Chapter 2 - Fundamentals

Chapter 2

Fundamentals

The computational models and engineering solutions proposed in this thesis for performing the dexterous manipulation and haptic exploration of surfaces follow the principles involved in human perception, cognition, and action.

As presented in Figure 2.1, the perceptual process starts a sequence of sub-mechanisms that work together to estimate a representation of the environment. The perceptual representation is then used to infer a reaction strategy to those stimuli coming from the environment.

Several models have been proposed to explain and describe the mechanisms involved in human perception and how they are integrated into global human behaviour. Humans perceive in order to act on the environment and, the actions performed with environment elements affect the perception of the environment: the so-called action-perception loop (Figure 2.1).

Although most of the time, the sensory signals are ambiguous and corrupted with noise, humans have a remarkable capability to create successful perceptual representations which they use to guide their actions [Ernst and Bulthoff, 2004]. To explain this capability, Hermann von Helmholtz proposed an approach to model the perception mechanisms (Figure 2.1), introducing a principle designed by unconscious inference [Westheimer, 2008]. The principle states that humans perceive a specific state of the environment, choosing the state which is most likely to have caused the pattern of stimulus that the human subject has received through the sensory apparatus. Additionally, although sensory data is diverse, it is not sufficient to uniquely determine what is perceived. Prior knowledge must be used, which introduces constraints to the process of inference from ambiguous sensory signals.

The next sections present the fundamentals of the formalism of probability theory used to model the state of robotic systems, human agents, and the environment (section 2.1). This chapter also presents the formalism of probabilistic grids (section 2.2), used in this thesis to represent the workspace surrounding the robotic and human agents. The information exchanged between the different modules of the methods proposed in this

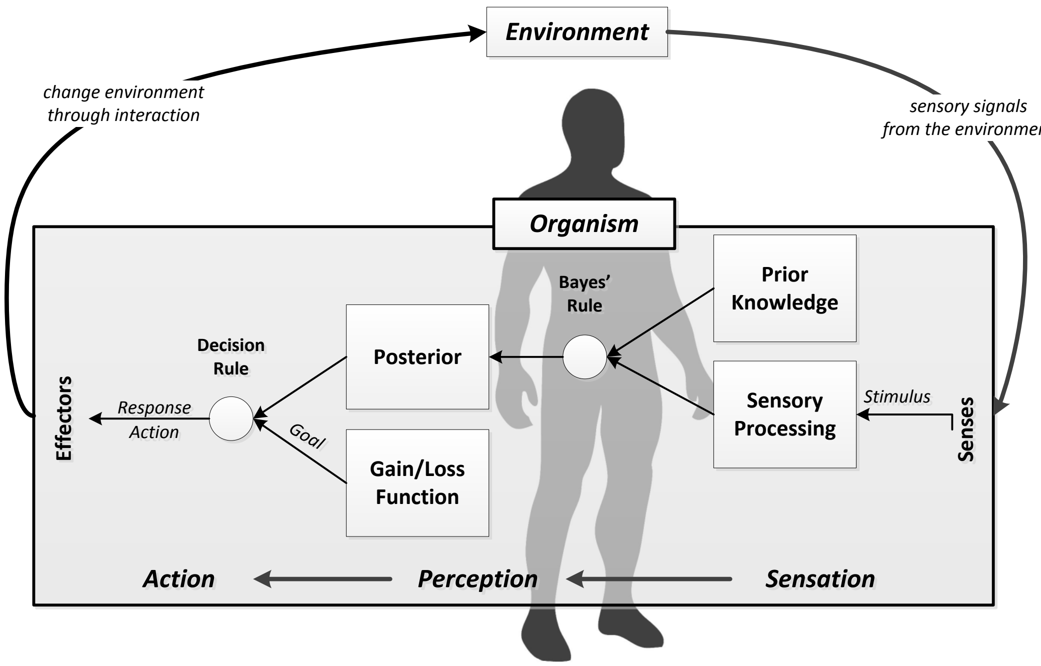


Figure 2.1: Representation of the fundamental mechanisms underlying the action-perception loop. Adapted from [Ernst and Bulthoff, 2004].

work is described using the formalism of information theory (section 2.3).

