UNIVERSITAT JAUME I
DPTO. DE LENGUAJES Y SISTEMAS INFORMÁTICOS

Improving Data Fusion in User Positioning Systems

Doctoral Thesis

Presented by Edith PULIDO HERRERA
Directors Dr. D. Ricardo QUIRÓS BAUSET
and Dr. D. Germán FABREGAT

Castellón, September 2009

Improving Data Fusion in User Positioning Systems

Edith PULIDO HERRERA

ENGLISH VERSION
Trabajo realizado bajo la dirección de
Dr. D. Ricardo QUIRÓS BAUSET
y el Dr. D. Germán FABREGAT
y presentado en la Universitat Jaume I
para optar al grado de Doctor por la Universitat Jaume I

Castellón, September 2009

Este trabajo ha sido parcialmente realizado dentro de los proyectos TECNOLOGÍAS PARA LA INTERACCIÓN VIRTUAL CON CONJUNTOS DE OBJETOS APILADOS DE ALTO CARDINAL (TIC2002-04166-C03-01) e INTERFACES AVANZADAS PARA CAMPOS DE LUZ AUMENTADOS (ALF) (TIN2005-08863-C03-01), subvencionados por la Comisión Interministerial de Ciencia y Tecnología.

Summary

A FAULT detection and correction methodology for user positioning systems based on Kalman filtering is presented in this thesis. To develop this methodology, various aspects of the design of positioning systems are taken into account, such as the dynamic models of the systems, the definition of the estimation technique, the available technology and the environment (to identify disturbances that can affect the systems).

In the first phase, the available technologies were identified, and the algorithms and the physical-mathematical models were defined for the positioning systems for both indoor and outdoor environments. The Dead Reckoning (DR) algorithm was included for this reason, because it can be applied in both environments. This algorithm allows the pedestrian's position to be obtained while s/he is walking. Given that one of the main parameters of DR is the pedestrian's step length, a careful analysis was carried out to determine it. Algorithms that allow, first, detection of a step and, second, calculation of its length are therefore presented here.

An integrated system based on UWB and inertial technologies (IMU) is proposed for indoor environments. This system uses the information about the step length to improve the information provided by the UWB system. The system that was defined for outdoor environments is a GPS–IMU system based on the DR algorithm. Data fusion is carried out by means of Kalman filtering for both systems. In the GPS–IMU–DR system, the errors of the azimuth bias and the step length are obtained by means of Kalman filtering, which allows the DR parameters to be corrected and, consequently, the pedestrian's position can be obtained with greater accuracy.

In the next phase, the fault detection and correction methodology is developed. This is based on the principles of causal diagnosis using the theory of possibility. This methodology is proposed in order to prevent the introduction of erroneous information into the Kalman filter. In order to carry this out, failure states of the sensor systems are defined and corrective measures are applied when one or more of those states are present. These states are defined taking into account the empirical knowledge of the behaviour of the system. The performance of the filters was also monitored. This consists in the continuous evaluation of their innovations and, in case of inconsistency, corrective measures are applied to the parameters of the filters.

The experiments, for the proposed systems, presented results that improved their initial response to a considerable degree. For the UWB–IMU system, an analysis that involves the detection of inconsistencies of the filter is presented, while for the GPS–IMU–DR system the fault detection and correction methodology is applied. In the UWB–IMU system, the corrective measures introduced into to the Kalman filter allowed both consistent filtering and a soft signal to be obtained, i.e. most of the reflections of the signal provided by the UWB system were eliminated. The results of the GPS–IMU–DR system indicated that by implementing the methodology developed here, consistent filtering was obtained and the values of the azimuth and the step length were corrected properly. As a result more exact pedestrian trajectories were obtained.

Keywords: pedestrian positioning systems, fault detection, Kalman filtering, causal diagnosis, chi-square test, step length, azimuth bias, UWB, IMU.

Acknowledgments

DURING this process, many people accompanied me to whom I would like to thank. I want to thank my mother Rosa for being so unconditional, for her love, for her continued support, because, thanks to her advice I fill of strength and for everything that I cannot express in a few lines. To my father Felipe, who is far, but I know that he has always been with us and that his great spiritual strength has accompanied me at all times, especially in the most difficult. To my sister and brother, Astrid and Felipe for always be there, for trying to teach me to see the things with more humor and be much more than a good friends. Thanks to you, my family, for believing in me and despite the distance, you were supporting me and willing to make sacrifices to help me to achieve this personal goal.

I thank my tutor professor Ricardo Quirós for giving me this wonderful opportunity and for his support to carry out the research. Also to thank my second tutor professor Germán Fabregat for his valuable advice and support in the different phases of the research.

I would like to thank Hannes Kaufmann for his cordiality and continued willingness to give me both technical and moral support, which were factors that led me to carry out and to complete this research in a more enjoyable way. I want to thank professor Christian Breiteneder and the guys of the Virtual Reality group of the TUWIEN. To my colleagues of office, Michael and Florian who always willing to resolve my doubts, I thank also Michael for his willingness to make more pleasant my stay in Vienna, Mathis for his kindness and his efforts devoted to teach me everything related to the UWB system and be so available for instructing me with the OpenTracker and Tomas for his help and kindness. In general, to the people of the IMS for offering their advice and helping me in different situations during my stay in Vienna.

To thank professor Axel Pinz for accepted me in his group and gave me all the facilities to work more comfortable and with freedom in Graz. To Miguel Ribo for his great patience, introduce me in the research world, teach me his designs and discuss with me new ideas. Also to the people in the EMT to make more pleasant my stay in Graz.

To professor Miguel Angel Perez for giving his support and friendship, and for his advice in particular in the writing of the thesis.

On the other hand, I would like to thank Manolo for his friendship and understanding, and to his family, for gave me a second home away from mine, making me feel like at home. Thanks to my great friend Janneth for helping me to find new ways, introduce me in the world of fuzzy logic and be a great adviser on the issue of Causal Diagnosis. I want to thank Justo and my friend Aldo because their support was decisive for my stay in Spain and for having placed their trust in me. I would like to thank all those friends (Sepp, Julieta, Miguel, ...) and family (Gilberto, Jaime, Josefa, Alicia, Margarita...) from Colombia and other places that offered their support in one way or another.

I want also to thank the members of the graphics group of the "Universitat Jaume I", Miguel, Mike, Carlos ... for giving a pleasant atmosphere to work and advise in the different phases of the doctorate, all the people of the DLSI department that collaborated me, and the translation team of the university.

To all reviewers who carried out the evaluation of my thesis, thanks for their opinions, recommendations and contributions.

Finally, to God despite the fact that I do not think have the appropriate words to express my infinite gratitude.

Contents

List of Figures

List of Tables

List of Abbreviation

CD	Causal Diagnosis
DR	Dead Reckoning
KF	Kalman Filter
EKF	Extended Kalman Filter
DFMS	Data Fusion of Multiple Sensors
GPS	Global Positioning System
HDOP	Horizontal Dilution of Precision
HGPS	High-sensitivity GPS
HMD	Helmet Mounted Display
NIS	Normalized Innovations Square
INS	Inertial Navigation System
FL	Fuzzy Logic
PVT	Position, velocity and time
FLS	Fuzzy Logic System
UPS	User Positioning System
UMTS	Universal Mobile Telecommunication System
IMU	Inertial Measurement Unit
UWB	Ultra Wideband

Nomenclature

β	Time constant of the Gauss-Markov process
δ	Scale factor of the weighted filter
ε_B	Bias error
ε_s	Step length error
ι	Normalized Square Innovation
$\mathscr{O}$	Set of observations
Φ	Transition matrix
π	Distribution possibility
ψ	Azimuth
ρ	Autocorrelation function
ϑ	Kalman filter innovation
A	Set of attributes
B_{SENSOR}	Sensor bias
$cons$	Index of consistency
H	Observation matrix
Hx	Campo magnético horizontal
Hy	Campo magnético vertical
P	Estimation error covariance
PE	East coordinate

PN	North coordinate
rel	Index of relevance
s	Step length
S	Innovation covariance
U	Universe of discourse
V	Velocity

CHAPTER 1

Introduction

LOCALIZATION systems perform tasks that are essential for applications related to mobile systems and such applications have changed to match the progress made by the technologies involved in such systems. About two decades ago, localization systems were used exclusively in robust structures such as airplanes, ships, vehicles, submarines, and so on. This was due to the fact that only large sensors existed. In the 80's a new type of sensor based on electromechanical technology (MEMS, Micro-Electro-Mechanical Systems) emerged and it became possible to develop miniature sensors such as accelerometers and gyroscopes among others. Consequently, the range of applications for localization systems has increased greatly since then. Given this scenario, the idea of developing systems for locating people in different environments arose and this is the area that this thesis will deal with.

The capability to locate people can be applied to a countless number of situations. Such applications include the Navigation Systems for Users (NSUs), which can ultimately be used for military purposes (e.g. location of the enemy or a soldier on the battlefield), rescue work (e.g. location of a firefighter inside a building) or tourism (e.g. a tourist guide for an area of interest). We can also apply them to security control situations in companies, for example, for detecting the presence of unauthorized users in certain areas, control of prisoners, location of personnel within buildings. Finally, the localization systems are also essential in virtual reality, augmented reality or in entertainment applications, such as video games.

1.1 Background

In general a localization system is an *integrated navigation system* composed of several subsystems that must perform the following tasks: *plan, record and control the* *movement of an object (user, robot, vehicle, etc.) from one point to another [16].* One of these subsystems is defined *the positioning system*.

The *positioning system* is established as the one responsible for providing information about *the object's position and orientation*. Since localization systems are strongly dependent on this information, the positioning system should be designed taking into account quality criteria such as *robustness, reliability, accuracy and autonomy, among others* [16].

With the aim to design and implement positioning systems with better quality characteristics, researchers have chosen to use different technologies and/or apply different methods that involve applying statistics and the control theory or the conceptual designs among others. In robotics and the vehicle location these methods have been intensively developed, while they have been been less commonly employed in user location systems. For this reason, these methods are considered to be an important field of exploration for user location systems and *the work that is presented here focuses on this line of research*.

The basic factors needed to define a positioning system, include the identification of the environment where the system is to work, and consequently the technologies that are available for use in that environment. As a result, systems comprising different technologies have been developed —these are known as *Hybrid Systems*— in order to achieve better quality parameters. For outdoor environments (parks, forests, roads, etc.) the most common navigation systems are made up of the global positioning system (GPS) combined with inertial or magnetic technology or other technologies and algorithms to improve the functioning of the system. In fact, the patent filed by Levi and Judd [17] (U.S. Patent US5583776) for a method user location addresses the issue under the principles of hybrid systems and the dead reckoning algorithm [1], and this became a reference for research that has been, and is still conducted in this field.

Recently, research has begun exploring the possibilities of localization in indoor environments (buildings). The main research interests for user positioning are focused on UWB (Ultra Wideband) technology, HSGPS (High-sensitivity GPS) and algorithms to determine the position. As in the case of outdoor systems, using hybrid systems is considered to be appropriate for indoor systems.

Once the environment is identified and the technologies available are chosen which will provide the needed information to locate the user, it should establish the way such information will be processed. This is the line of research followed in this thesis, since this process will be used to manage the information, thus ensuring that the user's position is known at all times, regardless of the difficulties arising from the environment, the technology or any other factor that could disrupt the proper functioning of the positioning system. To contextualize this, let us think about the following example: when a user walks through an area that is full of buildings or a thick forest, as a stand-alone system the GPS may run into difficulties, while an integrated system would provide additional information allowing such issues to be avoided or reduce the errors.

[1] Determination of the position, knowing the initial position, the distance traveled and the orientation.

1.2 Objective

From what has been outlined above, the question that arises is: *how should the information from different sources or sensors be processed in order to improve the quality parameters of positioning systems for users?*. As in the classic positioning systems a number of common issues arise from this question and must be considered, such as: *What are the circumstances by which the sources of information fail to work properly? What advantages and disadvantages can a certain combination have in the design? and how can that information be treated in order to exploit the advantages, in the best possible way, the advantages of each technology and reduce the impact of their disadvantages?*. The first two questions have been widely discussed and certain combinations have been clearly identified with their respective advantages and disadvantages. The last question, however, is an open issue that can be addressed in different ways, since no clearly defined framework can be identified. Turning to other fields of engineering, such as engineering control, we can find methods that seek to optimize any system. This thesis focuses on the application of optimization methods for the particular case of Positioning Systems for Users. Hence, the main objective for this thesis is:

To carry out research on alternative methods that help to improve the processing of positioning information in hybrid systems for users, taking into account the statistical behavior of the information to be processed as well as the empirical knowledge observed regarding the behavior of the system.

1.3 Methodology

Several aspects must be covered in order to achieve the main objective:

1. Definition of physical-mathematical models that allow user positioning systems to be analyzed properly in different environments.

2. Definition of algorithms for processing the information provided by the sensors.

3. Application of methods for the evaluation-correction of the algorithms.

4. Experimental evaluation of the algorithms and the methods of evaluation - correction.

Magnetic and inertial sensors, a GPS receiver and a UWB-based localization system can be used in the experimental evaluation depending on whether they are to be employed in outdoor or indoor environments. Different theories such as Kalman filtering, possibility theory and causal diagnosis, will be used for the evaluation-correction methods and algorithms. This is because they help to evaluate the systems for defining correction mechanisms while taking into account both heuristic and formal knowledge of the models and behavior of the systems, which allows for a better functioning of the systems to be achieved.

1.4 Document Outline

This thesis is structured as follows:

Chapter 2 includes related work on tracking systems and navigation systems. We consider two fields of applications mixed reality (virtual and augmented reality) and user navigation systems. These fields are taken into account because they are both strongly dependent on positioning information. In fact, they use different technologies to achieve almost the same goal, e.g. in mixed reality they use vision systems but also inertial technologies as in navigation systems. These aspects are analyzed and a final comparison between the needs for each field is performed. Moreover, the main needs that users require of navigation systems are also included.

Chapter 3 includes the models and algorithms proposed in this thesis. First, there is a description of the sensors used and of the mathematical and physical models based on the information available from the sensors for both indoor and outdoor environments. This chapter also includes the evaluation mechanism based on formal knowledge (such as Kalman filtering framework) and based on empirical knowledge (such as fuzzy logic framework) with specific subjects like possibility theory and causal diagnosis. These are fully defined in this chapter for the pedestrian positioning systems that have been proposed.

Chapter 4 describes the experimental and analytical results of the systems proposed in Chapter 3 for both indoor and outdoor environments .

Chapter 5 includes the conclusions and future work.

CHAPTER 2

Related Work

HERE, AN ENGLISH version of this chapter it is not presented, it only includes its summary at the end of the chapter to show the work that has been carried out. The Spanish version is presented below.

Estado de la Técnica de los Sistemas de Posicionamiento de Usuarios

USANDO COMO base la revisión bibliográfica, en este capítulo se presenta una reseña de las tecnologías —consideradas en esta investigación como las más relevantes— para la localización a pequeña escala (i.e. objetos o humanos) (sección 2.1), incluyendo un resumen comparativo de sus principales características (sección 2.1.6). Tras analizar las tecnologías relevantes se describen varios sistemas híbridos de localización (sección 2.2) dentro de dos ámbitos de aplicación: la realidad mezclada y los sistemas de navegación para usuarios. Por último, el capítulo introduce brevemente las técnicas de optimización en los sistemas de localización (para robots, vehículos, etc.) y presenta un resumen de las principales características de los sistemas de posicionamiento híbridos según su aplicación final (sección 2.2.3).

2.1 Tecnologías para Sistemas de Posicionamiento

Basándose en un estudio llevado a cabo por Welch y Foxlin [3] (siendo esta una buena referencia, dada la reconocida experiencia de estos dos autores en el campo del seguimiento del movimiento del cuerpo humano) se establecen las condiciones que debería cumplir un sistema de posicionamiento, el cual está encargado de aportar la información de posición y orientación de un objeto o usuario. Las condiciones que proponen estos autores están enfocadas al ámbito de aplicación de la realidad virtual y la realidad mezclada, dado que este no es el único campo de aplicación de los sistemas de localización, es necesario realizar una generalización. Partiendo de las principales características propuestas por estos autores, la generalización que se realiza es la siguiente, teniendo presente que la aplicación es a pequeña escala:

- Debe ser ligero y compacto, de forma que sea completamente cómodo para el usuario. En otras palabras, debe ser un dispositivo que permita al usuario llevar a cabo su actividad de forma normal.
- Debe ser autónomo. Es decir, que no requiera de infraestructura adicional en el entorno para su adecuado funcionamiento.
- Debe proporcionar información de la posición con los grados de libertad requeridos por cada aplicación. Por ejemplo, varias tareas en realidad mezclada requieren de 6 grados de libertad, mientras que en los sistemas de navegación en muchas ocasiones es suficiente con tener 2 grados de libertad.
- Debe tener una exactitud que permita realizar la tarea que requiera el usuario de la mejor manera posible. Los rangos de exactitud varían considerablemente según la aplicación.
- Su buen funcionamiento debe ser independiente de la velocidad o de las distancias recorridas por el objetivo.
- Debe ser compatible con el ambiente, es decir, resistente a cualquier fenómeno de este, como campos electromagnéticos, sonidos, etc.
- Debe funcionar bajo cualquier entorno (interior o exterior).
- Inalámbrico. Es una característica que implicaría mayor movilidad, que es en última estancia para lo que se espera usar un sistema de localización.
- Económico. El coste es otro factor que se espera sufra grandes reducciones con el avance tecnológico, permitiendo el uso de estas tecnologías por parte del ciudadano común, especialmente en el caso de los sistemas de navegación para usuarios.

A pesar del gran desarrollo tecnológico, en la actualidad no existe un único dispositivo que cumpla con todas las características anteriores. Por lo tanto, la solución alternativa es la combinación de diferentes tipos de tecnologías, con el fin de obtener un sistema de posicionamiento más completo.

A continuación se explicarán las tecnologías de posicionamiento más importantes basándose en los estudios de Foxlin [18] y de Retscher y Kealy [19]. La primera de estas referencias es un completo análisis del seguimiento del movimiento enfocado a los campos de la realidad virtual y aumentada, mientras que la segunda se enfoca a los sistemas de navegación para usuarios. Dado el avance técnico en ambos campos y el avance de las tecnologías de localización, con el paso del tiempo estos podrían utilizar dispositivos de localización similares para alcanzar sus respectivos objetivos.

2.1.1 Técnicas Acústicas

Los dispositivos acústicos detectan las ondas de ultrasonidos para medir la posición y orientación. Requieren de una infraestructura compuesta por: (1) disposición de transmisores de ultrasonido, (2) un mínimo de tres receptores de ultrasonidos y (3) un dispositivo para activar los emisores de la señal. Para calcular la posición y orientación se utiliza el *tiempo de vuelo (TOF, Time of Fly)*[1] o *coherencia de fase*.

En el TOF, la velocidad del sonido se usa para calcular el tiempo que transcurre entre el envío de un pulso emitido por el transmisor y la recepción del mismo. De esta manera, es posible calcular la distancia entre el transmisor y el receptor. Al tener tres transmisores se aplica triangulación y se obtiene una estimación de la posición. Para obtener la orientación, se instalan tres receptores en el objetivo. En otras palabras, *para estimar la posición y la orientación en 3D se requieren tres transmisores y tres receptores*[13]. Todos los sistemas comerciales emplean esta técnica, porque elimina el problema del multicamino. Su principal desventaja es la baja velocidad de actualización por el retardo que se sufre en la recepción de las señales.

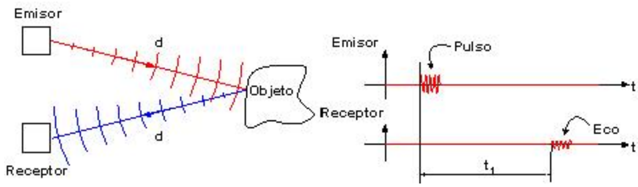
En la coherencia de fase, se calcula el ángulo de fase entre las ondas senoidales de las señales del transmisor y el receptor. Cuando la posición es constante se conoce la fase, pero cuando hay movimiento esta cambia. Actualmente este método no es empleado en los sistemas comerciales, debido a que con esta técnica, un sistema acústico es muy vulnerable a los efectos multi-camino producidos en ambientes acústicos ruidosos (i.e. ecos o reflexiones de la señal).

Los sistemas acústicos se aplican especialmente en entornos interiores, con un espacio de trabajo limitado, i.e. una habitación, ya que su infraestructura es compleja. Su ergonomía le permite al usuario libertad de movimiento dentro del volumen de trabajo, y facilita su integración con sistemas basados en otras tecnologías. Los principales inconvenientes de este tipo de sistemas son los siguientes:

- Aunque no son sensibles a los efectos electromagnéticos del entorno, sí a los fenómenos que afecten al sonido, como los ecos.
- Es una de la tecnologías que no funciona en ausencia de la linea directa de la señal entre el transmisor y el receptor.
- Finalmente, su exactitud se puede ver seriamente afectada, porque la velocidad del sonido es vulnerable a la temperatura, presión y humedad, lo que genera errores acumulables en el tiempo. Se han encontrado dos soluciones para este problema; la primera es habitual en los sistemas de realidad virtual y consiste en diseñar sistemas híbridos i.e. un sistema basado en visión y la técnica acústica.

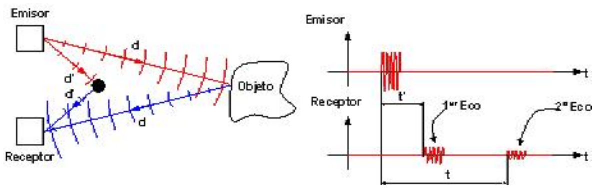
[1] Para estos sistemas sería más adecuado el término: *técnica impulso-eco.*

La segunda, se aplica en otros campos de la ingeniería (i.e. detección en submarinos (sonar-boyas), bancos de peces (sonar pesquero)) y consiste en controlar los errores mencionados realizando una corrección cada cierto tiempo, basándose en la detección de un objeto cuya distancia es conocida; estos principios son ilustrados en las figuras 2.1 y 2.2.



$$d = \frac{ct_1}{2}$$

Figure 2.1: Distancia de un objeto mediante la técnica impulso-eco. Tomado de [1].



$$d' = ct' \quad \Rightarrow \quad c = \frac{d'}{t'}, \quad d = ct \quad \Rightarrow \quad d = d'\frac{t}{t'}$$

Figure 2.2: Aplicando corrección de errores en sistemas acústicos mediante la técnica impulso-eco. Tomado de [1].

2.1.2 Técnicas Ópticas

La tecnología óptica propone una gran variedad de técnicas para la estimación de la posición y el seguimiento de objetos, pero algunas de ellas no son aplicables a la localización de usuarios. En esta sección se presentarán unicamente aquellas técnicas aplicables al problema planteado.

Los sistemas ópticos tienen como principio trabajar con la luz emitida o reflejada, mediante fuentes y detectores de luz. Las fuentes de luz se clasifican en pasivas y activas. Las fuentes activas emiten luz generada, como los diodos de luz infrarroja

(IR-LEDs) o los láser. Las fuentes pasivas son elementos que reflejan la luz, como los materiales retroflectivos.

Los sistemas ópticos utilizan sensores de imagen o de luz. Los sensores de imagen son cámaras CCDs [2], CIDs[3] y CMOS. Tienen la ventaja de determinar la posición de varios objetos en la misma imagen con exactitud, siempre que el objeto a localizar se distinga claramente en la imagen. Los sensores de luz son los fotosensores, fotocélulas, LDRs (fotoresistencias) y PMTs[4]. Utilizan fuentes de luz activas para garantizar una línea directa entre la fuente y el sensor.

Existen dos alternativas para instalar e implementar un sistema óptico. Estas son: *de fuera hacia dentro (outside-in)* y *de dentro hacia fuera (inside-out)*. La primera consiste en instalar una o varias cámaras en el techo o en la pared del espacio de trabajo y las estructuras de las fuentes de luz en el objeto a localizar; los sensores detectan la orientación de las estructuras de las fuentes de luz, y posteriormente se triangula con respecto a los ángulos de orientación de las cámaras más cercanas. En la segunda alternativa (*inside - out*) se instalan los sensores en el objeto a localizar, por ejemplo una cámara en un casco para realidad virtual (HMD) y se calcula la posición y orientación, con 6 grados de libertad aplicando técnicas basadas en la detección de marcas fijas en la escena [20].

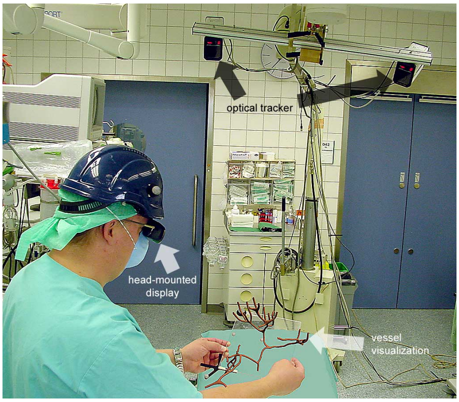


Figure 2.3: Sistema óptico empleado para ayuda quirúrgica con técnicas de realidad mezclada.

La tecnología óptica se utiliza habitualmente en entornos interiores; dentro de sus principales aplicaciones se encuentran los sistemas de realidad mezclada. Su principal ventaja es la precisión en la estimación de posición y orientación (en el rango de milímetros en posición y de décimas de grados en orientación). Estas características los hacen adecuados para sistemas de realidad mezclada en entornos interiores de dimensiones reducidas. No obstante, su aplicación en los sistemas de navegación es

[2] (charge-coupled device), dispositivo de cargas eléctricas interconectadas
[3] (charge-injection-device)
[4] (photomultipliers)

prácticamente nula, debido a las siguientes limitaciones:

- Un sistema óptico requiere de una infraestructura compleja que limita el espacio de trabajo. En los sistemas de navegación, el usuario debe poder recorrer grandes espacios no controlados, por lo que esta limitación recomienda evitar el uso de esta tecnología.

- No se ven afectados por ruidos del entorno como los electromagnéticos y los ecos, entre otros, no obstante, algunas sistemas ópticos si requieren ciertas condiciones de luz para su buen funcionamiento. También es imprescindible la linea directa de la señal entre el sensor y la fuente.

2.1.3 Técnicas Magnéticas

Esta tecnología será tratada en mayor profundidad, puesto que parte de los sensores empleados en las pruebas experimentales están diseñados bajo sus principios.

Dentro de esta tecnología se tiene la posibilidad de obtener la posición y orientación mediante sistemas avanzados que generan campos magnéticos, y también la posibilidad de usar dispositivos como brújulas electrónicas para obtener la orientación. La primera opción tiene una gama de aplicaciones entre las que se encuentran los sistemas de realidad mezclada, la medicina o los deportes, entre otras. En cuanto a la segunda opción, se considera una de las más populares en los sistemas de navegación. Aquí se hará énfasis en esta última; en [18], se encuentran en detalle los principios de la primera opción.

El campo magnético de la tierra forma parte de las primeras herramientas de la naturaleza usadas por el hombre para orientarse. Es modelado como un dipolo, cuyo norte difiere del norte geográfico o verdadero en un valor aproximado de 11.5°. Esto hace que exista una desviación angular entre los meridianos magnético y geográfico denominada *declinación magnética* D cuyo valor es variable en el tiempo y depende de la localización geográfica. Otro característica, es el ángulo que existe entre el campo magnético y la superficie de la tierra, el cual es denominado con su término en ingles “dip” o ángulo de inclinación I.

En teoría, el conocimiento de los componentes del campo magnético horizontal y vertical (Hy, Hx), bastaría para determinar el curso del movimiento —conocido como *Acimut* — aplicando la siguiente ecuación:

$$acimut = \arctan(Hy/Hx) \tag{2.1}$$

Existen algunas consideraciones que no permiten aplicar directamente la ecuación anterior, como que los componentes del campo medidos deben estar paralelos a la superficie terrestre, aspecto que en la práctica no se cumple. Por lo tanto, se aplica un método para eliminar los errores de inclinación que pueden ser importantes dependiendo del valor de I [21]. Este se lleva cabo obteniendo los valores de los ángulos *pitch* y *roll*, los cuales representan las inclinaciones del objeto. Los ángulos *pitch*, *roll* y *Acimut*, están definidos como los *ángulos de Euler*; estos ángulos están ilustrados en la figura 2.4. Para aplicar la corrección respectiva, además de conocer los valores de roll y pitch, es necesario conocer el valor del componente vertical del campo magnético Hz.

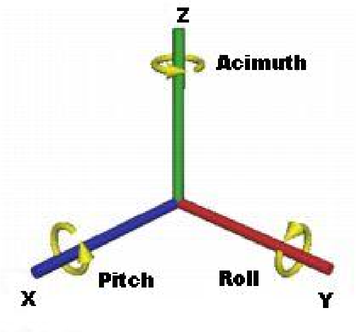


Figure 2.4: Ángulos Roll, Picth y Acimut.

Teniendo la información anteriormente mencionada, es posible trasladar los componentes del campo magnético al plano horizontal y obtener un valor corregido del acimut. Para ello se aplican las siguientes ecuaciones, con las que se obtienen los valores Hx_H y Hy_H, los cuales son Hx y Hy trasladados al plano horizontal.

$$Hx_H = Hxcos(a) + Hysin(a)sin(b) - Hzcos(b)sin(a) \tag{2.2}$$

$$Hy_H = Hycos(b) + Hzsin(b) \tag{2.3}$$

donde: Hx_H es el campo en X trasladado al plano horizontal, Hy_H es el campo en Y trasladado al plano horizontal, a es el ángulo de inclinación en el plano del pitch, b es el ángulo de inclinación en el plano del roll.

El acimut corregido se obtiene reemplazando Hx y Hy en la ecuación 2.1 por Hx_H y Hy_H, respectivamente, y restando el valor correspondiente de D.

En la práctica los sensores encargados de proporcionar la información de acimut, son las brújulas electrónicas; están compuestas por magnetómetros e inclinómetros. Lo magnetómetros obtienen los componentes del campo magnético de la tierra (H_X, H_Y, H_Z) y los inclinómetros proporcionan los ángulos de inclinación roll y pitch. Toda esa información va a procesadores (microcontroladores, DSP, etc.) que realizan las correcciones respectivas.

Un aspecto muy importante es que son sensores muy sensibles a los efectos magnéticos del entorno como los materiales férreos. Por ello, es necesario aplicar métodos que corrijan los errores introducidos en las lecturas del sensor. Las brújulas actuales incluyen métodos de calibración de fábrica, los cuales están clasificados en dos tipos: *hardiron* y *softiron*. Hardiron es el término empleado para identificar el fenómeno de las interferencias producidas por aquellos elementos que se encuentran en los alrededores de la brújula y que se comportan como un imán, teniendo como consecuencia que el campo que registrará el sensor es el campo magnético de la tierra con el campo magnético de dicho material superpuesto. Softiron se atribuye a las anomalías magnéticas cuando hay materiales ferromagnéticos en los alrededores del sensor, distorsionando el campo magnético de la tierra [22], [23]. La corrección en cualquiera de los casos es llevada a cabo "enseñándole" al sensor el entorno que le rodea. Esto es efectivo, sin embargo, no garantiza una completa corrección de los datos si el entorno varia demasiado, especialmente en el caso de la calibración softiron.

Normalmente la brújulas electrónicas son dispositivos pequeños de bajo coste, que pueden ser empleados tanto en entornos interiores como en entornos exteriores. No requieren infraestructura, puesto que, su única fuente son los campos magnéticos, las cuales no se ven afectadas por bloqueos, como sí es el caso de las anteriores tecnologías. Como ya se mencionó, son muy sensibles a los efectos magnéticos del entorno, siendo esta su gran desventaja. Se necesitan entornos libres de posibles elementos ruidosos para obtener todos sus beneficios; en caso contrario, se requerirán procesos de calibración y un procesamiento cuidadoso de la señal.

2.1.4 Técnicas Mecánicas

Es la tecnología más antigua utilizada para la localización o el seguimiento de partes del cuerpo de un usuario. Se dará una breve explicación de esta tecnología de acuerdo a dos principios físicos, en los que están basados los diseños actuales: —la cinemática y la dinámica —.

2.1.4.1 Cinemática

Dentro de esta tecnología se encuentran los sistemas biomecánicos en los que los menos sofisticados son estructuras articuladas colocadas sobre el objeto a localizar. Dichas estructuras están compuestas de elementos, que son conectados de tal manera, que es posible determinar la posición entre ellos. Sistemas más sofisticados disponen de cuerpos rígidos y sistemas de sensores que determinan la orientación y posición relativa entre los cuerpos rígidos. Esto incluye el conocimiento de las dimensiones del cuerpo rígido, sensores que miden los ángulos entre las uniones de los cuerpos rígidos y con esta información se realizan los cálculos correspondientes.

Un ejemplo clásico de esta tecnología en la realidad virtual, es el HMD (Head Mounted Display) diseñado por Sutherland y otros en 1968 [2] (figura 2.5). Su trabajo es pionero en el diseño de sistemas para seguir el movimiento de la cabeza y ha inspirado la investigación en este campo. Un HMD es un casco que lleva instaladas cámaras y sensores de posición para la visualización de entornos virtuales o reales.

Los sistemas diseñados bajo este principio tienen un entorno de trabajo muy limitado, por la pobre ergonomía para facilitar las labores del usuario. Por ello, su gama de aplicaciones está restringida a algunas aplicaciones en realidad mixta, medicina, simuladores, vídeo juegos, robótica, etc. Por otro lado, sí la ergonomía no es prioritaria,

(a)

(b)

Figure 2.5: Prototipo de HMD desarrollo por Sutherland [2].

son sistemas apreciados por su buena velocidad de actualización y baja latencia.

2.1.4.2 Dinámica

Son sistemas cuyo principio radica en la segunda ley de Newton y se les define como *sistemas de navegación inercial* (INS, Inertial Navegation System).

Los INS fueron sistemas empleados en un principio para barcos, aviones o misiles desde los años 50. Esto cambió alrededor de la década de los 90, debido al desarrollo experimentado por la tecnología *MEMs (Micro-Electro-Mechanical Systems)*, siendo posible diseñar sensores de bajo coste y tamaño reducido con los mismos principios de los INS. Este hecho permitió extender su gama de aplicaciones a la localización de objetos de menor tamaño, como individuos en los sistemas de navegación o, partes del cuerpo humano en los sistemas de realidad mezclada.

Los INS están compuestos por acelerómetros y giróscopos. Con ellos se adquiere información para seguir el movimiento lineal y el movimiento angular del sistema físico.

El giróscopo es un sensor que funciona según el principio de conservación del momento angular y detecta la velocidad angular. Una vez la velocidad angular es detectada, se realiza una integración para obtener la orientación. Existen varios tipos de error en estos sensores, de los cuales el error de *deriva* es especialmente relevante debido a sus consecuencias. Es un error que crece ilimitadamente con el tiempo, incrementado el valor de la orientación verdadera. Este error puede ser aceptable para dos tipos de giróscopos. El primer tipo de giróscopo es aquel que es empleado en las grandes infraestructuras (barcos, aviones, etc.) puesto que tienen una buena exactitud y por las condiciones de la aplicación, es posible realizar correcciones que dan resultados satisfactorios. El segundo tipo son los giróscopos de fibra óptica con una exactitud muy buena y de tamaño reducido, lo que permite su uso en los sistemas de navegación para humanos, no obstante, su alto coste no los convierte en una opción viable en la actualidad para la mayoría de las aplicaciones.

Los acelerómetros son sensores que miden de forma indirecta la aceleración lineal que sufre un sistema. Estos tienen una masa de prueba y detectan la fuerza ejercida sobre ella, de tal manera que al aplicar las leyes de Newton, obtienen la aceleración. Para calcular la posición, la aceleración es integrada dos veces. Al igual que los giróscopos esta integración lleva a tener un error de deriva en la posición, que crece de forma cúbica con el tiempo.

