**ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS**

**MASTER OF SCIENCE IN COMPUTER SCIENCE**

# The mobile phone as a platform for assisting the independent living of aging people

**Eleni Trouva**

**Supervisor: Professor George C. Polyzos**

Athens, June 2009

# Acknowledgements

This thesis was conducted at Athens University of Economics and Business during my master's studies. It would not have been possible without the guidance, extremely valuable support and contribution of my advisor Professor George Polyzos.

I would also like to thank Professor George Xylomenos for accepting the role of the external advisor. I am greatly indebted to Pantelis Fragoudis for his help and encouragement in every step of the process of this thesis.

Last but not least, I am especially grateful to my friends and my family for their love and patience through difficult times. Thank you.

# Abstract

Mobile phones or Personal Digital Assistants (PDAs) with advanced imaging and communication capabilities have become ubiquitous. In this work, we exploit these capabilities in the context of a Smart Home environment to monitor the activities and assist aging people. As a first step, we tackle the issue of using the PDA to track the position of a user moving within his home. Due to the importance of this problem in the realization of a Smart Home environment, numerous solutions have been proposed. In this thesis, we first discuss the technologies and the localization techniques proposed so far to achieve indoor positioning. Then, we explore the possibility of determining a user's location based on the camera images received from the PDA. In our system, the PDA is worn by the user as a pendant and images are periodically captured and transmitted to a Web server. The Web server returns the location of the user by comparing the received images with images stored in a database using off-the-shelf matching algorithms. The key advantage of the proposed solution lies in its ability to locate the user without the cost of additional hardware. In addition, with the purpose of assisting people with impaired vision, we evaluate the performance of an application that performs Optical Character Recognition (OCR) in photographs of medication boxes captured by smart phones or PDAs.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Motivation

The world population of people over the age of 65 is growing rapidly at a rate of 800,000 per month. As a consequence, the healthcare and medicare market already constitutes a major part of the world's economy and it will only expand in years to come. Recent advances in sensor technology, cellular networks and information technology promise to improve the well-being of the elderly by assisting them in their daily activities and monitoring their health status, thus enabling them to lead their lives to a larger extent independently from healthcare institutions and their caretakers.

However, the comparatively slow adoption of such systems indicates that there are certain factors prohibiting their acceptance and use. For example, most available home-care systems monitor the health of individuals suffering from diagnosed chronic diseases (heart disease, lung disorders, diabetes, etc.); as such, they depend on customised, costly, and difficult to use health monitoring equipment and they often confine patients to their homes. Less attention has been given to monitoring and maintaining the personal wellness of those elderly people who have not been diagnosed with any serious or chronic disease and who, therefore, wish or should be encouraged to live a normal life, both inside and outside their homes. The quality of life of this larger population depends upon and may decline as a result of combined and in some cases not easily measurable physical and psychological factors. Existing work in the direction of monitoring and improving their lives, however, is limited to a) interactive facilities for consultation with doctors, which, while simplifying some regular examinations, do not take advantage of the full potential offered by the latest advances in information and communication technologies, and b) satellite location tracking systems, which, while allowing people suffering from memory and orientation problems to be easily traced, do not otherwise exploit the wealth of data gathered by the location sensors.

While sudden changes and, more generally, unexpected patterns in the daily behaviour of elderly people are difficult to detect via medical data only, it is very likely that i) such unexpected patterns can be detected by combining machine learning and input from a larger variety of sensors, and ii) that they can be used as signs that a certain disease, illness or health condition is getting worse or that is about to occur. Of course, detecting these early signs is far from trivial, if we consider the uniqueness of each individual's personality, behaviour, and the diverse effects of different diseases upon this behaviour. Furthermore, early detection should not come at the cost of confining the elderly to their homes, nor should it rely on the assumption that elderly users will learn to control complex equipment.

This thesis is part of the work conducted at Mobile and Multimedia Laboratory of Athens University of Economics and Business for the ARCHANGEL project [1] . The aim of the ARCHANGEL project is to realise a holistic framework that will exhibit all the aforementioned principles and test its performance in a controlled trial setting. In this context, the project aims to design, implement, and validate a cost-effective, secure, adaptable and interoperable framework for learning and monitoring the daily behaviour and personal routines of the elderly using advanced sensor networking, machine learning, and controlled interaction with caretakers. In this work, we propose two separate systems, an indoor localization system and an application to assist visually impaired people that could be used in the framework of ARCHANGEL project. Both proposed solutions are based on the capabilities of today's mobile phones.

## 1.2 Thesis outline

This thesis is organised as follows:

Chapter 2 is a general overview of the Smart Home environment, discussing possible classifications and products and services provided by Smart Home solutions. At the end of the chapter we present related projects.

Chapter 3 is an overview of the basics of localization with the focus on indoor positioning. Applications, issues, demands and performance measures of indoor localization systems are discussed.

Chapter 4 introduces the technologies used to achieve indoor positioning, focusing on Bluetooth, Wi-Fi and image comparison. In addition, we review the common approaches for indoor localization, cell-based methods, range combining techniques and scene analysis.

Chapter 5 is a brief description of content-based image retrieval techniques and optical character recognition (OCR) technology.

In Chapter 6, we propose an indoor localization system that relies on image matching to locate a mobile user using the scene analysis technique. We present the system's architecture and evaluate the application performing experiments.

A proposal of an application to help people with visual impairment is done in Chapter 7. Again, we measure the performance of the proposed system, making suggestions for possible future enhancements.

Finally, in Chapter 8 the conclusion of this thesis is presented, summarizing its basic points.

# Chapter 2

# Smart Homes

## 2.1 Smart Homes for elderly

The proportion of the world's population of individuals over the age of 60 is expected to double by 2030 to twenty percent [12] [13] . This is the demographic trend for the next decades. The rapid growth of the elderly population and increase in life expectancy have led to new models of positive ageing where older adults are being empowered to lead fulfilling lives and adapt to degenerative changes to maintain functionality, autonomy and quality of life.

As older adults are living longer and more fulfilled lives, they desire to live as independently as possible. But independent lifestyles come with risks and challenges. Mobility and cognitive impairments among the elderly lead to functional decline. Interventions to improve function include both evidence-based nursing approaches and innovative technologies. Crucial to successful intervention is early identification of changing conditions that are precursors of impairments so that interventions can be offered at the earliest indications of need. Customized monitoring can enable the early detection of deteriorating health conditions, such as a shuffling gait, restless sleep, rapid change in activity level, or an unusual change in one's typical routine. To address these issues, researchers are developing "smart home" technologies to enhance residents' safety and monitor health conditions using sensors and other devices.

The terms smart homes, intelligent homes, home networking have been used for more than a decade to introduce the concept of networking devices and equipment in the house. According to the Smart Homes Association [47] [48] [49] the best definition of smart home technology is: the integration of technology and services through home networking for a better quality of living [3] .

Other terms that are related to smart homes are aware house, changeable home, attentive house and ambient intelligence. These terms are used to emphasise that the home environment should be able to respond and modify itself continuously according to its diverse residents and their changeable needs. For instance ambient intelligence is defined as a digital environment that is sensitive, adaptive and responsive to the presence of people. Ambient intelligence will encompass the home, car, clothing, work and public places. Reading this definition one could conclude that ambient intelligence is something that will happen in the far future. However there are many examples that ambient intelligence is something that is happening today. For instance cars adjust the settings of the chair and mirrors to a specific driver and large crowds are analysed by cameras and intelligent software to detect specific persons. With regard to the home environment there is a growing amount of digital equipment present in the home to support the residents. This doesn't mean that there are always the high-tech solutions that are present in the homes but also existing low-key technological devices which are readily available can make a contribution. In this sense smart homes are considered to be a first step in the process of creating a sensitive, adaptive and responsive home environment.

From a user perspective it is important to realise that residents consider their home to be a safe and comfortable place to live in. Sometimes technology is seen as an intruder in their safe environment, residents are afraid to lose control over their home. Some people even fear the use of technology in their home. So technology and the home environment are not naturally a perfect fit. With respect to smart home technology for elderly and disabled people another aspect is important. Some solutions are implemented to reduce the need to do things but it is also important to implement solutions that increase the participation of the resident in an activity. The goal of equipping the home environment with technology isn't just to automate all the tasks that are carried out by the residents. The objective in design is to provide tools and services that empower and enable people themselves to address their social, rational, and emotional needs. Equality, autonomy, and control are the goals of empowering design.

Furthermore technology is not the solution to create a perfect home environment but has the ability to make a useful contribution. The environment as a whole, including for instance social contacts and location of the home, is responsible for the overall satisfaction of the residents.

A smart home can enable elderly and disabled people to lead safe and independent lives in their own homes. A smart home can:

- Provide an environment that is constantly monitored to ensure the householder is safe (activity monitoring)

- Automate specific tasks that a householder is unable to perform (turning lights on or off)
- Provide a safe and secure environment (alerting the householder of potentially dangerous activities)
- Alert helpers or carers should the householder be in difficulties (through linking to a local community alarm scheme)
- Enable and empower the user
- Facilitate in the rehabilitation of householders (by giving prompts that be auditory and/or visual)

## 2.3 Hierarchical classes of smart homes

During the past years as a result of technical developments the possibilities of smart environments have risen tremendously. This increase in possibilities is physically visible in different types of realized smart homes. In order to classify smart homes, a functional perspective can be used. Since smart home technology is changing rapidly the functional perspective provides a better framework than a technological point of view.

With a focus on the functionality available to the user Aldrich [46] proposes five hierarchical classes of smart homes:

- Homes which contain intelligent objects - homes contain single, standalone applications and objects which function in an intelligent manner
- Homes which contain intelligent, communicating objects - homes contain appliances and objects which function intelligently in their own right and which also exchange information between one another to increase functionality
- Connected homes - homes have internal and external networks, allowing interactive and remote control of systems, as well as access to services and information, both within and beyond the home
- Learning homes - patterns of activity in the homes are recorded and the accumulated data are used to anticipate users' needs and to control the technology accordingly
- Attentive homes - the activity and location of people and objects within the homes are constantly registered, and this information is used to control technology in anticipation of the occupants' needs

Since broadband is becoming more widespread, available smart homes are shifting within the hierarchy from homes which contain intelligent, communicating objects to connected homes. Learning and attentive homes do exist but only in demonstration settings. The technology used in those homes is in most instances still experimental.

## 2.4 Products and services provided by Smart Homes

Within the field of smart home technology products and services play an important role in creating benefits for users. In general products and services can be divided into six categories:

1. Comfort
2. Energy management
3. Multimedia and entertainment
4. Healthcare
5. Security and Safety
6. Communication

The division is not strict, the different categories overlap. Mostly people don't choose just one category but several. The different categories strengthen each other leading to the fact that overall functionality of several categories combined is more than the sum of functionality of the independent categories. The following figure (Fig.1) shows an example of a smart home environment that combines products and services belonging to all the aforementioned categories.

