James: A Humanoid Robot Acting over an Unstructured World

Lorenzo Jamone, Giorgio Metta, Francesco Nori and Giulio Sandini
Lira-Lab, DIST
University of Genoa, Italy
{lorejam,pasa,iron,giulio}@liralab.it

Abstract—The recent trend of humanoid robotics research has been deeply influenced by concepts such as embodiment, embodied interaction and emergence. In our view, these concepts, beside shaping the controller, should guide the very design process of the modern humanoid robotic platforms. In this paper, we discuss how these principles have been applied to the design of a humanoid robot called James. James has been designed by considering an object manipulation scenario and by explicitly taking into account embodiment, interaction and the exploitation of smart design solutions. The robot is equipped with moving eyes, neck, arm and hand, and a rich set of sensors, enabling proprioceptive, kinesthetic, tactile and visual sensing. A great deal of effort has been devoted to the design of the hand and touch sensors. Experiments, e.g. tactile object classification, have been performed, to validate the quality of the robot perceptual capabilities.

I. INTRODUCTION

Flexible and adaptive behavior is not only the result of a smart controller but also of the careful design of the body, the sensors, and of the actuators. This has been shown to be true both in artificial and biological systems [1], [2], [3]. Specifically, this consideration has underlined the necessity of providing robots with a physical structure suited for interacting with an unknown external world. Within this framework, learning is the result of the continuous interaction with the environment and not the outcome of abstract reasoning. Knowledge about the external unstructured world is incrementally developed and becomes detailed only after extensive exploration. Following these new ideas, we are doomed to consider actions as the main tool for cognitive development, see for example [3]. When implementing such a general principle, two issues become apparent. On one hand, actions require the existence of a physical body (embodiment, [2]), which ‘safely’ interacts with a continuously changing, dynamic, unstructured and, in one word, “real” environment. On the other hand, the necessity of understanding the consequences of actions requires fine perceptual capabilities, i.e. the ability of retrieving information about the world (exteroception) and about the body (proprioception). Therefore, action and perception need to be considered as a unique tool for the development of the behavior of the robot, mutually affecting each other in a self-maintaining process of co-development [1]. Clearly, there is a direct relationship between the complexity of the action system and the quality of the data that the robot can have access to by means of action [3]. At the same time, actions can be improved because of the improvement of the perceptual system through learning and adaptation.

II. OVERVIEW OF THE PAPER

In the previous section we have discussed the importance of three fundamental concepts in humanoid robotics: embodiment, action, and perception. In the light of these ideas, the design and realization of a robot has become a research topic in itself. This paper focuses on the design of the humanoid robot called James (see Figure 1), primarily reporting the design choices which have been suggested/required by seriously considering embodiment, action, and perception. Some of the design aspects are related to the fact that the robot is intended to operate in the context of object manipulation, which will be the test-bed on which these new principles will be validated. James is a 22-DOF torso with moving eyes and neck, an arm and a highly anthropomorphic hand. In the next sections we cover the robot design, its actuation, and sensorization. Experiments are reported.

Fig. 1. The humanoid robot James.

The structure of the following sections has to be understood in terms of the three principles mentioned above: embodiment, action, and perception.

- Section III deals with the body, i.e. the physical structure of the robot. The robot structure is similar to that of humans, both in size, number of DOFs and range of movements. Such a complex structure is, in our view,
mandatory when studying complex tasks like manipulation. Specific attention has been devoted to the structure of the hand, which, through a necessary simplification retains most of the joints of a human hand. The complexity of the structure makes modeling a challenging problem. Therefore, the platform seems to be ideal for exploring the expressive potentialities of the principles described above.

- Section IV deals with the actuation system. As already pointed out, ‘safe’ interaction with the external world is a fundamental process. Therefore, when designing the robot, we decided to make it intrinsically compliant to act safely over an unknown and unstructured world, without damaging itself and the objects around it. This has been achieved by an accurate choice of the actuation system. Specifically, movements are made intrinsically compliant by using plastic belts, stainless-steel tendons which transmit torques from motors to joints, and springs at critical locations along the tendons. A great deal of attention has been focused on the actuation of the hand. Being impossible to actuate all the 17 joints\(^1\), an under-actuated mechanism was designed. Roughly speaking, some of the joints are weakly coupled in such a way that blocking one of them does not prevent the movement of the remaining joints.

