A Rapid Road Image Mosaic Method Based on Monocular Camera

Wunan Li\textsuperscript{a}, Yu Cao\textsuperscript{b,*}

School of Mathematics and Physics, Qingdao University of Science and Technology, No. 99 Songling Road, Qingdao, Shandong, China
\textsuperscript{a}+8618562789773, liwunan_nudt@hotmail.com, \textsuperscript{b}+8613156255568, caoyu.gfkd@gmail.com

ABSTRACT. As an important infrastructure, road has attracted a lot of attentions in recent years. Most of current road information extraction methods rely on either elaborate algorithms or complex equipment. This manuscript proposes a simple and rapid image mosaic method to extract road surface based on a monocular camera. First, key frames are extracted to remove redundant information from image sequences. To eliminate perspective effect, extracted key frames are transformed into top view of road images based on inverse perspective mapping algorithm with the help of an attitude tensor. Then, a coarse-to-fine registration strategy is proposed to align all transformed key frames into a unified coordinate system. Finally, considering the specificity of transformed images, a superposing and overlapping image fusion strategy is proposed to alleviate the effect of stitching seam. The experiments are conducted on a road with the size of 128×17 meters, and the experimental results demonstrate the effectiveness and efficiency of proposed method. In conclusion, the proposed method is simple, effective and efficient and can be applied in a wide range of applications such as large scale road surface diagnose.

1. INTRODUCTION

With the development of computational and sensing technologies, researchers have paid more and more attentions on intelligent transportation systems (ITS). ITS, which aims to achieve traffic safety and efficiency, is the integration of communication, control, and sensing technologies. As introduced in \cite{1}, ITS includes following components: advanced transportation management systems, advanced traveler information systems, advanced vehicle control systems, business vehicle management, advanced public transportation systems, and advanced urban transportation systems. As an indispensable component, vision-based devices have been broadly employed in recent years, because of their abilities in transferring information directly and cost of performance. Consequently, a large number of applications in ITS are implemented with vision-driven technologies, where input data are collected from visual sensors, and output is used for ITS related application \cite{1}.

Road, as an important component of ITS, has been attracted a lot of attentions in recent years. Wang \textit{et al.} \cite{2} proposes a road networks extraction method from very-high-resolution (VHR) aerial and satellite imagery. In this paper, a neural-dynamic tracking framework is proposed to extract road networks based on deep convolutional neural networks (DNN) and a finite state machine (FSM). Wei \textit{et al.} \cite{3} proposes a road structure refined convolutional neural network (RSRCNN) approach for road extraction in aerial images. Cheng \textit{et al.} \cite{4} proposes a cascaded end-to-end convolutional neural network (CasNet) to simultaneously cope with the road detection and centerline extraction tasks.
Wang et al. [5] reviews the technologies of road extraction from a remote sensing (RS) image in recent years. After analyzing different road features and road models, the authors classify road extraction methods into the classification based methods, knowledge based methods, mathematical morphology based methods, active contour model based methods, and dynamic programming based methods. Wang et al. [6] proposes an automatic road surface extraction method from mobile laser scanning point clouds of large scale complex urban environment, which modeled the road surface as 3D planes and the model parameters were estimated through the generalized projection based M-estimator (GPBM). Wang et al. [7] proposes a road boundaries detection method from mobile laser scanning (MLS) point clouds in an urban environment. The key idea of method is directly constructing a saliency map on 3-D unorganized point clouds to extract road boundaries.

