The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

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

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

***Abstract*— Perception and action are fundamental tasks for autonomous robots. Traditionnally, they rely on theoretical models built by the system's designer. But, is a naive agent able to learn *by itself* the structure of its interaction with the environment without any a priori information ? This knowledge should be extracted through the analysis of the only information it has access to: its high-dimensional sensorimotor flow. Recent works, based on the sensorimotor contingencies theory, allow a simulated agent to extract the geometrical space dimensionality without any model of itself nor of the environment. In this paper, these results are validated using a more sophisticated auditive modality. The question of multimodality fusion is then addressed by fitting up the agent with vision. Finally, preliminary experimental results on a real robotic platform are presented.**

I. Introduction

Perception is a key issue in the design of agents capable of acting autonomously in unknown and dynamic environments. Classical approaches of perception in mobile robotics generally rely on models of the environment, of the agent's morphology and of its sensors. Those models do not consider action as a component of perception but as its outcome, through a decision-making system.

This classical view, that we call "passive perception", has now been reconsidered for a while by many authors [3], [8] that propose instead an "active perception" in which action is a necessary component of perception. This approach is supported by experiments of Held & Hein [5] and Bach-y-Rita [2] that highlight the critical role of volontary action in the emergence of perception. The sensorimotor contingencies theory, proposed recently by O'Regan and Noe [8], argues that the experience of perception is not the activation of internal representations but the capacity to engage oneself in some structure of interaction with the environment. Taking inspiration from Poincaré's argumentation [13] on what he called *sensible space*, Philipona [10] proposed a mathematical formalism to explore this approach. In [12] he describes an algorithm allowing a simple simulated agent to estimate the dimension of the geometrical space in which it is immersed without any other information than its sensorimotor flux. In a recent paper [4], we proposed to pursue Philipona's simulation and the estimation of the geometrical space dimensionality in the specific case of the auditory sensorimotor flow. In this article, we extend this work to both auditory and visual modalities. In addition, we present some preliminary results with a real robot. Our goal is to validate the active perception approach in a more realistic experimental framework.

The article is divided into three parts. The first part is devoted to the theoretical background, gives an overview of the simulation and the description of the dimensionality estimation algorithm. In the second part, we present the results obtained with a bio-realistic audition, then with an elementary vision, and finally with both of them. We examine the sensitivity of the method with respect to the coding of the sensorimotor flow. Finally, the third part gives an evaluation of the proposed algorithm performance on a real robotic platform. Preliminary results are discussed.

II. Theoretical background and simulation overview

In this section, the formalism of the approach is introduced. First, the underlying theoretical background is recalled on the basis of our previous work [4]. Next, the proposed simulation, providing the sensorimotor flow of a totally naive agent, is depicted.

A. Recall of the formalization

The requisite notions and notations for the understanding of this paper are shortly introduced in this subsection. The key idea is based on Poincaré's intuition [13] dating from 1895. He suggested that the perception of the geometrical space dimensionality is based on a group of specific transformations called *compensable movements*. These transformations are described by Poincaré as the sensory consequences of a class of an agent's movements that can be compensated by some specific movements of the environment. As an example, the sensory consequences of an object moving away from an agent can be canceled out by an appropriate movement toward the object. According to Philipona's work on the formalization of Poincaré's intuition [11], these sensory consequences –lying in the sensory space– can be linked to the environmental and the body's states through the sensorimotor law $\Phi(.)$. In this sensory space, compensable movements consequences are located on two subspaces intersection: the sensory consequences subspaces obtained 1/ when only the agent moves, and 2/ when only the environment moves. In order to work on these two subspaces and their intersection, Philipona hypothesized that the sensory space is a differential manifold which can be approximated by its tangent space on a small enough neighborhood. As a consequence, the dimension of the compensable movements

A. Laflaquière, S. Argentieri and B. Gas are with UPMC Univ Paris 06 and ISIR (CNRS UMR 7222), F-75005, Paris, FRANCE name@upmc.fr
E. Castillo-Castenada is with CICATA-Instituto Politecnico Nacional, Cerro Blanco 141, Colinas del Cimatario, 76090, Queretaro, MEXICO ecastilloca@ipn.mx

