

Sensory Integration With Articulated Motion On A Humanoid Robot

Juan Rojas and Richard Alan Peters II

Center for Intelligent Systems
Vanderbilt University
Nashville, Tennessee, USA.

Email: Juan.Rojas@Vanderbilt.Edu, Rap2@Vuse.Vanderbilt.Edu.

Abstract

The paper describes the integration of articulated motion with auditory and visual sensory information that enables a humanoid robot to achieve certain reflex actions that mimic those of people. Such reflexes as reach-and-grasp behaviors can enable the robot to learn, through experience, its own state and that of the world. A humanoid robot with binaural audio input, stereo vision, and pneumatic arms and hands exhibited tightly coupled sensory-motor behaviors in four different demonstrations. The complexity of successive demonstrations was increased to show that the reflexive sensory-motor behaviors combine to perform increasingly complex tasks. The humanoid robot executed these tasks effectively and established the groundwork for the further development of hardware and software systems, sensory-motor vector-space representations, and coupling with higher-level cognition.

1. Introduction

Through sensing coupled with motion, people interact with, and learn about, their environments. Such interactions are developmental as well as purposeful. Infants, for example, manipulate objects and analyze them sensually, using vision for shape, color, pattern, and texture, touch for temperature, shape, and texture, smell for aromatic information, and taste for flavor [16]. Manipulation structures the sensory information. Since the laws of physics are constant, similar actions on similar objects yield similar results consistently. Over time, the integration of multimodal sensory information with motion grounds the infant in reality and provides a basis for understanding the world and his or her effects on it.

Robots, likewise, can learn to interact with the world through sensing and motion. Pfeifer has shown that sensory-motor state data can self-organize into vector space structures that, in effect, categorize the

world in terms of the robot's Sensory Motor Coordination (SMC) [23]. Thus the development of SMC grounds the robot and provides a foundation for it to interact purposefully with people and its environment.

The goal of our research is to create humanoid robots that can interact autonomously and socially within a human structured environment. This paper presents the results of combining sensory-motor reflexes that map auditory and visual input to camera-head motions and reach-and-grasp behaviors.

2. Research Test Bed

Researchers at Vanderbilt University have built a humanoid robot, ISAC, and a parallel, distributed software system, IMA, with which to perform research on the acquisition of intelligent behaviors.

2.1. The ISAC Robot

ISAC (Intelligent Soft Arm Control) is a stationary humanoid robot. It is an anthropomorphic upper torso equipped with a color stereo vision system, two independent pan-tilt units, two microphones, an infrared sensor array, two pneumatic arms and hands, speech output, and speech recognition capabilities. (Figure 1.) The robot was originally created to assist handicapped people while ensuring their security [13], [14].

2.2. The Intelligent Machine Architecture

The intelligent machine architecture (IMA) is an object-oriented software system that controls the robot through modular descriptions of its hardware, its tasks, and its environment [21], [22]. The architecture comprises two levels of abstraction. The higher level is a multi-agent network defined by *agents* and *relationships*. An agent is a software object that tightly encapsulates a *resource* (either physical or computational). A relationship is a software object that connects two or more agents. The lower level implements agents and relationships through

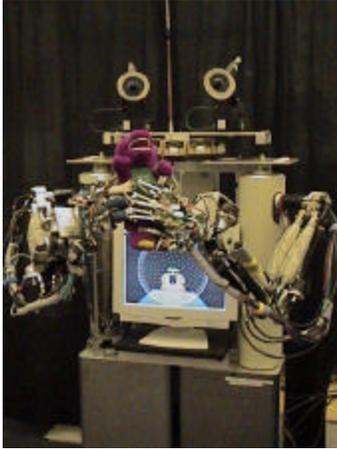


Figure 1. The humanoid Robot, ISAC.

component-object software modules that have an established communications protocol. (These are based on Microsoft DCOM, the distributed component object model service of Windows NT/2000.) Thus IMA is a network of parallel, concurrent software modules formed from simpler, reusable component objects.

IMA describes ISAC with respect to sets of resources:

- Physical: head, arms, and hands.
- Skill: sonic localization, visual tracking, reaching, and grasping.
- Task: Detecting audio cues, finding objects, and sensing objects.

Each resource is encapsulated as an agent and is connected dynamically to other agents by relationships. Agents can be created from component-objects or other agents by combining reusable subcomponents and their parameters. (See Figure 3.)

