MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

A MOBILE ROBOT FOR VISUAL MEASUREMENTS IN ARCHITECTURAL APPLICATIONS

Xavier Lebègue and J. K. Aggarwal

Computer and Vision Research Center
The University of Texas at Austin
Austin, Texas 78712, USA

ABSTRACT

This paper describes a method to perform distance measurements inside buildings using a single video camera mounted on a mobile robot. This algorithm is designed specifically to create architectural floor plans with accurate measurements. Several issues associated with architectural surveying are discussed in detail, such as the visibility of features and the accuracy in calibration. Results of the edge reconstruction in an indoor scene are presented and compared to an architect's drawing of the building. Applications of this method include the verification of constructed buildings, the update or creation of civil engineering CAD models, and the acquisition of data for architectural graphics simulations.

INTRODUCTION

The perception system of most mobile robots is geared towards navigation, where qualitative information is more important than a high accuracy. We propose here a different task: the automatic metrical surveying of a building using a single camera mounted on a mobile robot. The robot provides a sequence of images registered with odometric estimates of the motion between frames. The vision algorithm reconstructs the 3-D scene by using the known motion of the camera. In order to extract the most useful features for reconstruction, we concentrate on segments that have a particular orientation in the 3-D scene. In most buildings, there are three prominent 3-D orientations: the vertical and two horizontal 3-D orientations perpendicular to each other. This assumption is often used with indoor scenes [1, 2, 7, 8, 9], and will hold for the rest of this paper.

We begin by describing algorithms to extract useful line segments from each image and to reconstruct the 3-D scene. We then discuss several issues of particular importance for practical applications in architectural surveying, such as the visibility of segments in typical indoor scenes. The following section concentrates on accuracy issues and explains the calibration procedure for the camera and the robot. Finally, we provide a comparison between reconstructed edges and the corresponding architectural floor plan.

3-D RECONSTRUCTION

Line segments corresponding to particular orientations in 3-D are extracted from each image using a special line detector based on vanishing points [4]. The detection and interpretation process provides a 3-D *orientation* hypothesis for each 2-D segment. Vertical segments are of particular importance for floor plans, but other particular 3-D orientations can be processed as well. Segments that do not have any of the predefined 3-D orientations are ignored.

The 3-D *position* of segments is estimated from a sequence of images using a method based on Kalman filtering [5, 6]. The position of each 3-D segment is represented by an estimate and an associated Gaussian uncertainty. The floor plan can then be constructed from these estimates. In our approach, each 3-D orientation is treated separately. This is possible since the line detector indicates the 3-D orientation of the segments. Matching is therefore simplified, as is the complexity of prediction and update: all the computation is done in planes perpendicular to each one of the predefined 3-D orientations.

APPLICATION TO ARCHITECTURE

The application of our reconstruction algorithms to architectural surveying imposes a few practical constraints.

Field of view: In a typical indoor scene, there are relatively few *architecturally significant* features, such as room corners and doorways. Those features are usually far apart. However, the robot should see as many features as possible at any time. For this reason, we equipped the camera with a wide-angle lens (6 mm focal distance for a 2/3 inch CCD camera). The advantage is that more interesting edges can be tracked at once. In long and narrow corridors, segments can be observed not only from a distance, but also as the robot passes by them. Those segments are therefore acquired under very different angles, thus reducing the uncertainty in their reconstruction. With a wide-angle lens, the forward-looking camera can still see edges lying almost along a perpendicular to the direction of motion. For those edges, the precision of reconstruction is similar to that of a stereovision system with widely spaced cameras. The accuracy is then very high, assuming that the exact motion of the camera is known.

The drawback to using a wide-angle lens is the associated barrel distortion (see Figure 1). In the next section, we will explain how this distortion is calibrated and corrected.

Motion determination: We have assumed so far that the exact motion of the camera was given to the 3-D reconstruction algorithm. Although the robot's odometers provide a position and a heading, these measurements cannot be used alone, since odometers drift without bounds. Odometry may be adequate for estimating the motion between a few images, but it is insufficient for the long sequences necessary in mapping an entire building.

