To appear in IEEE Int. Conf. Intelligent Robots and Systems (IROS) 1997.

Humanoid As a Research Vehicle Into Flexible Complex Interaction

Yasuo Kuniyoshi and Akihiko Nagakubo
Humanoid Interaction Laboratory, Intelligent Systems Division,
Electrotechnical Laboratory, AIST, MITI, Japan

Abstract

This paper positions humanoid research as an approach to understanding and realizing complex real world interactions between a robot, an environment, and a human. As a first step towards extracting a common principle over the three term interactions, a concept of action oriented control has been investigated with a simulation example. The complex interaction view casts unique constraints on the design of a humanoid, such as whole body, smooth shape, non-functional-modular design. A brief description of ongoing design of ETL-Humanoid which conforms to the above constraints is presented.

1.Introduction

Following a couple of pioneering efforts ([1,2]), humanoid robot research has recently been started at several research groups including ours. Now that we have already seen some advanced prototype systems, it will not be justified any more for the humanoid researchers to keep claiming "we are just trying to build a human like robot". This paper presents our focus of research *using* a humanoid robot and shows how it constraines the humanoid design, which is currently under construction at our institute.

Our main focus is on the following points:

1. **Complex dynamical systems view.**
2. **Embodied cognition and emergent behavior.**
3. **Meaningful human-robot-environment interaction.**
4. **Open ended tasks and human like solutions.**
5. **Practical and versatile research platform**

The above points 1-4 are discussed further in the next section along with our current research directions. Section 3 presents one of our core ideas for dealing with complex interaction with an illustrating example. The top level specifications and the design decisions for our humanoid system are presented in section 4. They are the results of our best effort in finding optimal solutions and trade-offs between the above research points 1-4 along with the important pragmatic point 5, in light of the current state of the art of available component technology.

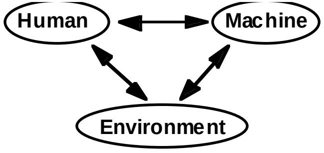

Fig. 1. Three term interaction.

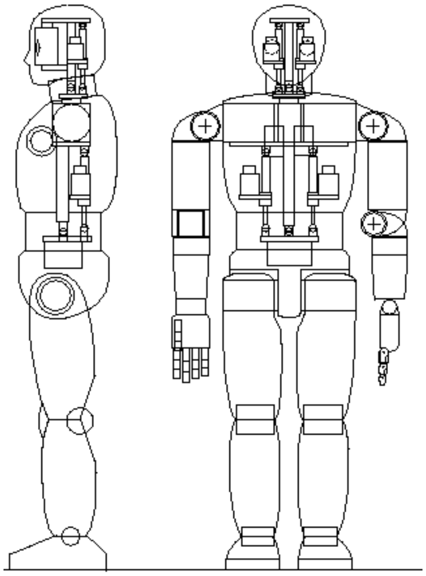

Fig. 2. Schematic plan of ETL Humanoid.

2.Humanoid Interaction Research

We call our project "Humanoid Interaction Research". It has triple meanings:

1. Research on **interaction through a humanoid robot**: Investigate the effect of having an anthropomorphic body (including the sensory and processing structures) on the nature of physical/cognitive interaction with the environment (*complexity*, *dynamical systems*, *embodiment*, *emergence*). How it can be theoretically modelled and/or exploited for control, learning, and useful functionalities. Of course the degree of similarity to real humans is quite crude at this stage, thus our current aim is to reveal global structures as a first approximation.
2. Research on **human-like interactions**: We are in quest for non-traditional task examples in *open ended* domains. For this purpose we avoid top-down rigorous (and abstract) definition and analysis of new tasks (in a traditional robotics sense). Rather, we take our inspirations from our observations about everyday human activity, along with scientifc findings from physiological, neuronal, psychological, and behavioral sciences; inspirations about new patterns of activity

and control/congnitive strategies to cope with them. This will be accompanied with implementation on our humanoid, and field testing in a real environment with ordinary objects and humans, which will then be observed by us to refine our inspirations and theories. This may be termed as an *ecological approach.*

3. Research **by interacting with humanoid robots**: Having a real full-fledged humanoid robot in front of a group of researchers gives them strong motivations for inventing new ideas and trying them out. It is intended to activate research efforts ranging from component technologies, theories, real applications and even philosophical issues.If the humanoid system serves as the focal point of all these research activities, we can expect that interdisciplinary research efforts and integration of various research achievements will be promoted which may lead to a new unified discipline in the long run.

In the following part of the section, each research issue listed in the previous section will be further extended, providing more support for the above research stance.

2.1.Complex dynamical systems

Our first goal is to investigate the global structure of humanoid interaction, therefore we tentatively ignore small localized tasks such as precise object manipulation with fingers and an arm. Instead, our current focus is on dynamic whole body motion in arbitrary postures with lots of contacts withthe real environment.

