2D Visual odometry based on Probability Data Association Filter

Bo Li[1,b], Shasha Dong[2,a], Liang zhang, Yinhua Xu

[1,] Troop 63956, BeiJing, China

[2]The Engineer Academy, Xuzhou, China

[a]beeflysbee@126.com, [b]efan1986@163.com

Keywords: visual odometry; probability data association (PDA) filtering

Abstract. This paper presents a novel approach based on probability data association (PDA) filtering for estimating a vehicle's trajectory in complex urban environments. We consider feature pairs acquired from consecutive frames as the measurements for the PDA filter to update the ego-motion vector of the camera. Compared to other feature based approaches, our approach presents a recursive filtering algorithm that provides dynamic estimation of ego-motion vector in the presence of falsely feature pairs. Experimental results show that this method provides good robustness under real traffic scenarios.

Introduction

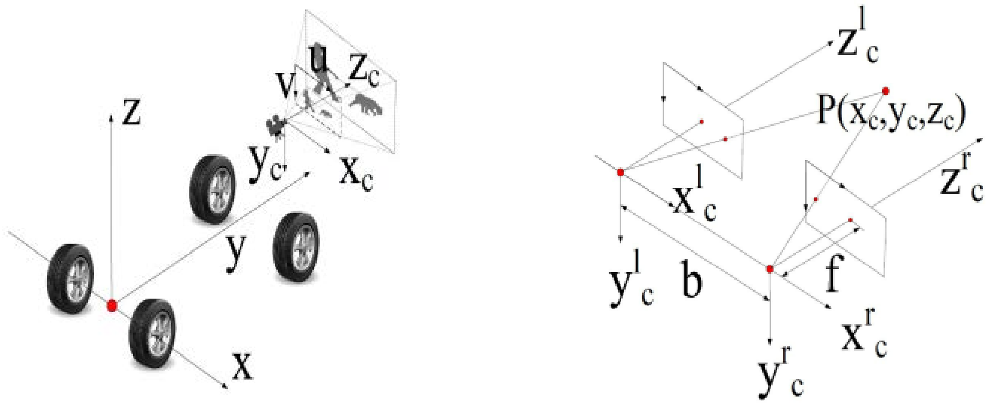
Visual odometry has been introduced in the last decade by using different techniques. At least two kinds of mechanics are proposed: techniques from Structure- from-Motion (SfM) and techniques from optical flow. The SfM technique works by finding good quality features in one frame and the corresponding features in the next frame, estimating displacements from these features and translating them to the motion of the camera. Optical flow is a different approach which focuses on the change in the brightness of the image, where this change in brightness results from the apparent motion in the image. This method is much simpler and computationally cheaper than the extraction and tracking of features [5]. However, the precision is not very good. Most researchers use the SfM technique to estimate the ego-motion vector of the camera. Singular Value De- composition (SVD) is a standard method to estimate the translation and rotation [6]. This method utilizes the whole feature pair set and often leads inaccuracy estimation since the outliers (falsely feature pairs and the features stemming from the moving targets are also taken into the estimation process). The RANSAC has been established as the standard method for motion estimation in the presence of outliers and it achieves its goal by iteratively selecting a random subset of the original data. Garcia [2] uses RANSAC to overcome these outliers in real traffic scenes. The limit of the both methods is as follows: Either SVD method (use the whole set) or RANSAC method (use the inliers set), the final feature pairs have the same weight with each other (if the outliers are still in the final set, they may influence the precision of the estimation since those feature pairs have the same weight). As a matter of fact, there should be an adaptive weight on each feature pair for estimating the ego-motion vector.

We propose a method for ego-motion computation based on the Probability Data Association (PDA) filter [1]. The PDA estimator avoids the association issues by utilizing the whole set of measurements with corresponding weights. As a result, the weight assigned to a given measurement corresponds to the probability that it came from that state.

This paper is structured as follows: Sec. 2 introduces details about the preprocessing phase. Sec. 3 describes the PDA filter and its implementation. Sec. 4 presents experimental results under real traffic scenes. Finally, the paper is concluded in Sec. 5.

