Adjusting Route Panoramas with Condensed Image Slices

Glenn Flora
Indiana Univ. Purdue Univ. Indianapolis
723 W. Michigan St., SL280
Indianapolis, IN 46202
grflora@cs.iupui.edu

Jiang Yu Zheng
Indiana Univ. Purdue Univ. Indianapolis
723 W. Michigan St., SL280
Indianapolis, IN 46202
jzheng@cs.iupui.edu

ABSTRACT
Route panorama (RP) has become an image medium to archive and visualize cityscapes along streets. Collecting data from a sampling line of a video camera mounted on a moving vehicle, the route panorama may suffer from the vehicle shaking and varying speeds due to uneven roads and busy traffic respectively. These may jitter the route panorama and change object widths. To produce a high quality route panorama, we extract spatial-temporal information in the video volume and rectify the route panorama. In this paper, we propose two condensed image slices to record traces of horizontal and vertical scenes during the vehicle motion. Instead of matching video frames that requires cost computation and large storage, we track the feature traces in the condensed image slices to remove jitter and adjust the local length of route panoramas. The method is effective and efficient in acquiring and rectifying long distance route panoramas.

Categories and Subject Descriptors
I.4.1 [**Image Processing and Computer Vision**]: Digitization and Image Capture – *scanning, imaging geometry, sampling.*

I.4.3 [**Image Processing and Computer Vision**]: Enhancement – *filtering, geometric correction, smoothing.*

General Terms:
Measurement, Design Experimentation, Theory

Keywords:
video stabilization, route panorama, camera motion, image representation, media editing

1. INTRODUCTION
In obtaining an extended image along a long street for city visualization, a route panorama has been extracted from a video sequence when the camera translates along a smooth path [1]. Different from local panoramas, the images taken at distributed viewpoints cannot be simply overlapped because of the disparities (or motion parallax). Two approaches so far have been proposed to generate such a route panorama — one is 1D slit scanning [1-7] (Figure 1) and another is 2D image mosaicing [8,9], which results in a parallel-perspective image and a multi-perspective image respectively.

We select the slit scanning approach to acquire route panoramas in a large urban area because of its efficiency. Although the slit scanning does not require inter-frame matching of video during view extension and can be generated in real time, it requires a well-controlled mechanic movement of the vehicle, which may not always be satisfied along a busy and bumping street. Another drawback is the deformation of shapes in the parallel-perspective images. To make the shape close to perspective projections, a dynamic slit selection scheme is also examined after 3D information is obtained [10,11]. On the other hand, the multi-perspective image takes the traditional approach to match consecutive frames and find the camera motion. Based on that, careful stitching by photomontage [12] or interpolation [13] is carried out to obtain high-resolution images. The drawback is the requirement of an equal distance of scenes from the camera path.

In measuring the vertical camera motion that causes instant jitters or local speed changes, we avoid image frame correspondence using feature extraction [13] and intensity [14]. Rather we select a method similar to [15], which rectifies the RP based on tracked structure lines. To overcome the shortcoming of [15] that the line tracking may not be interrupted in the RP due to the occlusion by vertical features, we propose a *condensed image slice* that records the intensity accumulation horizontally in the frames.

Similarly, Epipoloar Plane Images (EPI) have been used to display the feature traces that characterize the horizontal camera motion [16]. However, such an EPI is susceptible to the vertical camera shaking and feature traces are hard to track. To solve this problem, we propose another condensed image slice, namely *condensed motion slice*, which accumulates intensities vertically. It enhances the vertical features that play the most important role in the motion detection. By normalizing the length of RP locally according to the feature traces in the condensed image slice, we can scale the major scenes in the RP to have close aspect ratios as that in perspective images.

In the following sections, we first introduce the motion mode in capturing a route panorama. We propose condensed image slices in Section 3 and describe their properties. Section 4 explains the shaking removing and length adjustment based on the condensed image slices. Section 5 shows some experiment results followed with conclusions.

