Behavior Selection and Motion Modulation in Emotionally Grounded Architecture for QRIO SDR-4X II

Tsutomu SAWADA^{*1}, Tsuyoshi TAKAGI^{*2} and Masahiro FUJITA^{*1}

^{*1}Networked CE Development Laboratories Sony Corporation Shinagawa-ku Tokyo, Japan tsawada@pdp.crl.sony.co.jp *²Entertainment Robot Company Sony Corporation Minato-ku Tokyo, Japan

Abstract- This paper focuses on the design and evaluation of behavior module selection and motion modulation based on emotion in the Emotionally GrOunded (EGO) Architecture that is applied in autonomous robot QRIO SDR-4 X II. In the EGO architecture, a behavior module is selected based on homeostasis in order for the robot to regulate its internal state within a certain range. Each behavior module has a value called Activation Level (AL) that is composed of Motivation value (Mot) and Releasing value (Rel). Mot determines the degree that the instinct drives the behavior module. It is derived from internal state. Rel is the degree that reflects how much an external stimuli would satisfy an internal state as a result of the behavior. It is derived from the internal state and external stimuli. Each behavior module competes in AL to select a behavior module. Emotion is a key for motion modulation. It is derived from the internal state, its change, and the expected change of internal state associated with external stimuli in Long Term Memory. Through the implementation and experiments on QRIO SDR-4X II, we confirm the behavior selection and motion modulation processes in the EGO Architecture.

Keywords- Behavior Selection, Motion Modulation, Emotionally Grounded Architecture and QRIO SDR-4X II

I. INTRODUCTION

We have proposed Robot Entertainment as a new application of autonomous robots [1]. The aim of this proposition is to establish a new industry with autonomous robots and artificial intelligence. We earlier reported the development of a quadruped robot AIBO as a pet-type robot [2], a biped robot SDR-3X as a prototype for a motion entertainment robot that has capabilities of dancing performance etc. [3], SDR-4X as a prototype communication entertainment robot that has capabilities of interaction with a human and the environment [4], and a partner robot QRIO SDR-4X II (later referred to as just QRIO) as an extension of SDR-4X [5].

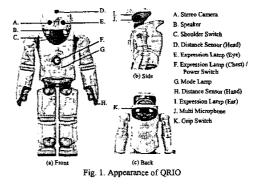
One of the significant technical achievements of QRIO is its integration of many advanced technical features, such as real-time motion control, face recognition, speech recognition, etc. A well-designed architecture is a key to integrating these technologies. Moreover, in general, the architecture is very important to develop numerous applications for motion and communication entertainment. A primary goal for QRIO is to be a partner robot. It is important that it behave spontaneously and emotionally.

From this viewpoint, we proposed the EGO (Emotionally GrOunded) architecture as a behavior and motion control architecture for autonomous robots [6]. The main strategy of EGO architecture is an ethological model [7]. Behavior control is based on homeostasis where the robot selects behaviors to regulate and maintain its internal state within a certain acceptable range. Behavior is a situationally-dependent motion sequence. Motion control is based on emotional modulation.

In this paper, details of the determination of a behavior module selection value based on homeostasis that considers not only the internal state of the robot but also an external stimuli, and also the emotional motion modulation process in the EGO architecture. Experimental results of our implementation on QRIO are then provided.

II. HARDWARE COMPONENT OF QRIO

Fig. 1 shows QRIO's appearance. It is 580 [mm] height, approximately 7 [kg] with battery and having 38 DOF. It is a stand-alone robot with three CPUs. The first is for audio recognition and text-to-speech synthesis. The second is for visual recognition, short- and long-term memory, and the behavior control architecture. The third is dedicated to motion control. Remote processing power and robot control is also available through wireless LAN.



III. SOFTWARE COMPONENT OF EGO ARCHITECTURE

In this section, the individual software components of the EGO architecture are briefly explained. Fig. 2 provides an overview. Please refer to [8] for more details on the EGO architecture.

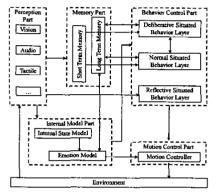


Fig. 2 Overview of the EGO architecture

A. Short Term Memory (STM)

STM integrates the results of perception. From audio perception, STM receives the result of not only speech recognition but also sound source direction by multi microphone localization. As for vision perception, STM can obtain the result of face recognition with its associated direction and distance computed from stereo vision. In the case that both, audio and visual directions, are same, STM merges the results to indicate that they are from the same user.

