BUILDING BLOCKS FOR AUTONOMOUS NAVIGATION USING CONTOUR CORRESPONDENCES

M. Pawan Kumar   *C. V. Jawahar*   *P. J. Narayanan*

Centre for Visual Information Technology
International Institute of Information Technology
Gachibowli, Hyderabad 500019. INDIA.

ABSTRACT

We address a few problems in navigation of automated vehicles using images captured by a mounted camera. Specifically, we look at the recognition of sign boards, rectification of planar objects imaged by the camera, and estimation of the position of a vehicle with respect to a fixed sign board. Our solutions are based on contour correspondence between a reference view and the current view. A mapping between corresponding points of a planar object in two different views is a $3x3$ matrix called the homography. A novel two-step linear algorithm for homography calculation from contour correspondence is developed first. Our algorithm requires the identification of an image contour as the projections of a known planar world contour and the selection of a known starting point. The homography between the reference view and the target view is applied to several real-life navigation applications, results of which are presented in this paper.

1. INTRODUCTION

This paper addresses problems encountered in visual navigation. Visual navigation deals with unsupervised, or minimally supervised, movement of automated vehicles with mounted camera(s) in a manner which serves a definite purpose. The purpose could be remote sensing, surveillance in hazardous environment or visual servoing [1, 2, 3, 4]. The sequence of images obtained using the camera are processed and decisions about the movement of the vehicle are taken automatically. Several approaches have been proposed for both indoor navigation and outdoor navigation. A comprehensive survey on this can be found in [5].

This paper addresses three problems in navigation, namely, recognition of sign boards which instruct the automated vehicle to perform specified actions, visualization by rectification of planar objects in the scene which are imaged by the camera, and estimation of the position of the vehicle with respect to a fixed object. Most approaches reported in the literature make use of explicit point correspondences or restricted geometric primitives such as parallel or perpendicular lines for navigation. Such features are not abundant and might not be available in most situations. We use correspondences of non-parametric, unstructured, contours for solving these problems. Our assumptions on the contours are also minimal, making this method applicable to a large number of situations.

Section 2 provides the necessary background in geometric formulation employed here. A linear homography calculation algorithm, needed for navigation using non-parametric contours, is presented in section 3. Section 4 discusses the solutions for the navigation problems. Several results are presented which prove the robustness of our method. We conclude the paper in section 5.

2. PROBLEM FORMULATION

Problems of object recognition, metric rectification and position estimation are important for visual navigation. In this paper, we attempt to solve these problems using contour correspondences. Non-parametric, arbitrary contours of objects provide rich information that has not been exploited sufficiently in the literature. They are practically more useful as no assumptions on their structure is made.

Let $\mathbf{x}^1[i]$ and $\mathbf{x}^2[i]$ be two contours represented as a sequence of points in homogeneous coordinates. If the contours were obtained from imaging the same planar object, then there is 3×3 matrix $\mathbf{H}$, known as the homography or the projective transform, such that

$$\mathbf{x}^1[j] = \mathbf{H}\mathbf{x}^2[k]$$

Note that j and k are not known or in other words, explicit point correspondence is not available. We present an algorithm to calculate the homography $\mathbf{H}$, exploiting the connectedness and the orderliness of contour points using projective invariants. Let p be some parameters in one view and let p' be the corresponding parameters in the other view. A function $I(\cdot)$ is called a projective invariant iff, $I(p) = I(p')$. Several cross ratios are invariant under projective transformation like cross-ratio of areas.

Cross-ratio of areas of five points: The cross-ratio of the areas of five points $\mathbf{x_1}$, $\mathbf{x_2}$, $\mathbf{x_3}$, $\mathbf{x_4}$, and $\mathbf{x_5}$ is defined in [6]

$$cr(\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}) = \frac{\triangle_{\mathbf{x_1 x_2 x_5}} \cdot \triangle_{\mathbf{x_3 x_4 x_5}}}{\triangle_{\mathbf{x_1 x_3 x_5}} \cdot \triangle_{\mathbf{x_2 x_4 x_5}}} \quad , \qquad (1)$$

where $\triangle_{\mathbf{x_i x_j x_k}}$ is the area of the triangle formed by points $\mathbf{x_i}, \mathbf{x_j}, \mathbf{x_k}$. This is invariant to general linear or projective transformations.