2. CAMERA MOTION MODEL FOR RP
The RP is captured via a video camera mounted to an automobile. The camera is directed sideways, perpendicular to the vehicle's direction of motion. The pixel data on a vertical line in the image frame are copied to an image memory from every video frame and the extended image forms a route panorama (Figure 1).

The motion of the vehicle can be described as a translation **V**(t) rotated by **R**(t) on the horizontal plane. By setting the coordinate system where the X-axis is parallel to the heading of the vehicle, the Y-axis is vertical, and the Z-axis is perpendicular to the X-axis which points towards the scene, the general motion of the slit can be described by six degrees of freedom (T_x, T_y, T_z, R_x, R_y, R_z) (Figure 2). T_x represents the forward motion of the vehicle. R_y occurs when the vehicle turns and is a necessary component of the RP. T_y represents the vertical translation of the camera. This can be neglected because small vertical displacements (<10cm) of the camera will have a limited image effect on distant objects within the scene; vertical disparities is small for distant scenes. T_z can be assumed zero for a four-wheeled vehicle and careful navigation by the driver. R_z, (roll) will be reduced by a long wheelbase. Finally, R_x (pitch) will disturb the acquisition of the RP by introducing jitters as the camera inclines or declines. As a result, all but two of components can be neglected by simplifying the motion of the camera by assuming that the vehicle follows an ideal curve along the horizontal plane by rotating around the vertical axis.

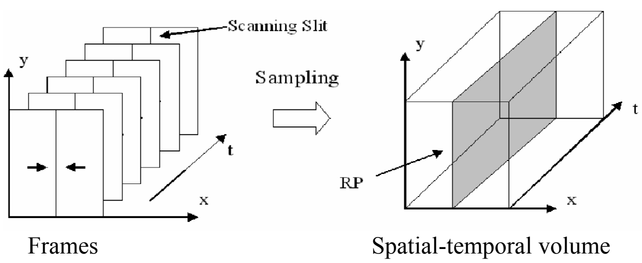

Figure 1 Route panorama acquisition as the vehicle moves (left) and its location in a spatial-time volume of video (right).

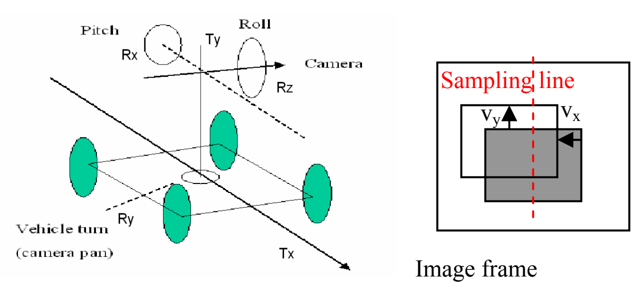

Figure 2 Camera motion with a vehicle and the image velocity due to shaking.

On the image frame, the optical flow of scenes yielded from the camera motion has mainly the horizontal image velocity disturbed by vertical image velocity from time to time. The horizontal component v_x is caused from T_x and its value is proportional to T_x and inversely proportional to the scene distance. The vertical component v_y is caused from T_y and R_x, particularly from R_x on distant scenes.

To remove jitters in an RP caused by vehicle shaking without doing inter-frame matching of video frames horizontal structure lines can be used as references [15]. By straightening a curved horizontal line in the RP, the corresponding section of the route panorama can be adjusted vertically to obtain a better image quality. However, there is no guarantee that horizontal structure lines can be found everywhere in the RP. Those lines can be interrupted easily by the occlusion of objects in front of them such as trees and poles. Some repetitive patterns such as windows on a building are difficult to track as a horizontal line, despite having a horizontal structure.

