國立交通大學

電控工程研究所

博士論文

基於 ZigBee 智慧型環境與移動式機器人之位置感知系統設計

Design of Location Aware Systems using ZigBee-based Intelligent Environment and Mobile Robots

研 究 生：林嘉豪
指導教授：宋開泰 博士

中 華 民 國 一百零二 年 七 月

基於 ZigBee 智慧型環境與移動式機器人之位置感知系統設計

Design of Location Aware Systems using ZigBee-based Intelligent Environment and Mobile Robots

研 究 生：林嘉豪　　Student: Chia-How Lin

指導教授：宋開泰 博士　　Advisor: Dr. Kai-Tai Song

A Dissertation
Submitted to Institute of Electrical Control Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in
Electrical Control Engineering
June 2013
Hsinchu, Taiwan, Republic of China

中華民國一百零二年七月

基於 ZigBee 智慧型環境與移動式機器人之位置感知系統設計

學生:林嘉豪　　指導教授:宋開泰 博士

國立交通大學電控工程研究所

摘要

位置感知機器人系統透過智慧型環境中特定物體、使用者或機器人本身之位置及這些位置所在的感測資訊提供適當的服務。ZigBee 無線感測網路由於成本低廉且功耗較低，適合用來佈建智慧型環境。本論文之主旨即在於利用 ZigBee 建構之智慧型環境，結合移動式機器人，提出一套能主動提供服務的位置感知服務系統。為了方便建置 ZigBee 無線網路之位置感知系統，本論文發展一套以機率為基礎的無線訊號強度定位演算法。此演算法只需要一次的校正即可應用於不同的環境，並維持一定的準確度。另一方面，為了提供更低廉的機器人自主導航解決方案，本論文提出以單眼視覺為基礎的地平面偵測方式進行環境中物體的偵測及物體距離量測，藉此可進而達成環境結構的估測，以利機器人避障控制。本方法僅需單一攝影機即可在移動平台上偵測其障礙物之距離，不需要使用特殊攝影機或者融合其他感測器資訊。本方法之特色在於融合了反向投影以及色彩分割以穩定的分類地面/非地面區域，可以更加穩定在不同特性的地面和障礙物環境中同時偵測出靜態與動態障礙物。在實際實施應用上本方法僅限制地面為平面，不限定環境中地面及障礙物材質或顏色特性、不需要特定標號、亦不需要先對環境或障礙物特性進行學習。本方法估測結果為 2D 平面上物體之距離與分佈地圖，因此除了可直接用以取代

傳統距離感測器，亦可根據所建構之環境地圖進行路徑規劃。本論文在整合實驗中展示此位置感知系統針對入侵者偵測和即時提供使用者服務兩個情境應用，以驗證所發展方法之有效性。

Design of Location aware Systems using ZigBee-based Intelligent Environment and Mobile Robots

Student: Chia-How Lin Advisor: Dr. Kai-Tai Song

Institute of Electrical Control Engineering
National Chiao Tung University

ABSTRACT

A location aware system provides location information of objects, users and mobile robots from sensor nodes deployed in the intelligent environment. The information can be used to support various intelligent behaviors of a service robot in day-to-day application scenarios. This thesis presents a probability-based approach to building a location aware system. With this approach, the inconsistencies often encountered in received signal strength indicator (RSSI) measurements are handled with a minimum calibration. By taking off-line calibration measurement of a ZigBee sensor network, the inherent problem of signal uncertainty of to-be-localized nodes can be effectively resolved. The proposed RSSI-based algorithm allows flexible deployment of sensor nodes in various environments. The proposed algorithm has been verified in several typical environments and experiments show that the method outperforms existing algorithms. The location aware system has been integrated with an autonomous mobile robot to demonstrate the an on-demand robotic intruder detection system.

To provide a low-cost autonomous navigation solution, we have developed monocular vision system to estimate distances between the robot and obstacles based-on inverse perspective transformation (IPT) in image plane. A robust image processing procedure is proposed to detect and segment drivable ground area within the camera view. The proposed method integrates robust feature matching with adaptive color segmentation for plane

estimation and tracking to cope with variations in illumination and camera view. After IPT and ground region segmentation, a distance measurement result is obtained similar to that of a conventional laser range finder for mobile robot obstacle avoidance and navigation. The merit of this method is that the mobile robot has the capacity of path finding and obstacle avoidance by using a single monocular camera. Practical experimental results on a wheeled mobile robot show that the proposed imaging system successfully estimates distance of objects and avoid obstacles in an indoor environment. Several interesting integrated experiments are presented in this thesis to demonstrate the effectiveness of the location aware robotic system in a home setting.

誌謝

攻讀博士學位的過程是未進入博士班前所無法想像的，這是一條辛苦而又漫長的自我成長之路。一路走來，由衷感謝我的指導教授宋開泰博士，感謝他多年來在我多次感到氣餒之時不斷給予鼓勵，使我終能順利到站，展開人生新的旅途；另一方面，在專業及論文寫作上的指導，不厭其煩的給我建議及修正方向，使我受益良多，也讓本論文得以順利完成。亦感謝論文口試委員－陳永平教授、莊仁輝教授、郭重顯教授及葉廷仁教授對於本論文的種種建議與指引，強化本論文的嚴整性與可讀性。

感謝學長戴任詔博士、蔡奇諡博士、韓孟儒博士對本論文的建議與討論，同時亦感謝過往相互鼓勵的實驗室碩士班學弟妹巧敏、信毅、格豪、允智、崇民、松峙、俊瑋、富聖、振暘、晉懷、濬尉、仕傑、煥坤、舒涵、科棟、奕文、哲豪、仕晟、建宏、上畯、章宏、昭宇、俊儒、京叡、依穎、明翰等在實務驗證時所提供的協助，以及生活上所帶來的樂趣。

另外，特別感謝我的父母，由於他們辛苦栽培，在生活上給予我細心地關愛與照料，使我才得以順利完成此論文；也感謝朋友們的加油打氣，在我最無助及失意的時候給予背後安定的力量。願能貢獻一己棉薄之力，成就有益家庭、社會之能事。

Contents

List of Figures

List of Tables

Chapter 1.

Introduction

1.1 Motivation

As a result of the progress of medical technology and the reduction in birthrate in recent years, the population ratio of the elderly has been increasing in many countries. Elderly-care services will soon become essential and cause social problems in the near future. Technology can play an important role in this issue by providing solution to reduce the man power needed for such labor-intensive work and increase the quality of life of the elderly.

One interesting area is the intelligent environment, which can gather user and environment information spontaneously via context aware systems. It also provides the connectivity to distribute information remotely via wireless sensor networks (WSNs) and internet. For instance, a family member can request or subscribe the images and location information of the elderly through e-mail or cell phone. The location of the elderly is then estimated using the wireless sensor module attached. The location server then collects the information, and sends to the video server. The video server controls all the cameras in the house and thus can select a proper camera to acquire images of the elderly according to the location information provided by the location server. Finally, all the information will be sent to the user B to ensure the exact situation of the elderly. The elderly can also ask for help via speech recognition system. Family members will be alerted using the same infrastructure. In practice, however, the coverage of the camera may not be complete, and even if the event is monitored, the system cannot provide any physical aid.

Another interesting area is the mobile robot. The development of robotic security and monitoring systems for indoor environments such as factories, offices or home settings has gained increasing attention in recent years [1-3]. A mobile robot can provide sufficient mobility and flexibility to help security guards which lowers the relative cost as long as the demand for adequate security is satisfied. Furthermore, it has the ability to help the elderly physically. However, robots are still far from practical for daily-life use in many circumstances, especially in a home setting. One critical reason is that they lack the capability to gather sufficient information of the environment in order to adapt their changes in real-life application scenarios. For instance, under the scenario of a user having an asthma attack, the robot should have the following capabilities. The robot should notice the incident immediately by monitoring the bio-physical situation of the user. After detecting the emergency situation, the robot should be able to plan a task to resolve the problem, that is, to bring the quick-relief medicines to stop asthma symptoms. In order to accomplish the planned task, the robot should be able to locate the medicines and user, and to navigate to their location autonomously. For an on-demand robotic system, a locational aware module is required to provide location data of objects of interest, users and of the mobile robot itself. This information supports various intelligent behaviors of a service robot in day-to-day scenarios. According to the reasons above, this thesis aims to investigate the location aware system using intelligent environment and mobile robots, and establish a practical on-demand robotic system to assist the security and healthcare of elderly at home.

Taking home security as an example, conventional surveillance systems have various limitations and shortcomings. A security agent employing conventional security systems has to dispatch security guards to respond to alarms. This creates a heavy burden for security agencies, especially considering a high portion of alarms are actually false alarms. Furthermore, it normally takes some time for the security guard to reach the alarm location to handle the

situation. Critical time may have already been lost by the time the security guards arrive on the scene. The combination intelligent environment and mobile robots can solve this problem. By deploying the wireless sensor network (WSN) and security cameras all over the guarding area, false alarm can be reduced. In addition to passive monitoring, the mobile robot can take place the security agent to check the alarm spots more actively and agilely.

1.2 Literature Review

1.2.1 Localization of ZigBee-Based Wireless Sensor Network

In the past decade, the production of various off-the-shelf devices for wireless sensor networks (WSNs) has progressed rapidly and such devices can support the use of on-demand robotic systems. WSNs are constructed using a series of sensor nodes deployed throughout the monitored environment. These sensor nodes are connected and communicate with each other over a wireless network such as wireless LAN, Bluetooth, ultra-wide band or ZigBee. This allows the use of free-ranging, autonomous service robots which operate in human centered environments to assist people. Because the robot has limited onboard sensing and computing abilities, integrating service robots with a WSN is desirable to enable robots to be used for more practical applications. A WSN can serve as a low-cost, distributed and easily deployed monitoring system which extends the limited sensing capability of the robots.

Unlike a stand-alone security robot, which is normally equipped with a variety of sensors but still has only a limited sensing range, on-demand security robots can obtain thorough sensory information of the guarded area and make it available on line. When integrated with the WSN, a security robot can acquire information from the entire sensed environment in real time. To enhance the prompt response of security robots to an intrusion situation, many researchers have studied the application of WSN and mobile robot techniques for use in security and monitoring [4-5]. WSN-based localization can estimate the location of a mobile

node with an initially unknown position in a sensor network using available a priori knowledge of positions of a few specific reference nodes in the network. However, most of these designs require overly complex and impractical installation procedures and suffer from problems with signal calibration/training. For example, some wireless sensor modules have to be placed in restricted locations such as on a ceiling [3]. Among off-the-shelf WSN technologies, ZigBee is a popular ad hoc network based on the standard IEEE 802.15.4. ZigBee Alliance [6] defines the specification of ZigBee for networks and higher layers. The ZigBee standard has many advantages for real-time positioning systems in terms of battery life and network size. It is widely used in low data rates, low power, and cost effective wirelessly networked products

Localization algorithms should meet the requirements of various hardware configurations such as signal transmission, power requirements, computational complexity, etc. These factors allow us to divide various approaches into three main categories [7]: range-free, range-based, and fingerprinting. Range-free localization [8] is based on the connectivity of the network. It does not require any special hardware and the content of the messages is received through simple operations. Range-based localization, on the other hand, estimates the distance between nodes using certain ranging techniques. The distance information can then be used to locate the position of unknown sensor nodes. Most range based localization algorithms adopt the received signal strength indicator (RSSI) technique to estimate distance based on the strength and path loss model of the signal which was received. After RSSI is used to estimate the distance, the second phase of localization performs computations based on the relative distance between nodes. Many strategies exist for location estimation such as multi-lateration [9] and min-max [10]. Unfortunately, indoor radio channels are unpredictable, because reflections of the signal against walls or ceiling may result in severe multi-path interference at the receiving antenna [11]. A straightforward way to overcome the inaccuracy of RSSI is to use enhanced or

additional radio hardware, such as multiple directional antennas [12]. However, the devices used in these solutions generally demand more energy and are much more expensive to deploy.

The fingerprinting approach or location pattern matching techniques [13] are based on the concept of identifying a specified position by relying on RSSI data received from nearby nodes. This approach uses two phases, a training phase and an estimation phase. In the training phase, the RSSI is measured at grid points in the area of interest. This information is used to estimate the propagation model parameters, which are employed later in the estimation phase. The accuracy of the calibration procedure depends on the number of points in the grid and the number of measures taken per point. Since the fingerprinting approach is basically a pattern matching and generalization technique, many researchers have applied state-of-the-art intelligent-computation approaches to resolve this problem. Oliveira *et al.* [14] combines the RSSI technique and the link quality indicator (LQI) using fuzzy logic and transferable belief model (TBM). Their results indicate that this type of combination metrics can refine the estimated distances. Gogolak *et al.*[15] preprocess the RSSI values to obtain the mean, median, and standard deviation of the data and uses these processed data to train a neural network in order to increase the accuracy of the localization system. Fang and Lin [16] proposed the discriminant-adaptive neural network (DANN), which extracts useful information into discriminative components (DCs) for neural network learning. The nonlinear relationship between RSSI and the object's position is then accurately constructed by incrementally inserting the DCs and recursively updating the weights in the network until no further improvement is required. In general, these techniques yield improved performance when the number of nodes is relatively large. However, they are very time consuming as the process requires exhaustive data collection, which is a practical barrier to its wider adoption.

Due to the long calibration time required for fingerprinting approaches, range-based methods are more suitable in applications where fast deployment is necessary. The most

challenging problem is how to overcome the irregularity of range estimation using RSSI. Ramadurai and Sichitiu proposed to merge all the data collected and estimate the probability distribution of WiFi radio frequency (RF) signal strength as a normal distribution function of distance [17]. While their results indicate that the actual position of a signal node is well bounded by the estimated position obtained despite ranging inaccuracies, their method is not suitable for an indoor environment because of the presence of multipath fading. With lower power, stronger shadowing and fading effects, it is even more difficult to find a proper distribution for ZigBee-based systems. Fig. 1-1 shows an example of the relationship between RSSI and distance using ZigBee modules. As the figure shows, the relationship between RSSI and distance is quite irregular, and a perfect mathematical model which adequately describes this type of distribution is difficult to find. In [18], a thorough free space loss model of RF signal is proposed. This model considers the irregularity seen in RSSI and the parameters of the RF hardware (antenna gain and efficiency, transmit power, etc.). However, this subtle

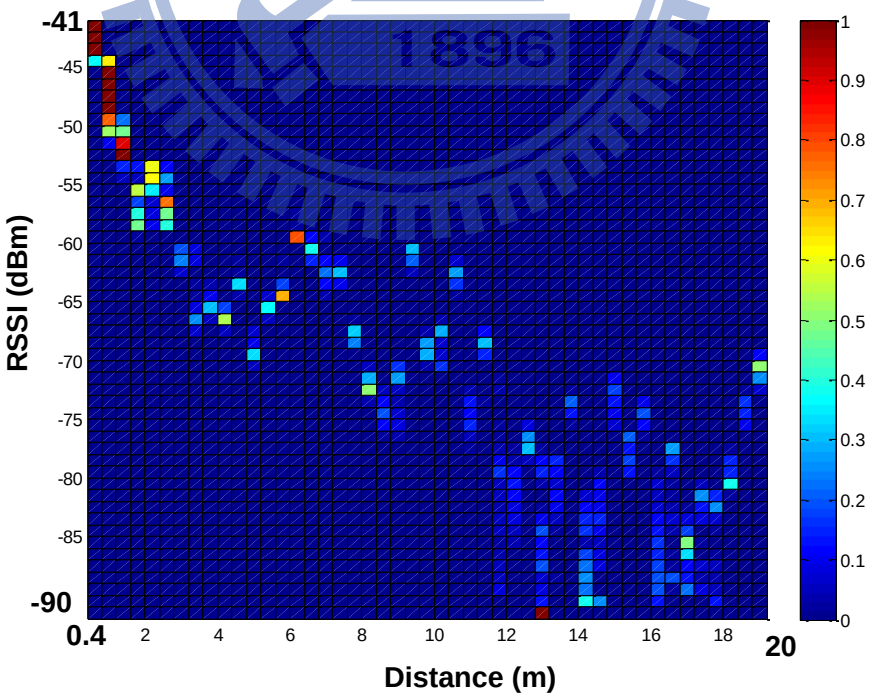


Fig. 1-1: An example of RSSI irregularity.

model assumes that only the direct signal reaches the antenna with the RSSI value, and may still fail in real life situations. Similar observations can also be found in many other attempts [19-21]. When the model fails to accurately describe the relationship between RSSI and distance, overall accuracy will decrease dramatically. Lee et al.[21] used both unscented Kalman filtering (UKF) and particle filtering (PF) to track a target tag with RSSI. The UKF assumes near-linearity and uni-modal Gaussians in the system while the PF does not. Their results show that the assumptions of UKF are very fragile and easily violated. Furthermore, while the tracking method itself is verified, several sets of parameters for the RSSI-model for different parts of the area need to be determined beforehand, to overcome the irregularity.

In summary, the currently available location-aware solutions, whether off-the-shelf or state-of-the-art approaches, suffer from the requirement of prior parameter training and retraining in different application areas. It is difficult or even impossible to build up a location aware system in a totally new environment. This motivated us to develop a new localization method which can provide a straightforward and practical solution for sensor-node calibration and provide WSN localization with acceptable accuracy. The proposed method requires no extra hardware. Most importantly, calibration is only needed once and provides adequate robustness in different environments.

1.2.2 Vision-based Mobile Robot Navigation

Obstacle detection and avoidance are essential for mobile robot navigation. Range-based systems such as ultrasonic sensors and laser scanners are most used sensing devices for environment detection. These ranging systems provide only distance information of the environment, making it difficult to distinguish between an obstacle and a target object, with which the robot would need to interact in many applications. On the contrast, vision-based environment detection systems provide rich information for mobile robot navigation, as long as

the illumination is sufficient in the environment. They have become promising alternatives considering the availability of low-cost image sensors and high-performance computing hardware. Various vision-based approaches, including omnidirectional cameras, stereo and RGB-depth cameras have been widely studied for environment detection, due to their wide view angles or the ability to estimate depth information. However, their specific hardware structures raise problems of price and system complexity and thus limit their applications. Monocular vision-based approaches, on the other hand, do not require special camera configuration. They are considered to be more suitable for cost-effective robotic applications.

Lacking depth information, many monocular vision-based approaches detect objects or obstacles with prior assumptions over the environment. For instance, color cue is commonly used to segment an obstacle or non-obstacle in an image frame [22]. Chen *et al.* [23] a applied moving object tracking techniques to find obstacles which have significant differences in motion. However, these approaches can only be applied to limited scenes. Optical flow, on the other hand, is a more general approach, which can be used to estimate time-to-collision of obstacles, or the environmental structure from motion (SFM) [24-25]. The computational load of optical flow is rather high, but the robustness against disturbance is still questionable. Recently, some researchers have proposed to use machine learning approaches to obtain depth or structural information using merely monocular cues from single view [26-27]. Although their results are quite impressive in many cases, the outcome is not predictable and stable enough.

Since detecting obstacles in a general manner is still challenging, many researchers propose to reduce the problem to ground detection in a road scene or indoor environments [28-31], with an assumption of ground planarity. The assumption assumes that the ground is always parallel to the robot's motion. Although this assumption may fail in certain environments such as wheelchair ramps or auditoriums, it is still valid in many common indoor

environments. Under this assumption, features on the ground can be modeled and tracked using projective invariants or homography matrix. Among many successful tools, Liang *et al.* [28] proposed to use reciprocal-polar (RP) image rectification and ground plane segmentation by sinusoidal model fitting in RP-space to segment the ground plane from a mobile robot's visual field of view. Furthermore, the height of off-ground plane features can also be measured. Zhou *et al.* [29-30]derived constraints that the homography of the ground plane must satisfy, and then used these constraints to design algorithms for detecting the ground plane by assuming that the camera is fixed on the robot platform and can at most rotate horizontally. Conrad and Desouza use EM Algorithm to classify the features, and achieve an almost perfect detection rate (over 99%) despite the relatively higher number of errors in pixel correspondence [31]. The feature detection and ground model can then be used directly to control the robot or further ground region detection.

As for ground region detection, pixel-level obstacle representation such as warping or inverse perspective transform (IPT) techniques has been widely studied in both stereo-based and monocular-based approaches [32-37]. This technique uses ground plane homography to warp one of the acquired images. Therefore, only pixels which correspond to the ground plane will match in both images. This discrepancy can then be detected by image subtraction and thresholding. Yamaguchi *et al.* [35] estimate relevant parameters for camera motion as well as the 3D road plane from correspondence points between successive images. The road region is determined by first estimating the ground homography matrix and warping the image of previous frames. In their method, while subtraction is a simple and effective method to detect pixel-wise floor regions, it suffers from some limitations. Due to the nature of subtraction method, textureless obstacles cannot be detected since no discrepancy can be observed even under erroneous projection. Similarly, part of the obstacles may be projected into the floor region and cause some floor areas to be recognized as obstacles. Nevertheless, textureless

obstacles are also difficult to detect using feature-based methods. These problems become more significant in indoor environments due to the common appearance of mono color walls, and need to be resolved for practical applications.

1.3 Research Objectives

The objective of this thesis is to develop a location aware service system which features a ZigBee WSN and a mobile robot. By integrating these technology, the location-ware system will be able provide not merely location information passively, but also provide on-demand services such as intruder detection and active health care. Take intruder detection as an example, if any intruders or abnormal conditions are detected, the state and location of this alert will be transmitted to the security robot and monitoring center on the WSN. The robot can navigate autonomously to the alarm location using the probability-based localization system. After the robot arrives on the scene, the onboard camera will transmit real-time images to the user via both WiFi and 3G networks. The security guards and end-users can therefore easily determine the exact situation in real time.

In order to deploy the WSN for location aware system, this thesis presents a novel probability-based approach to building a location aware system. With this approach, the inconsistencies often seen in received signal strength indicator (RSSI) measurements are handled with a minimum of calibration. By taking off-line calibration measurement of a ZigBee sensor network, the inherent problem of signal uncertainty of to-be-localized nodes can be effectively resolved. The proposed RSSI-based algorithm allows flexible deployment of sensor nodes in various environments.

This thesis also aims to find a robust method to detect ground plane region in an indoor environment. In particular, the method should be able to deal with textureless and mono color ground and obstacles. The determined ground plane will be further processed to provide range

information for mobile robot obstacle avoidance. The main concept is to integrate multiple visual cues to avoid limitations in image feature extraction and achieve robustness against illumination and camera view variations. First, robust feature extraction and matching methods are adopted for feature extraction and matching to increase both robustness and speed in the homography estimation step. Furthermore, a key-frame selection criterion is proposed to guarantee the estimation is proper and avoid the virtual plane problem caused by zero/near-zero camera translations. To observe ground region pixel-wise, we propose an adaptive self-segmentation and tracking algorithm to find the ground region according to both the feature points and color segments. Applying IPT on the ground-region-segmented image, results in a rectified image, similar to an occupancy grid map. The robot can then use distance information to prevent from colliding with obstacles. The constructed map itself can also be used for many other robot control purposes.

1.4 Organization of the Thesis

The remainder of this thesis is organized as follows: In Chapter 2, the probability-based localization method is first presented. The proposed algorithm guarantees a robust localization method under different environment. The localization result can be applied on robot as well as a user with the ZigBee module In Chapter 3, monocular visual navigation system is presented to provide navigation ability for the robot with merely a camera. Chapter 4 shows the hardware and other sub system implemented for establishing the intelligent environment and the mobile robot system. Experimental results of the proposed are reported and discussed in Chapter 5. Chapter 6 concludes the contributions of this work and provides the recommendations for future research.