IMA facilitates coarse-grained parallel processing because of the loose coupling afforded by message passing and because DCOM allows software objects on separate computers to be treated as if they were local to each other. Each agent acts locally based on its internal state and provides a set of services to other agents through various relationships. The resulting asynchronous, parallel operation of decision-making agents simplifies the system model at a high level.

Through encapsulation, IMA maintains robustness of operation within an evolving system because the network of agents is isolated from internal changes in any one agent. Essentially, the high level model provides a shell around the implementation level. That allows a programmer to experiment with intra-agent algorithms without changing the overall structure of the

system. Additionally, component mechanisms can be defined at both run-time and design-time allowing for a dynamic configuration.

3. Sensors

Vision and audio were used as the main sensory modalities for ISAC. Two Sony XC999 digital color cameras and two high fidelity microphones were used. Audio signals were used to direct the robot's attention and visual information was used to locate and track an object of interest.

3.1. Vision

ISAC's vision system was designed to be simple and efficient. Color and motion segmentation were used to extract information from the environment.

The color segmentation algorithm uses color histograms. There is a separate histogram for each object that the robot can identify. A histogram is specified by a probability ellipse in Hue-Saturation space. The latter were extracted from images of the object through statistical analysis. Histograms enable real-time comparisons with visual input to detect objects of interest [2]. Parameters μ_x , μ_y , σ_x , σ_y , and ρ are used to find the probability ellipse with an inner uniform distribution inside. (Figure 2.) The equation for the ellipse is:

$$\frac{1}{(1-r^2)} \left[\frac{(x-\mathbf{m}_x)^2}{\mathbf{s}_x^2} - \frac{(x-\mathbf{m}_x)(y-\mathbf{m}_y)}{\mathbf{s}_x\mathbf{s}_y} + \frac{(y-\mathbf{m}_y)^2}{\mathbf{s}_y^2} \right] = I^2 \quad (1)$$

where the center of the ellipse is at (μ_x, μ_y) . Parameter λ is the standard deviation of the data inside the ellipse and ρ is a correlation coefficient. The angle of rotation of the ellipse about the center, θ is defined by:

$$\theta = \tan^{-1} \left[\frac{1}{2r\mathbf{s}_x\mathbf{s}_y} \left(\mathbf{s}_y^2 - \mathbf{s}_x^2 \pm \sqrt{(2r\mathbf{s}_x\mathbf{s}_y)^2 + (\mathbf{s}_y^2 - \mathbf{s}_x^2)^2} \right) \right] \quad (2)$$

The color of interest is modeled through this statistical procedure in the 2-D HS color space. Pixel classification techniques check to see if a pixel value falls within the ellipse's boundary given some value for the standard deviation, λ . Each pixel in the image whose HS value falls within the ellipse, $\mathbf{a} < 1$, is marked as a foreground pixel, part of the target.

$$\mathbf{a} = \frac{1}{I^2(1-r^2)} \left[\frac{(x-\mathbf{m}_x)^2}{\mathbf{s}_x^2} - \frac{(x-\mathbf{m}_x)(y-\mathbf{m}_y)}{\mathbf{s}_x\mathbf{s}_y} + \frac{(y-\mathbf{m}_y)^2}{\mathbf{s}_y^2} \right] \quad (3)$$

A morphological opening — an erosion followed by a dilation with a 3x3 pixel cross — is done to eliminate small regions.

Motion segmentation is based on a thresholded, pixel-wise difference of successive frames. If the

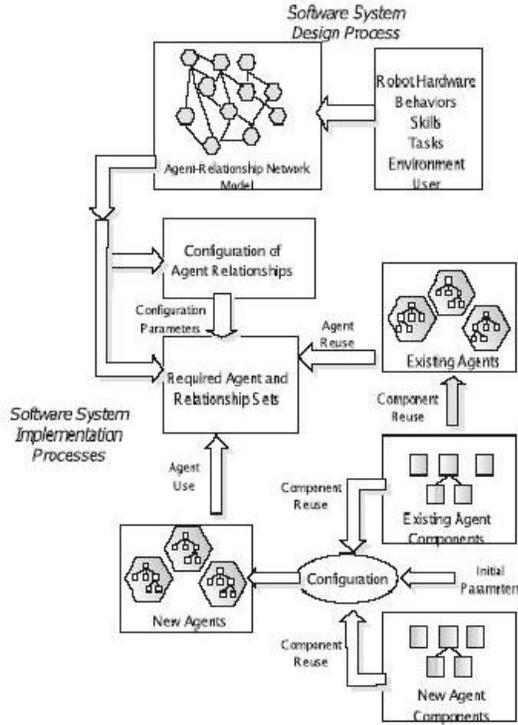


Figure 3. IMA Architecture.