2.1 Probabilistic modelling

In robotics, different formalisms (such as first-order predicate logic and probability theory) have been followed to represent knowledge and describe reasoning applications. An extensive review of these formalisms can be found in [Hertzberg and Chatila, 2008].

This work uses the probability theory to represent the knowledge of the state of the robotic system (and other agents) and its surroundings, following an approach analogous to the work [Knill and Richards, 1996] to model human perception mechanisms and reasoning. This formalism has been used extensively in robotics. The increasing interest in this formalism is related to its ability to deal with the incompleteness of the description of the system (inaccurate modelling, relevant effect of hidden variables) and uncertainty of the available data (multimodal noisy data) [Thrun et al., 2005], [Ferreira and Dias, 2014c]. A new generation of computer architectures [Faix et al., 2015] and programming languages [Lebeltel et al., 2004] is also being developed to optimize and generalize the implementation of the methods described using this formalism.

The probabilistic methods proposed in this thesis follow the principles of the Bayesian probability theory.

2.1.1 Bayes rule

In this thesis, several descriptors such as robot state, multi-modal sensor measurements, and surrounding environment state, are represented by continuous or discrete random variables. Each variable is defined for a specific domain (possible values). A random variable or logical operation of random variables is characterized by a probability density function (continuous variables) or probability mass function (discrete variables), which assigns a probability ($[0, 1]$) to each value of the domain of the random variable. The work [Chung and AitSahlia, 2012] presents an extensive introduction to the basic concepts of probability theory.

Let C denote a random variable and c denote a specific value of the domain of C. In this abstract formulation of a problem, C represents a potential cause of an event of interest E (with e being a specific value of this variable).

During the probabilist modelling of a problem, the random variables establish statistical independence relations between them. If E and C are considered independent, then E does not influence C. This type of influence is modelled by a conditional probability, as shown in equation 2.1 .

$$P(C|E) = P(C) \tag{2.1}$$

However, in robotics it is common that a random variable carries information about other random variables. Considering that assumption and returning to the example presented previously, equation 2.2 can be formulated.

$$P(C|E) = \frac{P(C,E)}{P(E)} = \frac{P(E|C)P(C)}{P(E)} \tag{2.2}$$

Equation 2.2 describes the Bayes rule. It expresses the relation between $P(C|E)$ and its inverse $P(E|C)$. C expresses the quantity to be inferred, using the knowledge of evidence E. The factor $P(C)$ represents the *prior probability distribution*, expressing the information available about C before the incorporation of the evidence E.

The probability distribution $P(C|E)$ is denoted as the *posterior probability distribution*. It describes the knowledge of C after integrating the data E and the *a-priori* information about C. The element $P(E|C)$ represents the *likelihood probability distribution*, which expresses the knowledge about how the variable C influences E. In robotics, this factor is also termed the *generative model*. The *likelihood probability distribution* is determined analytically, or it can result from a training period. The data acquired during this training period is used to learn the parameters of the probability distribution function $P(E|C)$. The literature [Ferreira and Dias, 2014b] presents different methods to perform the probabilistic learning of $P(E|C)$, such as *Maximum Likelihood* (ML) and *Expectation Maximization* (EM).

The factor $P(E)$ is a normalization constant. It guarantees that $P(C|E)$ sums up to 1, for all the domain of C. In some contexts, this parameter is not represented for simplicity purposes.

Additional details about the determination of $P(E|C)$ and $P(C)$ are given throughout this thesis as they are used to model each problem presented in this manuscript.

2.1.2 Bayesian inference

The *posterior probability distribution* $P(C|E)$ is used as a source of information to perform a decision about which state of C should be chosen. The work of [Ferreira and Dias, 2014a] presents different approaches to define the decision rule of the inference process.

In this thesis, the decision rule is formulated directly in the *posterior probability distribution* $P(C|E)$ by applying the *Maximum a-Posteriori* (MAP) principle.

