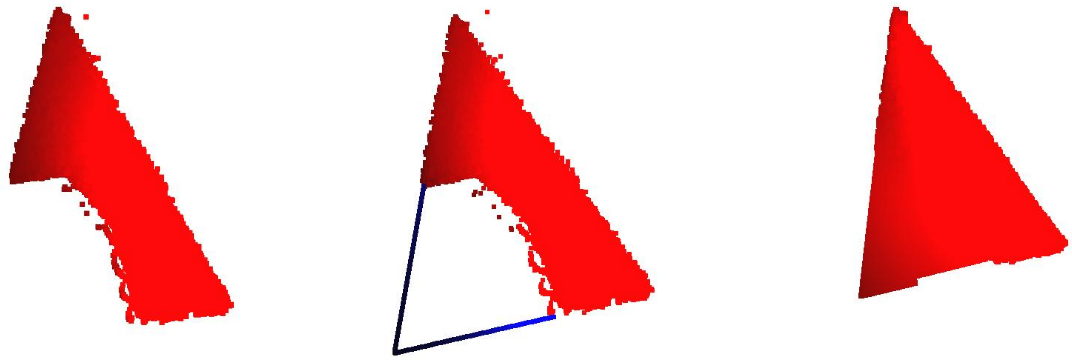

Figure 13. Reconstruction of a surface with rules of case B. Occluded surface (left), extended boundaries (center) and reconstructed surface (right).

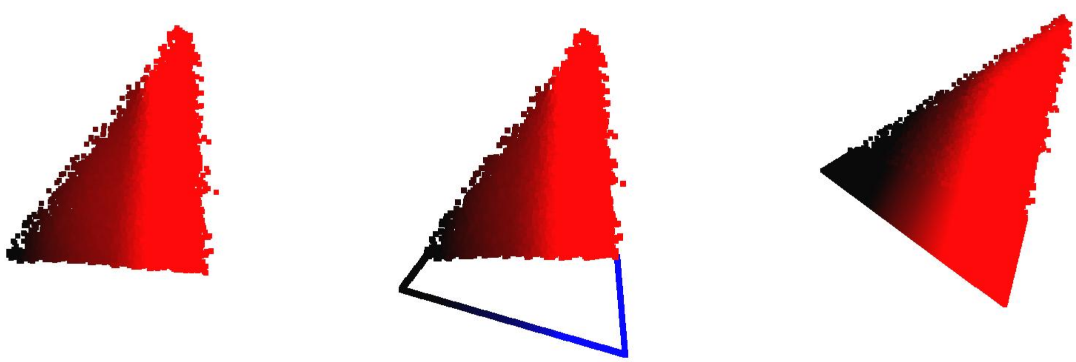

Figure 14. Reconstruction of a surface with rules of case C. Occluded surface (left), extended boundaries (center) and reconstructed surface (right).

