# Splitting up Panoramic Range Images into Compact 2[1/2]D Representations

**Angel Domingo Sappa**

Computer Vision Center, Edifici O Campus UAB, Bellaterra, 08193 Barcelona, Spain

**ABSTRACT:** This article presents a simple technique for splitting up a panoramic range image into a set of 2[1/2]D representations. The proposed technique consists of three stages. First, a spherical discretization map is generated. Second, main surface orientations are extracted together with their corresponding histogram of distances. Each one of these histograms is used to define the position of a projection plane as well as two associated clipping planes. Finally, data points bounded by clipping planes are mapped onto the corresponding projection plane defining a classical 2[1/2]D range image. This last stage—projection—is applied as many times as main orientations in the spherical discretization map. Experimental results with a panoramic range image are presented. © 2006 Wiley Periodicals, Inc. Int J Imaging Syst Technol, 16, 85–91, 2006; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.20069

**Key words:** range image processing; panoramic images; range image segmentation; 3D data representation

## I. INTRODUCTION

In recent years 3D computer vision has experimented a fast growth. The appearance of new sensors, which allow the obtainment of a large amount of three-dimensional information in a short time, and the need to process and represent these images efficiently, has given rise to new research topics in the 3D computer vision community. One of these topics is the 3D digital representation that has gained an important place in different fields [e.g., architectural (Cho et al., 2001), automotive engineering (Luo et al., 2000), robotics (Reed and Allen, 1997), computer animations (Di Fiore and Van Reeth, 2002), to mention a few]. Unfortunately, every day applications do not experiment the same fast evolution, so that most of the time new technologies need to be adapted in order to coexist with current ones. Moreover, in some cases, new technologies cannot be fully exploited, because end users do not have the right tools for filling the gap existing between new and current technologies. For instance, the classical planar drawings are preferred instead of 3D digital representations. Actually, planar drawings are still nowadays the "de facto" representations for many applications. In industrial environments or building works is not easy to find the appropriate tools to visualize and understand 3D digital representations. In this sense, this article presents a technique to generate automatically planar projections from 3D panoramic range images. It is based on the use of prior knowledge of a scanned scene, and it intends to be the bridge that link 3D digital representations with the classical planar drawing.

Scene prior knowledge has been extensively used for improving 3D representations. For instance, Cantzler et al. (2002a,b), propose to exploit features like parallelism and perpendicularity of walls for improving the structural quality of automatically acquired architectural 3D models. The use of semantic description of general indoor environments (i.e., architectural features as plane walls, ceilings, and floors) is proposed in Nüchter et al. (2003) for 3D indoor environments reconstruction with autonomous mobile robots. Fisher (2002) presents a domain knowledge based technique, which uses standard shapes and relationships, to solve or improve reverse engineering problems. The use of geometrical constraints is also exploited in Dick et al. (2000) for tackling the structure from motion problem. It uses geometric constraints, such as perpendicularity and verticality of walls, which are likely to be found in architecture. In the current work, we also propose to use geometrical constraint in order to split up a given panoramic range image into a set of planar representations. However, instead of enforcing parallelism, perpendicularity, or other kind of hard constraints or relationships between the surfaces [e.g., Dick et al. (2000); Cantzler et al. (2002a,b); Fisher (2002)], the proposed technique only assumes that buildings are defined by a set of planar walls.

Panoramic range images allow capturing the full geometry of big environments, with a high fidelity, in a short time. The required space to store all this information (images bigger than 500 MB) or the CPU power to process all these data is not a problem for the current technology. The only constraint for these panoramic data appears when it is necessary to print or represent all this 3D information in a single snapshot. The solution proposed in the current article is to split up the original panoramic range image into a set of easy to understand 2[1/2]D representations—indistinctly referred in this work as planar or 2[1/2]D representations.