The inferred value $\hat{c}$ of C is determined by selecting the argument of $P(C|E)$ which provides the highest value of probability, as presented in equation 2.3 .

$$\begin{aligned}
\hat{c} &= \arg\max_{c} P(C|E) \\
\hat{c} &= \arg\max_{c} \frac{P(E|C).P(C)}{P(E)} \\
\hat{c} &= \arg\max_{c} P(E|C).P(C)
\end{aligned} \tag{2.3}$$

2.1.3 Representing the Bayesian models

Throughout this thesis, several Bayesian models are formulated and described to provide a solution to the different challenges that are identified. In this manuscript, the Bayesian models are represented and characterized by using two different, but complementary, formalisms: Bayesian network and Bayesian program.

Bayesian network

A Bayesian network is a directed acyclic graph. The random variables of the Bayesian model are represented as nodes. The probabilistic (causal) relationships between pairs of

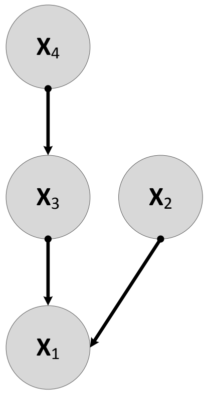


Figure 2.2: Graphical representation of a Bayesian model described by the random variables X_1, X_2, X_3, X_4 represented in the nodes. The causal dependencies are represented by the arrows (directed arcs).

random variables are represented as directed arcs. This representation approach provides an appealing visual description of the dependence relationship between random variables.

The dependence relationships expressed by the structure of a Bayesian network are used to simplify the formulation of the joint probability distribution function. These simplifications allow the design of efficient learning and inference algorithms based on simpler conditional probability distributions.

If the set of nodes which have arcs terminating at X_i is described by $parents(X_i)$, then equation 2.4 can be formulated. Let us consider a Bayesian model with N random variables $X_i, \ldots, X_N$ represented in a graph.

$$P(X_1, X_2, \ldots, X_N) = \prod_{i=1}^{N} P(X_i|parents(X_i)) \tag{2.4}$$

Equation 2.5 describes the joint probability distribution function for $N = 4$ and implementing the statistical causal dependence relations illustrated in the Bayesian network of Figure 2.2.

$$P(X_1, X_2, X_3, X_4) = P(X_1|X_2, X_3).P(X_3|X_4).P(X_2).P(X_4) \tag{2.5}$$

Bayesian program

The Bayesian program is a mathematical formalism and methodology used to organize and systemize the description of a Bayesian model. This approach facilitates the analysis and comparison of the properties of different Bayesian models and the respective

Bayesian Program
- Description
  - Specification
    - **Relevant variables:**
      Selection of the random variables to model the problem being studied.
    - **Decomposition:**
      The global joint probability distribution function is factored as the product of simpler probability distribution functions, expressing the conditional independence relations of the random variables.
    - **Parametric forms:**
      A custom mathematical function or a family of probability distribution functions (e.g. uniform, normal, Bernoulli, Poisson, Binomial) is associated to each of the factors identified during the decomposition stage.
  - **Identification:**
    The free parameters of the custom mathematical functions or of the families of probability distribution functions are estimated/learned from experimental data acquired during training sessions (or other sources of information).
- **Question:**
  This stage of the formalism describes the inference problem to be answered by the Bayesian model. The Bayesian model is queried concerning information about a given random variable. The answer provided by the Bayesian model (posteriori probability distribution function) is submitted to a Bayesian decision theory criterion (e.g. MAP)

(a)

Figure 2.3: Schematic description of the organization of the formalism used by a Bayesian program to describe a Bayesian model.

