Preprocessing of region of interest localization based on local surface curvature analysis for three-dimensional reconstruction with multiresolution

Wanjing Li, Rainer Schutze, Martin Böhler, Frank Boochs, Franck Marzani, Yvon Voisin

To cite this version:


HAL Id: hal-00637703
https://hal.archives-ouvertes.fr/hal-00637703
Submitted on 3 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Preprocessing of region of interest localization based on local surface curvature analysis for three-dimensional reconstruction with multiresolution

Wanjing Li
University of Burgundy
UFR Science and Technology
Laboratory LE2I BP 47870
UMR 5158
Dijon Cedex, 21078
France

and

University of Applied Sciences
Laboratory i3Mainz
Lucy-Hillebrand-Strasse 2
Mainz, 55128
Germany

Rainer Schütze
Martin Böhler
Frank Böochs
University of Applied Sciences
Laboratory i3Mainz
2 Lucy-Hillebrand-Strasse
Mainz, 55128
Germany

Franck S. Marzani
University of Burgundy
UFR Science and Technology
Laboratory LE2I BP 47870
UMR 5158
Dijon Cedex, 21078
France
E-mail: franck.marzani@u-bourgogne.fr

Yvon Voisin
University of Burgundy
Laboratory LE2I BP16
UMR 5158
Route des Plaines de l’Yonne
Auxerre Cedex, 89010
France

Abstract. We present an approach to integrate a preprocessing step of the region of interest (ROI) localization into 3-D scanners (laser or stereoscopic). The definite objective is to make the 3-D scanner intelligent enough to localize rapidly in the scene, during the preprocessing phase, the regions with high surface curvature, so that precise scanning will be done only in these regions instead of in the whole scene. In this way, the scanning time can be largely reduced, and the results contain only pertinent data. To test its feasibility and efficiency, we simulated the preprocessing process under an active stereoscopic system composed of two cameras and a video projector. The ROI localization is done in an iterative way. First, the video projector projects a regular point pattern in the scene, and then the pattern is modified iteratively according to the local surface curvature of each reconstructed 3-D point. Finally, the last pattern is used to determine the ROI. Our experiments showed that with this approach, the system is capable to localize all types of objects, including small objects with small depth. © 2009 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3156731]

Subject terms: intelligent 3-D scanner; region of interest (ROI) localization; surface curvature; adaptive pattern.

Paper 090178R received Mar. 13, 2009; revised manuscript received Apr. 27, 2009; accepted for publication Apr. 28, 2009; published online Jun. 23, 2009.

1 Introduction

Today, 3-D scanners are widely used for environment digitalization, such as the interior and exterior of buildings and historical heritage. To our knowledge, all of them scan the whole scene with uniform resolution. In the case of interiors of buildings, often the ceilings and the walls are rather flat, with occasionally some objects (such as lamps, picture frames, electric switches, windows, pipes) on them. Therefore, different approaches were proposed to simplify the representation of walls and roofs. Cruz et al.\(^1\) proposed integrating semantic knowledge into a 3-D point cloud. Johnston and Zakhor\(^2\) proposed scanning from the outside of buildings to get a 3-D point cloud, and then applying plan-fitting algorithms to estimate interior ceilings and walls. In fact, all these approaches can be regarded as a two-step 3-D reconstruction process: first, scan the building with high and uniform resolution; second, simplify the 3-D model.

We tried to solve the problem in another way—that is, to make the scanner scan the whole scene with different res-
olution for different parts. To achieve this purpose, we propose a preprocessing process of region of interest (ROI) localization, which can return two types of information: a point cloud representing the whole scene with multiresolution, and the positions of the ROI in the scene. The latter can then be used by the scanner so that the scanning would be done only in these regions during the acquisition process. By combining the point cloud issued from the preprocessing process and the one from the acquisition process, a point cloud with multiresolution can be obtained. Since during the acquisition process, the scanning is limited only in the ROI, compared to the classical way, which consists of scanning the entire scene, the whole scanning time would be largely reduced, and the obtained point cloud would contain much less data.

