

Estimating 3-D Movement of a Rigid
Object: Experimental Results

J.-Q. Fang* and T.S. Huang
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
1101 W. Springfield Avenue
Urbana, Illinois 61801

ABSTRACT

We report some preliminary experimental results on estimating general three-dimensional motion parameters of a rigid body from two consecutive image frames taken by a television camera. The approach is to first extract corner point coordinates to subpixel accuracy from grayvalue images, then match them between two frames, and finally solve a set of non-linear equations to obtain the motion parameters. Two factors affect the accuracy of solutions: The magnitude of the motion parameters and the errors in the raw data. These are discussed in the analysis of the experimental results.

I. Introduction

Three-dimensional motion estimation has applications in many areas, including target tracking and robotic vision. In addition, it is closely related to the problem of calibrating stereo cameras. In recent years, there have been a number of papers on this topic, see, e.g., [1,2]. However, we have not seen in the open literature work which reported the estimation of general 3-D motion parameters from real-world images. The present paper is a continuation of our previous work [3,4]. It gives some experimental results on estimating general 3-D motion parameters from a pair of consecutive image frames taken with a television camera. The approach is to first extract corner point coordinates to subpixel accuracy, then match them between the two image frames, and finally solve a set of nonlinear equations to obtain the motion parameters. In [3], some computer simulation results were given. Some experimental results on television images were presented in [4] where, however, the corner

point coordinates were determined to only pixel accuracy. The relatively poor numerical results of [4] prompted us to develop a subpixel corner detector and apply it to motion estimation, it is the main goal of the present paper to report some preliminary results along this direction.

II. Algorithms

Our approach to estimating 3-D motion parameters of a rigid object with small rotation from two successive images consists of the following steps:

- 1) Evaluating the gradient magnitude and direction of grayvalue images by the Sobel operator.
- 2) Detecting corner candidates by finding the local maxima of the magnitude of the gradient direction, the detail of which is given in Fang and Huang [4].
- 3) Distinguishing corner points from noise points by the Hough transform after detecting the corner candidates, also given in Fang and Huang [4].
- 4) Determining corner positions to subpixel accuracy.
- 5) Matching the sets of corner points in the two frames by a point pattern matching method.
- 6) Solving a set of non-linear equations to obtain the 3-D motion parameters using the coordinates of corresponding points as the raw data.

The algorithm for determining corner positions to subpixel accuracy is to first find edge point positions with subpixel accuracy using gradient magnitudes, then fit an edge line through these edge points, and finally intersect the edge lines to get a corner position. We will give the details of the algorithm in a future publication.

The match algorithm we used is based on the work of Rosenfeld [5], but we modified it to accommodate the cases where the scales of

* On leave from North China Institute of Computing Technology, Peking, China. His permanent address is "P.O. Box 619, Peking, China (PRC)".

the object in corresponding frames are different.

The basic non-linear equations we solved are of the form:

$$\begin{aligned} & (-F^2 + Y_i Y_i') \Delta x' + X_i' Y_i' \Delta y') \phi_1 + (X_i Y_i' \Delta x' \\ & - (F^2 + X_i X_i') \Delta y') \phi_2 + (X_i \Delta x' + Y_i \Delta y') F \phi_3 \\ & + F \Delta Y_i \Delta x' - F \Delta X_i \Delta y' + F^2 X_i' \phi_1 + F^2 Y_i' \phi_2 \\ & - F(X_i X_i' + Y_i Y_i') \phi_3 - X_i' \Delta Y_i F + Y_i' \Delta X_i F = 0 \\ & i=1,2,\dots,n \end{aligned} \quad (1)$$

where (X_i, Y_i) and (X_i', Y_i') are the coordinates of the i th point in image frame 1 and frame 2, respectively, and the unknowns are

$$\begin{aligned} x' &= \frac{F x}{z}, \quad y' = \frac{F y}{z}, \quad \phi_1 = n_1 \theta, \quad \phi_2 = n_2 \theta, \\ \phi_3 &= n_3 \theta \end{aligned} \quad (2)$$

The translation vector is (x, y, z) . The rotation is around an axis through the origin with directional cosine (n_1, n_2, n_3) . The rotation angle is θ . F is a constant. Algorithms for solving these equations (which are mainly the Newton and Levenberg-Marquardt methods) have been discussed in Ref. [3].

III. Results

General description

A major difficulty in estimating 3-D motion parameters from real-world images is that the non-linear equations (1) are very sensitive to errors in the raw data, i.e. the coordinates of point pairs. We have experienced this in a simulation study described in [3]. The situation was worse when we used a digital CID camera with low resolution as an input device. The point array of the camera was only 240 x 240 pixels. Some results of motion estimation from the CID camera images are listed in Table 1. Although these results contain fairly large errors (20%), they indicate the feasibility of our approach. The object in the images in our experiments was a robot arm holding a wooden cube. In Figure 1, we show some pairs of typical images we processed.

We processed dozens of picture pairs with a variety of motion. Only some of the results are listed in Table 1. In the table, "ideal" denotes ground truth. The cases "sub" and "pixel" denote the solutions of the motion equations obtained from the raw data with subpixel accuracy and pixel accuracy, respectively.

