Implementation of Bio-Inspired Vestibulo-Ocular Reflex in a Quadrupedal Robot

Ravi Kaushik, Marek Marcinkiewicz, Jizhong Xiao, Simon Parsons, and Theodore Raphan,
Member, IEEE

***Abstract*— Studies of primate locomotion have shown that the head and eyes are stabilized in space through the vestibulo-collic and vestibulo ocular Reflexes (VCR, VOR). The VOR is a reflex eye movement control system that stabilizes the image on the retina of the eye during head movements in space. This stabilization helps maintain objects of interest approximately fixed on the retina during locomotion. In this paper we present the design and implementation of an artificial vestibular system, which drives a fully articulated binocular vision system for quadrupedal robots to maintain accurate gaze. The complete robot head has 9 Degrees of freedom (DOF): pitch, yaw, and roll for the head and 3 DOF for left and right cameras. The SONY AIBO® Quadruped robot has been modified with additional hardware to emulate the vestibular system and the vestibulo-ocular reflex in primates**.

I. INTRODUCTION

Robot locomotion has been studied using a wide range of wheeled and legged robots [1]. Although wheeled robots move quickly, they can only move on smooth terrain and lack the versatility of legged robots in handling rough terrain. As a result, there has been a concerted effort within the robot community to understand the motion of legged robots. The motion of the biped and quadruped robots causes the head and the cameras mounted on them to pitch, yaw, and roll and linearly accelerate in three dimensions. The very fact that legged motion generates this kind of disturbance makes it difficult to keep the visual frame stable. Humanoid Robots and Quadruped robots are in need of a strong sense of awareness in terms of its position and its movements in space.

One approach to solving this problem is to develop an artificial vestibular system [2] and implement gaze stabilization techniques on robots that have a binocular camera system mounted on their heads and are capable of individual or common rotations [3-8]. To date, these implementations have been restricted to two dimensions. In [3] and [4] the design and development of an artificial vestibular system that integrates 3 uni-axial accelerometers and 3 uni-axial gyroscopes in order to obtain a six-axial sensory system is described. This vestibular system is implemented in a humanoid robot that has 7DOF from the neck up. The neck has 4DOF and the eye (Stereoscopic Vision System) has 3DOF. It consists of a common pitch for two eyes and two independent yaws for each camera. To represent the human eye motion, however, it is necessary to perform yaw, pitch and roll motions that consider the non-commutative property of the eye rotations as well as vergence eye movements for viewing close targets [2, 10]. That is, two 3D rotations in different order will result in two different orientations of the eye. Another approach [8] presented a design and development of a High Speed Binocular Camera Head. The two cameras mounted on the head were capable of Pan/Tilt motions. However, roll is not incorporated, and does not consider models of 3D rotations of eyes in space [2].

Recently, we have shown that an artificial system that senses head motion, i.e., an artificial vestibular system, can on average successfully stabilize the head oscillation during quadrupedal walking (AIBO, Sony) [9]. In the present work, we have designed and implemented a fully articulated binocular vision system embedded in the head of a quadrupedal robot and realized the Vestibulo Ocular Reflex (VOR) that compensates for head perturbations in three dimensions. This system allows us to stabilize gaze of the robot during locomotion. This system is unique in that each camera has three separate DOF: yaw, pitch, and roll. Each camera has been embedded in a gimbal system that can be then controlled, as is the eye orientation in three dimensions [10]. This would be important for stabilizing gaze of robots as they move about such as in RoboCup [9].

Manuscript received September 15, 2006. This work was supported in part by the National Science Foundation MRI Grant CNS-0520989NSF; National Institutes of Health under Grants DC05222, DC05204, EY04148, and a CUNY Collaborative Grant 80209-09.