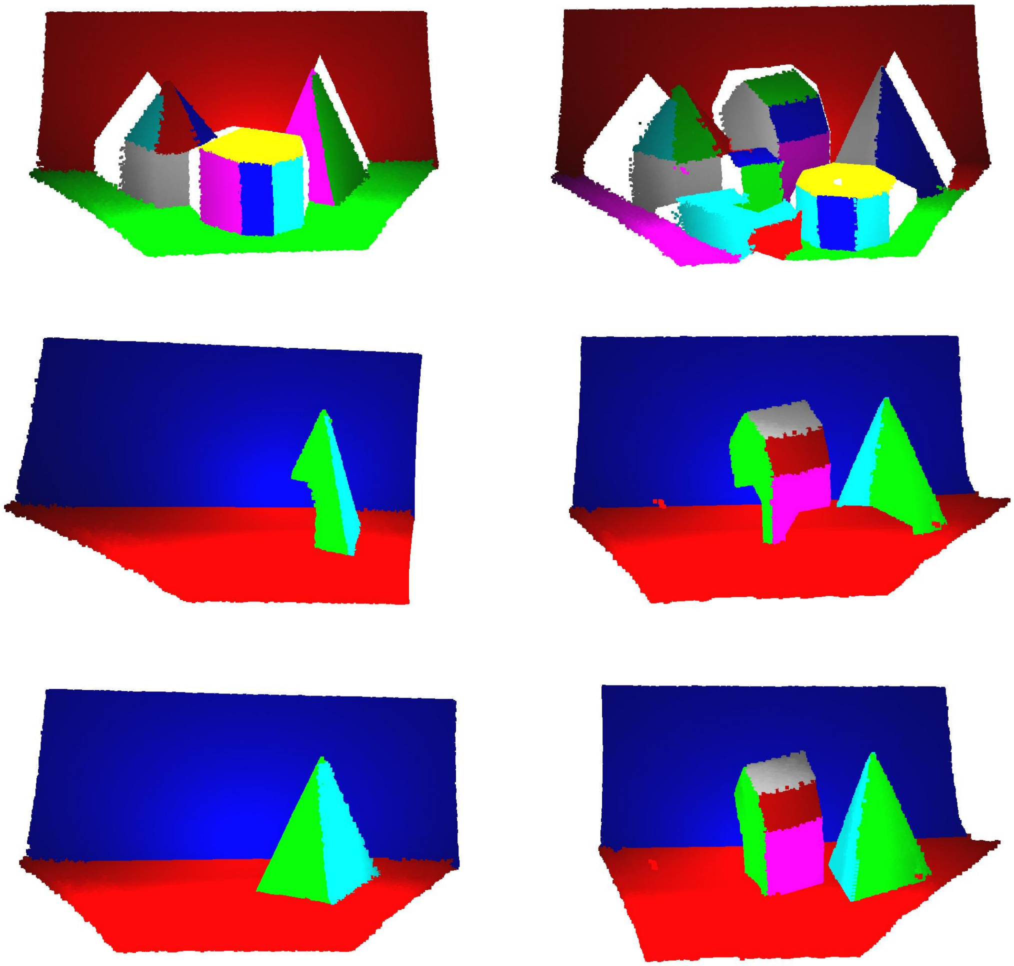

Figure 15. Examples of reconstructed objects: right-side pyramid of scenario 1 (top-left) and right-side pyramid and *little house* **of scenario2 (top-right). Pyramid pyramid before reconstruction (center-left) and pyramid after reconstruction (bottom-left),** *little house* **and pyramid before reconstruction (center-right) and** *little house* **and pyramid after reconstruction (bottom-right).**

[2] A.W. Fitzgibbon D.W. Eggert and R.B. Fisher. Simultaneous registration of multiple range views for use in reverse engineering of cad models. *Computer Vision and Image Understanding*, 69(3):253–272, 1998.

[3] P. Faber and R.B. Fisher. How can we exploit typical architectural structures to improve model accuracy? submitted for review.

[4] R.B. Fisher. http://www.dai.ed.ac.uk/daidb/ people/homes/rbf/camera/camera.htm.

[5] F.Stulp. *Completion of Occluded Surfaces, M.Sc Dissertation*. PhD thesis, Rijksuniversiteit Groningen, 2001.

[6] F.Dell'Acqua F.Stulp and R.B.Fisher. Reconstruction of surfaces behind occlusions in range images. In *Proc. 3rd Int. Conf on 3-D Digital Imaging and modelling (3DIM01), Quebec City, Canada*, pages 232–239, 2001.

[7] M. Devy H.Cantzler and R.B. Fisher. Quality enhancement of reconstructed 3d models using complanarity and constraints. sumbitted for review.

[8] L. Herault and R. Horaud. Figure-ground discrimination: a combinatorial optimization approach. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 9(15):899–914, 1993.

[9] U. Castellani S. Livatino and R.B. Fisher. Improving environment modelling by edge occlusion surface completion. sumbitted for review.

[10] C.K. Cowan P.G. Mulgahonkar and J DeCurtins. Understanding object configurations using range images. *IEEE Trans. Pattern Anal. and Mach. Intel.*, 14(2):303–307, 1992.

[11] K.K. Thornber S. Mahmud and L. R. Williams. Segmentation of salient closed contours from real images. In *Proceedings of the International Conference on Computer Vision*, 1999.

[12] J.M. Sanchiz and R.B. Fisher. Environment recovery by range scanning with a next-best-view algorithm. University of Edinburgh, to appear in Robotica, 2000.

[13] R.B. Fisher A.W. Fitzgibbon M. Waite M. Orr E. Trucco. Recognition of complex 3-d objects from range data. In *Proc. CIAP93*, pages 509–606, 1993.