



An improved vehicle panoramic image generation algorithm

Jindong Zhang^{1,2,3} · Xuelong Yin¹ · Jing Luan⁴ · Tong Liu¹

Received: 21 October 2018 / Revised: 14 April 2019 / Accepted: 10 June 2019 /
Published online: 25 June 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In order to reduce the traffic accidents caused by the blind area, vehicle panoramic view system has been paid more and more attention. However, the panoramic system is a complex and difficult system. In this paper, we propose an improved vehicle panoramic image generation algorithm. Several key technologies have been improved to ensure reliability and efficiency. First of all, we improve the spherical perspective projection algorithm (SPP) based on the scanning line idea and bilinear interpolation to rectification the fisheye image. Then the inverse perspective projection mapping of undistorted image is used to obtain a top view. In order to reduce computation, the method of manually selecting the target point is carried out. Finally, SURF algorithm is used to find the feature points between the bird's-eye view images around vehicle. We further put forward to utilize a RANSAC algorithm based on block matching to eliminate the mismatched points in the key point matching process. Experimental results indicate that our vehicle panoramic image generation method works efficiently. The proposed algorithm can effectively remove the serious distortion of fisheye lens, and generate a panoramic image around the vehicle in the end. It possesses good robustness, and can be widely used.

Keywords Vehicle panoramic image · Fisheye image correction · Inverse perspective mapping · Image mosaic

✉ Jindong Zhang
zhangjindong_100@163.com

¹ College of Computer Science and Technology, Jilin University, Changchun 130012, China

² Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

³ State Key Laboratory of Automobile Simulation and Control, Jilin University, Changchun 130025, China

⁴ College of Software, Jilin University, Changchun 130012, China

1 Introduction

In recent years, traffic accidents occur frequently, resulting in a large number of people died every year. The China Statistical Yearbook 2017 [20] shows that only in 2016 it occurred 212,846 traffic accidents. The number of dead and injured reached to hundreds of thousands and property losses amounted to more than one billion CNY, which caused a huge threat to people's life and property security [35]. According to the relevant statistics, part of traffic accidents are caused by the blind area of car, the main reason is that drivers cannot observe the surrounding environment completely driving which causes scraping, collision, rolling and so on [33]. To thoroughly eliminate the driver's visual blind area, panoramic vision system is becoming more and more important [5, 7, 31, 40].

Panoramic view system involves a lot of sophisticated technology. In recent years, with the rapid development of science and technology, more and more researches about key technologies of it have presented. For image correction, Zhang [38] put forward a simple and efficient calibration method in 2000. Both the internal and external parameters of a camera are calibrated by a two-dimensional plane checkerboard, then the distortion parameters of the lens are obtained. Finally, a distorted image can be corrected. It is a very classical approach. But for fisheye lens, because of its large distortion, Zhang's method may produce an unsatisfactory result. In 2003, Ying and Hu [36] proposed a fisheye lens correction method based on spherical perspective projection (SPP) constraint. The SPP model can be used to obtain the lens's radial and tangential deformation parameters to correct the distortion of the fisheye image. Tu et al. [25] presented a high precision two-step calibration method combining global and local ideas, which using a quartic polynomial to fit camera global projection model, then the image is split into blocks for local linear fitting optimization to reduce the error. In [8], Lai et al. adopted a distortion correction method based on Bayer format image. According to the characteristics of the Bayer image signal, they designed a befitting Hermite interpolation algorithm to realize the distortion correction of lens, which has realized on the FPGA device, and formed a real-time prototype system. Similarly, the real-time performance of lens correction was also studied in [26]. Turturici et al. optimized the process from three aspects, which contained algorithmic design, data transfer and the implementation of hardware and software. Based on the above idea, they put forward a lens correction scheme that can be applied to different types of lens, and can run on a low power embedded device in real-time. To avoid computing too many camera parameters, Yeh et al. [34] proposed an accurate and simple method to correct the fisheye image based on the relationship between the formation distance of the image and the incident angle. Bawa et al. [1] invented a general framework, which used a non-iterative method to correct a barrel distorted fisheye lens and is able to form a top-down undistorted image. The method is fast and computable.

In terms of image mosaic and fusion, research enthusiasm has not been weakened. The SIFT algorithm proposed by Lowe DG [14] is one of the most famous image stitching algorithms. It has a high precision but with a low running speed, so as another famous fast matching algorithm, SURF was proposed in [2]. On the basis of the two classical algorithms (SIFT and SURF), a large number of image mosaic algorithms are proposed. Such as BRISK and ORB algorithm [10, 23]. Cai et al. [3] improved the SURF algorithm by using the support vector machine to predict the determinant value of the Hansen matrix, which can remove some wrong feature matching points. The perspective scale invariant feature transform and artificial bee colony were introduced in [37], which made up for the deficiency of traditional matching algorithm in processing some large viewpoint changes images. Most of the image stitching

algorithms are running slowly. At the same time, the algorithms need to run on embedded devices in many scenes. So it is meaningful and indispensable to improve the image mosaic algorithm and utilize the parallel processing equipment [9, 29]. Zhou et al. [42] introduced a scheme for real-time mosaic of panoramic images. They transplanted the overall algorithm (image preprocessing, matching, stitching and fusion) to the FPGA device, whose high-efficiency calculating capability increased the performance of system greatly. In addition, Vourvoulakis et al. [28] increased the real-time performance of SIFT. They parallel processed the process of detecting and describing the feature points of SIFT on FPGA, and RANSAC algorithm was used to remove mismatching points, which guaranteeing the speed and accuracy. Dou et al. [4] reduced the SIFT extraction area by the fusion of wavelet filtering and Harris detection, which reduced the computation. In addition, He et al. [6] used Fast Fourier Transform Algorithm to convert images to frequency space, it was able to speed up the process of SIFT. The method run on GPU eventually. In [21], Dong et al. argued an image mosaic model based on BRISK algorithm. First, they got the rough matching points of image by BRISK algorithm, then the adjacent line segment matching was used, and finally acquired the exact matching of feature points by the probability statistical model. Victoria et al. [27] fused the SIFT and SURF based on fuzzy logic controller, and compared that with each original algorithm separately. It had the characteristics of low error and high noise resistance. For some complex stereoscopic images, a simple two-dimensional mosaic algorithm may not be applicable [15]. Yan et al. [32] expounded a mosaic method based on hybrid warping model. A unified homography transformation was performed on two images to be stitched, then the vertical gap between the two images was reduced through local refinement. Ultimately, an optimal stitching seam was found and stitched the images with the multi-frequency fusion algorithm. Unnatural stitching gaps may occur for images with different lighting. Zhang et al. [41] made the spliced image at the same illumination level by histogram matching. Then SIFT and RANSAC algorithms were used for matching key points roughly and removing false matching points, respectively. Ma et al. [16–19] have been conducting a series of in-depth researches on image registration methods in recent years, especially in feature point matching. They presented the locality preserving matching method, which is robust and fast. It is able to remove mismatches in feature matching, the base idea of which is the similarity of regional structure between the corresponding feature points [19]. Recently, with the advance in artificial intelligence, more and more researchers have applied it to feature extraction [11–13]. Rashid M et.al utilized deep convolutional neural network and SIFT point features achieve object detection and classification [22]. The neural networks methods have high degree of accuracy and also require higher time cost.

