MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11 - 13, 2002, Nara - ken New Public Hall, Nara, Japan

13—3

Extraction of Road Traffic Information from the Time Sequence of Aerial Images

Motoi Soga Hitoshi Saji[1]
Graduate School of Science and Technology, Shizuoka University

Abstract

In this paper, for analysis of road traffic information, we propose the method of extracting moving vehicles from time sequential images taken by a video camera on the airplane. The camera angles and positions are varied depending on the flight conditions, and hence, the locations of the road and vehicles in the images are varied. In this research, we use the template matching and perspective transformation for aligning the coordinates of the images taken from several points and angles. For extracting moving vehicles from matched images, we compute two differences from time-sequence of three images. Then, we extract the moving vehicles by the logical operation on the differences. This method enables the extraction of the vehicles on the road with one time search and can be applied to the real time analysis of the road traffic information on the wide area.

1 Introduction

Now the vehicle is important for our life. Comparing with other means of transportation, the flexibility of the movement is high, and it has various advantages. However it has many problems, such as the traffic accident and traffic jam. In particular, traffic jam is one of the unavoidable problems for a vehicle and has great influence on the life of people.

In this reason, road traffic information is necessary in the various scenes in the actual world. For example, when a wide-area natural disaster such as the big earthquake occurred, grasping the areas of the disaster and analyzing the situation are required. For reducing damage, the establishment of the route for the emergency vehicle is required. In addition, at the time of usual, many people need information of traffic jam on the road. In the present, this information is grasped using the sensor and the observation camera on the land. However these equipments are set at some space intervals and it is difficult to obtain the information of wide-area at once.

Furthermore, when the wide-area natural disaster is arisen, the sensors on the land for observing the traffic condition may be broken and it is difficult for the vehicles to find the passable road. To solve these problems, it is effective to grasp and analyze the information from the air.

Many methods using the aerial images are developed [1][2]. There is recognition of the subject using the various image features and map information [3]. In the analysis of traffic information, the motion of the vehicles on the straight road or the intersection are analyzed [4][5][6].

In this paper, for analysis of road traffic information, we propose the method of extracting and tracking the moving vehicles from time sequential images taken from the airplane. The aerial images are influenced by the geometrical change when it is taken by the video camera from the airplane. For reducing the influence, we first match the time sequence of the images taken from several points and angles using the perspective transformation [7].

2 Extraction Method

2-1. Summary

It is easy to extract the moving objects from the time sequence of images taken from the fixed camera. In particular, if the images are taken under the uniform lighting, we can extract the moving objects by computing difference between background image and each frame in the time sequence of images.

On the other hand, in this research, we extract the moving vehicles from the time sequence of aerial images taken by the video camera from the airplane. In this case, it is necessary to correct the geometrical changes from the position and angle of the camera.

In this research, we extract the moving vehicles after matching aerial images (Fig.1).

2-2. Matching Aerial Images

We use template matching and perspective transformation for matching the sequence of images. However, by the general template matching method, the correctness of the result depends on the template region. If the template includes moving vehicles, the result of the template matching is influenced by the motion of the vehicle. For obtaining stable matching results, we use voted block matching (VBM) method [8]. After obtaining the matching results, we align the whole coordinates of the images using projective transformation.

[1] Address: 3-5-1 Johoku, Hamamatsu-shi, Shizuoka 432-8011, Japan. E-mail: saji@cs.inf.shizuoka.ac.jp