result at a pixel is greater than the threshold, the pixel is marked as part of a moving object. Blurring and down-sampling are used to reduce image noise. The center of mass of the marked pixels, is taken as the image location of the moving object.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N F_{xi} \quad \& \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N F_{yi} \quad (4)$$

N is the total number of foreground pixels, F_x and F_y represent the number of segmented pixels across the rows and columns of an image. The method assumes that only one object is being analyzed at a time. If more than one object is segmented, the focus of attention will be somewhere between the objects.

3.2. Audio

A stereo pair of microphones is used for audio cueing of the camera head through sonic localization. The algorithm is based on a cross-channel ratio of sound energy between the right and the left audio channels. A sound intensity envelope is detected in both channels and passed through a lowpass digital Butterworth filter. The squared value of each signal at each instant is computed and used as a measure of energy:

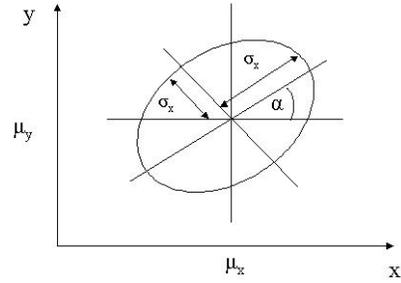


Figure 2. Geometry of the probability ellipse [2].

$$R_n = \frac{\bar{E}_{Ln}}{\bar{E}_{Rn}} \quad (5)$$

To determine the detection ratio, a sound source was measured at eleven different pan angle locations. The values of the ratios at each location were computed and were used as references to estimate the direction of the source of a sound. These values can be calibrated easily at any time to adjust the system for different ambient noise conditions. Upon the detection of a sound that exceeds a threshold in both channels, the ratio of channels energies and one of the 11 angular intervals is selected as the direction of the sound. The detected angle is made available to other agents in the system.

4. Motion

When a robot possesses knowledge about its environment, behavioral responses can take place. When the appropriate image processing techniques have taken place, a spatial location is computed using a neural net and passed to the pan-tilt unit. When the pan-tilt unit fixates on the desired object a Cartesian location in space is derived and used to reach towards the target and grab it.

4.1. Eyes

A pair of high-speed, accurate positioning pan-tilt units were used to move the cameras. The pan and tilt angles and the velocity and acceleration in each camera is independently controlled.

4.2. Eye Movements

The brain's oculomotor system uses three different types of eye motions to keep an object of interest in the fovea [3], [27]. In ISAC, saccades and smooth pursuits are the main type of eye movements [26].

ISAC's saccade function uses two modules: a map trainer and a command generator. The map trainer provides an adequate transformation for the

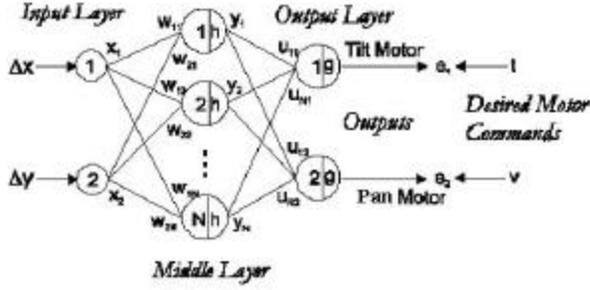


Figure 4. Feed-forward neural net [26].

command generator to issue accurate commands to the pan-tilt unit. A neural net was used to map image coordinates to motor coordinates because it is fast and accurate, and no camera calibration is needed. The training is done off-line by using a back-propagation learning algorithm. Each pan-tilt unit is trained individually. The inputs to the neural net are the x and y position of the gazed target, whilst the output nodes are those corresponding to the pan and tilt motions respectively. A hidden middle layer contains 25 units that are mapped through a bipolar hyperbolic tangent sigmoid function. Figure 4 is a diagram of the neural net. The command generator uses the map to produce motor steps corresponding to target-position inputs, which are sent to the camera head controller. Post saccade processing is used to correct the neural weights to minimize the error of the transformation.

