Chapter 1

Introduction

Finding a specific place or location one has never been to, or hasn't been to for a long time, is a common task that everybody encounters from time to time. Back in the days, almost everyone had a road map stowed away within the car's glovebox, ready to use, whenever needed. However, without a co-driver, reading the map and providing instructions, this was a quite cumbersome solution for getting from A to B. This situation changed in the year 2000, when the former military-only global positioning system (GPS) became freely available for civilian use. Now it was possible to locate objects anywhere on the earth, with an accuracy down to a few meters, using just a single receiver. Combined with digitized maps, this allowed for both, self-localization and navigation [DH10].

Starting from there, it only took several months for receivers to become both, significantly cheaper and smaller, and companies like *TomTom* or *Garmin* started developing products for motor vehicles, using digital maps from vendors such as *Tele Atlas* or *Navtech*. At first, navigation systems were either installed directly within a vehicle, or required fully featured hardware, like portable computers equipped with an external receiver. Yet, with the advent of *Personal Digital Assistants* (PDAs), containing even smaller GPS receivers and mass storage devices based on flash memory, navigation systems became portable. Today, almost every new smartphone is suitable for GPS-based navigation, using its built-in sensors, as well as a piece of software that includes the necessary maps and navigation algorithms.

Inexpensive receivers for GPS, new similar systems, like GLONASS, and (freely) available maps for almost every place on earth, lead to the ubiquity of navigation systems. Thus, their success was not only based on demand, but also on the availability of relatively affordable components for hardware, software, and low running costs. At least for the *customer*: Consumer hardware can be used for several years, and, depending on the vendor, map updates are either free of charge, part of an annual subscription, or charged per update.

For *providers*, however, the situation is different. Additionally to several billion dollars for the initial development and setup, the infrastructure behind GPS has to be kept up and running, costing additional millions – per day. While licensing fees, e.g. for access to increased accuracy, compensate for some of these costs, the remaining part is paid by the US government. Running costs for the *vendors* of navigation appliances can be expected to be much cheaper. However, due to rapid changes in infrastructure, they have to provide up-to-date mapping data, resulting in many companies charging for updates [Pac+95; DH10]

Due to ever-increasing globalization and transnational business connections, the need and wish to move or travel is constantly increasing. Besides getting to airports, train stations or company compounds, the buildings themselves often represent a navigation problem as well: Finding the correct terminal within an airport, the conference room in a large company, a room within a townhall, or the correct ward within hospitals, isn't always straightforward. With this in mind, localization and navigation *indoors* becomes of increasing importance as well.

However, while working perfectly for most navigation purposes, e.g. for cars, pedestrians and cyclists, currently available systems are unsuited, as both, the sensors and the typical map formats, are intended for *outdoor* use. For good location estimations, GPS relies on a direct line-of-sight between satellites and receiver, and older devices thus had to be installed on top of the car, in order to function properly. Similarly, the format of most digitized maps is focused on outdoor purposes, as the underlying data structures mainly use a two-dimensional representation of roads, lanes, and intersections, unsuited for modeling a building's interior.

Furthermore, when considering indoor environments, completely different use cases, besides typical navigation from A to B, arise as well. Starting from finding a specific product within a large supermarket, to the economy's interest in location-based services, e.g. placing ads for nearby products as well. Also covering cultural aspects, like guided tours through a museum, presenting useful information on exhibits, based on the visitor's current location and viewing direction. Depending on the building and intended use case, requirements can be completely different. This especially concerns the aspect of *localization accuracy*. While a coarse GPS location estimation is sufficient for a car driving along the motorway, it can be too erroneous for a slowly paced pedestrian, walking through an area with many small alleyways. The same holds true for localization indoors, where estimating the current whereabouts on a room-level scale might be sufficient for some intentions, like presenting information on nearby exhibits. For others scenarios, such as navigation, however, estimations should be as accurate as possible, for audible commands and visualizations given to the user, to be helpful instead of misleading.

Therefore, the question arises, how such a multi-purpose indoor localization and navigation system can be developed, and what criteria should be met for it to be valuable.