In this paper, we present an approach of integration of region of interest (ROI) localization into a 3-D scanner as preprocessing phase. In the frame of our work, an ROI stands for all the regions with high surface curvature, such as the border of an object or nonflat part of an object. To study the feasibility and efficiency of this approach, we tested it under an active stereoscopic system composed of two cameras and a LCD projector. At first, a regular point pattern is projected on the scene to get a rough 3-D point cloud representing the scene. The local surface curvature (LSC) of each 3-D point is then analyzed, and a new pattern is created automatically to project more points in the regions with high surface curvature. In this way, more 3-D points are retrieved in these regions and are added into the current 3-D point cloud. The iterative pattern adaptation process continues until the pattern reveals the best the ROI.

The remainder of this paper is organized as follows: Section 2 describes in detail the ROI localization process based on an iterative pattern adaptation process; the ROI determination method and results are presented in Sec. 3; and last, we draw conclusions in Sec. 4.

2 ROI Localization Process

2.1 System Initialization

Since the preprocessing approach was validated under an active stereoscopic system composed of two CCD cameras and one video projector, the cameras and the projector need to be calibrated to estimate their intrinsic (focal length, pixel size, etc.) and extrinsic parameters (the position and orientation related to the scene). These parameters are necessary to calculate the 3-D position of a point in the scene from images. For the calibration of the two cameras, we used self-calibration methods. As to the video projector, we applied the method of Faugeras.

The ROI localization begins by gathering sample points in the scene. A simple way is to project a regular point pattern via the video projector. Several factors, such as the size of the object and the distance between the scene and the projector, need to be taken into account to adjust the density of the pattern points so that at least one point will fall on each object. Otherwise, it may be impossible to localize the object. The two precalibrated cameras then take a pair of images of the scene; the 3-D coordinates of projected points in the scene can be calculated by using the technique of triangulation.
Vertex $V_i$ is considered as a neighbor of vertex $V$ if the edge $VV_i$ belongs to the mesh. Denote the set of neighbors of $V$ by $\{V_i| i=1, 2, \ldots, n\}$, the set of triangles containing $V$ by $\{T_i=\Delta(V_i, V, V_{(i+1)mod\ n})| i=1, 2, \ldots, n\}$, and the set of angles between $V$ and its two successive neighbors by $\{\alpha_i = \angle(V_i, V, V_{(i+1)mod\ n})| i=1, 2, \ldots, n\}$ (see Fig. 1).

The following Gauss-Bonnet–based formula is called the classical formula for estimation of discrete Gaussian curvature $K$ at vertex $V$ (Refs. 10–12) where $A_i$ is the area of triangle $\Delta(V_i, V, V_{(i+1)mod\ n})$:

$$K = \frac{2\pi - \sum_{i=1}^{n} \alpha_i}{\prod_{i=1}^{n} A_i}.$$

(1)

Instead of calculating the “real” area of each triangle, Meyer et al.\(^9\) consider that of the Voronoi region of each triangle, denoted by $A^{Mixed}$:

$$K = \frac{2\pi - \sum_{i=1}^{n} \alpha_i}{\sum_{i=1}^{n} A^{Mixed}}.$$

(2)

A simpler approach can be applied by ignoring the areas of the triangles,\(^8,13\) which is also called angle deflection:

$$K = 2\pi - \sum_{i=1}^{n} \alpha_i.$$

(3)

**Edge-based LSC estimation.** For a smooth surface patch, the Gaussian curvature provides reliable information for our application. However, in the case of a folded paper neighborhood, i.e., if a vertex has two neighbors on a same edge and others on two planes, it has a value of zero (see Fig. 2). In this case, the Gaussian curvature is not appropriate for our application. We therefore proposed an approach based on edges for the LSC analysis.

For each triangle, we calculate its normal vector (often simply called normal, and commonly denoted $N$). A plane passing through three points $\{(x_i, y_i, z_i)| i=1, 2, 3\}$ can be defined by the following equation, where $N=[a\ b\ c]^T$ is the normal vector:

$$ax + by + cz = d.$$

For each triangle, we calculate its normal vector $N$ of the triangles,\(^8,13\) which is also called normal vector.

$$\text{Fig. 1} \text{ Vertex } V \text{ and its neighborhood in 3-D Delaunay triangular mesh.}$$

![Fig. 1](image1.png)

$$\text{Fig. 2} \text{ The type of surface patch for which the Gauss-Bonnet scheme does not work.}$$

![Fig. 2](image2.png)

At vertex $V_i$, the minimal area of triangle, because if a vertex $V$ is situated either at the border of an object, or on the high surface curvature part of the object. In addition, it is necessary to define another threshold $A_{min}$ the minimal area of triangle, because if a vertex $V$ is situated at the border of an object, it may always have strong surface curvature. Therefore, it is necessary to determine at which moment the vertex $V$ should be ignored for new pattern design.