STM can also compute relative positions to detected objects (face and ball etc.) through kinematics. Therefore STM can store and recall results located outside of the limited view range.

B. Long Term Memory (LTM)

LTM associates the recognition results with an internal state. For example, LTM can associate an acquired name with an identified object or an identified voice, and change the internal state associated with a target object. Details of LTM are described in reference [9].

C. Internal State Model (ISM)

ISM maintains various internal state variables. It alters their values depending on the passage of time and incoming external stimuli. Basically, a behavior module is selected in order to keep these internal state variables within proper ranges. ISM is the core for generation of spontaneous behavior and response to external stimuli.

There is rank in ISM. The lower rank of an internal state variable is a real value that is grounded on a physical sensor. The higher rank of an internal state variable is a virtual value that is not grounded on a physical sensor. Others are a mixture of real and virtual values. As an example of a low-level internal state variable, NOURISHMENT is a real value grounded on the battery. A high-level internal state variable, SLEEP is a virtual value that represents a sleep-promoting substance. As a mid-level internal state variable, COMFORT is a virtual value ground on tactile sensors. It increases when QRIO's head is stroked and decreases on the passage of time.

D. Emotion Model (EM)

EM has 6+1 emotions, which are ANGER, DISGUST, FEAR, JOY, SADNESS, SURPRISE, and NEUTRAL, based on Ekmann's proposal [10]. Each emotion has an associated value. They are determined based on selfpreservation. The determination of self-preservation is composed of self-crisis and self-crisis expectation. The value of self-crisis is evaluated from external stimuli. Detail of this evaluation described in reference [9].

E. Situated Behavior Layer (SBL)

SBL controls behavior modules. Each behavior module has two basic functions, monitor and action.

Monitor function periodically and concurrently creates a value, which is called *Activation Level (AL)*, using internal state variables and external stimuli. It indicates how relevant the behavior is for the situation (e.g., observing an object and a sound event etc.). The details of this computation are described below.

A behavior module is selected by competition using AL. For example, greedy, that is maximum AL is selected, or soft max, that is larger AL is selected with larger probability, is used as a selection policy. Then the selected behavior module is given execution permission.

Availability of necessary resources for execution, e.g., head, arm, speaker, etc., are also considered during the competition. In the case where there is no resource conflict among behavior modules, all of them are given execution permission and then execute concurrently.

Availability of necessary resources for execution, e.g., head, arm, speaker, etc., are also considered in the competition. In the case where there is no resource conflict among behavior modules, all of them are given execution permission and then execute concurrently.

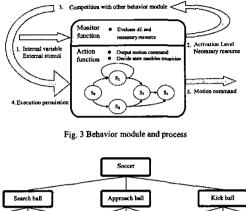
After a behavior module is given permission, the action function executes the behavior implemented as a state machine. Each node can output e.g. a motion command (designed motion command, walk command, and tracking command etc.) and can decide to state transition depending upon the given situation.

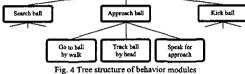
Fig. 3 shows a behavior module and associated process.

A tree structure is used to organize the behavior modules. An abstract behavior can be divided into concrete and multiple sub-behaviors. For example, as shown in Fig. 3, "Soccer" can be decomposed into "Search ball", "Approach ball" and "Kick ball", also "Approach ball" can be decomposed into "Go to ball by walk", "Track ball by head", and "Speak for approach" etc.

In the parent behavior module in the tree structure, a monitor function can also determine the AL through the child ALs instead of evaluation through the internal state

variables and external stimuli. The action function can also select a child behavior module instead of a motion command





SBL is organized in 3 modules, D-SBL (Deliberative SBL), N-SBL (Normal SBL) and R-SBL (Reflexive SBL). D-SBL realizes behavior control for deliberative behavior such as dialogue, N-SBL realizes behavior control for homeostatic behavior such as battery charge, and R-SBL realizes behavior control for quick responses like startle. The details of the SBL implementation are described in [11].

IV. EVALUATION OF ACTIVATION LEVEL FOR BEHAVIOR MODULE SELECTION

In this section, we focus on the evaluation of AL to realize homeostatic behavior in N-SBL. AL is composed of motivation value (*Mot*) and releasing value (*Rel*).

