Institut des Systèmes Intelligents et de Robotique
Autonomous and Mobile Integrated Systems / Perception Teams

Space Dimension Perception from the Multimodal Sensorimotor Flow of a Naive Robotic Agent

Alban Laflaquière, Sylvain Argentieri, Bruno Gas and Eduardo Castillo-Castenada

Context and Objectives

Traditionnal IA applied to Robotics commonly relies on models of the environnement, of the agent and of their interaction. We argue that all relevant information for the robot's behavior, usually captured in those models, can be actively extracted without any a priori knowledge. Indeed, a naive agent should be able to discover the structure of its interaction with the environnement through the analysis of its sensorimotor flow. This work proves the possibility to extract geometrical space features from a naive agent sensorimotor flow.

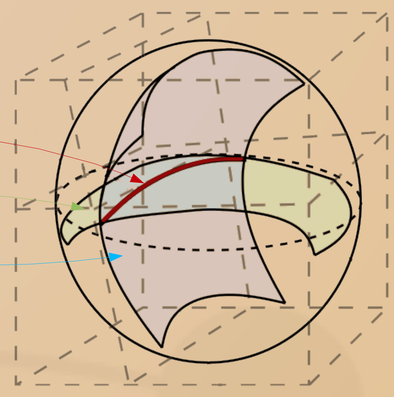
Theoretical background

According to H.Poincaré (1895), our perception of the space dimension is related to the space of « compensable movements ». This space lies at the intersection of two subspaces :

- The sensory variations generated when only the agent moves.
- The sensory variations generated when only the environment moves.

Knowing their dimensions (**p** and **e**) and the one of the whole sensory space (**b**), the dimension of the intersection is :

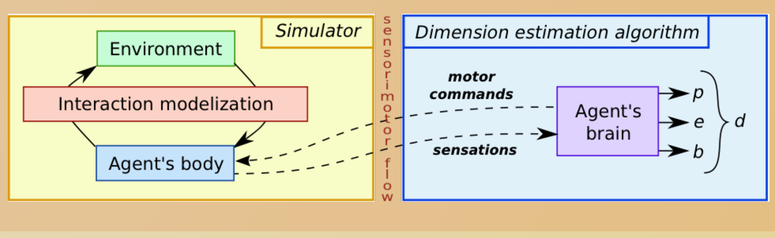
$$d = e + p - b$$



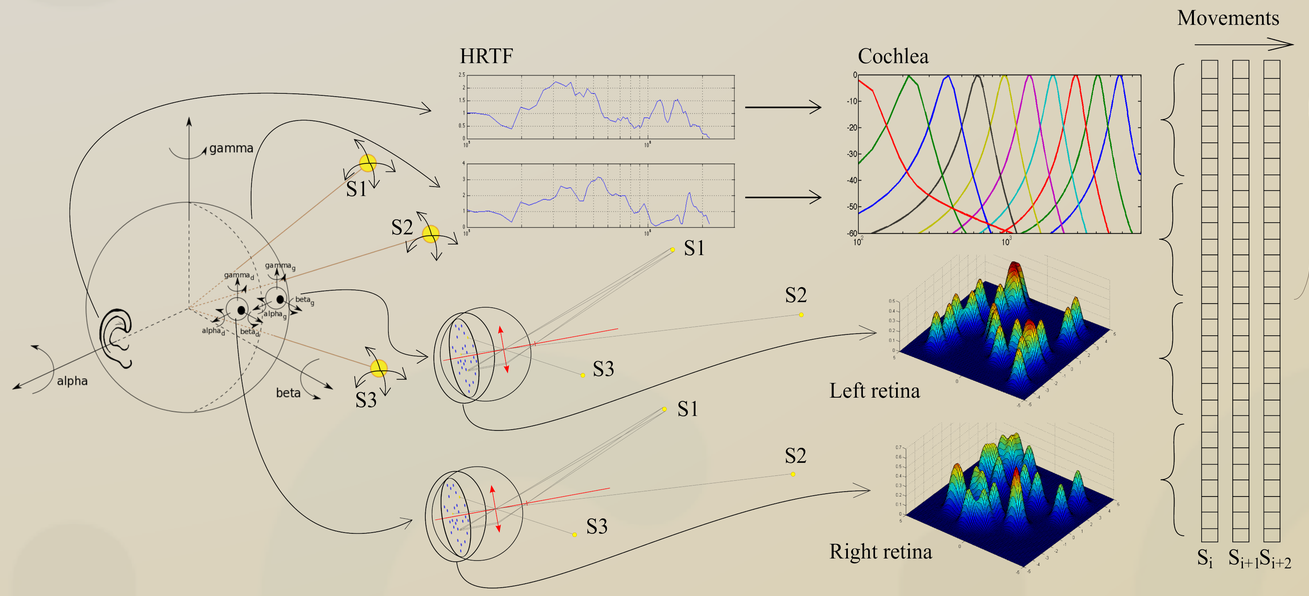
Sensory variations space

No a priori information about the agent nor its environment to evaluate **d**. The framework is thus split into two distinct parts :

- A simulator emulates the interaction between the agent and its environment.
- A dimension estimation algorithm determines **e**, **p**, **b** and **d** only through the sensorimotor flow of the agent.