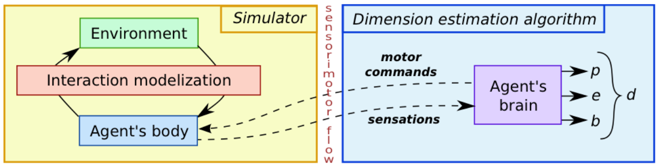

Fig. 1. Schematical representation of the approach: the simulator (left) relies on a modelization of the interaction between the agent's body and its environment; the agent's brain (right) has only access to the sensorimotor flow without any other a priori knowledge.

sensory consequences subspace can be deduced from the simple relation:

$$d = p + e - b, \tag{1}$$

where d denotes the dimension of the compensable movements sensory consequences subspace, p represents the sensory consequences subspace dimension when considering the movements of the agent only, e is the sensory consequences subspace dimension when considering the movements of the environment only, and b stands for sensory consequences space dimension when both the agent and the environment move.

Once the compensable agent's movements discovered, the Lie algebra can be exploited to distinguish between rotations and translations in the geometrical space [11] and then to extract the geometrical space dimensionality experienced by the agent. However, the forthcoming results only deal with the estimation of the dimension d.

B. Simulation overview

The whole simulation consists in two separate modules: a modelization of the interaction between the agent's body and its environment, and a dimension estimation algorithm accessing only the generated sensorimotor flow (see Figure 1). The proposed simulated agent is a simplified model of a human head that can rotate in pitch, roll and tilt (see Figure 2). It is fitted up with two ears, two independently mobile eyes, and is immersed in a three dimensional environment made up of several punctual light and sound sources. These sources live on a 1m-radius sphere centered on the head. Their position in the interaural frame is defined by two angles: θ, the azimuth between the sagittal plane and the source, and ϕ, the elevation between the transverse plane and the source. The head orientation with respect to the reference world frame is defined by the three angular parameters of pitch, roll and tilt (α, β, γ) which are independently controlled by three motors. The eyes orientations with respect to the head frame are defined by six angular parameters, corresponding to the pitch, roll and tilt of the left and right eyes, denoted $(\alpha_l, \beta_l, \gamma_l)$ and $(\alpha_r, \beta_r, \gamma_r)$ respectively. Each of these parameters is independently controlled by one motor. So, the whole agent configuration is fully controlled by 9 motors. Finally, the aforementioned punctual light and sound sources produce a continuous white light and a white noise of unitary variance.

The estimation of the dimension p relies on three successive steps:

- First, the agent executes a set of 40 movements around a working configuration, generated by applying random commands to its 9 motors; during this step, the environment remains static. For each movement, the sensory differences are computed by substracting the actual perception with the reference one measured in the working configuration. The resulting vectors S are then concatened in a sensory variation matrix M.
- Next, the singular values λ_i of this matrix are computed.
- Finally, the number of significant singular values corresponding to the matrix rank, and consequently to its intrinsic dimension p, is estimated by applying Philipona's method [12]. It consists in searching the index i producing the highest λ_i/λ_{i+1} ratio, i.e.

$$\dim = \arg\left(\max\left(\frac{\lambda_i}{\lambda_{i+1}}\right)\right). \tag{2}$$

This methodology is similarly used to estimate e and b, by considering respectively the environment moving randomly while the agent is static, and the environment and the agent both moving randomly.

III. Comparative Simulation Results

In this section, this estimation method is assessed by fitting up the agent with various sensitive capabilities: first a bio-realistic audition, next an elementary vision, and finally both of them. To conclude, the sensitivity of the method with respect to the coding of the sensorimotor flow is examined.

A. Simulation parameterization

The simulation is run for 100 different sources reference positions with respect to the head. For each trial:

- the initial motor commands are set to $0°$, and the sources initial azimuths and elevations are randomly drawn in the interval $[-45°; 45°]$,
- 2×40 movements around the reference position are randomly drawn, one for the environment, the other one for the agent.