The evaluation of *Mot*, *Rel* and *AL* is described using the following example behavior "Approach a target object for eating it". EGO architecture is based on ethological study. The example is for an agent to regulate the NOURISHMENT state variable. From the viewpoint of robotics, NOURISHMENT is interpreted as charge of battery, eating as pseudo-eating, that is charging-battery, and object as battery station.

A. Evaluation of Motivation value

The motivation value is the degree to which the instinct drives the behavior module. It is derived from internal state variables and is composed of instinct values.

An instinct value (Ins[i]) is designed for each specific internal state variable (Int[i]). Two examples for NOURISHMENT and FATIGUE are shown in Fig. 5 and can be interpreted as follows. The less nourishment there is, the larger is the instinct to eat it. Also, in the case of large nourishment, this instinct turns negative to realize a moderation or reduction in eating behavior. Fatigue has a negative effect. The more fatigue there is, the less the value of the instinct associated with it.

$$Mot = \sum W_{Mot}[i] \cdot Ins[i] \cdot$$
(1)

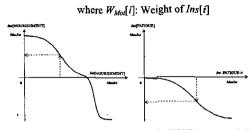


Fig. 5 Ins[i] versus Inf[i]in the behavior modules NOURISHMENT (left) and FATIGUE (right).

B. Evaluation of Releasing value

The releasing value is the degree regarding how much an external stimuli would satisfy an internal state as a result of the behavior. It is derived from an internal state variable and the external stimuli. It is composed of a satisfaction value and the expectation of satisfaction value.

A satisfaction value (*Sat*[*i*]) is designed for each specific internal state valuable. Examples for NOURISHMENT and FATIGUE are shown in Fig. 6.

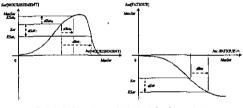


Fig. 6. Sat[i] against Int[i] in the behavior module

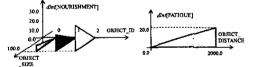


Fig. 7. Database about expectation of change in the internal state variable

To evaluate the expectation of satisfaction value (ESat[i]), the behavior module maintains a database on expectation of change in the internal state variable (dInt[i]) against the result of the behavior for the given external stimuli.

Fig. 7 is an example where the behavior module expects a change in NOURISHMENT and FATIGUE when an external stimuli (OBJECT_ID, OBJECT_SIZE, and OBJECT_DISTANCE) is obtained. It means that when an object is found which has OBJECT_ID = 1,

OBJECT_SIZE = 100, and OBJECT_DISTANCE = 200, NOURISHMENT would increase 20 and FATIGUE would increase 20 after approaching and eating the object.

ESat[i] and expectation of change in satisfaction value (dSat[i]) are shown in Fig. 6. They are interpreted as follows. When $dInt_0$ is determined by observing an object₀, the dSat[NOURISHMENT] is expected as positive. On the contrary, when $dInt_1$ is determined when observing another object₁, for example whose size is larger than object₀, the dSat[NOURISHMENT] is expected as negative due to overeating. dInt for fatigue is related to the distance of an observed object. The farther the distance is, the more dissatisfaction the agent receives.

Rel is evaluated by (2).

 $Rel = \sum W_{Rel}[i] \cdot (W_{dSar} dSat[i] + (1 - W_{dSar}) ESat[i]) \cdot$ (2)

where $W_{Rel}[i]$: Weight of $(W_{dSat}dSat[i]+(1-W_{dSat})ESat[i])$ W_{dSat} : Weight of dSat[i] against ESat[i]

C. Evaluation of Activation Level

AL is evaluated from Mot and Rel by (3).

$$4L = W_{Mot}Mot + (1 - W_{Mot})Rel$$
(3)

where W_{Mot}: Weight of Mot against Rel

Note that when there is no external stimuli for the behavior module, AL is set to 0, so that behavior module is never selected.

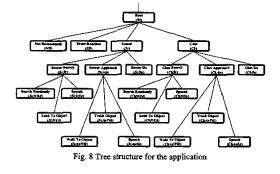
V. MOTION MODULATION BY EMOTION

LTM associates change in internal state with an observing object. EM controls emotions based on an internal state variable and the associated change in the internal state variable.

MC modulates motion with the current emotion by changing actuator speed, joint angle gain, posture, designed motion selection and LED color, etc.