Hay dos configuraciones de estos sensores para obtener la posición y la orientación en 3D, las cuales están ilustradas en la figura 2.6. En la primera configuración los sensores están montados sobre una plataforma, en la cual hay tres "gimbals"[5] con respecto a un marco de referencia. Esta configuración no es de interés para la localización de usuarios, puesto que es grande y pesada, por lo tanto, típicamente utilizada en las grandes infraestructuras.

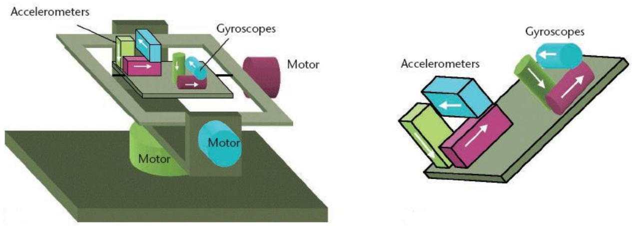


Figure 2.6: Concepto de los sensores inerciales. tomado de [3]

.

Por otro lado, la segunda configuración se denomina *Strapdown*[6], en la cual los sensores son adaptados de forma rígida al marco del objeto. Esta es la configuración típica de los sistemas actuales empleados en la localización de usuarios. Para obtener datos de posición y orientación en 3D, se instalan 3 acelerómetros y 3 giróscopos en tres ejes ortogonales, esto es, un acelerómetro y un giróscopo por cada eje. Las bases de algoritmo se ilustran en la figura 2.7.

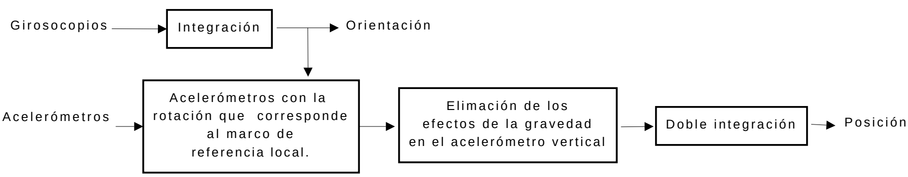


Figure 2.7: Bases de los sistemas strapdown. Adaptado de [3]

.

[5] El significado de "gimbal" tomado de Wikipedia es: un soporte pivotado que permite la rotación de un objeto alrededor de un eje.

[6] Se usa su término en ingles, puesto que no existe ninguna traducción al español

Los giróscopos y los acelerómetros son sensores con ventajas muy interesantes debido a que son *autónomos*, en otras palabras, al contrario de otras tecnologías no dependen de una fuente de la señal. Una consecuencia de ello es que su volumen de trabajo sea ilimitado, además no tienen problemas con posibles ruidos del entorno, lo cual los hace extremadamente ventajosos sobre otras tecnologías. Adicionalmente, en la actualidad se disponen de unidades muy pequeñas que permiten su adecuada integración con otras tecnologías. Otras de sus características son su alta velocidad de actualización y baja latencia.

Dado que los INS presentan limitaciones debido a los errores expuestos anteriormente, no son adecuados para usar como único sistema de posicionamiento, pero sí es posible realizar combinaciones con otras tecnologías de acuerdo a los requisitos de la aplicación.

En resumen, la tecnología basada en sistemas inerciales es considerada como una de las tecnologías que más promete y por ello, en sistemas híbridos, es una de las tecnologías base para obtener la posición y la orientación. Sin embargo, mientras los problemas de deriva no tengan una solución para los sensores de bajo coste, no pueden ser utilizados de forma única ni en los sistemas de posicionamiento de usuarios, ni en los sistemas de realidad mezclada.

2.1.5 Técnicas de Radio Frecuencia

Hay una gran diversidad de técnicas dentro de esta tecnología utilizadas para la localización de usuarios, pero pocas de ellas son adecuadas (por el momento) para la localización de personas con ciertos niveles de exactitud.

Aprovechando la infraestructura existente se localizan dispositivos dentro de las redes de telefonía móvil — GMS, UMTS, W-CDMA—. A pesar de su utilidad no son considerados adecuados para muchas de las aplicaciones de localización de usuarios, ya que su rango de exactitud es de 50 a 100 metros.

Otra sistemas de redes en los que se están llevando a cabo muchos esfuerzos de investigación, son los sistemas de redes inalámbricas WLAN. Al igual que en las redes de telefonía móvil, se pretende aprovechar la infraestructura ya existente para la localización de dispositivos.

El sistema de posicionamiento global basado en satélites GPS es el más popular y, en muchos casos, imprescindible para los sistemas de navegación, debido a que aporta información de la posición absoluta. Está compuesto por una constelación de 24 satélites que transmiten señales para determinar la posición, la velocidad y el tiempo (PVT) de los receptores. Un receptor debe detectar al menos 4 satélites para calcular dicha información y la exactitud alcanzada puede estar en el rango de metros. Para conseguir una mejor exactitud existe un sistema denominado DGPS (GPS diferencial), el cual a través de una red de balizas con posiciones conocidas a lo largo de la tierra, corrige las posiciones enviadas por los satélites.

El GPS es el sistema basado en satélites más común, sin embargo, otros países a parte de Estados Unidos han desarrollado o están desarrollando sistemas similares, como GLONASS (Rusia), Galileo (Europa) o COMPASS (China). Ya están comenzando a aparecer investigaciones en el uso de estos sistemas para el posicionamiento de usuarios [24].

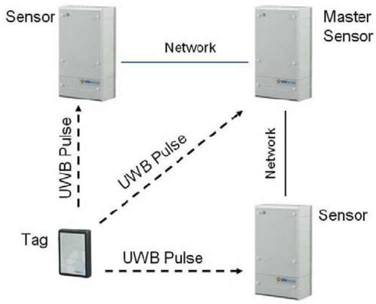


Figure 2.8: Sistema comercial basado en UWB, de la compañía Ubisense

.

Dado que el GPS es el sistema más explorado y conocido, sus desventajas están claramente identificadas. Estas pueden ser tratadas aplicando correcciones al hardware o software y esto se realiza en función de las restricciones de la aplicación. Por ejemplo, la exactitud exigida para la localización de un vehículo puede ser mucho menor que la exigida para localizar un individuo bajo labores de rescate, lo cual implicaría una aumentación de la información del GPS mediante la utilización de sensores adicionales. Dichas desventajas pueden ser la baja exactitud, el efecto multicamino y la falta de información debida a la atenuación u obstrucción de la señal de los satélites. Recientemente ha sido diseñado el *HSGPS (High Sensitivity GPS)*, con el fin de proporcionar información PVT en entornos interiores. En teoría debe funcionar cubriendo el mínimo de características técnicas que cubre el GPS, pero según las investigaciones de Mezentsev y otros [25], es un sistema que requiere del respaldo de otras tecnologías para trabajar, puesto que de manera independiente no es suficientemente fiable.

En los últimos años han surgido los sistemas basados en *Ultra-Wideband (UWB)*, para la localización en entornos interiores cubriendo un gran ancho de banda que llega hasta los GHz. Los investigadores tienen grandes expectativas alrededor de esta tecnología debido a la sencillez del hardware, las condiciones de ergonomía y aún más importante parece que no presenta el problema de multicamino, alcanzado unos rangos de exactitud aceptables para varias aplicaciones. Un ejemplo de esta tecnología se muestra en la figura 2.8, es un sistema comercial diseñado para localización de personas dentro de edificios y cuyos componentes son los que se indican en la figura. Las balizas son instaladas en posiciones conocidas, el usuario lleva un sensor consigo y mediante técnicas de triangulación el sistema calcula la posición y la trasmite a través de la red.

2.1.6 Comparación de las Tecnologías de Posicionamiento

En este apartado se presenta una comparación de las tecnologías teniendo en cuenta varios criterios importantes para la localización y resumida en la tabla 2.1.

La tabla ha sido realizada tomando como base la tabla de [16] para sistemas y técnicas para navegación de usuarios, pero ha sido adaptada a las tecnologías analizadas a lo largo de este capítulo.

Los criterios incluidos no son generalizados ya que no todas las tecnologías son útiles para las mismas aplicaciones. Por ejemplo, un sistema óptico puede ser muy útil en una sala de entrenamiento de cirugía, pero no como sistema de localización de un doctor dentro del hospital, en ese caso son útiles los sistemas de navegación para usuarios; en consecuencia la valoración se realiza según su entorno de aplicación.

Estos son algunos de los factores tenidos en cuenta para valorar las tecnologías:

- La exactitud varía según la tecnología y solo aplica para los campos en los que son útiles. Es decir, que se considera con exactitud baja o alta según sus espacios de trabajo. Por ejemplo, el GPS se considera aceptable porque localiza una persona dentro de rangos de 1-10 m dado que su entorno de trabajo se define a lo largo del globo terráqueo; mientras que los sistemas GMS o UMTS se consideran de exactitud baja ya que su rango no es menor a 50 m.

- La autonomía no fue valorada para varias de las tecnologías, porque a pesar de que un usuario disponga de dispositivos fáciles de llevar, se puede estar haciendo uso de una gran infraestructura, i.e. el GPS o el sistema UWB.

- La cobertura está asociada a la cantidad de espacio en la que es capaz de trabajar una tecnología. Por ello la cobertura del GPS se evalúa como alta, mientras que para los sistemas acústicos u ópticos se considera baja, ya que se utilizan normalmente en ambientes reducidos y cuidadosamente acondicionados para su uso.

En la tabla se diferencian tres grupos clasificados según el entorno de aplicación: (1) para entornos exteriores, (2) para ambos entornos: interiores, exteriores y (3) para entornos interiores. De esta manera se puede distinguir el ámbito de la valoración para cada una de ellas.

Como se observa, no hay una tecnología que cumpla con todos los criterios independiente del entorno, en otras palabras, no se puede garantizar información de posición bajo cualquier circunstancia basándose en una única tecnología.

2.2 Sistemas Híbridos

Como consecuencia del análisis de la sección anterior, la solución adoptada por gran parte de los diseñadores e investigadores es combinar tecnologías de forma complementaria para conformar *sistemas híbridos*. Es decir, se demanda la combinación de

Table 2.1: Comparación de tecnologías posicionamiento de pequeña escala

Sistema	Exactitud	Ergonomía	Autonomía	Cobertura
GPS	Aceptable	Alta	-	Alta
Galileo	Aceptable	Ata	-	Alta
HSGPS	Baja	Alta	-	Alta
GSM	Muy baja	Alta	-	Baja
UMTS	Muy baja	Alta	-	Baja
INS - cortos períodos de tiempo	Alta	Alta	Alta	Alta
INS - largos períodos de tiempo	Muy baja	Alta	Alta	Alta
Magnética	Alta	Baja	Muy baja	Baja
Óptica	Alta	Muy baja	Muy baja	Muy baja
Acústica	Alta	Baja	Muy baja	Muy baja
UWB	Aceptable	Alta	-	Alta

varias tecnologías (i.e. magnética y óptica o inercial y magnética) para obtener la información completa de la posición y la orientación de un sujeto u objeto. Esto tiene como objetivo incrementar las fortalezas individuales de cada tecnología y a su vez atenuar el efecto de sus debilidades.

En los siguientes apartados se describirán los principales prototipos o teorías desarrolladas bajo el principio de la combinación complementaria de diferentes tecnologías. Como ya se mencionó, han sido clasificados según dos campos de desarrollo, la realidad mezclada y los *Sistemas de Posicionamiento de Usuarios* (SPUs). Se puede decir que ambos campos tienen un objetivo en común, *cómo usar de manera conveniente el conocimiento de la posición dentro de una gran gama de aplicaciones i.e. labores de rescate, turismo o militar*.

2.2.1 Sistemas de Posicionamiento en la Realidad Mezclada

La realidad mezclada es uno de los campos de investigación con mayor perspectiva y considerada una de las tecnologías del futuro, por ello actualmente se están realizando grandes esfuerzos de investigación en este campo.

Más que sistemas de posicionamiento en un sistema de realidad mezclada se tienen sistemas de seguimiento del movimiento. En otras palabras, se necesita seguir el movimiento de un objeto (i.e. la cabeza o la mano de un usuario) de tal manera que las tareas puedan ser llevadas a cabo con éxito.

Previamente a mencionar varios de los trabajos más relevantes en este campo de investigación, se presenta el trabajo realizado por Neumann y Cho [4] con el fin de ilustrar la funcionalidad de la realidad mezclada. En este trabajo no se incluye el concepto de los sistemas fusionados como tal. Ellos presentan un sistema no híbrido completamente basado en visión; definen un método que se fundamenta en la siguiente información: (1) características de la posición, desde imágenes previas, (2) color, (3) discriminación de triángulos y círculos y (4) lista de posiciones de características previamente conocidas. En la figura 2.9 el vaso es anotado con un cuadrado y cuya posición se define mediante marcadores colocados de forma triangular. Este ejemplo ilustra como en un sistema de realidad mezclada lo que se busca es la interacción de objetos virtuales con el mundo real. A partir de este también se pueden señalar los principales problemas que tiene este campo como lo es la superposición de las imágenes virtuales sobre el mundo real de forma exacta, la visualización de las imágenes o la frecuencia de actualización de la información de posición.

Figure 2.9: Ejemplo de los resultados presentados por Neuman y Cho[4].

Este campo es especialmente dramático en este sentido comparado con otras aplicaciones, ya que grandes retardos o errores en la información de posición pueden ser detectados con facilidad por los usuarios. En otras palabras, un usuario inmerso en un entorno aumentado puede visualizar si las imágenes virtuales están mal posicionadas en el mundo real. Varios de estos problemas pueden ser solucionados en entornos controlados, pero en los entornos no controlados —como los entornos exteriores — los investigadores consideran que aún se está lejos de obtener resultados satisfactorios debido a la carencia del hardware adecuado [3].

En la sección 2.1.4, se presentó una breve descripción del trabajo de Sutherland. Su trabajo ha marcado un antes y un después en la realidad virtual, puesto que fue la primera aplicación publicada basada en la utilización de los conceptos de localización a pequeña escala, en su caso particular seguir el movimiento de la cabeza de un usuario (ver la figura 2.5). Sutherland es pionero en el inicio de la implementación de hardware en el seguimiento del movimiento de la cabeza de un usuario. Su trabajo fue muy interesante, ya que él y su grupo presentaron el uso del HMD, así como la introducción de un sistema acústico para seguir el movimiento de la cabeza del usuario, con el fin de llevar a cabo aplicaciones de realidad virtual. Pero, debido a las limitaciones tecnológicas de aquella época no se desarrollaron más trabajos hasta a mediados de los años 90 [26].

Uno de los prototipos más interesantes para aplicaciones de realidad aumentada fue publicado por Azuma y otros [5] a finales de los años 90, ya que presentaron un prototipo para trabajar en entornos exteriores. Como se observa en la figura 2.10, es un sistema híbrido conformado por un receptor GPS diferencial, una brújula electrónica, sensores de inclinación, giróscopos y un HMD. La fusión es llevada a cabo con un filtro de Kalman para sistemas lineales y su interés principal es reducir el problema del registro[7].

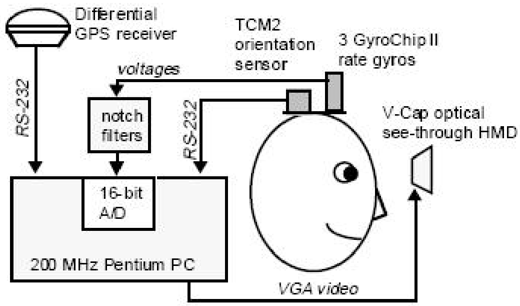


Figure 2.10: Configuración de los dispositivos utilizados en el prototipo presentado por Azuma [5].

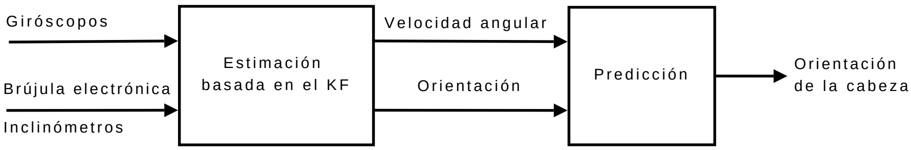


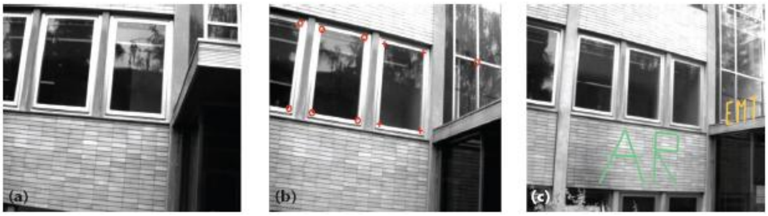
Figure 2.11: Fusión de datos y predicción del prototipo presentado por Azuma [5].

Dado que este fue un campo emergente a mediados de los años 90 también surgieron nuevos sistemas comerciales desarrollados a partir de investigaciones, como es el caso de la compañía Intersense. Ellos desarrollaron varios sistemas de sensores basados en la fusión de datos y la combinación de varias tecnologías. Uno de estos es un sistema que combina la tecnología inercial —un acelerómetro y un giróscopo — y la tecnología acústica —un conjunto de torres de transmisión de ultrasonidos y 3 receptores de ultrasonidos —. Ellos emplean la combinación con el fin de compensar los errores de deriva del sistema inercial y por otro lado apoyar el sistema acústico.

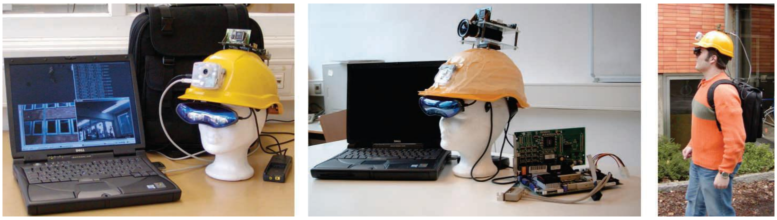
Con el paso de los años se han venido desarrollando diferentes sistemas híbridos, debido a que se ha comprobado la alta funcionalidad de la fusión de datos para incrementar el buen funcionamiento, exactitud y robustez de los sistemas. Uno de estos sistemas es el trabajo publicado por Ribo y otros [6]. Ellos presentan el diseño completo de un sistema híbrido para realidad aumentada, con su propio sistema inercial. Ellos afrontan el problema de la localización llevando a cabo la fusión de datos proporcionados por el sistema de visión y el sistema inercial mediante un filtro de Kalman

[7] Esto es, intentar mejorar la información visualizada por el usuario o permitir al usuario observar los objetos virtuales alineados con los objetos reales.

extendido (FKE). El sistema de visión trata de reconocer marcas naturales en el ambiente como esquinas, como se muestra en la figura 2.12.



(a)



(b)

Figure 2.12: Prototipo desarrollado por Ribo y otros [6].

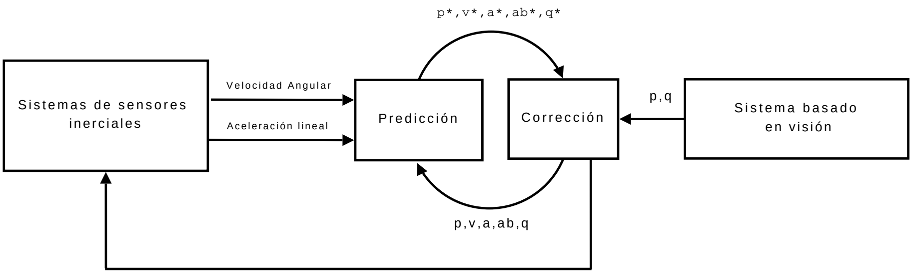


Figure 2.13: Fusión de datos en el prototipo desarrollado por Ribo y otros [6].

Posteriormente Foxlin y Naimark [7] desarrollaron otro sistema comercial híbrido basado en visión y tecnología inercial. En la figura 2.14 se observa que el usuario lleva los componentes en un salón donde existen marcadores que son detectados por el sistema de visión. Un sistema de visión tiende a desorientarse con giros o movimientos rápidos, en ese caso el sistema inercial es el sistema que determina la posición y la orientación, hasta que el sistema de visión se estabiliza de nuevo. De manera complementaria, el sistema de visión ayuda a corregir los errores de deriva del sistema inercial.

Desafortunadamente, uno de los mayores problemas en la actualidad es la localización en exteriores, se han llevado a cabo varias aproximaciones y hoy por hoy,

(a) (b)

Figure 2.14: Sistema VIS-Tracker de Foxlin&Naimark [7].

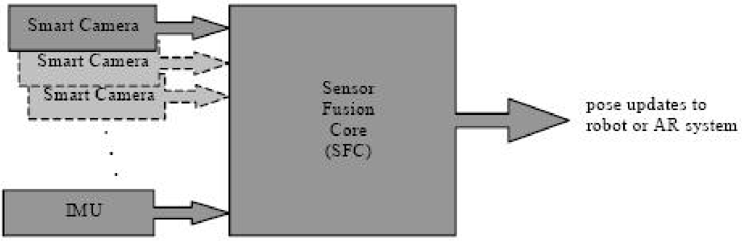


Figure 2.15: Arquitectura del sistema de Foxlin&Naimark [7].

varios grupos continúan buscando alternativas. Actualmente, se siguen desarrollando sistemas de realidad mezclada, pero más enfocados en la líneas de visualización y diseño de mundos virtuales entre otras, que no son de interés en este documento.

Como se observa en los diagramas 2.11, 2.13 y 2.15 los sistemas de realidad mezclada tienen dentro sus componentes básicos el sistema de visión y lo utilizan como parte de sus sistemas de localización. Los sistemas basados en visión funcionan correctamente siempre y cuando se disponga de referencias suficientes en el entorno, se tenga un sistema de gran capacidad computacional y el usuario no realice giros muy rápidos, ya que al sistema le cuesta recuperarse o referenciarse de nuevo bajo esta situación. Para solucionar el último inconveniente, la solución llevada a cabo ha sido la utilización de giróscopos o cualquier tipo de sensores que le permitan al sistema recuperar su orientación. Esto ilustra adecuadamente la combinación complementaria de diferentes tecnologías.

Por otro lado, y debido a la combinación, estos sistemas incluyen el concepto de fusión de datos, el cual llevan a cabo aplicando el filtrado de Kalman, debido a su reconocida eficiencia. En el estudio de [27] se realiza una comparación de diferentes técnicas de filtrado y predicción de la orientación para los sistemas de realidad virtual;

ellos definen diferentes parámetros que consideran importantes en la realidad virtual y los evalúan para cada uno de los métodos seleccionados para la comparación (FKE, filtro lineal invariante en el tiempo (LTI), el filtro de partículas (particle filter) y filtro de Kalman de lo desconocido (Unscented Kalman filter)) y concluyen que el FKE es un método adecuado y suficiente en los sistemas de realidad virtual.

Finalmente, cabe mencionar que la mejora del funcionamiento de los sistemas de realidad mezclada se debe en gran medida al hardware existente tanto para la parte de visualización como para la parte de localización. En otras palabras, a medida que los sensores actuales presenten menos errores y se vayan encontrando nuevas opciones menos sensibles a los cambios del entorno, las aplicaciones se incrementarán y el funcionamiento de los sistemas será mucho mejor.

2.2.2 Sistemas de Posicionamiento para Navegación u otras Aplicaciones

Los sistemas de navegación son menos restrictivos que los sistemas de realidad mezclada en lo que se refiere a niveles de exactitud. Su objetivo es detectar la posición del usuario de forma autónoma e independiente del entorno. En otras palabras, los SPUs no deberían depender de infraestructuras complejas y funcionar tanto en entornos interiores como en entornos exteriores. En este caso el término *infraestructuras complejas* hace referencia a que para el usuario debe ser un sistema sencillo de llevar, más que a la infraestructura que requiere la tecnología en si. Por ejemplo, ya se sabe que la infraestructura del GPS es compleja, pero el usuario final no nota este hecho, en cambio si lo notaría en un sistema óptico de realidad virtual ya que varios de los componentes deben ser colocados en su cuerpo.

Los estudios que se se presentan a continuación han sido clasificados según el tópico de desarrollo. En los trabajos reportados hasta la fecha han sido identificados los siguientes tópicos: prototipos y algoritmos que utilizan la dinámica del usuario (sección 2.2.2.1), prototipos autónomos (sección 2.2.2.2) y otras implementaciones (sección 2.2.2.3). Dentro de los prototipos y algoritmos han sido incluidos los trabajos que han utilizado los conceptos de sistemas híbridos, la fusión de datos o aquellas propuestas que proponen métodos para mejorar el funcionamiento de los SNUs; los prototipos autónomos son aquellos que se basan únicamente en la dinámica del usuario o en sistemas de navegación inercial; en otras implementaciones se incluyen algunos prototipos que combinan el procesamiento de los sistemas de posicionamiento de usuarios dentro de un sistema de realidad mezclada.

2.2.2.1 Algoritmos y Prototipos Integrando la Dinámica del Usuario

En un principio se exploraron las posibilidades para entornos exteriores, lo que llevó a la utilización del sensor más obvio para esta aplicación, como es el GPS. Por lo tanto, varios de los primeros SPUs están diseñados basándose en una combinación entre el GPS con otras tecnologías, para obtener mayor información o información redundante que incrementará la robustez de los sistemas. A nivel general, se plantea la utilización de sistemas híbridos con el fin de tener sistemas que aporten información en todo momento, es decir, sistemas que cumplan con esa función incluso si cualquiera de sus componentes tiene algún problema (i.e. cuando el GPS tiene problemas como los mencionados en la sección 2.1.5).

Partiendo de ese hecho, Levi y Judd [17] diseñaron un sistema compuesto por un GPS, acelerómetros, magnetómetros y un barómetro. En ese trabajo se pone en evidencia que la dinámica de un usuario puede ser usada para determinar su localización. Ellos demuestran que la aceleración lineal vertical presenta un patrón cuando el usuario camina. Utilizan este patrón para detectar la ocurrencia de los pasos y calculan la longitud del paso estableciendo una relación lineal entre la medida de la frecuencia del paso y el tamaño del mismo. La distancia recorrida se calcula de forma acumulativa a partir de las longitudes del paso. Con esto y el conocimiento del acimut, ellos aplicaron el algoritmo "Dead Reckoning" (DR)[8] (era el algoritmo empleado por los marineros para conocer la posición de sus barcos; en el capítulo 3 se hará su correspondiente descripción). En ese estudio se usan los principales conceptos con los que se han trabajado y se continua trabajando en los sistemas de localización de usuarios, de los cuales cabe destacar, la determinación del bias del acimut, la longitud del paso, la utilización del filtro de Kalman y la combinación de varias tecnologías.

A partir de esto, las investigaciones se han enfocado a encontrar diferentes alternativas para localizar personas planteando cuestiones como: ¿cómo mejorar los parámetros empleados en el algoritmo "Dead Reckoning"?, ¿cómo manejar la información redundante proporcionada por los diferentes sensores?, ¿cómo reducir los errores de las tecnologías?, ¿cómo mejorar el modelado de los sistemas? entre varias áreas que aún se están explorando o faltan por explorar en este campo. Basándose en estas cuestiones los trabajos que se citan a continuación tratan de resolver algunas de ellas.

Las investigaciones que se incluyen en este apartado tienen dentro de sus objetivos aplicar algoritmos que mejoren cualquiera de los componentes de los SPUs y no están presentadas en orden cronológico sino según sus objetivos. En general, la mayoría de prototipos desarrollados emplean la dinámica del usuario dentro de sus bases. En esta sección se incluirán aquellos cuyo objetivo principal es su uso en la navegación de usuarios, para otros fines son incluidos en la sección de otras implementaciones 2.2.2.3.

Con la intención de buscar una mejora a la técnica de determinar la longitud del paso de un usuario, Ladetto [28] presentó un estudio para determinar la longitud del paso y concluyó que puede ser representada como una función dependiente de la frecuencia y la varianza de la señal de la aceleración vertical, proponiendo así determinar la longitud del paso mediante regresión lineal y mejorando el resultado con un filtro de Kalman.

Un método para corregir los parámetros de un algoritmo DR es presentado por Jirawimut y otros [29] —el acimut y la longitud del paso —. El método consiste en estimar los errores de la longitud del paso y del bias del acimut mediante la imple-

[8] Determinación de la posición, conociendo la posición inicial, la distancia recorrida (longitud del paso) y la orientación.

mentación de un filtro de Kalman extendido y los utiliza para corregir el DR, cuando la señal del GPS está disponible. Cuando la señal del GPS no está disponible, los parámetros del DR son calculados usando las estimaciones anteriores.

Cho y Park [30] presentan un sistema de sensores MEMS con un GPS montado sobre el zapato del usuario. Proponen el uso de una red neuronal para determinar la longitud del paso y utilizan la información del GPS para re-entrenar la red una vez el sistema está funcionando.

Kim y otros [31] proponen un método para la detección del paso mediante el análisis de la zancada, el período del paso y la aceleración. El método está basado en el reconocimiento de los patrones de las aceleraciones vertical y horizontal de los sensores instalados en el tobillo del usuario. La configuración que emplean esta compuesta de un acelerómetro de un eje, un giróscopo de un eje y una brújula electrónica. Ellos determinan el patrón de la aceleración teniendo en cuenta el ángulo de inclinación de la pierna. El análisis que llevan a cabo es asumir que el ángulo de inclinación de la pierna varía de 30° a 50°, una aceleración horizontal 0,8 a 2,3g y una aceleración vertical de 0,6 a 2,0g. El ciclo de caminado lo definen en tres fases, las dos primeras las asocian con la "elevación" del pie y la última la definen cuando el talón toca el suelo. Teniendo en cuenta esto sugieren un algoritmo para la detección del paso y proponen una fórmula empírica para calcular la longitud de la zancada.

En el área de la fusión de datos Gabaglio y otros [32] presentan el uso de un filtro de Kalman centralizado para mejorar la información proporcionada por el GPS. Esto lo llevan a cabo modelando los errores del GPS, del acimut y del DR. El acimut es obtenido de dos fuentes: un giróscopo y una brújula electrónica. Esta combinación permite controlar los errores de deriva del giróscopo y las lecturas erróneas de la brújula electrónica cuando se ve sometida a interferencias electromagnéticas.

Ladetto y Merminod [33] proponen un sistema híbrido compuesto por: GPS, barómetro, brújula electrónica y giróscopos y el algoritmo DR. Este prototipo identifica si el movimiento del usuario es hacia delante o hacía atrás o hacía a la izquierda o hacía la derecha; determina el acimut combinando los giroscopios y la brújula teniendo en cuenta que: el giroscopio funciona adecuadamente en cortos períodos de tiempo y permite determinar el bias del acimut de la brújula cuando está es sometida perturbaciones magnéticas; determina la posición desde dos fuentes: el GPS y el DR. Un ejemplo de aplicación lo presentan en [34] cuya finalidad es facilitar la navegación de personas ciegas en entornos desconocidos para ellas.

Renaudin y otros [35] proponen un sistema compuesto por tecnología RFID y MEMS para la localización de usuarios en entornos interiores. Esta es una combinanción en la que se complementa la información del sistema MEMS con la información del sistema RFID, dicho sistema consiste en tener instalados dispositivos RFID en posiciones perfectamente conocidas a los largo del edificio de tal manera que cuando el usuario pasa por alguno de esos puntos corrige la información del algoritmo DR; la información es fusionada con un FKE.

Por otro lado, Weimann y otros [36] presentan un sistema compuesto por múltiples sensores, como un receptor Galileo, unidades de medida inercial, o radares. Este es el primer prototipo que incluye los radares en la localización de usuarios, además emplean esto para la localización en interiores.

2.2.2.2 Prototipos Autónomos

Actualmente se pueden encontrar en el mercado zapatillas que tienen instalados sistemas que permiten determinar la distancia recorrida por el caminante; este fue el resultado de una investigación realizada por [37]. Una continuación de esta investigación es presentada por Stirling y otros [38]. Es un sistema que se coloca en el pie del usuario, compuesto por acelerómetros y sensores magnéticos de bajo coste, para trabajar en modo DR. Para determinar la distancia recorrida mide la zancada, de acuerdo al método propuesto por [39] —consiste en integrar dos veces la aceleración 3D para obtener la distancia horizontal — y su orientación es obtenida a partir de los sensores magnéticos.

Foxlin [40] presenta un sensor inalámbrico para ser instalado en el zapato de un usuario. El sensor es un dispositivo que consta de una unidad de medida inercial y magnetómetros. A diferencia de los otros trabajos este autor no utiliza el DR para determinar la posición, sino que aplica los principios de navegación inercial, es decir deriva la aceleración dos veces para obtener la posición y deriva una vez la velocidad angular para obtener la orientación. Él realiza las correcciones respectivas al problema de deriva propio de la unidad de medida inercial (UMI) (sección 2.1.4.2). Estas correcciones son llevadas a cabo aplicando el método ZUPTs (Zero-Velocity Update, en español sería como la actualización a velocidad cero) cuando el pie toca el suelo. De tal manera que introduce esta corrección como pseudomedida a un filtro de Kalman extendido —el cual es el encargado de estimar los errores de orientación, bias de la velocidad angular y aceleración lineal, posición y velocidad —. La utilización ZUPTs *le permite al FKE corregir el error de la velocidad una vez dado el paso, de tal manera que elimina el error cúbico [40]* (generado en el calculo de la posición). De esta manera lo que este autor pretende es trabajar únicamente con el error lineal acumulado de los pasos registrados. Al igual que el presente trabajo, este autor opina que un sistema así puede ayudar en labores de realidad mezclada.

Moafipoor y otros [41] presentan un sistema difuso como sistema de posicionamiento basándose únicamente en el algoritmo DR. Ellos modelan de acuerdo a su conocimiento empírico los comportamientos del sistema bajo circunstancias como el usuario subiendo escaleras o caminando por zonas con cierta inclinación entre otras. Además tiene en cuenta características del usuario como peso o estatura.

2.2.2.3 Otras implementaciones

Como se mencionó al comienzo de esta sección, hay un cierto interés en la combinación de sistemas de posicionamiento de usuarios con los sistemas de realidad mezclada. Un ejemplo de ello es el trabajo de Kourogi y Kurata [42]. Ellos presentaron un método de localización en entornos interiores y exteriores, con el fin de utilizarlo en un sistema de realidad aumentada. El sistema esta compuesto por una cámara, un sensor inercial para seguir el movimiento de la cabeza, acelerómetros, magnetómetros y giróscopos. A diferencia de un SPU típico no obtiene la posición con un GPS, sino que al igual que en los sistemas de realidad mezclada, la obtiene a través de un sistema de visión incluyendo un análisis de la locomoción del ser humano (principio para los SPUs o del DR para usuarios). La orientación también es obtenida usando los principios de la realidad mezclada y de los SPUs, por un lado obtiene la orientación de la cabeza con un sensor inercial (realidad mezclada) y la orientación del torso del usuario con sensores autónomos (SPUs).

Newman y otros [43] utilizan un sistema de seguimiento óptico para localizar un usuario dentro de un edifico aplicando una metodología basada en la computación ubicua e implementandola para aplicaciones de realidad aumentada.

Kourogi y otros [44] presentan una nueva versión de su sistema eliminando el uso de una cámara como elemento de detección de la posición, pero la mantiene como sistema de visualización. Esta vez, implementan un sistema de identificación de radio frecuencias (RFID) e incluyen el uso del GPS para la localización en exteriores. Siguen manteniendo la aplicación de un sistema DR para dar mayor robustez al sistema en el caso en que falle el GPS o el RFID.

Finalmente, un compendio de algunas tecnologías para sistemas de posicionamiento en la computación ubicua es presentada por Retscher y Kealy [19]. Ellos presentan algunas tecnologías diferentes a las mencionadas en la sección 2.1 tales como Bluetooth. También presentan dos ejemplos, uno aplicado a un vehículo que lleva instalados sensores de localización y cuya información es fusionada mediante un filtro de Kalman. Este es un ejemplo de la importancia de la combinación de técnicas de localización para sistemas de realidad aumentada. Por otro lado presentan un ejemplo de la localización de un usuario dentro de un edificio. En resumen, sostienen que las tecnologías de localización emergentes para los sistemas de navegación pueden ser integradas para incrementar la información del GPS y el DR.

