

Map-Build Algorithm Based on the Relative Location of Feature Points

Cheng Zhao^{1,2}, Fuqiang Zhao^{1,2}, and Bin Kong^{1(✉)}

¹ Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China
bkong@iim.ac.cn

² University of Science and Technology of China, Hefei 230026, China
zhaoch93@mail.ustc.edu.cn

Abstract. Dynamic building road map is a challenging task in computer vision. In this paper we propose an effective algorithm to build road map dynamically with vehicular camera. We obtain top view of video by Inverse Perspective Mapping (IPM). Then extract feature points from top view and generate feature map. Feature map is the description of feature points' distribution and provides the information of camera status. By detecting camera status, we apply different degree of restraint geometric transformation to avoid unnecessary time cost and enhance robustness. Experiment demonstrated our algorithm's effectiveness and robustness.

Keywords: Road map · Inverse perspective mapping · Image stitching

1 Introduction

According to the China Statistical Yearbook 2016 [1], only in 2015 year, traffic accidents killed more than 58000 people in china. Vehicle manufacturers and drivers seek a more secure way of driving, such as Automatic Driving System and Advanced Driver Assistance System (ADAS). Rapid development of Autonomous vehicles and Advanced Driver Assistance System (ADAS) has put higher requirements on the application of computer vision technology, which can provide richer information and more reliable control system.

Over the past few decades, mature technical technology, such as GPS and preloaded digital map [2–5], has brought us great convenience. Although static map has played a significant role in the navigation and location fields, since the data of static map isn't real-time, Dynamical surroundings map can provide more accurate decision-making resources in real time. Under certain scenes, for example, pedestrians are present, dynamical map has advantage.

This project was supported by the Pilot Project of Chinese Academy of Sciences (XDA08040109), by the National Natural Science Foundation of China (No. 913203002) and Science and Technology Planning Project of Guangdong Province, China (No. 2016B090910002).

1.1 Related Work

Some technology has made considerable achievements. Simultaneous localization and map (SLAM) are a great Algorithm. ORB-SLAM [6–8] is one of the most representative algorithms. ORM-SLAM is feature-based Algorithm, consists of three parts, Tracking, Local Mapping and Loop Closing. In the tracking part, ORB-SLAM extracts ORB features of frame, and select key frames to avoid unnecessary redundancy. In the Local Mapping part, the collection of common ORB features set in two selected frames is used to estimate pose of camera. In turn, camera posture is the basis of calculating spatial location of feature points. Bundle adjustment is used to improve the accuracy of feature points. Collections of feature points construct a sparse three-dimensional space mode. The Loop Closing part is helpful to reduce the accumulative error when same scene is detected. SLAM can reconstruct 3D model dynamically, which is foundation of more secure and more convenient control system. Besides ORB-SLAM, there are other well-known algorithms, such as LSD-SLAM [9], PTAM [10].

SLAM system indeed has a wide range of applications, a very large number of experiments has verified the effectiveness of SLAM. However, in some special case of application, there are some weakness in typical SLAM algorithm. Firstly, SLAM algorithm extracts feature points from the whole frame, complex-shape environment can provide more robust feature points. SLAM can't only extract feature of monotonous flat road. That means extracted feature is too sparse to build road map. Secondly, the calculate cost of SLAM is expensive. For embedded device, this is a more serious problem, insufficient calculation time result in lost tracking of camera. Once system can't match the feature of previous key frame, the whole SLAM System get lost.

Inverse Perspective Mapping (IPM) is an alternate algorithm to obtain surrounding information. According to the perspective principle, IPM gets top view of video camera. IPM is a mature algorithm, IPM can get plane information with a small amount of calculation cost, but only 2-dimensional plane information is preserved. For road information building, it's enough, because we need to focus on the road plane sign. IPM is sensitive to shaking of vehicular camera. As a reliable algorithm, IPM has been implemented in many Automatic Driving Systemses and Advanced Driver Assistance System (ADAS). Such as measuring distance [11, 14], vanishing point estimation [12] and detecting [13, 15] (include lane line, obstacle, pedestrian and terrain, etc.).