A smooth pursuit keeps a target in the fovea continually using target position and target velocity. The latter is used to predict the future position of the camera and allow for a smoother trajectory. Two concentric regions of different radii are defined as: the fovea with a radius of F pixels, and the dead zone with a radius of D pixels, where, $F > D$. Smooth pursuit, also known as proportional tracking, occurs when the target is located outside the dead zone but inside the fovea. If the target exits the fovea, a saccade is issued to reach the target. The positional vector is the distance of the target from the center of the fovea and is defined as:

$$\vec{p} = (p_x, p_y) = (\Delta x, \Delta y), \quad |\vec{p}| = \sqrt{\Delta x^2 + \Delta y^2} \quad (6)$$

Similarly, velocity is described as:

$$\vec{v} = (v_x, v_y), \quad |\vec{v}| = \sqrt{v_x^2 + v_y^2} \quad (7)$$

Where, the units are in pixels per time unit. It follows that the motor commands, m_{PL} , m_{PR} , m_{TL} , and m_{TR} ,

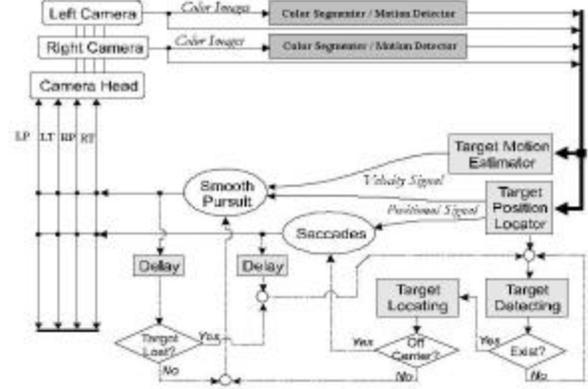


Figure 5. System integration for the visual system.

can be calculated by making use of the distance of the target and constant gains, $?_{PL}$, $?_{PR}$, $?_{TL}$, and $?_{TR}$, for the left pan, right pan, left tilt, and right tilt motors:

$$m_{PL,R} = k_{PL,R} \cdot \Delta x_{L,R}, \quad m_{TL,R} = k_{TL,R} \cdot \Delta x_{T,R}, \quad (8)$$

where, the L and R subscripts describe the left and right images respectively. To predict target position, let $?t$ be the time interval between consecutive image frames, let \vec{v} be the velocity of the target at time t , so the position of the target at the next frame, $t + ?t$, will be:

$$\Delta x_{t+\Delta t} = x_t + v_x \cdot \Delta t, \quad \Delta y_{t+\Delta t} = y_t + v_y \cdot \Delta t \quad (9)$$

This produces a smooth motion in the camera head when it tracks a specific target.

The camera head controller is open loop. Given that object trajectories vary dynamically in the environment, the trajectory cannot be precisely estimated. Overshoots of the controller may occur. To lessen this effect, $(?x_t, ?y_t)$ is low-pass filtered over t .

4.3. System Integration

The visual system combines all the modules previously mentioned. The images provided by the cameras are captured by an agent that encapsulates all the frame grabber functionalities. Then, objects are extracted from the images via color or motion segmentation. The center of mass is calculated and used by the saccade module to fixate on the target position. A velocity signal is also calculated and used by the smooth pursuit module. Both eye movements, experience a slight delay that accounts for the movement of the camera. Yet, data streams are passed as quickly as possible to keep the camera head on target. Saccades and smooth pursuits behave similarly to each other, see Figure 5.

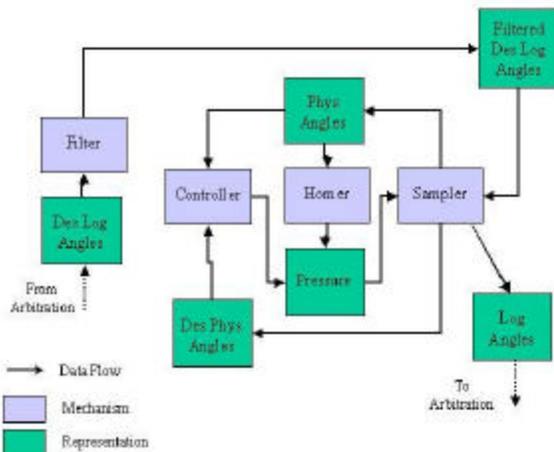


Figure 6. Arm Control Loop [1].