In our method, we focus on improving the key algorithms of generating vehicle panoramic image, which is able to effectively remove the serious distortion of fisheye lens and generate a panoramic image around the vehicle on embedded device compared to some existing technologies. The main contributions of this study are as follows:

- 1) The bilinear interpolation algorithm is used to optimize the SPP algorithm based on scanning line to corrected image. Meanwhile, we adopt a method by artificial marking to perform the inverse perspective mapping of the corrected image around the vehicle. The transformation matrix used is calculated by calibrating the selected points manually, which can reduce the time cost of system.
- 2) SURF, RANSAC and a weighted fusion algorithm are used for image mosaic and fusion, respectively. Which can balance the difference in the image seams caused by different

brightness. Furthermore, the methods are applied to the real environment by transplanting and running on IMX6Q (an embedded device).

The structures of the rest on the paper are organized as follows. In Section 2, the proposed method is expounded in detail. Section 3 summarizes the analysis of the experimental results. Conclusions and further work are placed in Section 4.

2 Proposed method

2.1 Image correction

Due to the short focal length and wide field of view of the fisheye camera [30], it is essential to correct the fisheye image captured before stitching [24, 39]. For the SPP algorithm [38], whose efficiency and performance are optimized in this paper.

2.1.1 Efficiency optimization

A suitable ROI can greatly improve the efficiency of the subsequent algorithm. Therefore, we extract the effective part of the image to be processed by selecting the ROI. A four directions scanning method (namely *ScanningMethod*) is proposed to determine the ROI in fisheye image. Due to the wide field characteristic of fisheye lens, for a fisheye image, only the near circle figure in which is the effective area. The edges and corners of the fisheye image are some redundant black areas. Hence, locating the effective area of the fisheye image before correcting is a feasible means to accelerate the efficiency. The specific descriptions are as follows (taking the left boundary as an example):

Step1: Initiate parameters

The brightness difference set according to experience: G . The left, right, top and bottom of the fisheye image are represented by *left*, *right*, *top* and *bottom*, respectively. $pixel_{max}$ and $pixel_{min}$ represent the maximum and minimum pixel values of a column, respectively. $pixel_{gray}$ is the gray value of the current pixel. The current pixel's RGB components are $pixel_R$, $pixel_G$ and $pixel_B$, respectively. $pixel_{center}$ and d are the center coordinates and radius of the fisheye image, respectively.

Step2: Scanning fisheye images from left to right column by column, $pixel_{gray} \leftarrow 0.3 \cdot pixel_R + 0.59 \cdot pixel_G + 0.11 \cdot pixel_B$

Step3: if $pixel_{gray} >= pixel_{max}$ then $pixel_{max} \leftarrow pixel_{gray}$

else if $pixel_{gray} <= pixel_{min}$ **then** $pixel_{min} \leftarrow pixel_{gray}$

Step4: **when** (Scanning a column of pixels to end) **do**

if $pixel_{max} - pixel_{min} >= G$ **then** *left* \leftarrow current column number

Step5: In the same way, the values of *right*, *top* and *bottom* can be calculated.

Step6: $pixel_{center} \leftarrow ((left + right)/2, (top + bottom)/2)$, $d \leftarrow \max(left - right, top - bottom)$.

2.1.2 Performance optimization

For the pixel mapping module in image correction. Nearest interpolation, bilinear interpolation and bicubic interpolation are common interpolation algorithms. Considering the processing speed and effect of the algorithm, bilinear interpolation algorithm is used to obtain the final corrected image. The detailed steps are as follows:

Step1: Initiate parameters

Set $image_{width}$, $image_{length}$ and Z are the width, length and height of projection plane, respectively. Loop control variables: i and j . $pixel_{image}$ is the pixel coordinate value of the fisheye image. $Pixel^*_{center}$ and $pixel^*_{image}$ represent the center coordinate and pixel values of undistorted image, respectively. $Pixel^*_{image}$ is an integer value of $pixel^*_{image}$. α and β are the weights of bilinear interpolation.

Step2: Get $pixel_{center}$ and d by *ScanningMethod*.

Step3: $pixel^*_{center.x} \leftarrow 1 \cdot Z \cdot pixel_{center.x} / \sqrt{(d \cdot d - pixel_{center.x} \cdot pixel_{center.x} - pixel_{center.y} \cdot pixel_{center.y})}$

$pixel^*_{center.y} \leftarrow 1 \cdot Z \cdot pixel_{center.y} / \sqrt{(d \cdot d - pixel_{center.x} \cdot pixel_{center.x} - pixel_{center.y} \cdot pixel_{center.y})}$

Step4: for $j < image_{width}$ && $i < image_{length}$ do

$pixel_{image.x} \leftarrow i - image_{length} / 2 + pixel^*_{center.x}$.

$pixel_{image.y} \leftarrow image_{width} / 2 - j - pixel^*_{center.y}$.

$pixel^*_{image.x} \leftarrow (d \cdot pixel_{image.x} / \sqrt{(pixel_{image.x} \cdot pixel_{image.x} + pixel_{image.y} \cdot pixel_{image.y} + Z \cdot Z)}) + d$.

$pixel^*_{image.y} \leftarrow d - (d \cdot pixel_{image.y} / \sqrt{(pixel_{image.x} \cdot pixel_{image.x} + pixel_{image.y} \cdot pixel_{image.y} + Z \cdot Z)})$

Step5: Round down to prepare for interpolation: $pixel^*_{image} \leftarrow (\text{int})pixel^*_{image}$.

Step6: $\alpha \leftarrow pixel^*_{image.x} - pixel^*_{image.x}$

$\beta \leftarrow pixel^*_{image.y} - pixel^*_{image.y}$

Step7: $pixel^*_{image} \leftarrow (1-\alpha) \cdot (1-\beta) \cdot (pixel^*_{image.x}, pixel^*_{image.y}) + \alpha \cdot (1-\beta) \cdot (pixel^*_{image.x} + 1, pixel^*_{image.y}) + (1-\alpha) \cdot \beta \cdot (pixel^*_{image.x}, pixel^*_{image.y} + 1) + \alpha \cdot \beta \cdot (pixel^*_{image.x} + 1, pixel^*_{image.y} + 1)$; end.

2.2 Inverse perspective mapping

After obtaining an undistorted image, there is a need to obtain a bird's-eye view (i.e., inverse perspective mapping). The process of inverse perspective mapping is mainly to transform a

perspective image into a bird’s-eye view image, which is also a process of homography transformation. So, it is necessary to know how an image is generated through a camera firstly. According to the image-forming principle of camera, the entire process from object to image consists of rigid transformation and perspective transformation. As shown in Eq. (1).

$$Z_C \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \cdot E \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \text{ with } A = \begin{bmatrix} f/d_x & 0 & u_0 & 0 \\ 0 & f/d_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, E = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

Where A and E are the intrinsic and extrinsic parameters of the camera, respectively. As mentioned above, A represents perspective transformation matrix and B represents rigid transformation matrix. u, v are the pixel coordinates of the image, X, Y and Z are the world coordinates, d_x, d_y represent the physical dimensions of image pixels on the horizontal and vertical axes, respectively. Z_C is the coordinate value in camera coordinate system, and u_0, v_0 represent the principal point of the image in pixel coordinate system. f is focal length, R and T are parameters matrices for rotation and translation, respectively.