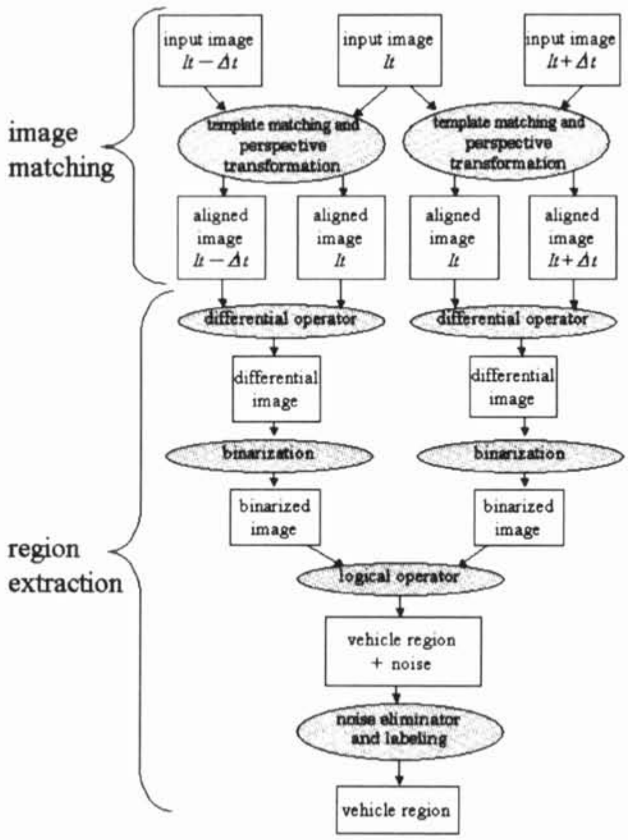

Fig.1 Outline of the extraction method

2-2-1. Template Matching using VBM

When a template is set up on the road in a frame, the moving vehicles may be inside the template. In this case, if we use the general template matching method, we may obtain wrong result of the matching, since the matching result is influenced by the motion of the vehicles. For this reason, we use VBM as the template matching method. VBM is the matching method using vote by block matching and the advantage of VBM is hard to be influenced by the occlusion.

First, VBM divides a template image into several blocks (Bs×Bs), and correlation function is used for matching of each block (Fig.2). By using correlation function, the result of the matching is not influenced by the uniform illumination changes. Next, VBM gives vote to the corresponding coordinates whose rate of matching is maximum on the object image. The Vote is given to coordinates of the vote space corresponding to each pixel in the object image. The coordinate whose sum of the vote to the whole template is maximum is obtained as the matching result.

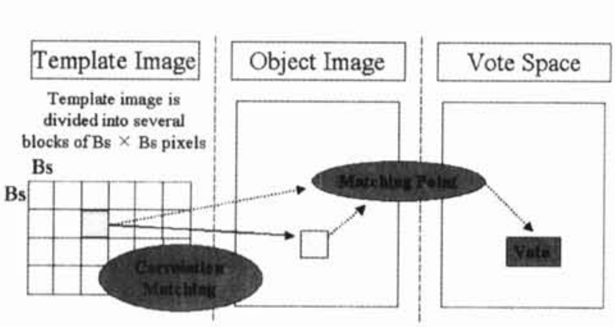

Fig.2 Voted Block Matching (VBM)

In this method, on the blocks of neighborhood of the vehicles in a template, the vote is given to wrong coordinates by the motion of vehicles. On the other blocks, the vote is concentrated on correct coordinates. By setting up the area of the template much larger than the area of the vehicles, the correct position can be obtained without influence of the motion of vehicle.

The images It-Δt, It, and It+Δt are the frames at t-Δt, t, and t+Δt. We define the feature point in the frame It and extract the region of the neighborhood as a template. Then we match frames It-Δt and It+Δt using the template.

2-2-2. Projective Transformation

By the result of template matching, the correspondent points between the coordinates (x_t, y_t) of feature points on the frame It and the coordinates $(x_{t+\Delta t}, y_{t+\Delta t})$ of feature points on the frame It+Δt is obtained. In this research, the coordinates between these frames are corresponded using the following equation of projective transformation.

$$x_t = \frac{S_1 x_{t+\Delta t} + S_2 y_{t+\Delta t} + S_3}{S_7 x_{t+\Delta t} + S_8 y_{t+\Delta t} + 1} \quad (1)$$

$$y_t = \frac{S_4 x_{t+\Delta t} + S_5 y_{t+\Delta t} + S_6}{S_7 x_{t+\Delta t} + S_8 y_{t+\Delta t} + 1} \quad (2)$$

Si are the projective transformation parameters (i= 1,2,...,8). The projective transformation parameters between two frames are computed by solving four and more feature points. By matching the coordinates of the frame It+Δt to the coordinates of the frame It using the obtained parameters, we align the whole coordinates of two frames. In the same way, we compute the projective transformation parameter between the frame It-Δt and It, and we match the coordinates of the frame It-Δt to the coordinates of It.

2-3. Extracting Moving Vehicles

We extract moving vehicles from sequence of three images. First, we compute the difference between corresponding pixels of the frame It and It-Δt. In the same way, we compute the difference between the frame It and It+Δt. Next, we binarize using different threshold of each differential image, then we compute logical operation using the binarized images. Finally, we extract vehicles using labeling operator and noise elimination.