The methods described above are either based on aerial and satellite images or based on 3d laser scanning point clouds. Compared with optical camera images, the data described above are not convenient to be acquired. To extract large scale road surface with optical camera, in this paper we propose a road image mosaic method based on monocular camera. The camera is erected with an angle on the roof of a car, and the image sequences of one side of road are continually acquired when the car is moving. In addition, an attitude sensor is also equipped beside the camera to get the real time poses of the camera. Usually, images sequences acquired by the camera are at high frame rate, which leads to a large overlapping rate between adjacent images. Thus, in this manuscript we propose to use key frames to remove redundant information in image sequences. Then, an inverse perspective mapping algorithm [8] is used to transform key frames from perspective view into top view with the help of an attitude tensor. To align all transformed key frames into a unified coordinate system, a coarse-to-fine registration strategy is proposed. In coarse matching stage, the SURF [9] feature points are extracted and described for each image, and then feature points are matched in the space of feature descriptor. Because this step only matches feature points based on local neighborhood information of feature point, many outliers exist inevitably. In fine matching stage, RANSAC (Random Sample Consensus) [10] is used to remove outliers, which considers the consistency of global geometric transformation between matched point pairs. Finally, considering the specificity of transformed images and the computational efficiency, a superposing and overlapping image fusion strategy is proposed to alleviate the effect of stitching seam.

The rest of this manuscript is organized as follows: Section 2 describes related work of the proposed method, and section 3 describes the proposed method. The evaluation and discussion is presented in section 4, and the concluding remarks are presented in section 5.

2. RELATED WORK
Image registration and image fusion [11-16] are two important steps of image mosaic technology. In the following content, we will make a brief review of these two steps.

2.1 Image Registration
Image registration is the critical step of image mosaic technology. As described in [17], image registration methods are mainly divided into three classes: frequency domain based methods, area based methods and feature based methods.

The frequency domain based methods transform images into frequency domain and the optimal transformation parameters are estimated using the property of phase correlation. The classical algorithm of this class is proposed in 1975 [18]. This method calculates the relative translation in frequency domain between image sequences. However, it only works when the translation transformation exists. Authors in [19] improves the phase correlation method and extends it to image registration with translational and rotational transformations. Afterwards, some image registration methods based on FFT (Fast Fourier Transform) [20, 21] have also been proposed. However, the FFT based method suffers from being sensitive to noise.

The area based methods mainly rely on computation between “window” of pixel values in the two images [17]. DI Barnea et al. first proposed a fast and simple algorithm based on sequence similarity
detection (SSDA) [22]. A. Rosenfeld proposed a template matching method based on cross-correlation function [23]. P. Viola utilizes interactive information as a measure function to achieve image registration [24]. However, the area based methods are dependent on the image brightness, and a significant mosaic error will occur if there is a large exposure difference among images. In addition, due to large parallax and low overlap rate, the serious mismatching, structural distortion and ghosting may take place when images are acquired from different viewpoints.

Compared with the two above-motioned classes, the feature-based registration methods are not only with higher accuracy and precision, but also with more robustness on real applications. A typical feature based image registration process is described as follows: first, feature points are extracted from two adjacent images; second, the matched point pairs are extracted based on feature descriptor matching methods or RANSAC-like methods; at the final step, the optimal transformation is estimated between the two images. Obviously, feature points extraction and description are the fundamental steps of feature based registration methods. In recent years, many classical feature point extraction and description method have been proposed, such as Harris corner detector [25], SUSAN [26], SIFT [27], SURF [9], FAST [28], BRIEF [29] and ORB [30]. Considering robustness, rotation invariance and computational efficiency, this manuscript utilizes SURF feature extraction and description method.

2.2 Image Fusion

After image registration, two consecutive images need to be stitched and blended together to create a seamless mosaicked image. In recent years, many researches focus on elimination of mosaic lines and ghosting parts while retaining original image information. According to the different gradations, image fusion algorithms can be divided into pixel level, feature level and decision level methods [17, 31, 32].

Among the three levels of image fusion algorithms, pixel level methods are the most mature and widely used one. Pixel level methods include direct averaging method, weighted averaging method, pyramid decomposition method and optimal seam-based method [32]. Direct averaging method directly averages pixel values of overlapping regions of two registered images. Weighted averaging methods assign different weights to different pixels based on a distance map. Pixels near the center of an image are weighted heavily and those near the boundaries are weighted lightly [17]. The pyramid decomposition methods highlight image details and smooth image transition by obtaining images at different scales [33]. The optimal seam-based method attempts to minimize the visibility of seams by looking for optimal seams in the joining boundaries between the images [34].

