

Refinement of Noisy Correspondence Using Feedback from 3-D Motion *

Yong C. Kim and Keith Price

Institute for Robotics and Intelligent Systems
Departments of Computer Science and Electrical Engineering
University of Southern California
Los Angeles, CA 90089-0273

Abstract

In an automated feature-based motion analysis of multiple frames, correspondence data are usually noisy and fragmented. We developed a technique that gradually refines the initial noisy correspondence data and links fragments of a single trajectory using feedback from 3-D motion estimation. First, 3-D motion parameters are estimated using the initial correspondence data. Then, each noisy trajectory is partitioned into subsets of points each of which conforms to the estimated motion. The best set is used as the input to the next motion estimation. This process is repeated and the gaps in the refined correspondence data are filled by guidance from the predicted motion. Test results for a standard real image sequence are presented.

1 Introduction

Motion analysis is one of the important research areas in computer vision. Correspondence problem is a key issue in feature-based motion analysis. Sethi and Jain [5] suggested a method of establishing the correspondence over multiple frames, which is based on the smoothness of motion.

We show that an improvement can be achieved with an integrated approach, where feature extraction, correspondence and motion analysis are performed in cooperative manner, exchanging information among the separate sub systems. This work is in the domain of feature-based motion analysis and concentrates on the use of 3-D motion to improve the matching of features and the resulting motion and structure estimation. Initial noisy correspondence data are gradually refined and fragments of a trajectory of a single feature are linked using feedback from the 3-D motion estimation.

*This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Office of Scientific Research under Contract No. F49620-90-C-0078. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

After a brief description of the system in section 2, refinement and linking of noisy correspondence are described in section 3. Finally, test results on real image data are presented, followed by a brief conclusion.

2 Description of Integrated System

The block diagram in figure 1 shows the subsystems and the feedback of information. The basic features are *regions* and *corners* extracted from the contour of the regions. Each image in the sequence is segmented into regions by a recursive splitting technique that uses the statistics of image attributes [4]. The segmentation of the current frame is guided by the result of the previous frame.

Feature matching is done in a hierarchical way to speed up computation and increase stability. Region correspondences are first established and corners are matched along the matched regions. Relaxation-based symbolic matching [2] is used for both region and corner matching. 3-D trajectory and structure are estimated with chronogeneous motion analysis [3], which can handle uniform acceleration with constant translation and rotation.

3 Refinement and Linking

The main sources of noise in correspondence data are inconsistent feature extraction and erroneous feature match. When one or two spurious points cause a trajectory to be partially faulty, part of such trajectories are still usable data for motion estimation. We developed a method where a trajectory is analyzed and spurious points are detected by clustering of points in the trajectory.

The noisy data are continuously refined in the feedback loop. First, motion parameters are estimated using the raw correspondence data. Then, each noisy trajectory is partitioned into overlapping subsets of points each of which conforms to the estimated motion. The best set is selected as the input for the next motion estimation. This selection process is repeated 3 or 4 times until a set of reliable correspondence is selected.

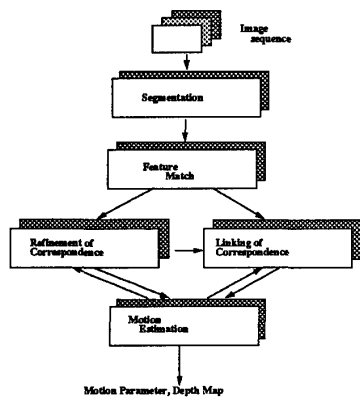


Figure 1: Overall Block Diagram

After the refinement, each point in a trajectory is marked as S (Selected) or D (Discarded). Refined correspondences are extended with the guidance from predicted motion. For each frame, a window is set which is centered at the predicted position of the point. The match strengths of all the corners in the window with respect to all the S corners in the trajectory are averaged. The best corner is selected as the one to be linked to the trajectory. Linked correspondences are refined in the same way since new spurious points may have been added in the linking process.

3.1 Clustering of Points in a Trajectory

Given a motion estimate M , and a trajectory of points (A B C D), point pairs (A B) and (C D) are defined to be compatible if C and D can be approximately obtained by extrapolating (A B) under M , and *vice versa*. Partitioning is based on *clustering of compatible point pairs*. Of the clusters with compatibility above a threshold, the largest set of points is chosen for the input to the next motion estimation.

3.2 Extrapolation from a Point Pair

Motion analysis produces the depth of each point feature. This estimated depth is not used in the 3-D extrapolation since it is obtained for the entire (possibly noisy) trajectory.

Given 2-D points P_i, P_j and an estimate of motion parameters, the 3-D points Q_i, Q_j which are projected onto P_i, P_j in the image plane are computed as follows.

