Phonotaxis Behavior in the Artificial Rat Psikharpax

Mathieu Bernard, Steve N'Guyen, Patrick Pirim
Brain Vision Systems
Paris, France
mathieu.bernard@bvs-tech.com

Bruno Gas, Jean-Arcady Meyer
ISIR, UPMC / CNRS
Paris, France

***Abstract*—Psikharpax is an artificial rat that aims at becoming as autonomous and adaptive as it's biological model. Its auditory system calls upon a pair of mobile pinnae, artificial cochlea, and a spike extraction algorithm. We show how this system endows the robot with a phonotaxis behavior that depends more on an efficient and robust peripheral auditive system than on an evolved sound source localization process. Moreover, this system also affords an emergent solution to the front-back ambiguity problem because the robot is able to retrace its steps if the sound source is localized in its back. In the absence of the pinnae such capacity is lost.**

I. Introduction

The development of the auditory system described herein takes place within the Psikharpax project [1], which aims at designing a bioinspired artificial rat. Indeed, both the morphology and the control architecture of this robotic platform (see Fig. 1) are inspired by those of the living rat. Complementary to visual and tactile modalities, one of our objectives is to conceive a complete bioinspired auditory system that will allow Psikharpax to mimic natural behaviors directly linked to hearing (approach or avoidance behavior, sound recognition, etc) and to send to higher level processes (such as navigation and action selection modules) an efficient and robust auditory information. Related approaches have been devoted to the echolocation system of bats [2] or to active and bimodal (vision and audition) aspects of perception for object tracking in humans [3]. Mimicking natural behaviors in natural environments requires real time and robust perception systems in order to make autonomy and reactivity possible on a robotic platform. Obviously a rat's auditory system is well adapted to these constraints and designing a bioinspired system becomes of evident interest. According to [4], we designed our model as a behavioral model with high level of abstraction. It does not aim to precisely fit biological data but reproduces the functional organization of the rat's auditory system.

This paper describes the phonotaxis behavior thus afforded to Psikharpax. Few bioinspired phonotaxic robots have been developed in the past. One such a system [5] emulates the neural circuitry underlying cricket phonotaxis. Another one exploits diffraction in a spherical head and require a model of it's acoustic properties [6]. Therefore, in the absence of more specific cues, we chose to conceive phonotaxis in rats as an iterative two step process. First, the sound source direction is estimated by computing the interaural energy (or intensity) difference [7] which provides an estimation of the side of the head where the sound is the loudest. This lateralization step calls upon the peripheral auditory system, especially the outer ear model. Then, a sensory-motor loop is used for control according to which the robot's head and body move towards the side were the source is detected.

Fig. 2 shows an overview of the complete system presented in this paper, which is organized as follow : section II describes in detail the peripheral auditory system (outer ear, cochlear model and spike extraction), section III presents the energy-based lateralization process and the sensory-motor loop devoted to the phonotaxis task. Different phonotaxis trajectories are presented in section IV. We also show that the robot is able to retrace its steps if the sound source is localized in its back. In the absence of the pinnae such capacity is lost. The paper concludes with a short discussion and provides indications for further work.

Figure 1. The artificial rat Psikharpax.

II. Peripheral Auditive System

The peripheral auditory system in mammals consists of elements upstream the auditory nerve, *i.e.* the outer, middle and inner ears. Its role is to filter and encode the input sound signal so as to provide a fine time and frequency representation allowing the brain to deal with noises, echoes and simultaneous sound sources [8]. Following Lyon [9], our approach is focused on the design of an efficient and robust peripheral auditory pathway whose first element is a pair of artificial pinnae. A sound wave, reflected by a given pinna, is captured by a microphone and send afterwards to a cochlear model. Finally, a spike extraction algorithm is applied to the output of the cochlea in order to increase the system's robustness to noise and the biological plausibility of this preprocessing. Moreover, by setting an important proportion of values to zero, spike generation decreases algorithmic complexity of higher level computations.

