MVA2009 IAPR Conference on Machine Vision Applications, May 20-22, 2009, Yokohama, JAPAN

Synchronization of Full-view Image and GPS Data for Route Map Building

Ying Hai
Graduate School of Engineering
Tottori University
t08t2004@faraday.ele.tottori-u.ac.jp

Shigang Li
Graduate School of Engineering
Tottori University
li@faraday.ele.tottori-u.ac.jp

Yuto Mizuno
Graduate School of Engineering
Iwate University
h16j57@cis.iwate-u.ac.jp

Abstract

In this paper we propose a method of synchronizing automatically the full-view image and GPS information which are recorded independently, respectively, for route map building. The full-view image sensor is composed of a pair of fisheye cameras. By synchronizing the two image streams of the pair of fisheye cameras a seamless full-view image is acquired; by synchronizing the full-view image with GPS information the full-view image stream is re-sampled uniformly so that the full-view image is registered to a digital map with approximately equal distance along routes. The above synchronization is achieved by matching the motion of these sensors. To estimate the motion between consecutive frames, a small motion model of spherical image is used. The experimental results show the effectiveness of our method.

1. Introduction

A driving map is usually plotted by lines and curves; it provides the geometrical information of roads. However, humans memorize scene mainly by visual information. To make a map which is easily understood to human visual information is needed to be added to the maps.

To build a large-scale route map including image information, a practical method is to register images to a geometrical digital map in terms of the GPS (Global Positioning System) information. To acquire a free-view-direction image, a full-view image is required. Therefore, to build full-view route map, we need to know the correct position information of the captured full-view image.

In this paper a full-view image sensor and a hybrid GPS sensor are installed on a vehicle as shown in Fig. 1. The full-view image sensor is mounted at the top of the vehicle; it consists of a pair of fisheye cameras which have a wider field of view than a hemisphere, respectively. All of these devices are usual commercial products. The data of these devices are recorded independently. Thus, to register full-view images to a digital map in terms of GPS information, we need to synchronize all the data as follows.

- Synchronize the two image streams of the pair of fisheye cameras of the full-view image sensor.
- Synchronize the full-view image data and GPS data.

In this paper we propose a method of automatically synchronizing these data based on motion information. To synchronize the two image streams, we compute the camera motion by a spherical stereo approach for each fisheye camera, respectively, and compare their motion under the same coordinate system; since their motion is caused by the same vehicle motion the synchronization can be determined by finding the maximum of the correlation values of their motion information. Since the azimuth angle of the vehicle can be acquired directly from our hybrid GPS sensor, their synchronization with full-view image sensor can be achieved by the similar method.

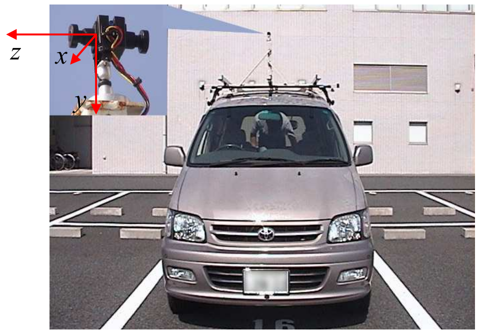


Figure 1. A full-view image sensor and a hybrid GPS sensor are installed on a vehicle to acquire data for route map building.

The rest of this paper is organized as follows. In the next section, we describe the related research. Section 3 describes the method of the synchronization of the pair of fisheye cameras of the full-view image sensor. Section 4 presents the synchronization of the full-view image and GPS data. The preliminary experimental results are given in section 5. Finally, we conclude and present future work in the last section.

2. Related Research

How to generate a visual-information contained map along route automatically is a challenge. To compress the video data captured along routes a compact representation called panoramic representation is proposed [1]; a cognitive map described by landmarks can be generated by analyzing the visual cues of the panoramic representation [2]; a 2D route map can be built as a graph-like representation which includes multiple route panoramic views [3]; further, combining the panoramic views with the GPS information a route map including the geometrical information of routes and the visual cues along route scenes can be acquired [4][5][6].

While the route panoramic view is compact distortions appear in it when the camera rotates, for example, at a junction since a panoramic view is a path-centered representation. Another approach is that the 3D information of the scan line is also acquired by a range sensor [7][8] or stereo vision processing [9] and the 3D model of route scene with texture is built. However, construction of exact 3D model of route scene is not easy.

