Stabilizing Route Panoramas

Jiang Yu Zheng
Department of Computer Science
Indiana University Purdue University Indianapolis
jzheng@cs.iupui.edu

Abstract

The Route Panorama (RP) has been proposed as a new digital medium to record and visualize cityscapes along a route. It is a compact, continuous and complete visual representation of scenes collected from a sequence of slit views with a camera moving along a smooth path. In real scene acquisition, a camera may suffer from vehicle shaking and the obtained route panoramas are jagged and waved. To improve the quality of route panoramas, we develop an algorithm to stabilize them. By referring the continuous linear features, we use median filters to smooth the 2D route panorama. By setting the slit properly with respect to the camera, we can reduce the influence from the vehicle shaking in the RP without matching consecutive video frames. We have rectified long distance route panoramas to a moderate level for virtual tours and visual navigation.

1. Introduction

The route panorama (RP) has been developed for scene registration and rendering along streets in a city. It is obtained from slit scanning of scenes with a camera mounted on a moving vehicle [2,3,8,12]. Different from local panoramic views [2-4,14], the route panorama archives scenes along streets in compact data sets so as to be widely transmitted over the Internet in real time. Figure 1(a) displays a segment of route panorama with jiggled parts.

Because a route panorama is obtained from a traversing of the street, it is an efficient approach of scene acquisition and even can be implemented in real time. The drawback, however, is the possible loss of image quality due to the connection of pixel lines taken at different time instances. Depending on the road condition, the camera may suffer from acute shaking. We must stabilize the route panorama to a certain degree so that it becomes an acceptable medium for virtual tours and navigation (Fig. 1(b)).

Motion compensation is a common way in stabilization of a video sequence. By matching consecutive images slightly deviated, motion vectors are extracted for rectifying wavy image sequences [9,13]. Mosaicing uses the same approach but stitches consecutive image stripes [10,11]. For a correct overlapping, the consecutive images must contain sufficient features to do matching. For a translating camera viewing scenes with different depths, a perfect overlay of images is theoretically impossible due to the inconsistent disparities. Image matching can only remove the camera rotation effectively. Extra morphing is required to fulfill the stitching.

We will stabilize 2D route panoramas by aligning slit views in accordance with continuous structure lines. Our approach handles much less data than a video sequence, retains the simplicity in the RP acquisition, and allows post-processing after the RP is obtained. The objective here is to remove zigzags and waves on structural lines in order to improve the appearance of RPs, rather than recovering the camera motion or object structures. A set of median filters will be applied to extracted linear features for the RP straightening and smoothing. If continuous structural lines exist in the scenes, our algorithm has the results approximately equivalent to those from the consecutive image matching because the same feature sources are used. Our algorithm is incapable of rectifying scenes only with texture and fragmentary features, e.g., an RP full of trees. A similar work has rectified aerial images from a push broom sensor [15,16].

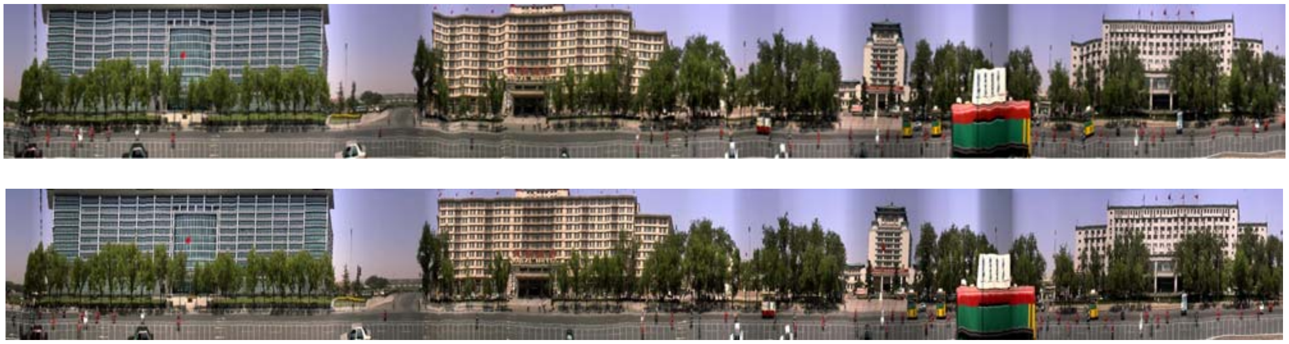

Fig. 1 A segment of route panorama taken from a bus before and after stabilization. (a) Upper image (b) Lower image.