In this way, the formula of inverse perspective mapping is able to be reasoned. The matrix named H is used instead of the complex transformation matrix in Eq. (1) to simplify it, which is called homography matrix. As shown in Eq. (2).

$$Z_C \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2}$$

In inverse perspective mapping, since both the original image and the result image are two-dimensional images, the value of Z coordinate in Eq. (2) can be set to zero. Since h_{33} in H is a known quantity as a homogeneous coordinate, H has eight unknowns, which requires eight sets of equations to be solved. A set of pixels can generate two equations, so at least four sets of pixels are needed. The specific calculation process of H will be described in detail later. After obtaining the H , it is necessary to determine the specific range of the inverse perspective mapping of the image. Because of the perspective effect, parallel lines in the world coordinate system will converge at one point in the image, namely vanishing point. According to experience, the inverse perspective mapping generally does not exceed the position of vanishing point. Therefore, the coordinates of vanishing point need to be calculated. Suppose there is a line in the world coordinate system (represented by L), whose direction is indicated by $N(N_x, N_y, N_z)$. If Q is a point on L , whose coordinate is (X_Q, Y_Q, Z_Q) , then the equation of L can be expressed as follows:

$$F(\varphi) = \varphi N + Q \tag{3}$$

Where φ is the coefficient, and as φ becomes larger, $F(\varphi)$ is closer to the point at infinity. Therefore, combining and calculating Eq. (2) and Eq. (3) to get Eq. (4).

$$Z_C \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} \varphi N_x + X_Q \\ \varphi N_y + Y_Q \\ \varphi N_z + Z_Q \end{bmatrix} \tag{4}$$

When φ is infinite, the vanishing point coordinate (represented by $(u_{+\infty}, v_{+\infty})$) will be obtained. It can be calculated according to Eq. (5).

$$\begin{cases} u_{+\infty} = \frac{h_{11}N_x + h_{12}N_y + h_{13}N_z}{h_{31}N_x + h_{32}N_y + h_{33}N_z} \\ v_{+\infty} = \frac{h_{21}N_x + h_{22}N_y + h_{23}N_z}{h_{31}N_x + h_{32}N_y + h_{33}N_z} \end{cases} \quad (5)$$

Assuming that H is known, $(u_{+\infty}, v_{+\infty})$ can be calculated. According to the vanishing point, the scope of the inverse perspective mapping will be determined. How to calculate H will be described below.

Homography matrix is the core of the inverse perspective mapping algorithm. As a common method, the camera calibration is able to obtain the parameters of H . However, this method depends on high accuracy when measuring external parameters. An improved inverse perspective mapping algorithm by artificial marking is carried out in this paper. First, a coordinate system model is established according to the plane where vehicle locates. As for each fisheye camera around the vehicle, four reference points are set up. Then, the coordinate values of the 16 points in the world coordinate system are manually measured multiple times, and the average value is taken as the final result. Similarly, the coordinate values in the image coordinate system corresponding to the reference points are also obtained. Finally, in each direction, the homography matrix of each fisheye camera can be obtained by the mapping relationship among the four pairs of points. Through four homography matrices, inverse perspective mappings of the four images around the vehicle will be completed.

To reduce the computation time, we build a same world coordinate system instead of creating it for each fisheye camera. The origin of the world coordinate is defined at the intersection between the line where the front and rear cameras are located and the line where the left and right cameras are located. The directions of the right camera and the front camera represent the x-axis and y-axis directions, respectively. For ease of observation and measurement during the actual operation, some special reference objects with distinct apex characteristics can be selected as reference points, such as black-and-white checkerboard. The specific implementations are as follows:

Step1: Initiate parameters

p_s and p_d are used to save pixel coordinates and world coordinates, respectively. H is the homography matrix. i_b indicate bird's-eye view image.

Step2: $p_s \leftarrow$ image pixel coordinates obtained by the fusion of measurement and calculation;

$p_d \leftarrow$ coordinates in the world coordinate system by artificial measurement.

Step3: Calculate H according to the mapping relationship between p_s and p_d .

Step4: Transform the corrected image to i_b by H .

2.3 Image mosaic

For the direction of each camera, the above process is called, resulting in a bird's-eye view around the vehicle. The SURF algorithm is applied to detect the feature points of

bird's-eye view images after this transformation. In order to achieve image mosaic, it is necessary to find the feature points which are matched between two images. Generally, Euclidean distance is used as a standard to measure similarity in the coarse matching process. In a pair of images that to be matched, one point to be processed is selected in one of them. And calculate the minimum and second minimum of Euclidean distance between this point and the point in another image. When the ratio of them is less than a threshold defined in advance (represented by r), two points are considered as a pair of rough matching points. For the mismatching that may be occurred during the rough matching process, The RANSAC algorithm based on block is used to eliminate these mismatches. In order to help speed up processing, the images to be processed are divided into blocks averagely. Only the image blocks containing the matching points are processed. A block with matching points is added to the dataset (represented by S) to be detected. RANSAC is preferentially applied to image blocks with more matching points in S . And select four pairs of points randomly to calculate a transformation matrix (represented by η). The point within a certain range (represented by ε) of the transformation matrix is called inner point. When the number of internal points reaches the maximum, η is optimal. Following this, the algorithm is used to match the four bird's-eye images, separately. The basic steps are as follows:

Step1: Initiate parameters

$S \leftarrow$ image blocks to be processed. A is an inner point set and A_g is the best case for A . $\eta \leftarrow$ a transformation matrix model. n is the number of iterations. m is an iterative control variable. $\varepsilon \leftarrow$ a certain threshold. δ indicates error.

Step2: select four pairs of matching points randomly from S , and calculate one model of η .

Step3: $\delta \leftarrow$ projection error between η and all data in S . **if** $\delta < \varepsilon$ **then** $A \leftarrow$ all data in S which satisfy $\delta < \varepsilon$.

Step4: **if** number of $A >$ number of A_g **then** $A_g \leftarrow A$.

Step5: **if** $m^{++} < n$ **then** go to Step2

else return η , **end.**

Because of the specificity of vehicle panoramic view system, the images around the car are captured from four fisheye cameras with different directions and areas. Therefore, different illumination may occur in different areas, resulting in uneven brightness of the images. In this case, there will appear obvious brightness differences in the fusion area. The weighted average method is used to deal with the case. The main idea is to determine the pixel values of the join points by the pixel values of the two images to be stitched together. It can greatly increase the correlation between stitching images. Usually, a coefficient (represented by K) is necessary, $K \in [0, 1]$. So how to calculate the pixel value (represented by $Pixel$) in the fusion area is shown in Eq. (6).

$$Pixel = K \cdot Pixel_1 + (1-K) \cdot Pixel_2 \quad (6)$$

Where $Pixel_1$ and $Pixel_2$ represent the pixel values of position at the seam on the two images to be stitched, respectively. $K = S_1 / (S_1 + S_2)$, S_1 and S_2 represent distances between the point in the fusion area and images to be stitched, respectively.

