Continuous Humanoid Interaction: An Integrated Perspective – Gaining Adaptivity, Redundancy, Flexibility – In One

Gordon Cheng, Akihiko Nagakubo, and Yasuo Kuniyoshi

Humanoid Interaction Laboratory
Intelligent Systems Division
ElectroTechnical Laboratory (ETL)
1-1-4 Umezono, Tsukuba, Ibaraki
305-8568 JAPAN
email:{gordon, nagakubo, kuniyosh}@etl.go.jp
http://www.etl.go.jp/~{gordon, nagakubo, kuniyosh}

Abstract. In this paper we take the view that human-like response can only be yielded through a richly integrated humanoid system. By taking such a view, we have confined ourselves to pursue the development of a richly integrated humanoid system. In this paper we present our initial effort in this pursuit, we present a humanoid system – an upper body humanoid robot – with active real-time stereo vision, an auditory system for spatial hearing, and proprioceptive systems, with a high performance motor control system.

The context in which we wish to establish our research is in the context of continuous humanoid interaction. Interaction with the environment, as well as interaction with people, all form part of this establishment. We present our approach to the problem of interacting with a continuum of multiple stimuli, while producing meaningful responses. We will show by using a relatively simple mechanism for integration, it is still possible to realise a vastly responsive system. Hence, providing a system that is adaptable through redundancy, and flexible for integration.

Our presentation includes a new humanoid robot system currently being developed for complex continuous interaction. An example of our humanoid robot in continuous interaction is presented. The system is able to track a person by sight in an unmodified environment, perform real-time mimicking of the upper body motion of the person, track a sound source (spatial orientation), and physical handling of the system in a compliance manner is also allowed. Each of the sub-systems of our humanoid is also introduced, with experimental results of each presented.

1 Introduction

Over the years robotics with AI research has taken many shapes and forms. We can identify a number of key issues: the aim to build machines that perform tasks; to understand how, and what mechanism will allow us to produce human-like responses[1]. We believe that humanoid research falls closely with both of these objectives, and that humanoid research will bring us even closer to the realisation of such goals.

Our own primary research objective however, falls closer with the goal of the investigation of mechanisms, which allow us to construct a robotic system that can yield human-like responses – in one completely integrated humanoid interactive system.

1.1 Humanoid Interaction

A humanoid[2] which takes the form of a human person, and should also respond in ways similar to a person. As it has been suggested, the form of a human will naturally induce human like response from other humans, and will deduce the expectation of human-like responses [1–4], for example see Figure 1. Figure 1 provides evidence of such a natural response by a human person with our humanoid robot system.

Placing our emphasis on humanoid interaction, we identified two types of interaction: active and passive.

Active: requires physical interaction with the environment, including humans, objects and the physical self. Such as being physically handled, or pushing itself up using its body and environment.

---

[1] the word “intelligent”, is not used here because the word “intelligent/intelligence” seems to be overused for various descriptions of various things. Therefore at this stage we wish to avoid the use of this term.

[2] **-oid** suffix forming adjectives and nouns, denoting form or resemblance (asteroid; rhomboid; thyroid). [Greek eidos form] — ©*Oxford English Dictionary*.

Fig. 1. Natural human like interaction: It is our belief that the form of our humanoid body will naturally bring about human like interaction. The evidence we present here is of a child, without prompting naturally reach for the lower arm of our humanoid robot. Purely due to the fact that the robot is shaped like a person.

Passive: no physical contact is necessary, such as hearing, seeing and smelling. Thus, pure observation only, like listening to or watching a person.

It is apparent that these interactions will require a rich number of abilities from a single integrated system. For example, with the abilities to hear, to see, and to physically move and be moved, formed part of such systems. This can be established as the base of our research, to initially provide a richly able system for diverse interaction.

Our earlier work purposed a number of attributes that should be encompassed by a humanoid interaction system [4]:

seamlessness which allows the system to interact with the environment in a continuous manner. The system should be able to interact continuously in a responsive and timely manner, thus without stopping or re-starting.

adaptivity/redundancy the mechanism used should provide the system with redundancy by being adaptive. The system should be able to adapt to failure of sub-systems or components, for example if one component fails another is able to function without any dependency.

flexibility additional components should be able to be introduced in a simple manner. Flexibility is an important attribute for the process of integration.

The mechanism we proposed, is through channelling all inputs into an integrated system, in a competitive manner. Allowing no inputs being discarded, thus allowing a continuous seamless flow of inputs and outputs to be produced (further discussion, is given in Section 9).

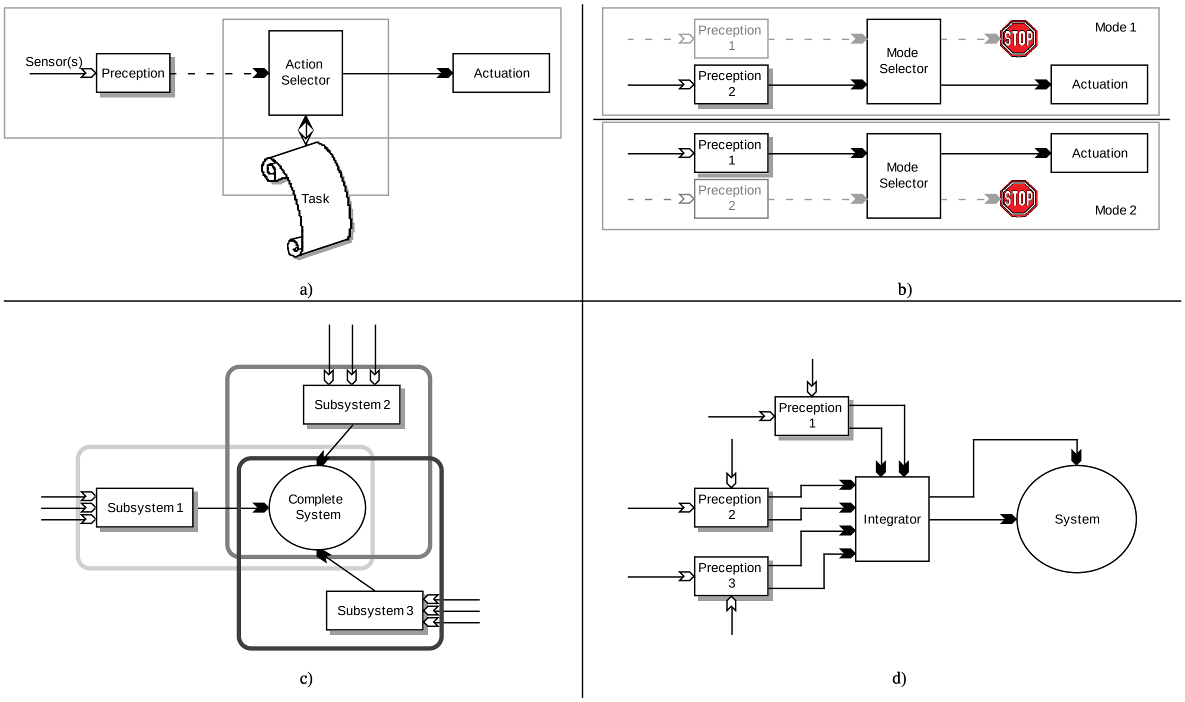


Fig. 2. Types of approach: a) Task-based/Task-oriented approach, b) Mode switching, c) Sub-systems – each assuming full control of a system, d) A form of an integrated approach.

From the above attribution of humanoid interaction, we arrived at the characterisation of an integrated humanoid system. We believe, an integrated humanoid system should possess a number of noticeable characteristics, and provide a number of benefits:

1. An integrated humanoid system, must have a rich sensory system available to it. And the processing must be processed in a concurrent and continuous manner, without any noticeable delay.
2. The system must be able to respond to the large and numerous types of inputs in various circumstances. And a wide range of responses must also be exhibited by the system.
3. Explicitly, an integrated humanoid system will ensure mutual human-like interaction is embraced, in a rich way: Physical, as well as non-physical in a seamless manner.
4. An integrated system will provide a different way in which problems can be solved, and to be solved. In the sense of humanoid – in a even more human like manner – for example, exploiting the rich sensory systems available, in dealing with the problem at hand.
5. Like humans, such a diverse system will yield a highly robust and redundant system, allowing it the ability to adapt to failure.
6. As like all endeavours, we are going to get some of this wrong, and having the availability of a large diverse system, will allow possibilities for further exploration. Thus, increasing the variation of possible solutions, and removing the boundary and limitation of common methodology.

Our ideas coincide with others, in the view that human-like attributes or characteristics can only be yielded from a complete embodied (human/human-like in this case) system [5–7, 2, 8].

Also, if we follow another line of thought, of [9] and [10], that a tool automatically enhances the way we think. Therefore, a humanoid robot will only increase our own understanding of human-like behaviours, and that, it will also increase our own intellectual abilities – so to speak. Along this line, a humanoid will change the way in which we would solve problems, it will allow us to match closely the way in which humans do things — “*in an interactive manner*”. Thus, allowing us to re-think and re-pose some of the problems that have been bound by the standard setup of a laboratory. By reposing their problem, the recent work of [11, 12] showed by exploiting sound and vision in one system, they were able to achieve better results for auditory segmentation and spatial localisation.

1.2 Types of Approaches

In order to clarify our discussion, we will begin by outlining some of the approaches taken by humanoid researchers, descriptions of some of these systems will be provided in Section 2. Categorically we have divided these approaches into four, as depicted in Figure 2:

1. Task-based/Task-oriented approach: these systems typically start with a task description, then procedurally derive a set of routines that are given to the system for later performance (e.g. WABOT-2, CENTAUR, P2 and P3, ARMAR, HRP). This approach can be limited by the available routines provided to the system, thus a large development effort is needed.
2. Mode Switching: a system is divided into individual sub-systems, a switcher or a selector is used to switch between each sub-system (e.g. HERMES, Hadaly-2 and WABIAN). Although switching may be the most straight forward and easiest way in which a system can be integrated. Nonetheless, one major drawback of this approach can be that, while the system is in one mode the system may miss to process, or discard important or critical information available to it.
3. Sub-systems having full control: in this approach each sub-system assumes complete control of the system (e.g. COG, DB). Thus, the interconnection and interaction between each sub-system is typically not considered. Due to this, this approach may eventually prevent a complete system to be derived as a whole.
4. Integrated view: should ensure that each sub-system can and are connected to the system, hence allowing a rich exploration of the interaction between all sub-systems to take place.