Table 2.2: Comparación entre GPS y Dead Reckoning. Tomada de [16].

Característica	GPS	Dead Reckoning
Método	Basado en Satélites	Autónomos
Posición	3D, absoluto	2D, relativa
Velocidad	Problemas en baja velocidad	Continua
Heading	Diferencias de posiciones	Continuas
Exactitud en corto tiempo	Aceptable	Muy alta
Exactitud en largo tiempo	Alta	Baja
Disponibilidad	Vista del cielo	Ilimitada
Otros problemas	Errores Aleatorios	Campos Magnéticos locales

2.2.3 Optimización en los Sistemas de Localización

En los SPUs se observa un continuo aumento en la exploración de nuevas ideas, para la aplicación de nuevas tecnologías para ambos entornos interiores y exteriores. No obstante, pocas propuestas en el área de la aplicación de técnicas de optimización han sido reportadas para los SPUs, un ejemplo de estas es el trabajo realizado por Moafipoor y otros [41] presentado en la sección de prototipos autónomos 2.2.2.2 y se diferencia porque trata la información de un SPU empleando la teoría de la lógica difusa. Por otro lado, lo contrario ha ocurrido en los sistemas de localización para robots, aviones o barcos, puesto que, las técnicas de optimización para estos sistemas han sido ampliamente estudiadas.

Las técnicas de optimización permiten hacer un uso eficiente de la información en caso de que no se disponga de modelos físicos exactos de los sistemas, la información sea ruidosa o la información esté incompleta. En los SPUs se pueden encontrar diferentes elementos que no pueden ser modelados completamente, por ejemplo, la longitud del paso, las perturbaciones del entorno o la dinámica del usuario, entre otros.

También es posible utilizar técnicas de optimización para diseñar sistemas de toma de decisiones, esto le permite a un sistema automatizar tareas como la realización de cálculos en diferentes escenarios, incluyendo aquellos en los que pueda existir fallo de uno o de varios de sus componentes. Esto conlleva a que inclusive bajo situaciones de falla, el sistema debería ser capaz de responder adecuadamente a las solicitudes del usuario.

En las secciones anteriores se ha mencionado la fusión de datos como una de las tareas fundamentales en los sistemas de localización. Esta tarea en los sistemas de navegación para robots, vehículos, barcos, etc. es comúnmente llevada a cabo mediante el filtrado de Kalman. Esto está prácticamente definido debido a la comprobada eficiencia de esta técnica [12]. A partir de esto, varias investigaciones se han enfocado en encontrar la manera más adecuada de utilizar esta herramienta para mejorar el funcionamiento de los sistemas.

Aprovechando que dentro del marco conceptual del filtrado de Kalman existen componentes que permiten observar si el funcionamiento de un filtro es el esperado, los investigadores han explorado esto para diseñar sistemas de detección de fallos, los cuales permiten identificar errores en el filtro o en los sensores. En otras palabras, diseñan sistemas de detección de fallos a partir del monitoreo del filtrado de Kalman. En los trabajos de Mirabadi y otros [45] y, Hwang y otros [46] (aplicación para un sistema de navegación para aviones) se adopta dicha metodología.

También existen estudios que mediante un cuidadoso análisis y un amplio conocimiento empírico del funcionamiento del sistema realizan implementaciones que combinan el filtrado de Kalman y la lógica difusa [47], [48], [49], [50], [51] y [52]; estos trabajos han sido implementados en sistemas de navegación de robots o vehículos.

En la tabla 2.3 se presenta un resumen en el que se incluyen aspectos generales a tener en cuenta en el diseño de los sistemas de posicionamiento en los dos ámbitos tratados en este capítulo —realidad mezclada y SPUs —. Varios de estos aspectos pueden ser mejorados aplicando técnicas de optimización, aquí se tratará desde el punto de vista de la detección de fallos y la selección de las medidas más adecuadas para obtener la posición final.

Table 2.3: Características de los sistemas de posicionamiento

Característica	Campo de Aplicación	
	Realidad Mezclada	SPU
Exactitud	Debe ser muy alta para no sufrir las típicas desventajas que experimenta el usuario [53].	Ciertos niveles de error pueden ser permitidos alrededor de centímetros, porque es una localización mas global.
Precisión	Muy alta.	Muy alta.
Disponibilidad	Depende de la aplicación.	De alta prioridad, normalmente son sistemas que se espera puede ser usados en cualquier momento y en cualquier ambiente.
Cobertura	Depende de la aplicación, pero si es ambientes exteriores extensos se requiere un alto nivel.	De alta prioridad.
Confiabilidad	Muy alta.	Muy alta.
Velocidad de actualización	Muy alta, aunque aún no se dispone del valor deseado.	Alta, es menos restrictivo que la realidad mixta.
Peso	Ligero, para ser llevado de forma cómoda por un ser humano.	Ligero, para ser llevado de forma cómoda por un ser humano.
Tamaño	Lo mas pequeño posible, para ser llevado de forma cómoda por un ser humano.	Lo mas pequeño posible, para ser llevado de forma cómoda por un ser humano.
Perturbaciones del ambiente	Depende de la aplicación.	Muy difíciles de controlar.
Coste	Su precio actual es elevado.	Se puede construir con hardware de bajo coste, pero el sacrifico sería el tiempo de implementación requerido.
Alimentación	Depende de la aplicación	Baterias con capacidad para muchas horas.

2.3 Resumen

Se presentaron los principales aspectos de las tecnologías utilizadas en el campo de la localización de objetos, así como una comparación de las mismas. La tecnologías incluidas fueron la acústica, la óptica, la magnética, la mecánica y radio frecuencias.

Posteriormente, se realizó un compendio de las investigaciones desarrolladas bajo el concepto de los sistemas híbridos (los cuales son sistemas compuestos por varias de las tecnologías mencionadas). Se hizo una distinción entre los sistemas híbridos según dos campos centrales de aplicación: la realidad mezclada y los sistemas de navegación para usuarios.

Las investigaciones y sistemas presentados en el campo de la realidad mezclada —más conocidos como sistemas de seguimiento — son prototipos que incluyen varios de los conceptos tratados en esta tesis, como la fusión de datos o la utilización de varias tecnologías de manera complementaria. En el caso de los sistemas de navegación para usuarios, los trabajos que se tuvieron en cuenta fueron clasificados según sus aportes o principios teóricos como aquellos que profundizan en la dinámica del usuario o la combinación de diferentes tecnologías. También fueron incluidos trabajos que realizan una combinación de los conceptos empleados en los sistemas de navegación para usuarios y los sistemas de seguimiento en la realidad mezclada, involucrando desde sistemas basados en visión hasta GPS o RFID dentro de un mismo sistema.

La optimización es uno de los principales temas a afrontar en esta tesis, por ello se incluye un apartado que trata este tema. Debido a que no se encontró literatura específica en el campo de la localización de usuarios, se realizó un pequeño compendió de los trabajos más relevantes en los sistemas de navegación clásicos, concretamente en aquellos que se enfocaban a buscar mejorar los resultados del filtrado de Kalman en la fusión de datos.

Por último, se presenta una tabla-resumen en la que se incluyen las principales características de los sistemas de localización para la realidad mezclada y los sistemas de navegación para usuarios.

2.4 Summary

The main aspects of the technologies used in the field of object location were presented together with a qualitative comparison. The technologies included were acoustic, optic, magnetic, mechanical, and radio frequencies.

Next, a compendium of related works on hybrid systems was given. Two fields of applications were taken into account: mixed reality and navigation systems for users.

The selected related works on hybrid systems (more commonly known as tracking systems) for mixed reality are prototypes that include several concepts dealt with in this thesis, such as data fusion or the combination of several technologies in a complementary way. For navigation systems for users, the works included were those that took an in-depth look at user dynamics or integrated systems based on several different technologies. Other works that use the whole philosophy from both fields, like [43] and [44], are also included; in these cases vision technology, GPS or RFID, among others are used in the same system.

Optimization is one of the main subjects to be addressed in this thesis, and this is why a section on the subject has been included. Due to the fact that there was no specific literature on location systems for users, a summary of the most relevant works on classic navigation systems was drawn up. More specifically this includes works that focused on achieving better performance from Kalman filtering, since this is the technique use to carry out the data fusion in this thesis.

Finally, a summary-table is presented. This includes the main characteristics (analyzed in a qualitative way) of the location system for both applications, that is, mixed reality and navigation systems for users.

CHAPTER 3

Localization

IN THIS chapter we outline the sensors that make up the UPS prototypes to be analyzed, as well as describing the physical-mathematical models of these systems, the algorithms used to determine the position and orientation, and the optimization mechanism.

3.1 Background

As was mentioned in the previous chapter, the discipline that covers many aspects of the design of systems with features of the UPSs is known as *Multisensor Data Fusion (MDF)*. This is defined as [54] *"the acquisition, processing and synergistic combination of the information provided by several sources and sensors to provide a better understanding of a phenomenon´´.*

In fields like robotics or industry (e.g. automotive), the application of this concept is seen to take place on a continual basis. A large amount of research in these fields is focused on optimizing systems by applying artificial intelligence techniques (e.g. neural networks or fuzzy logic), in addition to the conventional design developments. The purpose of such work is to obtain better characteristics, such as robustness, accuracy or reliability.

Research on UPSs started to become more intensive about a decade ago. In the literature review that was carried out it was observed that, in practical terms, the above mentioned techniques have received little attention from researchers. Any of them can be applied for different purposes within UPSs. For example, Renaudin et al.[35] use fuzzy logic to determine the posture of a person, while Moafipoor et al.[41] use the same technique to determine the parameters of a dead reckoning algorithm.

Taking into account everything that was outlined earlier, the design of a UPS would involve the following general tasks: (1) a careful analysis of the information provided by the sensors so as to be able to take maximum advantage of it; (2) data fusion by means of some estimation technique; and (3) analysis of the fused information so as to be able to apply inference mechanisms (in highly sophisticated process *Knowledge-Based System* are applied). In other words, for this study, the design tasks were classified as follows:

1. Selection of appropriate mathematical-physical models to obtain position and orientation.
2. Analysis of the information provided by the available sensors.
3. Selection and evaluation of an estimation technique.
4. Selection and evaluation of an inference mechanism.

Selection of appropriate mathematical-physical models to obtain position and orientation. A good understanding of the physical and the mathematical behavior of the systems is necessary. Therefore, this task includes all the models for both indoor and outdoor environments that are needed to know a user's position and orientation.
Analysis of the information provided by the available sensors. This involves analyzing the individual characteristics of each sensor and defining their functions within an integrated system.
Selection and evaluation of an estimation technique. The most efficient and widely used method of data fusion in navigation systems is *predictive Kalman filtering* [12], [55], which is outlined in Appendix A. This technique can be applied to fit the designer's needs or preferences, since it can be used to fuse all the data or simply to filter and correct part of the data. For this thesis, Kalman filtering will be implemented taking into account the environment and the information that is available —provided by the sensors. Moreover, issues concerning evaluating and correcting or tuning the filtering are also presented.
Selection and evaluation of an inference mechanism. This task covers issues to improve the performance of the whole system. In order to achieve this, the concept of *Knowledge-Based System* is considered. Here, *possibility theory*, *fuzzy logic* [56] and *causal diagnosis* [57] will be used because these allow decision-making and numerical information to be combined in a suitable manner. The basic theory of those subjects is presented in Appendix B.

3.2 Sensors

Today, radio frequencies, and magnetic and inertial technologies are considered to be the most suitable for *user positioning systems (UPSs)*. This is due to the fact that they offer a series of characteristics that make them more useful in personal localization, such as their levels of accuracy, the ergonomics, or because of the environment in which they work. For these reasons the following sensors are used:

- GPS Receiver
- Inertial Measurement Unit (IMU)
- UWB System

GPS Receiver: The GPS receiver used is an A1025 from Tyco Electronics, whose technical characteristics are summarised in Table 3.1. The format for sending data is given by *NMEA (National Electronics Marine Association)* sentences, an example of which is presented in Figure 3.1. The sentences contain all the information provided by the GPS system. In this work the information is extracted from the sentences $GPRMC, $GPGGA, $GPGSA and $GPVTG and refers to the speed, the altitude, the longitude, the HDOP and the time.

Item	Range	Unit
Accuracy	3	meters
Max. update rate	1	Hz
Supply	3.3/100	V/mA
Weight	5	g
Communication	RS232	—

Table 3.1: Technical Characteristics of GPS Receiver A1025 from Tyco Electronics

e.g. $GPGGA,152145.000,4805.8193,N,01132.2317,E,1,04,2.5,607.75,M,47.6,M,,*67		
(1)	$GPGGA	Vendor and message identifier
(2)	152145.000	Universal time coordinated (15h 21m 45.000s)
(3)	4805.8193	Latitude (48deg 05.8193min)
(4)	N	North (or S for south)
(5)	01132.2317	Longitude (011deg 32.2317min)
(6)	E	East (or W for west)
(7)	1	GPS fix valid (or 0 for fix not available or 2 for differential fix)
(8)	04	Four satellites in view (min 00, max 12)
(9)	2.5	Horizontal dilution of precision
(10)	0607.75	Antenna altitude above/below mean sea level (geoid)
(11)	M	Unit of antenna altitude: meters
(12)	47.6	Geoidal separation
(13)	M	Unit of geoidal separation: meters
(14)	<empty>	Age of differential GPS data, null field when DGPS is not used
(15)	<empty>	Differential reference station ID, null field when DGPS is not used
(16)	*67	Checksum

Figure 3.1: Sentence in format NMEA provided by the receiver GPS A1025 [8]

Inertial Measurement Unit The IMU used is the MTx device from Xsens consists of gyroscopes, accelerometers and magnetometers, which provide 3D information about linear accelerations, angular velocities and magnetic fields, respectively. This is an IMU that provides information mainly about orientation i.e. roll, pitch and yaw. Its technical specifications are summarised in Table 3.2.

System based on UWB The system used is from Ubisense and its configuration can be observed in Figure 3.2. This is a technology based on radio frequency and

Item	Value
Dynamic Range	All angles in 3D
Static Accuracy (roll/pitch)	<0.5°
Static Accuracy (azimuth)	<1.0°
Dynamic Accuracy	2° R.M:S
Update Rate	120 Hz

Table 3.2: Technical specifications of the IMU MTx from Xsens

it is a very comfortable for the user, since the user only has to carry a small light device that can be identified by the system, thus allowing it to triangulate in order to determine the user's position. This system provides the (X,Y,Z) coordinates with an accuracy of 50cm.

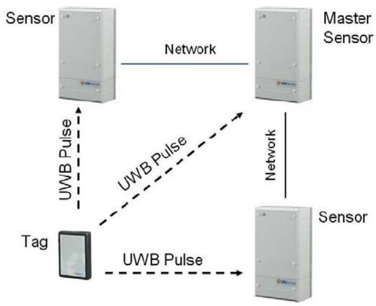


Figure 3.2: UWB system from Ubisense

3.3 Dynamics of the User

When a user walks it is assumed that they have a similar behavior to a *low dynamic 2D linear movement system [58]*, that is, the movement of a non-accelerated body. Therefore, the dynamics of a user walking is represented as:

$$\begin{aligned} X &= X_0 + V_X \Delta t \\ Y &= Y_0 + V_Y \Delta t \end{aligned} \tag{3.1}$$

where X is the coordinate in X; Y is the coordinate in Y; V is the speed; and Δt is the time interval.

To obtain the parameters of equation system 3.1, two very common methods can be used. The first is the *Dead Reckoning Algorithm*, which, as will be described later on, replaces speed by equivalent information instead of working with it directly as a parameter.The second method consists in determining the positions using the information provided by localization sensors like the GPS receiver.

These methods can be applied individually, which is not the most appropriate way, since it does not endow the systems with sufficient robustness and reliability. Therefore, they are applied in a complementary fashion (as in the case of traditional navigation systems for vehicles), which increases the functionality of a system and improves its characteristics.

In the following subsection, the DR algorithm is described, which requires a detailed analysis due to the techniques that should be used to determine the parameters needed to know the position of a user who is walking.

3.3.1 Dead Reckoning Algorithm

This is one of the oldest methods to calculate position, and is the one that has been commonly applied by sailors to localize their ships. Nowadays it continues to be a very popular method of localization and is used both in navigation systems and in other applications, for example, biomechanics. Its theoretical principle is: *obtain the position of an object by calculating the distance traveled by the object based on knowledge of the initial position, the speed, the orientation and the time.*

DR is applied to UPSs, assuming that the object to be analyzed is *a walking user* whose position can be determined by means of:

$$\begin{aligned} X &= X_0 + \sum_{j=1}^{i} s_j \ \sin(\psi_j) \\ Y &= Y_0 + \sum_{j=1}^{i} s_j \ \cos(\psi_j) \end{aligned} \tag{3.2}$$

where X is the coordinate in X and X_0 is the initial point; Y is the coordinate in Y and Y_0 is the initial point; s is the step length of the user and ψ is the orientation or azimuth.

This system of equations addresses the problem of the movement of a pedestrian directly, and uses a well-known initial position, *the step length* and the orientation to obtain the position of a pedestrian.

The critical parameters in DR are *step length* and *orientation*. Step length is determined by analyzing the user's locomotion when walking, while orientation is provided by the IMU.

The action of walking in humans is a process of locomotion in which the body rests first on one leg and then on the other [9]. Figure 3.3 illustrates the walking process and Rose and Gamble [9] identify two main phases in that process. The first has been called *the stance phase* and the second is *the swing phase*. This pattern is known as *the walking cycle*.

This analysis of the pedestrian's locomotion is very important for the UPSs, since it allows them to obtain information about *the occurrence of a step*. In other words, it allows them to ensure that a step has been taken and tells them they can proceed

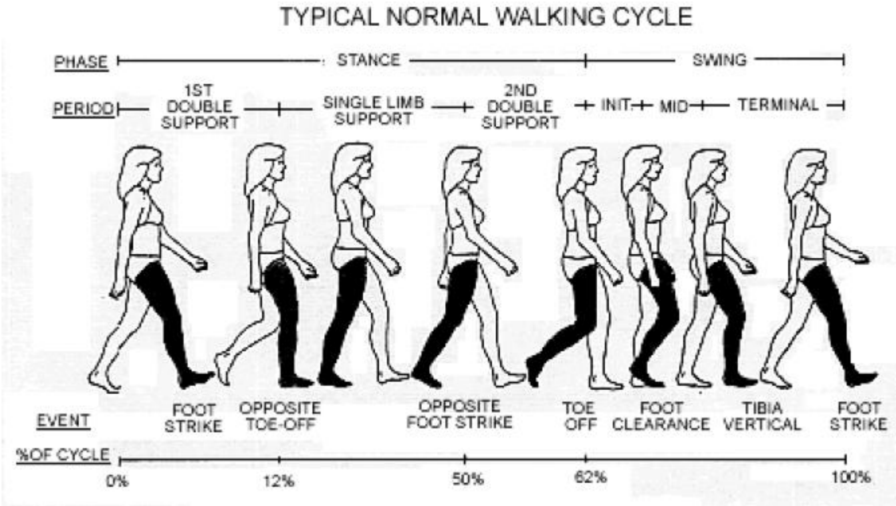


Figure 3.3: Standard walking cycle [9]

to calculate the step length. Researchers have proposed some empirical methods to detect the step and to calculate its length, and these are described in the following.

Step Detection Different proposals have been put forward for detecting the occurrence of a step, one of the most important being the pioneering proposal by Levi and Judd [17]. These authors analyzed the pattern presented by the vertical linear acceleration while a person is walking. In the stance phase the acceleration signal presents peaks and these are associated to the occurrence of a step. In other words, a peak should be detected when the user's foot hits the floor.

To detect a step, first it is necessary to know the thresholds of the acceleration peak[1] and the occurrence time period[2]. Once these values are known, the algorithm consists in: (1) taking a series of samples of the vertical acceleration; (2) evaluating the signal to detect possible peaks and applying the Fast Fourier Transformed (FFT) to obtain the signal frequency; (3) if a peak was registered, the thresholds both for the peak and for the period are evaluated; and (4) if the conditions have been satisfied the step is detected, otherwise the next series of data is taken.

Step Length This is a parameter that varies in time and thus has to be calculated continually. Different methods exist to obtain the step length, but most of them are empirical. Here the most common are mentioned.

- According to Levi and Judd [17] the step length can be defined as a linear

[1] Minimum value of the peak associated to the occurrence of a step.
[2] Minimum value that should exist between the occurrence of one step and the next.

function that depends on the step frequency and a predetermined step length.

$$s = s_0 + m(f - f_0) \tag{3.3}$$

where s is the step, f is the step frequency and m is the slope.

- Another method is proposed by Ladetto [28]. In his investigation, the step length can be calculated according to its frequency and the signal variance. This is performed by means of linear regression as indicated in the equation 3.4, where the parameters a, b, c, are the coefficients of linear regression and w is a Gaussian noise of mean zero.

$$s = a + bf + c + w \tag{3.4}$$

3.3.2 Methods to Calculate the DR Parameters

In this section the methods applied for detection the step and calculating of its length are described.

3.3.2.1 Step Detection Algorithm

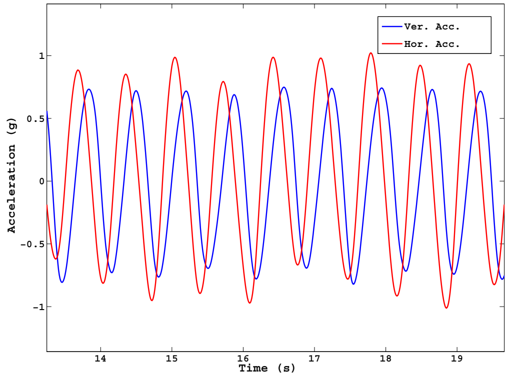
Figure 3.3 illustrated the different phases of the pattern when a person takes a step. Here, the cycle is considered to be divided in three phases, in which: (1) the heel leaves the ground, (2) the foot moves forward, and (3) the heel hits the ground. This pattern can be detected by analyzing the signal of the linear vertical acceleration [17]. To observe it, sensors can be located on the human body near the pelvis or on the foot. The first option is based on physiological studies that conclude that the center of gravity of the human body is located in this area [59]. Results of these studies also show that there is a standard walking pattern that is independent of the individual characteristics of each person (e.g. their weight or height).

Although the pattern of the vertical acceleration is appropriate for step detection, proposing new alternatives to improve this detection continues to be an issue of interest for the researchers. In this search for new alternatives in the works of Brannstrom [60], Kim et al. [31], and Kourogi and Kurata [42], the use of horizontal acceleration is suggested to increase the robustness of the procedure and, in the same way, increase the robustness of UPSs. In this research the same approach was adopted and therefore an analysis that includes both the vertical and the horizontal accelerations is presented in what follows.

Step detection is a task in which several empirical considerations must be taken into account. For this reason, several tests were performed and the detection algorithm was defined in accordance with the observations.

The sensor that provides the information about the vertical and horizontal accelerations is the IMU. This unit was firmly attached to a belt worn by the user, as illustrated in Figure 3.4(a). An example of the patterns of accelerations (after the signal has been processed by means of filters) is presented in Figure 3.4(b).

(a)



(b)

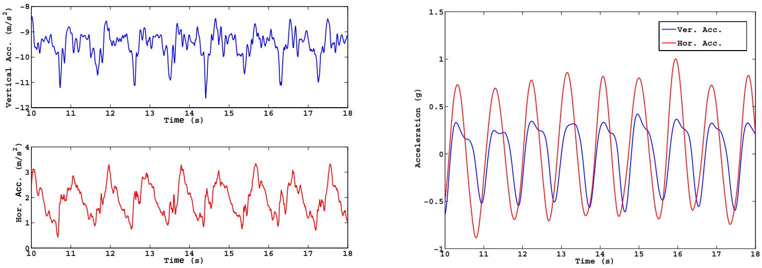
Figure 3.4: (a) A user with the set up to observe the accelerations' pattern (b) Accelerations' pattern

Figures 3.5, 3.6 and 3.7 are included as an example of the tests that were performed. The vertical and horizontal accelerations that are illustrated were measured under the following conditions: the user walked at three different speeds: slow, normal and quick. The results of these tests would later be used to determine the step length, the methodology for which is described in the next section. Here, they are presented to illustrate the fact that while the user maintains the standard walking conditions, the pattern of accelerations can be identified.

According to the analysis by Levi and Judd, in the vertical acceleration signal there should be a peak associated to the phase (3) of the step cycle (also known as the stance phase) of the step cycle. As can be observed in Figures 3.5a, 3.6a and 3.7a the accelerations are noisy signals, in which more than one peak can be found in this phase as has been mentioned in other studies and confirmed here. The most common way to solve this problem is to apply filters in order to eliminate unwanted peaks. Therefore, signals were processed through low-pass filters with a cut-off frequency of 50 Hz (the sampling frequency is 100Hz). The results of this process are shown in Figures 3.5b, 3.6b and 3.7b.

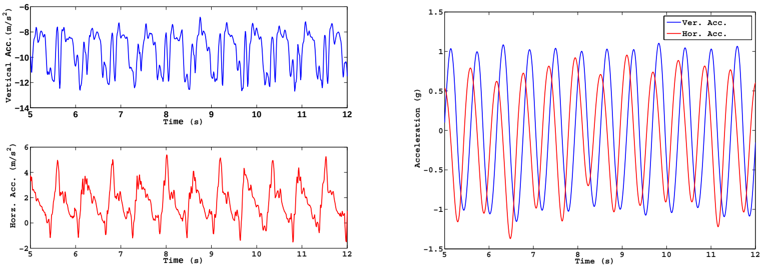
It is necessary to have a minimum sampling frequency [28] to obtain the pattern of accelerations illustrated in Figure 3.4b in an appropriate way. To observe this, tests were carried out with the following frequencies: 100Hz (sampling frequency of the IMU), which is enough to observe the expected pattern, and 40 Hz and 20 Hz, which are registered when data are processed through MATLAB. No notable differences were observed, and the cut-off frequency was simply adjusted for each case. In general expected patterns were obtained.

On the other hand, the values of the peaks have to be selected according to the



(a) (b)

Figure 3.5: Accelerations when the user walks very slowly

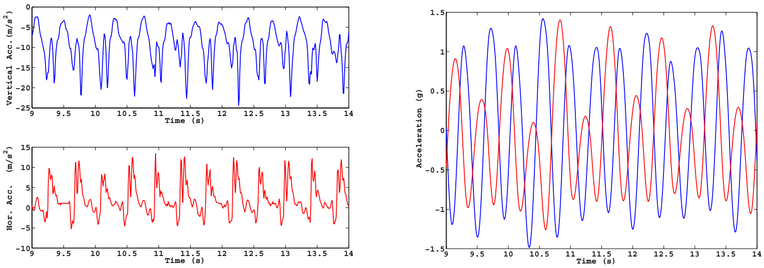


(a) (b)

Figure 3.6: Accelerations when the user walks normally

characteristics of each users, which means that no standard values can be determined to fit all cases. Another variable that ensures correct detection is *the step occurrence period*. The maximum and minimum thresholds of this variable are defined according to the limits established for the expected step frequencies (under standard walking conditions), which are between 0.5 Hz and 3 Hz [17]. The frequency can be determined by applying the (FFT) or by calculating the difference in times between two peaks. The FFT, however, has two disadvantages. One is that it requires a minimum number of samples of the signal and the other is that it assumes that the frequency of the pedestrian's walking is almost constant. For this reason, the second option was applied here, the step being considered as a time-varying parameter.

Assuming that the right conditions are given for the sampling frequency, the following would be fulfilled: initially two positive peaks are presented —one in the vertical acceleration and one in the horizontal acceleration — and in the third phase of the walking cycle, when the foot hits the ground, two negative peaks appear — again, one in the vertical acceleration and one in the horizontal acceleration. Additionally,



(a) (b)

Figure 3.7: Accelerations when the user walks quickly

the thresholds for the acceleration peak and for the step occurrence period are established in accordance with the calibration parameters obtained for the pedestrian. This information is then used to defined the detection algorithm shown in the diagram in Figure 3.8.

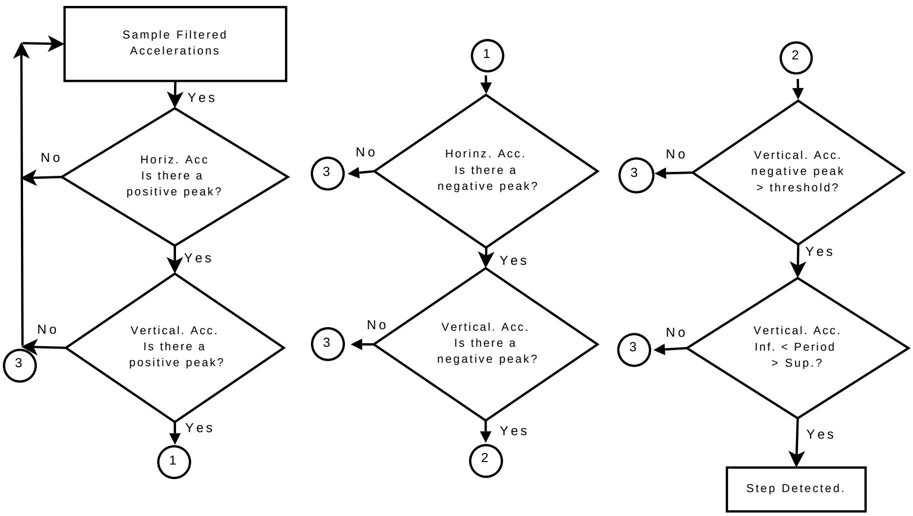


Figure 3.8: Step detection sequence

3.3.2.2 Technique to Calculate the Step Length

Step length is interpreted as the distance traveled by the foot during the step cycle. Several techniques have been proposed to calculate it. The most popular have been mentioned in 3.3.1 and are based on linear functions. Very recently the use of fuzzy logic [41] and neural networks [30] has also been suggested. This diversity of alternatives it is not only due to the search for methods of optimizing this problem, but also to the inherent difficulty involved in finding a solution to it.

Here the neural networks method is adopted to determine the step length. It is based on the fact that: (1) they have a capacity for learning, which is considered to be an advantage due to the variability of the step length, and (2) determining the step length is a very complex non-linear problem to modeling; neural networks are suitable for dealing with problems with these sorts of characteristics.

Neural networks are systems that need to be trained to find the solution to a problem; in other words, they have to be taught the appropriate behavior of the system under specific circumstances. Due to the characteristics of the problem in question, it is possible to apply *supervised training*. This training can be applied when it is known how the system will respond to a particular input. This means that it is necessary to have the exact data from the inputs to the system and the exact data from the outputs that correspond to these inputs.

If the network is trained correctly, it is not necessary to know the method or the model it uses to obtain the output data —it will simply provide the correct output data associated to the inputs. This facilitates the task of processing the problem, but it requires careful rigorous work on providing the training data (input-output pairs).

Step length can be represented as a function of the step frequency and the variance in acceleration. Hence, *the inputs to the network are defined as: the step frequency and the acceleration variance, while the step length is defined as the output of the network*.

The frequency is calculated as indicated:

$$fs_N = \frac{1}{Ts_k - Ts_{k-1}} \tag{3.5}$$

where Ts is the time associated to the occurrence of a step. The expected step length was calculated by dividing the distance traveled by the number of steps. To cover a wider range of possibilities, the training data are registered with the user walking at different speeds or rhythms: very slow, slow, normal, quick and very quick. This can also be applied to several situations, for example, when the user is walking on different slopes, moves backwards etc., but such variations have not been included here.

Figure 3.9 shows the frequencies and the variances for the different speeds that were used to train the network. The training data for the output will be presented later on.

Once the appropriate training data had been obtained, several tests were carried out with different network architectures. This was performed using the Matlab Toolbox for neural networks. Several characteristics were taken into account to find the network that presents the minimum error including the number of hidden layers, the number of neurons in the hidden layers, the activation function for the hidden layers and the activation function for the output layer (in this case there is a single output layer, since only the step length is expected as an output parameter).

After many trial-and-error tests, the network that presented the minimum root

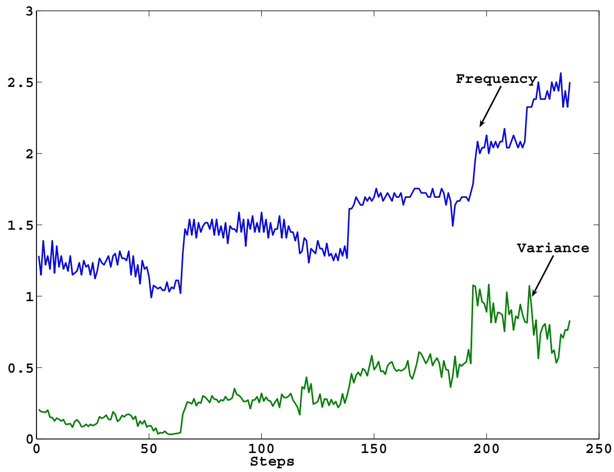


Figure 3.9: Training data for the network inputs (frequency and variance)

mean square error (RMS) has the following characteristics: it is a multilayer feed-forward network with five neurons, a sigmoid-like activation function in the hidden layer, and one output layer with a linear activation function. This configuration is illustrated in Figure 3.10.

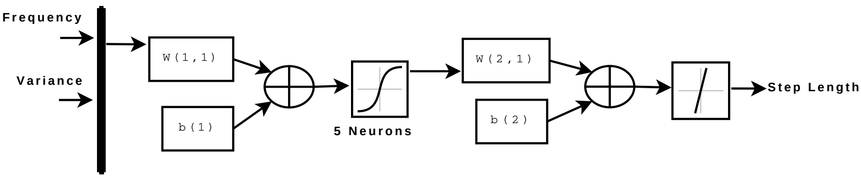


Figure 3.10: Neural Network

The weightings for the layers ($W\{1,1\}$, $W\{2,1\}$), with their respective biases ($b\{1\}$, $b\{2\}$), have the following values:

$$W\{1,1\} = \begin{bmatrix} -40.0823 & -2.8074 \\ 7.3081 & -32.079 \\ -20.0538 & -17.5692 \\ 50.7886 & 101.5925 \\ -15.3422 & -16.7743 \end{bmatrix},$$

$$W\{2,1\} = [-0.15468 \ -0.066323 \ -0.23578 \ 0.098932 \ -0.09178],$$

$$b\{1\} = \begin{bmatrix} 92.1913 \\ -2.4229 \\ 50.9007 \\ -117.1101 \\ 18.0086 \end{bmatrix}, \; b\{2\} = [1.4656]$$

A summary of the frequencies obtained according to the pedestrian's speed is presented in Table 3.3, together with the step lengths associated to these frequencies and the errors of the network. The table shows one of the typical characteristics of the step length, i.e. as the pedestrian increases his or her speed, the step length also increases.

Table 3.3: Results for the Calculation of the Step Length

Speed	Step Length (m)	Frequency (Hz)	Error (m)
Very Slow	0.94 – 1.006	0.99 – 1.389	0.01943 – 0.03372
Slow	1.04 – 1.08	1.266 – 1.587	−0.0290 – 0.0444
Normal	1.174	1.639 – 1.754	−0.0021 – 0.0096
Quick	1.408	1.961 – 2.174	−0.0088 – 0.0034
Very Quick	1.565	2.326 – 2.564	$1.25x10^{-4}$ – 0.0148

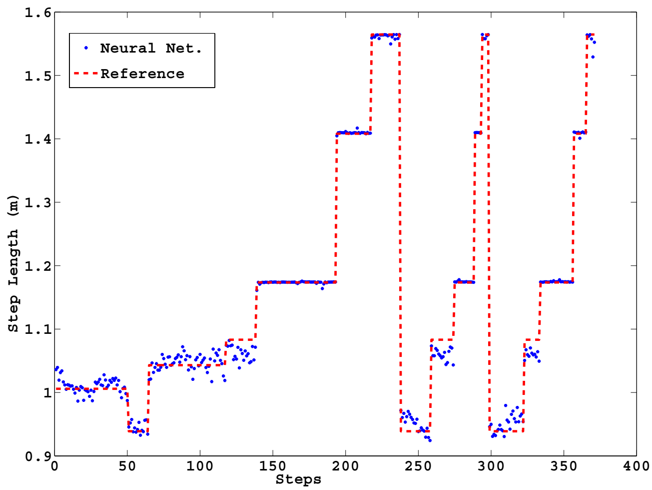


Figure 3.11: Results of the neural network

The results of the network are presented in Figure 3.11, which shows the data training for the step length and the step length provided by the network at different frequencies.

