Occlusion detection and surface completion in 3D reconstruction of man-made environments

Radka Pospíšilová
Masaryk University, Faculty of
Informatics
Botanická 68a
602 00, Brno, Czech Republic
xpospis5@fi.muni.cz

ABSTRACT

This paper presents the methods for the quality improvement of the man-made environments reconstruction. The 3D model is created from the set of images without any user interaction. Some generic assumptions about the scene geometry are used; for example continuity of surfaces, piecewise planar scene or frequent parallelism and orthogonality between planes and lines. This information is employed to reduce the ambiguity of the 3D reconstruction and for the occlusion recovery. Method for the completion of the occluded surfaces is presented. First, the position of missing parts of the scene is detected and then the gaps are filled by the most suitable shape.

Keywords
3D reconstruction, occlusion, surface completion.

1. INTRODUCTION

In this paper we focus on the quality improvement of the 3D reconstruction of man-made environments. The 3D model of the scene is created without any user interaction from the set of images. We investigate the possibilities of the completion ambiguous and unreconstructed surfaces. Concretely, we focus on full automatic completion of the gaps in 3D model, which was caused by occlusion. An occlusion is a covering of far surfaces by the closer ones. If the covered parts of the scene are visible in simply one image, we have only the information about texture and shape of their projection. If the covered parts are not visible on any photography, the surface is completely unknown.

The occlusion causes problems in most systems for scene reconstruction. It makes the reconstruction

process more erroneous and it creates holes in the 3D models, so the result of the reconstruction is visually unsatisfactory. Small holes in the surface are usually filled by interpolation, but if the object is indented, the interpolation gives very bad results. If the occlusion occurs in reconstruction of unknown scene, it is impossible to fill occluded surfaces because the geometry of the scene is unknown.

We focused on situation where we have some assumptions or information about the scene, so it is possible to use them. Man-made environments often satisfy some simple and useful assumptions and it is relatively easy to employ them in the reconstruction. In our system we used algorithm for 3D reconstruction which is able to reconstruct relatively exactly the neighbourhood of the gaps and then the best shape for the occluded surface is estimated.

Nowadays we develop the system for the reconstruction of piecewise planar scenes with straight lines. We integrated the algorithm of filling holes into the reconstruction process, but it is possible to use it for formerly completed models too.

2. RELATED WORK

Most of the existing systems for 3D reconstruction don't compute the shape of the occluded parts in the scene. Usually if the occlusion is detected, the

missing surface is completed by the interpolation. So the 3D points on occluded surface are computed from border points only. Most existing sophisticated approaches detect and solve the occlusion in range image from 3D scanner, for example [Cas02a] or [Del01a].

Occlusion Detection in Range Image
Some approaches only detect the occlusion and then compute the next best position of the range sensor [San99a]. The found position of the sensor is optimal for the holes completion.

The other methods also detect the occlusion and then they can complete some types of occluded surfaces. In [Del01a] the assumption on continuous surfaces is used and algorithm is focused on indoor planar scene. If some planar surface in the 3D model is splitted to the smaller parts by occlusion, then the optimal position of the plane is computed from these small parts. Then the parts are connected so the plane lies over the occluded gaps. Partly occluded non-planar surfaces (for example parts of cylinder or sphere) can be reconstructed too. When the shape and type of geometric object is recognized from the visible parts, the surface can be extended over the occluded parts.

The approach of Castellani et al. [Cas02a] is specialized on reconstruction of the planes, corners and edges of partially occluded simple-shape objects. Some types of larger occlusion can be solved too, for example object composited of one or more polygons and their parts. In Castellani's algorithm first the range image is segmented and the planes are divided into the background and foreground. Then the end-points of the edges between the foreground and the background are found and finally the surface and edges are reconstructed.

Occlusion and 3D Reconstruction from Images
If we reconstruct 3D model from the set of images, the occlusion can cause many problems. Especially many false 2D points are produced by edge detector. When the correspondences between the points are searched, all correspondences which contain these false 2D points are obviously incorrect, so the algorithm for the structure and motion estimation must be really robust due to a big amount of outliers. Another problem is a detection of the occlusion and finding of the borders of the obscured parts. For these two reasons there aren't many approaches focused on computing the occluded surfaces from images.