The feature level method extracts feature information for image mosaic. Compared to pixel level algorithm, feature level methods are with a more efficient computation. A typical feature level method is based on the wavelet transformation, which the image signal is decomposed into low-frequency contours and high-frequency details. Then, image fusion is implemented at each level based on common fusion methods such as weighted averaging method. At last, the final mosaicked image is obtained by inverse wavelet transformation [35].

The first step of decision level method is feature extraction [36], then discrimination and decision rules are selected to deal with the extracted feature information, and finally images are fused according to corresponding credibility. These kinds of methods have strong anti-interference ability but it cannot work without support from database and decision system.

3. METHOD

The principle of our proposed manuscript is shown in Fig. 1. The gray hexagonal area is the field of camera view on the road. In order to obtain a large-scale road image, the camera is erected with an angle on the roof of the car, and image sequences of one side of road are continually acquired while the car is moving. The Inverse Perspective Mapping (IPM) is used to transform the original image into the top view of the road surface with the help of attitude sensor [8, 37]. However, high frame rate of image sequence and large size of image (shown by the larger rectangular frame in the Fig 1) increases the computational burden of IPM algorithm. In order to improve the image mosaicking efficiency, the image needs to be cropped to preserve the area with higher resolution, and the cropped image is as
shown in the white rectangular area in Fig. 1. Then, the cropped images are aligned together and fused to get the final road surface image.

![Figure 1. Principle of road image mosaic system.](image1.png)

Fig. 2 shows the workflow of proposed rapid image mosaic algorithm. The proposed algorithm mainly contains four main steps: key frames extraction, image transformation based on IPM, coarse-to-fine image registration, and image fusion.

![Figure 2. Workflow of road image mosaic.](image2.png)

3.1 Key Frames Extraction

Usually, images captured by cameras are at a high frame rate, such as 6 frames per second in our experiments. Such a high frame rate leads to a large overlapping regions and a high information redundancy between adjacent images. Using all of captured images will lead to a large amount of computational burden. Thus in this manuscript we propose to use key frames to improve computational efficiency on the premise of no information loss. Each key frame is extracted at a given intersection, such as 12 in our experiments.

3.2 Inverse Perspective Mapping Transformation

In this manuscript, we denote an point in world coordinate system as \( Q(X_w, Y_w, Z_w) \), and denote the corresponding point in pixel coordinate system as \( q(u, v) \). Then the relationship between these two points is formulated as follows:

\[
Z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & t \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix}
\]

(1)

where \( u_0 \) and \( v_0 \) represent the coordinates of principle point relative to the origin in pixel coordinate system, \( Z_c \) is the normalization coefficient, \( d_x \) and \( d_y \) are the physical size of the pixel in the horizontal and vertical directions, \( f \) is the focal length of camera, \( R \) and \( t \) are the rotation and
translation of camera with respect to world coordinate system. The above equation can be further simplified as follows:

\[
\begin{pmatrix}
Z_e \\
u \\
v \\
1
\end{pmatrix} = M_1 M_O
\begin{pmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{pmatrix} = M
\begin{pmatrix}
X_e \\
Y_e \\
Z_e \\
1
\end{pmatrix}
\]

(2)

\[
M_1 = \begin{pmatrix}
\frac{1}{d_x} & 0 & u_0 & f & 0 & 0 & 0 \\
0 & \frac{1}{d_y} & v_0 & 0 & f & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0
\end{pmatrix}
\]

(3)

\[
\begin{pmatrix}
Z_e \\
u \\
v \\
1
\end{pmatrix} = M_1 M_O
\begin{pmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{pmatrix} = M
\begin{pmatrix}
X_e \\
Y_e \\
Z_e \\
1
\end{pmatrix}
\]

(4)

where \( M_1 \) denotes the inner parameter matrix, \( M_O \) denotes external parameter matrix, and the multiplication of \( M_1 \) and \( M_O \) represents the projection matrix \( M \). If the origin of camera coordinate system coincides with the origin of the world coordinate system, then only rotation matrix \( R \) can impact the external parameter matrix \( M_O \). We denote the world coordinate system as \( X_G \) and the camera coordinate system as \( X_S \). Then, \( X_G = R_{GS} X_S \), and \( M_O \) can be written as:

\[
M_O = \begin{pmatrix}
R_{GS} & 0 \\
0 & 1
\end{pmatrix}
\]