- Section V deals with sensors. Once again, we should stress the fact that sensors are fundamental to retrieve information about the external world and about the body of the robot. Perception is provided by visual, kinesthetic, proprioceptive and tactile sensors, some of which have been created specifically for James. We will describe the design and realization of a novel silicone-made tactile sensor. Touch is indeed critical in manipulation tasks.

Furthermore, to prove the crucial role of tactile sensing and proprioception in getting information about grasped objects, we describe a set of classification experiments on objects with different shape and/or softness.

### III. Physical structure

James consists of 22 DOFs, actuated by a total of 23 motors, whose torque is transmitted to the joints by belts and stainless-steel tendons. The head is equipped with two eyes, which can pan and tilt independently (4 DOFs), and is mounted on a 3-DOF neck, which allows the movement of the head as needed in the 3D rotational space. The arm has 7 DOFs: three of them are located in the shoulder, one in the elbow and three in the wrist. The hand has five fingers. Each of them has three joints (flexion/extension of the distal, middle and proximal phalanxes). Two additional degrees of freedom are represented by the thumb opposition and by the coordinated abduction/adduction of four fingers (index, middle, ring, and little finger). Therefore, the hand has a total of 17 joints. Their actuation will be described later in Section IV. The overall size of James is that of a ten-year-old boy, with the appropriate proportions for a total weight of about 8 kg: 2 kg the head, 4 kg the torso and 2 kg arm and hand together (the robot parts are made of aluminum and Ergal).

### IV. Actuation system

This section describes how compliance has guided the design of the actuation system. Special attention has been devoted to the hand actuation, which is crucial in manipulation tasks.

#### A. Overall actuation structure

The robot is equipped with 23 rotary motors (Faulhaber), most of which directly drive a single DOF. Exceptions are in the neck, where three motors are employed to roll and tilt the head, and in the shoulder, where three motors are coupled to actuate three consecutive rotations. Transmission of the torques generated by the motors is obtained through plastic toothed belts and stainless-steel made tendons; this solution is particularly useful in designing the hand since it allows locating most of the hand actuators in the wrist and forearm rather than in the hand itself, where strict size and weight constraints are present. Furthermore, tendon actuation gives a noteworthy compliance to the system\(^2\), easing the interaction over an unknown and unstructured world. Indeed, the elasticity of the transmission prevents damages to the external world and the robot itself, especially in presence of undesired and unexpected conditions, which are usual in the “real” world: small positional errors when the hand is very close to the target, obstacles in the path of the desired movements, strokes inflicted by external agents. Extra intrinsic compliance has been added by means of springs in series with the tendons to further help the coupling/decoupling of the fingers (see Figure 2).

#### B. Shoulder actuation structure

The design of the tendon driven shoulder was developed with the main intent of allowing a wide range of movements. The current design, consists of three successive rotations corresponding to pitch, yaw and roll respectively (see Figure 3 for details). The three motors that actuate these rotations are

\(^1\)All the motors for actuating the hand are located in the arm and forearm. Therefore, it was impossible to have one motor for each degree of freedom of the hand.

\(^2\)Stainless-steel tendons are extremely elastic especially when the applied tension exceeds a certain threshold.
located in the torso\textsuperscript{3}. Actuation is achieved by exploiting the design of tendons and pulleys. Special care has been taken in designing the abduction (yaw) movement. The abduction rotation is divided among two mechanically coupled joints (see Figure 4). The two joints correspond to a sequence of two rotations around two parallel axes. A mechanical tight coupling forces the two angles of rotation to be equal. This ‘ad hoc’ solution allows the arm to gain an impressive range of movement (pitch $\simeq 360^\circ$, yaw $\geq 180^\circ$, roll $\simeq 180^\circ$) and to perform special tasks (e.g. position the hand behind the head) at the expense of a multiplication of the required torque by a similar amount.

\textbf{C. Head actuation structure}

The head structure has a total of 7 degrees of freedom, actuated by 8 motors. Four of these motors are used to actuate the pan and tilt movements of the two independent eyes (see Figure 5 for a scheme of the tendon actuation). One motor directly actuates the head pan. The remaining three motors actuate the head’s two additional rotations: tilt and roll. These two rotations are achieved with an unconventional actuation system (see Figure 6). Each motor pulls a tendon; the tension of the three tendons determines the equilibrium configuration of the spring on which the head is mounted. The structure has an implicit compliance but it can become fairly stiff when needed by pulling the three tendons simultaneously.