Approaches with **model fitting** are the exceptions. We have a set of predefined models, where some properties of the models are known; for example the shape, color or texture. If we fit the models into the scene, it is possible to compute the small invisible parts of the surface. Model fitting is very useful especially in the reconstruction of an architecture object, because it is possible to complete the occluded parts of the columns, windows, doors, etc.

Another advantage of model fitting approach is more precise reconstruction of repeated models in the scene. For example buildings often contain several windows with the same shape, so it is possible to choose the best displayed windows and replace another distorted and occluded models by its. More information can be found for example in [Dic01a] or [Wer02b].

The occlusion must be detected also in the **dense depth estimation**. First, pictures are rectified so the epipolar lines are aligned horizontally. The corresponding points for each pixel are searched along the epipolar line. If the uniqueness of the correspondences is corrupted (one point has more then one corresponding point), then the occlusion probably occurs and space between these two points was displayed only in one image.

3. ASSUMPTIONS ON THE SCENE
The application of the assumptions on the scene geometry is very common in reconstruction of the man-made environments. Various assumptions and prior information can be used for calibration of cameras and for the quality improvement of the 3D model. Such prior information does not only solve the projective ambiguity in the reconstruction but also usually stabilizes the sensitive reconstruction process.

The reconstruction of architecture objects without any information is relatively complicated task. Searching for the point correspondences is often unreliable for some areas in the images. It is caused especially by the homogeneous region in the scene, repeated pattern, various lighting condition and occlusion. If we have only few images of the scene, then the problem is getting more serious.

Assumptions about the Architecture
Only simple and easy applicable geometric constraints we used for the reconstruction of the scene with the occluded parts. There is a list of employed assumptions:

- **Large coplanar regions** – man-made scenes usually contains coplanar regions. Large regions can be discontinued by wedge or protruded objects. These objects can be already coplanar.

- **Continuity of surface** – each plane has a boundary which is composed of the lines. If the occlusion doesn't occur then each border line has two planes which go through the line. If any border line lies only on one plane, probably the second plane isn't reconstructed, so the best position and orientation of the plane must be found.

- **Occluded surface doesn't change shape abruptly** – if this assumption is corrupted, the holes can't be reconstructed correctly.

- **Perpendicularity** of planes and lines.

- **Parallelism** of planes and lines.

- **Straight lines** – the scenes of this type usually contain big amount of straight lines. So we can assume that slightly curved lines are really straight and curvature is caused by the inaccuracy or distortion of the cameras.

- **Symmetry** – the symmetry along some axis frequently occurs in architecture. Usually it is suitable to search symetry along the axis, which is parallel to one of the dominant direction in the scene. The symmetry can be applied even in the situations where the other assumptions are useless because of a large area of the occluded surface. Unfortunately it is hard to detect the symmetry if the object is too complicated.

Figure 1. Example of architecture object (Kroměříž Chateau), which fulfils these assuptions.

4. 3D RECONSTRUCTION WITH GAPS COMPLETION

Our algorithm for reconstruction of the scene is inspired by the some methods which are described for example by Tomáš Werner [Wer02a], Baillard and Zisserman [Bai99a] etc. We extend and modify these existing approaches, so it is possible to fill the holes and ambiguous parts of the 3D model.

Points and Lines Reconstruction

First, feature points and lines are matched among the images. We reconstruct both lines and points because lines are not corrupted by the occlusion so much as

the points. On the other hand three views are necessary for the reconstruction of lines and reconstruction of the points needs only two views. Therefore we can obtain the positions of 3D points from those parts of the scene where the line reconstruction is impossible.

Trifocal tensor is used for 3D reconstruction of lines and for computation of camera matrix. It is estimated using the standard robust methods (see [Har03a] for full details). Now the camera matrices and a set of 3D points and lines must be transformed to metric reconstruction.

