

# A New Approach to High-accuracy Road Orthophoto Mapping Based on Wavelet Transform

**Ming YANG**

*Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China*

**Chunxiang WANG\***

*Research Institute of Robotics, Shanghai Jiao Tong University, No. 800 Dong Chuan Road  
Shanghai 200240, China*

**Fang CHEN, Bing WANG**

*Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China*

**Hao LI**

*The Robotics Laboratory, Mines Paris (ParisTech) / INRIA-IMARA Team, B.P.105,78153 Le Chesnay, France*

Received 10 July 2011; accepted 25 November 2011

## Abstract

Existing orthophoto map based on satellite photography and aerial photography is not precise enough for road marking. This paper proposes a new approach to high-accuracy orthophoto mapping. The approach uses inverse perspective transformation to process the image information and generates the orthophoto fragment. The offline interpolation algorithm is used to process the location information. It processes the dead reckoning and the EKF location information, and uses the result to transform the fragments to the global coordinate system. At last it uses wavelet transform to divide the image to two frequency bands and uses weighted median algorithm to deal with them separately. The result of experiment shows that the map produced with this method has high accuracy.

**Keywords:** orthophoto map; offline interpolation; image fusion; weighted median; wavelet transform

## 1. Introduction

Nowadays orthophoto map plays an important role in navigation, exploration and city planning. Meanwhile orthophoto map of urban environment is becoming a new hot demand. In urban environment, the environment is more complex and more informative. The road environment is the core of urban environment for its complexity, in which there are variety of static items and dynamic items, including road lanes, road signs, sidewalk, pedestrian and vehicles. The static

items are important for vehicles driving, navigation and driving safety<sup>[1][2]</sup>.

Orthophoto map produced by satellite and aerial photography such as google map cannot get high-resolution image information of the road, which make it impossible to extract and identify the road signs. Therefore the image map data vendors now use artificial way for road marking. It decreases the efficiency of map refresh.

In this paper, the authors use road-vehicle system to produce the road orthophoto map. Compared with ordinary orthophoto map, it has obvious advantage in image accuracy.

---

This work was financially supported by the National Natural Science Foundation of China (61174178) and the Shanghai Science and Technology Commission through the Expo Science and Technology Program(10dz0581100)

\* Corresponding Author: [wangcx@sjtu.edu.cn](mailto:wangcx@sjtu.edu.cn)

## 2. Overview of the mapping method

This paper presents a new mapping method: using road-vehicle system with camera of location system.

The method uses road-vehicle camera to capture road image information and uses global position system (GPS), inertial measurement unit (IMU) to get real-time location information, and then generates a serial of road orthophoto fragment with the image information and the camera parameters, at last mosaics the fragment according to their position information and fuses the huge amount of fragment to a whole map. The System implementation framework is shown in Figure 1.