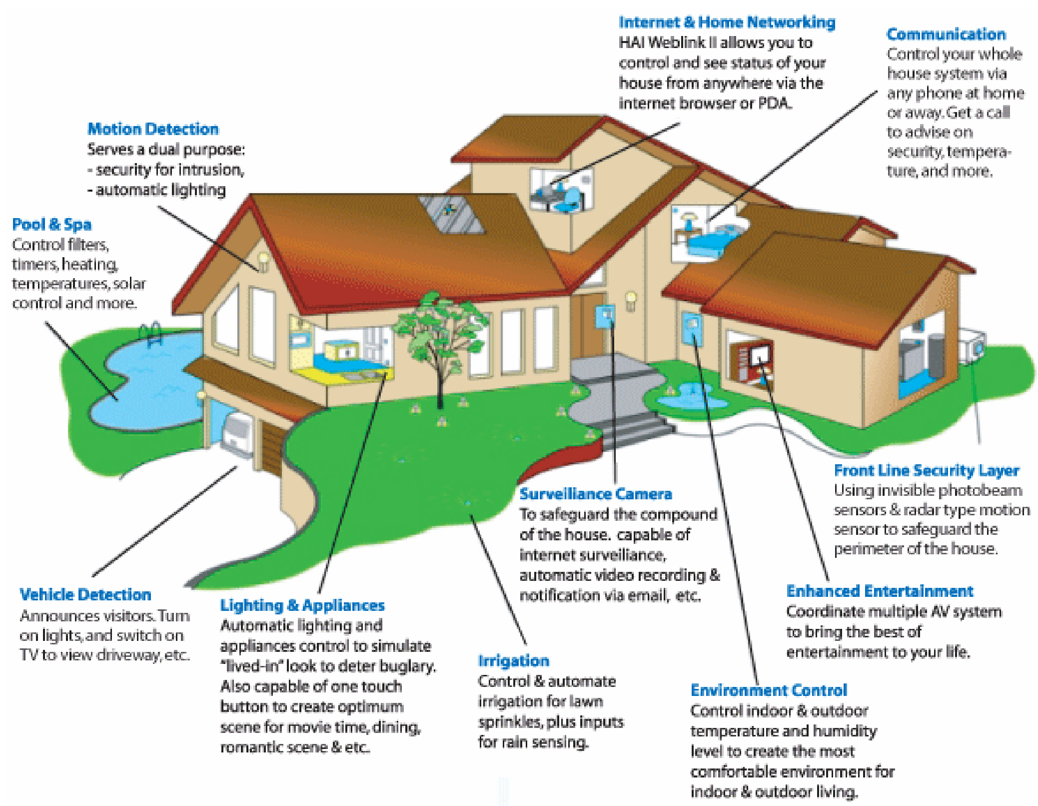


Fig. 1: A smart home solution[1]

It is impossible to give an extensive overview in this chapter, since new products and services are developing at such a rapid pace whilst by combining solutions new products and services are formed.

The emphasis in this section is on products and services belonging to the category healthcare, since we focus on smart home environment for elderly. We give some examples of solutions that fall within the different categories.

On a global level, healthcare solutions can be divided in the following categories:

- Active alarm systems: utilize remote emergency systems - usually telephone based - installed in the home of older persons
- Passive alarm systems: do not require the interaction of the person. For example, devices include sensors that are able to recognize the danger of a fall or a fire and send an emergency call automatically
- Remote support for care staff: include all kinds of telecommunication-based activities supporting the work of field staff

[1] Image courtesy of HomeControlPlus, http://www.homecontrolplus.net/solutions.html

- Remote support for family carers: includes all kinds of telecommunication-based activities supporting family carers
- Advanced services using video telephony: include remote monitoring and video-based alarm services
- Telemedicine

Alarm systems play a very important role in Smart Home applications. However, the systems, existing so far, are very often too complicated or not reliable enough. The key issue is the setting off of alarms. If, for instance, persons have an accident (fall, injury), a fire breaks out or a person suddenly becomes unconscious, then there is normally no time or possibility to operate a telephone or even an alarm button on their wrist or around their neck. Although there were many alarm systems on the market and several research projects have looked into this particular issue exist, there is still a large knowledge gap as to how persons react in a dramatic situation, be it panic or a collapse or simply because they are confused or have memory problems (forget that they are wearing an alarm button).

Probably the best solutions to overcome these problems are passive systems. Passive means an automatic control of vital functions (e.g. pulse, blood pressure, oxygen saturation), their evaluation and an automatic alarm being set off when the values of the vital parameters exceed predefined limits. The reliability of the alarm being set off can be dramatically increased when several different observations and decisions are combined. Among the measurement of vital parameters as described above, the person's activity (leaving and entering rooms, using water, electric light, TV and radio and many more) can give important additional information.

In order to avoid false alarms the receiver of the alarm (e.g. call centre or relatives) sends back a signal to the user for instance by telephone or via a message that appears on the alarm module worn on the wrist or around the neck. In the case of a false alarm, the user can answer the phone or press the button of the module indicating that no help is needed thus avoiding unnecessary attendance. If, on the other hand, the user does not react, it can be assumed that the alarm is serious.

One huge problem in healthcare is wandering and way finding. There are systems to detect where a client leaves the house where this would be inappropriate or dangerous for the client. These systems consist of magnetic contacts or pressure mats at/near hall door connected to local area (family) paging. The system, of course, does not restrict egress, but merely alerts a carer that the client has left. Similarly, for way finding (at night) lighting strips and passive infra-red light switches can assist and reduce the likelihood of a fall or disorientation around the house at night.

An area that receives an increasing amount of interest is tele-monitoring or personal health monitoring is based on the idea that persons can monitor themselves in their home using medical devices. Health care monitoring at home enables continuous measuring of physiological parameters. It is possible to embed sensors in different places or objects at home (e.g. in the furniture, electrical appliances), or to make them wearable by integrating them into clothing "Smart Shirt" or small apparel items such as watches and jewellery. By combining these wearable sensors with measurement devices embedded in home surroundings, advanced health monitoring may be achieved.

Recording of physiological and psychological variables in real-life conditions could be especially useful in management of chronic disorders or health problems; e.g. for high blood pressure, diabetes, anorexia nervosa, chronic pain, or severe obesity. Tele-monitoring could also be used to provide feedback about someone’s health in the form of behavioural feedback in order to prevent diseases.

Obviously tele-monitoring has many advantages for both, the patient and the medical institutions. The patient can stay at home and does not have the inconvenience, associated with a visit of the doctor or in the hospital, and the medical institution saves time because there is no need for spending time with routine work and in the hospital enormous costs for the bed and the care of the medical staff can be saved. In several tele-monitoring applications the data are not directly transmitted to the medical institutions but to a kind of ‘call centre’, which performs a first data evaluation.

## 2.5 Related projects

Several projects have introduced smart home technologies for aging people in the United States, Europe and Asia. The SmartBo project [50] in Sweden, for example, includes a two-room ground floor demonstration apartment containing technology for elders with mobility impairments and cognitive disabilities, such as dementia and developmental disability. Devices and sensors control lighting, windows, doors, locks, water outlets, electrical power, stove, as well as visual and tactile signalling devices, speech synthesizers, and Braille displays.

The PROSAFE project [51] in France examines devices and sensors that identify abnormal behaviour that can be interpreted as an accident and collect representative data on patients' nocturnal and daily activity. The experimental room has been designed to accommodate patients with Alzheimer’s disease. It is equipped with a set of infrared motion sensors connected to either a wireless or wired network.

The Hospital Without Walls [52] is a home telecare project in Australia that includes a wireless fall monitoring system in which patients wear small sensors that measure heart rate and body movement. A home computer records data from several sensors and uploads to a remote assessment center.

The Aware Home [53] developed by Georgia Tech in the USA is developed within two identical independent living spaces allow for controlled experiments with technology and enable inhabitants to live on one floor while demonstrating prototypes of assistive technologies on the other floor.

The University of Florida has built a 500-square-foot smart house that is designed assist and to provide medical care to Matilda, a life size mannequin [9] , [5] . The house implements devices including “a microwave that recognizes entrees and automatically determines how long to cook them” and devices to track the individuals location within the home. The house also uses devices to detect water on the floor and a camera that allows the person to view who is at the door and let them in using a cell phone. The Gator Tech Smart House in Gainesville [16] , developed by the Mobile and Pervasive Computing Laboratory at University of Florida, is a new construction house that's designed to assist older persons in maximizing independence and maintaining a high quality of life.

The CASAS Smart Home project [54] is a multi-disciplinary research project at Washington State University focused on the creation of an intelligent home environment. It approaches a smart home as an intelligent agent that perceives its environment through the use of sensors, and can act upon the environment through the use of actuators. The CASAS Smart Home aims at minimizing the cost of maintaining the home and maximizing the comfort of its inhabitants by adapting to information provided.

A home that uses the latest smart technology to give people with dementia and other serious long-term health conditions greater independence has been developed by the Bath Institute of Medical Engineering (BIME) in the School for Health at the University of Bath [6] [6] . This project has been designed to help people readjust to living on their own after a stay in hospital, and aims to reduce the risk of users being readmitted to hospital or going into long term care.

The Home Depot Smart Home is a 6,000-square-foot live-in research laboratory operated by Duke's Pratt School of Engineering [8] . The Home Depot Smart Home, part of a Duke Smart Home Program, creates a dynamic "living laboratory" environment that contributes to the innovation and demonstration of future residential building technology.

The Amigo project [10] is a joint effort of fifteen European companies and research organizations in mobile and home networking, software development, consumer electronics and domestic appliances. The Amigo project develops middleware that dynamically integrates heterogeneous systems to achieve interoperability between services and devices. For example, home appliances (heating systems, lighting systems, washing machines, refrigerators), multimedia players and renderers (that communicate by means of UPnP) and personal devices (mobile phones, PDA's) are connected in the home network to work in an interoperable way. This interoperability across different application domains can also be extended across different homes and locations.

The Smart Home project at Tampere University of Technology [11] is focused on building a smart home that resembles a normal 69m$^2$ flat with a bedroom, living room, kitchen, sauna and bathroom. The home is used for testing prototypes, usability testing, measurements and product demonstrations in the technologies of wired and wireless networks, lighting control, speech control and user interfaces (cellular phones, traditional buttons and switches, remote controls and www-based user interfaces).

# Chapter 3
# Indoor localization

## 3.1 Indoor localization

Localization refers to the task of determining the location of a traveller in a specified coordinate system, which is subject to topological constraints, using a mobile device carried by the traveller. Perhaps the most common example is terrestrial localization in the WGS84 reference frame using a mobile GPS receiver and the infrastructure of GPS satellites. Recent years have witnessed a rapid commoditization of GPS hardware and related products, as well as other similar technologies, so that outdoor localization is now inexpensive and accessible.

Indoor localization refers to the task of localizing a traveller, using a suitable mobile device, in typical indoor environments, such as home and office buildings, airports, and railway stations. Although there have been notable advances in this area as well, the progress in indoor localization has lagged that in outdoor localization.

In some aspects, indoor localization is simpler than outdoor localization. For instance, the geographical area covered is much smaller, and the expected speeds of travel are much lower. However, some other aspects make indoor localization much more challenging. For outdoor environments GPS provides an effective solution to determine location of GPS enabled mobile devices. Indoors and in crowded city streets however, the effectiveness of GPS is greatly diminished. Topological constraints in indoor environments are much more complex than those in outdoor environments. In addition to problems receiving reliable signals from satellites, or other outdoor beacons, in complex indoor environments, the requirements of indoor localization are also more stringent. For example, a vertical positioning uncertainty of several meters does not pose much of a problem for outdoor GPS applications such as route guidance for cars.

