

# A review of indoor localization technologies: towards navigational assistance for topographical disorientation

Jorge Torres-Solis, Tiago H. Falk and Tom Chau  
*Bloorview Research Institute & University of Toronto  
Canada*

## 1. Abstract

Indoor localization technologies hold promise for many ambient intelligence applications, including in-situ navigational assistance for individuals with wayfinding difficulties. Given that the literature on indoor localization is vast and spans many different disciplines, we conducted a comprehensive review of the dominant technologies. We propose a taxonomy of localization technologies on the basis of the measured physical quantity. In particular, we identified, radio frequency, photonic, sonic and inertial localization technologies as leading solutions in the field. For each selected technology, the fundamental scientific mechanisms for localization are explained, key recent literature appraised and the merits and limitations are discussed. Recommendations are made regarding the creation of context-aware systems that can be used to enhance a user's topographical orientation skills.

## 2. Acronyms

**AoA** angle of arrival  
**BLUPS** Bluetooth and Ultrasound Positioning System  
**DS/CDM** Direct Sequence Code Division Multiplexing  
**DToA** difference of time of arrival  
**GPS** Global Positioning System  
**HMD** Head Mounted Display  
**IR** Infrared  
**IrDA** Infrared Data Association  
**LOS** Line-of-sight  
**NLOS** Non-Line-of-sight  
**PDA** Personal Digital Assistant  
**RDS** root-mean-square delay spread  
**RF** Radio Frequency  
**RFID** Radio-frequency Identification  
**RSS** Received Signal Strength

**ToA** time of arrival  
**ToF** time of flight  
**TD** Topographical Disorientation  
**TV** television  
**UWB** Ultra-Wideband  
**WLAN** Wireless Local Area Network

### 3. Introduction

Topographical Disorientation (TD) refers to a family of deficits in environmental orientation and navigation. Aguirre and D'Esposito [3] provide a well-accepted taxonomy of TD, arguing that difficulties in wayfinding may arise as a result of the combination of different cognitive impairments. For example, it is well recognized that TD and spatial navigation deficits are common sequelae of brain injury [87, 74]. Individuals living with post-traumatic effects of brain injury are oftentimes faced with symptoms such as weak visual scanning skills, or deficits in complex attention, prospective memory or sequential processing [47]. These symptoms cause problems of interaction with and perception of the surrounding environment even several years post-injury [16, 72]. It has also been argued that deficits in topographical orientation can lead to spatial anxiety or wandering behaviours [11, 77, 87].

It has been suggested that wearable navigation technologies such as Global Positioning System (GPS) can be a useful wayfinding tool for individuals with cognitive impairments [7]. However, GPS signals have limited coverage indoors (e.g., [36, 40, 30, 61, 99, 38]). Given that patients spend significant periods of time indoors - be it in acute and tertiary care hospitals or, subsequent to rehabilitation, at home, schools, office buildings, shopping malls, long-term care facilities - identification of potential technologies for indoor navigational assistance is imperative. An initial survey of the literature has suggested that a diverse collection of candidate indoor localization technologies exists across many different disciplines. This diversity makes it difficult to grasp the potential of an existing technology for the rehabilitation of individuals with topographical disorientation.

Localization technologies are critical to emerging location-aware guidance systems and support services for individuals who have wayfinding difficulties due for example to low vision [69], stroke [86] and traumatic brain injury [6]. In particular, regarding indoor navigation systems for individuals with topographical disorientation, localization has often been human-mediated rather than automatic. For instance, Liu et al. evaluated the benefits of navigational tools in real indoor environments [46]. However, the location tracking and tool display decisions in their experiments were not automatic, but controlled by the experimenters. In similar vein, Sohlberg et al. [75] found that individuals with wayfinding difficulties secondary to brain injury responded well to speech-based auditory directions from a wrist-worn PDA navigation system. However, like Liu et al., the PDA's navigational instructions were transmitted by a human operator at a mobile computer. Undoubtedly, there is immense opportunity to explore the potential of automatic patient indoor localization technologies in the emerging fields of cognitive prosthetics and situated assistive technologies. As a consequence, the overarching goal of this review is to systematically organize the literature on indoor human tracking technologies, and to ascertain their feasibility for eventual use in the realm of TD rehabilitation.

## 4. Literature selection

We combed the literature for candidate localization technologies that could serve to create an assistive device for individuals with TD in indoor environments. In particular, peer-reviewed journal articles published in English between 2003 and 2009, inclusive, were sought from three different academic databases, namely: Compendex, Inspec and Geobase, using the keywords, "*indoor location*", "*indoor localization*", "*indoor tracking*" and "*indoor positioning*". After removing duplicate records, we arrived at 214 articles. To identify potential technologies applicable to the creation of a navigational assistance device that for individuals with topographical disorientation that offered accurate information in real time, the returned articles were subsequently screened according to the following inclusion criteria:

1. The article must focus on the development and experimental testing of a localization or navigation system: i.e., articles focusing on mathematical processing of localization data, or localization experiments in simulated environments were discarded.
2. The reported technology must:
  - (a) be usable indoors, within a building or a larger space, i.e., technologies used to track a capsule inside the human body, or a device within a single room were excluded;
  - (b) offer a localization accuracy of a mobile target within a 10 meter radius with a delay of 5 seconds or less;
  - (c) be applicable to humans, i.e., systems designed for vehicles, large objects, or objects that relied on a fixed pose or odometry measurements of a robot were excluded; and
  - (d) track and identify multiple humans concurrently.