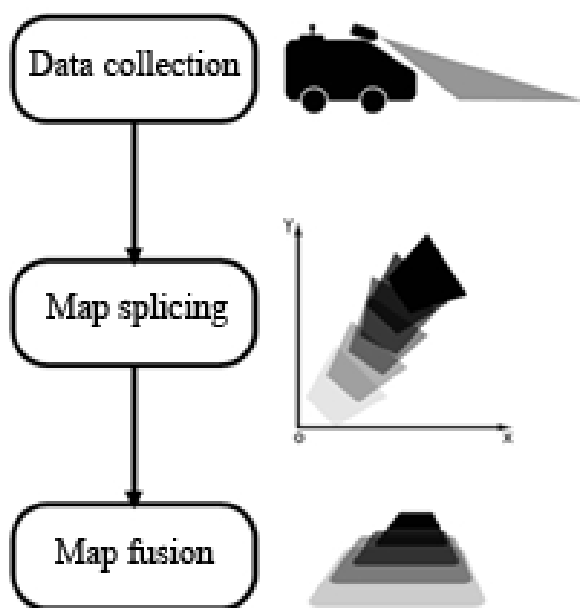


Fig. 1. System framework

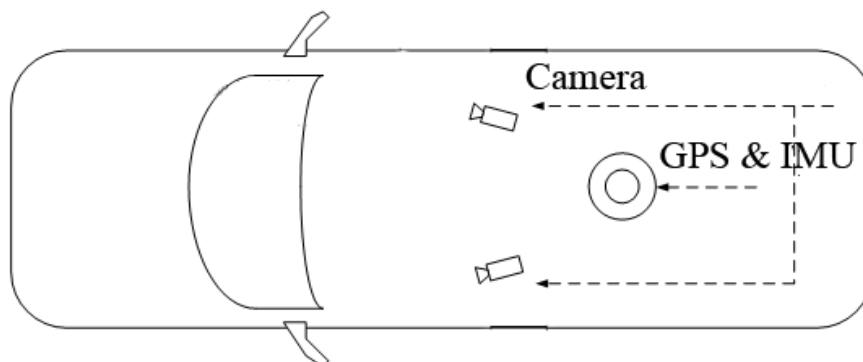


Fig.2. Sensor configuration of the data collection vehicle.

### 2.1. Overview of data collection

In the data collection system, a special data collecting vehicle is necessary. There are two slightly oblique camera equipped on the roof of vehicle whose view intersects in front of the vehicle. The data collecting vehicle is also equipped with GPS and IMU. The installation is shown in Figure 2. After the installation, the camera outside parameters should be calibrated. In the collection process the cameras cannot be moved and the collected image information and position information should be synchronized.

### 2.2. Offline interpolation algorithm

The map generation method has high requirements in the precision of data collecting vehicle's location.

The dead reckoning information is smooth enough but has accumulated error<sup>[3]</sup>. The fusion information from GPS and IMU with expand Kalman filter (EKF)<sup>[4][5]</sup> is more precise but it's not smooth enough, which causes the chasm of the result.

Therefore we use a new method called offline interpolation algorithm. It samples the EKF position information and plus a serial of dead reckoning information into each two adjacent EKF point. At last adjust the dead reckoning point and make them continuous. Suppose the two adjacent EKF point is  $(x_1, y_1, q_1)$  and  $(x_2, y_2, q_2)$ , the dead reckoning points in them is  $(u_1, v_1, t_1) \dots (u_m, v_m, t_m)$ , the result of the algorithm is:

$$\begin{cases} Rx_i = \frac{x_2 - x_1 - u_n + u_i}{n} * i + u_i - (x_1 - u_1) \\ Ry_i = \frac{y_2 - y_1 - v_n + v_i}{n} * i + v_i - (y_1 - v_1) \\ Rq_i = \frac{q_2 - q_1 - t_n + t_i}{n} * i + t_i - (q_1 - t_1) \end{cases} \quad (1)$$

Algorithm flow shown if Figure 3.

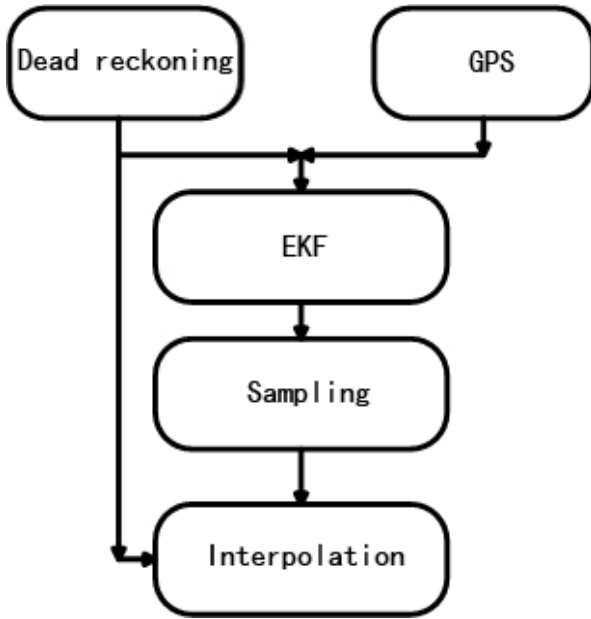


Fig.3. Flow of offline interpolation algorithm.

