

Research Article

3D Road Scene Monitoring Based on Real-Time Panorama

Yuezhou Wu,¹ Changjiang Liu,^{1,2} Shiyong Lan,¹ and Menglong Yang³

¹ National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610065, China

² Key Laboratory of Higher Education of Sichuan Province for The Enterprise Informationalization and Internet of Things, Sichuan University of Science & Engineering, Zigong 643000, China

³ School of Aeronautics & Astronautics, Sichuan University, Chengdu 610065, China

Correspondence should be addressed to Changjiang Liu; liuchangjiang@189.cn

Received 12 May 2014; Revised 30 July 2014; Accepted 6 August 2014; Published 21 August 2014

Academic Editor: Yuxin Zhao

Copyright © 2014 Yuezhou Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Road monitoring helps to control the regional traffic situation so as to adjust the traffic flow. Real-time panorama is conducive to timely treat traffic accidents and to greatly improve traffic capacity. This paper designs a 3D road scene monitoring framework based on real-time panorama. The system is the combination of large scale panorama, satellite map textures, and 3D scene model, in which users can ramble freely. This paper has the following contributions. Firstly, land-points were extracted followed by motion detection, then comotion algorithm was applied to land-points from adjacent cameras, and homography matrix was constructed. Secondly, reference camera was chosen and transformed to overhead viewpoint; subsequently multiviews were morphed to the same viewpoint and stitched to panorama. Finally, the registration based on high-precision GPS information between 2D road panorama and 3D scene model was also proposed. The proposed framework has been successfully applied to a large road intersection monitoring. Experimental results are furnished at the end of the paper.

1. Introduction

Independent camera only provides local information. As a result, it is difficult to make a reasonable judgment and immediate response to the global situation. Image mosaic can combine a set of images into a larger image with a wider field of view of the scene. It helps to improve observers' spatial awareness.

Over the years, numerous algorithms for image stitching have been developed. Image stitching is typically solved by finding global parametric warps to bring images into alignment [1]. 2D projective warps are parameterized by 3×3 homography, if the scene is planar or if the views differ purely by rotation [2, 3]. The warps of video images from multicameras were described in [4, 5]; however, they only focused on the images (infrared and visible images) of similar viewpoint. Lee et al. [6] introduced an automatic method to warp multiview images to a unified coordinate plane. They proposed time-space registration based on trajectory extraction and on the basis the homography matrices from images to the reference plane were estimated. However the trajectory

accuracy might influence fit precision. Reference [7] presented a method for ground plane estimation from image pairs and land-points, instead of static features such as color, shape, and contours, for image matching. Szlavik et al. [8] proposed a method for matching partially overlapping image pairs where the object of interest was in motion, namely, comotion algorithm. It was valid even if the motion is discontinuous and in an unstructured environment. Szlavik et al. [9] also applied comotion model to project a couple of camera images to the reference road plane.

Recently significant progress in the synthetic vision system (SVS) [10, 11], which is a computer-generated reality system, has been achieved. And that computer-generated image, namely, virtual image, is complementary to optical sensor-based vision. SVS supplies users with good immerse visualization. For example, virtual and reality technology has been studied in landmarks (horizon, runway) detection from an aircraft in low visibility conditions [12, 13].

In this paper, we propose a 3D road monitoring method based on real-time panorama. We have synthesized mosaicked images and static 3D scene model. Firstly, moving

targets were detected and land-points were extracted. Secondly, binary sequences in continuous time T were generated according to whether the pixel is land-point set or not. Subsequently land-points from adjacent cameras were matched according to similarity measure, resulting in homography transformation of adjacent cameras. Thirdly, choose a camera as a reference one and wrap the view to overhead view according to the world coordinates of ground control points (GCPs) and corresponding image coordinates. Thus with the transformation relationships between adjacent cameras, all the views were unified to overhead view. Finally, panorama was generated based on stitching lines. In the end of the paper, we proposed a novel panorama presentation of 3D road scene with real-time traffic situation. Specific process is shown in Figure 1.