R. K. Kaushik is with the Dept. of Computer Science, The Graduate Center, The City University of New York, 365 Fifth Avenue, NY, NY 10016. USA (e-mail: ravikaushik@sci.brooklyn.cuny.edu)

M. Marcinkiewicz is with the Dept. of Computer Science, The Graduate Center, The City University of New York, 365 Fifth Avenue, NY, NY 10016. USA (e-mail: marek@nsi.brooklyn.cuny.edu)

J. Xiao is with the Dept. of Electrical Engineering, The City College of the City University of New York, 138th Street and Convent Avenue, NY, NY, 10031 and Dept. of Computer Science, The Graduate Center, The City University of New York, 365 Fifth Avenue, NY, NY 10016. USA (e-mail: jxiao@ccny.cuny.edu).

S. Parsons is with the Dept. of Computer and Information Science Brooklyn College of the City University of New York, 2900 Bedford Avenue, Brooklyn, NY 11210 and Dept. of Computer Science, The Graduate Center, The City University of New York, 365 Fifth Avenue, NY, NY 10016. USA (email: parsons@sci.brooklyn.cuny.edu)

T. Raphan is with the Dept. of Computer and Information Science Brooklyn College of the City University of New York, 2900 Bedford Avenue, Brooklyn, NY 11210 and Dept. of Computer Science, The Graduate Center, The City University of New York, 365 Fifth Avenue, NY, NY 10016. USA (phone: 718-376-0503, fax: 718-951-4489 email: raphan@nsi.brooklyn.cuny.edu)

II. PRIMATE VESTIBULAR AND OCULOMOTOR SYSTEMS

In primates, the vestibular system is a biological acceleration sensor, which is embedded in the inner ears on both sides of the head, providing information about head movement in space [2]. This information is then utilized to stabilize gaze and orientation during locomotion in primates [11, 12].

A. Primate Vestibular System: The primate vestibular system consists of two organs namely semi circular canals (anterior, posterior and horizontal) and otoliths (utricle and saccule) (Fig. 1A).

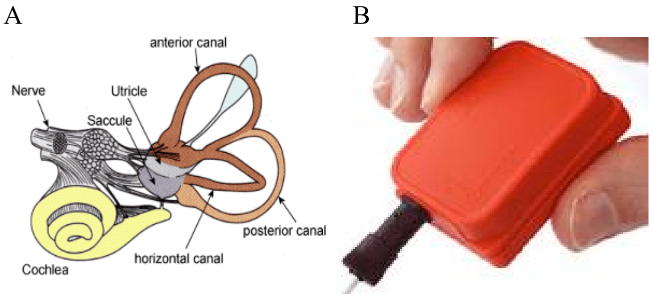

Fig. 1. (A) Human Vestibular System: The three semi-circular canals - anterior, posterior and lateral (horizontal) canals and the otoliths (utricle and saccule). (B) Artificial Vestibular System: MTx acceleration sensor. Outputs are calibrated 3D rate of turn, and 3D linear acceleration.

The semicircular canals sense angular acceleration through a set of approximately orthogonal fluid-filled canals (Fig. 1A) (endolymph), which are blocked by a membrane called the cupula. When the head rotates, the fluid pushes against the cupula activating hair cells that transmit the signal about head movement [2]. The viscosity of the fluid and the elastic properties of the cupula determine the dynamics of how the head movement is transduced by this canal system. The inertial properties of the fluid are relatively minor. The otoliths, which sense linear acceleration, accomplish this by activating hair cells that are embedded in membrane containing crystals.

The reflexes based on these sensors have been classified into an angular VOR (aVOR) and linear VOR (lVOR). The aVOR (rotation) generates eye movements incrementally for changes in head rotation to maintain stable gaze. The lVOR (translation) rotates the eyes to maintain fixation on a particular point in space. The lVOR also codes orientation of the eyes in space.

B. *Primate Oculomotor System:* Three pair of muscles, the medial and lateral rectus, superior and inferior rectus and the superior and inferior oblique control the eye movements (Fig. 2A). It has been shown that saccadic eye movements obey Listing's law and are controlled by two-dimensional signals, which are confined to the pitch-yaw axes of the head [10]. The aVOR does not obey Listing's law, and compensation occurs close to the axes of the head rotation.