1.1 Navigation within Buildings

Based on the previous aspects, it becomes clear that the topic of localization indoors is not solely related to sensors and achievable accuracy, but also to costs, for initial setup, maintenance over time, software and hardware required by the consumer, and by the system's operator. In case of localization indoors, the latter is unlikely to be a government, like it is for the GPS, GLONASS or Galileo, but more likely the owner of the building to deploy the system to, like an airport, hospital, supermarket or museum. This gives even more importance to the aspect of costs, as many public buildings that benefit from indoor localization, like townhalls or museums, typically are on a tight budget. Closely coupled with costs is the time required for setup and servicing, as they also arise per building, additionally dependent on its size. As known from other projects, the solution is a tradeoff between quality (accuracy), time and costs.

Similar aspects apply to the required building maps. As it is unlikely for a global company to create maps for every single building, where indoor localization could possibly be used, this data has to be supplied by the operator or a public community, dedicated to this task [Ope]. Furthermore, in contrast to maps for navigation outdoors, indoor maps can be rather eclectic, as they have to support buildings with multiple floors, elevators, escalators, and different types of stairs [EBS16; Elh+14]. Depending on the intended use case, they should also support adding semantic information, like room numbers, points of interest, and access restrictions or limitations. The latter is especially relevant to the disabled, who are unable to take stairs, or require additional audible information when visually impaired. These aspects can also affect the topic of navigation, as the shortest path towards the destination might not be the best solution for all pedestrians, especially not for those being handicapped or injured.

Based on the previously mentioned thoughts, a non-exhaustive list of requirements for indoor localization and navigation thus contains the following aspects:

- *Software and Hardware required by the consumer* should be as cheap as possible, with required components being small and always at hand, if possible.
- The system's *accuracy* must be sufficient for a pedestrian to be localized within the building, and to provide navigation guidance. Hereby, *sufficient* is not quantifiable, strongly depends on the intended use case, and the building's architecture, as narrow corridors with many adjacent rooms require a higher accuracy than e.g. large, open shopping malls.
- Time and costs for the *initial system setup* should be as low as possible. This includes costs for all necessary hardware components, time for their setup, and effort needed to provide a digital map of the building's floorplan.

- Time and costs for *maintenance* after the initial setup should be as low as possible. Ideally, the system is easily adaptable to architectural changes, like new/removed drywalls.
- *Partial failures of the infrastructure* should not completely disable the whole system, only may affect the provided accuracy.

Besides use case-dependent details, the question of suitable hardware components is the most critical. As existing positioning methods like GPS and GLONASS do rarely work indoors, other sensors are required to infer an absolute location. As of today, there is no established solution, and this matter is still open for new suggestions. However, to conform with previous discussions, it should not only be accurate, but also cheap, and easily available. Therefore, most ongoing research is targeted at *smartphones*, as they are ubiquitous, almost always at hand, and contain an increasing number of sensors [Tia+15; Gui+16; Ndz+17; Ye+14; Mou+15; Kir+18].

That is in contrast to outdoor navigation, where new platforms started to develop around the existence of a single sensor. For indoor localization and navigation, a desirable target platform is already available, and the question arises, whether it is suitable for the intended task. This lead to numerous new research topics, analyzing the suitability of certain sensors, that are installed within commodity smartphones. Most of them are adapted from previous research in different fields, where some sensor or component has already been proven helpful.

This e.g. covers velocity and heading, estimated from an accelerometer and a gyroscope, together providing the base for *dead reckoning* [ND97], which allows relative (incremental) location estimations, if initial whereabouts are known. This technique already underwent extensive research to adapt it from vehicles to pedestrians. Yet, the focus was mainly on *multiple* sensors, attached to different parts of the body, picking up leg movement and turning behavior of a pedestrian, well-suited for motion estimations [SD16; TS12; Goy+11]. With the rising interest in indoor localization, it began to be adapted to smartphone-only setups, where the orientation of the device has to be considered, when the pedestrian e.g. holds the smartphone upfront, looking at its screen while navigating through a building [PHP17; Yu+19; Kus+15].