4.4. Pneumatic Arms

The humanoid robot, ISAC, is actuated by pneumatic “McKibben artificial muscles”. These are made from an inflatable, tubular inner bladder sheathed with a nylon double helix weave that shortens lengthwise when expanded radially [15]. Those two main components are clamped with fittings at both ends, one of which contains an air intake. The nylon sheath holds constant the volume of the gas within the rubber tube.

Therefore, as they are inflated, the actuators contract along the axis of the tube. Similarly, as they deflate they expand along the axis. If one end is fixed, the other will move a load in an approximately linear trajectory [8]. The arms exhibit compliance since the pneumatic actuators operate on the basis of gas compressibility, and their inner bladders are elastic. Even if the gas pressure remains unchanged, an applied force that changes the length of the actuator produces a spring-like behavior in the rubber material of the bladder. That enhances the compliance of the actuator beyond the compressibility of the gas. Because of their constituent materials, McKibben air muscles are lightweight and have a characteristically high force to weight ratio [24].

To move the end-effector to a specific position, inverse kinematics computes joint-space angles that are converted to physical angles which are used to compute an output pressure. A sampler drives the pneumatic servo valves which control the pressure in the artificial muscles. See Fig. 6.

The controller paradigm was biologically inspired and created by S. Northrup [2001]. A rigorous study of EMG signals in the human arm was done in order to develop the paradigm. ISAC’s arms are modeled like those of humans with agonist-antagonist muscles. In

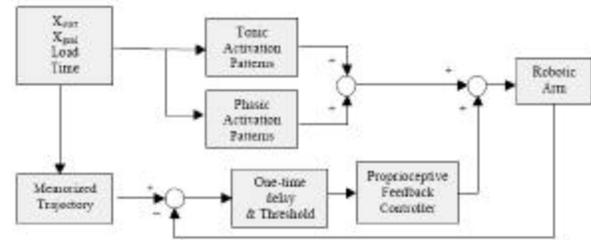


Figure 7. Block Diagram of Tonic-plus-Phasic Plus Feedforward Controller [20].

studying quick arm motions, Flanders observed that the arm experiences different forces in horizontal and vertical movements [9]. For horizontal motions, triphasic activation occurs. For vertical motions, triphasic activation is superimposed with tonic activation (the activation levels needed for quasi-static postural control). ISAC’s arms were modeled after that tonic-plus-phasic control paradigm.

Reaching motions may employ a variety of sensory modalities for motor control – visual and proprioceptive feedbacks are used by ISAC. However, if a fast motion directed toward a target is required, there is not enough time for visual or proprioceptive feedback loops to be effective. A feed forward control technique was chosen to overcome this problem. Northrup referenced the error efferent signals after the feedback could occur [20]. The controller compared the actual motion with the programmed one and if the difference exceeded an empirically calculated threshold, a feedback error controller would adjust the motion. A block diagram of the Northrup’s controller is shown in Figure 7. In it X_{start} , X_{goal} , are the initial and final goal positions, $Reaching\ Time$ is the duration of the movement, and $Load\ Size$ is the weight of the grasped object. For a given set of parameters, a motor program is known and is sent to the artificial muscles as a time sequence of the sum of the phasic and tonic activation levels. Also, after one hundred milliseconds the feedback loop was activated.

4.5. Hand

The hand is a in-house hybrid design that consists of a thumb and forefinger driven by an electric motor and pneumatically actuated distal fingers [7]. A PC controller card specifies the desired pressure on gas valves that either open or close the hand. The hand has proximity sensors to aid in grasping. They are photoelectric sensors located on the palm of the robot hand. One fires when an object is within 10cm; the other fires when an object is within 1 cm. Thus, the former warns of the approach of an object and the latter responds when the fingers can close on the object.

5. Sensory Motor Control

Sensory motor coordination (SMC) is the basis for purposeful action within a dynamic, loosely structured environment [5]. It is feedback loop through which an animal changes the environment by acting on it, and senses the changes. This close coupling of sensory signals and motor activity, can be described in terms of a *sense-act* paradigm. SMC is necessary for a robot that operates in a loosely structured environment. In ISAC each sensor is modeled as an independent agent other agents couple these to motor controllers to create reflex actions or basic behaviors [19]. Complex tasks can be performed through a composition or superposition of behaviors.