Chapter 2.

Probability-based Location Aware Design

2.1 Introduction

For an on-demand robotic system, a locational aware module is required to provide location data of objects of interest, users and of the mobile robot itself. This information supports various intelligent behaviors of a service robot in day-to-day scenarios. In this thesis, a novel probability-based approach to building up a location aware system is presented. In this approach, the uncertainties and inconsistencies often seen in received signal strength indicator (RSSI) measurements are handled with a minimum of calibration. By taking off-line calibration measurement of a ZigBee sensor network, the inherent problem of signal uncertainty of to-be-localized nodes can be effectively resolved. The proposed RSSI-based algorithm allows flexible deployment of sensor nodes in various environments. The proposed algorithm has been verified in several typical environments. The rest of this chapter is organized as follows: In section 2.2 the proposed probabilistic localization method is presented. The performance evaluation of the localization system are described and discussed in session 2.3.

2.2 Proposed ZigBee Localization Method

To overcome limitations posed by uncertainties of RSSI, we suggest a novel probability-based approach to estimating location by modeling the RSSI vs. distance relationship with discrete probability density functions. This approach aims to provide a straightforward method of describing different distributions without losing generality. While

most current approaches use mathematical equations to model the RSSI vs. distance relationship, the proposed method adopts the original RSSI vs. distance data to construct its own model. Furthermore, instead of using trilateration or fingerprint techniques, the location is estimated by accumulating a probability histogram observed from several reference nodes on a local probability map. This approach will be shown to have better tolerance against fluctuations and inconsistencies inherent with the RF signal. It is also flexible because it uses the available number of the RSSI measurements and the deployment of ZigBee nodes.

We first assume that the locations of all reference nodes are already known and the RSSI between each pair of nodes can be received. Fig. 2-1 shows a flow chart of the algorithm. It contains a calibration phase which models the RSSI vs. distance relationship and a localization phase which estimates the location using a filtered 2D probability map as follows.

2.2.1 Modeling of RSSI vs. Distance Relationship

The calibration phase estimates the distance between nodes based on the RSSI measurement. Instead of using a mathematical path-loss model, the RSSI vs. distance relationship is collected and modeled with a series of probability histograms, which record different distances measured under a fixed RSSI. The histogram therefore represents the discrete probability density function (*pdf*) of a given RSSI value. Let R and D denote the random variables of the RSSI reading and distance between nodes (in meters), respectively. Assume the model is tested from distance 0 to L meters at intervals of q meters. The *pdf* of the RSSI value equal to r can be defined as:

$$\mathrm{P}(D \mid R=\mathrm{r}) = \sum_{i=0}^{N-1} \mathrm{P}_D(\mathrm{d_i} \le D \le \mathrm{d_{i+1}} \mid R=\mathrm{r}) = \sum_{i=0}^{N-1} h_i \quad , \qquad (2.1)$$

where d_i represents the distance of the i^{th} interval, $N = L/q$ is the total bins in the histogram and h_i is the value of each bin. Fig. 2-2 shows an example of the probability histogram when RSSI

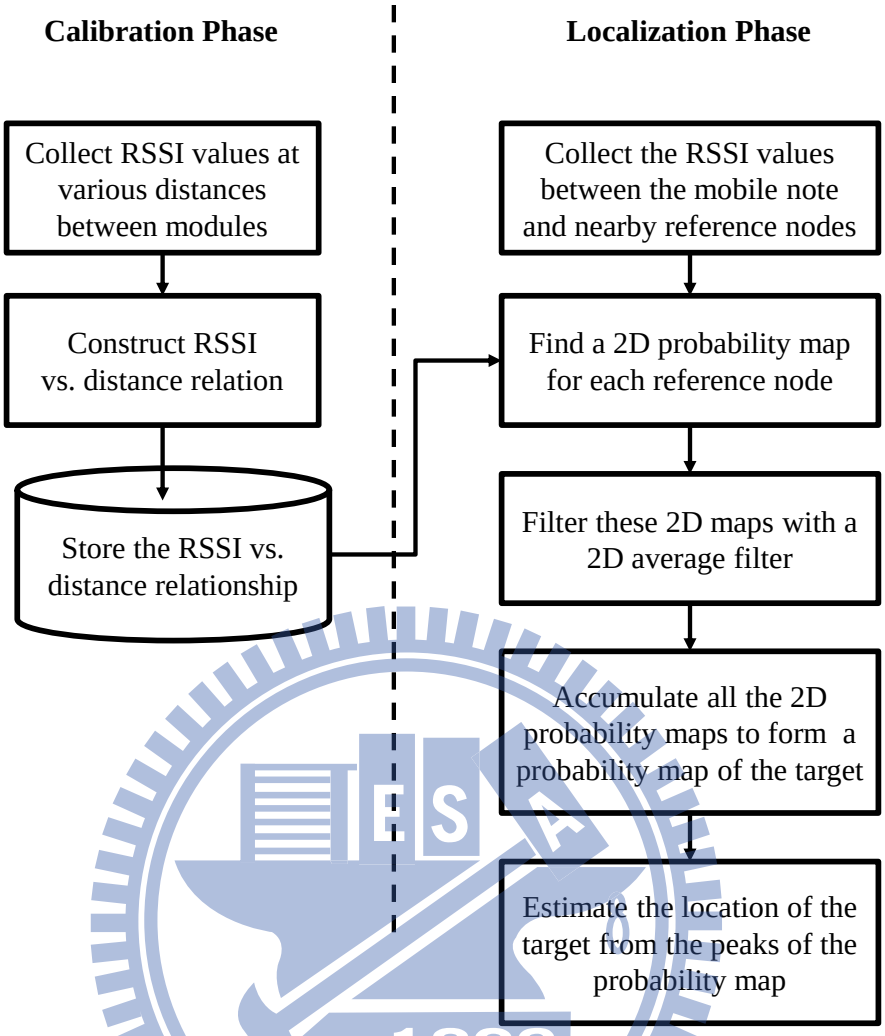


Fig. 2-1: The flow chart of the proposed localization algorithm.

= –88 dBm. This model can easily be implemented onboard the sensor node since it only needs a single look-up table. Clearly, the number of the bins in each histogram limits the resolution of the estimation of distance. This is, however, not critical since the RSSI value is also discrete in practice.

2.2.2 2D Probability Map

In the next step, the location of the mobile node is estimated using the RSSI values measured from several nearby reference nodes. The concept is to apply the two-dimensional multi-lateration method with the estimated *pdf*s. A trilateration method determines the location by finding the intersections of circles given from (1) the centers of the reference nodes; and (2)

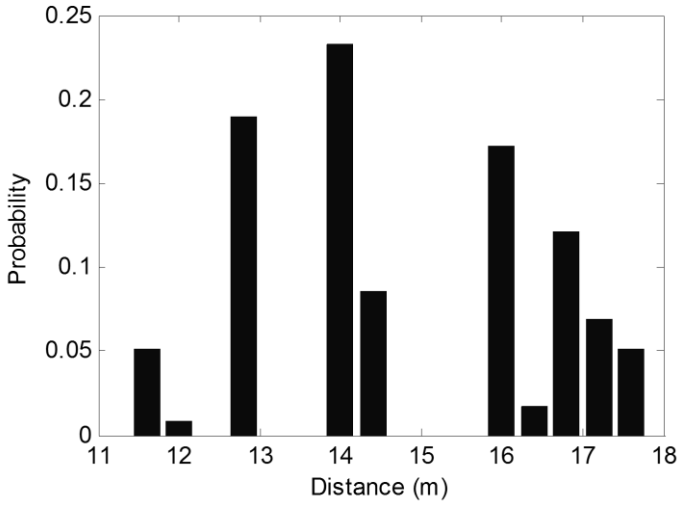


Fig. 2-2: Probability histogram of distance for RSSI = –88 dBm.

distance between the mobile node and several reference nodes (more than three). In the proposed algorithm, however, a mobile node may obtain several possible distance measurements from each reference node. The next step is to find the intersection sets of distance measurements with the largest probability from several reference nodes.

To simplify the calculations, the proposed algorithm estimates the 2D probability distribution directly on the 2D map quantized in grids of a fix size of K × K. Specifically, grids of the size N × M are spanned over the localization area, defined by (x,y), $1 \le x \le N$ and $1 \le y \le M$. The next step is to derive the location of the mobile node $\mathbf{x}_m = (x_m, y_m)$ from a RSSI vector $\mathbf{R} = (R_1, R_2 \ldots, R_i, \ldots R_K)$, containing the RSSI values received from K reference nodes with known positions $\boldsymbol{x}_{r,i}$, $1 \le i \le K$. For each reference node $\boldsymbol{x}_{r,i}$ with a RSSI value r measured at the mobile node, the 2D pdf $P(x,y,i)$ of the mobile node at position (x,y) can be given as:

$$\begin{aligned} P(x,y,i) &= \sum_{y=1}^{M}\sum_{x=1}^{N} P_{X,Y}\left(\mathbf{x}_m = (x,y) \mid R_i = r\right) \\ &= \sum_{y=1}^{M}\sum_{x=1}^{N} P\left(D = \left\|(\mathbf{x}_{r,i} - (x,y))\right\| \mid R_i = r\right) \end{aligned} \quad . \qquad (2.2)$$

Since the RSSI vs. distance model is known and modeled as a 1D *pdf*, the 2D probability distribution is equivalent to a repeat of the 1D *pdf* P(D |R = r) in (1) around the coordinates of the reference node from 0 to 360 degree. Fig. 2-3 shows a graphical example of the 2D *pdf* with one mobile node and one reference node at (0,0). The 2D *pdf* indicates the probability of the mobile node by means of a 2D probability map.

2.2.3 Filtered 2D Probability Map

While the localization method described in the previous section works under many circumstances, some practical issues still need to be resolved. Among them, the most important problem is the gap in the histogram created by the inconsistency of the RSSI values. For instance, in Fig. 2-2 (when RSSI = –88 dBm), the probability of the distance = 15 m is zero,

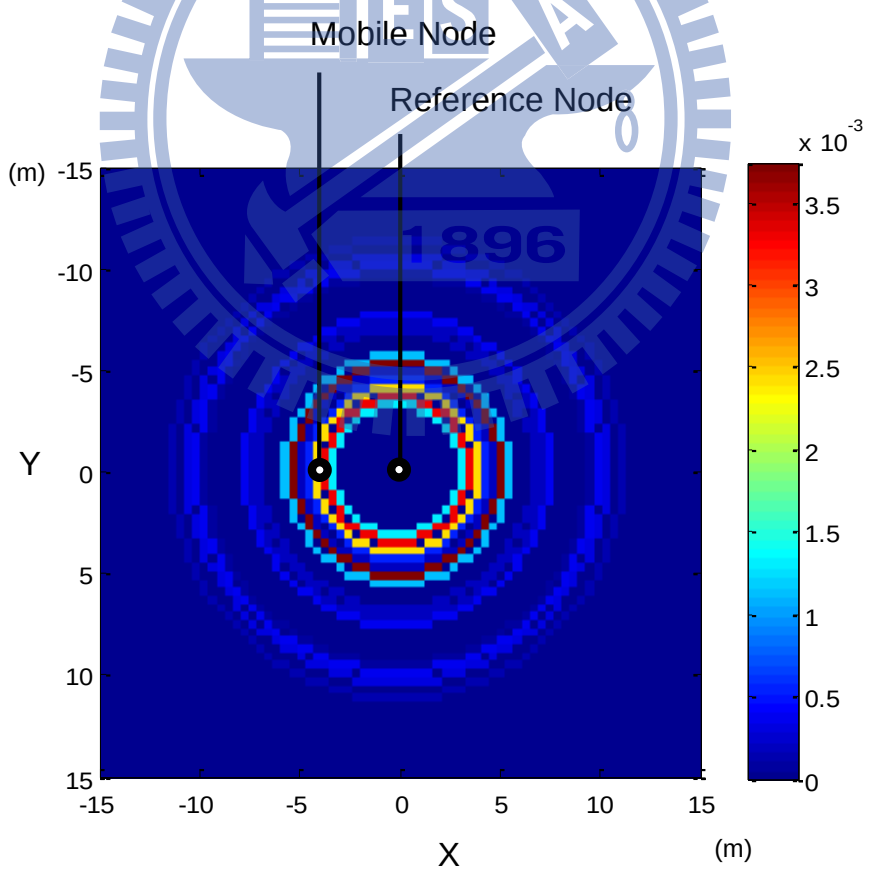


Fig. 2-3: The 2D pdf of a fixed RSSI reading between a reference node and a mobile node.

which is unlikely in reality, because of the irregular nature of the radio signal. To realize the goal that "the model only needs to be built once," we need to enhance the proposed model to tolerate these circumstances. Furthermore, for some WSN applications such as searching, it would also be beneficial if the proposed method can provide alternative locations when it fails on the first try. In order to resolve these issues, we treat the 2D probability map as an image and apply a 2D circular averaging filter H(x,y) with radius *r* to it. The idea is to smooth the map by filling the gaps and avoid any zero probability on the map by taking probabilities of nearby areas into account. The filter is defined as:

$$H(x,y) = \frac{1}{r^2\pi} \cdot \begin{cases} 1, \text{if } x^2 + y^2 \le r^2 \\ 0, \text{otherwise} \end{cases}, \tag{2.3}$$

where the value of *r* is adjusted according to the variance of the RSSI values. The filtered map is thus stated as:

$$\hat{P}(x,y,i) = P(x,y,i) * H(x,y). \tag{2.4}$$

As a result, the filtered map provides a relatively more robust result than the raw map (refer to the experiment section). The filtered map also provides other possible locations for the robot or human to search for the target, if they failed to find the target at the location with the highest probability.

2.2.4 Location Estimation

Finally, the location of the current mobile node will be estimated by accumulating all the 2D probability maps obtained from reference nodes such that:

$$P(x,y) = \sum_{i=1}^{K} \hat{P}(x,y,i). \tag{2.5}$$

The summation of each grid P (x,y) is analogous to the convolution of all the individual distributions. Fig. 2-4 shows an example of how maps obtained from different reference nodes are accumulated into the final map. As a result, the mobile node is located by the peak value of the probability map, even though the upper-right map suggests a false result. However, because of the natural irregularities of the RSSI, the results may lead to a map with several peaks. The location of the mobile node will not be determined by one of the peaks under those circumstances. Instead, the propose algorithm determines the location by estimating the geometric center of all the peaks such that:

$$\mathbf{x}_m^* = \frac{\sum_{p=1}^{P} \mathbf{x}_p}{P}, \tag{2.6}$$

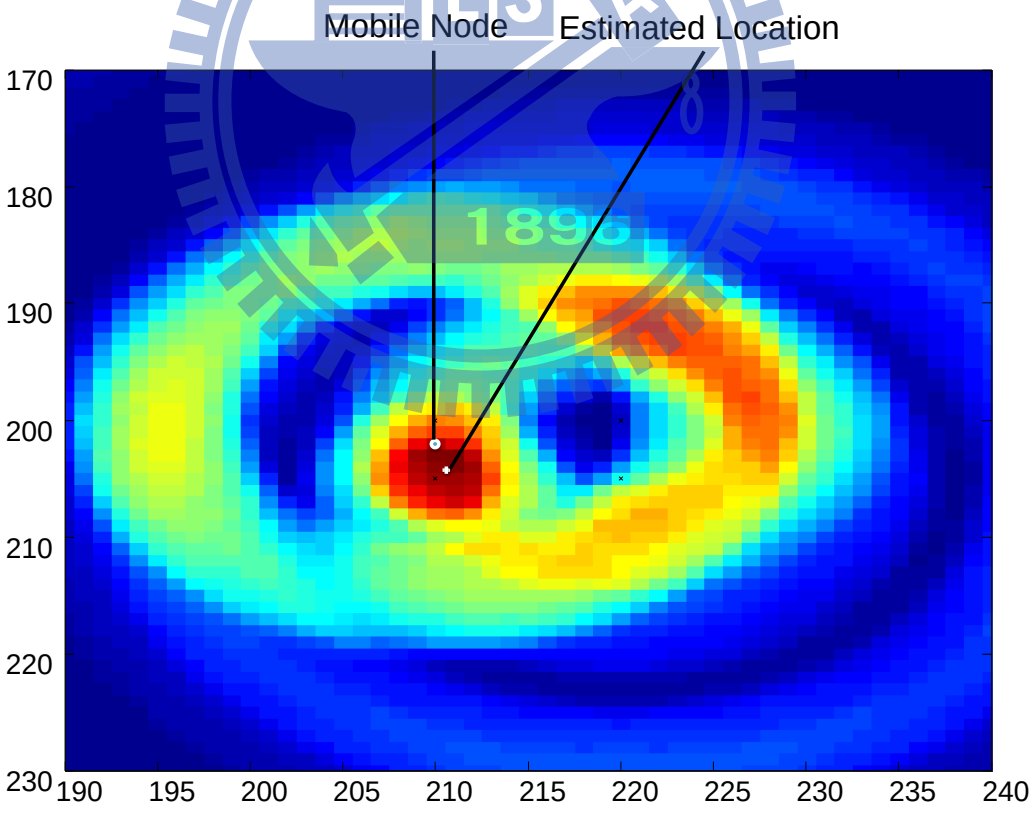


Fig. 2-4: An example of how filtered 2D maps obtained from different reference nodes are combined into the final map.

where $\boldsymbol{x}_p$ denotes location of the p^{th} peak on the 2D *pdf*, $1 \le p \le P$. Note that the multi-lateration method can now be treated as a special case of the proposed algorithm, where the *pdf* of each reference node is an impulse function.

2.3 Robustness Test of The Proposed Localization Algorithm

The experiment aims to test if the proposed method can handle ZigBee localization in different environments more robustly than other methods. Three localization methods: Min-max, ROCRSSI, and CC2431 are compared with the proposed method in this experiment. Min-Max is a widely-adopted range-based localization method [7]. The technique known as Ring Overlapping based on Comparison of Received Signal Strength Indication (ROCRSSI) is a commonly used range-free approach [8] to localization. The CC2431 location engine is embedded in a Texas Instruments CC2431 system-on-chip (Soc) solution for ZigBee/IEEE 802.15.4 wireless sensor networks. CC2431 implements Motorola's IEEE 802.15.4 radio–location solution exploiting a maximum likelihood estimation algorithm [6]. To compare the CC2431 location engine with the proposed method, we have adopted our modules with CC2431 chips to perform the experiment.

Fig. 2-5 shows the data flow of the experimental system. RSSI values of reference nodes were first estimated by the mobile node. The values were then sent to the observation node and finally transmitted to the host PC via a serial link. The collected RSSI values were then estimated with different algorithms and compared with ground truth. Three experiments have been performed in three different types of environments. Table 2-1 shows the configurations of these environments. The localization result using raw data as well as the filtered 2D map were individually tested in the experiment to further investigate the effectiveness of the proposed method. Fig. 2-6 , Fig. 2-7 and Fig. 2-8 provide photos and the locations of beacons in these environments. During the calibration measurements of the proposed method, RSSI values at

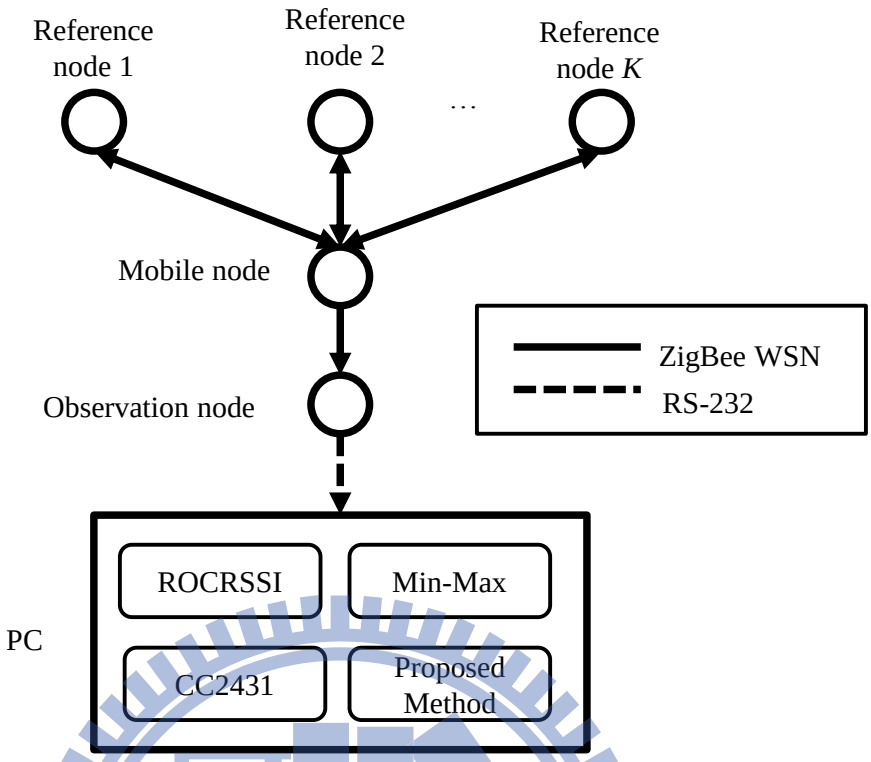


Fig. 2-5: Data flow of the ZigBee localization experiments.

various distances in several different environments of the lab building were recorded. In the measurement, the RSSI values between nodes were collected by moving a mobile node away from a reference node manually along a direction from a distance of 0 to 10.2 m with an interval of 0.3 m. The goal is to collect RSSI data under different circumstances. The calibrations of the other methods were performed only in Experiment Environment 1 (EE1). In the experiments, every ZigBee module was fully charged before the experiment, since the power of ZigBee sensor nodes affects RSSI and hence the location prediction.

Table 2-1 shows the overall localization results of EE1. The results show practically no difference in the localization error between the proposed method and the CC2431 localization engine. Both results outperform the Min-Max and ROCRSSI methods, which is similar to the results reported in [7]. Notice that the parameters of the signal propagation model used in the CC2431 localization engine are given to maximize the accuracy of the experiment, while our approach uses the sensor model established beforehand from other calibration places. In

Experiment Environments 2 (EE2) and 3 (EE3), the same RSSI models and parameters were applied as were used in EE1. Tables III and IV show the overall localization error of the experiments in EE2 and EE3, respectively. Since these two environments are different from the test in EE1, the RSSI-based localization models of the other methods are no longer suitable. As a result, other methods perform more poorly in the new environments. On the contrary, the proposed method already considers possible changes in the environment and preformed more robustly in different environments.