Yet, with dead reckoning providing information on *relative movements*, it is only suitable when initial whereabouts are known, and it is likely to fail over time, due to increasing errors. For actual indoor localization, hints on absolute whereabouts are mandatory. For this, former research on Wi-Fi-based location estimation [BP00] became of interest again. By using signal strength observations from nearby access points, it is possible to roughly estimate the distance towards them, and thus a coarse, absolute location information. This strategy also conforms with most aforementioned requirements: As of today, most public buildings are equipped with Wi-Fi, already containing the required infrastructure, and Wi-Fi is supported by all modern smartphones. However, besides these positive aspects, achievable accuracy is either too coarse,

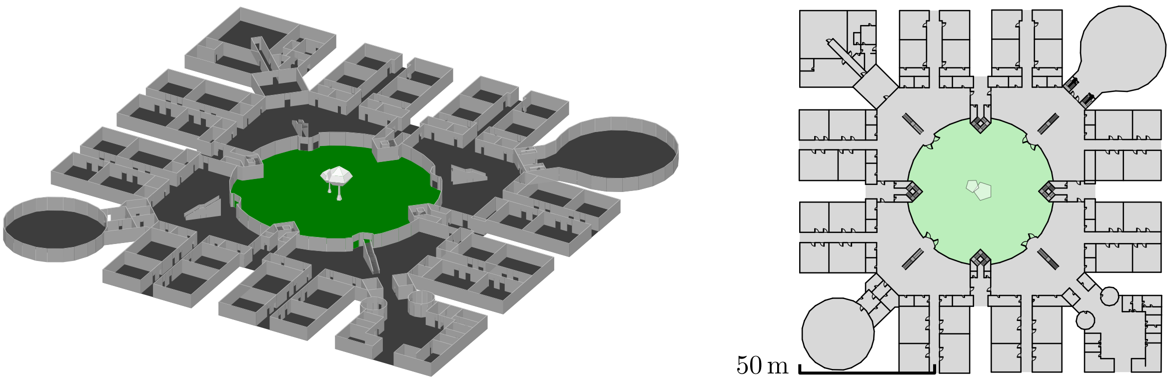
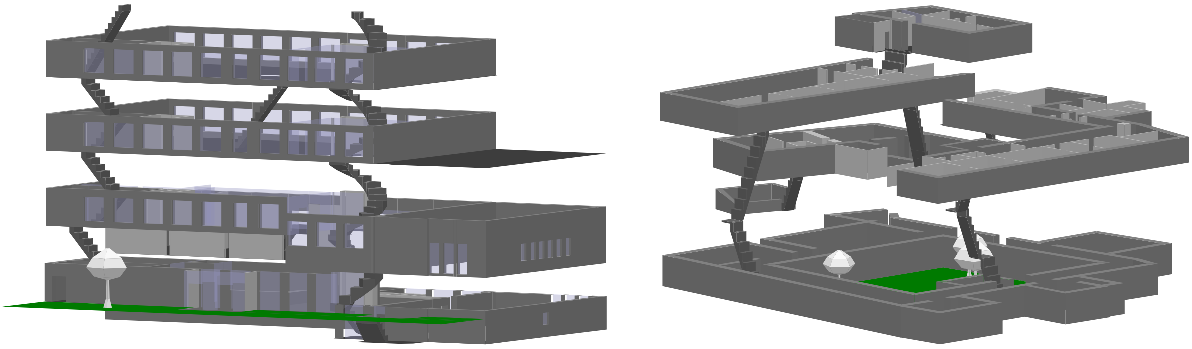


Figure 1.1: Example of a complex single-floor, with large open spaces and small adjacent rooms. First floor of the UAH building of the *University of Alcalá de Henares*, Spain.

or a manual and time-consuming setup is required beforehand. During the latter, accuracy is increased by actually measuring the behavior of the installed infrastructure's radio signals, throughout the whole architecture. Thus, this area is still undergoing extensive research.