3 Experiment and analysis

To verify the proposed method, all experiments are programmed on computer with 3.4 GHz Intel Core i7 CPU and 16GB RAM using Ubuntu 12.04.

Fig. 1 shows the correction result of fisheye image. Figure 1 (a) is the original image captured by fisheye camera. Figure 1 (b) and Fig. 1 (c) are the undistorted images by Zhang’s algorithm and our method, respectively. It can be seen that the correction effect in Fig. 1 (b) is not excellent, and there is still obvious distortion at the corners of the corrected image. Moreover, compared with the original image, Fig. 1 (b) greatly reduces the field of vision of the fisheye image. It will seriously affect the adjacent overlapping areas of the overlooking image after the inverse perspective transformation, and thus may increase the inaccuracy of image mosaic. Our method based on SPP is used to correct fisheye image, and optimize the corrected image through the bilinear interpolation. Figure 1 (c) has a wider view than Fig. 1 (b). Compared with Zhang’s algorithm, our method has better effect and the result image is smoother.

In order to calibrate the fisheye camera, we select an alumina chessboard as calibration board (size: 200 mm * 200 mm, square array: 12 * 9, side length of a square: 15 mm, precision error: ± 0.01 mm). Figure 2 is a part of the images taken by fisheye camera, which are to be calibrated.

The result of pixel error analysis for the calibration results is shown in Fig. 3, where Fig. 3 (a) represents the overall mean error. We can set the threshold to about 0.16 pixels, the images with a mean error of more than threshold are deleted as shown in Fig. 3 (b).

Fig. 4 shows the position relationship between the fisheye camera and calibration board. Figure 4 (a) and (b) are the extrinsic parameters visualizations centered on the camera and calibration board, respectively.

Camera parameters are obtained through calibration, as shown in Table 1.

The distortion correction of fisheye image is realized by using radial distortion and intrinsic matrix composed of focal length and principal point. In addition, the distortion of fisheye lens which are used in our system is too large, so the third-order radial distortion coefficient is employed. Compared with the general second-order coefficient, which can reduce the calibration errors and has better correction effects.

Then, we use the improved inverse perspective mapping algorithm to transform the corrected image. The resolution of corrected image is 720 * 576, so the points in the world coordinate system need to be processed according to Eq. (7).

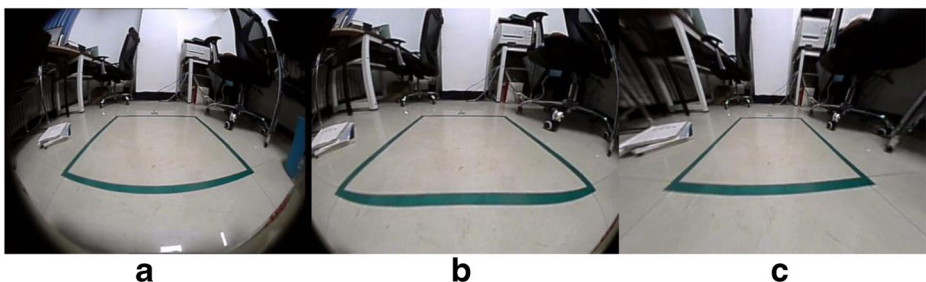


Fig. 1 Comparison between original image and undistorted image. **a** Original image. **b** The result of Zhang’s algorithm. **c** The result of our method

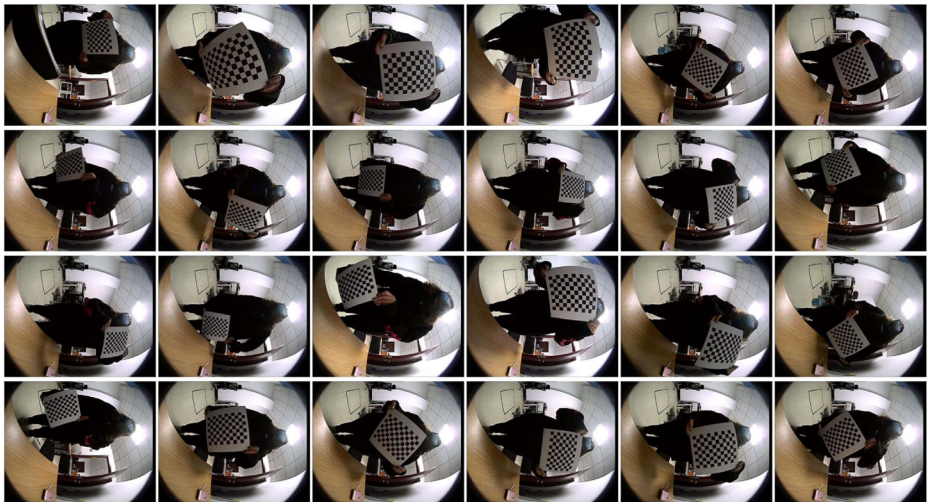


Fig. 2 Original images

$$\begin{bmatrix} X_{new} \\ Y_{new} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 720/2 \\ 0 & -1 & 576 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X_{old} \\ Y_{old} \\ 1 \end{bmatrix} \tag{7}$$

Where (X_{new}, Y_{new}) is the new coordinate, (X_{old}, Y_{old}) is the old world coordinate. After processing, the scale is reduced to half and final result is shown in Fig. 5 (c).

As shown above, the improved inverse perspective transform algorithm has a good overlooking effect. Compared with the manual measurement method based on internal and external parameters by camera calibration, the proposed method only calculates the homography matrix directly. It does not need to fix the camera on the car and use the instrument to measure the camera height, pitch angle, yaw angle and other

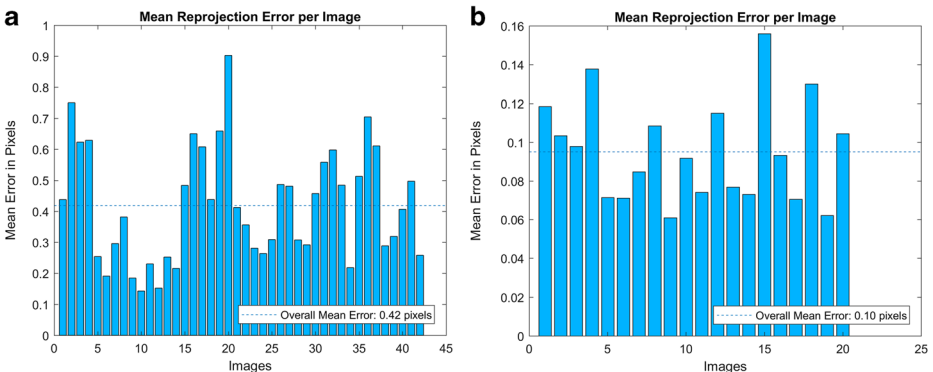


Fig. 3 The result of pixel error analysis. **a** The original overall mean error. **b** Delete the image with too much error