In this paper we'll present our initial work of an integrated humanoid robot system that takes an integrated view from the onset. Currently, we have produced a system that encompasses auditory processing, visual processing and proprioception processing abilities. Specifically, the vision processing for the detection of a person, with an attempt to infer the motion of the person has also been incorporated. In order to close the loop of humanoid interaction, an experiment of mimicking the motion of a sighted person has been realised. Our humanoid robot is able to visually track and mimic the person in real-time, in an unmodified environment. Other possible interaction includes physical, and auditory spatial hearing.

Section 2 presents some of the related work, outlining some of the current humanoid research efforts. Section 3 presents some of the design philosophy of our system. The current state of our system and some of the planned components are presented in Section 4. The processing and components of our humanoid system are then presented. Beginning with Section 5 the motor control of our system is presented. Section 6 presents the vision processing for our system. Section 7 present a spatial hearing system for our humanoid. The vestibular system for our humanoid is presented in Section 8. The initial way in which our system has been integrated is presented in Section 9.

In Section 10, we present an experiment of our initial attempt of the integration of a humanoid system, the presentation includes our humanoid copying the upper body motion of a person. A brief discussion of our future endeavours is presented in Section 11. A summary and conclusion is then given in Section 12

2 Research Landscape – Related work

In this section we discuss some existing humanoid projects and briefly outline some of their approaches to humanoid robotics research.

Although, this line of research in the development of human-like robots is not so new. One of the first[3] and most pronounced humanoids was the work of Waseda University. A human like robot, WABOT-2, plays a piano by sight reading music [13]. Their effort showed a great deal of engineering accomplishment back in the late 1980s. Producing a system which performed a highly skilled task. Their approach was pretty much task driven, in that the system must play the piano and read music. Other capabilities of the system were not considered as an issue.

In recent times, a large deal of attention has been placed on the humanoid named "COG". A robot which resided at the Massachusetts Institute of Technology, Artificial Intelligence Laboratory [14, 2, 15]. Their focus has mainly been towards the reproduction of human like cognitive abilities. Little of their earlier work was in addressing the issue of the production of body motion for their system, until the recent work of [16]. By carefully redesigning their existing mechanism, they have been able to produce some complex human like motions. Other sub-systems include the work of [17] for eye and face finding using vision. Their approach made every attempt in avoiding the use of task models, in deriving components for their system. In their development they have shown many subsystems that

---

[3] in modern time.

appeared to have close resemblance to human activities. However, each of their sub-systems made the assumption of complete control of a particular system. The integration of these rich components have not yet been made, and reaching an agreements between each of these components may prove to be difficult.

More recently, the system which is receiving the most amount of interest is that of Honda's humanoid systems, P2 and P3 [18]. Honda's key contribution was in producing biped locomotion that has not been achieved before. Such motions as autonomous walking or stair climbing have been produced. The impressive work of Honda have sparked a number of entities to undertake humanoid research, such as the efforts of the new projects of MITI[4] and KIST[5] (see below). The approach they have taken has been interesting, they set out to produce a system using whatever means possible, in order to provide the functionality/task that was specified, to walk. Other motions of their system were produced through a teleoperational interface. In their approach, each task or function is playback in the way that was developed – one at a time selected by an operator.

A recent project funded by the Ministry of International Trade and Industry(MITI) of Japan[6], the Humanoid Robotics Project (HRP) project, addresses the issues of searching for applications for humanoid systems [19]. Their aim is to evaluate the feasibility of current technologies, in producing humanoid systems which can be applied to a vast range of problems. The project is based on the development and integration of three key components: a humanoid robot, a remote cockpit and a virtual robot. The robot they are to employ is Honda's P3 humanoid robot. It has been proposed that in the first two years of this project, the establishment of a virtual environment and a virtual robot, will be used for the conceptualisation of various control methods. Such a control scheme as teleoperating a humanoid in performing complicated tasks. It is interesting to note that this project is approaching humanoid research from a different perspective, they are starting at a higher level. They are performing integration with existing systems, then evaluating how they will fit in an application domain and produce specific tasks. The switching for each task is then selected by a human operator in a closed operator-robot feedback manner [20].

The Korea Institute of Science and Technology (KIST) are also producing a humanoid system named, "CENTAUR" [21]. This system has an upper body of a person, and the lower half is horse like with four legs. In their system development approach, they have chosen to adapt open system technology. That is the components of their system are readily available, i.e. off-the-shelf. The work produced so far has been, teleoperation interface [22], and manipulate by motion planning of multiple-manipulators [23]. They have chosen two approaches, a close-loop approach and the planning-then-execution approach. However, it is still unclear how they have connected the two schemes.

Following on from their earlier work, Waseda University have been producing a number of humanoid systems, for example the humanoid robots Hadaly-2 and WABIAN (see [24] and [3]). Noticeably, two distinct topic of interest are under investigation, humanoid interaction and humanoid controls. Interaction with their systems can be via physical or non-physical contacts, such as by vision, voice or physical force. In the control of their humanoid they have manage to produced full body motions such as walking and dancing. In their organisation, they have chosen rather than to break up the system into subsystems, they divide into groups, each group pursuing development with a different objective. With this structure they are able to perform constant integration of their system, thus producing systems that are somewhat more coherent. We believe this is an attractive approach to complex system development. However, the connection between each sub-system has been divided in a way which the system switches between one sub-system to another. Thus, the switching may induce some latency to response or in some instances no response at all.

More recently, the Kawato Dynamic Brain Science research group acquired the humanoid robot manufactured by SARCOS Inc., named "DB". The actuation of the robot is based on hydraulics, making this system mobility extremely limited, due to the large pipe and pump required to operate the system. This group conducts research into motor learning, imitation and neural model-based control of a humanoid [25, 26]. Integration does not appear to be the focus of their research, therefore at this time no attempt has been made into the integration of their control schemes. Although this may be the case, many impressive demonstrations have been produced.

Humanoid as the perfect service robots has been the aim of many research groups. Recently a number of humanoid robots have been proposed, for example, two German group, one is the robot HERMES by the [27]. And the other named ARMAR, by the Forschungszentrum Informatik Karlsruhe (FZI) [28]. The work of the HERMES project is mainly based around the integration of off-the-shelf components in providing a human-like robot, with the equivalent capability of todays computer systems on a robot. A sweep of communication schemes have also been developed for HERMES, namely: by voice or email in a dialogue manner [29]. The ARMES project on the

---

[4] Ministry of International Trade and Industry of Japan.

[5] Korea Institute of Science and Technology.

[6] It is important to distinguish, that our research is not part of this new project, our project started independently in 1996, two years prior to this project.

other hand, has mainly been focused on the task of multiple manipulators, based planning and manipulation. Both of these systems provide a wheel-driven platform to allow mobility of their system.

The group at Vanderbilt University is also working towards a service orientated Humanoid robot, ISAC. Focus has been placed at the level of user interaction, such tasks as feeding of an elderly or disabled person has been considered [30]. A psychophysically based visual attention network has also been produced for this system [31]. A distributed software architecture has also been produced for this work [32]. The interesting aspects of their work is that they hold two views: one of engineering of a service robot and two, the utilisation of biologically inspired models in deriving sub-systems. However, these two separate views have yet to emerge as a complete system.

3 Design philosophy

Commonly, robotic systems have been designed for a particular task in mind, for example the work of [33], where the end-effector of the system is replaced when it is to manipulate a different object or perform a different task. This stands as the conventional view to system design, one module, and one function performing one task. The shortfall of this approach is that unexpected situations will ground a system to a halt, as the system is ill equipped to responded.

The design philosophy of our humanoid system is with the aim of producing a system that closely matches the capabilities of a human. A comprehensive view of the design philosophy was originally established in [1] and [34].

Having established capability as opposed to functionality as our main objective, we have built each of the components of our humanoid, achieving general movements like a person. Other human like features include, smooth outer surface, reasonable in size and weight, and backdrivability is also possible. These attributes make our humanoid ideal for physical interaction.

The design of our humanoid system was not to achieve any optimum task in particular, whereas the arm can be used for reaching, pushing, or for the support of its own body. Our design philosophy is well supported by the apparent evidences, that the structure of the human body is diverse in accomplishing a wide range of activities.

In terms of strength and power, in everyday situation the system should match the performance of a mature adult human, our aim is not to produce a super being. Our aim is just to produce a humanoid which can function as we do. The latest mechanical and technical details of our system can be found in [35].