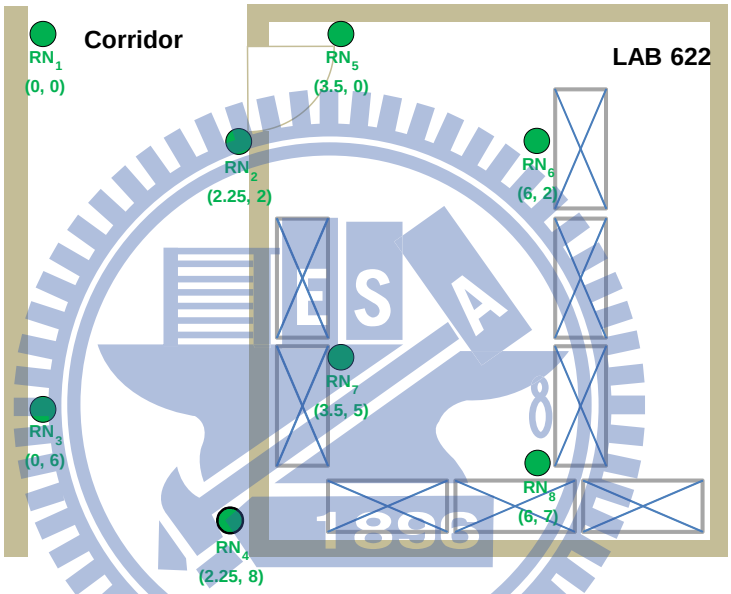
2.4 Summary

This chapter presents a novel method for localizing a ZigBee node in a WSN. The probability-based approach takes the uncertainties of RSSI-based distance estimation into account. The method is independent of the sensor node distribution and environment configuration and therefore can minimize the calibration burden in practical applications. The performance of the proposed algorithm has been evaluated by experiments in three different environments. The experimental results show that the average error of the implemented location aware system is 1.7 m, despite a much higher range uncertainties from raw RSSI values. This accuracy can effectively support service robotic applications, especially when properly combined with a vision system. For instance, the localization system can provide a possible area of a user calling the robot for service or when an emergency situation occurs.

Table 2-1: Configurations of Test Environments

Test Environment	Type	Test point #	Area/m^2	Beacon #
EE1	Room	26	49	8
EE2	Corridor	21	30	6
EE3	Hallway	60	600	8

(a)



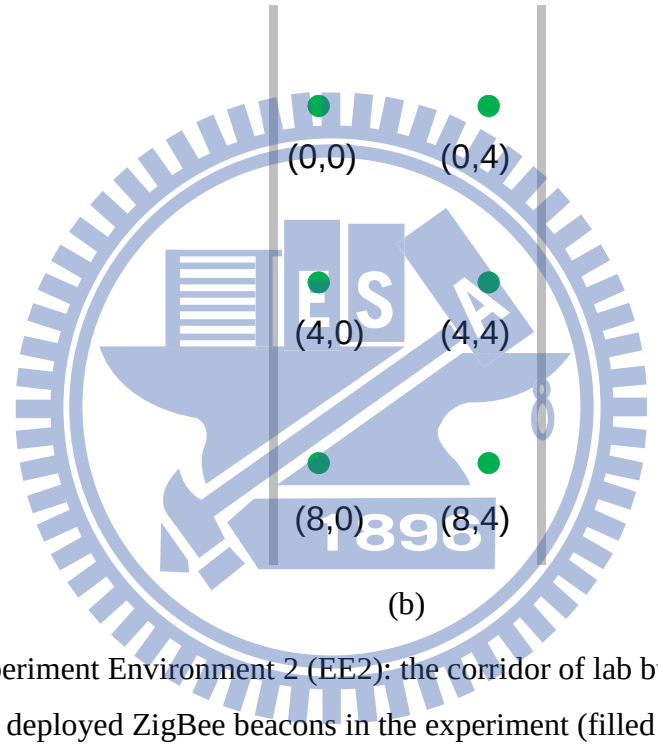
(b)

Fig. 2-6: (a) Experimental environment 1 (EE1): the lab 622 and corridor (b) the location of deployed ZigBee beacons in the experiment (filled circles).

Table 2-2: Estimation Error Comparision: EE1

Method	Average Error (m)	Max Error (m)	Standard Deviation
Min-Max	2.29	3.4	0.77
ROCRSSI	2.59	3.5	1.11
CC2431	0.91	3.25	0.64
Proposed (2D Map)	1.29	3.3	0.82
Proposed (with Filtered Map)	1.12	3.5	0.89

(a)



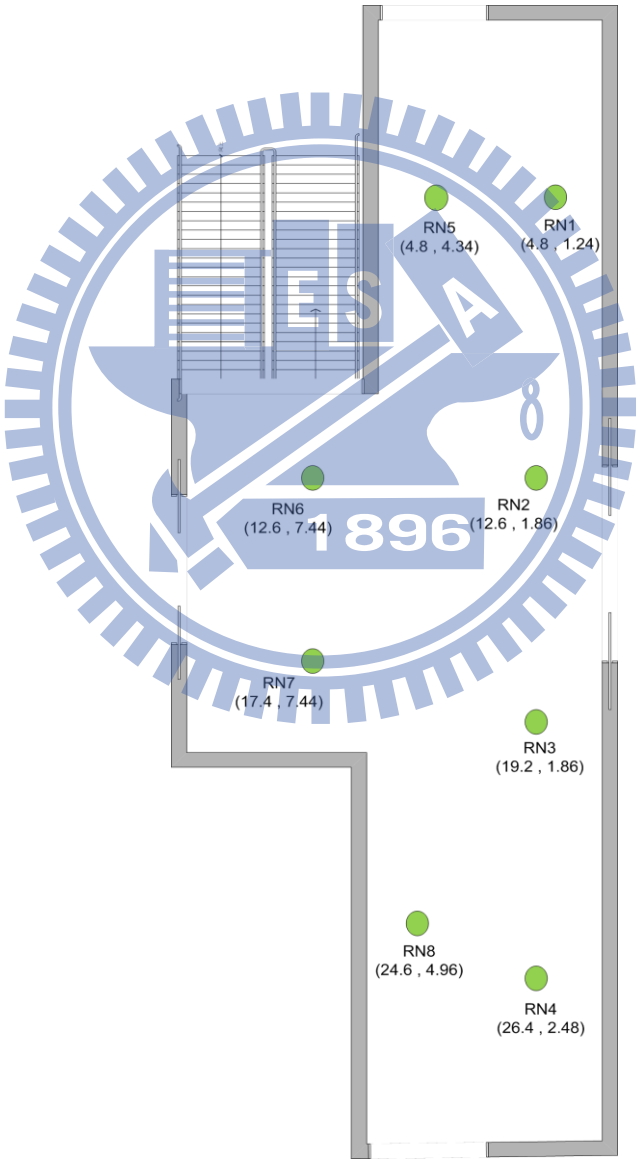
(b)

Fig. 2-7: (a) Experiment Environment 2 (EE2): the corridor of lab building, (b) the location of deployed ZigBee beacons in the experiment (filled circles).

Table 2-3: Estimation Error Comparision: EE2

Method	Average Error (m)	Max Error (m)	Standard Deviation
Min-Max	2.46	3.2	1.14
ROCRSSI	2.76	3.1	1.22
CC2431	1.82	3.5	1.5
Proposed (2D Map)	1.32	3	0.73
Proposed (with Filtered Map)	1.21	3.2	0.92

(a)



(b)

Fig. 2-8: (a) Experiment Environment 3 (EE3): the lobby of the lab building (b) the location of deployed ZigBee beacons in the experiment (filled circles).

Table 2-4: Estimation Error Comparision: EE3

Method	Average Error (m)	Max Error (m)	Standard Deviation
Min-Max	3.7	6.3	1.29
ROCRSSI	3.1	5.2	1.61
CC2431	2.2	7.58	1.74
Proposed (2D Map)	1.5	4.2	0.89
Proposed (with Filtered Map)	1.36	3.21	0.73

Chapter 3.

Monocular Visual Navigation Design

3.1 Introduction

Obstacle detection and avoidance are essential for mobile robot navigation. This chapter proposed a monocular vision-based obstacle avoidance design for mobile robot navigation. A novel image processing procedure is developed to estimate the distance between the robot and obstacles based-on inverse perspective transformation (IPT) in an image plane. A robust image processing solution is proposed to detect and segment a drivable ground area within the camera view. The proposed method integrates robust feature matching with adaptive color segmentation for plane estimation and tracking to cope with variations in illumination and camera view. After IPT and ground region segmentation, distance measurement results are obtained similar to those of a laser range finder for mobile robot obstacle avoidance and navigation. The merit of this algorithm is that the mobile robot can have the capacity of path finding and obstacle avoidance by using a single monocular camera. The rest of this chapter is organized as follows: The concept of the system is described in section 3.2. Section 3.3 goes through the detailed algorithm of ground plane and obstacle detection. The calibration experiment of distance measurement is presented in section 3.4. Section 3.5 shows the simulation result using synthetic images. Experimental results using mobile robot will be reported in Chapter 5.

3.2 System Overview

In order to estimate range information for mobile robot navigation control using merely monocular image sequence, the basic idea is to find the world coordinates of pixels belongs to the ground plane from the image plane by applying IPT. For common indoor environments, we can make the following assumptions: (1) the ground is a plane; (2) the robot moves parallel to the ground; and (3) the camera is fixed on the robot. The distance between the camera and the ground plane is thus fixed and the IPT of the ground plane can be determined beforehand. With the pinhole camera model, applying IPT can be simplified by using plane-to-plane homography [38], as shown in Fig. 3-1. There exists a 3 by 3 homography matrix **H** such that

$$\vec{p} = \mathbf{H}\vec{q}, \tag{3.1}$$

where $\vec{p}$ is the coordinate of the ground plane relative to the camera, and $\vec{q}$ the coordinate of image plane. The homography matrix **H** can be found from the relative coordinates of four points on the ground plane and their corresponding coordinates in the image. In the calibration phase, several sets of corresponding points are used to estimate a precise **H**. After calibration,

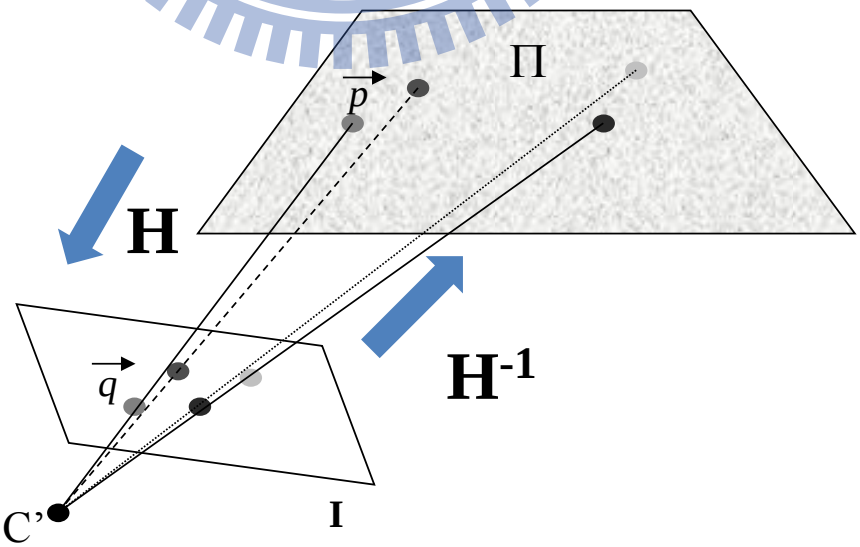


Fig. 3-1: Homography of coplanar points between image plane and ground plane.

the projected images can then provide the relative position of each pixel on the ground in world coordinates [33,34]. Off-ground objects, on the other hand, will be distorted after the transformation. A recalibration will be necessary if the position or angle of the camera is changed.

Due to the assumptions, IPT can only map pixels on the ground plane. In practical obstacle avoidance applications, pixels which belong to the ground plane must be found on the fly in order to determine distance to obstacles. The proposed system therefore aims to solve this problem via multiple visual clues. Fig. 3-2 illustrates the architecture of the proposed monocular visual navigation system (MonoVNS). In this design, two parallel procedures are launched simultaneously after the image preprocessing step. In the ground feature detection process, the system extracts all the features in current image at time step t, and matches them to features observed in previous N frames. N is determined by the frame number that can be processed in one second. According to the correspondence and pre-determined ground features, ground homography and features on the ground in the newly acquired image can be found. The system then uses the homography matrices among consecutive frames to classify all other features as on the ground or not. Meanwhile, in a parallel process, images are segmented into patches according to their color homogeneity. These segments are then classified into ground or off-ground regions by comparing with warped images. The detailed algorithm is described in Section 3.3.

Once the ground and off-ground areas are segmented, the coordinates of the ground area (relative to the camera) can be estimated by applying IPT to the acquired images. The images are virtually rectified with a downward-looking pose. The range information can then be obtained by a ray-casting operation. Ray casting is a procedure to mimic a laser range finder (LRF) using the rectified image. A set of virtual rays are casted from the camera center, one per pixel, until the pixel belongs to an off-ground area, as shown in Fig. 3-3. The distance of a

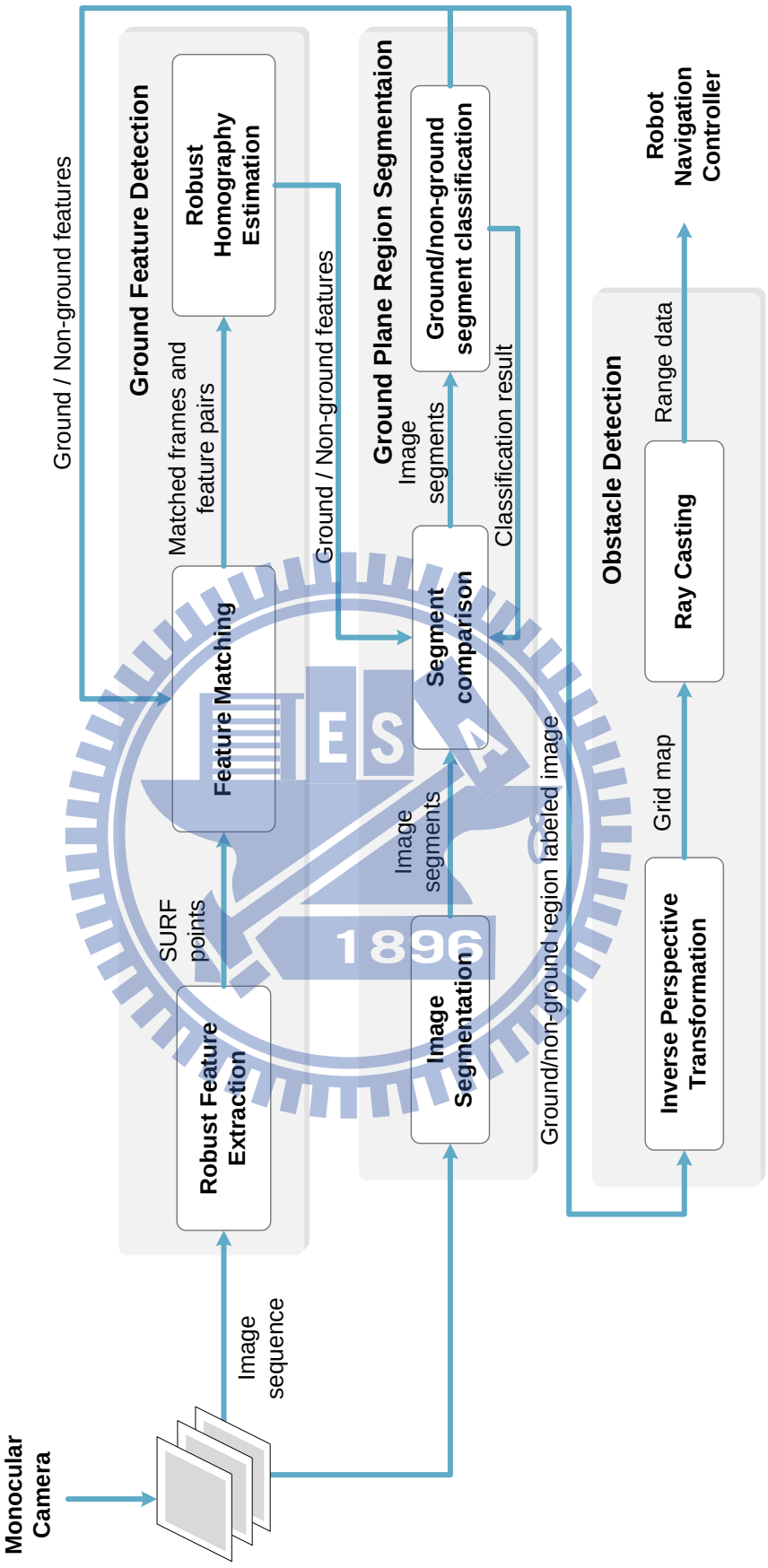


Fig. 3-2: The architecture of the proposed MonoVNS.

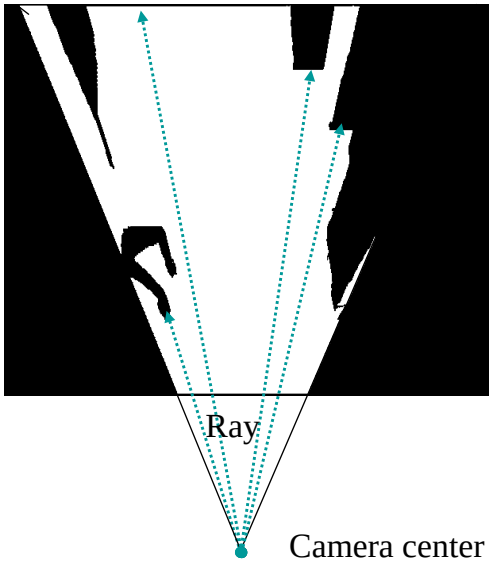


Fig. 3-3: An example of ray casting in an image after IPT. The white area indicates the ground region. The dark area indicates obstacles. The distances between the camera center and obstacles can be estimated by measuring the length of the rays.

ray to an object can thus be estimated by the two-norm of the pixel. The distances of these rays now represent the distances to obstacles, just like a typical LRF, but with a smaller view angle. This information can be applied to any LRF-based obstacle avoidance scheme, and make it to work in a similar manner.

3.3 Ground Plane and Obstacle Detection

3.3.1 Ground Feature Detection

In order to obtain robust matching results of corresponding points, the feature extractor should have properties such as invariance to rotation, scale, illumination and image noise. In previous related works, Harris and FAST [39] corner detectors have been widely adopted [28-31]. The robustness of these descriptors is relatively limited in practical applications. In last decade, there have been many advances in scale/rotation-invariant feature extractors, such as Scale-Invariant Feature Transform (SIFT) [40]. A An array of image gradient histograms has been used as a descriptor instead of a raw image patches. However, the heavy computation

load of SIFT algorithm makes it unsuitable for real-time navigation applications. In this work, a combination of robustness and execution speed has been investigated. In the current design, the FAST corner detector is first applied to find interesting feature points. These points will be further described with Speeded Up Robust Features (SURF) [41] for its superior robustness and execution speed. FAST provides a high speed corner detection performance, which can reach full-frame feature detection up to 400Hz. While the robustness can be preserved, SURF is about three times faster than SIFT in feature extraction step. SURF also provides an indexing value for fast matching purpose. Nevertheless, these methods may still fail under excessive motion blur or abrupt changes in view angle. The speed of the robot and therefore the camera motion is limited accordingly in this study.

Once features are found and matched, the next step is to determine whether these feature points are on the ground. Consider that a ground plane is projected into two views taken at different positions as shown in Fig. 3-4. With the pinhole camera model, two views of the same plane are related by a unique homography [20]. That is, for a plane $\Pi = [\mathbf{n}^T \ 1]^T$, a ray corresponding to a point p in image **I** and its corresponding point p' in image **I'** meet at point

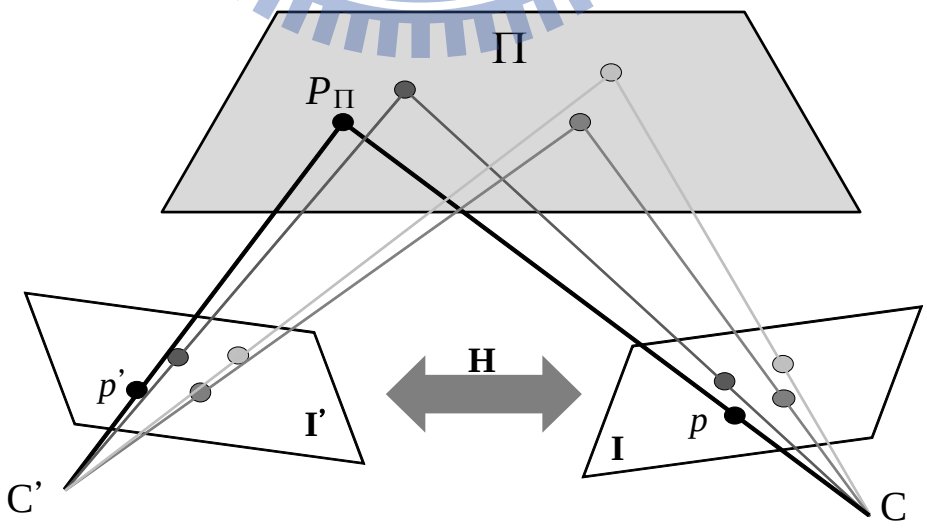


Fig. 3-4: Homography of coplanar points observed from two images.