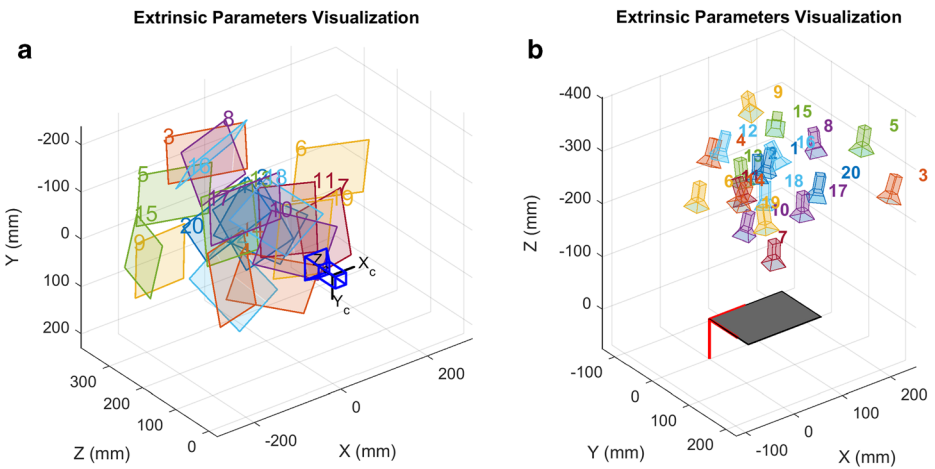


Fig. 4 The position relationship between the camera and calibration board. **a** Camera is the center. **b** Calibration board is the center

parameters artificially. Therefore, the measurement process of the improved method is simple and the error sensitivity is greatly reduced.

After overlooking transformation, it enters the image splicing stage of the whole system. In this paper, the method of controlling variable is used to compare the mainstream image stitching algorithm. And analyze them from three aspects: image blurring, illumination condition and algorithmic efficiency.

3.1 Image blurring

Manual processing of an image to change its blur strength, and recognize the feature points in the image through several popular algorithms. As shown in Fig. 6. Figure 6 (a) is a multi-level blurring process for the same image. From Fig. 6 (b) ~ (d), it is clear that SURF and SIFT algorithms can still accurately identify all kinds of feature points for a relatively blurred image, while BRISK algorithm can recognize most of the feature points, but it also loses some checkerboard corner points.

3.2 Illumination condition

The image is processed with different brightness, and the rest is the same as above. Figure 7 shows the result. Figure 7 (a) shows the brightness processing of the same image at different levels. From Fig. 7 (b) ~ (d), it can be clearly seen that SIFT algorithm

Table 1 Camera parameters

Intrinsic Matrix	Focal Length	Principal Point	Radial Distortion
$\begin{bmatrix} 248.44 & 0 & 0 \\ 0 & 255.84 & 0 \\ 352.42 & 278.69 & 1 \end{bmatrix}$	[248.44 255.84]	[352.42 278.69]	-0.33 0.14 -0.03

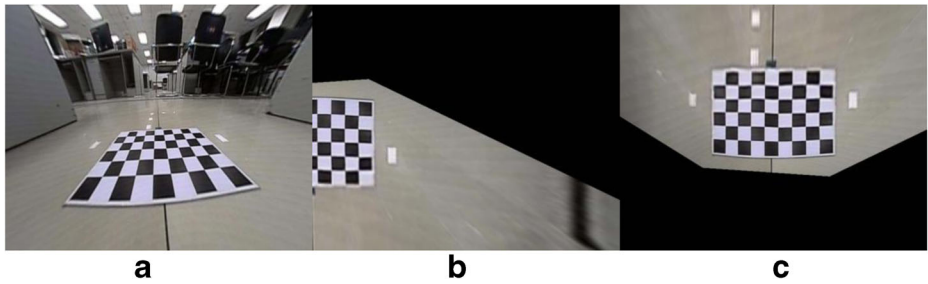


Fig. 5 Comparison between undistorted image and bird's-eye view image. **a** Original image. **b** Initial result of bird's-eye view image. **c** Improved result of bird's-eye view image

recognizes the most feature points for a dark image, followed by SURF, and it can still accurately identify all kinds of feature points. As for BRISK algorithm, some checker-board corner points are not recognized.

3.3 Algorithmic efficiency

In the vehicle panoramic view system, algorithmic efficiency is quite significant. It is not easy to achieve real-time effect when run all algorithms in an embedded system environment. Therefore, the algorithm with high speed should be preferred under the premise that stitching effect is acceptable. Additionally, for the three aspects above, the contrast results of different methods are shown in Fig. 8.

Where Fig. 8 (a) and (b) represent the number of feature points detected by different algorithms for the same image at different degrees of blur and brightness. Figure 8 (c) is the running time of the three algorithms. It can be seen from the results that the number of feature points detected will decrease with the decrease of image blurring. But at the same blurring, the number of feature points identified by SIFT and SURF is much higher than BRISK. When the brightness of image is

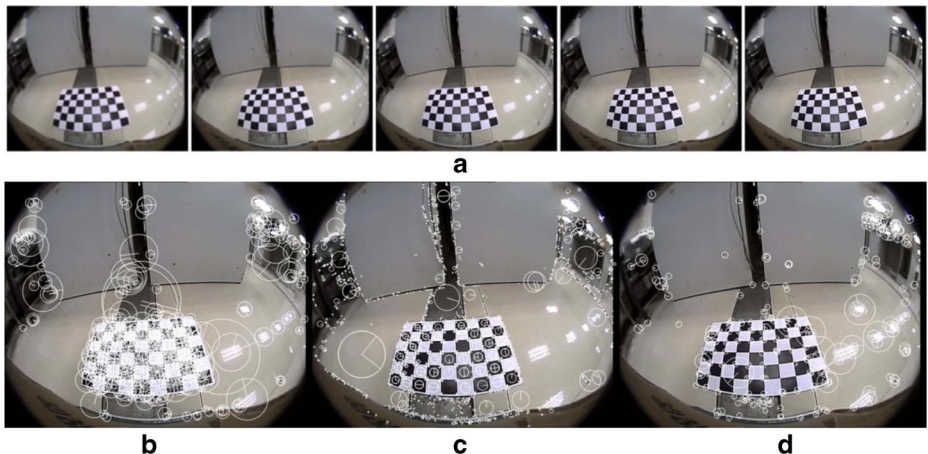


Fig. 6 The detection of feature points by different algorithms. **a** Five different blur strengths of an image. **b** By SURF. **c** By SIFT. **d** By BRISK