### 2.3. Map splicing

Upon completion of image acquisition and precise positioning of the vehicle, the authors use the inverse

perspective transform<sup>[6]</sup> method to transform camera image to road orthophoto fragment frame by frame, the result is shown in figure 4.

Then the vehicle location information is used to move, whirl and scale the image, at last projecting them to the global coordinate system.

Actually the authors use the methods reversely, choosing the point in global coordinate and transform it to vehicle coordinate with the position information, at last transform it to the image coordinate using inverse perspective transform to find original pixels.

Suppose the position of one point in the global coordinate system is  $(x, y)$ , the camera parameters of the inverse perspective transform is  $m_1$  to  $m_8$ , the position of the data collecting vehicle when the point collected is  $(u, v, q)$ , the position of the point in the vehicle coordinate system is:

$$\begin{cases} p = (x - u) * \cos q - (y - v) * \sin q \\ q = (x - u) * \sin q + (y - v) * \cos q \end{cases} \quad (2)$$

The position of the pixels in the image is

$$\begin{cases} r_x = \frac{\frac{m_3 - p}{m_8 * p - m_2} - \frac{m_6 - q}{m_8 * p - m_5}}{\frac{m_7 * p - m_1}{m_8 * p - m_2} - \frac{m_7 * q - m_4}{m_8 * p - m_5}} \\ r_y = \frac{\frac{m_3 - p}{m_7 * p - m_1} - \frac{m_6 - q}{m_8 * p - m_5}}{\frac{m_8 * p - m_2}{m_8 * p - m_5} - \frac{m_8 * q - m_5}{m_7 * p - m_4}} \end{cases} \quad (3)$$

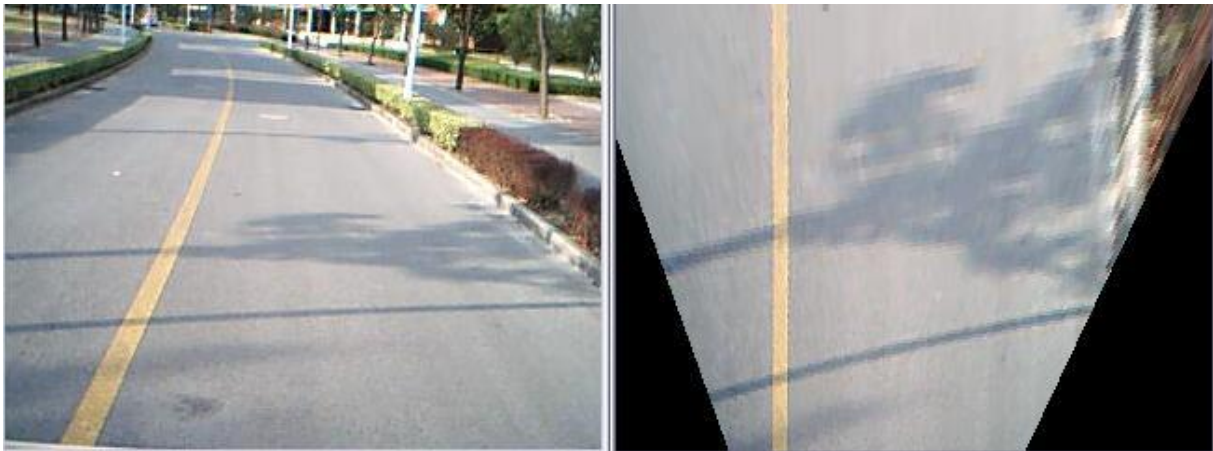


Fig.4. This is the result of inverse perspective transform.

### 3. Map fusion based on weighted median and wavelet transform

In the map generation system discussed in the paper, the work of map fusion is to integrate a large number of overlapping orthophoto fragment to a whole map.

The random location error of the fragments, the uncertain of the association between the fragments and the lack of an absolute reliable reference make the global alignment very difficult. The decrease of the local error will not cause the decrease of the global error. In many cases even increase it.

