Neural Processing of Auditory-tactile Sensor Data to Perform Reactive Behavior of Walking Machines

Poramate Manoonpong
Fraunhofer Institute
Autonomous Intelligent Systems (AiS)
Sankt Augustin, Germany
poramate.manoonpong@ais.fraunhofer.de

Frank Pasemann
Fraunhofer Institute
Autonomous Intelligent Systems (AiS)
Sankt Augustin, Germany
frank.pasemann@ais.fraunhofer.de

Joern Fischer
Fraunhofer Institute
Autonomous Intelligent Systems (AiS)
Sankt Augustin, Germany
joern.fischer@ais.fraunhofer.de

***Abstract*— Spiders can sense sounds in a frequency range between approximately 40 and 600 Hz by the use of hairs; they can detect e.g. the puff of wind of buzzing flies. On the contrary, scorpions use hairs as tactile sensors for obstacle avoidance. To integrate the advantages of both types of sensoric hairs, this article presents an artificial auditory-tactile sensor system, which combines the principles of the auditory hairs of spiders and the tactile hairs of scorpions, and investigates some neural techniques for processing these sensor signals. The different types of signals are discerned by recurrent neural networks in such a way that their output can generate different reactive behavior, like obstacle avoidance and tropism, of a walking machine. An evolutionary algorithm is applied to find an appropriate solution to this problem.**

I. INTRODUCTION

In nature, the biological auditory system plays an important role in the life of an animal. It is used for different tasks like prey-detection, communication and localization. For instance, the wandering spiders (*Cupiennius salei*) use their *Trichobothria* hairs at their limbs as an auditory sensor system. They detect low-frequency sound ranging from approximately 40 to 600 Hz, like the puff of wind created by buzzing flies [6], [16], which act as prey signals. As the auditory system, the tactile sensing system is necessary for insects and animals to avoid obstacles while wandering around or seeking food. The obvious example of tactile-hair sensors is the sensory hairs of scorpions. Scorpions have a very poorly performing visual system, which has difficulties to recognize obstacles; therefore they use their hair sensors on the pectines and chelicera [9], [15] to avoid obstacles while walking.

Analogs of these auditory and tactile hair sensor systems of spiders and scorpions can be useful in providing sensor information for a sensor-driven control system in wheeled robots as well as in walking machines. There still exist implementations of both types of sensors on real robots [1], [2], [4], [10], [11], [13], [17] but roboticists have not yet implemented these two sensor functions into one sensor system. The obtained sensor signals can be analyzed by either using a Fast Fourier Transform [3], [5] or by using diverse filter techniques. These methods are often too slow for generating a fast reactive action of machines, they are often too complex, and sometimes too expensive.

Here a simple combined auditory-tactile sensor is introduced together with its neural signal processor. It should enable autonomous walking machines to move around for in-door applications. The sensor shall protect the legs of walking machines from hitting obstacles, like chair or desk legs, and allow navigation based on sound tropism. This approach transfers knowledge of an artificial whisker system [14] with a real mouse whisker attached, hair-like, to a capacitor microphone, and modifies it for our auditory-tactile application. For data processing of the mixed sensor signals an evolutionary algorithm is used to develop the structure of an appropriate recurrent neural network, and to optimize its parameters, such that it can generate the desired behaviors.

The following section describes the construction and specifications of an artificial auditory-tactile sensor. Section 3 explains the neural networks for preprocessing of the sensor signals in order to recognize the two different inputs. The experiments and results are discussed in section 4. Conclusions and an outlook on future work are given in the last section.

II. AN ARTIFICIAL AUDITORY-TACTILE SENSOR

This sensor consists of a mini-microphone (0.6 cm diameter), a root and a whisker-shaped material taken from a whisker of a real mouse (4.0 cm long). This is shown in Fig. 1.

In order to build this sensor, the mouse whisker was inserted and fixed into a root by special glue and then a root was glued onto the diaphragm of a microphone with relatively hard glue. The circuit designed for this sensor is a simple microphone preamplifier with Vcc maximum at 5 V. The physical force of the whisker vibrates the membrane of the capacitor microphone, which results in a voltage signal.