Similarly, if the vehicle changes its speed during the translation, the image velocities on vertical lines vary. The route panorama changes its length locally with respect to the standard lengths as in the parallel-perspective projection. In the EPI, the traces of vertical lines are curved from straight ones, which give clues to normalize the scene lengths in the RP from the varied vehicle speeds. The problem now is that one EPI only reflects the motion of scenes at a single height. In addition, the shaking of the camera can easily disturb an EPI, and there is no guarantee that all the traces can be tracked stably. To solve these problems, we propose the idea of condensed image slices in section 3.

3. CONDENSED IMAGE SLICES

3.1 Horizontally Condensed Image Slice Displaying Shaking Components

In the spatial-temporal volume of a video, a center vertical slice of single pixels is cut from each frame and concatenated, thus creating the 2D RP (Figure 1). A condensed image slice proposed here, however, is an accumulated average of the data from each frame, as displayed in Figure 3. To acquire the condensed image slice, the data along each horizontal line of the frame is summed and averaged in real time as the frame is captured, and this value is represented as a point. This is repeated for all horizontal lines of each frame to create the single vertical line. The condensed image slice is finally created by concatenating the vertical lines from each frame.

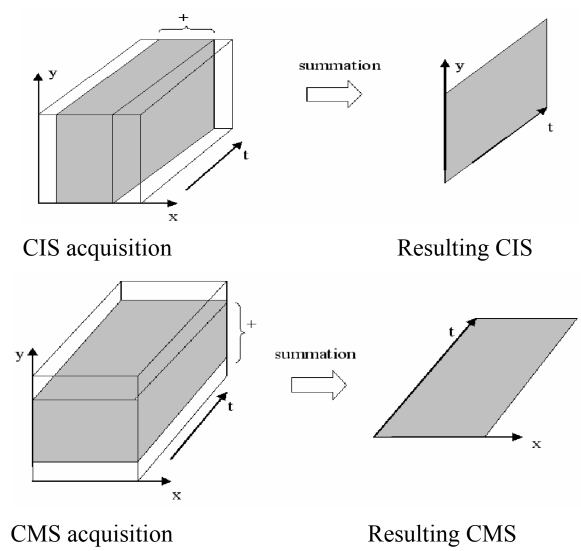

Figure 3 Condensed image slice (top) and condensed motion slice (bottom) in the spatial-temporal volume (left). The accumulation is done in real time as the RP is created.

As shown in Figure 4, vertical and slanted lines (e.g., on roofs) are blurred, but horizontal edges within the original scene will be amplified in the condensed image slice. The horizontal line is counted numerous times in the same position as the camera translates the scene, but the vertical line contributes to the horizontal accumulation equally at each height and thus leaves no traces. For example, if a portion of a horizontal line on a building is occluded by a small tree in the RP, the horizontal line will be brought to the front in the condensed image slice. This will allow for longer horizontal lines to be tracked and straightened, because the dominant horizontal features that determine the vertical flow will be distinct in the condensed image slice. Slanted long lines (e.g., on roofs) will not be mistakenly selected as a reference in rectifying the RP. Repetitive patterns not connected as horizontal lines in the RP still show their continuous traces in the condensed image slice.

3.2 Vertically Condensed Motion Slice

We also explored accumulating the image data in the vertical direction to produce a condensed motion slice. Each vertical line of pixels at each frame is averaged. This process yields a single horizontal line of intensities from each frame. The horizontal intensity lines from successive frames are concatenated along the time axis to create the condensed motion slice. Such an image slice shows the motion traces of major vertical scenes and can be used to correct aspect ratio distortion of objects in the RP caused by camera velocity variation. As shown in Figure 4, the condensed motion slice has recorded traces of vertical edges not influenced by shaking. We select the accumulative range at the middle of the image frame such that vertical features from major buildings along the streets can be captured.

The condensed motion slice represents the traversal of vertical scenes through the spatial-time volume where predominant vertical edges in the scene will be represented by edge traces in the condensed motion trace. The angle each of these edges makes with the time axis is dependent on vehicle velocity during RP acquisition and the depth of objects in the scene. An edge trace should be consistent or straight if the vehicle speed is constant, and its orientation depends on the edge distance from the camera path. The varied vehicle speed during the RP acquisition can be clearly seen within the condensed motion slice on curved traces.