\textbf{D. Hand actuation structure}

The main constraint in designing the hand is represented by the number of motors which can be embedded in the arm, assuming there is very little space in the hand itself. The current solution uses 8 motors to actuate the hand. Note that the motors are insufficient to independently drive every joint singularly. After extensive studies, we ended up with a solution based on coupling some of the joints with springs. This solution does not prevent movement of the coupled joints when one of them is blocked (because of an obstacle) and therefore results in an intrinsic compliance. Details of the actuation for each finger are given in the following.

- Thumb. The flexion/extension of the proximal and middle phalanges is actuated by a single motor. However, the movement is not tightly coupled. The actuation system allows movement of one phalanx if the other is blocked.
Two more motors are used to directly actuate the distal phalanx and the opposition movement.
- Index finger. The flexion/extension of the proximal phalanx is actuated by a single motor. One more motor is used to move the two more distal phalanxes.
- Middle, ring and little fingers. A unique motor is used to move the distal phalanxes of the three fingers (flexion/extension). Once again, blocking one of the fingers does not prevent the others from moving. One additional motor actuates the proximal phalanxes and, similarly, the movement transfer to the others when one of the fingers is blocked.

Finally, an additional motor (for a total of 8 motors) is used to actuate the abduction/adduction of four fingers (index, middle, ring and little fingers). In this case, fingers are tightly coupled and blocking one of the fingers blocks the others.

E. Motor control strategy

Low-level motor control is distributed on 12 C-programmable DSP-cards, some of which are embedded in the robot’s arm and head. Communication between cards and an external cluster of PCs is achieved by the use of a CAN-BUS connection. Robot movement is controlled by the user with positional/velocity commands which are processed by the DSPs which generate trajectories appropriately. Most of the motors are controlled with a standard PID control except for the shoulder, the neck and the eyes motors where mechanical constraints require the use of more sophisticated and coordinated control strategies. These specific control strategies fall outside the scope and size of this paper.

V. Sensors

The sensory system allows James to gather information about its own body and about the external world. The robot is equipped with vision, proprioception, kinesthetic and tactile inputs. As previously mentioned, the robot perceptual capabilities have been designed in line with the necessity of using it for manipulation.

Vision is provided by two digital CCD cameras (PointGrey Dragonfly remote head), located in the eyeballs, with the controlling electronics mounted inside the head and connected via FireWire to the external PCs.

The proprioceptive and kinesthetic senses are achieved through position sensors. Besides the magnetic incremental encoders connected to all motors, absolute-position sensors have been mounted on the shoulder to ease calibration, in the fingers (in every phalanx, for a total of 15) and in the two motors that drive the abduction of fingers and of the thumb. In particular, the hand position sensors have been designed and built expressly for this robot, combining magnets and Hall effect sensors. In order to understand the sensor’s structure and functioning, consider one of the flexion joints of the fingers, between a pair of links (phalanxes). One metal ring holds a set of magnets and moves with the more distal link (of the two we consider here). The Hall-effect sensor is mounted on the first link. Magnets are laid as to generate a specific magnetic field sensed by the Hall-effect sensor, monotonically increasing when the finger moves from a fully-bended position to a fully-extended one: from the sensor output, the information about flexion angle is derived. Furthermore, a 3-axis inertial sensor (Intersense iCube2) has been mounted on top of the head, to emulate the vestibular system.

Tactile information is extracted from sensors which have been specifically designed and developed for James, and have been realized using a two-part silicone elastomer (Sylgard 186), a Miniature Ratiometric Linear Hall Effect Sensor (Honeywell, mod. SS495A) and a little cylindrical magnet (grade N35; dim. 2x1.5 mm).

The sensor structure is shown in Figure 7: any external pressure on the silicone surface causes a movement of the magnet and therefore a change in the magnetic field sensed by the Hall-effect sensor, which measures indirectly the normal component of the force which has generated the pressure. An air gap between the magnet and the Hall-effect sensor increases the sensor sensitivity while keeping the structure robust enough. Two kinds of sensors have been realized, with a different geometric structure depending on the mounting location: phalangeal-sensors for some of the phalanxes and fingertip-sensors for the distal ends of each finger (Figure 8). The former present one sensing element, which can measure only a limited range of forces, even if with a high sensitivity. In practice, the response of the phalangeal sensors can be regarded as an on/off response. The fingertip-sensors, instead, have two sensing elements capable of sensing a larger range of stimuli.