In this paper we propose a new algorithm. We get top view by Inverse perspective mapping (IPM) in advance. Then we extract feature of the IPM of frame, generated feature map, and based on feature of map's characteristic we simply basic algorithm. In second chapter, we introduce the way to get top view, in third chapter, we discuss the feature map characteristic, and experiment detail is showed in fourth chapter. In last chapter, we sum up the whole algorithm.

2 Parameter Estimation System

Typically, Inverse perspective mapping (IPM) is used to obtain top view of frame, which get from fixed-posture camera. Top view is a scaled-down of real world plane. We can establish closed-loop negative feedback system, and define evaluation function help to

get geometric transformation. We can select 4 points of known size object in video frames and define relative displacement of responding point in IPM frames to get geometric transformation result. By changing different relative displacement approximate solutions. Using evaluation function obtains relative displacement in next time, until the distortion of IPM result is acceptable. Typically, we select points that lie on the edge of regular shape of object of real world. In urban road scene, lane line and traffic sign are eye-catching sign, and follow clear known traffic regulations. In our experiments, we select lane line and outline of bicycle sign (see Fig. 1). Once those points are selected, we never change it. Different selections of respond points in IPM result generate different top views, in this paper we call those top view result as IPM frames. To check the distortion extent, we should define an evaluation function, evaluation function is introduced in next section.

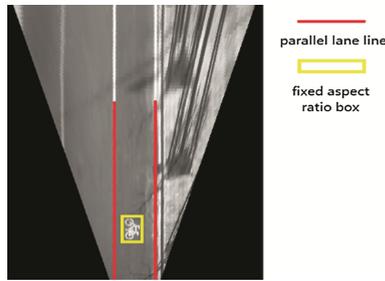


Fig. 1. IPM result of video frame. Red lane line is parallel. Yellow box's aspect ratio is 15:9 (Color figure online)

Our algorithm can be to show as follows

Algorithm 1

- 1: selected 4 target points $[u_i, v_i], i = 1,2,3,4$ at target plane
 - 2: selected 4 initial original points $[x_i, y_i], i = 1,2,3,4$ at original plane. Initialize the number of cycles *times* is 0
 - 3: solve the *trans*' variable
 - 4: get IPM frame of target plane by applying *trans*.
 - 5: **if** *times* is larger than its thread or evaluation function lower than the thread
 return *trans*,
 - else**
 Select neighborhood of target points as new target points. *times* increase 1.
 Go to step 3.
-

The computational complexity isn't important, since posture of vehicular camera is fixed, projective matrix parameters never change. This algorithm runs only once.

2.1 Evaluation Function

Another question is how to define evaluation function. Evaluation function is used to measure the distortion extent, the way of selection is customized. For example, in our experiment, we combine two factor to evaluation function, first factor is cross angle of two lane line, which should be 0. Line angle can be calculated by Hough transform [17] and trigonometric functions. Second factor is aspect ratio of bicycle's rectangular box, which should be R. So we define evaluation function as follows

$$\text{eva} = (\alpha * (\text{angle}_1 - \text{angle}_2))^2 + (\beta * (R_{\text{cal}} - R_{\text{real}}))^2 \quad (1)$$

3 Feature Map

3.1 Basic Algorithm

Before we discuss our algorithm, we should analysis the basic algorithm. According to basic algorithm image stitching, after we get a reliable transformation and apply it in the original frame to get IPM frame. For every IPM frame we detect SURF (speeded up robust features) feature points and extract SURF descriptor. For every two adjacent IPM frame, we match the respond SURF feature points. Then we find another geometric transformation between adjacent IPM frames. For this case, this geometric transformation should be similar or affine transformation. This geometric transformation is estimated by RANSAC or Squares Algorithms, since more than 3 pair matched feature points exist. At last, we apply those geometric transformation for a series of responding IPM frames one by one, map is reconstructed.