III. DESIGN OF ARTIFICIAL VESTIBULAR SYSTEM AND SENSING OF HEAD MOTION

The *MTx Sensor* by XSENS® Motion Technologies is a complete inertial measurement unit capable of providing 3D linear acceleration, 3D rate of turn and 3D magnetic field data (Fig. 1B). Static accuracy for Roll/Pitch is $<0.5^0$ and for heading (yaw) is $<1.0^0$. Dynamic accuracy is 2^0 RMS and angular resolution is 0.05^0. The angles are in 3D and updated at 120Hz.

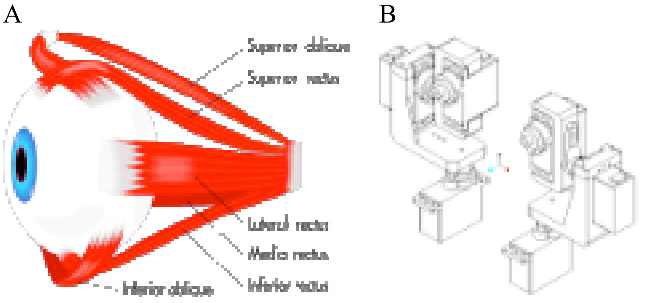

Fig. 2. (A) Eye with three pairs of Ocular Muscles: superior rectus, inferior rectus, superior oblique, inferior oblique, medial rectus, and lateral rectus. (B) Fick gimbal system for controlling the movements of the cameras in yaw, pitch and roll.

The orientation coordinate system of MTx sensor has a fixed coordinate frame with its X-axis pointing to the local earth's magnetic north (Field). Y-axis follows the right handed coordinate system pointing to west and Z-axis completing the vertical axis pointing up as shown in Fig. 3. The sensor fixed coordinate system is a right-handed coordinate system. The output is calibrated with sensor frame relative to the earth fixed frame. The output is drift compensated and runs on a proprietary sensor fusion algorithm developed by XSENS and can calculate absolute orientation in 3D space from miniature rate of turn sensors, accelerometers and magnetometers built into one unit.

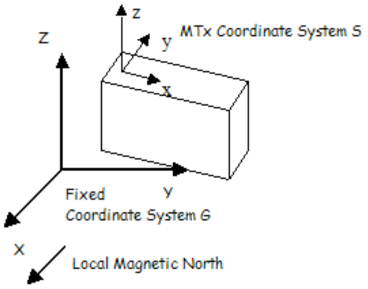

Fig. 3. Earth-Fixed Coordinate frame and Acceleration Sensor Coordinate frame

IV. DESIGN OF ARTIFICIAL OCULOMOTOR SYSTEM

A stereovision camera system mounted with three servos on each side forms the Ocular Servo Module (Fig. 2B, 4). The three servos are assembled as to form an Euler type gimbal system, following a Fick convention. The rotations are measured in terms of Yaw, Pitch and Roll. The Servos are from Hitec® and accepts pulse width modulated (PWM) inputs. The PWM signal range can be varied from 0.9ms to 2.1ms to position the servo shaft from -90^0 to $+90^0$ with 1.5ms PWM signal indicating the zero degree position. The camera is mounted on the Servo that provides roll motion.

A pair of prototype cameras from X10® is mounted on the Ocular Servo Module, one of which is shown in Fig. 4. The resolution is 640 × 480 pixels and is transmitted wirelessly to the receiver on the host computer. The transmitter and camera are powered by 12VDC from a DC-DC converter used to step up 5VDC to 12VDC from the NiMH battery.

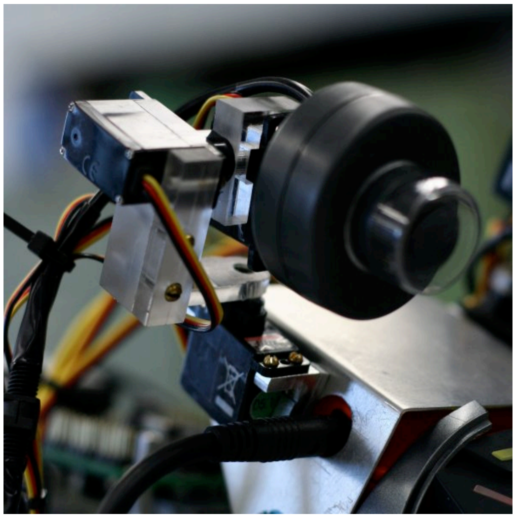

Fig. 4. The three-dimensional gimbal system and camera mounted on the robot head.