The robot is equipped with an odometer on its right and left driving wheels. In practice, translation measurements are very good because they are derived from the average of the two odometers. However, rotations (changes in robot heading) drift much faster because they rely on the difference between the two odometers. For this reason, we correct the odometric heading with vision every time an image is processed. By extracting the exact position of the vanishing points in the image, the orientation of the camera with respect to the scene is computed precisely. The heading of the robot is estimated, as well as the roll and pitch errors due to uneven floor surface. This technique prevents the heading from drifting without bounds. The position of the robot, obtained by integrating small displacements, can still drift because of errors in translations, but they build up much slower than errors in rotations.

Using this technique, the robot's positional error was reduced by several orders of magnitude after a rectangular trajectory of 125 meters around a building floor (see [5] for details).

ACCURACY AND CALIBRATION

For any practical application to architectural surveying, the precision of measurement is paramount. An accurate calibration procedure is necessary. In this section, we describe the calibration procedure for the intrinsic parameters of the camera, including the barrel distortion, as well as for the extrinsic parameters (camera-robot relationship).

The Optical axis: We first determine the intersection of the image plane and the optical axis, expressed in pixel units in the coordinate system of the frame grabber. The method we use is optical. After closing the iris of the camera, a low-power laser beam is shone through a perforated white screen into the center of the lens. The beam is partially reflected by the lens onto the screen. The goal is to align the laser beam and the optical axis of the lens. By carefully adjusting the pan and tilt of both the laser and the camera, the beam is made to reflect onto itself. At that point, an attenuating filter is placed in the beam and the iris is slightly opened. A bright spot becomes visible in the image, indicating the location of the optical axis on the retina. This method is not recommended with non-CCD cameras.

The projection model: The next calibration step consists of identifying the focal length and the distortion of the lens. The distortion is zero on the optical axis, and non-negligible near the borders of the image rectangle. As a simplification, we assume that the image is formed in two steps. Firstly, the scene is projected onto the image plane by a standard pinhole perspective projection. Secondly, the image is distorted in the image plane by a 2-D to 2-D function. This assumption means that the barrel distortion does not depend on the distance between the lens and the scene, and holds if the diameter of the lens is not too large.

The standard perspective projection is completely defined by the location of the optical axis in the image (determined in the previous subsection) and the focal length (to be determined later for points near the optical axis, where the distortion is negligible). The second step, a 2-D distortion in the image plane will then modeled by calibration. This model is used to correct the distortion in digitized images [3].

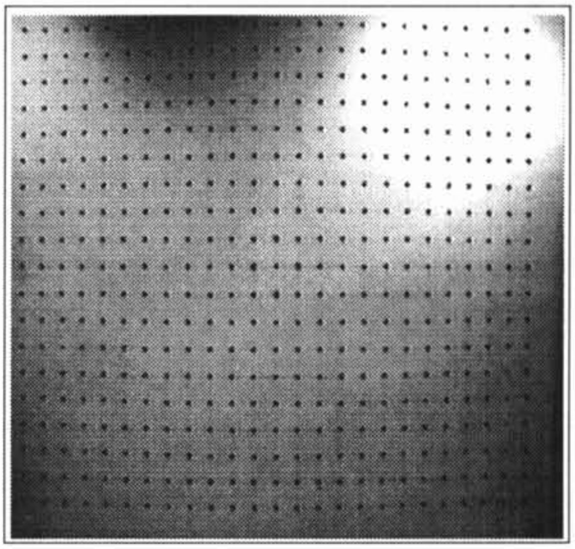

Figure 1: **The calibration pattern distorted by the wide-angle lens**

The calibration pattern: We first measure the location of a set of data points covering the image. To obtain the data points, we place the camera in front of a large calibration pattern on a wall (see Figure 1). The pattern consists of a grid of black dots placed every 10 cm. For best results, the distance between the camera and the pattern should approximate the typical distance to objects when the robot is in operation. In practice, this leads to a calibration pattern as large as the scene itself, if possible. The size of our pattern is 2.4 x 1.8 meters.