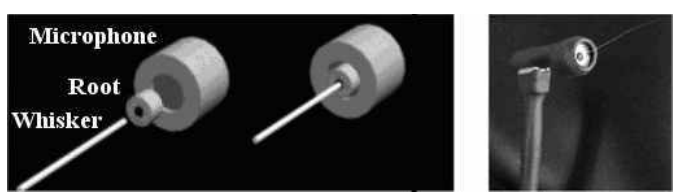

Fig. 1. The auditory-tactile sensor consisting of a whisker of a real mouse, a rubber root and a capacitor microphone. Left: assembly parts of a sensor; Right: the real sensor including a preamplifier circuit.

This signal is amplified by the microphone's integrated amplifier circuit, and is then sampled onto a sound card via a line-in of a computer for the purpose of recording and feeding it afterwards into the artificial neural network.

III. A SIMPLE NEURAL SENSORY SIGNAL PROCESSOR

The following approach to signal processing applies dynamical properties of recurrent neural networks. The standard additive neuron model with sigmoidal transfer function together with its time-discrete dynamics is given by

$$a_i(t+1) = B_i + \sum_{j=1}^{n} W_{ij}\ \tanh(a_j(t)) \quad i = 1, \ldots, n \quad (1)$$

where n denotes the number of units, a_i their activity, B_i represents a fixed internal bias term together with a stationary input to neuron i, and W_{ij} synaptic strength of the connection from neuron j to neuron i. The output of neurons is given by the sigmoid $o_i = \tanh(a_i)$. Input units are configured as linear buffers.

The desired network is divided into two subordinate networks, one for processing auditory signals to detect low-frequency sound, and one for processing tactile signals when the hair is sweeping over obstacles. Later the outputs of both networks will drive the corresponding reactive behavior of walking machines.

A. A low-pass filter for auditory signals

In order to create an auditory processing network, which is able to recognize frequency ranges of sound between 50 and 300 Hz, we first choose input signals of sine shape of 100 Hz and of 1000 Hz and map them to a voltage range between -1 and 1. To keep the problem simple we first use an ideal noise-free signal with constant amplitude. If a network is found which can distinguish between low-frequency (100 Hz) and high-frequency (1000 Hz) sounds, the next step of the experiment is to apply noisy sounds with varying amplitudes in a realistic environment to get a robust auditory processing network for low-frequencies.

First we utilize a single model neuron configured as a hysteresis element [7]; i.e., the network consists of an input neuron and a neuron with positive self-connection corresponding to a dynamical neural Schmitt trigger [12] (compare Fig. 2). Applying results from [12], we fix the weight $W_1 = 1$ from the input to the output unit, the bias term $(B = -0.1)$ and vary the self-connection weight W_2 of the output unit from 0 to 2.5 (see Fig. 2). For $W_2 = 2.45$ the network suppresses high-frequency sound of 1000 Hz, while low-frequency sound of 100 Hz passes through it.

By varying a weight W_2 of the self-connection of the output unit, one observes a splitting of the output signal, due to the hysteresis effect, which is different for different frequencies. This suggests that the hysteresis domain of a single neuron with self-connection can play an important role for the filtering of signals. To visualize this phenomenon, output versus input

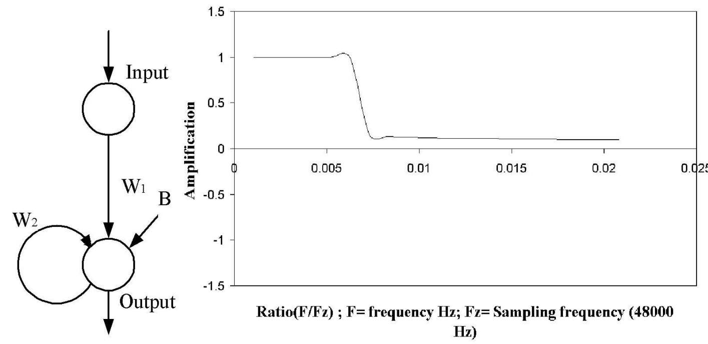

Fig. 2. Left: a simple network realizing a low-pass filter ; parameters are $W_1 = 1$, $W_2 = 2.45$ and $B = -0.1$. Right: the characteristic curve of this network with its cut off frequency at 300 Hz.

for low-and high-frequency signals are plotted in Fig. 3, and the different "hysteresis effects" can be compared with respect to the different strengths of the self-coupling.