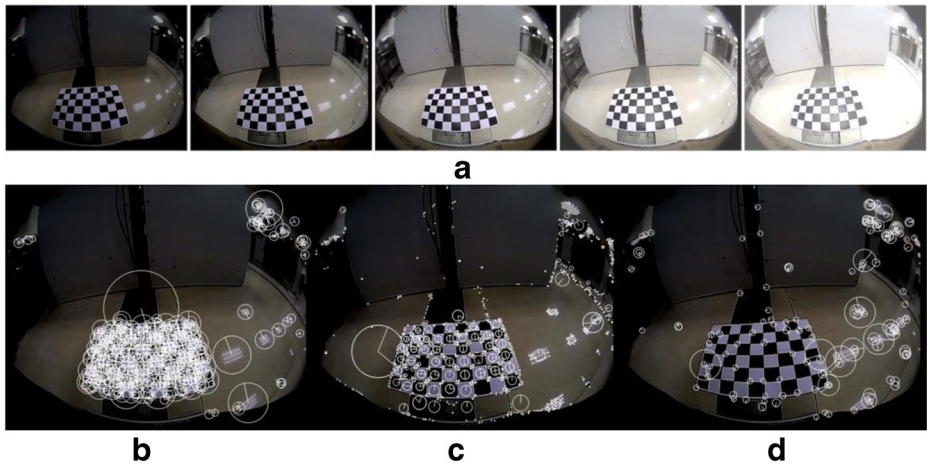


Fig. 7 The detection of feature points by different algorithms. **a** Five different luminance of an image. **b** By SURF. **c** By SIFT. **d** By BRISK

changing, SIFT, SURF and BRISK still detect a certain number of feature points and have a good stability. In terms of algorithmic efficiency, SIFT has the biggest number of feature points, but the execution speed is the slowest. SURF has a relatively large number of feature points and runs faster compared with SIFT.

Considering the generality of experiment, we used these methods to conducted experiments on the set of images with different features respectively. Analysis results are listed in the form of tables and graphs. The number of feature points detected by SIFT, SURF and BRISK are shown in Table 2.

For a more intuitive comparison of several methods, Fig. 9 shows the detection results of every methods, where horizontal axis is the number of images and ordinate represents the number of feature points. Black, red and blue lines in Fig. 9 represent the detection results by SIFT, SURF and BRISK, respectively.

From the above analysis, SURF is selected as the image mosaic algorithm on the basis of the comprehensive consideration of the detection effect and system time. In order to better reflect the effectiveness and practicability of the proposed methods, we have implemented the algorithms on PC and embedded platform, respectively. As shown below:

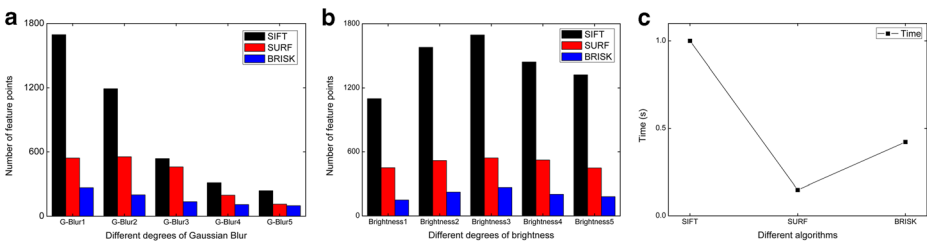
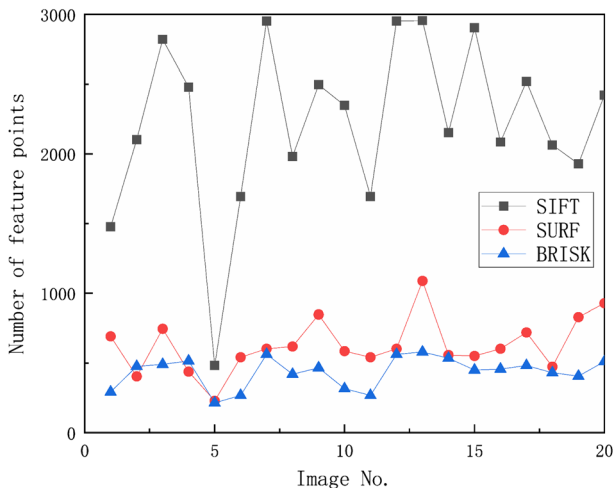


Fig. 8 Comparison of the different algorithms. **a** Image blurring. **b** Illumination condition. **c** Algorithmic efficiency

Table 2 The number of feature points detected by different algorithms

Image No.	Number of feature points		
	SIFT	SURF	BRISK
0	1477	692	292
1	2102	404	476
2	2821	746	491
3	2480	438	515
4	482	228	214
5	1695	542	267
6	3053	603	563
7	1981	619	419
8	2497	848	466
9	2348	586	316
10	1695	542	267
11	3053	603	563
12	2956	1088	580
13	2153	558	535
14	2905	551	450
15	2085	603	457
16	2520	720	482
17	2064	474	430
18	1928	830	406
19	2423	928	511

The results of the operation on PC are shown in Fig. 10. Four fish-eye images from the front, left, right and back of the vehicle form a 360-degree panoramic image through distortion correction, inverse perspective mapping, stitching and so on. Figure 10 (e) can clearly show the road information around the vehicle. Figure 11 shows the final result of vehicle panoramic view system which runs on IMX6Q. Display