VI. IMPLEMENTATION AND EXPERIMENTAL RESULTS

An application was implemented for the evaluation of value for behavior module selection and motion modulation by emotion as described previously, and experiments were conducted on QRIO. The tree structure of behavior modules for the application is shown in Fig. 8. In the following subsections, their details are described.



ł

A. Evaluation of Value for Behavior Module

In Fig. 8, Soccer (Sc) sub tree has 3 child behavior modules Soccer Search (ScSr), Soccer Approach (ScAp) and Soccer Do (ScDo). They evaluate AL based on an internal state variable and external stimuli. AL of Sc is the maximum AL among its children.

ScAp focuses on VITALITY and FATIGUE as internal state variables, and BALL_ID and BALL_DISTANCE as external stimuli.

AL is composed of Mot and Rel with $W_{Mot} = 0.4$.

Mot is composed of Ins[VITALITY] and Ins[FATIGUE], which are shown in Fig. 9 (a) and (b), with W_{Mod} [VITALITY] = 0.8 and W_{Mod} [FATIGUE] = 0.2.

Rel is composed of *dSat*[VITALITY], *dSat*[FATIGUE], *ESat*[VITALITY] and *ESat*[FATIGUE], which are shown in Fig. 9 (d) and (e), with W_{Rel} [VITALITY] = 0.8, W_{Rel} [FATIGUE] = 0.2 and W_{dSat} = 0.0.

Each *dint*[VITALITY], *dint*[FATIGUE] is estimated from BALL_ID and BALL_DISTANCE as shown in Fig. 10 (a) and (b).

ScSr focuses only on VITALITY as an internal state variable. It does not focus on external stimuli. Evaluation of AL is the same as for ScAp except for values whose index is FATIGUE. They are set to 0.

ScDo focuses only on VITALITY as an internal state variable and BALL_ID as an external stimuli. Evaluation of AL is same as ScAp except for values whose index is FATIGUE. They are set to 0. On the condition that the ball distance is not in the proper range for kick motion (0 - 400 [mm]), AL = 0. Note the distance is not used to evaluate Rel.

ScDo outputs kick motion commands in the action function. Motion commands for search and approach a ball are output in children behavior modules of *ScSr* and *ScAp*.

Chat (Ch) sub tree is composed of Chat Search (ChSr), Chat Approach (ChAp) and Chat Do (ChDo) and has the same structure as Sc except for internal state variable and external stimuli. INTERACTION and FACE_ID is specified instead of VITALITY and BALL ID.

Ins[INTERACTION], Sat[INTERACTION] and dInt[INTERACTION] are shown in Fig. 9 (c), (f) and Fig. 10 (c) respectively.

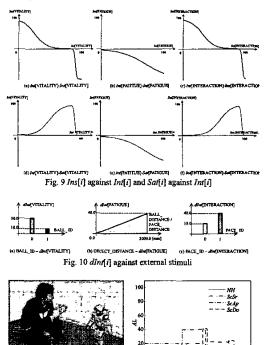
On the condition that the distance to the detected face is not in the proper range for interaction (100 - 500 [mm]), AL = 0.

Not Homeostasis (NH) is not a homeostasis behavior module, so AL = 10 constantly. It outputs an idle motion command like cock the head to one side, and tracking a face, etc. When AL of all homeostatic behavior modules are low (all internal states are satisfied), NH is executed.

Event Reaction (ER) does not output any motion command. When a reflexive event such as a clap sound comes, it keeps the physical resource by setting AL = 100to prevent a homeostatic behavior module from executing. A behavior module in R-SBL reacts to the event. It outputs a head command to look towards the clap sound source direction.

A parent behavior module selects its child behavior modules using a greedy policy based on the children's AL.

Fig. 11 (a) shows the appearance of the experiment.



20

20 40

40 60 time ([s]

> NH ScSr Sr

me / [s]

(d) AL in Experiment 3

(b) AL in Experiment 1

80

ល

Fig. 11 (b) shows the experimental result of change in AL that Int[VITALITY] = 20, Int[FATIGUE] = 10, $Int[INTERACTION] \approx 80$ at t = 0. In this condition, AL of ChSr, ChAp, and ChDo are negative at all times.