Planar representations can be used not only as a final description but they can also be used for other applications, such as to define the next position of the sensor, in order to scan the remains of the
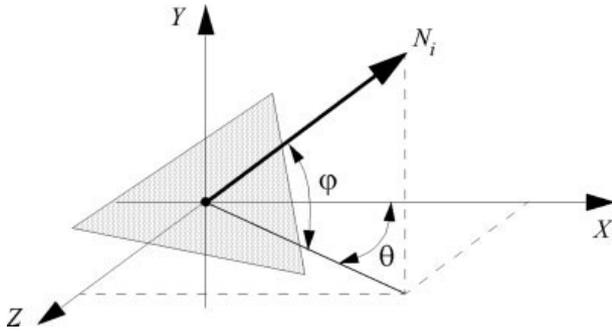
---

**Figure 1.** Sensor coordinate system and normal vector to a given triangle.

environment. This problem, known in the literature as the next-best-view problem [e.g., Reed and Allen (1997); Garcia et al. (1998); Klein and Sequeira (2000)], consists in computing the positions where the range sensor should be placed in order to acquire the surfaces of the objects present in the scene, minimizing the total amount of scans. As it will be illustrated in the experimental results section, the computed planar representations can be used to detect occluded or low-resolution areas, directly defining the next position of the sensor.

The proposed technique consists of three stages. First, a spherical discretization map (SDM) (Garcia et al., 1998) is generated; it will unveil the main surface orientations of the given panoramic scene. Second, main orientations are extracted together with their corresponding histogram of distances. A histogram of distances is computed for each main direction in order to define the position of the corresponding projection plane, as well as its two clipping planes. Finally, data points bounded by a couple of clipping planes are mapped onto the projection plane defining one of the sought 2[1/2]D representation. This last projection stage is applied as many times as main orientations in the spherical discretization map. Section II briefly introduces both, range sensors and images. Section III describes the SDM generation and main orientation extraction. The definition of projection planes and clipping planes, together with the generation of 2[1/2]D representations, are given in Section IV. Section V presents experimental results with an indoor panoramic scene.

## II. RANGE SENSORS AND IMAGES

According to the physical property utilized to generate 3D images, range sensors can be classified as passive or active systems. Passive systems do not interact with the objects of the scene, whereas active systems interact with the objects by projecting some kind of signal (e.g., laser, structured light) over them. The most representative passive 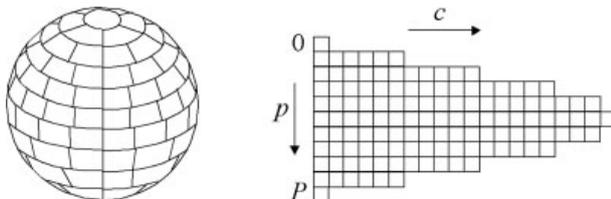sensors are the stereoscopic ones, which generate 3D images by triangulating selected points in a scene viewed at the same time by two conventional cameras (Faugeras, 1993). On the other hand, active systems include a wide variety of different sensors. For instance, some systems work by projecting patterns of structured light onto the scene—grids, strips, and elliptical patterns—and analyzing their corresponding deformation. Knowing the properties of both the camera and the projector, and their positions, the depth can be computed by triangulation [e.g., Fofi et al. (2003); Malassiotis and Strintzis (2005)]. The time-of-flight principle has also been used for developing active range systems. In this case, the sensor emit a light pulse into the scene and measure the distance by analyzing the reflected signal (time elapsed from the emission to the arrival to the sensor). Most of laser range scanners work with infrared lighting [(e.g., Hancock et al. (1998); Wang et al. (2002)].

Traditionally, range images are represented by means of a 2D array $R$, where each element $R(r,c)$ is a scalar representing a surface point of coordinates: $(x,y,z) = (fx(r), fy(c), fz(R(r,c))$ referred to a local coordinate system (Garcia and Sappa, 2003). Therefore, range images are also known in the literature as 2[1/2]D representations since their defining 3D data can be projected into a plane. Having in mind the 2[1/2]D nature of range images a large amount of work has been done during last two decades.