**2.2.3 Selection of candidate vertices**

A vertex $V$ of the current 3-D triangular mesh is selected as a candidate vertex for the design of a new pattern if the following two conditions are both satisfied:

1. $K > \tau$.
2. $A_i > A_{min}$ ($i=1, 2, \ldots, n$).

$\tau$ and $A_{min}$ are predefined thresholds. The use of high LSC threshold $\tau$ is essential, since it is necessary to project more points around a vertex only if it has strong surface curvature; in other words, if it is situated either at the border of an object, or on the high surface curvature part of the object. In addition, it is necessary to define another threshold $A_{min}$ the minimal area of triangle, because if a vertex $V$ is situated at the border of an object, it may always have strong surface curvature; therefore, it is necessary to determine at which moment the vertex $V$ should be ignored for new pattern design.
2.2.4 Design of new pattern
The new pattern designed for the next iteration is at first initialized as an image with black background, and then certain white pattern points are added in the image at specific positions.

In the case of the discrete Gaussian curvature estimation approach, for each candidate vertex \( V \) obtained previously, its corresponding pattern point \( P \) and those of all its neighbors are found by using the calibration parameters of the video projector; four white points are then added in the new pattern around the position of \( P \) at distance \( d/H^2 \) (see Fig. 3). To ensure that the pattern points fall closer and closer to the ROI, the current neighborhood information of the point \( P \) must be taken into account to determine the value of \( d \). We therefore set \( d = D/3 \), where \( D \) is the average distance between the pattern point \( P \) and all its neighbors based on the current 3-D mesh.

In the case of the edge-based LSC estimation approach, for each candidate edge \( E \), as shown in Fig. 4, the corresponding pattern points of the four vertices of the two adjacent triangles can be determined by using the calibration parameters of the projector; the positions of the two new pattern points are given by each triangle centroid (in black).

2.2.5 Stop condition
In some cases, by analyzing the distribution of pattern points, it is possible to localize the ROI at the second iteration. However, several more iterations are often necessary. Our new pattern design method ensures that the pattern points fall closer and closer to the ROI after each iteration; therefore, the triangles around these regions in the obtained 3-D triangular mesh become smaller and smaller. When the area of the smallest triangle is smaller than the predefined threshold \( A_{\text{min}} \), even if the vertex \( V \) still has strong surface curvature, it will not be selected as a candidate vertex for the new pattern design. In this way, for whichever LSC estimation approach, from the second iteration of the pattern adaptation process, the number of pattern points keeps increasing for several iterations, until it reaches a maximum. The iterative process stops then, and this last pattern is selected for ROI localization.

2.3 ROI Localization
In the context of image processing, morphology stands for a set of operations that process shape-based images. Morphological operations apply a structuring element \( S_e \) to an input image, with its center over each pixel, and create an output image of the same size. The value of each pixel in the output image is determined by both the corresponding pixel in the input image and its neighbors defined by \( S_e \). By choosing the size and shape of the neighborhood, it is possible to construct a morphological operation that is sensitive to specific shapes (such as line, disk, rectangle, and square) in the input image.

Dilation and erosion are two fundamental morphological operations. Dilation adds pixels to the boundaries of objects...
in the output image, while erosion removes pixels on object boundaries. The number and positions of pixels added or removed from the objects in an image depends on the size and shape of the applied structuring element $S_e$.

3 Results

It is well known that the smaller an object is, the harder it is to localize it. We therefore built a synthetic scene [see Fig. 5(a)] for the efficiency evaluation of our approach. In this scene, three small objects of different forms and sizes are situated at different positions and orientations. Object $A$ is a piece of coin (diameter: 21 mm, depth: 2 mm); object $B$ is a small box (length: 51 mm, width: 22 mm, depth: 10 mm); and object $C$ is a 5-in battery (length: 47 mm, diameter/depth: 14 mm).

We also tested our approach on a real scene—part of a wall in our office. The objective is to localize the two pipes whose diameter is respectively 16 mm and 9 mm and the handle of a curtain [see Fig. 5(b)].