Fig.3 shows that the hysteresis effects for high-frequency sound occur already for $W_2 = 0.25$, although it can not yet be observed for low-frequency sound. If W_2 is increased up to $W_2 = 2.45$ high-frequency sound is almost suppressed (low output) whereas the hysteresis effect for low-frequency sound switches the amplitude between almost saturation values. Increasing the self-connection up to $W_2 = 2.50$ also low-frequency sound is suppressed.

Because the bias term defines the base activity of the neuron, the amplitude of an high-frequency output is compensated and broken up between -0.804 and -0.998; eventually it will never rise above 0 again. In this situation, we suggest a low-pass filter function for a configuration with this specific bias (-0.1) and weight ($W_2 = 2.45$). The neural network behaves as a low-pass filter because the output amplitude of high-frequency sound stays around -0.9 while the output amplitude of low-frequency sound still oscillates between -0.997 and 0.998. More experiments and results will be demonstrated and described in the next section.

Having established that a single neuron is able to act as a low-pass filter for noise-free signals of constant amplitude, the next step is to derive a network, which behaves like a robust low-pass filter and which is capable to recognize low-frequency sound in a realistic environment. We improve the simple auditory network now by adding one self-connected hidden unit, and by adjusting again the weights. The final result, an advanced low-frequency detection network, is shown in Fig. 4.

The first synapse W_1 and the excitatory self-connection W_3 of the hidden unit reduce the amplitude of high-frequency sound. It becomes smaller than the amplitude of low-frequency sound. Afterwards the signals are again amplified by W_2. Then the bias term B together with the excitatory self-connection W_4 of the output unit shift the high-frequency signal to oscillate around -0.998 with a small amplitude. As result the network suppresses the high-frequency sound.