The features are restricted to straight lines. Our algorithm, however, uses a different computation model and works on more general continuous features. A strong tracking and feature selection function automates the robust shape smoothing in long route panoramas.

2. Route Panoramas

2.1 Vehicle motion and slit scanning

To acquire route panoramas to a significantly long distance, the camera is suitable to be fixed on the vehicle and experience the vehicle motion. A possible steady or active camera system is not as stable as the vehicle body, requires high computation cost, and cannot solve the latency problem in the control loop for a sudden shaking.

We decompose the camera motion into an ideal traversing motion and a disturbance from the vehicle shaking. A four-wheeled vehicle performs an ideal motion along a smooth curve including a straight line on a horizontal plane. This smooth curve can be described by an envelope of circular paths with a changing curvature. At each time instance, the vehicle has a translation ***V***(t) rotated by ***R***(t) on the horizontal plane (Fig. 2).

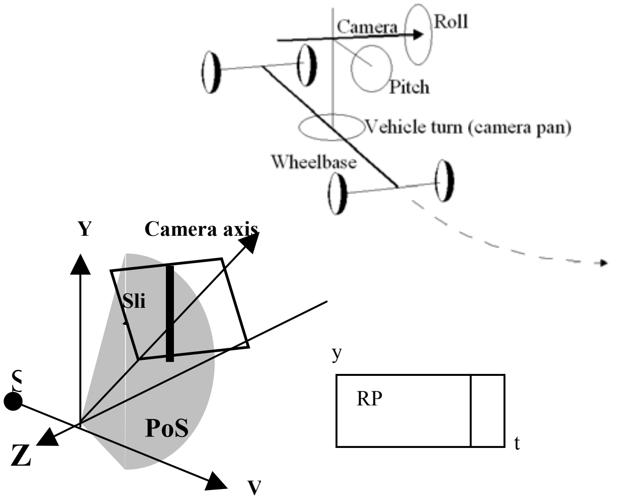

Fig. 2 Camera motion and Plane of Scanning

A vertical *plane of scanning* (*PoS*) is virtually located perpendicular to the vehicle heading direction. The camera can be mounted freely directing to high buildings or low ground, and its axis does not need to be in the *PoS*. The intersection of the *PoS* and the image plane determines a line. Its portion in the image frame is extracted as the slit for scanning scenes along the route. The general motion of the slit can be described by six degree of freedom (T_x, T_y, T_z, R_x, R_y, R_z). Among them, T_z is zero for a four-wheeled vehicle. T_x is the necessary forward motion that achieves scene scanning. R_y occurs when the vehicle moves on curved path and turns at a corner. Besides these three ideal components, T_y, R_x (pitch), and R_z (roll) disturb the RP. The roll R_z can be reduced if the vehicle has a long wheelbase. T_y is a limited displacement when the vehicle bumps over an uneven road. It will not affect the RPs largely because a small up and down translation of the slit will not generate a large vertical disparity on distant objects.

2.2 Shape characteristics in the RP

Instead of matching consecutive images for motion vectors, we examine the smoothness of features in the resulted RPs to infer the possible shaking of the camera. According to the analysis of shapes projected into the RP [2,5,6,12], vertical lines in the 3D space are scanned instantaneously by the *PoS* and appear vertical in the RP. This is invariant to the camera translation and rotation. The vertical lines will not be used as a clue in rectifying the RP.

For a linear camera path, a 3D non-vertical line appears as a hyperbolic curve (Fig. 3) if it changes depth from the camera path when it is scanned, and appears as a horizontal line if it is parallel to the camera path. A slant line in the 3D space also appears as a slant curve or line in the RP depending on its distance from the camera path.