(5)

Traditional IPM method assumes that \( R_{GS} \) is fixed in motion. However, the camera attitude is inevitably changing while the car is moving. Assuming \( R_{GS} \) unchanged will lead to large registration errors. In order to get the real-time pose of camera, an attitude sensor is used to get real-time attitude of camera, which is critical to guaranteeing the dynamic accuracy of IPM transformation. Given the real-time external matrix, the top view road image can be computed according to method in [8]. Fig. 3 shows the road image before and after IPM transformation. From Fig 5, we can observe that the transformed image eliminates the perspective effect. In addition, we observe that the resolution varies from regions. This is because the camera is equipped with tilt, thus road surface closer to the camera owns a higher resolution.

![Figure 3. Images before and after IPM transformation.](image)

### 3.3 Rapid Road Image Registration based on a Coarse-to-Fine Registration Strategy

The main steps of rapid road image mosaic are as follows:

Firstly, to ensure that the finally mosaicked image is with a high resolution, regions with low resolutions after IPM transformation must be cropped out. In addition, the image sequences are usually acquired with a high frame rate, which ensures the quality of final image. For each cropped image, feature points are extracted and described by SURF algorithm.
Secondly, a coarse-to-fine method is proposed to deal with registration between two consecutive images. At the coarse stage, matched point pairs are extracted based on descriptor of SURF algorithm. Actually, the SURF descriptor is a local feature descriptor, which means each point is described based on their local neighborhood information. Because of the high discriminative ability, SURF based image registration algorithm has been successfully applied in many domains [38, 39]. However, the discriminative ability is degraded in texture less areas which is usual in road surface. Thus, there exists many false matched point pairs in this stage and an outlier removal algorithm is needed. RANSAC algorithm is a general parameter estimation approach designed to cope with a large proportion of outliers in the input data. Comparing to other model estimation algorithms, RANSAC can deal with data set with more than 50% outliers [10]. In this manuscript, we proposed to use RANSAC algorithm to remove outliers and estimate the model parameters. In this stage, images are transformed into top view, which makes rigid transformation more suitable than perspective transformation to model the geometric relationship between images after IPM. Thus, in this manuscript, we propose to use RTS (rotation, translation and scaling) transformation model in RANSAC algorithm. We assume the set of matched point pairs in two consecutive images is \( \{(x_i, y_i); (x'_i, y'_i)\}, i = 1, 2, 3...n \), where \( n \) is the number of matched point pairs. For RTS transformation model, the relationship between two matched points is described as follows:

\[
\begin{bmatrix}
  x \\
  y \\
\end{bmatrix} = s \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta \\
\end{bmatrix} \begin{bmatrix}
  x' \\
  y' \\
\end{bmatrix} + \begin{bmatrix}
  dx \\
  dy \\
\end{bmatrix}, i = 1, 2, 3...n
\]

(6)

where \( s \) stands for the scale change factor, \( \theta \) indicates the angle at which image is rotated, and \( (dx, dy) \) is the translation. Since \( s, \theta, dx, dy \) are the only unknown quantities in (6), thus only two pairs of matching feature points are needed to calculate all the parameters in model. After estimating parameters, the images are transformed to be registered together.

### 3.4 Superposing and Overlapping based Image Fusion

In this step, an image fusion strategy is adopted to achieve seamless image mosaicking. Traditional fusion methods such as pyramid decomposition method always have a large computational burden, which is not suitable for real-time applications. In addition, comparing with usual images, our images have following two factors which limit the effectiveness of traditional fusion methods:

#### 3.4.1 Black Areas

Because of the limitation of camera view, there exist black areas in cropped images after IPM transformation as shown in Fig. 3. This means no information is acquired in black areas. However, traditional fusion methods are usually based on pixel information of overlapping regions.