Two points should be explained. The reason for extracting SURF feature points instead of other feature points is compromise of calculating cost and accuracy, other feature is also feasible. But selected feature should avoid to generating too much response points that lie the edge of background. Another point is why those geometric transformations should be similar or affine transformation. Different from popular algorithm like Autostitch [16], All IPM frame should be at same plane, we should concentrate on displacement in same plane and rotation. Accord to our experiment, similar transformation is good enough, affine transformation lead in the distortion.

Basic algorithm is a viable option, but part of this algorithm can be optimized. Based on our experiment, we witness 3 weaknesses for basic algorithm. Firstly, error is mainly attributable to the progressive effect. Every inaccurate geometric transformation effect subsequent work. Secondly, for undulating area or shaking of camera result in inaccurate matching result. Refer to first reason, unreliable matrix estimated by inaccurate matching feature pair worsen performance of map. Thirdly, it's a great idea in theory that built map from transformation, which estimated at matched feature. However, when vehicle travelled straightly, calculation for every two IPM frame is wasteful of time, since two IPM frame have only a single direction of the displacement difference. Moreover, estimated transformation increased progressive error.

For first problem, progressive error is unavoidable, but progressive error suggest two clues: IPM frame should be as standard as possible, that is duty of IPM stage; another

clue is we should preserve original shape of IPM frame, this is another explanation that similar transformation is better choose than affine transformation. For second and third problem, if we detect the status of vehicular camera in advance, we can simply the basic algorithm by applying different degree of restraint geometric transformation to avoid unnecessary time cost and enhance robustness. When vehicle travelled straightly, estimated transformation has negligible angle, instead, using fixed orientation similar matrix is effectiveness and robustness enough. When the status of vehicular camera is unstable, the inaccurate matched feature points should be abandon.

3.2 Feature Map and Auxiliary Conception

Based on the above analysis, we propose a new algorithm to simply the process by detecting current status of vehicular camera. Before we show our algorithm, we should introduce some auxiliary conception.

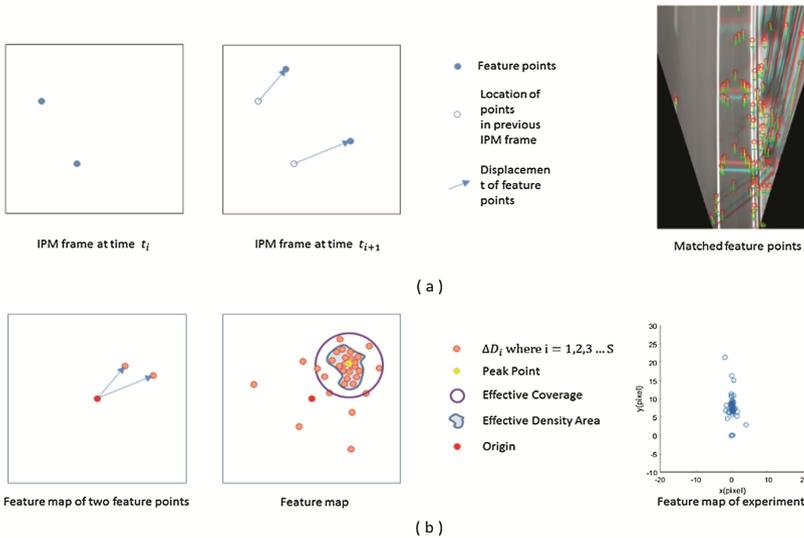


Fig. 2. Feature Map and auxiliary conception. (a) The displacement of two matched feature points in adjacent IPM frame. (b) Obtain feature map and the value of auxiliary conception