Recently, new technologies have given rise to panoramic range sensors. As can be appreciated in Figure 4, the main drawback of panoramic range images is that there is not a single projection plane. Note that this figure only presents different views of a panoramic range image (i.e., a dense cloud of 3D data points), although visually look like wire-frame representations. Two different options can be adopted for sorting out this problem. On one hand, new techniques, specifically proposed for handling panoramic range images, could be developed [e.g., Hancock et al. (1998); Wang et al. (2002); Hirahara and Ikeuchi (2004)]. For instance, Hancock et al. (1998) propose a robust method to register a panoramic range image and an omni-directional image. It works by matching horizontal and vertical edge histograms of the two images. Another approach for handling panoramic range images have been proposed in Wang et al. (2002). It addresses the problem of automatic reconstruction of real environments. Finally, Hirahara and Ikeuchi (2004) presents a technique for vehicle detection, using panoramic range images, based on cluster analysis. Alternatively, instead of developing new techniques for handling the whole 3D information at the same time, another option could be to split up the original panoramic range image into a set of classical 2[1/2]D range images, so that most of current range image processing algorithms could be directly used. The technique proposed in this work follows this second option.
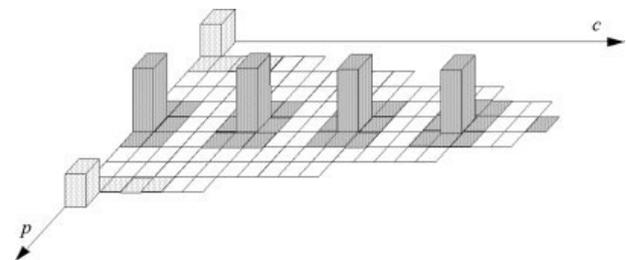


**Figure 2.** Example of a spherical discretization map with 20 cells along the equator, leading to 11 parallels and a total of 126 cells.



**Figure 3.** Illustration of an orientation histogram and main surface orientation extraction.
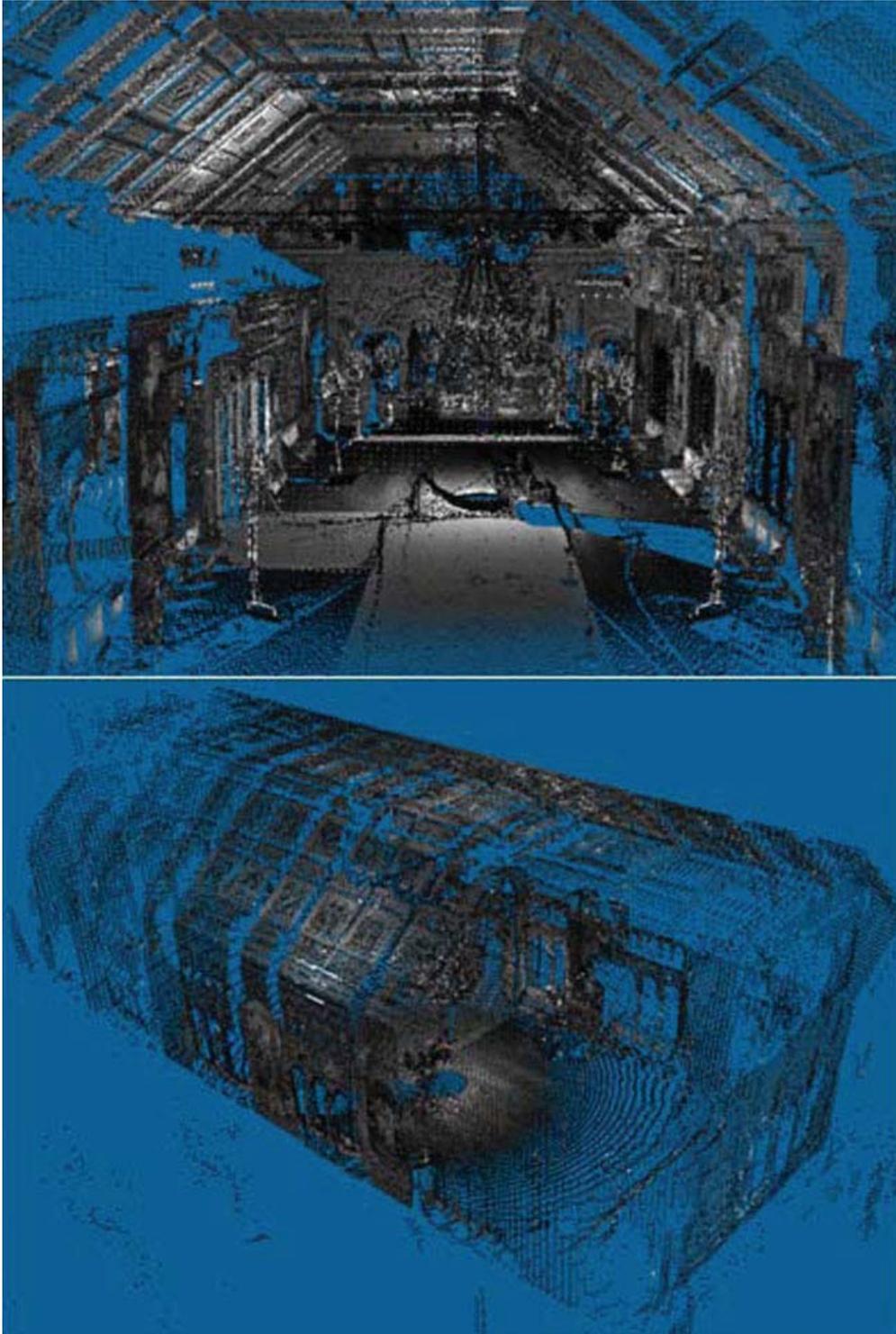
**Figure 4.** Inside view of a panoramic range image (top). Overview of a low resolution representation of the previous panoramic range image (bottom). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com]
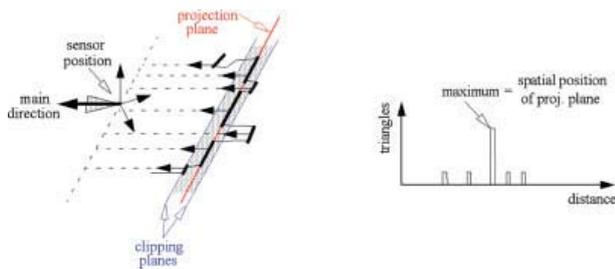
**Figure 5.** Section of surface associated with a main direction (left). The corresponding distance histogram (right). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com]