Motion parameters

An important discovery in our experiments is the big influence of the magnitudes of motion parameters upon the accuracy of the solution of the non-linear equations (1).

1) A large angle of rotation caused slow convergence or even divergence of the iteration in solving the equations (1), whereas too small an angle also made the solution less accurate. The promising range of rotation angle in our experiments appeared to be 2-4 degrees.

2) There was mutual interference between the parameters of rotation and those of translation. General speaking, the estimates of rotation parameters were more accurate than those of translation parameters. But large translation degraded the results of rotation parameters. The reason of this phenomenon seems to be that the 2-D projection of a rotation in the 3-D space, say around the y axis, is similar to that of a translation in the x direction. This is likely to confuse human vision as well. So a moderate amount of translation (for example, x' and y' less than 0.5) tended to yield good results in our experiments.

3) A large translation in the z direction (which made a big difference in scale between two frames) helped to obtain good results of motion parameters.

In summary, the magnitudes of the motion parameters have a big impact on the accuracy of the solution of equations (1), because these parameters determine the characteristics of the non-linear equations (1). Investigation of this influence can be helpful in deciding on the time intervals between snapshots.

Accuracy of raw data

Another factor which affected the accuracy of the solution is the accuracy of the raw data, i.e., the coordinates of corner points in the two frames. For each image pair, we computed using the raw data with sub-pixel accuracy and those with pixel accuracy. In Table 1, we list these results in adjacent lines. Comparing these, one concludes that the results in the "subpixel" case are generally better than those in the "pixel" case. In theory, the results in the "subpixel" case should be much better than those listed in Table 1, but in practice, there are other factors which degrade the estimation accuracy. The distortion of lenses and the physical deformation of the corners of objects are two examples.

IV. Conclusion

We have reported in this paper some first experimental results on estimating general 3-D

motion parameters of a rigid body from two consecutive image frames taken with a television camera. The approach is to first extract corner points, then match the corners in the two frames, and finally solve a set of non-linear equation iteratively to obtain the motion parameters. It was found that the solution to the non-linear equations was very sensitive to errors in the corner point coordinates, and that determining these coordinates to subpixel accuracy improved the motion parameter estimation results.

Acknowledgement

This work was supported by National Science Foundation Grant MCS-8206926.

REFERENCES

- [1] Dreschler, L., and W.-H. Nagel, "Volumetric Model and 3-D Trajectory of a Moving Car Derived from Monocular TV-Frame Sequences of a Street Scene", Proc. IJCAI-81, pp. 692-697.
- [2] Roach, J.W., and J.K. Aggarwal, "Determining the Movement of Objects from a Sequence of Images", IEEE Trans. PAMI-2, Nov. 1980, pp. 554-562.
- [3] Fang, J.-Q., and T.S. Huang, "Solving Three Dimensional Small Rotation Motion Equations: Uniqueness, Algorithms and Numerical Results", submitted to Computer Graphics and Image Processing.
- [4] Fang, J.-Q., and T.S. Huang, "A Corner Finding Algorithm for Image Analysis and Registration", Proceedings AAAI-82, Pittsburgh, Pennsylvania, Aug. 18-20, 1982, pp. 46-49.
- [5] Panade, S. and A. Rosenfeld, "Point Pattern Matching by Relaxation", PR-12 (1980), pp. 269-275.

Table 1. Results of estimation of 3-D motion parameters

Name	Case	x'	y'	n_1	n_2	n_3	(radians)	z
dr4-de31	ideal	0.58	0.00	0.00	1.00	0.00	-0.052	11.7
	sub	0.39	0.10	-0.11	-0.95	0.28	0.065	
	pixel	0.05	0.06	-0.04	-0.996	0.08	0.087	
cr11-cr2	ideal	0.44	0.00	0.00	1.00	0.00	-0.061	-5.9
	sub	0.07	0.08	0.05	0.994	-0.09	-0.056	
	pixel	0.05	0.31	0.21	0.76	-0.61	-0.072	
fr12-fr2	ideal	0.00	-0.25	0.00	1.00	0.00	-0.0698	-10
	sub	-0.13	-0.33	-0.01	-0.998	0.06	0.0683	
	pixel	0.18	-0.08	0.10	-0.99	-0.06	0.0838	
cr42-cr5	ideal	0.00	0.66	0.00	1.00	0.00	-0.0698	8
	sub	-0.05	0.28	0.31	-0.92	-0.21	0.078	
	pixel	0.60	-0.29	0.00	1.00	0.00	-0.061	
gr1-gr21	ideal	0.60	-0.29	0.00	1.00	0.00	-0.061	-10.7
	sub	0.52	-0.26	-0.00	-0.998	-0.06	0.052	
	pixel	0.09	0.15	0.56	-0.75	-0.33	0.039	
gr12-gr3	ideal	0.44	-0.24	0.00	1.00	0.00	-0.0698	-12.8
	sub	0.13	-0.06	0.20	-0.97	-0.10	0.051	
	Pixel	0.03	-0.24	-0.04	-0.99	-0.14	0.046	

Figure 1. Results of matching point patterns. The 2 points of a matched pair are labeled with the same number.