3. HOMOGRAPHY CALCULATION

We develop a novel two step algorithm for homography calculation from contour correspondences. In the first step, two invariant lines are obtained given one point correspondence. In the next step, a large number of correspondences over the contour are computed thereby making homography calculation stable.

Invariant Lines: We use an approach similar to the one used in finding polygonal approximation in a projective invariant approach [7]. Let $\mathbf{x}$ be the given point in the first view. Let $\mathbf{y}$ be the point adjacent to $\mathbf{x}$ on the contour. For obtaining the invariant lines, a measure of deviation of a point $\mathbf{z}$ from the line joining $\mathbf{x}$ and $\mathbf{y}$, which is invariant to projective transformation, is found. Let $\lambda(\mathbf{z})$ be that measure of deviation. An invariant line can be found by checking for $\lambda(\mathbf{z})$ for each point on the contour starting from the given point. We continue traversing the contour until there is a deviation. By traversing clockwise once and then anticlockwise the next time from $\mathbf{x}$, two lines are obtained on the contour. Since $\lambda(\mathbf{z})$ is invariant to projective transformation, following this procedure on two different images results in two pairs of invariant lines which intersect at the given point.

One such measure $\lambda(\mathbf{z})$ is the ratio of cross-ratio of areas. Let $\mathbf{t_1}, \mathbf{t_2}, \mathbf{t_3}$ be any three points. Define $\lambda(\mathbf{z})$ as the ratio of two cross ratios. We get

$$\begin{aligned} \lambda(\mathbf{z}) &= \frac{cr(\mathbf{x}, \mathbf{t_1}, \mathbf{t_2}, \mathbf{t_3}, \mathbf{y})}{cr(\mathbf{x}, \mathbf{t_1}, \mathbf{t_2}, \mathbf{t_3}, \mathbf{z})} \\ &= \frac{\triangle_{\mathbf{x t_3 z}} \cdot \triangle_{\mathbf{x t_2 y}}}{\triangle_{\mathbf{x t_2 z}} \cdot \triangle_{\mathbf{x t_3 y}}} \end{aligned}$$

Let len_{ij} be the distance from point i to j and $dist_i^{jk}$ be the perpendicular distance of the point i from the line joining j and k. Then using $\triangle_{ijk} = 1/2 * len_{ij} * dist_{jk}^i$, we rewrite

$$\begin{aligned} \lambda(\mathbf{z}) &= \frac{len_{\mathbf{xt3}} \cdot dist_{\mathbf{z}}^{\mathbf{xt3}} \cdot len_{\mathbf{xt2}} \cdot dist_{\mathbf{y}}^{\mathbf{xt2}}}{len_{\mathbf{xt3}} \cdot dist_{\mathbf{y}}^{\mathbf{xt3}} \cdot len_{\mathbf{xt2}} \cdot dist_{\mathbf{z}}^{\mathbf{xt2}}} \\ &= \frac{dist_{\mathbf{z}}^{\mathbf{xt3}} \cdot dist_{\mathbf{y}}^{\mathbf{xt2}}}{dist_{\mathbf{y}}^{\mathbf{xt3}} \cdot dist_{\mathbf{z}}^{\mathbf{xt2}}} \qquad (2) \end{aligned}$$

The value of $\lambda(\mathbf{z})$ provides us with the required measure of deviation. It is clear that if $\mathbf{z}$ lies on the line joining $\mathbf{x}$ and $\mathbf{y}$ then $\lambda(\mathbf{z}) = 1$ because then,

$$\frac{dist_{\mathbf{z}}^{\mathbf{xt3}}}{dist_{\mathbf{z}}^{\mathbf{xt2}}} = \frac{dist_{\mathbf{y}}^{\mathbf{xt3}}}{dist_{\mathbf{y}}^{\mathbf{xt2}}}$$

The measure $|1 - \lambda(\mathbf{z})|$ provides an estimate of the deviation of $\mathbf{z}$ from the line joining $\mathbf{x}$ and $\mathbf{y}$. These invariant lines can be used to obtain large number of correspondences on the contour.