Simulation results

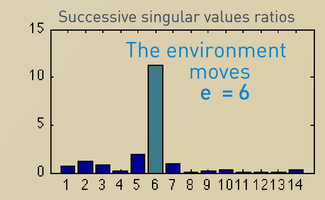


Sensory variation matrix M of size 120x40 : (20 cones by eye + 40 filter by cochlea) x (40 random movements around a working point).

Each space dimension is determined by looking at the number of significant singular values :

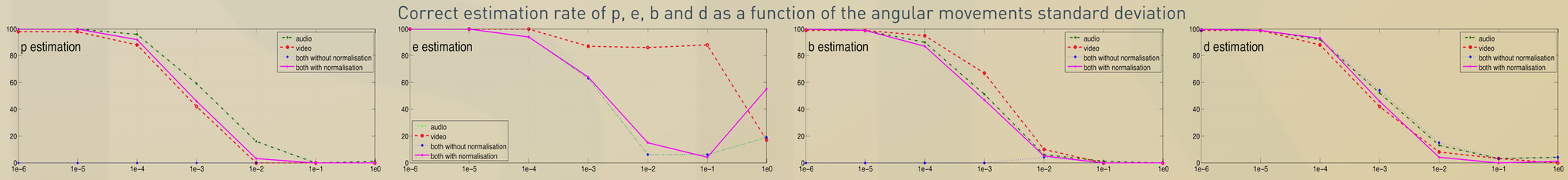
$$[\lambda_1 ... \lambda_n] = SVD(M)$$

$$\dim = \arg\left(\max\left(\frac{\lambda_i}{\lambda_{i+1}}\right)\right)$$



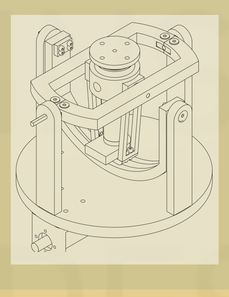
The estimated values **e**, **p**, **b** and **d** are consistent with the expected dimensions. The multimodal algorithm is efficient whatever the nature of the sensations (visual, auditive or both) but is only adapted to very small movements because of its linear processing of the data.

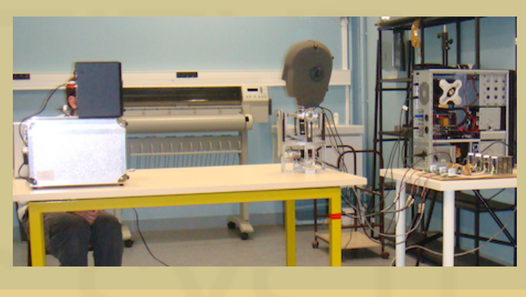
Correct estimation rate of p, e, b and d as a function of the angular movements standard deviation

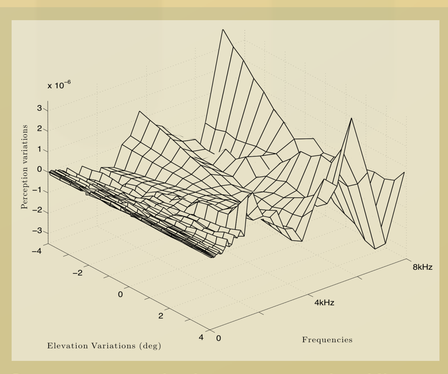


Preliminary experimental results

- A KU100 dummy head with two high-quality microphones embedded inside two realistic ears.
- A mechanical neck reproducing the human neck capabilities.
- A static white noise source.

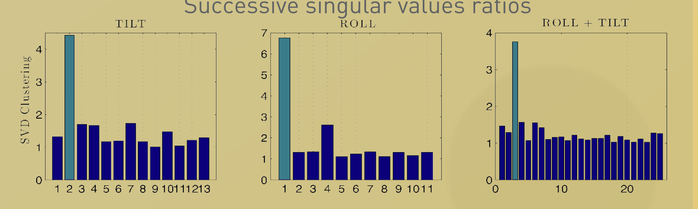






Sensory variations are significant even for small movements (here for the left ear during a roll movement of amplitude 10°) .

- **Expected dimension p** for a tilt or a roll movement is 1, and 2 when both are performed.
- **Estimated dimension p** is 2 for a tilt movement, 1 for roll movement and 3 when both are performed.



Estimation is partly wrong because of fairly large movements but remains consistent with the experiment as **p(tilt) + p(roll) = p(tilt + roll)**.

Relevance of the approach and perspectives

The long term goal of this approach is the development of new perception algorithm for mobile Robotics. A naive agent should be able to explore and characterize the structure of its interaction with the environment through action. The extracted knowledge would not be biased by any engineer design and would be directly relevant for the robot's action. This developmental stage would then lead to much more adaptive and autonomous robots. The intrinsic low-level and multimodal aspect of the approach is also interesting for research in information fusion. The next step consists in the development of new tools for the analysis of the sensorimotor flow. Non-linear processing would allow us to overtake the linear hypothesis that restrains the movements' amplitude of the agent.

A.Laflaquière, S.Argentieri and B.Gas are with UPMC Paris 06 and ISIR (CNRS UMR 7222), F-75005, Paris, FRANCE
name@isir.upmc.fr

E.Castillo-Castenada is with CICATA-Instituto Politecnico Nacional, Cerro Blanco 141, Colinas del Cimatario, 76090, Queretaro, MEXICO
ecastilloca@ipn.mx

This work was conducted within the France/Japan **BINAAHR** (BINaural Active Audition for Humanoid Robots) project funded by the French National Research Agency