Therefore the problem of map fusion can be summarized as the information selection among huge amount of hard-aligning and overlapping image.

#### 3.1. Weighted median algorithm

In this paper, we use the weighted median algorithm as the base of map fusion. The algorithm sorts the points at same position according to their gray, and selects the variable related to the distance to the camera as the

weigh, at last selects the point closest to the half of the sum of the weight.

The serial of the sorted points is  $p_1$  to  $p_n$  and their weight is  $w_1$  to  $w_n$ . The point selected at last is the one minimized the value of  $R_i$

$$R_i = \left| \frac{w_1 + \dots + w_n}{2} - (w_1 + \dots + w_i) \right| \quad (4)$$

The median algorithm overcomes the fuzzy of mean algorithm. For example there is a yellow point at position A in the real world and the camera collected the point 10 times. From the result of the error only 7 points collected were located to A and another 3 were located to B or C. The result of mean algorithm of A is not the same as it should be but the result of median algorithm is correct. However if the error increase and less than 5 point were located at A, the yellow will be cleared.

Therefore the median algorithm can select information more correctly if the error is limited.

In order to limit the error we use weighted median algorithm and include the distance between the point and the camera as the main parameter of the weight.



Fig.5. result of weighted median algorithm and the contrast to the result of mean algorithm and weighted mean algorithm.

The point closer to the camera has higher resolution and less rotation error. Its weight should be higher.

After the experiments we choose the weight:

$$w = \begin{cases} 10 * \frac{a^2}{d^2} & (d > a) \\ 10 & (d \leq a) \end{cases} \quad (5)$$

$d$  represents the distance and  $a$  is a parameter.

The median algorithm need the majority of the error of the pixels be limited and weighted median algorithm relaxes the requirements that the error of the pixels close to the camera should be limited. In order to meet the requirements the authors limit the vision of the camera and make the proximal of the vision as close as possible to the camera.

The result of weighted median map fusion is shown in figure 5. Figure 5-a, 5-d, 5-g shows the result of mean algorithm, 5-b, 5-e, 5-h shows the result of weighted mean and 5-c, 5-f, 5-i shows the result of weighted median. Obviously 5-b, 5-e and 5-h is a little better than 5-a, 5-d, 5-g. 5-c, 5-f, 5-i is much better than the others.

### 3.2. Wavelet transform introduction

In the processing of map fusion, lack of road lane and mark is much important, so some detail and grain information should be well kept in order to make the result more realistic and easy-recognized. In existing mean fusion, HSI image fusion<sup>[7]</sup>, Laplace image fusion<sup>[8]</sup> and wavelet transform fusion, wavelet transform distinguishes the structure and grain information of the image and treats them separately in order to keep more grain information.

Image fusion base on wavelet transform is the hot issue in today's image fusion area. It uses multi-resolution analysis and mallat fast algorithm, using the directional character and non-redundancy of wavelet decomposition to treat images on different feature domain, picking up structure and grain information effectively. Image with an  $L$  level wavelet transform, will have  $(3L + 1)$  layers of sub-band, including the low-frequency sub-band C and high-frequency sub-band H, V, D. The low-frequency sub-band mainly consists the structure information and the high-frequency sub-band consists the detail and grain information<sup>[9]</sup>. Algorithm flow is shown in Figure 6.

