A BAYESIAN METHOD FOR TRIANGULATION AND ITS APPLICATION TO FINDING CORRESPONDING POINTS

Anand S. Bedekar and Robert M. Haralick

Dept. of Electrical Engg., Box 352500,
University of Washington,
Seattle, WA 98195-2500.

E-Mail : {anand,haralick}@ee.washington.edu

ABSTRACT

In this paper, the problems of finding corresponding points from multiple perspective projection images, and estimating the 3-D points from which these points have arisen, are addressed. The problem of finding corresponding points is formulated as a hypothesis verification problem. Given a set of 2-D points, one from each of N perspective projection images, under the hypothesis that the points are projections of the same 3-D point, the co-ordinates of the 3-D point are estimated. The triangulation problem – the problem of estimating the co-ordinates of a 3-D point, given its projections in N perspective projection images – is posed as a Bayesian estimation problem, taking into account the uncertainties in the observed image points and the camera parameters. Based on the Bayesian estimate of the triangulated point, a statistical test is derived for verifying the hypothesis that the given set of image points is in correspondence. For finding N-tuples of corresponding points from N perspective projection images, this test can be used on each N-tuple of points to verify the hypothesis that that N-tuple of points is in correspondence, selecting those N-tuples that pass the hypothesis test. Experiments are described for characterizing the distance of the 3-D point estimated by the Bayesian triangulation from the true 3-D point, and characterizing the misdetection and false alarm rates of this method of finding corresponding points.

1. INTRODUCTION

The problems of finding corresponding points from multiple images, and of estimating the 3-D points that gave rise to them, arise in the construction of a 3-D model for the scene being observed. "Corresponding points" are points that are projections of the same 3-D point. The problem of estimating the co-ordinates of a 3-D point, given its projections in N images, is called the "triangulation problem". This paper addresses the problems of finding corresponding points from multiple perspective projection images, and triangulating them to estimate the 3-D points from which these points have arisen, given estimates of the camera parameters and the covariance matrices of these estimates.

The triangulation problem has been extensively analyzed in the photogrammetry literature [3][4]. The classical approach to triangulation is a least-squares approach, involving minimizing the sum of squared residual distances of the observations from the projections of the triangulated point. Recent work [5] has addressed the issue of finding triangulation methods that are invariant to certain kinds of transformations. These triangulation methods start with the assumption that the given 2-D points are guaranteed to be perspective projections of the same 3-D point. When the correspondences between the 2-D points are not known, one can make a hypothesis about a particular set of 2-D points being in correspondence, and triangulate to get the 3-D point under this hypothesis. This hypothesis should then be verified, based on how good the estimate of the 3-D point is. However, none of the above methods address the issue of verifying this hypothesis.

Recently, there has been considerable interest in the problem of finding corresponding image points with the aim of getting a good triangulation. Scott and Longuet-Higgins [6] proposed a method for finding corresponding points from a pair of images based on a proximity matrix that involves Gaussian weighted distances between features. They use the eigenvectors of this matrix to construct a pairing matrix, which minimizes the inner product of the pairing matrix and the proximity matrix. Cheng *et al* [7] proposed two methods that find corresponding points and do triangulation. One approach is based on the work of Scott and Longuet-Higgins [6]. They construct a proximity matrix using the squared distances between the observed 2-D points and the projections of a "pseudo-triangulated" point, that is, a point that minimizes the 3-D distance to the rays projected back through the image points. The other method poses the correspondence problem as that of finding a maximum matching in a bipartite graph in which the nodes are the points from two images. This problem is then solved as a network flow maximization problem by assigning an appropriate weight to each edge based on a similarity function of the points that are connected by the edge. However, none of these methods takes into account the uncertainty in the locations of the image points and the camera parameters with an explicit noise model. Also these methods are restricted to finding corresponding points and triangulating them from just two images.