For a curved camera path, non-vertical lines appear as smooth curves in the RP. We can prove that their equations are envelopes of hyperbolic curves with changing parameters. Eventually, they are smooth and non-vertical in the RP.

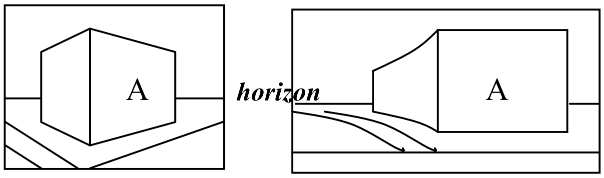

Image Route panorama

Fig. 3 Shape deformation in the RP from the image.

3. Route Panorama Stabilization

3.1 Local camera shaking estimation

To extract the camera (slit) pitch $R_x(t)$ based on the non-vertical features, we first extract edges in the RP along the time axis (Fig. 4). Some of them are not explicitly from horizontal structures, e.g., edges on leaves or boundaries of trees. Next, we track edges horizontally and count a structural line if the tracked length is over a threshold. In order to guarantee the quality of lines, we examine color and contrast at edge points and search forward steps for possible connections of edges.

These lines may still come from a slanting structure such as a bridge or a roof. Making them horizontal in the RP will destroy other structures there. To obtain a reliable pitch at any time instance, we use multiple features since a common jagging of features at a certain instance indicates an explicit pitch disturbance. From each non-vertical line $L_i=(t, y_i(t))$ lasting for a certain length, the change in its y coordinate is converted to the deviation of the assumed camera pitch,

Fig. 4 Horizontal edge detection in an RP and filtered result with straight lines

$$\Delta_i(t) = \arctan\left(\frac{y_i(t)}{f}\right) - \arctan\left(\frac{y_i(t-1)}{f}\right) \qquad (1)$$

At an instance t, we pick up the median value of all the deviations Δ_i, i=1,2,…k, i.e., by

$$\Delta(t) = median[\Delta_1(t), \Delta_2(t), \ldots , \Delta_k(t)] \qquad (2)$$

as the deviation of the camera (slit) pitch at time t. If there is not sufficient number of edges detected at time t, we set $\Delta(t)=0$ and no structure correction will be carried out there. This can avoid miss-assignment of deviation from an unexpected slanting structure.

3.2 Removing joggling by median filter

For the obtained camera deviation sequence $\Delta(t)$ $t\in[0, T]$ over the entire RP, we use the second median filter to remove those abrupt changes in the sequence. As Fig. 5 depicts, a window with size w is shifted along the RP for filtering the deviation $\Delta(t)$. The window size is as large as 1000 pixels to avoid noise along an even longer route panorama of 80000 pixel lines. At either end of the RP, the window shrinks automatically to stay in the RP. A new camera pitch sequence $\Delta'(t)$ is then produced by

$$\Delta'(t) = median[\Delta(t-w/2), \ldots\ldots, \Delta'(t+w/2)] \qquad (3)$$

The difference between the smooth $\Delta'(t)$ and the original $\Delta(t)$, i.e.,

$$E(t)= \Delta'(t) - \Delta(t) \qquad (4)$$

is the shaking in the camera pitch. We use its converted value to the y coordinate to shift the vertical pixel lines in the RP in order to align the continuous features smoothly. Because of a sufficiently long window for the median filter, our approach can remove small and sharp disturbances in the pitch sequence, and large waves as well on structures with many horizontal lines.

3.3 Straighten structural lines

At this stage, a joggled RP has been smoothed according to the composite effect of multiple features. To straighten the solid structure in the RP, we focus on long lines to apply the third median filter. The lines modified in the RP are retraced and uneven parts are further filtered. That is, for any tracked line $L'_i(t)=(t,y'_i(t))$ in the modified RP, we use a window of size λ (e.g., $\lambda=300$) to filter it so that a possible modification $\varepsilon_i(t)$ is obtained at t, i.e.,

$$\varepsilon_i(t) = median[y'_i(t-\lambda/2)\ldots y'_i(t+\lambda/2)] - y'_i(t) \qquad (5)$$

The window is scalable near the ends of the line to produce results. Then, for each t, we compare lengths of all the existing lines $L'_i(t)$, $i = 1, 2, \ldots, k$ there, and find the modification factor from that of the longest line, i.e.,

$$\varepsilon(t) = (\ \varepsilon_m(t) \mid \ \|L'_m(t)\|=\max(\|L'_i(t)\|, i=1,2,\ldots k\ )) \qquad (6)$$

We use this factor obtained from the most reliable line to modify the RP along the time axis. This improves the shape of architectures.