Fifty three articles met such initial criteria. Such articles were subsequently scanned for alternate localization technologies that were referenced three or more times and that were not selected in the initial search. Eleven additional articles were included in this manner, totaling sixty four articles for consideration in the present review.

### 4.1 Taxonomy of localization technologies

The location of an object in space is determined by measuring a physical quantity that changes proportionally with the position of the object of interest. The present review is structured in terms of the measured physical quantity. The selected articles were divided into six main categories based on the physical quantity measured, namely, (1) radio frequency waves, (2) photonic energy, (3) sonic waves, (4) mechanical energy (inertial or contact), (5) magnetic fields, and (6) atmospheric pressure. Each physical quantity grouping can be further subdivided according to the underlying hardware technology. Figure 1 summarizes this two-tiered taxonomy. Note that the latter two phenomena (magnetic and atmospheric) have been collapsed into one category, named, "Other" due to low article counts in these areas.

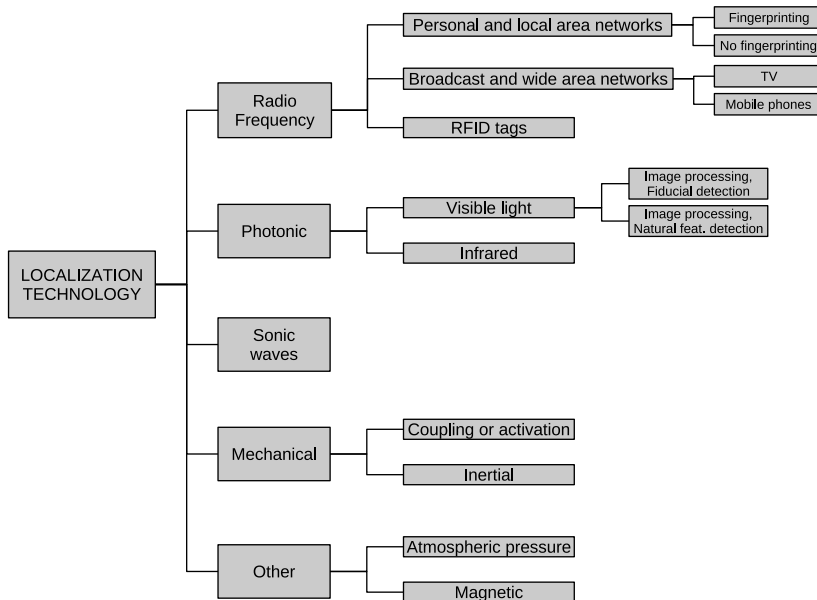


Fig. 1. Taxonomy of indoor localization technologies by measured physical quantity and hardware technology

Where appropriate, articles are further differentiated by principal localization technique. For this third level of classification, the following localization technique definitions are provided, expanding on those proposed by Hightower and Borriello [29]:

**Triangulation** is a family of methods that include lateration, angulation and variations thereof. Lateration refers to the calculation of the position of the human subject based on his relative distance to several previously-known fixed points in space. Such distances are commonly obtained indirectly by measuring parameters that are proportional to distance. Time of flight and power attenuation of a radio signal are common indirect distance metrics [20]. Angulation refers to the calculation of the position of the subject using the angles of arrival of signals emitted from fixed points in space. [88, 80].

**Proximity** refers to a class of methods which establish the presence of the human subject in the vicinity of a sensor, which alone has limited sensing range and analysis capabilities. The proximity of the subject can be detected through physical contact, presentation of a device such as a magnetic band to an appropriate reader or through the monitoring of a physical quantity in the vicinity of the sensor, for instance, a magnetic field.

**Scene analysis** involves the monitoring of a wide area around the subject of interest from a specific vantage point. The commonly deployed sensors have broad coverage area and range. Examples include ceiling-mounted video cameras or passive infrared (PIR) sensors.

**Dead reckoning** refers to the usage of sensors that provide location updates, calculated using information about a previously-estimated location. Position estimation is commonly based on accelerometry and gyroscopy.

The ensuing review of literature will adhere closely to the taxonomy depicted in Figure 1. For each physical phenomenon, we will briefly present the general principles of localization, review articles in the relevant subcategories and comment on their relative merits and limitations. We will conclude the article with recommendations of indoor localization technologies suitable for addressing the development of assistive devices for individuals with TD.

## 5. Radio Frequency

An electromagnetic wave is the energy generated by an oscillating, electrically charged particle in space. The generation of electromagnetic waves is known as a radio frequency emission.

Solutions in this category estimate the location of a mobile target in the environment by measuring one or more properties of an electromagnetic wave radiated by a transmitter and received by a mobile station. These properties typically depend on the distance traveled by the signal and the characteristics of the surrounding environment.

As depicted in Figure 2, most of the articles in this survey describe a Radio Frequency (RF) localization system. These articles can be further subcategorized according to the underlying hardware technology as listed below.

1. Personal and local area networks, including technologies such as IEEE 802.11, Ultra-Wideband (UWB), ZigBee, or Bluetooth, either as the sole localization technology [12, 92, 17, 43, 20, 64, 66, 42, 41, 53, 79, 28, 22, 96, 18, 67, 4, 95, 88, 23, 83, 9, 57] or as a contributing technology within a hybrid solution [98, 13, 100, 49, 60, 2].
2. Broadcast and wide area networks, including networks designed for localization purposes, such as the GPS and the Global Navigation Satellite System (GNSS) [67], and broadcast networks not originally intended for localization purposes, such as cellular phone networks [73, 25] and television broadcast signals [62, 63].
3. Radio-frequency Identification (RFID) tags [45, 35, 81]
4. Radar [65, 71]

Each flavour of RF localization is reviewed below.