From Figure 3.11, several aspects can be noted. First, there is a certain amount of overlap of step lengths with values between 1.006 m and 1.08 m, since they have

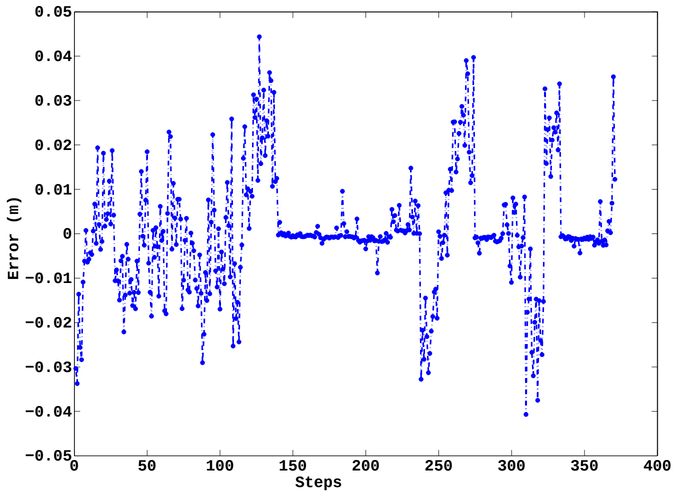


Figure 3.12: Error of the neural network

similar frequencies registered between 1.266 Hz and 1.389 Hz. This aspect generated a conflict in the determination of some lengths, which resulted in an especially high error between steps 118 to 130, 259 to 274 and 323 to 333. In spite of this drawback, the error that was detect does not exceed 4cm and these values are therefore acceptable for DR. On the other hand, it is observed that for step lengths of 0.94 m and 1.04 m identification was very good with lower errors.

Second, it is observed that for normal, quick and very quick speeds, the frequencies are clearly defined, and the different step lengths can be perfectly distinguished for each of them. Hence, the error is very low, the minimum value being near zero and the maximum value is 1.4 cm.

3.4 Models and Algorithms

In this section we will defined the mathematical-physical models for the systems to be analyzed and also algorithms to determine the position, for both indoor and outdoor environments.

3.4.1 Indoor Environments

The analysis presented below deals with the problem of localizing a user in indoor places from two points of view. The first consists in using the DR algorithm as a stand-alone method, while the second involves the use of a complementary system composed of an IMU and information provided by the UWB-based localization system.

3.4.1.1 The DR Algorithm

The system of equations 3.2, which is the one that governs the DR algorithm does not consider the errors of the sensors, which depend on the kind of technology.

As the sensor that will provide the information is the IMU, its errors should be taken into account. This unit can be affected by sources of magnetic interference or by its own errors, which can give rise to errors in the readings that can be particularly critical in the case of the azimuth. On the other hand, the neural network could also present errors which can be due to its failing to recognize the frequency or the variance or to errors in the network itself. Regardless of the cause of the errors, they should be included for the analysis as follows:

$$\check{s}_k = s_k + \varepsilon_s \tag{3.6}$$

$$\check{\psi}_k = \psi_k + \varepsilon_\psi + \delta \tag{3.7}$$

where ε_s and ε_ψ represent the errors associated to the step length and the azimuth, respectively, and δ is the Gaussian error of mean zero. These variables replace the corresponding variables in 3.2.

A system based on this method is less robust than complementary systems, because it does not have any additional information with which to carry out an appropriate correction of the errors. It requires other sources such as maps, GPS or some other system that allows correction of the errors to be carried out on a continuous basis.

3.4.1.2 System based on UWB-IMU

In this section we present the model for an integrated system based on the UWB and inertial technologies. The system is made up of the Ubisense system (which is based on UWB) and the IMU (a system of sensors based on inertial technology).

The dynamic model for user localization is represented by means of the following equations:

$$\begin{aligned} X_k &= X_{k-1} \ + \ s \ \cos(\psi_k) \\ Y_k &= Y_{k-1} \ + \ s \ \sin(\psi_k) \end{aligned} \tag{3.8}$$

where the coordinates are X and Y, s is the step length and ψ is the azimuth.

As can be observed, this model allows the user's position to be obtained by combining information from DR and from the UWB system. It is defined as *complementary system*, since it uses the step length as a *support or information that is redundant* to the data provided by the UWB system. Nevertheless, if there is no information about the step length (due to a failure in step detection or to failure to register the user's movements), the positions continue to be determined by the UWB system.

These equations are similar to the classic equations of movement of a low dynamic system, but instead of a completely time-dependent system what is obtained is a system that could depend on the step occurrence to update the position.

Since the information from the different parameters is obtained from different sources (UWB, IMU) data fusion must be performed. This will be carried out by applying a discrete Kalman filter (KF), the theory of which is detailed in Appendix A.1. This filter requires the definition of two models: *the state model* and *the observation model*. The general equation for the state model in the KF is A.1 and the approach that is carried out for this system is:

$$x_{k+1} = \Phi_k x_k \tag{3.9}$$

where x is the state vector and Φ is the transition matrix .

For this system two state models have been defined. The first one is a model whose transition matrix and whose state vector are represented by 3.10 and by 3.11, respectively:

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-\beta_\psi \Delta t} \end{bmatrix} \tag{3.10}$$

$$x_k = [X_k \; Y_k \; \varepsilon_\psi \;]^T \tag{3.11}$$

ε_ψ is the estimate error of the azimuth and is modeled as a Gauss-Markov process of first order. This type of process has an autocorrelation function that tends to zero as $\tau \rightarrow \infty$ and is defined by[3]:

$$R_x(\tau) = \sigma^2 e^{-\beta|\tau|} \tag{3.12}$$

where σ^2 and $1/\beta$ are the mean square and time constant of the process. Then the azimuth is determined as:

$$\psi = \psi' + \varepsilon_\psi \tag{3.13}$$

where ψ' is the azimuth obtained from the IMU.

This state model is considered to be more useful for an analysis of the sensor signals, and also when the user is not in continuous movement. The second model (which is presented later on) includes the case of the user in movement.

As regards the observation model (the general equation for which is A.2), the approach that is carried out is given as follows:

$$y_k = H_k x_k \tag{3.14}$$

[3] Definition taken from [12]

in which the observation vector y_k is defined as :

$$y_k = \left[\tilde{X}_k\ \tilde{Y}_k\ \tilde{\varepsilon_{\psi_k}}\right]^T \tag{3.15}$$

where the first two components represent the (X,Y) measure provided by the UWB system, while $\tilde{\varepsilon_\psi}$ is determined as:

$$\tilde{\varepsilon_\psi} = \psi_{true} - \psi^{'} \tag{3.16}$$

and the observation matrix matrix —denoted by H — is given by:

$$H = [I]_{3\times 3} \tag{3.17}$$

The equations of the user dynamics (3.8) are used to defined a new model that represents the movement of a pedestrian in the best possible way. The new state vector and the new transition matrix are thus represented by:

$$x_k = [X_k\ Y_k\ \psi_k\ s_k]^T \tag{3.18}$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 & \cos(\psi_k) \\ 0 & 1 & 0 & \sin(\psi_k) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.19}$$

As regards the observation model its respective observation vector and H matrix are given by:

$$y_k = \left[\tilde{X}_k\ \tilde{Y}_k\ \tilde{\psi}_k\ \tilde{s}_k\right]^T \tag{3.20}$$

$$H = [I]_{4\times 4} \tag{3.21}$$

Finally the algorithm of the KF illustrated in the diagram A.2 is applied and its results are presented in Chapter 4.

3.4.2 Outdoor Environments

The sensor that is typically employed for outdoor localization is the GPS, since it provides information about the absolute position. Nevertheless, it is a system that can give unacceptable errors when its signal is obstructed or attenuated, or it is affected by the “multipath” phenomenon. To get round this problem, the same principle as the one employed in most research studies (both in this type of applications and in robust structures, such as robots) is applied. This consists in defining a system that behaves better than the GPS system on its own.

The solution that is often adopted is to apply the concept of multiple sensors (section 3.1), which for outdoor navigation systems implies combining GPS with other systems (INS, IMU, etc.). In this work we have defined *a system based on DR which integrates the information from the GPS and the IMU*.

The information from the GPS can be used in two ways. First, it can be improved by applying the Kalman filtering and, second using the Kalman filter allows the corrected DR parameters [29], i.e. the step length and the azimuth, to be improved. To apply the Kalman filter, the user dynamics should be defined. This is represented by:

$$\begin{aligned} PE_k &= PE_{k-1} + V_{E_k}\Delta t \\ PN_k &= PN_{k-1} + V_{N_k}\Delta t \end{aligned} \tag{3.22}$$

where PE is the east position, PN is the north position, V is the velocity and Δt is the time interval.

To apply the second technique, the step length and the azimuth are modeled in a similar way to that described in [29]:

$$\tilde{s_k} = s_k + \varepsilon_s \tag{3.23}$$

$$\tilde{\psi_k} = \psi_k + B + \mu \tag{3.24}$$

In 3.23: $\tilde{s}$ is the measured step, s is the true step and ε_s is the error of the step. In 3.24: $\tilde{\psi}$ is the measurement of the azimuth, ψ is the true azimuth, B is the true bias and μ is the Gaussian noise of mean zero.

B is given by:

$$B = B_{SENSOR} + \varepsilon_B \tag{3.25}$$

where B_{SENSOR} is the sensor bias and ε_B is the bias error .

The estimation technique used are the discrete Kalman filter (KF) and the extended Kalman filter (EKF). The KF will be used to estimate the position from the GPS measurements. The job of the EKF will be to estimate the errors of the step length and of the azimuth from the information obtained from the GPS and the IMU.

According to the concepts presented in A.1 and A.2 (the sections which includes the concepts of the KF and the EKF) the state and observation models are then defined, together with other basic components needed to design each filter.

The general state equation for linear systems, A.1, is used to produce the following approach to the state model:

$$x_{k+1} = \Phi x_k + w_k \tag{3.26}$$

Given that the KF will estimate the position, the state vector is defined as follows:

$$x = [PE\ PN\ V]^T \tag{3.27}$$

where PE, PN, V represents the east and north coordinates, and the velocity, respectively.

The transition matrix is defined as follows:

$$\Phi = \begin{pmatrix} 1 & 0 & V\sin(\psi_k)\Delta t \\ 0 & 1 & V\cos(\psi_k)\Delta t \\ 0 & 0 & 1 \end{pmatrix} \tag{3.28}$$

where ψ is the true azimuth.

The observation vector is given by:

$$y = \left[\tilde{PE}\ \tilde{PN}\ \tilde{V}\right]^T \tag{3.29}$$

which components are the measures obtained from the GPS. To complete the models for this filter, the observation matrix is defined as follows:

$$H = [I]_{3\times 3} \tag{3.30}$$

The general state equation for non-linear systems, A.15, is used to produce the following approach to the state model:

$$x_{k+1} = \Phi x_k + w_k \tag{3.31}$$

Given that the EKF will estimate the errors of the DR parameters, the components of the state vector are:

$$x = [\Delta E\ \Delta PN\ \varepsilon_s\ \varepsilon_B]^T \tag{3.32}$$

where the last two components represent the errors indicated in equations 3.23 and 3.25; ΔE y ΔPN are determined as follows [61]:

$$\begin{aligned} \Delta E &= PE_k - PE_{k-1} \\ \Delta N &= PN_k - PN_{k-1} \end{aligned} \tag{3.33}$$

The matrix of state transition Φ is given by:

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-\beta_s \Delta t} & 0 \\ 0 & 0 & 0 & e^{-\beta_B \Delta t} \end{pmatrix} \tag{3.34}$$

As can be observed in the transition matrix, the errors of the step length and of the azimuth have been modeled as Gauss-Markov processes [33] [29]. β_s and β_B are the time constants for each process, i.e. the error of the step length and the azimuth error.

In the EKF framework, the Jacobian matrix of the non-linear function f of the state model should be evaluated, as presented in A.2 by applying the equation A.19.

The observation model is represented as in A.16 and the observation vector is given by:

$$y = \left[\Delta\tilde{E}\ \Delta\tilde{N}\ \tilde{s}\ \tilde{\psi}\right]^T \tag{3.35}$$

$\Delta\tilde{P}E$ and $\Delta\tilde{P}N$ are calculated using the measures from the GPS as is indicated in 3.33, while $\tilde{s}$ is the measure of the step length obtained from the neural network and $\tilde{\psi}$ is the measure of the azimuth obtained from the IMU.

Continuing with the theory outlined in A.2, the non-linear function h of the observation model is defined as:

$$h = \left[\Delta E\ \Delta N\ S_{GPS} - \varepsilon_s\ \ atan\left(\frac{\Delta E}{\Delta N}\right) - B\right]^T \tag{3.36}$$

S_{GPS} is the step length, which is calculated from the GPS information, as the distance between two consecutive points:

$$S_{GPS} = \sqrt{(PE_k - PE_{k-1})^2 + (PN_k - PN_{k-1})^2} \tag{3.37}$$

Lastly the H matrix is given by:

$$H = \frac{\partial h(x)}{\partial x}\bigg\|_{x=\hat{x}^-} \tag{3.38}$$

The information processing algorithm is illustrated in Figure 3.13.

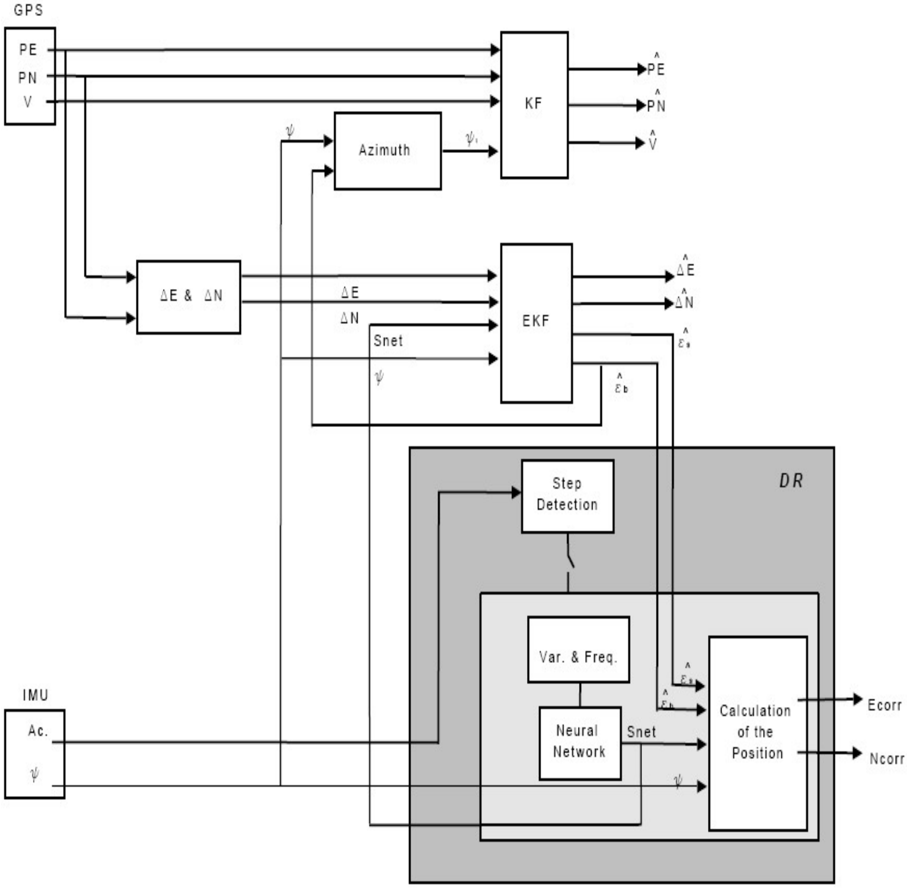


Figure 3.13: Data flow diagram for the outdoor system

3.4.3 Correction Mechanism

The Kalman filtering should be analyzed to ensure good functioning and thus to determine the validity of the results, which depends on several factors (see section A.3). The ideal situation is to be perfectly familiar with the system model, the parameters of which should have normal distributions; the true initial values for matrices R and Q should also be known.

First, in practice it is difficult to obtain a model that shows the behavior of the whole system and it is not always possible to include models of all the interferences that affect a system. Second, we often have very little or no knowledge at all of the values of Q and/or R. As a result, the statistical conditions of the variables may not be fulfilled, thus violating the basic principles of filtering.

Here, two aspects are considered in order to determine the validity of the filter. Which are based on the theory presented in Appendix A.3.

The analysis presented in Appendix A.3 deals with the concept of *consistency of the Kalman filtering* by studying *the innovations* defined in A.28, such as the difference between the predicted state by the filter and the measurement. In short, *the filter innovations should fulfill two conditions i.e. they should be white and have a mean of zero*. These conditions are evaluated with equations A.30, A.31 and A.32 and if their values are within the allowed limits, the Kalman filtering is considered to be acceptable. Nevertheless, if any of those conditions are not fulfilled, checks should be performed to seek the possible causes of such a result and the factors that are affecting the poor performance of the filtering should be corrected.

On the other hand, different techniques have been proposed to prevent these situations from occurring. Here, only some of them will be taken into account, while this topic is addressed in greater depth in [55], [62] and [12]. The techniques considered are: (1) scaling Q matrix, (2) the weighted Kalman filter, and (3) the χ^2 test.

Scaling Q matrix

This algorithm, which was originally proposed Hu et al. [63] and was adapted by Hide [64] as is indicated in the equation 3.39. This consists in simply applying a scale factor to the Q matrix. As this solution is based on the definition of weighting P^-, in fact what happens is that when applying the scale factor to Q P^- is also being scaled:

$$P^- = \Phi P \Phi^T + \lambda\,(Q) \tag{3.39}$$

If the scale factor λ is greater than 1, the new measurement will have more weight on the estimation.

Weighted Kalman Filter

This method has been applied in the field of navigation in robotics [48] [49] and was proposed by Anderson [65]. It addresses the case where R and Q are unknown and it assumes them to be exponential weighted values. By assuming these matrices in this way, in fact we are actually weighting the filter in general. The resulting equations for the filter, which have been adapted from [62], are then presented.

The matrices R and Q are assumed to be:

$$R_k = R\delta^{-2(k+1)} \tag{3.40}$$

$$Q_k = Q\delta^{-2(k+1)} \tag{3.41}$$

$$\delta \geq 1 \tag{3.42}$$

From [65], the weighted error covariance matrix P is defined as:

$$P_k^{\delta} \triangleq P_k \delta^{-2k} \tag{3.43}$$

After carrying out the corresponding transformations, the resulting equations for the prediction and correction phases are:

$$\hat{x}_k^- = \Phi_k \hat{x}_k \tag{3.44}$$

$$P_k^{\delta -} = \delta^2 \Phi_k P_k \Phi_k^T + Q_k \tag{3.45}$$

$$K_k = P_k^{\delta -} H_k^T \left(H_k P_k^- H_k^T + R_k / \delta^2 \right)^{-1} \tag{3.46}$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - H_k \hat{x}_k^- \right) \tag{3.47}$$

$$P_k = \left(I - K_k H_k^T \right) P_k^{\delta -} \left(I - K_k H_k^T \right)^T + K_k R_k K_k^T \tag{3.48}$$

The aim is to give more weight to the last measurements [62], thereby avoiding the use of erroneous values from previous estimations [66] and facilitating the process of obtaining a consistent filter.

χ^2 *Test*

This method consists in the continuous evaluation of the values of the normalized quadratic innovations, which can be evaluated for each variable or for a group of variables. Here, we perform an analysis for each variable, with the aim of making it easier to identify possible faults. The idea is to improve the results of the filtering by eliminating those erroneous values from the fusion.

The normalized quadratic innovations should be governed by the χ^2 distribution so that variables that do not fulfill this condition will be corrected and/or eliminated from the fusion.

3.5 Optimization

In Chapter 2 it was pointed out that in navigation systems for vehicles, robots, etc. a lot of research has been conducted in the area of data fusion optimization. Most research addresses the problem by combining evaluation techniques for Kalman filtering with AI techniques such as fuzzy logic or neural networks.

Many researchers use fuzzy logic to manipulate some or all of the parameters of the Kalman filtering in accordance with the criteria of an expert, the aim being to decrease the errors regarding the estimated position [47, 48, 49, 50, 67, 51]. Other proposals include fuzzification of the variables and training of the systems so as to introduce better references into the filter [52]. There are also proposals based entirely on fuzzy logic and/or neural networks, e.g. Hiliuta et al. [68] propose an adaptive neuro-fuzzy inference system (ANFIS). Most proposals agree that the most important issue is to avoid, as far as possible, failure in navigation systems due to incorrect estimations of positions.

3.5.1 Fuzzy Logic. General Aspects

Problem-solving can be accomplished using two types of knowledge, *subjective knowledge* and *objective knowledge*. The former is associated with the information that is difficult or impossible to model with traditional mathematical methods, while the latter is associated with well–defined formulations, e.g. mathematical models [69]. With an appropriate combination of those types of knowledge, we can expect to solve any problems in a process, especially when there is no well-defined mathematical formulation to describe the behavior of the system.

Subjective knowledge is widely used in engineering design, mainly in subsystems or systems of a process where decision-making is required. In practice, the expert is responsible for making decisions. Taking into account such aspects, L. Zadeh [70] developed a framework called *Fuzzy Logic*, which *" aims at modeling the imprecise modes of reasoning that play an essential role in the remarkable human ability to make rational decisions in an environment of uncertainty and imprecision "* [71]. For example, designers often use linguistics expressions like: "the pressure is high", "the temperature is lower", and fuzzy logic allows this type of expressions to be quantified and, at the same time, the process can be automated.

Appendix B.1 offers a summary of the theoretical foundations of fuzzy logic. These concepts include fuzzy sets, basic operations, membership functions, fuzzy rules and systems based on fuzzy logic. In fuzzy logic, the concept of the degree of compatibility is introduced. This concept is represented in membership functions, which have a range within the continuous real interval $[0,1]$ (see B.1). In other words, while in crisp sets we say $A \in B$ or $A \notin B$, in fuzzy sets we would say A has a $degree\ of\ compatibility = x$ (e.g. 0.8) in B.

Zadeh [56] also developed another framework based on fuzzy sets called the *Theory of Possibility*. This theory can be seen as a *non-classical uncertainty theory* [72] (which is also used in this thesis). It involves a similar concept to the membership function called *possibility distribution (π)*. The difference between π and the membership function consists in the fact that while the first works with the more or less possible values of a variable, the second requires a precise or exact information. Appendix B.2 includes a summary of the theoretical foundations of this framework.

Taking into account the points outlined above, regarding the tendency to optimize navigation systems for robust structures by combining Kalman filtering and fuzzy logic, here this philosophy is introduced for the case of UPSs. Possibility theory is suitable for UPSs because they usually have noisy and incomplete information. Moreover, a fault identification system could also be designed based on a proper evaluation of Kalman filtering. In this work, *"Causal Diagnosis"* was chosen to carry out that task. Causal diagnosis is based on the possibility theory and its theoretical foundations are presented in the following paragraphs.

3.5.2 Causal Diagnosis

Zadeh [56] developed a framework based on fuzzy sets called the *Theory of Possibility*. This framework focuses mainly on problems involving uncertainty, imprecision and incomplete information. Later, Dubois et al. [57] proposed a framework, which is based on the theory of possibility and causal relations [4] with the aim of finding faults or malfunctions in systems.

This section is adapted from [57]. In that work, Dubois et al. define the framework of causal diagnosis based on the theory of possibility under conditions of uncertainty. In this framework the problem of finding possible causes that describe the current observations is addressed.

Causal diagnosis (CD) is based on *causal knowledge* and for this reason the expert's knowledge is an essential part of this framework. Expert knowledge gives information about the more or less possible or impossible values for the parameters of a system. The information provided by the expert is processed in order to identify the possible causes of the current observations [73].

It is important to have a fault detection system that not only triggers the alarms but also helps to find the causes of the failure so that corrections can be carried out, if possible in real time. The general idea is to identify malfunctions by means of expert knowledge and observations (reading sensors).

There are three methodologies of diagnosis that depend on the nature of the knowledge that is represented [57]:

- *Model-based diagnosis*: Diagnosis is carried out by means of consistent models, since it is assumed that a model of the normal behavior model of the system is available. Therefore, inconsistencies are detected. It is useful, for example, to trigger alarms.

- *Abductive diagnosis*: The models of abnormal behavior of the systems are known. With this diagnosis is important to find the malfunctions that explain the observed manifestations. It is focused on finding particular faults.

- *Heuristic Diagnosis*: It is an ad hoc methodology which uses deduction to find failures. Examples of these systems are expert systems based on fuzzy logic.

Dubois et al. [57] propose using the concepts of Boolean causal relations, which describe the symptoms associated to each cause, as well as the symptoms that are not certainly caused. With this perspective the diagnosis problem is addressed in two ways: *diagnosis based on consistency and based on relevance*. The first method rejects the faults that are incompatible with the current observations, while the second method is focused on a small group of failures that *possibly* explain the current observations.

[4] *"Boolean causal relations can describe the set of symptoms for each cause which are certainly caused, as well as the set of symptoms which are not certainly caused"*[57].

3.5.2.1 Basic Concepts

Let us define a set of m possible failures $\mathscr{F} = f_1, \cdots, f_m$ and a set of n observable attributes $\mathscr{A} = A_1, \cdots, A_n$. The attributes can be Boolean, discrete or continuous.

As was mentioned before, CD uses the theory of possibility. In this case *the possibility distributions are defined as the restrictions of the more or less possible or plausible values from the attributes*.

The problem of diagnosis in CD refers to the disorders or malfunctions that may have caused the observed symptoms. This is dealt with from the point of view fuzzy causal relations. The simplest form of causal knowledge is given by binary relations linking causes directly with their symptoms. On the other hand, the fuzzy causal relations between causes and their effects are modeled with degrees of compatibility (membership degrees) attached to the observations or to the observed effects when a fault is present.

Formally, π_i^f denotes the possibility distribution that is a restriction of the values of the attribute A_i when a failure f is present. U_i is the domain of the A_i defined in the continuous real interval $[0,1]$. In this case π_i^f includes the knowledge about possible values of the attribute when a failure is present, so that the following is fulfilled:

$\pi_i^f(u) = 1$, means that it is possible to observe $A_i = u$ with the current failure.

$\pi_i^f(u) = 0$, means that $A_i = u$ is impossible to observe with the current failure.

The causal representation is made by fuzzy rules, as follows: IF *the failure f is present,* THEN, A_1 *is* a_1 *and ...* A_n *is* a_n.

The observations can be uncertain and imprecise and are defined as fuzzy sets, denoted by $\mathscr{O}$. These have an associated possibility distribution denoted by $\mu_{\mathscr{O}_i} : U_i \rightarrow [0,1]$, which defines the more or less possible values of the observations of the attributes.

When an attribute is not observed in the presence of a failure, this can have any value $U_i : \forall u \in U_i,\ \pi_i^O(u) = 1$. When a failure has no known effects on A_i then $\forall u \in U_i,\ \pi_i^f(u) = 1$.

Analysis based on Consistency This focuses on the detection of failures that are consistent with the observations. The extent to which the observations are consistent with the symptoms associated to a failure is defined as *the index of consistency*, denoted by *cons*:

$$cons_f = \min_{i=1,\cdots,m} sup\{min\left(\pi_i^O(u), \pi_f^i(u)\right) : u \in U_i\} \tag{3.49}$$

This represents the possibility of the presence of a failure with the current observations. In other words, the higher *cons* is, the more likely the presence of a failure will be; conversely if $cons_f = 0$ then at least one of the attributes is not compatible with whatever could have caused f.

With this point of view, the objective is to reject failures that are inconsistent with the observations.

In some circumstances the index of consistency is not enough to clearly discriminate the failure that is the consistent with the observations. Sometimes this is due to a lack of knowledge about the effects on the attribute. When this happens, it is necessary to apply a more restrictive index.

Analysis based on Abduction The relevance of a failure is established when it is sure that the effects are consistent or coherent with the current observations. That is to say, refining the failures obtained in the analysis of consistency should make it possible to establish the failure or failures that certainly explain the current observations. This is obtained by calculating *the index of relevance*, denoted by *rel*:

$$rel_f = \min_{i=1,\cdots,n} inf\{\pi_i^O(u) \rightarrow \pi_i^f(u) : u \in U_i\} \tag{3.50}$$

The most suitable method to solve 3.50 is to use the Dienes' implication [73]. The solution is $a \rightarrow b = max(1-a,b)$, whose value is 0 only when $a = 1$ and $b = 0$. On the other hand, $rel = 1$, if and only if each fuzzy observation is within the core[5] of the corresponding symptom, thereby ensuring that all the symptoms of a failure are present. In consequence, the analysis rejects a failure, if and only if it is incompatible with the observed symptoms.

Implication 3.50 fulfills the condition that $cos_f \geq rel_f$, so this analysis only takes into account those failures that are certainly present in order to explain the symptoms.

When the index of consistency clearly identifies a failure and the index of relevance is 0, we can say that the consistency analysis is sufficient. When the index of consistency is 0 for all failures, it means that the knowledge is incomplete and there is an undefined failure or it is normal behavior of the system [74].

Finally, Figure 3.14 shows how the concepts of consistency and relevance are used to discriminate failures.

3.5.3 Methodology for Fault Detection and Correction

Using Kalman filtering as the method to carry out data fusion means that a number of aspects must be taken into account: *"(1) the filter is designed for random processes, therefore, processes with other characteristics have to be analyzed carefully; (2) the Kalman filter in its more primitive form does not adapt to changing conditions and does not have a self-learning mechanism"* [12].

[5] Set of elements with a degree of compatibility equal a 1.

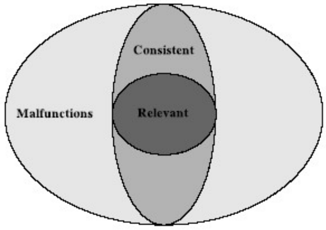


Figure 3.14: Scope of the concepts of consistency and relevance in casual diagnosis using the possibility theory. Figure from [10]

The Kalman filter is an optimum estimator if the error of the sensors is modeled properly. Thus the performance filter is affected when error characteristics of the sensor change with respect to the error models. Likewise, the noise of the measurements and the process must be white and with a mean of zero. As was mentioned in the section dealing with correction mechanism (see Section 3.4.3), the consistency test has to be carried out to determine these conditions.

In practice, the environment conditions where UPSs work are usually variable, and so it is important to take steps to guarantee the quality of the data fusion. Hence, a two-phase an evaluation–correction methodology is proposed. The first phase tries to avoid the introduction of measurements with higher errors (e.g., big reflections in the signal) or measurements that, according to the empirical knowledge, are not appropriate. This is accomplished by applying the principles of causal diagnosis and a correction subsystem is also included. The second phase consists in applying Kalman filtering evaluation mechanisms. These mechanisms are combined with the principles of fuzzy logic. As in the first phase, a correction subsystem is also implemented.

Diagram 3.15 shows the phases of the methodology. As can be seen, the diagnosis subsystem is responsible for detecting failures or anomalies. When a failure is detected, a correction subsystem is called, before introducing information into the Kalman filter.

The filtering evaluation subsystem is responsible for detecting inconsistent data; when the subsystem detects inconsistent data, the correction subsystem is called. In the following, each subsystem of the methodology is described.

3.5.3.1 Subsystem based on Causal Diagnosis

The causal diagnosis framework is considered a suitable method for identifing different behaviors of the measurement system, which would also allow an evaluation of the measurements to be carried out before introducing them into the Kalman filter. Some of the characteristics of causal diagnosis that make it suitable for the measurement systems are:

- It allows advantage to be taken of the empirical knowledge of the systems.

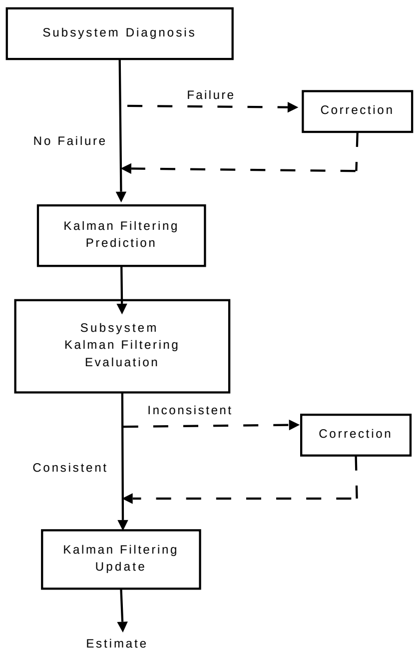


Figure 3.15: Diagram of the evaluation–correction methodology

When we have information about the more or less possible values of the parameters, it becomes possible to define possibility distributions that include those values. For example, the velocity of a user —while walking in a standard manner — can be defined within a range.

- It makes it possible to work with different types of knowledge, that is to say, with complete or partial knowledge or total ignorance about some attributes. This allows possibility distributions to be defined when there is no certainty about the values of an attribute (under certain contexts).

- The levels of uncertainty of the attributes are taken into account. This makes it an appropriate method for systems that have to work in noisy environments, such as the GPS.

- It allows normal behavior and abnormal behavior to be defined. Abnormal behavior corresponds to those situations where a failure or an anomaly is present.

- It is considered a good method for discriminating failures because it uses two criteria to select a failure: consistency and relevance.

- It is not a method that involves complex operations; the most common are maximums and minimums. Therefore, its computational costs should not be high.

The DC subsystem is used to identify states in the measurement system. This system is made up of the GPS and the step detection algorithm. These components are selected because there is enough empirical knowledge to apply it within the DC framework. Additionally, a correction subsystem is implemented to make the appropriate adjustments where a failure is present.

The basis design procedure of this subsystem is described as follows:

1. The attributes of the system are defined.
2. The sets of observations are defined with their respective levels of uncertainty.
3. The possibility distributions are defined for each attribute.
4. The normal and abnormal behaviors are defined.
5. The consistency and relevance indexes are calculated to identify the state of the system.
6. Once the state has been determined, the correction subsystem is called if there is a failure state.

Attributes

The attributes are defined as the parameters obtained from the GPS measurements and the step detection algorithm, and are: $step$, velocity (V), difference in East (ΔE), difference in North (ΔN), step length (SL_{GPS}) and azimuth (ψ_{GPS}). SL_{GPS} and ψ_{GPS} are obtained as indicated in 3.37 and the first term of the component $(1,4)$ of vector 3.36, respectively; the difference here is that they are calculated using the measurements from the GPS, not according to their estimations.

Observations

Observations are the intervals between the uncertainties of the parameters.

$step$ is a binary attribute whose values are 0 and 1 and, therefore, it has no uncertainty. For the rest of the parameters, the observations for each one are represented by:

$$\mathscr{O} = [O_j - \ell_j, O_j + \ell_j] \tag{3.51}$$

where j is the observed parameter (e.g.V, ΔE, etc.) and ℓ is its level of uncertainty. ℓ may be defined by an expert or by using a partition or the standard deviation; in this thesis the second option was chosen. The standard deviation was obtained from several tests performed under normal conditions.

Possibility Distributions

The forms of π for the attributes were: singleton, triangular, ascending and descending, and trapezoidal.

