14

Imitation-Based Task Programming on a Low-Cost Humanoid Robot

Jacopo Aleotti and Stefano Caselli
University of Parma,
Italy

1. Introduction

Humanoid robots are complex service platforms with anthropomorphic features, specifically designed for close interaction with humans. Conventional programming strategies are hardly applicable to humanoids due to the high number of degrees of freedom that must be coordinated concurrently. Therefore, exploiting humanoids' potential in service tasks remains an elusive goal. One of the most promising techniques for dealing with humanoid robots is programming by demonstration, which allows even unexperienced users to easily interact with the robot based on the teaching by showing or imitation paradigm. In particular, the ability to imitate human gestures and follow task-relevant paths are essential skills for legged humanoid robots, as they provide the fundamental techniques for physical human-robot interaction. This chapter investigates the potential of imitation in programming humanoid motor skills. As target platform, we have adapted a Robosapien V2 (RSV2), a low-cost small humanoid available in the toy market. The chapter focuses on the teaching of basic, humanoid-relevant skills such as body postures and walking paths. We have explored and combined multiple sensing sources to capture human motion for imitation purposes, namely a dataglove, an electromagnetic motion tracker, and a monocular vision system for landmark recognition. The imitation approach illustrated in this chapter is rather general, even though its implementation is constrained by limitations of RSV2 and by sensor inaccuracies. In particular, the chapter reports successful experiments on gesture imitation, including arms motion as well as upper body and head movements. The gesture repertoire learned by the robot can serve both as a body language for understanding human requests in human-robot interaction and as a set of primitives which can be combined for programming more complex tasks. We believe that a deep assessment of a low-cost humanoid robot is extremely important for the robotic research community since the technological requirements and the costs to develop more advanced humanoid robots still prevent them to become broadly available. Currently, most high-end humanoids are developed as prototypes platforms under the supervision of important private companies. Therefore, low-cost humanoid platforms such as RSV2 provide an exciting and affordable opportunity for research in humanoid integration in service tasks.

2. Background and related work

The development of humanoid systems is a tremendous challenge for robotics research. Many academic laboratories as well as industrial companies are devoting substantial efforts and resources in building advanced humanoid robots. Several authors (Schaal, 1999; Mataric, 2000; Breazeal & Scassellati, 2001) have pointed out that the principal challenges in building effective humanoid robots involve different aspects such as motor actuation, perception, cognition and their integration. Typically, complex mechanical systems must be regulated and a high number of degrees of freedom must be controlled. Study of appropriate materials, motors, power supplies and sensors is also required. Moreover, humanoid systems require advanced algorithms and software solutions to achieve learning skills for autonomous operations. The development of natural user interfaces for human robot interaction is also required since one of the main goals is to overcome the traditional approaches of robot design by creating robots able to interact with humans in everyday life. However, the technological requirements and the costs to develop effective humanoid robots still prevent such advanced systems to become available to the public. Few low-cost humanoids have been designed for research purposes. One of them is Robota (Billard, 2003), a small humanoid doll used as educational toy for robotic classes and playing vision-based imitation games. Recently, several low-cost humanoid platforms have appeared in the consumer market. Besides being considered as high quality toys for personal entertainment, these devices can also be exploited for research projects. One of such robots is Robosapien V2 (RSV2) developed by WowWee and released at the end of 2005. RSV2 has been chosen for the experimental evaluation reported in this chapter as it is one of the most promising low-cost robots available in the market. Our research on imitation is motivated by the fact that robot programming by demonstration is an effective way to speed up the process of robot learning and automatic transfer of knowledge from a human to a robot. As it is well known that robot programming using traditional techniques is often difficult for untrained users, especially in the context of service robotics, programming by demonstration provides an intuitive solution by letting the user act as a teacher and the robot act as a learner (Ikeuchi & Suehiro, 1994; Zöllner et al., 2002). The need for a high-level approach in robot programming is even more apparent with humanoid robots, where conventional programming of a high number of degrees of freedom would be clearly unacceptable.

In this work we present a system for robot programming by demonstration where a Robosapien V2 is programmed to imitate human gestures and walking paths. In particular, we have explored the idea of augmenting an entertainment humanoid with multiple sensors to accomplish rather complex imitative tasks. The system combines a dataglove, a motion tracker and a monocular vision system. The first contribution of the chapter is the experimental evaluation of various imitation games where the robot follows the gestures of a human teacher. A second contribution is the proposal of an algorithm for walking imitation. The chapter shows how humanoid walking paths can be synthesized by fitting a parametric curve with data observed in human-demonstrated trajectories, thereby providing an effective technique for path imitation. Walking paths can be further adapted by the robot based on online visual guidance.

