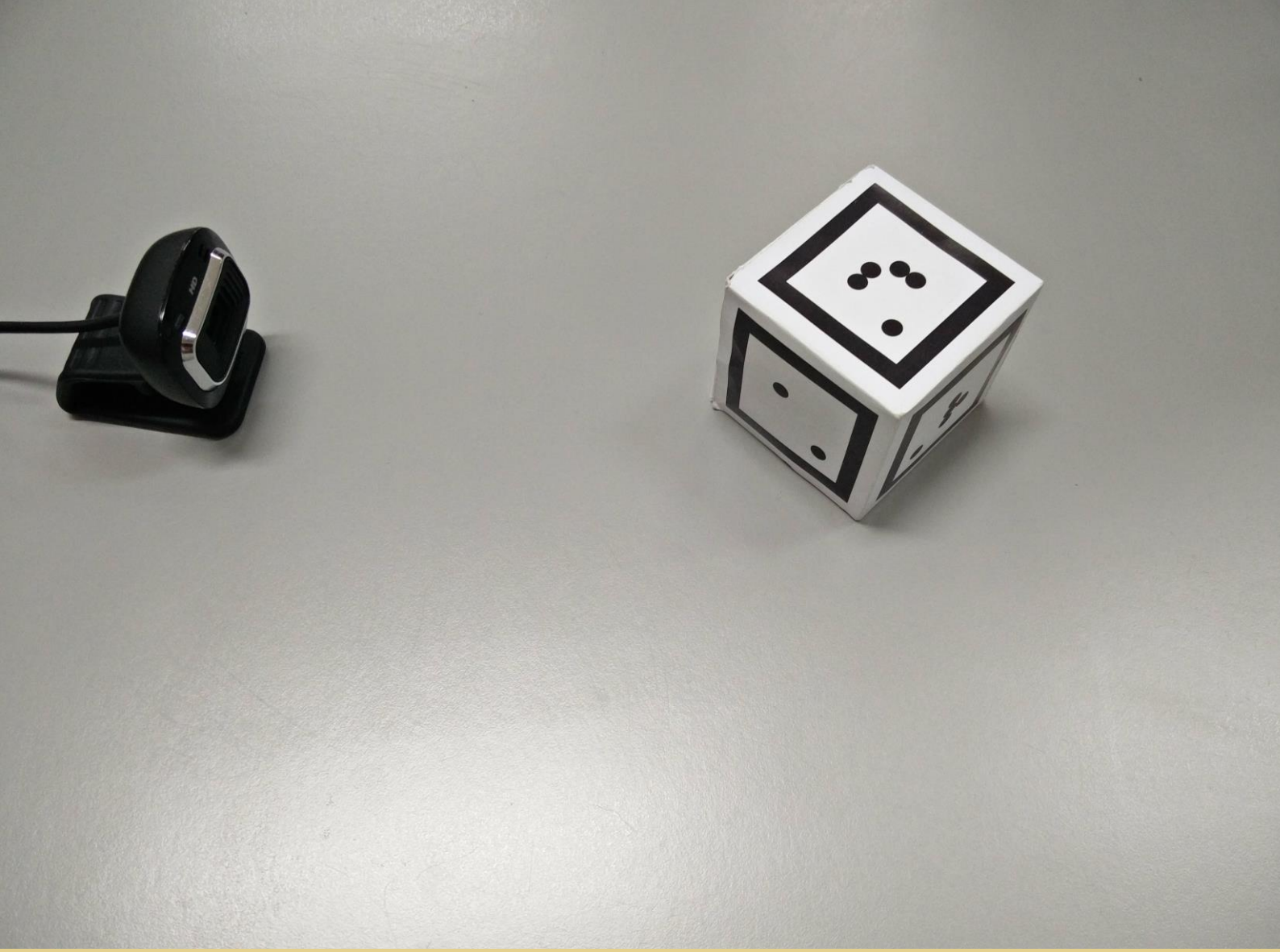

Nuno Rafael Pereira Gouveia

# Manipulation of Tactile Perception of Objects in Mixed Reality

September 2017

· U C ·

Universidade de Coimbra

Department of Electrical and Computer Engineering,
Faculty of Sciences and Technology, University of Coimbra,
3030-290 Coimbra, Portugal.

A Dissertation for Graduate Study in MSc Program
Master of Science in Electrical and Computer Engineering

# Manipulation of Tactile Perception of Objects in Mixed Reality

Nuno Rafael Pereira Gouveia

Research Developed Under Supervision of:
Prof. Paulo Jorge Carvalho Menezes
Master Bruno André Santos Patrão

Jury:
Prof. Rui Pedro Duarte Cortesão
Prof. Jorge Nuno de Almeida e Sousa Almada Lobo
Prof. Paulo Jorge Carvalho Menezes

September 2017

# Acknowledgements

I would like to start by thanking my advisers Professor Paulo Jorge Carvalho Menezes and Bruno André Santos Patrão for steering me in the right the direction throughout the development of this work. I would also like to acknowledge my colleagues at the Mobile Robotics Laboratory in Institute of Systems and Robotics, for the great atmosphere and support provided.

A special thanks to all my friends, your friendship and guidance aided me through this journey and made me into what I am today.

Finally, I must express my profound gratitude to family especially my mother, sister and grandmother, for providing me with inexhaustible support and continuous encouragement throughout my years of study. Thank you!

# Abstract

After the resurgence of virtual reality as a research topic due to, among other advancements in technology, the improvement of head-mounted displays. Mixed reality falls in, naturally, as the next step in human, computer and environment interaction, blending the physical world with the digital one.

Heretofore, human interaction with computers occurs through an assortment of devices, which range from trackballs to controllers. Within mixed and virtual reality environments, most applications purely rely on proprioception for object interaction, which causes the manipulation challenging, in most cases. On the other hand, the operation of traditional input methods sound inadequate for the interaction in these type of environments.

This dissertation proposes the use of a real object, referred to as instrumented object, with pose tracking, passive haptics and wireless capabilities. The aforementioned systems determines the object's pose, by fusing a visual pose estimation granted by an external camera, with the pose handed over by strapdown inertial navigation employing a MEMS inertial sensor.

The visual pose estimation is performed through the tracking of a set of fiducial markers placed on the surface of the instrumented object. The pose tracking in the inertial system is performed by double integration of the measured acceleration in the correct orientation frame, after gravity removal. The attitude tracking is performed by a complementary filter, which combines data from the gyroscope, accelerometer and magnetometer for a orientation estimate. In other to certify the performance of the inertial sensor, an analysis was conducted illustrating the sensor disturbances and a calibration was implemented to lessen these effects.

The work hereby presented, concluded in the development of a functioning prototype. It provides an improvement for tactile stimulus which as been neglected until now, therefore enhancing user's immersion, contributing this way as a complementary tool for adoption within virtual and mixed reality scenarios.

**Keywords:** Virtual Reality, Mixed Reality, Instrumented Object, Visual Tracking, Inertial System, Complementary Filter.

# Resumo

Após o renascimento da realidade virtual como uma área de estudo e pesquisa devido a, entre outros avanços na tecnologia, o desenvolvimento nos capacetes de realidade virtual. A realidade mista é vista como uma natural sucessora, na interacção entre homem, computador e ambiente, misturando o mundo real com o mundo digital.

Até aos dias de hoje, a interacção com computadores era feita através de uma miscelânea de dispositivos, que variam desde comandos a ratos. Com a chegada de ambientes de realidade mista e virtual, a interacção com objectos baseia-se, de maneira acentuada, em conceitos como o de propriocepção, o que na maioria dos casos torna a manipulação dos mesmos desafiante. Por outro lado, o uso de periféricos de entrada tradicionais parece inadequado.

Esta dissertação propõe a utilização de um objecto, normalmente apelidado de objecto instrumentado, com capacidades hápticas passivas, de seguimento de posição e de comunicação sem-fios. O sistema anteriormente descrito, determina a sua posição através da combinação de uma estimação visual obtida por uma câmara exterior e de um sistema inercial colocado no seu interior.

A estimação visual é feita através da identificação da posição de marcadores visuais colocados no exterior do objecto. O seguimento da posição do sistema inercial, é obtido através da dupla integração da aceleração no sistema referencial correcto após subtracção da gravidade. A orientação do sistema inercial é obtida através de um filtro complementar, que combina dados do acelerómetro, giroscópio e do magnetómetro. De maneira a certificar o desempenho do sensor inercial, foi conduzida uma análise demonstrando as perturbações do sensor e, em seguida, feita uma calibração ao mesmo para esbater estes efeitos.

O trabalho apresentado culminou no desenvolvimento de um protótipo funcional. Este possibilita o estimulo de um dos cinco sentidos até aqui um pouco negligenciado, o tacto, aumentando desta forma a imersão do utilizador, contribuindo como uma ferramenta complementar ao desenvolvimento de cenários de realidade virtual e mista.

**Palavras-Chave:** Realidade Virtual, Realidade Mista, Objecto Instrumentado, Seguimento Visual, Sistema Inercial, Filtro Complementar.

# Contents

# List of Figures

# Chapter 1

# Introduction

In contrast with the past, current development in Virtual Reality (VR) technology is growing at an unprecedented pace. Notwithstanding, more than half a century has passed since Ivan Sutherland acquainted the idea of Ultimate Display [1], in which the participant is completely immersed in a fabricated world. The aforementioned world, can resemble a real-world environment or exceed its limits, with the boundaries exclusively drawn by our imagination. Additionally, twenty years were required for Jaron Lanier to coin the term Virtual Reality, gathering the different concepts presented until then. This started a wave of euphoria to accomplish Sutherland's vision, that fainted few years after. In 2012, the Oculus Rift Head-Mounted Display project, initiated on Kickstarter and posteriorly bought by Facebook, Inc., acted as a spark reigniting the topic and the movement recouped attention.

Mixed Reality (MR), is a broader concept, that involves blending real and virtual worlds, as proposed by Paul Milgram and Fumio Kishino [2]. Presumably, the most well know example of this area is Augmented Reality (AR), compounding all scenarios where a real environment is enhanced by virtual objects.

Although, VR and MR worlds provide a better understanding of three-dimensional shapes and spaces, proper interaction is still difficult to accomplish, despite of the research in interaction techniques [3, 4]. Humans depend upon physical constraints and haptic feedback for interaction in the real world. Virtual environments, generally, neglect any form of sensory feedback other than visual, predominantly relying on notions such as *proprioception* [5], the sense of the relative position and orientation of one's own parts of the body, for manipulation and interaction with the virtual environment.

## 1.1 State-of-the-art

By virtue of the improvements in display technology and computer hardware and software, especially the development of new three-dimensional graphic techniques, the impact of virtual reality and related technologies, on visual perception and simulation has profoundly changed the fashion people take advantage of computers. For instance, users immerse themselves in virtual worlds to experience events and attain enlightenment infeasible in conventional manners. Such as the idea of telepresence in teleoperation tasks, aiming at placing the user in the center of the task execution. A couple of techniques are commonly employed, depending on the intended approach, the user is either embodied in robotic character with his actions directly represented on the environment [6], or as an embarked pilot or driver, controlling a vehicle [7].

Concerning the embodied scenario, the robotic character, must be perceived as the own body, in a manner that neighbouring elements in the surroundings must be perceived as being adjacent to one's body. Furthermore, the controls must be intuitive and easily manipulable. Regarding the embarkment approach, the perception provided differs from the previously characterized. Typically, accommodating a virtual representation of the user, commonly entitled as avatar, as the vehicle operator placed on the inside, in a sort of virtual cockpit which may include a set of appropriated controls, e.g., joysticks, steering wheels to name a few.

Considering both aforementioned immersive systems, as well as, other application scenarios there is a compel for the integration of several interaction devices, where users can interact with the objects of the virtual environment in a natural way and can obtain the real-time experience and feelings of the physical environment by operating the objects in the virtual environment. This demands research and development of different tracking technologies in the interest of achieving fruitful experiences.

### 1.1.1 Motion Tracking Technologies

Tracking devices consent the systems to monitor the position and orientation of a selected part. In VR/MR systems these devices are widely adopted for tracking body parts, for instance, in HMDs this information defines the user's viewpoint in the virtual world, and determines which part of it should be rendered to the visual display.

There are two methods of approaching tracking, a passive or remote approach, in which the

target is monitored from a distance, and active/*in loco* one, where the monitoring device is attached to the target. Nowadays, the latter represents the most frequent technique, mainly due to efficiency.

The efficiency of tracking devices, in virtual environments, depends to a large degree on whether the movements are synchronized to the virtual world actions. Additionally, factors as sensitivity, environmental interference among other can limit tracking performance. Current tracking devices are based on acoustic, inertial, magnetic, mechanical, optical or radio frequency technology. A succinct presentation of each of these approaches and some of their limitations follows.

### Acoustic Tracking

Acoustic trackers utilise the transmission and sensing of sound waves. Commercially available acoustic ranging systems operate employing the principal of time-of-flight of an ultrasonic pulse. On the other hand, Sutherland built a continuous wave ultrasonic tracker, in order to complement his mechanical head tracker for his original HMD [8]. Although not experiencing latency, the system suffered from *multipath*, meaning that the signal measured often comprise the sum of the direct signal and the reflected signals. This is partially overcame by pulsed time-of-flight acoustic systems, by waiting until the first pulse arrives, which guaranteed to have arrived via the direct path except if the signal is blocked.

### Inertial Tracking

Inertial Navigation Systems (INSs) application was widely spread, in airplanes, submarines, ships among others [9], before their late usage in interaction devices for virtual environments. This later application only emerged with the arrival of microelectronic mechanical systems (MEMS) inertial sensors. Initially, INS were built utilizing a stable platform aligned with a fixed reference frame, equipped with gyroscopes and motors in a feedback loop to uphold the reference. Position was obtained by double integration of the accelerometers values, attached to the platform, after gravity compensation.

Nowadays, systems operate in a distinct way [10], disposing of the mechanical platform and measuring orientation by integrating angular-rates from three gyroscopes, positioned orthogonally. For position, three linear accelerometers measure the acceleration vector in body-frame,

which is then rotated to the navigation-frame through the established orientation, that can then be gravity-compensated and double-integrated in the same way as before. Strapdown MEMS inertial sensors exceed the previous solution, being less prone to error, owing that to the lack of mechanical gimbal stabilization while having a smaller form factor.

### Magnetic Tracking

Magnetic trackers rely on measurements of the local magnetic field vector at the sensor, employing Hall effect sensors, or by measuring variations in the magnetic field through current induced in electromagnetic coils. A sensor unit is embedded with three orthogonal magnetic sensors. A couple of approaches can be adopted for tracking, inducing excitations with a multicoil base station [11] or, alternatively, take advantage of earth's magnetic field to estimate heading. The multicoil solution, works by sequentially energize each of the source coils and measuring the corresponding magnetic field in the sensor. Pose estimation can be inferred with atleast three excitations.

### Mechanical Tracking

Mechanical systems, arguably the simplest approach, assess position and orientation by a direct physical linkage between the a reference point in the environment and the target. Ordinarily, a light-weight arm with potentiometers or encoders placed at the joints measure the changes in pose related to the reference point. This approach can provide very precise and accurate pose estimates for a single target, but only over a relatively small range of motion.

Boom tracked displays by Fakespace Labs [12] represents a great example of a mechanical-based tracking, employing counterweights to balance the load and for 3D pose tracking, another noteworthy example follows in Sutherland's original HMD work [8].

### Optical Tracking

Optical sensing comes in two variants, sensors can be placed with a *outside-looking-in* approach or *inside-looking-out* approach. Systems like Optotrak Certus [13] use the first technique, where a set of cameras are located at a fixed position in the environment and triangulate the pose of a set of markers placed on the desired target. The alternative design solution, is presented in

systems like VRvana Totem HMD [14], consisting in a pair of front facing cameras mounted on the headset for pose tracking. Optical trackers in general present good performance. However, they suffer from occlusion, ambient light and infrared radiation also adversely affect optical tracker performance.

#### Radio and Microwave Tracking

Most radio tracking systems operate on the principle of time-of-flight range finding, similarly to the acoustic systems earlier described. The waves travel faster, making the task of timing the flight duration, with sufficient precision, more demanding. Radio and microwaves are widely used in navigation systems, such as Global Positioning System (GPS) [15], and various airport landing aids and radar systems, but have little to none investigation in tracking human motion.