2-3-1. Differential Operation

In the two images whose coordinates are aligned, we compute differential values of the pixels of images. We compute sum of the difference of RGB values in each pixel.

By this process, we compute the difference of the whole images from the sequence of images. As a result, we obtain two differential images from two sets of time-sequential aligned images.

2-3-2. Logical Operation

The error by using the projective transformation is appears in the whole image. When we extract moving vehicles, the coordinates of the background in the image have great errors. Moreover, the differential value between background and moving vehicles is large, however the differential value between moving vehicles is small. Therefore we use "logical and image" and "logical or image" obtained from images binarized by different threshold values.

- Logical and

We extract the pixel whose values are different on both of two differential images. By using the low threshold, we obtain the image including the region of the moving vehicles and background noise.

- Logical or

We extract the pixel whose values are different on either of two differential images. By using the high threshold, we obtain the image including the neighborhood of the moving vehicles.

2-3-3. Noise Elimination and Labeling

We obtain the regions of vehicles by applying "logical and operation" between "logical and image" and "logical or image" computed above. We apply labeling method and noise elimination method (shrink and expand) to the regions. Then small region of noise is rejected. Finally, by gathering adjoining regions, we obtain the regions of the moving vehicles

3 Experimental Results

We extract the moving vehicles using the technique of the preceding chapter. We processed the 30 images (320x240pixels, 256levels, frame rate 5/30 second). We set up four feature points at equal intervals on the frame It, and extracted template of 80×60 pixels around the feature points. Then we matched It-Δt and It+Δt using the template. From this result, we aligned the coordinates between the frames using the projective transformation, and computed difference between them (Fig.3, Fig.4). Finally, we extracted vehicles from the result of computing difference using logical operation, labeling method, and noise elimination. Fig.5 shows the regions of the vehicles with white rectangles.

Furthermore, we experimented two cases, in which the threshold of the binarization of "logical and" and "logical or" is fixed or varied optimally in each image. The results are shown in Table 1 and Table 2.

- Threshold of logical and is fixed to10.
- Threshold of logical or is fixed to 50.

	Extraction Result of the Vehicles(66totals)		Incorrect Extraction	
	A Number of Vehicles	Rate of Extraction(%)	Building	8
I	59	89	Shadow	6
II	6	9	Etc	42
Total	65	98	Total	56

I : extraction with accuracy
II : one vehicle is extracted as two regions

Table.1 Result in the case threshold is fixed

- Threshold of logical and is varied from 5 to 10.
- Threshold of logical or is varied from 50 to 95.

	Extraction Result of the Vehicles(66totals)		Incorrect Extraction	
	A Number of Vehicles	Rate of Extraction(%)	Building	2
I	63	95	Shadow	4
II	2	3	Etc	5
Total	65	98	Total	11

I : extraction with accuracy
II : one vehicle is extracted as two regions

Table.2 Result in the case threshold is varied

Table.1 shows that the rate of extraction of vehicles is high, however many error regions are extracted. By fixing the threshold, background noise cannot be reduced. Table.2 shows that all items are improved in accuracy by varying the threshold of the binarization,

We describe the cause of the errors as follows:

- Errors of aligning by the projective transformation
- Extraction of the vehicle running slowly
- Extraction of the shadow region
- Extraction of the high building

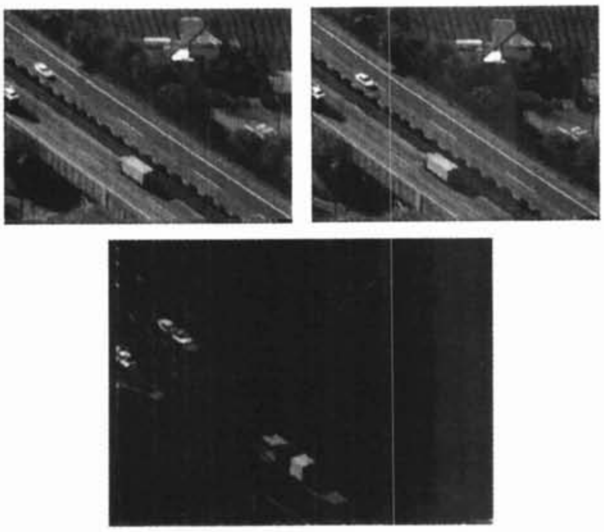

Fig.3 Differential image (Upper image is the input image. (Left is It-Δt. Right is It.) Lower image is the differential image between It-Δt and It.)

