2

Representation of 3D Space and Sensor Modelling within a Probabilistic Framework

Between stimulus and response there is a space. In that space is our power to choose our response. In our response lies our growth and our freedom.
Man's Search for Meaning, *Viktor Frankl (1946)*

For the wise man looks into space and he knows there are no limited dimensions.
Lao Tzu (500-600 BC)

2.1 Introduction

For living organisms, perception can be defined as a set of *cognitive processes*, in the sense that it consists in the processing of sensorial data in order to generate essential information with the purpose of building a coherent and useful representation of the surrounding world. Perception has been paramount for living beings, its importance having propagated from supporting the original primal objective of survival up to the more recent evolutionary purpose of promoting social interaction.

Now, *cognition* (from the latin *cognoscere*, which means "to know", "to conceptualise" or to "recognise"), although always implying a relatively complex process, can be unconscious. This is, in fact, the case in most of the cognitive processes underlying perception in living beings. However, as can be easily understood from this line of thought, being unconscious does not rule out that these processes involve "reasoning" (even if done automatically).

Therefore, when modelling perception, we are in fact establishing how a cognitive system reasons when relating abstract concepts – or *symbols* – with raw sensor data. This problem has been studied for decades now in Artificial Intelligence and Robotics, under many different names, one of the most well known being the "symbol grounding problem" [35; 16]. Thankfully, as was shown in Chapter 1, plausible reasoning and Bayesian inference address incompleteness and uncertainty, that so often make perception as a symbol grounding problem difficult, inherently.

Unfortunately, the world that is to be analysed by a perceptual system is dynamic: not only the intrinsic properties of objects in the environment are susceptible to changes through time, but also their spatial disposition. This means that the world cannot be perceived with a single glance. Consequently, for the raw sensory data arising from observing objects to be robustly

J.F. Ferreira and J. Dias: *Probabilistic Approaches for Robotic Perception*, Springer Tracts in Advanced Robotics 91,
DOI: 10.1007/978-3-319-02006-8_2,

related to symbols, the spatial properties of a perceptual scene must, at least implicitly, also be assigned a representation.

With this in mind, we have structured this chapter as follows: firstly, we will discuss how space might be referenced and represented; next, we will show how to build probabilistic sensor models taking spatial representation explicitly into account; finally, we will briefly address the difference between detecting and recognising objects, and give an inkling of how to address this problem within a probabilistic framework.

2.2 The Reference of Representation – Egocentric *vs.* Allocentric

Sensor data will most certainly relate to physical entities, each of which placed in specific locations in the observer's surroundings. The most important and immediate associations that humans and other animals make when trying to make sense of the incoming sensory data are precisely spatial associations, since these generally have imminent significance. Firstly, in simple, primal tasks such as navigation, it is not always as important to know *what* is exact nature of the objects one is observing (i.e. if the road one is following is made of tarmacadam or cobblestones, or if an obstacle one has to avoid is a rock or a chair) as it is to know *where* these objects are. On the other hand, categorising the nature an object is more than often related to where the object is located (for example, when we are interested in getting "the third cup from the left").

A *reference frame* allows the representation of locations of entities in space. It is defined by a geometrical *origin* and a set of (ordinarily, but not necessarily, orthogonal) directions, called *axes*, one per each spatial dimension to be represented.

The "reference of spatial representation", as a point of view from which the observer perceives the world, has been extensively studied in several contexts [25; 8; 21; 34]. There are two different broad ways of qualifying reference frames depending on this point of view:

- **Egocentric:** This is related to point-of-view *centred on the observer* (*ego* being the latin word meaning "self"; i.e. dependent of observer pose). Usually directly or indirectly sensor-referred.
- **Allocentric:** This refers to "other than self" (from the greek word *allos*, meaning "other"). In other words, it refers to any abstract or concrete entity other than observer.

The broad classification of allocentric reference frames is then used with different, more specific meanings depending on the author. For example, it is sometimes used to signify "exocentric" or "geocentric", i.e. the so-called "world reference frame", often abbreviated as *wrf*, which is centred in an abstract origin and usually assumed static. In other cases, it is used to designate