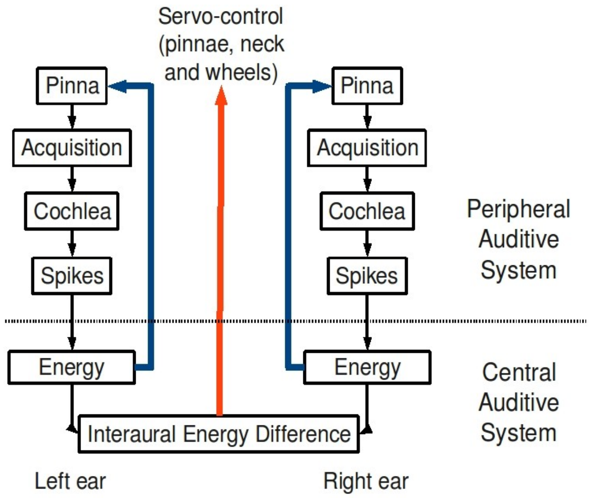

Figure 2. Overview of the complete auditive system allowing phonotaxis behavior. Blue arrows (pinnae) refers to a maximization sensory-motor loop and orange arrow (both neck and wheels) to a minimization sensory-motor loop.

A. Outer ear : an artificial pinna

The outer ear of mammals is analogous to a complex filter which depends of the direction and spectral components of the sound source and of the acoustical properties and movements of the body (including torso, head, pinna and ear canal). This outer ear provides spectral cues allowing the brain to compute the elevation of a sound source. Based on physiological and psychophysical studies, robotic heads with artificial pinnae are already used [10, 11] to investigate and simulate these effects. Moreover, sensor asymmetry could be exploited by robotic systems, as in barn owls, to solve the front/back ambiguity [12].

The system reported in this paper implements a pair of artificial pinnae, as shown in Fig. 3. Each pinna is an ellipsoid section supported by a servomotor. A sound signal coming from the environment, reflected by the pinna, is captured by a microphone oriented towards the center of the pinna. Finally the analogic stereo signal is digitalized by an embedded sound card and send by a wireless connection to a remote computer. As discussed in section IV, this outer ear model provides efficient intensity cues for sound source lateralization.

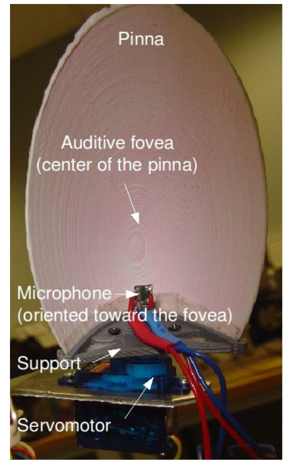

Figure 3. The mobile pinna of Psikharpax.

B. Inner ear : a cochlear model

Taking place in the inner ear, the cochlea is the organ of hearing. It is the place where mechanical vibrations are converted into spike trains and transmitted to the brain through the auditory nerve. The input sound signal is separated into several frequency-band channels, each auditory nerve fiber representing the activity around a particular frequency. This frequential selectivity of the cochlea is spatially organized: from the base to the apex of the cochlear membrane, the selectivity decreases from highest to lowest frequencies. This tonotopic organization is found in all levels of the auditory pathway, especially in subcortical centers. Moreover, the cochlea has an efferent innervation, which allows active adaptation in terms of frequential selectivity and gain control. Reference [8] provides a recent review about these issues.

There exist several cochlear models in the literature witch try to reproduce internal (both passive and active) phenomena of the cochlea, *e.g.* membrane elasticity or fluid mechanics [13]. Such models are precise but far too complex for any robotic perception system. Another popular model is based on gamma tone filters [14]. The auditory system presented here uses the Lyon's cochlear model [15, 16], an efficient black-box cochlear model : it doesn't model the internal physiology but aims to reproduce the output of the cochlea thanks to a two dimensional filter bank, as showed in Fig. 4. First, receiving a digital signal from the middle ear, low pass filters (the cutoff frequencies gradually decreases) reproduce the cochlear tonotopy by decomposing the signal into frequency bands. Then, the resonators at the second level of this filter bank allow the frequential selectivity of each channel thanks to a peak transfer function. Finally, a compression step allows an automatic gain control which reproduces the active behavior of the cochlea. The Lyon's cochlear model reproduces in a simple way the most important aspects of the cochlea. However, in comparison with the unidimensional input signal, the outputs of this model are n continuous signals, each corresponding to one of the n channels. In order to reduce the computational complexity of this model and to increase its robustness to noise, we added to it a spike extraction algorithm. Thus, each local extremum of each output channel is considered as a spike if it is superior to a given threshold. All values which are not detected as spikes are set to zero. Consequently, this threshold has a practical importance : it allows the robot to be more or less deaf to its environment and to suppress the motor and background noises. Finally, the output of our peripheral auditory system is a multichannel spike train which represents the auditory nerve information flow.