P_Π in plane Π. Therefore, if a set of coplanar points p_i with homogeneous image coordinates $(x_i, y_i, 1)^T$ are found, and their correspondences $\{p_i \leftrightarrow p'_i\}$ in two images are also found, there exists a 3 by 3 homography matrix **H** such that

$$p'_i \rightarrow p_i = \mathbf{H}p'_i \tag{3.2}$$

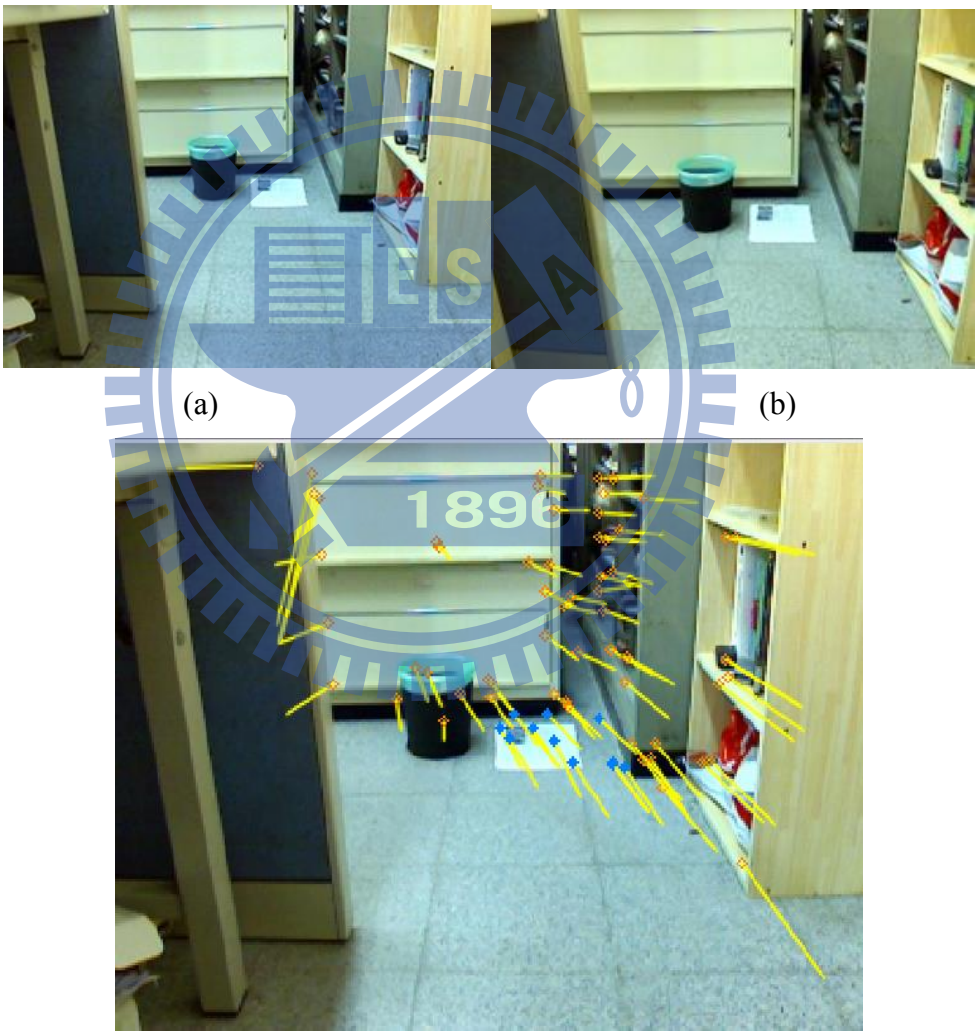
$$\mathbf{H} = \alpha\mathbf{K}(\mathbf{R} + \frac{1}{d}\mathbf{t}\mathbf{n}^T)\mathbf{K}^{-1} \tag{3.3}$$

where **K** is the intrinsic camera parameter matrix, **R** is the rotation matrix, **t** is the translation vector, α is the scale factor, and d is the distance between the plane and the camera. To determine **H**, four non-degenerated corresponding points are required since each point correspondence provides two independent constraints. In the proposed system, the features on the ground are initially determined, i.e., a subset of p_i is known as on the ground plane. The homography of the ground plane can thus be determined initially. To further reduce possible matching errors, we apply RANdom SAmple Consensus (RANSAC) to eliminate outliers and robustly determine **H** [42-43]. Note that the **H** matrix here is only relevant to find a robust plane relationship between two frames. It does not affect the homography matrix applied in IPT computation. Once the homography matrix is determined, the rest of the corresponding points can also be determined if it is on the ground by using the back projection technique such that:

$$p'_i \in \begin{cases} \text{ground, if} \left\| p'_i - \mathbf{H}p_i \right\| < threshold \\ \text{off - ground}, otherwise \end{cases} . \tag{3.4}$$

The factor of determining the threshold depends on how accurate the ground feature classification should be. Since the mobile robot can run over certain small obstacles, there is a small range of tolerance for the feature classification. This facilitates the mobile robot to navigate on somewhat uneven surfaces. A fixed value is determined beforehand in the current

implementation. Fig. 3-5 illustrates an example of ground feature classification from two images. Note that homography estimation has its limitation. For instance, in case of near-zero camera translation, i.e., $\mathbf{t} \cong 0$ in (3), no information on coplanarity can be inferred since the plane normal **n** can be arbitrary [30]. This condition must be checked since in this case all the points will be determined as ground. A simple test is to determine whether all observed points conform to the same homography. This test will fail only when there is just a single scene plane visible, or when camera intrinsic parameters are changed. Both of these cases can be



(a) (b)

(c)

Fig. 3-5: An example of ground feature detection. (a) and (b) are images taken in different views. (c) is the classification result of the matched features observed from (a) and (b). The features with blue label indicate features on the ground, while the features with orange label are features off the ground.

avoided in the current design.

3.3.2 Ground Region Segmentation

In the Section 3.3.1 feature points which belong to the ground plane are determined. Pixels that are not selected as feature points, however, are still needed to be classified if they are on the ground. In this work, this problem is resolved by segmenting the image into regions beforehand and classifying these regions on the ground afterwards.

The quality of segmentation depends not only on the algorithm used, but also the targeted application. In this design, the segmentation method should be able to distinguish objects from image frames in a fast and robust way under various environmental conditions. In the current design, an image frame is segmented by adopting multi-scale Mean-shift algorithm in HSI (Hue-Saturation-Intensity) color space. Mean-shift algorithm [44] is a nonparametric, iterative clustering technique which does not require prior knowledge of the number of clusters and constrain the shape of the clusters. It is therefore suitable for unsupervised color segmentation. The proposed multi-scale Mean-shift algorithm is summarized as follows:

Step 1. Choose a search window with proper bandwidth.

Step 2. Choose the initial location of the search window.

Step 3. Compute the mean value (centroid) of the data in the search window.

Step 4. Move the center of the search window to the mean location computed in Step 3.

Step 5. Repeat Step 3 and 4 until convergence.

Note that in step 1, the bandwidth of the kernel needs to be determined. In particular, the image is under-segmented when the bandwidth is too large, and over-segmented when the bandwidth is too small. In this work, proper bandwidth is determined dynamically, according to the frequency analysis results of the image [45] . For instance, a clustered image such as

crowds often indicates larger energy in high frequency, and requires a smaller bandwidth value. On the contrary, a simple image such as a white wall, will give a larger bandwidth value. Additionally, since the Hue values of pixels are unstable under low intensity and saturation, these *colorless* pixels should be segmented separately. In the current implementation, the hue value of a pixel is set to 2, if this pixel has a saturation value<0.1 or an intensity value <0.1.

The proposed segmentation algorithm takes 0.1 to 0.5 seconds to process an image of the size of 640 480 pixels. In order to further boost the speed for real-time applications, images are scaled down to one-tenth beforehand. Small objects may be neglected in this case. Therefore, the proposed algorithm estimates the purity of each segment with original resolution, and segments again if needed. The modified method is on average 10 times faster than that using the original scale.

3.3.3 Obstacle Detection

After ground/off-ground features are determined, and the image is segmented into regions, it is now possible to classify the ground region in a pixel-wise manner. As mentioned in Section 1, many previous works warp the image by using the homography matrix and calculating the Sum of Absolute Differences (SAD) between the warped or rectified image and current image. However, it is difficult to determine a proper value of threshold since the SAD value is correlated to the environment. Furthermore, homogenous obstacles may be neglected. The proposed system determines if the region is on the ground according to the displacement and the distortion of each segment separately. Similar techniques have been used in stereo vision when estimating disparity maps [46]. To do so, corresponding segments need to be found. While segments contained already matched feature points can be found easily, homogenous segments are matched using both feature points and its color distribution. The

proposed segment matching method considers finding the maximum of the overall likelihood function of multiple cues, that is,

$$L(a \mid b) = L_{color}(a \mid b) L_{motion}(a \mid b) \quad . \tag{3.5}$$

The color likelihood model is defined in a way to estimate the similarity between the color histogram of two segments. Bhattacharyya coefficient has been widely used to determine the similarity of two discrete probability distributions [47]. Within the interval [0, 1], the larger the coefficient is, the more similar the two histograms are. Assuming two objects a and b with color histogram $h_a = \{h_{a,1} \ldots h_{a,N}\}$ and $h_b = \{h_{b,1} \ldots h_{b,N}\}$, the color cue likelihood function between two object is therefore defined by

$$L_{color}(a \mid b) = \rho(h_a, h_b) = \sum_i \sqrt{h_{a,i} h_{b,i}} \quad . \tag{3.6}$$

The motion likelihood model is defined in a way to estimate the Euclidian distance between two segments. Assuming the centroid of segment a is (x_a, y_a), and the predicted centroids of (x_a, y_a) b is (x_b, y_b), the motion cue likelihood function between two object is therefore defined by

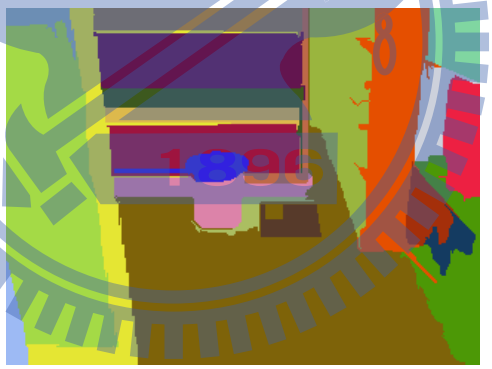
$$L_{motion}(a \mid b) = \frac{1}{\exp(\|(x_a, y_a) - (x_b, y_b)\|)} \quad . \tag{3.7}$$

Those segments which fail to be matched will be labeled as undetermined. Finally, each matched segment is then assigned an initial probability value 0.5. The probability value of each segment is updated on the basis of SAD. These segments are then classified as ground or off-ground. Both static and moving obstacles can be observed. Fig. 3-6 shows both the result of segmentation and ground plane labeling based on the images in Fig. 3-5. Note that in the image a piece of paper on the ground is classified as ground as expected.

In summary, The proposed ground plane detection algorithm is able to segment ground regions from two images. The feature classification step determines the features on the ground planes, while also finding the homography matrix **H** between the ground planes in two images. The second image can thus be rectified with **H**. These two images are also segmented using color homogeny. These segments are matched with multiple cues, and finally determined as ground or off-ground. In practice, the key image for matching is not merely obtained from a single frame, but also a collection of tracked features and segments among N frames that can be processed in previous one. Therefore, this method can be applied even when the robot is rotating, or the camera is temporally occluded.

3.4 Calibration Experiment of Distance Measurement

In this experiment, distance measurement of the proposed MonoVNS was calibrated. The

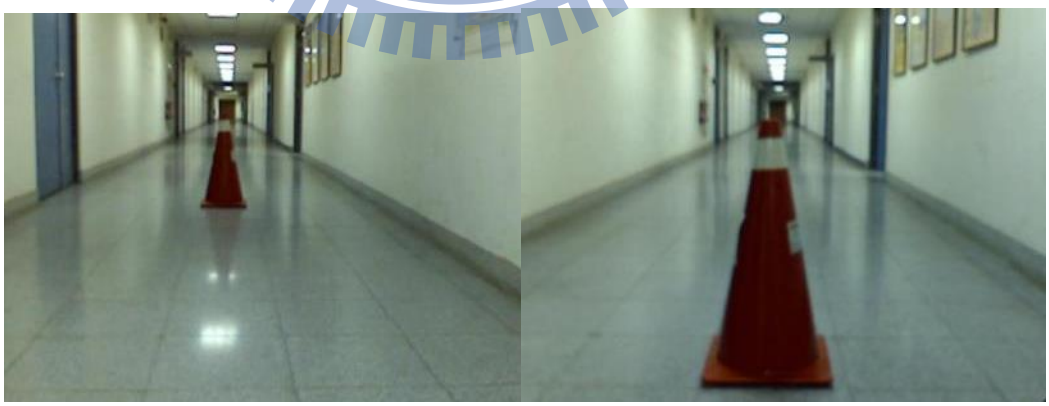


(a)

(b)

Fig. 3-6: (a) Color segmentation result and (b) ground region labeling result based on the images in Fig. 3-5.

experiment was performed with a static object on the ground, while the camera was positioned to take snapshots and measure the distance to the object, as shown in Fig. 3-7. The snapshot was taken from a distance from 1.6m to 6m, at the interval of 0.4 m. The estimated distance was compared with the ground truth values. To examine the influence of scaling in segmentation step, we experimented with both the original images and the scaled images. The IPT homography was initially estimated by manually assigning known points on the ground within three meters to the camera. The acquired original image was 320×240 pixels, while the scaled image was 32×24 pixels. The average processing time using the original image is 251ms, while that of using the scaled image is 190ms. The processing time varies depending on the extracted features and the complexity of input imagery. Table I shows the calibration results. The measurement error is within 3% in a range of 4.8m, even if the image is scaled down to one-tenth. The reasons of the error are mainly due to the imperfect result of segmentation and the quantization error in IPT. These errors are especially significant when the object is far away and cannot be distinguished due to limited resolution of the imagery. Note that the minimal and maximum sensing ranges are both related to the height and the tilt angle of the camera.



(a) (b)

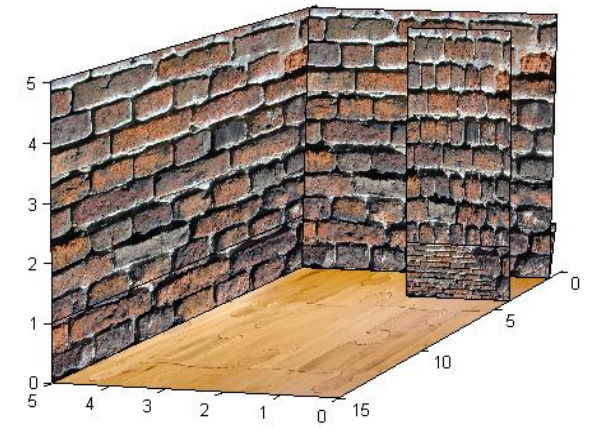
Fig. 3-7: The obstacle used in the experiment. (a) the robot is 6m away from the obstacle (b) the robot is 1.6m away from the obstacle.

Table 3-1: Experimental Result for Distance Measurement by Using the proposed MonoVNS

Ground truth(m)	Estimated distance(pixels)		Estimated distance(m)		Error(m)		Percentage error (%)	
	Original	Scaled	Original	Scaled	Original	Scaled	Original	Scaled
1.6	45	4	1.63	1.6	0.03	0.0	1.94	1.7
2	100	10	1.97	2.0	0.03	0.0	1.41	0.1
2.4	133	13	2.42	2.4	0.02	0.0	0.66	0.0
2.8	158	15	2.87	2.7	0.07	0.1	2.57	1.8
3.2	176	17	3.26	3.2	0.06	0.0	2.01	1.2
3.6	193	19	3.68	3.6	0.08	0.0	2.32	1.0
4	206	20	4.04	3.9	0.04	0.1	0.92	2.5
4.4	220	22	4.45	4.5	0.05	0.1	1.09	1.6
4.8	232	23	4.83	4.8	0.03	0.0	0.55	0.4
5.2	238	23	5.02	4.8	0.18	0.4	3.38	8.1
5.6	243	24	5.19	5.1	0.41	0.5	7.25	8.8
6	247	24	5.33	5.1	0.67	0.9	11.1	14.9

3.5 Simulation results of Synthetic Images

In order to test the proposed system under different environments, a computer synthetic scene as shown in Fig. 3-8(a) was established by using Matlab to examine the classification rate with a known ground truth. The synthetic scene contains three walls, with the same texture. In each trial of the experiment, the path of the robot is randomly generated to observe 10 consecutive views. The texture of the walls also changed randomly in each trial. Fig. 3-8 (b) shows an example of the classification result. The performance of the algorithm is evaluated by calculating the false detection rate of the number of features and the area of regions. False detection includes false positive- the features or regions on the ground are determined as off-ground, and true negative- the features or regions off-ground are determined as ground. The experiment after 100 trials shows that the maximum false detection rate of features is 0.96%. The average false detection rate of features is 0.84%. The experiment also shows that



(a) (b)

Fig. 3-8: The synthetic room experiment. (a) The structure of the environment. (b) A classification result. The red zone indicates the region correctly labeled as ground, while the yellow region indicates false detection.

the maximum false detection rate of region is 2.91%. The average false detection rate of region is 2.08%.

3.6 Summary

A monocular vision-based robot navigation system has been developed by combining IPT and ground region segmentation. Useful steps in image processing have been developed and integrated into the ground plane detection algorithm to enhance the robustness of the system. To fulfill the requirement of real-time performance, we propose to combine a corner detection algorithm with SURF feature descriptor such that a scaled down image can give acceptable accuracy in distance measurement. Experimental results show that the system has adequate

accuracy of 1% for distance measurement in a range of 6 meters. A local grid map can thus be built with single camera for robot navigation. The proposed MonoVNS can be an alternative to popular laser range finders and RGB-D sensing devices like Kinect for robot navigation applications. The merit of this approach is to have both obstacle avoidance and path finding functions with a low cost monocular camera.

Chapter 4.

Location Aware System Design

4.1 Introduction

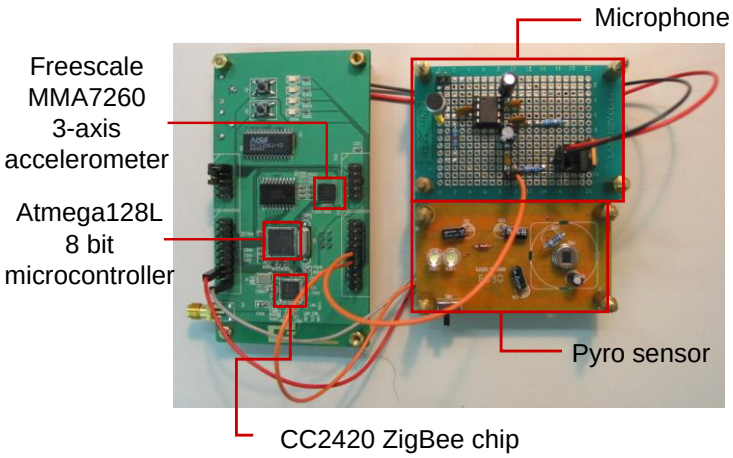
In this chapter, the detailed implementation of intelligent environment and mobile robot is described, including the hardware and software module. The implemented system aims to provide intruder detection and user monitoring services. The implemented system will be applied in the experiment reported in Chapter 5.

4.2 Intelligent Environment

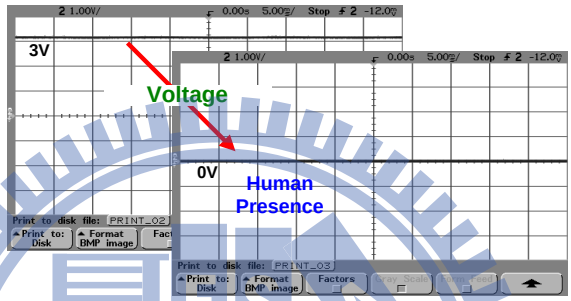
The developed intelligent environment consists of three main sub-systems: ZigBee-based intrusion detection system, multimedia Human-machine interface and OSA platform. The intrusion detection system can sense abnormal states around the sensor and report the location. The video server controls all the cameras in the house to acquire images. The selected camera is able to take pictures from the elderly according to the location information provided by the location server. These information will be sent to users through an Open Service Access (OSA) Gateway[48]. In addition to passive monitoring, the elderly can also ask for help via the speech recognition system. His/her location can be found and image can be sent to the user B in real time.

4.2.1 Intrusion Detection Sensors

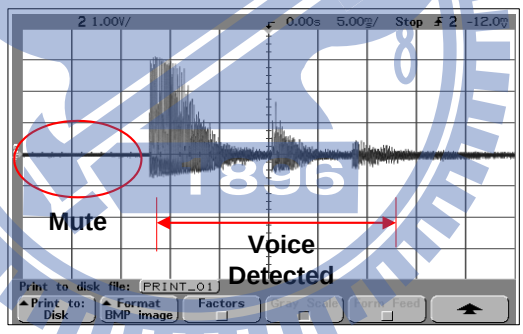
We designed several useful sensor modules for intrusion detection to verify the effectiveness of the proposed algorithm [49-50]. Fig. 4-1 shows such an intruder detection



(a)



(b)



(c)

Fig. 4-1: Intruder detection module (a) hardware of the module (b) signal of the pyro sensor (c) signal of the microphone.

sensor module. It consists of an 8-bit microcontroller Atmega128L (Atmel Corp.), a ZigBee chip CC2420 (Texas Instruments), and an onboard Freescale MMA7260QT triaxial accelerometer (Freescale Semiconductors). Two other sensor modules, a pyro sensor and a microphone sensor, were also designed and integrated into the experiments. The details of these sensor modules are described below:

a) Triaxial accelerometer

The triaxial accelerometer is applied to detect vibrations. Individual magnitudes of the dynamic acceleration from X, Y and Z axes are first isolated. The summation of these magnitudes can thus be used to detect abnormal vibrations during an intrusion or collision.

b) Pyro sensor

We adopted off-the-shelf pyro-electric sensor components. The pyro sensor can detect the presence of humans in a certain area by a change in the voltage level of sensor output. As Fig. 4-1 (b) shows, if the voltage level is close to zero, the sensor will be triggered.

c) Microphone

A simple microphone was integrated into the ZigBee module for abnormal sound detection. The rising edge of the filtered microphone signal will trigger the interrupt of the chip and can be used to determine loud sounds (Fig. 4-1). The trigger level is set at +2.01 V. A loud sound can be caused by a falling object resulting from an intrusion or the breaking of a window.

If any sensor is triggered, a message including the current time, the ID of the triggered sensor, and the location of the node will be sent to the robot. Since the robot may be far from the location, an adaptive route selection algorithm is applied to send the message by forwarding data packets to the robot via selected sensor nodes [51].

4.2.2 Visualization System

The coordinates estimated and provided by ZigBee sensor are in format of numbers, which is difficult for human users to understand. Therefore in practice a usable system needs to represent the results in a more comprehensible way. The most common method nowadays is to use figure or images such as floor plan or map. However, in order to monitor the current state of the elders, the located elderly will be actually shown on the physical image obtained from a

pan-tilt camera. To achieve this feature, one needs to apply camera calibration technique. Using camera calibration, the projection between world and image coordinates can be obtained. In order to increase the coverage area of the camera, the pant-tilt ability of the camera is also necessary. The system first calibrate the camera and build the camera model, which represents the mapping from the world coordinate to the image plane and the pan-tilt movement of the camera.

The calibration process starts from the geometric camera calibration procedure, which aims to estimate the intrinsic and extrinsic parameters of a static pin-hole camera model. The implemented algorithm is the linear estimation method given in [38]. Assumed that the target environment observed by a camera and the image position $p_i=(u_i,v_i)^T$, of n feature points with known world coordinate vectors $Pi=(x_i,y_i,z_i)^T$ are matched in the image, i =1…n .The projection between world and image coordinates can then be modeled by

$$\mathbf{p}_i = \frac{1}{z_i}\mathbf{M}\mathbf{P}_i \quad \text{,where } \mathbf{M} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}, \mathbf{p}_i = \begin{pmatrix} p_i \\ 1 \end{pmatrix}, \mathbf{P}_i = \begin{pmatrix} P_i \\ 1 \end{pmatrix} \tag{4.1}$$

M is decomposed to **M**=**K**(**R** **t**), where (**R** **t**) represents the extrinsic parameters, including rotation matrix **R** and translation vector **t**. **K** represents intrinsic parameters, including skew and aspect ratio. Collecting the constraints from (2) associated with n pairs points yields a system of 2n homogeneous linear equations in 12 coefficients of the matrix **M**-namely, $\underline{\mathbf{P}}\cdot\mathbf{m}=0$,where

$$\underline{\mathbf{P}} = \begin{bmatrix} \mathbf{P}_1^T & \mathbf{0}^T & -u_1\mathbf{P}_1^T \\ \mathbf{0}^T & \mathbf{P}_1^T & -v_1\mathbf{P}_1^T \\ \cdots & \cdots & \cdots \\ \mathbf{P}_n^T & \mathbf{0}^T & -u_n\mathbf{P}_n^T \\ \mathbf{0}^T & \mathbf{P}_n^T & -v_n\mathbf{P}_n^T \end{bmatrix}, \text{ and } \mathbf{m} = [\mathbf{m}_1^T \ \mathbf{m}_2^T \ \mathbf{m}_3^T] \tag{4.2}$$

As long as $n \geq 6$ (since $\underline{\mathbf{P}}$ is a $2n$ by 12 matrix), the homogeneous linear least-squares can be exploited to compute the unit vector **m** that minimize $|\underline{\mathbf{P}}\mathbf{m}|^2$ and thus estimate **M**. In this system, the *n* correlation points are marked manually from the captured image scene with a GUI interface. To increase the accuracy and avoid outliers, the Random Sample Consensus (RANSAC) algorithm is also implemented. This algorithm picks n points from a large set of sample points randomly and the estimate the matrix iteratively until the fitting error is below the expected criterion. In the experiment, the average error is around 2 pixels.