The first attribute considered is $step$. Due to its binary characteristics in this case two singleton possibility distributions are assigned: *SD* y *SND*. *SD* represents the event *step occurrence*, so when $step = 1$: $SD = 1$. *SND* represents the event *no step occurrence*, so when $step = 0$: $SND = 1$. This is shown in Figure 3.16(b). As a fuzzy rule: "IF $step = 1$ THEN $SD = 1$ and $SND = 0$", or, "IF $step = 0$ THEN $SD = 0$ and $SND = 1$".

The linguistic labels assigned to the attributes V, ΔE, ΔN, SL_{GPS} and ψ_{GPS} in normal conditions are: *Valid, Normal and Acceptable*. *Valid* corresponds to the first three attributes, while *Normal and Acceptable* are assigned to SL_{GPS} and ψ_{GPS}, respectively.

The previous conditions that are needed to define the possibility distributions are: (1) the user is in movement and walks forward; (2) the user walks with a normal and a fast frequency; and (3) the possibility distributions of Figure 3.16(c) are defined for the case when the user walks in one of those directions, that is to say, those possibility distributions are only applied to the parameter associated with the direction of the user (East or North).

Once the more or less possible values of the attributes have been fully identified, under normal conditions, the next step is to define the behavior under other conditions. By defining other possibility distributions it becomes possible to establish gradual changes between the normal behavior and abnormal behavior. This is achieved by defining the following distributions: *slightly high* and *slightly low* for intermediate conditions, while for abnormal behavior, the following are defined: *high not valid* and *low not valid* for ΔE y ΔN; *low abnormal* y *high abnormal* for SL_{GPS} and *high unacceptable* and *low unacceptable* for ψ_{GPS}. All distributions are illustrated in Figure 3.16.

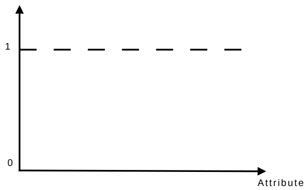
States

In the original CD methodology, the concept of disorders set is used; nevertheless, in this thesis the concept used is *state*. The intention is to define states for normal behavior and states of failure for abnormal behavior. Within the framework of CD, the disorders are defined by taking into account a set of manifestations of the observable attributes, which are established by means of possibility distributions. Here the same procedure is adopted to define the states, that is to say, the possibility distributions (illustrated in Figure 3.16) of each attribute will be combined to define each state. From this, the state of completely normal behavior of the system is established as follows:

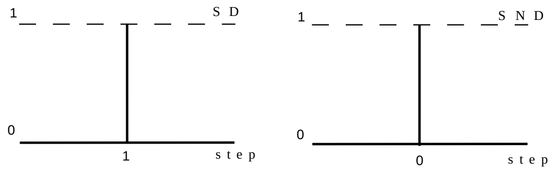
e_1: "$step$ is *SD*" and "V is *Normal*" and "ΔE is *Valid*" or "ΔN es *Valid*" and "SL_{GPS} is *Normal*" and "ψ_{GPS} is *Normal*"

To define the states, the following aspects were taking into account:

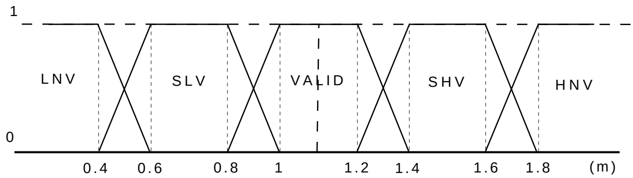
- If there is no detection of the step, that is, $step = 0$ which implies that π_{SND} is



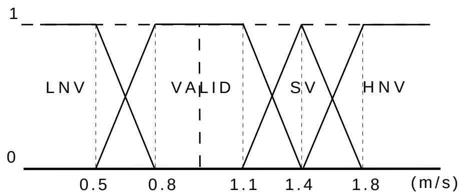
(a) $\pi_{attribute}$ for the label *unknown*



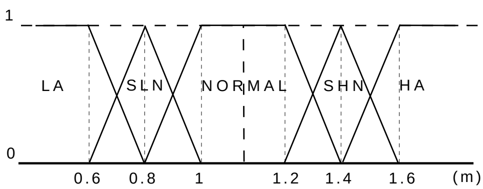
(b) π_{SD} and π_{SND}



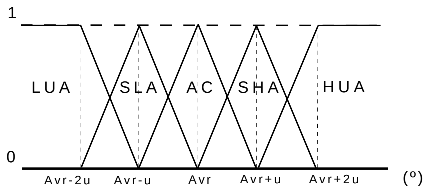
(c) $\pi_{\Delta E}$, $\pi_{\Delta N}$. Notations: LNV: Low Not Valid; SLV: Slightly Low Valid; SHV: Slightly High Valid; HNV: High Not Valid



(d) π_V. Notations: LNV: Low Not Valid; SV: Slightly Valid; HNV: High Not Valid



(e) π_{SL}. Notations: LA: Low Abnormal, SLN: Slightly Low Normal, SHN: Slightly High Normal; HA: High Abnormal



(f) π_ψ. Notations: LUA: Low Unacceptable, SLA: Slightly Low Acceptable, Ac: Acceptable, SHA: Slightly High Acceptable; HUA: High Unacceptable

Figure 3.16: Possibility distributions for the attributes

activate and if V is *Normal*, there is a failure state in the step detection algorithm or in the data reception of the IMU.

- When ΔE or ΔN is SHV or SLV, a total ignorance of the attributes SP_{GPS} and ψ_{GPS} is established. In consequence, the possibility distributions for such attributes take unity as the value of the degree of compatibility (see Figure 3.16(a)). It is defined in this way because neither ΔE or ΔN belongs to the label *Valid*, which means that the effects on the parameters obtained from them are unknown.

The failure states are classified as “main failure sates and secondary failure sates”. The main failure states are those in which the GPS parameters used in the Kalman filters have the following labels: *HNV* and *LNV* and are called ΔE failure, ΔN failure and V failure.

The following are examples of the defined state:

e_m: “*step* is *SD*” and “*V* is *Slightly Normal*” and “ΔE is *Unknown*” and “ΔN is *Valid*” and “SL_{GPS} is *Unknown*” and “ψ_{GPS} is *Unknown*”.

e_{m+1} : “*step* is *SD*” and “*V* is *Normal*” and “ΔE is *Slightly High Valid*” and “ΔN is *Unknown*” and “SL_{GPS} is *Unknown*” and “ψ_{GPS} is *Unknown*”.

ΔE Failure: “*step* is *SD*” and “*V* is *Normal*” and “ΔE is *High Not Valid*” and “ΔN is *Unknown*” and “SL_{GPS} is *Unknown*” and “ψ_{GPS} is *Unknown*”.

Step Detection Failure: “*step* is *SND*” and “*V* is *Normal*” and “ΔE is *Valid*” and “ΔN is *Unknown*” and “SL_{GPS} is *Normal*” and “ψ_{GPS} is *Acceptable*”.

Calculation of the indexes *cons* and *relv*

To calculate the consistency (*cons*) and relevance (*relv*) indexes, it is neccesary to ensure the observations are members of the sets defined with the possibility distributions (see Figure 3.16). Membership of an observation to a set is defined when the degrees of compatibility are above 0.5. The following example illustrates the calculations performed to obtain *cons* and *relv*.

Let us suppose that the observation of ΔN belongs to the set with the label *Slightly High Valid (SHV)*. First, *cons* is calculated. To obtain it, Equation 3.49 has to be resolved for one simple attribute:

$$cons(O_j, \pi_m) = sup\{min\left(\pi_j^O(u), \pi_m^j(u)\right) : u \in U_j\} \tag{3.52}$$

where j refers to the attribute (e.g. ΔN), m refers to the possibility distributions (also, called manifestations) over the respective label (e.g. SHV):

$$cons(O_{\Delta N}, \pi_{SHV}) = sup\begin{cases} min\left(\pi_{O_{\Delta N}}(x_1), \pi_{\Delta N}^{SHV}(x_1)\right) \\ \\ min\left(\pi_{O_{\Delta N}}(x_2), \pi_{\Delta N}^{SHV}(x_2)\right) \end{cases} = p_2 \tag{3.53}$$

From the observation sets (see Equation 3.51), it is obtained that:

$$x_1 = (\Delta N_k - \ell_{\Delta N}),\ x_2 = (\Delta N_k + \ell_{\Delta N}) \tag{3.54}$$

This calculation is performed for each attribute, and finally Equation 3.49 is applied to discriminate the current state.

It is possible to have several states with $cons = 1$, which implies that $relv$ must be calculated. The index of relevance is obtained by resolving Equation 3.50 for one attribute, as follows:

$$rel(O_j, \pi_m) = inf\{\pi_j^O(u) \rightarrow \pi_j^m(u) :\ u \in U_j\} \tag{3.55}$$

$$relv(O_{\Delta N}, \pi_{SHV}) = inf \begin{cases} max\left((1 - \pi_{O_{\Delta N}}(x_1)), \pi_{\Delta N}^{SHV}(x_1)\right) \\ \\ max\left((1 - \pi_{O_{\Delta N}}(x_2)), \pi_{\Delta N}^{SHV}(x_2)\right) \end{cases} = p_1 \tag{3.56}$$

As for the index of consistency, $relv$ must be calculated for each attribute, in order to finally resolve Equation 3.50. Thus, the state of the system is determined. Figure 3.17 shows the interval observations, and $cons$ and $relv$.

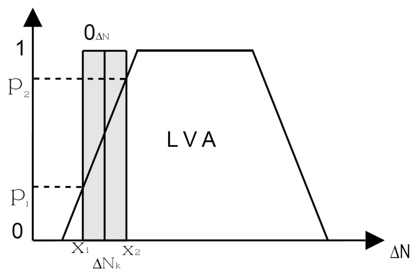


Figure 3.17: Example of calculations of the consistency and relevance indexes

Correction Subsystem

In accordance with the observation models of the Kalman filters, the parameters required from the GPS are: PE, PN, ΔE, ΔN y V. Therefore, in keeping with Diagram 3.15, if the discriminated state is a failure state for ΔE or ΔN and/or V, then this means that it is necessary to carry out adjustments to introduce the data from the observation vectors of the Kalman filters.

The parameters are continuously being evaluated and the system always stores the last values labeled as “Valid”. This is useful to gain information about the average from each of the parameters and so be able to use it when the state of failure occurs.

As was mentioned earlier, there are two types of failures: main and secondary. The main failures are associated to the parameters required by the Kalman filters, while the secondary failures are associated with SP_{GPS} and ψ_{GPS}. Step detection failure is considered to be a main failure, since there is something wrong with the detection algorithm or with the sensor.

In the case of failure of ΔE or ΔN, the average —obtained from the last N epochs to the current instant— is used to corrected the coordinates. From this, the current coordinate is defined as: $(PE_{k-1} + \Delta E_C)$ or $(PN_{k-1} + \Delta N_C)$ (the subscript C refers to corrected value). Regarding the other parameters, they simply take the value of the calculated average.

SL_{GPS} and ψ_{GPS} are stored in order to have that information available any time; it could be useful, for example, when other systems fail. To correct ψ_{GPS} the following must be taken into account: (1) the average is calculated while the user continues in the same direction and the gyroscope does not detect any turn, and (2) if there is a direction greater than 90^o, the average is ignored, because it is not valid for the current instant. For other trajectory forms (e.g. curves) other possibility distributions are required.

3.5.3.2 Filter Evaluation Subsystem

The methodology developed in this subsystem combines concepts from: (1) consistency test for Kalman filtering, (2) fuzzy logic, and (3) adaptive methods for the Kalman filtering. The consistency test is described in Appendix A.3. In short, it consists in calculating the normalized innovations squared (NIS) (see A.30) and the autocorrelation function (see Equation A.31). From (2) the concept of the membership function is used and in (3) the matrices R y Q are adjusted taking into account concepts of the adaptive methods for the Kalman filtering.

The way to apply the consistency test (also known as χ^2 test) was outlined in Section (3.4.3). This consists in evaluating the normalized innovation squared (denoted by ι_k) and the autocorrelation function (ρ_j) would be evaluated for each component of the innovation vector (ϑ) (that is, individually) for each filter. For example, for the EKF there are 4 components, since the innovations to be evaluated are: ΔE, ΔN and the bias and the step-length errors. Hence, there is a one degree of freedom. The χ^2 test consists in evaluating whether the NIS for η degrees of freedom is within the confidence limits, which is usually 95% or 97% of the χ^2 distribution. The autocorrelation function must fulfill $\rho_j \in \pm 1.96/\sqrt{N}$ [75] (N is the number of samples).

To evaluate the NIS, the confidence limits of the χ^2 distribution are defined in a gradual way. This is accomplished by specifying a membership function as illustrated in Figure 3.18. The interval (which ranges from 90% to 99%) gradually changes between the degrees of compatibility $[1,0]$.

It was also pointed out that the matrices R and Q could be adjusted to improve the filter performance, in case it tends to diverge. Therefore, this is taken into account in the correction subsystem in the second phase (see Diagram 3.15).

A technique based on scaling Q was introduce into the correction mechanism. To obtain the scale factor for Q, Hu et al. [63] proposed the following equation:

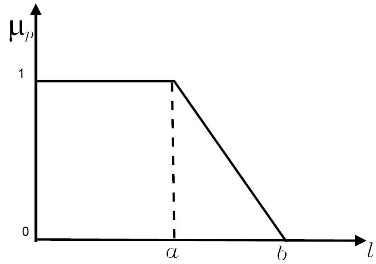


Figure 3.18: Membership function to evaluate the NIS. μ_p represents the membership functions for each parameter p (e.g. ΔE); a and b represent the confidence limits (90% and 99%) of the χ^2 distribution for η degrees of freedom

$$E_Q = \frac{\vartheta_k^T \vartheta_k}{\frac{1}{N} \sum_{j=k-N+1}^{k} \vartheta_j^k \vartheta_j} \tag{3.57}$$

where N is the number of previous epochs to the current instant k. From this, Hide et al. [64] performed an adaptation that is the one assumed for this thesis:

$$E_{Q_{p_k}} = \frac{\vartheta_{p_k}^T \vartheta_{p_k}}{fct_p} \tag{3.58}$$

where fct is an empirical value obtained in accordance with the expected sum square of the innovations [64]. p represents the component of the innovation vector to which the scale factor will be applied.

Matrix R components are adjusted as follows [76]:

$$R_{p_k} = \frac{1}{N} \sum_{j=k-N+1}^{k} \vartheta_{p_j} \vartheta_{p_j}^k - H_{p_k} P_k^- H_{p_k}^T \tag{3.59}$$

where N is the number of the previous samples and ϑ corresponds to the component of the innovation vector associated to the parameter p.

Finally, a general diagram showing the subsystems from the two phases is illustrated in Figure 3.19.

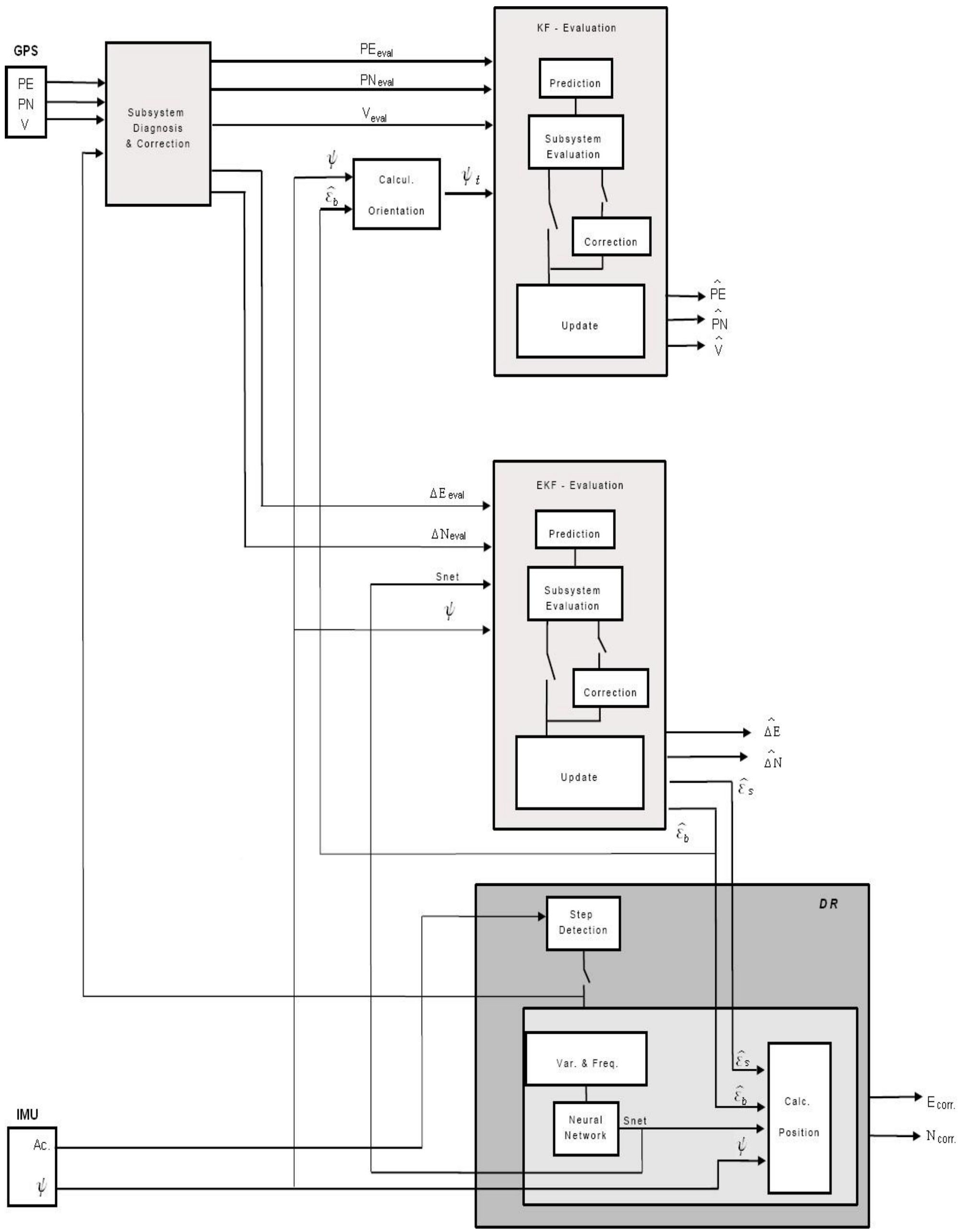


Figure 3.19: Diagram of the system with fault detection and correction methodology

3.6 Summary

First, the sensors used to define the UPS for indoor environments and the UPS for outdoor environments were described. Then, the dynamics of the user while walking was studied to determine his position with the DR algorithm. The algorithms to detect the step and to calculate the step length were also presented. Step detection is based on the patterns of the vertical and horizontal linear accelerations while the user walks, and the step length was obtained with a neural network, whose inputs are the step frequency and the acceleration variance. The results of the network were summarised for different velocities of the user: slower, slow, normal, fast, and faster.

Later, the location algorithms for a system for indoor environments and for a system for outdoor environments were defined. The first system is a combination of UWB-IMU and the step length. The second system is a combination of GPS-IMU and DR. The models for the Kalman filtering (as a method of data fusion) were then defined.

For the UWB-IMU system we used a discrete Kalman filter (KF), whose state model is similar to the DR equations, but in this case instead of the accumulative distances, the step length is used to improve the estimated position (X,Y).

For the GPS-IMU-DR system, two Kalman filters were defined: the first is a KF and the second one is an extend Kalman filter (EKF). The KF estimates the position (E,N) and the velocity of the user. The EKF estimates ΔE, ΔN, and the bias and the step length errors. The models of the EKF were defined in such a way that the information from the GPS can be used to correct the sensor bias so as to obtain the right azimuth. At the same time, if the neural network has problems in calculating the step length, the model allows the errors to be estimated based on information from the GPS. Kalman filtering is used for two purposes: (1) to have additional information about the position, and (2) to correct the DR parameters. The final position is determined with the DR and when this has problems, the position can be determined with the KF.

On the other hand, the filter must be evaluated to guarantee the quality of the information and the proper performance of the filter. For this reason, a summary of corrective methods (very often used in navigation systems) is presented. This includes methods that are considered suitable for this thesis, such as: scaling Q, weighting the filter and using the χ^2 test to evaluate each component of the innovation vector. Matrices Q and R affect the performance of the filter, which is why the methods presented here take into account the fact that, in practice, it is difficult to know their right values, and this in turn means that these matrices have to be tuned.

An entire methodology for evaluating and correcting the outdoor system information is presented. This consists in a complete analysis of everything from the measurement system to the output of fused data. This methodology has two phases: the first consists in evaluating a measurement system made up of the GPS and the step detection algorithm, while the second consists in tracking the filter to avoid the divergence.

The principles of Causal Diagnosis are used to evaluate the measurement system. These allowed a diagnosis system to be designed to identify normal and abnormal states. The diagnosis system determines the state of the measurement system by taking into account two fundamental criteria from casual diagnosis: consistency and relevance, which leads to a robust method to detect faults. When a failure is detected, the system applies corrective measures. The correction consists in using the average of the last valid data to correct or replace the current values. In general this prevents large errors from being introduced into the filter in accordance with the empirical knowledge.

In the second phase, tracking the filter involved concepts such as the χ^2 test, membership functions, and correction methods for Kalman filtering. First this subsystem evaluates the individual components of the innovation vector in accordance with the χ^2 test. The degrees of validity of the test are established in a membership function that contains the values for the confidence limits of the χ^2 distribution. When these parameters do not pass the test, a correction method is applied. Here the selected method was to tune the matrices Q and R by applying techniques such as *innovation-based adaptive estimation* [76].

CHAPTER 4

Analysis and Results

4.1 General Considerations

THE EVALUATION of each of the systems covered in Chapter 3 was performed with a similar analysis, and they therefore share some common aspects:

- ✓ The average NIS (A.31) is taken as a global observation criterion of the good performance of the filters.
- ✓ The innovation vectors were defined taking into account the information provided by each sensor [51], that is, an innovation vector was associated to each sensor (for the UWB-IMU system) and to each variable (for the outdoor system).
- ✓ For the NIS the degrees of freedom to be considered depend on the information provided by each sensor, e.g. for the UWB system there are 2 degrees of freedom, as it provides information about X,Y. The average NIS were defined by the number of samples M taken in each experiment and the information provided by the sensor, e.g. the UWB system would have $2 \cdot M$ degrees of freedom.
- ✓ In cases where correction mechanisms were used, such as scaling the matrix Q, scale factors were defined for each variable e.g. the scale factors for X,Y would be $E_{\Delta E}$, $E_{\Delta N}$.

The main prototypes are the UWB-IMU and GPS-IMU systems. For the first system an analysis was conducted that did not include the complete methodology, while for the second system the fault detection and correction methodology (section 3.5) was applied. In both systems some correction is needed regardless of the method used.

4.2 Indoor Environments

This section presents the experiments that were carried out and the results for the DR algorithm and for the system for indoor environments (see section 3.4.1), in addition to the analysis of those results.

4.2.1 DR Algorithm

This system is based on the model presented in section 3.4.1.1. The experiments designed to evaluate it were carried out on the ground floor of one of the University buildings.

The information needed to calculate the DR parameters was acquired by the IMU sensor system, which makes use of gyroscopes, accelerometers and magnetometers to obtain the acceleration, angular velocity and orientation values. The IMU was installed as illustrated in Figure 3.4 and the user was asked to walk along the halls of the building.

As a stand-alone positioning system, with DR there are no support systems or redundant information — just the IMU information. This is clearly illustrated in Figure 4.2, where the initial angle is correct but the system leads the user out of the building unless adjustments are made. This meant that several measures had to be taken in order to improve the results of the DR:

- The initial angle must be known. Due to the high degree of disturbance affecting the readings of magnetic sensors, it was not possible to know the initial angle by the IMU. Therefore, the initial angle was obtained using a calibrated electronic compass.

- Several known positions were defined. To make corrections along the path followed by the user, the corners where he was to turn were taken as reference points.

- The turns were identified by taking into account the angular velocity provided by the gyroscopes. While the user keeps a straight path the angular velocity does not undergo any major changes, but when the user turns its value increases considerably, as shown in Figure 4.1.

- To calculate the step length error: (1) expected distances are stored for each of the benchmarks, and distance traveled are calculated from the DR; (2) the difference between those distances is calculated to obtain the error; and (3) this is divided by the number of steps in order to obtain an approximate value of the step error, so that it can be included in the next steps.

- The error at the azimuth was evaluated in each line. Determining that error is a very complex task because there are a countless number of causes of interference inside the building, such as metal doors, pipes, and so forth. Modeling such variable errors in time and distance is very complicated and because there was no support system to correct the azimuth, the CAD information was used as an aid. This attempted to correct the course of the user's path.

- To facilitate the analysis, the user tried to walk in straight lines.

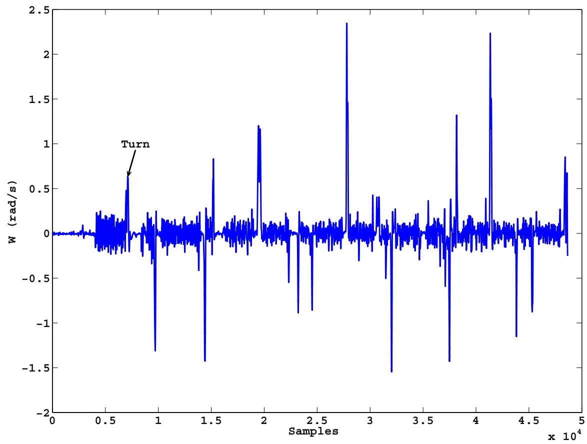


Figure 4.1: Angular velocity, while the user walks

The result of applying these corrective measures is illustrated in Figure 4.2. Because the only reference points were the corners and due to the high burden of electromagnetic disturbance inside the building, the route sometimes has big errors, some examples of which are indicated by the arrows on the chart.

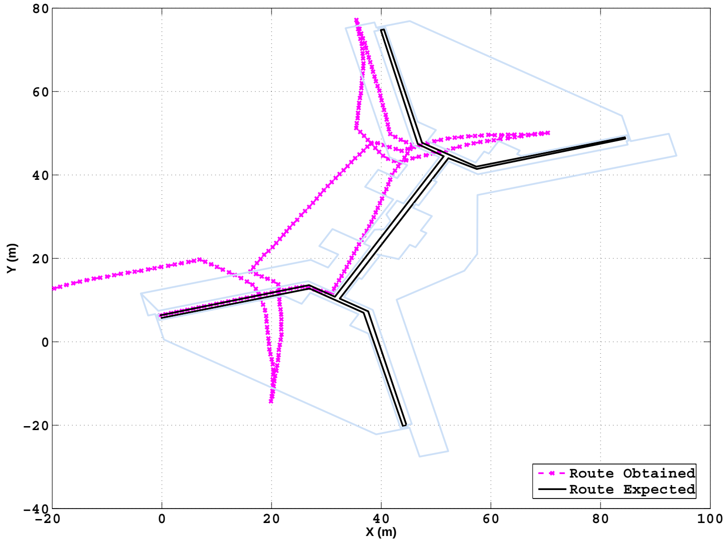
4.2.1.1 Discussion

The measures used here require databases with known positions to make continuous comparisons and this would call for other techniques, such as map-matching, which are not evaluated in this investigation. The results could be improved if the reference positions were closer together, for example, every 10 meters or less.

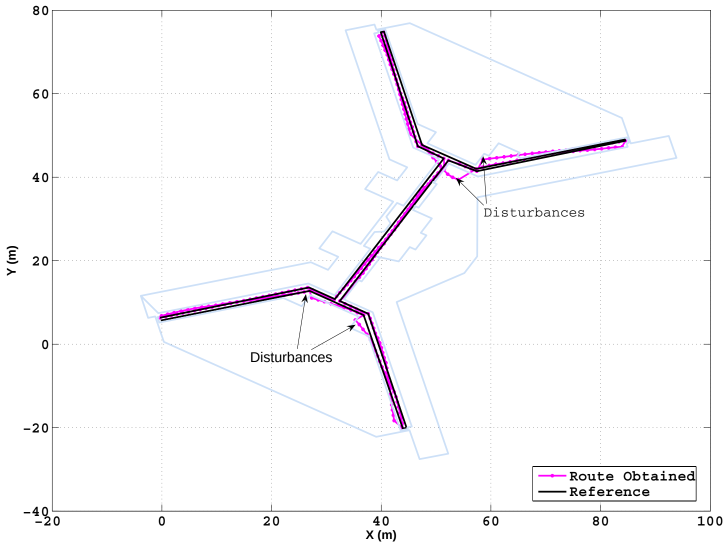
Hence, the calculations show the main advantage and main disadvantage of DR as the only positioning method. In other words, DR is a good support or redundant system to other more sophisticated systems, but it cannot work alone. An alternative may be to use the IMU in order to have an inertial navigation system (INS), but it has the same calibration problem because it requires an excellent modeling of the sensor bias which generates drift errors and can cause unacceptable results.

The step length with the proposed neural network provides a very low error, which ranges 0.1623% and 2.039% of the total distance traveled in several tests.

Determining the azimuth is a very complex issue requiring the combined systems, some designers suggesting the combination of a gyroscope and an electronic compass



(a) DR without corrections



(b) DR with corrections

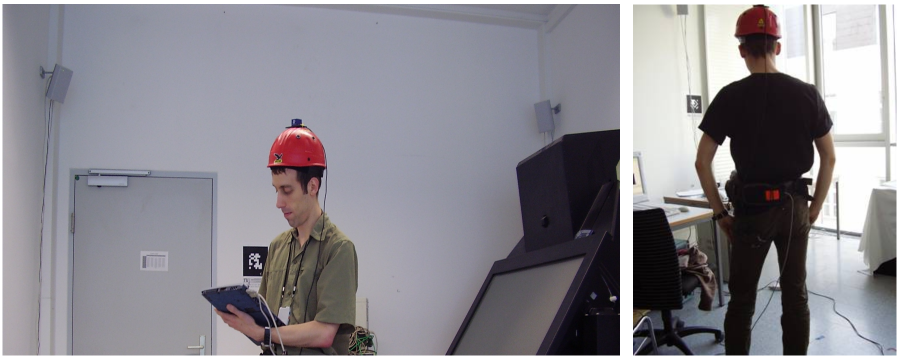
Figure 4.2: Results for the stand-alone DR

[33]. This is a feasible in environments that are free of disturbances so that the electronic compass can be calibrated and thus obtain appropriate values for modeling the bias of the gyroscope. When the compass readings are clearly affected, however, predicting the bias is no simple matter. Other research suggests using other technologies such as radar [36].

In the search for good accuracy, DR alone is not an appropriate solution. Therefore, more reliable, robust systems that help determine the azimuth in a direct or an indirect way and increase the accuracy of the measurements of the sensors are a *priority*.

4.2.2 UWB-IMU System. First Approach

The sensors setup for this system consisting of the UWB system and the IMU is illustrated in Figure 4.3. As can be noted, the user wears a belt on which the IMU is installed (4.3(b)) and he also carries an identification device from the UWB system around his neck (4.3(a)). Figure 4.3(a) shows two of the four beacons belonging to the UWB system installed in the room, which receive and transmit the signal from the device that the user takes through the network.



(a) Ubisense UWB System. (b) IMU arrangement.

Figure 4.3: Setup of the system that locates the user

The data were collected by a PC using the OpenTracker software and were later post-processed and analyzed with Matlab. OpenTracker was developed at the Vienna and Graz Universities of Technology, in Austria, and can be used in real-time applications. It defines the interface with the sensors by using modules implemented in C++. The various modules are called up with parameters defined in an XML file.

The room in which the experiments were carried out measured $4 \times 7\ m^2$ and, bearing this in mind, the experiments were set up as follows:

1. The user tried to keep still for between 1 and 3 minutes. This was performed repeatedly in different positions around the room.

2. The user was asked to walk along a predefined route.

To find the standard deviation of the UWB system, tests were carried out prior to defining the values of the covariances to be introduced into the Kalman filtering. On the other hand, IMU errors were assumed to be those provided by the manufacturer.

The following are the experiments that were conducted for this system using the models described in section 3.4.1.2.

4.2.2.1 Model A

First, the initial values of the error covariance matrices R and Q should be defined, as well as the priori estimation error covariance P. To define the components of the diagonal of the matrix Q —denoted by $\left[Q_x \; Q_y \; Q_{\varepsilon_\psi}\right]$— initial values of the first two elements were assumed to be $\left(10^{-3}, 10^{-3}\right)$, while the third element —Q_{ε_ψ}— was obtained by [12]:

$$Q_k = \sigma^2 \left(1 - e^{-2\beta\Delta t}\right) \tag{4.1}$$

where $\sigma^2 = 1$.

The initial values of the components of the diagonal of P were assumed to be $[1,\ 1,\ 1]$. To determine the initial values to the components of the diagonal of R —denoted by $diag\left[\sigma_x^2,\ \sigma_y^2,\ \sigma_{\varepsilon\psi^2}\right]$ — previous tests were conducted to observe the behavior of the sensors in the area. From this the most appropriate initialization values were: $\sigma_x^2 = 0.01664$, $\sigma_y^2 = 0.0942$, $\sigma_{\varepsilon\psi}^2, = 1$.

For the Gauss-Markov process that models the azimuth error, the value assumed for the time constant is high because there is no change in the environment [29] and the time interval Δt is given by updating the UWB system.

Simulations were then carried out for the model defined by equations 3.11, 3.10, 3.15 and 3.17. Since this model does not represent the movement of the user adequately, but rather situations in which there is an absence of movement by the user, experiments were carried out for fixed positions.

Five positions in the room were defined and data collected for each of them were the (X,Y) and azimuth. Subsequently these data were processed using the KF and evaluated according to the criteria set out in section 3.4.3 and Appendix A.3. The result was generally a consistent filter, that is, from a statistical point of view the innovations were correct. In other words, there are two conditions that the innovations must satisfy: being white and having of a mean of zero, which is evaluated by equations A.31 and A.32. To declare the filter as consistent, the autocorrelation function (A.32) had to be within the limits allowed and the average normalized quadratic innovations (A.31) were to be within the 95% confidence limits of the χ^2 distribution.

The NIS evaluation was performed separately for each of the sensor systems (as already mentioned in the general considerations). That is, two NIS factors were identified, one for the UWB system and one for the azimuth error. The degrees of freedom for the average NIS of the UWB system are $2 \cdot M$ and $1 \cdot M$ for the IMU. The maximum limits of the variables are indicated in Table 4.1. Moreover, the evaluation of the autocorrelation function was performed using the innovation vector.

NIS can be improved by manipulating the values of Q and R, since one of the most frequent causes of inconsistency of the KF is that the values of these matrices are inadequate. Since the consistency criteria are satisfied, there was no need to make any changes, especially in the matrix R. Nevertheless, in the "position 5" there filter is inconsistent. This was solved by scaling the matrix Q. The scale factor for each parameter was derived empirically and this technique was seen to improve the results of the autocorrelation.

Table 4.1 shows a comparison of the expected and obtained values before and after correction for both the average NIS and the autocorrelation values. This simply illustrates the consistency of the filter and in the few cases where this was not the case, consistency was achieved by applying the technique in question.

Table 4.1: Results of the χ^2 test for the UWB-IMU system in fixed positions

Item	NIS UWB			$NIS_{\varepsilon\psi}$			AutoCorr.		
	NoCorr.	Corr.	Conf. Lim. (95%)	No Corr.	Corr.	Conf. Lim. (95%)	No Corr.	Corr.	Limits
1	0.1542	0.1551	1.1064	0.6353	0.6353	1.1518	0.1058	0.1055	±0.1252
2	0.4920	0.4951	1.1101	0.2335	0.2335	1.1571	0.0575	0.0594	±0.1284
3	0.4802	0.2705	1.1078	0.2705	0.2705	1.1538	0.0615	0.0605	±0.1257
4	0.6068	0.6037	1.1129	0.7525	0.7525	1.1611	0.1187	0.1182	±0.1315
5	0.7019	0.7130	1.1260	1.8088	1.0557	1.185	0.0938	0.0779	±0.1465

Figures 4.4, 4.5 and 4.7 illustrate the signals from one of the positions. The functionality of the filter can be seen quite clearly by the presence of a smoother signal.