A straight method is to register full-view image directly to a geometrical digital map in terms of GPS information as the popular Google street-view tool. In this paper we use this approach. A full-view image is captured by a pair of fisheye cameras; the full-view image to be registered is re-sampled uniformly from the recorded video stream based on the GPS information. As described in section 1, we need to synchronize the full-view image sensor and GPS data because these data are recorded independently by different devices with different time stamps.

For the synchronization of multiple cameras, multiple stationary video cameras are synchronized temporally with overlapping views [10]. In [11] it is assumed that the cameras are stationary and take the images of the same scene from various viewpoints, in which there are moving objects such as human in motion. However, in our case the pair of fisheye cameras point two opposite sides there are almost no overlap between the views of the two fisheye cameras.

In this paper the pair of fisheye cameras of the full-view image sensor is very close, and their relative position is fixed and calibrated using the method of [12]. When the vehicle moves the motion consistency should be maintained between the pair of fisheye cameras. We use this cue to synchronize the pair of fisheye cameras. The similar method also is used to synchronize the image sensor and GPS sensors by using the motion information measured from the hybrid GPS sensor.

3. Synchronization of the pair of fisheye cameras of the full-view image sensor

First we describe the idea of the synchronization the full-view image sensor. Next we give the motion estimation method of the fisheye camera based on a small motion model.

3.1. Idea of the synchronization of the full-view image sensor

The full-view image sensor used in our research is shown at the top-left corner of Fig. 1. We choose it due to its simple structure, and it is easy to construct it by commercial parts. Although a single camera full-view image sensor is proposed in [12], a special lens is required.

Suppose that the matrix of the relative pose of the two cameras pointing forwards and backwards, respectively, is R_{rf} which is calibrated beforehand. Since the distance of pair of fisheye cameras is very small, the origins of the two camera coordinates are assumed to be the same. When the vehicle moves we compute the two camera motion, respectively. Since the camera translation can only be estimated apart from the scale factor from an image sequence we use the estimated rotation for the synchronization. Let the estimated camera rotation be R_f and R_r of the pair of fisheye cameras, respectively. We have the following equation.

$$R_f = R_{rf} R_r \tag{1}$$

Since the image streams of the pair of fisheye cameras are recorded independently there may be a time delay, Δt, between them. That is

$$R_{f(t+\Delta t)} = R_{rf} R_{r(t)}. \tag{2}$$

Thus, the task of the synchronization of the full-view image sensor is to determine this time delay, Δt.

Suppose two rotation sequences, $R_{f(t)}$ and $R_{r(t-\Delta t)}$ ($t = 1, \cdots, n$), are computed for the pair of cameras, respectively. By computing the pitch (*x*), yaw (*y*) and roll (*z*) angles from the rotation matrices we have three angle sequences for each fisheye camera.

For example, suppose the roll angle sequences of the pair of fisheye cameras are $X_f(x_{f(1)}, x_{f(2)}, \cdots, x_{f(n)})$ and $X_r(x_{r(1)}, x_{r(2)}, \cdots, x_{r(n)})$, respectively. Then the correlation ρ_X between two X_f and X_r is defined as

$$\rho_X = \frac{\operatorname{cov}(X_f, X_r)}{\sigma_{X_f}\sigma_{X_r}}, \tag{3}$$

where $\operatorname{cov}(X_f, X_r)$ is the covariance of X_f and X_r, σ_{X_f} and σ_{Xr} are standard deviations of X_f and X_r, respectively.

Considering the delay time, Δt, between $X_{f(t+\Delta t)}$ and $X_{r(t)}$ the correlation of roll angle can be defined as

$$\rho_X(\Delta t) = \frac{\operatorname{cov}(X_{f(t+\Delta t)}, X_{r(t)})}{\sigma_{X_{f(t+\Delta t)}}\sigma_{X_{r(t)}}}. \tag{4}$$

The correlation of pitch and yaw, $\rho_Y(\Delta t)$ and $\rho_Z(\Delta t)$, can be computed by the same way. Instead of evaluating the correlations of the three angles, respectively, we compute the rotation correlation, $Corr(t)$, of the two cameras as the product of the correlation of the three angles as follows.

$$Corr(t) = \rho_X(\Delta t)\rho_Y(\Delta t)\rho_Z(\Delta t) \tag{5}$$

Thus, the synchronization is computed by maximizing the above correlation function.

$$\arg\{\max_{\tilde{t}}\{Corr(\tilde{t})\}\} \tag{6}$$

Next, we explain how to compute the fisheye camera rotation from a fisheye image sequence.