Generating Correspondences: We define

$$\mu(\mathbf{z}) = dist_{\mathbf{z}}^{\mathbf{xt3}} / dist_{\mathbf{z}}^{\mathbf{xt2}}$$

which remains invariant under projective transformation up to scale i.e., the ratio of the values of $\mu(\mathbf{z})$ in two views is a constant for all points $\mathbf{z}$. Note that $\mu(\mathbf{z})$ is the ratio of the distance of a point from two lines.

Using the two invariant lines that are obtained, $\mu(\mathbf{z})$ is calculated for all points on the contour. Given two views of the same object, two sequences of real numbers which are related by a common scale factor are obtained. By obtaining this common scale, a large number of correspondences on the contour are also be obtained. That is, the best match for each point on the contour given its two views is found by comparing the value of $\mu(\mathbf{z})$s.

We prove that $\mu(\mathbf{z})$ is invariant to projective transformation up to scale. Let $\mathbf{l_1}, \mathbf{l_2}$ be the two invariant lines. Let $dist_i^l$ be the perpendicular distance of the point i from the line l. From equation 2,

$$\lambda(\mathbf{z}) = c * dist_{\mathbf{z}}^{\mathbf{l_1}} / dist_{\mathbf{z}}^{\mathbf{l_2}}$$

where $\quad c = dist_{\mathbf{y}}^{\mathbf{l_2}} / dist_{\mathbf{y}}^{\mathbf{l_1}}$

In another view where the corresponding entities of $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{l_1}, \mathbf{l_2}$ are $\mathbf{x}', \mathbf{y}', \mathbf{z}', \mathbf{l}'_1, \mathbf{l}'_2$, we get

$$\lambda(\mathbf{z}') = c' * dist_{\mathbf{z}'}^{\mathbf{l}'_1} / dist_{\mathbf{z}'}^{\mathbf{l}'_2}$$

where $\quad c' = dist_{\mathbf{y}'}^{\mathbf{l}'_2} / dist_{\mathbf{y}'}^{\mathbf{l}'_1}$

Since $\lambda(\mathbf{z}) = \lambda(\mathbf{z}')$, implies $\mu(\mathbf{z}) = c'/c * \mu(\mathbf{z}')$ under projective transformation. In other words, $\mu(\mathbf{z})$ remains invariant under projective transformation up to scale (the scale being c'/c).

To obtain the common scale factor, consider two sequences $\mathbf{x_1}[i]$ and $\mathbf{x_2}[i]$ related by a scale factor 's'. Their Fourier transform $\mathbf{X_1}$ and $\mathbf{X_2}$ are also related by the same scale factor. The scale is obtained by comparing the corresponding frequency values in the two Fourier transform sequences. This is a robust method for finding the scale and similar approaches have been used for various applications such as object recognition [8]. Having obtained a large number of correspondences on the contours, the homography relating the two views is computed by the DLT algorithm [6].

Once the homography calculation from contour correspondence is obtained, various navigation problems can be solved using it as an intermediate step.

4. NAVIGATION APPLICATIONS

We now present solutions to the problems of signboard recognition, metric rectification of planar objects and estimation of position of the vehicle.

4.1. Signboard Recognition

Given the reference view of various signboards and the distorted view of one of them, the problem is to identify that reference signboard which results in the distorted view by imaging.