After we get matched SURF feature point pairs, from two IPM frames, we mark former IPM frame as f_1 , later IPM frame as f_2 , matched SURF feature points of f_1 as fp_1 , matched SURF feature points of f_2 as fp_2 , the size of feature points is S . Every feature point in fp_1 is marked as fp_{1i} , responding feature point in fp_2 is marked as fp_{2i} , (fp_{1i} and fp_{2i} are matched feature points pair, where $i = 1, 2, 3 \dots S$). For every matched feature point pair (fp_{1i} and fp_{2i}), we define ΔD_i is relative displacement of fp_{2i} and fp_{1i} .

$$\Delta D_i = fp_{2i} - fp_{1i}, \text{ where } i = 1, 2, 3 \dots S \tag{2}$$

We collect all ΔD_p , remark those ΔD_i at original point of new plane, we define this new plane as feature map (see Fig. 2). The feature map can provide richer information about status of vehicular camera.

There are 4 auxiliary conceptions to introduce. Peak point is location of the highest density of points. Effective coverage is circle, its center is peak points and its radius is customized d . Effective proportion is the ratio of the number of ΔD_i in effective coverage to the sum of matched feature point pair S . effective density area is irregular shape that included peak points and this area's density exceeds certain threshold (see Fig. 2). Those conceptions are used to help to simply the basic algorithm.

In this paragraph, we discuss the distribution of ΔD . If we make a basic assumptions: IPM frame is a scaled-down version of top view of real world (see Fig. 3). The points of different color are evenly distributed features of IPM frame, the gray box are IPM frame at times t_i , red box is IPM frame at time t_{i+1} when vehicle travel straightly, green box is IPM frame at time t_{i+1} when vehicle turn, blue box is IPM frame at times t_{i+1} when status of vehicular camera is unstable.

We define the feature point's location in real world coordinate is (x_{real}, y_{real}) . For the IPM frame at time t_{i+1} , the relative location of feature point is (x_b, y_b) , the origin is (x_{bo}, y_{bo}) . Relative location is the location of feature point in IPM frame coordinate system. For the IPM frame at times t_{i+1} , When vehicle travelled straightly, the relative location of feature point is (x_r, y_r) , the origin is (x_{ro}, y_{ro}) ; When vehicle turn left, the relative location of feature point is (x_g, y_g) , the origin is (x_{go}, y_{go}) ; When vehicle travel straightly, the relative location of feature point is (x_p, y_p) , the origin is (x_{po}, y_{po}) .

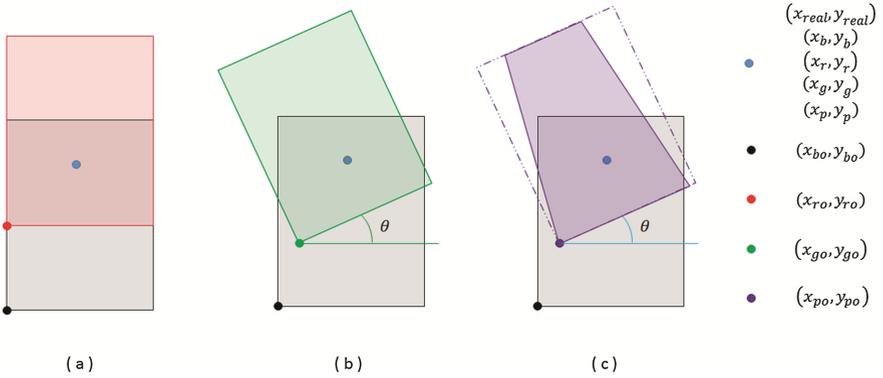


Fig. 3. Ideal model. (a) Vehicle travel straightly. (b) Vehicle turn left. (c) Camera status is unstable (Color figure online)

When vehicle travelled straightly, we have

$$x_{real} = x_b + x_{bo} = x_r + x_{ro} \tag{3-1}$$

$$y_{real} = y_b + y_{bo} = y_r + y_{ro} \tag{3-2}$$

So the distribution of ΔD should be a point.