The second step is to estimate the influence brought by the pan-tilt movement. In the proposed system a pan-tilt camera model, as shown in Fig.4, is used. A rotation around the x-axis, $\mathbf{R}_x$, and the y-axis, $\mathbf{R}_y$, correspond to tilt and pan respectively. With this model, the projection between world and image coordinates can then be modified as:

$$\mathbf{p}_i = \frac{1}{z_i}\mathbf{M}\mathbf{R}_y\mathbf{R}_x\mathbf{P}_i \tag{4.3}$$

The choice of this simplified camera model is a tradeoff for straightforwardness over accuracy. The model assumes that the center of rotation of the camera is fixed and coincides with the lens center of projection during operation, which many cameras may violate such an assumption. However, in this application, the deviation of the center is negligible.

To validate the camera calibration result, eight positions in the experimental room were assigned to emulate the result of the positioning engine. As shown in Fig. 4-2, given the exact coordinate, the camera successful turns and marks the person (in the bathroom) on the image. Thus the localized person can be visualized and the picture can be transferred through the OSA platform. To further exam the result, an experiment of user tracking combining the position engine experiment also performed. The user carried the beacon moved around room, and stopped at one of the eight position assigned in Fig. 4-2. The position engine will estimate the location of the user and the camera will move toward the user using the given coordinate. The

result showed that the user can always be captured by the camera, as shown as shown in Fig. 4-3.

4.2.3 Speech Recognition system

Considering that elderly members are generally few in a family, say around 3 persons, the mode of the speech recognition system is set as speaker-dependent. The command and user index are pre-defined integers to represent the recognition results of the input speech signals.

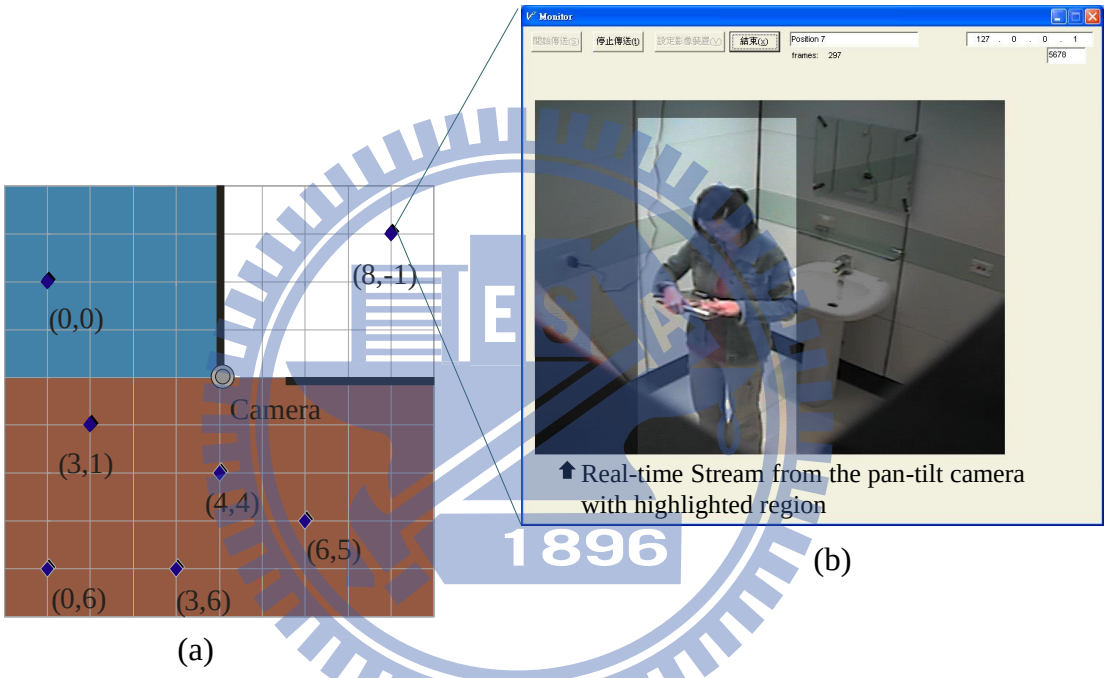


Fig. 4-2: An example of the "target marking" of the visualization system. (a)The map of the experiment environment. (b)The graphic user interface and the view of the pan-tilt camera at coordinate (8,-1)

Fig. 4-3: Snapshots from the experimental result of the "target marking" of the visualization system. (a) the person was at the coordinate (0,0) (b) the person was at the coordinate (0,6) (c) the person was at the coordinate (8,-1)

The recognition algorithm is based on the dynamic time warping (DTW) method [52]. The procedure of the speech recognition is divided into three parts.

The first part is the process of speech information sampling, which is the setting of gain control and sampling frequency when analog signals transforms to digital signals. The second part is the extraction of speech features. The endpoint detection process can determine the location of real speech signals by the short time energy detection and zero-crossing rate detection. The started 128 samples are used to determine the threshold value of the energy detection and zero-crossing rate detection. When the real speech signals are determined, they are divided into a 16ms length frame. Each frame has an overlap region. In order to do the encroachment of the high frequency spectrum, the pre-emphasis work is adopted. Considering the continuity of signals at two sides of a frame; Hamming window is used on every frame. It can hold the signals at mid parts and press down the signals at two sides. The Mel Frequency Cestral Coefficients (MFCCs) is extracted for every frame. A feature vector represents each frame. The DTW recognizer compares the input feature vectors with the reference speech and gets a minimum matching error as the recognition result. The third part is the process of feature matching by the dynamic time warping method. The speech recognition system is implemented in a TI TMS320VC5402 DSK board.

4.2.4 OSA platform

The OSA API is an open, standardized interface for applications to use the capabilities of a network without owning it or knowing its technology. It consists of a framework, which is in charge of access control and service discovery; and some Service Capability Features, which map to network capabilities. It is specified and standardized in the Joint API Group, with participation of 3GPP, ETSI and the Parlay Group: a single API for the whole developer community.

Parlay/OSA [48] enables operator and 3rd party applications to make use of network functionality through a set of open, standardized interfaces. This leads to several advantages such as shorter time to market (TTM) for applications, network independency, etc. Parlay/OSA Gateway consists of several Service Capability Servers (SCS): functional entities that provide Parlay/OSA interfaces towards applications. Each SCS is seen by applications as one or more Service Capability Features (SCFs): abstractions of the functionality offered by the network, accessible via the Parlay/OSA API. Sometimes they are also called services. The Parlay/OSA SCFs are specified in terms of interface classes and their methods.

In this system, the OSA gateway is connected with the location and visualization system in order to extend the service to mobile network. Fig. 4-4 shows components applied in this system. The OSA application server receives the request from users using mobile terminals, and request data from the OSGi gateway at home. The data can then be sent to the user through two different paths: by e-mail via the internet, or by MMS via cellular network.

4.3 Mobile Robot System

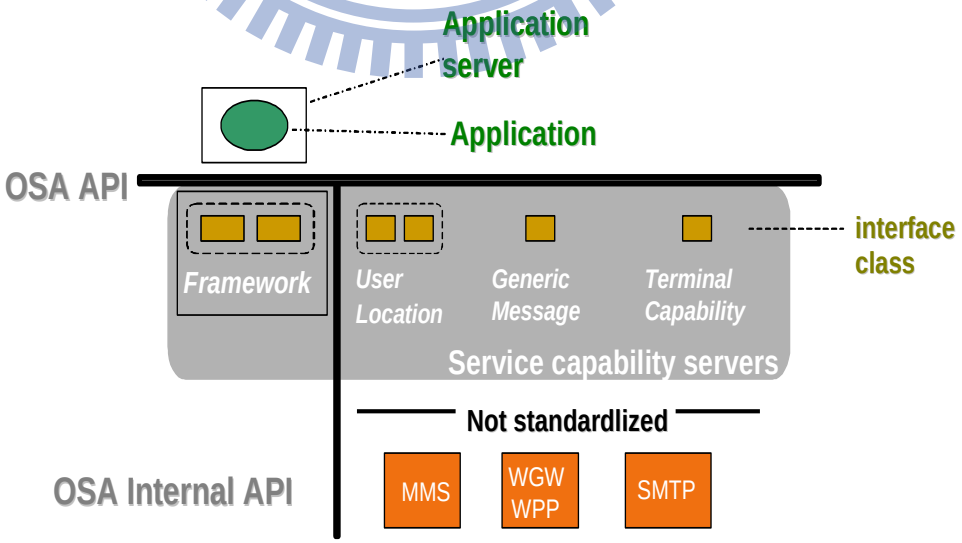


Fig. 4-4: OSA network services applied in the proposed system

The proposed location aware system is integrated and tested with an on-demand intruder detection system which features a ZigBee WSN and a mobile robot. If any intruders or abnormal conditions are detected, the state and location of this alert will be transmitted to the security robot and monitoring center via the WSN. The robot can navigate autonomously to where the alarm takes place using the probability-based localization system. After the robot arrived at the spot, the onboard camera will transmit real-time images to the user via both WiFi and 3G networks. The security guards and end-users can therefore easily determine the exact situation in real time.

4.3.1 Agent-based Control System

An agent-oriented robot control (ARC) architecture [53] was adopted to handle the multi-modal information fusion and motion control of the robot. ARC is a distributed hybrid system, which couples reactive and local deliberative agents. In this architecture, a router manages the inside communication of the robot, while other components handles the real-time performance of the robot behaviors. Each individual agent is responsible for sensing, actuating, and executing a set of functions. A task manager will integrate all the information and guide each agent with proper parameters.

4.3.2 Mobile Robot Localization

In the intelligent environment, several reference nodes were deployed beforehand in the environment to localize the robot using the localization algorithm in Chapter 2. Their positions were stored in a database on the robot's onboard computer. Two kinds of messages are defined for the ZigBee WSN: the measurement-demand and signal-report. The measurement-demand message is used for a mobile node to request the RSSI measured from reference nodes. This message is broadcasted to all the nodes which are able to receive it. Furthermore, a counter is included in this message to keep track of various measurement demands. The signal-report

message is used by the reference nodes to report the measured RSSI values. In summary, location estimation is performed using the messages and the following process:

1) A measurement-demand message is broadcasted to sensor nodes from a mobile node.

2) Each sensor node measures RSSI at the time it receives the packet. Subsequently, it transmits the RSSI, counter number and the mobile node ID to the robot.

3) The robot collects all the data, separates each by the ID and counter. If three or more RSSI values with the same counter are observed, the location of the robot can be estimated and updated using the proposed algorithm. In the current implementation, the entire procedure can be accomplished within 600ms.

4.3.3 Autonomous Navigation System

Autonomous navigation in a dynamic changing environment is essential for a robot patrolling from one place to another. In the proposed system, this task is achieved by a behavior-fusion approach adopted from [53, 54]. Fig. 4-5 illustrates the architecture of navigation system using behavior fusion. We designed three primitive behaviors for autonomous navigation in an indoor environment. These behaviors are goal seeking behavior, wall following behavior, and obstacle avoidance behavior. Goal seeking behavior is treated as attempting to move toward the direction of the target. Wall following behavior is defined as maintaining the same distance from the nearest wall. The obstacle avoidance behavior is designed so that the robot moves in a direction away from any obstacles. These behaviors will each output a set of desired wheel velocities.

The concept of pattern recognition was adopted to fuse the output of navigation behaviors. The pattern recognition technique was used to map the environmental configuration, obtained from ultrasonic and infrared sensors onboard the robot, to the fusion weights of three navigation behaviors. The concept can be expressed as “When the environment configuration

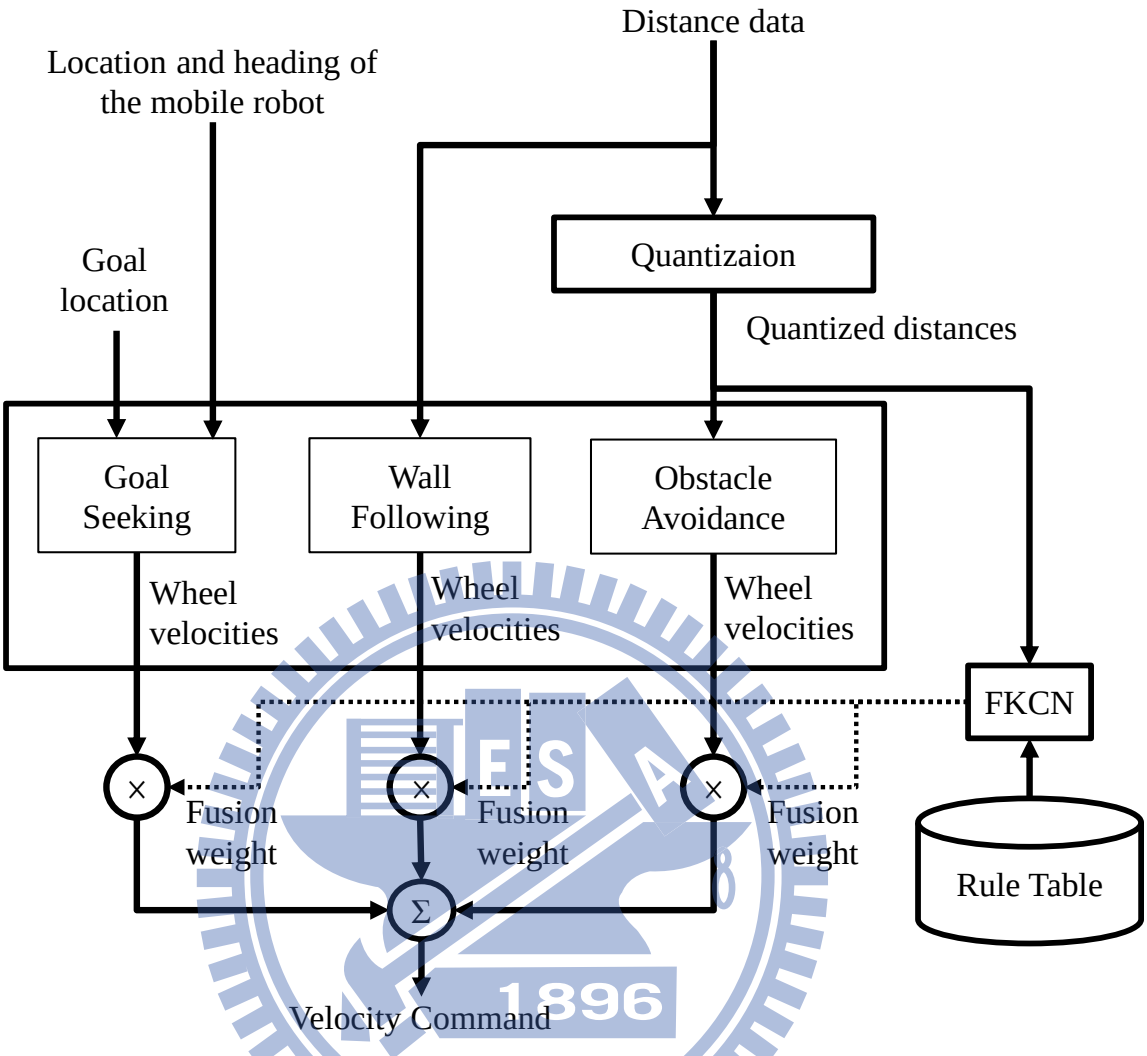


Fig. 4-5: System architecture of autonomous navigation using behavior fusion design..

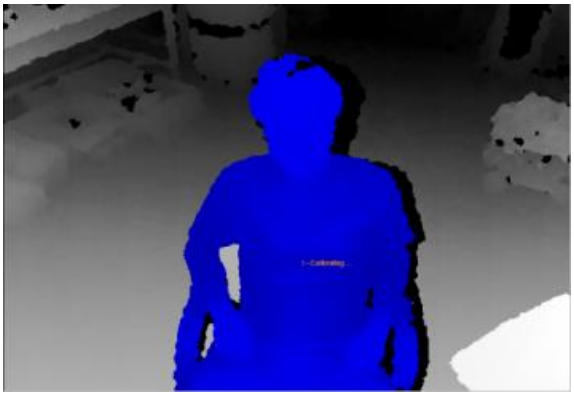
of surroundings is similar, the fusion weights of navigation behaviors should be similar." In this work, the Fuzzy Kohonen Clustering Network (FKCN) is applied to map the environmental patterns to fusion weights [54] . In this design, ten prototype patterns defined by the quantized distances to objects around the robot were set manually to represent typical environmental configurations for indoor navigation. Suitable fusion weights for each navigation behavior corresponding to these patterns were assigned in the rule table of FKCN. When the robot navigates through the environment, the sensors will obtain current range data around the robot. Then the FKCN works to generate proper fusion weights for each navigation behavior, corresponding to the immediate environmental sensory data pattern. As a result, the

mobile robot can navigate to the desired location without colliding into any objects in the environment. In this application, the positions of both the robot and the goal will be determined by the location aware system.

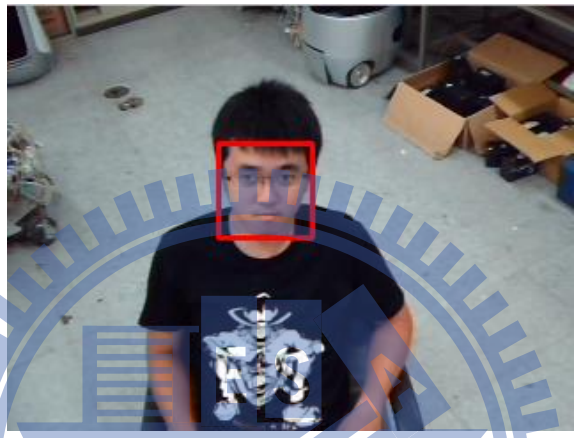
4.3.4 Human Detection and Tracking

The current design of location aware system is based on the received signal strength indicator (RSSI) of ZigBee wireless sensor networks. However, due to limited accuracy of the RSSI-based localization method, the robot cannot rely on merely RF information. A second method is required as the robot approaches to the user. In this implementation, a vision-based human detection and tracking system is integrated with the RSSI-based localization system to allow the robot approach the user more precisely.

In the current design, the human detection system includes two schemes: the body detection and face detection. Both methods are exploited to search and track the user as the robot approaches the user, as shown in Fig. 4-6. If a user is detected, the location of the user can be estimated by combing the detection result and the localization method proposed in Chapter 2. The robot is therefore able to track the user. In particular, the face detection can ensure that the user is facing the robot. For a robot with RGBd camera, the body detection method is used in this design based on the implementation of OpenNI[55]. The body detection system can not only identify the body, but also able to track the skeleton of the body from the depth image of Kinect. The posture of the human can therefore be recognized and the robot can adopt its pose to fit the user while delivering an object. The face detection method, on the other hand, is adopted from the functions in OpenCV[56]. The skin color segmentation is applied before the detection step to further increase the robustness



(a)



(b)

Fig. 4-6: Examples of the human detection system: (a)human body detection, (b)face detection

4.4 Summary

This chapter describes hardware and software modules of the location aware system developed and implemented in this thesis. The developed intelligent environment consists of three main sub-systems: ZigBee-based intrusion detection system, multimedia Human-machine interface and OSA platform, which provides the user to be alerted and monitor the house via camera or the robot from their mobile devices. The implemented mobile robot system is able to autonomous localize itself, navigate the environment, and detect human during the navigation to provide service.

Chapter 5.

Experimental Results

5.1 Introduction

In this chapter, several experimental results using the proposed method and implementation from Chapter 2 to 4 are presented and discussed. In Section 5.2, the experiment of Human-machine-Interface in Intelligent environment shows how the human can monitor the intelligent environment via speech command. The monocular visual navigation system is demonstrated in section 5.3. The experiment of intruder detection application which integrates the location aware system and intelligent environment is shown is section 5.4. Finally, The experiment of robot-on-demand service is presented in section 5.5.

5.2 Human-Machine-Interface in Intelligent Environment

In order to evaluate the capability and practicability of the system, the system is tested in a sample room. The pan-tilt camera DCS-5300G from D-Link was installed in the middle of the room. Nine ZigBee modules were placed around the room, while three users carrying the ZigBee beacons. The integrate system has been tested to exam the scenarios described in the first section: location request and “asking-for-help”. In the first scenario, several people with ZigBee beacons move around some room simultaneously. A user then requests to watch these people from the cell phone. The camera successfully moves to the target person and the user retrieves the target image for each request successfully. In the second scenario, several people wore a headset microphone, as shown in Fig. 5-1. They took turns saying their name, to pretend they were in a situation to ask for help. The speech recognition system successfully

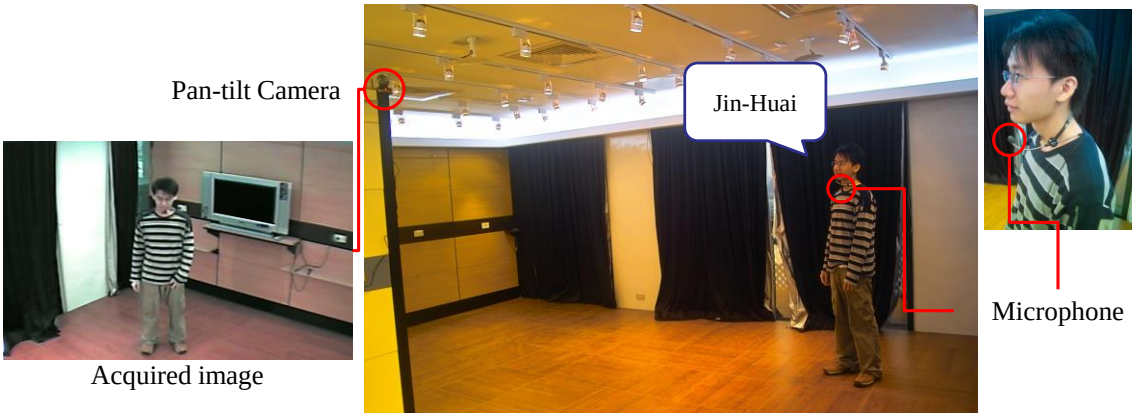


Fig. 5-1: An example of the “asking-for-help” scene

recognized these people, translated the messages to the location server, and launched the camera to rotate and capture image of the person. The images successfully transmitted to other users’ cell phones and PC via e-mails.