. Fig. 11 Appearance of Experiment and results of AL

- EH ···ChSr ···ChAp - ChDo

(a) Appearance of Experiment

20 30 time / [s]

(c) AL in Experiment 2

10

At first QRIO searches randomly for a ball in AL[ScSr]= 20, QRIO finds the ball with BALL_ID = 0 (red ball) at t = 36, then AL[ScAp] increases to 40 because of dSat[VITALITY] for the ball and the robot starts to approach the ball. QRIO reaches the distance where kick motion is effective at t = 55, and then AL[ScDo] increases to 42 and kicks the ball. After kicking the ball, Int[VITALITY] is satisfied to a level of 50 at t = 57. Ins[VITALITY] is still large in this condition. Then QRIO approaches and kicks the ball again (t = 59 - 78, t = 79 - 84 for each). Finally Int[VITALITY] is fully satisfied (80), and NH is executed after t = 85.

Fig. 11 (c) and (d) show the experimental results of AL for Int[VITALITY] = 20, Int[FATIGUE] = 10, Int[INTERACTION] = 20 at t = 0.

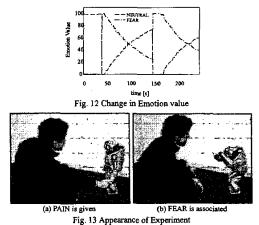
In experiment 2, QRIO detects a clap sound at t = 16during approaching the ball in AL[ScAp] = 40. Then ScApis interrupted and ER is executed in AL[ER] = 100 at t = 16- 27. The behavior module in R-SBL outputs a motion command to turn toward the sound source direction. At t =25 QRIO finds a face whose FACE_ID = 0 (person A), then AL[ChAp] increases to 27. Because AL[ScAp] is still larger than AL[ChAp], QRIO ignores the face and resumes the approach from t = 28 and kicks the ball at t = 44.

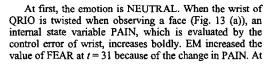
On the contrary in experiment 3, QRIO detects a clap sound at t = 17, and ScAp is interrupted by ER. At t = 32QRIO finds a face whose FACE_ID = 1 (person B), then AL[ChAp] increases to 45 which is larger than AL[ScAp]. QRIO now approaches the face, suspending its previous approach to the ball and chats with him at t = 48. After the chat, Int[INTERACTION] is satisfied (60), and AL of ChSr, ChAp, and ChDo turn negative. QRIO now looks toward the ball, and then approaches and kicks it at t = 81.

B. Motion Modulation by Emotion

In the condition that all internal state variables are satisfied, NH is executed. Its idle motion command is modulated by emotion.

Fig. 12 shows an experimental result of the specific changes in the emotion values NEUTRAL and FEAR. Emotion of QRIO is represented as the one having maximum value.





this time LTM learns a connection between the face whose $FACE_{ID} = 0$ and the associated change in PAIN.

After that, the value of FEAR decreases and the value of NEUTRAL increases depending on the passage of time. Eventually, the value of NEUTRAL turns larger than FEAR at t = 94 and QRIO takes self-possession back.

Afterwards, when QRIO again observes the same face, LTM associates a change in PAIN, and EM increases the value of FEAR without getting real PAIN at t = 145. Then an idle motion command of NH is modulated in MC by changing the expressed motion command from a cock head to one side command to the fear expression command (Fig. 13 (b)).

VII. SUMMARY AND DISCUSSION

In this paper, we focus on the determination of the value for behavior module selection and motion modulation by emotion in the EGO architecture for autonomous robots. Through implementation and experiments on QRIO, we confirm our approach.

In behavior selection, AL is evaluated using the internal state variable and external stimuli, and dInt is estimated from external stimuli from a database. In the present state, dInt in database is designed heuristically and not actually grounded. We consider, however, that the problem would be clarified by employing learning. After a behavior module obtains an external stimuli and is executed, internal state variables really change. By feedback the change in internal state variable for the external stimuli and learning it in the database, the estimation of dInt grounds on the robot's real internal state. We will implement and experiment the learning architecture as future work.

There are some existing approaches that similar to EGO architecture.

Regarding to behavior selection, our point is to define the AL based on the expected change of internal state variables through external and internal factors.

In Blumberg's approach [12], the value of behavior, which corresponds to our AL, is also computed from the combination of external factors and internal factors. However, the external factors are directly computed from the defined relevance of the stimuli in the situation.