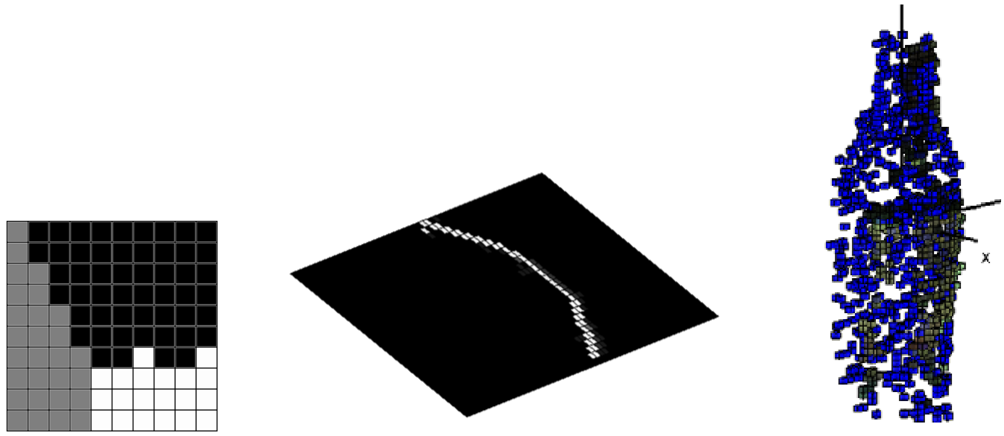
computational implementations.

This methodology represents the Bayesian model as it follows an approach consisting of different stages, which are described in Figure 2.3.

2.2 Probabilistic grids

In chapters 6 and 7 of this thesis, the environment surrounding the agent (robotic system or human) performing the haptic exploration is represented by a bi-dimensional probabilistic grid. The workspace is divided into a uniform grid of square cells.

A property of interest of the environment is associated to each cell with coordinates (i, j) and described by a random variable $X_{(i,j)}$. This approach considers that the repre-



(a) (b) (c)

Figure 2.4: Example of probabilistic grids used in robotics research. Probabilistic grid representing: a) the occupancy of a 2D environment [Rocha et al., 2005]. b) the haptic discontinuity between two regions of a surface [Martins et al., 2014]. c) the 3D shape of a hand explored object. Occupancy state is fused with color information from an artificial vision system [Faria et al., 2010a].

sentations of the cells of the grid are independent from each other.

The representation framework allows the integration/fusion of multi-modal data using probabilistic modelling techniques; this approach can deal with uncertainty of the data sources. The grid structure also provides the ability to represent heterogeneous environments (e.g. spatial discontinuities of the property of the environment being represented).

In several previous works (e.g. [Rocha et al., 2005], [Faria et al., 2010a], [Elfes, 1989]), the bi-dimensional probabilistic grids were typically used to represent the state of the workspace regions as *empty* or *occupied*. In this work, the state $X_{(i,j)}$ of each cell represents multivalued properties of the workspace regions, such as category of material.

The state $X^k_{(i,j)}$ of each cell is updated at each time iteration k by integrating new sensory measurements $Z^k_{(i,j)}$ acquired at that region of the workspace. At the time instant $k = n$, the cell (i,j) has integrated n sensory measurements $\mathbf{Z}^n_{(i,j)} = (Z^1_{(i,j)}, Z^2_{(i,j)}, \ldots, Z^n_{(i,j)})$. At that time instant, the state of each cell of the probabilistic grid is described by equation 2.6.

$$P(X^n_{(i,j)}|\mathbf{Z}^n_{(i,j)}) = \frac{P(Z^n_{(i,j)}|X_{(i,j)}).P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})}{P(\mathbf{Z}^n_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})} = \Theta.P(Z^n_{(i,j)}|X_{(i,j)}).P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)}) \quad (2.6)$$

The parameter Θ is a normalization constant. Consecutive sensory measurements $Z^k_{(i,j)}$ and $Z^{k-1}_{(i,j)}$ are considered independent.

According to equation 2.6, the updated representation of the state of each cell of the grid $P(X^n_{(i,j)}|\mathbf{Z}^n_{(i,j)})$, after a new sensory measurement, is given by $P(Z^n_{(i,j)}|X_{(i,j)})$ and $P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})$.

The factor $P(Z^n_{(i,j)}|X_{(i,j)})$ represents the likelihood probability distribution function, which expresses the sensor measurements model. It models the knowledge available of how the sensor measurements are affected by the possible state of the cell/workspace (i,j).