The angular standard deviation of the movements is set to 10^{-6} degrees for both the agents' motors $[\alpha, \beta, \gamma, \alpha_l, \beta_l, \gamma_l, \alpha_r, \beta_r, \gamma_r]$ and the sources positions $[\theta, \phi]$. The number of sources in the environment is set to 3.

At the end of the simulation, 3 sets of 100 sensory variation matrices M are computed: one when only the agent moves, the other one when only the environment moves, and the last one when both move together. Each matrix is the concatenation of 40 column-vectors S made up of K perception coefficients.

B. Auditory modality

1) Modality description: As mentioned in the introduction, audition has already been exploited in our previous paper [4]. The same bio-realistic audition system is proposed in this work. It is based on real recordings of Head-Related Transfer Functions (HRTF) provided by the CIPIC database [1], and on a biologically plausible model of the cochlea from [9], [14]. More precisely, one cochlea is

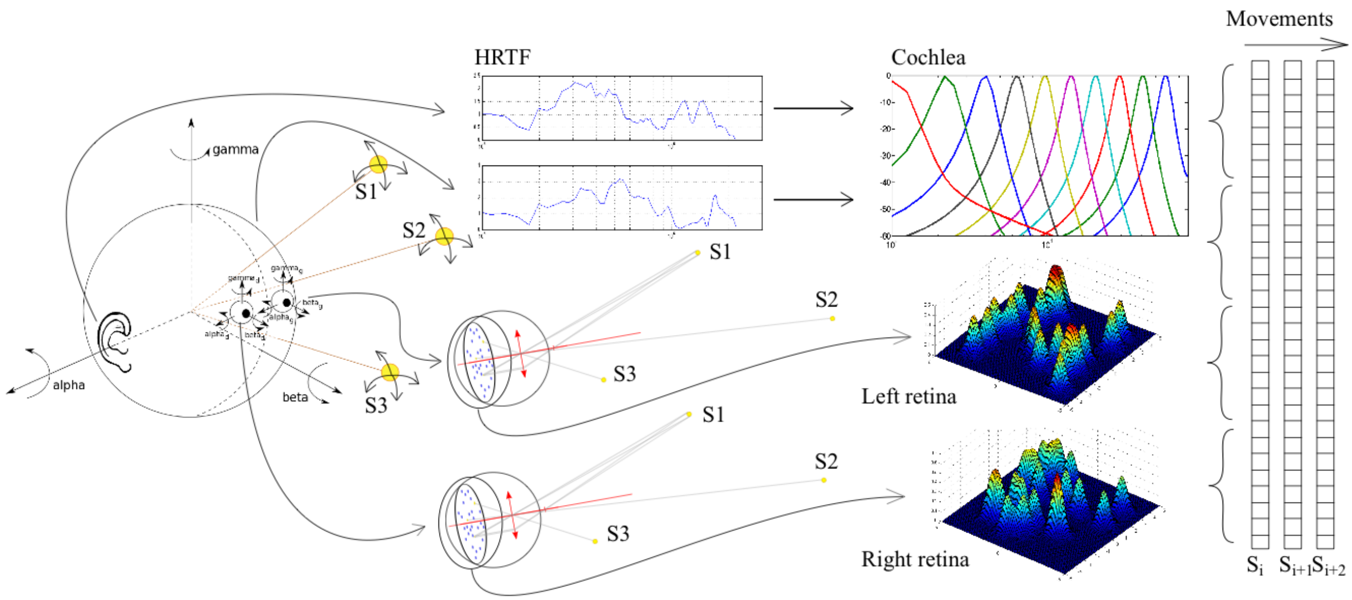

Fig. 2. Overview of the system: the head, represented on the left, is fitted up with an auditory modality (top) and a visual modality (bottom). The sound signals are first filtered by the physical structure of the head and the external ears (using HRTF) and next by a model of the cochlea. The filters output energies of both ears constitute the auditory part of the sensory vector S. The sources lights are projected on the retina of both eyes through the pupils. The set of cones excitations generated on the retina constitute the visual part of the sensory vector S. For different environmental configurations and/or motor commands, the sensory vector S varies. The scheme of the eye is taken from [12].

modeled by 40 gammatone filters, which leads to 40 energy coefficients per ear, computed on 60ms-length signals. Consequently, the dimension of the total auditory sensory vector is set to $K = 80$.