The complex body (complicated shape, highly redundant DOFs and sensors) and the complex surrounding (including human bodies) give rise to extreme complexity (which we call 'supercomplexity') of physical interaction, sensory data, and motor control. Moreover, it is accompanied by uncertainties due to friction in contact motion, mechanical/sensory uncertainty, unpredictable properties and behavior of surrounding objects, etc.

In such cases, the collapse of traditional model based approach becomes more drastic than with simple mobile robots. Even motion control or state recognition alone require new approach exploiting, for exapmle, global structure of the interaction dynamics, global entrainment[5], on-line stability analysis, and attentional mechanisms.

These points will be discussed further in section 3, with an illustrating example.

2.2.Embodied cognition and emergent behavior

The effect of having a body on cognition and behavior is often termed as 'embodiment', usually accompanied with the notion of 'emergence' in this context.

A behavior emerges as a result of interaction of the robot with the world through its body[3].

Embodiment in perception can be understood as the statistical spatio-temporal structure imposed upon the perceptual data by the body[6]; Sensors are affixed to the body, having particular geometric arrangements, and as the robot moves, the sensors also move along, experiencing particular spatio-temporal pattern of data due to the robot motion relative to the environment, which is actually guided by the sensory data.

Exploiting such spatio-temporal structure in sensory data, learning is made easier compared to general purpose learning paradigm[6].

We believe that introducing a highly complex body will cast a new light on the above understanding; If the robot has a highly redundant mobility, it can choose a variety of different motor constraints during similar behavior. As opposed to a simple robot case, a humanoid has a high degree of versatility in actively choosing and imposing various spatio-temporal structure by controlling its internal degrees of freedom.

This idea may constitute a portion of what J. J. Gibson called "perceptual system", which searches for "affordances" and picks them up as "invariance within variation"[4].

Similarly to the cognition, an effect of having a complex body on emergent behavior can be that it may actually *simplify* and add strong *robustness* to the way the robot deals with the highly complex physical interactions, if it exploits its body (internal DOFs) appropriately.

Intelligence could be explained as an interactive emergence between the emergent behavior and embodied cognition, as they are mutually dependent and generative, and both share a common body.

2.3.Meaningful human-robot-environment interaction

The ultimate theme of our humanoid interaction research is the "meaningful three term interaction structure" (Fig. 1); Understanding and realizing the complex real world interactions among a human, the environment, and a robot, in a unified framework. This is fundamental for true human-robot symbiosis.

Our current working hypotheses are the following:

1. Correct understanding will be approached only by always viewing the three terms (two agents and the shared environment) as a whole. Disecting it into individual terms/interactions will lead to ill-posed problems.
2. There will be common principles governing apparently different interactions as long as they share the three term structure.
3. The above hypotheses imply that we can start by investigating several different interactive examples, all with the three term structure, and extract at least some simple versions of the principles.

Three of the principles will be about (1) 'auto-segmentation of global dynamics', (2) 'points of interaction between different interactions'(aka. 'attention'), and (3) 'interactive emergence of meanings'[14] . Discussions in section 3 tries to capture one aspect of the above principles.

The different forms of three term interactions include the following:

1. **Action generation and action recognition** are forwad/inverse mappings between a 'meaning' and a 'complex spatio-temporal pattern'. However, differnt instances of real actions with the same meaning never

have exactly identical set of features. Thus recognition requires dynamic attention control for focussing on meaningful spatio-temporal features [10]. And generation should allow each instance of the same action to vary and conform to each situation while assuring the equivalence of the overall meaning[8]. Thus generation would also require a focussing mechanism (points of intervention, see section 3).
Various 'meaning's can be attributed to one instance of an action.In order to ground the meaning, we need to resort to the interaction between the observer (who also acts) and the actor (who also observes the other). The meaning must be shared in generation and recognition and between different agents via interaction. As a whole, this forms the three term structure.

2. **Mutual imitation and learning to imitate** is the most basic and the most important mode of the structure discussed above[9]. Imitation is doing the same thing through observing the other's actions.
Note that 'same' here is undefined. But due to the nature of imitation, an agent can close a loop by 'mutual imitation', and find a fixed point purely in behavioral terms without resorting to an external definition of similarity.
Each agent can also learn by trial and error in imitating the other better and better, interleved with the mutual imitation, and this can eventually lead to the interactive emergence of a shared meaning.
3. **Teleoperation and semi-autonomy**: Even a teleoperated robot must have some degree of autonomy in the real world (ex. climbing a rubble-heap with a teleoprated biped robot). If human control and autonomous control are statically separated, the flexibility and robustness of the overall man-machine system will be quite limited. Dynamic selection of human intervention in terms of levels, times, and modes are desirable. How can this be achieved without breaking the stability and consistency of both the robot and human behavior?
4. **Coopearation, communication, and teaching** all share the similar three term structure and the intervention or interaction of interactions problems. Dynamically establishing cooperative behavior requires matching and overlapping of individual behaviors [7]. Situated communication requires shared attention. Since the attention is also the focus of individual behavior, this often implies behavior merger, which is very similar to cooperation. Teaching has similar aspects but with more complicated higher order structure.