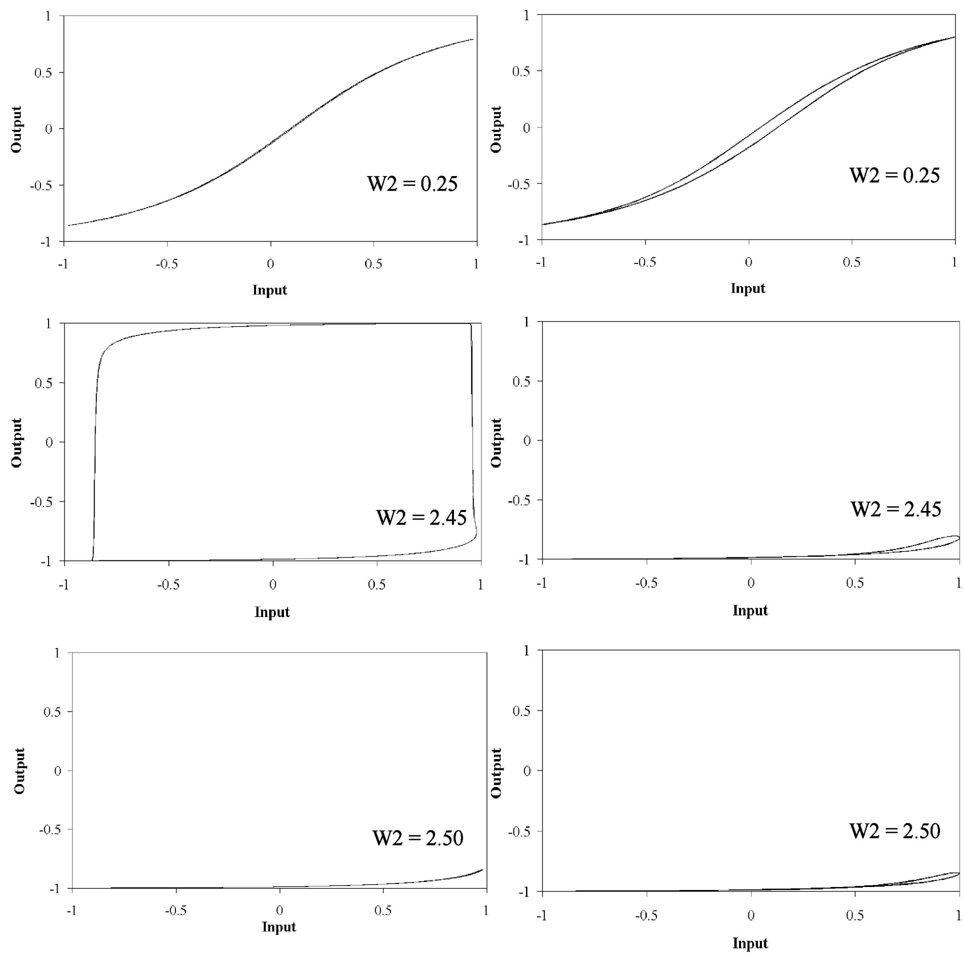

Fig. 3. Comparison of the "hysteresis effects" between input and output signals of high-and low-frequency sounds for W_2= 0.25, 2.45 and 2.50, respectively. Left: low-frequency sound (100 Hz); Right: high-frequency sound (1000 Hz).

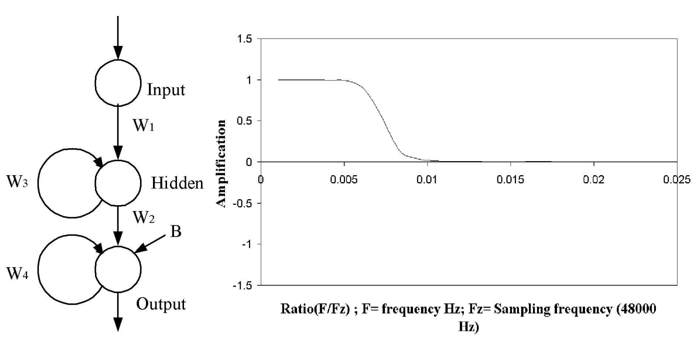

Fig. 4. Left: a simple and robust network for auditory processing, performing as a low-pass filter for noisy signals with varying amplitude. The bias term B is equal to -6.7 and all weights are positive, W_1 = 0.01, W_2 = 32, W_3 = 1 and W_4 = 0.27. Right: the characteristic curve of the network displaying the cut off frequency at 300 Hz.

B. Processing tactile signals

To process the signals from the tactile channel of the sensor, we apply the ENS^3-algorithm (Evolution of Neural Systems by Stochastic Synthesis [8]) to evolve an appropriate neural network. At the beginning only one input and one output unit without connections are given. The ENS^3-algorithm then increases or decreases the number of an synapses and hidden units throughout the evolutionary process, and optimizes the parameters at the same time, until the output signals are good enough for a reasonable solution. The fitness function F is chosen in such a way that evolution minimizes the square error between target and output signals; i.e., it is defined by

$$F = \sum_{t=1}^{N} \left(1 - (target(t) - output(t))^2\right) , \qquad (2)$$

where N is the maximal number of time steps, usually set to $N \approx 20000$. The target signal gives a +1 if a tactile signal is presented, and -1 in all other cases. This is exemplified in Fig. 5, where on the left the real sensor signals are shown, and on the right a corresponding target signal is depicted.

The resulting network at 800 generations had a fitness value of F = 0.6, which is sufficient to recognize the tactile signals. It consists of 2 hidden units and 7 synapses as shown in Fig.6.

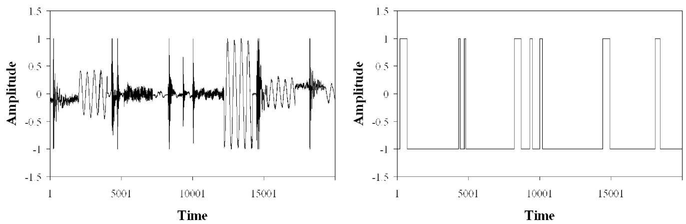

Fig. 5. Left: real signals coming from the physical sensor. Right: the corresponding target function.