In this paper, the problem of finding corresponding points is formulated as a hypothesis verification problem. Given

Proceedings of the 1995 International Conference on Image Processing (ICIP '95)
0-8186-7310-9/95 $10.00 © 1995 IEEE

a set of N points, one from each of N perspective projection images, under the hypothesis that the set of points is in correspondence, the optimal estimate of the 3-D point is computed. The triangulation problem is formulated as a Bayesian maximum *a posteriori* estimation problem, assuming the random perturbations in the observed 2-D points and the camera parameters to be Gaussian distributed with known covariance matrices. Based on this maximum *a posteriori* estimate of the 3-D point, a statistical test is derived for verifying the hypothesis that the given image points are projections of the same 3-D point. To find N-tuples of corresponding points from N perspective projection images, a method is proposed that does the Bayesian triangulation and hypothesis verification test on each N-tuple of points, and accepts those that pass the hypothesis test.

2. A BAYESIAN FORMULATION OF THE TRIANGULATION PROBLEM

The triangulation problem – the problem of estimating the co-ordinates of a 3-D point, given its projections in N perspective projection images, and estimates of the parameters of the N cameras – can be posed as a Bayesian estimation problem, as follows.

Data:

- 2-D points $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N$ are observed in perspective projection images $I_1, \ldots, I_N$ respectively.
- Estimates $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N$ of the parameters of the cameras of images $I_1, \ldots, I_N$ respectively are given.
- The observed points are the result of random perturbations on the perspective projections of a 3-D point $\tilde{\mathbf{q}}$ in the respective images, i.e.

$$\hat{\mathbf{x}}_i = P(\tilde{\mathbf{q}}, \theta_i) + \xi_i, i = 1, \ldots, N$$

  where $P(\tilde{\mathbf{q}}, \hat{\theta}_i)$ denotes the perspective projection of the 3-D point $\tilde{\mathbf{q}}$ in a camera with parameter vector θ_i. The random perturbations $\xi_i, i = 1, \ldots, N$ are assumed independent of each other. $\xi_i \sim N(\mathbf{0}, \Sigma_{\hat{\mathbf{x}}_i}), i = 1, \ldots, N$, where $\Sigma_{\hat{\mathbf{x}}_1}, \ldots, \Sigma_{\hat{\mathbf{x}}_N}$ are given.
- The 3-D point $\tilde{\mathbf{q}}$ is considered as a random variable with a certain *a priori* density $p(\tilde{\mathbf{q}})$.
-

$$\hat{\theta}_i = \theta_i + \eta_i, i = 1, \ldots, N$$

  where $\eta_i \sim N(\mathbf{0}, \Sigma_{\hat{\theta}_i}), i = 1, \ldots, N$. $\Sigma_{\hat{\theta}_1}, \ldots, \Sigma_{\hat{\theta}_N}$ are given. $\eta_1, \ldots, \eta_N$ are independent of each other.
- The true camera parameters $\theta_1, \ldots, \theta_N$ are considered as random variables with independent *a priori* densities $p(\theta_1), \ldots, p(\theta_N)$ respectively.

Problem:
Estimate $\mathbf{q} = (x, y, z)^T$ to maximize

$$p(\mathbf{q} \mid \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N)$$

i.e. to maximize

$$p(\mathbf{q}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N)$$

The method of solution is outlined below. The details are given in [1].

Under the assumptions that the covariance matrices of the camera parameters have small diagonal entries, and that the perturbations in the camera parameters are independent of each other and of the perturbations in the co-ordinates of the observed points, it is shown in [1] that

$$p(\mathbf{q}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N) \approx p(\mathbf{q}) \prod_{i=1}^{N} p(\hat{\mathbf{x}}_i \mid \mathbf{q}, \hat{\theta}_i) p(\hat{\theta}_i)$$

It is also shown in [1] that, to a first order approximation,

$$p(\hat{\mathbf{x}}_i \mid \mathbf{q}, \hat{\theta}_i) = N(P(\mathbf{q}, \hat{\theta}_i), \hat{\Sigma}_i(\mathbf{q}, \hat{\theta}_i))$$

where

$$\hat{\Sigma}_i(\mathbf{q}, \hat{\theta}_i) = \Sigma_{x_i} + \left[\frac{\partial P}{\partial \theta}(\mathbf{q}, \hat{\theta}_i)\right] \Sigma_{\hat{\theta}_i} \left[\frac{\partial P}{\partial \theta}(\mathbf{q}, \hat{\theta}_i)\right]^T$$