4.1.1. Upgrade to metric reconstruction

Our system must be able to make quality 3D model from relatively small amount of images; therefore the geometric constraints are employed in camera calibration and metric rectification process. Man-made environments usually contain three dominant directions in the scene and these directions are perpendicular. First, the principal directions (represented by points at infinity) are computed from their images – vanishing points. The plane at infinity is identified and the projective reconstruction is updated to affine. Then the model is transformed so that the principal directions are perpendicular. Finally, the scaling in each principal direction is computed using assumption on the square pixels.

Plane and Polygon Fitting

Some approaches detect planes and polygons directly in the cloud of 3D points, which are reconstructed using dense depth estimation. However these methods are inaccurate for 3D points, which are situated near occluded parts of the scene. Therefore we use the different plane fitting algorithm which uses plane-sweep method.

In our approach we have used two methods for plane fitting, it is plane-sweep algorithm and RANSAC [Fis81a]. Details about plane-sweep method are described in [Bai99a]. We use the reconstructed feature lines and points for the detection of planes. Big advantage of the plane-sweep method is an ability to determine the scene plane even in the absence of any reconstructed points of that plane. On the other hand if two views of the scene are too different, the method can fail. It is usually caused by significant movement and rotation of the camera or by reflection. Second problem is detection of homogeneous planes without texture because comparison of the points on such planes is very inaccurate. Therefore it is necessary to verify the reconstructed planes and to combine plane-sweep method with RANSAC.

4.2.1. Plane-sweep

In several next sections we introduce the plane-sweep method and we show the extension of this method by searching the boundary of the plane. First the position of the plane is searched along each reconstructed 3D line. All possible positions along each line are explored and optimal angle for position of plane is chosen. If the line at infinity is used for searching, then all parallel planes in given direction are searched.

(a)   (b)

(c)

(d)
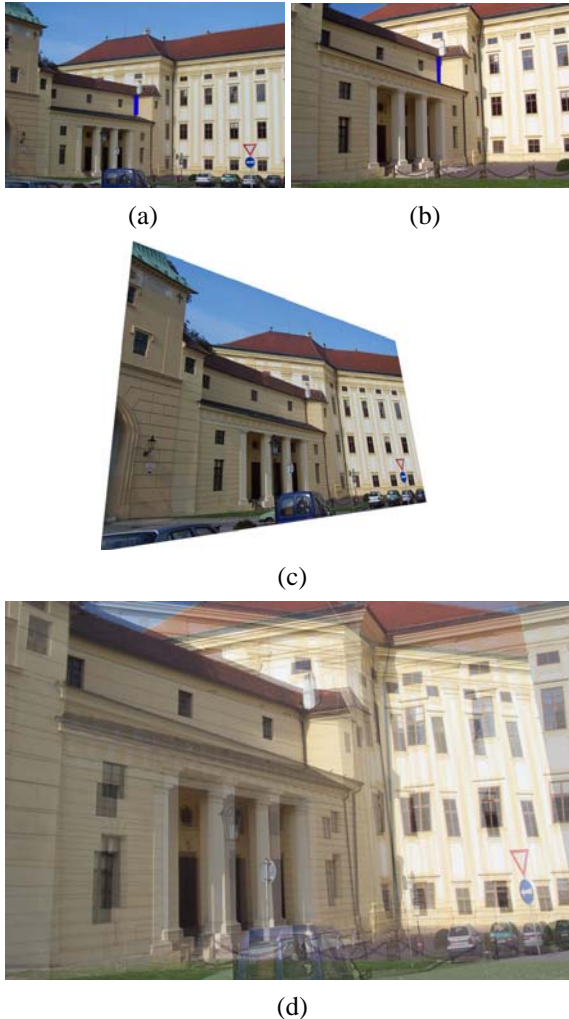
Figure 2. Plane-sweep applied along the line which is marked by blue color in the image (a) and (b). Picture (c) was created by application of homography on image (a). This homography correspond to real position of the plane on the left of the blue line. Image (d) is obtained as a picture (b) which is superimposed by the picture (c). It can be seen that only this plane is "focussed".