However, inside a building, that uncertainty translates into uncertainty in the floor of the building. More generally, topological constraints in indoor environments are much more complex than those in outdoor environments. For these and other reasons, prior work on indoor localization has looked to several alternative methods, such as the use of visual markers that are detected by ultrasonic signals, RFID, Bluetooth and 802.11.

## 3.2 Applications of indoor localization

While the need for estimating the physical position of humans and objects has accompanied mankind since its first steps, it becomes apparent that in the future, the importance of localization will even grow. New breakthroughs, particularly in information technology, sensors and micro-electronics, enable ever new applications that demand location information at an ever growing accuracy or lower cost. Many of these applications are situated indoors.

The range of traditional and emerging applications requiring indoor position data is manifold. A large classic part is in the industrial and logistics sector. Automated or self guided vehicles and robots are instructed to transport goods in an automated production facility or warehouse. In order to navigate autonomously or by remote control, they must localize themselves in their surroundings. Besides these traditional applications, indoor localization has experienced additional efforts in recent years with the appearance of ubiquitous computing. In various scenarios objects and goods must be located or tracked, e.g. in an industrial or medical environment. Additionally, people localization enables a large number of applications and services. Customers and employees can be observed, intruders detected and patients tracked. In public places like subway stations, safety concerns can be addressed by an emergency guidance system. Another application that has raised much interest in recent years is ad hoc sensor networks. Besides distributed sensing it allows for location aware routing which may improve network efficiency and lower power demands. Some applications of indoor localization are summarized in the following table (Fig. 2).

| Industrial | Logistics | Commercial | Medical | Smart house | Public places | Tourism | Sensor networks | Target | Purpose |
|---|---|---|---|---|---|---|---|---|---|
| ☒ | ☒ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | AGVs and robots | Navigation in production facility or warehouse |
| ☒ | ☒ | ☒ | ☒ | ☒ | ☐ | ☐ | ☐ | Goods | Finding goods/objects |
| ☐ | ☐ | ☒ | ☐ | ☐ | ☐ | ☐ | ☐ | Customers | Profiling customers' habits, shopping assistance by navigation |
| ☐ | ☐ | ☐ | ☐ | ☒ | ☐ | ☐ | ☐ | People | Smart house climate control, convenience features, intruder detection, electronic home arrest monitoring |
| ☐ | ☐ | ☐ | ☒ | ☐ | ☐ | ☐ | ☐ | Patients | Finding patients |
| ☒ | ☐ | ☒ | ☒ | ☐ | ☐ | ☐ | ☐ | Employees | Observation and profiling |
| ☐ | ☐ | ☐ | ☐ | ☐ | ☒ | ☒ | ☐ | People | Emergency guidance systems |
| ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☒ | Nodes | Location aware routing, distributed sensing |

Fig. 2: Applications of indoor localization[2]

Nowadays there is an immense interest in the field of Location-based services (LBS). Location-based services make use of location information about an entity to offer services based on this information. For example, services can use either the location of the user and offer certain services to the user based on the user's location, or they can use the location of one entity and offer services to another entity. An example of the first case could be a Personal Digital Assistant (PDA)-based tour guide that gives information about the sights in the user's current location. An example of the other case could be a system that keeps track of where your co-workers are, so that they can be easily located, should you need to meet them.

## 3.3 Indoor localization choices

Localization process involves several issues and characteristics that need to be defined in order to develop a localization system. Each choice on these issues delivers certain strengths and weaknesses and thus, it may be suitable in one application over the other.

### 3.3.1 Physical position and symbolic location

Physical position is the position obtained from the system using physical co-ordinates such as those obtained from a GPS device. The GPS device usually provides the user

[2] Image courtesy of Holger Linde - On aspects of Indoor Localization [39]

with co-ordinates in the form of lateral and longitudinal degrees as well as altitude with respect to the sea. On the other hand, a symbolic location is a location that has a meaning in its context. For example, a symbolic location system could inform users that they are in the living room instead of giving them co-ordinates. Usually, a physical positioning system can provide the co-ordinates to another interface that would query a database of symbolic locations in order to emulate a symbolic location system.

### 3.3.2 Absolute and relative positioning

In an absolute positioning system, if two mobile devices are located at the same spot, their position will be the same (for example, in a GPS system). On the other hand, if a relative positioning system is used, the device will report its location with respect to a specific reference point or in case of locating other devices, with respect to itself. As the case with physical position and symbolic location, either systems; absolute or relative, can be augmented to simulate the other, though with limited success. In case of an absolute positioning system, the system could emulate relative positioning by searching for a reference point and modifying the reported location to be with respect to that reference point. In case of relative positioning, one could use the knowledge of the position of other reference points in order to determine the absolute location.

### 3.3.3 Location Computation

The location computation parameters define whether the client is the one that calculates the location using a specific algorithm or that it is a centralized server that manages this calculation. This choice would depend upon the size as well as the designated power consumption of the client. If the client computes its location (for example using triangulation from GPS signals), it would need to have sufficient computational power in order to calculate its position in adequate time. This computational power would translate as more electronic power consumption. On the other hand, a device that relies on an external server to locate it by broadcasting a beacon signal or receiving the locations directly from the server would not need much computational power but would mean that the server would have to undertake more responsibilities which might affect the number of clients that could be served simultaneously. On another note, having the server compute the clients' locations and possibly saving them raises privacy concerns.

## 3.4 Demands and performance measures for indoor localization systems

In order to judge an implementation, its performance parameters must be analyzed. In this section we discuss possible performance measures for indoor localization systems.

One of the major performance parameters of a localization system is its accuracy. This term indicates how far the estimated location deviates from the true location. Accuracy is often confused with precision. Precision is a measure of the reproducibility of a measurement. Therefore, when a statement is made about the resolution of a localization system, accuracy is commonly supplemented by precision (e.g. 20 cm accuracy over 90% of the time). For the applications mentioned above, the required localization accuracy ranges roughly from ±5 cm (robot positioning) to about ±1m (guidance systems). In some cases, even less accuracy is sufficient, for instance in people detection within a room.

Cost is a critical issue, particularly in applications where numerous units have to be equipped with localization technology, such as ad hoc sensor networks. It is therefore reasonable to use inexpensive standard off-the-shelf components whenever possible. Besides the mobile targets, infrastructure, deployment and maintenance cost must be considered.

Power consumption must be considered, especially when a large number of battery driven mobile units are involved (e.g. mobile phones).

Another point is scalability. On the one hand, the system's outer borders must be easily extendable (e.g. by adding further rooms to the covered area). On the other hand, the system must be able to cope with an increasing number of mobile targets that need be located concurrently.

The estimation cycle length, i.e., the time between two location updates, must be considered in some cases. For instance, autonomous robots will act blindly until the next position update. If the update cycle is too long, they can only operate reliably at a very low speed in order to avoid collisions.

Robustness is another aspect that need be mentioned. This includes the level of immunity against interference from other sources, but also from ambient variations such as temperature and humidity etc.

Particularly in those systems where active tags are carried around by humans, device size must be taken into account. If these are large and heavy, the positioning system might be turned down by its users for concerns of convenience.

At last, the utilized technology must be compatible with human beings that work or live within range of the localization system. For instance, acoustic range measurement using artificial sound sources is unacceptable. Many domestic animals can perceive sound at frequencies much higher than humans. Hence they might be affected by systems using ultrasound.

# Chapter 4

# Technologies and techniques for indoor localization

## 4.1 Technologies for indoor localization

A variety of systems have been proposed to achieve positioning depending on different technologies such as Bluetooth, Wi-Fi, RFID, infrared or ultrasound signals, radio frequency signals, GSM/UMTS, GPS and also image matching through scene analysis. An overview of the common real time locating systems is sketched in Fig. 3. Each of these methods has its strengths and weaknesses and also delivers different results in varying circumstances (outdoor - indoor positioning, topology, etc.).

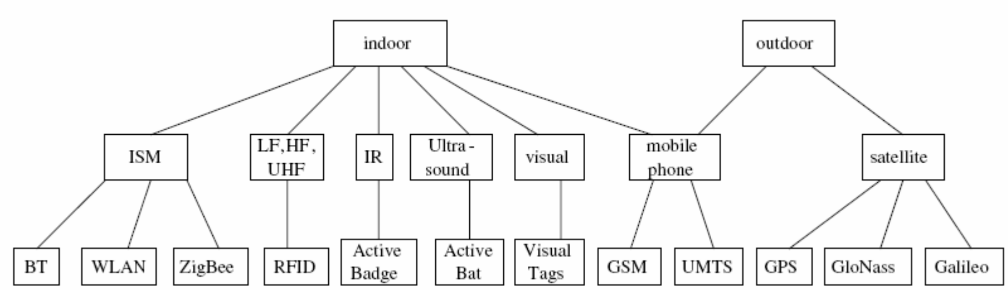


Fig.3: Classification of localization systems[3]

[3] Image courtesy of J. Oliveira Filho, A. Bunoza, J. Sommer, W. Rosenstiel, Self-Localization in a Low Cost Bluetooth Environment [35]

In this work we concentrate on achieving indoor localization in the framework of a Smart House designed for aging people. In such a system the cost factor is much more than important, as a less expensive system could potentially become widespread. For that reason we concentrate on technologies that require no complicated infrastructure, such as Bluetooth, Wi-Fi and mobile-phone cameras that can be used in developing a low cost localization mechanism. In the following section we focus on these technologies, summarizing their strengths and weaknesses with respect to indoor localization.

### 4.1.1 Indoor localization using Bluetooth

Bluetooth is a stable, inexpensive, and mature technology upon which a location aware infrastructure can be built [27] [28] [30] [31] [32] [33] [35] . The Bluetooth technology is one of the most promising and cost-effective candidates to provide short-range wireless connectivity for mobile users. Most modern mobile phones feature Bluetooth for short range communication, so this method is cost-effective to deploy. Key features of this technology are the robustness, the low complexity, the low power consumption and the short range.

If Bluetooth is enabled on a mobile phone, it can detect and identify other Bluetooth enabled devices within range limits. If we build up an infrastructure of pre-positioned devices (fixed nodes), that are periodically scanning their proximity for other devices, we can evaluate which fixed nodes can recognize other mobile devices, and furthermore, can deduct their location. For a fixed node can only detect devices in a limited range (typically about 10m) the position of mobile devices can be pinpointed to quite a small area. An important advantage of this method is that it allows the mobile device to determine its location while remaining anonymous, unidentified to the beacons or other nearby devices, by performing the scanning itself.

Moreover, some of the Bluetooth signal parameters, such as RSSI, Link Quality, Received and Transmit Power Level, could be used to achieve indoor localization through the scene analysis technique.

The Bluetooth standard was devised for short-range (few meters ≈10m) communication and therefore most Bluetooth devices have very limited range. The short range, compared to alternatives such as 802.11, has some disadvantages: For instance, a much larger number of devices must be deployed in order to provide adequate coverage. However, the short range is also a significant advantage. One of the problems identified by prior work on 802.11-based localization is that it is extremely difficult to accurately judge the distance from a beacon, using signal strength or other properties, due to complex signal propagation artifacts. With the short range of Bluetooth devices, this problem is significantly mitigated. In effect, if a mobile device detects a Bluetooth beacon then it is very likely that the beacon is only a few meters away. In contrast, an 802.11 beacon may be detected even if it is two buildings away. So, we could take advantage of this feature to yield an indoor localization scheme that is inexpensive and easy to deploy. Another benefit of Bluetooth is that the transceivers are extremely inexpensive, costing less than 3 USD in bulk. In addition, Bluetooth transceivers are designed for low power consumption.