2. Transformation of Multiview Based on Land-Points Comotion Method

The multiple views would be unified to the same viewpoint. In this paper 2D projective transformation brings all the images aligned to a common ground plane with overhead view. There are three main steps.

- (1) Homography relations are implemented via land-points comotion statistical maps between adjacent cameras.
- (2) Referenced image is projected to the ground plane on the basis of the world coordinate of GCPs and the corresponding image coordinates.
- (3) From steps (1) and (2), all the views are converted to the overhead view.

2.1. Calculation of Land-Points. Assuming the moving blobs has been detected before the calculation of land-points, land-point [7] is the point of central axis projected to the ground. The calculation of land-point is as follows.

- (1) Binarization of video images using hybrid Gauss background algorithm: moving blobs are denoted as

$$O = \{(x, y) \mid f(x, y) = 255\}. \quad (1)$$

- (2) Trace the contour of O , name the contour C , and only the blobs whose area of external polygon is greater than T_a (≥ 10) pixel² are counted.
- (3) The external rectangle of C is denoted as

$$R = \{(x, y) \mid x_{\min} \leq x \leq x_{\max}, y_{\min} \leq y \leq y_{\max}\}. \quad (2)$$

- (4) Line l_j perpendicular to the x -axis $x = x_j$, $x_{\min} \leq x_j \leq x_{\max}$; the distance from point $P_i(x_i, y_i) \in O$ to line l_j is

$$D(P_i, l_j) = |x_i - x_j|. \quad (3)$$

- (5) Calculate the abscissa of land-point

$$x_{\text{mid}} = \arg \min_j \left\{ \text{median}_{P_i \in O} (D(P_i, l_j)) \right\}. \quad (4)$$

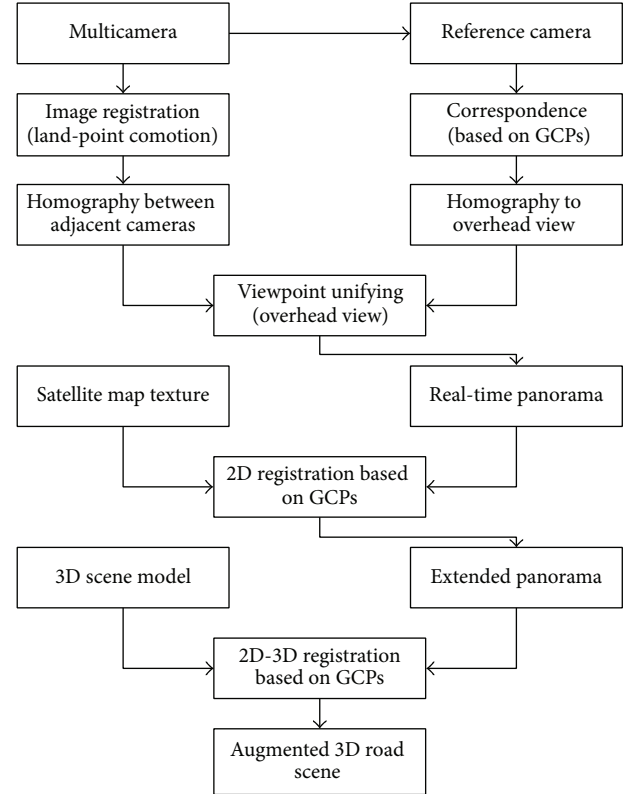


FIGURE 1: Flowchart of 3D road scene monitoring based on real-time panorama.

If there are more than one x_j meeting

$$x_j = \min_j \left\{ \text{median}_{P_i \in O} (D(P_i, l_j)) \right\}, \quad (5)$$

then x_{mid} is the average of these quantities.

- (6) Get land-point $(x_{\text{mid}}, y_{\text{max}})$.

The land-point of moving blobs is illustrated in Figure 2.