$$\Delta D = (x_r - x_b, y_r - y_b) = (x_{bo} - x_{ro}, y_{bo} - y_{ro}) \tag{4}$$

When vehicle turned left or right,

$$x_{real} = x_b + x_{bo} = x_g \cos\theta - y_g \sin\theta + x_{go} \tag{5-1}$$

$$y_{real} = y_b + y_{bo} = x_g \sin\theta + y_g \cos\theta + y_{go} \tag{5-2}$$

So we have

$$x_g = \sin\theta(y_{real} - y_{go}) + \cos\theta(x_{real} - x_{go}) \tag{6-1}$$

$$y_g = \cos\theta(y_{real} - y_{go}) - \sin\theta(x_{real} - x_{go}) \tag{6-2}$$

So we can obtain ΔD as follows

$$x_g - x_b = \sin\theta(y_{real} - y_{go}) + \cos\theta(x_{real} - x_{go}) - (x_{real} - x_{bo}) \tag{7-1}$$

$$y_g - y_b = \cos\theta(y_{real} - y_{go}) - \sin\theta(x_{real} - x_{go}) - (y_{real} - y_{bo}) \tag{7-2}$$

Since $y_g = y_{real} - y_{go}$ and $x_g = x_{real} - x_{go}$

$$x_g - x_b = \sin\theta y_g + \cos\theta x_g - x_b \tag{8-1}$$

$$y_g - y_b = \cos\theta y_g - \sin\theta x_g - y_b \tag{8-2}$$

So the distribution of ΔD point is stretched. Stretching extent is related to yaw angle θ and feature points' location. When $\theta = 0$, the distribution of ΔD reduced to a point, equate to Eq. 4.

Analysis of unstable status is impossible, various reason effect the distribution of ΔD . The distribution is unpredictable, but the degree of aggregation is reduced inevitably.

If vehicle travelled straightly, ΔD_i , where $i = 1, 2, 3 \dots S$ are highly concentrated. The effective proportion is close to 1, Effective density area keep in small range. In this case the displacement of peak point is displacement of vehicular camera, peak point location is related to speed of vehicle. If vehicle turned, the effective proportion kept in high level, effective density area expanded. In this case the basic algorithm should be applied. If status of vehicle is unstable, the effective proportion is kept in low level, the matched feature point pairs should be abandon. So our algorithm is as follows.

Algorithm 2

- 1: extracted SURF feature of IPM frame
 - 2: match SURF feature of every two IPM frame.
 - 3: generate feature map
 - 4: obtain auxiliary conception
 - 5: **if** effective proportion is high
 - if** Effective density area keep in small range
 - Set geometric transformation as fixed parameters similar matrix, its displacement is displacement of peak point
 - else**
 - Estimated geometric transformation from matched feature point by RASANC.
 - else**
 - Abandon this matched **go to** step7
 - 6: warp geometric transformation into two IPM frames.
 - 7: Get next IPM frame, **go to** step 1.
-

3.3 Implement of Algorithm

The implement of feature map and auxiliary conception is customized. In our experiment, we use $3 * 3$ pitch to cover ΔD_i , where $i = 1, 2, 3 \dots S$ in feature map, feature map became mountain-like area. This area's peak is peak point, if this area's peak is connected plateau. The plateau's geometric center is the peak point. If this area has more than one peak, we use the highest connected cross-section's geometric center as peak point.

Effective proportion is a criterion to measure aggregation degree of feature points. If effective proportion of a certain threshold, the matched feature points pairs should be abandoned.