In the light of the previously characterized tracking technologies, we can conclude that no single technology or device can overcome all the problems that arise for every tracking application. However by acknowledging the key necessities of the system, a noteworthy solution can be attained, employing one or more of the available technologies. The preferred approach relies on a combination of optical and inertial technologies, capable of complementing each other in the presence of hindrances, pursuing a reliable tracking solution, that can be achieved with the appropriate estimation approach.

### 1.1.2 Estimation Techniques

Concurrently with the use of sensors, measurements errors will be present. In order to lessen them and obtain a correct measurement, multiple sensors data and fusion through filters is the ordinary approach. Additionally, position and orientation sensors have numerous limitation and usually require a rigid setups. Hence, for position and orientation estimation through accelerations and angular rates a estimation process is required.

In this fashion, Kalman Filter, Particle Filter and Complementary Filter are briefly deliberated.

### Kalman Filter

Kalman filter constitutes an algorithm for linear systems that infers an estimate from a series of inaccurate and uncertain measurements. It iteratively estimates a prediction and correction of the state, taking into account the expected values, the previous values of the state and the information provided from the sensors. For a state-space model, the real state $\mathbf{x} \in \Re^n$ of the system can be granted by the linear stochastic difference equiation 1.1 [16].

$$\mathbf{x_k} = A\mathbf{x_{k-1}} + B\mathbf{u_k} + \mathbf{w_k} \tag{1.1}$$

Where, $\mathbf{x_k}$, represents the state vector at step k, $A$ is the state transition model which is applied to the previous state $\mathbf{x_{k-1}}$ and $B$ is the control-input model which is applied to the control vector $\mathbf{u_k}$. The measurements, $\mathbf{z} \in \Re^m$, represented in equation 1.2, where $H$ is the observation model matrix and $\mathbf{z_k}$ is the measurements vector at step k.

$$\mathbf{z_k} = H\mathbf{x_k} + \mathbf{v_k} \tag{1.2}$$

The variable $\mathbf{w_k}$ and $\mathbf{v_k}$ represent the process and measurements noise respectively. They are accepted as independent to each other, white and with normal probability distributions. Where $Q$ and $R$ constitute the process and measurement noise covariances respectively.

$$p(w) \sim N(0, Q), \qquad p(v) \sim N(0, R) \tag{1.3}$$

The filter estimates the state of a process employing a form of feedback control. Kalman equations are divided in two groups: prediction (1.4) and correction (1.5). Where, $\mathbf{P}$ constitutes the error covariance which is computed in both the prediction and update states and $\mathbf{K}$ is the Kalman gain.

$$\begin{aligned} \hat{\mathbf{x}}_\mathbf{k}^- &= \mathbf{A}\hat{\mathbf{x}}_{\mathbf{k-1}} + \mathbf{B}\mathbf{u_k} \\ \mathbf{P}_\mathbf{k}^- &= \mathbf{A}\mathbf{P_{k-1}}\mathbf{A^T} + Q \end{aligned} \tag{1.4}$$

$$\begin{aligned} \mathbf{K_k} &= \mathbf{P}_\mathbf{k}^-\mathbf{H^T}(\mathbf{H}\mathbf{P}_\mathbf{k}^-\mathbf{H} + R)^{-1} \\ \hat{\mathbf{x}}_\mathbf{k} &= \hat{\mathbf{x}}_\mathbf{k}^- + \mathbf{K_k}(\mathbf{z_k} - \mathbf{H}\hat{\mathbf{x}}_\mathbf{k}^-) \\ \mathbf{P_k} &= (I - \mathbf{K_k}\mathbf{H})\mathbf{P}_\mathbf{k}^- \end{aligned} \tag{1.5}$$

Kalman filter is widely employed in numerous applications, ranging from engenieering in aircraft navigation to economics [17]. Due to its reputation, extensions and generalizations where presented, such as extended Kalman filter [18] and unscented Kalman filter [19] considered for nonlinear systems. The main hindrance of Kalman-based approaches in a few applications constitutes the computational complexity requirements, caused by multiple matrix operations, presented in (1.4) and (1.5).

### Particle Filter

Particle filters represent a set of algorithms, with a recursive implementation based on the Monte Carlo method, since they concern the employment of random samples to approach the probability distribution aimed to estimate. Particle filters approaches, implement the prediction-update transition of the filtering equation by directly using a particle algorithm [20].

Considering a probability distribution $p(x)$, with unknown samples yet proportional to another distribution $\pi(x)$ easily assessed. The samples from the distribution are represented by a set of particles, $x^{(i)} \sim q(x), i = 1, \cdots N$, each particle has a likelihood weight assigned to it, $w^{(i)} \propto \frac{\pi(x^{(i)})}{q(x^{(i)})}$, that represents the probability of that particle being sampled from the probability density function. A weighted approximation to $p$ can be produced by

$$p(x) \sim \sum_{i=1}^{N} w^{(i)} \delta(x - x^{(i)}) \tag{1.6}$$

Prior to particle filtering methods became renowned, the Kalman filter was the standard method for solving state space models. The Kalman-based approaches can be applied to either solve a linear Gaussian state space model, or when the linearity or Gaussian conditions do not hold, the solution can be achieved through its variants. However, for highly non linear and non-Gaussian problems they fail to provide a reasonable estimate. Particle filtering techniques offer a satisfactory alternative method. Applications can be seen in a wide range of fields, such as object tracking [21] or neuroscience [22] to name a few. Regardless, their computational power demands are even higher than Kalman-based solutions.

**Complementary Filter**

Complementary filters constitute a set of algorithms that fulfil the demand for low computational power filters with frequency filtering properties for linear systems. The approach resides in applying a low-pass filter to some low-frequency signals and high-frequency ones pass through a high-pass filter, merging both in the final phase. A simple example can be pictured in the following figure 1.1, where $x$ and $y$ are noisy measurements of some signal $\bar{Z}$ and $z$ is the estimate of $\bar{Z}$ composed by the filter, $G(s)$ can be regarded as a low-pass filter, to filter out the high-frequency noise of $y$. Complemented by $[1-G(s)]$, a high-pass filter that filters the low-frequency noise in $x$.

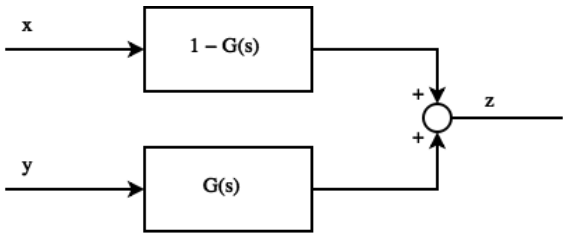


**Fig. 1.1:** Basic representation of a Complementary Filter

Typically, complementary filters disregard adjustability, nevertheless, variants such as complementary filter developed by Mahony *et al.* [23], and gradient descent, sugested by Madwick *et al.* [24] represent implementations capable of effectively adapt to inconsistencies.

After considering the most renowned estimation approaches, lets consider the solutions already present in the consumer market seizing a broader understanding of potentialities of the consummated systems.

### 1.1.3 Input Devices for Consumers

At the present time, there are some input devices available for purchase of the shelf with tracking capabilities appropriate for virtual environment usage. In spite of the fact that the main purpose of some of these systems not being object manipulation, they worth being considered in the ambiance of this work.

### Game controllers

Most game controllers were not intended for object interaction in a virtual environment. Despite that, researchers and game developers took advantage of them being the first kind of controllers that offered the possibility of motion tracking presented to consumers. Controllers like the Nintendo's Wii remote, PlayStation Move or Razer's Hydra present technology that allow position tracking and navigation, attributes that can be adopted for object manipulation.

The Nintendo Wii remote [25], regularly addressed as wiimote, features a three-axes accelerometer empowering the ability to sense acceleration while also providing an orientation estimate. Additionally, it is equipped with an optical sensor that combined with the console's sensor bar grant inside-looking-out position tracking. Subsequently, an extension was released, Wii MotionPlus, with a three-axes gyroscope increasing the precision of the orientation tracking. Besides tracking, the remote provides audio and haptic feedback.

Developed by Sony, the PlayStation Move [26] incorporates an accelerometer and a gyroscope for attitude tracking. The controller also contains a magnetometer for drift correction. The orb at the head of the controller, dynamically selects its color based on the environment, in order to distinguish from it. The colored light serves as an active marker, allowing for position tracking in combination with a camera, PlayStation Eye.

Razer Hydra [27], previously acknowledged as Sixsense TrueMotion, represents a noteworthy application of a magnetic tracking solution. The absolute position and orientation of the controllers is determined by reading a weak magnetic field created by the base station.

### Virtual Reality Controllers

As consumer available HMD's where released, specialized controllers for interaction follow along. There are two relevant areas of development representing this type of input devices exist, mobile VR and standard VR. Between the two aforementioned areas, a trend can be easily identified. Mobile input devices are generally designed with one-hand use in mind, meanwhile, for standard VR a two-handed approach is commonly employed. The systems worth consideration are, Daydream controller and Samsung's Gear VR controller, for the mobile area, and Oculus Touch and HTC Vive controller for standard VR.

The Daydream controller [28] is developed by Google, for the Daydream mobile VR platform.

The controller is equipped with three sensors, an accelerometer, a gyroscope and a magnetometer, that are employed for attitude tracking.

The Gear VR controller [29] came to life as a result of a joint partnership by Samsung and Oculus, for the Gear VR headset. The release of the controller granted a more interactive and immersive experience, prior to that, the interactions where made by either utilizing a bluetooth gamepad or controls on the headset. For tracking, only attitude is considered with an array of sensors composed by an accelerometer, a magnetometer and a gyroscope.

Oculus Touch [30] two mirror-image controllers attain high precision tracking through the same Constellation tracking system as the one of the Rift HMD. Constellation is an optical-based outside-looking-in tracking system, that performs position tracking by distinguishing the infrared led markers placed on the devices from the background environment, though an external camera. The attitude tracking is done by inertial measurement units (IMUs) located inside the devices. Furthermore, the system is designed to provide natural resting position while holding the controller, while also sensing finger position for hand gesture capabilities provided through a matrix of sensors mounted inside.

HTC Vive controller [31] mutually developed by Valve and HTC, employs an equal tracking system as the Vive HMD called Lighthouse. With resemblance to Constellation, attitude tracking is provided by IMUs, contrastingly an inside-looking-out system based on lasers is adopted for position tracking. The laser-based system is composed by two base stations, each of them comprising an infrared beacon and two spinning laser emitters, these work as reference points allowing for the three-dimensional position tracking to be performed.

Furthermore, two meaningful systems exist worth mentioning, although not being available for consumers, HTC Vive tracker [32] and Tactical Haptics [33]. HTC Vive tracker benefits from the aforementioned Lighthouse tracking, while allowing it to be attached to any object. The later, Tactical Haptics, can be connected to any of the standard VR tracking systems previously addressed enabling them with touch feedback, conveying motion and force.

## 1.2 Objectives and Outline

Taking into consideration each and every theme heretofore addressed, the purpose of this work is to develop an instrumented object enabling a new way for object manipulation within virtual and

mixed reality worlds, improving the user experience by awarding him with full haptic feedback. The proposed instrumented object, it's an everyday cube-shaped object, which comprises force sensing, and wireless navigation capabilities at an inexpensive cost.

Following earlier research in virtual and mixed reality immersion techniques demonstrated in a series of fields, such as psychological therapy [34] or remote operation [6, 7] to name a few. The device developed in this thesis is intended to be used as a complementary tool for experimental scenarios, increasing the immersion they present.

The key contributions of this work are the following:

- An instrumented object with grasping force sense capabilities;
- An UDP based communication protocol capable of being used for several instrumented objects in the same network;
- A serious game application with the helps to improve visual-motor coordination.

This dissertation is organized in the following way:

- Chapter 2: Pose estimation and tracking of objects. Relates to the approach employed for tracking the instrumented object proposed in this thesis. Firstly, a brief overview of the visual estimation techniques capable of tracking is presented, followed by the implementation employed. Secondly, the inertial approach is examined, calibration requirements are presented and both implementations for attitude and position are granted. Finally, the fusion of both the tracking methods is addressed and their results presented.
- Chapter 3: Connecting an Object to the IoT. Exposes the integration of the instrumented object in the network is addressed, presenting the developed communication protocol for it's integration.
- Chapter 4: Development of Prototype for Motion Tracking and Force Sensing. Addresses to the design evolution and improvements along the development of the instrumented object, culminating in the presentation of the final design. Furthermore, some implementation scenarios for the use of the developed system are conferred.
- Chapter 5: Conclusions and Future Work. It presents the final conclusions from the work and suggested future work in order to further improve the solution, as well as, new research and application scenarios.

# Chapter 2

# Pose estimation and tracking of objects

Taking into consideration Virtual and Mixed Reality, the term immersion is often times tied to it. Immersion is a dual-natured phenomenon, it can mean both representational immersion or participatory immersion. The first one, relates to a imitative representation of the real world, a depiction of an environment employing high graphical execution that comes close to photo-realistic. The latter, represents the interaction between users and environment, that we must endeavour to be bidirectional, much the same way they performs in real life. Habitually, the representational aspect is accurately characterized, on the contrary, the participatory aspect is often overlooked and could see some improvements.

In an effort to improve the previously stated, we propose an improvement of interaction through the use of passive haptics [35]. Passive-haptic devices are real objects, either rigid or deformable, which grant feedback to the user exclusively by their shaped, texture or other native properties. In distinction to the active-haptic systems, which feedback is furnished by the computer. In pursuance of employing passive haptics in immersive virtual environments, the elected object requires at least one of the following, to be externally tracked or bear embedded tracking capabilities.

The developed solution employs an hybrid approach, chosen among the previously mentioned tracking methods 1.1.1. It combines a visual and an inertial approach for a more reliable tracking, which will be further explained in the following sections.

## 2.1 Visual Pose Estimation

One of the underlying goals of computer vision is to discern properties that are inherent to a scene by analysing one or several images of this scene. Within this paradigm, a fundamental procedure is the determination of the position and orientation of the sensing system (the camera) with respect to objects in the scene. The contrary, can assuredly be accomplished visual estimation techniques for object tracking have been around since the development of the first cameras, among the approaches are feature matching from previously known model [36] or by marker tracking [37]. Our proposed approach identifies the real world object and estimates its position and orientation using a combination of visual markers and a camera [38].

In this section a concise explanation of the camera model used is presented, followed by, a couple of pose estimation approaches. Afterwards, a description of our pose estimation technique is presented, concluding with a succinct interpretation of the end results obtained.

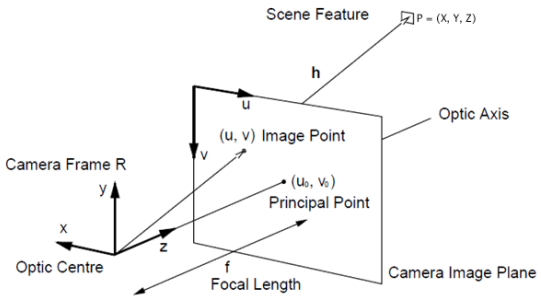


**Fig. 2.1:** Pinhole camera model.

Our system makes use of the pinhole camera model, illustrated in Figure 2.1, where a scene view is formed by projecting three-dimensional points into the image plane using a perspective transformation.

$$s\mathbf{p} = K[R|t]\mathbf{P}, \qquad s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2.1}$$

where,

- $(X, Y, Z)$ represent the coordinates of a 3D point in the world coordinate frame;
- $(u, v)$ are the coordinates of the projection point $\mathbf{p}$ in pixels;

- $K$ is the matrix of intrinsic parameters;
- $(u_0, v_0)$ is a principal point that is usually the image center;
- $f_x, f_y$ are the focal lengths expressed in pixel units.

Accordingly, if an image from the camera is scaled by a factor, the previous parameters should be scaled by the same factor. The matrix of intrinsic parameters, $K$, is independent from the scene pictured, thus, unless the focal length, $(f_x, f_y)$, is altered, it can be employed indefinitely upon estimation. The extrinsic parameters matrix is represented by the joint rotation-translation matrix, $[R|t]$, which is used to characterize the camera motion around a static scene or the other way around. The preceding matrices, can be obtained in a number of ways, which will be briefly described.