In contrast to navigation outdoors, there is not yet a single sensor that solves the problem formulation with sufficient accuracy. Instead, research tends towards employing combination of multiple components, each of which providing a contribution to the overall result. Besides the two mentioned examples for relative and absolute estimations, various other sensors, such as the camera, magnetometer or barometer, which are also found within smartphones, can thus be of interest as well [HB08; Shu+15; Mur+14].

Alongside sensors, where some components already seem established, mapping still requires extensive research. In outdoor navigation, a graph data structure is ideal to model rivers, roads and interconnections, for both, displaying and routing. Considering indoor use cases, however, there is not yet a clear best-candidate among potential data structures [ARC12]. Indoor environments are less restrictive and often inhomogeneous, ranging from narrow hallways with multiple adjacent rooms, to large open spaces, as can be seen in figure 1.1 and 1.2. This scalability must be supported by the chosen model, including minor details where needed, yet without requiring too much memory. Furthermore, the map has to provide all the semantic information that might be required for some sort of sensor component. Additionally, multiple floors and their interconnections, like stairs, escalators or elevators, are also a strong requirement. Not to mention *editability*, as the map has to be generated for each and every building, with support for including future architectural changes. The problem of creating a 3D representation of such a multistory building has already been solved by computer graphics [KSS17]. Yet, determining whether a particular movement is possible, calculating the shortest path towards a room or point of interest, correctly including stairs and elevators, all while being computationally efficient, still is a topic of active research.



Deutsches Hutmuseum Lindenberg, Germany *RothenburgMuseum*, Germany

Figure 1.2: Two complex multi-floor buildings. While the left one is stacked almost evenly, the right one is irregular in size, shape, and floor-level. The distance between floors was increased for visualization.

Mentioned earlier, the floorplan not only serves as a visualization to the user, it contributes valuable information as well. The map within car navigation systems is also used to compensate uncertainties of the GPS, e.g. by placing the virtual car onto the nearest road. Additionally, when the car drives through a tunnel, and the GPS signal is lost, the last known velocity and heading can be used to continue predicting the car's whereabouts, based on the underlying mapping information. Similar aspects apply to localization and navigation indoors, where the map is used to denote possible movements, limit impossible movements, and to prevent the impact of sensor uncertainties and errors. For example, assuming two subsequent absolute location observations to be ten meter apart from each other. Such a change in location is likely, when both locations refer to the same floor, and several seconds have passed between the two sensor observations. Similarly, such a change is unlikely, when e.g. only one second has passed, or both locations belong to two different floors, and neither stairs nor elevators nor escalators are nearby. By combining assumptions on *pedestrian walking behavior* and information provided by the floorplan, probabilities for potential location changes can be inferred.

Aforementioned aspects lead to the requirement for a technique, which fuses all available information, to derive the overall result. As every sensor provides its own point of view, there is no straight-forward solution of combining all observations. Especially in case of sensors indicating relative location changes, restrictions of the floorplan should be included to rule out physically impossible movements. Furthermore, every single component is subject to different types of errors that must be considered as well. The overall task thus is to determine the *most likely* whereabouts, based on all sensor observations, assumptions, and the building's floorplan. Depending on the complexity of the latter, and the number of sensors, this task can exceed the capabilities of embedded devices, and represents an extensive research topic on its own [Gus10].

Based on all presented thoughts and requirements, the research objective of this work is formulated within the following.

1.2 Research Objective

In contrast to outdoor navigation, where most devices were developed around a single sensor, with its accuracy sufficient for most use cases, as of today, pedestrian indoor localization relies on multiple sensors, with the smartphone representing a desirable target platform. The goal of this work is to derive a scalable system for pedestrian indoor localization and navigation, targeting this platform. Thus, the focus is solely on smartphones, the sensors available within, and to build a system that is suitable for most use cases, easy to set up and maintain. Neither requiring large amounts of time, nor cost for setup and infrastructure. While considering solely sensors and infrastructure available as of today, the discussed system is intended to be scalable, allowing for easily including new sensors in the future. For the use case of localization and navigation, the smartphone is expected to be held upfront by the pedestrian, e.g. looking at navigational advice, presented on the device's screen. This aspect is relevant to certain sensors and corresponding coordinate systems, discussed throughout the course of this work.