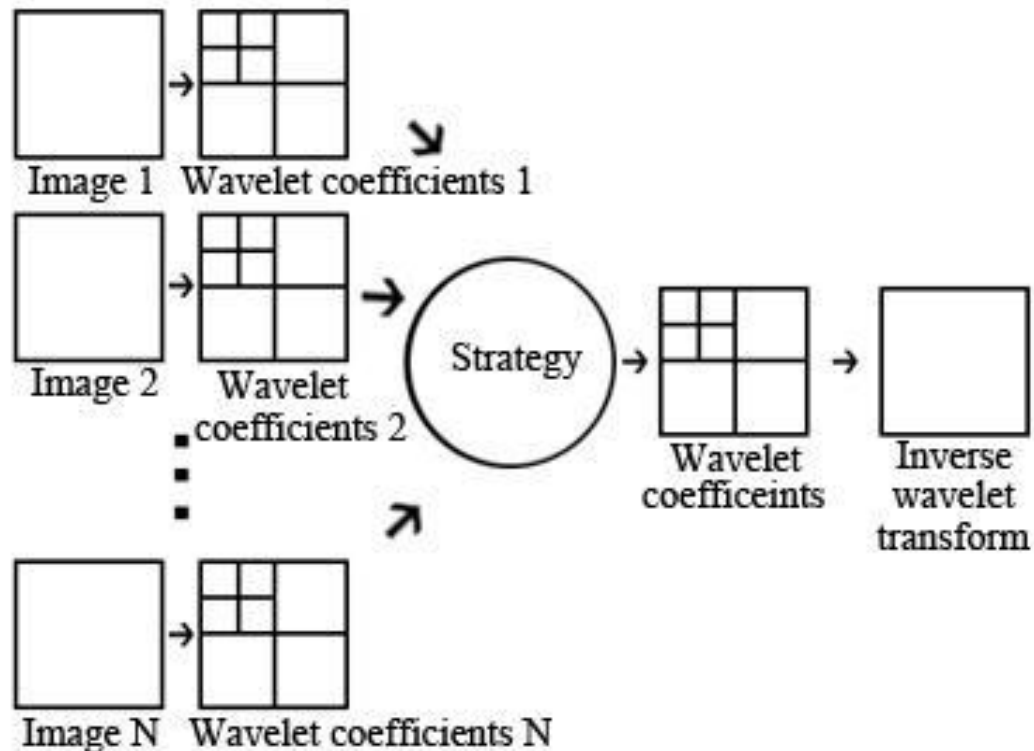


Fig.6. flow of wavelet transform

### 3.3. Wavelet transform strategy

The wavelet transform strategy of low-frequency and high-frequency sub-band is the core<sup>[10]</sup> of wavelet transform. Different fusion strategy brings different result.

Existing strategy such as local gradient average strategy, local variance strategy and local energy strategy fuses the map in different sub-band according to the statistics information, which is not as strong as the weight information according to the pixels distance.

According to the result in chapter 3.2, weighted median image fusion is effective, so the authors choose it as the low-frequency fusion strategy.

The high-frequency fusion strategy is important for grain and detail information. After the experience the authors choose a low-pass-filter plus the weighted median algorithm.

The result of wavelet transform is shown in figure 7. Figure 7-a and 7-c show the result of weighted median algorithm and figure 7-b, 7-d show the map at the same position with wavelet transform using the weighted median strategy. The yellow lane in 7-a and 7-c is blur because of the position error. But in figure 7-b and 7-d, the yellow-lane is well filled. Meanwhile the other part of the map such as the grey road is not affected by the wavelet transform.

## 4. Experiment result

Figure 8 shows the wide range experiment result in Shanghai Jiao Tong University. Figure 8-a is the thumbnail of the whole map and 8-b is the same position in Google Map. Figure 8-c shows the enlarged map and 8-d is the same position in Google Map but it can't be enlarged anymore. Actually the real resolution of the result is 0.0187 meters per pixel. It's narrowed for the limit of paper space.

Figure 8-e and 8-f show some detail in the map. 8-e is a road which has been collected twice from different direction. It shows that the way of map fusion can fuse the map with different illumination angle. 8-f is a road sign, it's clear enough for road marking.

There are errors exist because of the bumpy in road and the camera error. Figure 8-g shows the distorted result.

Summary the mapping method has good performance in the wide range experiments.