Kismet [13] uses "affective tag" in its emotion systems. The affective tag is used to determine "emotional response", not to select goal-oriented behavior itself. On the other hand, in the EGO architecture, association with external stimulus is used for AL as well as emotion variables.

In WAMOEBA's approach [14], they define some internal variables for homeostasis regulation or self-preservation function. The mechanism causes to elicit four basic pseudo-hormones, which determine "performance" of motor and sensor devices. WAMOEBA's goal is to self-organize its behaviors without explicit descriptions by a designer. The behaviors look very simples such as avoidance, approach, and tracking.

Regarding to emotional expression, it is important to evaluate the system how it is effective. Some research works evaluate it subjectively by SD method using questionnaire [15]. Not only the viewpoint of system but also appearance and degree of freedom, QRIO with EGO architecture would be the most humanlike robot and potential of physical expression would be large. The evaluation will be described in next paper.

ACKNOWLEDGMENT

We greatly appreciate Dr. Ronald C. Arkin in Georgia Institute of Technology, and all researchers and engineers for QRIO in Sony Corporation.

REFERENCES

- Fujita, M. and Kageyama, K., "An open architecture for robot entertainment", International Conference on Autonomous Agents, pp. 435-440, 1997.
- [2] Fujita, M., and Kitano, H., "Development of an quadruped robot for robot entertainment", Autonomous Robots, Vol. 5, pp.7-18, Kluwer Academic, 1998.
- [3] Ishida, T., Kuroki, Y., Yamaguchi, J., Fujita, M., Doi, T.T., "Motion entertainment by a small humanoid robot Based on OPEN-R", IEEE/RSJ International Conference on Intelligent Robots and Systems, 2001, pp. 1079-1086.
- [4] Fujita M. et. al., "A small humanoid robot SDR-4X for entertainment applications", International Conference on Advanced Intelligent Mechatronics, 2003, pp.938-943.
- [5] M. Fujita, K. Sabe, Y. Kuroki, T. Ishida, and T.Doi, "SDR-4X II: A small humanoid as an entertainer in home environment", International Symposium of Robotics Research, Siena, Italy, Oct. 2003.
- [6] M. Fujita and et al., "An autonomous robot that eats information via interaction with human and environment," IEEE International Workshop on Robot-Human Interactive Communication, Bordeaux and Paris, France, pp. 383–389, Sept. 2001.
- [7] R. Arkin, M. Fujita, T. Takagi, and R. Hasegawa, "Ethological and emotional basis for human-robot interaction," Robotics and Autonomous System, vol. 42, 2003, pp.191-201.
- [8] M. Fujita, Y. Kuroki, T. Ishida, and T.Doi, "Autonomous behavior control architecture of entertainment humanoid robot SDR-4X" IEEE/RSJ International Conference on Intelligent Robots and Systems, LA, USA, 2003, pp.960-967.
- [9] F. Tanaka, K. Noda, T. Sawada and M. Fujita, "Associated Emotion and its Expression in an Entertainment Robot QRIO", IFIP International Conference Entertainment Computing, Eindhoven, NZL, Sep. 2004, in press.
- [10] Ekman, P., and Davidson, R. J., "The nature of emotion", Oxford University Press, 1994.
- [11] Y. Hoshino, T. Takagi, and M. Fujita, "Behavior description and control using behavior module for personal robot" IEEE International Conference on Robotics and Automation, Luisiana, USA, Apr. 2004, pp. 4165-4171.
- [12] B. M. Blumberg, "Old Tricks, New Dogs: Ethology and Interactive Creatures", PhD thesis, MIT, 1997
- [13] C. L. Breazeal, "Designing Sociable Robots", MIT Press, 2000
- [14] T. Ogata and S. Sugano, "Emotional Communication between Humans and the Autonomous Robot WAMOEBA-2 (Waseda Amoeba) Which has the Emotion Model", JSME Int. Journal, Series C: Mechanical Systems Machine Elements and Manufacturing, Vol.43, No.3, pp. 568-574, 2000.
- [15] T. Kanda, H. Ishiguro, T. Ono, M. Imai and Ryohei Nakatsu, "Development and Evaluation of an Interactive Humanoid Robot "Robovi"", IEEE Int. Conf. on Robotics and Automation, Washington D.C., USA, May 2002, pp.1848-1855