Let $Q_i = K_i P_i$ and $Q_j = K_j P_j$. Then $R_{i,j} Q_i + T_{i,j} = Q_j$, where $T_{i,j}$ and $R_{i,j}$ are the translation vector and rotation matrix from frame i to j , respectively. Let $A = [R_{i,j} P_i | -P_j]$. Then,

$$[K_i K_j]^t = -(A^t A)^{-1} A^t T_{i,j}$$

If either K_i or K_j is negative, the point pair (P_i, P_j)

is not used. The 3-D position of the point, Q_k , at an arbitrary frame k , is obtained from $R_{i,k} Q_i + T_{i,k}$.

4 Results

The algorithm has been tested for standard sets of real image sequences. We present the results for the Rocket Field sequence [1]. Of the 30 frames of the sequence, we used the first 15 frames.

The initial and final corner correspondences are shown in figure 2. Figure 2 (a) represents the initial noisy correspondence, where some of the noisy trajectories move in random direction. The refined data are linked and refined again as in figure 2 (b). The reconstructed trajectories of the objects for which ground truth is provided are shown in figure 2 (c). The start/end frame for those matched corners are arbitrary and thus the end points of each trajectory may not point to any physical object on the underlying image, which is the first frame of the sequence.

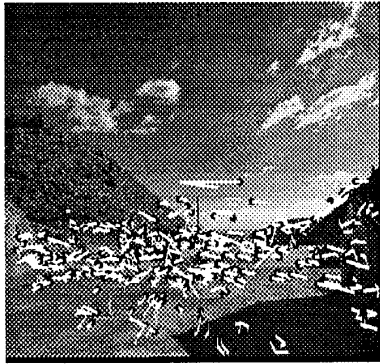
Figure 3 (a), (b) show the top view of objects with ground truth, estimated from the initial noisy correspondence and the final correspondence data, respectively. The trajectory of the camera is numbered as -1. For comparison, the top view computed from ground truth is shown in figure 3 (c).

5 Conclusion

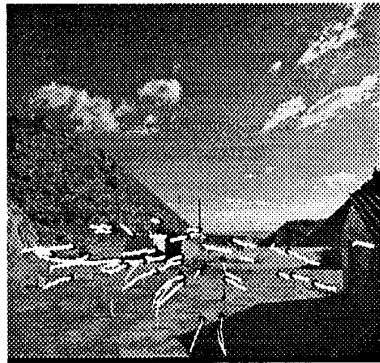
We presented a feedback approach to improve the matching of features in the domain of feature-based motion analysis for multiple frames. We showed that initial noisy correspondence data are gradually refined and fragments of a single feature trajectory are linked using the feedback from 3-D motion estimation. We have applied this method to real image sequences of standard data set and showed that the motion and structure estimation is improved with the improved correspondence.

References

- [1] R. Dutta, R. Manmatha, L. Williams, and E. Riseman. A data set for quantitative motion analysis. In *Proceedings of the Conference on CVPR*, pages 159–164, 1989.
- [2] O. D. Faugeras and K. Price. Symbolic description of aerial images using stochastic labeling. *IEEE Transactions on PAMI*, 3(6):633–642, Nov. 1981.
- [3] W. Franzen. Structure and motion from uniform 3D acceleration. In *Proceedings of the Workshop on Visual Motion*, pages 14–20. IEEE Computer Society, Oct 7–9, 1991.
- [4] R. Ohlander, K. Price, and R. Reddy. Picture segmentation by a recursive region splitting method. *Computer Graphics and Image Processing*, 8:313–333, 1978.
- [5] I. K. Sethi and R. Jain. Finding trajectories of feature points in a monocular image sequence. *IEEE Transactions on PAMI*, 9(1):56–73, Jan. 1987.



(a) Initial correspondence

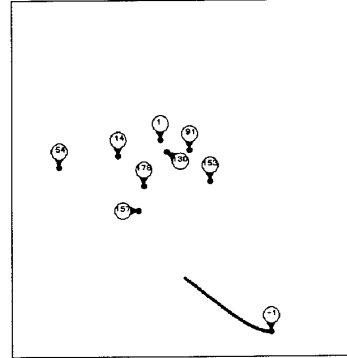


(b) Final correspondence

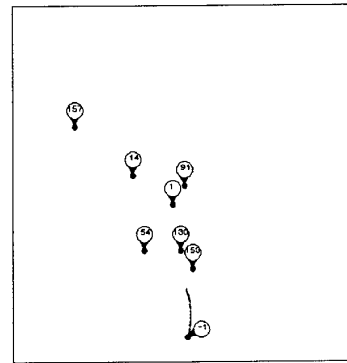


(c) Reconstructed from final match

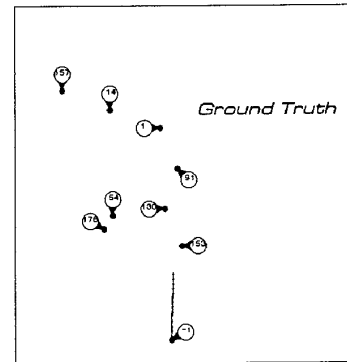
Figure 2: Improvement of correspondence



(a) Top view from initial match



(b) Top view from final match



(c) Top view from ground truth

Figure 3: Reconstructed trajectory and top view of objects with ground truth.