We first need to ensure that the image plane is parallel to the pattern, so that the perspective projection will not interfere with the 2-D distortion in the image plane. To achieve this constraint, we place a mirror flat on the pattern. We then adjust the pan and tilt of the camera until the reflection of the camera by the mirror coincides with the location of the optical center in the image plane. To facilitate this operation, we digitally superimpose cross hairs centered on the the optical axis over the live image from the camera. Using the cross hairs, any roll in the camera's orientation is also eliminated at this stage. Finally, the image of the pattern is digitized and the data points are automatically extracted.

The focal length: The next intrinsic parameter to be determined is the focal length, expressed in horizontal and vertical pixel units. The focal length is computed for points lying close to the center of the image, where the barrel distortion is negligible. The distance D between the camera and the calibration pattern is physically measured, and expressed in millimeters. Let L be the distance (in millimeters) between two points close to the center of the pattern, and l the distance (in pixels) between their perspective projections. Neglecting the barrel distortion, the focal length f is given in pixels by $f = lD/L$. Since the image pixels are not necessary square, the focal length needs to be expressed both in horizontal and vertical pixel units.

Barrel distortion: At this point, all the parameters for a conventional perspective projection have been determined. However, most points of the pattern do not project exactly to their theoretical location on the image plane because of the barrel distortion. We automatically measure the resulting 2-D displacement vectors in the image plane, as the difference between the theoretical and observed coordinates of projected dots. The 2-D distortion function is estimated by linear interpolation between data points [3]. Digitized images can then be corrected to eliminate barrel distortion. The intensity of every pixel in the corrected image is found by looking up the intensity at the appropriate location in the original distorted image. In order to speed-up the distortion correction, a look-up table is computed at the time of calibration to translate corrected coordinates into raw coordinates.

Extrinsic parameters: We then measure the rotation and translation of the camera relative to the robot's coordinate system. The translations are physically measured. The roll is forced to zero by aligning cross hairs digitally superimposed on the image to the horizontal axis of the calibration pattern, as described previously. Next, the pan and tilt are adjusted until the robot is able to back away from the calibration pattern in a straight line while keeping the same point of the pattern under the cross hairs of the optical axis. This process results in zero pan and tilt. If a non-zero pan or tilt is desired, the camera is rotated from the zero-angle position by a controlled angle.

RESULTS

Figure 2 shows a short sequence of images acquired by the robot. The corresponding line segment images are given in Figure 3. The result of the reconstruction of vertical edges using these four images is presented in Figure 4, and compared to the architectural floor plan of the building. Vertical edges are indicated by small crosses. The last position of the robot is represented by a small square in the middle of the corridor. The grid has a one-meter spacing. The door in the top right corner was closed, and some segments on it are reconstructed. A few segments with no architectural interest are picked up. For example, two of them lie on the wall at the bottom of the figure. The segments in the top-left corner of the figure appear to have been shifted to the right. After measuring the actual dimensions in the scene, we found that the corner of the corridor was built approximately 5 cm to the right of the position shown on the architectural drawing. The robot therefore correctly indicated this discrepancy, although it overestimated it slightly.

CONCLUSION

We proposed a method for the task of surveying buildings using mobile robots. We addressed several important issues resulting from the particular requirements of architectural applications, and we presented results obtained in the reconstruction of indoor scenes. Finally, we compared those results to the architect's drawing of the floor plan. Improvements to our approach are possible in several areas, to further increase the precision of measurements and to produce CAD-like building maps. We believe that automatic architectural surveying using mobile robots is a feasible task in the fairly near future. Applications will include the verification of constructed buildings, the update or creation of civil engineering CAD models, the acquisition of data for architectural graphics simulations, and the creation of maps for use by other robots.