2.4.Open ended tasks

From the application point of view, the most outstanding characteristic of a humanoid in contrast with traditional robots is that it is *not* task specific and its application is *open ended;* it can be anything that a human could do. This is an important point to note because if we assume any particular, well defined task, as in the traditional robotics, then we will end up with a task specific optimal solution which most likely does not take the form of a humanoid.

One of our aims is to explore this space of open ended tasks using our humanoid as a vehicle, which may lead to a different robotics methodology, e.g. starting with a particular *form* and explore *indefinite tasks*, rather than starting with a particular task and explore for an optimal form.

Imitation, discussed earlier, is one example of an open ended task; A humanoid which can imitate any (but coarse) human actions encountered in various situations (to test flexibility), and can learn to imitate.

2.5.Practical and versatile research platform

Our humanoid system design attempts to facilitate all the above issues. Equally important constraints are the practicality, versatility and availability as a research platform.

In order to facilitate various research approaches, the software platform should have clear and open application interface with a built-in mechanism which assures interoperability among application modules . Also, both the hardware and software must be available after a very short development period, and they must be reliable. Parts of our design is described in section 4.

3.Segmentation of Global Dynamics and Points of Intervention

In this section, we first present a simulation example of a dynamic whole body humanoid motion. In the discussions that follows, the example is used as a special instance of our novel control framework for complex physical interaction called 'action oriented control'. This framework conforms to the 'three term interaction' view discussed earlier, and the example captures or implies some aspects of the 'common principles': Auto segmentation of global dynamics, points of intervention, and interactive emergence of meaning.

3.1.Example: Rising action

Fig. 3 shows a simulation sequence of dynamic whole body motion, which illustrates our idea. Initially the humanoid is lying flat on the floor (a), first, it swings up its legs (b), and throws them down to roll up on its feet (c). Due to inertia it continues to roll forward until it hits the ground by both hands and becomes globally stable (d). Then it slowly moves the center of mass above its feet by gently pushing the ground with its hands (d), and finally stands up (f).

These snapshots are exerpts from a real continuous output from a general purpose 3D dynamics simulator (ADAMS). It is therefore a faithful simulation of a possible physical phenomenon. Given the task defined by its initial state (a), and its goal state (f), there are ininite number of solutions which connects the two states. In order to find and choose a solution and to realize it, a typical traditional robotics approach might take the following strategy: (1) Explicitly model the underlying physics in detail. (2) Solve and choose the optimal solution in terms of energy, sometimes in combination with other cost functions. (3) Design a controller which tightly controls the system state to follow the desired solution (state space trajectory).

In the example sequence, there are two problems with this approach; (1) Complex, deformable body shape and multiple physical contacts, combined with dynamic motions make the explicit modelling extremely difficult. (2) The search space is extremely large and complex, making it difficult to reach an optimal solution. (3)It is quite difficult to continuously observe the state variables and control the relative motion of the highly complex articulated body involving multiple point contact.

We propose an alternative approach called 'action oriented control'; It does not require an explicit detailed model, it does not seek for energy-optimal solution, and it does not control the system state so tightly.

The motion in the presented simulation has been achieved extremely simply by just specifying several basic (key) poses and arranging them at appropriate time intervals in an open-loop fashion, all chosen intuitively by the experimenter without any detailed analysis or a large number of trials and errors. Small fluctuations in the poses and their timings do not affect the global behavior. Remember that the motion was realized in a full fledged dynamics simulator.