The position of the new plane is detected by similarity measure between first image and second image, which is transformed by homography. Homography is point-to-point mapping between the images and it is induced by the real 3D position of the plane. So the homography determines the correspondence between the points and their local neighbourhood is compared and correlation of the points is computed. The sum of the correlation scores is compared for all possible directions and plane is placed to that position, where the similarity score function has the maximum. We use the same similarity function as a Baillard in [Bai99a] and we use only key points for similarity computation too. These key points are obtained by applying an edge detector with very low threshold on the gradient.

Boundary Detection

One of the most serious problems in plane-sweep method is a lack of reconstructed 3D lines. It is necessary to identify the whole boundary for each plane and it means detection of all visible lines and the estimation of some occluded lines in the scene.

Usually only about 30-50 % of the important feature lines is found for average intended scene during the lines reconstruction. Edge detector can't find all borders, because they can be blurred or obscure. Moreover some borders are occluded on some photos so the reconstruction of the line can't be done. We show the method for detection of these missing lines.

4.3.1. Searching for new lines by planes intersections

If the line can't be reconstructed by the trifocal tensor, the simplest and most precise method is to find it as an intersection of two planes. This method was first used in [Bai99a].

### 4.3.2. Creation of boundary from similarity values

If occlusion didn't occur and all planes are sharp, then it is possible to find all missing lines as the plane intersection. However occlusion and presence of blurred planes makes the detection of lines more complicated. In this case the only possibility is to obtain the line from correlation values. It is obvious that this procedure produces lines with lower accuracy.

The plane-sweep method is very sensitive to displacement of the lines, and the lower accuracy of line can cause failure of the next plane detection. However it is not significant problem, because lines detected only from correlation values lie on planes, where the next plane is occluded or unreconstructable. So the plane can't be found even along the accurate line.

4.3.2.1 Computation of rough border

When we want to estimate the boundary of plane, it is necessary to include correlation scores of all pixels (or regular samples of pixels) in computation. For each pixel the weight values are computed from the gradient of the pixel and it defines its significancy. The key points have the weight equal to 1 and for the pixels in homogeneous region the weight tends to zero.

Then the image of correlation values is segmented into the parts of plane and remaining parts. Before the segmentation it is necessary to smooth the correlation values. The thresholding with iterative threshold selection is used for segmentation and then rough borders between regions are computed. If the closed boundary inside the plane is found, it is excluded out of the plane.
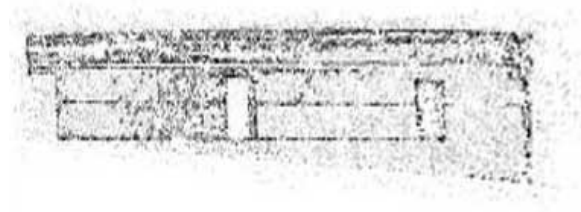
Figure 3. Correlation values for identified plane are represented by pixel darkness. The magnitude of the similarity scores is modified by the gradient of the pixel.

4.3.2.2 Computation of exact position of the border
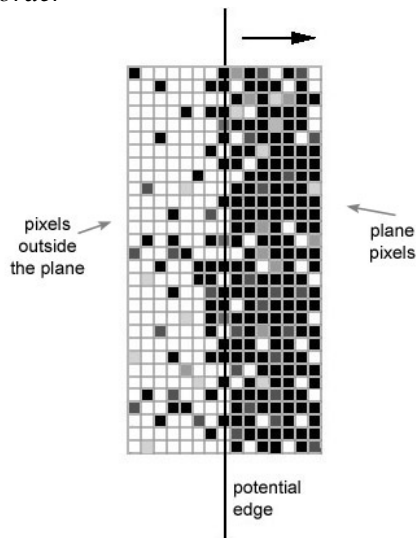
Figure 4. Searching for the optimal edge, white points are pixels with zero correlation; black points are pixel with maximal correlation.

Found borders are usually too thick so it is necessary to compute their exact position. If the real line (reconstructed by trifocal tensor) exists in this thick edge, it is set as a border of the actually reconstructed plane. If such line isn't available, its position must be computed from the correlation values only. The thick border is divided into straight segments and in each segment the optimal position of the border is computed. We choose the outline which divides the local neighborhood in the best way into a region densely covered by points with high correlation and a region with low correlation values. Lines are searched in the direction parallel to the actuallly tested border segment.