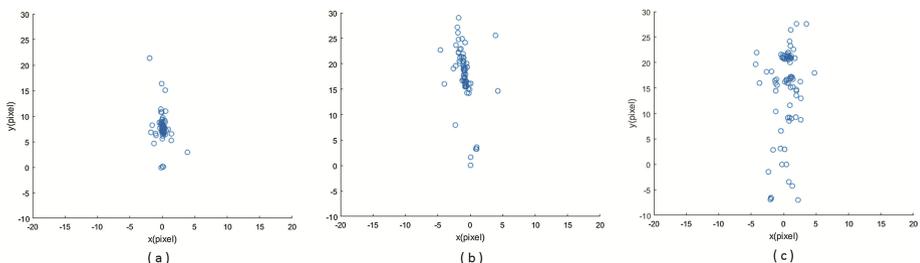


Fig. 4. Feature map. (a) When the vehicle travel straightly. (b) When the vehicle is on a bend. (c) When the status of vehicle is unstable

Effective density area's threshold is apposite proportion of value of peak point. We use outline of effective density area to measure angle changing extent roughly. If the effective

density area is close to circle or square, so car travel straightly, or the car is on a bend (see Fig. 4).

4 Experiment

In order to evaluate the performance of our algorithm, a database with 5 videos is tested resource, which consist of 2570 raw images. Experiment vehicle is a vehicle mounted CCD video camera. All images are RGB images with $640 * 480$ resolutions. All program is implemented by MATLAB on PC.

For the geometric transformation of IPM process. As the Eq. 1, We define $\beta = 25$, acceptable error is 0.5 degree angle and 2% loss in aspect ratio, so the thread is 0.5. Other reasonable choose is also acceptable. Loop upper limit is 1000. One of the original frame and its IPM is to show as follows (see Fig. 5). The IPM results are gray scale images with $1061 * 1523$ resolutions. The Geometric transformation is as follows.

For algorithm 2, in order to obtain the effective proportion, we select 30 pixels as effective coverage's radius, threshold is 80%. For effective density area, the proportion of value of peak point is 70%.

Figure 6 shows effective proportion of 5 video frames. 80% is thread, when effective proportion of feature map is less than 80%, matched feature points is abandon. From video frame, we witness that status of those frame's video is unstable. Experiment confirmed positive correlation between aggregation extent of ΔD and vehicular camera.

Figure 7 shows angle of 5 video frames. The upper part of fig are angle of estimated geometric matrix parameters. The lower part of fig is angle of our algorithm. When the angle of geometric matrix is tiny and fixed direction geometric matrix is used to avoid unnecessary progressive error and enhance robustness. When yaw angle is big enough, namely the vehicle is turn left or right, geometric matrix is estimated from basic algorithm.