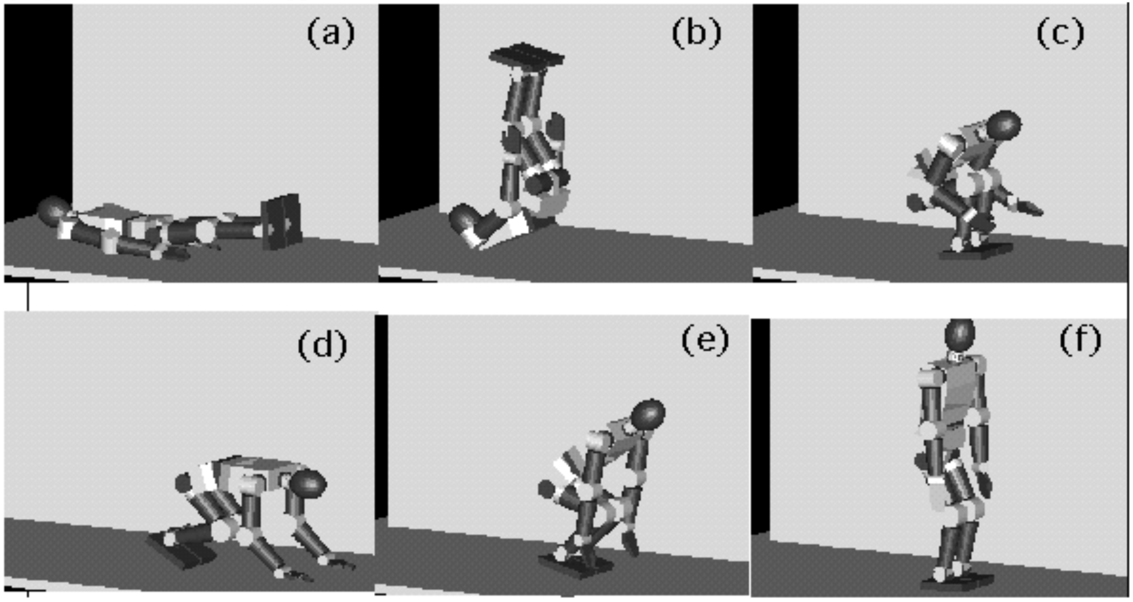

Fig. 3 "Rising" action sequence. Output from a dynamics simulator.

3.2.Segmentation and transition in global dynamics

An important question to be raised here is "**why did it work**?" Why were those key poses and the timings appropriate for the desired motion? How could the experimenter find them? Why it was robust against the deviations in control parameters? -- These are the fundamental questions which lead us to the underlying principles of controlling complex motion. The other arguments, such as arguing how we can construct a general controller, or a learning algorithm that automatically find those poses, may rather be an implementational detail compared to this fundamental question.

Our control strategy in the above example is characterized by the following guidelines:

1. The controller (human, in this case) imagines how he/she would achieve the task with his/her own body. Just focus on the global, qualitative sequence. Often the controller can think of many alterative paths towards the same goal. Currently, the controller chooses the *cognitively easiest* path which consists of fewer qualitative steps and least ambiguous.
2. Let the natural dynamics take over when it is stable and unambiguous. Do not specify the trajectory. This means the motion control is quite loose.
3. Intervene only when the system state is in vicinity of the boundaries (and branching points) between different dynamics. The intervention *kicks* the system into a desired adjacent dynamics either by applying appropriate impact-like force or locally modifying the dynamics by a posture change.
4. Each dynamics is realized by taking a particular posture at an appropriate timing, and following the planned qualitative sequence step by step.

Just like a tumbling dice, the humanoid overpassed the peaks and got dragged towards the bottom of the potential field formed by the body and the ground. But unlike the dice, the humanoid is able to change the global structure of the potential, and shift from one dynamics to another desirable one, by changing its posture and pushing the ground at appropriate times.

It is important to note the following points:

1. **Attention**: This type of control succeeds only when the controller can appropriately detect and focus on the *boundaries* between different dynamics, or potential wells. This is quite different from a traditional view of motion control whose central concern is the stability *within* a fixed dynamics. The *boundaries* correspond to the *segmentation points* notion in qualitative action recognition [10]. Both in action recognition and control, the most important task of the recognizer or the controller is to detect and focus onto the segmentation points. This establishes a direct mapping between the real motions and the high level action classes. Then recognition and control becomes quite simple; detecting a qualitative change in the focussed low level features in recognition, and exerting coarsely controlled force or roughly changing the posture in control.

   Remeber that the controller first *imagined* how he/she would move him/her-self, then outlined a qualitative sequence, which was used for coarse qualitative control. The duality and mutual dependencies between action generation and action recognition is important here. And note that the controller, robot, and the environment formed the 'three term interaction' structure, and the attention on segmentation points, in a sense, supported the transfer of skills (interaction of each with the environment) from the controller to the robot.

2. **Intra-/Inter-dynamics**: The natural dynamics itself gives some adaptiveness to the behavior, because the attractor dynamics stabilizes the system from perturbations. The same principle is used in previous research on dynamic motion control based on limit cycle attractor dynamics [5,11].Howerver, there is a fundamental difference between the action oriented control and the limit cycle dynamics methods. The previous methods focus on maintenance of a fixed dynamics, whereas our method focuses on dynamically creating, and switching between, different dynamics. The two approaches can be combined, complimenting each other; Our method takes the system through a tour through global dynamics structures and the fixed dynamics methods assure the robustness of the trajectory within each local dynamics.

3. **Purposive**: The criteria for choosing alternative paths should not necessarily be optimizing efficiency. In a complex real world system, it makes more sense to choose a reliable and/or a computationally tractable path rather than an *optimal* path which is very difficult to define and to achieve. For example, in the presented sequence of Fig. 3, it would be more *efficient* to achieve a convergence of the system state (balanced) at the posture (e) directly following (b). But it would require a tight coordination and accurate dynamic control of the whole body by observing the precise dynamics variables on-line. Instead, an alternative strategy was adopted in the presented sequence; Deliberately increased the rolling velocity so that the system divergingly overpass the delicate state (c), reaching the state (d), which is robustly achieved. This state is so reliable and has a unique set of state values (e.g. velocity being zero), that the controller does not have to check or measure the state. Thus it can initiate a new step regardless of a previous trajectory, reliably achieving the state (e). This implies that local instability can be useful in the context of achieving a global action. It will be useful in various situations including walking and object grasping. Another useful strategy will be to deliberately perturb the state in order to disambiguify it, which is otherwise difficult to identify.

