# Supporting Indoor Navigation by Using Image Recognition Technique on Mobile Devices

Shan-Chih Lin I

Kuo-Feng Hwang Geeng-Neng You

Abstract—With the rapid progress of mobile technology, applications of mobile devices about geographic navigation support to help people move around have been extensively explored during the last decade. However, most of these applications are focused on outdoors. This study presents a novel way to support indoor navigation. We explore various approaches to support indoor navigation and employ an image recognition technology called Scale Invariant Feature Transform (SIFT) to construct a prototype system. The prototype can be easily installed on smartphones for supporting indoor navigation. We did a field test on a real occasion of guiding new meeting attenders around the designated floor of a building. Most users with relevant devices had utilized the system and delivered precious feedbacks. The general responses about the usage of the system are acceptable and interesting. But this kind of test and users' suggestions for improvements need to be further explored.

*Keywords*—mobile navigation, indoor navigation, scale-invariant feature transform (SIFT), image recognition

### I. Introduction

Recently, the fast-progressed technologies, combined with the popularity of smart phones and mobile devices, mobile applications were explored in various directions. One of the successful example is the navigation support tools for these devices' users moving around in some unfamiliar area by using the built-in GPS(Global Positioning System). However, most of these applications are for navigating outdoors. For indoor navigation support, different approaches are required due to the limitation of current GPS technologies. Most of these approaches utilized advanced technologies such as RFID (Radio IDentification)[1], Frequency Wi-Fi(Wireless Fidelity)[2], and QR-code(Quick Response Code).

Department of Multimedia Design. National Taichung University of Science and Technology, Taiwan

Shan-Chih Lin

Kuo-Feng Hwang

Gemg-Neng You

On the other hand, Image processing technologies had been developed over decades. It had been useful in applying to various fields such as face recognition, video motion detection and intelligent traffic systems. One of the successful image processing techniques called SIFT (Scale Invariant Feature Transform) is proposed by David Lowe in 1999. SIFT has been applied to many computer vision problems with its capability to detect and describe local features[3]. SIFT algorithm's robustness with respect to the correct matching of SIFT features has been widely explored and utilized[4]. Thus, our study will employ the SIFT technology to support mobile device users in indoor navigation.

# II. Related studies

# A. Navigation Support for Mobile Devices

In outdoor environment, GPS has widely been used in the navigation and positioning. In indoor environment, however, due to the weakening of the satellite signals by various barriers such as cement walls and other building materials, the applicability of GPS is limited. The accuracy of positioning satellite signals is another concern. Therefore, alternatives on the positioning of users in indoor environment had been explored. Several current approaches to develop positioning tool are reviewed in the following:

#### a. Wi-Fi

The Wi-Fi-based indoor positioning method exploits the strength and orientation of signals from Wi-Fi devices. The positioning algorithm uses signals from several (3 or more) nearby Wi-Fi APs (Access Point) for the computation. Theoretically, this mechanism could achieve high positioning accuracy. There are studies on its application such as the facilitation of guiding maps for Indoor Navigation[5]. However, due the limited to strength of the Wi-Fi signals. It might often be necessary to set up sufficient number of Wi-Fi APs in order to achieve pervasiveness. The arrangement of large number of APs is not only increasing the establishing cost but also imply the problem of wireless bandwidth congestion.

#### b. RFID

RFID technology applies reader and tag for developing positioning system. The general approach is by using the reader to acquire information in the preset Tags. The information in the tag will be retrieved and calculated to position the location of the user after comparison with the database. Some



### International Journal of Advances in Image Processing Techniques– IJIPT Volume 2 : Issue 2 [ISSN : 2372-3998]

### Publication Date : 30 October, 2015

studies have shown the potential of position-related applications[6]. When a user uses mobile device to sense RFID tags, the system will return data about the user's location while the RFID tags are placed in relevant places of a building. Increase the number of RFID tags could improve the positioning accuracy but also cause the rise of setup costs. The other disadvantage is the increase of the number of RFID tags means signal interference is inevitable in some occasions as this technology uses radio wave transmission.

### c. NFC

NFC (Near Field Communication, NFC) includes a set of standards for mobile devices to facilitate applications within very short distance. NFC mobile devices needs antennas for the transmission of radio communication with each other. A typical application is in the classrooms and offices of a school campus. Teachers and students can access and trade information at close range through pre-established NFC Tags [7].

### d. Mobile camera

Most mobile devices are equipped with cameras. Research on the application of these built-in cameras and image processing technology has yielded a versatile of useful tools [8]. With the matching of camera-captured input image and the pre-stored image database, a more flexible form of vision-based localization could be achieved. An example application is the application of this technology in the guidance tool of a museum[9].