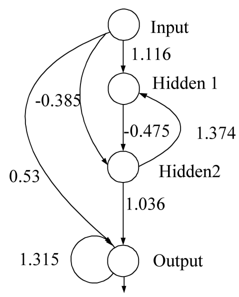

Fig. 6. A network processing tactile sensor signals. It filters the low-frequency sound. Its output signal follows the tactile signal, which has a high frequency around 2000 Hz.

To explain the detailed function of the network will go beyond the scope of this article.

C. The integrated network

One now can combine the advanced auditory processing network and the tactile processing network to obtain an integrated network which is able to distinguish between low-frequency sound and tactile signals. This network, consisting of one input unit, 3 hidden units and 2 output units, is shown in Fig.7. It is active at output 1 and oscillating between approximate 0.998 and -0.997 if low-frequency sound is recognized, and it is active at output 2 if a tactile signal is recognized. Otherwise both outputs are inactive.

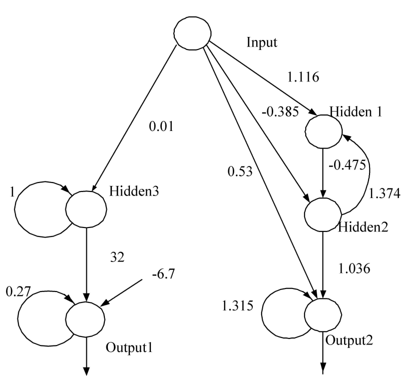

Fig. 7. The combined network recognizes low-frequency sound around 100 Hz (Output 1), and the tactile signal around 2000 Hz (Output 2).

IV. Experiments and Results

To test the capability of the auditory-tactile processing network, several experiments have been carried out. The input signal is recorded via an artificial auditory-tactile sensor and the output signals are sampled through the line-in of a sound card at a sampling rate of 48 kHz. The network is applied to a 1 GHz personal computer.

The experiments are divided into 3 parts. First, we feed constant amplitudes of noise-free input signals (Fig. 8a) into the simple auditory network (see Fig. 2) as well as into the advanced auditory network (see Fig. 4) . Then the same procedure is done with the noisy signals recorded via a sensor from a realistic environment (Fig. 8b) and finally both networks are tested with the signals obtained from the sensor installed on a walking machine's leg (Fig. 8c). Second, we test a tactile processing network by applying the tactile signal recorded from sweeping a sensor over the object (Fig.9). Third, we experiment with an artificial auditory-tactile processing network by feeding in mixed signals between low-frequency sound and the tactile signal (Fig. 10).

Fig. 8 shows that the simple auditory network is able to recognize the low-frequency signal when the signal is noise-free with constant amplitude. For the the noisy signals, the advanced auditory network is more robust and it is sufficient to detect the sound with a high enough amplitude. Furthermore, the advanced auditory network is also able to filter noise coming from the motors of a walking machine during walking and standing. Therefore, we integrate the advanced auditory network into the tactile processing network for signal processing of the sensor.

The output signal (see Fig. 9) from the network proves that discrete time dynamical systems as well as an evolutionary algorithm are able to construct the tactile processing network. The output signal is shifted to around –0.77 when the tactile signal is not present.

V. Conclusions

An auditory-tactile hair sensor was presented which consists of a mini-microphone with an integrated pre-amplifier, and a real mouse-whisker attached on it. In a couple of experiments it was shown that a simple recurrent neural network can discern between low-frequency auditory signals coming from this sensor, and higher frequency signals related to tactile information. Part of the network has been developed by an evolutionary algorithm to derive a network which is robust against real world noise. The output signals from this network will then be used to drive the reactive behavior of a walking machine controlled by an evolved recurrent neural network. For instance, the tactile signals should generate negative tropism, and the low-frequency sound a positive tropism so that the machine follows a sound source but avoids obstacles. Thus, with the auditory-tactile hair sensor together with the developed network implemented on a walking machine the described set-up should function in analogy to the sensor systems of spiders and scorpions.