4.Humanoid System for Complex Interactions

So far we have described our 'three-term' complex interaction view and our approach. They introduce a certain suit of design costraints which makes our humanoid unique compared to other existing humanoid systems.The suit includes whole body free mobility assuming frequent multiple contacts, anti functional-modular design, and transparent data accessibility.

Our design of the ETL-humanoid is a result of finding the best trade-offs between these constraints and the currently available technology.

4.1. Design constraints

1. **Versatility/Non-Modularity**: The design avoids attribution of any fixed specific task oriented functionality to any part of the body. Thus, for example, the power of the fingers and the arms is not limited to what is required for ordinary object handling in a standing or seated position. Instead they are strong enough to hold onto a grip and support the body while lifting itself. And the mobility range of the legs are much wider than required for stable walking. Also the speed of each DOF are set so that it can do fast dynamic motion. In general, we do not assume a particular posture or a particular class of tasks.

   As discussed earlier, our goal is to capture the global strucutre of human interaction and its complexity, robustness and flexibility. So, we ignore highly skilled or precise manipulation and other localized skills, and try to cover a broad range of normal human perceptual and/or motor ability in a typical everyday life. Non functional-modular design is important because the whole body always functions as a unity, and this redundancy is the source of complexity, robustness and flexibility. A functional decomposition/optimization will presuppose particular strategies for particular tasks and situations, depriving the freedom of dynamically exploiting extra available redundancy . Some examples of exploiting the extra redundancy are, when you are carrying a large item in your arms, you might use your elbow/shoulder to push the door open, and yor mouth to pick up a sheet of paper on the cargo, and of course use the whole body (muscles) to lift and control the cargo.

As a comparison, the advanced prototype from Honda adopts a functional modular design; it consists of the lower body which is optimized for biped walking, and the upper body which assumes that the stable upright posture is robustly maintained.

Another example, the SARCOS humanoid, is specifically designed for anthropomorphic gesture generation, but only in the standing posture (bolted to the floor). It is not designed for tight sensory/motor interaction with the world.

2. **Anthropomorphic structure**: The overall physical structure (the dimensions, shape, and kinematics) is made as close as possible to humans. Since we focus on embodiment, and our approach partly relies on intuitive interaction strategies of humans, the global anthropomorphic structure is essential.

   It has a head, two eyes, and a neck, two arms and hands, five fingers on each hand, a torso and a waist, two legs and feet. It is planned to be approximately 1.50[m] tall and the total weight of 60[kg] (including all the internal electronics, excluding the add-on battery system).

   The mechanical degrees of freedom are made as close as possible to humans with some omissions due to trade-offs with performance and compactness: The active DOFs for the arm with the wrist, the leg with the foot, the waist, the neck, the eye, and the hand (five fingers) are 6, 6, 3, 3, 3, 5, accordingly, the total DOFs being 46.

   The robot is not intended to have a super-human nor aspecialist's physical performance. The performance target is that of an average (or weaker) person in normal everyday situations.

3. **Free motion and contacts**: The entire body is not affixed to the ground or any other external structure. Moreover, the system design does not assume any particular posture. So the system can move around freely, take any posture (upright, laying down, or upside down, and so forth) and is still functional. The external surface of the whole body is made as smooth as possible, without steep corners, protrusions, and wires (except for a thin 'tail' for external connection). Thus it can make contacts with extrenal objects including humans at any part(s) of its body. The smooth shape also facilitates attachment of tactile sensors, 3D visual recognition, and obstacle avoidance.

4. **Safety/Affinity**: The overall size and weight (1.5[m], 60[kg]) are restricted to be close to or less than average adults in order to reduce the danger when moving in proximity of people, and to promote affinities. Also, in order to reduce the risk of injury to surrounding people, the smooth external surface and pinch proof joint mechanisms are adopted whereever possible.

5. **Integrity**: Our current goal is *not* just building an anthropomorphic mechanism. Rather, our emphasis is on software development, real world experiments, and theoretical developments *using* the humanoid as a whole. Innovation of component technologies is not our first concern except where essential for the system integration. Therefore, priorities are given to integrity (whether a component integrates well with the rest of the system), usability (whether it facilitates easy and productive use for users), and availability (whether it is already or close to available, not delaying the whole development).

6. **Transparentreal-time data access on a standard open architecture**: A humanoid system is essentially an integrated system and most likely shared by researchers from various different fields. Unusual non-standard systems will introduce undesirable barrier to the researchers. Therefore, the control/sensing system is designed to have a standard interface and open architecture.

   Transparent, real-time access to all the sensory/motor data is essential for our research. The low level communication and software architecture is designed to facilitate the above.

7. **Availability**: The basic mechanical components are selected from commercially available assortment. We basically avoid development of component technology. This is for the sake of reducing the development time and maintenance needs, because we plan to build the entire system in a very short period (2 years) in order to concentrate on software development and experiments as soon as possible.