In order to realize the silicone parts two different molds have been used (Figure 9 and Figure 10), employing the same building procedure. Two different covers for each mold can be seen in the figures: cover A, with little cylindrical bulges to produce suitable holes for the magnets, and cover B, with rabbets on some edges to create the already mentioned air gaps. Molds are first filled with some viscous silicone (mixture of the two parts, base and curing agent), covered with cover A and left drying under progressive heating: this latter process lasts 24 hours at room temperature and then, in the oven, 4
hours at 65°C, 1 hour at 100°C and 20 minutes at 150°C. Afterward, magnets are inserted, some more silicone is poured, and molds are covered with cover B and left drying again. When ready, silicone parts are fixed on the fingers with a sealing silicone, next to the Hall effect sensors previously glued in the desired positions.

The *fingertip-sensor* characteristic curve is reported in Figure 11, where the thicker line is the mean value over series of several measurements and the vertical bars are the standard deviation. The fact that the standard deviation is reasonably low is due to the robustness of the mechanical structure, which is robust to pressure applied at different positions and from different directions, even if with substantially different intensities.

The non-linear response of the sensor is a feature because of its similarity with the log-shaped response curve observed in humans, whose sensitivity decreases with the stimulus intensity [4]. Furthermore, the minimum force intensity detected by the sensor (less than 10 grams for *fingertip-sensors* and about 1-2 grams for *phalangeal-sensors*) is very low, an aspect which has been pursued as the main feature of the device during the design and development of its mechanical structure. Finally, the use of the soft silicone shows an intrinsic compliance and increases the friction cone of the grasping forces. Silicone adapts its shape to that of touched object without being subject to plastic deformations, easing the interaction with unknown environments.

12 tactile sensors have been mounted on the James’s hand: 5 *fingertip-sensors*, one for each finger, and 7 *phalangeal-sensors*, two on the thumb, ring finger and middle finger, and one on the index finger, as shown in Figure 12.

All the 32 Hall-effect sensors of the hand (employed in tactile and proprioceptive sensing) are connected to an acquisition board, mounted on the hand back and interfaced to the CAN-BUS as for the DSP-based control boards. The acquisition card is based on a PIC18F448 microcontroller, an ADC and multiplexer, and a 40-cable connector, which holds the 32 signals and provides the 5 Volt DC power supply to the sensors. To wire such a considerable number of devices.
in such a little space, with a reasonable resiliency (required because of the constant interaction of the hand with the external environment), we have chosen a very thin stainless-steel cable, coated in Teflon, with a 0.23 mm external diameter. Moreover, in order to further increase system robustness, cables are grouped into silicone catheters along their route between different sensors and toward the acquisition card.

VI. Experiments

Experiments have been carried out to validate the design and quality of proprioceptive signals to gather useful information about grasped objects. Proprioception can be further combined with touch in object clustering and recognition showing a notable improvement. The importance of proprioception in object clustering and recognition has been proved both in human subjects [5] and in artificial systems [6]; our aim here is to show how tactile information improves the system performance, providing a richer haptic perception of grasped objects. A Self Organizing Map (SOM) has been employed to cluster and display data. The object representation is simply given by the data acquired by the proprioceptive and tactile sensors of the hand, which provides information about object shape and softness.

The experiment is divided into three stages: data gathering, SOM training and SOM testing. During the first stage, six different objects (bottle, ball, mobile phone, scarf, foam blob and furry toy) have been brought metaphorically to the robot’s attention. James performed several times the same pre-programmed action of grasping and releasing the objects. A simple motor grasping strategy turned out to be effective because of the overall design of the hand. Data provided by proprioceptive and tactile sensors have been acquired every 50 ms, recorded, and organized in 25,000 32-sized vectors, each representing the haptic configuration of the hand in a specific time instant. Afterward, data have been fed for 50 epochs to the SOM, which is formed, in our case, by a main layer (Kohonen layer) of 225 neurons, organized in a 15x15 bi-dimensional grid, with hexagonal topology, and an input layer of 32 neurons, that receive the recorded vectors of hand configurations. During this training stage, main layer neurons self-organize their position according to the distribution of the input vectors, creating separate areas of neurons responding to similar stimuli.

Finally, the neural network has been tested with both dynamic and static inputs, to see how the grasping action is represented and to check the classification and recognition capabilities of the trained system. Results are shown in Figure 13 and Figure 14.