## III. SURFACE ORIENTATION

This section describes the approach used for computing main surface orientations from a SDM. Let $R(r,c)$ be a panoramic range image with $R$ rows and $C$ columns, $r \in [0,R)$, $c \in [0,C)$; in contrast to classical range images each array element contains three values representing a surface point of coordinates $(x,y,z)$, referred to a coordinate system attached to the sensor. Initially, from the given range image a trivial triangulation is computed by linking all the points horizontally and vertically, and by dividing the obtained cell choosing one of the diagonals (triangles defined by edges longer than a user defined threshold are discarded, considering that they are linking noisy data or a surface discontinuity). Then, a unitary normal vector $N_i$ is computed for each one of the obtained triangles. Additionally, each triangle has associated its distance to the sensor position $\delta_i$. Triangle distance is computed as the average distance of its three defining points.

In a previous version Sappa (2004) surface orientation information is obtained by using two orthogonal histograms, horizontal and vertical histograms. The horizontal histogram, ranging from 0 to 360°, is used to keep the orientation, of a $XZ$ projection (Fig. 1), of every triangle's normal vector defining the surfaces of the scene. On the other hand, the vertical histogram, ranging from 0 to 180°, is used to keep the angle orientation between the $Y$ axis (Fig. 1) and every triangle's normal vector. Although useful, the main drawback of that compact approach appears when the sensor coordinate system is not correctly oriented with the main surfaces of the scene (i.e., the $XZ$-axes of the coordinate system attached to the sensor should be parallel to the floor). Have in mind that the sensor works in a plug-and-play basis; in other words, it does not require an infield calibration process or specific set-up, so a constraint such as parallelism between $XZ$-axes and floor makes it difficult to use the previous proposal (Sappa, 2004). Actually, in most of the cases we do not have control to the sensor set-up during the scanning process. To avoid this problem, speeding up at the same time the set-up process, the use of a histogram from a SDM is proposed in the current work.