4.2. Overall Design

The requirements for free mobility and anthropomorphic structure/performance result in an extremely challenging mechanical design problem. All the actuators and mechanical components must be contained within the body, which has very limited internal space due to the dimensions, shape and articulations. Moreover, even if we have to allow some external connections they should not hinder the free mobility significantly.

Electric motors are used all over the system. A miniature drive circuit is placed near each motor, which communicates motor commands and sensor readings via an internal data network. This specially designed network reduces the amount of internal wiring down to three types; motor power, circuit power, and serial digital communication.

Considering the internal space limitation and requirements in research use, the main power unit and the main computer system are placed externally. Only two types of external connection, power and communication, are required thanks to the internal network and drivers, minimizing the undesired effect on mobility.

In order to realize a short-term wireless independent operation for experimental and demonstration needs, a compact on-board battery system and an embedded computer system are also under development.

The entire system is scheduled to be complete by thesecond quarter of1998. In the following, we describe the parts completed so far (more parts will be described in the final version of this paper), the arm, the hand, the internal network, and the external computer system.

4.3. Mechanical Subsystem

Given a task of building an anthropomorphic mechanism capable of as powerful and dextrous motions as those of humans, an experienced mechanical engineer might immediately see the limitations of available actuators and mechanical components, and set out to invent a special mechanism.

We took a different approach; Within everyday situations in an ordinary modern human life, humans rarely exert their maximum physical performance. And if it is within such a range, we claim that it is possible to achieve a comparable mechanical performance with a simple orthodox design. This contributes to an important space factor trade-off between the mechanical complexity and embedded circuitry and devices for motor control and sensing.

It is also important to note that the mechanical performance is only a *potential* ability. An appropriate motion control, such as discussed in the previous section, is essential for achieving a human-like motion performance. Our trade-off reflects this balance between the mechanical potential and control requirements.

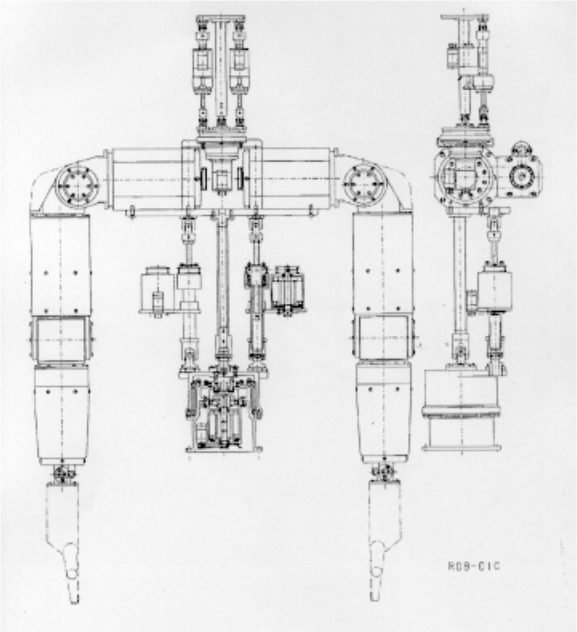

Fig. 4 Mechanical overview of the upper body

4.3.1.Arms

A typical serial link mechanism with revolute joints is adopted. Total 6 DOFs; 3 in the shoulder, 2 in the elbow, and 1 in the wrist. (One wrist DOF was ommitted due to a performance/space trade-off.) Joint structures are pinch-proof. In order to cope with contact motions, it is desinged to facilitate force control. Also, it is assumed that there are situations where an arm must support the whole body, so the high enough joint torques are chosen.

A drive unit consists of an AC servo motor and a harmonic reduction gear. Low mechanical stiffness is allowed for the sake of weight reduction, but backlashes are avoided in order to assure stable force feedback. Therefore the arm is not suitable for a stiff position control but capable of a medium response rate force control. Force sensors are placed in each joint instead of a 6-axis wrist mounted sensor, because we assume frequent use of whole arm contact operations.

4.3.2.Hands and Fingers

Each hand has five fingers in an anthropomorphic configuration. Each finger has three joints driven together by one active degree of freedom actuator adopting *coupled differential mechanism*[12]. We assume that the coupled DOFs are effective in most of the typical situations, excluding special skills.

A clear advantage of this mechanism is that it requires only moderate space for the actuators and drive circuits.

4.3.3.Torso, Neck, and Eyes

The torso and neck each has 3DOFs implemented as combinations of parallel links and revolute joints. Non functional-modular constraint is adopted also here.They are stronger than is necessary for supporting/ orienting the upper structures in the upright posture.

Each eye has 3 orthogonal rotational DOFs; pan, tilt and cyclotorsion. The last DOF is atypical in robotic eye systems, however, human eyes have this DOF, and it gives extra advantage for computer vision algorithms.