Acknowledgments

This research was supported in part by the DoD Joint Services Electronics Program through the Air Force Office of Scientific Research (AFSC) Contract F49620-89-C-0044, and in part by the Army Research Office under contract DAAL03-91-G-0050.

References

[1] S. T. Barnard. Interpreting perspective images. *Artificial Intelligence*, 21(4):435–462, November 1983.

[2] P. Belluta, G. Collini, A. Verri, and V. Torre. Navigation by tracking vanishing points. In *Working Notes of AAAI Robot Navigation Symposium*, pages 6–10, Stanford University, March 1989.

[3] D. A. Butler and P. K. Pierson. A distortion-correction scheme for industrial machine-vision applications. *IEEE Trans. on Robotics and Automation*, 7(4):546–551, August 1991.

[4] X. Lebègue and J. K. Aggarwal. Detecting 3-D parallel lines for perceptual organization. In *Proc. Second European Conf. on Computer Vision*, pages 720–724, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag.

[5] X. Lebègue and J. K. Aggarwal. Extraction and interpretation of semantically significant line segments for a mobile robot. In *Proc. IEEE Int. Conf. Robotics and Automation*, pages 1778–1785, Nice, France, May 1992.

[6] X. Lebègue and J. K. Aggarwal. Semantically significant line segments for a mobile robot... ROBOTEX. In *Video Proc. IEEE Int. Conf. Robotics and Automation*, Nice, France, May 1992.

[7] P. Olivieri, M. Gatti, M. Straforini, and V. Torre. A method for the 3D reconstruction of indoor scenes from monocular images. In *Proc. Second European Conf. on Computer Vision*, pages 696–700, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag.

[8] T. Shakunaga. 3-D corridor scene modeling from a single view under natural lighting conditions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):293–298, February 1992.

[9] M. Straforini, C. Coelho, M. Campani, and V. Torre. The recovery and understanding of a line drawing from indoor scenes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):298–303, February 1992.