Preprocessing Phase

In this paper, we use difference of Gaussians for blob detection to estimate the ego-motion. Interesting points are extracted from 2 consecutive frames both left cam- era and right camera then matched from those 4 images as the feature pairs to estimate the ego-motion vector (translation and rotation). Here the feature pairs are the same pixels between consecutive frames in vehicle coordinates (see Fig. 1(a)), the mapping between the vehicles coordinates and the image coordinates is by using the homograph [3].



(a) Top view of the scene (b) Same scene from the camera view

Fig 1. different coordinates systems

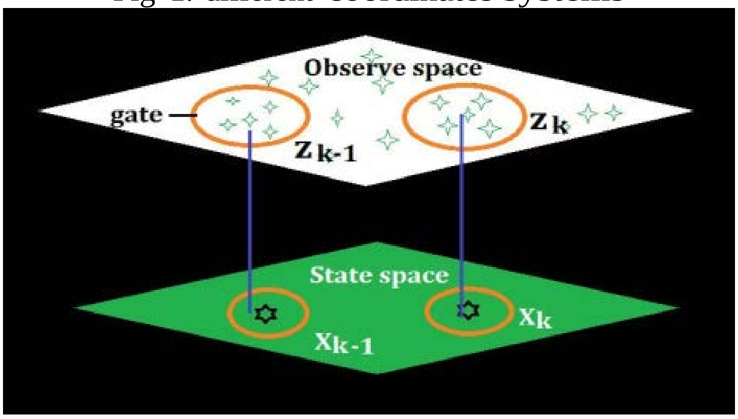


Fig 2. Single target multiple observations

Estimation phase

The Probabilistic Data Association Filter (PDAF) is an alternative approach to update the target's state by using all of the validated measurements with different weights instead of using only one measurement and discarding the others [1]. Fig. 2 shows the PDAF in state space and observes space.

In this paper, we use the PDAF to calculate the ego- motion vector in visual odometry domain. The measurements are the ego-motion vectors that calculated from the feature pairs (more details are in Sec. 3.3). The PDAF calculates the association probabilities for each measurement to estimate the motion state. When there existing falsely matched feature pairs, their corresponding motion vectors' measurements may have low association probabilities.

The Stochastic Model

The dynamic of the target is modeled by the equation

$$x(k) = Fx(k-1) + \omega(k) \tag{1}$$

Where the process noise vector$\omega(k)$ is assumed to be white Gaussian with zero mean and covariance Q, F is the process matrix. The measurement system is modeled as follows:

$$z(k) = Hx(k) + \zeta(k) \tag{2}$$

Where the measurement noise $\zeta(k)$ is taken to be white Gaussian with zero mean and covariance R, H is the measurement matrix.

The PDA Filter

The PDAF is able to make use of several observations $Z_k = \{z_i(k)\}_{i=1}^{m_k}$for each time instant, where m_k is the total number of observations for time k. The cumulative set of measurements up to time k is $Z^k = \{Z_j\}_{j=1}^{k}$Define the events:

$$\theta_i(k) = \{z_i(k) \text{ is the target observation at time } k\} \tag{3}$$

$$\theta_0(k) = \{\text{ all observations are clutter}\} \tag{4}$$

And the probability of each event

$$\beta_i(k) = P\{\theta_i(k)|Z^k\}, i = 0,1,\cdots,m_k. \tag{5}$$

As the events $\{\theta_i(k)\}_{i=0}^{m_k}$ are exhaustive and mutually exclusive, we have $\sum_{i=0}^{m_k}\beta_i(k) = 1$. Moreover, invoking the total probability theorem, the filtered state estimate can be expressed as

$$\hat{x}(k|k) = \sum_{i=0}^{m_k}\hat{x}_i(k|k)\beta_i(k) \tag{6}$$

Where $\hat{x}_i(k|k)$ is the updated state estimate conditioned on $\theta_i(k)$, which is given by the standard Kalman Filter as

$$\hat{x}_i(k|k) = \hat{x}_i(k|k-1) + K(k)\tilde{z}_i(k) \tag{7}$$

Where $\tilde{z}_i(k) = z_i(k) - \hat{z}(k|k-1)$ is the corresponding innovation, and $K(k)$ is the standard Kalman gain. The error covariance associated with the updated state estimate is defined as

$$P(k|k) = \{[x(k) - \hat{x}(k|k)][x(k) - \hat{x}(k|k)]^T|Z^k\} \tag{8}$$

And can be evaluated by

$$P(k|k) = \beta_0(k)P(k|k-1) + [1 - \beta_0(k)]P^c(k|k) + \tilde{P}(k) \tag{9}$$