Function C_E is maximized over all position of lines:

$$C_E = \left(\sum_{Cor(P_i) \geq T} W\big(Cor(P_i)\big) + \sum_{Cor(O_i) < T} W\big(T - Cor(O_i)\big) \right)$$
$$-$$
$$\left(\sum_{Cor(P_i) < T} W\big(T - Cor(P_i)\big) + F \sum_{Cor(O_i) \geq T} W\big(Cor(O_i)\big) \right),$$

where P_i are pixels on plane, O_i are pixels outside the plane, $Cor(X)$ is the correlation value of the point, T is the threshold, which divides pixels to high correlated and low correlated and W is weight of pixels, which depends on the gradient. Because minimizing of the outliers has a higher precedence than maximizing the high correlated plane pixels, the weighting coefficient F is used to tune the algorithm; typical values are $F \in [1, 2]$.

Problems can arise if the actually reconstructed plane forms angle near 180° with the adjacent plane or if no keypoints are present near the border. In this situation the boundary of region can be detected wrongly.

Verification of the line position

The position of each border line is verified by its re-projecting to the images. The level of gradient is explored and the border line is accepted or rejected. Simultaneously the edges are divided into several types (see section 4.5.1) because the correct classification of edges is essential for the reconstruction of the occluded parts.

If the level of gradient for some edge is low in all images, it is possible to search the edge in the next N suboptimal solutions of cost function C_E. If no other suitable border is found, the first detected line is used.

Occlusion Detection

During the reconstruction of the planes we often meet the situation where some border line belongs only to one plane or some lines are not connected to other lines. It can be caused by various reasons; the edge detector doesn't find the border line, the occlusion appears or the plane fitting algorithm fails. Another eventuality can be presence of non-planar object in the scene. For the completion of „invisible" parts it is necessary to classify all edges to several groups.
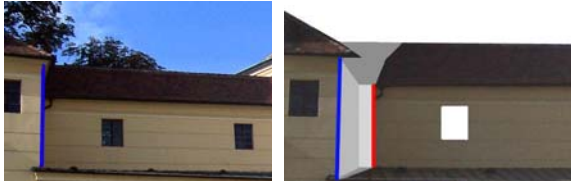
(a)   (b)

Figure 5. Picture (a) is view with the occlusion; in picture (b) the occluded parts are completed. The blue colour denotes self-occluded edge and red is false edge. These two lines have same image in picture (a). The red edge is computed from two photos during boundary estimation.

4.5.1. Edge classification

We can distinguish tree main types of edges:

Normal edge – It creates boundary for more than one plane, so the occlusion doesn't occur in the neighbourhood of this edge.

Self-occluded edge – If one plane occludes the other plane and these two planes are connected by common edge, we denote this edge as **self-occluded**. We can easy recognize this type of edge because during reconstruction it belongs only to one plane and when the edge is verified, it is displayed in more than one image. For estimation of the plane behind the self-occluded edge the assumption on the continuous shape can't be used, because the shape of the surface is usually changed abruptly on this edge.

False edge – We can recognize the unreconstructed parts of the plane by the presence of the **false edges.** False edges don't really exist in the scene; their images in photos are caused by the occlusion. All false edges are computed only from correlation values. The edge is classified as false if it has image in at most one photo. The false edges can't have same position in more then one photo because the location of the edge depends on direction of camera, so it is necessary to exclude too similar views from edge classification.

If the occlusion was caused by other parts of the scene, then the image of false edge is coincident with some self-occluded edge in just one photo. If false edge hasn't corresponding self-occluded edge then its presence could be caused by:

- **Covering by unreconstructed objects** – for example the buildings can be covered by branches of a tree. This results in the holes in the scene which are bounded by irregular false edges without corresponding self-occluded edge.

- **Errors in reconstruction** – some parts of the planes are not reconstructed, because the plane detection algorithm failed.

If we can't reconstruct the object in the foreground, evidently it is impossible to distinguish two last situations – covering by unreconstructed object and errors in the reconstruction.