The method adopted for recognition is as follows. Calculate the homography relating this distorted view with all the reference views. In doing so, a large number of correspondences are obtained on the contour in each case. In all cases but one, solving for the homography with these correspondences using DLT algorithm would result in very high error values. This is because we are trying to find the homography between two views of two different planar objects (which doesn't exist). We identify the distorted view as that signboard which results in the least error.

For testing, we used twenty views of 8 signboards having different contours on them and tried to recognize them correctly using the above method. Figure 1 shows some examples of the images that were used for testing. Table 1 shows the results of the experiments. A high level of accuracy was achieved thereby verifying the robustness of the system.

Sign Board Model	Accuracy (%)
1	100
2	100
3	94.73
4	100
5	100
6	100
7	89.47
8	100

Table 1. Accuracies obtained for various sign boards using our recognition approach

4.2. Metric Rectification of Planar Objects

The mounted camera on the automated vehicle is restricted in motion and the path it follows is also fixed. Therefore, it will not be able to capture the fronto-parallel view of various planar objects present in the scene. The process of obtaining fronto-parallel view from the projectively transformed view is called metric rectification.

The images of billboards on the road often contain letters which when imaged are transformed by the same projective transform as the entire billboard. However, the shape or contour of these letters without the pure projective and pure affine transformations is known. Using the homography calculation method from contour correspondence, the transformation between the contour in the image and the original contour is calculated. Applying the inverse of this transformation results in a fronto-parallel view of the sign board.

Figure 2 shows the distorted and the fronto-parallel view of one such billboard. The fronto-parallel view of this is obtained using the above method. We used the 'X' in the billboard for rectification.

(a) (b)

Fig. 2. Projectively distorted (a) and geometrically corrected (b) images. Planar boundary of 'X' is used for rectification.

4.3. Estimation of Position of Automated Vehicle

This problem involves calculating the relative position of the vehicle with respect to a fixed object. It is used in various applications, such as visual servoing. We need to know exactly how much and in which direction to move in order to reach a particular location. For this problem, we assume that the intrinsic parameters of the mounted camera, such as focal length and principle point, are known. This could easily be achieved by using any of the known camera calibration techniques [9].

The location is specified using a signboard containing a known contour. The camera captures a transformed image of the contour on the sign. Using the above method, the homography $\mathbf{H}$ relating the contour in the image to the contour in the world is calculated. The problem then deals with determining the external parameters of the camera using $\mathbf{H}$.

Let $\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}]$ be the perspective matrix describing the

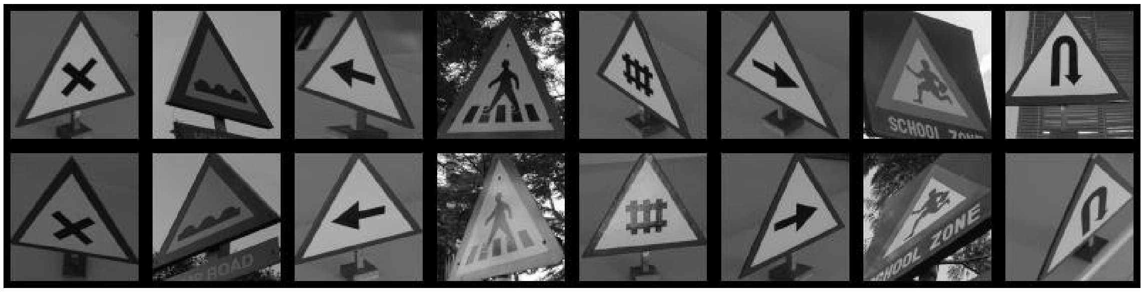

Fig. 1. Examples of images used in signboard recognition experiment

camera used. Here $\mathbf{K}$ contains the known intrinsic parameters of the camera. The rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$ represent the position of the camera. $\mathbf{R}$ and $\mathbf{t}$ are obtained as follows. $\mathbf{H} = \lambda\mathbf{KL}$ where λ is an unknown scale factor and $\mathbf{L}$ contains the first two columns of $\mathbf{R}$ and $\mathbf{t}$. An estimate of λ can be obtained as the average of the norms of the first two columns of $\lambda\mathbf{L} = \mathbf{K}^{-1}\mathbf{H}$. Once we obtain $\mathbf{L}$, the first two columns $\mathbf{R}$ and $\mathbf{t}$ can be obtained. The third column of $\mathbf{R}$ is the cross product of the first two columns of $\mathbf{R}$. Thus, the position of the vehicle with respect to the sign is calculated. Figure 3 shows the initial, intermediate and final positions of the vehicle. As can be seen, the vehicle successfully reaches the position specified by the sign.