For the synthetic scene, with the same initial pattern, we applied respectively in the pattern adaptation process the four preceding formulas: classical, Meyer’s, angle deflection, and dihedral angle. Figure 6 shows the scene at the

![Figure 5](image1.png)

Fig. 5 Two scenes: (a) synthetic scene with different types of objects; (b) real scene with a curtain handle and two pipes of different orientation.

![Figure 6](image2.png)

Fig. 6 Images of the synthetic scene acquired by the left camera, at the 2nd, 3rd, and 4th iterations: (a) with the classical formula, $\tau = 3 \times 10^{-4}$; (b) with Meyer’s formula, $\tau = 3 \times 10^{-4}$; (c) with the angle deflection formula, $\tau = 0.35$; (d) with the dihedral angle formula, $\tau = 20$ deg.
second to fourth/fifth iterations, by applying each formula respectively.

It is easy to notice that with the classical formula and Meyer’s formula, there is no difficulty recognizing the borders of objects. However, it is hard to recognize the flat part of objects. Actually, more and more points are projected on the objects, despite the geometric characteristics of the object surface. This is caused by the fact that the classical formula and that of Meyer take into account the areas of the surrounding triangles. Therefore, for a given vertex with weak LSC, if at the next iteration its neighborhood changes because of other vertices, its estimated LSC may become higher since the areas of its surrounding triangles decrease. As a result, these two approaches are considered inappropriate for our application.

Starting from the same initial pattern, by applying the angle deflection formula and dihedral angle formula, better results are obtained. However, compared to the angle deflection formula, the dihedral angle formula presents more advantages: the pattern adaptation is more precise, and much fewer points were projected (see Table 1). This can be explained by our new pattern design strategy: for the Gauss-Bonnet–based LSC estimation approaches, since the estimation is based on vertices, in order to catch the ROI, it is necessary to add at least four points around each candidate vertex, whereas for the edge-based approach, only two points need to be added, since the two adjacent triangles associated to the candidate edge reflect the local ROI. Therefore, compared to the Gauss-Bonnet–based LSC estimation approaches, the dihedral angle formula leads to less pattern points.

Another advantage presented by the dihedral angle formula is that it works for all types of neighborhood, whereas the angle deflection formula does not work for a folded paper neighborhood. This was proven by some experiments on the real scene presented previously. We can see from Fig. 7(b) that with the dihedral angle formula, we encountered no difficulty localizing the handle and the two pipes, whereas with the angle deflection formula, the vertical pipe is ignored [see Fig. 7(a)].

The reason is illustrated in Fig. 8. Actually, in the initial 3-D triangular mesh, the vertices representing the small pipe have the folded paper neighborhoods. By using the angle deflection formula, they all have zero LSC; therefore, no point was added around them during the rest of the iterative pattern adaptation process.

Figure 9 shows respectively, for the synthetic and real scene, how the number of pattern points evolves during the pattern adaptation process by using the four LSC estimation approaches. We finally abandoned the classical formula and

<table>
<thead>
<tr>
<th>Table 1 Number of points projected at each iteration by using the four LSC estimation approaches.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Iteration</strong></td>
</tr>
<tr>
<td>Classical</td>
</tr>
<tr>
<td>Meyer’s</td>
</tr>
<tr>
<td>Angle deflection</td>
</tr>
<tr>
<td>Dihedral angle</td>
</tr>
</tbody>
</table>

![Fig. 7](image1.png)  
**Fig. 7** Images of the real scene with two pipes and a curtain handle, acquired by the left camera, at the 2nd, 3rd, and 4th iterations: (a) with the angle deflection formula, $\tau$=0.35; (b) with dihedral angle formula, $\tau$=20 deg.

![Fig. 8](image2.png)  
**Fig. 8** The initial 3-D triangular mesh of the scene, with the vertical pipe in big zoom.
Meyer’s formula at the end of the third iteration for the synthetic scene, because the number of pattern points increases too much, and they do not reveal pertinently the ROI. Therefore, they were not tested for the real scene. By using the angle deflection formula and dihedral angle formula, however, for both the synthetic and the real scene, we can see that the number of pattern points increases from the second iteration, until a maximum value is reached at the

![Fig. 9](image1.png)  
**Fig. 9** The curve of the pattern point number during the pattern adaptation process: (a) the synthetic scene; (b) the real scene.

![Fig. 10](image2.png)  
**Fig. 10** Projected patterns for the synthetic scene with the dihedral angle formula: (a) 2nd to 4th iterations; (b) 5th to 7th iterations.
fifth iteration, and then the number drops down. Therefore, for both the synthetic and the real scene, it was the pattern image at the fifth iteration that was selected for the ROI localization.