### 5.1 Personal and local area network-based solutions

#### 5.1.1 Fingerprinting-based localization solutions

Most articles included in this section propose the localization of mobile targets using a two stage process [98, 17, 64, 66, 42, 53, 9, 57, 92, 43, 79, 22, 96, 18, 67, 4, 95, 83]. The first stage consists of an off-line radio scene analysis (i.e., performed prior to localization) in which a

mobile station extracts radio fingerprints, i.e., features from one or more metrics of the radio signal measured at predefined points in the environment. These radio fingerprints are proportional to the distance between the mobile receiver and the emitting station. Common

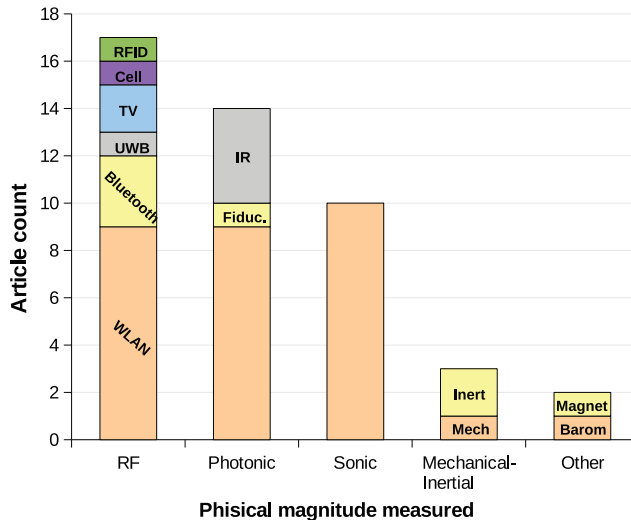


Fig. 2. Distribution of articles by physical quantity measured

metrics include the direction or angle of arrival (AoA), Received Signal Strength (RSS), or time of flight (ToF) of the incoming radio signal [52]. A radio map or database of fingerprints is created, storing signal feature values at each location along with the corresponding spatial coordinates. Some authors propose automated or assisted radio map creation techniques, exploiting the characteristics of the environment, such as the spatial configuration and the material composition of the environment [79, 34, 98].

The on-line stage comprises the active localization process where the mobile receiver extracts a fingerprint of the radio signal at an unknown location. Localization is commonly achieved by proximity techniques, i.e., finding the closest match between the features of the received radio signal and those stored in the radio map [98, 17, 64, 66, 42, 22, 18, 67, 4, 57, 92]. More accurate localization can be achieved using a triangulation-like process, in which several candidate locations (each with a fingerprint bearing some resemblance to that of the received signal) are geometrically combined to provide an estimate of the receiver location in space [43, 53, 79, 67, 95, 83, 9]. Algorithms deployed in the selection of the closest match or matches from the radio map include: 1) nearest neighbours techniques and variations thereof [17, 95, 9]; 2) Bayesian statistical matching [98, 92, 43, 64, 66, 67, 83, 57]; 3) maximum likelihood estimation [79, 22]; 4) correlation discriminant kernel selection [42, 53]; and neural networks [18, 4].

Some fingerprinting techniques also provide coarse estimates of orientation, for example, 4 different orientations [92, 9]. The radio map is created with a user transporting the mobile device. At a given location, a fingerprint is recorded at each possible orientation. Since the

human body affects the propagation of radio signals, the fingerprint generated for each orientation will be different [92, 9].

Fingerprinting localization accuracy is commonly down to a few meters (i.e. within 3 metres 90% of the time [42] and within 3 metres 91.6% of the time [53]). The lowest number of base stations used to create the feature space of a fingerprint was one [98]. In other words, fingerprinting seems to provide reasonable localization accuracies without excessive hardware requirements. The most pressing challenge however is the non-stationarity of the radio map. This is reflected as differences in the measured signals during the on-line and off-line phases at the same exact location. The time-variant nature of the radio map can be attributed to radio signal propagation effects induced by dynamic aspects of the environment such as the presence or absence of people, elevators, moving doors and other environmental changes [45, 42, 9].

As a particular case, in [67], the authors proposed an indoor and outdoor hybrid localization system, combining GPS and WLAN localization technologies. The authors proposed handing off the localization responsibility between GPS and WLAN depending on their availability. Fingerprinting was done through a GPS-on-line stage, collecting positional information of the access points within a nearby building and geo-referencing these measurements with the information obtained through GPS. After this fingerprinting process, indoor WLAN positioning was achieved by estimating a point located among a set of the most probable locations. These locations were predefined by the user - i.e. copy room, cafeteria. The histogram of the RSS measurements was used to determine the location that best matched the histogram received. If the histogram did not closely match any known locations, a centroid algorithm estimated the location of the user from the locations of previously geo-referenced access points and probable nearby locations. The place detection algorithm, which relied on an almost perfect fingerprint match, yielded room level localization accuracy, while the WLAN localization algorithm, using the centroid of several probable locations, yielded an accuracy of approximately 30 metres.

### 5.1.2 Non-fingerprinting-based solutions

RF-based localization can also be achieved without *a priori* analysis of the radio properties of the environment (i.e., without development of a radio map). Four of these articles, all of them based on UWB radio signals, rely on signal triangulation as the sole localization technique [28, 88, 23, 41], while in [20] localization is achieved by proximity and scene analysis.