A SDM allows a relatively uniform discretization of a unitary sphere with a simple way of mapping orientations to cells. This technique has been proposed in Garcia et al. (1998) and consists in dividing a unitary sphere into a fixed number of parallels $P$. Then each parallel is divided into a number of cells that is proportional to the area covered by that parallel, the latter being approximated by the length of circumference of the parallel. The aim is that the equator has the maximum number of cells while the poles have a single cell.

SDMs are defined by a predefined number (multiple of four) of cells along the equator, CE. From it, $P+1$ parallels are defined, with $P = \mathrm{CE}/2$. Given a certain parallel $p$, its corresponding elevation angle is $\varphi(p) = \pi/2(4p/\mathrm{CE} - 1)$. From the latter, the number of cells that belong to a parallel $p$ is $\zeta(p) = 1$ if $\cos(\varphi(p)) = 0$ (i.e., $p$ is equal to 0 or $P$ and corresponds to a pole) and $\zeta(p) = \lfloor \mathrm{CE} \cos(\varphi(p)) \rfloor$ otherwise. It is easy to show that $\cos(\varphi(p))$ is the ratio between the length of circumference of the parallel at elevation $\varphi(p)$ and the length of circumference of the equator. The orientation angle of a given cell $c$ that belongs to a parallel $p$ is obtained as $\theta(p,c) = 2\pi c/\zeta(p)$.

Conversely, given an elevation angle $\varphi$, $-\pi/2 \leq \varphi \leq \pi/2$, and an orientation angle $\theta$, $0 \leq \theta \leq 2\pi$, the corresponding parallel $p$ is obtained as $p(\varphi) = \lfloor (\varphi + \pi/2)P/\pi \rfloor$, whereas the cell inside $p$ is calculated as $c(\varphi,\theta) = \lfloor \zeta(p(\varphi))\theta/2\pi \rfloor$. According to that, the pole at $\varphi = \pi/2$ is mapped to parallel $P$, the equator at $\varphi = 0$ to parallel $P/2$ and the pole at $\varphi = -\pi/2$ to parallel 0. As illustrated in Figure 1, a given triangle contributes with a vote to a SDM's cell according to its normal vector orientation $(\theta, \varphi)$.

The resolution at which the sphere is discretized only depends on the number of cells along the equator. In our implementation, all the SDMs are defined with 20 cells along the equator. This leads to a discretization of the whole sphere into 11 parallels and a total of 126 cells. The number of cells for each parallel is: 1, 6, 11, 16, 19, 20, 19, 16, 11, 6, 1 (Fig. 2).

Every normal vector contributes with a vote to the corresponding cell. To compute a more precise orientation, besides keeping the number of votes, every cell of the SDM also keeps the sum of the normal vectors that voted for that cell. Later on this value will be used to compute the representative orientation of that cell.

After computing the orientation histogram the main directions are extracted by using a two steps iterative process (Fig. 3). First, the cell with the maximum number of votes is detected $C_{(p,c)}$—global maximum. That value indicates a predominant direction in the panoramic view. Second, from that cell a neighborhood propagation process is applied $C_{(p \pm i, c \pm j)}$. This propagation depends on the resolution of the SDM and consists in labeling as belonging to the same surface those triangles with a similar orientation (neighbor cells in the SDM). In the current implementation the eight neighbor to the cell, containing that main direction, have been selected, ($i,j \in \{0,1\}$). All those $C_{(p \pm i, c \pm j)}$ cells are removed from the histogram and the process starts again by detecting the new global maximum, over the cells left in the histogram. This iterative process is applied until the value of the new maximum in the SDM is below a user defined threshold, or until a predefined number of maximums have been found (in the example presented in Fig. 4, the second option
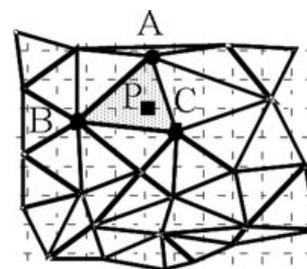