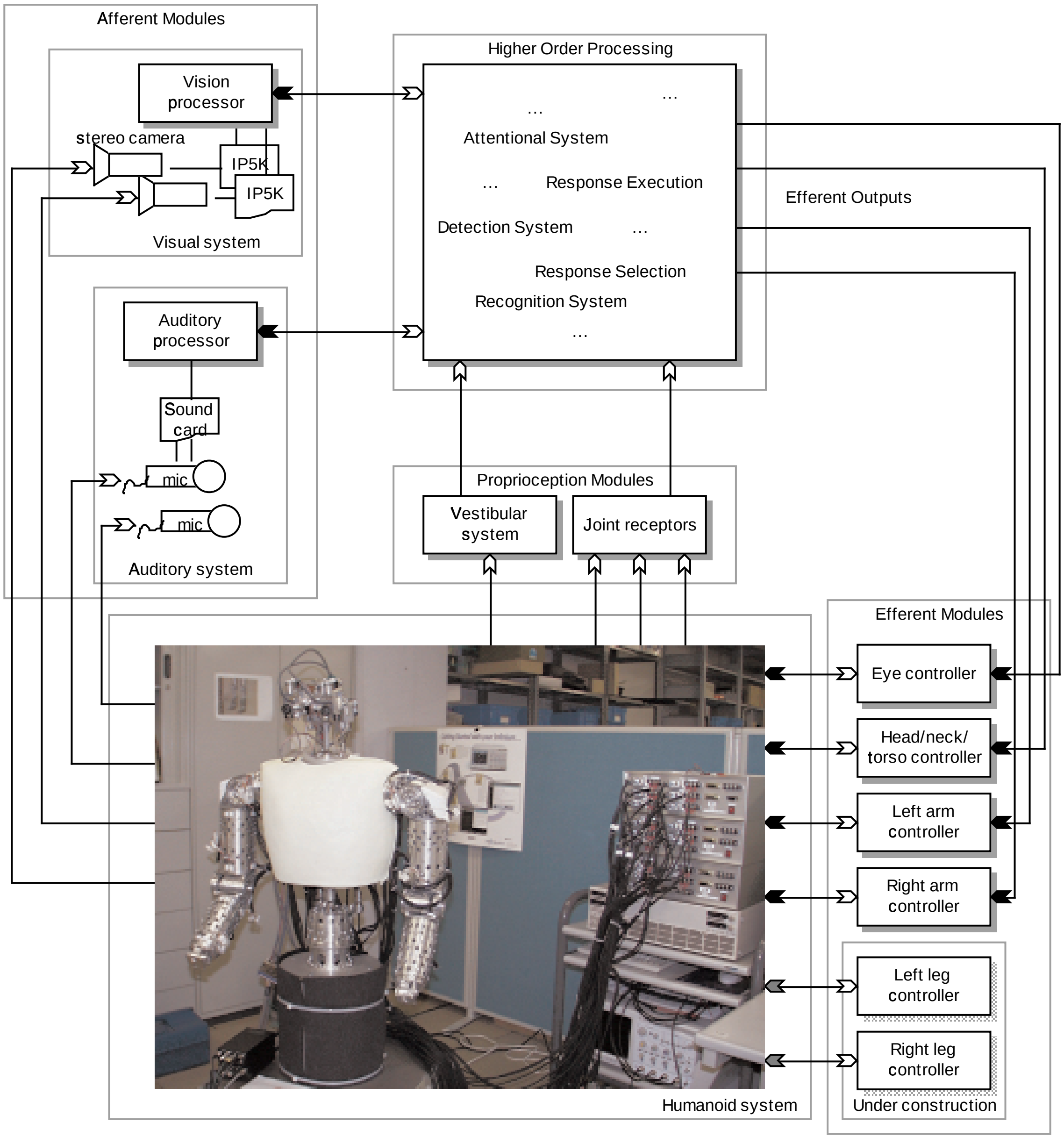
In contrast, to the approaches that have been taken in other systems (see section 2). In our approach to system development, an emphasis is placed on total integration. As soon as any part of the system is available, it is integrated. Adopting a constant integration philosophy provides one of the key point in undertaking humanoid research, moving toward the development and understanding of a fully integrated system such as a human. It is apparent that in day to day situations humans are governed by a number of integrated factors influencing outcomes of their activities. Our hypothesis in approaching development in this manner is that, by moving toward full integration from the start enhancement of the system will as whole be more meaningful. From the view of components, this will allow components to be trial and evaluated in a much more coherent fashion. And also, the arrival of solutions will be far more consistent with the system as a whole.

A positive effect of this is constant integration, no components of the system will be left un-utilised at any stage of development. Preventing it from being disregarded at a later time due to technical advancement. This paper further describe our initial attempt at this merger.

4 System configuration – ETL-Humanoid

In our current phase of development, the upper body of our humanoid robot has been completed: two arms, head and torso, as shown in Figure 3. The system has been designed to be compact and light-in-weight while retaining similar size and proportion as an average person. The final system will be 160cm in height and 60kg in weight. The upper body of our system provides 24 degrees of freedom: 12 d-o-f for the arms, 3 d-o-f for the torso, 3 d-o-f for head/neck and 6 d-o-f for the eyes. Other parts of the body are still under construction. Further discussion of this system is given in [1, 34, 35]. The processing for our system is currently performed over a cluster of seven desktop PCs connected to the robot via external cables, see also [4].

We use one PC equipped with two Hitachi IP5005 vision processor for the vision processing of our system. Another PC equipped with a SoundBlaster$^{\text{TM}}$ card is used for the auditory processing. One PC is also used for the vestibular system. Currently, the connections between these PCs is via an isolated 100Mbs Ethernet network. In the next generation of our system, we plan to embed a large portion of these components into a high speed interconnected computational network (see Section 4.1).



Limb	Degrees of freedom(DOF)
Arms	6×2
Torso	3
Head and neck	3
Eyes	3×2

Fig. 3. ETL-Humanoid (codename – 'JACK'): In its current form, the upper portion of the body has been completed (head, eyes, arms and torso). Each of the processing sub-systems are illustrated. From left to right: Afferent Modules, processes the external inputs from the outside world; through to the Higher Order Processing, providing such sub-systems as Response Selection, etc. These processes influence the internal sub-systems as well as the overall response of the system. The effect produced by the system is via the Efferent Modules.

The configuration we have established, has been deliberately designed for concurrent actuation and processing sensory inputs. Thus reducing the processing cycle of each sub-system, and further allowing the possibility of real-time and rich interaction to take place.

As shown in Figure 3, we have divided the controls of our humanoid system into separate limbs. Each arm having it's own controller, with the torso and the head sharing one controller. The motors control of the body limbs have been made possible via a set of embeddable in-bodied LAN networks, as shown in Figure 3. Each controller manages one LAN, controlling one limb, see for full technical detail of this in-bodied LAN system [34, 36, 35].

In the current version of the LAN, up to 8 nodes can be connected. Each node is responsible for providing low-level feedback information; motor position, encoder, temperature and motor current feedback. Each node also has facilities for two additional A/D channels, allowing further sensory inputs. All nodes can be made to operate simultaneously. Although for testing purposes, currently the nodes are off-board, physically they have been kept to a bare minimum, allowing them to be small and compact in anticipation for future integration.

Each of the in-bodied LAN networks are interfaced via the usage of individual PCs, providing all necessary computation, such as current, velocity and position controls. The PC also acts as a joints server, allowing a transparent service to be provided to any of the networked PCs. Thus, allowing a wide number of control topology to be connected and trial easily.

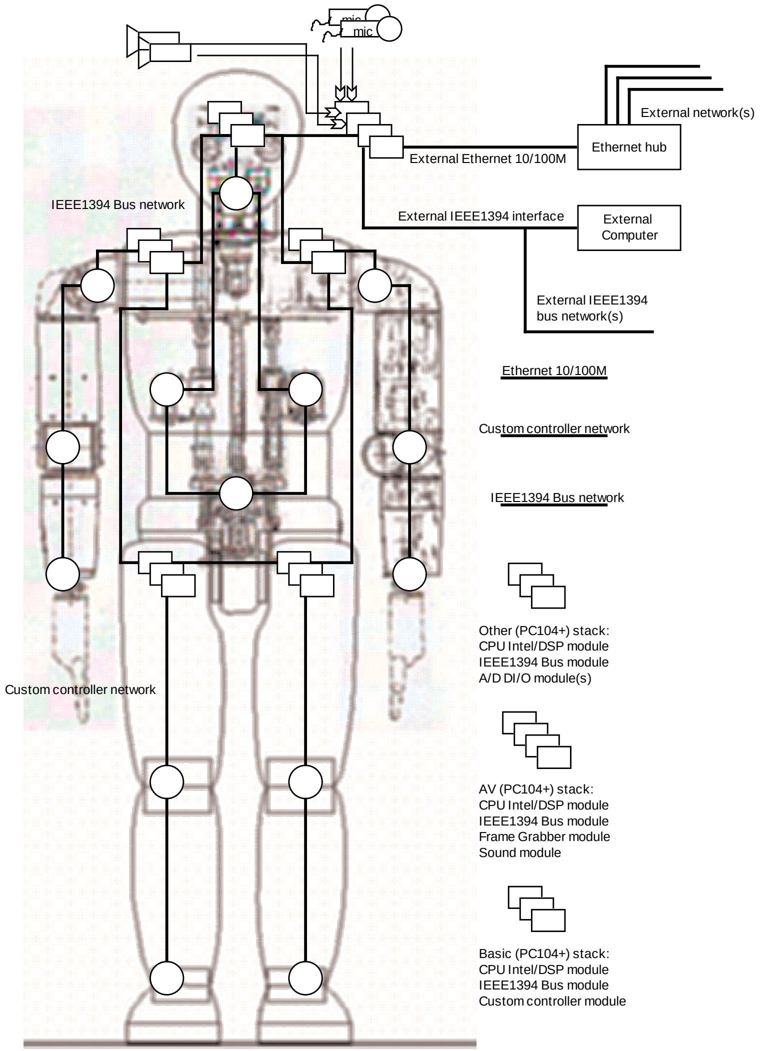


Fig. 4. New Hardware Topology: This new hardware architecture allows high speed data transfer, the design is based on industrial standards, rather than proprietary components. This will allow further expansion easier, allowing the benefit of further development to flow with the advancement of standard technical components.

4.1 Toward Autonomy – Next Generation of Computing

Another goal which should be considered for a humanoid system, is the attribute of "autonomy"[7]. *Autonomy* or *autonomous* is often described as the meaning of governing of one's self — self-governing. Autonomy plays a key role in the way in which we interact with the world, and the way in which problems are to be solved. Therefore, we believe that we should move in the direction of attributing a humanoid system with autonomy.