Indoor localization based on triangulation of radio waves is a non-trivial problem because the transmitted signal can suffer obstructions and reflections. As a consequence, Non-Line-of-sight (NLOS) conditions emerge. In the presence of NLOS conditions, the radio signal can travel to the receiver through a non-direct path, giving rise to erroneous distance estimates.

To overcome these problems, the use of UWB radio signals has become the most novel solution in radio frequency-based solutions. The properties of ultra-wide band, short duration pulses mitigate the propagation problems associated with multi-path radio propagation. The most representative example is the system proposed by Venkatesh and Buehrer. They introduced a triangulation localization system based on impulse UWB radio signals [88]. They suggested that the statistical parameters describing the distribution of the received root-mean-square delay spread (RDS) serve as the best discriminant estimator

between Line-of-sight (LOS) and NLOS signal propagation. This means that the statistical parameters defining the RDS of the received signal can be compared against a predefined rule set to determine if the signal was received via a direct or indirect path. Subsequently Venkatesh and Buehrer tracked a mobile station through 71 predefined locations within a building, achieving localization accuracies ranging from 1 centimetre to 2 metres. As another example of RF triangulation-based schemes, Krejcar and Cernohorsky presented a localization system [41] that relied on the triangulation of RSS metrics. Room-granularity accuracies were reported but further details of the triangulation or localization schemes were not revealed.

Finally, a system based on a combination of scene analysis and proximity techniques using a Bluetooth ad-hoc network was presented in [20]. Bluetooth inquiry signals were used for localization. In inquiry mode, a bluetooth device inquires about neighbouring bluetooth stations. This inquiry process consists of scanning for devices in the vicinity, using a sequence of different power levels. Low power levels will detect devices in close proximity while high power levels will include devices that are located farther away, providing coarse distance estimates in this fashion. This approach requires a fixed or "anchor" node which establishes the position of nearby mobile nodes. Subsequently, the localized nodes can establish the position of other undetected mobile nodes in their vicinity, creating an ad-hoc localization network. The reported localization error was 1.88 meters.

## 5.2 Broadcast and wide area network solutions

### 5.2.1 Solutions based on television signals and cellular networks

The solutions in this section are based on RF infrastructure transmitters that cover a wide area, and while not originally designed for localization, can be adapted to provide indoor localization services. In particular, two such technologies were identified: television (TV) broadcast signals [62, 63] and cellular phone networks [73].

Rabinowitz and Spilker [62, 63] proposed the use of synchronization signals already present in the Advanced Television Signal Committee (ATSC) standard for compliant digital TV signals. As the signal properties and geometrical arrangement of the TV broadcast network have been designed to penetrate indoors, they offer significantly greater indoor coverage than GPS-based solutions. Implementation of a localization solution would require no modification of the existing broadcast signal. To overcome the inherent lack of synchrony between stations due to clock imperfections, Rabinowitz and Spilker deployed a fixed reference station that transmitted an offset correction signal. A mobile station then calculated the ToF of the signal, and subsequently the distance to each TV broadcast station. As the positions of the broadcast stations were fixed and known, the position of a mobile receiver could be estimated. At least three visible transmitting stations were required for triangulation purposes. Rabinowitz and Spilker presented experiments in indoor environments in which they obtained a mean localization error ranging from 10 to 23 meters depending on the environment.

Hu et al. developed a method for the localization of mobile phones inside a cell [73] using fingerprinting techniques. Invoking a method for the automated creation of a radio fingerprint of the cellular signal, Hu et al. argued that granular localization can be achieved in indoor environments by statistically matching the fingerprint of the received signal with a record in the radio map. The authors emphasized that the localization accuracy in this case



is highly dependent on the size of the cell and the characteristics of the environment. This localization solution can be considered a combination of scene analysis (the off-line phase) and proximity techniques. Unfortunately, the improvement achieved in localization accuracy over conventional cell-ID localization was not reported.

### 5.3 Solutions based on RFID tags

An RFID system is commonly composed of one or more reading devices that can wirelessly obtain the ID of tags present in the environment. The reader transmits a RF signal. The tags present in the environment reflect the signal, modulating it by adding a unique identification code [45, 59]. The tags can be active, i.e., powered by a battery, or passive, drawing energy from the incoming radio signal. The detection range of passive tags is therefore more limited.

Lionel et al. [45] proposed a localization system named LANDMARC, using active tags. Reference tags were located in known, fixed positions in the environment. The reader was also situated in a fixed position. To locate a mobile tag, the reader scanned through 8 different power levels for tags in the vicinity. When a mobile tag was detected, the receiver compared the power returned by the reference tags and the mobile tag, determining the closest reference tags using a nearest neighbour algorithm. The position of the mobile tag was determined by triangulating the position of the nearest reference tags. The authors reported a localization accuracy of 2 meters, 75% of the time. The maximum localization delay was of 7.5 seconds.

Jia et al. [35] proposed a hybrid radio and vision based system which used RFID tags and a stereo camera for robot navigation purposes. To estimate the location of the mobile unit, RFID tags were used to mark walls and obstacles within the environment. The RFID detector comprised of a directional antenna, which yielded the general direction of the RFID tags detected. When a tag was in proximity of the robot, it obtained images from the stereo camera to estimate the distance to the obstacle marked by the RFID tag. Subsequently, the ID of the tag was compared against entries in a tag location database to determine if it belonged to a fixed landmark (i.e. RFID tag fixed to the wall) or an obstacle prone to change its position in the environment (i.e. a chair or a human). If the tag belonged to a fixed object or location, the information from the camera, combined with the directional orientation obtained from the RFID tag, were used to estimate the distance and orientation of the robot with respect to the tag. The localization accuracy was of 8.5 centimetres. Although this localization system was designed for a robot, the constraints on the robot's movement and posture were minimal. Therefore, this system might be adapted to human localization, at the risk of accuracy reduction.