Where

$$\tilde{P}(k) = K(k)\left[\sum_{i=1}^{m_k}\beta_i(k)\tilde{z}_i(k)\,\tilde{z}_i^T(k) - \tilde{z}(k)\tilde{z}^T(k)\right]K^T(k) \tag{10}$$

And

$$P^c(k|k) = (I - K(k)H)P(k|k-1), \tilde{z}(k) = \sum_{i=1}^{m_k}\beta_i\,(k)\tilde{z}_i(k) \tag{11}$$

The association probabilities are given by

$$\beta_i(k) = \frac{e_i}{\sum_{i=0}^{m_k}e_i} \tag{12}$$

Where

$$e_i = \exp\left(-\frac{1}{2}\tilde{z}_i^T(k)S^{-1}(k)\tilde{z}_i\right), i = 1,\cdots,m_k \tag{13}$$

with $S(k)$ being the innovation covariance, PG being the probability that the target originated measurement falls within the validation gate, PD being the probability that the correct measurement is detected, and cnz being the volume of the nz dimensional unit hyper sphere (c1 = 2, c2 = π, c3 = 4π/3, etc.). γ is the threshold for the validated measurements by defining the following validation region

$$\{z_i: \tilde{z}_i^T(k)S^{-1}(k)\tilde{z}_i(k) \le \gamma\} \tag{14}$$

Implementation Details. In this paper the ego-motion vector is considered as the target state while the dynamic process of the Kalman filter is defined as follows

$$X(k) = [\alpha(k), \dot{\alpha}(k), \eta(k), \dot{\eta}(k), \psi(k), \dot{\psi}(k)]^T \tag{15}$$

$$F = \begin{bmatrix} 1 & \Omega & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Omega & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \Omega \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{16}$$

where **Ω**, (α, η) and ψ are the interval, translation and rotation parameters between consecutive frames.

The measurement is modeled as follows

$$z(k) = [\alpha(k), \eta(k), \psi(k)]^T \tag{17}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{18}$$

In this paper, the measurements of the ego-motion vector are acquired as follows:

Suppose at frame k, the coordinates of the associated feature are x(k), y(k) in vehicle coordinates. According to Euler's rotation theorem the relationship between the features is as follows

$$\begin{bmatrix} x(k+1) \\ y(k+1) \end{bmatrix} = \begin{bmatrix} \cos\psi(k) & -\sin\psi(k) \\ \sin\psi(k) & \cos\psi(k) \end{bmatrix}\begin{bmatrix} x(k) \\ y(k) \end{bmatrix} + \begin{bmatrix} \alpha(k) \\ \eta(k) \end{bmatrix} \tag{19}$$

We cannot directly acquire the ego-motion vector from each feature pair since the above equations are underdetermined system of equations. Since 2 feature pairs can calculate the ego-motion vector as a measurement for the PDA filter, there is a total number of $C_{n(k)}^{2}$ measurements for the PDA filter at frame k according to the principle of permutation and combination, where $n(k)$ is the number of feature pairs. The PDA filter calculates the association probabilities for each of the $C_{n(k)}^{2}$ measurements instead of using only one measurement and discarding the others.

The reasons of selecting 2 feature pairs to acquire a measurement are due to the computational complexity and the diversity of measurements.

The process noise and the measurement noise are acquired from the covariance of the difference between the true value (IMU) and the estimated value (nonlinear least square method). P_D, P_G and γ are also predefined.

Experimental Evaluation

The visual odometry algorithm described in this pa- per has been implemented with the public dataset [4]. All sequences correspond to real traffic conditions in urban environments.

As can be observed, we also use the SVD (with the whole feature pairs), RANSAC (based on the least square method) matlab toolbox [7] and Kitt's approach [4] to estimate the ego-motion vector. The feature pairs for the SVD, RANSAC and PDA are the same on consecutive frames while Kitt estimates trajectory by bucketing fast features and filtering (Kitt's code is available in dataset). There is also a high precision trajectory (acquired from the IMU) to compare the performance of each method on Fig. 3.