5.3 Mobile Robot Obstacle Avoidance

This experiment aims to test if the proposed system can be used to replace a common laser range finder and accomplish an autonomous navigation task. The proposed system has been implemented on an experimental mobile robot, as shown in Fig. 5-2. An Industrial PC (IPC) was used for image processing and robot control. All the motors are velocity-servoed by using a DSP-based motion control board. The navigation system described in Section 4.3.3 is adopted in this experiment, as shown in Fig. 5-3. The robot localized itself by using odometry. The navigation task was achieved by behavior fusion of two navigation behaviors, namely, obstacle avoidance and goal seeking. The goal seeking behavior works to drive the mobile robot towards the direction of the target. The obstacle avoidance behavior guides the mobile robot to move along the direction away from any obstacle to prevent from any possible collision on the way to the target. The main challenge of the proposed vision-based system is that the original laser scanner provides a distance canning of 240-degree while the adopted web camera, however, can provide only 60-degree scanning, as shown in Fig. 5-4, due to the camera’s field of view. Therefore, the robot may not move too fast to collide into obstacles

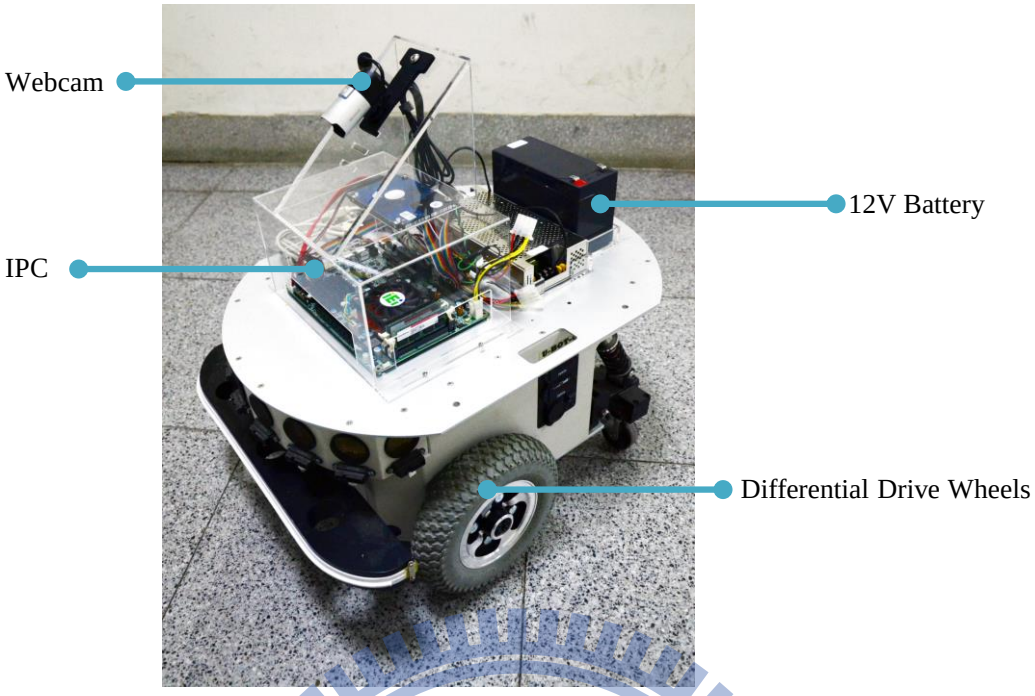


Fig. 5-2: The mobile robot used in the navigation experiment.

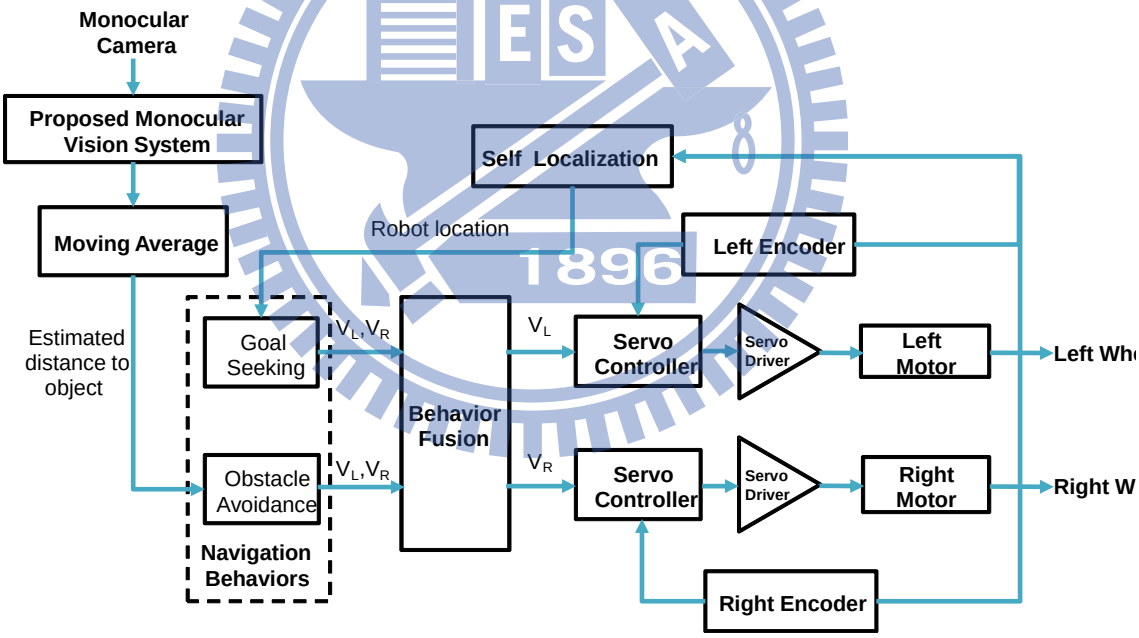


Fig. 5-3: The navigation design of the experimental mobile robot

which appear in the blind zone. In the experiments, the threshold value in (3-4) was set to 2 pixels, considering the current implementation for the laboratory environment.

The experimental results are shown in Fig. 5-5 and Fig. 5-6. In the experiment, the robot was programed to navigate from coordinates in meters (0,0) to (13,1) via (2,2.5) and (12,2.5).

The initial speed of the robot was 25cm per second. The recorded trajectory in Fig. 5-5 shows that the robot successfully avoided several static and moving obstacles and reached the destination. Note that the robot is able to detect a moving pedestrian, as shown in Fig. 5-6 (f), Fig. 5-5. (f) and Fig. 5-7 (a), and also detect the homogeneous ground region, as shown in Fig. 5-6 (h), Fig. 5-5. (h), and Fig. 5-7(b). These two scenarios validate the effectiveness of the proposed obstacle detection method mentioned in Chapter 3.

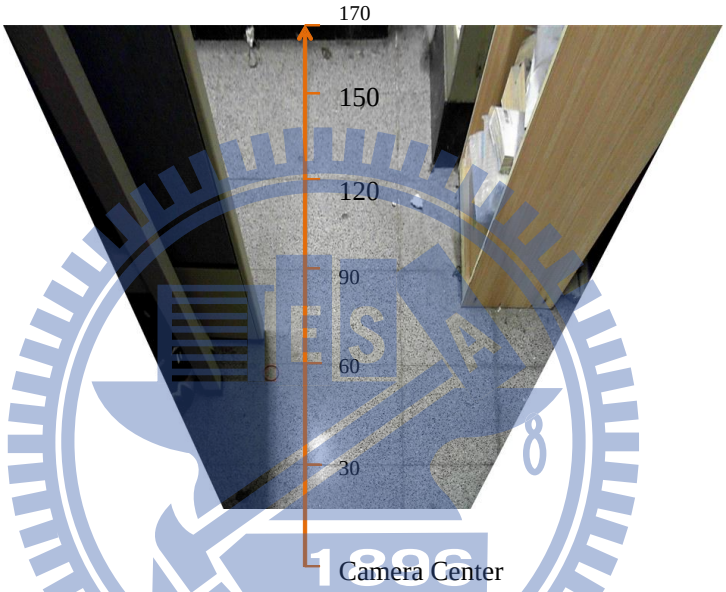


Fig. 5-4: The scale of the view from the mobile robot after IPT.

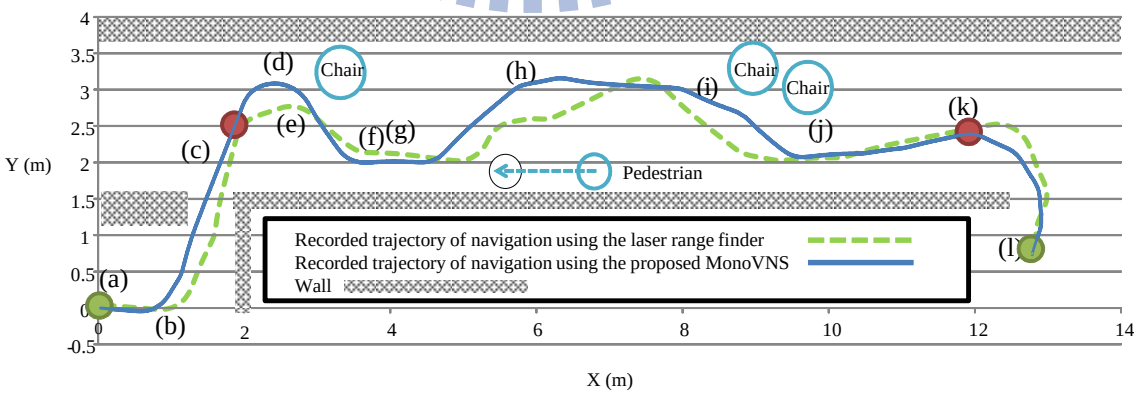
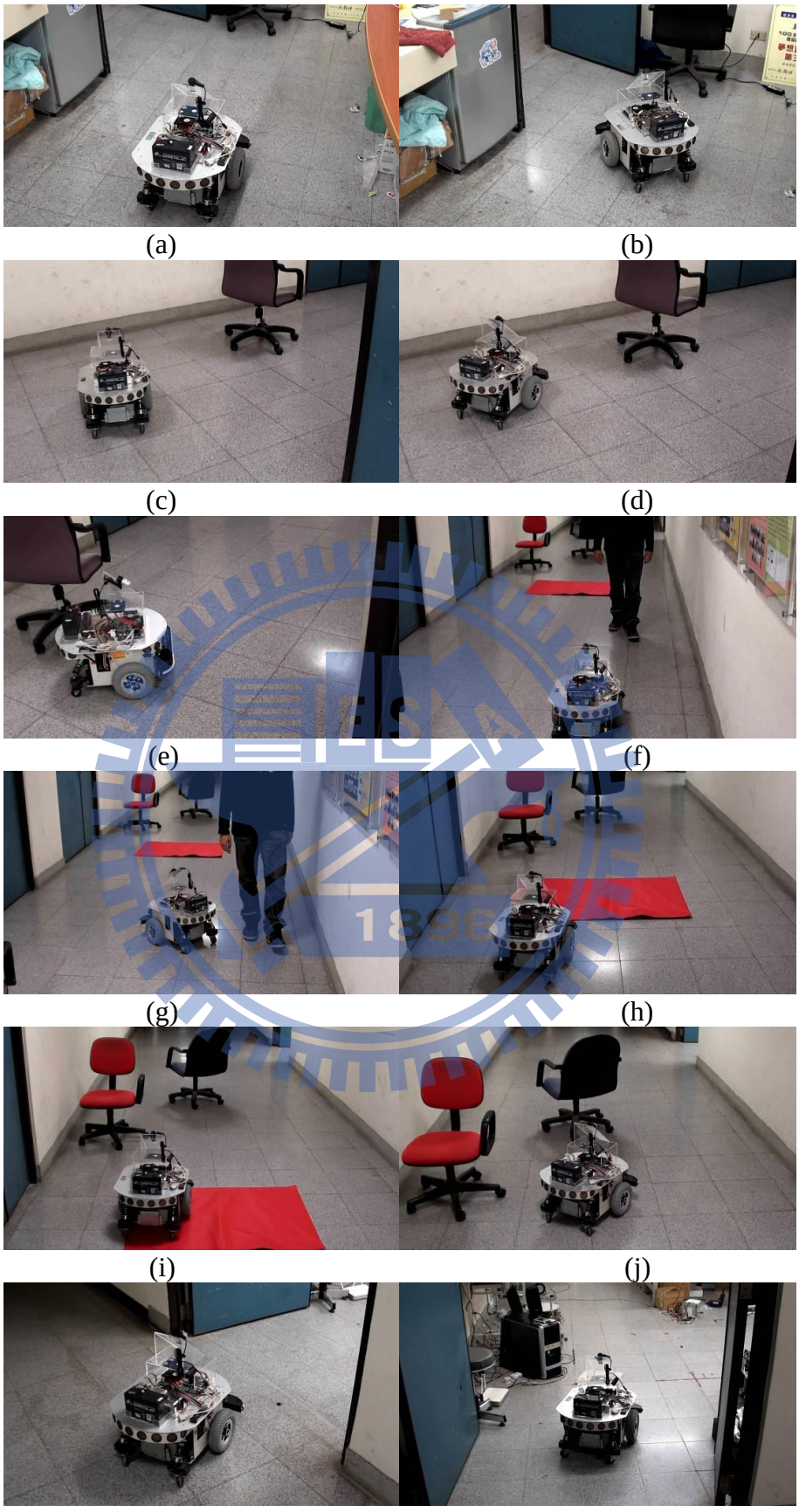


Fig. 5-5: The recorded trajectories of navigation experiments using the proposed MonoVNS and the URG-04LX-UG01 laser range finder. Label (a)-(l) represent the position of the robot in Fig. 5-6. (a)-(l).



(a) (b)

(c) (d)

(e) (f)

(g) (h)

(i) (j)

(k) (l)

Fig. 5-6: Snapshots from the navigation experiment.

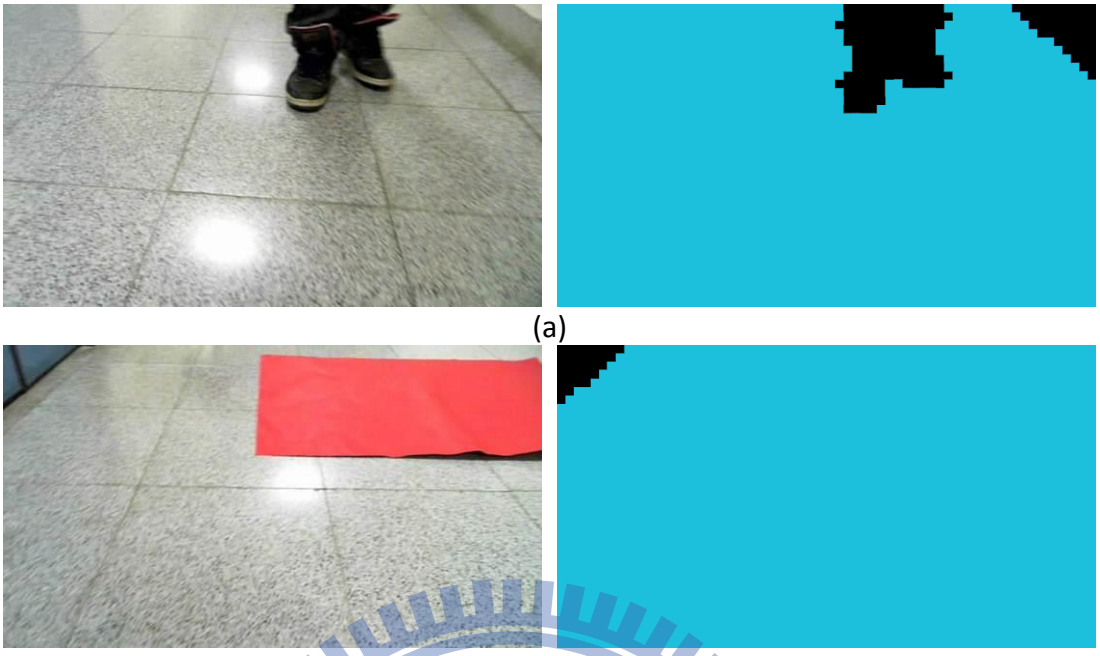


Fig. 5-7: Experimental results of ground detection and pedestrian detection

In theory, the robot is possible to achieve speeds higher then 25cm/s under current implementation (with 190ms processing time on average). However, the narrow view angle would make it dangerous to take turns and therefore 25cm/s was set for safety. To compare with a common laser range finder, we have repeated the experiment using a URG-04LX-UG01 to experiment in the same environment with the same navigation algorithm. The range of the laser range finder is 20 to 5600mm. The scanning angle of the laser range finder was limited as 60 degrees (the same condition with the monocular camera) so that the results can be fairly compared. The trajectory of the experiment is also shown in Fig. 5-5. The time of navigation using laser range finder was 38 seconds, while that of using the proposed method is 43 seconds. The distance travelled using laser range finder is 16.9m, while that of using the proposed method is 17.3m. The experimental results show that the both sensors give similar performance under the same view angles.

5.4 Intruder Detection Application

The experiment aimed to demonstrate the application of a WSN-based location aware system and mobile robot for intruder detection. Fig. 5-8 shows the system architecture of the experiment. Three intruder sensor modules were placed in the environment, including one microphone and two pyro sensors as well as a 3-axis accelerometer. The sensing data from the intruder sensor modules were first collected and processed on the modules independently. A location aware system was formed using sensor modules, and other ZigBee nodes deployed in the environment. The system can provide the locations of the nodes installed on the sensor modules and the robot. If an intruder is detected, the ID of the module and related sensor data will be transmitted to the mobile robot via the WSN. The current position of the robot and the triggered module are also determined by the location aware system at the same time. Meanwhile, the mobile robot will receive the detection result from the WSN, navigate to each

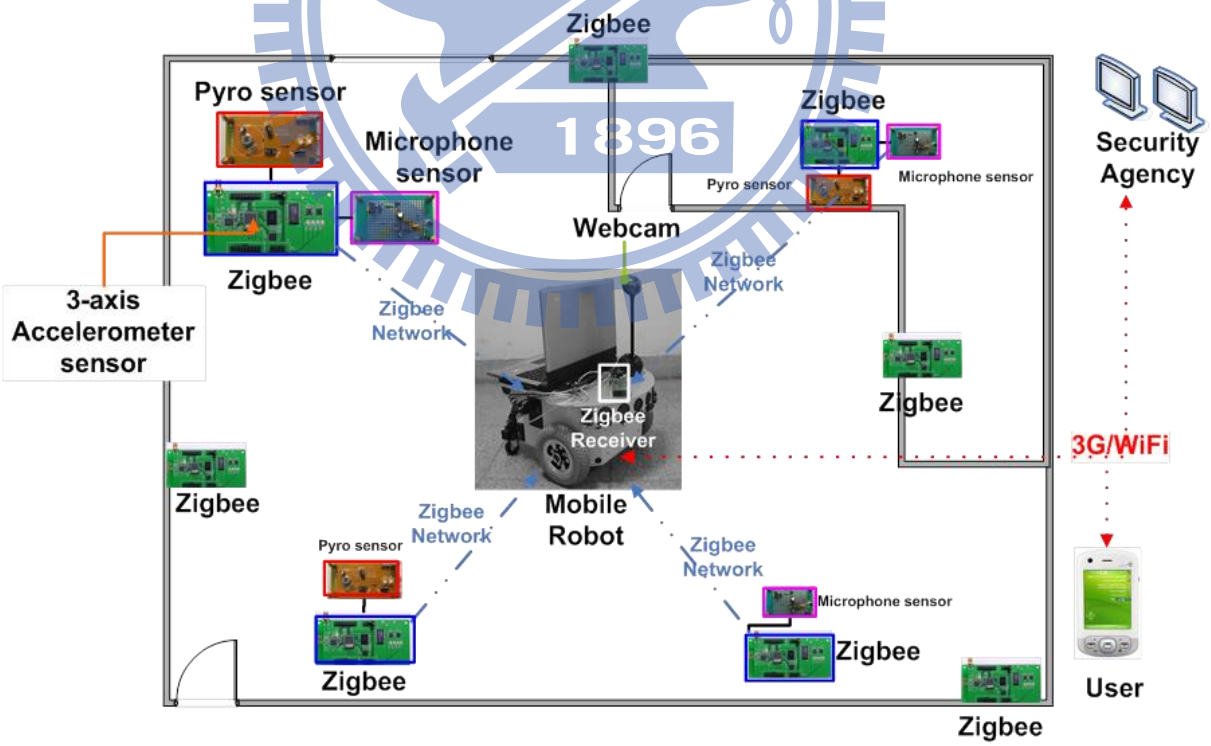


Fig. 5-8: The hardware arrangment of an intruder detection system.

of the alarm locations, and take real-time images using the robot's webcam. The security agency and users can easily access both the detection result and real-time video transmitted by the robot on their portable devices via both WiFi and 3G networks. Users can also remotely control the robot to confirm the circumstances in detail.

Fig. 5-9 shows the execution flowchart of the intrusion detection system. Once a sensor module is triggered by the detection of a possible intruder, the ID of that module will be logged and added to the patrol schedule. Additional redundant triggers will then be ignored. Once the robot arrives at the alarm location, it will send both a short message and real-time

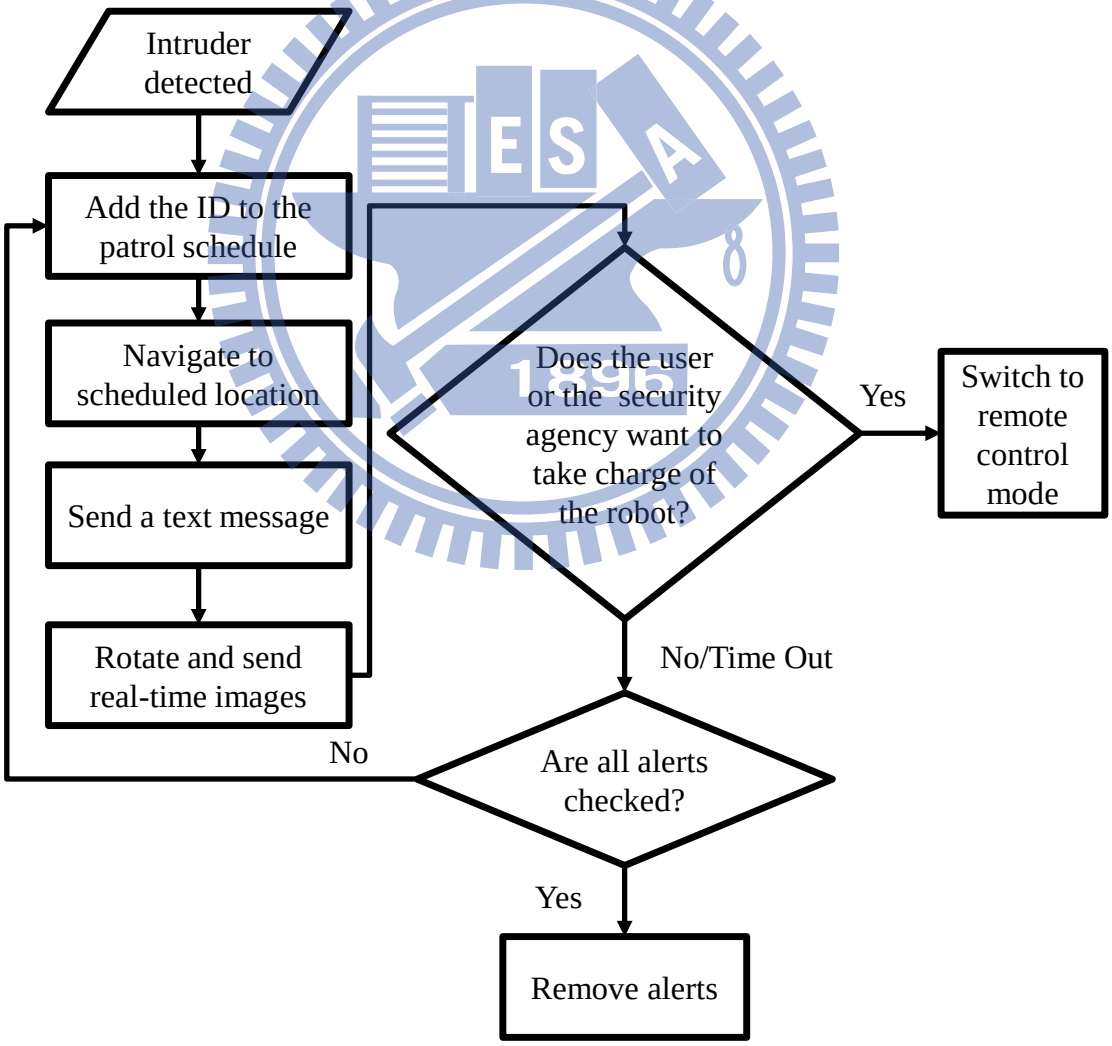


Fig. 5-9: Flowchart of the proposed intruder detection system

images to the security agency and end-users. The operators can then take control of the robot if necessary. Otherwise, the robot will move to the next alarm position. In the experiment, three different cases were tested in one script, including situations that mimicked a broken window by hand clapping (Event 1), a person entering the room (Event 2), and a thief sneaking into the room (Event 3), see Fig. 5-10. Fig. 5-11 illustrates the experiment using snapshots. Fig. 5-12 shows the actual trajectory of the robot during the experiment. During Event 1, a person clapped his hand and triggered the alarm, as shown in Fig. 5-10 (a) and Fig. 5-11 (a). The robot received the alarm signal and moved to where the corresponding surveillance location had been designated by the ZigBee sensor module (Fig. 5-10 (b)-(c), Fig. 5-11 (b)-(c) and Fig. 5-12 (a)-(c)). During Event 2, a person entered the room, which triggered the pyro sensor near the door. The robot reacted by moving toward the door (Fig. 5-10 (d)-(f), Fig. 5-11 (d)-(f) and Fig. 5-12 (d)-(f)). Finally, during Event 3, a person acting as a thief sneaked inside and wanted to steal a notebook. He triggered the pyro sensor and the robot once again approached the notebook and took pictures of the thief (Fig. 5-10 (g)-(i) and Fig. 5-11 (g)-(i), and Fig. 5-12(g)-(i)). Fig. 5-10 and Fig. 5-11also show the images received by a remote user. Since these images were taken without any recognizing algorithm and tracking mechanism, the system cannot guarantee to take any certain parts of the intruder.