Figure 13 shows the state of the SOM main layer after training. Each small diagram accounts for the values of the corresponding weights of the neuron in that particular position. It is worth noting that nearby neurons show similar curves, which gradually change moving across the network. The path indicated by the black arrows represents the temporal sequence of the neurons activated during a ball grasping. Neurons lay on a continuous path, because of the gradual changes in sensor outputs during a continuous movement. Figure 14 shows the response to static inputs: the spikes on both graphs indicate the number of activations of the neuron in that position after the presentation of inputs describing hand haptic configuration when grasping a particular object or when free and open. It is worth noting that different objects are well classified and similar objects lay in adjacent regions.

Another experiment has been carried out to prove the special importance of tactile stimuli, focusing on situations where the use of proprioception alone may lead to ambiguous results. A first SOM has been trained only with proprioceptive information. A second SOM was instead trained with both tactile and proprioceptive inputs. Figure 15 demonstrates that tactile sensing improves the system performance: indeed, the grasp of a scarf and a bottle is only distinguished by the second neural network. In fact, with these two objects the final hand posture is very similar, as a consequence of the fact that the two objects have approximatively the same shape. Therefore, the only way to distinguish the two objects is by the use...
of tactile information. The two objects have indeed different material characteristics (softness) and therefore only the use of touch can sense their difference.

VII. FUTURE WORK

Future work is already planned, concerning both the arm and the hand control. In particular, adaptive modular control and the exploration of grasping strategies will be investigated.

Remarkably, Bizzi and Mussa-Ivaldi [7] have proposed an interesting experimental evidence supporting the idea that biological sensorimotor control is partitioned and organized in a set of modular structures. At the same time, Shadmehr [8] and Brushers-Krug [9] have shown the extreme adaptability of the human motor control system. So far, adaptation has been proved to depend on performance errors (see [8]) and context related sensory information (see [10]). Following these results, a dynamic control law which consists of the linear combination of a minimum set of motion primitives have been examined. Interestingly, the way of combining these primitives can be parametrized on the actual context of movement [11]. Even more remarkably, primitives can be combined in approx- imately the same way forces combine in nature. Adaptation (i.e. by means of combining primitives with different mixing coefficients) can arise both from error-based updating of parameters (where the error represent the difference between the desired and measured position and velocity of the joints) and from previously learned context/parameters associations (even combining some of them in a new and unexperienced context can be seen as the combination of some experienced contexts).

As far as the hand movement is concerned, in the experiment described in this paper, the grasping action used, even if effective, is still very simple: fingers are just driven to a generic closed posture, wrapping the object by means of the elastic actuation system. Following cues from existing examples [12], future development include the implementation of a system for choosing the best grasping action among a series of basic motor synergies, combining some of them, in trying to maximize the number of stimulated tactile sensors and hence obtaining a more comprehensive and effective object perception. Of course, in order to show this ability of evaluation and improvement of performances (i.e. move the grasped object toward the unstimulated sensors) a topological map of the sensors within the hand is needed. Following the developmental approach, we want this map to be autonomously learned by the robot by random interactions with the environment. It is reasonable to assume [13] that this could be possible by analyzing the correlation between channels from the recorded data as provided by touch and proprioception, exploiting multisensory integration as suggested by generic neural science results[14].

VIII. CONCLUSIONS

We have described James, a 22-DOF humanoid robot equipped with moving eyes and neck, a 7 DOF arm and a highly anthropomorphic hand. The design of James has been guided by the concept of embodiment, material compliance and embodied interaction. Since James is intended to operate in a manipulation scenario, special care has been taken on the design of novel silicone-made tactile sensors. These sensors have been designed, realized, tested and integrated within the James’s anthropomorphic hand. Hence, to prove the importance of proprioceptive and tactile information, experiments of classification and recognition of grasped objects have been carried out using a Self Organizing Map to cluster and visualize data. As expected, results have shown the improvements offered by tactile sensing in building object representations, enabling the system to classify and recognize objects according to their softness and elasticity. Shape information provided by the measured posture of the grasping hand are indeed not sufficient to discriminate two objects with similar size but different softness. Finally, the future development of efficient and flexible grasping strategies has been outlined.

ACKNOWLEDGMENT

The work presented in this paper has been supported by the ROBOTCUB project, funded by the European Commission through Unit E5 “Cognition”. Moreover, it has been partially supported by NEUROBOTICS, a European FP6 project (IST-2003-511492).

REFERENCES