Alternatively, the factor $P(X^{n-1}_{(i,j)}|\mathbf{Z}^{n-1}_{(i,j)})$ describes the state of the cell (i,j) at the previous time iteration $n-1$. It encodes a complete summary of all past integration of sensory data by that cell of the grid.

2.3 Information theory and entropy

Several chapters of this thesis use random variables to model the proposed approaches and transfer information between the different modules. This PhD thesis uses Shannon entropy [Shannon, 2001a] to quantify the information encoded by a probability distribution function $P(X)$ of a random variable X. The formulation for discrete random variables is presented in equation 2.7.

$$H(X) = E[-log(X)] = \sum_x -P(x)\log_2(P(x)) \tag{2.7}$$

Higher values of entropy express lower levels of information (e.g. uniform probability distribution). Lower values of entropy encode higher levels of information (e.g. certain event).

Different approaches for calculating the entropy of continuous probability distribution functions are presented in [Gelfand and Yaglom, 1993].

Bibliography

[Chung and AitSahlia, 2012] Chung, K. L. and AitSahlia, F. (2012). *Elementary probability theory: with stochastic processes and an introduction to mathematical finance*. Springer Science & Business Media.

[Elfes, 1989] Elfes, A. (1989). Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57.

[Ernst and Bulthoff, 2004] Ernst, M. O. and Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in cognitive sciences*, 8:162–169.

[Faix et al., 2015] Faix, M., Mazer, E., Laurent, R., Othman Abdallah, M., Le Hy, R., and Lobo, J. (2015). Cognitive computation: A bayesian machine case study. In *Cognitive Informatics Cognitive Computing (ICCI*CC), 2015 IEEE 14th International Conference on*, pages 67–75.

[Faria et al., 2010] Faria, D., Martins, R., Lobo, J., and Dias, J. (2010). Probabilistic representation of 3d object shape by in-hand exploration. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 1560 –1565.

[Ferreira and Dias, 2014a] Ferreira, J. and Dias, J. (2014a). Bayesian decision theory and the action-perception loop. In *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*, pages 121–145. Springer International Publishing.

[Ferreira and Dias, 2014b] Ferreira, J. and Dias, J. (2014b). Probabilistic learning. In *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*, pages 147–167. Springer International Publishing.

[Ferreira and Dias, 2014c] Ferreira, J. F. and Dias, J. (2014c). *Probabilistic Approaches to Robotic Perception*, volume 91 of *Springer Tracts in Advanced Robotics*. Springer.

[Gelfand and Yaglom, 1993] Gelfand, I. and Yaglom, A. (1993). Amount of information and entropy for continuous distributions. In Shiryayev, A., editor, *Selected Works of A.*

N. Kolmogorov, volume 27 of *Mathematics and Its Applications*, pages 33–56. Springer Netherlands.

[Hertzberg and Chatila, 2008] Hertzberg, J. and Chatila, R. (2008). Ai reasoning methods for robotics. In Siciliano, B. and Khatib, O., editors, *Springer Handbook of Robotics*, pages 207–223. Springer Berlin Heidelberg.

[Knill and Richards, 1996] Knill, D. C. and Richards, W., editors (1996). *Perception as Bayesian inference*. Cambridge University Press, New York, NY, USA.

[Lebeltel et al., 2004] Lebeltel, O., Bessire, P., Diard, J., and Mazer, E. (2004). Bayesian robot programming. *Autonomous Robots*, 16(1):49–79.

[Martins et al., 2014] Martins, R., Ferreira, J. F., and Dias, J. (2014). Touch attention bayesian models for robotic active haptic exploration of heterogeneous surfaces. In *Proceedings of 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, pages 1208–1215, Chicago, USA. IEEE.

[Rocha et al., 2005] Rocha, R., Dias, J., and Carvalho, A. (2005). Cooperative multi-robot systems: A study of vision-based 3-d mapping using information theory. *Robotics and Autonomous Systems*, 53(3-4):282–311.

[Shannon, 2001] Shannon, C. E. (2001). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1):3–55.

[Thrun et al., 2005] Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.

[Westheimer, 2008] Westheimer, G. (2008). Was helmholtz a bayesian? *Perception*, 37(5):642–650.