2.2. Matching of Land-Points Based on Comotion Method. In our approach no prior knowledge is needed, and the method also works well in images of randomly scrambled motion. It is difficult to realize the registration among images with different position and angle of view. In this paper, the road surface served the reference plane so as to solve the image registration with comotion statistics maps [4, 5]. The hybrid Gauss background model with a fine performance in shadow removal was constructed to detect the foreground of camera video. For any pixel at t th frame, a binary variable $\text{seq}_\Lambda(i, j, t)$ indicates whether the pixel belongs to foreground. Continuous $\text{seq}_\Lambda(i, j, t)$ ($1 \leq t \leq T$) construct sequence vector $\text{seq}_\Lambda(i, j)$ [6],

$$\text{seq}_\Lambda(i, j) = [\text{seq}_\Lambda(i, j, 1), \dots, \text{seq}_\Lambda(i, j, T)], \quad (6)$$

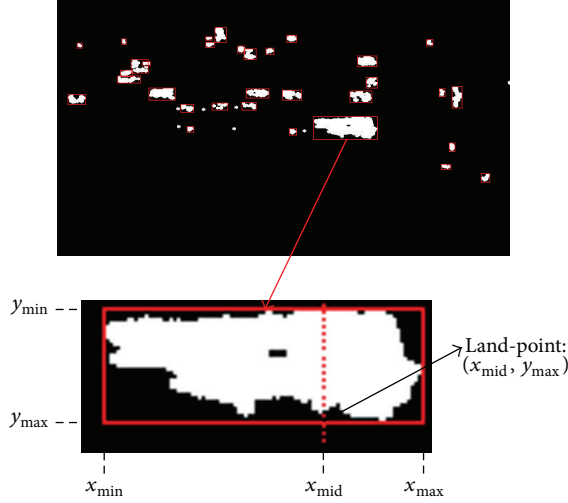


FIGURE 2: Land-point of moving blob.

where

$$\text{seq}_\Lambda(i, j, t) = \begin{cases} 0, & \text{if } (i, j) \text{ is background at } t\text{th frame} \\ 1, & \text{if } (i, j) \text{ is foreground at } t\text{th frame,} \end{cases}$$

$$1 \leq i \leq W_\Lambda, 1 \leq j \leq H_\Lambda, \Lambda \in \{A, B\}. \quad (7)$$

The aforementioned is illustrated in (6). Among them, subscript B is a reference image and A is a matched one. Let W and H denote the width and height of image, respectively. And T is the length of the video to participate in the registration.

Hence coded sequence seq can be divided into two categories: valid motion point and noise point with almost no or frequent movement. Let t_s be static point threshold and let t_n be the noise point threshold. Subsequently characteristic function $\alpha_\Lambda(i, j)$ is defined as follows:

$$\alpha_\Lambda(i, j) = \begin{cases} 0, & \text{if } \|\text{seq}_\Lambda(i, j)\|_1 < t_s \text{ or} \\ & \|\text{diff}(\text{seq}_\Lambda(i, j))\|_1 > t_n \\ 1, & \text{else,} \end{cases} \quad (8)$$

where $\|\cdot\|_1$ is the L_1 -norm and $\text{diff}(\cdot)$ is the difference of vector, for a vector \mathbf{v} is $[v(2) - v(1), \dots, v(n) - v(n-1)]$.

Point (i, j) will be removed from candidate matching pairs if $\alpha_\Lambda(i, j) = 0$. Here similarity measure of the binary sequence is depicted as L_1 -norm. Thus from (8) perfect matching objective function can be written as follows:

$$\text{map}(A(i, j)) = \arg \min_{(p, q) \in U_B} \|\text{seq}_A(i, j) - \text{seq}_B(p, q)\|_1, \quad (9)$$

$$(i, j) \in U_A, (p, q) \in U_B,$$

with

$$U_\Lambda = \{(i, j) \mid \alpha_\Lambda(i, j) = 1, 1 \leq i \leq W_\Lambda, 1 \leq j \leq H_\Lambda\},$$

$$\Lambda \in \{A, B\}. \quad (10)$$

In general, pixel (i, j) which is valid moving point in image A based on (9) matches the moving point (p^*, q^*) in image B if and only if the L_1 -norm of $\text{seq}_A(i, j) - \text{seq}_B(p^*, q^*)$ is minimum.