**Fig. 9** Detection results of the different algorithms for different images

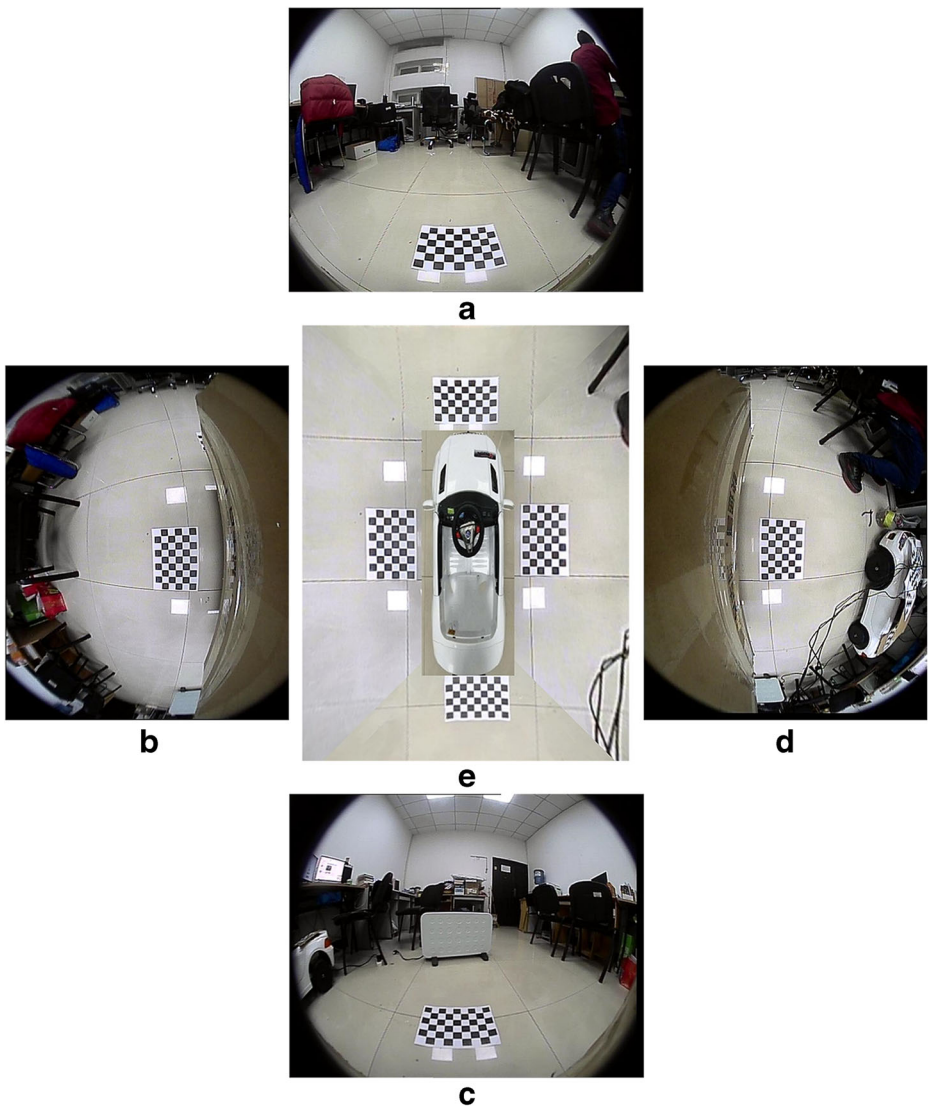


Fig. 10 System running on PC. **a** Front camera. **b** Left camera. **c** Rear camera. **d** Right camera. **e** panoramic image

screen is divided into two parts, the left is the original image captured by fisheye camera and the panoramic image is on the right.

It is easy to see from Figs. 10 and 11 that the proposed methods can be well applied to the vehicle panoramic view system. Furthermore, since vehicle panoramic view system contains abundant environmental information around the car, it is possible to integrate various functions based on which. Such as lane departure warning system, parking assistant system and so on. If the panoramic image includes lane lines and parking spaces, by the extraction and recognition of which, a highly integrated and low-cost driving assistant system will be formed. It can greatly increase the safety of driving.

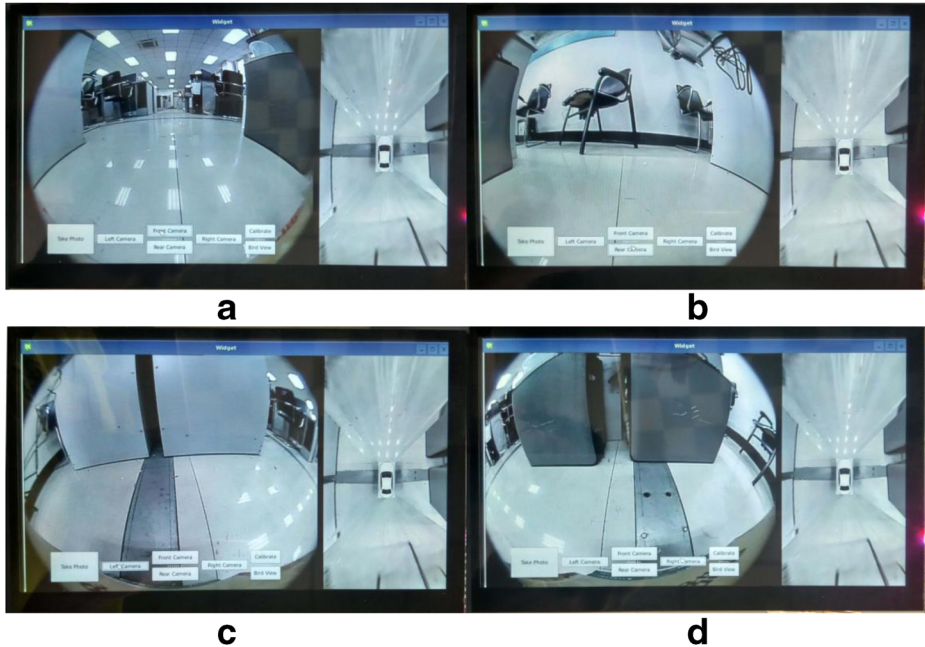


Fig. 11 System running on IMX6Q. **a** Front camera & panoramic image. **b** Rear camera & panoramic image. **c** Left camera & panoramic image. **d** Right camera & panoramic image

4 Conclusion

An improved vehicle panoramic image generation algorithm is present in this paper. There are some improvements for image correction, inverse perspective mapping and image stitching. Statistics and analysis show that it has high accuracy and feasibility. Moreover, the proposed method has good reliability and robustness under different blurring and illumination. The method is transplanted and running on IMX6Q, which can accurately and effectively display the panoramic view around the vehicle for driver to improve the driving safety. In the future, we will enhance the real-time performance of the method and apply it in the real environment.

Acknowledgements This work was supported by the National Key Research and Development Program of China (Grant No. 2017YFB0102500); the Natural Science Foundation of Jilin province (Grant No. 20170101133JC); the Korea Foundation for Advanced Studies' International Scholar Exchange Fellowship for the academic year of 2017-2018, and Jilin University (Grant No. 5157050847, 2017XYB252).

References

1. Bawa VS, Kumar K, Kumar V (2017) A generalized non-linear method for distortion correction and top-down view conversion of fish eye images. *Image Anal Stereol* 36(2):141–150
2. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: *European conference on computer vision*. Springer, pp 404–417
3. Cai C, Wang P, Liang Y-h (2016) Fast image stitching based on improved SURF. In: *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, IEEE, pp 411–416