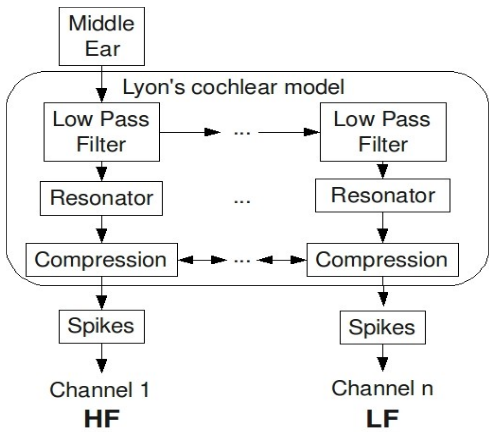

Figure 4. Overview of the inner ear model.

III. Azimuthal Lateralization and Phonotaxis

In the description of the peripheral auditory system, we focused on the passive features of the Psikharpax's auditory system. As we can see in Fig. 1, this system has different actuators (mobile pinnae, neck and wheels). Their control is included in a sensory-motor loop that allows active sound source localization and phonotaxis behavior. This section presents the computation of the interaural energy difference used for the lateralization of sound sources. Then, the sensory-motor loop allowing pinnae movements, head-tracking and phonotaxis behavior is introduced.

A. *Interaural energy difference*

Suppose a binaural system with two sensors $x_1(t)$ and $x_2(t)$, and a sound source $s(t)$ in a noisy environment. We have :

$$x_1(t) = s(t) + n_1(t) \text{ and } x_2(t) = \alpha . s(t+\Delta t) + n_2(t) \quad (1)$$

where $n_i(t)$ is the noise captured by the sensor i. α represents the attenuation, or interaural level difference (ILD), of the signal from the first to the second microphone. Δt represents the time-lag, or interaural time difference (ITD), between the signal's arrivals at the two sensors. The neural basis of both ITD and ILD computations have been located in the superior olivary complex of mammals [8]. There exists also many binaural hearing models using these cues for sound source localization [17, 18]. According to [7], we chose to base our lateralization model on the ILD estimation. The energy computed represents the cochlear activity for a given time interval, which is correlated to the current loudness of the auditory scene. In this way, by computing the interaural energy difference, we estimate the ILD. For a given C-channels spike train, the current energy is expressed as :

$$E = \sum_{t=0}^{T} \sum_{n=1}^{C} s_n(t)^2 , \quad (2)$$

where $s_n(t)$ is the spike value of the channel n for the sample t and T is the sampled time between two motor commands. Finally, the interaural energy difference is simply computed as :

$$\Delta E = E_{left} - E_{right} . \quad (3)$$

This difference is afterwards compared to a threshold and the dominant sound source is localized in the front, the left or the right of the robot.

B. *Pinnae servoing*

Fig. 2 shows that mobile pinnae are directly controlled by energy maximization, which allows each pinna to move towards the direction were the loudest sound is perceived. Pinnae are controlled independently, but synchronously, from the neck and the wheels. Under the hypothesis of a directional selectivity of the pinnae, periodic auditory jerks aim to increase the signal to noise ratio thanks to active energy maximization. For a given sensor, the pinna servoing system is a two-step process based on the successive values of the energy E, each T samples. Firstly, to initialize the process, a little rotation is done randomly towards the left or the right side. Secondly, as in the inner ear model described above, spikes are extracted from $E(t)$. A spike is emitted if the sound's loudness sharply increases or decreases. If the current energy is decreasing, the rotation is inverted. Finally, to maintain the stability of this system, rotations are prevented and the pinnae are centered in case of too low energy. Movements of pinnae's motors produce noise suppressed by thresholding the spike generation.

C. *Phonotaxis sensory-motor loop*

Inspired by a taxis control framework [19], the phonotaxis sensory-motor loop is based on the estimation of the sound source direction given by interaural energy difference. At first, the head makes a few degrees rotation towards this source. Then, the wheels are controlled by a velocity vector. The velocity component of this vector *(i.e.* the speed of the robot) is set to a constant value while the directional component is the same as the horizontal head orientation : if the head is oriented 45° to the left, the robot will move 45° towards the left. The same process is repeated at each iteration (*i.e.* every T samples), which is clearly sufficient for the phonotaxis task, as described below.