5.5 Robot-on-Demand Experiment

This experiment aims to demonstrate how to accomplish the robot-on-demand service using location aware system and mobile robot. In order to provide practical services in the demonstration, several adaptive behaviors such as autonomous navigation, mobile manipulation and grasping, object delivery to the robot control system. The architecture of the robot control system is shown in Fig. 5-13. These behaviors are designed to safely manipulate the robot according to the status of users and the environment. All the outputs of these

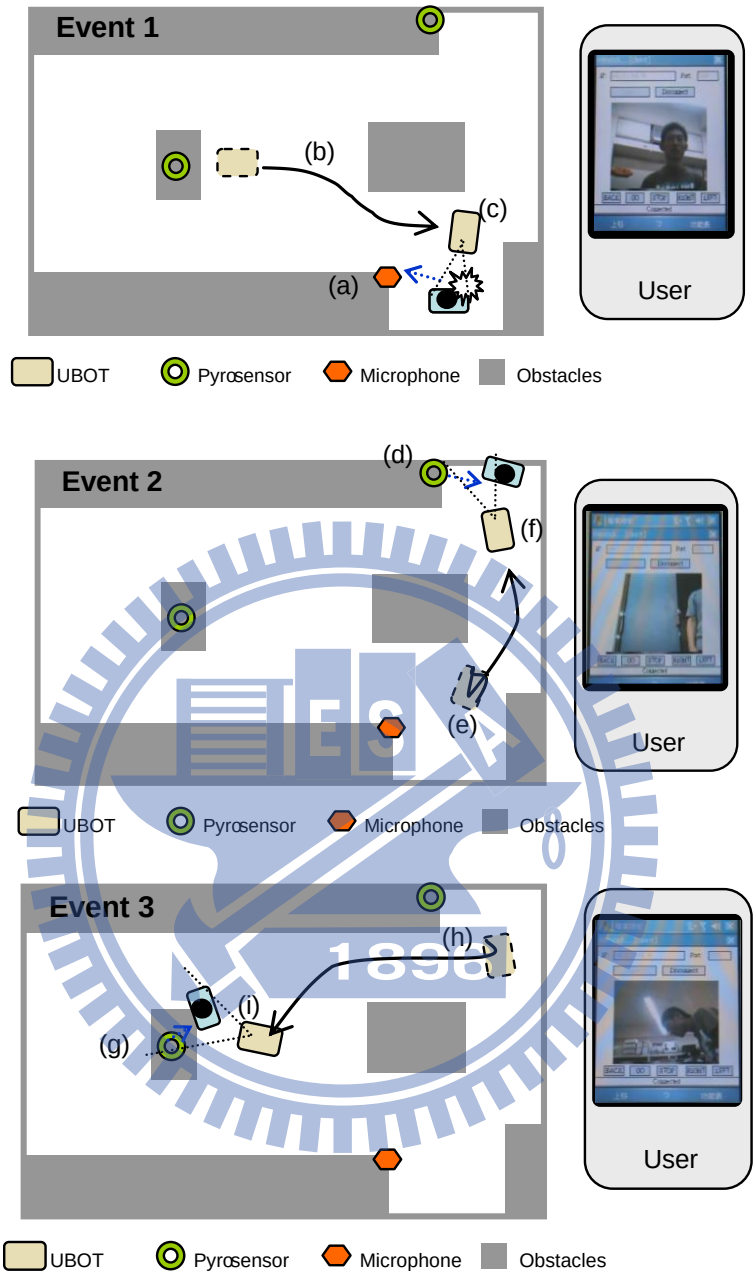


Fig. 5-10:The process and the images acquired by the robot in the intruder detection experiments: (a) a person claped his hands near the microphone; (b) the robot received the alert and moved to the person; (c) the robot took the person's image and sent it to the user; (d) a person walked into the room from the doorway; (e) the robot received the alert and moved to the person; (f) the robot took the person's image and sent it to the user; (g) a person (acting as a theif) intruded into the room; (h) the robot received the alert and moved to the person; (i) the robot took the image and sent it to the user.

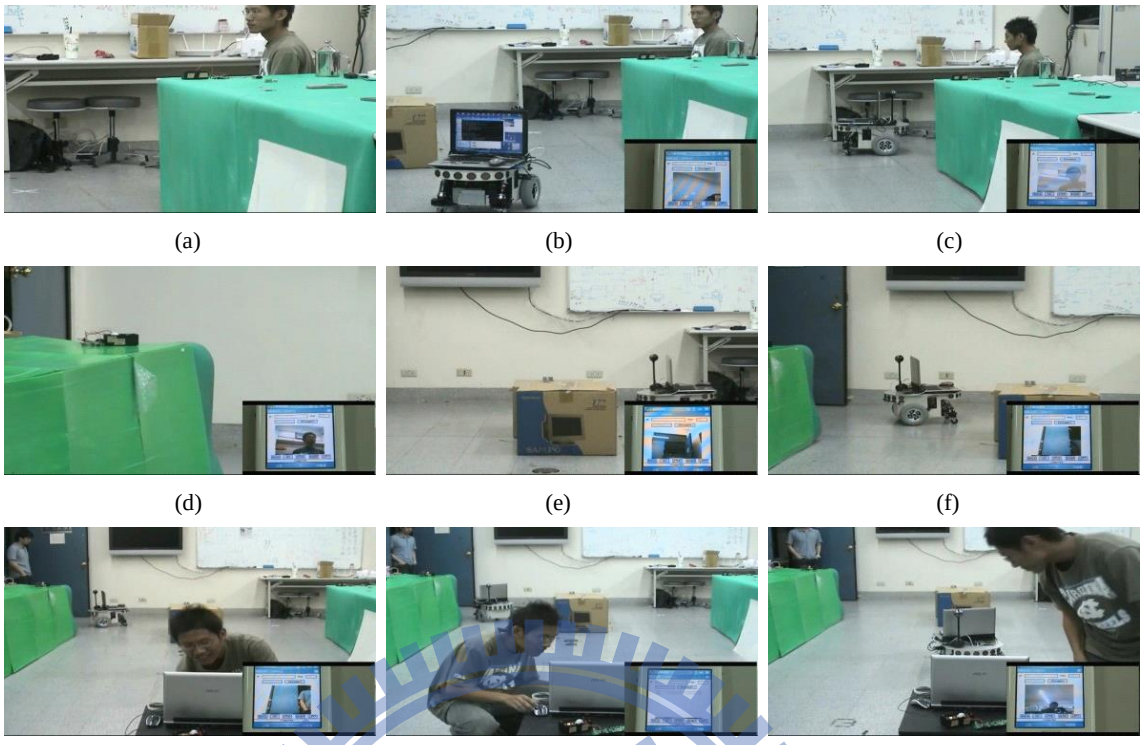


Fig. 5-11: Snapshots of the intruder detection experiment corresponding to Fig. 5-10.

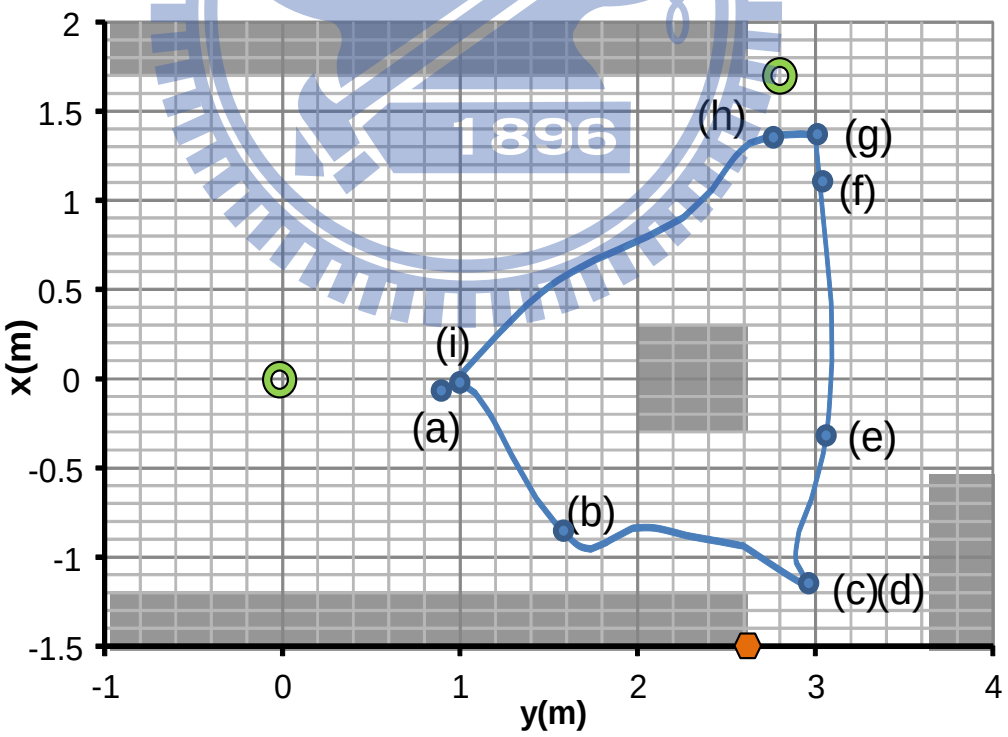


Fig. 5-12: The robot trajectory in the intruder detection experiment. The mark (a)-(i) represent the corresponding location of the robot in 5-10 (a)-(i)

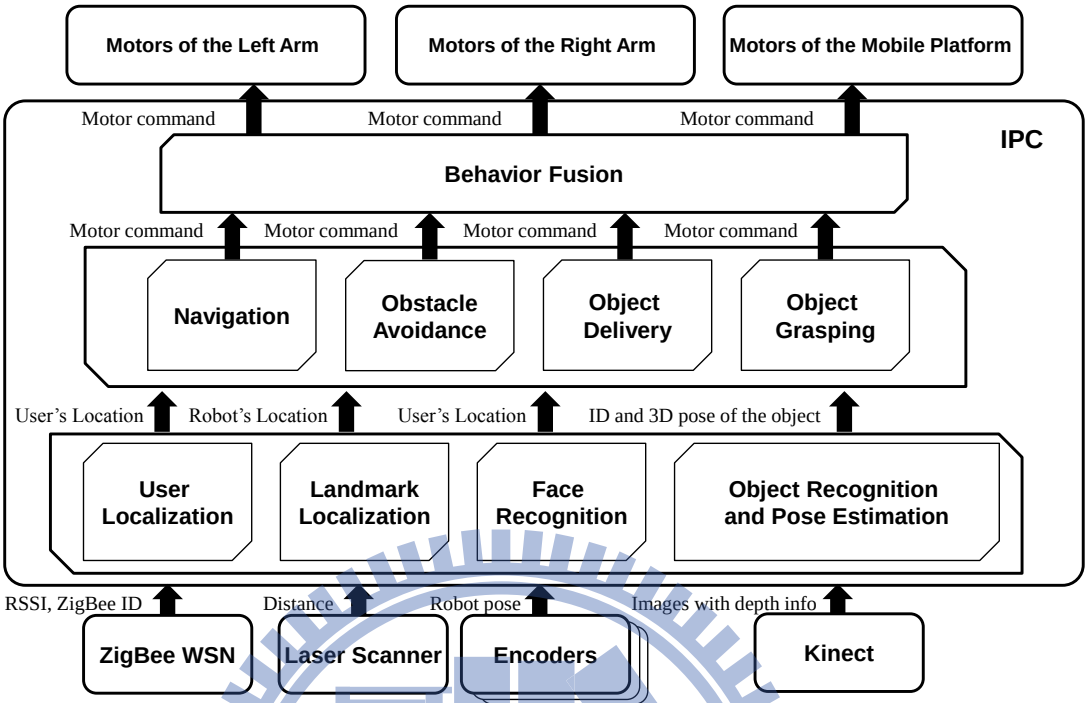


Fig. 5-13:System Architecture of the dual arm mobile manipulator.

behaviors are collected and fused to determine the ultimate motion of the robot using a fuzzy neural network. The behavior-fusion design of this work is extended from the method presented in section 4.3.3. The robot on-demand behavior integrates location aware, face detection and the autonomous navigation system of the mobile robot [57].

To demonstrate various complex tasks for daily-life assistance in a home setting, a multi-purpose mobile robot is also required. In this experiment, a dual-arm manipulator was installed on an omni-directional mobile platform, as shown in Fig. 5-14 (a). A recent photo and the hardware configuration of the mobile manipulator are shown in Fig. 5-14 (b). A Kinect camera is mounted on the pan-tilt platform for vision-based grasping and human detection. An on-board industrial PC (IPC) is responsible for analyzing the sensory data and task planning and control of the robot behaviors. A laser scanner was put at the front of the robot for obstacle avoidance. The semi-humanoid design not only provides a human-friendly looking, but also makes it more suitable to work in an environment originally structured for human. Harmonic

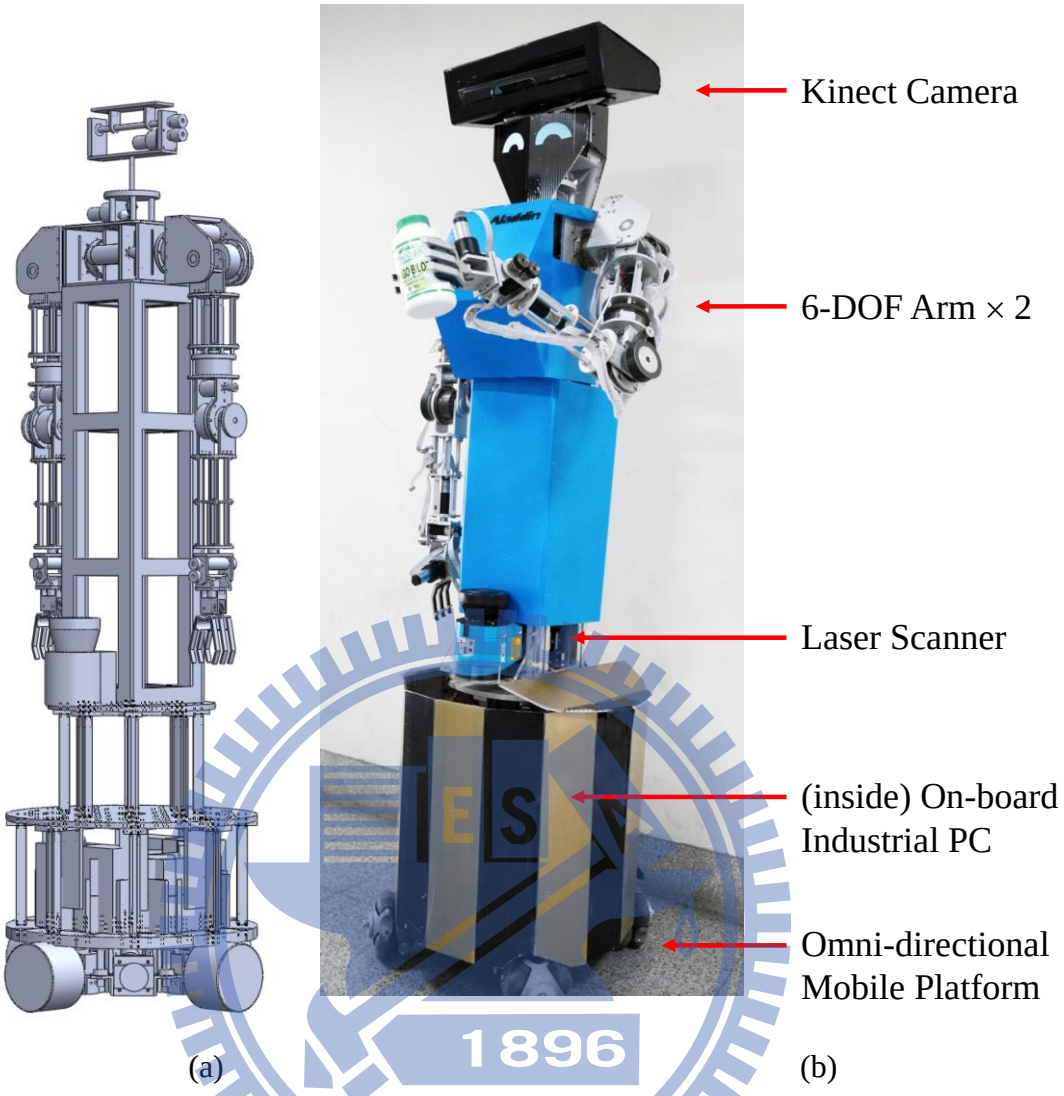


Fig. 5-14:The dual-arm mobile manipulator; (a)the mechanical design of the mobile manipulator; (b) a recent photo of the robot and the key components of the mobile manipulator.

drives are employed at shoulder and elbow for smoother motion. The EPOS positioning controllers from Maxon were used to control the motors of each joint. The position commands were transferred to qc values as input to the controller. It will make the arm move to the assigned position. Current gripper design aims to mimic the function of human palm and fingers, making it capable of handling various objects. The gripper is driven by motor and reduction gears in order to provide accepeable grip force. The mobile platform is provided with four omni-directional wheels, each driven by a self-contained motor. The platform gives the

robot omni-directional degrees of freedom for navigation and vision-based grasping in narrow space. The detailed motion model of platform was presented in [57].

The effectiveness of the proposed methods is demonstrated by performing practical daily assistive tasks in which the robot interacts with a user for handling objects in the environments. There are two phases in the experiment. In phase 1, the robot shows how it can deliver an object to a user. In phase 2, it is demonstrated that the robot can detect if the user is under critical situation, in this case, a heart attack. The robot can then come to help the user by bring him the medicines. In order to both monitor the location and the poses of the user, the user carries two ZigBee wireless sensor modules during the experiment, as shown in Fig. 5-15. A human pose recognition module [11] is tied on his waist, and a pulse oximeter (SpO_2) is mounted on his finger [12]. With these two sensors, the robot can monitor the pose, heart rate, and oxygen saturation in blood of the user. Six additional ZigBee modules are installed in the environment to form the ZigBee sensor network for transferring the sensor data and localize the user. The positions of these modules and the estimated user positions are shown in Fig. 5-17 and Fig. 5-19. The experiment is described in the following.

Phase 1: Object delivery

In the beginning of the experiment, the user was sitting on in the living room reading the newspaper, as shown in Fig. 5-15 The robot was at standby mode, as shown in Fig.8(a). Meanwhile, a parcel delivery man came in, notifying that he has a package to deliver. Since the user was reading newspaper, he commanded the robot to take the package for him. The robot therefore moved to the front door and received the package from the delivery man, as shown in Fig. 5-16 (b). It is noticed that although the delivery man moved the object during the grasping process, the robot could still grasp the object by visual servoing. After that, the robot moved toward the user according to estimated location of the user to hand over the parcel. During its way to the user, the human detection system recognized that the user was doing exercises. So

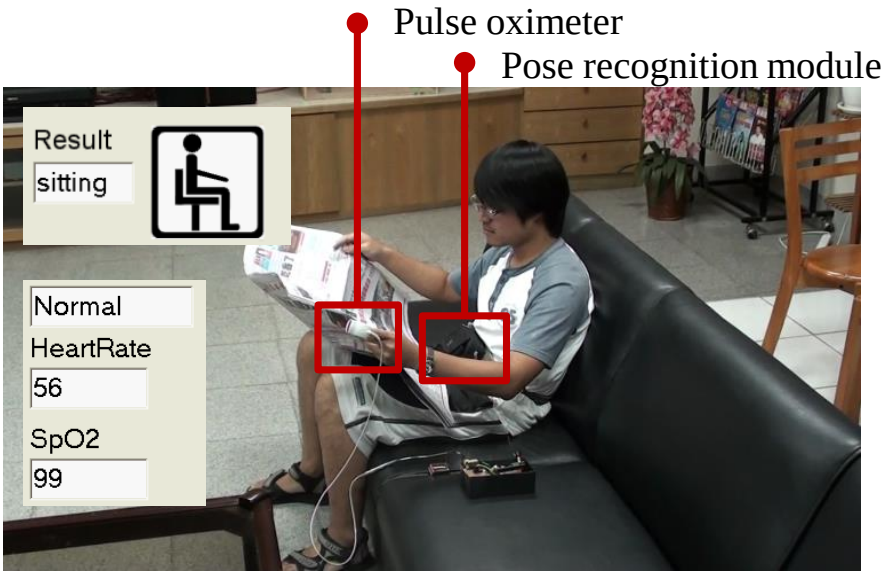


Fig. 5-15: The user wearing two sensor modules in the experiment.

the robot gradually adjusted its movement and stopped in front of the user. Then, the robot delivered the object to the user, as shown in Fig. 5-16 (c). Finally, the robot headed back to its rest zone, as shown in Fig. 5-16 (d), where it corrected its position estimation using a landmark in the environment. Fig. 5-17 shows the trajectory of the robot and the user during phase 1 of the experiment.