### 2.1.1 Pose estimation for non-planar objects

Non-planar objects pose can be estimated by solving the Perpective-n-Point problem, the problem was initially composed for evaluating the camera position and orientation in relation to an unmoving scene. However, it can also be employed inversely to acquire the pose of an object in relation to a stationary camera. The intent of Perpective-n-Point camera pose determination, commonly known as the PnP problem, is to constitute the relative position between the camera and scene from $n$ acknowledged correspondences of 3D reference points and image points. The first algebraic solution, employs a three-point algorithm and was proposed in 1841 by a German mathematician, with multiple variants developed ever since. Fischler and Bolles [39], introduced a distance-based definition, widely used in the computer vision, concurrently with a RANSAC paradigm to detect outliers in the data. They also coined the problem PnP. Horaud *et al.* [40] presented a P4P approach supported by a transformation-based definition. The aforementioned definitions, the distance-based and the transformation-based, were extensively used in the studies of the PnP problem being broadly present in the literature [41, 42]. Recently, both definitions where mathematically proven to be equivalent by Wu *et al.* [43].

The PnP problem is commonly solved using least-squares techniques, these require the computation of numerical solutions. In order to achieve a stable solution, a numerous among of points are required inherently amplifying the complexity of the computation. For these reason, researchers pursued to ascertain the minimum number of points required for a solution. A finite

number of solutions is only granted for 3 or greater number of points. The use of three points, requires additional information to guarantee the uniqueness of the solution. Four points ordinarily suffice for uniqueness outside of certain known critical configurations in space, as demonstrated by Thompson [44]. A possible approach for the pose determination follows.

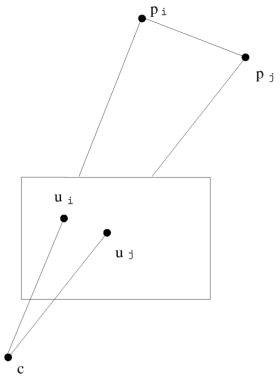


**Fig. 2.2:** Basic geometry of the camera pose determination for each pair of correlations between 3D reference points and their image.

Given a calibrated camera at $c$ and $n$ correlations between 3D reference points $p_i$ and their image $u_i$, each pair of correlations $p_i \leftrightarrow u_i$, $p_j \leftrightarrow u_j$ provides a constraint on the unknown camera-point distances $x_i = \| p_i - c \|$ and $x_j = \| p_j - c \|$, pictured in figure 2.2:

$$d_{ij}^2 = x_i^2 + x_j^2 - 2x_i x_j \cos\theta_{ij} \tag{2.2}$$

where $d_{ij} = \| p_i - p_j \|$ is the known distance between the two arbitrary reference points and $\theta_{ij}$ is the angle to every pair of this points from the camera's optical center. The cosine of the angle is directly computed from the coordinates of the image points and the calibration matrix, $K$, of the camera following:

$$\cos\theta_{ij} = \frac{u_i^T C u_i}{(u_i^T C u_i)^{1/2}(u_j C u_j)^{1/2}} \tag{2.3}$$

where $C = (KK^T)^{-1}$. This formula is derived from the direction vector of the viewing line $K^{-1}u$ for the given image point $u$. This quadratic constraint can be rewritten as

$$f_{ij}(x_i, x_j) = x_i^2 + x_j^2 - 2x_i x_j \cos\theta_{ij} - d_{ij}^2 = 0 \tag{2.4}$$

For $n = 4$, a vastly constrained system of six polynomials $f_{ij}(x_i, x_j) = 0$ is achieved for

the four unknowns $x_i, i = 1, \cdots, 4$. One elementary approach is to consider subgroups of three of the four points, solve the polynomial equations for each subgroup and to conclude discover the shared solution. After acquiring the camera-point distances, $x_i$, those are converted into the camera-centered 3D coordinates $\hat{p_i} = x_i K^{-1} u_i$ of the reference points in space. The rotatition, $R$, is obtained with least-square approach in closed-form using quaternions [45], the determination of the translation, $t$, and the scale result from the estimate of the rotation.

### 2.1.2 Pose estimation for planar objects

For planar objects the pose estimation can be assessed from the decomposition of the homography matrix. The planar homography is a non-singular linear association among points on planes. Images of points on a plane in one perspective are associated to corresponding image points in another perspective by a planar homography using a homogeneous representation. This is a projective relation since it exclusively relies on the intersection of planes with lines. The homography achieved by a plane is distinct up to a scale factor and is determined by 8 degrees of freedom. It can be estimated from the matching of 4 points or lines in two perspectives. Every matching pair contributes to two constrains, fixating two degrees of freedom. Furthermore, the degrees of freedom can be fixed by corresponding other parametric and non-parametric curves or contours present in the image. Alternative properties in the image, for instance texture or color can also be employed to compute the planar homography [46]. Succeeding a feasible estimation of the pose using the homography is presented.

$$s_i \begin{bmatrix} x_i^{'} \\ y_i^{'} \\ 1 \end{bmatrix} \sim H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{2.5}$$

Considering a couple of given groups of bi-dimensional corresponding points, $x_i \leftrightarrow x_i^{'}$. The perspective transformation, $H$, is computed through a least-square approach, minimizing the back-projection error, $\epsilon$, where $h_{ij}, i = 1 \cdots 3, j = 1 \cdots 3$ are the indices of the homography matrix.

$$\epsilon = \sum_i \left( x_i^{'} - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 + \left( y_i^{'} - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \right)^2 \tag{2.6}$$

By applying single value decomposition to $Q = U\Sigma V^T$, where $H_i, i = 1, 2, 3$ are the columns

of the homography matrix and $K$ the matrix of intrinsic parameters, the rotation matrix $R$ and the translation vector $t$ can be estimated as presented in Equation 2.7.

$$Q = \left[ \frac{K^{-1}H_1}{\|K^{-1}H_1\|_2} \middle| \frac{K^{-1}H_2}{\|K^{-1}H_1\|_2} \middle| \frac{K^{-1}H_3}{\|K^{-1}H_1\|_2} \right]$$
$$T = \frac{K^{-1}H_3}{\|K^{-1}H_1\|_2}; \qquad R = \Sigma V^T \tag{2.7}$$

### 2.1.3 Pose estimation of markers using the planar approach

In our approach, we use bi-dimensional square-shaped markers with a previously established size to obtain absolute position and orientation estimations, similar approaches are extensively present in literature with multiple aspirations [47, 48, 49]. The markers have a white background, delimited by a couple of distinct sized squares with the space between them coloured in black. Within the delimited area two important sections exist. First, the orientation point, closer to the boundaries assists in the detection of the marker's orientation, solving any ambiguity. The second, represents the identification section, that uniquely identifies each marker, consisting of 3 rows with 8 points in total, corresponding to a total of $2^8 = 256$ different representations. The described marker with a few pertinent zones highlighted, can be pictured in the following figure.

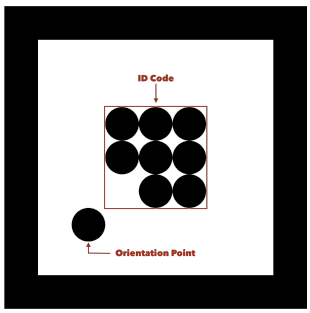


**Fig. 2.3:** Example of a bi-dimensional marker employed with pertinent zones highlighted

Taking into consideration both pose estimation methods formerly introduced, the homography decomposition approach (2.1.2) was preferred, for representing a more reliable solution considering an equal extent of points. This can be achieved due to the added constraints of being a planar approach. Since our approach uses square markers of know size, we are able to compute the homography between the model and the marker on the image frame. After thresholding of the input image from the camera, our approach looks for regions whose contours can be fitted by four line segments and inwards present the same shape, basically a square inside a square representing the marker, pictured in figure 2.4. The coordinates of both the four vertices of the inside

and outside square, are employed for computing the homography matrix and sequentially the rotation and translation. Following, the orientation point of the marker, pictured in Figure 2.3, is sought. Depending on its location, the initial rotation and translation might be revised, after reordering the vertices of the inside and outside squares. Finally, the identification code of the marker, which allow to differentiate the markers, is decoded and saved.

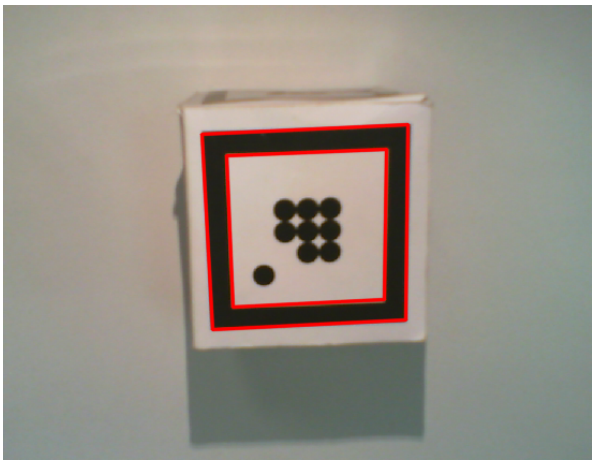

**Fig. 2.4:** Example of marker detection with inside and outside squares zones highlighted

In order to access the reliability of the pose estimation of the system, a simple series of tests were performed, where the visual marker was placed at a known distance of the camera in multiple orientations. The position estimation performed greatly for the intended application scenario, proving to be a reliable solution, presenting an error of $\leqslant 2mm$ in the position estimation and a $< 1^{o}$ in the orientation.

Figure 2.5, presents two datasets of the pose estimation obtained through the aforementioned method, while performing the same motion. We can quickly picture the difference between the two dataset, in one of them the marker was occluded during its tracking, leading to the lack of estimation during that period of time, while on the other the marker was always visible. The presented visual tracking method, represents a satisfactory tracking solution, however as every visual tracking system, it suffers from a collection of known problems such as occlusion, illumination, fast movement of the tracking object, to name a few. In an effort to produce a more reliable solution, a complement to this approach is presented in section 2.3.

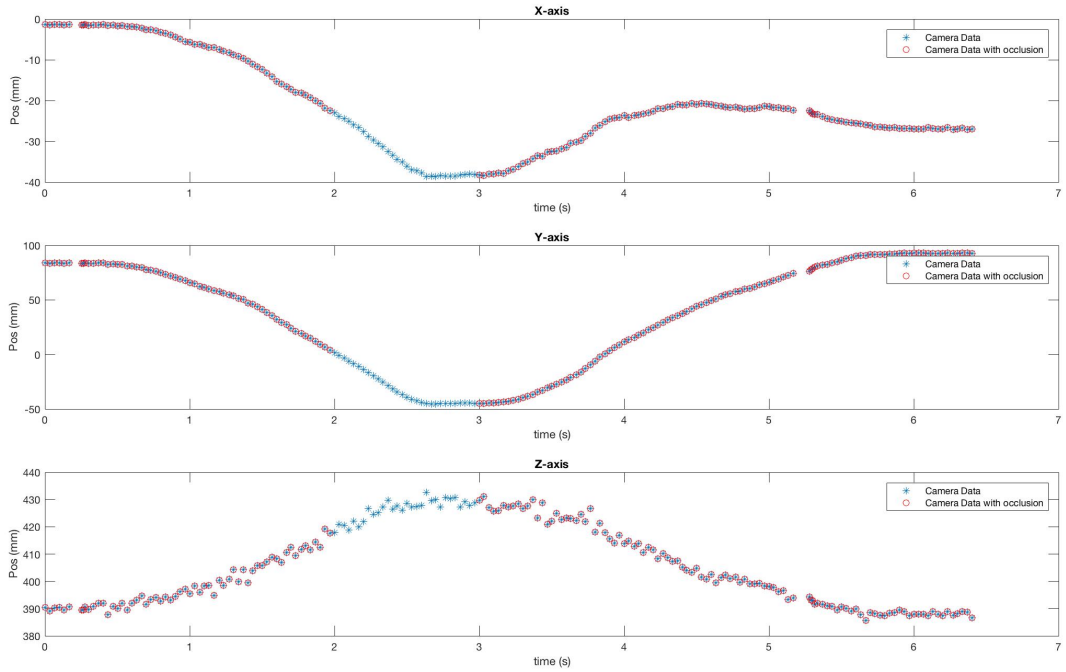


**Fig. 2.5:** Postion estimation computed by the presented approach with and without occlusion.

## 2.2 Strapdown Inertial Tracking

The operation of inertial tracking revolves around the laws of classical mechanics. Newton's laws tell us that the motion of a body continues consistently in a straight line unless a disturbance by an external force is applied on the body. Additionally, this disturbing force will produce a measurable acceleration on the body. Given the capability of measuring the acceleration, the alterations on the system are possible to infer. Usually, inertial system comprise three orthogonally-mounted accelerometers, each capable of sensing acceleration in a single direction. Additionally, to navigate with respect to the frame of reference, it is fundamental to record the direction of the accelerometers. Orientation of the accelerometers is tracked through the use of gyroscopic sensors that attain the rotational motion of the body.

Hence, inertial tracking is a navigation approach in which measurements obtained by accelerometers, gyroscopes (and in some cases magnetometers) are utilized to trail the position and orientation of an object relative to a known starting point, orientation and velocity. Our system uses a strapdown approach, being firmly mounted on the object, as a result sensor measurements are made in the navigation system's frame rather than in the frame of reference.

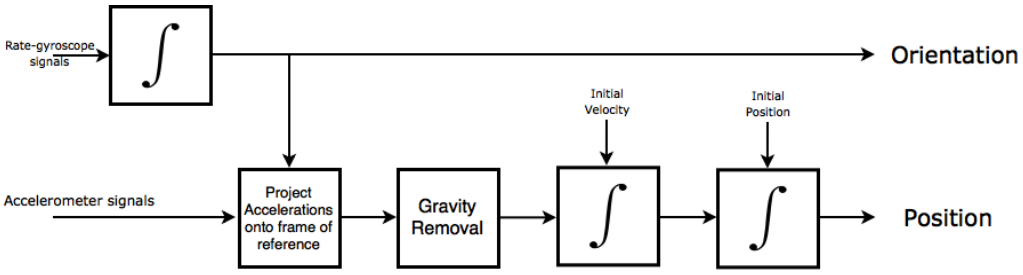


**Fig. 2.6:** Generic representation of an Inertial Navigation System.

### 2.2.1 Inertial Measuring Unit

An Inertial Measuring Unit (IMU) is a multi-axis sensor on a chip, generally based on MEMS technology. The system is capable of measuring acceleration, angular rate, and sometimes the magnetic field surrounding the body. These chips represent a extremely low-cost, lightweight and compact solution. Regardless, they display some flaws, the measurements they provide are relatively noisy, can present offset bias, different scale factors and non-orthogonality between axes. These hindrances led to the advancement of highly rigorous calibration and manufacture procedures, providing IMU approaches which are extremely accurate, however becoming extremely expensive. Depending on the application and rigour required, straightforward calibration methods may suffice, providing an affordable and reliable solution. Following, the calibration approaches employed, will be presented.

#### 2.2.1.1 Gyroscope Calibration

Gyroscopes measure the rate of rotation of a body, using Coriolis acceleration effect on vibrating masses to detect inertial angular rotation. The Coriolis force, acts in a direction, that is perpendicular to both the axis of vibration and the axis about which the rotation is applied [50].

As formerly outlined, MEMS technologies suffer from some weaknesses. Gyroscopes generally, present a non-zero value at rest, referred as bias, representing a fixed value to the sensor's output response, possibly displaying the system as to be rotating when it is indeed stationary. The result is a constant accumulation of angle measurement error over time. Additionally, scale-factor errors contribute to measurement corruption, but only when motion occurs.