4. Dou J, Qin Q, Tu Z (2018) Robust image matching based on the information of SIFT. *Optik* 171:850–861
5. Gao Y, Lin C, Zhao Y, Wang X, Wei S, Huang Q (2018) 3-D surround view for advanced driver assistance systems. *IEEE Trans Intell Transp Syst* 19(1):320–328
6. He Y, Deng G, Wang Y, Wei L, Yang J, Li X, Zhang Y (2018) Optimization of SIFT algorithm for fast-image feature extraction in line-scanning ophthalmoscope. *Optik* 152:21–28
7. Ji X, Zhang G, Chen X, Guo Q (2018) Multi-perspective tracking for intelligent vehicle. *IEEE Trans Intell Transp Syst* 19(2):518–529
8. Lai S, Xiong Z, Chen L, Tan X, Zhang M (2014) Real-time distortion correction of fish-eye lens based on Bayer image signal. *Opt Rev* 21(2):162–173
9. Lee J-H (2017) Panoramic image stitching using feature extracting and matching on embedded system. *Trans Electr Electron Mater* 18(5):273–278
10. Leutenegger S, Chli M, Siegwart R (2011) BRISK: binary robust invariant scalable keypoints. In: 2011 IEEE international conference on computer vision (ICCV) Ieee, pp 2548–2555
11. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: recognizing complex activities from sensor data. In: Twenty-fourth international joint conference on artificial intelligence
12. Liu Y, Zhang L, Nie L, Yan Y (2016) Rosenblum DS fortune teller: predicting your career path. In: Thirtieth AAAI conference on artificial intelligence
13. Liu Y, Nie L, Liu L, Rosenblum DS (2016) From action to activity: sensor-based activity recognition. *Neurocomputing* 181:108–115
14. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
15. Lu Y, Hua Z, Gao K, Xu T (2018) Multiperspective image stitching and regularization via hybrid structure warping. *Comput Sci Eng* 20(2):10–23
16. Ma J, Zhao J, Tian J, Yuille AL, Tu Z (2014) Robust point matching via vector field consensus. *IEEE Trans Image Process* 23(4):1706–1721
17. Ma J, Qiu W, Zhao J, Ma Y, Yuille AL, Tu Z (2015) Robust L2E estimation of transformation for non-rigid registration. *IEEE Trans Signal Process* 63(5):1115–1129
18. Ma J, Zhou H, Zhao J, Gao Y, Jiang J, Tian J (2015) Robust feature matching for remote sensing image registration via locally linear transforming. *IEEE Trans Geosci Remote Sens* 53(12):6469–6481
19. Ma J, Zhao J, Jiang J, Zhou H, Guo X (2019) Locality preserving matching. *Int J Comput Vis* 127(5):512–531
20. NBoSo C (2017) China statistical yearbook. China Statistics Press, Beijing
21. Qiang D, Jinghong L, Chao W, Qianfei Z (2017) Image mosaic algorithm based on improved BRISK. *J Electron Inf Technol* 39(2):444–450
22. Rashid M, Khan MA, Sharif M, Raza M, Sarfraz MM, Afza F (2018) Object detection and classification: a joint selection and fusion strategy of deep convolutional neural network and SIFT point features. *Multimed Tools Appl*:1–27
23. Rublee E, Rabaud V, Konolige K, Bradski GR (2011) ORB: an efficient alternative to SIFT or SURF. In: 2011 international conference on computer vision, pp. 2564–2571
24. Tiwari U, Mani U, Paul S, Nath N (2015) Non-linear method used for distortion correction of fish-eye lens: comparative analysis of different mapping functions. In: 2015 international conference on man and machine interfacing (MAMI), IEEE, pp 1–5
25. Tu B, Liu L, Liu Y, Jin Y, Tang J (2013) High precision two-step calibration method for the fish-eye camera. *Appl Opt* 52(7):C37–C42
26. Turturici M, Saponara S, Fanucci L, Franchi E (2014) Low-power DSP system for real-time correction of fish-eye cameras in automotive driver assistance applications. *J Real-Time Image Proc* 9(3):463–478
27. Victoria BA, Jorge RSA (2017) Manuel PHL SIFT-SURF commutation using fuzzy logic to image mosaicking. In: 2017 14th international conference on electrical engineering, computing science and automatic control (CCE). IEEE, pp 1–6
28. Vourvoulakis J, Kalomiros J, Lygouras J (2017) FPGA accelerator for real-time SIFT matching with RANSAC support. *Microprocess Microsyst* 49:105–116
29. Vourvoulakis J, Kalomiros J, Lygouras J (2018) FPGA-based architecture of a real-time SIFT matcher and RANSAC algorithm for robotic vision applications. *Multimed Tools Appl* 77(8):9393–9415. <https://doi.org/10.1007/s11042-017-5042-x>
30. Wu J, Huang F, Hu W, He W, Tu B, Guo L, Ou X, Zhang G (2019) Study of multiple moving targets' detection in fisheye video based on the moving blob model. *Multimed Tools Appl* 78(1):877–896. <https://doi.org/10.1007/s11042-018-5763-5>
31. Xiong Z, Ying J, Zhang R (2011) Research of bird's-eye panoramic view for vehicle parking. 2011 International Conference on Multimedia Technology, IEEE:456–459
32. Yan W, Hou C, Lei J, Fang Y, Gu Z, Ling N (2017) Stereoscopic image stitching based on a hybrid warping model. *IEEE Transactions on Circuits and Systems for Video Technology* 27(9):1934–1946

33. Yang Z, Zhao Y, Hu X, Yin Y, Zhou L, Tao D (2018) A flexible vehicle surround view camera system by central-around coordinate mapping model. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-018-6744-4>
34. Yeh Y-T, Peng C-K, Chen K-W, Chen Y-S, Hung Y-P (2014) Driver assistance system providing an intuitive perspective view of vehicle surrounding. In: *Asian conference on computer vision*. Springer, pp 403–417
35. Yin X, Zhang J, Wu X, Huang J, Xu Y, Zhu L (2018) An improved lane departure warning algorithm based on fusion of F-Kalman filter and F-TLC. *Multimed Tools Appl*:1–20. <https://doi.org/10.1007/s11042-018-6762-2>
36. Ying X-H, Hu Z-Y (2003) Fisheye lense distortion correction using spherical perspective projection constraint. *Chinese journal of computers-chinese edition* 26(12):1702–1708
37. Zeng Q, Wang Y, Liu J, Liu S (2017) A matching algorithm for large viewpoint changes images. *Optik* 137: 268–278
38. Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 22
39. Zhang Y, Zhou S (2017) Study on methods for fish-eye image correction based on spherical projection model. In: *2017 5th international conference on Frontiers of manufacturing science and measuring technology (FMSMT 2017)*. Atlantis Press, Paris, pp 848–854
40. Zhang B, Appia V, Pekkucuksen I, Liu Y, Umit Batur A, Shastry P, Liu S, Sivasankaran S, Chitnis K (2014) A surround view camera solution for embedded systems. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp 662–667
41. Zhang J, Chen G, Jia Z (2017) An image stitching algorithm based on histogram matching and SIFT algorithm. *Int J Pattern Recognit Artif Intell* 31(04):1754006
42. Zhou W, Liu Y, Lyu C, Zhou W, Peng J, Yang R, Shang H (2016) Real-time implementation of panoramic mosaic camera based on FPGA. In: *2016 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, 2016 IEEE, pp 204–209

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Jindong Zhang received the MSc and PhD degree in computer science and technology from Jilin University, in 2009. He is on faculty of Jilin University as an associate professor of Computer Science and Technology, Key Laboratory of Symbol Computation and Knowledge Engineering of the Ministry of Education, and State Key Laboratory of Automobile Simulation and Control. His primary research interests include computer vision, machine learning, embedded system and applications in intelligent vehicle. He has co-authored more than 30 research publications in peer reviewed journals and international conference proceedings.



Xuelong Yin received the BSc degree in Software Engineering from Jilin University in 2016. He is currently a postgraduate student in College of Computer Science and Technology, Jilin University. His main research interests include image processing and computer vision.



Jing Luan received the BSc degree in Software Engineering from Jilin University in 2015. She is currently a postgraduate student in College of Software, Jilin University. Her main research interests include image processing and computer vision.



Tong Liu is currently a postgraduate student in College of Computer Science and Technology, Jilin University. His main research interests include image processing and computer vision.