The following table summarizes the advantages and disadvantages of Bluetooth technology with respect to indoor localization:

| Advantages | Disadvantages |
|---|---|
| • Most modern mobile phones are BT enabled<br>• Inexpensive (cheaper than wireless LAN components and adapters)<br>• Short range, higher accuracy<br>• Low power consumption | • Short range, need for more access points to cover the same area<br>• Low data rate, slower update of position<br>• RF Band is shared, possible interferences |

Table 1: Advantages and disadvantages of Bluetooth technology

### 4.1.2 Indoor localization using Wi-Fi

The IEEE 802.11 wireless Ethernet standard is becoming increasingly popular and has been deployed in many indoor environments. With the prevalence of wireless LAN infrastructure as well as the ever decreasing cost of wireless access points, the use of wireless access points as a basis for location determination becomes an attractive option.

Wi-Fi location determination systems use the popular 802.11 network infrastructure to determine the user location without using any extra hardware [33] [40] [42] [43] [44] [45] . The only additional hardware requirement for such a system is a Wi-Fi card (the receptor), if it is not present, although many modern mobile phones are Wi-Fi enabled. Most indoor localization systems are based on the fingerprinting technique as cell-based methods rely in more short ranged technologies, such as Bluetooth. However, a simple discovery system using Wi-Fi could be developed, by reducing the transmission power.

In order to estimate the user location, wireless Ethernet devices measure signal strength of received packets. This signal strength is a function of the distance and obstacles between access points and the receptor. The system might need one or more reference points (access points) to measure the distance from. Triangulation on signal strength from multiple access points could be the most natural technique to be applied but unfortunately, in indoor environments, the wireless channel is very noisy and the radio frequency (RF) signal can suffer from reflection, diffraction and multipath effect, which makes the signal strength a complex function of distance. To overcome this problem, Wi-Fi location determination systems uses a priori radio map (wireless-map), which captures the signature of each access point at certain points in the area of interest. Systems that make use of a wireless-map work in two phases: training phase and estimation phase. During the training phase, the system constructs the wireless-map. In the estimation phase, the vector of samples received from each access point is compared to the wireless-map and the "nearest" match is returned as the estimated user location. However, the accuracy of this method usually depends on a meticulous calibration procedure that consists of physically moving a wireless client to many different known localizations, and sometimes orientations, inside a building.

The following table summarizes the advantages and disadvantages of Wi-Fi technology with respect to indoor localization:

| Advantages | Disadvantages |
|---|---|
| • Higher data rates (11Mbps), faster update of position<br>• Longer range (approx. 100 meters), need for less access points to cover the same area<br>• Relatively inexpensive hardware (though more expensive compared to BT)<br>• Use of the existing wireless LAN infrastructure, no additional equipment (except Wi-Fi adapter for user's terminal, if needed) | • RF Band is shared, possible interferences<br>• Not all modern mobile phones are Wi-Fi enabled (fewer than BT) |

Table 2: Advantages and disadvantages of Wi-Fi technology

### 4.1.3 Indoor localization using mobile phone's camera

Camera-equipped mobile phones are being put to many uses these days. Following a similar approach to scene analysis, the mobile phone's camera could be used in order to achieve indoor localization [38] [55] .

Locating a user is performed by matching the image of what he/she is actually seeing with a database of different images taken for the place considered. The camera-phone, worn by the user, captures images periodically and sends them to a web server. The web server has a database of images with their corresponding location. Upon receiving an image, the web server compares it with stored images, and based on the match, estimates user's location.

This localization approach does not have any infrastructure requirements, as a digital camera that can be found is almost all today's mobile phones. The only cost is that of building an image database.

The following table summarizes the advantages and disadvantages of using a mobile phone's camera for the purpose of indoor localization:

| Advantages | Disadvantages |
|---|---|
| • Minimized infrastructure requirements<br>• Low cost<br>• Low complexity | • Construction of an image database with user involvement for the tagging process<br>• Low accuracy in varying lighting conditions |

Table 3: Advantages and disadvantages of using a mobile phone's camera

## 4.2 Localization techniques

In this section we overview the most popular techniques for indoor localization using Bluetooth, Wi-Fi and image matching through mobile phone's camera. These are cell-based methods, range-combining techniques and scene analysis or fingerprinting.

### 4.2.1 Cell-based methods

Cell-based methods [29] determine the location of the traveller based on only the visibility of beacons, without using any distance, angle or signal parameters measurements. Localization is based on the knowledge of the limited range of each of the beacons, allowing the traveller to be localized to the region of intersection of the ranges of all visible beacons. Given the problems with many of the other methods, cell-based methods are quite popular for Bluetooth, as well as for RFID and infrared (IR) technologies.

An example of how the cell-based localization approach works is shown in Fig.4. Consider five beacons, A, B, C, D, and E, with ranges of varying shapes and sizes. Now suppose the traveller is at some location from which beacons B, C, and D are visible, while beacons A and E are not visible. We may conclude that the traveller is located in the shaded region. Note that we use information on both visibility and non-visibility of beacons to determine the region of the traveller. So, if we assume that the location and range of each beacon in an indoor environment is known, then it is simple to determine the location of a travelling mobile device based on the visibilities of beacons using the method illustrated by the example above.

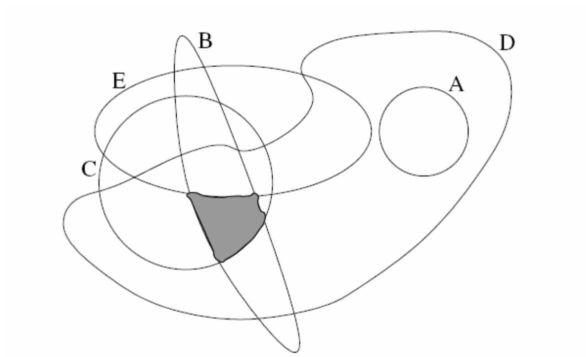


Fig. 4: Cell based localization[4]

An important consideration in cell-based methods is the shape and size of each cell, i.e., the range of each beacon. In outdoor applications, it is often reasonable to assume cells of a regular shape, such as an ellipsoid. However, in indoor applications, such an assumption is not realistic due to channelling and other artifacts of buildings and their contents. For example, it is common for signals to travel along hallways for large distances, but to attenuate rapidly in a transverse direction due to intervening walls and

[4] Image courtesy of Sudarshan S. Chawathe, Beacon placement for indoor localization using Bluetooth[29]

equipment. Similarly, interior windows and doorways, and details such as the construction material, all significantly affect the range of beacons in different directions.

### 4.2.2 Range-combining techniques

In range-combining techniques we use signal characteristics information and perform calculations to determine the location of an entity in relation to some other infrastructure devices. Examples of range-combining techniques are triangulation, trilateration, multilateration, time-of-flight and attenuation.

Triangulation requires the measurement of angles between the line connecting two beacons and the line of sight from each of the beacons to the traveller (Fig. 5). It is therefore not suitable for Bluetooth and other radio frequency technologies, but may be successfully applied for a system of visual beacons. The related rho-theta method, based on measuring the distance and angle to a single beacon, suffers from a similar drawback for RF technologies. We note that sometimes the term triangulation is used as a synonym for trilateration, described next, but the two are quite different.

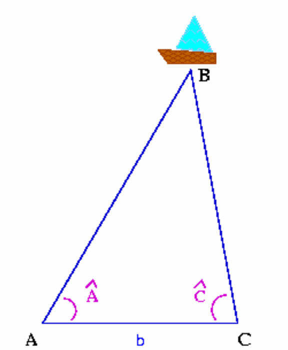


Fig. 5: Triangulation example[5]

Trilateration falls under the category of triangulation where the geometric properties of triangles are used to compute the objects' locations. It requires the computation of the distance of the traveller from each of three beacons (Fig. 6). The traveller's position is then uniquely determined as the point of intersection of three circles, each centered at one of the beacons. The radius of each of these circles is the computed distance of the traveller from that beacon. Prior work has explored the use of received signal strength parameter as a proxy for distance, but the results are not encouraging. In general, the

[5] Image courtesy of Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Triangulation.png

correlation between distance and signal strength is not sufficiently high because of a variety of radio artifacts, as well implementation features such as automatic power management by the hardware based on signal strength. By very carefully controlling the transmit power management features, it may be possible to obtain a mean absolute positioning accuracy of 1.2 m.

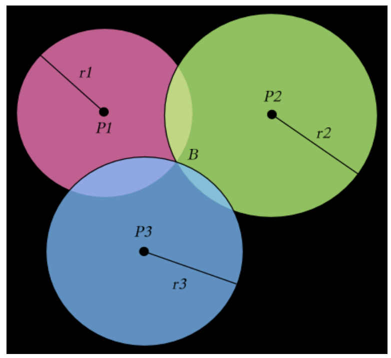


Fig. 6: Trilateration example[6]

Multilateration-based localization is based on measuring the time intervals between the transmission of a pulse from the traveller and its and reception at multiple receivers. Such measurements require features not supported in the low-cost commodity devices. Further, even if such features are available, the Bluetooth specification permits a clock jitter of 10 microseconds, which translates into a measurement error of roughly three kilometres, making it unsuitable for most localization applications. However, multilateration is an attractive alternative when using slower signals, such as sound.

Time-of-flight applies by recording the time it takes for a signal to propagate from a transmitter to a receiver at a known velocity. Since this method involves time measurement, a specific clock resolution must be attainable by the system according to the timing technology used. Ultrasound based systems do not require a very good resolution as the waves travel at the speed of sound so the differences to be measured are usually in the milliseconds range. On the other hand, systems using RF signals such as WLAN and Bluetooth would need to have a much higher clock resolution since radio waves travel at the speed of light so the time difference is usually in the nanosecond range which is 6 times less in the order of magnitude than the ultrasound based technologies.

[6] Image courtesy of Wikimedia Commons, http://commons.wikimedia.org/wiki/File:Trilateration.png

In attenuation we take advantage of the relation that the intensity of the emitted signal is inversely proportional to the square of the distance from the source. Using this relation and given the signal strength at the source, the measured signal strength could be used to estimate the distance of the object from the source. Measuring the distance using attenuation is usually less accurate than using time-of-flight especially in environments with many obstructions which cause a lot of signal reflection.

### 4.2.3 Scene analysis – Fingerprinting

The scene analysis technique depends on analyzing the "scene" to obtain features that are easily compared and represented. A scene can be described either by images (image matching technique) or by signal property readings at that location (Bluetooth, Wi-Fi). Signal properties that have been used with the fingerprinting technique in prior work are Link Quality, Received Signal Strength Indicator, Transmit Power Level, Response Rate and Signal-to-Noise ratio.

Scene analysis can be static or differential. In static scene analysis, the observed features are looked up in the database that maps objects to locations. In differential scene analysis, the difference between successive scenes is tracked. Those differences will imply that the user has moved. If features are known to be at specific locations, the observer can compute its location relative to them. The advantage of such system is that it can act independently from an external server by storing the database locally which means that less power is required for data transmission and the privacy of the user could be maintained. The disadvantages include the need for a pre-built lookup database of the features that might need to be rebuilt if the features of the scene change.