To provide an accurate aspect ratio, similar to the perspective projection, the consecutive vertical columns should perfectly cover the scene surfaces without overlapping or skipping, i.e. at the just-sampling depth [1]. Therefore, the traces should make a 45 degree angle with the time axis within the condensed motion slice.

4. SHAKING RECTIFICATION AND RP NORMALIZATION

4.1 Tracking and Smoothing Features

The rectification of the RP is essentially tracking horizontal feature traces in the condensed image slice to obtain shaking parameters with respect to the horizontal line, and applying the deviation correction parameters to the corresponding RP for the realignment of vertical pixel columns. We use the median filtering to smooth jagged lines locally and straighten waved structures globally in the RP.

First, we track traces in the condensed image slice after horizontal edge detection with a low threshold. An abrupt vertical shaking jitters many horizontal lines. The deviations of them from linear are collected and median filtered at each position to produce a common deviation. With the sequence of the common deviations along the path, we use the second median filter to obtain an ideal trace that is not jittered. The displacements of the deviation sequence form the ideal trace are applied to the RP for adjusting the jagged portions by raising or lowering the entire corresponding column of pixels. This operation can automatically stabilize small jitters in the RP.

Second, we straighten the waves appropriately by following long horizontal traces in the condensed image slice. Long horizontal liens are extracted and straightened. The changes to the condensed image slice are then applied to the RP to rectify waved shapes due to gradual agitation of the road surface. Figure 5 represent a segment of a RP and the corresponding condensed image slice.

Figure 4 Examples of condensed image slice (right) and condensed motion trace (left) with range of summation.

4.2 Tracking Features in Condensed Motion Slice for Length Normalization

We normalize the curved traces in the condensed motion slice in order to scale the RP at the corresponding locations. This can solve the problem of the varied vehicle speed. Also, we scale the width of RP in sections according to the trace orientation of major buildings in the condensed motion slice. This can improve the aspect ratio of scenes with respect to the perspective projection.

First, detect the edges of the traces based on the strong gradient magnitude. For each position in the condensed motion slice, we straighten curved traces by finding the line of best fit on the most prominent edge. The angles of the gradient along the newly straightened traces are compared to the expected angle of the trace, which is a value set tentatively for straightening. Such an expected angle is determined from the angle of nearby traces that have been straightened so far.

The RP is then scaled horizontally segment by segment based on disparities to make the angle of the traces $\pi/4$. Figure 6 shows such a result of RP normalized in length locally so that major scenes in it have close aspect ratios as what are in the perspective images. The improvement is most seen from the more natural width of the buildings and trees.

5. CONCLUSIONS

This work removes the shaking and waving on route panoramas due to camera agitation during image acquisition. It also normalizes the aspect ratio of major scenes in the route panorama. Specifically, it traces longer horizontal edges by amplifying the horizontal structures and blurring the vertical structures, which allows us to follow horizontal line despite possible occlusions. Also, if camera velocity varies while capturing the route panorama, segments of the scene will be expanded or compressed as the velocity increases or decreases respectively. The condensed motion trace shows the velocity without influenced by camera shaking. This work allows us to correct these imperfections by examining the feature trace to create high quality route panoramas.

6. REFERENCES

[1] J. Y. Zheng, S. Tsugi, Panoramic Representation for Route Recognition by a Mobile Robot. Int. Journal on Computer Vision, 9(1), 55-76, 1992.

[2] R. Gupta, R. Hartley, Linear Push-Broom Cameras, IEEE Trans. PAMI, 19(9), 963-975, 1997.

[3] J. Y. Zheng, S. Tsuji, Generating Dynamic Projection Images for Scene Representation and Understanding, Journal of Computer Vision and Image Understanding, 72(3), 237-256, Dec. 1998.

[4] P. Rademacher, G. Bishop, Multiple-center-of-projection Images. SIGGRAPH, 199-206, 1998.