Maximizing $p(\mathbf{q}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N)$ is the same as minimizing $-\ln p(\mathbf{q}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N)$. If the 3-D point $\mathbf{q}$ and the camera parameters $\hat{\theta}_i$ are *a priori* uniformly distributed, the problem reduces to finding $\mathbf{q}$ to minimize

$$\begin{aligned} &\epsilon(\mathbf{q}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N) \\ &= \sum_{i=1}^{N} \Big\{ (\hat{\mathbf{x}}_i - P(\mathbf{q}, \hat{\theta}_i))^T \left[\hat{\Sigma}_i(\mathbf{q}, \hat{\theta}_i)\right]^{-1} (\hat{\mathbf{x}}_i - P(\mathbf{q}, \hat{\theta}_i)) \\ &\quad + \ln \left|\hat{\Sigma}_i(\mathbf{q}, \hat{\theta}_i)\right| \Big\} \end{aligned} \tag{1}$$

The minimization can be done iteratively by a steepest descent procedure. The details of the computation of the gradient, the initial guess and the length of the step to be taken at each iteration are in [1].

3. FINDING CORRESPONDING POINTS

It is shown in [1] that the maximum *a posteriori* estimate of the 3-D point with uniform prior densities for the 3-D point $\mathbf{q}$ and the camera parameters $\theta_1, \ldots, \theta_N$ is equivalent to the maximum likelihood estimate for $\mathbf{q}$ under the following noise model :

$$\hat{\mathbf{x}}_i = P(\mathbf{q}, \hat{\theta}_i) + \xi_i, i = 1, \ldots, N$$

where ξ_i is distributed as $N(\mathbf{0}, \hat{\Sigma}_i(\mathbf{q}, \hat{\theta}_i))$. It is also shown that at the estimated 3-D point $\hat{\mathbf{q}}$, the function

$$\begin{aligned} &\delta(\hat{\mathbf{q}}, \hat{\mathbf{x}}_1, \ldots, \hat{\mathbf{x}}_N, \hat{\theta}_1, \ldots, \hat{\theta}_N) \\ &= \sum_{i=1}^{N} (\hat{\mathbf{x}}_i - P(\hat{\mathbf{q}}, \hat{\theta}_i))^T \left[\hat{\Sigma}_i(\hat{\mathbf{q}}, \hat{\theta}_i)\right]^{-1} (\hat{\mathbf{x}}_i - P(\hat{\mathbf{q}}, \hat{\theta}_i)) \end{aligned} \tag{2}$$

is distributed as a χ^2_{2N-3} random variable, where N is the number of image points used for the triangulation. This can be used as a statistic for a test of a given probability of type I error to verify the hypothesis that the given set of 2-D points are in correspondence. Let $\delta_{\min}$ be the value of this test statistic at the estimated minimum. Let x_α be such

Proceedings of the 1995 International Conference on Image Processing (ICIP '95)
0-8186-7310-9/95 $10.00 © 1995 IEEE

that $P(\chi^2_{2N-3} > x_\alpha) = \alpha$. If $\delta_{\min} > x_\alpha$, the hypothesis that the N points are in correspondence is rejected.

To find corresponding N-tuples of points from N perspective projection images, the following method can be used. For each N-tuple of points containing one point from each image, under the hypothesis that those N points are in correspondence, the Bayesian triangulation and hypothesis verification are done. If an N-tuple passes the test, that set of points is selected as a correspondence. The threshold for the hypothesis test can be set to limit the probability of rejecting a correct correspondence to a given value.

4. EXPERIMENTS AND RESULTS

The results of some of the experiments for characterizing the performance of the Bayesian triangulation and point correspondence methods are given here. More details are in [1].