In (9) the nearest neighbor searching is implemented with ANN (approximate nearest neighbor) library [14] developed by Mount and Arya. Define the dimension of searching space as T and Minkowski norm to be L_1 -norm. ANN supports kd-trees and box-decomposition data structures, which greatly improves the efficiency of high dimensional feature matching.

2.3. *Homography between Adjacent Cameras.* According to the camera model, the relationship between world coordinate and image pixel coordinate is

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \mid \mathbf{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (11)$$

$$= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix},$$

where α_x , α_y , u_0 , and v_0 are called the internal camera parameters, \mathbf{R} and \mathbf{t} are called the external parameters, and Z_c is depth coefficient.

When the world point falls in the $z = 0$ plane, (11) is cast as

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix}. \quad (12)$$

Customarily (12) is rewritten as

$$\lambda_i \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad (13)$$

where $\mathbf{H} = (h_{ij})_{3 \times 3}$ is called a homography.

Based on (9), we get the corresponding land-point pairs between images A and B to estimate 8 parameters of the projective transformation model; pixel (x'_i, y'_i) is the destination pixel in image B corresponding to the source pixel (x_i, y_i) in image A . Substitute $\{x_i, y_i\}$, $\{x'_i, y'_i\}$ in (13) for the land-point pairs; then the homography \mathbf{H} is computed such that

$$\sum_i \left[\left(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left(y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \right]. \quad (14)$$

In many practical situations this assumption is not valid because land-points with noise are mismatched [15]. To improve the robustness of the transformation model, random sample consensus (RANSAC) procedure is applied to the estimation of homography matrix [16].

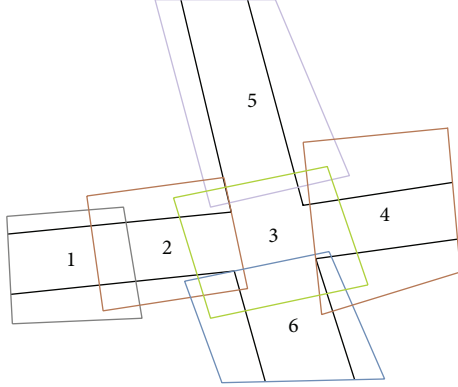


FIGURE 3: Visual field of each camera.

2.4. *Projection of the Reference Camera Based on GCPs.* Given the center of visual range of camera i C_i ($i = 1, 2, \dots$) and the center of the panorama C_0 , then the reference camera i^* is chosen via

$$i^* = \arg \min_i |C_i - C_0|. \quad (15)$$

As shown in Figure 3, camera 3 is the reference camera satisfied with (15).

The view of reference camera will be projected to the ground plane based on GCPs by means of measurement. That is to say, the view has been transformed to overhead view till now. For any quadrangle $ABCD$ on the ground, the intersection of edges AC and BD is the origin of the world coordinate and x -axis parallels AB pointing right. The coordinate system and length of each edge are as shown in Figure 4.

Given $|AB| = a$, $|BC| = b$, $|CD| = c$, $|DA| = d$, $|BD| = e$, and $|AC| = f$, according to simple geometric relations, the vertices of quadrilateral $ABCD$ are as below [17]:

$$\begin{aligned} x_A &= -\frac{\alpha \tan \beta}{\tan \alpha + \tan \beta}, \\ y_A &= -x_A \tan \alpha, \\ x_B &= a - x_A, \\ y_B &= y_A, \\ x_C &= x_B - b \cos \angle B, \\ y_C &= y_B - b \sin \angle B, \\ x_D &= x_A + d \cos \angle A, \\ y_D &= y_A - d \sin \angle A. \end{aligned} \quad (16)$$

In engineering application, so as to get better visual effect, similarity transformation is necessary:

$$x'_i = s_x x_i + \Delta x, \quad y'_i = s_y y_i + \Delta y, \quad i \in \{A, B, C, D\}, \quad (17)$$

where proper translation Δx and Δy makes the coordinate nonnegative, and scaling coefficient determines the height from viewpoint to the ground.