[5] J. Y. Zheng, Digital Route Panorama, IEEE Multimedia, 10(3), 57-68, 2003.

[6] S. Seitz, J. Kim, Multiperspective Imaging IEEE CGA, 23(6), 16-19, 2003.

[7] J. Y. Zheng, Y. Zhou, P. Mili, Scanning Scene Tunnel for City Traversing, IEEE Trans. Visualization and Computer Graphics, 12(2), 155-167, 2006.

[8] Z. Zhu, A. R. Hanson, 3D LAMP: a new layered panoramic representation, ICCV, 723-730, 2001.

[9] A. Agarwala, et al. Photographing Long Scenes with Multi-Viewpoint Panoramas. ACM Transactions on Graphics, 25(3), 853-861, August 2006.

[10] A. Roman, G. Garg, M. Levoy, Interactive Design of Multi-Perspective Images for Visualizing Urban Landscapes, IEEE Conf. Visualization 2004, 537-544, 2004.

[11] D. Zomet, D. Feldman, S. Peleg, D. Weinshall, Mosaicing New Views: the crossed-slits projection, IEEE Trans. PAMI, 741-754, 2003.

[12] Z. Zhu, A. R. Hanson, E. M. Riseman, Generalized Parallel-Perspective Stereo Mosaics from Airborne Video. IEEE Trans. PAMI, 26(2), 226-237, 2004.

[13] J. Weng, Y. Cui, N. Ahuja, Transitory Image Sequences, Asymptotic Properties, and Estimation of Motion and Structure, IEEE Trans. PAMI, 19(5), 451-464. 1997.

[14] S. Srinivasa, Extracting Structure from Optical Flow Using the Fast Error Search Technique, IJCV, 37(3), 203-230, 2000.

[15] J. Y. Zheng, Stablizing Route Panorama, 17th ICPR, 4, 348-351, 2004.

[16] Z. Zhu, G. Xu, X. Lin, Efficient Fourier-Based Approach for Detecting Orientations and Occlusions in Epipolar Plane Images for 3D Scene Modeling. IJCV, 61(3): 233-258, 2005.

[17] J. Y. Zheng, X. Wang, Pervasive Views: Area Exploration and Guidance Using Extended Image media, ACM Multimedia Conference. 05, 986-995, Singapore, 2005

Figure 5 Stabilizing jitters, sections of RP and CIS before shaking removing (top) and after rectification (bottom).

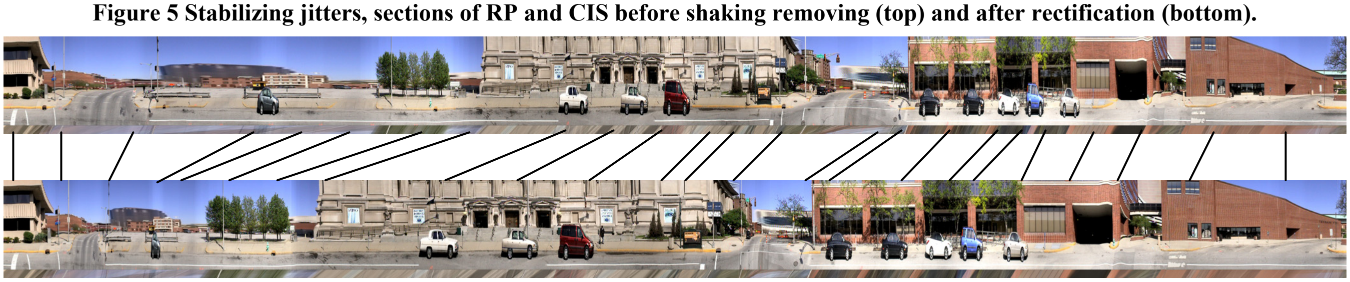

Figure 6 Local length normalization of the RP based on the condensed motion slice before (top) and after (bottom) normalization