With GPS being unavailable indoors, Wi-Fi is considered the main component for absolute location information, as required infrastructure is available within most buildings where localization or navigation are a benefit, and it is supported by most of today's smartphones [BP00; YA05; Roo+02; Liu+12]. Yet, with the expected accuracy being insufficient for navigation, additional sensors are required. Here, the focus is on well-known dead reckoning techniques that are adapted for use on smartphones. This e.g. covers the smartphone being held upfront by the pedestrian, therefore applying required compensation techniques. Besides, additional sensors, such as the barometer and magnetometer, will also be considered, providing further information to increase the overall accuracy, without affecting setup, costs or maintenance. As discussed, every sensor component is subject to different types of errors that have to be handled accordingly. Therefore, the focus is on *probabilistic* approaches, including all sensor observations based on their likelihood. That is, for every individual component, a probabilistic model will be derived, describing the likelihood of some whereabouts or movements, from every sensor's point of view.

Not only relevant for visualization purposes, but also for limiting impossible movements or for providing routing information to derive the best path towards some destination, the building's floorplan represents the second major research objective. Conforming with sensors and aforementioned aspects, probabilistic movement models will be derived, where the floorplan is used to describe potential and unlikely pedestrian movements.

The information from individual smartphone sensors is combined by *sensor fusion*, based on *recursive density estimation* [MU49; Mar51; Sär13]. This is used to determine the globally most likely whereabouts, based on *all* sensors observations since starting the estimation process.

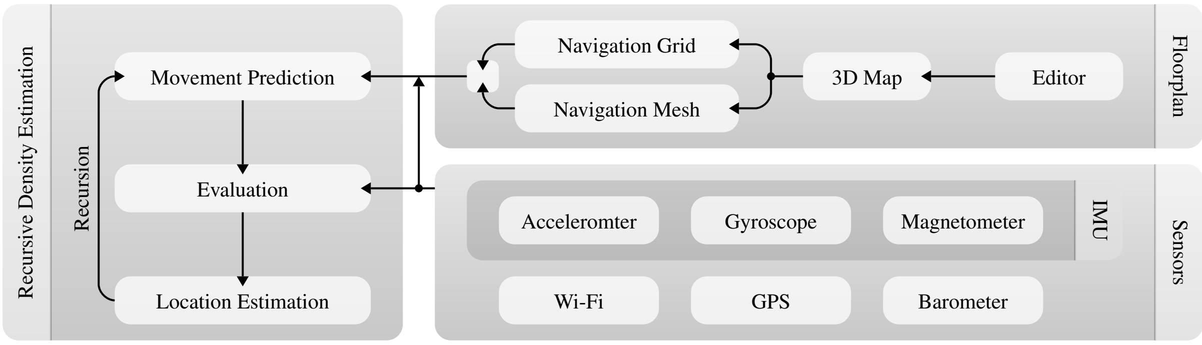


Figure 1.3: Brief overview of the overall system. Floorplan and Sensors represent the main source of information, combined via recursive density estimation, determining the most likely whereabouts.

By considering the *history* of all sensor observations, *relative* location information, like aforementioned dead reckoning, are supported as well, and results are refined over time. Throughout this process, the floorplan will be included, used to e.g. filter impossible movements that would cross a wall or other obstacles. To include individual errors, chances and similar, all required calculations are given on a probabilistic basis.

Figure 1.3 provides an overview of the overall system, its individual components, and the way they interact with each other. This figure is intended to provide a brief impression on the global research objective, without going into details of each and every component. As can be seen, the sensors and the building's floorplan represent the two main sources of information, combined via recursive density estimation. Both, sensors and floorplan, are intended to be interchangeable, with the ability to include new sensors and spatial models, scaling with new future components. To get an impression on the impact of choosing some specific data structure, two different spatial floorplan models, as well as their advantages and disadvantages, will be discussed. This also addresses the topic of how to include *semantic* information, e.g. to label a room, or to include additional information, useful for routing or people with special needs.