There is evidence to suggest that SMC serves as a foundation for higher-level learning. In [23], Pfeifer reported that sensory data and concurrent motor control information recorded as a vector time-series formed clusters in a sensory-motor state-space. He noted that the state-space locus of a cluster corresponded to a class of motor action taken under specific sensory conditions. In effect, the clusters described a categorization of the environment with respect to sensory-motor coordination (SMC).

An exemplar of an SMC cluster corresponds at once to a basic behavior as defined by Brooks [4] and to a competency module in a spreading activation network [17]. The latter is a specific example of a more general class of topological, action-map representations of an environment [18], which can be controlled by discrete-event dynamical systems (DEDS) [11], with transition probabilities given by Markov decision processes. If the state-space is parameterized by time, the clusters are collections of trajectories and an exemplar is a single representative trajectory through the space.

Thus, if a robot is controlled through an environment to complete a task while recording its SMC vector time-series, the result is a state-space trajectory that is smooth during the execution of a behavior but that exhibits a corner or a jump during a change in behavior (an *SMC* event). From this, a DEDS description of the task can be formed as a sequence of basic behaviors and the transitions between them. The task is learned in terms of the robot's own sensors, actuators, and morphology.

Thus, the basic information obtained from the environment can be used to deliberate based on some goal and sequence of tasks to execute. As demonstrated in [6], robots can learn from their own experience by constructing models of the dynamics of its own SMC data. Through teleoperation a set of trajectories that cover the extreme points of the robot's workspace were learned. Then the task could be

executed autonomously by the robot under differing conditions.

With this research in mind, the humanoid robot, ISAC, was given a set of basic behaviors. These enabled ISAC to emulate human reactions to sensory information. From this, two results could be achieved: to react to the environment in a human-like fashion, and to through experience perform autonomous behaviors. A description of the demonstrations implemented for ISAC now follows.

6. Agents

The IMA architecture allows for the implementation of multiple agents. Each agent's functionality is a result of the interactions of their multiple subcomponents. A total of six agents were used: the Sound Agent, the Camera Agent, the Head Agent, the Hand Agent, the Arm Agent, and the Trajectory Agent.

The Sound Agent was used to find the direction of a sound source near the robot. To be detected the sound had to exceed a predetermined threshold value. This was done to ignore ambient noise. The pan angle to the sound source was computed in the horizontal plane so that the camera head could respond directly. The angles were quantized in intervals of 15 degrees with an origin at the perpendicular bisector of the camera head baseline.

A Camera Agent captured images at a rate of about 10Hz. There was one agent for each camera in the pair. Depending on the task at hand, a color or a motion segmentation algorithm was selected. Each new image was segmented per the algorithm, the center of mass was calculated to be used as the focus of attention (FOA) and was made available to other agents. The Head Agent, which serves as the "brain" of the sensory-motor demonstrations, read the FOA from both Camera Agents through a proxy connection which was used to trigger a saccade or smooth pursuit behavior. Subsequent to the motion, the pan and tilt angles of the PT units were then used to calculate the 3-D Cartesian position (with respect to ISAC's base frame) of the fixated object.

The Hand Agent could produce two different grasp mechanisms: a fast grasp and a slower but more precise grasp. The former did not use the motored fingers, whereas the latter did. The choice of grasp is selectable and can be registered in the finite state machine (a simple DEDS) that controlled the grasp. The hand's proximity sensors were used to prevent false (empty) grasps.

The Arm Agent computed the kinematics that articulated the appropriate motion in the arm. Goal

locations could be passed to the arm in two forms: as Cartesian coordinates or joint-space angles.

The Trajectory Agent generated the Cartesian paths for the end-effector and periodically updated the Arm Agent. Motions were computed from given starting and ending locations (intermediate points for specific routing could be used as well), the duration of the trajectory, and the parameter type (either joint angles or Cartesian coordinates). The starting point used was the arm's home position; the ending point was a Cartesian provided by the Head Agent.

The appropriate connection of these six autonomous agents gave the robot a variety of reflex-actions.

7. Demonstrations

Four demonstrations were implemented to elicit various behaviors from ISAC. The demonstrations began with a low degree of complexity and increased with each consecutive routine. The goal of the demonstrations was to produce basic behaviors by achieving sensory fusion and articulating basic motion in the robot.

The first demonstration displayed a relationship between sonic cues and attention by using the Head Agent and the Sound Agent. The Sound Agent output the angle of localized sound cues to the Head Agent. The latter generated a pan motion towards the given locations. The demonstration successfully emulated audio cueing, the animal behavior of attending to a distinct noise in the environment.