#### 3.4.2 Different Angles

Because images are taken under different view angles when the car is moving, the occlusion area caused by objects above the road surface differs in each image. Fusion based on current methods will make the content in overlapping regions blurry.

Therefore, in this manuscript we propose a method of superposing and overlapping pixels directly to accomplish image fusion. If the overlapping area in the latter image does not contain any road information, such as the black area in Fig. 3, it will be automatically discarded, otherwise it will be used to overlie the corresponding areas of the front image. This process is iterated until all images have been processed. The proposed fusion method is simple, effective, and more importantly efficient. This makes the proposed method can be applied in real time applications such as real-time road surface disease diagnosis.
4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Experimental Setup
As shown in Fig 4, the camera and the attitude sensor are mounted and fixed on the top of the car. The camera is tilted downward, and the height from ground to the camera is about 3.62 meters. The optical axis of camera is perpendicular to the traveling direction, and the road surface on one side of the car is scanned. The attitude sensor is used to monitor the camera pose in real time, which is used to transform captured images dynamically according to inverse perspective mapping algorithm.

![Equipment setup.](image)

The left image in Fig. 5 shows one of original road images captured by our equipment. The size of captured images is 640 × 480 pixels, as shown in the left of Fig 5, and the image after IPM transformation is 378×756 pixels, as shown in the right of Fig 5. As described in section 3.2, the resolution of transformed image varies according to the distance to the camera. Thus, in the experiment only one square area with high resolution is cropped out from transformed image, and the size of it is 230×230 pixels, as shown in the right of Fig 5. To evaluate the mosaicking performance of proposed algorithm, a section of cable is laid on the road surface and the performance can be evaluated by judging the continuity of the cable in the final fused image.

![Road surface image before and after IPM.](image)

4.2 Experimental Results
The experimental results are conducted on two different situations. The first is the car moving along a straight line and the second is the car moving along a curve line. In the first experiment, the images are captured at 6 frames per second. The whole capturing time consumes about 102.5 seconds and a total of 615 images are obtained. According to the distance traveled by the car and the consumed time, the averaged running speed of the car in this experiment is about 1.25 m/s. Since the camera only travels about 0.2 m forward during the time step of two adjacent images, the overlapping area between two adjacent images is very large. Thus key frames are extracted for every 12 frames of images to improve the computational efficiency, and a total of 51 key frame images are selected.
The mosaicking result is shown in Fig. 6. The size of image after mosaicking is 6118×813 pixels, and it corresponds to the actual size of road surface which is about 128×17 meters. As observed in Fig 6, the cable curves laid on the road surface are continuous and smooth. The image mosaic result of second experiment is shown in Fig. 7. As observed in Fig 7, the cable in the mosaicking image is still continuous and smooth. Thus, the proposed algorithm still achieves good performance when the car is moving along a curve line.

The algorithm of this manuscript is implemented based on MATLAB platform on a machine with a core of AMD Athlon64 X2 Dual 4600+ clocked at 2.4GHz and 1G RAM. The whole processing time takes about 77 seconds. The image registration algorithm takes about 60 seconds, and the key frame extraction and image fusion take about 17 seconds.

4.3 Discussion
In this subsection, the geometric relationships of each image relative to the first image are analyzed. Fig. 8 shows the relative positional relationship between images calculated by the proposed coarse-to-fine registration algorithm using the cropped image sequences. The curve in Fig 8 consists of 615 points. The horizontal axis represents the travelling distance of each image relative to the first one, and the vertical axis represents the deviation of accumulated distance relative to the first image. The coordinate of each point represents the location displacement of its corresponding image relative to the first image. From the end point of the curve, it can be inferred that travelling distance of the car is about 128.3 meters. Fig. 9 shows the relative rotation angles between images calculated by the proposed registration algorithm. The horizontal axis represents the point index, and the vertical axis represents accumulated angle relative the first image.

Figure 6. Image mosaic result along with a straight line.

Figure 7. Image mosaic result along with a curve line.

Figure 8. Location displacement of image sequence against the first image.
Figure 9. Rotation angle of image sequence against the first image.

Figure 10. Image mosaic result based on perspective projection model.