3.2. Motions estimation of fisheye camera

A fisheye camera has a wider field of view than a hemisphere. The conventional pin-hole camera cannot be used directly. Here, we use a spherical camera model to compute the camera motion by mapping the point of a fisheye image to a spherical image as in [12].

Since we compute the camera rotation from an image stream which is recorded in 30ftp the motion between consecutive frames is small. Therefore, we use the small motion model of stereo [13] to compute the camera motion, but modify it with a spherical camera model.

Suppose a point, *P*, is observed at two different places in space. Let the coordinates of the projection point on the first spherical image and the coordinates of the projection point on the second spherical image be *m* and m', respectively. The relationship between the two spherical images can be represented by a rotation matrix, *R*, and a translation vector, *t*. For the vectors *m*, m' and *t* in the same plane, we can expressed it as

$$m \cdot [t \times (Rm')] = 0 \tag{7}$$

Let us now turn our attention to infinitesimal displacements. We consider a moving camera with translational velocity *v* and rotational angular velocity *ω*. For small motion, we have

$$\begin{cases} t = \delta t v \\ R = Id + \delta t[\omega_\times] \\ m' = m + \delta t \dot{m} \end{cases} \quad (8)$$

Substituting in equation (5) and neglecting all terms of order two or greater in δt yields we have

$$m^T([v_\times][\omega_\times])m - (m \times \dot{m}) \cdot v = 0 \quad . \quad (9)$$

(9) is a nonlinear equation of v and ω. If v or ω is assumed to be known the equation becomes linear. Here we use the following iteration method to estimate v and ω.

Algorithm: solving linearly translational velocity v and rotational velocity ω

If the vehicle is moving, the following computation begins.

1. Suppose ω_o =0.
2. Compute v in terms of (9) by using known ω_o .
3. Compute ω_n in terms of (9) by using the computed v.
4. Let $\omega_o = \omega_n$.
5. If the following error is small than a threshold, T_e , output v and ω as answer. If not, go to 2.

$$\sum_{i=1}^{N}\left(m_i^T([v_\times][\omega_\times])m_i - (m_i \times \dot{m}_i) \cdot v\right)^2 \le Te \quad (10)$$

Where N is the number of the feature points.

Using the computed rotation angle of a consecutive image sequence of the pair of fisheye cameras we compute the delay time Δt for the synchronization as described in subsection 3.1.

4. Synchronization of full-view image sensor and hybrid GPS sensor

The hybrid GPS sensor used in our research consists of GPS receiver and inertial sensor, and is connected with the speedometer of the vehicle. It outputs the position, the velocity, and the azimuth angle of the vehicle by the rate of 4 fps. Here we synchronization of the full-view image sensor and the hybrid GPS sensor based on the measured azimuth angle as follows.

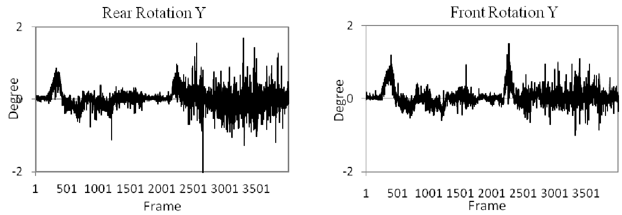
- Carry out interpolation for the azimuth angle from the hybrid GPS data so as to acquire the same data rate as the image stream.
- Accumulate the estimated angle on the ground between the neighboring frames of the image sequence captured by the full-view image sensor.
- Compute the correlation of the interpolated azimuth angle, $\theta_{GPS(t)}$, of the hybrid GPS sensor and the accumulated angle, $\theta_{\mathrm{Image}(t+\Delta t)}$, on the ground from the full-view image sensor by the same method as (4).

After synchronizing the full-image data and the hybrid GPS data, we can re-sample the full-view image stream based on the synchronized GPS information so that the neighboring frames registered to digital map have approximately the same distance.