Figure 4.6 illustrates the quadratic innovations of the UWB system. No points are seen to go beyond the confidence limits (5.99) of the χ^2 distribution. Meanwhile, Figure 4.8 illustrates the quadratic innovations for the azimuth error. It can be seen how they decrease once the values of the filter have been adjusted.

In accordance with the tests performed the model selected for the azimuth error has a good performance, if any of the parameters of consistency are not fulfilled, acceptable values for these parameters are easily achieved with some minor adjustments.

Moreover, it was also noted how the filter operation depend on proper introduction of the measures. In cases where the sensors were most affected by large disturbances which could not be modeled, the changes in the variables of the proposed filter were drastic. In practice, for example, if the user is in a position where the UWB system has important problems involving signal reflection, the filter will not work properly if its parameters are not adjusted to this situation. This implies that such phenomena must be taken into account for a high level of functionality of the filter.

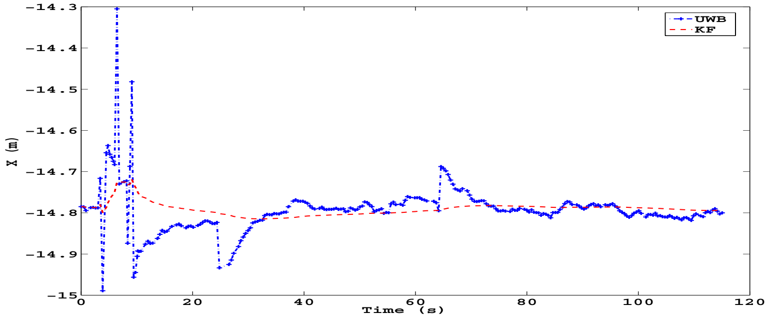


Figure 4.4: Results for the X coordinate

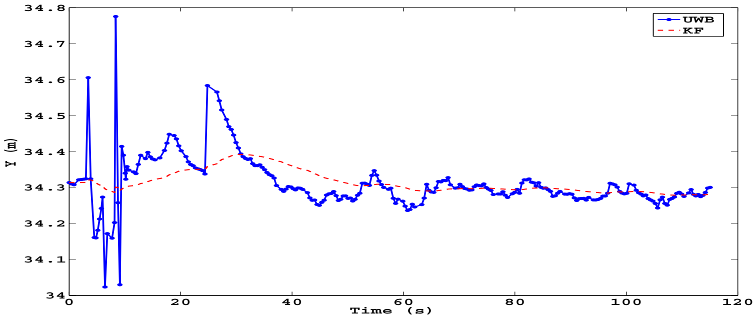


Figure 4.5: Results for the Y coordinate

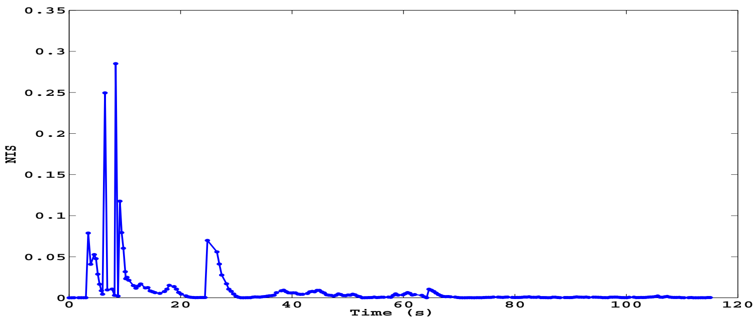


Figure 4.6: NIS for (X,Y)

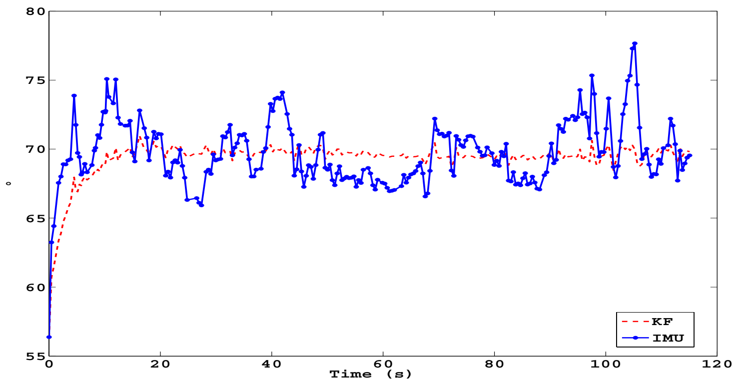
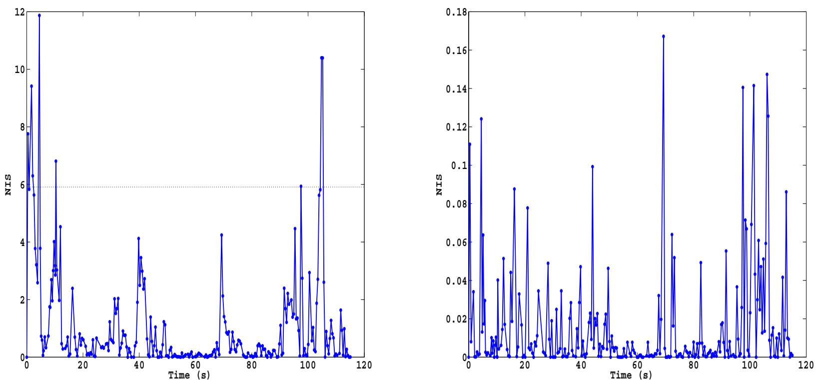


Figure 4.7: Results of filtering for the azimuth



(a) NIS with the initial assumed values.

(b) NIS after corrections

Figure 4.8: NIS for the azimuth error

4.2.2.2 Model B

This section addresses the second model proposed in section 3.4.1.2, which is represented by the set of equations 3.18, 3.19, 3.20 and 3.21. To observe the behavior of the model, the second phase of the proposed experiments were carried out, that is, the user was asked to walk around predefined points in the room. In this case the step length was known, and this was why the neural network algorithm was not applied, also bearing in mind that it was very short trip.

A similar procedure that described for model A was performed. The initial values for Q and R were assumed in accordance with the knowledge about their errors. This first resulted in non-consistent filter, that is, from a statistical point of view the innovations were incorrect. This means that either the autocorrelation function ((A.32) was outside the permitted limits or the average of the normalized quadratic innovations (A.31) were beyond the 95% confidence limits of the χ^2 distribution.

The methodology consisted of applying two correction techniques. The first manipulated the values of Q and R in order to identify the values that allow consistent filtering and the second was the same technique used to improve the performance of model A. This involves applying a scaling factor for Q, which in turn scales the matrix P^-.

Table 4.2 summarizes the results achieved by implementing the above techniques. “Test 1” refers to the implementation of the filters with the initially proposed values, “test”2 was carried out using the technique and for “test 3”, the second technique was applied.

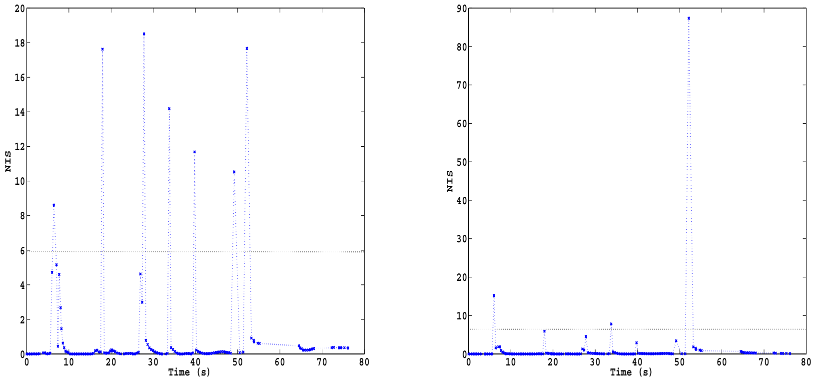
Table 4.2: Results of the χ^2 test for the UWB-IMU system when the user walked

Test	NIS UWB		NIS$_\psi$		AutoCorr.	
	Result	Conf. Lim. (95%)	Result	Conf. Lim. (95%)	Result	Limits
1	1.0923		49.5061		0.4630	
2	1.0241	1.1424	0.2287	1.2036	0.1615	±0.1651
3	1.1018		0.9146		0.1615	

At first, the innovations of the filter were not white and the NIS of the azimuth error were unacceptable. From this, the R values for each parameter were treated on an individual basis in the first technique, but also taking into account the fact that each value could affect the rest of the variables. In other words, X,Y depend on the value of the step length as well as the value of the azimuth. Since the azimuth error was a critical parameter, its component in Q had to be increased (in the order of 10^4). In the second technique, a Q factor had to be applied for the azimuth error of a similar order to that used in the first technique. The other factors changed according to the behavior of the variable; for X and s this factor was small, while for Y it was in the order of tenths. This led to increases, such as Q for Y and the azimuth, and decreases, as in Q for X and s. Regardless of the changes introduced in any of the techniques, the results of the average NIS were within expected values and white innovations were achieved.

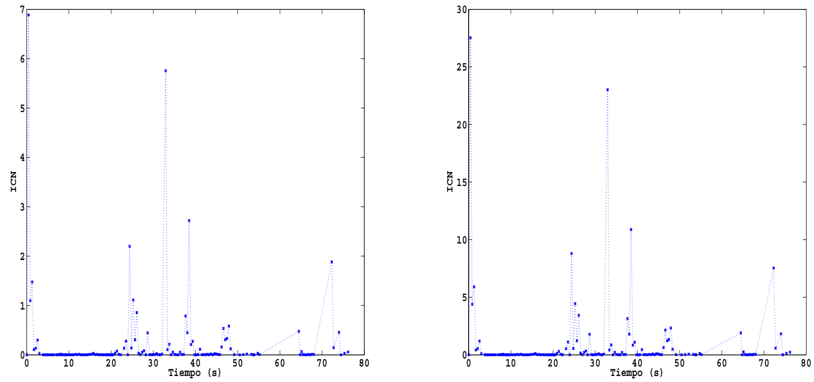
The NIS resulting at each instant for each technique are illustrated in Figure 4.9. On applying the first technique it was observed that there are more points beyond the limits than in the second technique. For azimuth error, there are some points outside the limits in both cases. This behavior is acceptable, as only a few points lie outside the bounds.

Finally, a comparison of the routes is shown in Figure 4.11. The corrected route was obtained with the first technique and there were no significant differences between the two techniques. It should be noted that, on improving the values of Q and/or R, mitigation of the reflections becomes obvious and this is particularly apparent at the ends of the line. This may be logical considering that location of the beacons was at the limit and the system might have trouble triangulating.



(a) NIS with changes in Q and R. (b) NIS scaling Q.

Figure 4.9: NIS for (X,Y)



(a) NIS with changes in Q and R. (b) NIS scaling Q.

Figure 4.10: NIS for azimuth error

4.3 Outdoor Environments

The system used for analysis in outdoor environments consisted of a GPS receiver and an IMU. In this case the user carries a laptop which the sensors are connected to via

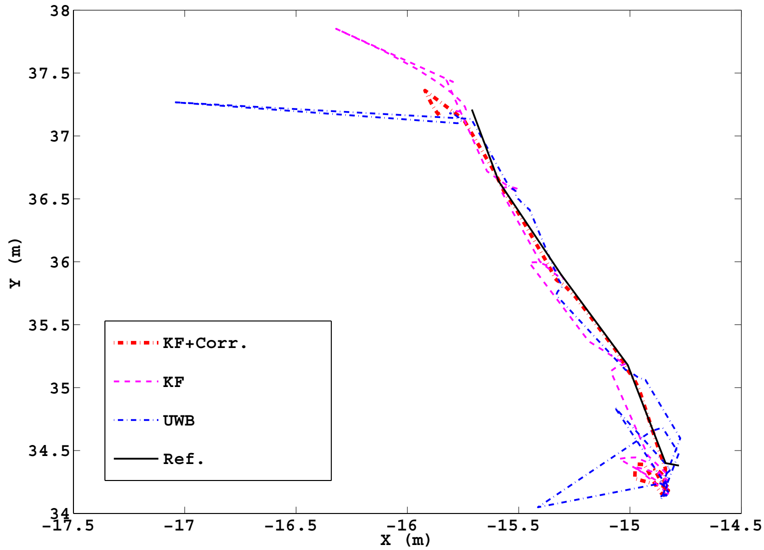


Figure 4.11: Comparison of results for the route by the user

RS-232 and USB, respectively.

The information obtained from the GPS is the latitude, longitude, speed, time and HDOP (Horizontal Dilution of Precision). The information provided by the GPS is in NMEA format and, before filtering, it was converted into the UTM (Universal Transverse Mercator) system, which gave the information in meters. The information that is obtained from the IMU is the same as that presented in the UWB-IMU system. This will therefore be an *GPS–IMU system based on the DR algorithm*.

To observe the behavior of the integrated system the user was asked to walk along a route, which was set up in one of the parking lots at the university. The route was approximately 335.2 m in length and consisted of straight lines and 90° turns. The information was collected by the laptop and then processed by MATLAB. In this system the information was processed by applying the following methods:

- Kalman filtering was applied as described in section 3.4.2 and the assessment methodology was the same as that used for the UWB-IMU system, that is to say, the criteria in Appendix A.3 were evaluated.

- The fault detection and correction methodology put forward in section 3.5 was applied.

In the following subsections, the analysis for each of those techniques is presented.

4.3.1 Kalman Filtering for the GPS-IMU System

As was indicated in the section 3.4.2, there are two sources of position information: from the KF or by DR. The KF estimates the E,N coordinates and the velocity of the user. In this thesis the final user's position will be determined by the DR algorithm, as shown in diagram 3.13. The DR parameters are corrected continually, this is why the DR is considered more reliable.

To properly determine the user's position with the DR, the models for correcting the step length and the azimuth were developed. Diagram 3.13 indicated that the EKF is responsible for obtaining the step length and the bias errors. Thus, this is a correction method for the DR parameters as proposed in [29]. Here, the parameters used for the correction are different to those employed in that research. In other words, while they suggest obtaining the step length error by evaluating the step period and getting the azimuth from the east and north speeds of the GPS, here the step length is obtained from the GPS as shown in equation 3.37. This information is then also used to calculate the azimuth, which is reflected in component (1,4) of vector 3.36.

To carry out the tests, first the initial values of Q and R for the filters (KF and EKF) were assumed. The initial values of Q did not allow a plausible functioning of the system, hence those values were adjusted by means of a trial and error process. Therefore the initial values of Q for each filter were assumed to be: $Q_{KF} = diag\left(10^{-}3\ 1\ 10^{-2}\right)$, $Q_{EKF} = diag\left(10^{-3}\ 10^{-3}\ 1.2422\times10^{-4}\ 0.0124\right)$. Regarding the values of R for V and the step length and bias errors were fixed in the tests of the filters. Meanwhile two techniques were employed to choose the values of the R for E,N, the first consisted in assuming values according to the presumed error of the GPS system, while the second was that applied in [29], which takes into account HDOP:

$$\sigma_E^2 = \sigma_N^2 = \left(\frac{HDOP}{\sqrt{2}}\sigma_\rho\right)^2 \tag{4.2}$$

where σ_ρ is the pseudorange error. This technique allows for additional information of the errors of the GPS. As regards the IMU, the errors are those suggested by the manufacturer. Nevertheless, R values also needed adjustments to obtain a plausible trajectory. Therefore the initial values for R were assumed to be: $R_{KF} = diag\left(3^2\ 3^2\ 0.05^2\right)$, $R_{EKF} = diag\left(0.1^2\ 0.1^2\ 0.1^2\ 10^2\right)$.

Regarding the tests, it was observed that the neural network had some problems when its conditions of training change. Even so, the step length provided by the network was within the range of values that corresponding to the user, but its errors were higher than the errors obtained with the original conditions of training —this could have been due to the slope of the path or the type of terrain. Then using the step length of the neural network (with these new conditions) would be similar to having a predefined step length, whose errors would be determined with the methodology proposed (see section 3.4.2).

The user walked along the route illustrated in Figure 4.12 and stopped at every turn —it can be observed that the user walked only few meters of the last line. As was mentioned, the idea is observing the performance of the filter with little knowledge of the values of R and Q. Therefore, in this section the fault detection and correction methodology was not applied, in order to observe the performance of the system. On other hand, there was a fault of the IMU at the end of the first line, therefore, for this situation, the step length was assumed to be the average of the last n values (e.g. $n = 10$).

Figure 4.12 shows the trajectories obtained, which were determined from: the GPS, the DR without any correction (including the bias, but not its errors) and the DR with its parameters corrected with the errors estimated by the EKF. The place —where the tests were carried out — is a typical outdoor area with some metal posts or cars, which could affect the functioning of the IMU. Those aspects and wrong parameters of the filters are sources of errors, which can lead to have inconsistent filters and wrong trajectories. For example, this can be observed in the first and third lines where the trajectories undergo a deviation. In the last line the GPS signal is unavailable, thus the position was obtained with the DR.

Figure 4.13, on the other hand, presents the azimuth obtained. This figure includes the azimuth obtained from the GPS, the one provided by the IMU and the corrected azimuth, the model of which is detailed in section 3.4.2. The bias is determined by initially assuming value of zero and then estimating it during the user's first steps; this bias is attributed to the offset of the body [17]. On the other hand, the turns were also checked by observing the angular velocity.

When the user is still, the system determined the azimuth taking into account the last errors estimated while the user walked. This is obvious, since while the user is still, the GPS information cannot be used in the models proposed. As the angle shows, the model for determining the bias is appropriate, but when the GPS information is wrong the results of the filter are affected.

Figure 4.14 presents the NIS for the innovations of the KF. In spite of data of the KF were not selected to determinate the final position, the validity of its results must be guaranteed, since it can be used as a last resort. As it is observed the velocity presents NIS with unacceptable values, in other words, several values fall outside the bounds —3.84 for 95% and 5.02 for 97%. This must be corrected because V affects the values of E and N.

Figure 4.15 illustrated the NIS of the innovations of the EKF. As it can see several points lie outside the limits, therefore the filter is inconsistent, despite the initial adjustments of the Q and R values.

According to the results, it is a system that requires adjustments to obtain a better trajectory and a consistent filter. In the following section the fault detection and correction methodology (see section 3.5) is applied to correct the results obtained in this section.

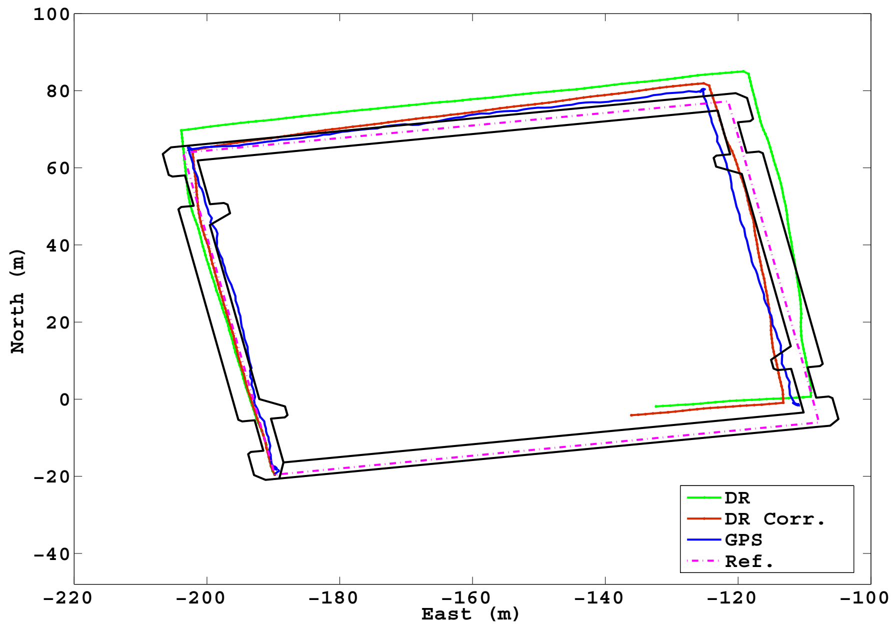


Figure 4.12: Comparison of results for the route by the user. The paths are obtained from GPS, DR without any correction, DR with its parameters corrected with the errors estimated by the EKF

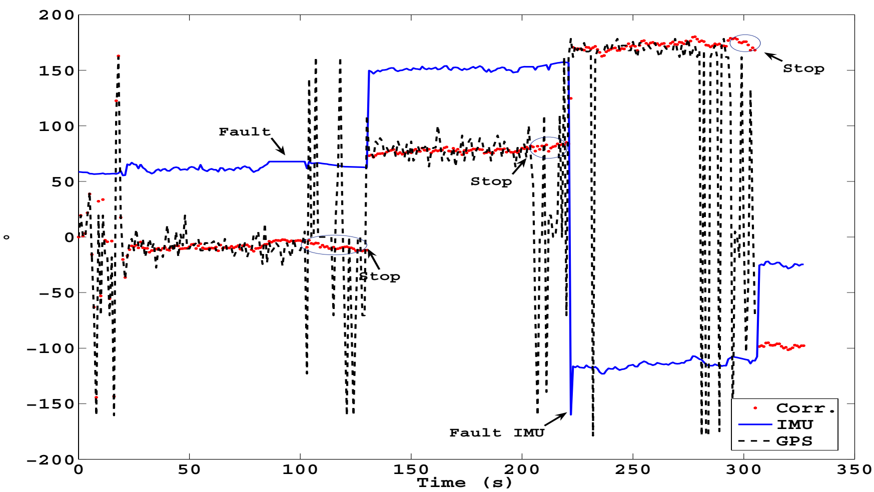


Figure 4.13: Azimuth obtained from: GPS, IMU and the corrected azimuth

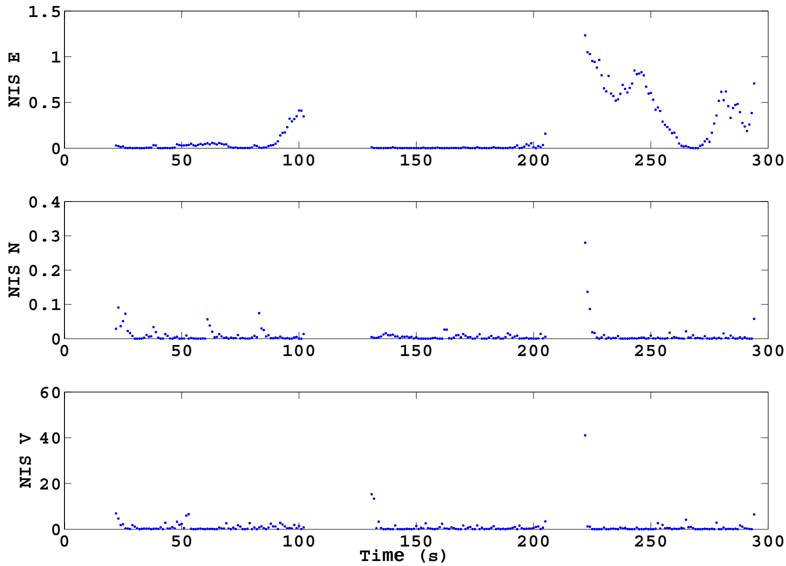


Figure 4.14: NIS obtained for the innovations of the KF

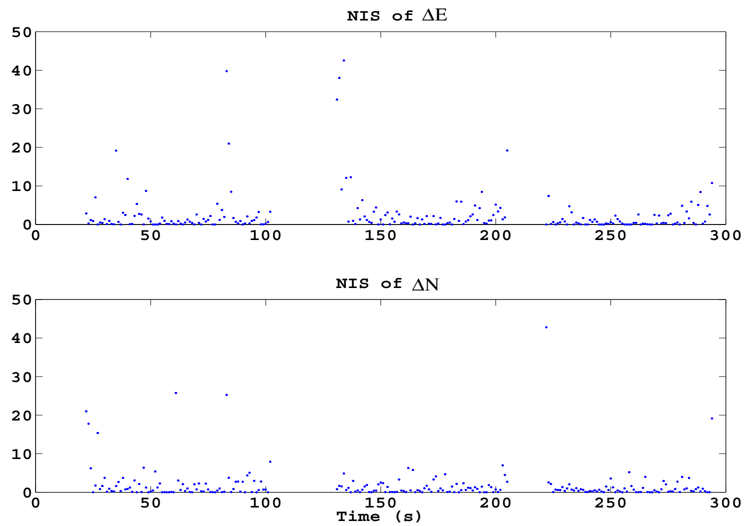
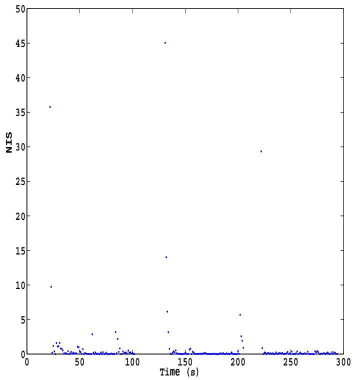
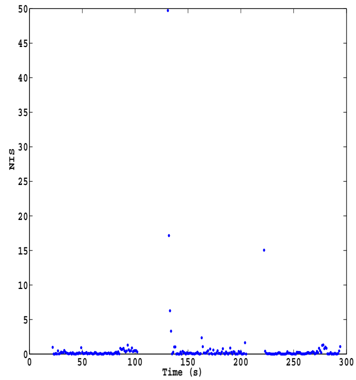


Figure 4.15: NIS obtained for the innovations ($\Delta E, \Delta N$) of the EKF



(a) NIS for the innovations of ε_s



(b) NIS for the innovations of ε_B

Figure 4.16: NIS obtained for the innovations of the errors of the DR parameters

Table 4.3: Results of the consistency tests for the KF and EKF for the GPS-IMU system

Item	NIS	Confid. Limit. (95%)	AutoCorr.	Lim. (95%)
E	0.1617	1.1585	0.9806	±0.1295
N	0.0087		0.8550	
ΔE	2.5012		0.1714	
ΔN	1.88651		−0.0196	
ε_s	0.8603		0.4672	
ε_B	0.6276		0.3877	
$Innov_{FK}$	1.1453		0.5443	
$Innov_{FKE}$	6.0770		0.1938	

4.3.2 Fault Detection and Correction Methodology

The following is the analysis for the same experiment (as presented in previous section), which was conducted using the methodology described in subsection 3.5.3.

Here, adjustments for the matrices Q and R were not carried out, therefore, the initial values of their components were assumed to be: $Q_{KF} = diag\left(10\ 10\ 10^{-3}\right)$, $Q_{EKF} = diag\left(10^{-3}\ 10^{-3}\ 1.2422\times 10^{-4}\ 0.0124\right)$, $R_{KF} = diag\left(3^2\ 3^2\ 0.05^2\right)$, and $R_{EKF} = diag\left(0.1^2\ 0.3^2\ 0.1^2\ 1^2\right)$.

In the first stage, failures in E, N and in the step detection were identified. At the end of the first line there was a problem with the reception of the data of the IMU, this introduced a 9% of failures, because there was not information about the azimuth or the step detection while the user even continued walking. When this failure appears the step length is assumed as the average of last values as reported by the neural network.

First the results obtained for the azimuth are illustrated in Figure 4.17. As can be observed the diagnosis system eliminated the peaks of the azimuth obtained from the GPS (presented in Figure 4.13). It was also taken into account that the user made a stop in each corner previous to make the turn and begin to walk again, this is why the azimuth in the stops is corrected with the last errors estimated from the EKF and the GPS data are not used, because as was mentioned in the previous section the GPS information is not adequate for this situation.

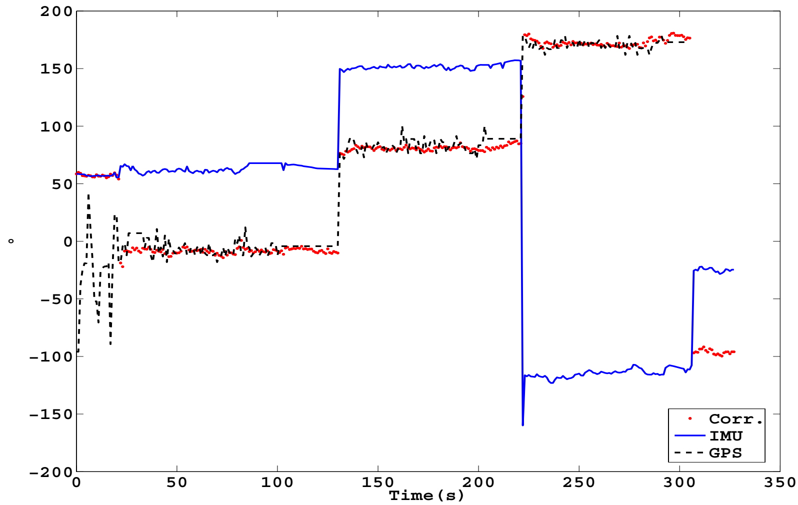


Figure 4.17: Azimuth obtained from: GPS, DR without correction and DR with the parameters obtained with the fault detection & correction methodology

The determination of the bias was not established until the user begins to walk; this can be seen at the beginning of the first line, whose angles are uniquely determined by the IMU. Once the bias is estimated, is corrected in accordance with the errors estimated from the EKF for obtaining the corrected azimuth. In the same manner that the step length was determinated while there was the problem with the reception of data from the IMU (at the end of the first line), the azimuth was defined as the average of last data and given that the information of the GPS did not indicate any change of direction this one was kept. As it can see at the beginning of the first line, the azimuth of the IMU has certain troubles to be stabilized which is due to the interferences of the environment —cars and metallic posts around—, even they were not enough strong as to need other techniques for determining the azimuth. In the fourth line there was no signal of the GPS, but the system —as in previous analysis— makes a correct estimation of the azimuth. Generally, the determination both of the azimuth as provided by the GPS once the data is corrected and the corrected azimuth for the DR are considered as satisfactory.

The trajectory as obtained by applying this methodology is illustrated in Figure 4.18 (red line). At first, we observe that the obtained trajectory is more approximated to the trajectory of reference than trajectories of GPS and DR without applying any correction to bias nor to the step length. In spite of slight interferences of the environment, the azimuth is better defined than in previous section, which is observed both in the first line and in the third line (see 4.19). Finally, it is observed how in the few meters as walked by the user in the fourth line the system executes a good approximation of the azimuth when the GPS signal was unavailable.

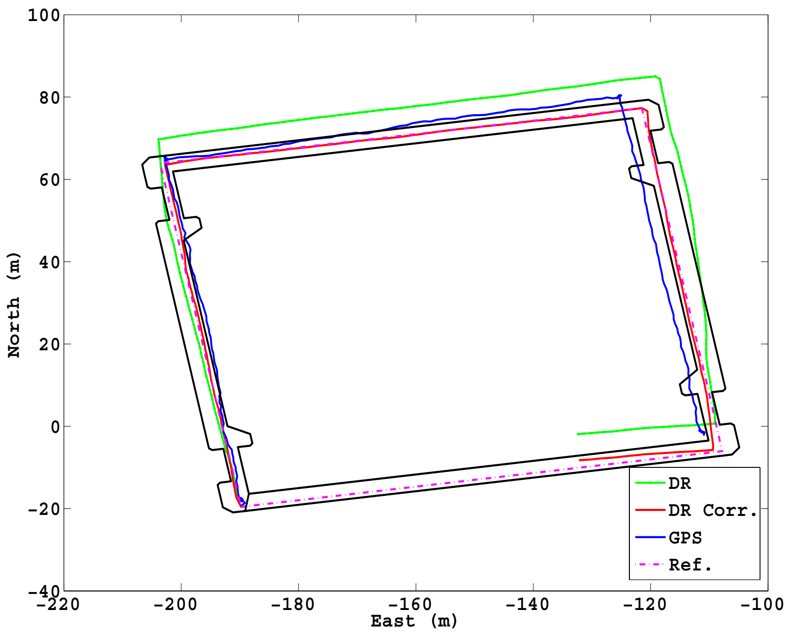


Figure 4.18: User trajectories obtained with GPS, DR without correction and DR with the whole methodology

The corresponding adjustments to the covariance matrices R and Q were applied to the variables which presented inconsistency according to the χ^2 test. In accordance with the different tests carried out the values of Q and R not only permit to obtain a consistent filter but also are a determinant factor in results of the azimuth and final trajectory. The adjustments of R and Q were only applied in the moment when the system gave the alarm of inconsistency and uniquely to the variable presenting such inconsistency. As a result of this process the NIS obtained for each variable did not exceed the value of the confidence limit, e.g. for 95% confidence limit is 3.84. This is

Table 4.4: Results of the consistence tests for the KF and EKF or the GPS-IMU system by applying the complete methodology

Item	NIS	Grd. Compat. (0.7)	AutoCorr.	Lim. (95%)
E	0.0040	1.1585	−0.0972	±0.1295
N	0.0057		−0.1286	
ΔE	0.8256		0.1275	
ΔN	0.2094		0.0116	
ε_s	0.0192		−0.0702	
ε_B	0.0381		−0.0663	
$Innov_{KF}$	0.3686		0.0645	
$Innov_{EKF}$	1.0351		−0.0325	

Table 4.5: Relative error of the distance traveled in each line of the route

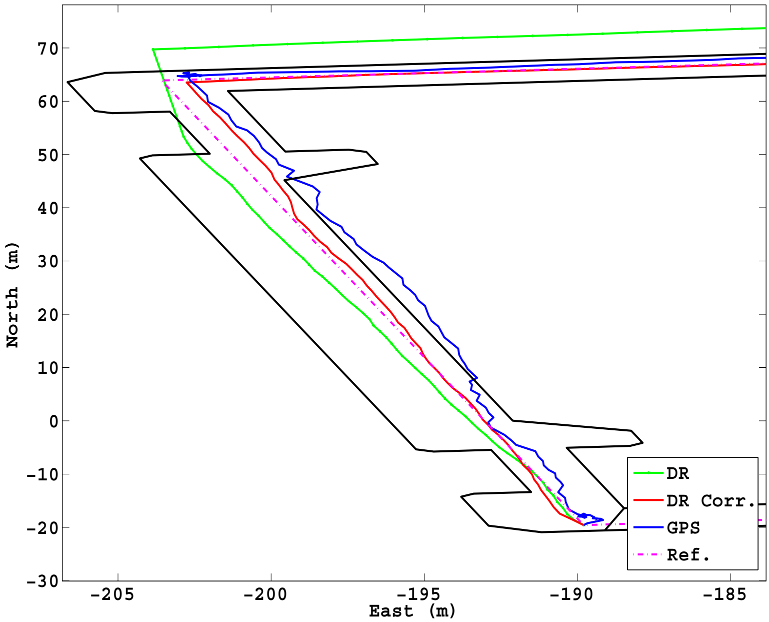
Line	DR No corr . (%)	DR-FKE (%)	DR Corr. (%)
1	8.0481	1.6496	0.3838
2	4.1359	4.1359	0.2361
3	1.8199	0.3655	0.4089

illustrated in Figures 4.20, 4.21 and 4.22.