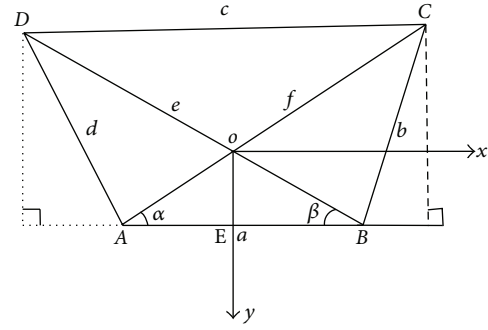


FIGURE 4: World coordinate system.

TABLE I: Parameters of GPS.

Item	Property value
Heading	0.05 degrees (4 m baseline)
Pitch	0.10 degrees (4 m baseline)
Position accuracy	0.3 m
Speed accuracy	0.02 m/s

Suppose the coordinates of 4 points from (16), (17), and the corresponding image points substitute $\{x'_i, y'_i\}$ and $\{x_i, y_i\}$ in (13); then the view of reference camera has been transformed to overhead view.

2.5. *Viewpoints Unifying.* Now suppose there are three cameras. Let G_i be the homography transforming the current viewpoint to overhead one. Suppose camera 2 is the reference camera; then

$$\begin{aligned} G_1 &\cong G_2 H_{12}, & H_{12} &\cong H_{21}^{-1}, \\ G_3 &\cong G_2 H_{32}, & H_{32} &\cong H_{23}^{-1}, \end{aligned} \quad (18)$$

where H_{21} and H_{23} are obtained in Section 2.3 and G_2 in Section 2.4.

3. 3D Road Scene Based on Multiview Panorama

2D multiview panoramas from multiple cameras installed top near the road were generated in real time. 2D panoramic road texture, satellite map texture, and 3D scene model were fused. The process is referred to as 2D-3D image registration. The whole data architecture has been shown in Figure 5.

3.1. *Registration of Panorama and Satellite Map Texture.* GPS geodetic coordinate transformation was introduced in [18, 19]. Suppose point of panorama corresponds to GPS geodetic coordinate and the position precision (Table 1) meets the demand of measurement.

The integrated navigation system XW-ADU5630 was used to obtain the GPS information of panorama's key-points, and the pixels of satellite map texture corresponding to the key-points also have GPS information. Therefore it is easy to

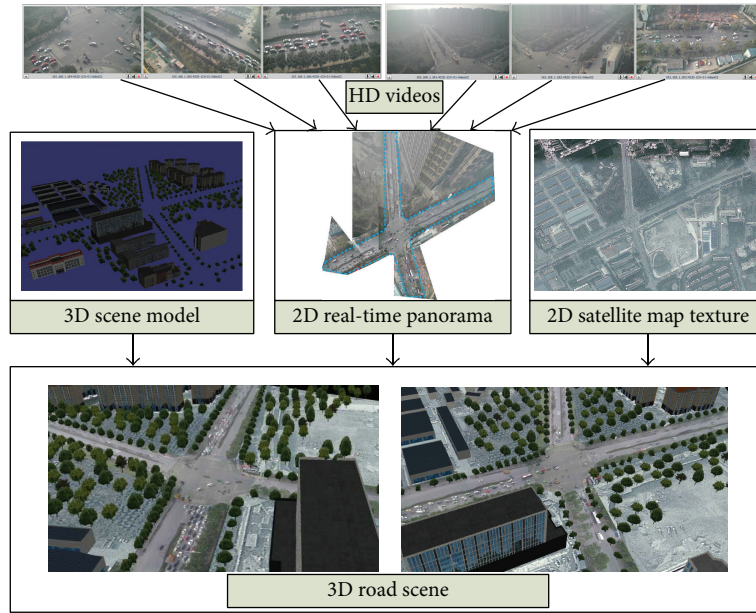


FIGURE 5: Data architecture of 2D and 3D image registration.

match the panorama with satellite map texture based on GPS information.