Algorithm for Reconstruction of Piecewise Planar Scene

First, the empty list **P** of planes in the scene is created. Each plane contains reference to the lines, which create the boundary of the plane. Edges are stored in four lists:

- E_{single} contains normal edges, which don't belong to any plane.

- E_{normal} contains edges, which are assigned to more then one plane.

- E_{false} contains false edges.

- $E_{self-occ}$ contains self-occluded edges.

4.6.1. Algorithm

1. Metric 3D reconstruction of the points and lines is computed.

2. The list of all reconstructed lines E_{single} is created, lines are sorted descending by length and the lines at infinity are inserted in the first positions of the list.

3. The first edge from E_{single} list is taken and the line is searched along this edge (section 4).

   - If there isn't any plane, the edge is removed from E_{single} list and if the line isn't at infinity, it is moved to the $E_{self-occ}$ list. New plane along the next edge is searched.

   - If the plane is found, it is added to P list. If number of plane for actual edge is equal to 2 and if line isn't at infinity, then the edge is added to the E_{normal} and removed from E_{single} list.

4. The approximate positions of the edges are found (section 4.3.2.1). If some existing lines

have suitable positions then they are chosen as the border lines and the shape of boundary and number of plane for the edge is corrected.

5. If the boundary of the plane isn't closed, the intersection with other plane is searched. If new line is found, it is added to E_{single} list and the boundary of these planes is corrected by this edge. If the E_{single} list isn't empty, go to 3.

6. If E_{single} list is empty, missing borders are computed for each plane in **P**, which doesn't have closed boundary. Position of these borders is computed from the correlation values (sections 4.3.2.1, 4.3.2.2). Again the next planes are searched along each new border line. The type of the line is recognized (section 4.5.1).

7. If no more edges can be found, then the reconstruction process is finished. Now the occluded and unreconstructed surfaces must be filled.

5. OCCLUSION RECOVERY

The way of completing covered parts is different for the object which is occluded by the other object or by itself. Further it depends on the shape of the neighbourhood of holes.

If the occluded parts are too complex, it is often impossible to recovery the shape. But it's not surprising because only very little information is available about the occluded part.

Searching for Correct Boundary

(a)   (b)
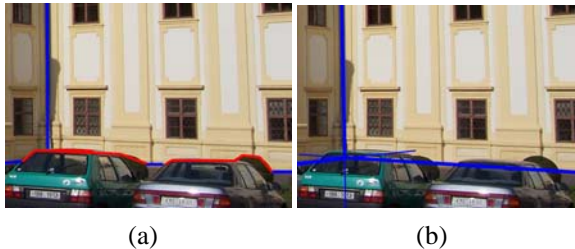
Figure 6. In picture (a) the planes have boundary from normal (blue) and false (red) edges. In picture (b) the false edges are removed and normal edges are extended over the gap and corners are found.

The most common situation is incomplete boundary of plane, where some border edges are false or missing. If the border lines are false, then the real existing lines must be extended over the gap. Sometimes new border line must be added using geometric assumptions about the scene. Finally the false edges are removed from the boundary.

Searching for New Planes

If we have two self-occluded edges with common point, then these two edges determine the position of the occluded plane.

Completion of the plane is more difficult if we have only one self-occluded edge for identification of the plane position. In this situation the coordinates of the plane must be determined only from geometric assumptions (usually the perpendicularity, symmetry or parallelism). Remaining border lines of new plane are computed by intersection with other planes or again by geometric assumtions.

Searching for New Planes without any reconstructed line

If some small wedge or protruded objects appear on the larger plane (for example the window on the wall), then it is usually imposible to reconstruct any line of this object. This situation is recognized as a gap in the plane, which is bounded by normal and self-occluded edges (wedge objects) or normal and false edges (protruded objects). Without at least one reconstructed 3D line of new plane we can't use the plane-sweep method for plane detection, so we only use the geometric assuptions and the new plane is searched as a parallel in the set of reconstructed 3D points. Other small planes, which connect the large plane and its gap, are perpendicular to them.