Fingerprinting is a type of scene analysis and consists of two stages:

The first stage is the **training stage.** At specific points called reference points within the environment, a chosen signal property at each point is recorded and saved in a database. For the fingerprinting database to be accurate, a significant number of reference points should be made available which increases the time needed for the training phase. In the case of the image matching technique, a set of photographs for each reference point should be captured and saved in the database.

The second phase which is the **matching stage** and the actual positioning stage where the object measures the chosen signal property at its current position and then, using a matching algorithm, the measured value is matched with vales stored in the database in order to obtain an approximation on the user's current position. The matching could be done on the client side which requires that the database of fingerprints is stored locally, thus increasing the storage and memory requirements on the client. The other option is matching on the server side which increases the load on the server and possibly the load on the network. The algorithm by which the location of a user could be determined could either be deterministic or probabilistic. When we use images as fingerprints, image retrieval algorithms are used in order to find the best match from the constructed database. A more detailed description of these techniques follows in the next chapter.

# Chapter 5

# Content-based image retrieval

# & Optical character recognition

In the following chapters of this thesis we propose two applications in the context of a Smart Home environment for elderly. The first one is an indoor localization system that uses images captured by the mobile phone's camera to determine the user's position. The second application, proposed to help visually impaired people, performs optical character recognition on photographs of medication boxes, attempting to discover their identity. Due to the above mentioned applications, it is important to discuss further the subjects of content-based image retrieval and optical character recognition technology.

## 5.1 Content-based image retrieval

Content-based image retrieval (CBIR), also known as query by image content and content-based visual information retrieval is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases[24] [25] [26] . "Content-based" means that the search will analyze the actual contents of the image. The term "content" in this context might refer to colours, shapes, textures, or any other information that can be derived from the image itself. When it is not possible to rely on metadata of images such as captions or keywords, image retrieval can be based on colour, texture or shape in the content of an image.

Retrieving images based on colour similarity is achieved by computing a colour histogram for each image that identifies the proportion of pixels within an image holding specific values (that humans express as colours). Current research is attempting to segment colour proportion by region and by spatial relationship among several colour regions. Examining images based on the colours they contain is one of the most widely used techniques because it does not depend on image size or orientation. Colour searches will usually involve comparing colour histograms, though this is not the only technique in practice.

Texture measures look for visual patterns in images and how they are spatially defined. Textures are represented by texels which are then placed into a number of sets, depending on how many textures are detected in the image. These sets not only define the texture, but also where in the image the texture is located. Texture is a difficult concept to represent. The identification of specific textures in an image is achieved primarily by modelling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality may be estimated.

Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out. Shapes will often be determined first applying segmentation or edge detection to an image. In some cases accurate shape detection will require human intervention because methods like segmentation are very difficult to completely automate.

In the proposed indoor localization system we used three off-the-shelve algorithms for image retrieval. The first two algorithms retrieve similar images based on colour (AForge.NET framework [17] [20], Eye.Open image processing library [18] ), while the third (Image Recognition project [19] ) combine image retrieval based on shape and colour.

## 5.2 Optical Character Recognition (OCR)

OCR is a field of research in pattern recognition, artificial intelligence and machine vision [24] . Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic translation of images of handwritten, typewritten or printed text into machine-editable text.

Though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques. Optical character recognition (using optical techniques such as mirrors and lenses) and digital character recognition (using scanners and computer algorithms) were originally considered separate fields. Because very few applications survive that use true optical techniques, the OCR term has now been broadened to include digital image processing as well.

In recent years, there has been increased interest in OCR, driven by a number of factors. Search engines have raised the expectation of universal access to information on-line, and cheap networking and storage have made it technically and economically feasible to scan and store the books, newspapers, journals, and other printed materials of the world.

Early OCR systems required training (the provision of known samples of each character) to read a specific font. "Intelligent" systems with a high degree of recognition accuracy for most fonts are now common. Some systems are even capable of reproducing formatted output that closely approximates the original scanned page including images, columns and other non-textual components.

Commercial OCR engines have traditionally been optimized for desktop use–scanning of letters, memos, table form processing and other end-user documents, while some other engines have been optimized for special applications like bill scanning, address scanning, signature verification and reading aid for the blind.

# Chapter 6

# Implementation of an indoor localization system based on mobile phone's camera

We propose an indoor localization mechanism that relies image matching to locate a mobile user (scene analysis technique). The benefit of this approach lies in its ability to locate the user without the cost of additional hardware.

## 6.1 Overview

With this application we explore the possibility of determining user's location indoors based on what the camera-phone "sees". The camera-phone is worn by the user as a pendant (Fig. 7), which captures images periodically and sends them to a web server over FTP. The web server maintains a database of images associated with their corresponding locations. Upon receiving an image, the web server compares it with stored images, and based on the match, estimates user's location. We accomplish this with off-the-shelf image matching algorithms, by tailoring them for our purpose. We built an image database for a small apartment with nearly two hundred pictures to account for real-life issues such as varying heights of the users, different angles that may correspond to the same image, etc. The key advantage of using this approach is that it does not require any infrastructure. Neither custom hardware, nor wireless access points are required. Physical objects do not have to be "tagged" and users do not have to carry any device apart from what they usually do: a mobile phone. The only cost involved is that of building an image database.

Figure 7: User wearing the phone as a pendant[7]

## 6.2 Related work

The scene analysis approach has been previously used in the robotics field, where a robot is equipped with a digital camera. Image taken by the camera is matched against a database of images to identify the location.

One of the earliest projects that used camera images for indoor localization was the Easy Living project proposed by Microsoft [55] , [37] . In this work, a set of three cameras were used to capture 3D images and use them for scene analysis. The problem with this approach was its high cost and high processing power to process the 3D images, in addition to the requirement of extra hardware that are not available everywhere.

In [56] , the idea of indoor localization using image matching has been proposed in an adaptive tourism application, where, for each image taken by the user, a set of interest points is computed and matched with the interest points of all images stored in the database. The image with the highest match is the one chosen for recognizing the object the user is looking to.

In the system proposed in [38] , the user is assumed to wear a camera-phone as pendant that captures images periodically. Images are sent to a central server that maintains a database of images and a matching algorithm is run. The probability of success is in the average of 50% for corner-level accuracy.

An extended system of the above is proposed in [36] , where an alternative algorithm attempts to improve the accuracy of the matching algorithm while maintaining the processing overhead and the matching time at acceptable levels.

[7] Image courtesy of N. Ravi, P. Shankar, A. Frankel, A. Elgammal and L. Iftode, Indoor Localization Using Camera Phones[38]

## 6.3 Experimental environment, hardware and programming tools

We tested this application in the area of a small apartment around 50m$^2$. The mobile device that was used in our experiments was an HTC Touch Cruise (HTC P3650), operating on Windows Mobile 6 Professional. The camera on this mobile phone is a 3.0 megapixel camera. All the desktop and mobile applications were developed in Microsoft Visual Studio 2008 using Microsoft .NET Framework 3.5 SP1 in C#. The database for the needs of our experiments was built using Microsoft SQL Server 2008.

## 6.4 Localization approach

Fingerprinting is chosen as the localization approach as it is the only technique that can be used with a mobile phone's camera (Fig. 8). Generally, fingerprinting has a higher accuracy then other approaches such as range-combining techniques especially in indoor environments with multi path effects.

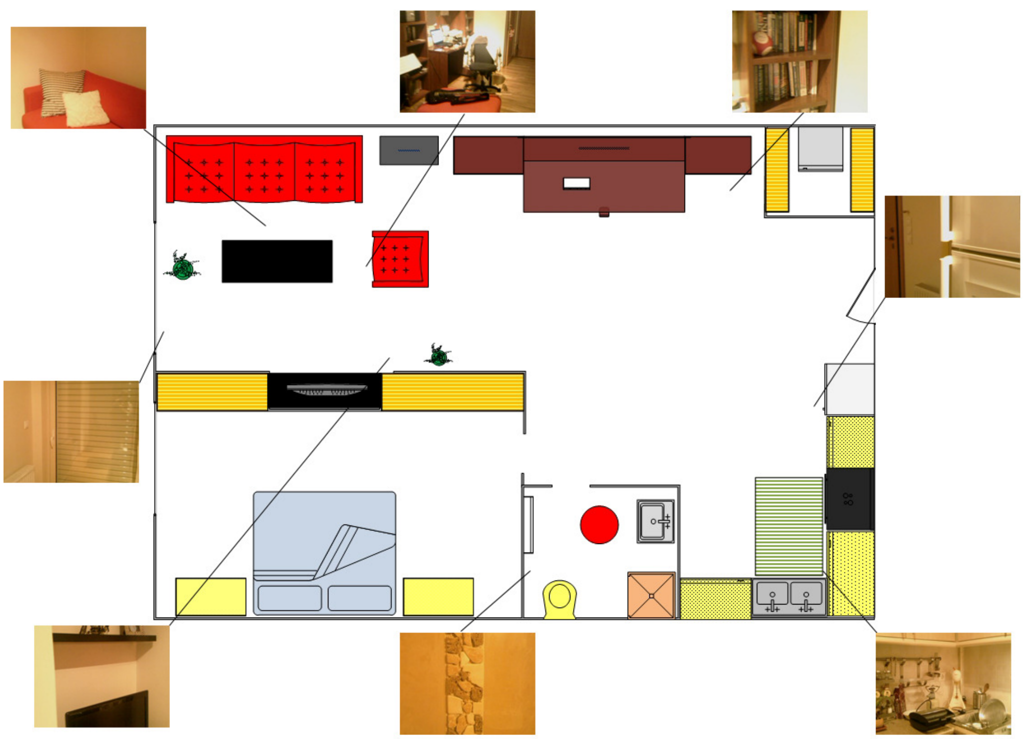

Fig. 8: Scene analysis using photographs

The steps followed for achieving localization are:

1. Create database with images and tag them manually
2. Take a query image and send it to the server
3. Compare it with the images stored in the database, using off-the-shelve algorithms for image comparison
4. Each algorithm assigns a weight to every image in the database which reflects the degree of similarity
5. If the weight of the best match is less than a certain threshold, discard the query image
6. Else return the location of the image that matches the query image with maximum weight

The first step belongs to the training stage of the fingerprinting technique, whereas the rest steps to the matching stage.

### 6.4.1 Training stage

To have better results, it is important to build an extensive database of images. A C# client runs on the phone and captures photographs (Fig. 9) periodically as the mobile user walks around and then sends the images to a web server through FTP.