Applying the set of correspondences (GCPs) to (13), satellite map texture is registered to panorama. Therefore fusion of panorama and satellite map texture can be easily achieved. Panorama with satellite map texture is called extended panorama.

3.2. 2D-3D Image Registration. In our system, all the views are projected to the common road plane, which is the region of interest (ROI) of extended panorama. As a result there is much perspective deformation in top buildings above the ground. Therefore valid road area needs to be clipped out and registered to 3D model (Figure 12). Therefore observers can toggle 3D rendering engine to view the scene freely.

When we align the points on ground in 3D model and corresponding points in extended panorama, 2D-3D image registration is completed. A similarity transformation (or more simply similarity) is conducted so as to align the 3D model and 2D panorama. The similarity can be written more concisely in block form as

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \cong \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \quad (19)$$

A similarity in (19) has seven degrees of freedom and can be computed from four point pairs, that is, four corresponding GCPs. The GCPs are manually picked from ground plane of 3D scene model and 2D extended panorama.

4. Results and Discussion

Our experiments are conducted on the workstation with an Intel Xeon E5-2650 CPU and two graphic cards (NVIDIA GTX Titan Black). Six Samsung web cameras with 2.0-Megapixel are set up at the top of a building next to the road, monitoring a large crossroad. In the experiment, six cameras with three types of focal length (4 mm, 6 mm, and 8 mm) were used.

4.1. Abridged General View of Generation of Extended Panorama

(1) *Motion Detection.* The hybrid Gauss background model was applied to detect moving blobs (foreground). The results are demonstrated in Figure 6. The moving blobs, especially cars, have high accuracy to generate nice matching pairs.

After comotion process, land-points comotion statistics map [9] is shown in Figure 7. The higher the statistics value at a given position the brighter the corresponding pixel in the image.

(2) *Overhead View.* Four points easily to be measured on the ground were chosen, and the corresponding image coordinates were picked manually as shown in Figure 8.

(3) *Panorama and Satellite Map Texture Registration.* Figure 9 is the result of merging ROI of real-time traffic panorama and 2D satellite map texture. It is shown in Figure 9 on the left. On the right there are 6-channel real HD video images.

(4) *Viewpoint Roam.* The 3D model is driven based on OSG. Users can observe the road at any viewpoint. In Figure 10, we



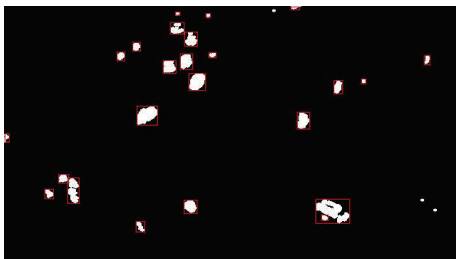
(a)



(b)



(c)



(d)

FIGURE 6: Moving blobs detection: (a) visible image from camera 2; (b) the moving blobs of Figure (a); (c) visible image from camera 3; (d) the moving blobs of Figure (c).

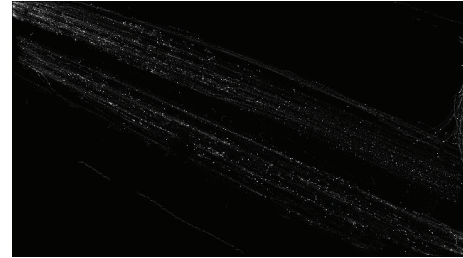
give two screenshots from two viewpoints. It helps observers to inspect the traffic situation in any direction and react promptly in emergency.

4.2. Experiments on Other Roads. Figure 11 shows the experiments on other two roads based on the method proposed. One is the panorama of civic center from 6-channel videos. The other is the panorama of a plaza from 2-channel videos. The polygon region in red dotted lines is mosaicked image.