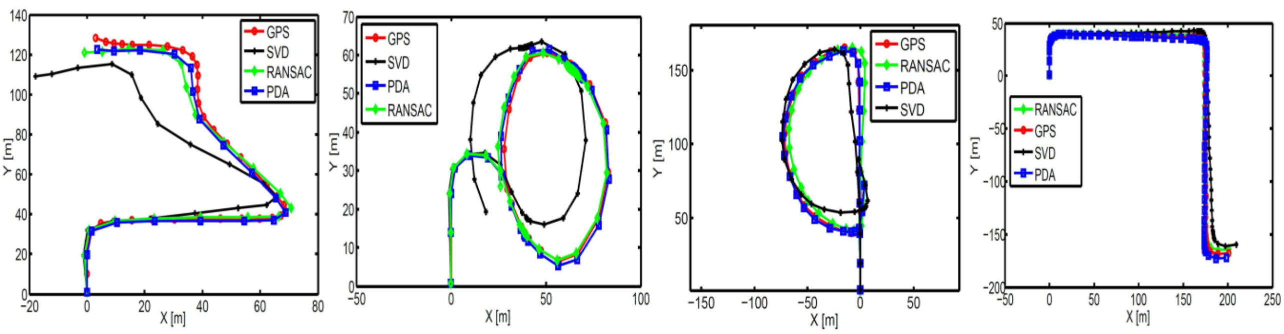


Fig 3. Result

The benefits of the PDA filter are as follows:

The PDA filter uses all the measurements with different weights to estimate ego-motion vector, which ensures the optimal under Bayes framework. Compared to the PDA filter, other methods don't concern about the weight issue and may even lose information. Here are some situations that may influence the performance of those methods: First, the goal of the RANSAC is to remove the outliers from the whole set then estimate the parameters, which may discard useful information. Second, if the features pairs contain a large number of outliers then they may fail. However, the PDA filter avoids the above situations by combining weighted measurements to estimate the state (the"outliers" have a lower weight according to the PDA filter.

As a matter of fact, our approach has its own limit. Since the filter cannot directly acquire the measurements from each feature pair, there is a number of $n(k)$ converted measurements for the PDA filter at each frame. The complexity of PDA growths exponential as the feature pairs increased. However, it is not the fault of PDA filter and the future work is to find a solution to reduce the complexity.

Conclusion

Visual odometry in urban scenes is challenging due to a large amount of outliers. The clutter from falsely matched features causes the results to deviate from the real status. This paper presents an approach of PDA filtering under Bayes framework. In comparison to earlier works, this contribution is among the first to apply a PDA filter to visual odometry in a real traffic scenes. The PDA filter makes use of all the feature pairs' in- formation with different weights instead of using only a part of the information and discarding the others, which ensures the completeness of information and optimal of estimation under Bayes framework. In additional, the PDA filter only use the converted measurements from the feature pairs instead of tracking those features over multiple frames, which is more adaptable in visual odometry system.

The evaluation results show that the PDA filter achieves accuracy and robustness under different scenes.

References

[1] Y. Bar-Shalom and T. E. Fortmann. Tracking and Data Association. Academic Press, San Diego, 1988.

[2] R. Garcia, M. Sotelo, I. Parra, D. Fernandez, and M. Gavilan. 2d visual odometry method for global positioning measurement. In Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on, pages 1~6, oct. 2007.

[3] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.

[4] B. Kitt, A. Geiger, and H. Lategahn. Visual odometry based on stereo image sequences with ransacbased outlier rejection scheme. In Intelligent Vehicles Symposium (IV), 2010 IEEE, pages 486 –492, june 2010.

[5] C. McCarthy and N. Barnes. Performance of optical flow techniques for indoor navigation with a mobile robot. In IEEE International Conference on Robotics and Automation, ICRA '04, pages 5093 – 5098, Vol. 5, May 2004.

[6] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 13(4):376 –380, apr 1991.

[7] M. Zuliani. Ransac toolbox for matlab. http://www.mathworks. com