**Figure 6.** Triangles projected over the projection plane and the uniform sampling. ($P_x = uA_x + vB_x + wC_x$; $P_y = uA_y + vB_y + wC_y$; $P_z = uA_z + vB_z + wC_z$).
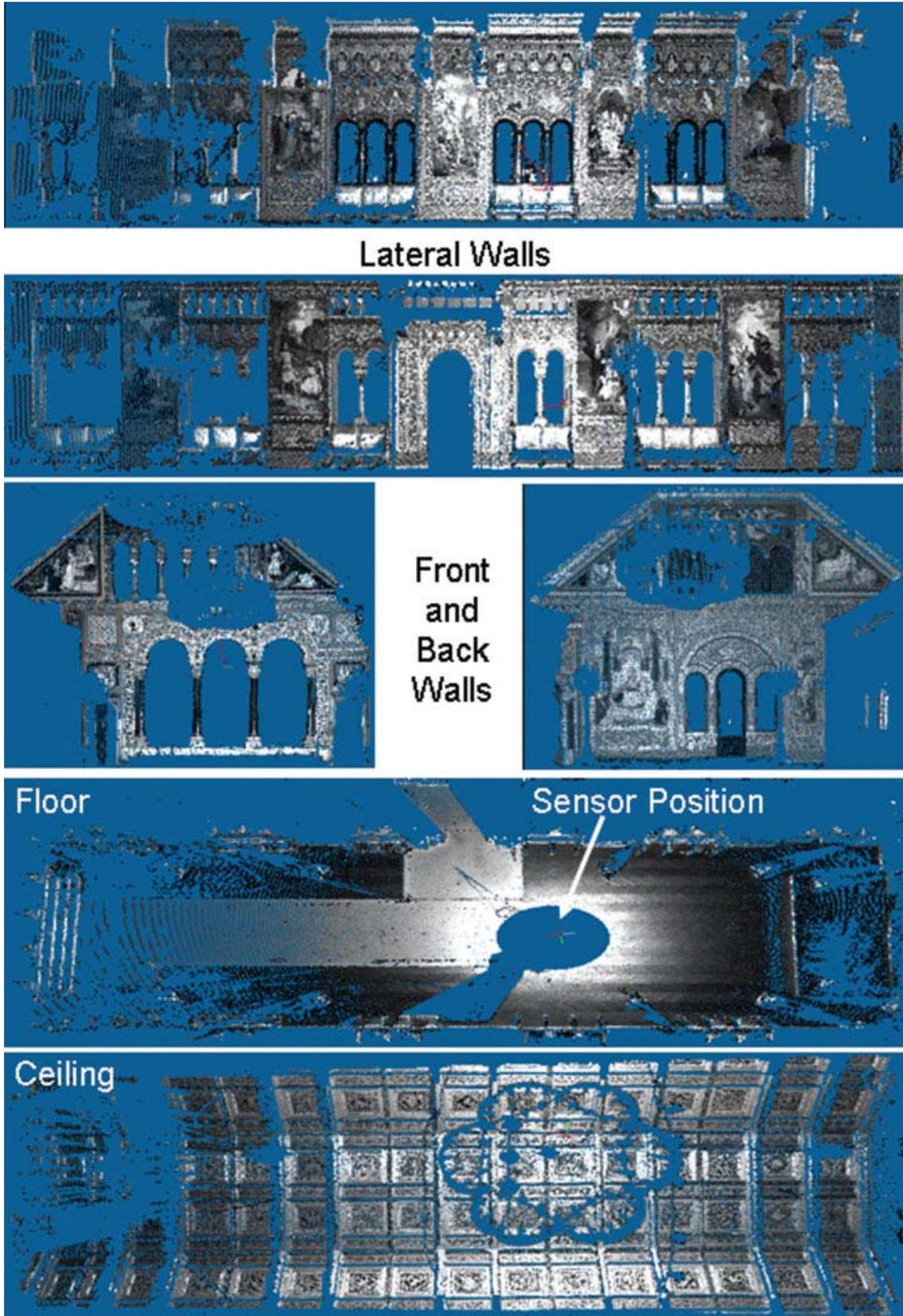
**Figure 7.** Set of planar projections automatically obtained from the panoramic range image presented in Figure 4. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com]

## IV. 2[1/2]D REPRESENTATION

The outcome of the previous stage is a set of surface orientations that describes the structure of the digitized scene. Now it is necessary to define the spatial position of projection planes—one projection plane for each main direction—as well as two clipping planes associated with each one of them.