This interpretation of MEMS gyroscopes leads to the model of this sensor in Equation 2.8. There, the real angular velocity, $\omega_{measured}$, is the sum of the real angular velocity, $\omega_{real}$, plus the

gyroscope offset bias, $\omega_{bias}$, multiplied by the scaling factor, $K_g$.

$$\omega_{measured} = K_g(\omega_{real} + \omega_{bias}) \tag{2.8}$$

Thus, in order to calibrate the gyroscope the offset bias and the scale-factor must be determined. Averaging the gyroscope output while holding the device in a constant position is a simple method for estimating bias errors. The Allan Variance Curve for a gyroscope helps in analysing the trade-off between test time that is, length of average and bias accuracy. Typically, scale-factor errors characterize using a servo-motor, which employs an optical encoder for precise rate control. In this approach, gyroscopes rotate at known rates while providing output measurements. The described system also allows for axis misalignment correction, Equation 2.9 represents the model for this rectification, where $T_g$ represent a misalignment between sensor and body axis.

$$\omega_{measured} = T_g K_g(\omega_{real} + \omega_{bias}) \tag{2.9}$$

Although this approach is effective, a simpler approach exists. Assuming bias-error correction, the integration of a gyroscope provides another mechanism for observing scale factor. In this case, the scale-factor error is the ratio of the measured angle to the actual angle displacement. In the following **??**, an example of the described calibration method is presented, for comparison reasons the scale factor is also applied to the raw data.

#### 2.2.1.2 Magnetometer Calibration

Broadly used magnetometers, are Hall-effect transducers with a magnetic concentrator. An optimal magnetometer, should consistently provide the direction of the Earth's magnetic North. However depending on the location, the magnitude of the field and in addition the angle between the magnetic North and the geographic North also varies. Notwithstanding, considering the application at hands the aforementioned hindrances do not represent a problem. On the other hand, the multiple local magnetic fields encountered in the surrounding present as a complication [51]. These undesirable magnetic fields can be divided in two types of distortions: *soft iron* and *hard iron*.

The *hard iron* distortions are regularly easier to compensate, if the working environment does not change, they are characterized by generating a constant offset in the measurement readings.

These distortions emerge from permanent magnets, ferrous materials, any material that as a permanent field, and that preserves their pose relative to the sensor. The *soft iron* distortions appear when a magnetically soft material is present in the surroundings of the sensor, most indoor environments contain appliances that create or distort magnetic fields. In this case the bias, depends of the sensor's orientation and position on the environment and they main not keep consistent over time. Figure 2.7 present a two-dimensional representation of the effects in magnetic field measurements, produced by the aforementioned distortions.

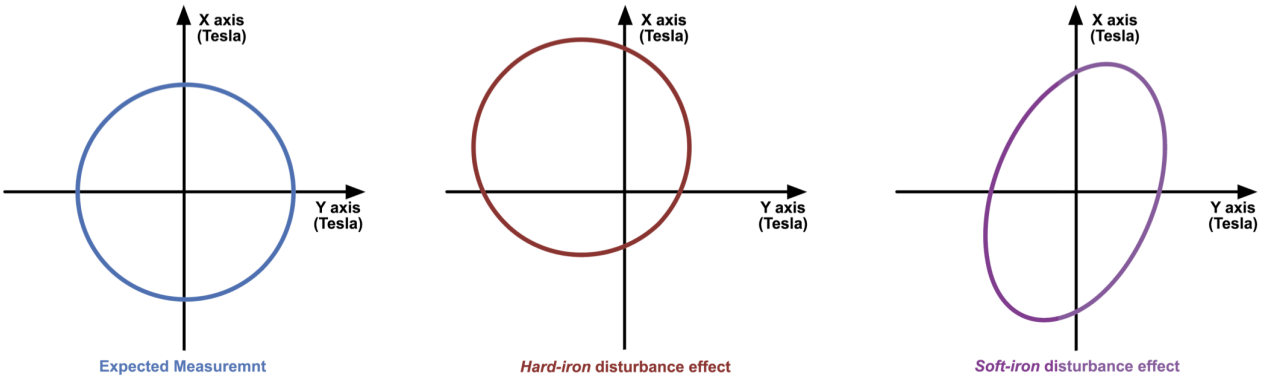


**Fig. 2.7:** Two-dimentional representation of the disturbances to Earth's magnetic field readings.

Considering the description of the sensor and disturbances, the magnetometer can be modelled as Equation 2.10. The *soft iron* distortions are modelled as a $3 \times 3$ matrix ,$K_s$, the *hard iron* distortions as a constant offset bias, $m_h$ and $R_m$ is a rotation matrix required to align the magnetometer with the axes of both gyroscope and accelerometer, IMUs frequently show different orientations for both.

$$m_{corr} = R_m.K_s.(m_{raw} - m_h) \tag{2.10}$$

Several procedures and algorithms have been proposed to perform the calibration [52, 53, 54], that diminishing the distortions. We implemented a calibration proposed by Merayo *et al.* [55], based on the linear least square estimator. Taking into consideration, that any vector can be expressed as a linear combination of an independent set of vectors, the intrinsic reference-frame system of a specific sensor can be defined as follows:

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = T \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \gamma_1 & \gamma_2 & 0 \\ \zeta_1 & \zeta_2 & \zeta_3 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \tag{2.11}$$

The relation between the orthogonal and non-orthogonal bases, $T$, represented in Equation 2.11 is linear. Accordingly, the components are related also by a linear transformation, that is the inverse of the matrix $T$ transposed. In the Equation 2.11, $\omega_i$ represent the orthogonal basis and $\mu_i$ are the measurements in the non-orthogonal one. The measurements, $m_i$, are centred by the offset vector, $O_i$, and normalised by the sensitivities, $s_i$.

$$\mu_i = s_i(m_i - O_i) \tag{2.12}$$

Therefore, the relation between the physical magnetic field in the intrinsic reference-frame system and the measurements is obtained by combining equations 2.11 and 2.12, giving:

$$\begin{bmatrix}\omega_1\\ \omega_2\\ \omega_3\end{bmatrix} = [T^{-1}]^T \begin{bmatrix}s_1(m_1 - O_1)\\ s_2(m_2 - O_2)\\ s_3(m_3 - O_3)\end{bmatrix} = A \begin{bmatrix}m_1 - O_1\\ m_2 - O_2\\ m_3 - O_3\end{bmatrix} \qquad A = \begin{bmatrix}s_1 & -\frac{\gamma_1}{\gamma_2}s_2 & \frac{\gamma_1\zeta_2 - \gamma_2\zeta_1}{\gamma_2\zeta_3}s_3\\ 0 & \frac{1}{\gamma_2}s_2 & -\frac{\zeta_2}{\gamma_2\zeta 3}s_3\\ 0 & 0 & \frac{1}{\zeta_3}s_3\end{bmatrix} \tag{2.13}$$

The implemented calibration was compared with a more simple approach. This approach computed the *hard iron* distortion, $m_h$, by averaging the recorded the maximum and minimum in each of the axis. Subtracting the computed average from the measurement data centres the response surface. The compensation *soft iron* disturbances, $K_m$, was performed by a scale factor calculated from the ratio of the average of the maximum and minimum samples in the considered axis and the average of all axis. This is just a simple orthogonal rescaling granting some additional correction for scale bias.

$$m_h = \begin{bmatrix}\frac{max(m_x) + min(m_x)}{2}\\ \frac{max(m_y) + min(m_y)}{2}\\ \frac{max(m_z) + min(m_z)}{2}\end{bmatrix}; \qquad s = \begin{bmatrix}\frac{max(m_x) - min(m_x)}{2}\\ \frac{max(m_y) - min(m_y)}{2}\\ \frac{max(m_z) - min(m_z)}{2}\end{bmatrix}; \qquad K_m = \frac{\bar{s}}{s} \tag{2.14}$$

Figure 2.8 present an example of the both previously mentioned calibration methods, as well as the raw measurement readings from which the calibration was computed. The data, can be collected in a couple of fashions, either by doing figure-eight patterns or by circular motions while tilting the sensor. As expected, both calibrations present clear improvements in regards to the raw data. The employed calibration approach, demonstrates a closer attempt to the perfect origin-located circle in all the axis.

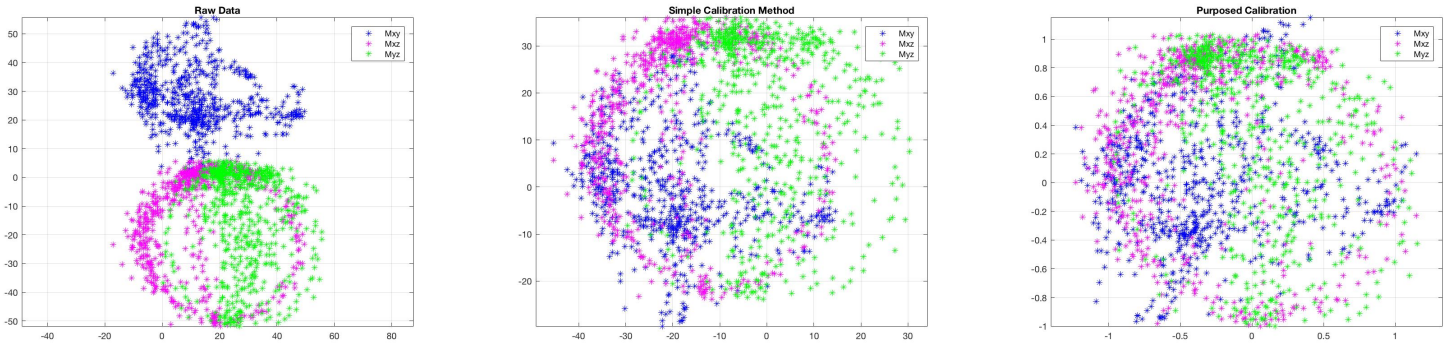


**Fig. 2.8:** Left: Raw data measured from the magnetometer; Center: Simple calibration; Left: Implemented Calibration.

### 2.2.2 Attitude Tracking

The computation of attitude is particularly critical in a strapdown system, it comes with no surprise that this topic has been subject of vast research [56, 57, 58]. The attitude of most inertial tracking system, relative to the frame of reference, generally called global frame, is tracked by mathematical integration of the angular velocity $\omega = [\omega_x, \omega_y, \omega_z]$, acquired by the system's rate-gyroscopes.

The quaternion attitude representation, at step $t$, that describes the rate of change of the global frame relative to the body frame ${}^{B}_{G}\dot{q}_{\omega,t}$ can be determined from Equation 2.16. ${}^{B}_{G}\tilde{q}_{t-1}$ is the quaternion representing the previous estimate of the orientation.

$$ {}^{B}\omega = [0 \quad \omega_x \quad \omega_y \quad \omega_z] \tag{2.15} $$

$$ {}^{B}_{G}\dot{q}_{\omega,t} = \frac{1}{2} {}^{B}_{G}\tilde{q}_{t-1} \otimes {}^{B}\omega_t \tag{2.16} $$

Accordingly, the orientation estimate, ${}^{B}_{G}\tilde{q}_{\omega,t}$, is obtained by integrating the quaternion derivate, as noticed in Equation 2.17, where $\delta t$ representes the sampling time.

$$ {}^{B}_{G}\tilde{q}_{\omega,t} = {}^{B}_{G}\tilde{q}_{t-1} + {}^{B}_{G}\dot{q}_{\omega,t} \cdot \delta t \tag{2.17} $$

Inferring the INS attitude only considering the gyroscope measurements will expose the approach to errors. The error, will proliferate through the system generating accumulated drift

in the integrated signal, representing a problem for attitude estimation and, more severely, to the position one. This prompt the commitment for the adoption of a solution with increased robustness. Acknowledging the aforementioned estimation techniques in 1.1.2, considering the kind of sensors and application used at this work, these approaches are widely employed. Regardless of Kalman and Particle filters producing superior solutions for the intricacy, they have computational requirements that are incompatible with the capabilities of the system. Therefore, the complementary filter was contemplated for the development.

#### 2.2.2.1 Complemetary Filter

The major ambition of the complementary filter is to optimally combine the attributes of each sensor, soothing the measurement errors in the estimates, through a low-pass filter and his complementary high-pass filter. Taking into consideration an IMU system, it contemplates merging the low frequency signals from the accelerometer and the magnetometer with the high frequency signals from the gyroscope. The applied implementation is based on the Mahony *et al.* [23, 59, 60] approach, which has low computational requirements and uses a porportional-integral(PI) feedback for the angular error ${}^{B}\omega_{\epsilon,t}$. The implementation of the filter can be effortlessly perceived high-level block diagram in Figure 2.9.

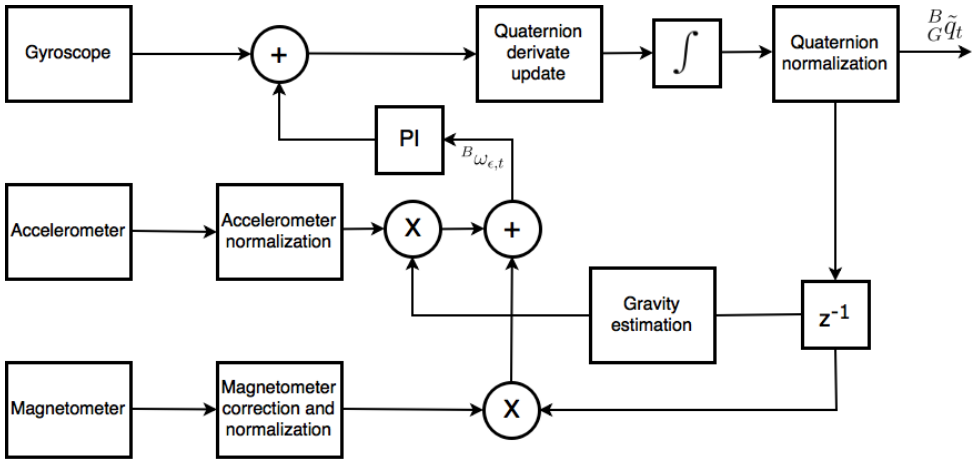


**Fig. 2.9:** Block diagram of the Mahony filter.

The angular error, ${}^{B}\omega_{\epsilon,t}$, represents an approximation between the measured orientation and the predicted one. Acquiring it, requires the calculation of the cross product between the gravitational field direction measured by the accelerometer, ${}^{B}\hat{a}_t$, and the predicted one, ${}^{B}\hat{g}_t$, plus the magnetometer measurements, ${}^{B}\hat{m}_t$, and the predicted one, ${}^{B}\hat{b}_t$, as follows.

$$^{B}\omega_{\epsilon,t} = {}^{B}\hat{a}_t \times {}^{B}\hat{g}_t + {}^{B}\hat{m}_t \times {}^{B}\hat{b}_t \tag{2.18}$$

$$^{B}\hat{g}_t = {}^{B}_{G}\tilde{q}^{*}_{t-1} \otimes {}^{G}\hat{g} \otimes {}^{B}_{G}\tilde{q}_{t-1} \tag{2.19}$$

$$^{B}\hat{b}_t = {}^{B}_{G}\tilde{q}^{*}_{t-1} \otimes {}^{G}\hat{b} \otimes {}^{B}_{G}\tilde{q}_{t-1} \tag{2.20}$$

The cinematic equation for the attitude of the sensor Equation 2.21, is acquired through a PI gain of the angular error described in Equation 2.22. The gain $k_p$ and $k_i$ represent the proportional and integral gains, respectively. The proportional gain regulates the frequency values that establishes the boundaries of the gyroscope sensor relevance against the accelerometer and magnetometer sensors data. The integral gain rectifies the gyroscope offset drift. The error is proportional to the sine of the angle between the measured directions and the expected ones (cross product). The feedback of the error through the PI controller into the gyroscope measurements, regulates the estimated orientation to accompany the reference vectors and, in this manner, the gyroscope offset and gyroscope offset drift is diminished.

$$^{B}_{G}\dot{q}^{*}_{t} = \frac{1}{2}{}^{B}_{G}\tilde{q}^{*}_{t-1} \otimes ({}^{B}\omega_t + \phi_t) \tag{2.21}$$

$$\phi_t = k_p \cdot {}^{B}\omega_{\epsilon,t} + k_i \cdot \int {}^{B}\omega_{\epsilon,t} \tag{2.22}$$

In order to assess the validation of the developed approach, the Invensense MPU-9250 motion tracking sensor was utilized for the evaluation. The sensor comprises a Digital Motion Processor$^{\text{TM}}$ (DMP$^{\text{TM}}$), capable of producing an orientation estimation adopted for comparison against the advanced solution. The values employed for the proportional and integral gains where computed in a previous application presenting an identical system configuration [61]. The results from both approaches can be observed in Figure 2.10, DMP$^{\text{TM}}$ represented in blue and the developed complementary filter expressed in red.