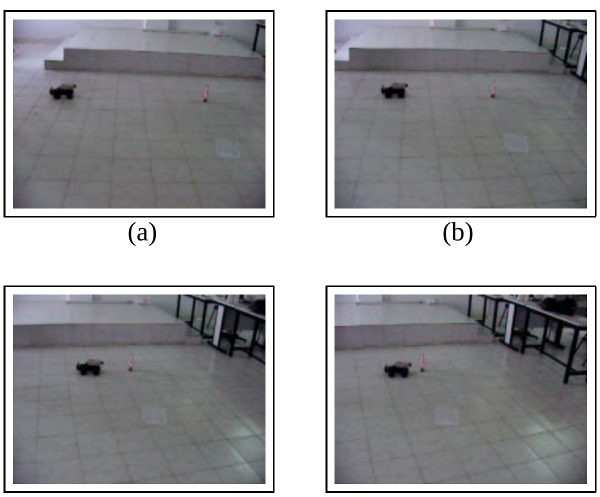

Fig. 3. Initial(a), intermediate(b and c) and final(c) positions of the vehicle. The vehicle is parked at the place indicated by the sign.

5. CONCLUSION

In this paper, we presented solutions for some of the visual navigation problems using contour correspondences. Line and point correspondences are fewer and harder to obtain; but arbitrary contours are found everywhere. At the centre of our method is an algorithm for computing the homography between a reference view and a target view in question. The algorithm requires the identification of a unique starting point in all views. We are currently working on making the method independent of the required one point correspondence. Other problems associated with navigation, such as avoiding obstacles, also need to be explored.

6. REFERENCES

[1] B. Espiau, F. Chaumette, and P. Rives, "A new approach to visual servoing in robotics," Tech. Rep. 542, IRISA, July 1990.

[2] S.Y. Harmon, "The ground surveillance robot (gsr): An autonomous vehicle designed to transit unknown terrain," *IEEE Journal of Robotics and Automation*, vol. 3, pp. 266–279, 1987.

[3] C. Taylor and D. Kriegman, "Vision-based motion planning and exploration algorithms for mobile robots," *IEEE Tran. on Robotics and Automation*, vol. 14, pp. 417–427, 1996.

[4] L. Lucido, J. Opderbecke, V. Rigaud, R. Deriche, and Z. Zhang, "An integrated multiscale approach for terrain referenced underwater navigation," *International Conference on Image Processing*, vol. B, pp. 633–636, 1996.

[5] G.N. DeSouza and A.Kak, "Vision for mobile robot navigation: A survey," *IEEE Trans. on PAMI*, vol. 24, pp. 237–267, 2002.

[6] R. Hartley and A. Zisserman, *Multiple View Geometry*, Cambridge University Press, 2000.

[7] M. Pawan Kumar, Saurabh Goyal, Sujit Kuthirummal, C. V. Jawahar, and P. J. Narayanan, "Discrete contours in multple views: Approximation and recognition," *Image and Vision Computing(To Appear)*, 2004.

[8] Sujit Kuthirummal, C.V. Jawahar, and P.J. Narayanan, "Planar shape recognition across multiple views," in *Intl. Conf. on Pattern Recognition*, Aug. 2002, pp. 482–488.

[9] Michael Grossberg and Shree Nayar, "A general imaging model and a method for finding its parameters," in *International Conference on Computer Vision*, 2001, pp. 108–115.