4.3. Comparative Experiments. A multisource images (visible image, satellite map texture, and 3D scene model) fusion



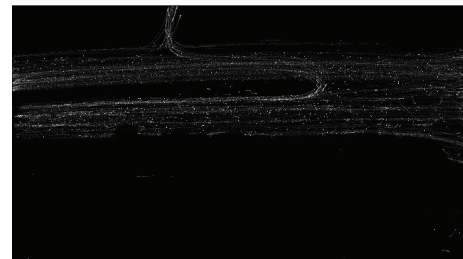
(a)



(b)



(c)



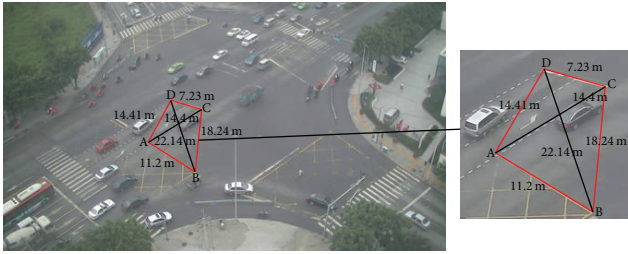
(d)

FIGURE 7: Comotion statistics maps for land-points: (a) visible image from camera 1; (b) comotion statistics map of camera 1; (c) visible image from camera 2; (d) comotion statistics map of camera 2.

method was proposed. By applying the method in [6], we have found the regions in the neighborhood of tall buildings have much projective deformation, as shown in Figure 12. The polygon region in red dotted lines is mosaicked image from real-time video images. And the polygon region in sky blue dotted lines is ROI of panorama, which is clipped out in our method as shown in Figure 12.

5. Conclusion

This paper implements a reasonable integration of 2D multiview overhead panorama, 2D satellite map texture, and



(a)

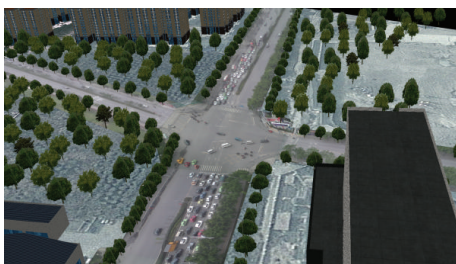


(b)

FIGURE 8: Overhead view: (a) visible image from camera 3; (b) overhead view of Figure (a).



FIGURE 9: Fusion of multiview panorama and satellite map texture.



(a)

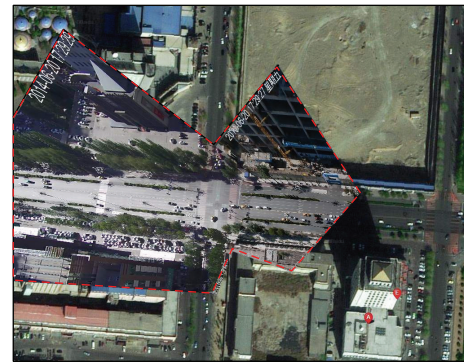


(b)

FIGURE 10: Effect of augmented reality 3D road scene: (a) from viewpoint 1, (b) from viewpoint 2.



(a)



(b)

FIGURE 11: Mosaicked images of two arterial roads.

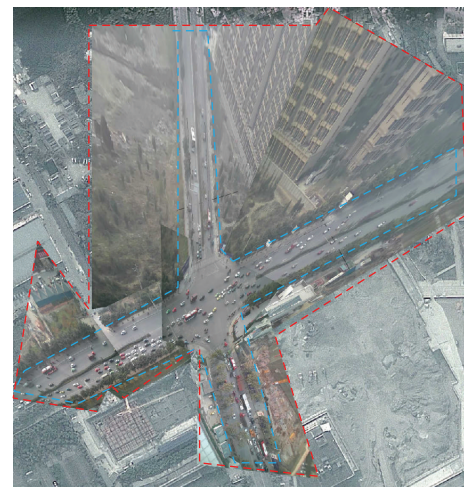


FIGURE 12: Mosaicked images in [6].

3D model to synthesize 3D road scene. The 3D road scene monitoring system has been successfully applied to traffic supervision. The fusion method is innovative and enables the observers to have the sense of immersion. Besides, 3D road scene OSG driven can be translated, rotated, and scaled as one wishes.