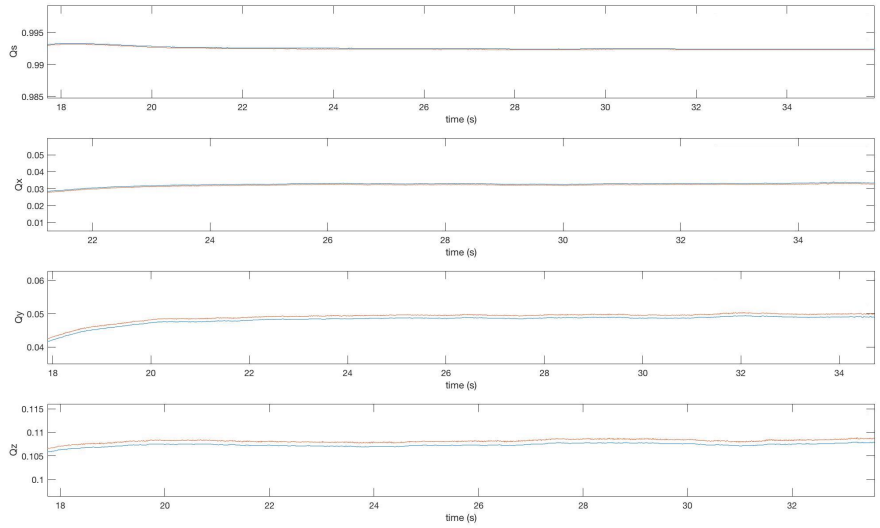


**Fig. 2.10:** Comparison between Complementary Filter developed and Orientation of the DMP.

### 2.2.3 Position Tracking

Regarding indoor low computational devices, the three-dimensional position estimate is still a intricate matter, since Global Positioning System (GPS) sensors cannot be used indoors. The employment of accelerometers, bears considerable inconvenience, considering the increase in error produced by the double integration of the measurements. There are two main sources of error, the orientation error bias along with the double integration of noise/bias. Accounting the aforementioned concerns, the position estimate from double integration can only obtain acceptable results if performed on small periods of time [62, 63, 64].

Commonly, position estimation from an accelerometer can be performed, using Equations (2.23 - 2.25). Where, the corrected acceleration vector in the global frame, ${}^{G}\tilde{a}_t$, is determined from the measurement of the accelerometer, ${}^{B}\tilde{a}_t$, without the gravity vector, ${}^{G}a_g$, that concedes the accelerations conceived from motion to be accounted for. Consequently, an integration of this assessment, Equation 2.24, produces the velocity of the sensor in the global frame, ${}^{G}\tilde{v}_t$, presupposing the precedent velocity is possessed. Lastly, the integration of the velocity provides the position, ${}^{G}\tilde{p}_t$ in the same manner. The $\delta t$ represent the duration of two sequential time steps from the accelerometer sampling.

$$ {}^{G}\tilde{a}_t = {}^{B}_{G}\tilde{q}_t \otimes {}^{B}a_t \otimes {}^{B}_{G}\tilde{q}_t^{*} - {}^{G}a_g \tag{2.23} $$

$$^{G}\tilde{v}_t = {}^{G}\tilde{v}_{t-1} + {}^{G}\tilde{a}_t \cdot \delta t \tag{2.24}$$

$$^{G}\tilde{p}_t = {}^{G}\tilde{p}_{t-1} + {}^{G}\tilde{v}_{c,t} \cdot \delta t \tag{2.25}$$

Considering a constant bias, $^{G}a_\epsilon$, on the accelerometer readings, according to Equation 2.26, will develop into an error in position, $^{G}p_\epsilon$, that increases quadratically with time, where $n$ is the step number. The bias must be corrected by the means of a calibration which, ideally, should be performed over time to correct any bias drift, the adopted approach will come to further enlightenment in 2.3.

$$^{G}p_{\epsilon,t+n} = \frac{1}{2}{}^{G}a_\epsilon \cdot (n \cdot \delta t)^2 \tag{2.26}$$

In order to admission the approach developed for modest sensors, a motion tracking sensor from Invensense MPU-9250 was considered for testing. The sensor comprises an accelerometer, a gyroscope, a magnetometer and a DMP$^{TM}$ capable of processing complex algorithms. The DMP$^{TM}$ contains digitally programmable low-pass filters (for the gyroscope and the accelerometer), an assessment was performed to select the best filter configuration in the interest of lessen the measurement errors. The measurements where performed with the sensor place on a table in a resting position. In Figure 2.11 the result for the low-passed DC component of the signal is presented.

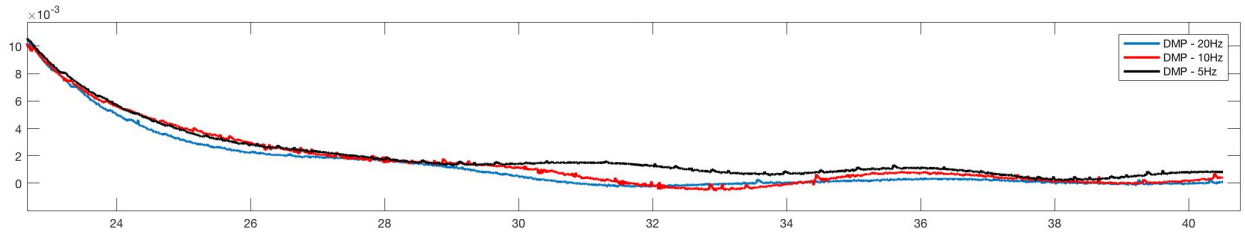


**Fig. 2.11:** DMP low-pass filter assessment example.

Furthermore, in Figure 2.12 an example of accelerometer readings without the gravity vector and bias offset is displayed. The outcome appears fairly close to the presented values on the sensor datasheet, which state a $\pm 60mg$ for the $X$ and $Y$-*axis* axis and $\pm 80mg$ for the $Z$-*axis*, suggesting a respectable estimation.

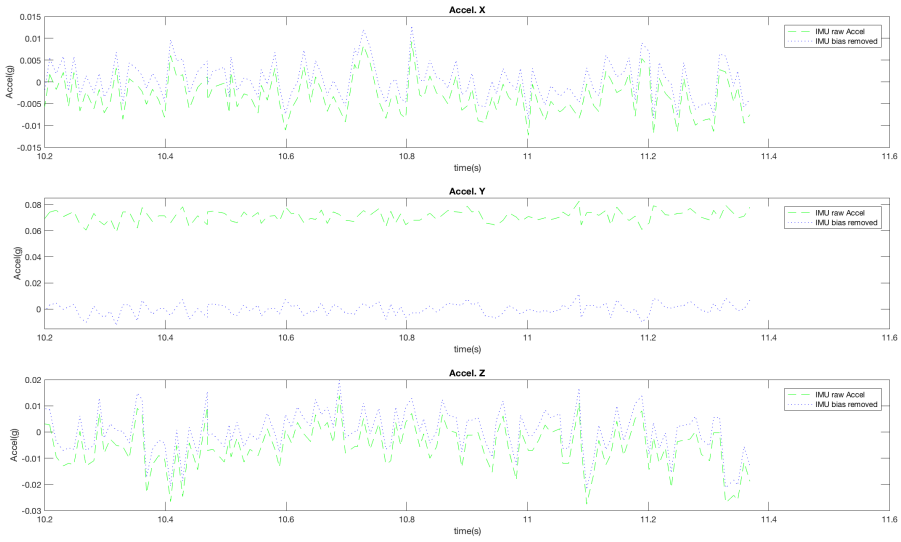


**Fig. 2.12:** Accelerometer raw reading and gravity and bias removal example.

Finally, in Figure 2.13 a comparison between the provided DMP$^{TM}$ orientation estimation, the proposed complementary filter solution and the raw measurements is displayed. Similarly to the previously displayed in Figure 2.10, the distinctness between both algorithms is modest.

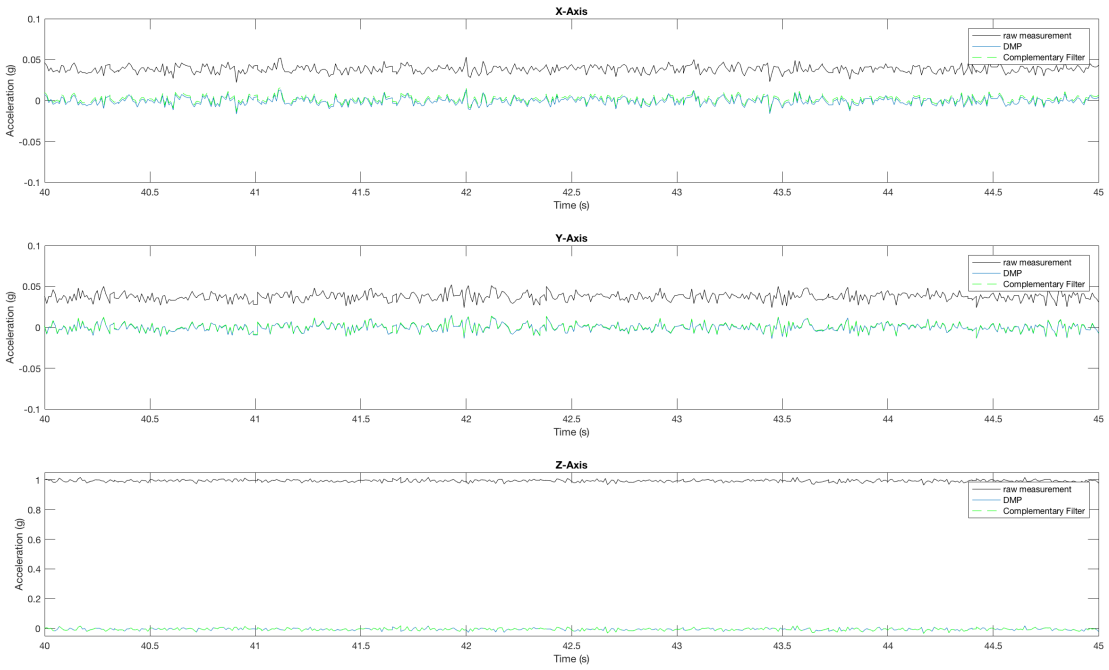


**Fig. 2.13:** Comparison between Complementary Filter developed, DMP orientation estimation and raw measurements.

Integrating white noise will produce what is commonly called as "random walk", it represents an error on the estimated position with zero mean, but which variance widens over time, proportional to $(n \cdot \delta t)^{\frac{3}{2}}$ [65]. Orientation errors induce incorrect projections of the acceleration signals onto the global coordinates frame, contributing to the integration of the acceleration in an inaccurate direction, likewise as the gravity acceleration not being accurately removed, which addresses back to the bias dilemma.

All these errors propagate and become eventually to accumulate in the position estimate. Considering all the error factors collectively, it is easily understandable that the position estimate is truly demanding, especially for long periods of time. In section 2.3 an approach with added reliability for position estimation during long periods of time is considered.

## 2.3 Visuo-Inertial Fusion for Object Tracking

Considering that both aforementioned approaches present some hindrances, combining both approaches seems a great alternative for a more reliable pose estimation. The sensor fusion technique allow for the estimation to produced at a higher rate than the visual system sample rate alone. Several approaches associating both visual-based and inertial estimation are widely present in literature [66, 67, 68].

In order to perform the suggested sensor fusion approach, both tracking systems should adopt the same frame of reference in order for the system to function correctly. The designated frame of reference was the same as the camera, since it preserves its position during the time the system is being operated, forging a less computational demanding procedure.

Equation 2.27 presents the equation employed for the change of referential in quaternion form, where ${}^{G}q_{cam}$ represents the camera reference frame and the ${}^{S}q_{imu}$ the inertial sensor reference frame ($S$).

$$ {}^{S}_{G}q_{dif} = {}^{G}q_{cam} \otimes {}^{S}q^{*}_{imu} \tag{2.27} $$

Taking into account the change in the reference frame, Equation 2.23 is inadequate for the approach in consideration, and requires to be rewritten accounting for the newly adopted reference frame, remodelling into Equation 2.28. With this amendment in mind, velocity and position estimations can still be applied as illustrated before, in equations 2.24 and 2.25 respectively.

$$ {}^{G}\tilde{a}_t = {}^{G}q_{res} \otimes ({}^{B}\tilde{a}_t - {}^{B}\tilde{a}_{bias,t}) \otimes {}^{G}q^{*}_{res}; \qquad {}^{G}q_{res} = {}^{S}_{G}q_{dif} \otimes {}^{S}q_{imu} \tag{2.28} $$

Regarding the estimate of the accelerometer bias, ${}^{B}\tilde{a}_{bias}$, as previously stated, it should be performed over time, in order to, account for changes in the evaluation. In Figure 2.14 a high-level block diagram of the employed procedure is conferred. The approach avails from the current and previous camera position estimations to ascertain if movement occurred between frames. Considering there was none, the current acceleration, ${}^{B}\tilde{a}_t$, (with the gravity formerly extracted), is saved in a circular buffer. Once the buffer is full, the average of the sample is computed and the bias for each axis is updated.

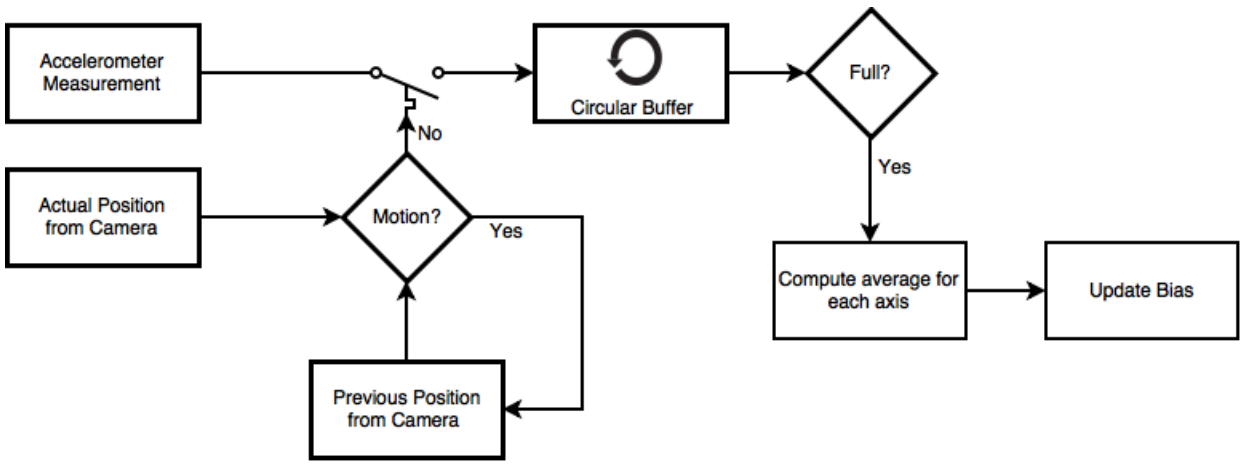


**Fig. 2.14:** Diagram of the accelerometer bias estimation.

In Figure 2.15, a high-level block diagram of the overall hybrid system is presented. Considering that both tracking approaches function at different sample rates, the inertial technique being more than three times faster than the visual one, two different proceedings must be contemplated in order to fully understand the system structure. The visual update, where no reference frame transformation is required, holds added emphasis since its estimation for position is more reliable and trustworthy. This update is responsible for correcting any drift that might have occurred during former inertial updates. Additionally, the inertial update must also be contemplated, this update allows for a faster update in position between visual update, beyond being employed in occlusion circumstances.