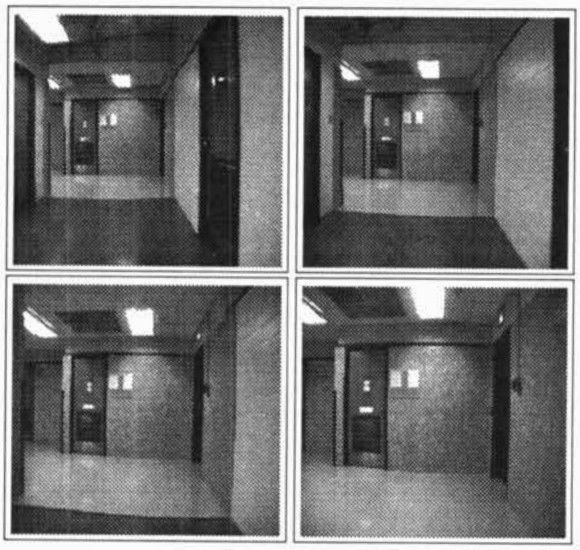

Figure 2: **The four images used in the reconstruction**

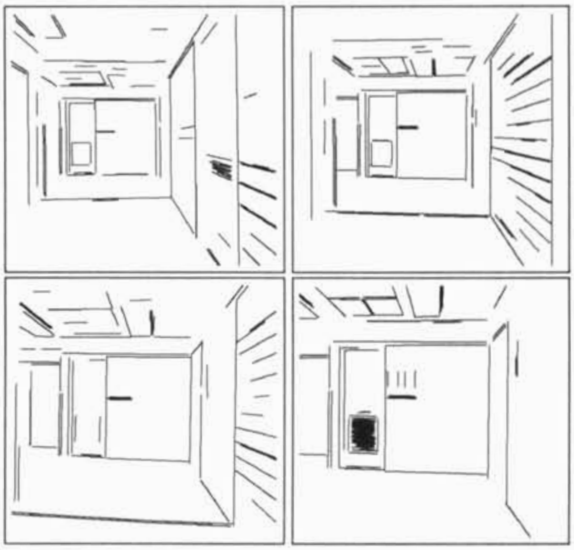

Figure 3: **The corresponding line segment images**

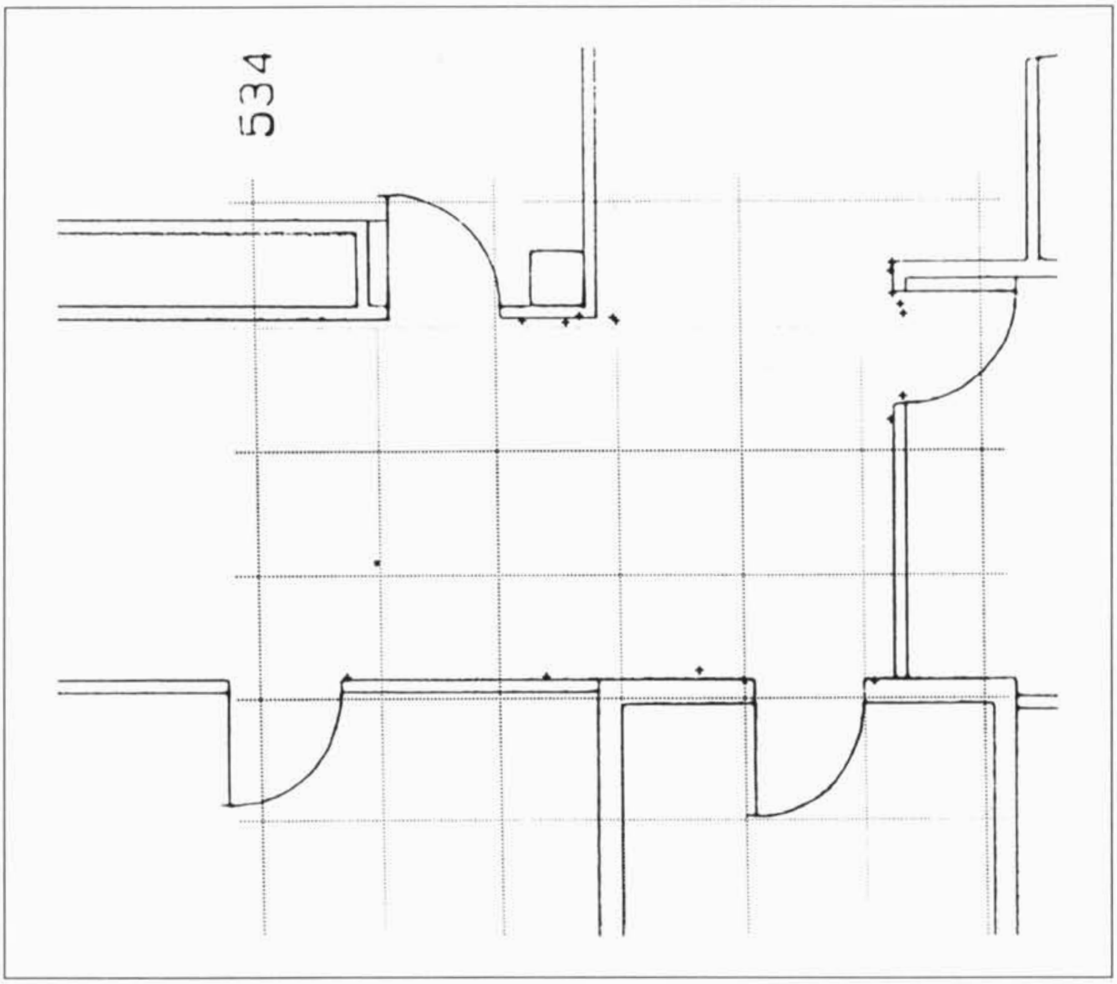

Figure 4: **The reconstructed vertical edges overlaid on the floor plan**