This decision is supported by comparing visually all the pattern images projected during the iterative process. As an example, Fig. 10 lists the patterns projected from the second to the seventh iteration for the localization of the ROI of the synthetic scene, by using the dihedral angle formula. It is obvious that at the fifth iteration, when the pattern points’ number reaches to the maximum value, the distribution of pattern points best reveals the ROI: both borders and nonflat parts of objects are indicated. The same phenomenon was observed by using the angle deflection formula. We therefore always select the pattern with the maximum number of points. By applying several morphology operations, the ROI can be marked out.

Figure 11 shows the results of ROI localization for both synthetic and real scenes, by applying a structuring disk respectively on four selected pattern images. Obviously, by applying the dihedral angle formula, the ROI was better localized.

4 Conclusions
We presented a concept of integrating a preprocess of ROI localization into a 3-D scanner, and we studied its feasibility and efficiency under an active stereoscopic system.

The core of the ROI localization is the iterative pattern adaptation process. The local surface curvature estimation plays a very important role during the whole process. We have tested three Gauss-Bonnet–based Gaussian curvature estimation approaches, among which the angle deflection formula led to the best results. However, in the case of a folded paper neighborhood, it does not work well. We therefore proposed an edge-based LSC estimation approach: the dihedral angle formula. Our experiments showed that this method works well for all types of 3-D triangular mesh. Compared to the angle deflection formula, the dihedral angle formula presents other advantages: during the iterative pattern adaptation process, fewer points are projected, so less time is needed; and the pattern-based ROI localization is much precise.

This concept is especially useful for the scanning of building interiors, because the ceilings and the walls are often flat, and it is therefore a waste of time to scan them with high precision. However, if there are some objects on the wall, it may be necessary to scan the objects precisely. Our solution consists in integrating the ROI localization approach into a 3–D scanner as preprocessing process, which would help the scanner to localize rapidly the objects, so that precise scanning would be done only in these regions instead of in the entire scene. Thus, the obtained 3-D point cloud would be multiresolution, which represents pertinently the scene.

Acknowledgments
This work is supported by the European Social Funds, University of Applied Sciences in Mainz (Germany), and the...
Regional Council of Burgundy (France). The contribution of Mr. Burkhard Tietz (member of i3mainz) to our work is highly appreciated.

References


Wanjing Li is currently a PhD student at both the University of Burgundy in France and the University of Applied Sciences in Mainz, Germany. Her research work focuses on the stereovision system, 3-D scene reconstruction, and structured light.

Rainer Schütze received his diploma in geoinformatics from the University of Applied Sciences in Mainz, Germany, in 2005. He now works as a research assistant in the Department of Geo-Information and Surveying at the University of Applied Sciences in Mainz. His main research field is software development in the field of digital photogrammetry with the main focus on bundle adjustment.

Martin Böhler received his engineering diploma in 2003. He has since been working as a research associate at the i3mainz (Institute for Spatial Information and Surveying Technology at the University of Applied Sciences), Germany. Software development in the field of digital photogrammetry and image processing has been his main area of research. His recent work focused on the development and implementation of algorithms for active stereoscopic systems.

Frank Boochs became a member of the scientific staff at the Institute of Photogrammetry in Bonn in 1978 and received his PhD degree there in 1984. His research interests were image matching and remote sensing. In 1987, he became a member of the scientific staff at the Institute for Applied Geodesy in Frankfurt. His work focused on digital image processing with spatial context. Since 1993, he has been working as a professor in the Department of Spatial Information and Surveying at the University of Applied Sciences in Mainz. He is also cofounder of the i3mainz laboratory.

Frank S. Marzani received his PhD degree from the University of Burgundy, Dijon, France in 1998. He is currently an associate professor in computer science at the University of Burgundy and member of the Laboratory LE2i (Laboratoire Electronique, Informatique et Image), UMR CNRS 5158, of the University of Burgundy. His research interests include both multispectral images and range sensing systems by structured light.

Yvon Voisin received his PhD degree in 1993 from the University of Franche Comté, France. He is currently a full professor at the University of Burgundy, France. He is also a member of the Image Processing Group of the LE2i. His research interests are 3-D reconstruction and motion analysis. He is also working on the application of artificial vision in industry.