Finally, in [81], Tesoreiro et al. introduced a localization system based on proximity to RFID tags. In this system a museum visitor used a personal digital assistant (PDA), which served as an automated museum guide. To estimate the position of the user, the PDA obtained the ID of RFID tags in the vicinity. Each tag was associated with an exhibit in the museum. The ID of the detected tag was subsequently transmitted to a server, which returned information about the exhibit in proximity to the user. This information was displayed in the PDA screen. The localization accuracy of this system was not reported. However, the accuracy was related to the density of tags in the environment.

#### 5.4 Solutions based on a radio frequency radar

Roehr et al. [71] presented an extension to the conventional frequency-modulated continuous-wave radar. The particular characteristic of this system was that it used two-way radio communication. Both fixed and mobile units were capable of transmitting and receiving a frequency modulated signal with a 5.8 GHz carrier. The fixed and mobile clocks were synchronized before distance and velocity estimations could be calculated. Once the units were synchronized, the fixed unit emitted a signal, which arrived at the mobile station. Subsequently, the mobile station sent a reply to the fixed station, which was synchronized using the signal just received from the fixed station. The round trip time of the signal was used to calculate the distance between the fixed and the mobile stations, while the frequency deviation was used to estimate the velocity of the mobile unit. The experimental setup included one experiment within an office building, where distances ranging from 5 to 25 metres were measured with the radar system. These distance measurements were then compared against the measurements obtained with a laser range finder. This experiment was designed to test line-of-sight conditions only and yielded measurements with a deviation of less than 3 centimetres when compared to laser range finder measurements.

In [65], the authors presented a radar indoor human tracking system, which exploited the Doppler effect of moving objects and micro-Doppler signal features that are particular to human movements. Various movements of the human body were classified based on the Doppler features received by the radar system. Such features were obtained from a joint time-frequency analysis, using the short-time Fourier transform and the reassigned joint timefrequency transform. The authors also proposed a scheme for tracking several human beings concurrently; the Doppler separation effect between moving humans was exploited, while target differentiation was realized using an antenna array that formed directed beams. Three antennas and two frequencies were required for multiple target tracking. Finally, the authors presented experiments to determine the range of the target, using frequency diversity, and studied the effects and errors introduced by the attenuation due to walls in the environment. The authors concluded that the properties of human movements can be exploited for localization purposes and that human localization using radar technologies is challenging within an indoor environment. They suggested a few techniques that might improve indoor localization, such as the placement of the radar as far away from the walls as possible.

#### 5.5 Limitations

The propagation of RF signals in indoor environments poses a central challenge. Certain materials within the indoor environment affect the propagation of radio waves. For example materials such as wood or concrete attenuate RF signals, while materials such as metals or water cause reflections, scattering and diffraction of radio waves. These effects lead to multi-path radio wave propagation, which encumbers accurate calculation of the distance between the transmitter and the receiver [88, 91, 73, 42, 99, 79, 71, 52]. Several authors have proposed techniques to compensate for these inaccuracies by automatically generating radio maps which consider the structure of the building [79, 73, 34]. However, a comprehensive model of all the materials in a complex environment such as a health care facility or a patient's home is a non-trivial problem.

The propagation of radio waves are adversely affected by changes to the physical environment such as the rearrangement of furniture, structural modifications or movement

of personnel within a building. Clinical settings are under constant change; elevators, personnel and large metallic structures such as beds and wheelchairs are constantly moving through the building. In these environments, the radio properties are highly dynamic, and a radio map captured at a certain point in time cannot be used reliably for localization without accounting for these dynamic changes [45, 43, 92, 9].

Interference and noise are often-mentioned challenges [42, 84]. Although some solutions operate within a reserved radio band [62, 73, 63], most of the research is conducted on open spectrum bands. This means that these solutions must account for the increased risk of interference due to other systems sharing the same frequency bands of the radio spectrum [43]. Finally, the usage of radio transmitting devices is often times restricted in critical areas of most healthcare facilities, according to recommendations made by the Association for the Advancement of Medical Instrumentation (AAMI) [1] and other standards or regulatory bodies. These restrictions limit the deployment of localization systems based on non-broadcast radio waves to specific, non-patient care areas, e.g, waiting rooms.

Localization technologies based on RF technologies can be attractive due to the ubiquity of certain infrastructural technologies, such as wireless data networks, that may already be present in the facilities. Care must be taken, however, in evaluating the impact of the physical environment on the RF localization technologies, as the solution may be rendered non-operative in certain clinical settings.

## 6. Photonic energy

Light refers to the phenomena of electromagnetic radiation at wavelengths within the visible range, which extends approximately between 380 and 750 nanometres. Photonic energy refers to the energy carried by electromagnetic radiation in this wavelength range, known as visible light, or in its lower or upper vicinity, known as ultraviolet and infrared light, respectively. A photon is the minimum possible discrete amount of light energy.

Solutions in this category rely on the photonic energy received from infrared or visible light emissions or reflections, to estimate the position of an object in space.

The articles selected for this section can be distinguished based on the sensor required to estimate the location of the tracked object. Some articles proposed the use of cameras to locate a mobile subject or device via image processing [58, 5, 94, 24, 21, 93, 97, 89, 39, 70, 31, 48, 10]. In contrast, some other articles present localization solutions based on non-image processing devices [15, 56, 55, 14, 33, 90].