2) Dimension estimation with audition: Applying the method described in II-B, the algorithm estimates that p is equal to 3, e is equal to 6, b is equal to 6 and thus that d is equal to 3 in 100% of the trials (see Figure 3 for a standard deviation of 10^{-6}). As expected, these values are coherent with the simulated system. Indeed, p is equal to the number of independent parameters needed to describe the sensory variations when only the agent moves. These correspond to the 3 head orientation parameters influencing its hearing. Similarly, e is equal to the number of independent parameters needed to describe the sensory variations when only the sources move. Actually, this corresponds to the 6 position parameters of the 3 sources when only the environment moves. Nevertheless, such an insight intuition can not be performed on the dimension b insofar as only the difference between $(p+e)$ and b is geometrically meaningful. Finally, the resulting estimation of the dimension d of the compensable movements sensory consequences subspace is 3. This evaluation is consistent, as 3 independent rotations are geometrically allowed by the simulated system, leading to 3 independent compensable movements. Note that these results –even though trivial– are only obtained from the sensorimotor flow, without any a priori knowledge. As illustrated in Figure 1, the agent does not make use of any model of itself nor of the environment.

Moreover, the influence of the angular movements standard deviation has also been evaluated. As shown in Figure 3, the algorithm performance decreases for standard deviation larger than about 10^{-5} degrees. This loss of performance can be explained by the underlying manifold hypothesis (see II-A and [12]) that allows a linear approximation of the sensory space by its tangent space on a *small enough* neighborhood. Performing too large movements causes sensory variations outside of this neighborhood, and thus deteriorates the aforementioned linear hypothesis. In such a case, the SVD-based method is not relevant anymore, and an appropriate nonlinear dimension estimation method should be developed.

C. Visual modality

1) Modality description: The proposed visual modality is similar to the one implemented in [12]. The modelization captures the main aspects of visual perception without being biologically realistic. More precisely, the eyes are made up of a lens and a square retina dotted with 20 light-sensitive cells or cones. The light emitted by the sources is projected on the retina through the lens. Each cone stimulation E_i is then obtained with:

$$E_i = a\,\frac{\exp\left(-\,\mathrm{dist}(cone_i, proj)^2\right)}{\mathrm{dist}(eye, source)^2}, i \in [1, 20], \quad (3)$$

where a is the cones sensitivity, arbitrarily set to 10^{-3}. Equation (3) exhibits a Gaussian function of the distance $\mathrm{dist}(cone_i, proj)$ between the i^{th} light-sensitive cell $cone_i$ and the light projection, inversely proportional to the squared distance $\mathrm{dist}(eye, source)$ between the light source and the eye. For each cone, the total sensation is the sum of all the stimulations generated by each light source projection. So, each eye generates 20 sensations and consequently the dimension of the total visual sensory vector is $K = 40$. Note that the excitation function (3) has been chosen arbitrarily and any other function could be considered.

2) Dimension estimation with vision: Considering this new visual modality instead of the auditory one, the same experimental conditions as in §III-A are applied. Moreover, the cones distributions on the retinas are randomly drawn at each trial. The proposed algorithm estimates that $p = 9$, $e = 6$, $b = 12$ and $d = 3$ for 100% of the trials (see Figure 3 for a 10^{-6} degrees standard deviation - note that the p estimation strictly reaches 99%). Even if the values of p and b are different from their estimation with audition only, they remain coherent with the agent's new perceptive capabilities. In details, p is now equal to 9 as the 3 head orientation parameters together with the 2×3 eyes orientation parameters influence the vision when only the agent moves. In the same way, the b value is modified but the dimension d finally remains the same. Such a result explicitly demonstrates that the agent explore the *same* geometrical space, but with different sensing capabilities, still without any a priori information. As already outlined before, the performance of the algorithm decreases when the angular standard deviation of the movements increases, see Figure 3. This phenomenon can be explained by the same reason than already outlined in §III-B.2.