V. SERVO CONTROLLER

A servo controller SSC-32 from Lynxmotion® technologies is used to provide PWM inputs to the servos mounted on the AIBO head. The PWM signals have a resolution of 1μS for accurate positioning and extremely smooth moves. The range extends from 0.5ms to 2.50ms for a range of 180° rotation by the servomotor. It provides a DB9 input and supports RS-232 for serial communication. The baud rate is set at 115.2kbps on the Servocontroller. It accepts a set of commands to set the pulse width of a channel. The host computer communicates with a pair of bluetooth transceivers to the servo controller. The angular position derived from the MTx acceleration sensor output is processed and converted into appropriate hexadecimal code to generate the precise PWM signals. This PWM signal positions the servos in real time. The power is supplied onboard the AIBO® using NiMH rechargeable batteries rated 5V, 720mAH.

Fig. 6A shows the unencumbered AIBO and Fig. 6B shows the AIBO with servo controller, MTX sensor, and gimbal system for positioning the cameras mounted on the autonomous robot.

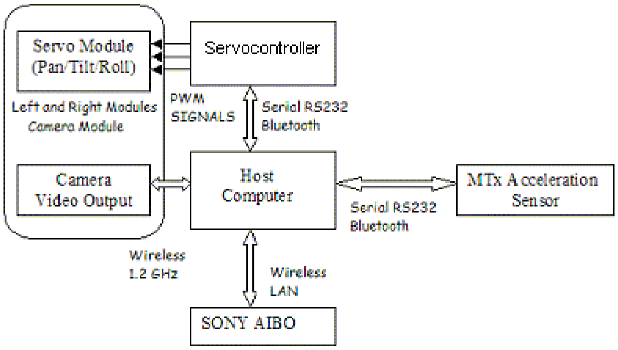

Fig. 5. Block Diagram. Ocular Servo Module receives input from Servocontroller SSC-32 (Lynxmotion® Technologies) and Host Computer. Communication channels with various peripheral modules are shown.

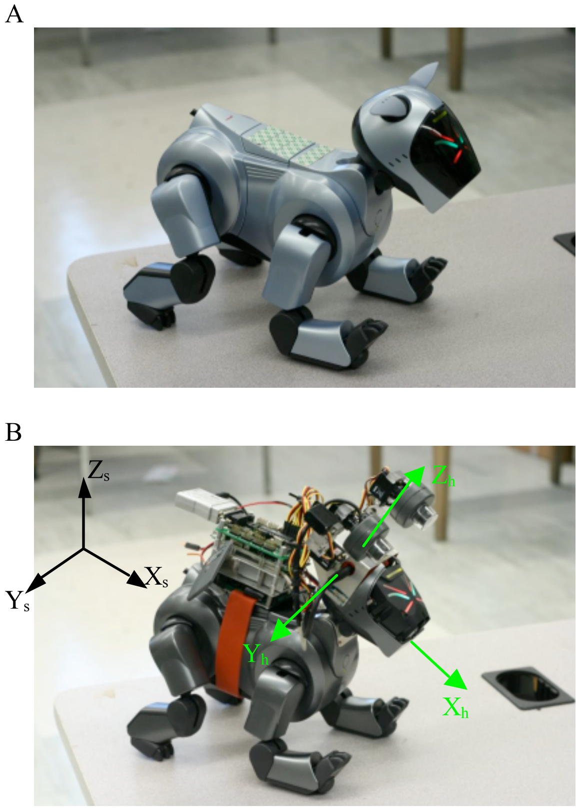

Fig. 6. (A) An unencumbered Sony AIBO robot. It has 3DOF of the head and is capable of communicating over a wireless interface with a computer. (B) The modified Sony AIBO robot. An X10® camera system mounted within a Fick gimbal system that is controlled by Servocontroller mounted on the back of the robot.