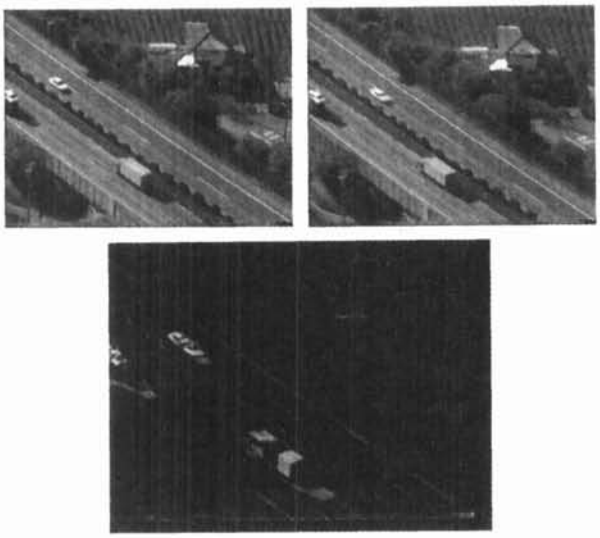

Fig.4 Differential image (Upper image is the input image. (Left is It. Right is It+Δt.) Lower image is the differential image between It and It+Δt.)

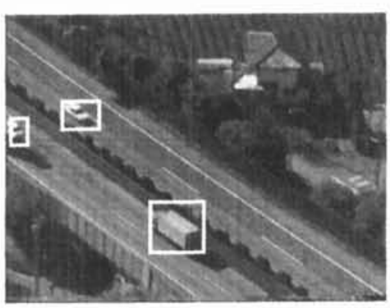

Fig.5 Extraction result of the vehicles

4 Conclusions

We extracted moving vehicles correctly from the time sequence of aerial images. The proposed method aligns the aerial images only from the time-sequence of them, and hence, this method does not need the background images or the map information. Therefore, this method enables the extraction of the vehicles on the road with one time search and can be applied to the real time analysis of the road traffic information on the wide area.

The future works are described as follows:

- Tracking moving vehicles
- Improvement in accuracy of aligning the coordinates of the images
- Improvement in accuracy of extracting moving vehicles
- Extraction of regions hidden by some buildings
- Verification by many images

References

[1] T.Matsuyama, and M.Nagao, "Structural Analysis of Aerial Photographs," Information Processing Society of Japan, Vol.21, No.5, pp.468-480, 1980.

[2] Y.Chang, M.Takagi, "The Geometric Correction about Meteorological Satellite NOAA Imagery," The Transactions of the Institute of Electronics, Information and Communication Engineers, Technical Report of IEICE, Vol.IE92-80, 1992-11.

[3] Y.Ogawa, S.Kakumoto, and K.Iwamura, "Extracting Regional Features from Aerial Images Based on 3-D Map Matching," The Transactions of the Institute of Electronics, Information and Communication Engineers, Vol.J81-D-II, No.6, pp.1242-1250, 1998.

[4] S.Kamijo, Y.Matsushita, K.Ikeuchi, and M.SAKAUCHI, "Occlusion Robust Tracking Utilizing Spatio-Temporal Markov Random Field Model," The Transactions of the Institute of Electronics, Information and Communication Engineers, Vol.J83-D-II, No.12, pp.2597-2609, 2000.

[5] T.Fuse, H.Yoda, and E.Shimizu, "Auto-Recognition of Vehicle Maneuvers Based on Spatio-Temporal Clustering Approach," The Infrastructure Planning Committee, Japan Society of Civil Engineers, Vol.24, No.2, pp.715-718, 2001.

[6] E.Shimizu, T.Shidoji, and T.Fuse, "A New Technique for Vehicle Tracking on the Assumption of Stratospheric Platform," Journal of the Japan Society of Photogrammetry and Remote Sensing, Vol.38, No.4, pp.53-58, 1999.

[7] T.Toriu, T.Endoh, and A.Tamekiyo, "An Image Interpolation Method Based on Perspective Transformation," The Transactions of the Institute of Electronics, Information and Communication Engineers, Technical Report of IEICE, Vol.IE94-14, 1994-05.

[8] F.Saitoh, "Robust Image Matching for Occlusion Using Vote by Block Matching," The Transactions of the Institute of Electronics, Information and Communication Engineers, Vol.J84-D-II, No.10, pp.2270-2279, 2001.