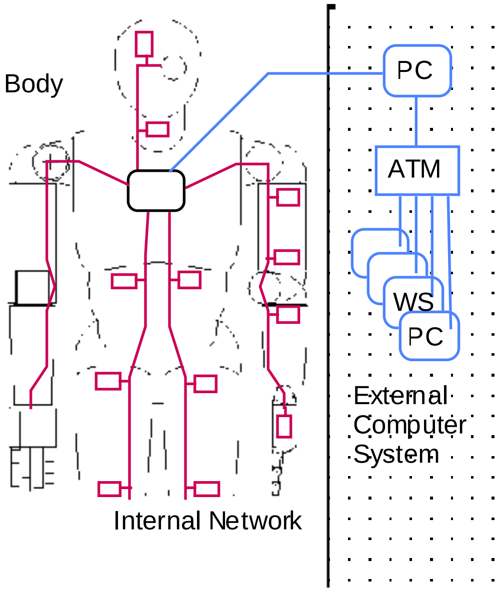

Fig. 5 Information processing subsystem

4.4.Data Processing Subsystem

4.4.1. Humanoid Internal Data Network

The internal network supports high speed serial digital commuication which transfers motor commands and sensor readings between all the motor drivers and the central computer. It is essential for reducing internal wirings.The overall throughput guarantees 1 [ms] constant cycles.

It is tempting to place many processing units distributed throughout the body. However, we decided against it based on the following considerations:

1. The need for transferring raw data to other modules, or a high level process intervening low level motor control can never be eliminated; With traditional manipulator control methods it is mandatory that force/position data be concentrated to one control process at servo rate. More generally, a free access to raw data/commands is a prerequisite for motion control experiments and flexible behavior architectures.
2. It is possible to guarantee a fast enough response via the internal network transfers, because the amount of data transfer is actually quite small. Also, the total amount of computation for all the joints (typically assigned to local processors) is manageable by a single high-performance micro processor.
3. Therefore, the optimal design is to eliminate distributed processors (which tend to be weaker and inflexible due to space limitation) and make the local nodes as small and simple as possible, doing all the computation at a cluster of powerful processors on the other end of the network.

The required communication bandwidth is roughly estimated as follows: Let a servo cycle be 1 [ms], the amount of data transfer per each motor driver per cycle be 64 [bits] (16 [bit] each for position, velocity, force, and command), and the number of motors be 46. Then the required effective bandwidth is: 64[bits] * 1[kHz] * 46[DOFs] < 3[Mbps], which is surprisingly small.

Considering the state of art of digital networks (1Gbps devices are commercially available), there is enough margin for accomodating all the delays and extending the network to a remote external processor system. (Image data from the eyes are treated separately, which will be reported elsewhere.)

Our current design assigns one network segment for each limb and the torso, total 5, in a star topology. This multiplies the bandwidth margin. In the near future, we plan to have tactile sensors all over the body, and the required total bandwidth for them is estimated as several to some dozen Mbps.

For the physical media, optical fibers and copper wires have been considered. Because fibers have strict limitations on bending (radius and cycles) and also require larger transceiver devices, we adopted a copper wire. One disadvantage of this medium is its sensitivity to electrical noises, which can be serious in a machine embedded network like our case. We have developed a communication protocol which maintains a global robustness in noisy environments, which will be presented elsewhere.

This design is the result of trading rapidly developing network devices for a space factor, openness and flexibility. The physical locations of the processors do not matter as long as the communication bandwidth is high enough. This sets the basis for the above trade and our benefit will increase as the device technology advances. Our current design assumes 20[Mbps] operation with FPGA technology. Scaling up to 100[Mbps] or more will require custom VLSI technology but will bring a new class of flexibility and be quite beneficial.

4.4.2.External Computer System

Although we plan to embed several micro processor boards at the root of the internal network in the body for low-level or housekeeping control and limited autonomous operations, the main computation will be done on an external computer system connected via the extension of the internal network. This gives us a great flexibility in configuring the *brain* of the humanoid. An important point is that we can continuously upgrade the external computer system by introducing the latest available technology.

Our current design of the external computer system consits of a cluster of multiple high end PC's interconnected by a 155Mbps ATM LAN. PC's are low cost, high performance, rapidly catch up with device technology. A properly configured bare ATM transfers small data chunks(53 octet fixed length cells) with constant high speed and guaranteed small transmission delay, which is suitable for real time control.

As for the operating system, we are currently testing two alternatives, (1) CHORUS/OS[1], a distributed real time unix[2] and (2) LynxOS[3], a real time unix. The CHORUS/OS became one of our options because it supports real time processing, Unix API, distributed processing, accessibility to the source code. LynxOS is our another option because of its reliability, completeness and performance, though it does not have a comparable distributed processing functionality as CHORUS/OS. A network transparent object-oriented robot programming environment is currently under development on the two platforms. It is based on inter-object message passing model which will be a subset of a standard message protocol CORBA. This way we reserve a connectivity among modules written by researchers, and to other operating systems and commercial application software.

Similar to the mechanical design decisions, here we have avoided development of a custom parallel computer or a custom real time operating system. These mandate vast amount of effort for design, implementation and optimization, which can critically delay the completion of the overall system, and often results in lagging behind the frontier of commercial technology.