The servo controller, the bluetooth transceiver connected to it, two wireless video transmitters, a bluetooth transceiver for the acceleration sensor and two 4 cell battery packs powering all the units are stacked at the back of the AIBO (Fig. 7)

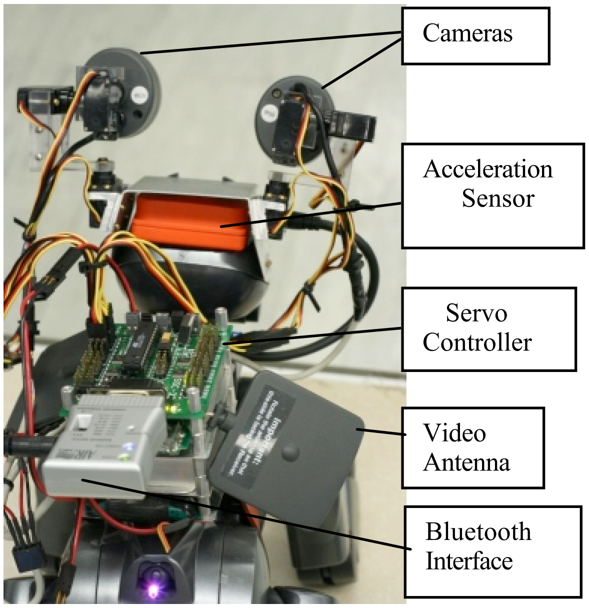

Fig. 7. View from the back showing the cameras, acceleration sensor, servo controller, bluetooth interface, and video antenna.

The MTx acceleration sensor provides low-level RS232 communication. The user can set the configuration settings like sample frequency, input output synchronization, baud rate and data output modes. The binder connector of the acceleration sensor is connected to an AIRCABLE® serial bluetooth transceiver. The baud rate and other parameters can be set to provide efficient transmission between the host computer and the acceleration sensor. Another bluetooth transceiver is connected to the servo controller to communicate with the host computer. The video cameras transmit the streaming images using the third party X10® transceiver at 2.4Ghz.

The Sony AIBO ERS-210 quadrupedal robot was a suitable platform for implementing the VOR system, since the ERS-210 has 3DOF in the head making it possible to fully compensate for head motion during quadrupedal locomotion. The control of the robot has also been the object of much study in its relationship to RoboCup [9]. The robot has an onboard MIPS processor and is able to communicate with a host computer through an 802.11b wireless TCP/IP network.

VI. IMPLEMENTATION OF COMPENSATORY OCULAR MECHANISMS

The head moved in space relative to a coordinate frame defined by $X_hY_hZ_h$ (Fig. 6B). The naso-occipital axis (X_h), the interaural axis (Y_h) and the axis out of the top of the head (Z_h) formed a right-handed coordinate frame for the head. The acceleration sensor was embedded in the head approximating the location of the vestibular system in primates and was aligned with the head coordinate frame (Fig. 7). The Ocular Servo Module (OSM) was built as a gimbal system according to a Fick convention. The movements of the cameras were programmed to rotate in a compensatory fashion relative to the head movement based on the information obtained from the acceleration sensor. It therefore mimicked the high frequency angular VOR of the primate [2] and was implemented as follows [10]:

Rotations of the robot head can be given as a sequence of rotation matrices given by:

$$R_{xyz} = R_x(\psi)\cdot R_y(\theta)\cdot R_z(\phi) \tag{1}$$

where

$$R_x(\psi)=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\psi & -\sin\psi\\ 0 & \sin\psi & \cos\psi\end{bmatrix}\quad R_y(\theta)=\begin{bmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{bmatrix}\quad R_z(\phi)=\begin{bmatrix}\cos\phi & -\sin\phi & 0\\ \sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix} \tag{2}$$