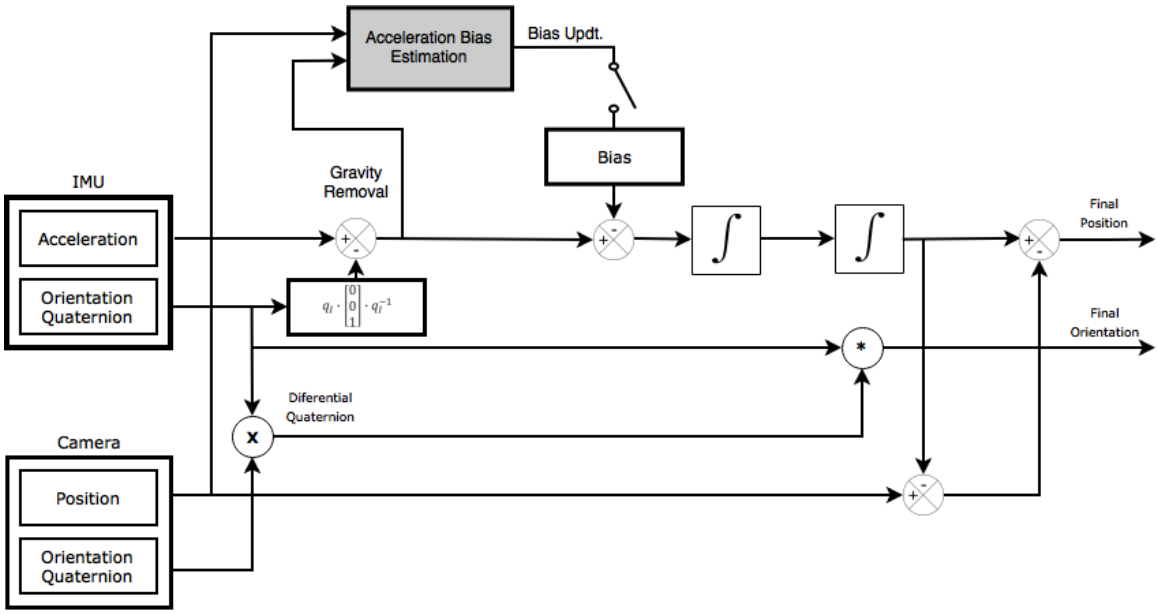


**Fig. 2.15:** Diagram of fusion of estimates for pose inference.

In Figure 2.16 an example of the performance of the system is displayed. The experiment consisted in handling the instrumented object while covering all the visual markers detectable by the camera, moving it and placing it back, as close as possible to the initial position, uncovering the markers. The aim of this experiment was to evaluate the system's performance of dealing with occlusion, a commonly faced problem in visual-based tracking. The results where appealing, although presenting minor inconsistencies during movement where only inertial data was present.

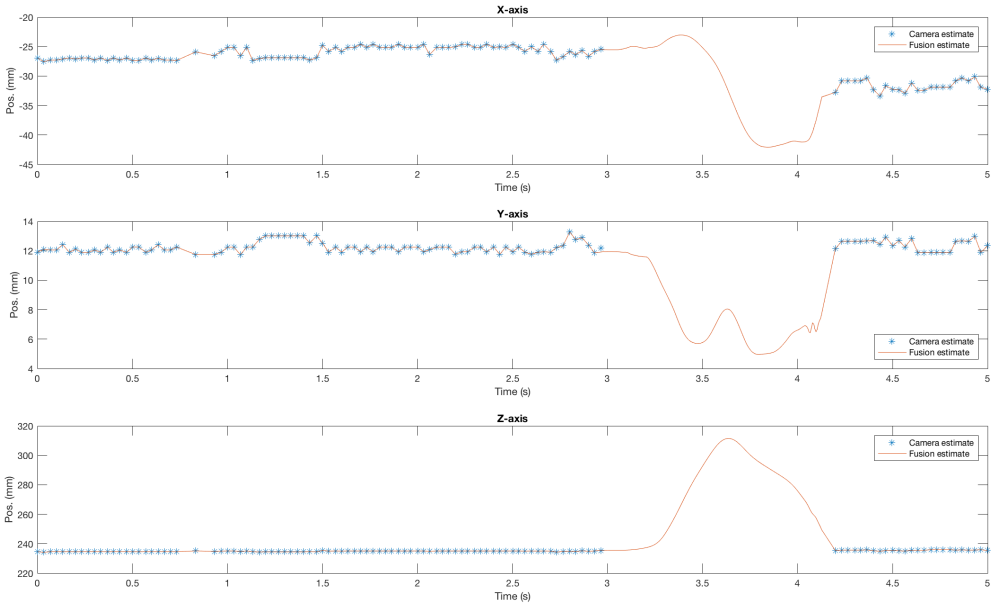


**Fig. 2.16:** Example of hybrid approach compared to the vision-based system.

# Chapter 3

# Connecting an Object to the IoT

The next wave in the era of computing will be outside the realm of the conventional desktop, we are witnessing the dawn of a Internet of Things (IoT) era. The term "Internet of Things" is, in fact, attributed to the Auto-ID Labs at MIT, which investigates in the field of networked RFID and emerging sensing technologies [69]. In 2005, International Telecommunication Union (ITU) released a report [70] formally proposing the concept, empowering the promise of a world of networked and interconnected devices.

Conventionally, IoT refers to the networked interconnection of everyday objects, which are frequently equipped with various sensors enabling data acquisition for diversified purposes, commonly denominated as *smart devices*. IoT will expand the ubiquity of the Internet by incorporating each and every object for interaction through embedded systems, leading to a highly distributed network of devices that communicate with each other, cooperating to attain a common objective. Although, being in a early stage IoT presents a wide range of applications, from modern agriculture to public security to name a few [71].

In the interest of enabling the proposed instrumented object and the forthcoming devices integration in this new vision, an elementary communication protocol was developed, enabling the communication for multiple *smart devices*, in the network. The designed protocol is established upon the User Datagram Protocol (UDP), introduced in 1980 by David Reed, which represents a simple transport layer protocol for client/server network applications, based on Internet Protocol (IP), and is the principal alternative to Transmission Control Protocol (TCP). This method of transmission, when compared to TCP, allows for less data overhead and delay. Since, the data packets are smaller to begin with and, additionally, it allows for individual packets to be dropped,

not using acknowledgements or retransmissions that may reduce the speed of the delivery of real-time data.

In regards to our protocol, for classification purposes, every device has an identifier that categorizes it by type of interaction or intent of use, as displayed on Table 3.1. This allows for multiple devices connected to the same wireless network to be easily distinguished, facilitating the communication of each application to the desired device or set of devices.

| Device Category | Instrumented Objects | Hand-Trackers | Segway$^{TM}$ Simulator | General |
|---|---|---|---|---|
| Letter associated | C | H | S | G |
| Hex representation | 0xF1 | 0xF2 | 0xF3 | 0xFF |

**Table 3.1:** Example of applications and identifier for the developed protocol

Considering the introductory set-up of a *smart device*, when firstly powered on it will try to reconnect to a previously saved Access Point, if unsuccessful (or no previous network was saved) it moves into Access Point mode, generating a wireless network for initial configuration. By utilizing any Wi-Fi enabled device with a browser, and connecting to the Access Point will lead to a configuration portal, where all the available wireless networks will be presented for selection, as well as, a field for inserting the corresponding password. When successfully connected to a network the *smart device* will await for a computer connection, otherwise, if unsuccessful, it will move back to Access Point mode expecting to be reconfigured.

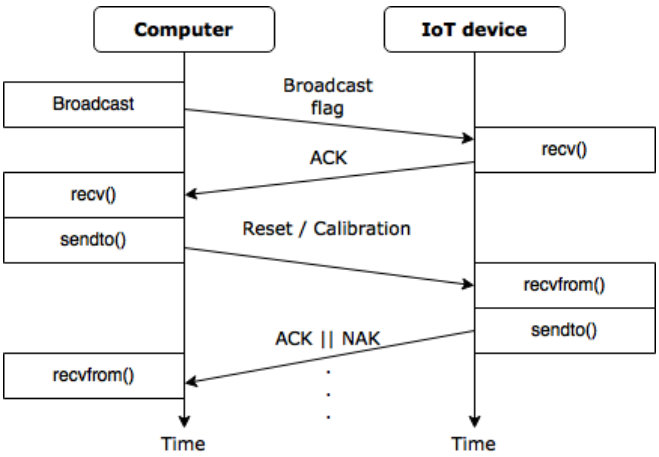


**Fig. 3.1:** Diagram of the initial communication between the computer and IoT device.

The ordinary employment of the formulated protocol, initiates with the computer sending a broadcast message to the network, as pictured in figure 3.1. This message can either be directed

at a specific class of devices or generic (regarding all devices), depending on the applications requirements. Each accessory responds to the broadcast, disregarding if its being used or not. Succeedingly, the computer sends a message that can represent each of two: a reset/calibration, if the device requires such before data acquisition, or merely to check the readiness of the appliance. The reply is conditioned by availability of the device, when already in use a rejection message is sent, otherwise it sends an acknowledge changing its state of accessibility, bounding to that connection.

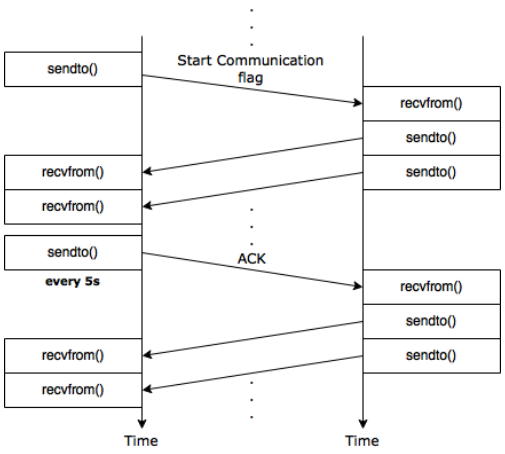


**Fig. 3.2:** Diagram of the data acquisition between the computer and IoT device.

Following the availability request, the application can start the data acquisition in a couple of manners, the standard mode or the debug mode, which provides additional information not required for the adequate functioning of the application. The system can halt the data acquisition at any given time through a specific message, and proceed the data acquisition posteriorly if desired.

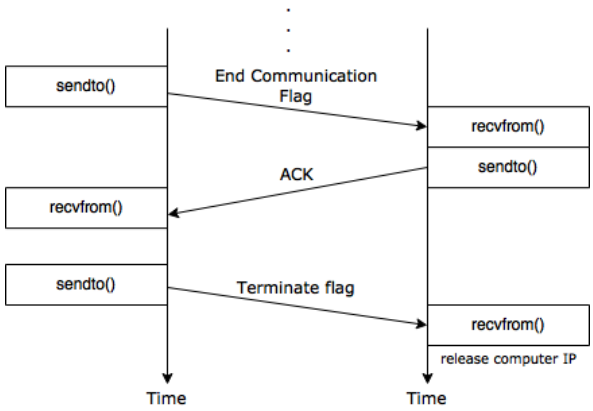


**Fig. 3.3:** Diagram representing the conclusion of the communication between the computer and IoT device.

The connection will only be concluded in two specific ways, if the computer sends a termination message unbounding the device, making it available for a new interaction, or if during fifteen seconds the system is unable to provide to the *smart device* with a message representing that the connection is still up. This probably expresses that the application crashed, since the system sends a acknowledgement of connection every five seconds, an example of this can be pictured in figure 3.2. The messages for starting, stopping and terminating the communication flow, will solely be considered if transmitted by the system to which the device is bounded to, otherwise they are disregarded. A list of the available commands with their description can be observed in Table 3.2.

| List of Commands | Hex representation | Description |
|---|---|---|
| E | 0x14 | Halts data transmission. |
| D | 0x16 | Start data transmission in debug mode, providing additional data not required for normal functioning. |
| C | 0x11 | Checks for device availability, binding the smart device to the application if possible (performing a calibration if needed). |
| S | 0x12 | Start data transmission in normal mode. |
| SM | 0x13 | Start data transmission in normal mode, with additional features (enabling additional sensors or other). |
| T | 0x10 | Finishes the connection, unbinding the smart device. |

**Table 3.2:** Available commands for the developed protocol

For evaluating the performance of the developed communication protocol, it was tested on the network of Instituto de Sistemas e Robótica(ISR). The solution presented a delay of $\simeq 50\mu s$ in direct message data transmission. While on the broadcast message at the start of the communication, a higher delay is presented, $\simeq 21ms$, due to uniqueness of message in consideration. Even though the transmission of the broadcast message is slow, that produces no obstacle for the communication, since it is only used initially, to establish the communication between the device and the application. Since this is an UDP based protocol, the number of dropped packets during communication were also evaluate, considering a $5min.$ window for the test on the same network, there where a total of 2 dropped packets. The results confirm that the developed protocol presents a satisfactory and reliable solution for the task at hand.

# Chapter 4

# Development of a Prototype for Motion Tracking and Force Sensing

In this chapter we present the prototype developed during this dissertation. Firstly, an overview of the proof-of-concept prototypes and the decisions made during their development are presented. Following, the end design decisions for of the instrumented cube is conferred. Finally, a couple of applications for the advanced system are described, both regarding to the instrumented cube as the main interaction method, one being more focused on assessing its capacities, the other employs it to achieve a therapeutical purpose.

## 4.1 Hardware Implementation

### 4.1.1 Proof-of-Concept Prototypes

Prior to the development of the fully functional concept for the instrumented cube, in order to demonstrate the viability of the proposed solution a couple of proof-of-concept prototypes where developed. The early designs allowed a better understanding of the solution's limitations, permitting future improvements for the coming iterations. The initial prototype consisted of, a paper cube with six fiducial markers, one on each of the cube's faces, covering the visual tracking component. Additionally, filling in for the inertial tracking element an ESP-01 module based on ESP8266 microcontroller unit, made by AI-Thinker, capable of connecting to a Wi-Fi network, complemented with an inertial sensor. The inertial sensor prefered is the MPU-9250, developed

by Invensense, a system in package that combines two chips: the MPU-6050 comprised by a 3-axis accelerometer, a 3-axis gyroscope and a Digital Motion Processor$^{TM}$ with some built in motion algorithms and the AK8963 a 3-axis magnetometer. The aforementioned solution can be pictured in figure 4.1. In the conducted experiments using the previously alluded solution, the results provided fulfilled our expectation, stimulating us to further improve on the concept.

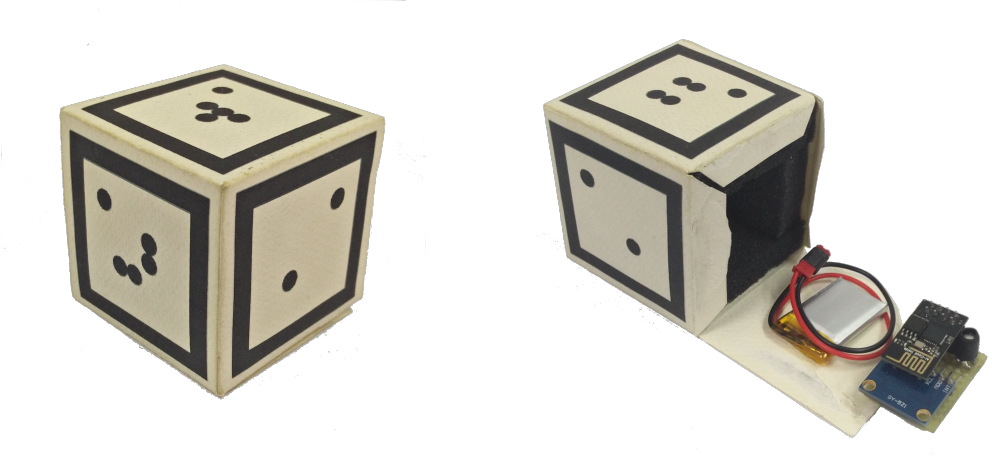

**Fig. 4.1:** Initial proof-of-concept of instrumented cube.

In the second iteration of the prototype some adjustments where produced. The cube construction was composed of paperboard for added strength, since the paper solution presented ephemeral partially due to, the weight of the inertial solution placed inside, as well as, the grasping forces applied by the users will handling the object. Furthermore, an assessment to the use of conductive foam as a force sensor was performed. The conductive foam sensor is assembled as shown in figure 4.3, using two conductive stripes surrounding the conductive foam, that are isolated by a couple of plastic strips or any other electrical insulator.