Fig. 5. Original frame from video and IPM result

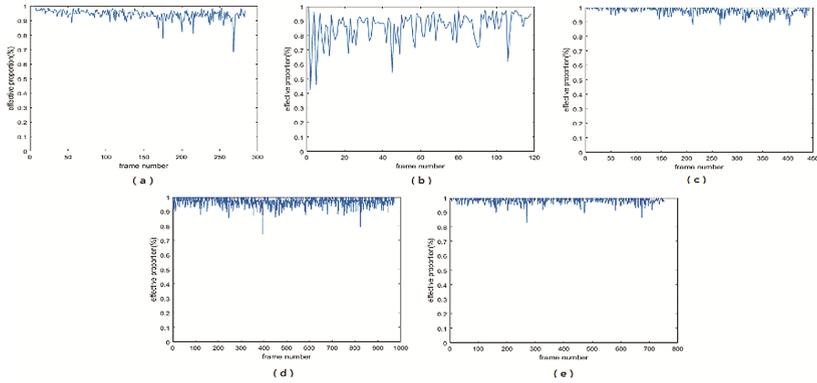


Fig. 6. Effective proportion

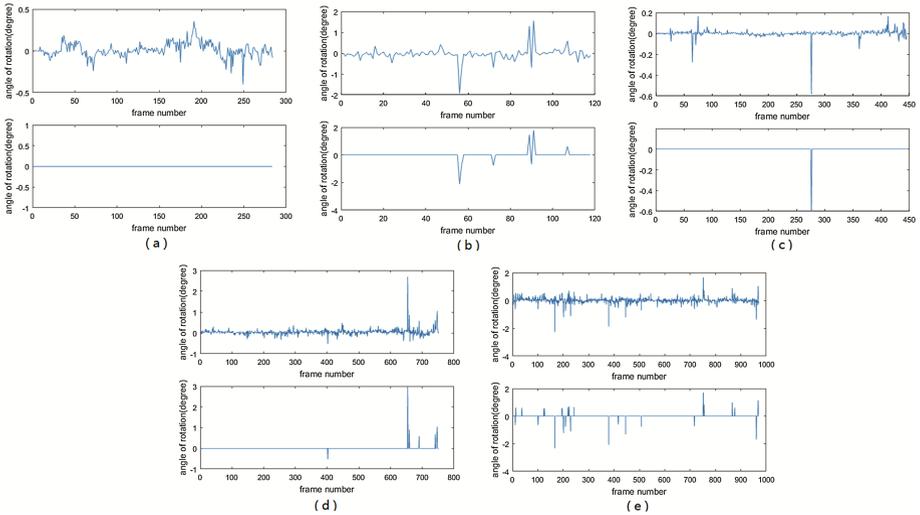


Fig. 7. Angle of rotation

Experiment result showed as follows (see Fig. 8 and Table 1). Figures 8(a) and (b) show the different of video 1, Figs. 8(c) and (d) optimization algorithm shows better performance than basic algorithm obviously. Basic algorithm is great, but many factors effect the performance, under massive input and unstable status of vehicular camera, progressive error lead in unavoidable the distortion. Our algorithm provided more reliable result with lower time cost, structure of road map is reconstructed.

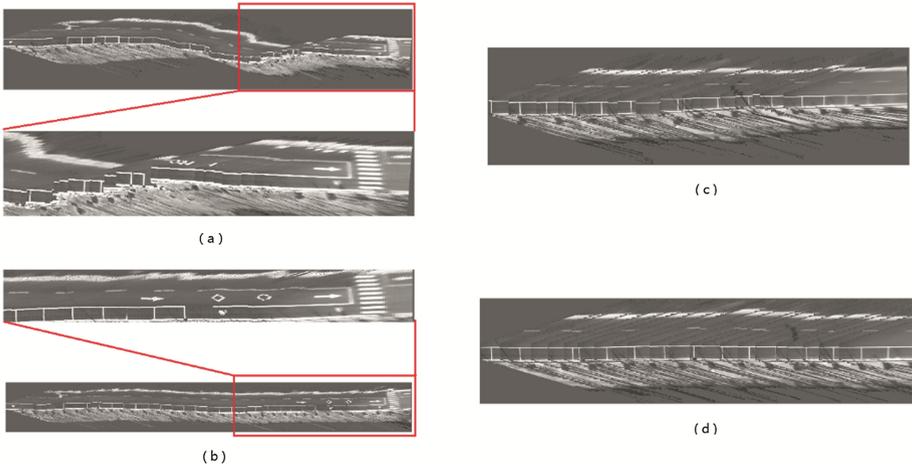


Fig. 8. Experiment result. (a) Basic algorithm and the distortion part of video 1. (b) Our algorithm and improved result of video 1. (c) Basic algorithm and the distortion part of video 2. (d) Our algorithm and improved result of video 2

Table 1. The average time cost of algorithm2

Frame number	Basic algorithm	Our algorithm
284	103.14 s	91.59 s
118	37.24 s	30.19 s
445	82.76 s	90.80 s
752	299.20 s	263.28 s
971	661.03 s	591.67 s

5 Conclusion

In this paper, we propose a fast and robust dynamical map building algorithm based on IPM and feature map. We establish feedback algorithm to obtain geometric transformation matrix of IPM. According to IPM frame’s characteristic, that IPM frame is a scaled-down version of top view’-eye of real world. Parallelism and aspect ratio is used to checking the accuracy of geometric transformation. In second stage of algorithm we define feature map and some auxiliary conception to help to classify status of camera or vehicle. When vehicle travelled straight, fixed orientation geometric transformation is used to avoid unnecessary and inaccuracy calculation. Experiment results show that the proposed algorithm can provide reliable and effective result. Future work will include search for relationship between of effective density area and estimated geometric transformation from matched feature points.