A projection plane is orthogonal to the corresponding main direction (actually to the representative orientation resulting from the normal vectors that voted into that cell) and placed at a distance to the sensor coordinate system defined by a distance histogram. The distance histogram is computed by using the distances $\delta_i$ corresponding to those triangles whose normal vector is contained into the SDM's cell of that considered main direction. For each main direction its distance histogram is computed and the corresponding global maximum is extracted. These global maximums define the spatial position of the projection planes. In addition, each projection plane has associated two parallel clipping planes placed at a user defined distance backward and forward to the projection plane. Figure 5 (right) illustrates a distance histogram associated with a main direction and a surface distribution, both presented in the section showed in Figure 5 (left) (thick lines correspond to triangles with a normal vector contained into the SDM's cell of the main direction).

After defining the spatial position of a projection plane all those triangles bounded by the corresponding clipping planes are projected onto the projection plane. Next, a uniform sampling is computed through that planar projection generating a new 2[1/2]D representation—classical range image. The coordinates of each one of the points of that new range image are obtained by a linear interpolation of the three points of the intersecting triangle (Fig. 6). This process—projection and uniform sampling—is applied over each main direction. The result of this last stage is a set of 2[1/2]D range images associated with each of the computed planar projections.

## V. EXPERIMENTAL RESULTS

Several panoramic range images have been tested (both indoors and outdoors) showing good results. Panoramic range images have been obtained by using the Imager 530 scanner developed by the Z+F company. This sensor allows a scanning rate up to 625,000 points per second with a panoramic field of view in the horizontal direction and a field of view of $135°$ in the vertical direction (more technical details about the Imager 530 scanner are given at the company's Web page: www.zf-uk.com). The full geometry of a big environment is captured, with a high fidelity, in a short time.

Figure 7 presents snapshots of the six 2[1/2]D range images obtained by processing the panoramic range image presented in Figure 4. In this case a SDM with 20 cells along the equator was used. SDMs with a higher resolution are not necessary since in general there is a big difference between main surface orientations. Furthermore, as mentioned in Section III, every cell keeps the sum of the normal vectors that voted for that cell. Therefore, a precise resulting orientation is computed independently of the size of the SDM cells. The number of rows and columns of every 2[1/2]D range image was defined by preserving the density of points in the area of the projection plane nearest to the sensor. After applying a regular sampling of that area the projection plane is sampled with the same resolution. This process is applied independently over every projection plane since the sensor may be is not equidistant to symmetric walls.

Note that the density of points, in the results presented in Figure 7, decreases with the distance to the sensor position (floor, ceiling, and lateral walls). Density of points in the front and back walls is almost uniform since they are not so large and are placed far away to the sensor. By studying occluded areas, the next position of the sensor can be easily defined (left part of the ceiling and floor show a low density of points; in addition, front and back walls show areas occluded by the chandeliers that need further scan).

## VI. CONCLUSIONS AND FURTHER IMPROVEMENTS

This article presents a simple technique to generate 2[1/2]D representations from a panoramic range image. It works automatically without parameters tuning (except the distances between the clipping planes) and can be used for indoor as well as outdoor scenes. It is useful not only to visualize the scanned surfaces of a panoramic range image but also to define the next position of the sensor or detect occluded areas. An improvements over a previous version (Sappa, 2004) is presented, which results in a less-constraint technique.

Further work will include the study of point density as well as occluded areas detection in order to compute automatically the position where the range sensor should be placed at the next scan.