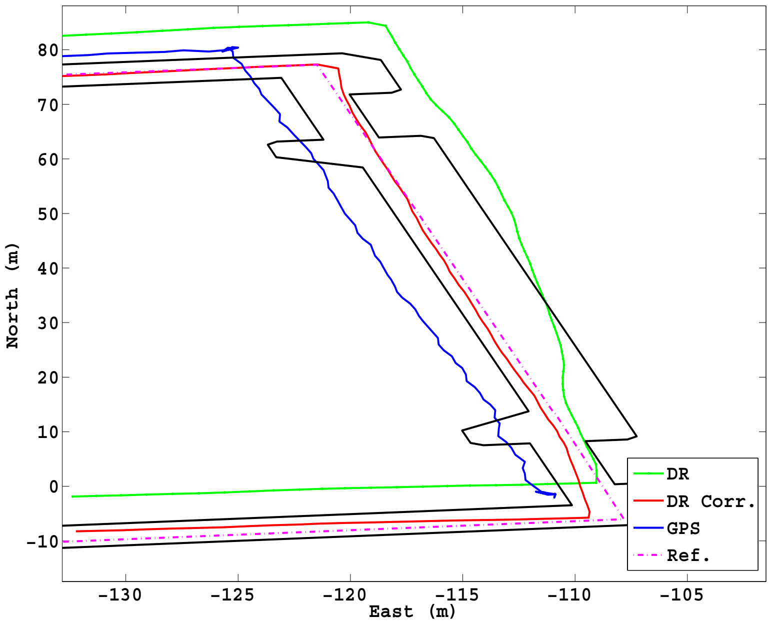
To observe the global performance of the filtering, the average NIS and the auto-correlation function were calculated. Table 4.4 shows the results obtained. It can be seen that all the parameters fulfill the conditions as demanded of whiteness and the NIS were within the allowed limits.

Finally, Table 4.5 includes the relative errors of the distances traveled in each line, for each one of the positioning methods. That is to say, for the DR without correcting its parameters, the DR combined with the EKF (as obtained in the previous section) and the DR obtained with the fault detection and correction methodology. The results indicate a notorious improvement with regard to the two first methods especially in the two first lines.

In short, the presented results indicate that it is very important determining corrected values for the covariance matrices Q and R, since Kalman filtering is strongly dependent on their values. The filtering will provide information to notoriously improve the results of the system, when the values of Q and R are adequately determined. For example, with a correct adaptation of the parameters of the filtering, the disturbances as observed in the third line are nearly imperceptible in the final trajectory of the corrected DR. On the other hand, introducing corrected data of the GPS to the Kalman filtering (in the first phase of the methodology), facilitated also the work of the filters.



(a)



(b)

Figure 4.19: User trajectories obtained with GPS, DR without correction and DR

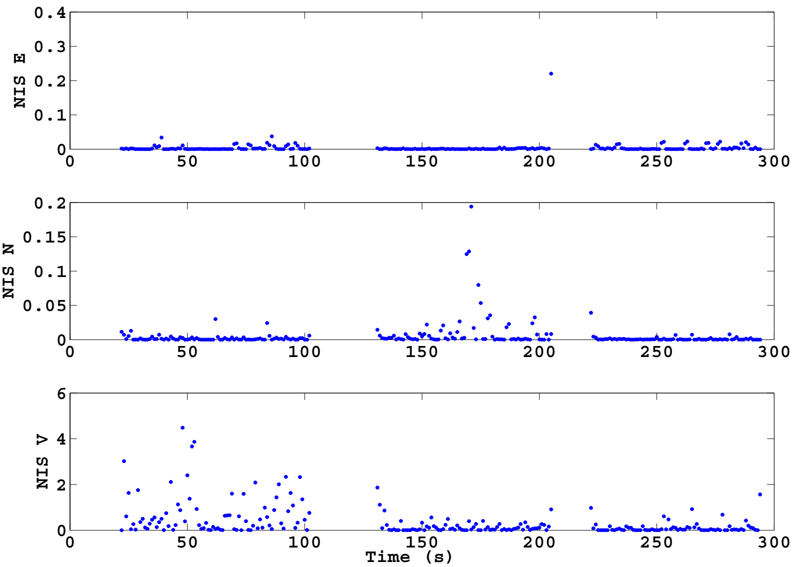


Figure 4.20: NIS of the components from the innovation vector of the KF for the GPS-IMU system by applying the complete methodology

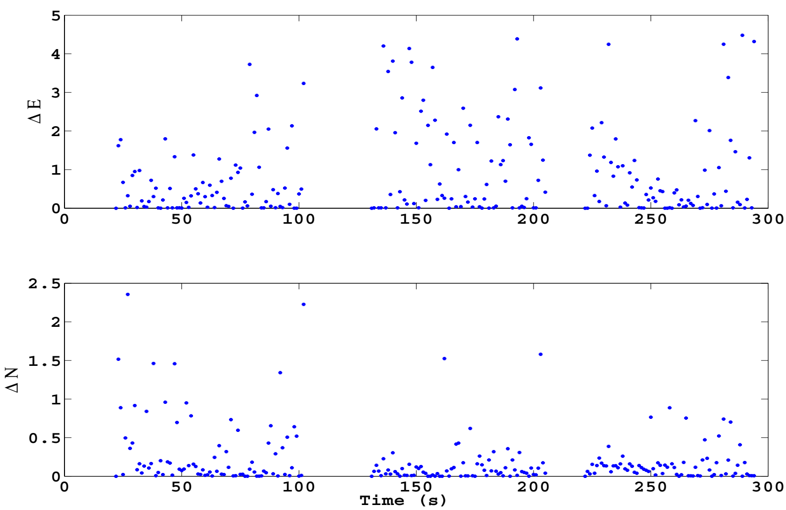
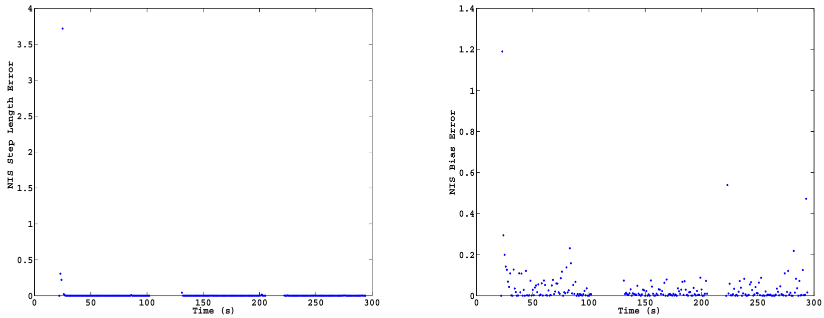


Figure 4.21: NIS of ΔE and ΔN by applying the complete methodology



(a) NIS of the step length error

(b) NIS of the bias error

Figure 4.22: NIS of the innovation of the errors of the DR parameters

4.4 Summary

In this chapter the results have been presented for the DR algorithm as a stand-alone positioning system, the UWB-IMU system and the GPS-IMU system based on the DR algorithm. The results obtained for the first system indicated that continuous calibration of the DR parameters — the step length and the azimuth — is needed.

The UWB-IMU system is analysed for the case when the user walks along a short route. In accordance with previous knowledge about the errors of the parameters, the initial values of the matrices R and Q were assumed. With such values the filter was inconsistent according to the χ^2 tests. This fact indicates the need to apply some correction mechanism and, hence, the evaluation mechanism was applied, that is, the values of matrices Q and/or R were adjusted. To do this, first, trial–and–error tests were carried out in order to obtain the values for R and Q that would allow a consistent filter to be obtained; second, the "scaling the matrix Q" method was applied. The NIS were present with each technique. With both techniques a consistent filter was obtained, which also eliminated the reflections of the signal, as is presented in the resulting route. The tests indicated how a filter evaluation must be performed, and the pertinent corrective measures applied, with little knowledge about the values of R and Q.

For the analysis of the "GPS-IMU system based on the DR algorithm" two methods were taken into account. The first consists in applying Kalman filtering without the methodology for fault detection and correction in order to observe its performance and to see whether the estimations of the errors of the DR parameters help obtain an improved user trajectory. The second method includes the complete methodology for fault detection and correction.

Filtering was inconsistent for the first method because many values of the NIS fall outside the allowed limits (according to the χ^2 test). In addition, the trajectory obtained had some problems with orientation because of perturbations from the environment. In spite of this, there is a correction of the route of the DR. Nevertheless, this trajectory could be acceptable or not depending on the accuracy that is required. Regardless of this fact, the correction to the filter must be carried out, because its basic conditions must be kept in order to guarantee the validity of the results.

Different states of the system were identified with the methodology proposed. Thus, it was possible to identify states of failure (by means of the CD subsystem) that were due to wrong measurements from the GPS or problems with data reception from the IMU. When fault detection was applied, those data were corrected before introducing them into the filters. As a result, several errors were eliminated, e.g. the azimuth obtained from the GPS did not present the peaks that appear when the subsystem was not used.

Regarding the mechanism for detecting inconsistencies in the filters, the results showed that the NIS decreased when the matrices R and Q were adjusted according to the concepts of adaptive Kalman filtering. Therefore, consistent filters were obtained, that is to say, their innovations had a mean of zero and were white. The complete methodology made it possible to obtain a better trajectory with smaller relative errors for the distance travelled than the DR without correction or the GPS.

In general, the use of the complete methodology presented a solution that covers the different stages of data fusion from reading the sensors to the statistical evaluation of the estimations and, as was presented in the results, the methodology corrected both the inconsistencies of the filters and the user's trajectory.

The methodology might be used as an off-line tool for fault detection and if the computational costs do not alter the performance of the systems in real time, it may be applied to obtain more reliable, versatile and robust systems.

CHAPTER 5

General Discussion

THE PROBLEM of user location has been dealt with from the point of view of the evaluation of the information to be processed. This research was conducted on several issues, such as models, prototypes, and Kalman filtering, among others.

First, the DR algorithm was evaluated as a stand-alone pedestrian positioning system that uses elements that make it applicable to positioning in any environment. In this algorithm the step length and the orientation must be determined carefully. To do this, the following tasks were carried out: (1) step detection; (2) determination of the step length; and (3) applying corrective measures for the step length and the orientation. Here, the algorithm selected to establish the step detection was based on the identification of the vertical and horizontal acceleration patterns. In the tests, those patterns were identified for different user velocities (e.g. slow, normal, quick). A neural network was trained to obtain the step length, its inputs being step frequency and acceleration variance. The tests were carried out in a very noisy indoor environment. This was done to observe the performance of the algorithm and, as was expected, it does not work properly as a stand-alone system. Some ad-hoc corrections were established for these experiments; however, as the environment was quite noisy, the tests were not good enough. To correct the parameters of the DR properly, a more sophisticated method was proposed for application to the outdoor system.

A positioning system based on UWB-IMU was proposed. The analysis of this system was based on the discrete Kalman filter. Since the analysis is focused on the evaluation of the filter, then little knowledge of the measurement and process error covariances was assumed. A careful evaluation of the filter was conducted with the consistency test. The orientation was modelled as a first–order Gauss–Markov process. The consistency test involved evaluating each sensor and each component of the innovation vector. This made it possible to observe the behaviour of each variable of the system. In order to carry out the corrective methods, scale factors were applied to matrices R and Q taking into account the fact that they are very important in order to achieve better performance from the filter.

A fault detection and correction methodology was developed for a pedestrian positioning system for outdoor environments. This system consisted of a GPS and an IMU, and was based on the DR algorithm. Its information was processed with optimisation methods for data fusion in multisensor systems. The optimisation was addressed taking into account both empirical and formal knowledge. The empirical knowledge was processed by applying causal diagnosis using the theory of possibility. The formal knowledge was represented by Kalman filtering. Some of the main issues involved in the methodology are: (1) an evaluation of the measurements was established with the framework based on the empirical knowledge; (2) data fusion was carried out with Kalman filtering; (3) the χ^2 test was conducted to observe the performance of the filtering; and (4) a decision-making system was implemented to decide what corrections had to be made. There was sufficient knowledge to define the possibility distributions for the parameters required for the EKF (such as ΔE or ΔN). The failure states of the GPS were identified by applying the basis of causal diagnosis and, in each identification, the decision-making system was executed. The corrections were made taking into account the information from previous epochs and by calculating the average with the last valid data. Once the information to be introduced into the filter had been selected, the prediction was made and the χ^2 test was carried out. In the same way as in the case of the indoor system, the χ^2 test was applied for each component of the innovation vector. The values of the confidence limits in the range (90% - 99%) were introduced into a membership function so as to have different levels of validity. After this process was completed, the system allowed the process to be updated in order to carry out data fusion and/or apply the corresponding correction.

5.1 Discussion

The techniques summarised above to determine the parameters of the DR are robust and reliable. For example, as was expected, once the neural network was trained and if the training conditions were kept, its errors were lower, e.g. in the worst case, its error was 2.039% over the total distance travelled.

When the analysis for the indoor systems was performed [77], the relevant bibliography did not show the UBW–IMU combination as it is dealt with here. The system was not tested in wide areas, although the models proposed should be valid for long routes as well.

The models proposed for the GPS–IMU system based on the DR algorithm had been designed to use the information provided by the sensors and they were slightly different to the models presented in other research. Using the information from the GPS to improve the error estimations in the step length and the bias improved the route considerably. This system is a clear example of how a complementary method can be used, e.g. the position obtained with the DR was corrected by an EKF (which uses GPS-IMU information to estimate the errors of the DR parameters). All these factors, from exploiting the sensor information to correcting the errors and obtaining additional information, incremented the versatility of the system. On the other hand, if some sources failed, the system had other options available with which to guarantee the position, e.g. the estimates from the KF.

The diagnosis system based on the principles of causal diagnosis makes it easier to process the information. That is to say, the advantage of causal diagnosis and possibility theory was exploited to treat uncertain or incomplete information, and characteristics presented in the positioning systems. To do this, the values of the wrong measurements were reduced and it was possible to identify failure states, e.g. erroneous data from the GPS. This produces an increment in the robustness of the system without implying a high computational cost, since it does not require the calculation of complex operations. It is worth noting that, when a wider scope is available, the method will be a very useful tool.

Another method to raise the quality of the information was to evaluate of the Kalman filter. Evaluating each component of the innovation vector individually is useful, as it helps to detect any wrong information. That is, when erroneous data from one of the parameters is received from any of the sensors, the system does not eliminate all the information provided by the sensor, but instead it processes the right information according to the χ^2 test and applies corrective measures to the data that did not pass the test. This also allows a fault detection algorithm to be developed for both the filter and the sensors (e.g. the case of faults not included in the causal diagnosis, such as magnetic disturbances).

In its first phase this methodology succeeded in eliminating the wrong data and, in its second phase, with the appropriate evaluation, it was possible to ensure the validity of data fusion. The final results show the positions obtained, which were closer to the expected routes.

In general, regardless of the design of the positioning system the methodology developed here for the outdoor system can also be applied to other systems (e.g. indoor systems) because its bases are not designed exclusively for outdoor systems. In other words, the methodology can be adapted to several types of sensor systems, e.g. RFID with orientation sensors.

5.2 Contributions

To achieve the objective of the thesis, several issues were analysed. These covered Kalman filtering and the quantification of the empirical knowledge to automate the decision–making process related to the failures in UPSs. This process has led to the following contributions in the area of UPSs:

State of the Art
The literature review was carried out taking into account two fields of applications for location systems for users: mixed reality and navigation systems. This makes it possible to determine some trends to fuse techniques in both fields and to fill certain gaps in the second field, such as the development of fault detection systems.

Causal Diagnosis using the Theory of Possibility for UPSs
The author is not aware of any other research conducted to date that includes causal diagnosis to design fault detection systems for UPSs. This could be very useful, especially if there is a large amount of empirical knowledge to define normal and abnormal models of the systems.

Prototype proposed for indoor environments
When the research on indoor environments was being carried out, no work had been published on the combination of a UWB system and an IMU for UPSs. This is a very promising combination and also provides all the information needed to implement the models described in this thesis.

Fault Detection and Correction Methodology
The development of this methodology provides a framework that allows wrong data to be eliminated in accordance with the point of view of the empirical knowledge. This was conducted by identifying failures using a method based on the principles of causal diagnosis using the possibility theory. Conversely, it was also possible to select the parameters that indicate problems, according to the evaluation from the Kalman filter, and corrective measures were applied for those parameters. This whole process allowed a system to take advantage of the information in a suitable way, that is to say, it keeps the correct information and corrects the information that is incorrect. The proposed methodology was tested in a *GPS-IMU system based on the DR algorithm*, the results of which showed better results for DR. This is achieved thanks to the correct assessment of the GPS information, as this can be used even in cases in which the GPS has high errors (e.g. reflections of the signal). Under such circumstances, these could not be used in the correction of the parameters of the DR without applying any technical assessment to the GPS measurements, and hence the importance of the design of more robust systems.

5.3 Publications related with this Thesis

To date, the original publications based on this work are the following:

> *"Analysis of a Kalman Approach for a Pedestrian Positioning System in Indoor Environments"*
> E. Pulido, R. Quirós, H. Kaufman
> Lecture Notes in Computer Science 4641. The 13th International Euro-Par Conference European Conference on Parallel and Distributed Computing, (Euro-Par 2007), 931-941, Rennes (France), August, 2007.

> *"Assessment of Step Determination in a GPS/Compass/IMU System for Personal Positioning."*
> E. Pulido Herrera, H. Kaufmann, R. Quirós

20th International Technical Meeting of the Satellite Division of the Institute of Navigation The Institute of Navigation (ION GNSS 2007), 1508–1515, Fort Worth, Texas, September, 2007

“Pedestrian Localization System Based on Kalman Filter and Fuzzy Logic”
E. Pulido Herrera, M. Ribo, A. Pinz, R. Quirós, G. Fabregat.
Proc. of the 5th IEEE Conference on Sensors, 1183-1186, Daegu, (Korea), October, 2006.

“Uso del Filtro de Kalman Extendido en Sistemas de Localización de Usuarios”
E. Pulido Herrera, R. Quirós.
XIII Seminario Anual de Automática, Electrónica Industrial e Instrumentación, (SAAEI 2006), September 2006.

“Wearable System for the Localization of Shipping Containers in a Maritime Terminal”
E. Pulido Herrera, M. Ribo, A. Pinz, R. Quirós, G. Fabregat.
Proc. of the 23rd. IEEE Instrumentation and Measurement Technology Conference, (IMTC 2006), 611-616, April, 2006.

“Un Sistema para la Aumentación de Información en Objetos Apilados.”
E. Pulido, R. Quirós Bauset, G. Fabregat and M. Rubio Rojas.
6ta Semana Geomática 2005.

“Sistemas de Seguimiento.”
E. Pulido Herrera and G. Fabregat.
I International Workshop on Information Technologies for Container Terminals ITCT 2004.

5.4 Future Work

There are several lines of research that can be explored within this field the following being but a few examples.

Currently, there is a great amount of interest, both scientifically and commercially, in the development of new location systems for indoor environments. Therefore, the implementation of the methodology proposed in systems comprising different technologies such as RFID, UWB or wireless networks can have a great impact in many sectors that use location systems.

A hardware system based on the methodology outlined here could be designed with fuzzy processors to evaluate the advantages of both the system and the methodology over the traditional methods.

On the other hand, the change of environment was not discussed in this work and should be evaluated in order to have a complete positioning system capable of operating in any environment.

On a more theoretical level, another way to address the changes in the sensor errors (due to changes in the environment or the introduction of sources of error) is the use of adaptive filters and banks of multiple models. Therefore, it would be interesting to carry out a comparison between the methodology presented in this thesis and algorithms such as adaptive filtering Kalman by calculating their computational cost and the errors.

Causal diagnosis can be applied to different levels of the system, thus making it possible to design a complete system that includes the variables of the GPS system (pseudoranges, number of satellites, etc.), the detection of magnetic disturbances either through analysis of magnetic fields or the angular velocities, and other options.

Finally, in the field of mixed reality, it would be interesting to conduct an exploration of these methodologies to see how a system of this type can be referenced and how it might be useful to reduce the costs in terms of the databases that require the storage of images for vision systems.

Bibliography

[1] M. A. PÉREZ, J. C. ÁLVAREZ, AND J. C. CAMPO. "Instrumentación Electrónica". Ed. Paraninfo (2004).

[2] I. E. SUTHERLAND. A Head-Mounted Three Dimensional Display. In "Proc. Fall Joint Computer Conference", volume 33, part 1, pages 757–764. Thomson Books (1968).

[3] G. WELCH AND E. FOXLIN. Motion Tracking: No Silver Bullet, but a Respectable Arsenal. *IEEE Computer and Graphics* **22**(6), 24–38 (2002).

[4] U. NEUMANN AND Y. CHO. A Self-Tracking Augmented Reality System. In "Proc. ACM Virtual Reality Software and Technology", pages 109–115 (1996).

[5] R. AZUMA, B. HOFF, H. NEELY III, AND R. SAFARTY. A Motion - Stabilized Outdoor Augmented Reality System. In "Proc. IEEE Conference on Virtual Reality.", pages 252–259 (1999).

[6] M. RIBO, P. LANG, H. GANSTER, M. BRANDNER, C. STOCK, AND A. PINZ. Hybrid Tracking for Outdoor Augmented Reality Applications. *IEEE Computer Graphics and Applications* **22**(6), 54 – 63 (2002).

[7] E. FOXLIN AND L. NAIMARK. VIS-Tracker: A Wearable Vision-Inertial Self-Tracker. In "Proc. IEEE Conference on Virtual Reality", pages 199–206 (2003).

[8] TYCO ELECTRONICS. "GPS Firmware A1021 and A1025. User's Manual" (2004).

[9] J. ROSSE AND J. G. GAMBLE. "Human Walking". Williams & Wilkins (1994).

[10] O. DE MOUZON, X. GUÉRANDEL, D. DUBOIS, H. PRADE, AND S. BOVERIE. Online possibilistic diagnosis based on expert knowledge for engine dyno test benches. In "Artificial Intelligence Applications and Innovations, AIAI", pages 435–448 (2004).

[11] P. S. MAYBECK. "Stochastic Models, Estimation, and Control". Academic Press (1979).

[12] R. G. BROWN AND P. Y. C. HWANG. "Introduction to Random Signals and Applied Kalman Filtering". Jhon Wiley & Sons (1997).

[13] B. D. ALLEN, G. BISHOP, AND G. WELCH. Tracking: Beyond 15 Minutes of Thought (2001).

[14] O. DUARTE. Sistemas de Lógica Difusa - Fundamentos. *Ingeniería e Investigación, Universidad Nacional de Colombia* **42**, 22–30 (1999).

[15] D. CAYRAC, D. DUBOIS, AND H. PRADE. Handling Uncertainty with Possibility Theory and Fuzzy Sets in a Satellite Fault Diagnosis Application. *IEEE Transactions on Fuzzy Systems* **4**(3), 251–269 (1996).

[16] K. MACHEINER. Performance Analysis of a Commercial Multi-Sensor Pedestrian Navigation System. Master's thesis, Graz University of Technology (2004).

[17] R. W. LEVI AND T. JUDD. Dead Reckoning Navigational System using Accelerometer to Measure Foot Impacts. U.S. Patent US5583776 (1996).

[18] E. FOXLIN. "Handbook of virtual environments: Design, implementation, and applications", chapter Motion tracking requirements and technologies, pages 163–210. IMahwah, NJ: Lawrence Erlbaum Associates, Inc. (2002).

[19] G. RETSCHER AND A. KEALY. Ubiquitous Positioning Techonologies for Modern Intelligent Navigation Systems. *The Journal of Navigation* (59), 91–103 (2006).

[20] A. PINZ M. RIBO AND A. FUHRMANN. A new Optical Tracking System for Virtual and Augmented Reality Applications. In "Proc. IEEE Instrumentation and Measurement Technology Conference, IMTC", volume 3, pages 1932–1936. IEEE (2001).

[21] M. J. CARUSO. Applications of Magnetoresistive Sensors in Navigation System. *Sensors and Actuators* pages 15–21 (1997).

[22] M. J. CARUSO. Applications of Magnetic Sensors for Low Cost Compass Systems. In "Proc. IEEE Position Location and Navigation Symposium", pages 177–184 (2000).

[23] MICROSTRAIN, INC. "3DM-GX HardIronCalibration" (2005).

[24] F. WEIMANN, P. TOMÉ, A. WAEGLI, K. AICHHORN, O. YALAK, AND B. HOFMANN-WELLENHOF. SARHA-Development of a Sensor-Augmented GPS/EGNOS/Galileo Receiver for Urban and Indoor Environments. In "Geomatic Week" (2007).

[25] O. MEZENTSEV, J. COLLIN, H. KUUSNIEMI, AND G. LACHAPELLE. Accuracy Assessment of a High Sensitivity GPS Based Pedestrian Navigation System Aided by Low-Cost Sensors. In "11th Saint Petersburg International Conference on Integrated Navigation Systems", pages 156–164 (2004).

[26] R. T. AZUMA. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* **6**(4), 355–385 (1997).

[27] A. VAN RHIJN, R. VAN LIERE, AND J.D MULDER. An Analysis of Orientation Prediction and Filtering Methods for VR/AR. In "Proc. IEEE Virtual Reality 2005.", pages 67– 74 (2005).

[28] Q. LADETTO. On Foot Navigation: Continuous Step Calibration using both Complementary Recursive Prediction and Adaptive Kalman Filtering. In "GNSS ION", pages 1735–1740 (2000).

[29] R. JIRAWIMUT, P. PTASINSKI, A. GARAJA, F. CECELJA, W. BALACHANDRAN, J.B. FISHER, AND H. HONDA. A Method for Dead Reckoning Parameter Correction in Pedestrian Navigation System. *IEEE Transactions on Instrumentation and Measurement* **52**(1), 209–215 (2003).

[30] S. Y. CHO, C. G. PARK, AND J. G. LEE. A Personal Navigation System Using Low-Cost MEMS/GPS/Fluxgate. In "ION 59th annual meeting" (2003).

[31] J. W. KIM, H. J. JANG, D-H. HWANG, AND C. PARK. A step, stride and heading determination for the pedestrian navigation system. *Journal of Global Positioning Systems* **3**(1–2), 273–279 (2004).

[32] V. GABAGLIO, Q LADETTO, AND B. MERMINOD. Kalman Filter Approach for Augmented GPS Pedestrian Navigation. In "Global Navigation Satellite System, (GNSS)" (2001).

[33] Q. LADETTO AND B. MERMINOD. Compass and Gyroscope Integration for Pedestrian Navigation. In "9th Saint Petersburg International Conference on Integrated Navigation Systems" (2002).

[34] Q. LADETTO AND B. MERMINOD. In Step with INS. Navigation for the Blind, Tracking Emergency Crews. *GPS World* **13**(10), 30–38 (2002).

[35] V. RENAUDIN, O. YALAK, P. TOMÉ, AND B. MERMINOD. Indoor Navigation of Emergency Agents. *European Journal of Navigation* **5**(3) (2007).

[36] F. WEIMANN, G. ABWERZGER, AND B. HOFMANN-WELLENHOF. A Pedestrian Navigation System for Urban and Indoor Environments. In "Proc. 20th International Techinal Meeting of the Satellite Division of the Institute of Navigation ION GNSS'07", pages 1380–1389 (2007).

[37] K. FYFE. Motion Analysis System. US Patent 6513381 (1999).

[38] R. STIRLING, K. FYFE, AND G. LACHAPELLE. An Innovative Shoe-Mounted Pedestrian Navigation System. In "Proc. Conf. European Navigation (GNSS)". Austrian Inst. of Navigation (2003).

[39] K. SAGAWA, I. HIKARU, AND S. YUKATA. Non Restricted Measurements of Walking Distance. In "Proc. of the IEEE Internationsla Conference on Systems, Man and Cybernetics.", volume 3, pages 1847–1852 (2000).

[40] E. FOXLIN. Pedestrian Tracking with Shoe-Mounted Inertial Sensors. *IEEE Computer Graphics and Applications* **25**(6), 38–46 (2005).

[41] S. MOAFIPOOR, D. A. GREJNER-BRZEZINSKA, AND C. K TOTH. A Fuzzy Dead Reckoning Algorithm for a Personal Navigator. In "Proc. 20th International Techinal Meeting of the Satellite Division of the Institute of Navigation ION GNSS'07", pages 48–59 (2007).

[42] M. KOUROGI AND T. KURATA. Personal Positioning based on Walking Locomotion Analysis with Self-Contained Senbsors and a Wearable Camera. In "Proceedings. The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003.", pages 103– 112. IEEE Computer Society (2003).

[43] J. NEWMAN, M. WAGNER, M. BAUER, A. MAC WILLIAMS, T. PINTARIC, D. BEYER, D. PUSTKA, F. STRASSER, D. SCHMALSTIEG, AND G. KLINKER. Ubiquitous Tracking for Augmented Reality. Technical Report, Vienna University of Technology (2004).

[44] M. KOUROGI, N. SAKATA, T. OKUMA, AND T. KURATA. Indoor/Outdoor Pedestrian Navigation with an Embedded GPS/RFID/Self-contained Sensor System. In "Proc. 16th International Conference on Artificial Reality and Telexistence.", pages 1310–1321. Springer (2006).

[45] A. MIRABADI, N. MORI, AND F. SCHMID. Fault Detection and Isolation in Multisensor Train Navigation Systems. In "UKACC International Conference on Control'98", pages 969–974. IEE (1998).

[46] D-H. HWANG, S. H. OH, S. J. LEE, C. PARK, AND C. RIZOS. Design of a Low-Cost Attitude Determination GPS/INS Integrated Navigation System. *GPS Solutions* **9**, 294–311 (2005).

[47] S. R. SWANSON. A Fuzzy Navigational State Estimator for GPS/INS Integration. In "IEEE Position Location and Navigation Symposium", pages 541 – 548 (1998).

[48] J. Z. SASIADEK AND Q. WANG. Sensor Fusion based on Fuzzy Kalman Filtering for Autonomous Robot Vehicle. In "Proc. of the IEEE International Conference on Robotics and Automation", volume 4, pages 2970–2975 (1999).

[49] J. Z. SASIADEK, Q. WANG, AND M. B. ZEREMBA. Fuzzy Adpative Kalman Filtering for INS/GPS Data Fusion. In "Proc. of the IEEE International Symposium on Intelligent Control (ISIC)", pages 181–186 (2000).

[50] S-T. ZHANG AND X-Y. WEI. Fuzzy Adpative Kalman Filterinf for INS/GPS. In "International Conference on Machine Learning and Cybernetics", volume 5, pages 2634– 2637 (2003).

[51] F. CARON, E. DUFLOS, AND P. VANHEEGHE. Introduction of Contextual Information in a Multisensor EKF for Autonomous Land Vehicle Positioning. In "IEEE Proc. Networking, Sensing and Control.", pages 592–597 (2005).

[52] W. ABDEL-HAMID, T. ABDELAZIM, N. EL-SHEIMY, AND G. LACHAPELLE. Improvement of MEMS-IMU/GPS Performance Using Fuzzy Modeling. *GPS Solutions* **10**, 1–11 (2006).

[53] K. MEYER, H. L., APPLEWHITE, AND F. A. BIOCCA. A Survey of Position Trackers. *Presence* **1**(2), 173–199 (1992).

[54] P. K. VARSHNEY. Multisensor Data Fusion. *Electronics & Communication Engineering* (Dic. 1997).

[55] Y. BAR-SHALOM, X-R. LI, AND T. KIRUBARAJAN. "Estimation with Applications to Tracking and Navigation". Jhon Wiley & Sons (2001).

[56] L.A. ZADEH. Fuzzy Sets as a Basis for a Theory of Possibility. *Fuzzy Sets and Systems* (1), 3–28 (1978).

[57] D. DUBOIS, M. GRABISCH, O. DE MOUZON, AND H. PRADE. A Possibilistic Framework for Single-Fault Causal Diagnosis under Uncertainty. *Int. J. General Systems* **30(2)**, 167–192 (2001).

[58] A. KELLY. A 3D State Space Formulation of a Navigation Kalman Filter for Autonomous Vehicles. Technical Report CMU-RI-TR-94-19, Carnegie Mellon University (May. 1994).

[59] M. P. MURRAY. Gait as a total pattern of movement. *American Journal of Physical Medicine* **46(1)**, 290–333 (1967).

[60] F. BRANNSTROM. Positioning techniques alternative to gps (2002).

[61] V. GABAGLIO. "GPS/INS integration for pedestrian navigation". PhD thesis, EPFL Lausanne (2003).

[62] F. L. LEWIS. "Optimal Estimation". Jhon Wiley & Sons (1986).

[63] C. HU, W. CHEN, Y. CHEN, AND D. LIU. Adaptive Kalman Filtering for Vehicle Navigation. *Journal of Global Positioning Systems* **2**(1), 42–47 (2003).

[64] C. HIDE, T. MOORE, AND M. SMITH. Adaptive Kalman filtering algorithms for integrating GPS and low cost INS. In “Symposium of Position Location and Navigation PLANS”, pages 227–233 (2004).

[65] B. D. O. ANDERSON. Exponential Data Weighting in the Kalman-Bucy Filter. *Information Science* **5**, 217–230 (1973).

[66] G. ABDELNOUR, S. CHAND, S. CHIU, AND T. KID. On-Line Detection & Correction of Kalman Filter Divergence by Fuzzy Logic. In “Conf. American Control”, pages 1835–1839 (1993).

[67] A. TIANO, A. ZIRILLI, AND F. PIZZOCCHERO. Application of Interval and Fuzzy Techniques to Integrated Navigation Systems. In “IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th”, volume 1 (2001).

[68] A. HILIUTA, R. LANDRY, AND F. JR. GAGNON. Fuzzy Corrections in a GPS/INS Hybrid Navigation System. *IEEE Transactions on Aerospace and Electronic Systems* **40(2)**, 591–600 (2004).

[69] J. M. MENDEL. Fuzzy logic systems for engineering: a tutorial. In “Proceedings of the IEEE”, volume 83, pages 345–377 (1995).

[70] L.A. ZADEH. Fuzzy Sets. *Information and Control* **8**, 338–353 (1965).

[71] L.A. ZADEH. Fuzzy logic. *Computer* **1**, 83–93 (1988).

[72] D. DUBOIS AND HENRI PRADE. “Fundamentals of Fuzzy Sets”. Springer (2000).

[73] O. DE MOUZON, D. DUBOIS, AND H. PRADE. Using consistency and Abduction Based Indices in Possibilistic Causal Diagnosis. In “The Ninth IEEE International Conference on Fuzzy Systems” (2000).

[74] O. DE MOUZON, D. DUBOIS, AND H. PRADE. Causal Diagnosis in Engine Dyno Test Benches: A Possibilistic Treatment. In “Proc. 15th German-French Institute of Automation and Robotics Conference-Intelligence Control and Diagnosis (IAR-ICD)”, pages 27–32 (2000).

[75] R. K. MEHRA. On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Automat. Contr.* **AC-15**(2), 175–184 (1970).

[76] A. H. MOHAMED AND K. P. SCHWARZ. Adaptive Kalman Filtering for INS/GPS. *Journal of Geodesy* **73**(4), 193–203 (1999).

[77] E. PULIDO, R. QUIRÓS, AND H. KAUFMAN. Analysis of a Kalman Approach for a Pedestrian Positioning System in Indoor Environments. *The 13th International Euro-Par Conference European Conference on Parallel and Distributed Computing, Euro-Par 2007 LNCS* **4641**, 931–941 (2007).

[78] R. E. KALMAN. A New Approach to linear Filtering and Prediction Problems. *ASME Transactions Journal of Basic Engineering* **82**(1) (1960).

[79] R. C. DORF AND R. H. BISHOP. "Modern Control Systems". Prentice Hall (2005).

[80] Y. BAR-SHALOM AND THOMAS E. FORTMANN. "Tracking and Data Association". Academic Press Inc. (1988).

[81] H. W. SORENSON. Least-Squares Estimation: from Gauss to Kalman. *IEEE Spectrum* **7**, 63–68 (Jul. 1970).

[82] FUZZYTECH. urlhttp://www.fuzzytech.com/.

[83] GENSYM. urlhttp://www.gensym.com/.

[84] D. DUBOIS AND H. PRADE. The three Semantics of Fuzzy Sets. *Fuzzy Sets Syst.* **90**(2), 141–150 (1997).

[85] W-H. STEEB. "The nonlinear workbook". World Scientific Publishing Company (2005).

[86] G.L.S. SHACKLE. "Decision, Order and Time in Human Affairs". Cambridge University (1961).

[87] D. DUBOIS AND H. PRADE. Fuzzy Sets in Approximate Reasoning: A Personal View. pages 3–35 (1996).

[88] R. PINO, A. GÓMEZ, AND N. MARTÍNEZ. "Introducción a la Inteligencia Artificial: Sistemas expertos, redes neuronales artificiales y computación evolutiva". Universidad de Oviedo (2001).