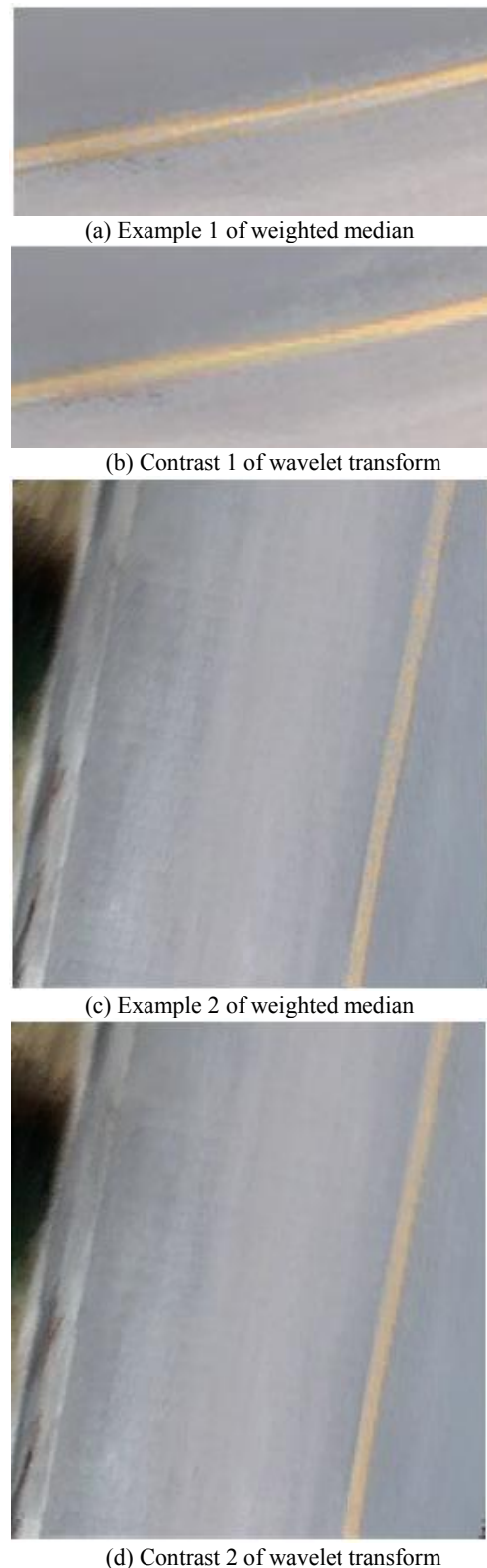
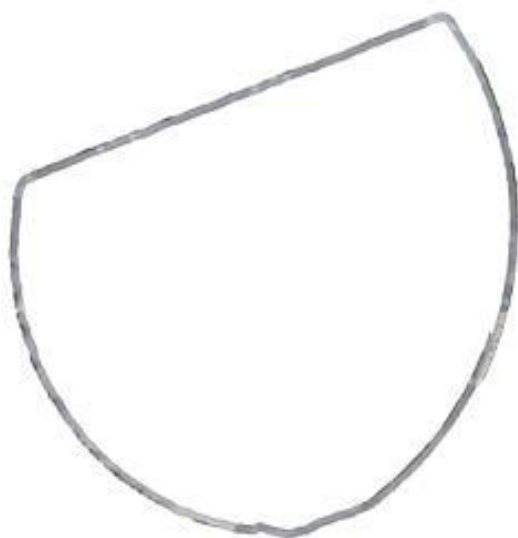


Fig.7. result of wavelet transform map fusion and the contrast of weighted median map fusion.