Methods based on image analysis are all of the scene analysis persuasion. The classical model of computer vision-based location detection consists of 4 main stages [5]

1. Image acquisition, normally through a video camera
2. Segmentation of the image and extraction of relevant features
3. Selection of the closest match or matches of the detected features against the entries of a database of features (e.g. edges or fiducials). This process typically involves mathematical transformations of the spatial relationships between the features detected to account for the variability in scale, rotation or luminance of a given scene.

4. Computation of the pose of the camera. This consists of estimating the position and orientation of the camera that could have given rise to the observed image, assumed to be a distorted and rotated version of a database image. The selection of the closest match is achieved as in the previous step.

We can subsequently distinguish between localization systems that rely on image processing that exploits natural environmental features and systems that rely on the detection of a predefined synthetic pattern (i.e. fiducial) in the environment.

Consequently, the articles included in this section were organized in three main groups: 1) image processing and natural feature extraction, 2) image processing and fiducial recognition, and 3) non-image processing sensors. This organization is presented in the following sections.

### **6.1 Image analysis, natural feature extraction and recognition**

Natural feature extraction refers to the mathematical processing of an image to extract a set of numerical values that uniquely represent that image. Features of an image are commonly selected from its colour histogram [21] or from structural edges and their spatial relationships [5]. A reduced subset of these features is subsequently defined to uniquely identify the image.

#### **6.1.1 Mobile camera systems**

Articles grouped in this section consider solutions where the camera was carried along by the subject or device being located. In most cases the localization process involves two stages, analogous to the two-stage fingerprinting process described for wireless localization solutions. In the off-line stage, images of the environment are captured at predefined locations. Each image is processed to extract its unique features. Subsequently, the extracted features are stored in a database along with the associated camera position and orientation, at which the image was captured.

In the on-line stage, the camera captures an image, and image features are subsequently extracted. These features are compared to the entries in the feature database, either directly (to define a crude location of the camera), or through spatial transformations that yield the best match between the features in question and those in the database (to obtain more accurate location estimates). Such transformations reflect the differences in between on-line and offline image capture locations. Once the best matching set of database features is identified, the corresponding database entry for camera position and orientation is used to estimate the absolute location of the camera.

Articles adopting a mobile camera system include [5, 94, 24, 21, 93, 70, 31]. It must be noted, however, that these articles propose the indoor localization of robots. Nevertheless, the algorithms used in such articles are not strictly dependent on properties of the robots and can be easily adapted to human localization. Three distinctive examples are presented in this section.

Observing the rectilinearity of human-made indoor environments, Aider, Hoppenot and Colle [5] proposed a localization system based on a mobile monocular vision system which exploited straight line environmental features. The maximum localization error reported

during their experiments was 20 centimetres, and the maximum angle estimation error reported was 2.5 degrees.

Frontoni and Zingaretti [21] used a single colour camera to detect coloured areas of an image and their spatial relationships, which were, in turn, used as features. Through a rotation invariant feature transformation, they estimated the most probable location of the camera from a previously created database of features. The authors reported indoor localization errors less than 50 centimetres, 94% of the time.

In [31] the authors presented a navigation system for a humanoid robot, which can be extended to the case of human walking. Location determination was accomplished in two stages. In the off-line stage, images were captured while following the route between two points. The on-line stage involved autonomous robot navigation between two arbitrary points while capturing images. To achieve localization, an algorithm correlated freshly captured images with images in the route database. In this way, the unprocessed raw image was considered as a large set of features. Correlation analysis yielded the position deviation between the learned route and the current position. Temporary occlusions were detected as sudden drops in the correlation. The maximum position deviation reported during an experimental 17 metre long route, was 0.9 metres. This represents a deviation of 5.3% of the walked distance.

### **6.1.2 Fixed camera systems**

When the camera or cameras of the system are mounted in fixed locations in the environment, the structural features of the building cannot serve as discriminant factors for localization purposes. Instead, features of the object being tracked must be used. In this way, if the salient features of the object or objects appear in the field of view of the camera, the location of the object or person can be calculated with respect to the camera's fixed position. The position of the object of interest in the environment is estimated based on its position within the captured image, and the spatial distribution of its salient features.

Although several candidate articles discussed solutions for tracking the position of a person using cameras fixed in the environment, few considered tracking a person in a building. Instead, most of the articles presented a solution in which the tracking process was limited to the field of view of a single camera. Therefore, only articles [58, 97, 89, 48, 10] were considered as an indoor tracking solution based on our inclusion and exclusion criteria.

In [58], the authors proposed a system that detected elements of the scene which were not part of the static environment. Through segmentation, the colour histograms of the region or regions of interest were obtained, along with a vertical colour average to estimate a general vertical orientation of the object detected. In this way, the algorithm accounted for the global colour scheme of the clothes of a person while standing up. A camera, located at the entrance of the test environment, was used to create an initial colour model of each person. Fifty subjects were tracked within an office environment using surveillance cameras in this fashion. Visual overlaps between cameras allowed constant tracking of subjects. The authors reported correct user recall values of 87.21% with an availability of 73.55%.

A similar approach was presented in [97]. The authors used colour-based features assuming vertical differentiation of colour regions of a human figure while standing up. To account for movement between cameras with no visual overlapping coverage, the authors proposed the creation of a connected graph that represented the areas covered by each camera. The edges of the graph denoted physical connections of the areas covered by the field of view of

each camera. This connected graph was used in the probabilistic modeling of the movement pattern and traffic constraints of the user, to improve recognition and tracking accuracy.