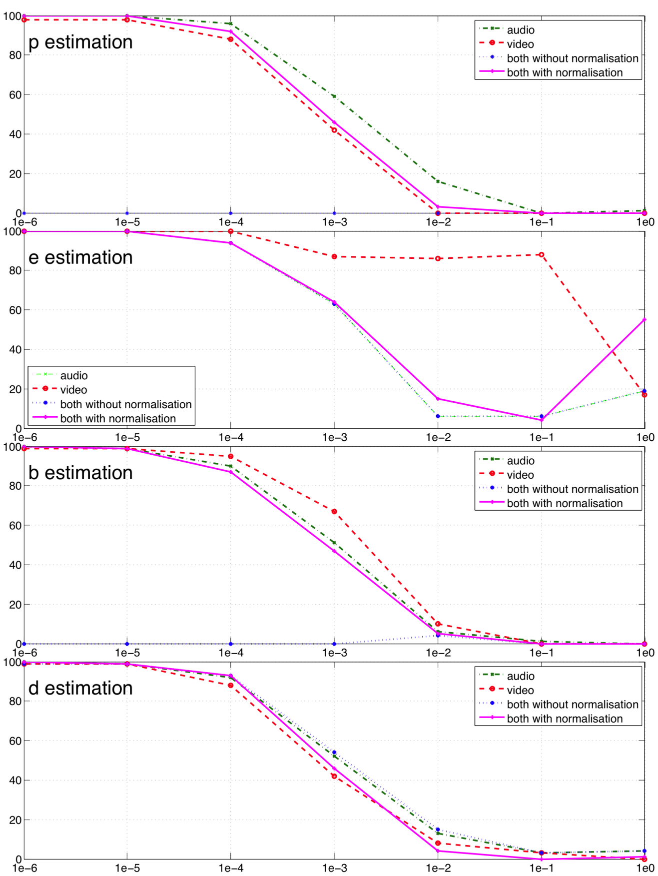

Fig. 3. Correct estimation rate of the dimensions as a function of the angular movements standard deviation of the agent or/and the sources. Note the dotted graph for the estimations of p and b are near 0.

D. Multimodality

In this subsection, the consequences of the fusion of audition and vision are investigated. In what follows, the experimental conditions are the same than before.

1) Dimension estimation with both modalities: Considering the simulated system, now including the two modalities, the expected estimated dimensions are the following. First, p should be equal to 9, insofar as there are still 9 parameters influencing the perception when only the agent moves. Next, e should be equal to 6 as there are once again 2×3 parameters acting on the perception when only the environment moves. Finally, b should be equal to 12 in order to obtain a dimension d of 3, as there are still 3 independent rotations allowed in the proposed system. Surprisingly, the estimation performance rate is 0% for p and b, while being 100% for e and d, see Figure 3. More precisely, the algorithm evaluates that $p = 3$ and $b = 6$ in 100% of the 100 trials. Interestingly, this set of dimensions corresponds to the agent fitted up with only audition. Thus it appears that the visual sensations do not have any influence on the dimensions estimation. Indeed, the sensations variations generated by the two modalities do not have the same order of magnitude: visual sensations are 10^3 times lower in average than the auditory ones. As a result, from a numerical point of view, the SVD has some difficulties taking into account the contribution of visual stimulations. Note that this amplitudes mismatch is only a consequence of the modalities arbitrary implementation, and not an intrinsic property of the interaction between the agent and its environment.

2) Sensory outputs normalization: As a solution, each row of the sensory variation matrix M is normalized so as to exhibit a unitary standard deviation. Concretely, this means that the outputs of all the cones and cochlear filters are now comparable in terms of variation. Interestingly, such a preprocessing seems biologically plausible according to the neuronal intrinsic plasticity [6]. After this normalization step, the algorithm is able to estimate the expected dimensions of p, e, and b. Importantly, the d value obtained without normalization was already 3: from this point of view, the sensory normalization sounds useless. But one has to keep in mind that this result involves only the auditory modality, without any multimodal fusion. Finally, the algorithm performance still decreases when the angular movement standard deviation increases (see Figure3), as already outlined in §III-B.2.

