

# A Platform for Mobile Image Recognition and Mobile Mapping in Local Based Services

Jinsuk Kang and Chung Jae Young

**Abstract**—We have developed a prototype image recognition system capable of identifying landmarks (typically, city buildings and structures) from photographs captured on camera phones. The image recognition process runs server-side, as a web service accessible over GPRS from a web enabled phone or Smartphone. Photographs of landmarks are pre-loaded into a central database, which is then queried by users submitting images from their camera phones. The image recognition algorithm is robust to variations in both illumination and point of view. In this research, the effective method was suggested for the transmission of the geographic information acquired by camera attached to a Smartphone such as position data, attitude data, and image data in the wireless internet environment in real-time.

**Index Terms**—Local based services, mobile mapping, mobile device, mobile image processing.

## I. INTRODUCTION

The progress of computer science and information technology, including the Internet, has rapidly accelerated the spread of the personal computer, and some mobile devices such as handheld PCs (HPCs), PDAs (Personal Digital Assistants) have become widespread aspect of a continuous growth of the Internet Technology business [1]-[3]. This allows people to use various services without limits of time and space. Machines like mobile phones and Smartphone with electronic control systems work by using small-sized OS (Windows CE, Embedded Linux, pSOS, etc) in ROM, not like a hard disc of normal PC, RAM, etc. are called embedded systems [4].

Amid this, the term LBS (Location Based Services) or mobile GIS (Geographic Information System) is emerging. LBS is a wireless information service that uses the location information of a mobile user. The convergence of multiple technologies including GIS, internet, wireless communication, location determination, and portable device has given rise to the LBS. To provide various information based on the location of mobile user, LBS requires GIS functionality for mapping user's current location and user's interest area through the wireless internet. This location information is acquired by the location determination technology. The geographic information created with the location information is represented on the screen of a portable device [5], [6].

In this paper, we propose to use man's geographic recognition in addition to a GPS for mapping objects. With

the help of man's ability of recognizing placement of objects and geographical relations between objects, location information given by a GPS is effectively compensated and utilized. This method improves quality of a resulting map.

## II. IMAGE CAPTURE OF SMARTPHONE

### A. System Concept

Therefore, we propose a solution, which is completely dedicated to address the objects (landmarks, city building and structures) extraction by optimising all the steps of the sensor integration, including hardware, data acquisition, and algorithmic processing. Mobile Mapping Systems have been using image sequences for a long time since it is an essential part of the concept. However, progress toward the automation of the image sequence processing has been slow for two reasons. First, economy; the actual feature extraction represents approximately less than 20% of the overall cost, and therefore, the financial motivation is weak. Second is the varying image scale, which makes the object recognition task quite difficult in the feature-rich object space.

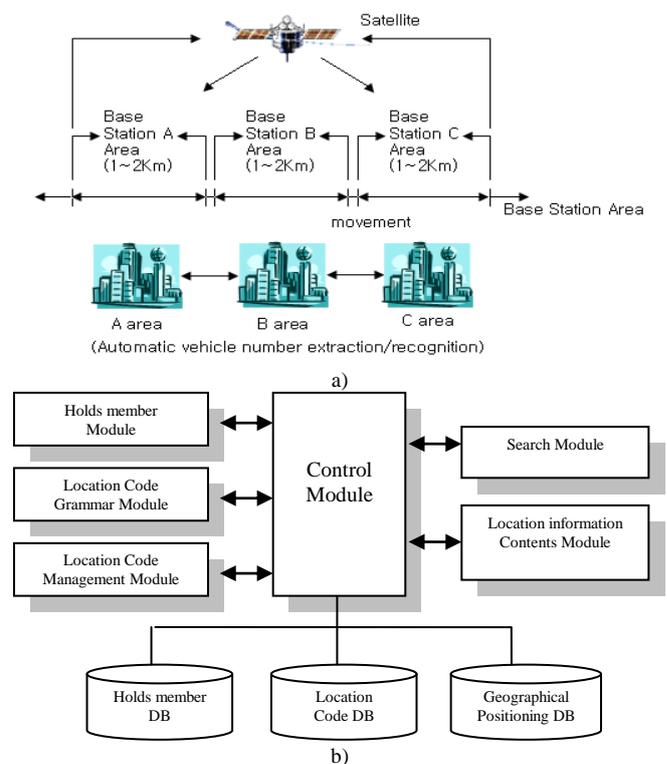


Fig. 1. Generic model of the dedicated mapping system.