Pursuing a slightly different approach, [89] proposed a distributed sensor network based on image processing. The robustness of the solution relied on overlapping the visual field of several cameras, and distributing computational processing for localization purposes among the elements of the distributed network. Features were extracted using a principal components analysis of features obtained by differencing consecutive segments of the image. The field of view overlap of different cameras allowed robustness against occlusions. The authors demonstrated the feasibility of human tracking in a crowded setting. The localization accuracy, however, was not reported.

In a similar vein, [48] proposed a system with 4 cameras with partially overlapping coverage areas. Colour and non-colour features were used to account for areas of interest in the image that yielded colour and illumination information, respectively. These cameras were calibrated to estimate the 3D position of the user within the 2D image using a projective algorithm. Using an evidential filter, particles represented the probability of a user being in a certain location. Experiments were conducted with multiple users navigating a room concurrently, without deliberately avoiding occlusions. Three hundred particles were used to represent each user. In some cases the users wore similar clothes. Some of these users walked out of the field of view of some of the cameras intermittently. The reported localization accuracy of a single user using the 4 cameras was 0.15 m, while using a minimum of 30 particles for the evidential filter.

Finally, the Easyliving project [10] proposed the usage of image-based localization systems to provide context awareness within intelligent environments. The project was based on ceiling-mounted stereo vision cameras capable of estimating the anatomical posture of humans or the orientation of objects in the environment. The identification of the subject being tracked was based on colour and structural features. The localization accuracy was not reported.

## 6.2. Image analysis, fiducial markers

Fiducial image detection differs from natural feature detection in that the image processing algorithms are designed to detect predefined, synthetically created, patterns in the environment. These patterns are called fiducial markers [51, 19].

In principle, localization algorithms based on fiducial recognition are very similar to localization based on natural feature recognition. However, as the properties of the image to be detected are constrained, there is no need for an *a priori* stage to extract the features of the fiducial. A database is still required to determine the location of the camera relative to a reference frame, fixed with respect to the environment. In this case, however, the database can be created automatically by storing a numeric ID associated with the fiducial along with its coordinates in the environment [94]. The ID of the fiducial can be encoded within the fiducial image, analogous to how the image of a bar code represents a numeric ID. Some reported advantages of fiducial recognition over natural feature recognition are reduction of computational requirements, improved detection accuracy and resilience to noise artifacts [51].

In [94], the authors proposed a position detection system for robot navigation in indoor environments. They conducted a simple experiment using a fiducial marker as reference. They presented an image analysis technique based on homography to obtain the relative

position of the camera with respect to the fiducial marker. Although developed for a robot, the localization technique is not constrained to intrinsic properties of the robot, and can be applied to human localization.

Kim and Jun [39] introduced a wearable indoor localization system composed of a portable computer, a head mounted display (HMD) and a camera. Localization was achieved through image processing, combining fiducial markers and natural visual feature extraction. Localization with fiducials was achieved through an open source library called ARToolkit [37]. The localization algorithm detected a synthetic visual pattern in the image captured by the camera. Using affine transformations, the authors estimated the distortion and scaling of the fiducial due to angle of view and capture distance. With this information, the exact position (i.e. distance and orientation) of the camera with respect to the marker could be determined. The authors modified ARToolkit, adding an adaptive illumination thresholding algorithm and an algorithm for natural scene feature recognition, which determined the location of the person when there were no fiducial markers in view. Natural feature recognition was achieved by analyzing the color and hue histograms for a sequence of frames. The authors defined a "location" as a sequence 64 consecutive frames. Hence, each location was defined by a high dimensional space, namely a sequence of 64 colour and 64 hue histograms. A Linear Discriminant Analysis (LDA) was applied to reduce the dimensionality of the features which defined each location. The first five LDA coefficients of each frame were used as descriptive features of that frame. The sequence of features in consecutive frames defined a location. An off-line navigation phase facilitated the creation of a database of natural features of the building. To estimate the location of the user, the vector of features obtained from the last 64 frames, defining the current location, were compared against the features in each entry of the database. The difference between a feature vector captured by the camera in the on-line stage against feature vectors stored in the database were quantified by the Euclidean distance. If the Euclidean distance was lower than an unspecified threshold, the location of the person was defined as the matching database coordinates. In this fashion, the system yielded the general location of the user in a continuous fashion. The exact location of the user was determined using fiducial markers in the field of view of the camera. The location information obtained by both visual systems was presented in a virtual map, along with instructions to complete a predefined path through the HMD. The localization accuracy of the system was not reported.

### 6.3 Other photonic sensors

Articles included in this section make use of non-image capturing Infrared (IR) sensors [15, 56, 55, 14, 33, 90].

In [15], the authors presented an IR proximity-based localization system, which provided museum visitors with useful information about exhibits in each hall. For this purpose, IR emitters were installed in the ceiling of the door frames of every room. Each emitter transmitted a unique ID using the Infrared Data Association (IrDA) protocol. The visitor carried a Personal Digital Assistant (PDA) with an infrared port. The PDA contained a database of visual and textual information of the exhibits, as well as maps of the museum. Upon reception of a new ID, the PDA automatically presented the map of the corresponding hall. While in the hall, a graphical user interface in the PDA helped the visitor to obtain information about a particular exhibit. The authors noted some problems while deploying

this localization system, in particular, noise and reflections of IR signals. The user localization granularity of this system was down to the scale of a room.