E. Coding-independence

In this subsection, the coding-independence of the method is illustrated by modifying the way the sensory information is encoded.

Firstly, the method sensibility to information redundancy is assessed. This is achieved by altering the size K of the sensory vector S to K', through a random mixing matrix of size $K' \times K$, $K' \in [10, 200]$. Notice that the matrix M is normalized as in §III-D.2 before applying the mixing matrix, so that the sensory variations are comparable. For a movements standard deviation of 10^{-6} degrees, the algorithm performance stays unchanged as long as $K' \geq 13$, but radically drops down to 0% if $K' < 13$. Actually, because of the successive singular values ratio-based method (2), at least 13 sensations are required to determine that $b = 12$. Secondly, the influence of the coding is analyzed for the auditory modality. Until now, this step consisted in the extraction of 2×40 energy coefficients, constituting the auditory part of the sensory vector S. These auditory features are now replaced by 2×20 MFCC coefficients, or by 2×40 interaural level and phase difference cues. Using these two different sensory features, the algorithm performance remains 100%, with or without the visual modality. All these examples illustrate the sensory coding-independence of the approach. A similar study could be performed on the motor side of the agent to emphasize its motor coding-independence too. Indeed, modifying the motor or sensory coding has no effect on the structure of the geometrical space the system is embedded in, but only on the way the sensory space is explored. Logically, the estimated dimensions p, e, b and d should then be unchanged, regardless of the coding. In fact, such coding modifications can only affect the maximum movements standard deviation insuring their correct evaluation. More precisely, p value could be affected solely by changes in the agent morphology. In the same way, e value could exclusively be altered by environmental modifications, and d only by the geometrical space dimensionality. Of course, any of these changes would lead to a different value of b according to (1).

IV. Preliminary experimental results

In this section, the proposed algorithm performance is evaluated on a real robotic platform to experimentally validate the underlying sensorimotor approach on real acoustic signals. The developed platform is first described and preliminary experimental results are then presented.

A. Experiment overview

1) The robotic platform: The proposed testbed is composed of two elements:

a) Measuring unit: it is made up of a KU100 dummy head from Neumann, equipped with two high-quality balanced microphones embedded inside two ears imitating the human pinnae (see Figure 4 (right)). Its outputs are simultaneously sampled and acquired with a National Instruments PCI acquisition card.

b) Movement unit: it is a simple physical structure allowing the 3 natural rotations of a human head: flexion (forward and backward), bending in the frontal plane and rotation in the transversal plane. It relies on three MAXON DC servomotors, three capacitive incremental encoders and a stainless steel structure as shown in Figure 4 (left). Since very few studies [16], [15] on accelerations, velocities and joint range of a human neck are available in the literature, they have been determined by using a motion capture system with 9 active markers located on a person's torso and head.

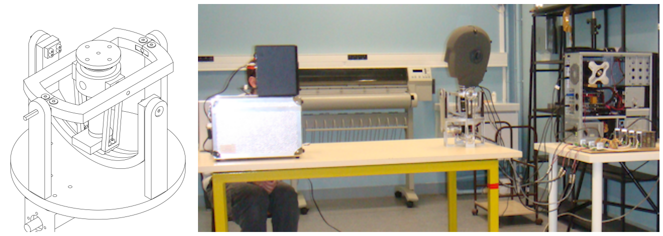

Fig. 4. System overview: (Left) The 3 DOF mechanical system reproducing the three rotations of a human head; (Right) Experimental setup, from the loudspeaker on the left to the acquisition card on the right.

2) Experimental setup: The records take place in the ISIR robotics hall, a highly reverberant and noisy room. First, a loudspeaker is laid at about 1m from the front of the head, diffusing a white noise signal (see Figure 4 (right)). Next, the head is rotated at one of the randomly chosen reference positions $P_1 = (\gamma_1, \alpha_1, \beta_1) = (-13°, 5°, -15°)$ or $P_2 = (\gamma_2, \alpha_2, \beta_2) = (0°, 0°, -30°)$ (see Figure 2). Then, the head performs small successive movements around the three rotation axis, with a 0.5 degrees angular step. At each position, the sounds acquired in the two ears are digitally converted with a sampling frequency $f_s = 44.1$kHz and recorded during 5s. Each signal is then processed by 40 cochlear filters identical to the ones exploited in §III-B, leading to a total sensory vector of length 2×40. The variations of these coefficients with respect to their values at the reference position are then concatenated in a matrix M to be analyzed using the method described in §II-B. Note that for the moment, this experiment allows only the estimation of the dimension p as the head is mobile but not the environment yet.