The merit and weakness of each of these technologies has been documented in previous studies[10]. To sum up the reviewing, we provide perspectives in terms of equipment cost and accuracy in table1.

TABLE 1 COMPARISONS OF CURRENT INDOOR NAVIGATION

technology	Accuracy within area	Setup cost		
Wi-Fi	2-100m	High		
RFID	5cm-5m	High		
NFC	20cm	Low		
Mobile camera	1cm-5m	Low		

# **B.** Image Recognition

Current progress of technologies in the handling of multimedia data such as images, sound, video and other media had made digital equipment more effective than traditional textual, printed media to convey information. Furthermore, digital multimedia data are easier to manipulate and facilitate real-time interaction. However, the recognition of patterns in digitized data, such as sound, speeches and image in specific can be challenging. Through years of development, the pattern recognition techniques have evolved from statistical approaches into a versatile of paradigms such as machine learning, neuronetworks as well as genetic algorithm. The scope of application is even more versatile[11]. Typical applications could be found in word recognition applied to real-time language translation, voice recognition application on a mobile device for speech queries, image recognition applications in video surveillance systems and camera face recognition. This section focus on the aspect of image recognition, which is in the field of computer vision to extract image features to locate or identify the object.

The SIFT (Scale-invariant feature transform, SIFT) is currently one of the methods commonly used to solve the object recognition problem. It is a computer vision algorithms developed by Lowe in 1999[3]. It is based on the concept of the scale-invariant feature image pyramid and multiple feature points. SIFT algorithms has the advantage in its ability to robustly resist image manipulation such as rotation, scaling, distortion, lighting changes, masking and other effects. Our study employ the application of SIFT to find the unique characteristics of the image in order to overcome a variety of changes when users upload images in different situations.

The main processes of SIFT can be divided into four steps. The first step is to establish the scale space. The second step is to find the feature points. The third step is characterized by points to give direction. And the final(fourth) step is characterized by the establishment of sub-narrative. Details of the steps are illustrated in the following:

### Step 1: Create a layer of the pyramid (Pyramid layer)

These three steps establish the scale space. Firstly, shrink the length and width of the original image according to into a variety of geometric sizes, such as 1/2, 1/4, 1/8 etc., until a certain minimum size. It can be viewed as stacked-up like a pyramid of images, as shown in figure1. Thus it is called a layer of the pyramid.



Figure 1 Pyramid layer

Secondly, evolve into Gaussian pyramid (Gaussian Pyramid) for finding feature points. Use I (x, y) to depict the original image. Gaussian blur function is expressed as  $G(x, y, n\sigma)$ , where  $\sigma$  is the scale of size. The Gaussian Pyramid is a result of blurring caused by the filter of the original image in the following formula:

 $L(x, y, n\sigma) = I(x, y) \times G(x, y, n\sigma)$ 

Next, apply the above Gaussian blur procedures to each layer of patterns. Each pattern will be performed at least 4 times. The increase of its variables, such as



 $\sigma$ ,  $2\sigma$ ,  $3\sigma$ , ., will result in more blurred patterns, The establishment of a Gaussian pyramid is shown in Figure 2.



Figure 2 Gaussian pyramid (Gaussian Pyramid)

The final stage to establish the scale space is to obtain Difference of Gaussian images(DOG). The scale space is obtained from the results of subtracting the neighboring Gaussian pyramids  $L(x, y, n\sigma)$  to facilitate the next step to find points of difference, as shown in Figure 3.



Figure 3 Gauss difference image

Step 2: Find feature points

Ways to find SIFT feature points is by using the maximum value test. After obtaining the DoG images, feature points will be identified as local minima/maxima of the DoG images across scales. We not only test a single DOG layer(eight comparisons), but also DOG images of neighboring layers(nine comparisons each). And hence a total of 26 comparisons to determine the maximum or minimum points among all compared pixels, the extreme points were then marked, as shown in Figure 4.



Figure 4 maximum or minimum points test

After obtaining extreme point, to gain a more accurate characterization, we apply Hessian corner detection algorithm for the corner filtering of each extreme point. The Hessian corner filtering feature is for acquiring image corner, thereby extreme point edge of images can be deleted. The points left can be regarded as the key feature point.

Step 3: Acquire directional feature points

In order to the find the direction of feature points. We use the feature points as the center and calculate all angle and intensity within a radius of  $3 \times 1.5\sigma$ . With the accumulated eight directions., the highest will be the main direction, see figure 5.