Experiments were done to validate the distribution of the test statistic of the hypothesis test, to characterize the mean distance of the 3-D point estimated by the Bayesian triangulation from the true 3-D point, and to characterize the misdetection and false alarm rates of the point correspondence method. For these experiments the geometry and the camera parameters of model board 2 of the RADIUS data set were used. The camera parameters and their covariance matrices were estimated by a multi-image camera calibration procedure citeken-calibration. 3-D points were generated uniformly from the volume occupied by the model board, and projected into the cameras. The projected points and the camera parameters were perturbed by noise of a given distribution, and input to the Bayesian triangulation and hypothesis test procedure.

The distribution of the test statistic of equation (2) is χ^2_{2N-3} independent of the noise in the 2-D points. Figure 1 shows that for each number N of cameras, by choosing the

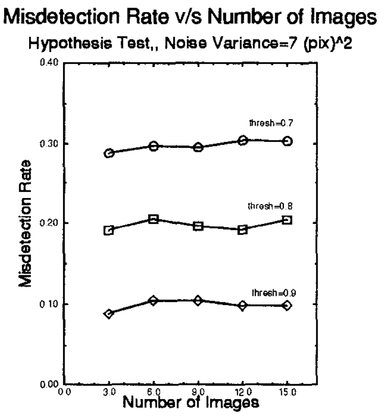

Figure 1: Misdetection rate of hypothesis test v/s Number of images

threshold as the appropriate quantile of the χ^2_{2N-3} distribution, the desired misdetection rate of the hypothesis test can be achieved. Figure 2 shows the variation of the mean and standard deviation of the distance of the estimated 3-D

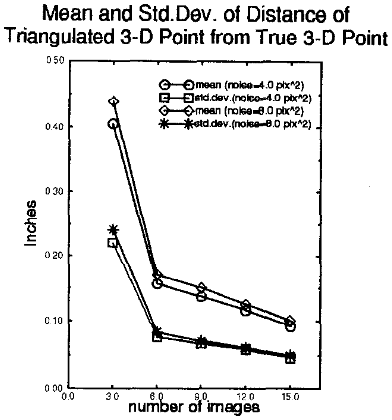

Figure 2: Mean and standard deviation of triangulated 3-D point from true 3-D point v/s Number of cameras (Model board 2 of the RADIUS data set)

point from the true 3-D point as a function of the number of cameras used for the Bayesian triangulation. The model board is about 5 feet$\times$4 feet, and is about 20 feet from the camera. The mean distance of the triangulated 3-D point from the true 3-D point is about 0.4 inch using 3 cameras, and decreases to about 0.1 inch using 15 cameras.

To characterize the misdetection and false alarm rates of the point correspondence method, 20 3-D points were uniformly generated from the volume of model board 2 of the RADIUS data set, and projected into 3 cameras. The projected points and the camera parameters were perturbed and input to the point correspondence method, which does the Bayesian triangulation and hypothesis test on each triple of points, accepting those that pass the test. 10 such trials were done for each value of the variance of the noise in the image points. The variation of the misdetection and false alarm rates of this method are shown in figures 3 and

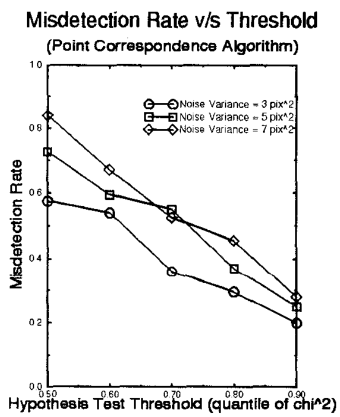

Figure 3: Misdetection rate of point correspondence method v/s Threshold

4 respectively, as functions of the threshold used for the

Proceedings of the 1995 International Conference on Image Processing (ICIP '95)
0-8186-7310-9/95 $10.00

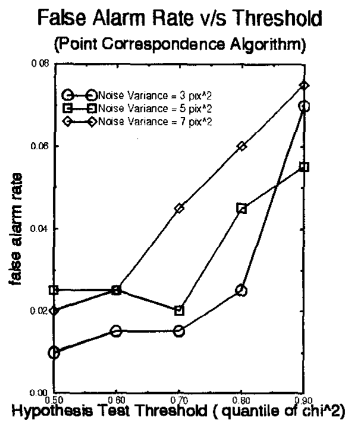

Figure 4: False alarm rate of point correspondence method v/s Threshold

hypothesis test.