Substituting Eqn. 2 into Eqn. 1, we obtain the rotation matrix for the eye relative to the head as:

$$R_{xyz}=\begin{bmatrix}\cos\phi\cdot\cos\theta & \cos\phi\cdot\sin\theta\cdot\sin\psi-\sin\phi\cdot\cos\psi & \cos\phi\cdot\sin\theta\cdot\cos\psi+\sin\phi\cdot\sin\psi\\ \sin\phi\cdot\cos\theta & \sin\phi\cdot\sin\theta\cdot\sin\psi+\cos\phi\cdot\cos\psi & \sin\phi\cdot\sin\theta\cdot\cos\psi-\cos\phi\cdot\sin\psi\\ -\sin\theta & \cos\theta\cdot\sin\psi & \cos\theta\cdot\cos\psi\end{bmatrix} \tag{3}$$

The incremental change in head rotation was obtained from the acceleration sensor as the angular velocity vector represented by ω, in rad/sec. The incremental axis of rotation of the head at any given time t is therefore given by

$$\hat{n}=\frac{\omega}{\|\omega\|} \tag{4}$$

where $\hat{n}$ is a unit vector in the direction of the incremental rotation.

Φ_{Inc} is the incremental angle of rotation about the axis of rotation and is given by:

$$\Phi_{Inc}=\|\omega\|\cdot\Delta t \tag{5}$$

where Δt is the time (10mS) between the two sensor readings.

The resulting incremental axis of rotation and incremental angle of rotation is fed into the Rodrigues Formula (An efficient method to compute rotations).

$$R_{Inc}=\begin{bmatrix}\cos\Phi+n_1^2\cdot(1-\cos\Phi) & n_1\cdot n_2\cdot(1-\cos\Phi)-n_3\cdot\sin\Phi & n_1\cdot n_3\cdot(1-\cos\Phi)+n_2\cdot\sin\Phi\\ n_2\cdot n_1\cdot(1-\cos\Phi)+n_3\cdot\sin\Phi & \cos\Phi+n_2^2\cdot(1-\cos\Phi) & n_2\cdot n_3(1-\cos\Phi)-n_1\cdot\sin\Phi\\ n_3\cdot n_1\cdot(1-\cos\Phi)-n_2\cdot\sin\Phi & n_3\cdot n_2\cdot(1-\cos\Phi)+n_1\cdot\sin\Phi & \cos\Phi+n_3^2(1-\cos\Phi)\end{bmatrix} \tag{6}$$

Eqn. 6 provides the incremental rotational matrix representing the change in angles of the head in 3 dimensions. The new position of the servomotors at any instant of time is obtained by determining the incremental changes for the Euler angles, each of which is associated with a particular motor. The incremental Euler angles are obtained from Eqn. (6) as:

$$\psi_{new} = \psi_{old} + \tan^{-1}\frac{r_{23}}{r_{33}} \qquad (7)$$

$$\theta_{new} = \theta_{old} + \sin^{-1} r_{13}$$

$$\phi_{new} = \phi_{old} + \tan^{-1}\frac{r_{12}}{r_{11}}$$

where the parameters r_{11}, r_{12}, r_{13}, r_{23}, r_{33} are the associated matrix elements for a particular head orientation in Eqn. (3). The command to rotate the servomotors is converted PWM signals that orient the cameras. The necessary conversion is made from Euler angles obtained in degrees to PWM signals (ranging from 0.5 ms to 2.5 ms representing -90° to +90°).

VII. EXPERIMENTAL RESULTS

To test the "head-eye" compensation of the simulated aVOR, we executed a software program that results in an oscillating head movement about the yaw axis. We also studied the behavior of the system as the robot walked quadrupedally. The eye and head movements were monitored and the angular position commands to the cameras were plotted as a function of time. When the head was commanded to oscillate about yaw as shown in Fig. 6, there were predominantly yaw components of head rotation in the spatial coordinate frame (Fig. 7). There were also small roll and pitch oscillations. There are a number of reasons for this. First, while there was a command to rotate about yaw, this was done by the AIBO motor and the actual motion of the head may have some pitch and roll (Fig. 7). Our sensor detected this. This is a property of the AIBO motor since it could not be absolutely aligned with the axes of the artificial vestibular sensor. Also, motors, which drive the AIBO head, do not have sufficient torque to overcome the moment of inertia of the head for large velocities. We have used head velocities that were limited to 35 degrees/sec. This led to small roll movements of less than ±1°. Thus, what we think is a pure yaw rotation may induce both pitch and roll in the coordinate frame of the sensors. This effect was dependent on the head velocity command and we were able to reduce the pitch and roll by generating head movements at lower speeds. The commands to rotate the cameras were compensatory and followed the head motion (Fig. 8). Similar compensatory commands were given to the cameras as the AIBO robot walked over-ground (Fig. 9 and 10). There are no specific sensors that gave us feedback on the actual position of the cameras. This could be obtained by processing the visual signals, but is beyond the scope of this paper and will be considered in a separate study.