References

1. China: China Statistical Yearbook 2015 (2016)
2. Nedeveschi, S., Popescu, V., Danescu, R., et al.: Accurate ego-vehicle global localization at intersections through alignment of visual data with digital map. *IEEE Trans. Intell. Transp. Syst.* **14**(2), 673–687 (2013)
3. Gruyer, D., Belaroussi, R., Revilloud, M.: Map-aided localization with lateral perception. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings, pp. 674–680. IEEE (2014)
4. McLaughlin, S.B., Hankey, J.M.: Matching GPS records to digital map data: algorithm overview and application. National Surface Transportation Safety Center for Excellence (2015)
5. Lu, M., Wevers, K., Van Der Heijden, R.: Technical feasibility of advanced driver assistance systems (ADAS) for road traffic safety. *Transp. Planning Technol.* **28**(3), 167–187 (2005)
6. Mur-Artal, R., Montiel, J.M.M., Tardos, J.D.: ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE Trans. Robot.* **31**(5), 1147–1163 (2015)
7. Gálvez-López, D., Tardos, J.D.: Bags of binary words for fast place recognition in image sequences. *IEEE Trans. Robot.* **28**(5), 1188–1197 (2012)
8. Mur-Artal, R., Tardos, J.D.: ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. arXiv preprint [arXiv:1610.06475](https://arxiv.org/abs/1610.06475) (2016)
9. Engel, J., Schöps, T., Cremers, D.: LSD-SLAM: large-scale direct monocular SLAM. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) *ECCV 2014*. LNCS, vol. 8690, pp. 834–849. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10605-2_54
10. Klein, G., Murray, D.: Parallel tracking and mapping for small AR workspaces. In: 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2007, pp. 225–234. IEEE (2007)
11. Jiang, G.Y., Choi, T.Y., Hong, S.K., et al.: Lane and obstacle detection based on fast inverse perspective mapping algorithm. In: 2000 IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 2969–2974. IEEE (2000)
12. Nieto, M., Salgado, L., Jaureguizar, F., et al.: Stabilization of inverse perspective mapping images based on robust vanishing point estimation. In: 2007 IEEE Intelligent Vehicles Symposium, pp. 315–320. IEEE (2007)
13. Jiang, G.Y., Choi, T.Y., Hong, S.K., et al.: Lane and obstacle detection based on fast inverse perspective mapping algorithm. In: 2000 IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 2969–2974. IEEE (2000)
14. Gupta, A., Merchant, P.S.N.: Automated lane detection by K-means clustering: a machine learning approach. *Electron. Imaging* **2016**(14), 1–6 (2016)
15. Fritsch, J., Kühnl, T., Kummert, F.: Monocular road terrain detection by combining visual and spatial information. *IEEE Trans. Intell. Transp. Syst.* **15**(4), 1586–1596 (2014)
16. Ma, B., Zimmermann, T., Rohde, M., et al.: Use of Autostitch for automatic stitching of microscope images. *Micron* **38**(5), 492–499 (2007)
17. Mukhopadhyay, P., Chaudhuri, B.B.: A survey of Hough Transform. *Pattern Recogn.* **48**(3), 993–1010 (2015)
18. Torr, P.H.S., Zisserman, A.: MLESAC: a new robust estimator with application to estimating image geometry. *Comput. Vis. Image Underst.* **78**(1), 138–156 (2000)