[89] E. N. SÁNCHEZ AND A. Y. ALANIS. "Redes Neuronales". Pearson-Prentice Hall (2006).

Appendix A

Filtrado de Kalman

ESTA TÉCNICA fue desarrollada por R. E. Kalman en 1960 [78]. Es un trabajo muy relevante que ha servido como apoyo en el desarrollo de muchas investigaciones en la teoría de control.

Hasta el momento es considerado el método más adecuado para llevar a cabo la fusión de datos en diferentes procesos, y cuya eficiencia y confiabilidad han sido demostradas en varias investigaciones [11], [12]. Sigue siendo la principal herramienta de fusión en sistemas de localización de satélites, aviones, robótica, entre otras. Aunque sus principales aplicaciones en un principio fueron los sistemas de navegación, su aplicabilidad se ha extendido a sistemas eléctricos, economía y a cualquier proceso que requiera predicción, filtrado y suavización de datos. A continuación se presenta la teoría del *filtro de Kalman discreto (FKD)* y de una de sus variaciones el *filtro de Kalman extendido (FKE)*.

A.1 Filtro de Kalman Discreto

Maybeck [11] define en términos sencillos el filtro de Kalman (FK) como: *un algoritmo de procesamiento de datos recursivo y óptimo*, el cual es llevado a cabo dentro del contexto de trabajo ilustrado en la figura A.1. Se podría decir que, el filtro procesa medidas para estimar los valores de las variables del sistema [11], teniendo en cuenta:

1. El conocimiento de la dinámica del sistema y de la medida.

2. La descripción estadística de los ruidos del sistema, errores de la medida y la incertidumbre de los modelos dinámicos.

3. Cualquier información disponible acerca de las condiciones iniciales de las variables de interés.

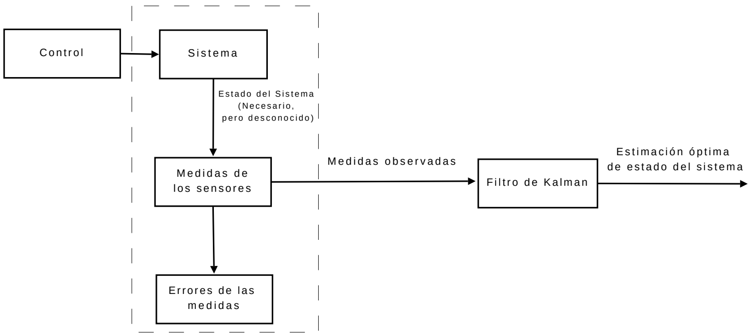


Figure A.1: Contexto de trabajo del FK. Adaptada de [11].

En términos más formales, "el filtro combina las medidas de los dispositivos con el conocimiento a priori de las variables del sistema, para producir una estimación de dichas variables, minimizando el error en términos estadísticos [11]". En otras palabras, el filtro lleva a cabo una propagación de funciones de densidad de probabilidad para problemas que pueden ser descritos a través de un modelo lineal, en el cual los ruidos del sistema y de la medida son gaussianos y blancos [1].

En las siguientes líneas, se presentarán las ecuaciones básicas del filtro de Kalman, las cuales han sido adaptadas en su mayoría de [12].

Kalman empleó la teoría de control en *modelos de variables de estado*[2] para problemas de predicción y filtrado lineal. Él analizó el caso particular de *los sistemas dinámicos lineales variables en el tiempo*.

Partiendo de dicha teoría, dos modelos son definidos: *el modelo de un proceso aleatorio (a estimar)* o *modelo de estado* y *el modelo de la observación o medida*. El primero está definido por:

$$x_{k+1} = \Phi_k x_k + \mathbb{B}_k u_k + w_k \tag{A.1}$$

donde, x_k, es *el vector de estado*, Φ, es la matriz de transición —matriz que relaciona el estado x_k, con el estado x_{k+1}—, $\mathbb{B}$, es la matriz que relaciona la entrada de control opcional u, con el estado x, y w, es el ruido relacionado con el sistema.

El *modelo de observación* está definido por:

$$y_k = H_k x_k + v_k \tag{A.2}$$

donde, y_k, es *el vector de observación* en el instante k, H, es la matriz que relaciona el estado x_k a la medida y_k en condiciones ideales, y v, es el ruido relacionado con la observación.

[1] Su valor no esta correlacionado en el tiempo.

[2] *Las variables de estado describen la futura respuesta de un sistema dando: la actual respuesta, las entradas de excitación y las ecuaciones que describen la dinámica* [79].

w y v son blancos, no están correlacionados, y tienen una distribución de probabilidad gaussiana de media cero:

$$E\left[w_k w_i^T\right] = \begin{cases} Q_k, & i = k \\ 0, & i \neq k \end{cases} \tag{A.3}$$

$$E\left[v_k v_i^T\right] = \begin{cases} R_k, & i = k \\ 0, & i \neq k \end{cases} \tag{A.4}$$

$$E\left[w_k v_i^T\right] = 0, para\ todo\ k\ e\ i. \tag{A.5}$$

Siendo, *Q la matriz de covarianza de error del sistema* de valor conocido y R, *la matriz de covarianza de error de la observación* de valor conocido.

A partir de los modelos anteriores y de la teoría de estimación lineal, Kalman dedujo una serie de ecuaciones recursivas que conforman un algoritmo compuesto por dos fases, denominadas fases de *proyección* y *corrección*.Están definidas así por qué en la fase de proyección se obtienen los valores a priori para el estado $\hat{x}_k^-$[3] y la covarianza de error de la estimación P_k^-; mientras que en el fase corrección se obtiene una estimación mejorada a posteriori $\hat{x}_k$, con la introducción de la observación o medida [62]. La deducción de las ecuaciones que se mencionan a continuación se puede encontrar en, [11], [62], [80], [12], [55].

La *covarianza de error de estimación P*, está asociada al error de estimación e. Para los momentos a priori y posteriori se definen las respectivas covarianzas de error de estimación, P_k^- y P_k, como:

$$e_k^- = x_k - \hat{x}_k^- \tag{A.6}$$

$$P_k^- = E\left[\hat{e}_k^- \hat{e}_k^{-T}\right] = E\left[\left(x_k - \hat{x}^-\right)\left(x_k - \hat{x}_k^-\right)^T\right] \tag{A.7}$$

$$e_k = x_k - \hat{x}_k \tag{A.8}$$

$$P_k = E\left[e_k e_k^T\right] = E\left[\left(x_k - \hat{x}\right)\left(x_k - \hat{x}_k\right)^T\right] \tag{A.9}$$

La determinación de los valores a priori del estado y la covarianza P, en la fase de predicción, se realiza mediante:

$$\hat{x}_k^- = \Phi_k \hat{x}_k \tag{A.10}$$

$$P_k^- = \Phi_k P_k \Phi_k^T + Q_k \tag{A.11}$$

Una vez obtenido lo anterior, se procede a relacionar la observación con la estimación a priori. Esto fue deducido desde el punto de vista de densidad condicional[4], lo que significa que la estimación del vector de estado esta condicionada a las observaciones disponibles[5] con el fin de obtener el *error mínimo cuadrático medio (EMCM)*.

[3] ^, representa estimación, _ representa la mejor estimación a priori en t_k

[4] Una deducción detallada de las ecuaciones es presentada en [11].

[5] Principio de la teoría de estimación.

Para conseguir esto, la estimación a posteriori *"es representada como una combinación entre la estimación a priori y la diferencia poderada entre la actual medida y la predicción de la misma"[13]*, dado por:

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - H_k \hat{x}_k^- \right) \tag{A.12}$$

El factor de ponderación es K_k y es conocida como *la ganancia de Kalman*. La cual es determinada mediante:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k \right)^{-1} \tag{A.13}$$

Finalmente, realizando las transformaciones correspondientes entre, A.2, A.9 y A.12 se obtiene la covarianza de error a posteriori P_k:

$$P_k = (I - K_k H_k) P_k^- \tag{A.14}$$

Las ecuaciones A.10, A.11, A.12, A.13 y A.14, representan el filtro de Kalman y conforman el algoritmo de ciclo infinito, ilustrado en en la figura A.2. Como se ilustra, el filtro debe ser inicializado, lo cual está representado por $\hat{x}_{k-1}^-$, P_{k-1}^-.

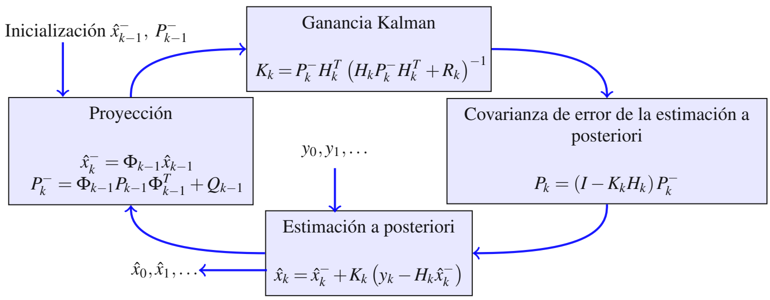


Figure A.2: Flujo de Datos en un Filtro de Kalman. Adaptado de [12], [13]

A.2 Filtro de Kalman Extendido

Cuando los sistemas son no-lineales, el problema de la estimación es llevado a cabo mediante una variación del FK, conocida como el *filtro de Kalman extendido (FKE)*. Una diferencia importante es que mientras que el FK es considerado como un filtro estimador de estado óptimo, el FKE es considerado un *"estimador ad hoc [13]"*, en el que se pierden las propiedades gaussianas de las variables aleatorias durante el proceso.

En este caso el sistema dinámico está representado por:

$$x_{k+1} = f(x_k, u_k, w_k) \tag{A.15}$$

$$y_k = h(x_k, v_k) \tag{A.16}$$

donde:

$x \in \mathscr{R}^n$, es el vector de estado, cuyos componentes son las variables estocásticas del sistema,

f, función no lineal que representa la dinámica del sistema,

$u \in \mathscr{R}^m$, es el vector de control de entrada,

w, vector del error del estado,

$y \in \mathscr{R}^r$, es el vector de medida,

h, función no lineal que representa la observación,

v, vector del error de la observación.

La idea básica es linealizar un sistema no-lineal, mediante la expansión en series de Taylor de los modelos de estado y de observación. Esto es, en la predicción del estado $\hat{x}_k$, la ecuación A.15 se expande en series de Taylor alrededor de la última estimación $\hat{x}_{k-1}$ [55], en la cual u es ignorada para facilitar los cálculos. De manera similar se procede con y_k, deduciéndose así las ecuaciones que representarán el sistema no-lineal [13]:

$$x_k = f(\hat{x}_{k-1}) + F_{k-1}(x_{k-1} - \hat{x}_{k-1}) + w_{k-1} \tag{A.17}$$

$$y_k = h(\hat{x}_k) + H(x_k - f(\hat{x}_k)) + v_k \tag{A.18}$$

donde, F y H son matrices jacobianas:

$$F = \frac{\partial f}{\partial x} \tag{A.19}$$

$$H = \frac{\partial h}{\partial x} \tag{A.20}$$

De manera similar al FK, mediante un análisis más complejo de los errores (no presentado aquí) [13] se obtienen las ecuaciones recursivas que conformarán el filtro, las cuales también están repartidas en las fases de predicción y corrección. Las ecuaciones resultantes para la fase predicción de la estimación de estado a priori y la covarianza de error de estimación a priori están dadas por:

$$\hat{x}_k^- = f(\hat{x}_{k-1}) \tag{A.21}$$

$$P_k^- = F_k P_{k-1} F_k^T + Q_{k-1} \tag{A.22}$$

En cuanto a la fase de corrección las ecuaciones resultantes para la ganancia de Kalman, la covarianza de error de estimación a posteriori y la estimación de estado a posteriori son:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k \right)^{-1} \tag{A.23}$$

$$P_k = \left(I - K_k H_k \right) P_k^- \tag{A.24}$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - h\left(\hat{x}_k^- \right) \right) \tag{A.25}$$

Las ecuaciones de las fases de predicción y corrección conforman el FKE, cuyo flujo de datos está ilustrado en la figura A.3. Se puede observar que a diferencia del FK, en el FKE la predicción de la matriz de covarianza de error de estimación P es calculada usando la derivada parcial de f y la matriz de la ganancia del filtro K es calculada usando la derivada parcial de h.

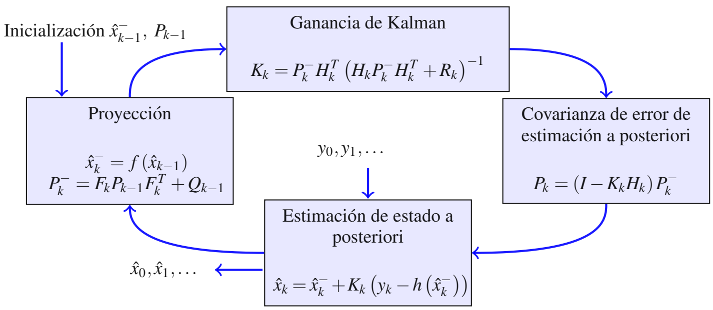


Figure A.3: Flujo de Datos en un Filtro de Kalman Extendido

A.3 Evaluación del Filtrado de Kalman

El filtrado de Kalman es considerado una las herramientas más poderosas para obtener la respuesta óptima de un sistema. Sin embargo, cuando no se encuentra bajo las condiciones adecuadas para un correcto funcionamiento presenta errores no deseados. De aquí que sea necesario comprobar la *convergencia o consistencia* del filtro.

Existen varias causas que pueden llevar al filtro a divergir. Dentro de estas posibles causas de error se tienen ([12],[55]): errores de redondeo, errores de modelado, observabilidad y errores de programación.

Un estimador converge cuando tiende al valor verdadero de la variable que está estimando. Esto es correcto si se está haciendo referencia a un estimador de parámetros constantes, pero en el caso de sistemas dinámicos se tiene además de la estimación, una covarianza asociada a ella [55]. Considerando esto, se debe aplicar el concepto de *convergencia estocástica* [62], en la cual una secuencia de valores $Y_1, Y_2, \ldots$ *converge en probabilidad* a Y, si:

$$P\left(|Y_n - Y| > \varepsilon \right) \rightarrow 0 \; para \; n \rightarrow \infty, \; \varepsilon > 0 \tag{A.26}$$

Se dice que hay *divergencia* o *no convergencia* del filtro cuando *la covarianza del error de estimación P comienza a ser de manera indebida mucho más pequeña que el error de la estimación [81].* En otras palabras, las siguientes definiciones se deberían satisfacer:

$$\begin{gathered} E\left(x_k - \hat{x}_k\right) = 0 \\ E\left[\left[x_k - \hat{x}_k\right]\left[x_k - \hat{x}_k\right]^T\right] = P_k \end{gathered} \tag{A.27}$$

El análisis de la consistencia del filtro se va realizar mediante la evaluación de las innovaciones del filtro, las cuales están representadas por:

$$\vartheta_k = y_k - \hat{y}_k \tag{A.28}$$

donde, $\hat{y}_k$ en el FK es igual a $H_k x_k$, y en el FKE es igual a $h\left(\hat{x}_k^-\right)$ y ϑ_k tiene una covarianza asociada dada por:

$$S_{k+1} = H_{k+1} P_{k+1} H_{k+1}^T + R_{k+1} \tag{A.29}$$

En general la calidad del filtrado es evaluada de acuerdo a los siguientes criterios (basados en las innovaciones), denominados *criterios de consistencia* [55]:

1. La innovaciones deben tener una media de cero.

2. Las innovaciones deber ser blancas.

Sí estas propiedades se cumplen el filtro tiene un buen desempeño. En el caso contrario se deben aplicar métodos para obtener resultados más consistentes.

Si el filtro es consistente, se cumpliría que el *innovación cuadrática normalizada (NIS[6]o ICN)* estaría dado por:

$$\iota_k = \vartheta_k^T S_k^{-1} \vartheta_k \tag{A.30}$$

donde, ε_k debe tener una distribución chi-cuadrado (χ^2) con η grados de libertad. El valor de η está determinado por la dimensión del vector de observación.

Para comprobar la condición (2) en tiempo real se evalúa la *"innovación cuadrática normalizada en un tiempo-promedio* [7], dada por:

$$\bar{\iota} = \frac{1}{N} \sum_{k=1}^{N} \vartheta_k^T S_k^{-1} \vartheta_k \tag{A.31}$$

donde, $N\bar{\varepsilon}$ tiene una distribución χ^2 con $N\eta$ grados de libertad, con N como el número de muestras.

La condición (3) en tiempo real se comprueba aplicando la *autocorrelación en un tiempo promedio para las innovaciones*[8] [55], dada por:

[6] Abreviatura de su término en ingles, "Normalized Innovation Squared"
[7] Del término en ingles *time-average normalized innovations squared.*
[8] Del término en ingles, *Time-Average Autocorrelation.*

$$\rho_j = \sum_{k=1}^{N} \vartheta_k \vartheta_{k+j} \left[\sum_{k=1}^{N} \vartheta_k^2 \sum_{k=1}^{N} \vartheta_{k+j}^2 \right]^{-1/2} \tag{A.32}$$

Dicha condición es aceptada si $\rho_j \in \pm 1.96/\sqrt{N}$ [75]. Cuando las condiciones anteriormente expuestas no se cumplen, se debe analizar el por qué el filtro no es consistente. Cuando se han eliminado causas, como errores en la programación o en el redondeo, la falta de consistencia pueden ser atribuida a problemas de modelado del sistema o a la mala definición de los parámetros iniciales.

APPENDIX B

Lógica Difusa y Teoría de la Posibilidad

LA LÓGICA DIFUSA (LD) es una teoría propuesta por L. Zadeh [70]. Él define un marco conceptual en el cual la base del conocimiento es tratada desde el punto de vista de su naturaleza lingüística, con el cual *"trata de modelar modos imprecisos de razonamiento, los cuales juegan un rol importante en la toma de decisiones"* [71]. En otras palabras, se define un *conocimiento lingüístico* que parte de la forma de razonar de los humanos, en la cual nosotros atribuimos características lingüísticas a las variables —por ejemplo, el café está caliente o frío — y en función de ello tomamos decisiones. Esta aproximación permite mejorar diseños y encontrar soluciones a una gran cantidad de problemas.

Los alcances de la lógica difusa a nivel tecnológico han sido muy grandes, lo cuales han llevado a desarrollar tanto hardware como software especializados basados en sus principios [82]. Tanto en el campo científico como en el industrial se encuentran una gran cantidad de aplicaciones que cubren diseños en la ingeniería de control para aplicaciones industriales, así como otras áreas en las que modelar el conocimiento de un experto no es posible a través de métodos formales. Por ejemplo, la LD ha sido empleada en diseños de sistemas de control para sistemas de potencia o para sistemas de detección de fallos en telefonía móvil, incluso en sistemas de predicción de las bolsas de mercado [83], entre otras.

A partir de la teoría de la lógica difusa, Zadeh desarrolló el marco conceptual para situaciones en las que la información es *incompleta, incierta o ambigua*, al cual denominó *Teoría de la Posibilidad* [56]. A continuación se hace una presentación de los fundamentos teóricos de dichos marcos conceptuales.

B.1 Fundamentos de la Lógica Difusa

La teoría de la LD y de los conjuntos difusos es extensa por ello, aquí se resumen solamente los aspectos concernientes y comúnmente usados en diseños de ingeniería.

A partir de la teoría de conjuntos clásica y de la lógica de proposiciones, Zadeh desarrolló los principios de la LD. Mientras que en la teoría de conjuntos "concretos o clásicos", se define la pertenencia a un conjunto únicamente con dos valores de verdad —verdadero o falso— en la teoría de *conjuntos difusos*, los valores o grados de "pertenencia o membresía" a un conjunto están definidos dentro del rango de valores reales continuo $[0,1]$.

A nivel semántico, los grados de membresía son tratados en términos de *similaridad, preferencia e incertidumbre* [84]:

- Similaridad: Se emplea especialmente en análisis conglomerado (cluster) o análisis de regresión. Consiste en que el grado de la función de membresía a un elemento es asignado de acuerdo a su proximidad con el elemento prototipo de un conjunto difuso.

- Preferencia: Empleado para la programación lineal y el análisis de decisiones. Los conjuntos difusos contienen los objetos más o menos preferidos (o valores de decisión), por lo tanto el grado de membresía indica la intensidad de preferencia en favor de un objeto.

- Incertidumbre: En este caso se dispone de un conocimiento acerca de los valores más o posibles que puede tener una variable en un contexto determinado, para los cuales se establecen los grados de posibilidad de dichos valores. En otras palabras, se tiene un conjunto difuso que contiene los valores más o menos plausibles de una variable, los cuales están adecuadamente ordenados de acuerdo a su plausibilidad.

A continuación se mencionan las definiciones básicas más empleados en los sistemas basados en la lógica difusa.

Se define un universo de discurso U, el cual contiene los valores de una variable x. El *grado de verdad o compatibilidad* del valor de una variable dentro de un conjunto está definido por una *función de membresía o pertenencia*.

Definición 1. *Un conjunto difuso F dentro de una colección de objetos U está caracterizado por una función de membresía $\mu_F(x)$, cuyo rango de valores está en el intervalo real continuo $[0,1]$, denotado por:*

$$F = \{(x, \mu_F(x)) \mid x \in U\} \tag{B.1}$$

donde, x denota un elemento generico de los objetos en U.

La definición de la función de membresía tiene como fuente principal el criterio del diseñador, quien la define basándose en su experiencia o conocimiento empírico del comportamiento del proceso. Dicha función puede tener varias formas; dentro de las más comunes están la singleton, la triangular, la trapezoidal, entre otras (más detalles se pueden encontrar en [69]). Una ilustración típica de las funciones de membresía se presenta en la figura B.1; estás son tres funciones de membresía para tres conjunto difusos (bajo, mediano y alto) que definen la estatura de una persona.

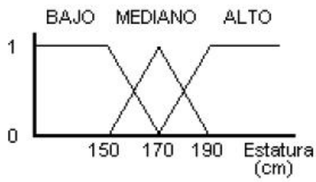


Figure B.1: Ejemplo de funciones de membresía [14].

De la misma manera que en la teoría clásica de conjuntos se definen operaciones entre los subconjuntos de un universo, en la teoría de conjuntos difusos esta tarea es llevado a cabo. Existe una gran diferencia entre estas teorías, ya que mientras en la primera, las operaciones son realizadas con los elementos de los subconjuntos, en la segunda son realizadas con las funciones de membresía que caracterizan cada subconjunto. Las operaciones más comunes son *la unión*, *la intersección* y *el complemento*.

Definición 2. *Teniendo dos conjuntos difusos A y B, la unión está definida a través de sus funciones de membresía:*

$$\mu_{A\cup B}(x) = max\left[\mu_A, \mu_B\right] \tag{B.2}$$

Definición 3. *Teniendo dos conjuntos difusos A y B, la intersección está definida a través de sus funciones de membresía:*

$$\mu_{A\cap B}(x) = min\left[\mu_A, \mu_B\right] \tag{B.3}$$

Definición 4. *El complemento de un conjunto difuso está dado por:*

$$\mu_{\bar{A}} = 1 - \mu_A \tag{B.4}$$

Otro aspecto a resaltar de la LD es el uso del concepto de *implicación*, cuyo operador viene definido desde la teoría de proposiciones. El uso de este operador permite establecer relaciones de implicación entre varios conjuntos difusos pertenecientes a diferentes universos discursivos. La definición de estas relaciones permite trabajar con uno de los conceptos más relevantes de la LD, denominado *proposiciones "IF-THEM" o SI-ENTONCES*. En las cuales SI representa *un antecedente* y ENTONCES representa *un consecuente*. Como en los conceptos explicados anteriormente, la implicación entre dos conjuntos difusos también se define a través de sus funciones de membresía.

Definición 5. *La implicación entre dos conjuntos difusos A y B está dada por:*

$$\begin{gathered}\mu_{A\to B} \triangleq min\left[\mu_A(x), \mu_B(x)\right] \\ o \\ \mu_{A\to B} \triangleq \mu_A(x)\,\mu_B(x)\end{gathered} \tag{B.5}$$

En la mayoría de las aplicaciones que la emplean la LD, los sistemas que se desarrollan se denominan *Sistema Basados en Lógica Difusa (SBLD)*. Un SBLD típico está compuesto por: *difusor*, *reglas difusas*, *máquina de inferencia* y *concresor*; esto se ilustra en la figura B.2.

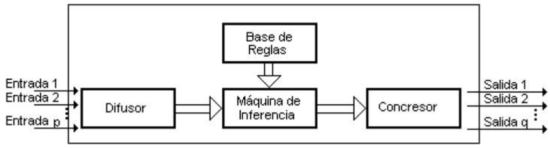


Figure B.2: Estructura de un sistema de LD [14].

Difusor: Los valores concretos de las variables deben asociarse a los grados de compatibilidad dentro de sus respectivos universos y conjuntos difusos. Es decir, se establece la relación entre las variables lingüísticas y sus respectivos grados de membresía.

Reglas Difusas: Son las proposiciones SI-ENTONCES que definen las reglas del sistema, las cuales representan las relaciones causa-efecto (SI *x* es A ENTONCES *y* es B). Normalmente estas reglas son determinadas según el criterio del experto, quien las define teniendo en cuenta las variables de estado del sistema, criterios experimentales o matemáticos, entre otros.

Máquina de Inferencia: Establece la forma de interacción entre las reglas, es decir, de acuerdo a la entrada del sistema difuso, se obtendrá una salida determinada. Para llevar a cabo esto, hace uso de las herramientas de la teoría difusa explicadas a lo largo de esta sección, como las operaciones de conjuntos difusos y la relación de implicación. Las reglas que son tenidas en cuenta son aquellas cuyo grado de membresía es distinto de cero.

Concresor: Dependiendo de los requerimientos del sistema, las salidas en un SLD pueden ser difusas o pueden ser valores concretos. Este componente, permite definir un valor concreto cuando sea necesario. Existen diferentes técnicas para realizar dicha conversión, entre las más populares están centro de gravedad, método de la altura, primero del máximo, último del máximo y media de los máximos. En los sistemas de control, el más popular es el centro de gravedad [85].

Definición 6. *El centroide o centro de gravedad de una serie difusa, $\tilde{A}$, con una función de membresía $\mu_{\tilde{A}}$, $x \in X$, está definido para el caso discreto por:*

$$c := \frac{\sum_{x \in X} x \mu_{\tilde{A}}}{\sum_{x \in X} \mu_{\tilde{A}}} \tag{B.6}$$

Finalmente, los *sistemas de lógica difusa basados en reglas (SLDBR)* están clasificados en dos tipos: *el método de Mamdani* y *el método Takagi-Sugeno-Kang (TSK)* [69]. En el método de Mamdani, las reglas difusas son de la forma: ***SI*** *x_1 es A Y x_2 es B* ***ENTONCES*** *Z es C*; mientras que en el método TSK son de la forma: ***SI*** *x_1 es A y x_2 es B* ***ENTONCES*** $Z = f(x1, x2)$. Esto significa, que en el TSK no se necesita un concresor, puesto que la consecuencia es una valor dado por una función. Su desventaja sería que representa de forma menos natural el conocimiento humano, pero por otro lado, es más eficiente a nivel computacional.

B.2 Teoría de la Posibilidad

Zadeh [56] definió un marco conceptual basado en los conjuntos difusos para tratar problemas con *información incompleta con incertidumbre*, denominado *Teoría de la Posibilidad*. Con esta teoría, se propone solucionar problemas de acuerdo a la información disponible, planteando soluciones más desde un punto de vista de posibilidades que de probabilidades. En consecuencia, algunos autores la definen como una *teoría de la incertidumbre no-clásica* [72]. Dentro de las áreas de aplicación se encuentran la representación del conocimiento, la toma de decisiones en sistemas con incertidumbre o el diagnóstico etc.

El término *posibilidad* comprende los siguientes significados dentro de esta teoría [72]: *viabilidad, plausibilidad y consistencia*. La viabilidad hace referencia a la facilidad de solucionar un problema bajo ciertas restricciones; la plausibilidad hace referencia a la tendencia de que ciertos eventos ocurran; y la consistencia hace referencia a que no exista contradicción en la información. Otro aspecto que tratan Dubois y otros [72] es el concepto epistémico de la plausibilidad, originalmente definido por Shackle [86]. Él introduce la idea de calcular grados de sorpresa, entonces al asignar dichos grados a un evento de acuerdo al conocimiento de un agente, se indiciaría en que medida la ocurrencia del evento es contradictoria con el conocimiento del agente. De aquí que estos autores definan la plausibilidad como *carencia de sorpresa* de la ocurrencia de un evento, la cual involucraría también la consistencia de la información.

A cada evento se le asigna una *"distribución de posibilidad"*. Zadeh define esta distribución como la función de membresía de un conjunto difuso. Sin embargo, no puede ser interpretada como en la teoría de la lógica difusa, expuesta en la sección anterior. En la teoría en cuestión, la distribución de posibilidad representa aquella información incompleta que se tiene acerca de un atributo, cuyos valores son solamente aquellos que se consideran posibles, mientras que en los conjuntos difusos esa información es precisa.

A continuación se presentan de manera formal conceptos básicos de la teoría de la posibilidad, como la distribución de posibilidad y las medidas de posibilidad y necesidad.

Zadeh establece que una variable X está asociada a una distribución de posibilidad π_X. Paralelamente define un universo discursivo U que representa el estado de los hechos, es decir los posibles valores que puede tener la variable X. La distribución de posibilidad π_X representa *"el estado de conocimiento, distinguiendo lo que es plausible de lo que no, lo que es el curso normal de las cosas de lo que no lo es o lo que sorprende de lo que es esperado"* [72]. π_X en U es un mapeo desde U al intervalo $[0,1]$ asignado a los valores de la variable X. π_X representa restricciones flexibles, las cuales limitan la información de los posibles valores de X de acuerdo a la información disponible, como se indica [87]:

$\pi_X(u) = 1$, significa que $X = u$ es totalmente posible o plausible.

$\pi_X(u) = 0$, significa que $X = u$ es imposible.

donde u es el valor tomado por X.

En términos generales plantea que las distribuciones de posibilidad son definidas en términos de plausibilidad, la cual está vinculada a la certeza. Por lo tanto, si U ha sido bien definido se debe cumplir que al menos uno de sus elementos debe tener un grado de posibilidad de 1, $\exists u, \pi_X(u) = 1$ [87].

Dentro de este marco conceptual son definidas dos formas de conocimiento parcial extremas:

conocimiento completo: para algún u_0, $\pi_X(u_0) = 1$ y $\pi_X(u) = 0, \forall u \neq u_0$

(solamente $x = u_0$ es posible)

completa ignorancia: $\pi_X(u_0) = 1$, $\forall u$ (todos los valores en U son posibles)

La tendencia de X de pertenecer a un subconjunto A de U, puede ser establecida mediante las definiciones de *las medidas de posibilidad y de necesidad*, denotadas por Π y N, respectivamente.

$$\Pi(\mathrm{A}) = sup_{u\in \mathrm{A}} \pi_X(u) \tag{B.7}$$

expresa la consistencia de la declaración $X \in A$.

$$\mathrm{N}(\mathrm{A}) = 1 - \Pi(\mathrm{A}) \tag{B.8}$$

expresa que los valores más altos de la distribución de posibilidad deben encontrarse dentro de los elementos de A. En otras palabras, expresa la certeza que se tiene de que el valor de X pertenezca a A.

En la figura B.3 se ilustran las condiciones para la posibilidad y la necesidad, las cuales cumplen:

- A es más cierta a medida que su complemento $\overline{\mathrm{A}}$ es menos consistente.
- π_X está normalizada, lo que significa que $max\left(\Pi(\mathrm{A}), \Pi\left(\overline{\mathrm{A}}\right)\right) = 1$
- $\Pi(\mathrm{A}) = \Pi\left(\overline{\mathrm{A}}\right)$ corresponde al caso de ignorancia absoluta.
- $\Pi(\mathrm{A}) = 0$, ciertamente falso.

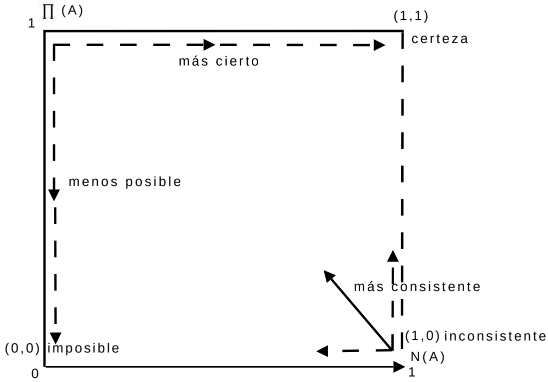


Figure B.3: Convenciones de la teoría de la posibilidad, tomado de [15].

APPENDIX C

Redes Neuronales

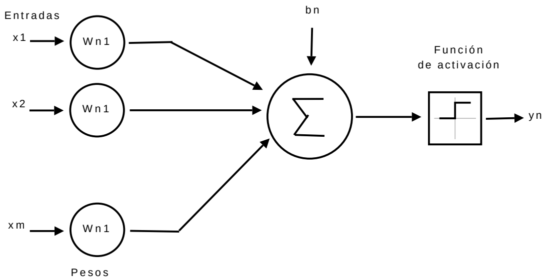
Las redes neuronales al igual que la LD son un método que busca encontrar soluciones a problemas altamente complejós y que con las técnicas tradicionales no pueden ser tratados. Parten del principio de funcionamiento del cerebro humano, el cual es muy eficaz solucionando problemas difusos o problemas que requieren un procesamiento complejo de la información. Para ello, su desarrollo se ha llevado a cabo imitando el comportamiento de redes de neuronas biológicas.

La definición más popular para una red neuronal es de Robert Hetch-Nielsen [88]: *Sistema de computación que consta de un gran número de elementos simples, muy interconectados, que procesan la información respondiendo dinámicamente frente a unos estímulos externos*. Dichos elementos son denominados *neuronas*.

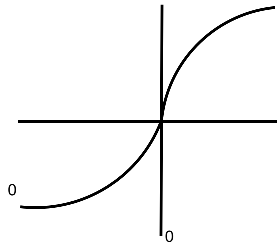
Las neuronas son las encargadas de transformar las señales de entrada en una salida y cuyo modelo se ilustra en la figura C.1a. Dicho modelo está compuesto por unos valores que ponderan las entradas, un elemento sumador, el factor de corrección y la función de activación. Cada uno de esos componentes tiene una formulación matemática que no es expuesta aquí [89]. La función de activación puede tener varias formas como la escalón, lineal a tramos o sigmoidal; la función sigmoidal es presentada en la figura C.1b.

Por otro lado, para establecer las interconexiones entre las neuronas, las redes tienen arquitecturas compuestas por tres tipos de capas —entrada, intermedia y salida —. Una arquitectura típica es la que se ilustra en la figura C.1c.

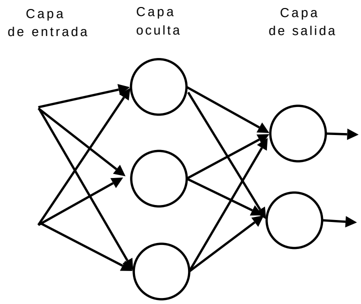
Una característica que hace interesantes a las redes neuronales, es su *capacidad de aprendizaje*, con la cual puede modificar sus parámetros y generar respuestas de acuerdo al ambiente. Esto implica que la red debe ser entrenada y las respuestas van a ser adecuadas sólo con un correcto entrenamiento. Por otro lado, cuando se tiene información de las entradas y de sus correspondientes salidas, se puede encontrar la



(a) Modelo de una Neurona



(b) Función de Activación Sigmoidal



(c) Red Neuronal Multicapa

Figure C.1: Conceptos de las redes neuronales.

solución de un problema sin necesidad de entrar a detallar el modelo, ya que podemos tratar la red neuronal como una caja negra, que proporcionará la salida que corresponda a la entrada dada.