The motivation to provide a fully autonomous humanoid system has urged us to come up with a new internal design of our computational system, that can provide enough computing power for the next generation of our system. Figure 4 provides the current design of a new embeddable computing system. The requirement for this system have been to allow compactness, expendability, flexibility for future usage. Due to the speciality of the original custom control LAN, it is to be retained. However, the computing system for the communication with and within these LAN will be moved on-board, embedded inside the humanoid. Thus, this will bring us closer to a truly autonomous humanoid robot system.

5 Motor Control – Humanoid Motion

The design of our current motor control system provides the flexibility of control of motors at various levels: position (angular), velocity and current (force). These schemes run with a conventional **P**roportional **I**ntegral controller. The controls of each joint can be commanded via any of the above schemes, in a flexible way. The motor can be controlled in a mixed fashion. For instance, in the current experiment, the motion of the arm is driven at both velocity and current level. The motion of the arm is controlled via velocity, but once no motion is required, the arm is commanded to fall into a zero current loop. Hence, allowing the arm to be free and compliant for physical interaction, and human-like ballistic motion can also be achieved.

The motion produced by our humanoid system closely resembles the motion of human movement – ballistic like motion. The underlying control is based on a human movement model, physically it is controlled via a mix of velocity and force control. The velocity of each motor is given by Equation (1).

$$V_j(t) = V_{max} \frac{log_{10}(U_i(t) + 1.0)}{2.0} \tag{1}$$

where j is the motor joint number. i is the index for each i_{th} output. $U_i(t)$ is the i_{th} output at instant t given by the integrator. $V_j(t)$ is the velocity at j_{th} motor joint.

Self-regulating motion is achieved through the monitoring of the encoder at each joint. The joint limits are set in two ways, *a priori* at the start, and through physical interaction. Physical interaction, is done by taking advantage of the compliance of the system while it is not in motion. A person may physically move the robot, the system monitors this movement. The new limit of the joints is determined by the upper most position reached.

5.1 Physical Interaction

Various kinds of human like responses have been included into our system, for instance the reflex of sudden withdraw of a limb when excessive force is being asserted. The same response will also occur when the limit of a limb has been reached. Due to the light-weight design of our system, the complianceness during physical handling of the system was easily achieved through a simple zero-current control loop at the joints.

To demonstrate the physical characteristic of our system, physical interaction experiment was performed. Figure 5 presents our humanoid physically interacting with a person, the system is being man handle by the person in a compliant manner. The torso motion of our system has also been shown here, the configuration in our design closely matched the torso of humans. The same design has also been applied to the neck of our humanoid. The kinematics and dynamics of our torso and neck is similar to the actuated manipulator described in [37, 38]. A full description of the configuration of our system is presented in [35].

---

[7] **autonomous** *adj.* **1.** *Govt.* **a.** self-governing; independent; subject to its own laws only. **b.** pertaining to an autonomy. **2.** independent; self-contained; self-sufficient; self-governing. **3.** *Philos.* containing its own guiding principles. **4.** *Biol.* existing as an independent organism and not as a mere form or state of development of an organism. **5.** *Bot.* spontaneous. — ©*The Macquaire Dictionary.*

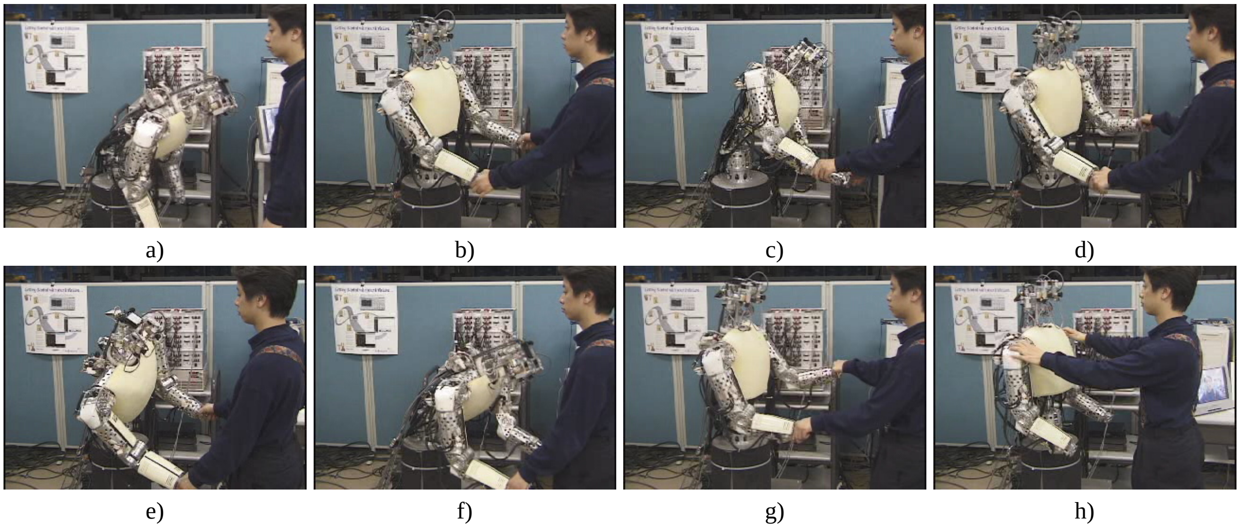

a) b) c) d) e) f) g) h)

Fig. 5. Physical Interaction: Our humanoid system is light in weight, and has a similar range of motion as a human. This example shows our system physically being man handled by a person. The range of the torso is especially depicted. a) Torso fully extended forward, b) Upright, c) Full extension to the left, d) Physical handling of the robot, e) Full extension to the right, f) full extension to the front, g) Light-weight and physical compliance, h) Physically guiding the system.


a) b)

Fig. 6. Visual cues: a) Motion Cue: Detection of a person moving, b) Contextual cue: Human Upper Body Detection

5.2 Coordinated motion

Further experiments of physical interaction was also realised through the use of coupled neural oscillator, coordinated arms and torso motion was achieved by our system. A similar experiment has also been reported by [16]. Exploiting the physical coupling of the two arms with the torso coordinated motion was produced. For example, synchronised motion of both arms swinging up and down was realised through the forward and backward rocking motion of the torso. And an alternating arm motion (similar to the motion of a runner) was realised through the twisting motion of the torso. Physically the system retained the property of complianceness during this experiment, while allowing itself to be physically handled.

5.3 Coupled Motion – Torso/Neck

The motion coupling of the eyes and neck has been widely studied and determined to be useful, e.g. the Vestibular-Ocular (VOR) and Opto-kinetic Reflexes (OKR) [39–41]. However, little study has been made to the coupling of motion of other parts of the human body. From our own observation and studies, we hypothesis that the motion of the head and the torso is based on a delay between the two. Thus, when the neck moves the torso will follow with smaller motion but with a delay, and the neck will also compensate this motion with an equivalence counter motion. The motion will stop once the neck and torso have reached a neutral position of comfort. The end effect of this coupled motion, yielded a smoother and greater range of motion for the whole system, for example smoother tracking of a target.

6 Seeing

The human visual receptor is the single most developed and heavily utilised part of our sensory system, it provides a wide range of information which can be used to determine various detail within an environment. For example, detection of people and their corresponding motions. Currently, this part of our system provides the following capabilities: motion detection for the initiation of attention in a scene (see Figure 6 a), skin detection is currently being used for the detection of a person (see Figure 6 b). By using these two components our humanoid is able to visually track the movement of a person. Disparity and depth perception have also been included into our system. Figure 7 provides a summary of the vision processing stages of our system. The processing is performed at real-time, Table 1 provides the processing cycle time of these key stages of the vision processes. The processing is well within the 30Hz video cycle, with a total processing time of 16.951ms. Allowing interaction to be performed at real-time.

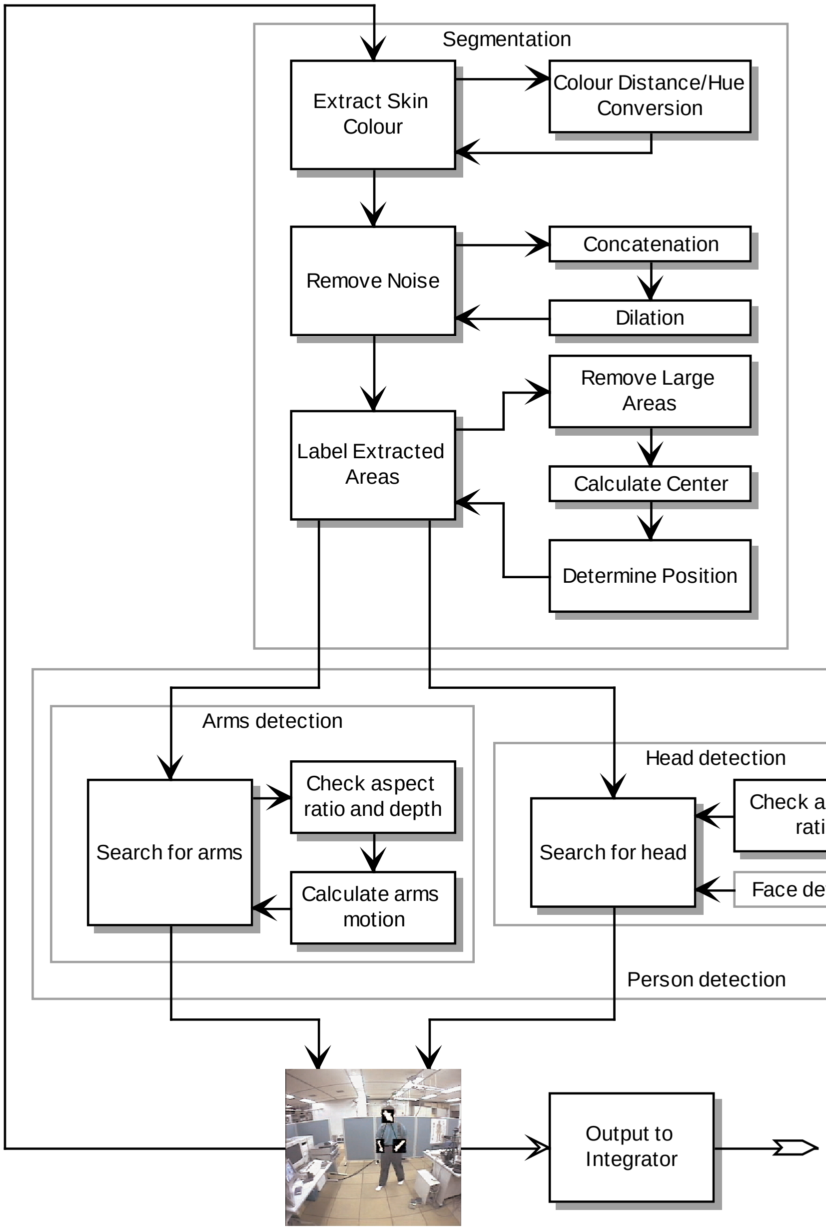


Fig. 7. Vision processing: Important stages of the vision processing performed by our system is shown. The total processing time required is only 16.951ms, see Table 1.