The point correspondence method was used to find sets of corresponding points from the corners detected by a corner detector [8] from the RADIUS images. A groundtruth database of about 20 corresponding triples of detected corners from three images was prepared. The correspondences detected by the point correspondence method were compared against these groundtruth correspondences. Figure 5(a) shows the misdetection rate, and figure 5(b) shows the false alarm rate of the point correspondence method on the detected corners, as functions of the threshold of the hypothesis test.

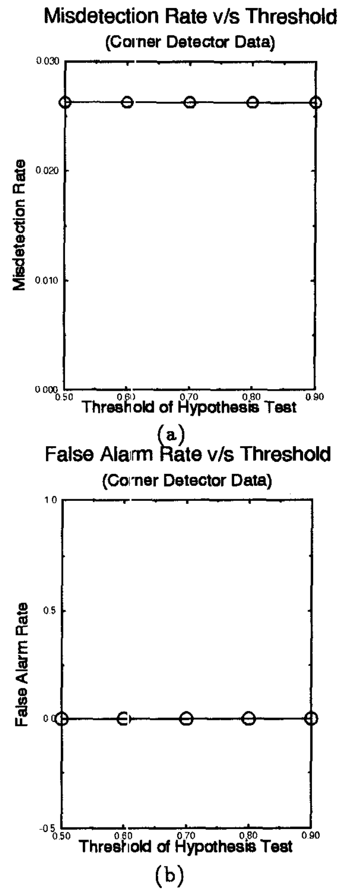

Figure 5: Misdetection and false alarm rates of the point correspondence procedure on the corner detector data: (a)Misdetection rate v/s Threshold of the hypothesis test, (b)False alarm rate v/s Threshold of the hypothesis test

5. SUMMARY

In this paper, a Bayesian formulation of the triangulation problem is given. For a given set of N 2-D points, one from each of N perspective projection images, under the hypothesis that all the points are projections of the same 3-D point, the optimal Bayesian estimate of the 3-D point is computed. Based on this estimate a statistical test is given for verifying the hypothesis that the given points are in correspondence. For finding N-tuples of corresponding points from N perspective projection images, this hypothesis test is done on each N-tuple of points, accepting those that pass the test as N-tuples in correspondence. The mean distance of the 3-D point estimated by the Bayesian triangulation from the true 3-D point, and the false alarm and misdetection rates of this method for finding corresponding points, are characterized.

6. REFERENCES

[1] Anand Bedekar. Finding corresponding points based on Bayesian triangulation. Master's thesis, Dept. of Electrical Engg., University of Washington, 1995.

[2] R.M.Haralick and L.G.Shapiro. *Computer and Robot Vision*, volume II. Reading, MA:Addison-Wesley, 1992.

[3] H. Schmid. An analytical treatment of the problem of triangulation by stereophotogrammety. *Photogrammetria*, (3):66–77 and 91–116, 1956.

[4] G.H.Schut. An analysis of methods and results in analytical aerial triangulation. *Photogrammetria*, (1):16–32, 1957.

[5] R.I.Hartley and P.Sturm. Triangulation. In *Proc. ARPA Image Understanding Workshop*, pages 957–966, 1994.

[6] G.L.Scott and H.C.Longuet-Higgins. An algorithm for associating the features of two patterns. In *Proc. Royal Society, London*, volume B244, pages 21–26, 1991.

[7] Y-Q.Cheng, R.T.Collins, A.R.Hanson, and E.M.Riseman. Triangulation without correspondences. In *Proc. ARPA Image Understanding Workshop*, pages 993–1000, 1994.

[8] X.Zhang, V.Ramesh, R.M.Haralick, and Anand Bedekar. A Bayesian corner detector: Theory and performance evaluation. In *Proc. ARPA Image Understanding Workshop*, pages 703–715, 1994.

Proceedings of the 1995 International Conference on Image Processing (ICIP '95)
0-8186-7310-9/95 $10.00