In this paper, we focused mainly on the generation of 2D panorama and the presentation of 3D road scene. In future work, we will research on intelligent video analysis and traffic situation analysis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National High Technology Research and Development Program (863 Program) of China under Grants 2012AA011804-1 and 2013AA013802, in part by the Open Research Fund of Key Laboratory of Higher Education of Sichuan Province through the Enterprise Informationalization and Internet of Things under Grant 2013WZJ01, and in part by the Scientific Research Project of Sichuan University of Science and Engineering under Grants 2014PY08 and 2014RC02.

References

- [1] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape preserving half projective warps for image stitching," June 2014.
- [2] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2nd edition, 2003.
- [3] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving dlt," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 2339–2346, IEEE, 2013.
- [4] R. Ai, Z. Shi, D. Xu, and C. Zhang, "A line mapping based automatic registration algorithm of infrared and visible images," in *Proceedings of the 5th International Symposium on Photoelectronic Detection and Imaging (ISPDI '13)*, p. 89072, International Society for Optics and Photonics, 2013.
- [5] X. W. Zhang, Y. N. Zhang, T. Yang, X. G. Zhang, and D. P. Shao, "Automatic visual-thermal image sequence registration based on co-motion," *Acta Automatica Sinica*, vol. 36, no. 9, pp. 1220–1231, 2010.
- [6] L. Lee, R. Romano, and G. Stein, "Monitoring activities from multiple video streams: establishing a common coordinate frame," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 758–767, 2000.
- [7] M. Hu, J. Lou, W. Hu, and T. Tan, "Multi-camera correspondence based on principal axis of human body," in *Proceedings of the International Conference on Image Processing (ICIP '04)*, pp. 1057–1060, October 2004.
- [8] Z. Szlávik, L. Havasi, and T. Szirányi, "Estimation of common groundplane based on co-motion statistics," in *Image Analysis and Recognition*, vol. 3212 of *Lecture Notes in Computer Science*, pp. 347–354, Springer, 2004.
- [9] Z. Szlávik, T. Szirányi, and L. Havasi, "Video camera registration using accumulated co-motion maps," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, no. 5, pp. 298–306, 2007.
- [10] T. Schnell, Y. Kwon, S. Merchant, and T. Etherington, "Improved flight technical performance in flight decks equipped with synthetic vision information system displays," *International Journal of Aviation Psychology*, vol. 14, no. 1, pp. 79–102, 2004.
- [11] A. L. Alexander, C. D. Wickens, and T. J. Hardy, "Synthetic vision systems: the effects of guidance symbology, display size, and field of view," *Human Factors*, vol. 47, no. 4, pp. 693–707, 2005.
- [12] C. J. Liu, Y. Zhang, K. K. Tan, and H. Y. Yang, "Sensor fusion method for horizon detection from an aircraft in low visibility conditions," *IEEE Transactions On Instrumentation and Measurement*, vol. 63, no. 3, pp. 620–627, 2014.
- [13] C. J. Liu, Q. J. Zhao, Y. Zhang, and K. K. Tan, "Runway extraction in low visibility conditions based on sensor fusion method," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1980–1987, 2014.
- [14] D. M. Mount and S. Arya, ANN: library for approximate nearest neighbour searching, 1998.
- [15] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [16] S. Choi, T. Kim, and W. Yu, "Performance evaluation of ransac family," *Journal of Computer Vision*, vol. 24, no. 3, pp. 271–300, 1997.
- [17] T. K. Koo and Y. B. Aw, "A three-dimensional visualization approach to traffic accident mapping," *Photogrammetric Engineering & Remote Sensing*, vol. 57, no. 7, pp. 921–925, 1991.
- [18] D. H. Maling, "Coordinate systems and map projections for gis," in *Geographical Information Systems: Principles and Applications*, pp. 135–146, John Wiley & Sons, 1991.
- [19] R. Reulke, S. Bauer, T. Dring, and F. Meysel, "Traffic surveillance using multi-camera detection and multi-target tracking," in *Image and Vision Computing New Zealand*, pp. 175–180, 2007.