Typically, earlier experimental methods have worked in post-processing mode and required non-negligible user interactions. The objects in our approach, however,

Manuscript received August 13, 2014; revised October 31, 2014.

The authors are with Information and Management Research Consortium (IMRC), South of Korea (tel.: +82-2-2183-1167; fax: 82-2-2183-1168; e-mail: jskang01@imrc.or.kr, c603jy@chol.com).

represents a much more forgiving task from the automated image processing view because due to the special sensor arrangement the image scale changes are limited and the image contents are rather well defined. Traditional Mobile Mapping Systems (MMS) work with forward- or side-looking cameras, while our system uses a down- looking camera, with an image sensor plane almost parallel to the road surface.

This way the image scale changes very slightly and there is an almost constant scale along the vehicle trajectory. To compensate for the smaller footprint, the image covers the road area of about that of a vehicle; however, the image capture rate should be increased. The object contents of the images are rather simple and predictable, such as the line marks, the primary interest to us, surface texture variations, cracks, potholes, skid marks, etc. Therefore, extracting features from a well-defined set of possible objects from an almost constant scale imagery constitutes a much more formidable task compared to the generic MMS approach. Consequently, a large number of proven computer vision methods can be successfully applied for object extraction. Figure. 1 shows the generic model of the dedicated mapping system.

**B. Coordinate Measurement of Digital Image**

*1) Coordinate system of digital image*

Fig. 2 shows is the target coordinates measured from digital image are related to pixel coordinate system, the pixel edge at the uppermost left hand side is the origin and has the horizontal line parallel to x-axis and vertical line parallel to y-axis. Their origin must be moved to the location corresponding to position of center, in other words, the main point. The transformation from pixel to image coordinates is determined by going through the image coordinate transformation.

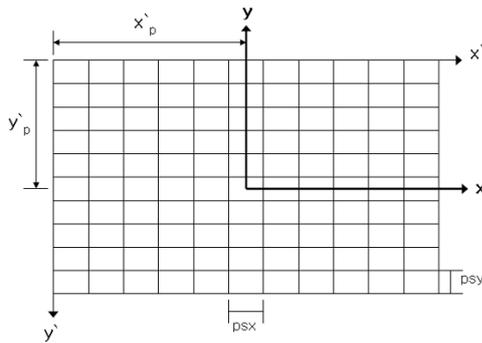


Fig. 2. Pixel coordinate system and photo coordinate system.

Here:

y: image coordinates,

x', y': pixel coordinates

x'\_p, y'\_p: position of main point at pixel

psx, psy: pixel interval of x and y

*2) Coordinate measurement of digital image*

The center of target images can be determined by calculating the center of gravity and first moment while center folding technique is one of approaches to determine the position of sub-pixel toward circular or elliptical target images. [7]. The initial position of targets acquired by bi-level images are placed within a small rectangular around target

and a suitable threshold value process is used in order to lower the background from the target. In this window, the target image is calculated using equation (1).

$$x_c = \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^m j \cdot g_{ij}, \quad y_c = \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^m j \cdot g_{ij} \tag{1}$$

$$M = \sum_{i=1}^n \sum_{j=1}^m g_{ij}$$

$g_{ij}$  is the shading intensity of pixel located at  $i^{\text{th}}$  horizontal line and  $j^{\text{th}}$  vertical line and has the value of 0 or 1. But because inconsistent backgrounds form unbalance of target image, they can create derivation of pixel as the measured results, size of window, position and deflection point get changed [8]. Fig. 3 shows vertical section diagram of target image in a round Smartphone camera of 5mm in diameter. Accordingly, higher intensities of the target were given as greater weight in the calculation in order to reduce the effect from the background.