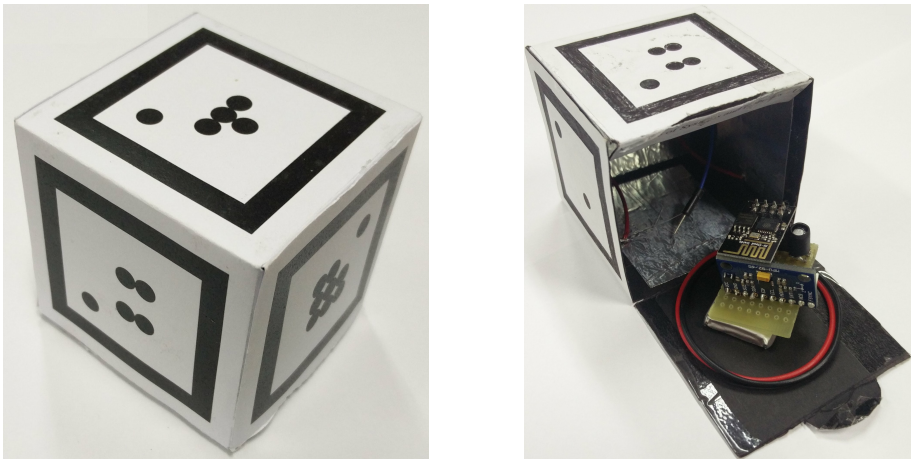

**Fig. 4.2:** Second iteration in the development of instrumented cube.

The sensor takes advantage of the varying resistance of the conductive foam with pressure, by using a voltage divider with a previously selected resistance the force applied can be inferred upon. Although providing an acceptable solution, the fact that the compressing and decompression measurements are dissimilar for an equivalent pressure, as pictured in figure 4.3, disowned it as a forthcoming approach.

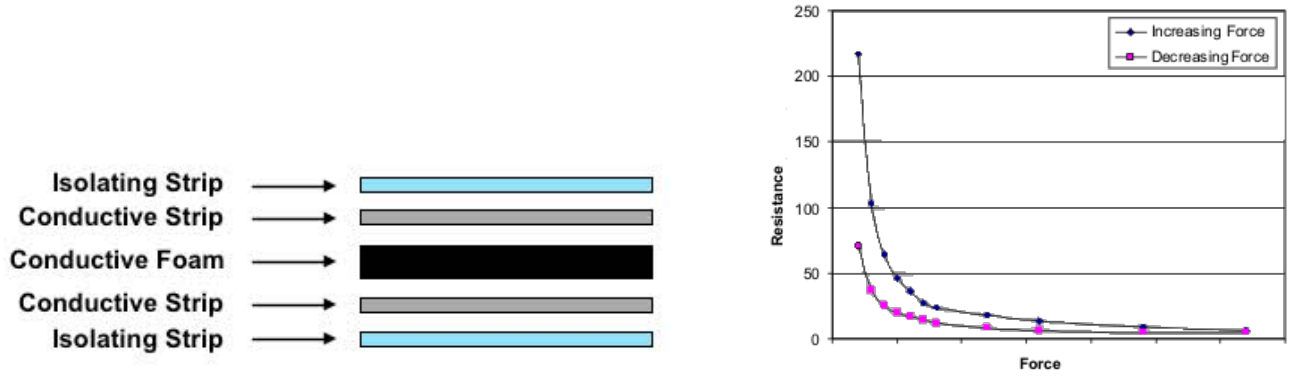


**Fig. 4.3:** Left : Mock-up of the assembly of foam sensor; Right: Conductive foam performance graph.

### 4.1.2 Instrumented Cube Prototype

In pursuance of delivering a reliable and durable solution, a 3D printable cube-shaped container was sketched, the design is presented in figure 4.4, all the perforations and supports for the sensors and the microcontroller where included in the design. The visual markers will be placed posteriorly to the placement of the force sensors, since those are placed on the exterior of the container.

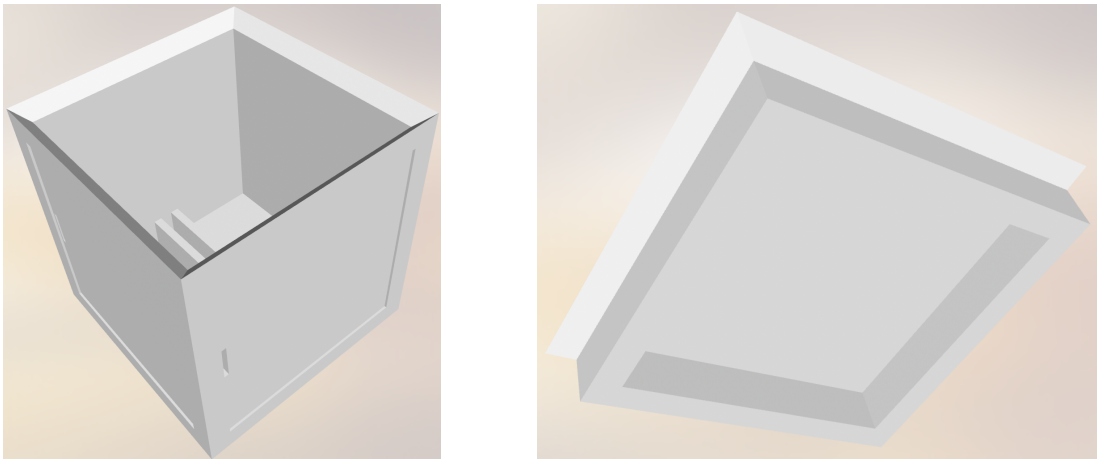

**Fig. 4.4:** Printable 3D sketch of the cube.

The ESP-01 module was exchanged for the Wemos D1 module in behalf of the presence of some extra features, the microcontroller unit remains unaltered being based on the ESP8266. The additional features rage from an on-board USB-to-UART bridge, a built-in voltage regulator, an analogue port and several supplementary modules. From the available modules, a battery shield was employed complementing the system with the ability to recharge the LiPo battery through a micro-B USB port.

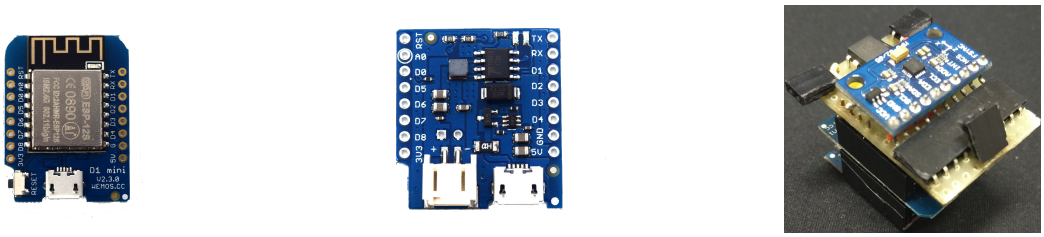

**Fig. 4.5:** Left: Wemos D1 microcontroller module; Center: wemos battery module; Right: full system with inertial sensor assembled.

Regarding the force sensing capabilities of the developed solution, three force sensing resistors FSR-406 from Interlink Electronics where adopted, being placed one on each of the axis for the greatest surface coverage. The force sensing resistors operate in a similar way to a potentiometer, varying the resistance with the force applied. Typically, a voltage divider is employed for measurement, as presented in figure 4.6, we designated a $RM = 10k\Omega$ due to the hardware limitation relating to the maximum admissible voltage of the analogue port. Since just an individual analogue input is accessible, in order to acquire force information from all three axis, the integration of a 4-channel multiplexer, the HCF4052B from STMicroelectronics was employed. The readings where performed sequentially and the selection of the sensor to read was carried out by two digital pins connected to the selection ports of the multiplexer.

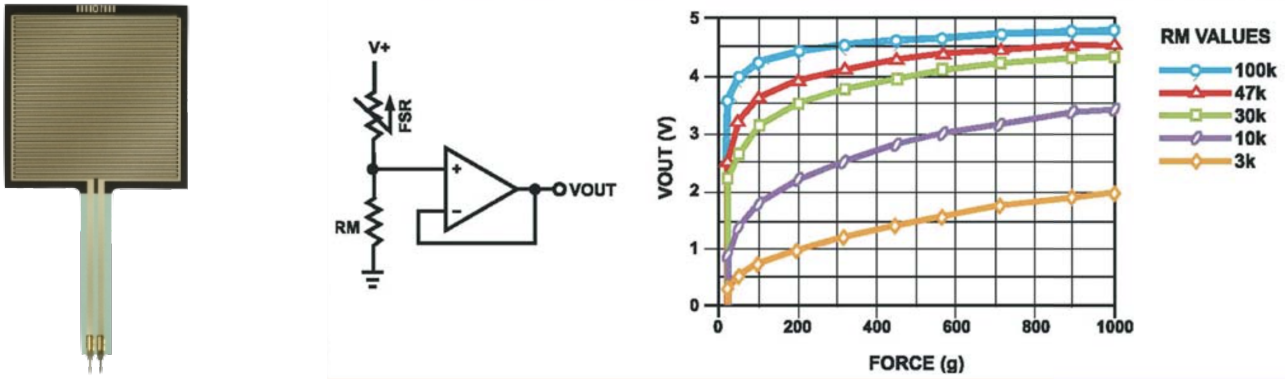


**Fig. 4.6:** Left: Interlink Electronics FSR 406; Center: Schematic of sensor usage in voltage divider; Right: Graph relating the voltage output to the force applied for multiple resistances.

## 4.2 Developed Applications

In this section two developed applications in the OpenAR platform, that embrace the instrumented object, will be presented. The first one, aims at fully testing each of the capabilities of the system, as well as, recording information to be computed later on. While the second, is a third-person-view runner-based game applied to therapeutical purposes, aspiring to improve the user's visual-motor coordination.

### 4.2.1 Motion Tracking

The Motion Tracking application was created in the OpenAR, an internal development platform design primarily for Augmented and Virtual Reality environments, with built-in open source libraries such as, OpenCV, OpenGL, to name a few. The main purpose of the following implementation was to assess on each of the capabilities of the developed instrumented cube. As displayed in figure 4.7 which represents the start-screen of the motion tracking application, there are several modes produced to individually evaluate each of the functionalities of tracking system, as well as, to generate datasets for future analysis.

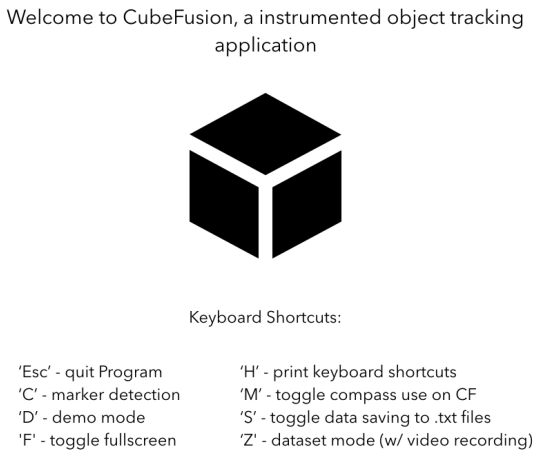


**Fig. 4.7:** Start-screen of the application.

In the following figure 4.8, the most meaninful tracking approaches are presented. On the left, an example of the use of the marker-based pose estimation approach, presented in the 2.1, can be pictured with two markers being detected and the inferred position shown on the top right corner of the screen. On the right, the visuo-inertial fusion estimation used for tracking the

instrumented object is introduced, with a virtual cube overlaying the instrumented cube. The use of this application concedes a greater perception of the system performance compared to a uniquely dataset-based post processed manner, bearing an easier and quicker way of perceiving possible inaccuracies to be rectified. This application presents an adequate base for further development, having the necessary base tools for the employment of the instrumented object in other applications.

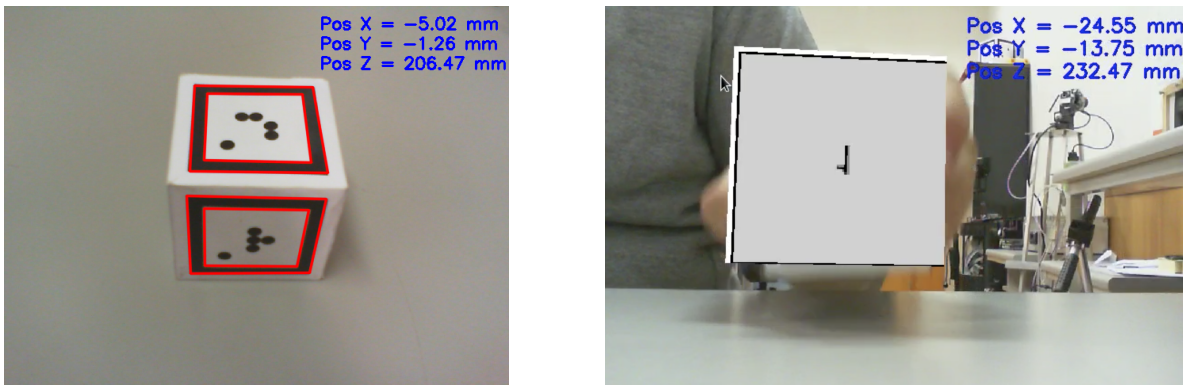


**Fig. 4.8:** Left: Example of the marker-based visual pose estimation approach; Right: Example of the visuo-inertial fusion pose tracking estimation.

### 4.2.2 PuzzleTime

PuzzleTime [72] is a serious game, that serves as a tool for improving visual-motor coordination, especially for people young age. It consists in a third-person-view runner-based game, where the character is placed in a three-dimensional virtual environment, and requires the help of the player to overcome some obstacles in order to, successfully complete the level. The game comprehends a few levels, pictured in the figure 4.9 bellow, that slowly increase the level of difficulty, conceding the user the opportunity for a gradual improvement of the visual-motor coordination.

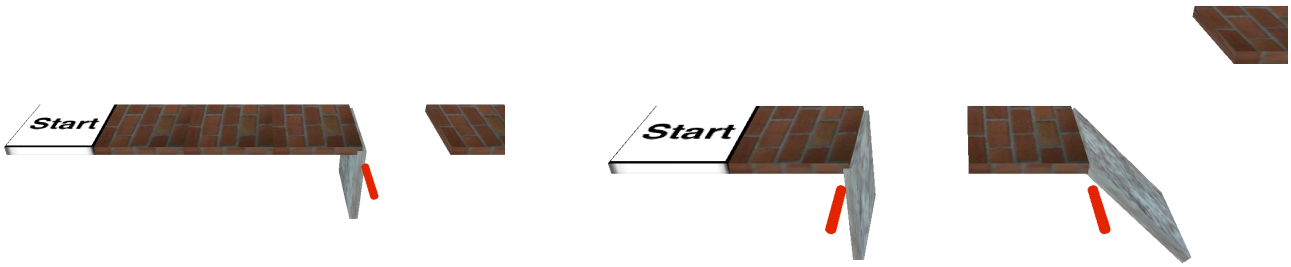


**Fig. 4.9:** Example of the different levels developed for the game.

The typical obstacles present in the game consist in holes or harmful terrain, preventing the character to reach the end of the level. The player may help the character to overcome them by placing a bridge, generating a new unrestrained passage. In order to place the linking tile, the user must held the instrumented cube with a specific pose, evidently a margin of error is granted. The cube does only connected to one tile at a time, avoiding possible distraction differing from the present obstacle to overcome.

The use of the instrumented cube in this application, awards the player with thorough control during handling, in addition to the main ambition of improving the visual-motor coordination, providing an overall pleasanter interaction for the user.