At this stage, it is not clear what kind of processing model is suitable for the entire humanoid software system. That is

---

[1] CHORUS/OS is a trade mark of Chorus Systems.

[2] UNIX is a trade mark of X/Open Co.Ltd.

[3] LynxOS is a trade mark of Lynx Real-Time Systems, Inc.

indeed one of our research goals. Therefore, our current design makes the least but important commitment to the programming model; Real time, distributed object oriented programming with clear interfaces. With this, we as a group of researhers can try various processing models (algorithms, data structures and scheduling) for various parts of the system in a loosely coupled mannar while maintaining the overall integrity and real time performance.

5.Concluding Remarks

We have positioned our humanoid research as an approach to understanding and realizing flexible complex interactions between robots, environment and humans.

We claim that a humanoid robot is an ideal tool for the above research; First of all it introduces complex interactions due to its complex structure. It can be involved in various physical dynamics by just changing its posture without need for a different experimental platform. This promotes a unified approach to handling different dynamics. Since it resembles humans, we can start by applying our intuitive strategy and investigate why it works or not, as exemplified in a previous section. Moreover, it motivates social interactions such as gestural communication or cooperative tasks in the same context as the physical dynamics. This is essential for three term interaction, which aims at fusing physical and social interaction at fundamental levels.

As a first step towards extracting a common principle over three term interactions, the concept of action oriented control has been investigated with a simulation example.

The complex, three term interaction view casts unique costraints on the design of a humanoid, such as whole body, smooth shape, non functional-modular design. A brief description of ongoing design of ETL-Humanoid which conforms to the above constraints is presented.

We are also investigating other aspects of the humanoid interaction, such as adaptive imitation [13].

These efforts will hopefully manifest as a first-order approximation of a new understanding and control methodology of the three term interaction in the future.

Acknowledgements

The presented research is supported by “COE Global Informatics” project and “Emphasized Basic Research” project funded by Science and Technology Agency of Japanese government. Special thanks to ETL Director-General Koichiro Tamura, Global Information Processing Laboratory Leader Kazuhito Ohmaki, Director of Intelligent Systems Division Hideo Tsukune, staffs of research planning office and staffs of robotics group for their support. The described research is a collective effort by the members and collaborators of Humanoid Interaction Laboratory including Luc Berthouze, Seiichi Miyakoshi, Sebastien Rougeaux, Kenji Konaka, Youichi Ishiwata, and Fuminori Saito.

References

1. J. Yamaguchi and A. Takanishi. Development of a Humanoid Robot - Design of a Biped Walking Robot Having Antagonistic Driven Joints Using Nonlinear Spring Mechanism. *Proc. 1st Int. Symposium on HUmanoid RObot*s, pages102-111, 1996. 3.
2. R. Brooks and L. Stein. Building Brains for Bodies. MIT AI Memo #1439, August 199
3. M. Mataric, Behavioural Synergy Without Explicit Integration. Proc. AAAI Spring Symp. Integrated Intelligent Architectures, Stanford Univ., Sigart Builletin 2, no.4, pages 130-133.
4. J. J. Gibson. The Ecological Approach to Visual Perception. Houghton Mifflin, 1979.
5. G. Taga, Y. Yamaguchi, H. Shimizu. Self-Organized Control of Bipedal Locomotion in Unpredictable Environment. Biol. Cybern., vol. 65, pages 147-159, 1991.
6. C. Scheier and R. Pfeifer. Classification as Sensory-Motor Coordination: A Case Study on Autonomous Agents. In Advances in Artificial Life. Proc. Third European Conf. Artif. Life (ECAL), pages 656-667, 1995.
7. Y. Kuniyoshi. Behavior Matching by Observation for Multi-Robot Cooperation. In G. Giralt and G. Hirzinger (eds). *Robotics Research -- The Seventh International Symposium*, pages 343-352, Springer, 1996.
8. Y. Kuniyoshi, M. Inaba, and H. Inoue. Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. Robotics and Automation*, Vol.10, No.5, 1994.
9. Y. Kuniyoshi. The science of imitation --- towards physically and socially grounded intelligence ---, RWC Symposium, Technical Report TR-94001, RWC, 1994.
10. Y. Kuniyoshi and H. Inoue. Qualitative Recognition of Ongoing Human Action Sequences. In *Proc. IJCAI93*, pages 1600-1609, 1993.
11. A.A. Rizzi and D.E. Koditschek. Further progress in robot juggling: The spatial two-juggle. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 919--924, 1993.
12. S. Hirose. Connected Differential Mechanism and Its Applications. *Proc. 2nd Int. Conf. on Advanced Research*, pages 319-326, 1985.
13. L. Berthouze, P. Bakker and Y. Kuniyoshi. Learning of Oculo-Motor Control: a Prelude to Robotic Imitation. In *Proc. IROS*, 1996.
14. H. Hendriks-Jansen, Catching Ourselves in the Act. MIT Pr. 1996.