IV. Phonotaxis experiments

As shown in the previous section, the phonotaxis task is accomplished by a sound source lateralization process coupled to a simple sensory-motor loop. The present section presents the phonotaxis trajectories realized in a laboratory environment. In order to investigate the outer ear influence on the phonotaxis behavior, we also present the trajectories obtained with the pinnae removed.

A. Phonotaxis trajectories in normal conditions

A sound source – a constant loudness white noise – diffused by a speaker, is placed in the proximity of the robot. The sample frequency used for capture was 44.1 kHz and a 32 channels cochlea was used. Microphones were previously calibrated. Fig. 5 shows that Psikharpax is able to move towards a sound source from different positions, even if it is localized in his back (black plots on Fig. 5). In this case, the robot makes a turn back to the source. In this way, the phonotaxis solves the front-back ambiguity as a side effect. Nevertheless, if a sound source is placed exactly in the back of the robot, the corresponding energy difference (above the threshold) is so small that the source is localized in front of the head. This problem is solved when the head is turned by a few degrees so as to let the interaural energy difference become significant. It is such an asymmetry that is created and exploited by the sensory-motor loop and makes it possible to detect if the sound source comes from the back or the front of the robot, so as to achieve the phonotaxis task.

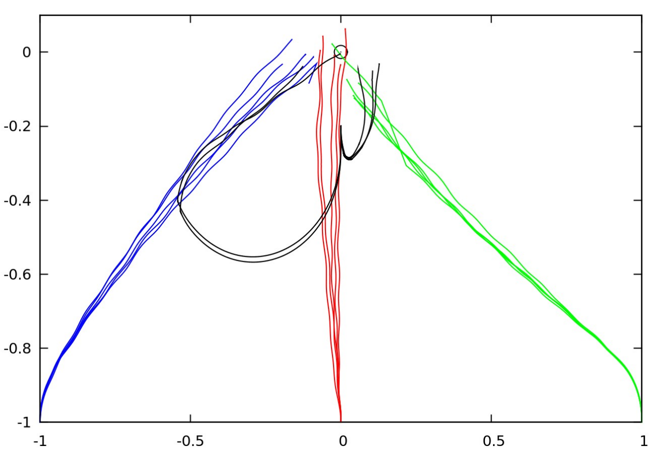

Figure 5. Phonotaxis trajectories from different start situations : front (red), left (blue), right (green) and back (black). Sound source position at (0, 0). Axes in meters.

B. Outer ear influence on phonotaxis trajectories

The same experience as in Fig. 5 was performed with a single difference : the pinnae were removed and the microphones were oriented towards the front of the head. As Fig. 6 shows, even if the pinnae are removed, the phonotaxis is achieved correctly. Nevertheless, the trajectories towards a side source are less direct. The major interest of this experience is to show that removing the pinnae disables the front-back disambiguation: even if an asymmetry is created by moving the head, the energy difference remains lesser than the threshold. In this case, the robot trajectory is a line from the starting point towards the opposite side of the sound source, as shown by the arrow in Fig. 6. This fact illustrates the interest of the outer ear model for the phonotaxis task: it generates a bigger energy difference when a sound is coming from a side of the head.

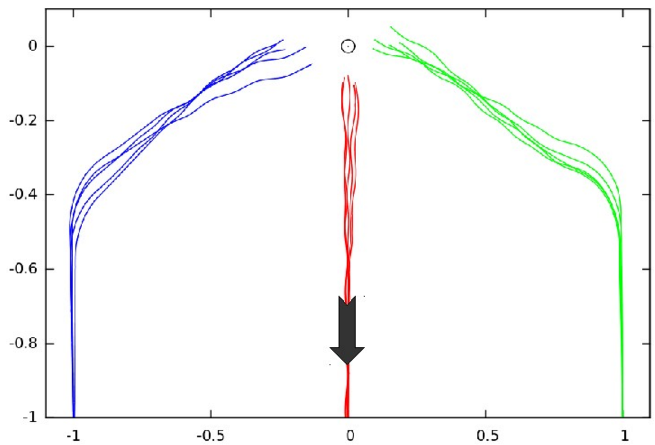

Figure 6. Same experience as in Fig. 5, with pinnae removed and microphones oriented toward the front of the head. Sound source position at (0, 0). Axes in meters. Trajectories from back start situation are represented by the black arrow, see text for details.