The second and third demonstrations built on the first one but used simple visual object recognition. In one demonstration, objects were recognized by color and in the other, by motion. Once a sonic cue was detected and attended to by the system, an event was triggered to activate the segmentation algorithm. The center of mass of the object was calculated and used by the head controller to fixate on the object and to track it if the object moved from the fixation point.

The final demonstration increased the complexity by performing a reach-and-grasp behavior after the object had been detected and attended to. The Head Agent computed a Cartesian location in space for the object of interest. This location was then used as the goal position for the arm. The Trajectory Agent created a motion path from the current position of the arm to the goal position. Once the reach motion was completed, the proximity sensors in the hand checked for the presence of the object, if it was detected, the grasp was executed. Single reach-and-grasp behaviors were successfully achieved for selected objects. Additionally, multiple reach-and-grasp reflex actions were achieved by updating the robot's goal and issuing a desired number of grab commands.

7.1. Limitations

The demonstrations were inherently limited by their reactive nature. The behavior sequences exhibited by ISAC were strict responses to sensory input. Currently the system lacks integration with higher-level commands, thus ISAC can only perform instinctive but not deliberate actions. For example, it cannot deliberately shift gaze to another relevant area in the visual field, nor can it decide if it should inhibit a reach-and-grasp motion if it were not appropriate.

Also, integration and cooperation between auditory and visual information has not been established beyond audio cueing. The humanoid cannot simultaneously use information from both inputs to visualize the environment; rather, its sensory input stream is sequential. Auditory inputs trigger visual inputs, which in turn trigger motion in the arm.

8. Conclusion and Future Work

The humanoid robot, ISAC, successfully used sensory data environment to articulate motion reflexively. It reliably detected the presence of people and objects in the world and displayed sensory-coupled reactions that grossly approximate those of a human baby. In the presence of audio cues in the environment, the robot was able to pan its cameras toward them to search of objects of interest. This sets up ISAC to be in a situation where it can learn more about an adjacent person's intentions. When the robot was directed to recognize specific objects (with characteristic colors) the humanoid effectively panned, fixated, and tracked them. This step allows ISAC to analyze potential salient features of the object of interest. Finally, vision enabled simple reach-grasp behaviors to occur. Through reaching and grasping, the robot has gained the ability to fetch objects and do further analysis. These reflex actions, therefore, have set the stage for further research in robot-environment interaction and learning and, perhaps, more meaningful human-machine interactions.

Much progress is pending in software and hardware systems, in sensory-motor associations, and in cognitive control. A more functional and precise sound acquisition and analysis system would be helpful as would speech output and speech recognition. Much work remains to be done on the vision system to increase its functionality for better integration with articulated motion, so that more sophisticated experiments in SMC can be performed. Similarly, a more sensitive hand and arm with better control mechanisms are needed. Adaptive control also is being implemented.

Finally, the additions of attention and higher level commands guided by decision-making systems would push the robot to a state where cognition and intelligent

behavior can be explored and developed. Efforts to produce cognitive behaviors are being pursued in other labs [9],[25]. This area of study promises technological advancement and exciting breakthroughs that will drive the effort to develop a cognitive and social humanoid robot.