Processes	Time taken (ms)
Initialisation	0.101
Skin Colour Extraction	5.135
Noise filtering and Labelling	4.608
Calculate Labelled Areas	6.608
Search Head	0.495
Search Arms	0.004
Total	16.951

Table 1. Vision processing execution cycle time

The basic information used by our system is as follows: the position of the head in the scene and the motion of the arms. Figure 7 shows the stages of our vision processing. The detection of an upper body of a person from the environment is based on segmentation of skin colour. The process of skin colour detection is based on Colour distance and Hue extraction. The overall process is as follows:

1. segmentation is made between the environment and the person.
2. select head, based on a set of attributes (e.g. colour, aspect ratio, depth etc.)
3. extract the arms from the scene based on a set of attributes (e.g. colour, aspect ratio, depth, etc.)

Position tracking of these extracted features is based on error filtering of the past locations. Once the segmentation and extraction have been performed, the motion is then determined. The motion is calculated by taking the derivative of the visual information of the arms being tracked. The derivatives calculated for the arms are then pass on to be used for motor control. The position of the head in the scene is used to orient the robot toward the person.

Each of these positions taken from both of the video cameras are sent, as action vectors to the integrator of the system. In addition, an associated activation signal for each vector is also forwarded. The activation is determined by an accumulative counter of the existence of the tracked item.

7 Hearing – Spatial hearing

From the inspiring notion that at the earliest stages of life, a baby is able to detect the direction of a sound source without the use of sight [43]. In terms of humanoid interaction, sound can play a great role in providing a number of cues, such as initiating focus of a moving target before a sighting can be made.

The spatial hearing of our humanoid is based on the interaural difference processing proposed by Psychophysical studies of human hearing [44]. Figure 8 shows the steps involved in our initial spatial hearing processes of our system. The processing steps start by simultaneously sampling the left and right channels, a Fast Fourier Transform (FFT) is then perform on each channel producing an energy spectrum for each channel. The difference between the spectra of the two channels are then used to determine the direction of the sound source.

In our current examination of auditory response for our humanoid, we performed the experiment of left and right discrimination, by allowing our system to center on a sound source by minimising the differences of the two channels. Figure 9 shows an example of this process after connecting it to the rest of the system. The robot is allowed to servo to a sound source (a “baby rattle”), simply by moving to the direction with the greater spectrum. Thus, this simple interaction allows our humanoid to track a sound source without the need of any complex models.

For integration purposes, the output of the sound source is sent to the integrator as an action vector, indicating its magnitude and direction. The action vector carries the accumulative counter value of the existence of the sound source, used as the activation signal.

8 Vestibular system

The vestibular system has commonly been associated with the Vestibular-Ocular (VOR) and Opto-kinetic Reflexes (OKR), allowing the stabilisation of the eyes while the head moves. A number of researchers also embrace such mechanism, mostly for the determination of ego-motion for an active head/eyes system [45, 2]. We also believe this is an essential mechanism that need to be incorporated into our system, however we believe the vestibular system will service additional purposes for our humanoid system, such as:

- the stabilisation of the eyes while the whole body is in motion, as well as the head when in motion (as suggested by [45, 2]),

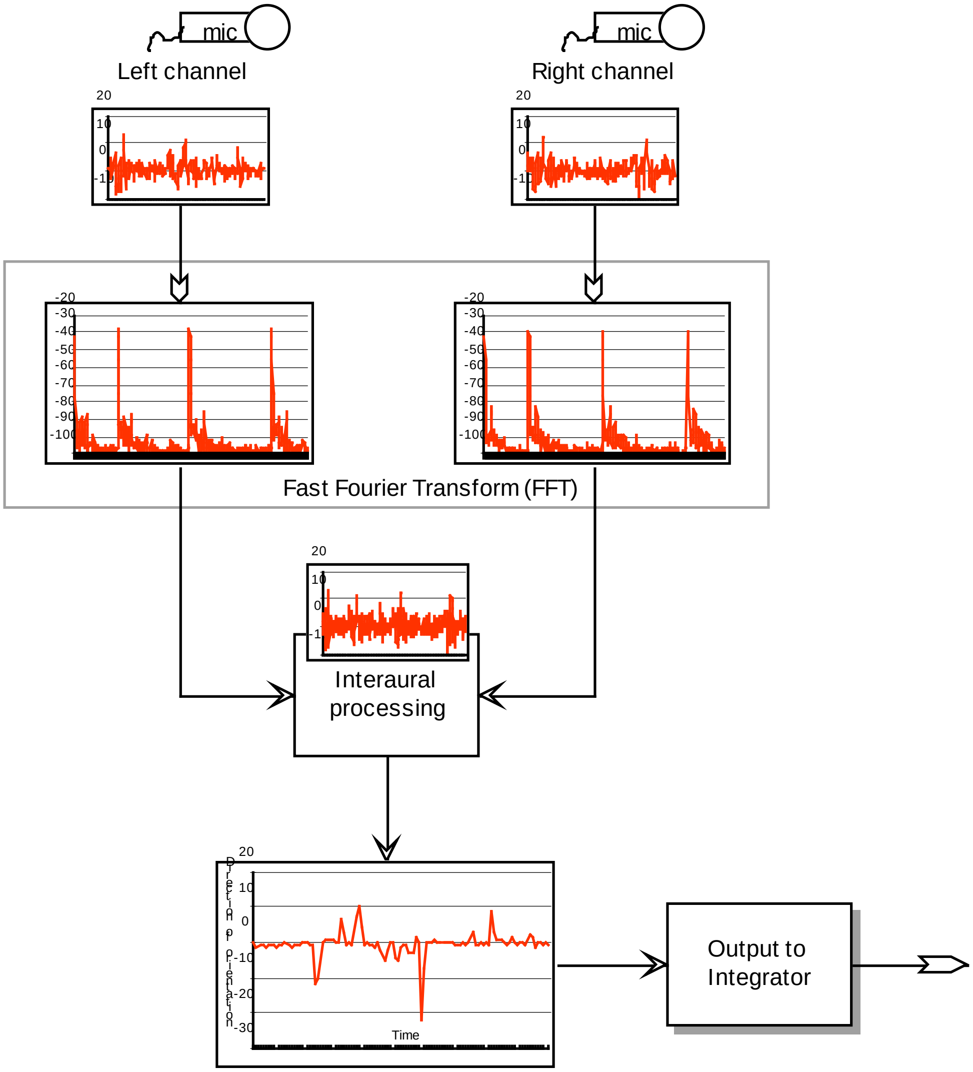


Fig. 8. Auditory processing: Stages of the auditory processing performed by our system is shown. First, simultaneous sampling of the left and right channels is made, A Fast Fourier Transform (FFT) is then perform on each channel. The difference between the two channel is then used as the output for spatial hearing response.

- will provide information for balance and locomotion for our humanoid system [46, 18, 41], and
- providing additional information for the auditory spatial orientation, without the presumption of a flex posture.

We are currently using an 3-axes accelerometer and 3 gyroscopes for x, y, and z-axis. The gyroscopes provides the rotational information of the system. Whereas the accelerometer provides the system with 3 dimensional translational information of the system. The vestibular system is controlled via one PC equipped with a 32kHz A/D module, currently the information is sent to the rest of the system via 100Mb ethernet network.

9 Integration – Putting it together

With the attributes we have outlined in Section 1, we arrived to a way in which competition between sensory-cues should be used in order to yield a interactive system. This idea has been suggested by a number of interdisciplinary departments in human science [39, 47, 10, 2, 41]. The integration process allows each of the components in the system to run in parallel, concurrently competing or cooperating for the each response, as shown in Figures 3 and 10.