A combination of scene-analysis and triangulation is presented in [56]. In this solution, a unique ID was modulated by the IR emitter. The carrier frequency used for modulation was changed in a cyclic way, from low to high frequencies. As the attenuation properties of an infrared signal are frequency-dependent, ID's modulated at lower frequencies can be successfully detected farther away from the emitter. However, the power of the IR signal decays in a nonlinear fashion with distance. To characterize this power decay, the authors obtained off-line measurements of the received signal in the vicinity of the emitter, in a similar fashion to fingerprinting for RF signals. The authors measured the ID detection success rate in 10 cm concentric regions, at steps of 5 degrees, repeating this process for different modulating frequencies. Instead of creating a database of the signal in each point, the authors modeled the decay of the signal with an equation that was dependent of the orientation of the receiver, the distance between receiver and emitter, and the modulation frequency. Using this approach the system achieved a maximum localization error of 10 centimetres, within an area of 5 square metres. Although a building-wide experiment was not conducted, the usage of a unique ID per transmitter would allow the deployment of multiple emitters in a cellular arrangement to cover large areas.

Petrellis et al. [55] presented a localization system consisting of two IR emitters fixed in the environment, and two receivers installed on a mobile unit. The transmitters were mounted facing each other on the walls of a corridor, while the receivers on the mobile unit faced away from each other. Each emitter transmitted a series of unique cyclic data patterns, modulated at a carrier frequency of 38 Hz. The opposite-facing arrangement of the receiving sensors facilitated detection of user orientation and, at the same time, discrimination between signals received via direct path and reflected signals. Reflection rejection was enhanced using a predictive model, which took into consideration the immediate previous location and orientation of the user. Since the system relied on predictions of future position and orientation of the mobile device, the estimation rules constrained the maximum linear and angular speeds at which the system was reliable. Given that the emitters transmitted unique sequences, the authors proposed a cellular-like spatial emitter arrangement to cover extensive areas. The system was sensitive to moving personnel and other objects that caused reflections although compensation algorithms reduced such effects. The localization accuracy was not reported.

Cheok and Li [14] presented a localization system which made use of existing fluorescent lamps in a building. A user carried a wearable computer instrumented with a photo-detector and a gyroscope. Each fluorescent lamp emitted an ID, associated with the location of the lamp. The ID was encoded through pulse frequency modulation. The encoded information did not introduce perceptible illumination effects. The photo-detector and gyroscope were mounted on a cap, worn by the user. The gyroscope served to obtain orientation information. The wearable computer was carried on a vest, worn by the user. The area illuminated by two lamps transmitting different location ID's could not overlap, since this would cause interference, resulting in the inability to detect the encoded signal. Whenever available, location information was presented to the user via the wearable computer. Such information was overlaid on the user's visual field through a HMD. The reported localization accuracy was within three to four meters. Although this accuracy is



dependent on the lamp and user heights, these numbers were not provided. The minimal separation between lamps was of 2.31 metres.

iGPS, commercialized by Metris (formerly ArcSecond) [33], is a triangulation-based localization system for tracking assets, personnel or any other mobile elements. To accurately estimate the position of a receiver, a pair of eye-safe IR laser emitters radiate a signal using two different wavelengths, while an infrared strobe provides a reference signal. Using signals from 3 or more transmitters, the receiver calculates and transmits its position to a central data collecting station. In order to estimate the orientation of a solid body, two or more receivers are attached to it. The iGPS system claims to offer sub-centimetre accuracies.

Scientists from Olivetti Research designed the Active Badge location system, which consisted of small badges that transmitted a unique ID via IR emitters [90]. The badges were worn by people, who could then be located when in the vicinity of a receiving station. The badge transmitted a unique identification code every 10 seconds. The system offered sub-room accuracy, and was used to redirect phone calls for personnel.

#### 6.4 Limitations

Common problems reported for photonic sensor localization are the ambient noise in the form of light or thermal radiation [5, 56, 97], signal reflections [15, 44, 56] and in the case of image processing solutions, illumination variability [58, 89, 78, 97].

In image processing solutions, ambient noise is usually overcome by image filtering techniques. In the case of IR sensing, effects of ambient noise can be mitigated by using a combination of different modulation frequencies [56, 33].

Another problem commonly mentioned in the surveyed articles is the occlusions caused by dynamic elements of the environment [5, 24, 21, 89, 10]. For instance, in the experiments performed in [56], the introduction of new objects or humans was specifically avoided during the experiments. In order to reduce the risk of visual occlusions by humans and objects, solutions comprising building-mounted equipment commonly install the detection equipment on the ceiling [15, 90]. Another way of reducing occlusion is to deploy sensors with overlapping coverage areas [58, 89, 33].

However, clinical settings and public indoor areas like shopping malls are oftentimes densely populated. Consequently, occlusion conditions can emerge frequently even with ceiling mounted sensors. Therefore, if a photonic-based system is to be considered for localization purposes, it may be advantageous to simultaneously invoke a secondary tracking system to assist in the localization process during periods of optical occlusion.

In the case of laser based-solutions, only class 1 laser devices should be used, which are classified as "eye-safe" by the IEC 60825-1 standard [32]. In clinical settings, however, special care must be taken with even class 1 laser devices, to ensure that no harm will be caused by concentrated light on the skin or eyes of light-sensitive patients.

Finally, we should emphasize the privacy issues that may arise as secondary to a localization solution. Health care facilities operate under the precepts of information non-disclosure to protect the privacy of personnel, patients and clients. This is an important consideration when a localization system is designed to capture images of the environment, as such images can reveal important information about the person wearing the system or about the patients and health care personnel in the vicinity. Frequently, image processing mobile localization devices are designed to send the captured image to central,

## Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