B. Preliminary results

1) Noise and perception variations: First of all, two successive records of *silence* have been exploited to obtained the SNR of the experiment at each cochlear filter output (see Figure 5 (left)). It is 28dB for the lowest frequencies, while it reaches up about 70dB for the highest ones.

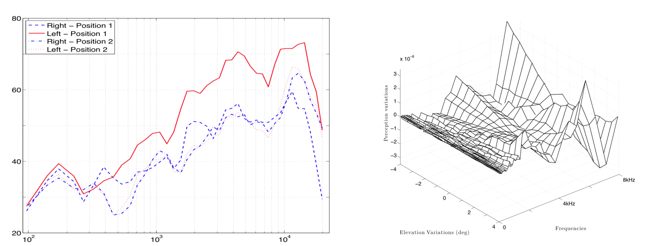

Fig. 5. (Left) Signal-to-Noise Ratio of the experiment as a function of the frequency. The SNR has been computed at the cochlear filters outputs. (Right) Variations of the 40 cochlear energy coefficients as a function of the frequency and of the position around the reference position for the roll movement (left ear).

In the same vein, the order of magnitude of the 80 energy coefficients variations has been evaluated. These values are shown in Figure 5 (right), for the roll movement around the reference position P_2 and considering only the left ear. The highest frequencies exhibit the most significant perception

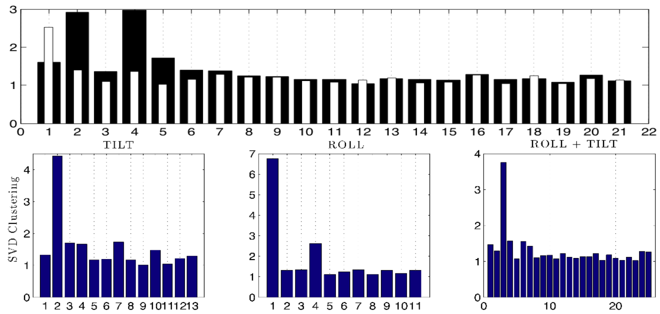

Fig. 6. (Top) Successive singular values ratio for the azimuth movement (ref. point P_1), with (white) or without (black) normalization; (Bottom) Successive singular values ratios for the tilt and roll rotations (ref. point P_2), independently or jointly considered.

variations, which appear quite linear with respect to the roll parameter for such small movements.

2) Experimental p estimation: Two successive preliminary experimental results are proposed. They both concern the estimation of the dimension p, obtained by moving the dummy head while the environment remains static. First of all, only the perception variations around the azimuth axis are investigated, for the reference position P_1. The expected dimension estimation is $p = 1$ since only one parameter is needed to describe the sensory variations. After 23 angular steps have been executed, the variations vectors S of the left and right cochlear energy coefficients are concatenated in the matrix M to be analyzed. The resulting successive singular values are shown in Figure 6 (top) without (black bars) and with (white bars) the standard deviation normalisation of the matrix M, see §III-D.2. The maximum value index with the normalization is 1, which corresponds to the expected estimation $p = 1$. The second experiment involves the two remaining roll and tilt rotations, performed at the reference position P_2. The resulting singular values as well as their successive ratios are exhibited in Figure 6 (bottom). For each axis, the expected estimation value is $p = 1$. Nevertheless, the maximum value index for the tilt movement is 2, while the evaluation is correct for the roll axis. This can potentially be explained by the fairly large movements performed during the experiment, degrading the linear approximation of the sensory space by its tangent space. Nonetheless, the estimated dimension obtained by simultaneously considering the two angular movements is 3: although the anticipated value should be 2, this result seems encouraging as it corresponds to the sum of the two individually estimated dimension.