Our approach require only two features from the above components:

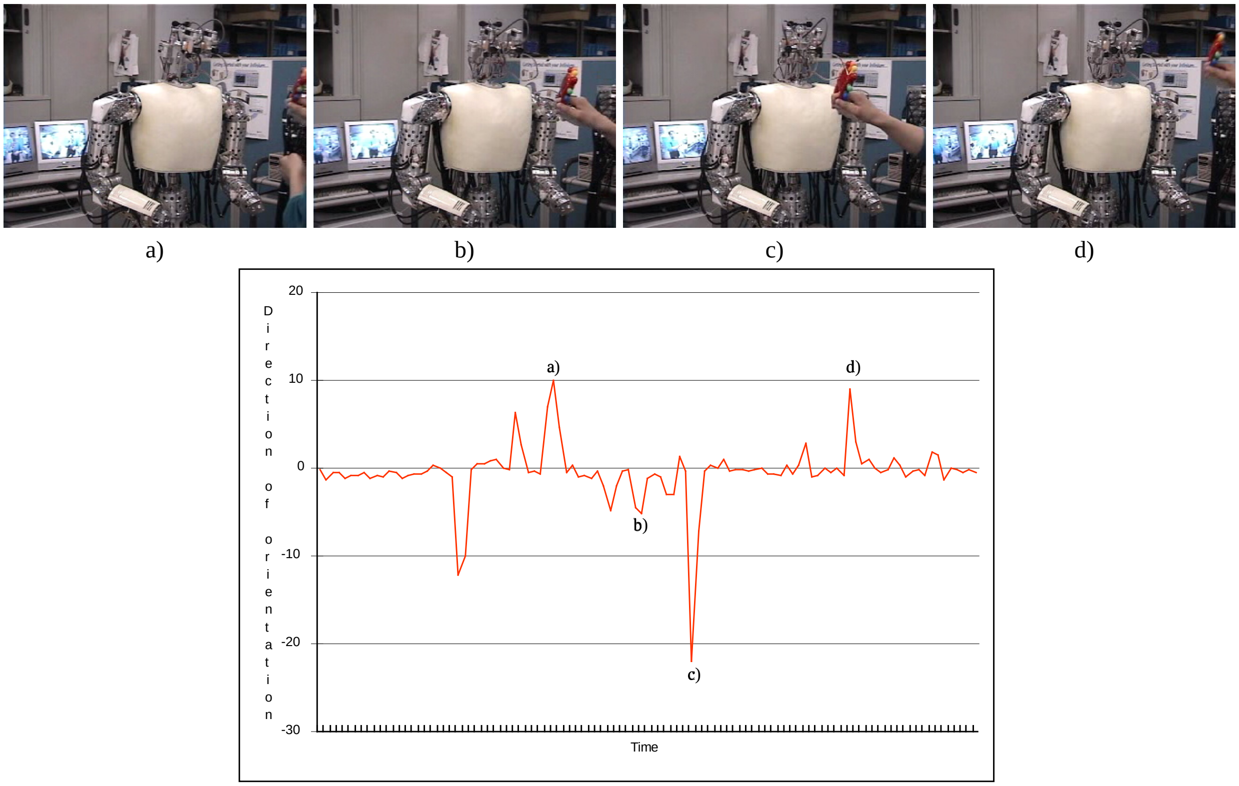

a) b) c) d)

Fig. 9. Auditory Tracking: Our system is able to perform auditory servoing, these diagrams show our humanoid servoing toward the sound source of a “baby rattle”. Visual tracking was deliberately set with a low gain, in order to allow auditory tracking to take over the system.

action vector providing the magnitude and direction of a given input. e.g. a vector can be used to represent the relative action of the arm, positive for up, negative for down. Its speed being represented by its magnitude.

activation potential provides temporal duration of its associated **action vector**, representing the degree presence or the absence of a particular stimulus/cue given by the vector, determining its reliability, i.e. confidence.

Inspired by the generality of a biological neural system. The key and central idea of this *Basic integrator* must be applicable across many levels, at both the sensory level and the actuation level, as a neuron would be in a biological system.

For our current investigation we introduce Equation (2), as our *Basic integrator* for use throughout our system. As discussed, the important properties of this integrator is that it is model-free, it can be used at many levels, from sensory processing to the final output of the system.

$$U_i(t) = \frac{\sum_k \alpha_k(t) a_k(t) v_k(t)}{\sum_k a_k(t)} \tag{2}$$

where k is the index for each relevant input. i is the index for each i_{th} output. $U_i(t)$ is the i_{th} output vector at instant t. $a_k(t)$ is the activation potential of the k_{th} input at instant t. v_k is the k_{th} input vector. $\alpha_k(t)$ is the parameter which allows the alteration of the strength of a particular input.

Currently we are exploring the parameter $\alpha_k(t)$, which was introduced for the alteration of the overall system behaviour. This is inspired by the daily interaction of a person. Influences from sensory systems tend to be alter based on some selective occasions, depending on the *mood* of an individual at that particular time. Many other factors also comes into play, a well know phenomenon exhibited by a person, is the decay in response to a continuous stimulus over some duration of time [39, 41]. It is believed that emotional factors do come into play, in determining the responses of our daily lives [47, 48]. Therefore, we believe that this parameter will allow us to further explore some of these issues.

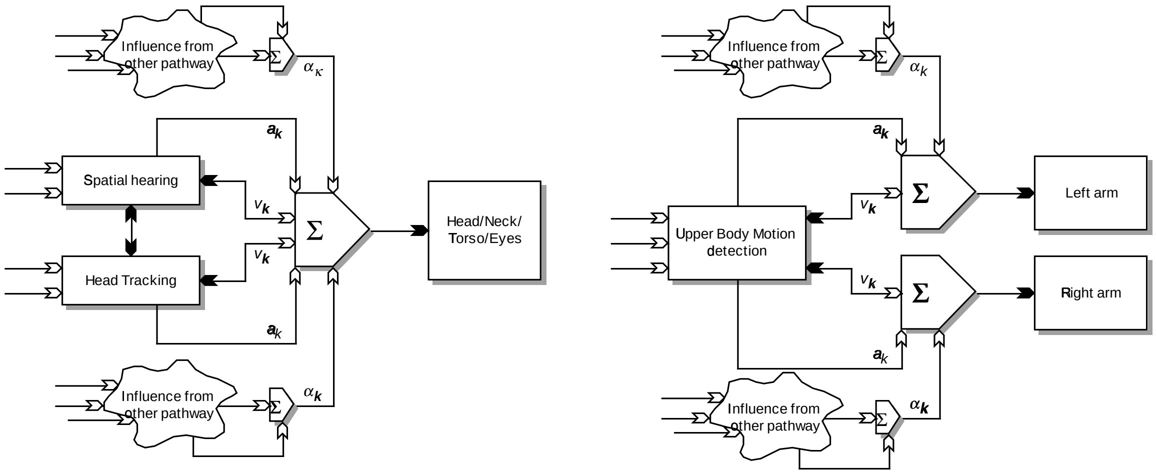


Fig. 10. Sensory pathways: this diagram shows conceptually the pathways between perceptions and actions. These simplified pathways are contextually connected.

9.1 Sensory Pathways

At this stage we have divided a simple sensory pathway for our system, the overall interaction of the system is produced through the external sensors and the flow of these pathways. A simplified representation of these pathways are given in Figure 10. As discussed, these pathways are be utilised in a parallel manner allowing the overall response of the system to be yielded through the interaction with the outside world.

10 Continuous Humanoid Interaction – Mimicking

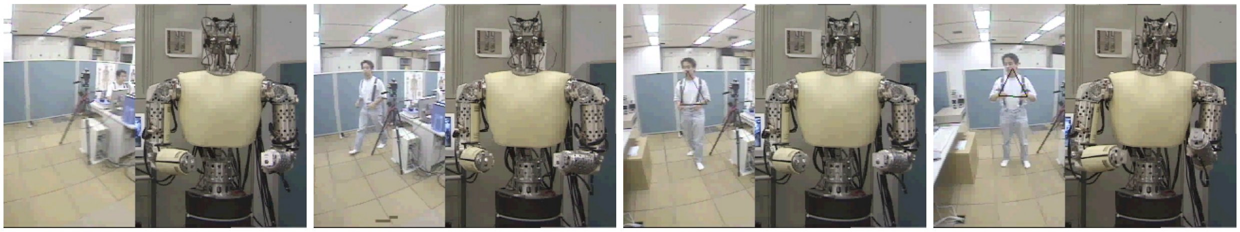
Figure 11 shows the experimental results of our system. The humanoid robot first orientates toward a person, tracking the person. Then once the upper body has been fully sighted, i.e. head and two arms have been detected, the humanoid robot mimics the upper body motion of the person. When the upper body can not be fully sighted, the robot continuously tracks the person. When system loses sight of the person, then an auditory response to a sound made by the person regains the attention of the robot. The system continues to track and mimic in a robust and continuous manner. These snapshots were taken from an experiment lasting over 4 minutes. Some experiments and demonstrations of this system have lasted continuously over 20–30 minutes. For details of the integration of this system see [4]. Figure 12, shows the sensory inputs traces of this experiment. Figure 12 a) shows the tracked inputs of the upper body of the person. Figures 12 b) and c) shows the motion of the left and right arms respectively. The transition between various interactions are also shown in the traces. That is, start “tracking a person” ⇒ “lost” track of the person ⇒ “tracking a person” again ⇒ “lost” track of the person again. These transitions further show the continuum of the integrated system we have derived.

11 Further work

Further work will involve higher order abilities, such as the incorporation of memory systems: associative memory, motor memory and short/long-term memory. Furthermore learning scheme(s) will need to be investigated: on-line learning, as well as off-line learning. As mentioned emotional factors for humanoid systems will also need to be considered. Connecting together all of these with the rest of our system, will attribute to the majority of our future work.

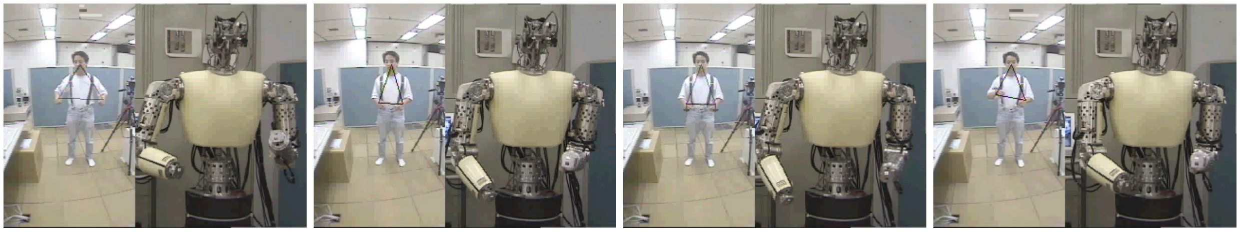
12 Summary and Conclusion

In this paper we presented the current version of our integrated humanoid system, a 24-d-o-f upper body humanoid robot system. The sensory systems available in our humanoid was also presented. Sensory systems include: an auditory system, a real-time stereo vision system, a proprioceptive system and a vestibular system. Our 24-d-o-f humanoid system is actuated by a set of high performance motors, and can be controlled in a flexible manner.