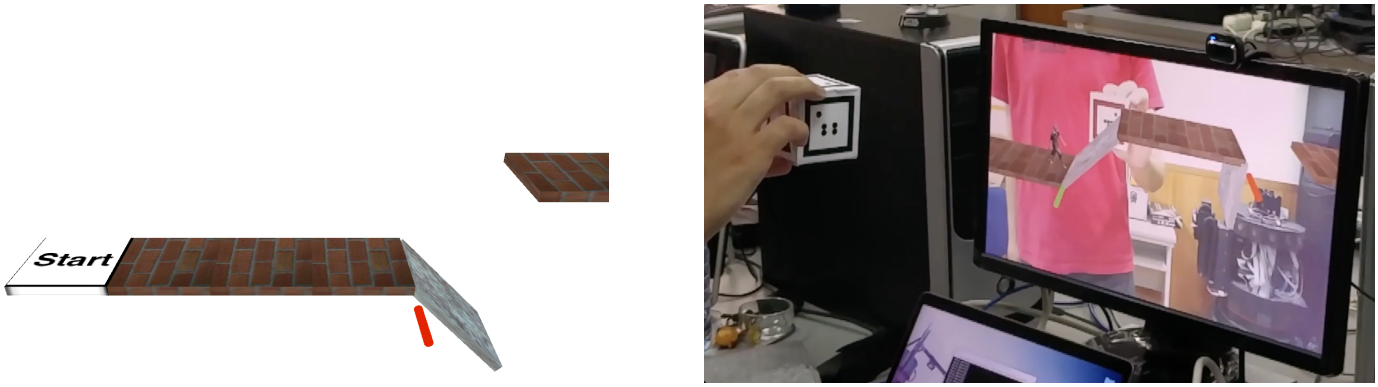


**Fig. 4.10:** Left: Current level map at start; Right: User playing the developed game.

# Chapter 5

# Conclusions and Future Work

Ongoing virtual and mixed reality system allege to provide thorough immersion, even supposing there is little to none to improve upon in the visual aspect, the other senses have been widely disregarded, in the majority of the systems. Considering the extensively employed proprioception technique for object interaction, it is absent of physical constrains and haptic feedback, creating burdensome interaction for an effortless task in the real world. The work developed in this dissertation attempts to revert this tendency, by providing a straightforward interaction for handling and manipulating a virtual object, by the means of instrumented real object. In the interest of delivering a reliable system, a position and orientation tracking approach based on a hybrid system, relying on visual and inertial tracking technologies, was employed. Visual pose estimation was achieved by the means of fiducial markers located on the object's surface. While, the inertial tracking was performed by an strapdown INS, computing double integration of the acceleration in the correct orientation frame, provided by the complementary filter developed. A final economical prototype, with force sensing capabilities and wirelessly connected, through a conceived elementary communication protocol, was accomplished, depicting a complementary tool for virtual environments, as well as, other applications.

Regarding future work, there exist a few features that can be improved upon to extend the capabilities of the proposed system. Regarding the software, the hybrid solution fusing algorithm can be enriched, a extended Kalman-based or particle-based approaches must be considered. Additionally, the accelerometer calibration performed over time could benefit from the integration of the previously mention estimation techniques or other better suited. Lastly, a low power mode with tap-to-wake feature could be contemplated, further extending the time between recharges.

Taking into consideration the hardware, a LED status indicator could present as a beneficial approach to monitor the battery level. Furthermore, the amount of force sensor could double allowing more thorough estimation of the grasping force, as well as, presenting as a possibility for detecting when the object is placed at rest. Lastly, a multiple camera setup could be considered providing a wider working space, while reducing frequency of problems such as, occlusion and illumination. Concluding, there are several applications scenarios where this approach can be used such as, the therapeutical use for grasping force in patients that suffered from a stroke, use as controller in varying stressful scenarios using the force sensing capabilities to relating the stress levels with grasping force, visuomotor coordination improvements other than the already addressed among others.

# References


[1] Ivan E. Sutherland. The ultimate display. *International Federation of Information Processing IFIPS Congress*, 2:506–508, 1965.

[2] Fumio Milgram, Paul; Kishino. A Taxonomy Of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D(12):1321–1329, 1994.

[3] Mark R Mine. Virtual Environment Interaction Techniques. *University of North Carolina, Chapel Hill, NC*, 1995.

[4] Doug A. Bowman and Larry F. Hodges. Formalizing the Design , Evaluation , and Application of Interaction Techniques for Immersive Virtual Environments. *Visual Languages and Computing*, 10(1):37–53, 1999.

[5] Mark R Mine, Frederick P Brooks, and Carlo Sequin. Moving Objects in Space : Exploiting Proprioception In Virtual-Environment Interaction. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*, pages 19–26, 1997.

[6] Luis Almeida, Bruno Patrao, Paulo Menezes, and Jorge Dias. Be the robot: Human embodiment in tele-operation driving tasks. *IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions*, pages 477–482, 2014.

[7] J. G. Sanchez, B. Patrão, L. Almeida, J. Perez, P. Menezes, J. Dias, and P. Sanz. Design and Evaluation of a Natural Interface for Remote Operation of Underwater Robots. *IEEE Computer Graphics and Applications*, 2015.

[8] Ivan E Sutherland. A head-mounted three dimensional display. *Proceedings of AFIPS*, pages 757–764, 1968.

[9] C. S. Draper, W. Wrigley, G. Hoag, R.H. Battin, J. E. Miller, D. A. Koso, A. L. Hopkins, and W. E. Vander Velde. Apollo Guidance and Navigation. 1965.

[10] Seong-hoon Peter Won, Farid Golnaraghi, Wael William Melek, and Senior Member. A Fastening Tool Tracking System Using an IMU and a Position Sensor With Kalman Filters and a Fuzzy Expert System. *IEEE Transactions on Industrial Electronics*, 56(5):1782–1792, 2009.

[11] D. Roetenberg, P. J. Slycke, and P. H. Veltink. Ambulatory position and orientation tracking fusing magnetic and inertial sensing. *IEEE Transactions on Biomedical Engineering*, 54(5):883–890, 2007.

[12] Fakespace Labs. Boom Displays Available at: http://www.fakespacelabs.com/tools.html.

[13] Norther Digital Inc. Optotrak Certus Available at: https://www.ndigital.com/products/.

[14] VRvana Totem Available at: https://www.vrvana.com/.

[15] Paramvir Bahl and V. N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, 2000.

[16] Greg Welch and Gary Bishop. An Introduction to the Kalman Filter. *SIGGRAPH 2001*, 2001.

[17] I. Strid and K. Walentin. Block Kalman filtering for large-scale DSGE models. *Computational Economics. Springer*, pages 277–304, 2008.

[18] Eric Foxlin, Michael Harrington, and George Pfeifer. Constellation $^{TM}$ : A Wide-Range Wireless Motion- Tracking System for Augmented Reality and Virtual Set Applications. *SIGGRAPH '98 Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998.

[19] John L Crassidis, F Landis Markley, Control Systems, and Engineering Branch. Unscented Filtering for Spacecraft Attitude Estimation. *Journal of guidance, control, and dynamics*, (August):536–542, 2003.

[20] Pierre Del Moral. Nonlinear Filtering: Interacting Particle Resolution. *Markov Processes and Related Fields*, 2(4):555–580, 1996.

[21] Alexander Krull, Frank Michel, Eric Brachmann, Stefan Gumhold, Stephan Ihrke, and Carsten Rother. 6-DOF Model Based Tracking via Object Coordinate Regression. *Computer Vision - ACCV 2014*, 2014.

[22] I Salimpour and H Soltanian-Zadeh. Particle filtering of point processes observation with application on the modeling of visual cortex neural spiking activity. *Neural Engineering, 2009. NER '09. 4th International IEEE/EMBS Conference*, pages 718–721, 2009.

[23] Robert Mahony, Dep Eng, Dep Eng, and Tarek Hamel. A coupled estimation and control analysis for attitude stabilisation of mini aerial vehicles . *Australiasian Conference on Robotics and Automation*, 2006.

[24] Sebastian O H Madgwick, Andrew J L Harrison, and Ravi Vaidyanathan. Estimation of IMU and MARG orientation using a gradient descent algorithm. *IEEE International Conference on Rehabilitation Robotics*, 2011.

[25] Nintendo Wii Remote Available at: https://store.nintendo.com/.

[26] Sony Interactive Entertainment PlayStation Move Available at: https://www.playstation.com/en-us/explore/accessories/vr-accessories/playstation-move/.

[27] Razer Hydra Available at: https://www.razerzone.com/.

[28] Google Daydream Controller Available at: http://adcardboard.com/daydream.html.

[29] Samsung Gear VR controller Available at: https://www.oculus.com/gear-vr/.

[30] Oculus Touch Available at: https://www.oculus.com/rift/.

[31] HTC Vive Controller Available at: https://www.vive.com/eu/product/.

[32] HTC Vive Tracker Available at: https://www.vive.com/eu/accessory/.

[33] Ashley L Guinan, Markus N Montandon, Andrew J Doxon, and William R Provancher. Discrimination Thresholds for Communicating Rotational Inertia and Torque using Differential Skin Stretch Feedback in Virtual Environments. *IEEE Haptics Symposium (HAPTICS)*, (c):277–282, 2014.

[34] B. Patrão, P. Menezes, and P. Castilho. Development of Mixed Reality Systems to Support Therapies. *Doctoral Conference on Computing, Electrical and Industrial Systems (DoCEIS)*, 2016.

[35] Robert W Lindeman, John L Sibert, James K Hahn, and The George. Hand-Held Windows : Towards Effective 2D Interaction in Immersive Virtual Environments. *IEEE Virtual Reality*, pages 205–212, 1999.

[36] Radu Horaud. Methods for Matching 3-D Objects Perspective Views. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9, No.*, (3):401–412, 1987.

[37] Jun Rekimoto. Matrix : A Realtime Object Identi cation and Registration Method for Augmented Reality Jun Rekimoto Sony Computer Science Laboratory Inc . Phone : + 81-3-5448-4380 Mail : rekimoto@csl.sony.co.jp. *Proceedings of Asia Pacific Computer Human Interaction 1998 (APCHI'98)*, 1998.

[38] B. Patrão. *Biblioteca para desenvolvimento de aplicações de Realidade Aumentada com base em marcadores binários*. PhD thesis, DEEC-FCTUC, Universidade de Coimbra, 2011.

[39] Martin A Fischler and Robert C Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[40] Radu Horaud, Bernard Conio, and Oliver Leboulleux. An Analytic Solution for the Perspective 4-Point Problem. *Computer Vision, Graphics, Image Processing*, 47(1):33–44, 1989.

[41] Long Quan and Zhongdan Lan. Linear N-Point Camera Pose Determination æ. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(8):774–780, 1999.

[42] Adnan Ansar and Kostas Daniilidis. Linear Pose Estimation from Points or Lines. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(5):578–589, 2003.

[43] Yihong Wu and Zhanyi Hu. PnP problem revisited. *Journal of Mathematical Imaging and Vision*, 24(1):131–141, 2006.

[44] E. H. Thompson. Space Resection: Failure Cases. *Photogrammetric Record*, pages 201–204, 1966.

[45] B. K. P. Horn. Closed Form Solution of Absolute Orientation Using Unit Quaternions. *J. Optical Soc. Am.*, 5(7):127–135, 1987.

[46] L. Lucchese. A hybrid frequency-space domain algorithm for estimating projective transformations of color images. *n Proc. International Conference on Image Processing*, pages 913–916, 2001.

[47] Ulrich Neumann, Suya You, Youngkwan Cho, Jongweon Lee, and Jun Park. Augmented reality tracking in natural environments. *International Symposium on Mixed Realities*, pages 1–24, 1999.

[48] H Kato, M Billinghurst, I Poupyrev, K Imamoto, K Tachibana, and M I C Laboratories. Virtual Object Manipulation on a Table-Top AR Environment. *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000).*, pages 111–119, 2000.

[49] Fabio Zünd, Mattia Ryffel, Stéphane Magnenat, Alessia Marra, Maurizio Nitti, Mubbasir Kapadia, Gioacchino Noris, Kenny Mitchell, Markus Gross, and Robert W Sumner. Augmented Creativity: Bridging the Real and Virtual Worlds to Enhance Creative Play. *SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications*, pages 21:1—-21:7, 2015.

[50] J. Seeger, M. Lim, and S. Nasiri. Development of high-performance, high- volume consumer MEMS gyroscopes. *Proc. Tech. Dig. Solid-State Sensors Actuators Microsystems Workshop*, pages 61–64, 2010.

[51] W. Storms, J. Shockley, and J. Raquet. Magnetic Field Navigation in an Indoor Environment. *Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS)*, pages 1–10, 2010.

[52] D. Gebre-Egziabher, G.H. Elkaim, J. D. Powell, and B. W. Parkinson. Calibration of strapdown magnetometers in magnetic field domain. *Journal of Aerospace Engineering*, 19(2):87–102, 2006.

[53] J. F. Vasconselos, G. Elkaim, C. Silvestre, P. Oliveira, and B. Cardeira. A geometric approach to strapdown magnetometer calibration in sensor frame. *Navigation, Guidance and Control of Underwater Vehicles*, 2008.

[54] T. Ozyagclicar. Calibrating an ecompass in the presence of hard and soft-iron in- terference. *Tech. Rep. Freescale Semiconductor Ltd*, 2012.

[55] V; Merayo, J, M, G; Brauer, P; Primdah, F; Petersen, J, R; Nielsen, O. Scalar calibration of vector magnetometers. *Measurement Science and Technology*, 11(2):120–132, 2000.

[56] B. T. Costic, D. M. Dawson, M.S. Queiroz, and V. Kapila. A Quaternion-Based Adaptive Attitude Tracking Controller Without Velocity Measurements. *Proceedings of the 39th IEEE Conference on Decision and Control, 2000.*, 2000.

[57] Tarek Hamel and Robert Mahony. Attitude estimation on SO(3) based on direct inertial measurements. *Proceedings IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, (May):2170–2175, 2006.

[58] Peter Kazanzides Hongliang Ren. Investigation of Attitude tracking using an integrated inertial and optical navigation system for hand-held surgical instruments. *International Conference on Control, Automation and Systems*, 17(2):290–293, 2014.

[59] Mark Euston, Paul Coote, Robert Mahony, Jonghyuk Kim, and Tarek Hamel. A Complementary Filter for Attitude Estimation of a Fixed-Wing UAV. *EEE/RSJ International Conference on Intelligent Robots and Systems*, pages 22–26, 2008.

[60] Robert Mahony, Senior Member, Tarek Hamel, and Jean-michel Pflimlin. Nonlinear Complementary Filters on the Special Orthogonal Group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, 2008.

[61] Tiago Catarino. Development of a Hand-Tracker: Wireless Solution based on Inertial Sensors. 2016.

[62] K. Seifert and O. Camacho. Implementing Positioning Algorithms Using Accelerometers. *Application Note Freescale Semiconductor*, pages 1–13, 2007.

[63] David Vincent. Accurate Position Tracking Using Inertial Measurement Units. *PNI Sensor Corporation White Paper*, (February), 2013.

[64] Pedro Neto, J Norberto Pires, and Anónio Paulo Moreira. 3-D Position Estimation from Inertial Sensing : Minimizing the Error from the Process of Double Integration of Accelerations. *ECON 2013-39th Annual Conference of the ndustrial Electronics Society*, pages 4026–4031, 2013.

[65] O. J. Woodman. An Introduction to Inertial Navigation. *University of Cambridge*, (696):1–37, 2007.

[66] Miguel Ribo, Markus Brandner, and Axel Pinz. A Flexible Software Architecture for Hybrid Tracking. *Journal of Robotic Systems*, 21(2):53–62, 2004.

[67] Yaqin Tao, Huosheng Hu, and Huiyu Zhou. Integration of Vision and Inertial Sensors for 3D Arm Motion Tracking in Rehabilitation. *The International Journal of Robotics Research*, 26(6), 2007.

[68] Neda Parnian and Farid Golnaraghi. Integration of a Multi-Camera Vision System and Strapdown Inertial Navigation System (SDINS) with a Modified Kalman Filter. *Special Issue Instrumentation, Signal Treatment and Uncertainty Estimation in Sensors*, pages 5378–5394, 2010.

[69] I. Bose and R. Pal. Auto-ID: managing anything, anywhere, anytime in the supply chain. *Communications of the ACM*, 48(8):100–106, 2005.

[70] Itu. The internet of things. *ITU Internet Reports - Executive Summary*, pages 1–28, 2005.

[71] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: a survey. *Computer Networks*, 54(15):2787–2805, 2010.

[72] Nuno Gouveia, Bruno Patrão, and Paulo Menezes. Puzzle Time - VR Runner. *Exp.at'17 International Conference*, pages 121–122, 2017.