To summarize, the focus of this research is on deriving a smartphone-based pedestrian indoor localization and navigation system, enabling to localize *oneself* within a building, e.g. for navigating to a desired destination. This is achieved by adapting existing techniques to this use case, combining the information form several smartphone sensors with movement prediction based on the building's floorplan, by using probabilistic sensor fusion. Other use cases, such as localizing all pedestrians currently residing within a building [Xu+13], are not covered by this work. Also excluded are topics that are related to indoor localization, but not to pedestrians, like determining the current location of some equipment within a large industrial compound [Nuc+04; Kar+17]. Furthermore, the focus is solely on ubiquitous components. Special hardware for accurate localization indoors, such as ultra-wideband [FG02], is thus not considered.

1.3 State of the Art

This section provides a brief overview on the current *state of the art*, concerning the main topics identified during previous remarks and the research objective. More detailed overviews, and *related work* from other researchers, are given within each of the chapters, and individually for every topic.

While indoor localization and navigation became of increasing interest to researchers during the last decade, there is no standardized solution yet. Even when referring solely to smartphone-based systems, the sensors used, the way they are integrated and combined, the required infrastructure, and the underlying spatial models for the floorplan, if used, are completely varying. Most systems refer to some sort of probabilistic setup, combining individual components, based on likelihoods. However, the scale of integration, that is, the number of sensors that are combined, and the degree of additional information added, like the floorplan, is significantly varying. Often, limited fusion techniques are applied, being computationally efficient, but unable to fully include all available information, such as obstacles, or the pedestrian's desired destination [Tia+15; Hel+13; Ndz+17; NRP16; EBS16; Zha+18b].

Probabilistic Sensor Models As mentioned, core components of the system are sensors, providing information on whereabouts or movements. While the latter can be performed using solely dead reckoning, that is, starting from a known location with incremental updates based on detected movements, this also leads to incremental errors [Ser28]. These errors eventually were considered, estimating the likelihood for certain whereabouts, and their changes over time [Goy+11; Li+12]. Yet, the degree of considered information varies significantly. While some works consider only two sensors and their respective uncertainties, others include additional observations from other components, and further assumptions, affecting the way the probabilistic models are defined and handled [Hel+13; KGD14; Tia+15]. As shown by others, and discussed in a later chapter, probabilistic sensor models that consider prior information, such as the floorplan, can mitigate growing uncertainties, and increase the quality [NRP16; Kna17].

Probabilistic Wi-Fi Localization With Wi-Fi representing an infrastructure already available within most public buildings, it is also part of many indoor localization and navigation systems. Yet, implementations often rely on a complex and time-consuming setup procedure, conducting fine-grained measurements throughout the whole building, to estimate the behavior of radio signal propagation, required for inferring potential whereabouts [Men+11; YWL12; Zha+18b]. These initial measurements can later be compared against readings from the pedestrian's smartphones, to determine the best matching one, representing the current whereabouts. This variant of localization is rather discrete, and based on the density of these initial measurements. While interpolation techniques exist, they suffer from various drawbacks, and come with a computational overhead, often exceeding the capabilities of embedded devices [Par62]. Furthermore, resulting accuracy comes at the cost of setup and maintenance times, whenever the architecture or Wi-Fi infrastructure is modified. When on a tight budget, different approaches are required.

These are e.g. given by describing radio signal behavior, using some sort of model [SR92; PC94; JLH11]. Similarly to the initial measurements approach described above, the model's predictions can then be compared against current readings from the smartphone. However, as the model is typically able to perform this prediction for any location within the building, it is continuous, and does not require for additional interpolation. Yet, for every prediction model several parameters are required to describe the behavior of radio signals. The prediction quality thus not only depends on the accuracy of the model itself, but also on the chosen parameters [Sey05; Hee+11]. For use cases where a reduced accuracy is sufficient, empiric values can be chosen, allowing for a fast deployment and adaption to infrastructural changes.