Figure 5 SIFT direction detection

To achieve the invariance of image features, the characteristic center point of the image is rotated with an angle  $\theta$  according to the determination of the main direction. The next step is to establish the characteristics of the information.

Step 4: Establish feature descriptor

To establish the feature descriptor of the image, the surroundings of the feature point is decomposed into  $4 \times 4$  blocks, with feature point as the center. And for each block, the accumulated strength of each of the eight directions for each block is calculated. The final result is the  $4 \times 4 \times 8 = 128$  feature values for each feature point as shown in figure 6.

惫	寮	衆	☀
衆	≽	楽	$\mathbf{k}$
≽	举	×	₩
≽	*	×	*

Figure 6 SIFT 128 sub-feature narrative generation

Many researches on SIFT-related applications exploit the dynamic learning capability of the feature extraction in its algorithm. In a typical research of applications, the SIFT image processing technology is used with dynamic timing correction for motion analysis, i.e. to find out the trajectory of an object in continuous images[12]. This approach is also adopted in our study.

# II. System prototyping

## **System Architecture**

This study aims at applying SIFT image recognition for the support on navigation with mobile devices. We believe it is suitable to take advantage of the scale



### International Journal of Advances in Image Processing Techniques– IJIPT Volume 2 : Issue 2 [ISSN : 2372-3998]

invariant feature to cope with the scaling and rotation effects resulting from the movement of users. Following the exploration and feasibility test of SIFT algorithm, our prototype is designed to have two subsystems, namely, the administrators system and the mobile user system, a brief diagram of the architecture of the prototype is shown in figure 7.





# A. System Design

The detailed description of the design of the two parts of our indoor navigation system is illustrated in the next two subsections.

### 1 Administrator System Design

Our indoor navigation prototype provide functions for the administrating of the system. It has the image capturing feature(camera mode) for building the initial data base for mobile device users. It has editing facilities for managing the image database to allow administrators to quickly and easily update the required information. The diagram for depicting the architecture of administrator interface is in shown in figure 8.



Figure 8 Architecture of administrator system interface

The description of the successive steps of the administrator interface is show in figure 9.



Figure 9 Administrator System Interface

After entering the system, an adding data page is shown in 9(b), with features for adding image data. By obtaining photos through camera, administrators could enter the two text input field PictureName and PicContent with name and a simple specification of the image content, and appoint the picture to its indoor location map's coordinate points, and finally click the Upload button to add data to the database. To check the existing data in the database contents, click on the upper right Data button to switch to the profile page as shown in 9(c). The updating page in 9(d) has similar function and appearance as adding data page.

### 2 Mobile User System Design

The mobile user system architecture is comprised of three pages, the query navigation target page, the navigation page, and the destination page. A details page cab be switched back and forth with he destination page, as shown in figure 10.





Figure 10 Mobile users' system architecture

The sequences of the mobile users' operation are depicted in figure 11.



Figure 11 Mobile device users' interface

First, the mobile user enter the system with instructions as shown in 11(b), then enter the destination in figure 11(c). To facilitate positioning the next page will ask the user to take pictures, as shown in 11(d). The image is upload to the server for comparison directly after taking the picture, as shown

in figure 11(e). A successful match will show the users with details page to guide them to the destination as shown in 11(f).

# III. System Testing and Discussion

We implement a field test in a real environment with the guidance of the attendees, the subjects, of a meeting that were taken place in a building of our campus. The System Usability Scale (SUS) is adopted for measuring users' response. It provides a quick and reliable tool for measuring the usability. Subjects were asked to install the indoor navigation system in their mobile devices beforehand. 15 out of the 40 attendees had used the indoor navigation support but only twelve of them successfully complete the test and filled out the SUS questionnaire and feedback suggestion. The statistical summary of the SUS survey result is shown in table 2.

Table 2 SUS	system	usability	scale	score
-------------	--------	-----------	-------	-------