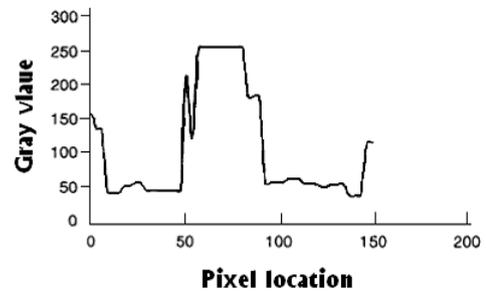


Fig. 3. Vertical section diagram of a target image.

As mentioned previously, because the inaccuracy in determining position of center using equation (2) has been proven, the accurate position of center can be determined by an improved formula in which the intensity weight, in other words the brightness, is considered can be determined.

$$x = \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^m j \cdot g_{ij} \cdot w_{ij}, \quad y = \frac{1}{M} \sum_{i=1}^n \sum_{j=1}^m j \cdot g_{ij} \cdot w_{ij}, \tag{2}$$

$$M = \sum_{i=1}^n \sum_{j=1}^m g_{ij} \cdot w_{ij}$$

$g_{ij}$  is the shading intensity value of each pixel and  $w_{ij} = g_{ij}$ .

*3) Two dimensional coordinate transformation*

Affine conversion (Fig. 4) is used for rotation of coordinate axis, non-rectangular rotation and correction of linearly polarized wave followed by the change of axis scale like formula.

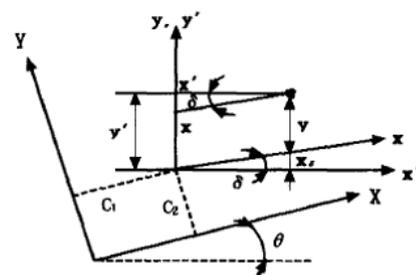


Fig. 4. Affine conversion.

$$X = ax + by + c, Y = dx + ey + f \quad (3)$$

The above formula 3 is derived from following formulae

$$x' = x \cos \delta = x, y' = y + x \sin \delta = y + x \delta \quad (4)$$

Formula 4 can be expressed as determinants of formula 5.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \delta & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = M_\delta \begin{bmatrix} x \\ y \end{bmatrix} \quad (5)$$

And can be expressed as formula 6, formula 7, formula 8 by considering axis conversion and parallel movement.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \delta & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_x x \\ \lambda_y y \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (6)$$

$$X = \lambda_x (\cos \theta - \delta \sin \theta) x - (\lambda_y \sin \theta) y + c_1 \quad (7)$$

$$Y = \lambda_y (\sin \theta - \delta \cos \theta) x - (\lambda_y \sin \theta) y + c_2$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ f \end{bmatrix} \quad (8)$$

The formula 8 can be expressed simply as formula 1. It is generally used in photo measurements and has six parameters. Therefore, it requires at least three base control points.

### III. MAPPING TO CONTEXT MANAGEMENT SYSTEMS

#### A. Mobile Embedded Mapping

Our architecture for the mobile mapping in the embedded environment is composed of GPS data handling module, mapping module, and map data which are interconnected between mobile device and GPS receiver (Fig. 5).

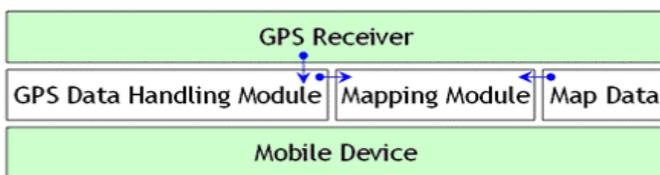


Fig. 5. Architecture of mobile embedded mapping in the geocomputing platform.