Hereafter, prior work on humanoid imitation and motion learning and is discussed. Many authors have investigated the problems of human body tracking and mapping of body postures to humanoid robot. The paper by (Bandera et al., 2004) proposes a method for real-time estimation of upper body postures based on computer vision and an inverse kinematic model of the robot. In (Riley et al., 2003) a similar approach was used incorporating 3D vision and a kinematic model of the human teacher. A single camera system for data acquisition with a particle filtering technique for parameter estimation was adopted in (Menezes et al., 2005). Other works were mainly focused on the problem of adaptation of human motion to humanoid motion with kinematics constraints such as joint and velocity limits (Pollard et al., 2002) or ground contact conditions for humanoid locomotion (Matsui et al., 2005). In (Matsui et al., 2005) a method was proposed for the generation of natural motions in an android by measuring the robot posture at its visible surfaces and comparing it to the posture of a human teacher. The work in (Inamura et al., 2001) focused on mimesis learning using primitive symbol observation with Hidden Markov Models to abstract the dynamics of human motion and to generate natural motion patterns.

In (Inamura et al., 2005) an imitation strategy was adopted based on attention points and intent imitation. A neural learning system was proposed in (Kuniyoshi et al., 2003) for learning motion patterns from the observation of self robot movement. A mirror neuron model was implemented in (Ito & Tani, 2004) for imitative interactions. In (Shon et al., 2006) a nonlinear regression algorithm was used for mapping motion capture data from a human actor to a humanoid robot. Calinon et al. (Calinon et al., 2006) presented a method for extracting the goals of a demonstrated task and determining the best imitation strategy that satisfies the goals. As proposed in (Nakazawa et al., 2002; Nakaoka et al., 2005), studies were also conducted for learning and reproducing human dances through human observation.

Our work is among the first attempts to investigate the capabilities of RSV2 for humanoid research. The proposed imitation mechanism is strongly constrained by the limitations of the toy robot. The work in (Behnke et al., 2006) is the only one that has considered the use of Robosapien for research purposes. The authors have developed an augmented version of Robosapien V1. A Pocket PC and a color camera were added to the robot to make it autonomous and its capabilities were tested for basic soccer skills.

The problem of tracking and following demonstrated navigational routes has been investigated mainly in the context of wheeled mobile robots. Dixon and Khosla (Dixon & Khosla, 2004) presented a system for learning motor skill tasks by observing user demonstrations with a laser range finder. The system is able to extract subgoals and associate them with objects in the environment generalizing across multiple demonstrations. The work in (Hwang et al., 2003) proposed a touch interface method to control a mobile robot in a supervised manner. The algorithm extracts a set of significant points from the user-specified trajectory, produces a smooth trajectory using Bezier curves and allows on-line modification of the planned trajectory in a dynamic environment. In (Hon Nin Chow et al., 2002) a reactive sensor-motor mapping system is described where a mobile robot learns navigational routes and obstacle avoidance. In (Wanitchaikit et al., 2006) a self-organizing approach for robot behaviour imitation is presented. A demonstrator mobile robot collects visual information about the environment; then the movement features are presented to the imitation engine. Tang et al. (Tang et al., 2001) proposed a vision based autonomous navigation system for mobile robots in an indoor environment by teaching and playing-back scheme. The system memorizes a sequence of omnidirectional images using them to compute the trajectory to track the taught route. Arakawa et al. (Arakawa et al., 1995) developed a trajectory generation method for wheeled vehicles. The algorithm uses Bezier curves for smoothly reducing the deviation from the guideline. Morioka et al. (Morioka et al., 2004) focused on the problem of human-following for a mobile robot in an intelligent environment with distributed sensors. The control law is based on a virtual spring model.

Several interesting research projects have considered the use of humanoid robots for navigation and autonomous mapping of indoor environments. The work in (Michel et al.,

2005) presented an approach to autonomous humanoid walking for the Honda ASIMO in the presence of dynamically moving obstacles. The system combines vision processing for real-time environment mapping and footsteps planning for obstacle avoidance. A similar approach was presented in (Kagami et al., 2003). In (Michel et al., 2006) an online environment reconstruction system is presented. The system utilizes both external sensors for global localization, and on-body sensors for detailed local mapping for the HRP-2 humanoid robot.

The chapter is organized as follows. Section 3 describes the experimental set-up including a description of the augmented RSV2 and the input devices. Sections 4 and 5 describe the adopted gesture and path imitation techniques. Section 6 discusses the trend toward low-cost humanoid and imitation systems. The chapter closes in section 7 discussing the obtained results.

3. Experimental set-up

3.1 Robosapien V2

The humanoid robot used in this work is a Robosapien V2 developed by WowWee and shown in figure 1 (left image). It is a low cost toy robot expressly designed for the consumer market. RSV2 is 60cm tall and is driven by 12 DC motors. It has realistic joint movements and onboard sensors. Some RSV2 technical specifications are reported in Table 1.

Length	43cm
Width	32cm
Height	60cm
Weight	7Kg
Body batteries	6 D
Brain batteries	4 AAA

Table 1. RSV2 specifications.