5. Experiments

In this experiment, a full-view camera and a Pioneer GPS receiver, GPS-MIZZ, are mounted on a vehicle as show in Figure 1. The full-view image sensor consists of a pair of fisheye cameras which 185° field of view, respectively. Two sample images captured by one of the pair of fisheye cameras are shown in Fig. 2. The red and blue points are the detected feature points for computing camera motion. The two estimated rotation angle sequences from the image sequences of the pair of fisheye cameras, respectively, are plotted in Fig. 3.

Figure 2. Two neighboring sample images captured by one of the pair of fisheye cameras.

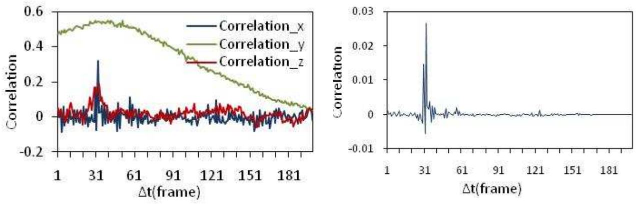


rear camera | front camera

Figure 3. The two estimated rotation angle sequences from the image sequences of the pair of fisheye cameras, respectively

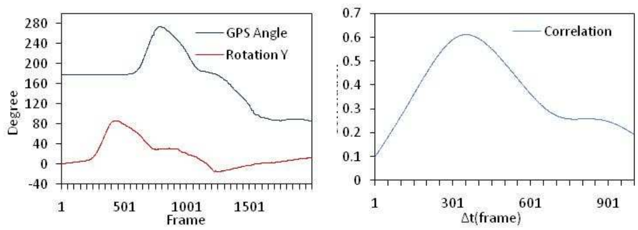
The correlation of the three angles, the pitch (x), yaw (y) and roll (z), is shown the Figure 4(a). A distinguished peak appears in the product of the correlation of the three angles, as shown in Fig. 4(b). The synchronization is determined at frame 32.

The azimuth angle from the hybrid GPS data and the accumulated yaw angle of the rotation between full-view image frames are shown in Figure 5(a). The computed correlation between the both is shown in Fig. 5(a) where the synchronization is determined at frame 354.



(a) (b)

Figure 4. (a) The correlation of the three angles, roll (x), pitch (y) and yaw (z). (b) A distinguished peak appears in the product of the correlation of the three angles.



(a) (b)

Figure 5. (a) The azimuth angle from the hybrid GPS data and the accumulated angle of the rotation between full-view image frames. (b) The computed correlation between the both.

The re-sampling result of the full-view image stream is shown in Figure 6, where the red dots indicate the position of the captured full-view image stream according to the hybrid GPS information and the blue dots indicate the position of the full-view image re-sampled in terms of the synchronized information. Since the velocity of the vehicle may change due to urban traffic situation the density of the red dots varies greatly along the route. By re-sampling uniformly the full-view image streams in terms of the displacement of the vehicle the full-view image is registered to the digital map with approximately equal distance between neighboring frames along routes.

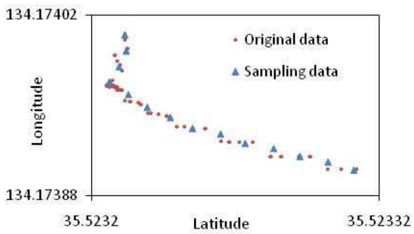
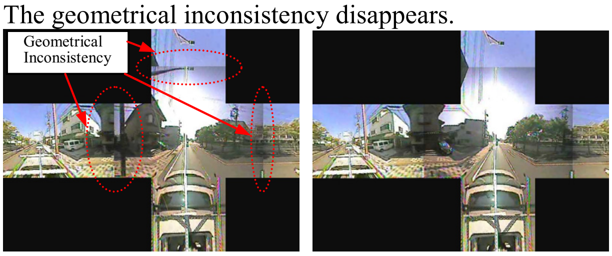


Figure 6. The re-sampling result of the full-view image stream. The red dots indicate the position of the captured full-view image stream according to the hybrid GPS information and the blue dots indicate the position of the full-view image re-sampled in terms of the synchronized information.