GPS data handling module gets data from GPS receiver, and mapping module visualizes the GPS coordinate on a map using the map data embedded in the mobile device. As a stand alone system, the overall structure may be similar to that of major vendor's mobile GIS products such as ESRI ArcPad, MapInfo MapX Mobile, and Pocket Systems PocketGIS. This Mapping Module has an MVC (Model-View-Controller) architecture whose goal is to separate the application object (model) from the way it is represented to the user (view) from the way in which the user controls it (controller), thereby allowing flexibility and reusability. The role of model, view, and controller is addressed in Table I, and the relationship between model, view, and controller objects is shown in Fig. 6.

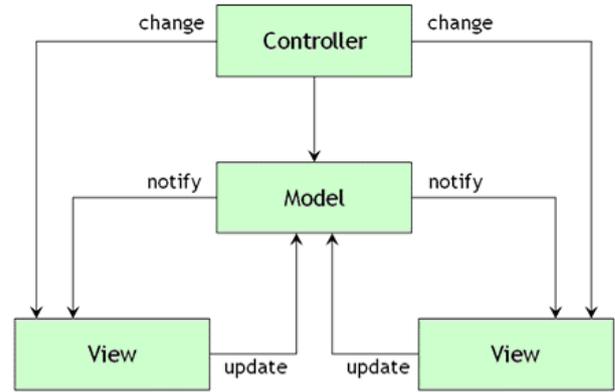


Fig. 6. Architecture of MVC (model-view-controller).

TABLE I: ROLE OF MODEL, VIEW, AND CONTROLLER IN GRAPHICS PROGRAMMING

Model	The model object knows about all the data that need to be displayed. It also knows about all the operations that can be applied to transform that object. However, it knows nothing whatever about the GUI, the manner in which the data are to be displayed, nor the GUI actions that are used to manipulate the data. The data are accessed and manipulated through methods that are independent of the GUI.
View	The view object refers to the model. It uses the query methods of the model to obtain data from the model and then displays the information.
Controller	The controller object knows about the physical means by which users manipulate data within the model. In a GUI for example, the controller object would receive mouse clicks or keyboard input which it would translate into the manipulator method which the model understands.

#### B. Performance of the Automated Image Sequence Processing

To assess the feasibility of automated line extraction with 3D positioning and consequently its real-time realization, a rich set of the potential image processing functions was developed in a Visual Embedded C++ programming environment. Fig. 7 shows the overall dataflow and processing steps, which will be illustrated in more detail later. In short, the real-time image processing is feasible due to a simple sensor geometry and the limited complexity of the imagery collected.

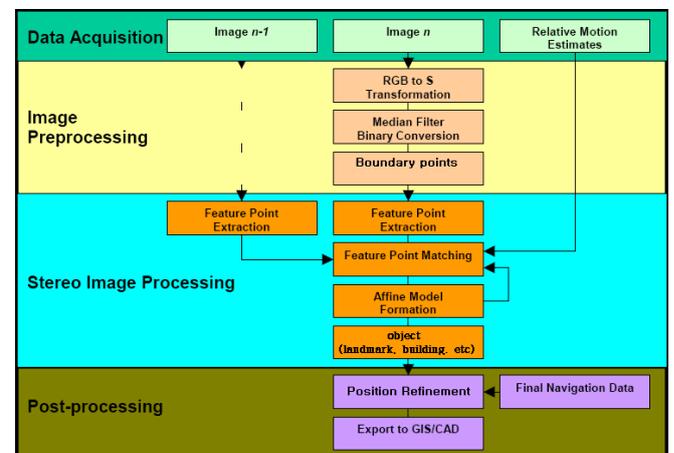


Fig. 7. Real-time image processing and post-processing workflow.

#### C. Positioning Performance

The important part of a still image taken by a Smartphone

camera is the algorithm to detect the objects. There are two general methods in use: one to detect objects using brightness information, and the other to identify characteristics by edge detection and Hough transformation [9]. Figure 8 shows the overly sensitive to the environment, with a lower recognition rate when there are noises around. The second uses vertical and horizontal components in the landmarks, with lowered recognition rate and longer processing time in case of damage or noise in plate edge, which is not proper for real time processing.