However, for most setups, a compromise between both techniques represents a viable trade-off, with sufficient accuracy and fast setup times, thus being the focus within this work.

Building Floorplans and Probabilistic Movement Prediction With the floorplan representing an important component of every localization and navigation system, not only for visualization but also for limiting impossible movements and routing, it is part of many state of the art systems. Yet, as there is no standardized format for indoor floorplans, and many spatial representations are suitable [Led06; Yan06; Wu10; ARC12], different approaches have established over time, most of which limited to a specific use case.

Simple 2D setups e.g. describe each floor with lines that can be used for intersection tests, to determine impossible walks [EBS16]. This, however, is not suitable for most buildings, as they consist of multiple stories. Therefore, 2.5D setups were derived, created by stacking multiple 2D floors, with a discrete connection in between [GF06]. Yet, these setups suffer from various drawbacks. On the one hand, intersection tests are costly, thus requiring some sort of pre-calculated approximation for use on embedded devices [Köp+12; NRP16]. On the other hand, due to the *discrete* interconnection, changing floors requires some sort of heuristic or additional sensor information. Besides, this also yields a reduced user experience in visualization.

For both, visualization and prediction, actual 3D representations thus are preferred. To be suited for use on smartphones, the spatial model should be conservative in use of memory. Viable is e.g. a polygonal representation of the walkable surface [WH08], or some other type of primitive [BJK05]. Referring to the aforementioned problem of costly intersection tests, the 3D spatial model should also be able to quickly determine whether two whereabouts are connected or separated by an obstacle, and, if navigation is desired, the shortest path in between.

Independent of the chosen inclusion and spatial representation, the floorplan must be defined in some way or another. Besides manual creation, crowd-based approaches can be suitable, e.g. determining the walkable area by recordings from hundreds of pedestrians, refined over time [AY12]. Yet, this only allows for a coarse representation, not ideal for visualization purposes.

Alternatives are e.g. given by robots equipped with a laser-scanner, recording the building's interior to derive a 3D representation [SCI13; Hes+16], using several panoramic images to estimate depth [CF14], or scanning the blueprint and using algorithms to derive walls, doors, stairs and similar [Liu+17]. However, dependent on the chosen strategy, expensive hardware might be required, stairs are not supported, or semantic information, like room numbers, still has to be added manually.

The quality of the resulting floorplan strongly depends on the chosen technique and the building's architecture. The same holds true for the time needed to acquire all required information. A manual setup, using some sort of editor, thus also is a viable choice.

Sensor and Information Fusion As identified earlier, individual sensors and information should be fused together, including the history of all observations, to derive the globally best solution, based on all previous inputs. Ideally, individual uncertainties are included as well, to decide how trustworthy each information is. The domain of sensor/information fusion, also referred to as *recursive density estimation*, is well-researched, both, analytically and experimentally. Initial analytical approaches were limited to linear and Gaussian problems only [Kal60]. While this is sufficient for some setups, such as basic inertial predictions [Meh70], or general tracking approaches [CHP79], for more complex problems, such as indoor localization and navigation, including the building's floorplan, it is not.

When relaxing some requirements, and slightly modifying the analytical process, nonlinear problems are supported as well [SSM62]. Concerning indoor localization, these changes add support for basic parts of the overall system, like step-detection and tracking [Goy+11; Jim+12; Gar+16]. Yet, more complex information, such as a building's floorplan, can still not be included, as it is impossible to describe the impact of walls, stairs, and similar, on a purely analytical basis.

For this, non-analytical variants were developed, *approximating* the recursive density estimation problem via *simulations* [Del96; LC98; Del98; IB98]. In doing so, they also support discrete and discontinuous problems, like a wall abruptly blocking all movements. However, they either come at the cost of reduced accuracy, or require significantly more computations, as the approximation's quality depends on the number of simulations [CGM07]. Nevertheless, with the steady increase in computational power, they became viable even for use on embedded devices, such as smartphones.