Monocular Vision based Independently Moving Feature Detection using Image Correspondences

Sijong Kim\textsuperscript{1}, Jungwon Kang\textsuperscript{2} and Myung Jin Chung\textsuperscript{3}

\textsuperscript{1,2}Department of Electrical Engineering, KAIST, Daejeon, 305-701, Korea (Tel: +82-42-350-5429; E-mail: sunday10am@gmail.com, kctown99@gmail.com)

\textsuperscript{3}Department of Electrical Engineering, KAIST, Daejeon, 305-701, Korea (Tel: +82-42-350-3429; E-mail: mjchung@ee.kaist.ac.kr) * Corresponding author

Abstract: In this paper, we propose a independently moving feature detection algorithm using monocular vision. The proposed algorithm can detect independently moving feature points from all of the feature points by geometric constraints. The geometric constraints consist of epipolar constraint and trifocal constraint. The proposed algorithm can be implemented in real time. Therefore, it can be directly apply to mobile robots and intelligent vehicles. We test our algorithm with real outdoor datasets.

Keywords: monocular vision, image correspondence, independently moving feature detection

1. INTRODUCTION

Detecting of moving objects is important to the development of several applications of mobile robots and intelligent vehicles. If the mobile robots can detect moving objects then they can avoid collision with moving obstacles and other mobile robots. A variety of approaches exist for detecting moving parts of the scene using image correspondences [1] [2]. J. Klappstein and et al. has discussed the detection limits of independently moving objects utilizing available constraints existing for static 3D points. For detection limit analysis, they consider the three most frequent kinds of motion in traffic and analyze that. Several papers deal with moving object detection algorithms using stereo vision [3] [4] [5]. In general, the approaches using stereo vision can detect the moving objects efficiently. Because they can use the 3D information. For more robust detection, some papers cope with sensor fusion methods that exploit the stereo vision and the laser scanner [6] [7]. Image based approaches are prone to errors due to digital quantization and triangulation. To overcome these problems, they use sensor fusion methods.

In this paper, our goal is to detect the independently moving feature point from all of the feature points. We adopted geometric constraints for error criteria. We use the real outdoor datasets that contain pose information. Therefore, we only focus on detecting independently moving feature points.

The remainder of this paper is organized as follows. Section 2 presents the proposed algorithm. Section 3 describes experimental results. Conclusions and further work are discussed in Section 4.

2. PROPOSED MOVING FEATURE DETECTION ALGORITHM

2.1 Overview

An overview of our algorithm is shown in Fig. 1. The proposed algorithm consists of two modules. The first module estimates a distance error $e_1$ over two views using epipolar constraint. The second module estimates a distance error $e_2$ over three views using trifocal constraint. Finally, we can detect the independently moving feature points using these two errors. If both of the distance errors of tracked feature point is higher than predefined thresholds $\lambda_1$ and $\lambda_2$ then the feature is determined by independently moving feature.

2.2 Moving Feature Detection with Epipolar Constraint

The geometrical relations over three views can be seen in Fig. 2. The motions of interframe and the results of tracking by KLT are given. Fig. 2 illustrates the epipolar constraint. The epipolar constraint signifies that the viewing rays of a static 3D point must meet [8]. A moving features violate the epipolar constraint. Therefore, we can compute the distance error $e_1$ in the second view as in (1) where $t_2 = \begin{bmatrix} a & b & c \end{bmatrix}^T$ is the epipolar line in the second view and $p_2 = (u_2, v_2)$ is the corresponding
The sequence that we chose for the experiments taken from a moving vehicle in real outdoor environments [9]. The trifocal tensor plays an analogous role in three views to that played by the fundamental matrix in two views [8]. It encapsulates all the geometric relations between three views that are independent of scene structure. If we set the camera matrices for the three views to \( P_1 = [I|0] \), and \( P_2 = [A|a_4], P_3 = [B|b_4] \) then we can calculate the set of three matrices \( \{T_1, T_2, T_3\} \) constitute the trifocal tensor as in (2) where \( A \) and \( B \) are 3 x 3 matrices, and the vectors \( a_i \) and \( b_i \) are the \( i \)-th columns of the respective camera matrices for \( i = 1, \ldots, 4 \).

\[
T_i = a_i^T b_4 - a_4 b_i^T
\]

(2)

Using the trifocal tensor, we can obtain the transferred point \( \hat{p}_3 \) in third view using relations involving point \( p_1 \) in the first view and the epipolar line \( l_2 \) in the second view as in Eq. (3).

\[
\hat{p}_3 = \left( \sum_i p_i^T T_i \right) l_2
\]

(3)

Finally, we can compute the distance error \( e_2 \) in the third view as in (4) where \( \hat{p}_3 = (\hat{u}_3, \hat{v}_3), p_3 = (u_3, v_3) \).

\[
e_2 = \|\hat{p}_3 - p_3\| = \sqrt{(\hat{u}_3 - u_3)^2 + (\hat{v}_3 - v_3)^2}
\]

(4)

3. EXPERIMENTAL RESULTS

We tested our algorithm with the Karlsruhe dataset that contains both vehicle pose information and image sequences taken from a moving vehicle in real outdoor environments [9]. The sequence that we chose for the test consists of 2579 frames. Each frame is an image of 1344 x 391 pixels.

Fig. 4(a) shows the frist image among three input images. At this time, only the independently moving objects are a man who are crossing a road and a man on the bicycle. Fig. 4(b) shows the extracted features. These features are used for KLT tracking. Fig. 4(c) shows the detected independently moving features with epipolar constraint. The distance error \( e_1 \) of these features are higher than predefined thresholds \( \lambda_1 \). Therefore, these features are determined by independently moving features. Among the detected independently moving features with epipolar constraint, some features also violate the trifocal constraint. Fig. 4(d) shows the detected independently moving features with both constraint. Through this process, we can detect the independently moving feature point from all of the feature points.

4. CONCLUSION

In this paper, we have presented an algorithm for independently moving feature detection using monocular vision. Any pixel point in the second image matching the point in the first image must lie on the epipolar line. Trifocal tensor can be used to transfer points from a correspondence in two views to the coresponding point in a third view. Using these constraints, we computed two distance errors \( e_1 \) and \( e_2 \). Finally, we detected the independently moving feature points using these two errors. We have tested our algorithm with real outdoor dataset.

Future challenges include a scheme for feature clustering. It is necessary to cluster the independently moving features into moving objects for generating more effective detection results.

5. ACKNOWLEDGEMENTS

The Authors gratefully acknowledge the support from UTRC(Unmanned technology Research Center) at KAIST (Korea Advanced Institute of Science and Technology), originally funded by DAPA, ADD.

REFERENCES


Fig. 4 The results of independently moving feature detection on Karlsruhe Dataset.