9. References

- [1] Alford, A., M. Cambron, A. Srikaew, *ISAC Operation Manual*, CIS Technical Report, Vanderbilt Univ., Nashville, TN, 1999.
- [2] Barile, J. B., *Color Segmentation for Focus of Attention in an Active Trinocular Camera Head*. Masters Thesis, Vanderbilt Univ., Nashville, TN, 1997.
- [3] Bear, M. F. & B. W. Connors, "Exploring the brain", *Neuroscience*, Williams & Wilkins, 1996.
- [4] Brooks, R. A., "A robust layered control system for a mobile robot", *IEEE Journal of Robotics and Automation*, 1986, vol. RA-2, pp. 14-23.
- [5] Cambron, M., and R.A. Peters II, "Sensory motor control for grasping in a humanoid robot". *Proceedings 2000 IEEE International Conference on Systems, Man and Cybernetics*, Nashville, TN, 2000.
- [6] Campbell, C. L., *Learning through Teleoperation on Robonaut*, Masters Thesis, Vanderbilt Univ., Nashville, TN, 2003.
- [7] Christopher, J.L., *A PneuHand for Human-like Grasping on a Humanoid Robot*, Masters Thesis, Vanderbilt Univ., Nashville, TN, 1998.
- [8] Daerden, F., D. Lefeber, "Pneumatic artificial muscles: actuators for robotics and automation", *European Journal of Mechanical and Environmental Engineering*, 2002, 47(1): 10-21.
- [9] Flanders, M., J. J. Pellegrini, and S. D. Geisler, "Basic features of phasic activation for reaching in vertical planes", *Experimental Brain Research*, 1996, 110, pp. 67-79.
- [10] Fitzpatrick, P., G. Metta, L. Natale, S. Rao and G. Sandini, "Learning about objects through action - initial steps towards artificial cognition", *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, 2003.
- [11] Huber, M., R. A. Grupen, "A Hybrid Discrete Event Dynamic Systems Approach to Robot Control", UMass Computer Science technical report, No. 96-43, Univ. of Massachusetts, Dept. of Computer Science, Amherst MA, 1996.
- [12] Huber, M., *A Hybrid Architecture for Adaptive Robot Control*, PhD Thesis, Univ. of Massachusetts, Amherst MA, 2000.
- [13] K. Kawamura, R. A. Peters II, D. M. Wilkes, W. A. Alford, T. E. Roger, "ISAC: Foundations in Human-Humanoid Interaction", *IEEE Intelligent Systems*, 1996, vol. 15, no. 4, pp. 38-45.
- [14] Kawamura, K., R. A. Peters II, S. Bagchi, M. Iskarous, and M. Bishay, "Intelligent robotic systems in service of the disabled", *IEEE Transactions on Rehabilitation Engineering*, 1995, 3.1:14-21.
- [15] Klute, G., B. Hannaford, "Fatigue Characteristics of McKibben Artificial Muscle Actuators", *Proceedings IROS*, B. C., Canada, 1998, pp. 1776-82.
- [16] Lamb, M. E., M. H. Bornstein, D. M. Teti, *Development in Infancy: An Introduction*. Mahwah, N.J. Lawrence Erlbaum Associates, Inc., 2002, pp. 25-30.
- [17] Maes, P. and R. A. Brooks, "Learning to Coordinate Behaviors", *Proc. Eighth National Conf. on Artificial Intelligence*, 1990, vol. AAAI-90, pp. 796-802.
- [18] Mataric, M. J., "Integration of Representation Into Goal-Driven Behavior-Based Robots", *IEEE Transactions on Robotics and Automation*, 1992, vol. 8, no. 3, pp. 304-312.
- [19] Mataric, M. J., "Behavior-Based Robotics as a Tool for Synthesis of Artificial Behavior and Analysis of Natural Behavior", *Trends in Cognitive Science*, 1992, vol. 2, no. 3, pp. 82-87.
- [20] Northrup, S., N. Sarkar, and K. Kawamura, "Biologically-Inspired Control Architecture for a Humanoid Robot", *IEEE International Conference on IROS*, Hawaii, 2001.
- [21] Pack, R.T, *IMA: The Intelligent Machine Architecture*, PhD Thesis, Vanderbilt Univ., Nashville, TN, 2003, pp. 51-80.
- [22] Pack, R. T., D. M. Wilkes, and K. Kawamura, "A Software Architecture for Integrated Service Robot Development", *IEEE Conf. On Systems, Man, and Cybernetics*, Orlando, FL., 1997, pp. 3774-3779.
- [23] Pfeifer, R., and C. Scheier, "Sensory-motor coordination: the metaphor and beyond", *Robotics and Autonomous Systems, Special Issue on "Practice and Future of Autonomous Agents"*, 1997, 20.2-4, pp. 157-178.
- [24] Rojas, J. L., "Sensory Integration to Produce Articulated Motion on a Humanoid Robot", Master Thesis, Vanderbilt Univ., Nashville, TN, 2004, pp. 42-43.
- [25] Scassellati, B., "Foundations for a Theory of Mind for a Humanoid Robot", PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, 2001.
- [26] Srikaew, A., "A Biologically Inspired Active Vision Gaze Controller", PhD Thesis, Vanderbilt Univ., Nashville, TN, 2000.
- [27] Wandell, B. A., *Foundations of Vision*, Sinauer Associates, Inc., 1995.
- [28] Xinyu A., "Navigation via Sensory Motor Coordination", Master Thesis, Vanderbilt Univ., 2003.