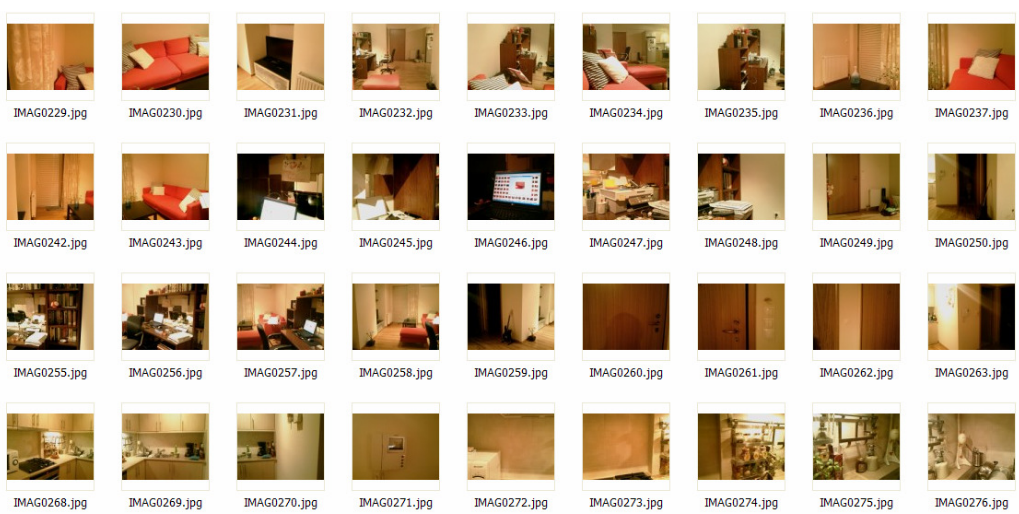


Fig. 9: Examples of photographs used in the training stage

The images are tagged with location manually afterwards and their location information along with some of their EXIF metadata is stored in a database (Fig. 10).

| | Column Name | Data Type | Allow Nulls |
|---|---|---|---|
| ▶🔑 | path | varchar(50) | ☐ |
| | locationX | smallint | ☐ |
| | locationY | smallint | ☐ |
| | date | smalldatetime | ☑ |
| | width | smallint | ☑ |
| | height | smallint | ☑ |
| | megaPixels | float | ☑ |
| | hDPI | smallint | ☑ |
| | vDPI | smallint | ☑ |
| | camera | varchar(50) | ☑ |
| | | | ☐ |

Fig. 10: The database table that stores the image information

To facilitate the tagging process, we decided that it would be more convenient to tag an image with the exact location on a house map rather than with a descriptive sentence. For that reason we constructed a house map of the apartment and we built an application where the user can click with his mouse at the exact location to tag an image. The location stored in the database is in the form of Cartesian coordinates (X,Y) referring  to the house map area. An example of this application is shown in Fig. 10.

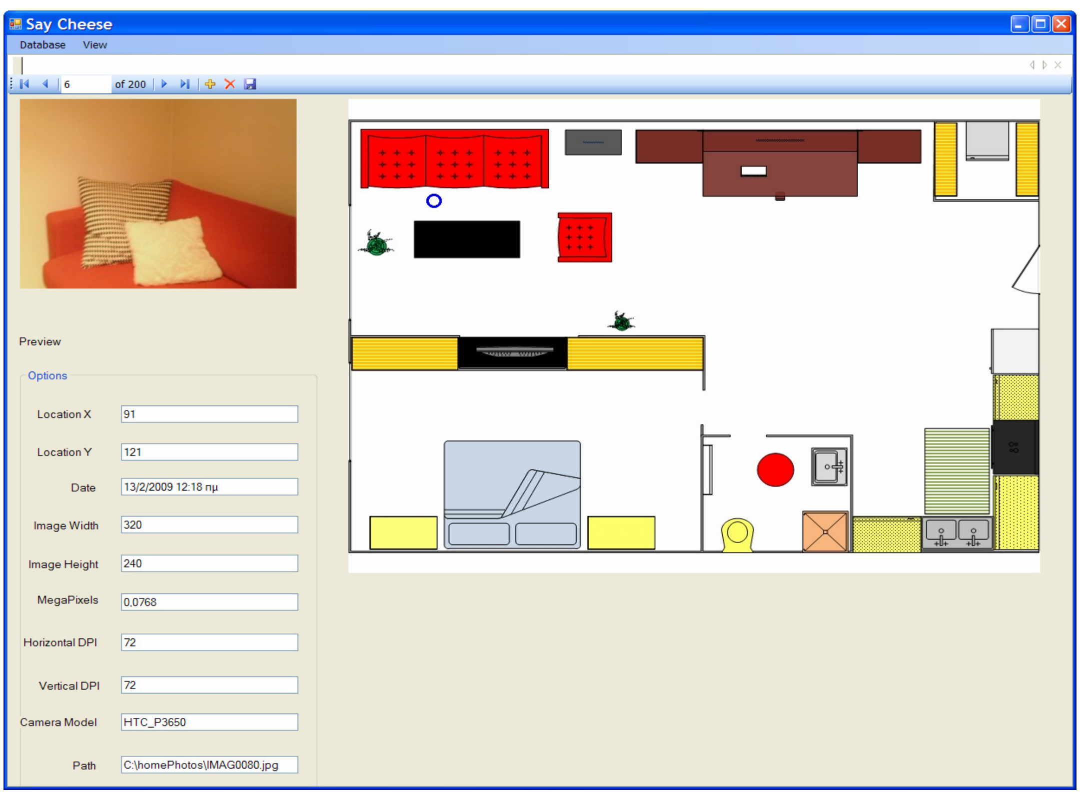


Fig. 11: The tagging process

The process of tagging images with location could be partially automated, by using a speech recognition interface on the phone, so that the database creator can tag images by announcing her location while pictures are taken.

### 6.4.2 Matching stage

After the creation of the database, we can proceed to the matching stage of the scene analysis technique, where we can test our localization system.

For the image comparison we use three off-the-shelf, open source algorithms for image retrieval developed in C#. The first algorithm is part of the AForge.NET framework [17] which provides an exhaustive template matching algorithm. The algorithm performs complete scan of source image, comparing each pixel with corresponding pixel of template. The second algorithm is part of the Eye.Open image processing library [18] and finds similar images using RBG horizontal and vertical projections. The third algorithm is found in the Image Recognition project [19] at CodePlex Open Source Community and performs image retrieval using edge detection and colour recognition.

When the web server receives a query image, it compares it with the images stored in the database. Every image included in the database is assigned a weight by each algorithm which reflects the degree of similarity between the two images. The similarity level is a float number in the interval [0,1], where 1 is the maximum degree of similarity between two images and thus, exact match. The total weight of an image is calculated as a linear combination of the weights assigned by each algorithm. If the weight of the best match is less than a certain threshold value, the query image is discarded. This is necessary to prevent wrong location updates from being sent to the client. Once the weights of the images in the database with respect to the query image are known, we can return the location of the image that matches the query image with maximum weight. A screenshot of the application that performs the matching stage is shown in Figure 12.


Fig.12: The matching stage

The query image is shown at the up-left corner of the application. We use a combo-box to return the top n matches from the image database, where $n = 1, \dots, 10$ . We draw the location returned on the house map as an ellipse, using different colours for each algorithm (AForge.NET, Eye.Open image processing, Image Recognition and linear combination of all algorithms).

At the right side of the application we print in a textbox the results of the matching stage. For this reason we defined the 1meter accuracy as success in the task of localization. To describe a matching result with the term "Success" or "Fail" programmatically, we tested our application with some images that we used in our training set. The database contains the exact location for these images in the form of Cartesian coordinates. In the returning matches we discarded the top match with similarity level equal to 1, which of course is the same as the query image and kept the second best match. The Euclidean distance between the query image and the second best match gives the accuracy of the location calculation, as the application behaves as the query image was never in the training set.

## 6.5 Evaluation

In order to evaluate the proposed indoor localization system, we created a database of 200 photographs of the apartment, following the method described in the matching phase of the fingerprinting technique. In the photographs of the training set we tried to capture every place on the house and especially the positions that match with daily life's activities (for example living room - watching television, kitchen - cooking, bathroom - taking a bath, etc). We tested our application using a set of 100 query images and examined the results derived from the weighted combination of the three image retrieval algorithms. As mentioned above, if the system's returned location was within 1 meter from the actual position of the user, the result was described as success, differently as failure. The user's location was determined correctly with an accuracy of 69%. The following table shows the accuracy of all algorithms used in the application in determining the user's position within 1 meter accuracy.

| AForge.NET | Eye.Open | Image Recognition | Linear Combination |
|---|---|---|---|
| 46% | 48% | 61% | 69% |

Table 4: Success rates - 1m accuracy

Although the algorithms used do not achieve a very good performance, we observe that the linear combination of all algorithms results in a better accuracy. In addition, we checked the success of our application in achieving room-level accuracy. For that purpose we added a room description (living room, bedroom, bathroom, kitchen, office) to each photograph. Table 5 shows the results for room-level accuracy:

| AForge.NET | Eye.Open | Image Recognition | Linear Combination |
|---|---|---|---|
| 62% | 56% | 72% | 80% |

Table 5: Success rates - Room level accuracy

Although the achieved accuracy is a point that needs further improvement, the proposed application and the experiments conducted prove that an indoor localization system based on photographs captured by mobile phone's camera is feasible. This is undoubtedly a first step towards low cost indoor monitoring with the purpose to discover patterns in daily behaviour of elderly and provide automated services within a Smart Home environment.

## 6.6 Future work

In this section we propose some ideas in order to improve the accuracy of this localization system. A first step to derive a better estimate of the user's current location is to use a history-based approach based on the past locations of the user. In this approach, the web server monitors the trajectory of the user and tracks his location in order to extract ordinary routes. Of course, to follow this approach we should alter the mobile application to capture more pictures in short intervals. When the server receives a query image, it looks at the last n-1 query images. The current location of the user is the one that maximizes the probability of seeing the n query images in the shortest period of time.

Moreover, checking the strengths and weaknesses of each image retrieval algorithm used in this application could result in finding a better weighted combination that optimizes the localization results.

Another idea to achieve better results is to combine this application with another localization system. For example, as we already take advantage of the mobile phone's capabilities, we could use the Bluetooth or Wi-Fi infrastructure that exists in most modern mobile phones.

Finally, we observed that the performance of this indoor localization system is extremely limited by the varying illumination conditions that occur inside a house during the day. It would be interesting to test this application with next generation cameras that capture higher quality photographs.

# Chapter 7

# Implementation of an application to help people with impaired vision

We propose an application to help people with impaired vision. The application performs Optical Character Recognition (OCR) on photographs of medication boxes, captured by a mobile phone's camera.

## 7.1 Introduction

A rapidly increasing proportion of the aging population experiences eye problems that make simple daily tasks difficult or impossible, even when wearing glasses or contact lenses.

Globally, in 2002 more than 161 million people were visually impaired, of whom 124 million people had low vision and 37 million were blind. However, refractive error as a cause of visual impairment was not included, which implies that the actual global magnitude of visual impairment is greater. Visual impairment is unequally distributed across age groups. More than 82% of all people who are blind are 50 years of age and older, although they represent only 19% of the world's population. The risk of severe eye problems has been found to increase significantly with age, particularly in those over age 65. Also, visual impairment is not distributed uniformly throughout the world, as more than 90% of the world's visually impaired live in developing countries (statistics and more info on visual impairment at [20]  [22] [23]

According to the Eye Diseases Prevalence Research Group, without improvements in preventing and treating eye diseases, the number of blind or visually impaired Americans 40 and older will grow from 3.4 to 5.5 million in 2020 [21] .

Considering the above facts, it becomes clear that society should shift towards people with impaired vision. Technological advancement in the field of robotics is able to provide devices and techniques to help people with physical or sensorial restrictions. The proposed application is an attempt to create a vision-assistance system for people who experience problems in vision.