Figure 11. Image mosaic result based on RTS model.

The comparison between perspective projection model and RTS model is also conducted in this subsection. The mosaicking image based on perspective projection model is shown in Fig 10, and the corresponding image based on RTS model is shown in Fig 11. From Fig 10, we can observe that there exists large scale change between images, although the seams between images are smoother compared to Fig 11. This is because perspective projection model has more freedom than RTS model, whereas RTS model restricts the transformation between images to rotation, translation and scaling.

5. CONCLUSION AND FUTURE WORK
This manuscript proposes a rapid image mosaic method used to extract road surface from monocular camera. The proposed method uses key frames to remove redundant information, uses a coarse-to-fine registration method to estimate model parameters, and a fast superposing and overlapping method is used for image fusion. The experimental results demonstrate that our method is robust enough under different traffic scenes. Thus, the proposed method is simple, effective and efficient enough for practical applications such as road surface diagnosing. In the future work, we consider extending the proposed method to a system which setting up cameras on both sides of the car to simultaneously scanning the road surface.
FOUNDING
National Natural Science Foundation of China (NSFC) (61705261).
China Postdoctoral Science Foundation (2014m562649).

REFERENCES
[1] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. J. I. T. o. I. T. S. Chen, “Data-driven intelligent transportation systems: A survey,” vol. 12, no. 4, pp. 1624-1639, 2011.
[2] J. Wang, J. Song, M. Chen, and Z. J. I. J. o. R. S. Yang, "Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine," vol. 36, no. 12, pp. 3144-3169, 2015.
[3] Y. Wei, Z. Wang, M. J. I. G. Xu, and R. S. Letters, "Road structure refined CNN for road extraction in aerial image," vol. 14, no. 5, pp. 709-713, 2017.
[4] G. Cheng et al., "Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network," vol. 55, no. 6, pp. 3322-3337, 2017.
[5] W. Wang et al., "A review of road extraction from remote sensing images," vol. 3, no. 3, pp. 271-282, 2016.
[6] H. Wang, C. Wang, Y. Chen, W. Yang, and J. Li, "Extracting road surface from mobile laser scanning point clouds in large scale urban environment," in 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014, pp. 2912-2917: IEEE.
[7] H. Wang et al., "Road boundaries detection based on local normal saliency from mobile laser scanning data," vol. 12, no. 10, pp. 2085-2089, 2015.
[8] M. Bertozz, A. Broggi, A. J. I. Fascioli, and v. computing, "Stereo inverse perspective mapping: theory and applications," vol. 16, no. 8, pp. 585-590, 1998.
[9] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in European conference on computer vision, 2006, pp. 404-417: Springer.
[10] M. A. Fischler and R. C. J. C. o. t. A. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," vol. 24, no. 6, pp. 381-395, 1981.
[11] Y.-S. Chen and Y.-Y. Chuang, "Natural image stitching with the global similarity prior," in European Conference on Computer Vision, 2016, pp. 186-201: Springer.
[12] Z. Huiqing, Z. Jingli, and D. Ruyong, "A fast image matching research based on MIC-SURF algorithm," in The 27th Chinese Control and Decision Conference (2015 CCDC), 2015, pp. 542-547: IEEE.
[13] K. Lin, N. Jiang, L.-F. Cheong, M. Do, and J. Lu, "Seagull: Seam-guided local alignment for parallax-tolerant image stitching," in European Conference on Computer Vision, 2016, pp. 370-385: Springer.
[14] X. Jin et al., "A survey of infrared and visual image fusion methods," vol. 85, pp. 478-501, 2017.
[15] S. Li, X. Kang, L. Fang, J. Hu, and H. J. I. F. Yin, "Pixel-level image fusion: A survey of the state of the art," vol. 33, pp. 100-112, 2017.
[16] B. Zitova, J. J. I. F. Flusser, and v. computing, "Image registration methods: a survey," vol. 21, no. 11, pp. 977-1000, 2003.
[17] D. Ghosh, N. J. J. o. V. C. Kaabouch, and I. Representation, "A survey on image mosaicing techniques," vol. 34, pp. 1-11, 2016.
[18] H. Xie, N. Hicks, G. R. Keller, H. Huang, V. J. C. Kreinovich, and Geosciences, "An IDL/ENVI implementation of the FFT-based algorithm for automatic image registration," vol. 29, no. 8, pp. 1045-1055, 2003.
[19] E. De Castro, C. J. I. T. o. p. a. Morandi, and m. intelligence, "Registration of translated and rotated images using finite Fourier transforms," no. 5, pp. 700-703, 1987.
[20] X.-x. Wu, B.-l. Guo, and J. Wang, "Octa-log-polar Fourier transform for image registration," in 2009 Fifth International Conference on Information Assurance and Security, 2009, vol. 1, pp. 601-604: IEEE.
[21] B. S. Reddy and B. N. J. I. t. o. i. p. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," vol. 5, no. 8, pp. 1266-1271, 1996.