## REFERENCES

H. Cantzler, R.B. Fisher, and M. Devy, Improving architectural 3D reconstruction by plane and edge constraining, Proc British Machine Vision Conf, Cardiff, UK, September, 2002a, pp. 43–52.

H. Cantzler, R.B. Fisher, and M. Devy, Quality enhancement of reconstructed 3D models using coplanarity and constraints, Proc Symp for Pattern Recognition, Zürich, Switzerland, September, 2002b, pp. 34–41.

Y. Cho, C. Hass, K. Liapi, and S. Sreenivasan, Rapid visualization of geometric information in a construction environment, Proc IEEE Fifth Int Conf on Information Visualisation, London, UK, July, 2001, pp. 31–36.

A. Dick, P. Torr, and R. Cipolla, Automatic 3D modelling of architecture, Proc British Machine Vision Conf, Bristol, UK, September, 2000, pp. 372–381.

F. Di Fiore and Van F. Reeth, Employing approximate 3D models to enrich traditional computer assisted animation, Proc IEEE Computer Animation, Geneva, Switzerland, June, 2002, pp. 183–190.

O. Faugeras, Three-Dimensional computer vision: A geometric viewpoint, MIT Press, Cambridge, MA, 1993.

R.B. Fisher, Applying knowledge to reverse engineering problems, Proc IEEE Geometric Modeling and Processing, Riken, Japan, July, 2002, pp. 149–155.

D. Fofi, J. Salvi, and E. Mouaddib, Uncalibrated reconstruction: An adaptation to structured light vision, Pattern Recogn 36 (2003), 1631–1644.

M.A. Garcia and A. Sappa, Efficient generation of discontinuity-preserving adaptive triangulations from range images, IEEE Trans Syst Man Cybern B Cybern 34 (2004), 2003–2014.

M.A. Garcia, S. Velázquez, and A. Sappa, A two-stage algorithm for planning the next view from range images, Proc British Machine Vision Conf, Southampton, UK, September, 1998, pp. 720–729.

J. Hancock, D. Langer, M. Hebert, R. Sullivan, D. Ingimarson, E. Hoffman, M. Mettenleiter, and C. Froehlich, Active laser radar for high-performance measurements, Proc IEEE Int Conf on Robotics and Automation, Leuven, Belgium, May, 1998, pp. 1465–1470.

K. Hirahara and K. Ikeuchi, Detection of vehicles in panoramic range image, Proc IEEE Int Conf on Robotics and Automation, New Orleans, LA, April, 2004, pp. 84–89.

K. Klein and V. Sequeira, The view-cube: An efficient method of view planning for 3D modelling from range data, Proc IEEE Workshop on Applications of Computer Vision, California, USA, December, 2000, pp. 186–191.

R. Luo, J Tzou, and Y. Chang, The integration of 3D digitizing and LCD panel display based rapid prototyping system for manufacturing automation, Proc IEEE IECON 2000, Nagoya, Japan. October, 2000, pp. 1255–1260.

S. Malassiotis and M. Strintzis, Robust real-time 3D head pose estimation from range data, Pattern Recogn 38 (2005), 1153–1165.

A. Nüchter, H. Surmann, K. Lingemann, and J. Hertzberg, Semantic scene analysis of scanned 3D indoor environments, Proc Vision, Modeling, and Visualization, Munich, Germany, November, 2003, pp. 215–221.

M. Reed and P. Allen, A robotic system for 3D model acquisition from multiple range images, Proc IEEE Int Conf on Robotics and Automation, Albuquerque, New Mexico, April, 1997, pp. 2509–2514.

A. Sappa, Automatic extraction of planar projections from panoramic range images, Proc IEEE Int Symp on 3D Data Processing, Visualization and Transmission, Thessaloniki, Greece, September, 2004.

C. Wang, H. Tanahashi, Y. Sato, H. Hirayu, Y. Niwa, and K. Yamamoto, Registering panoramic range data and omni-directional color image based on edge histograms, Proc Int Conf on Pattern Recognition, Quebec, Canada, August, 2002, pp. 355–358.