## 7.2 The idea

Our aim is to help people with limited sight, for example not being able to read the labels on medication boxes. The basic idea is to make a mobile phone or another device recognise the correct medication box that a user should take according to his medication list and schedule. With the term "recognise" we imply a system that is able to "see" the medications labels and "notify" the user when the appropriate medication is reached.

In order to achieve the identification of the correct medication, we can apply Optical Character Recognition (OCR) on the photographs medication boxes. Having succeeded in the identification, the user could be easily notified by an audio alarm.

## 7.3 OCR on photographs of medication boxes

For the part of the artificial sight needed as described in the previous section, we can use Optical Character Recognition (OCR) to extract text and layout information from images. We developed an application that performs OCR on photographs of medication boxes, captured by a mobile phone's camera. An example of the application is shown in the following figure.

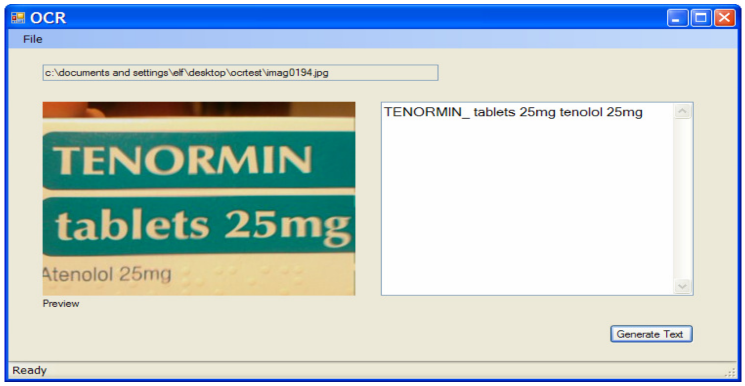


Fig. 13: Performing OCR on a photograph of a medication box

Optical Character Recognition is achieved with the help of Microsoft Office Document Imaging (MODI) object model. MODI is a Microsoft Office application that supports editing documents scanned by Microsoft Office Document Scanning. It was first introduced in Microsoft Office XP and is included in later Office versions including Office 2003 and Office 2007. One of the MODI features is an OCR engine that it allows users to produce editable text from scanned documents, interpreting pixel-based images to layout and text elements.

## 7.4 Optimizing the results

We tested our application with a number of photographs captured by a HTC P3650 smart phone. The results derived from our experiments were discouraging, so we did several efforts to optimize the performance of our application.

A well-known method that can be used to increase the image quality and thus, the OCR results, is performing image processing techniques to the photographs. We used Adobe Photoshop CS2 to apply image processing functions to photographs of medication boxes in order to achieve a better performance of the OCR. The image processing techniques that we performed are:

- Contrast adjustment
- Level adjustment
- Grayscale
- Sharpness filter
- Threshold adjustment

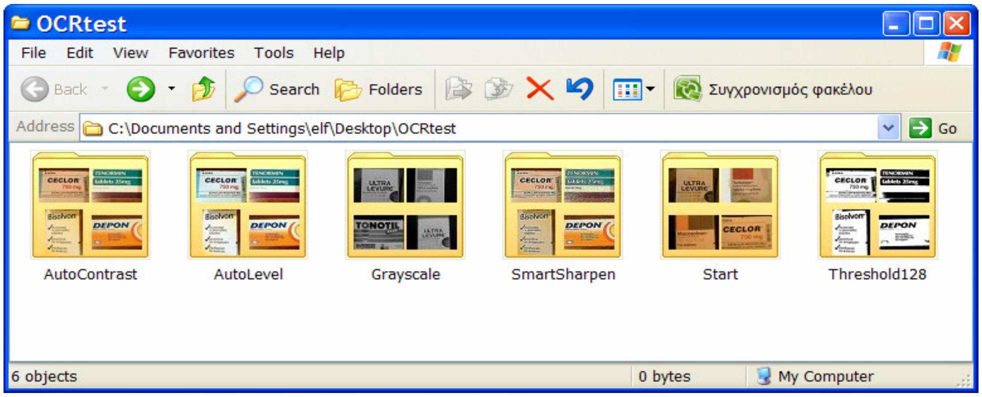


Fig. 14: Use of image processing techniques to optimize the OCR results

Apart from performing image processing techniques to the photographs of medication boxes, we decided to optimize the results by taking advantage of an a priori known medication list. For this reason we added to our project an XML file (Fig. 14) containing the names of the medications used in our experiments.

```xml
<?xml version="1.0" encoding="utf-8" ?>
<MEDICATION_LIST>
  <MEDICATION>
    <NAME>CECLOR</NAME>
  </MEDICATION>
  <MEDICATION>
    <NAME>TENORMIN</NAME>
  </MEDICATION>

       .
       .
       .

  <MEDICATION>
    <NAME>Bactrimel</NAME>
  </MEDICATION>
  <MEDICATION>
    <NAME>ULTRA LEVURE</NAME>
  </MEDICATION>
</MEDICATION_LIST>
```

Fig. 15: A medication list written in XML

With this modification, we perform OCR on the photographs of medication boxes and in the text produced we search for the exact names using the medication list. If none of the medication names is found, we calculate the similarity between the OCR text and each one of the medication names and return the best result. The similarity level between the two strings is calculated with an off-the-shelve algorithm based on the Levenshtein distance.

We observed that a major factor of low performance was that a photograph of a medication box contains much more text information except from the name of the medication (Fig. 15). To overcome this obstacle, we tried to isolate only the words with the maximum font size, which usually is the exact name of the medication.

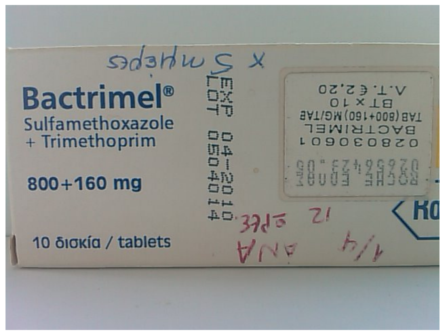


Fig. 16: Text information on a photograph of a medication box

## 7.5 Evaluation

In order to evaluate this application, we intentionally have chosen to capture photographs during different hours of the day, which lead to a variety of lighting conditions, as it would have happened with the use of this system in real time environment. Also intentionally, we have chosen medication boxes with similar market names (e.g. Tenormin - Tonotil), as this could be a possible situation in a person's medication list.

We captured with the mobile phone's camera 150 photographs of medication boxes. In the next step, we applied the image processing techniques mentioned in the previous section to each photograph. For the needs of the evaluation process, we modified the application to be able to scan the contents of a folder and write the output (OCR results or medication name matching) to a Microsoft Office Word document.

At the first stage of this experiment we performed OCR on the photographs and printed the output of the OCR engine to the Word document. We consider that the application identifies correctly the name of the medication if the output string contains the exact market name. At the second and third stage, we took advantage of the XML medication list, printing the best match as the result. The difference between the last two stages is that in the third stage we isolate the text with maximum font from the OCR output and according to that text find the best match from the medication list. The results of the conducted experiment are shown in the following table:

| | | Raw images | Auto contrast | Auto level | Gray scale | Smart Sharpen | Threshold 128 |
|---|---|---|---|---|---|---|---|
| 1 | No use of medication list | 45% | 51% | 49% | 46% | 27% | 51% |
| 2 | Use of medication list | 59% | 68% | 64% | 59% | 48% | 71% |
| 3 | Use of medication list & isolation of maximum fonts | 54% | 67% | 60% | 56% | 41% | 69% |

Table.5: Success rates for the three stages of the experiment

Regarding these results we made several observations. We can see that the success rate in identifying the medication name using the original photographs as captured with the mobile phone's camera (raw images) is very low at 45%. Even with the use of the a priori known medication list the success ratio never raises above 59%. This fact indicates that the use of image processing techniques is absolutely necessary to achieve better accuracy.

In addition, we observed that all the image processing techniques, with the Sharpness adjustment being the only exception, resulted in better performance. Especially the contrast and the threshold adjustment gave us the best success ratios, reaching at 71%.

Finally, the use of the medication list apparently helped in a better estimation of the medication names in the tested photographs. What could be described as a bad surprise is the fact that our final try to optimize the results with the isolation of the maximum fonts failed. Although achieving a much better performance than the experiment's first stage, the isolation of the maximum font text resulted in worse identification accuracy in comparison with the experiment's second stage.

The most reasonable explanation for the low performance is that although the commercially available OCR systems are capable of reading text information, they are merely designed to process typical "clean" documents, scanned under good illumination conditions. That is why they usually fail to read text messages embedded in natural scene images, where the text background is textured or heavily cluttered, text fonts may differ in size, orientation and alignment, lighting conditions are uneven.

To confirm this we examined carefully the OCR output for the photographs of the third stage of the experiment and realized that in some cases graphics or textures on the medication boxes were falsely interpreted as maximum font text. We also tried to fed the OCR engine with photographs of articles from newspapers captured by the mobile phone's camera and discovered that it certainly achieved a higher level of accuracy.

## 7.6 Future work

In this section we provide some suggestions and changes for the proposed application in order to improve its performance.

As occurred from the experiments conducted in the evaluation process, the development of a technique to exclude the graphics and textures and keep only the text from the photographs is very important. We found that there is related work conducted [57] [58] [59] [60] [61] [62] [63] [64] [65] [66] , focusing on a similar technique referred as "text localization in natural scenes". It is quite possible that following such an approach could result in a far better accuracy.

As it happened with the proposed indoor localization system, we observed that the quality of the mobile phone's photographs depends heavily on the varying lighting conditions, limiting the performance of the overall application. It would be interesting to test this application with next generation cameras that capture higher quality photographs.

Finally, although we already demonstrated how such an application could be an aid tool, we would like to see this system tested in real time environment from people who cope with vision problems in order to be able to better estimate its' usability.

# Chapter 8

# Conclusion

## 8.1 Conclusion

Mobile phones have long outgrown their intended mandate of voice connectivity. Today smart phones and Personal Digital Assistants (PDAs) function as composite communication and computing devices. Regarding their special characteristics and capabilities, we make great effort to develop new applications to improve people's daily lives to the most degree.

In this thesis we proposed two applications with the purpose to assist aging people to live a more independent life in the context of a Smart Home environment. The first application is an indoor localization system that uses images captured by a smart phone to track a user moving inside his home. It is based on the fingerprinting technique and achieves location determination using image matching. The second application aims to help people dealing with visual impairment. It performs Optical Character Recognition in photographs of medication boxes captured by a smart phone and tries to identify the market names of the medications.

The experiments we conducted to evaluate the proposed systems indicated that accuracy is a point that needs further improvement. However, the major contribution of this work is that it proved that the development of applications aiming to help people can be easily achieved by exploring the capabilities of the existing technology with no additional cost.