[22] D. I. Barnea and H. F. J. I. t. o. C. Silverman, "A class of algorithms for fast digital image registration," vol. 100, no. 2, pp. 179-186, 1972.

[23] A. Rosenfeld, Digital picture processing. Academic press, 1976.

[24] P. Viola and W. M. J. I. j. o. c. v. Wells III, "Alignment by maximization of mutual information," vol. 24, no. 2, pp. 137-154, 1997.

[25] C. G. Harris and M. Stephens, "A combined corner and edge detector," in Alvey vision conference, 1988, vol. 15, no. 50, pp. 10-5244: Citeseer.

[26] S. M. Smith and J. M. J. I. j. o. c. v. Brady, "SUSAN—a new approach to low level image processing," vol. 23, no. 1, pp. 45-78, 1997.

[27] D. G. J. I. j. o. c. v. Lowe, "Distinctive image features from scale-invariant keypoints," vol. 60, no. 2, pp. 91-110, 2004.

[28] D. G. Viswanathan, "Features from accelerated segment test (fast)," ed: nd, 2009.

[29] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in European conference on computer vision, 2010, pp. 778-792: Springer.

[30] E. Rublee, V. Rabaud, K. Konolige, and G. Orb, "An efficient alternative to SIFT or SURF," in Proceedings of International Conference on Computer Vision, pp. 2564-2571.

[31] R. J. N. P. Szeliski, "Image alignment and stitching: a tutorial, foundations and trends in computer graphics and computer vision," vol. 2, no. 1, p. 120, 2006.

[32] L.-h. CAI, Y.-h. LIAO, D.-h. J. C. T. GUO, and Development, "Study on image stitching methods and its key technologies [J]," vol. 3, 2008.

[33] A. Pandey and U. C. Pati, "A novel technique for non-overlapping image mosaicing based on pyramid method," in 2013 Annual IEEE India Conference (INDICON), 2013, pp. 1-6: IEEE.

[34] M. El-Saban, M. Izz, A. Kaheel, and M. Refaat, "Improved optimal seam selection blending for fast video stitching of videos captured from freely moving devices," in 2011 18th IEEE International Conference on Image Processing, 2011, pp. 1481-1484: IEEE.

[35] H. Li, B. Manjunath, S. K. J. G. m. Mitra, and i. processing, "Multisensor image fusion using the wavelet transform," vol. 57, no. 3, pp. 235-245, 1995.

[36] S. Liu, P. Du, and S. J. Y. X.-I. o. R. S. Chen, "A novel change detection method of multi-resolution remotely sensed images based on the decision level fusion," vol. 15, no. 4, pp. 846-862, 2011.

[37] M. Bertozzi and A. J. I. t. o. i. p. Broggi, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," vol. 7, no. 1, pp. 62-81, 1998.

[38] P. Lukashevich, B. Zalesky, S. J. P. R. Ablameyko, and I. Analysis, "Medical image registration based on SURF detector," vol. 21, no. 3, p. 519, 2011.

[39] M. Teke and A. Temizel, "Multi-spectral satellite image registration using scale-restricted SURF," in 2010 20th International Conference on Pattern Recognition, 2010, pp. 2310-2313: IEEE.