V. Conclusion and perspectives

In this paper, an active and model-free feature extraction approach from a high-dimensional sensorimotor flow has been presented. It has been applied to the estimation of the geometrical space dimensionality by a naive simulated agent fitted up with a simple visual modality and a bio-realistic auditory modality. The algorithm validation in simulation allows an experimentation on a real platform for which some preliminary results have been introduced. The main limitation of the method consists in the restricted range of movements ensuring a good estimation of the dimension. Non-linear dimension estimation tools should be developed to overcome this issue: non-linear mapping and curvilinear component analysis methods [7] are being evaluated for this purpose. More generally, we are currently working on the extraction from the sensorimotor flow of new features which could be more relevent to the emergence of behaviors. Beside, this active approach proves to be of high interest as it requires no prior modelization of the interaction between the agent and its environment. Moreover, the sensors sophistication, in terms of information richness, would not rise as an obstacle anymore. It is particularly hopeful in mobile robotics where designers mostly use simple sensors whose modelization remains accessible.

Acknowledgment

This work was conducted within the French/Japan BINAAHR (BINaural Active Audition for Humanoid Robots) project under Contract n°ANR-09-BLAN-0370-02 funded by the French National Research Agency.

References

[1] V.R. Algazi, R.O. Duda, R.P. Morrisson, and D.M. Thompson. The cipic hrtf database. *Proceedings of the 2001 IEEE Workshop on Applications of Signal Processing to audio and Acoustics*, pages pp. 99–102, 2001.

[2] P. Bach-y Rita, C. C. Collins, F. Sauders, B. White, and L. Scadden. Vision substitution by tactile image projection. *Nature*, 221:963–964, 1969.

[3] A. Berthoz. *Le Sens du Mouvement*. Odile Jacob, février 1997.

[4] C. Couverture and B. Gas. Extracting space dimension information from the auditory modality sensori-motor flow using a bio-inspired model of the cochlea. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2742 – 2747, 2009.

[5] R. Held and A. Hein. Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, 56(5):872 – 876, 1963.

[6] P. Joshi and J. Triesch. Rules for information maximization in spiking neurons using intrinsic plasticity. In *International Joint Conference on Neural Networks*, 2009.

[7] J.A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Information Science and Statistics Series, Springer, 2007.

[8] J.K. O'Regan and A. Noë. A Sensorimotor Account of Vision and Visual Consciousness. *Behav. Brain. Sci.*, 24(5):939–973; discussion 973–1031, 2001.

[9] R.D. Patterson and J. Holdsworth. A functional model of neural activity patterns and auditory images. *Advances in Speech, Hearing and Language*, 3:547–563, 1996.

[10] D. Philipona. *Développement d'un cadre mathématique pour la notion de dépendances sensorimotrices, dans le contexte d'une théorie de l'expérience sensorielle*. PhD thesis, Ecole Normale Supérieure de Cachan, janvier 2008.

[11] D. Philipona, J. K. O'Regan, and J.-P. Nadal. Is There Something out There? Inferring Space from Sensorimotor Dependencies. *Neural Comp.*, 15(9):2029–204, 2003.

[12] D. Philipona, J.K.n O'Regan, J-P. Nadal, and O. J-M.D. Coenen. Perception of the structure of the physical world using unknown sensors and effectors. *Advances in Neural Information Processing Systems*, 15, 2004.

[13] H. Poincaré. L'Espace et la Géométrie. *Revue de métaphysique et de morale*, 3:631–646, 1895.

[14] K. Robenson, R.D. Patterson, and J. Holdsworth. Complex sounds and auditory images. In *Processings of Auditory Physiology and Perception*, volume 9, pages 429–446, 2004.

[15] R.S. Sodhi. Evaluation of head and neck motion with the hemi-spherical shell method, international journal of industrial ergonomics. *Journal of Industrial Ergonomics*, 25(6):683–691, 2000.

[16] J. W. Youdas. Normal range of motion of the cervical spine: an initial goniometric study. *Physical Therapy*, 72(11):770–780, 1992.