Phase 2: Emergency handling by fetching a drug can

In the beginning of phase 2 of the experiment, the user was sitting on the sofa unpacking the package. The robot was in standby mode at the rest zone. Suddenly, the user had a heart attack, as shown in Fig. 5-18(a). This situation was simulated by un-mounting the pulse oximeter on the user's finger, and the heart rate data became the maximum value. The robot noticed the situation right away via the ZigBee sensor network and moved toward the user to assist. When the robot arrived, it offered both of its hands to comfort the user, as shown in Fig. 5-18 (b). Although the user tried to react, he eventually laid down again due to severe illness. The robot therefore turned back to take the medicine on the cupboard, as shown in Fig. 10(c), and handed it to the user, as shown in Fig. 5-18 d). Finally, the heart rate of the user was back

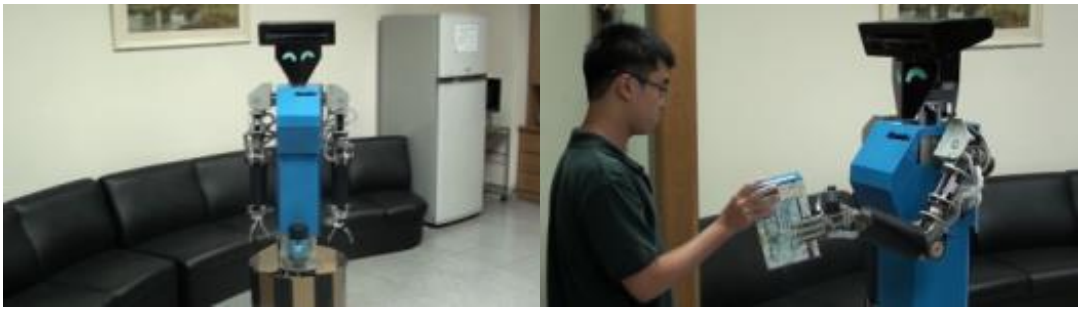
to normal, and the robot headed back to its rest zone again. Fig. 5-19 shows the trajectory of the robot and the user during phase 2 of the experiment.

The experiment shows that the ZigBee WSN successfully provides the approximated location of the user for the robot to approach and search for the exact location of the user. Furthermore, the experiment shows the robot can accomplish two object-grasping-and-delivery tasks, despite very different situations of the object and the user.

5.6 Summary

This chapter first evaluates the basic components of the location aware system: the intelligent environment and the mobile robot. The experiments show the intelligent environment can locate the user via the ZigBee location deisgn. The multimedia HMI facilitates a user to visualize the located target, and command the system by speech. The experiment of monocular visual navigation system shows the robot is able to reach the desired goal using a single camera. The result show that the proposed method is able to distinguish moving obstacles and homogenous region on the ground.

Several integrated applications of the proposed location aware system have also been demonstrated, such as intruder detection and robot-on-demand service. For the intruder detection application, three types of intrusion sensors have been installed and tested for the robotic patrolling system. This system can easily adopt other off-the-shelf sensor modules, while at the same time retaining compatibility with most existing security solutions. Our initial experiments show that the proposed system can effectively detect intruders and transmit their snapshots to portable devices at a remote site in real time. For the robot-on-demand service, the experiment shows that the robot is able to locate the user if an emergency occurs, and complete the task safely and robustly.



(a) (b)

(c) (d)

Fig. 5-16: Snapshots of phase 1 experiment; (a)The robot was at standby mode; (b) The robot took the package from the delivery man;(c) the robot delivered the package to the user; (d) The robot went back to the rest zone.

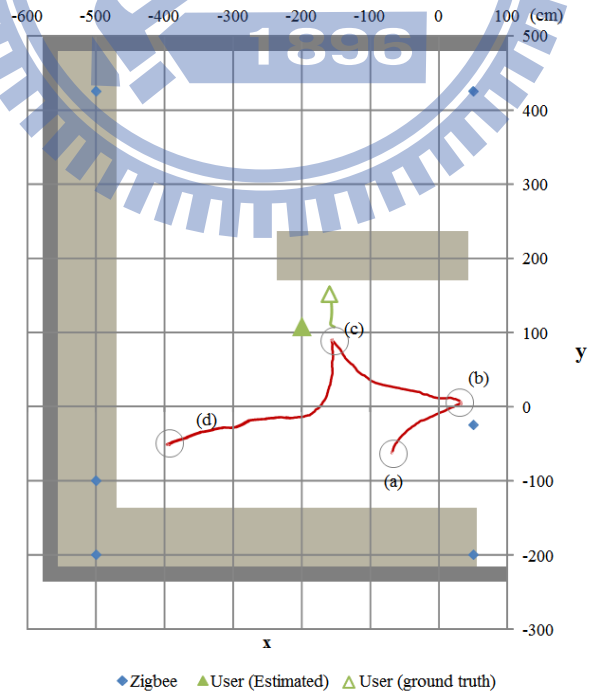


Fig. 5-17: Recorded trajectory of the robot and the user in phase 1 experiment.

(a) (b)

(c) (d)

Fig. 5-18: Snapshots of phase 2 experiment; (a)The the user had a heart attack; (b) The robot offered its hands to comfort the user;(c) the robot took a drug can from the cupboard; (d) The robot handed the drug to the user.

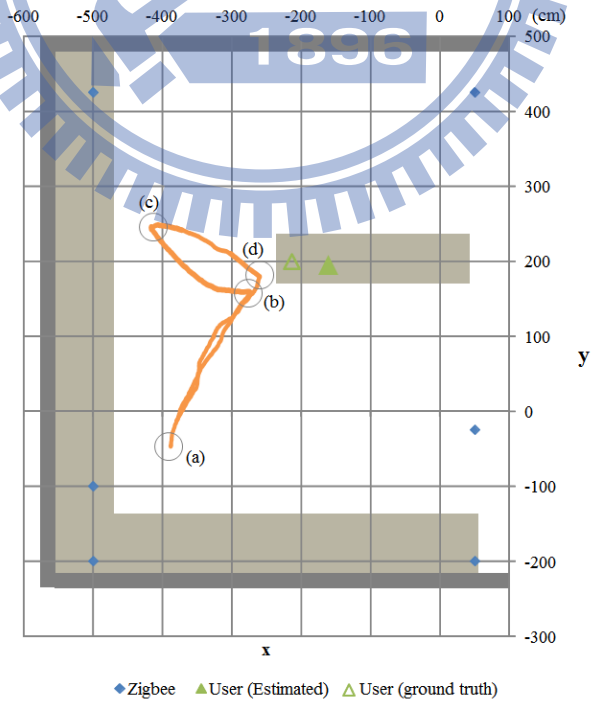


Fig. 5-19: Recorded trajectory of the robot in phase 2 experiment

Chapter 6.

Conclusion and Future Works

6.1 General Conclusions

This thesis proposes a location aware robotic system which features a ZigBee WSN and a mobile robot. By integrating these technologies, the location-ware system will be able provide not only location information passively, but also provide on-demand services such as intruder detection and active health care.

To provide a faster deployment of intelligent environment, we present a novel method for localizing a ZigBee node in a WSN. The probability-based approach takes the uncertainties of RSSI-based distance estimation into account. The method is independent of the sensor node distribution and environment configuration and therefore can minimize the calibration burden in practical applications. The performance of the proposed algorithm has been evaluated by experiments in three different environments. The experimental results show that the average error of the implemented location aware system is 1.7 m, despite a much higher range uncertainties from raw RSSI values. This accuracy can effectively support service robotic applications, especially when properly combined with a vision system. For instance, the localization system can provide a possible area of a user calling the robot for service or when an emergency situation occurs. The robot can then find the alarm area and monitor it with its vision system. Use of a mobile robot to detect an intruder has been demonstrated. Three types of intrusion sensors have been installed and tested for the robotic patrolling system. This

system can easily adopt other off-the-shelf sensor modules, while at the same time retaining compatibility with most existing security solutions.

For a low-cost robot navigation solution, a monocular vision-based robot navigation system has been developed by combining IPT and ground region segmentation. Useful steps in image processing have been developed and integrated into the ground plane detection algorithm to enhance the robustness of the system. To fulfill the requirement of real-time performance, we propose to combine a corner detection algorithm with SURF feature descriptor such that a scaled down image can give acceptable accuracy in distance measurement. Experimental results show that the system has satisfactory accuracy of 3% for distance measurement in a range of 4.8 meters. A local grid map can thus be built with single camera for robot navigation. The proposed MonoVNS can be an alternative to popular laser range finders and RGB-D sensing devices for robot navigation applications. The merit of this approach is to have both obstacle avoidance and path finding functions with a low cost monocular camera.

The integration experiments show that the proposed system can effectively detect intruders and transmit their snapshots to portable devices at a remote site in real time. The system can also provide on demand service, which robot shows up right on the spot and at the right time. The integration of ZigBee WSN and human detection system guarantees that the robot can always find the user. The experiment shows that the robot behaviors can adapt to the situation of both the user and the environment in order to complete the task safely and robustly.

6.2 Future Works

Some directions for future study are recommended below:

For the localization environment, a distributed version will be designed and implemented. The position of the reference node will be saved and sent to the mobile node in the localization

phase. Location-aware computation can be performed on the ZigBee module. Furthermore, the proposed method estimates the location from only a set of RSSI values without using previous RSSI values. Thus it is also interesting to examine the performance of the target tracking with the proposed method. Furthermore, the system can be applied to a “never-lost” system, while not only human, but also any object which is RF-tagged in the house. They can be found through location aware system and visualized on a TV set, for instance.

For the MonoVNS, we will work on reducing the system execution time in order to use high resolution imagery. The computation intensive parts of the proposed method are feature extraction and matching. It can be further improved by using other feature descriptors or by using a GPU to accelerate. In the future, we will also investigate solutions for a wider camera view angle, for instance, by using a wide-angle lens or pan-tilt platform.

For the integrated location aware system, the mobile robot and the intelligent environment should be able to cooperate more thoroughly. For instance, the robot should access the camera in the environment for self-localization and human tracking. The object with ZigBee modules or RF-tagged should also be fetched by the robot.

Bibliography

[1] R. C. Luo, T. Y. Lin, K. L. Su, “Multisensor based security robot system for intelligent building,” *Robotics and Autonomous Systems*, vol. 57, No. 3, pp. 330-338, 2009.

[2] B.J. You, M. Hwangbo, S.O. Lee, S.R. Oh, Y.D.Kwon and S. Lim, “Development of a home service robot 'ISSAC',” in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Nevada, USA, Oct. 2003, pp.2630-2635.

[3] R. Borja, J.R. de la Pinta, A. Á lvarez and J.M. Maestre, “Integration of service robots in the smart home by means of UPnP: A surveillance robot case study,” *Robotics and Autonomous Systems*, vol. 61, No. 2, pp. 153-160, 2013.

[4] C. Williams, Y. Kwon Cho and J.H. Youn,“Wireless Sensor-Driven Intelligent Navigation Method for Mobile Robot Applications in Construction,” in *Proc. of ASCE International Workshop on Computing in Civil Engineering*, Pennsylvania, USA, 2007, pp.493-498.

[5] K. Han, J. Lee, S. Na and W. You “An Ambient Robot System Based on Sensor Network: Concept and Contents of Ubiquitous Robotic Space,” in *Proc. of International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, Pennsylvania, USA, 2007, pp.155-159.

[6] ZigBee Specification, ZigBee Alliance, 2008. Available: http://www.zigbee.org/

[7] G.Zanca, F. Zorzi, A. Zanella and M.Zorzi, “Experimental Comparison of RSSI-based Localization Algorithms for Indoor Wireless Sensor Networks,” in *Proc. of the Workshop on Real-World Wireless Sensor Networks* , Glasgow, Scotland, 2008, pp. 1-5.

[8] C. Liu, T. Scott, K. Wu and D. Hoffman, “Range-free sensor localisation with ring overlapping based on comparison of received signal strength indicator,” in *International Journal of Sensor Networks*, No.5/6, pp. 399-413, 2007.

[9] A. Savvides, H. Park and M. B. Srivastava, “The n-hop Multilateration Primitive for Node Localization Problems,” *Mobile Network and Applicatons*, vol. 8, no. 4, pp. 443–451, 2003

[10] K. Langendoen and N. Reijers, “Distributed Localization in Wireless Sensor Networks: a Quantitative Comparison,” *Computer Networks*, vol. 43, no. 4, pp. 499–518, 2003

[11] F. Gustafsson and F. Gunnarsson, “Mobile Positioning Using Wireless Networks: Possibilities and Fundamental Limitations Based on Available Wireless Network Measurements,” *IEEE Signal Processing Magazine*, Vol. 22, pp. 41-53, 2005.

[12] G. I. Wassi, C. Despins, D. Grenier and C. Nerguizian, “Indoor Location Using Received Signal Strength of IEEE 802.11b Access Point,” in *Proc. of Canadian Conference on Electrical and Computer Engineering*, Saskatoon, Canada, 2005, pp.1367-1370.

[13] K.T. Song, C.Y. Tsai, F.S. Huang, J.W. Hong, C.Y. Lin , C.W. Chen and Z.S. Lin, “Development of the Robot of Living Aid: RoLA,” in *Proc. of 2008 IEEE International Conference on Automation and Logistics* , Qingdao ,China, Sep. 1-3, 2008 , pp. 443-448 .

[14] T. Oliveira, M. Raju and D. P. Agrawal, “Accurate Distance Estimation Using Fuzzy based combined RSSI/LQI Values in an Indoor Scenario: Experimental Verification, ” *Network Protocols and Algorithms Journal, Special Issue on 2nd IEEE Smart Communication Protocols and Algorithms*, vol. 4, No. 4, pages 174-199, 2012.

[15] L.Gogolak, S. Pletl and D. Kukolj, “Indoor Fingerprint Localization in WSN Environment Based on Neural Network,” in *Proc. of IEEE 9th International Symposium on Intelligent Systems and Informatics*, Subotica, Serbia, Sep. 8-10, 2011, pp.293-296.

[16] S. H. Fang and T. N. Lin, “Indoor Location System Based on Discriminant-Adaptive Neural Network in IEEE 802.11 Environments,” *IEEE Transactions on Neural Networks*, vol.19, no.11, pp. 1973-1978, 2008.

[17] V. Ramadurai and M. L. Sichitiu, “Localization in Wireless Sensor Networks: A Probabilistic Approach,” in *Proc. of the 2003 International Conference on Wireless Networks*, Las Vegas, NV, June 2003, pp. 275–281.

[18] R. Vaughan and J. Andersen, Channels, *Propagation and Antennas for Mobile Communications*, Institution of Electrical Engineers, 2003.

[19] J. Graefenstein and M.E. Bouzouraa, “Robust Method for Outdoor Localization of a Mobile Robot Using Received Signal Strength in Low Power Wireless Networks,” in *Proc. of 2008 IEEE International Conference on Robotics and Automation*, Pasadena, CA, USA, 2008, pp. 33-38

[20] H. Cho, M. Kang, J. Park, B. Park and H. Kim, “Performance Analysis of Location Estimation Algorithm in ZigBee Networks Using Received Signal Strength,” in *Proc. of the 21st International Conference on Advanced Information Networking and Applications Workshops*, Niagara Falls, Canada, May 2007, pp. 302–306.

[21] K. Lee, A. Oka, E. Pollakis and L.H. Lampe, “A Comparison Between Unscented Kalman Filtering and Particle Filtering for RSSI-based Tracking ,” in *Proc. the 7th Workshop on of Positioning Navigation and Communication*, Dresden, Germany, March 11-12, 2010, pp.157-163.

[22] S. Lenser and M. Veloso, “Visual Sonar Fast Obstacle Avoidance Using Monocular Vision,” In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, Nevada, USA, 2003, pp. 886- 891.

[23] C. Chen, C. Cheng, D. Page, A. Koschan and M. Abidi, “A Moving Object Tracked by a Mobile Robot with Real-Time Obstacles Avoidance Capacity,” in *Proc. of the 18th International Conference on Pattern Recognition*, Washington, D.C., USA, 2006, pp.1091-1094.

[24] N. Ohnishi and A. Imiya, “Dominant Plane Detection from Optical Flow for Robot Navigation,” *Pattern Recognition Letter*, vol. 27, pp. 1009–1021, 2006.

[25] Y. Kim and H. Kim, “Layered Ground Floor Detection for Vision Based Mobile Robot Navigation,” in *Proc. of IEEE Int. Conf. on Robotics and Automation*, New Orleans, Louisiana, USA, 2004, pp. 13-18.

[26] J. Michels, A. Saxena and A.Y. Ng, “High Speed Obstacle Avoidance using Monocular Vision and Reinforcement Learning,” in *Proceedings of the Twenty-first International Conference on Machine Learning (ICML)*, Bonn, Germany, 2005, pp. 593-600.

[27] B. Micusik, H.Wildenauer and M. Vincze, “Towards Detection of Orthogonal Planes in Monocular Images of Indoor Environments,” in *Proc. of IEEE Int. Conf. on Robotics and Automation*, Los Angeles, USA , 2008, pp. 999-1004.

[28] B. Liang, N. Pears and Z. Chen, "Affine Heigh Landscapes for Monocular Mobile Robot Obstacle Avoidance", in *Proc. of the 8th Int. Conf. on Intelligent Autonomous Systems*, Amsterdam, The Netherlands, 2004, pp.863-872.

[29] J. Zhou and B. Li, “Homography-based Ground Detection for a Mobile Robot Platform Using a Single Camera,” in *Proc. of the 2006 IEEE International Conference on Robotics and Automation*, Orlando, Florida, 2006, pp. 4100-4105.

[30] J. Zhou and B. Li, "Robust Ground Plane Detection with Normalized Homography in Monocular Sequences from a Robot Platform," in *Proc. of the 2006 IEEE International Conference on Image Processing*, Atlanta, GA, USA, 2006, pp. 3017-3020.

[31] D. Conrad and G. N. DeSouza, "Homography-Based Ground Plane Detection for Mobile Robot Navigation Using a Modified EM Algorithm," in *Proc. of the 2010 IEEE International Conference on Robotics and Automation*, Anchorage, Alaska, USA, 2010. pp. 910-915.

[32] A. Mittal and S. Sofat, "A Robust and Efficient Homography Based Approach for Ground Plane Detection, " *BVICAM's International Journal of Information Technology (BIJIT)*, vol.4, no. 2, 2012.

[33] N.Simond and M. Parent, "Obstacle Detection from IPM and Super-Homography," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, CA, 2007, pp.4283-4288.

[34] S. Wybo, R. Bendahan, S. Bougnoux, C. Vestri, F. Abad and T. Kakinami, "Improving Backing-Up Manoeuver Safety with Vision-Based Movement Detection," *Intelligent Transport Systems, IET*, vol., no. 2, pp.150-158, 2007.

[35] K. Yamaguchi, A. Watanabe, T. Naito, and Y. Ninomiya, "Road Region Estimation Using a Sequence of Monocular Images," in *Proc. of the 19th international Conference on Pattern Recognition*, Tampa, Florida, USA, 2008, pp.1-4.

[36] E. Fazl-Ersi, and J.K. Tsotsos, "Region Classification for Robust Floor Detection in Indoor Environments," in *Proc. of the 6th International Conference on Image Analysis and Recognition (ICIAR)*, Halifax, Canada , 2009, pp. 717-726

[37] F. Bonin-Font and A. Ortiz, "Building a Qualitative Local Occupancy Grid in a New Vision-Based Reactive Navigation Strategy for Mobile Robots," in *Proc. of IEEE*

International Conference on Emerging Technologies and Factory Automation (ETFA), Palma de Mallorca, Spain, 2009, pp. 1511-1514.

[38] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, New York, NY, 2003.

[39] E. Rosten and T. Drummond, "Fusing Points and Lines for High Performance Tracking," in *Proc. of 10th IEEE International Conference on Computer Vision (ICCV 2005)*, Beijing, China, 2005 , pp. 1508-1515.

[40] D. G. Lowe, "Distinctive Image Features from Scale-invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp.91-110, 2004.

[41] H. Bay, A, Ess, T. Tuytelaars and L.V. Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol.110, pp. 346-359, 2008.

[42] O. Kähler and J. Denzler, "Detecting Coplanar Feature Points in Handheld Image Sequences," in *Proc.* of Conference on Computer Vision Theory and Applications (VISAPP), Barcelona, Spain, 2007, pp. 447-452.

[43] E. Serradell, M. Ö zuysal, V. Lepetit, P. Fua and F. Moreno-Noguer, "Combining Geometric and Appearance Priors for Robust Homography Estimation," in *Proc. of European Conference on Computer Vision* 2010, Greece, 2010, pp.58-72.

[44] D. Comaniciu, V. Ramesh and P. Meer, "The Variable Bandwidth Mean-shift and Data-Driven Scale Selection," in *Proc. of Eighth Int'l Conf. on Computer Vision*, Vancouver, British Columbia, Canada, July 2001, pp. 438-445.

[45] C. H. Lin, Y. T. Chen and K. T. Song, "A Visual Attention and Tracking Design for a Robotic Vision System," in *Proc. of 2009 IEEE Workshop on Advanced Robotics and Its Social Impacts*, Tokyo, Japan, 2009, pp. 30-35.

[46] L. Hong and G. Chen, "Segment-Based Stereo Matching Using Graph Cuts," in *Proc.* of *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, D.C., USA, 2004, pp.74-78.

[47] A. Bhattacharyya, "On a Measure of Divergence Between two Statistical Populations Defined by Probability Distributions," *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.

[48] " OSA/Parlay Overview," [Online document], 2002 Sep 7, [cited 2006 February 20], Available HTTP: http://www.openapisolutions.com/brochures/OSAParlayOverview.pdf

[49] K. T. Song and C. W. Chen, "Multi-Person Pose Recognition Using a ZigBee Sensor Network," in *Proc. of 17th IFAC World Congress*, Seoul, Korea, July 6-11, 2008, pp. 14976-14981.

[50] C.H. Lin, S.H. Yang, H.T. Chen and K.T. Song, "Mobile Robot Intruder Detection Based on a Zigbee Sensor Network," in *Proc. of 2008 IEEE International Conference on Systems, Man and Cybernetics (SMC 2008)*, Singapore, Oct. 12-15, 2008, pp. 2786-2791.

[51] S. H. Yang and K. T. Song,"An Adaptive Routing Protocol for Health Monitoring with a Sensor Network and Mobile Robot," in *Proc. of IEEE IECON 2010*, Phoenix, AZ, USA, 2010, pp. 2187 – 2192.

[52] C. H. Lin, S.P. Kuo, Y. J. Kuo, K.T. Song and Y.C. Tseng, "Visualization Design for Location aware Services," in *Proc. of IEEE International Conference on Systems, Man and Cybernetics*, Taipei, Taiwan, 2006, pp.4380-4385.

[53] C. H. Lin and K.T. Song, "An Interactive Control Architecture for Mobile Robots," *International Journal of Robotic and Automation*, vol. 28, pp. 1-12, 2013.

[54] C. H. Lin, S. Y. Jiang, Y. J. Pu and K.T. Song, "Obstacle Avoidance of Mobile Robots Using a Monocular Camera and Ground Plane Detection," in *Proc. of 2010 IEEE/RSJ Int. Conf. on Intelligent Robots Systems* Taipei, Taiwan, Oct. 2010, pp. 3706-3711.

[55] OpenNI, www.openni.org (last accessed 24/8/2011).

[56] OpenCV, opencv.willowgarage.com (last accessed 24/8/2011).

[57] K.T. Song, C.H. Lin, S.C. Tsai, S.C. Hung, Y.C. Liao and C.C. Wu, "Development of Robot-on-Demand Behaviors Based-on Visual Tracking and Wireless Sensor Network," in *Proc. of 2011 International Conference on Service and Interactive Robots (SIRCon 2011)*, Taichung, Taiwan, 2011, pp. 261-266.