The effect of the synchronization of the pair of the fisheye cameras of the full-view image sensor is shown in Figure 7(a) and (b). Figure 7(a) shows the result of a full-view image generated from the images captured by the pair of fisheye cameras without synchronization. The distinguished geometrical inconsistency can be observed along the boundary of the two hemispherical views. Figure 7(b) shows the result generated with synchronization. The geometrical inconsistency disappears.



(a) (b)

Figure 7. The effect of the synchronization of the pair of the fisheye cameras of the full-view image sensor for generating a full-view image.

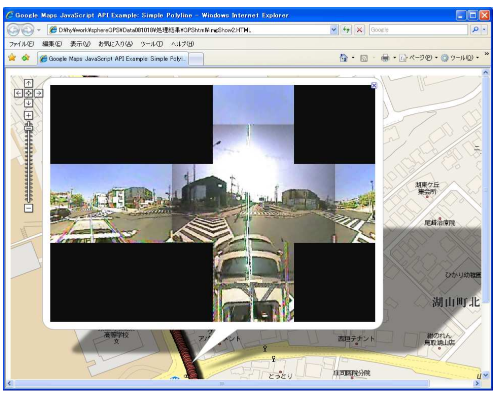


Figure 8. A full-view image is displayed at the clicked position of the Google map.

Finally, we also register the uniformly re-sampled full-view image in the Google map by using the open Google Map API and release in [14]. A sample image of the result is shown in Figure 8 where a full-view image is displayed at the clicked position of the Google map.

6. Conclusions

In this paper we propose a method of building full-view route map from the independently recorded information of the two fish-eye cameras and the hybrid GPS receiver in the circumstance of vehicle-mounted. There are mainly the following contributions.

- In the conventional multiple camera synchronization method it assumes that overlap views exist among these cameras. In this paper we synchronize a pair of fisheye camera by the motion consistency without the assumption of overlap views.
- We also propose a method of synchronizing the full-view image and the hybrid GPS data so as to register the full-view image along routes with approximately equal distance between neighboring frames.

The results of experiment show that the synchronization was effectively solved. It simplified the process of full-view route map building.

Smoothly playing full-view image sequence continuously on the map is a future work to do.

References

[1] J.Y. Zheng and S. Tsuji, “Panoramic representation for route recognition by a mobile robot”, IJCV, 9(1), pp.55-76, 1992.

[2] S. Li and S. Tsuji, “Selecting Distinctive Scene Features for Landmarks”, Proceeding of IEEE 1992 International Conference on Robotics & Automation, pp.53-59, 1992.

[3] S. Li, A. Ochi, Y. Yagi and M. Yachida, “Making 2D Map Based upon Route Scenes”, Autonomous Robots, Vol.8, No.2, pp.117-128, April 2000.

[4] S. Li and A. Hayashi, “Robot Navigation in Outdoor Environments by Using PGS Information and Panoramic Views”, Proceeding of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.570-575, 1998.

[5] J.Y. Zheng, “Route panorama”, IEEE Multimedia, 10(3), pp.57-68, 2003.

[6] J.Y. Zheng and S. Li, Employing fish-eye camera for scene tunnel scanning, ACCV, pp.106-117, January, 2006.

[7] C. Frueh and A. Zakhor, “3D model generation for cities using aerial photographs and ground level laser scans”, CVPR, 2001.

[8] J. Hu, S. You and U. Neuman, “Approaches to large-scale urban modeling”, IEEE Computer Graphics and Applications, 23(6): 62-69, 2003.

[9] N. Cornelis, K. Cornelis and L.V. Gool, “Fast compact city modeling for navigation pre-visualization”, CVPR, 2006.

[10] Anthony Whitehead , Robert Laganiere , “Temporal Synchronization of Video Sequences in Theory and in Practice”, IEEE Workshop on Motion and Video Computing, Vol.2, pp. 132-137,2005

[11] Ushizaki.M et al. “Video Synchronization Based on Co-occurrence of Appearance Changes in Video Sequences”, Proc of ICPR, Vol.3, pp. 71-74,2006

[12] Shigang Li, “Monitoring around a vehicle by a full-view spherical image sensor”, IEEE Trans. On Intelligent Transportation System, Vol.7 No. 4, pp.541-550, 2006.

[13] D.A. Forsyth and J. Ponce, Computer Vision A Modern Approach, Prentice Hall, 2003.

[14] http://www.skybird.ele.tottori-u.ac.jp/data/mva2009/googleMap.html