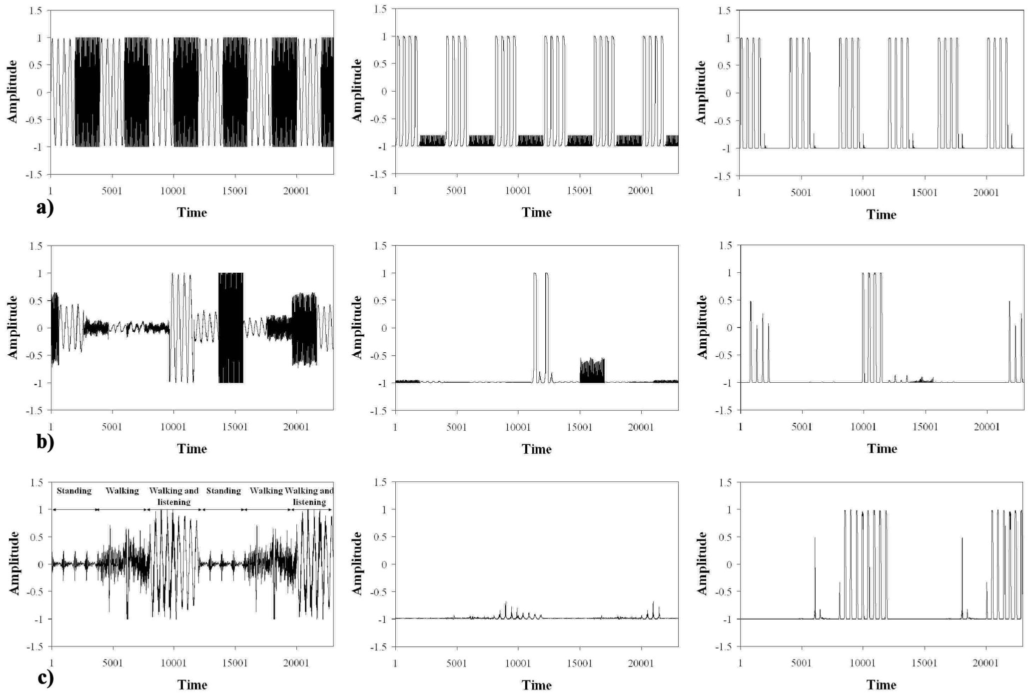

Fig. 8. Left: the input signals for the auditory networks. Middle: the corresponding output signals from the simple auditory network. Right: the corresponding output signals from the advanced auditory network. All figures have the same scale in x-axis and y-axis.

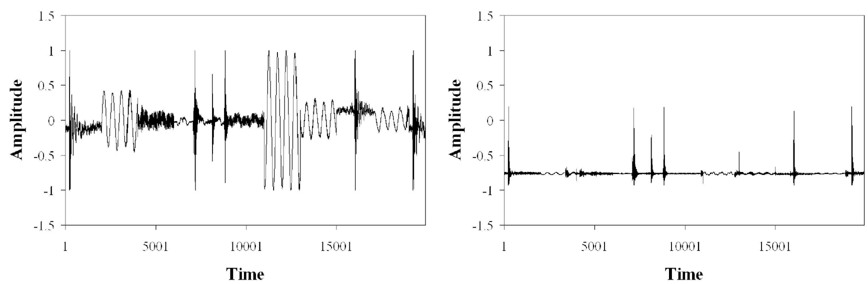

Fig. 9. Left: mixed input signals with sound at 100 Hz and the tactile signal. Right: the response of the network to the tactile signal. Both figures use the same scale in x-axis and y-axis.

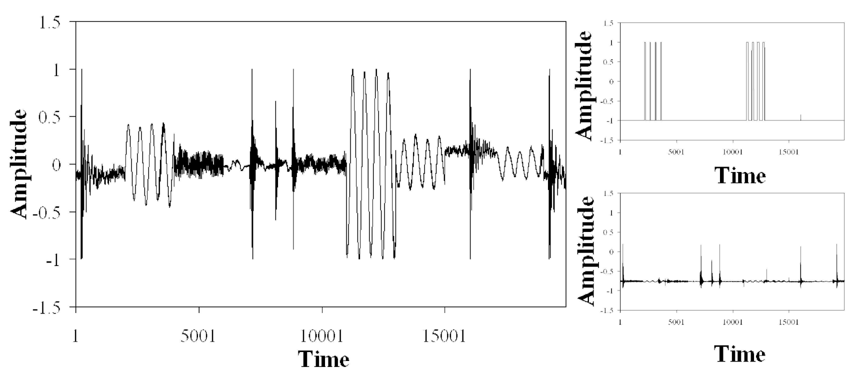

Fig. 10. Left: mixed input signals, sound at 100 Hz and tactile signals, for the network. Right: signals at output 1 (upper right) and 2 (lower right); they are active only for sound and tactile signals, respectively. All other signals are suppressed. All figures use the same scale in x- and y-axis.