VIII. CONCLUSION

In this work, we designed and implemented a lightweight binocular ocular system that has been mounted on a gimbal system capable of rotations similar to the eyes. The ocular system has also been successfully mounted on a commercially available Sony AIBO robot on which an acceleration sensor has also been mounted in the head of the robot. The software has been developed to implement a rudimentary compensatory aVOR. The acceleration sensor performs an excellent role as an artificial sensor in providing a 3D angular velocity vector. The artificial aVOR reacts to the head movements in all three dimensions and sends command signals that compensate any small movements. The novelty of this approach is that the artificial vestibular sensor was implanted in the head and that the artificial eyes, i.e. cameras, were implemented in Fick gimbals. As such, each camera can rotate in three dimensions as is the case for the eyes in humans and other animals. Moreover, our algorithm to rotate the eyes in response to head movements follows the compensatory equations for rotations in three dimensions. This will insure that yaw, pitch and roll movements of the head while locomoting will be compensated by the counter movements of the camera and maintain stable images. Such an implementation has not been accomplished in any robot to date.

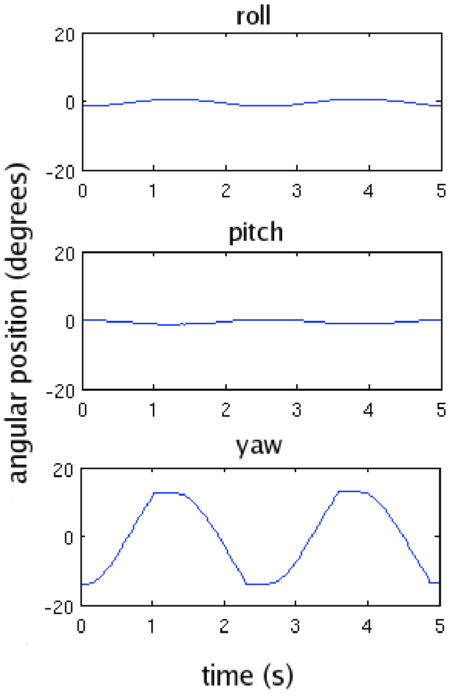

Fig. 7 Head Position in space during head rotation about a head yaw axis while the head is tilted down

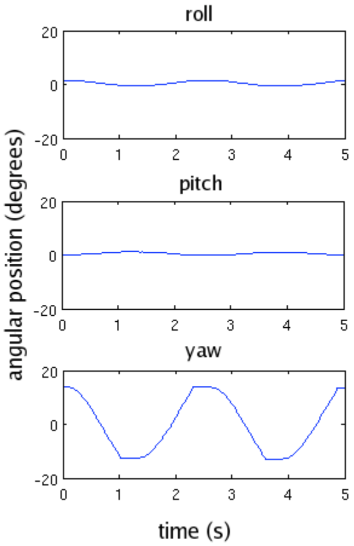

Fig. 8. Compensatory eye movements during head rotation about a head yaw axis while the head is tilted down

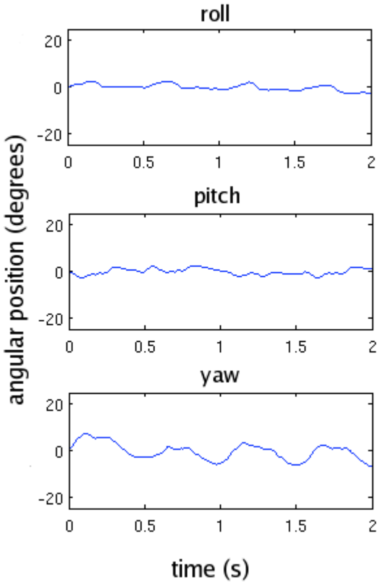

Fig. 9 Head Movements during quadrupedal locomotion of the AIBO robot.