SUS syst	tem usabilit	y scale sco	tre conver	sion results		-				
sample number	Arithmetic mean-	Median	Min	Maximum	Standard Deviation					
12-	54.79-	53.75-	47.5	65	4.94-					
SUS sys	tem usabilit	y scale of	the data-	1 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		i.				
subject	Arithmetic mean-	Standard Deviation	akewaess	kurtosis-	5-po	5-point Likert scale percentage -				The second highest
					17	2	3-	4-	5	-
1	2.92	0.900	-0.712-	0.533-	8%	17%	50%	25%	0%	25%-
2	2.67-	0.778-	0.719	-0.792	.0%-	50%	33%	1756-	0%	17%
3.	3.00-	0.426	0.000	5.500	0%	8%	83%	8%	0%	8%6-
4	1.83-	0.937-	0.383	-1.931-	50%	17%-	33%	0%	0%	0%-
5-	2.75-	0.965	-0.136-	-0.770	8%	33%	33%	25%	0%	25%
6-	2.42	1.165	-0.174	-1.523	33%	8%	42%	17%	0%	17%
7.	2.67-	1.062-	-0.559-	-0.309	17%	17%-	50%	17%-	0%	17%
8.	2.92	0.900-	0.185	-1.865	0%	42%	25%	33%	0%	33%
9.,	3.25-	0.754	-0.478-	-0.868	0%	17%	42%	42%	0%	42%
10-	2.83	0.835-	-0.771-	1.148-	8%	17%	58%	17%-	0%	17%
18間-	2.73	0.872	-0.154	-0.088	13%	23%	45%	2045	0%	20%

The average, 54.79, is acceptable for an new innovative prototype[13]. Yet it also implies that further exploration and improvement need to be done.

# **IV. Concluding Remarks**

This study proposes an innovative approach used in indoor navigation system, the field test of this system shows the subjects are interested in the prototype. However, the SUS questionnaire showed that the usefulness of the system can still be improved. For example, an immediate modification of the prototype has been done from users' feedbacks by providing a more dynamic route indication in the details page of the mobile user interface as depicted in figure 12.



### International Journal of Advances in Image Processing Techniques– IJIPT Volume 2 : Issue 2 [ISSN : 2372-3998]

#### Publication Date : 30 October, 2015



Figure12 dynamic route indication

# References

- Koutsou, A. D., Seco, F., Jiménez, A. R., Roa, J. O., Ealo, J. L., Prieto, C., & Guevara, J. "Preliminary localization results with an RFID based indoor guiding system." WISP 2007. IEEE International Symposium on Intelligent Signal Processing, 2007.
- [2] Lassabe, F., Canalda, P., Chatonnay, P., Spies, F., & Charlet, D. "Refining WiFi indoor positioning renders pertinent deploying location-based multimedia guide." AINA 2006. 20th International Conference on Advanced Information Networking and Applications. Vol. 2. IEEE, 2006.
- [3] Lowe, D. G. "Distinctive image features from scale-invariant keypoints." International journal of computer vision 60.2, 2004: 91-110.
- [4] Mikolajczyk, K., & Schmid, C. "A performance evaluation of local descriptors." IEEE 2005. Transactions on Pattern Analysis and Machine Intelligence, 27.10, 2005: 1615-1630.
- [5] Hansen, R., Thomsen, B., Thomsen, L. L., & Adamsen, F. S. "SmartCampusAAU--An Open Platform Enabling Indoor Positioning and Navigation.", 2013 IEEE 14th International Conference on Mobile Data Management (MDM) Vol. 2. IEEE, 2013: 33-38.
- [6] Lin, C. Y., Chang, C. Y., & Hsu, H. H. (2011, October). "A RFID-Based Personal Navigation System for Multi-story Indoor Environments." In Broadband and Wireless Computing, Communication and Applications (BWCCA). IEEE, 2011: 460-465.
- [7] Hammadi, O. A., Hebsi, A. A., Zemerly, M. J., & Ng, J. W.. "Indoor localization and guidance using portable smartphones. "In Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 03. IEEE, 2012: 337-341.
- [8] Kawaji, H., Hatada, K., Yamasaki, T., & Aizawa, K. "Image-based indoor positioning system: fast image matching using omnidirectional panoramic images." In Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis. ACM, 2010: 1-4.
- [9] Möller, A., Kranz, M., Huitl, R., Diewald, S., & Roalter, L. "A mobile indoor navigation system interface adapted to vision-based localization." In Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia. ACM, 2012.
- [10] Muthukrishnan, K., Lijding, M., & Havinga, P. Towards smart surroundings: Enabling techniques and technologies for localization. Springer Berlin Heidelberg, 2005.
- [11] Bishop, C. M. Pattern Recognition and Machine Learning, Springer, 2007
- [12] Lin, Y. S., Chang, S. M., Tsai, J. C., Shih, T. K., & Hsu, H. H. "Motion analysis via feature point tracking technology." In Advances in Multimedia Modeling. Springer Berlin Heidelberg, 2011: 296-303.

[13] Bangor, A., Kortum, P., & Miller, J. "Determining what individual SUS scores mean: Adding an adjective rating scale." Journal of usability studies 4.3, 2009: 114-123.