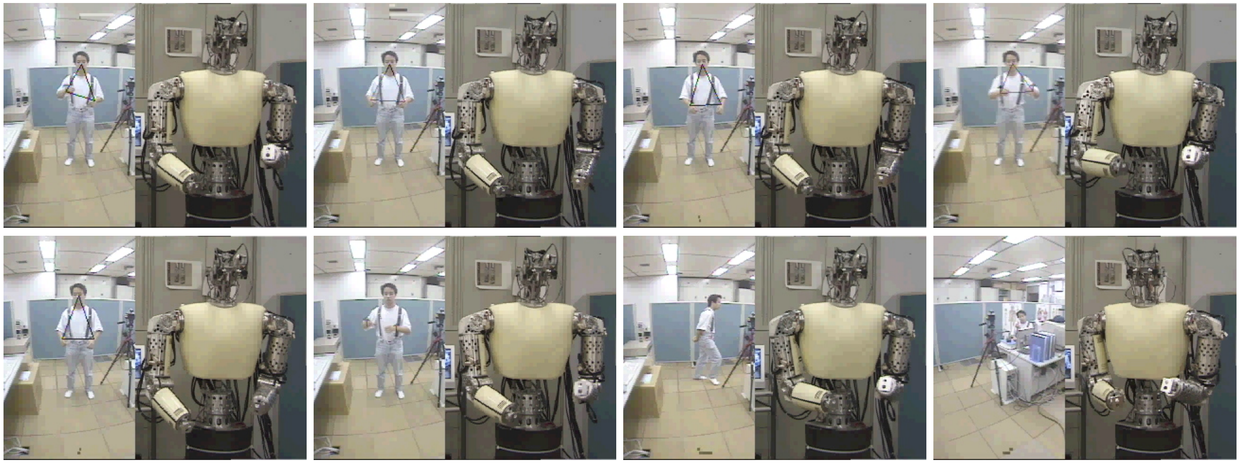


Detect and track a person

Notice upper body

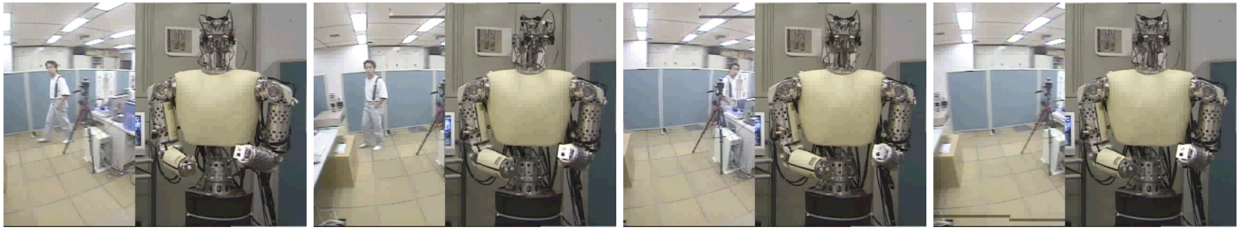


Start Mimicking



Stop Mimicking

Track person

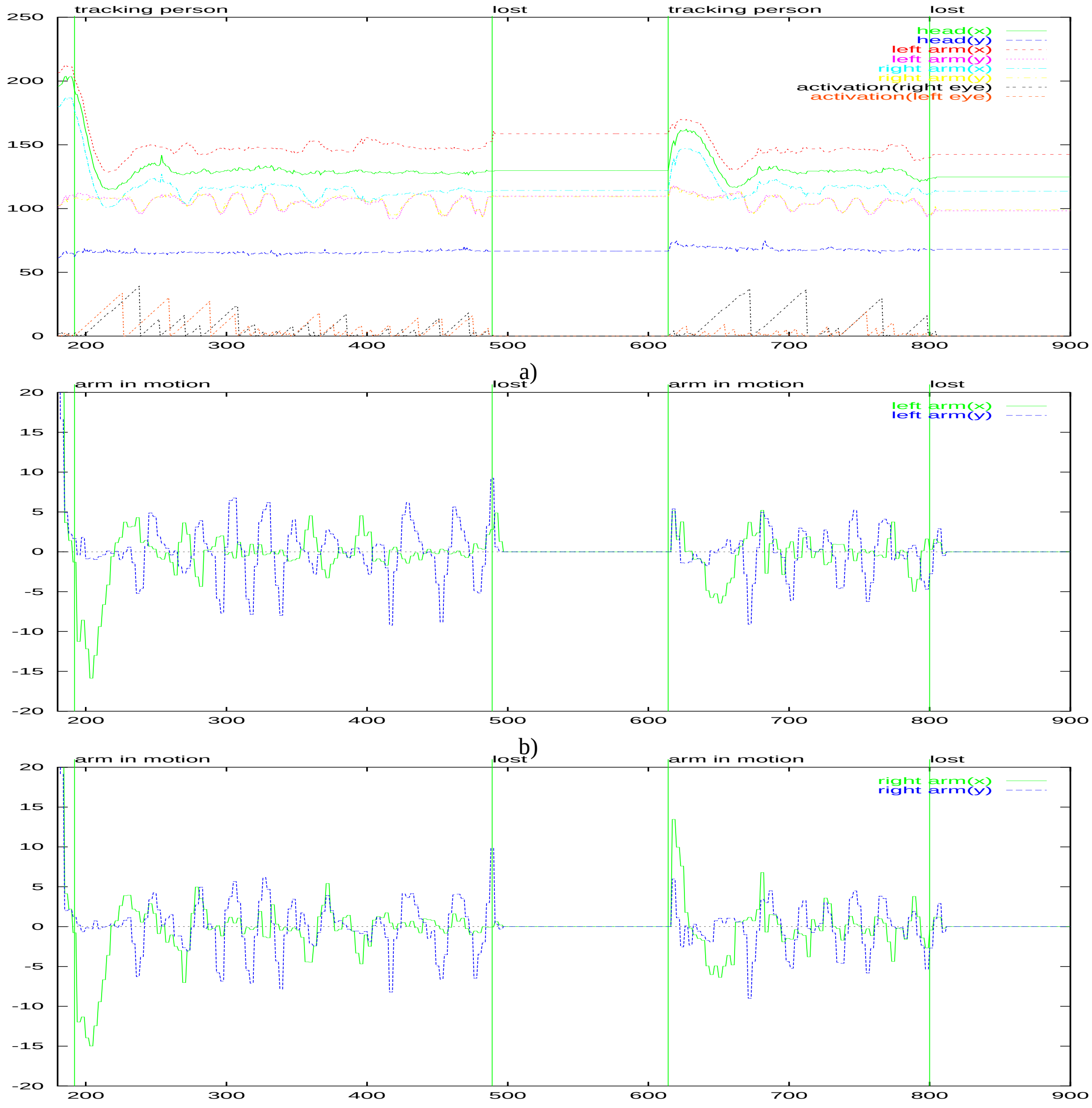


Person, out of sight

Regain attention via sound

Restart Mimicking

Fig. 11. Humanoid Interaction: Visual Attention and Tracking, Auditory Attention and Tracking, Mimicking of a person. This experiment shows the detection of a person, while estimating the upper body motion of the person. Detection of sound, allows our system to determine the spatial orientation of the sound source – spatial hearing. The output response of this system, allows our humanoid to orientate/servo toward the person by keeping track of the person. The humanoid produces the same upper body motion as the person, thus mimicking the person. The system performs the visual processing at 30Hz in stereo, auditory sampling is performed at 44kHz.



a)

b)

c)

Fig. 12. Mimicking Traces: a) shows the inputs movements of all parts being tracked, b) shows the left arm motion of the humanoid, c) shows the right arm motion of the system. The four segments on each graph shows the continuos form of interaction, person entering and exiting the scene and returning to interaction again.

The distributed configuration of our hardware architecture for the support of such an integrated system was also presented. The system architecture provides a way in which concurrent computations can be realised, allowing processes to be interconnected and interacted with each other in a responsive manner.

The current progress of our humanoid robot system has been briefly discussed. Details of the design and motivation of our current and future system was given as part of our presentation. The various approaches to humanoid research was also discussed, their benefits and drawbacks were highlighted. Our view of an integrated humanoid system was taken into perspective. This perspective for the integration of a complete humanoid system was asserted throughout the work presented.

Our approach to integration embraced a number of factors:

- A task model was not taken as the prime purpose of development, thus the system is not bound to any preprogrammed task.
- No switching between modules or sub-systems was allowed, ensuring that all data flowing through the system are used in influencing the outcome of the system.
- Each sub-system does not assume to have full control of the system, all sub-systems are integrated as a whole.

- In producing a system which has a rich and diverse set of sensory capabilities for interaction.
- It has enforced us to produce a system that can produce a greater range of responses.
- Allowing us to take a wider range of view in solving problems.
- Allows a way in which redundancy can be exploited. Thus, providing a robust system.

We showed by using a simple competitive mechanism, that were able to produce an integrated humanoid interactive system that is adaptive due to the redundancy, and flexible nature of this mechanism. By taking an integrated view, we were able to produce a system that gained adaptivity, redundancy and flexibility in a robust system.