IX. FUTURE WORK

The VOR in primates has been studied extensively and has been shown to play an important role during locomotion. The vestibulo-ocular reflex and vestibulo-collic reflexes implemented previously will be used to study their role in quadrupedal robot locomotion. Also, we plan to use image processing techniques to confirm the extent of stabilization of the images from the camera during locomotion.

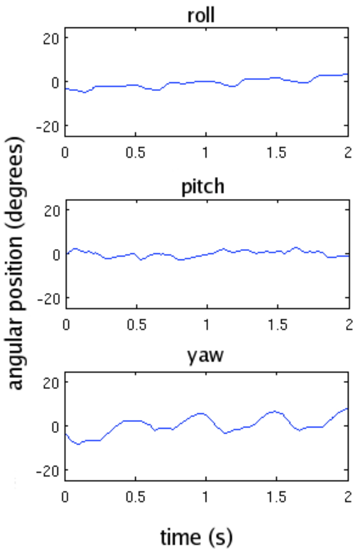

Fig. 10. Eye movements during quadrupedal locomotion of the AIBO robot.

X. ACKNOWLEDGEMENT

We thank Mr. Louis Tundis for providing the mechanical design and structure to the Ocular Servo Module, mounting acceleration sensor and for his innovative ideas for building and mounting the various components. We also would like to thank Mr. William Morris and Ms. Leah Kelley for providing drawings and helpful inputs to the project.

REFERENCES

[1] G. Bekey, *Autonomous robots*. Cambridge: MIT Press, 2005.

[2] T. Raphan and B. Cohen, "The vestibulo-ocular reflex (VOR) in three dimensions," *Exp. Brain Res.*, vol. 145, pp. 1-27, 2002.

[3] F. Patane, C. Laschi, H. Miwa, E. Guglielmelli, P. Dario, and A. Takanishi, "Design and development of a biologically-inspired artificial vestibular system for robot heads," presented at IEEE Conference on Intelligent Robots and Systems, Japan, 2004.

[4] G. Asuni, G. Teti, C. Laschi, E. Guglielmelli, and P. Dario, "A robotic head neuro-controller based on biologically-inspired neural methods," presented at IEEE Conference on Robotics and Automation, Spain, 2005.

[5] D. J. Coombs and C. M. Brown, "Intelligent gaze control in binocular vision," presented at IEEE Conf., 1990.

[6] F. Panerai, G. Metta, and G. Sandini, "Learning visual stabilization reflexes in robots with moving eyes" *Neurocomputing*, 2002.

[7] T. Shibata and S. Schaal, "Biomimetic gaze stabilization.", Neural Computing, Word Scientific, 1999.

[8] T. Tsujita, A. Konno, and M. Uchiyama, "Design and development of a high speed binocular camera head," presented at IEEE Conference, 2005.

[9] M. Marcinkiewicz, M. Kunin, S. Parsons, E. Sklar, and T. Raphan, "Towards a methodology for stabilizing the gaze of a quadrupedal robot," presented at RoboCup International Symposium, Bremen, Germany, 2006.

[10] T. Raphan, "Modeling control of eye orientation in three dimensions. I. Role of muscle pulleys in determining saccadic trajectory," *J. Neurophysiol.*, vol. 79, pp. 2653-2667, 1998.

[11] T. Imai, S. T. Moore, T. Raphan, and B. Cohen, "Interaction of the body, head, and eyes during walking and turning," *Exp Brain Res*, vol. 136, pp. 1-18, 2001.

[12] T. Raphan, T. Imai, S. T. Moore, and B. Cohen, "Vestibular compensation and orientation during locomotion," *Ann N Y Acad Sci*, vol. 942, pp. 128-38, 2001.