Algorithm for occlusion recovery

1. For each plane all identified edges (both inner and outer border lines) are revised:

   • The sequence of edges is checked.

   • If the sequence is corrupted or if boundary contains false edges, then real edges must be extended and false edges are removed.

   • If the closed boundary is created only by false edge, it can be removed from the reconstruction.

2. New planes along the self-occluded edges are searched and for each new plane existence of intersections with other planes is tested. If the intersection is found, then the new edge is added.

3. For planes with the gaps the best shape is estimated for wedge and protruded object.

4. All planes without closed outer boundary are removed from the reconstruction.

5. If the occluded object was displayed in one image, then the texture can be extracted. If this part of the scene doesn't have an image, the surface is textured by blurred texture of the neighbourhood of the gap.

6. CONCLUSIONS

We have presented a method for occlusion detection and recovery in reconstruction of piecewise planar scene. The method is based on fitting planes in the cloud of reconstructed points and lines. If there isn't enough information for fitting the plane in common way, then the type of occlusion is estimated and the most suitable shape of the plane is determined according to neighbourhood of the plane and assumptions on the scene.

The big advantage of this method is ability to estimate occluded surfaces in full automated reconstruction without any high-level information about the scene. However, there are still limitations and some parts of method must be improved. Indeed, the method strongly depends on border detection and occlusion classification, but these parts of method can be unreliable due to lightning condition, weak contrast etc. We still work on improvement of these parts of algorithm, but the problems with bad light condition, contrast, shadows or reflections are common in all computer vision applications and they can't be solved definitively.

7. FUTURE WORK

We have focused only on the 3D reconstruction of piecewise planar scene, however various non-planar objects occur frequently in the man-made environments. If such an object is in the scene, the reconstruction will be incorrect. So we plan to extend this approach by reconstruction of arcs and some curved surfaces.

Further, system of geometric constraints can be extended by new, user defined information. These constraints could be added before the reconstruction or interactively in the process of reconstruction, when the algorithm requires additional information for ambiguous surfaces.

Algorithm for detection of plane shape is reliable for textured planes, but if the texture is missing then the borders are inaccurate. This problem can be partially solved by the extension of algorithm by texture and color analysis. Results of this analysis can then be used for chaining of the planes and polygons, for detection of the objects outside the plane and for the better estimations of the texture of occluded surfaces.

8. ACKNOWLEDGMENTS

9. REFERENCES

[Bai99a] Baillard, C. and Zisserman, A. Automatic reconstruction of piecewise planar models from multiple views. Conference on Computer Vision and Pattern Recognition, pp. 559-565, June 1999.

[Cas02a] Castellani, U., Livatino, S., and Fisher, R. B. Improving environment modelling by edge occlusion surface completion. Int. Symp. on 3D data processing, visualization and transmission (3DPVT), Padova, Italy, pp. 672-675, 2002.

[Del01a] Dell'Acqua, F. and Fisher, R.B: Reconstruction of surfaces behind occlusions in range data. Conf. on 3D digital imaging and modeling (3DIM), Montreal, Canada, pp. 232-239, 2001.

[Dic01a] Dick, A. R., Torr, P. H. S., Ruffle, S. J. and Cipolla, R. Combining single view recognition and multiple view stereo for architectural scenes. In Int. Conf. Computer Vision, pp. 268-280, 2001.

[Fis81a] Fischler, M.A. and Bolles, R.C. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM, 1981

[Har03a] Hartley, R., Zisserman, A. Multiple View Geometry in Computer Vision, Cambridge University Press, ISBN: 0521540518, 2003.

[San99a] Sanchiz, J.M., Fisher, R.B. A next-best-view algorithm for 3D scene recovery with 5 degrees of freedom, BMVC, pp. 163-172, 1999.

[Wer02a] Werner, T. and Zisserman, A. New techniques for automated architectural reconstruction from photographs. European conference on computer vision, Copenhagen, Denmark, pp. 541-555, 2002.

[Wer02b] Werner, T. and Zisserman, A. Model selection for automated reconstruction from multiple views. British Machine Vision Conference, Cardiff, UK, pp. 53-62, 2002.