The automatic rectification may fail when continuous deviations are not available due to lack of tracked lines. Foe example, a mild wave is observed on a building with repetitive windows. The wave is longer than piece wised line segments on the repetitive windows, and the deviation is not distinct on each line. In such a case, we throw an elastic curve (snake) that approaches the piece-wised edges. The snake then starts to straighten so that the RP is self-adjusted and the repetitive patterns are aligned straightly.

The final verification is the vertical position of the horizon in the RP, even it might not be visible at some locations. The horizon should not be sheered acutely because of the local adjustment of structural features. We apply a *shear* operation on the RP to make the horizon horizontal.

4. Experiments and Discussion

We have experimented capturing route panoramas for many miles along many streets (Fig. 5). The vehicle moves at a speed of 20 mile/ph. The RP can have a height of 720 pixels and an unlimited length. Virtual tours have been also designed by connecting route panoramas from maps. We apply our algorithms on the input RPs before they are uploaded to the website. During the stabilization, different sizes of edge filters are applied for various densities of features in the RPs. For the three median filters, we use a quick selection algorithm (simplified quick sort) to achieve a fast computation.

The alignment of slit views is implemented at pixel level. Therefore, there is no vertical blurring that drops the image resolution. As compared with the conventional stabilization using 2D images, this approach can achieve the same level of alignment since it uses horizontal features in the image localization. Our approach works not as good as the conventional ones at a forest area, which will not influence the perception of those non-structural features in the RP.

Another point to clarify is that our approach aligns structural lines smoothly in the image space, rather than in a physical space. It keeps the RP in a regular window approximately after stabilization. The conventional approach may lead the rectified view out of the window, either because the elevation changes along a road, or because the error in the motion estimation is accumulated. This is not in favor of the registering, filing, transmission, and display of scenes in a uniformed and efficient way. Our objective is to provide recognizable scenes to viewers in a multimedia system.

5. Conclusion

This work achieves an important step in the presentation of route panoramas taken by a camera on a moving vehicle. We remove zigzags and waves in the route panoramas. The shaking of the vehicle over uneven roads disturbs the camera motion along an ideal smooth path. This affects the image quality of route panorama comprised of slit views at consecutive positions along the path. This work focuses on the non-vertical structural lines in the scenes and realigns the vertical slit views in the RP to straighten the continuous lines. We have improved the quality of the route panorama to a significantly long distance, for which conventional algorithms tracking image sequence may not fulfill because of their tremendous requirement of storage and computational cost.

References

[1] H. Baker, R. Bolles, “Generalizing epipolar-plane image analysis on the spatial-temporal surface” CVPR pp.2-9 (1988).

[2] J. Y. Zheng, S. Tsuji, Panoramic representation of scenes for route understanding, 10th International Conference on Pattern Recognition, Vol. 1, pp.161-167, 1990

[3] J. Y. Zheng, S. Tsuji, Panoramic Representation for route recognition by a mobile robot, International Journal of Computer Vision, (1992), Vol.9, No.1, pp.55-76.

[4] S. E. Chen, L. Williams, Quicktime VR: An image-based approach to virtual environment navigation, SIGGRAPH95, pp. 29-38, 1995

[5] J. Y. Zheng, S. Tsuji, Generating Dynamic Projection Images for Scene Representation and Recognition, Computer Vision and Image Understanding, Vol. 72, No. 3, pp. 237-256, 1998.

[6] R. Gupta, R. I. Hartley, "Linear Push broom Cameras", IEEE PAMI, Vol. 19, No. 9, 1997, pp. 963-975.