VI. ACKNOWLEDGEMENTS

For the ISEE (Integrated Structure Evolution Environment) software platform for the evolution of recurrent neural networks (RNN) substantial support was provided by Keyan Zahedi, Martin Hülse, Björn Mahn and Steffen Wischmann. The authors would like to thank also Manfred Hild for valuable suggestions and discussions.

REFERENCES

[1] A. D. Horchler, R. E. Reeve, B. H. Webb and R. D. Quinn, "Robot phonotaxis in the wild: A biologically inspired approach to outdoor sound localization," *11th International Conference on Advanced Robotics (ICAR'2003)*, Coimbra, Portugal, 2003.

[2] B. Webb and R. R. Harrison, "Integrating sensorimotor systems in a robot model of cricket bahavior," *Sensor fusion and Decentralised Control in Robotic Systems III*, Boston, 2000.

[3] D. Coulter, *Digital Audio Processing*, CMP Books, KS, USA, 2000.

[4] D. Jung and A. Zelinsky, "Whisker-Based Mobile Robot Navigation," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 497–504, 1996.

[5] E. Oran Brigham, *The Fast Fourier Transform and its Applications,* Prentice-Hall, NJ, USA, 1988.

[6] F. G. Barth and Geethabali, "Spider Vibration Receptors Threshold Curves of Individual Slits in the Metatarsal Lyriform Organ," *Journal of Comparative Physiology. A*, vol. 2, pp. 175–186, 1982.

[7] F. Pasemann, "Dynamcis of a single model neuron," *International Journal of Bifurcation and Chaos*, vol. 3, no.2, pp. 271–278, 1993.

[8] F. Pasemann, U. Steinmetz, M. Hülse and BrunoLasa, "Robot Control and the Evolution of Modular Neurodynamics," *Theory in Biosciences*, pp. 311–326, 2001.

[9] F. T. Abushama, "On the behaviour and sensory physiology of the scorpion *Leirus quinquestriatus*," *Animal behaviour*, vol. 12, pp. 140–153, 1964.

[10] H. Hirose, S. Inoue and K. Yoneda, "The whisker sensor and the transmission of multiple sensor signals," *Advance Robotics*, vol. 4, no.2, pp. 105–117, 1990.

[11] M. Fend, S. Bovet and V. Hafner, "The Artificial Mouse - A Robot with Whiskers and Vision," *Proceedings of the 35th International Symposium on Robotics (ISR)*, Paris, 2004.

[12] M. Hülse and F. Pasemann, "Dynamical Neural Schmitt Trigger for Robot Control," *J. R. Dorronsoro(Ed.): ICANN 2002* , LNCS 2415, Springer Verlag Berlin Heidelberg New York, pp. 783–788, 2002.

[13] M. Kaneko, N. Kanayama and T. Tsuji, "Active Antenna for Contact Sensing," *IEEE Trans. On Robotics and Automation*, vol. 14, no.2, 1998.

[14] M. Lungarella, V. Hafner, R. Pfeifer and H. Yokoi, "An Artificial Whisker Sensor for Robotics," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2931–2936, 2002.

[15] P. H. Brownell, "Prey dection by the sand scorpion," *Scientific American*, vol. 251, p. 94, Dec 1984.

[16] R. Hergenröder and F. G. Barth, "Vibratory signals and spider behavior: How do the sensory inputs from the eight legs interact in orientation?," *Journal of Comparative Physiology.A*, Springer-Verlag, pp. 361–371, 1983.

[17] T. Chapman and B. H. Webb, "A neuromorphic hair sensor model of wind-mediated escape in the cricket," *International Journal of Neural Systems*, vol. 9, pp. 397–403, 1999.