A number of experiments were also presented in this paper, those include: physical interaction, our humanoid allowing itself to be physically handle by a person. Auditory interaction, that allows our humanoid to servo toward a sound source present in its environment. A combined and integrated experiment of our humanoid mimicking the upper body of a person, while visually tracking the person was also presented. The experiment also allows physical contact to be made by the robot and the person, and auditory response was also available at the same time. The processing for these interaction were done continuously, without discarding any data inputs presented to the system. Human ballistic like motion was made possible through the high performance actuator motors available within our humanoid system.

We believe by taking an integrated view from the onset, into the investigation of humanoid research, that it will allow us to explore further possibilities for the field of humanoid robotics, that have yet to be fulfilled.

Acknowledgments

We wish to acknowledge kindly the support of the Center of Excellence (COE) program and the Brain Science program funded by the Science and Technology Agency (STA) of JAPAN and the ElectroTechnical Laboratory (ETL), Tsukuba, JAPAN. We wish also to thanks Youichi Ishiwata, Takashi Kimura, Hiroshi Ohkawa and Olivier Stasse for their assistance with this project.

References

1. Yasuo Kuniyoshi and Akihiko Nagakubo. Humanoid Interaction Approach: Exploring Meaningful Order in Complex Interactions. In *Proceedings of the International Conference on Complex Systems*, 1997.
2. Rodney A. Brooks, Cynthia Breazeal, Matthew Marjanović, Brian Scassellati, and Matthew M. Williamson. The Cog Project: Building a Humanoid Robot. In *IARP First International Workshop on Humanoid and Human Friendly Robotics*, pages I–1, Tsukuba, Japan, October 26–27 1998.
3. S. Hashimoto *et. al.* Humanoid Robots in Waseda University – Hadaly-2 and WABIAN. In *IARP First International Workshop on Humanoid and Human Friendly Robotics*, pages I–2, Tsukuba, Japan, October 26–27 1998.
4. Gordon Cheng and Yasuo Kuniyoshi. Complex Continuous Meaningful Humanoid Interaction: A Multi Sensory-Cue Based Approach. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 3, pages 2235–2242, San Francisco, U.S.A., April 2000.
5. Masanao Toda. *Man, Robot and Society*. Martinus Nijhoff Publishing, 1982.
6. Andy Clark. *Being There: Putting Brain, Body, and World Together Again*. MIT Press, 1997.
7. R. C. Arkin. *Behaviour-Based Robotics*. MIT Press, May 1998.
8. Rolf Pfeifer and Christian Scheier. *Understanding Intelligence*. The MIT Press, 1999.
9. Richard Gregory. *Mind in Science*. Combridge University Press, 1981.
10. Daniel C. Dennett. *Kinds of Minds*. Science Masters series. Basic Books, 1996.
11. K. Nakadai, T. Lourens, H. G. Okuno, and H. Kitano. Active audition for humanoid. In *Proceedings of 17th National Conference on Artifcial Intelligence (AAAI-2000)*. AAAI, 2000. To be presented.
12. M. Rucci, J. Wray, and G. M. Edelman. Robust localization of auditory and visual targets in a robotic barn owl. *Robotics and Autonomous Systems*, 30:181–193, 2000.
13. I. Kato. Wabot-2: Autonomous Robot with Dexterous Finger-Arm. In *Proceedings of IEEE Robotics and Automation*, volume 5 of *2*, 1987.
14. Rodney A. Brooks and Lynn Andrea Stein. Building Brain for Bodies. Technical Report A.I. Memo No. 1439, Massachusetts Institute of Technology, August 1991.
15. Rodney A. Brooks, Cynthia Breazeal (Ferrell), Robert Irie, Charles C. Kemp, Matthew Marjanović, Brian Scassellati, and Matthew M. Williamson. Alternative Essences of Intelligence. In *National Conference on Artificial Intelligence 1998-1999*. American Association for Articial Intelligence, 1998.
16. M. M. Williamson. Rhythmic robot control using oscillators. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'98)*, 1998.

17. Brian Scassellati. Eye Finding via Face Detection for a Foveated, Active Vision System. In *National Conference on Artificial Intelligence 1998-1999.* American Association for Articial Intelligence, 1998.
18. Kazuo Hirai, Masato Hirose, Yuji Haikawa, and Toru Takenaka. The Development of Honda Humanoid Robot. In *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, pages 1321–1326, Leuven, Belgium, May 1998.
19. Hirochika Inoue. A Platform-based Humanoid Robotics Project. In *IARP First International Workshop on Humanoid and Human Friendly Robotics*, pages I–1, Tsukuba, Japan, October 26–27 1998.
20. Hirochika Inoue *et. al.* HRP: Humanoid Robotics Project. In *Humanoids2000*, Boston, USA, September 7–8 2000. In Review.
21. Young-Jo Cho, Bum-Jae You, Sang-Rok Oh, Mun Sang Kim, ye Suan Hong, and Chong Won Lee. A Design of a Compact Network-Based Controller Mounted on a Humanoid Robot. In *IARP First International Workshop on Humanoid and Human Friendly Robotics*, pages VII–1, Tsukuba, Japan, October 26–27 1998.
22. Sooyong Lee, Munsang Kim, and Chong-Won Lee. Design of a New Master Device for controlling a Humanoid Robot, *Centaur*. In *Proceedings of the Second International Symposium on Humanoid Robots*, pages 161–167, October 8–9 1999.
23. Munsang Kim, Sungchul Kang, and Sooyong Lee. Design and Control of a Humanoid Robot CENTAUR. In *Proceedings of the Second International Symposium on Humanoid Robots*, pages 96–101, October 8–9 1999.
24. Atsuo Takanishi, Tadao Matsuno, and Isao Kato. Development of an Anthropomorphic Head-Eye Robot with To Eyes-Coordinated Head-Eye Motion and Pursuing Motion in the Depth Direction. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'97)*, 1997.
25. Stefan Schaal. Is Imitation Learning the Way to Humanoid Robots? *Trends in Cognitive Sciences*, 1999.
26. G. Tevatia and S. Schaal. Inverse Kinematics for Humanoid Robots. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 3, pages 294–299, San Francisco, U.S.A., April 2000.
27. Rainer Bischoff. Advances in the development of the humanoid service robot HERMES. In *Proceedings of the Second International Conderence on Field and Service Robotics*, pages 156–161, Pittsburgh, PA, August 1999.
28. T. Asfour, K. Berns, and R. Dillmann. The Humanoid Robot ARMAR. In *Proceedings of the Second International Symposium on Humanoid Robots*, pages 174–180, October 8–9 1999.
29. Rainer Bischoff. Natural Communication and Interaction with Humanoid Robots. In *Proceedings of the Second International Symposium on Humanoid Robots*, pages 121–128, October 8–9 1999.
30. K. Kawamura, D.M. Wilkes, T. Pack, M. Bishay, and J. Barile. Humanoids: Future Robots for Home and Factory. In *Proceedings of the First International Symposium on Humanoid Robots*, pages 53–62, October 30-31 1996.
31. Joseph A. Driscoll, Richard Alan Peters II, and Kyle R. Cave. A Visual Attention Network for a Humanoid Robot. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'98)*, Victoria, British Columbia, Canada, October 13–17 1998.
32. R. A. Peters II, D. M. Wilkes, D. M. Gaines, and K. Kawamura. A Software Agent Based Control System for Human-Robot Interaction. In *Proceedings of the Second International Symposium on Humanoid Robots*, page Late Paper, October 8–9 1999.
33. Atsushi Konno, Koichi Nagashima, Ryo Furukawa, Takuro Noda Koichi Nishiwaki, Masayuki Inaba, and Hirochika Inoue. Development of a Humanoid Robot *Saika*. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'97)*, volume 2, pages 805–810, 1997.
34. Yasuo Kuniyoshi and Akihiko Nagakubo. Humanoid As a Research Vehicle Into Flexible Complex Interaction. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'97)*, volume 2, pages 811–819, 1997.
35. Akihiko Nagakubo, Yasuo Kuniyoshi, and Gordon Cheng. Development of a High-Performance Upper-Body Humanoid System. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'00)*, 2000. To be presented.
36. Youichi Ishiwata and Akihiko Nagakubo. In-body LAN and external computer system of ETL-Humanoid. In *Robotics Society of Japan Annual Conference*, 1998. (In Japanese).
37. Kok-Meng Lee and Dharman K. Shah. Kinematic Analysis of a Three-Degrees-of-Freedom In-Parallel Actuated Manipulator. *IEEE Journal of Robotics and Automation*, 4(3):354–360, June 1988.
38. Kok-Meng Lee and Dharman K. Shah. Dynamic Analysis of a Three-Degrees-of-Freedom In-Parallel Actuated Manipulator. *IEEE Journal of Robotics and Automation*, 4(3):361–367, June 1988.
39. J. A. Scott Kelso, editor. *Human Motor Behavior: An Introduction.* Lawrence Erlbaum Associates, Publishers, 1982.
40. Vladimir M. Zatsiorsky. *Kinematics of Human Motion*. Human Kinetics, 1998.
41. Richard A. Schmidt and Timothy D. Lee. *Motor Control and Learning: A Behavioural Emphasis*. Human Kinetics, third edition, 1999.
42. Brian C. J. Moore, editor. *Hearing*. Academic Press, 1995.
43. D. Wesley Grantham. *Spatial Hearing and Related Phenomena*, chapter 9. In Moore [42], 1995.
44. Jens Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localization*. The MIT Press, revised edition, 1999. Second printing.
45. Francesco Panerai. *Integration of inertial and visual information in binocular vision systems*. PhD thesis, Department of Communication, Computers and Systems Science. University of Genova, Genova, Italy, 1998.
46. M. Vukobratović, B. Borovac, D. Surla, and D. Stokić. *Biped Locomotion: Dynamics, Stability, Control and Application*. Scientific Fundamentals of Robotics 7. Springer-Verlag, 1990.

47. Antonio R. Damasio. *Descartes' Error: Emotion, Reason and the Human Brain*. Avon Books, 1994.
48. Joseph LeDoux. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Touchstone, 1996.