V. Conclusion and Further Work

This paper presented a bioinspired phonotaxic system which allows an autonomous mobile robot to move towards a sound source localized in its proximity. The movements of pinnae, neck and wheels are controlled by an energy driven sensory-motor loop. As shown in the previous section, Psikharpax's pinnae provide good cues for lateralization by improving the interaural energy difference. Moreover, the pinnae allow the robot to turn back if a sound source is localized in its back. Traditional approaches in binaural sound source localization are more interested by temporal or spectral cues, allowing a qualitative estimation of the sound source direction. By restricting our model to intensity cues, active aspects of perception become the more important determinants of sound source lateralization and phonotaxis. Contrary to the few works dedicated to bioinspired robotic phonotaxis, we based our approach on a robust, generic and efficient peripheral auditory system. Robustness is afforded by the spike-generating process which renders the robot more or less deaf to motor and background noises. Although such preprocessing may appear too complex with respect to the simple binaural computations a priori needed by phonotaxis, this - relative - complexity is intended to serve as a basis for future higher-level computation, such as sound source segregation and recognition.

Acknowledgment

We thank Alexis Rodet and Christophe Grand who respectively implemented the robot's cochlear model and its pinnae. This work has been funded by the EC Integrated Project ICEA (Integrating Cognition, Emotion and Autonomy), IST-027819-IP.

References

[1] J. Meyer, A. Guillot, B. Girard, M. Khamassi, P. Pirim and A. Berthoz. “The Psikharpax project: towards building an artificial rat,” Rob. Autnom. Syst., Vol. 50, 2005, pp. 211-223.

[2] H. Peremans, and J. Reijniers, “The CIRCE head: a biomimetic sonar system,” ICANN, 2005, pp. 283-288.

[3] H. Okuno, K. Nakadai, T. Lourens and H. Kitano, “Sound and visual tracking for humanoid robot,” Appl. Intell., Vol. 20, 2004, pp. 253-266.

[4] B. Webb, “Can robots make good models of biological behaviour ?,” Behav. Brain Sci., Vol. 24, 2001, pp. 1033-1050.

[5] R. Reeves and B. Webb, “New neural circuits for robot phonotaxis,” Philos. Transact. A Math. Phys. Eng. Sci., Vol. 361, 2003, pp. 2245-2266.

[6] S. Andersson, A. Handzel, V. Shah and P. Krishnaprasad, “Robot phonotaxis with dynamic sound-source localization,” IEEE Rob. Autom. Int. Conf. on, Vol. 5, 2004, pp. 4833-4838.

[7] L. Smith, “Using IIDs to estimate sound source direction,” ICSAB, From animals to animats, MIT Press, 2002, 60-61.

[8] R. Warren, Auditory Perception : An Analysis ans Synthesis, 3rd ed., Cambridge University Press, 2008.

[9] R. Lyon, “Computational models of neural auditory processing,” IEEE Trans. Acoust. Speech Signal Process., 1984, Vol. 9, pp. 41-44.

[10] J. Hornstein, M. Lopes, J. Santos-Victor and F. Lacerda, “Sound localization for humanoid robots - building audio-motor maps based on the HRTF,” IROS, 2006, pp. 1170-1176.

[11] V. Algazi, R. Duda, D. Thompson and C. Avendano, “The CIPIC HRTF database,” IEEE ASPAA Workshop., 2001, pp. 99-102.

[12] O. Nobutaka, F. Souichiro, and S. Shigeki, “Sound source localization with front-back judgement by two microphones asymmetrically mounted on a sphere,” J. Multimedia, Vol. 3, 2008, pp. 1-9.

[13] E. Givelberg and J. Bunn, “A comprehensive three-dimensional model of the cochlea,” J. Comput. Phys., Vol. 191, 2003, pp. 377-391.

[14] A. Katsiamis, E. Drakakis and R. Lyon, "Practical gammatone-like filters for auditory modeling", J. on Audio, Speech, and Music Processing, 2007, pp. 12.

[15] R. Lyon, “A computational model of filtering, detection, and compression in the cochlea,” IEEE Trans. Acoust. Speech Signal Process., Vol. 7, 1982, pp. 1282–1285.

[16] M. Slaney, “Lyon's cochlear model,” Apple Technical Report, Vol. 13, 1988.

[17] J. Blauert, Spatial Hearing : The Psychophysics of Human Sound Source Localization, 2nd ed., MIT Press, 1997.

[18] A. de Cheveigné, “Time domain processing in the auditory system,” Proc. ICONIP Conf., 1998., pp. 1327-1332.

[19] I. Rano, “On taxis for control and its qualitative solution on mobile robots,” Proc. IAV Conf., 2007.