[7] S. Li, and A. Hayashi, Robot navigation in outdoor environments by using GPS information and panoramic views, Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 570-575, 1998

[8] S. Ozawa, M. Notomi, and H. Zen: A Wide Scope Modeling to Reconstruct Urban Scene, Proc. ISPRS Commission V, International Symposium on Real-Time Imaging and Dynamic Analysis, pp.370-376, 1998

[9] Z. Zhu, G. Xu, Y. Yang, Jesse S. Jin, Camera stabilization based on 2.5D motion estimation and inertial motion filtering, IEEE International Conference on Intelligent Vehicles, Oct 28-30, 1998

[10] S. Peleg, B. Rousso, A. Rav, Acha, A. Zomet, Mosaicing on adaptive manifolds, IEEE PAMI, Vol. 22, No. 10, pp. 1144-1154, Oct. 2000.

[11] Z. G. Zhu, E. Riseman, A. Hanson, Parallel-perspective Stereo Mosaics, 8th ICCV 2001, Vol. I, pp. 345-352.

[12] J. Y. Zheng, Digital Route Panorama, IEEE Multimedia, July-Sept. pp. 57-68, 2003.

[13] Y. S. Yao and R. Chellappa, Selective Stabilization of Images Acquired by Unmanned Ground Vehicles, IEEE Trans. Robotics and Automation, Vol. RA-13, pp.693-708, Oct. 1997.

[14] H. Ishiguro, M. Yamamoto, S. Tsuji, “Omnidirectional stereo”, IEEE Trans. PAMI, Vol. 14, No. 2, pp.257-262, 1992.

[15] C. Lee, H. J. Theiss, J. S. Bethel, E. M. Mikhail, Rigorous Mathematical Modeling of Airborne Pushbroom Imaging Systems, ISPRS Journal of Photogrammetry and Remote Sensing Vol. 66, No. 4, pp. 385-392, 2000.

[16] C. Lee, J. S. Bethel, Georegistration of airborne hyperspectral image data, IEEE Trans Geoscience and Remote Sensing, Vol. 39, No. 7, pp. 1347-1351, 2001.

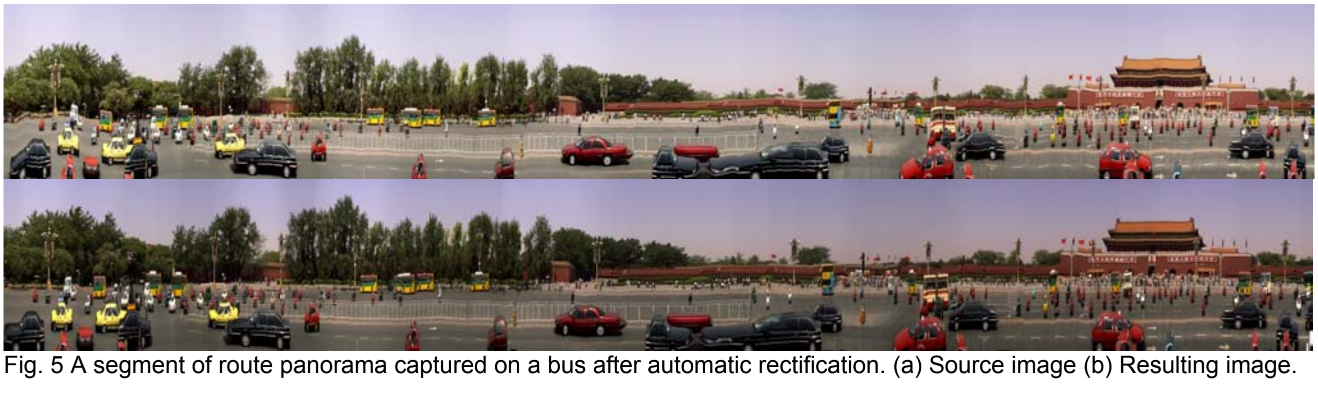

Fig. 5 A segment of route panorama captured on a bus after automatic rectification. (a) Source image (b) Resulting image.