# References

[1] ARCHANGEL Project, Mobile Multimedia Laboratory - Athens University of Economics and Business, http://mm.aueb.gr/research/archangel/

[2] Tiresias.org, "What are smart homes and how to make them accessible for people with disabilities and the elderly", http://www.tiresias.org/research/guidelines/smart_home.htm

[3] Patrick R.W. Roe, "Towards an inclusive future: Impact and wider potential of information and communication technologies", http://www.tiresias.org/cost219ter/inclusive_future/index.htm

[4] Thomas R. Eng, "The eHealth Landscape: A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care", http://www.informatics-review.com/thoughts/rwjf.html

[5] Wikipedia - Home automation for the elderly and disabled, http://en.wikipedia.org/wiki/Home_automation_for_the_elderly_and_disabled

[6] ScienceDaily (Jan. 24, 2007), Smart Way Of Living For People With Dementia, http://www.sciencedaily.com/releases/2007/01/070124104419.htm

[7] Bath Institute of Medical Engineering, BIME Projects, http://www.bath.ac.uk/bime/projects/dc_projects.htm

[8] Duke Smart Home Program, http://www.smarthome.duke.edu/home/index.php

[9] University of Florida - Gator Tech Smart House, http://www.icta.ufl.edu/gt.htm

[10] Amigo - Ambient Intelligence for the Networked Home Environment, http://www.hitech-projects.com/euprojects/amigo/amigo.htm

[11] Tampere University of Technology, Department of Electronics, Smart Home Project, http://www.ele.tut.fi/research/personalelectronics/projects/smart_home.htm

[12] George Demiris, Marjorie Skubic, Marilyn Rantz, James Keller, Myra Aud, Brian Hensel, Zhihai He, "Smart Home Sensors for the Elderly: A Model for Participatory Formative Evaluation "

[13] Gilles Virone, Norbert Noury, Jean-Pierre Thomesse, Vincent Rialle, Jacques Demongeot, "A home health information system based on the can fieldbus", 5th IFAC Conference On Fielbus Systems and Their Applications, Aveiro, Portugal, July 7-8, 2003

[14] MedPac, "Report to the Congress: New Approaches in Medicare", Chapter 7 - Information Technology in Healthcare, June 2004

[15] George Demiris, Marilyn J. Rantz, Myra A. Aud, Karen D. Marek, Harry W. Tyrer, Marjorie Skubic, Ali A. Hussam, "Older adults' attitudes towards and perceptions of 'smart home' technologies: a pilot study", MED. INFORM., Vol. 29, no. 2, p. 87-94, June 2004

[16] Sumi Helal, William Mann, Hicham El-Zabadani, Jeffrey King, Youssef Kaddoura, Erwin Jansen, "The Gator Tech Smart House: A Programmable Pervasive Space", IEEE Computer Society, March 2005

[17] AForge.NET Framework, http://code.google.com/p/aforge/

[18] CodePlex Open Source Community, Similar images finder - Image processing in C# and RGB projections, http://similarimagesfinder.codeplex.com/

[19] CodePlex Open Source Community, Image Recognition , http://www.codeplex.com/IR/

[20] Tiresias.org, Visual impairments, http://www.tiresias.org/accessible_ict/visual.htm

[21] Allaboutvision.com, "Statistics on Eye Problems, Injuries and Diseases", http://www.allaboutvision.com/resources/statistics-eye-diseases.htm

[22] Wikipedia - Visual impairment, http://en.wikipedia.org/wiki/Visual_impairment

[23] American Foundation for the Blind (AFB), "Special Report on Aging and Vision Loss", September 2008, http://www.afb.org/Section.asp?SectionID=15&DocumentID=4423

[24] Wikipedia - Optical Character Recognition, http://en.wikipedia.org/wiki/Optical_character_recognition

[25] Wikipedia - Computer Vision, http://en.wikipedia.org/wiki/Computer_vision

[26] Wikipedia - Content-based Image Retrieval, http://en.wikipedia.org/wiki/Content-based_image_retrieval

[27] Markus Jevring, “Automatic Management of Bluetooth Networks for Indoor Localization”, Faculty of Electrical Engineering, University of Twente, August 2008

[28] Mortaza S. Bargh, Robert de Groote, “Indoor Localization Based on Response Rate of Bluetooth Inquiries”, Telematica Instituut, MELT’08, ACM, September 2008

[29] Sudarshan S. Chawathe, “Beacon Placement for Indoor Localization using Bluetooth”, Department of Computer Science, University of Maine

[30] Albert Huang, Larry Rudolph,” A Privacy Conscious Bluetooth Infrastructure for Location Aware Computing”, Massachusetts Institute of Technology

[31] A. K. M. Mahtab Hossain, Wee-Seng Soh, “A comprehensive study of Bluetooth signal parameters for localization”, Department of Electrical & Computer Engineering National University of Singapore, 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC’07)

[32] Oliver Keiser, Philipp Sommer, “BlueLocation II - A Localization Infrastructure for Bluetooth Enabled Mobile Devices”, Swiss Federal Institute of technology, January 2006

[33] Frank Reichenbach, Dirk Timmermann, “Indoor Localization with Low Complexity in Wireless Sensor Networks”, Institute of Applied Microelectronics and Computer Engineering, Rostock - Warnemuende, Germany

[34] Suguna P. Subramanian, Jurgen Sommer, Stephen Schmitt, Wolfgang Rosenstiel, “SBIL: Scalable Indoor Localization and Navigation Service”, University of Tubingen, Department of Computer Engineerin

[35] Julio Oliveira Filho, Ana Bunoza, Jurgen Sommer, Wolfgang Rosenstiel, “Self-Localization in a Low Cost Bluetooth Environment”, Wilhelm - Schickard - Institut fur Informatik, Department of Computer Engineering, University of Tubingen, Germany, LNCS 5061, pp. 258–270, Berlin 2008

[36] Rimon Elias, Amal Elnahas, “An Accurate Indoor Localization Technique Using Image Matching”, Computer Science and Engineering Department, The German University in Cairo, Cairo, Egypt

[37] Barry Brumitt, Brian Meyers, John Krumm, Amanda Kern, and Steven Shafer, “EasyLiving: Technologies for Intelligent Environments”, Microsoft Research, Redmond, USA, HUC 2000, LNCS 1927, pp. 12-29, 2000

[38] Nishkam Ravi, Pravin Shankar, Andrew Frankel, Ahmed Elgammal and Liviu Iftode, “Indoor Localization Using Camera Phones”, In The Seventh IEEE Workshop on Mobile Computing Systems and Applications, pages 1–7, 2006

[39] Holger Linde, “On Aspects of Indoor Localization”, Faculty of Electronics and Informatics, University of Dortmund, August 2006

[40] Andrew M. Ladd, Kostas E. Bekris, Algis P. Rudys, Dan S. Wallach, Lydia E. Kavraki , “On the Feasibility of Using Wireless Ethernet for Indoor Localization”, IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, VOL. 20, NO. 3, JUNE 2004

[41] Kamol Kaemarungsi, Prashant Krishnamurthy, “Modeling of Indoor Positioning Systems Based on Location Fingerprinting”, School of Information Science, University of Pittsburgh, IEEE INFOCOM July 2004

[42] Dik Lun Lee and Qiuxia Chen, “A Model-Based WiFi Localization Method”, The Hong Kong University of Science and Technology, Department of Computer Science and Engineering, Hong Kong, INFOSCALE 2007 June 6-8, 2007, Suzhou, China, ACM 2007

[43] Chin-Heng Lim, Yahong Wan, Boon-Poh Ng, Chong-Meng Samson See, “A Real-Time Indoor WiFi Localization System Utilizing Smart Antennas”, IEEE Transactions on Consumer Electronics, Vol. 53, No. 2, MAY 2007

[44] Anindya S. Paul and Eric A. Wan, “Wi-Fi based indoor localization and tracking using sigma-point Kalman filtering methods”, OGI school of Science and Engineering, Oregon Health and Science University, IEEE 2008

[45] Ahmed Ali Sabbour, “WiGuide: Indoor System for LBS”, Faculty of Media Engineering and Technology, German University in Cairo, July 2007

[46] Aldrich, F.K., “Smart Homes: Past Present and Future”, p. 17- 39, London: Springer-Verlag UK, 2003

[47] Van Berlo, “Smart home technology: have older people paved the way?” Gerontechnology, p 77-87, 2002

[48] Van Berlo et al., “Design Guidelines on Smart Homes. A COST guidebook.”, 1999

[49] Van Berlo, Felibaum K., “Smart Home Technology: Useful Applications for older people”. 1999 - In: Bühler, C.and H. Knops: Assistive Technology on the Threshold of the New Millennium. Assistive Technology Research Series Vol. 6, p.508-51 2. Amsterdam

[50] Gerhard Elger, Barbro Furugren, "SmartBo - An ICT and computer-based demonstration home for disabled people”, TIDE July 1998

[51] Eurosafe PROSAFE project, http://www.childsafetyeurope.org/csi/eurosafe2006.nsf/wwwVwContent/l3prosafeproject.htm

[52] Hospital Without Walls, http://www.hospitalwithoutwalls.com/

[53] Aware Home, Georgia Institute of Technology, http://awarehome.imtc.gatech.edu/

[54] CASAS Smart Home project, http://ailab.wsu.edu/casas/

[55] J. Krumm, "Multi-camera Multi-person Tracking for Easy Living", In The third IEEE International Workshop on Visual Surveillance, pages 3–10, New Jersy, 2000

[56] H. Bay, "Interactive Museum Guide", In Proceedings of UBICOMP, Tokyo, Japan, September 2005

[57] V. Gaudissart, S. Ferreira, C. Thillou and B. Gosselin, "SYPOLE: a mobile assistant for the blind," Proc. 13th European Signal Processing Conference, EUSIPCO 2005, Antalya, Turkey, 2005, pp. 1–4.

[58] S. Ferreira, C. Thilou and B. Gosselin, "From picture to speech: an innovative application for embedded environment," Proc. of the 14th ProRISC Workshop on Circuits, Systems and Signal Processing, ProRISC'03, 2003.

[59] N. Ezaki, M. Bulacu and L. Schomaker, "Text detection from natural scene images: towards a system for visually impaired persons," Proc. 17th IEEE International Conference on Pattern Recognition, ICPR'04, 2004, pp. 683–686.

[60] N. Ezaki, K. Kiyota, B.T. Minh, M. Bulacu and L. Schomaker, "Improved text-detection methods for a camera-based text reading system for blind persons," Proc. 8th International Conference on Document Analysis and Recognition, ICDAR'05, 2005, pp. 257–261.

[61] E.D. Haritaoglu and I. Haritaoglu, "Real time image enhancement and segmentation for sign/text detection," Proc. IEEE International Conference on Image Processing, ICIP'03, 2003, pp. II993–II996.

[62] J. Zhang, X. Chen, A. Hanneman, J. Yang and A.Waibel, "A robust approach for recognition of text embedded in natural scenes,", Proc. IEEE International Conference on Pattern Recognition, ICPR'02, 2002, p. 30204.

[63] J. Gao and J. Yang, "An adaptive algorithm for text detection from natural scenes," Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'01, 2001, pp. II84–II89.

[64] X. Chen and A.L. Yuille, "Detecting and reading text in natural scenes," Proc. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'04, 2004, pp. II366–II387.

[65] X. Chen and A.L. Yuille, “A time-efficient cascade for real-time object recognition with applications for the visually impaired,” Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR’05, 2005, p. 28.

[66] Marcin Pazio, Maciej Niedzwiecki, Ryszard Kowalik, Jacek Lebiedz, “Text Detection System for the Blind”, EUSIPCO, Poznan 2007