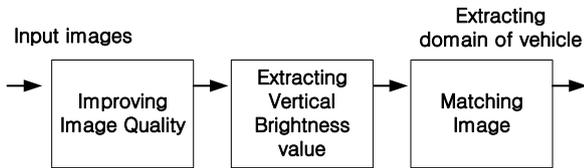


Fig. 8. A Block diagram of preprocess stage.

#### IV. EXPERIMENTAL SCENARIOS, RESULTS AND EVALUATION

If image measurements from GPS/INS are to be used, the result of exterior orientation must be actually applied in mapping and editing operations while photographing.

It should be noted that the verification process of aerial mapping result is the most crucial process at the GPS/INS based digital photogrammetry. To this end, this study has compared the mapping result by the ways of AT results, plotter, and orientation, which is from basis of the Analytical raw map produced by the existing AT result.

In this application, a mobile user can see the current location is pinpointed on the land parcel map according to the GPS coordinate and radius extent assigned. If the user clicks on a specific parcel, the information of the parcel such as parcel type, area, rent, selling, and register notes show in Fig. 9. The application of field collection can assist mobile field work through the real-time notification of current location, on-site input of attribute value, and the compatibility with desktop GIS package. The application of cadastral information service is useful for the purchasers/renters of house or land through the real-time notification of current and on-site inquiry of cadastral information.

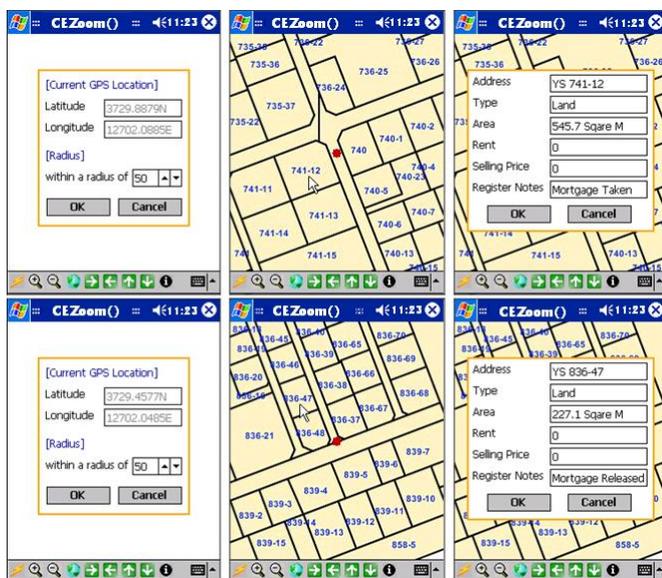


Fig. 9. Screenshot of cadastral information service.

If we examine Table II, the difference of about 76cm has been created with analytical & digital plotter by conventional AT for the horizontal error and error of 66cm in analytical plotter and 89cm in digital plotter have been created when AT results (conventional AT vs GPS/INS AT) have been set as the standard. Because the tolerable horizontal error while producing a 1/5,000 map by current public measurement operation standard is 3.5m (within 7mm according to map distance), it adequately satisfies the accuracy requirements without any problems in practical usage. In vertical error, the difference of about 45cm has been created with analytical & digital plotter (using conventional AT) and error of 28cm in analytical plotter and 33cm in digital plotter have been created when AT results (conventional AT vs GPS/INS AT) have been set as the standard.

And when descriptive horizontal and vertical error by analytical plotter and digital plotter are compared regardless of AT results and locating methods, The analytical plotter is showing better results than digital plotter.

Also, as we can find out in Table II, we can see that the horizontal error is shown as greater than vertical error. The reason is estimated as descriptive error being included to some extent because the horizontal position of check point at 1/20,000 aerial photograph is relatively more difficult to make out than the vertical point installed on the ground.