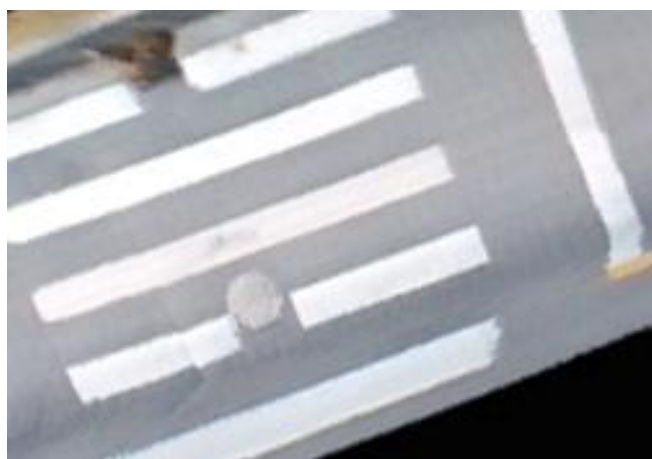




(a) The whole map



(b) Contrast in Google map



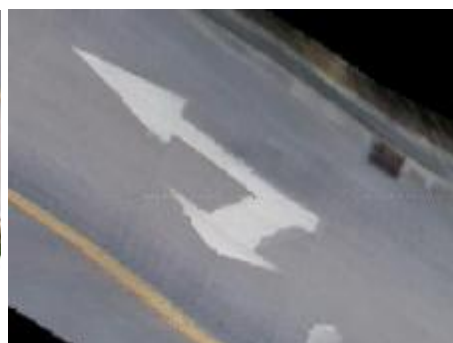
(c) Example of the detail



(d) Contrast of the same position



(e) Example of the road collected twice



(f) Example of the road sign



(g) Example of the error

Fig.8. experiment result in SJTU

## 5. Conclusion and outlook

Existing orthophoto map based on satellite photography and aerial photography is not precise enough for road marking. This paper presents the method of using camera and vehicle location information for road orthophoto mapping. It uses offline interpolation algorithm for location and uses inverse perspective transform for map splicing. In the part of map fusion, the method use wavelet transform to distinguish the low-frequency and high-frequency sub-band, and then use weighted median algorithm to treat them separately. The result of experiment shows that the map produced with this method has high accuracy.

In the future, the authors will use laser radar to detect obstacle and pedestrians in order to remove them from the map. On the other hand the moving shadow detection<sup>[11]</sup> can be used to improve image quality.

## 6. References

1. Wuhong Wang, Fuguo Hou, Huachun Tan, Heiner Bubb, A framework for function allocation in intelligent driver interface design for comfort and safety, *International Journal of Computational Intelligence Systems*, 2010 3(5),531-541.
2. Wuhong Wang, Xiaobei Jiang, Shuangcheng Xia, Qi Cao, Incident tree model and incident tree analysis method for quantified risk assessment: an in-depth accident study in traffic operation, *Safety Science*, 2010 48(10),1248-1262.
3. Charles Vlcek and Patricia McLain, GPS/Dead Reckoning for Vehicle Tracking in the "Urban Canyon" Environment, in *IEEE-IEE Vehicle Navigation & Information System Conference (Ottawa-VNIS'93)*, A34.
4. R. E. Kalman, A New Approach to Linear Filtering and Prediction Problems, in *Transactions of the ASME-Journal of Basic Engineering*, 82 (Series D), 35-45.
5. Edward J. Krakiwsky and Clyde B. Harris, A Kalman Filter for Integrating Dead Reckoning and, Map Matching and GPS Positioning, CH2675-7/88/0000-0039 (1988), 39-46.
6. Li Hao and Yang Ming, A Nonlinear Inverse Perspective Transform Based Method for Camera Distortion Calibration, in *Journal of Shanghai Jiao Tong University* (vol. 42, No. 10, Oct 2008), 1736-1739.
7. Vivek Maik, Jeongho Shin and Joonki Paik, Pattern selective image fusion for multi-focus image reconstruction, *CAIP 2005, LNCS 3691 (2005)*, 677-the thumbnail 684.
8. Jorge Nuñez and Xavier Otazu, Multiresolution-Based Image Fusion with Additive Wavelet Decomposition, *IEEE transaction on geosciences and remote sensing*, (VOL. 37, NO. 3, MAY 1999), 1204-1211.
9. H.Li, B. S.Manjunath and S. K.Mitra, Multisensor Image Fusion Using the Wavelet Transform, *Graphical Models and Image Processing (Volume 57, Issue 3, May 1995)*, 235-245.
10. Gonzalo Pajares and Jesús Manuel de la Cruz, A wavelet-based image fusion tutorial, *Pattern Recognition (Volume 37, Issue 9, September 2004)*, 1855-1872.
11. Sohail Nadimi and Bir Bhanu, Moving Shadow Detection Using a Physics-based Approach, in *1051-465U02( 2002)*, 701-704