TABLE II: COMPARISON TABLE OF 1/5,000 MAPPING RESULTS

	Digital Plotter Tradition AT		GPS/INS AT	
	Analytical Plotter	Digital Plotter	Analytical Plotter	Digital Plotter
Horizontal error	Ref. Data	0.76	0.66	0.89
Vertical error	Ref. Data	0.45	0.28	0.33

#### V. CONCLUSION

This paper has discussed the design and implementation of a geocomputing platform for LBS and modular mobile mapping. This geocomputing platform targets LBS application development, and focuses on modular mobile mapping as a core element of LBS. Modular mobile mapping which support LBS consists of the functionalities of location information processing, mobile embedded mapping, and mobile internet mapping which play a unique role in the LBS application development. Modular mobile mapping becomes not only a core element of LBS, but also a useful application for mobile GIS. In this paper, we proposed to use human geographic recognition in addition to a GPS for mapping objects. Location information by a GPS is effectively compensated and supplemented by human geographic recognition of objects because these two have different characteristics. A mobile map-making learning system has been developed to utilize both types of location information. From the evaluation of the system, importance of geographic information based on human recognition, and usefulness of the proposed method to improve the quality of a map were recognized. Though the direct evaluation of educative effects

has not been conducted, the system is thought to be useful in mobile map-making learning. Making relationships between objects is common to organizing learning objects or knowledge objects. To support organizing knowledge in mobile learning, more consideration on relationships between objects will be needed. Adding collaborative learning feature to the system will also be our next step.

#### ACKNOWLEDGMENT

This research was supported by basic science research program through the NRF funded by MEST (No.2010-0020985). and also supported by the Technology Commercialization support program, Ministry of Agriculture, Food and Rural Affairs (No.113038-03-1-HD020).

#### REFERENCES

- [1] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Trans. on Neural Networks*, vol. 4, pp. 570-578, July 1993.
- [2] R. Li, K. P. Schwarz, M. A. Chapman, and M. Gravel, "Integrate GPS, INS, and CCD camera for rapid GIS data acquisition," *GIS World*, pp. 41-43, vol. 7, no. 4, 1994.
- [3] C. K. Toth and T. Schenk, "Feature-based matching for automatic image registration," *ITC Journal*, pp. 40-46, 2004.
- [4] J. Feldman and S. Czuckerberg, "Notebook system," US Patent No. 5553959, 1996.
- [5] A. John *et al.* (2001). Open eBook Publication Structure 1.0.1: Recommended Specification. Technical report. Open eBook Forum. [Online]. Available: <http://www.openebook.org/>
- [6] J. Chen and T. A. Clarke, "The automatic recognition, location and labelling of targets in digital," *International Archives of Photogrammetry and Remote Sensing*, Washinton, ISPRS Commission V, vol. xxix, part B5, pp. 686-693, 1992.
- [7] W. Benning., "Mobile mapping by a car driven survey system," in *Proc. International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation*, Vienna, Austria, 1998.
- [8] G. Ganci and M. R. Shortis, "The use of digital photogrammetry for large scale metrology," in *Proc. 5th South East Asian and 36th Australian Surveyors Congress*, vol. 1, pp. 247-261, Singapore: ASC. 2003.
- [9] J. Chen and T. A. Clarke, "The automatic recognition, location and labelling of targets in digital," *International Archives of Photogrammetry and Remote Sensing*, Washinton, ISPRS Commission V, vol. xxix, part B5, 1992.



**Jinsuk Kang** received his B.S. degree in information engineering from Cheju National University, Jeju, Korea, in 1999 and his M.S. and Ph.D. degrees in computer engineering from Cheju National University, Jeju, Korea, in 2001 and 2005, respectively. From 2006 to 2009, he was with the University of Incheon, Korea, as a research professor. From February 2009 to March 2010, he worked with Chungbuk National University, Korea, as a visiting professor. Since March 2010, he has been with the Jangwee Research Institute for National Defence, Ajou University, Suwon, Korea, where he is currently a research professor. His research interests include the areas of multimedia, computer vision, human-computer interaction, mobile computing and embedded system, etc.