RSV2 is fully controllable by a remote infrared controller and has capabilities which make it suitable for entertainment but also for research purposes. The main functionalities of the robot include true bipedal walking with multiple gaits, turning, bending, sitting and getting up. RSV2 can also pick up, drop and throw small objects with articulated fingers. The robot is equipped with an infrared vision system, a color camera, stereo sonic sensors and touch-sensitive sensors in its hands and feet. These sensors can be used to trigger simple reactive behaviors to environmental stimuli through preprogrammed motion sequences, which can be stored in the onboard memory. For example RSV2 can recognize and react to primary colors and skin tones. Moreover, it is able to react to sounds, to track close moving objects with its head and to avoid obstacles while walking. The locomotion of RSV2 is achieved by alternatively tilting the upper body and moving the leg motors in opposite directions. The lateral swinging movement of the upper body generates a periodic displacement of the center of mass between the two feet. To achieve fully autonomous capabilities RSV2 must be augmented with external computing power. There are two possible ways for augmenting RSV2. The first method is hacking the on board electronics to get access to the motor and sensor signals. The second method is to bypass the remote controller with an infrared transmitter connected to a processing unit. This strategy is less intrusive but it only allows transmission of commands to the robot, therefore it also requires external sensors like a camera, as proposed in [3]. In this work we chose a similar approach. We adopted a USB-UIRT (Universal Infrared Receiver Transmitter) device, which is connected to a host PC. The device is able to transmit and receive infrared signals. Reception has been used to learn the IR codes of the RSV2 remote controller. As a result, the IR codes of the motor commands for the RSV2 have been decoded and stored in a database.

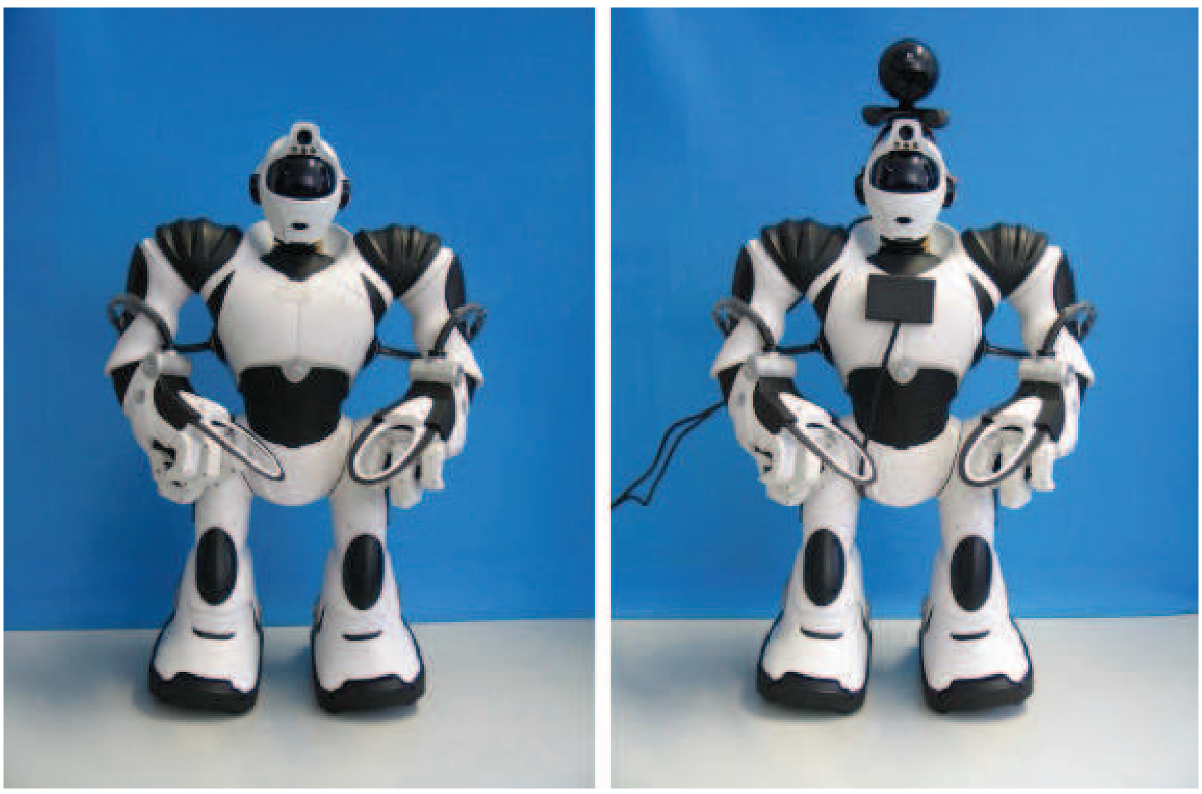

Fig. 1. Original RSV2 (left image) and augmented RSV2 (right image).

3.2 Sensor devices

The experimental set-up of the proposed system comprises a CyberTouch glove (by Immersion Corporation), a FasTrack 3D motion tracking device (by Polhemus, Inc.) and a monocular vision system. The CyberTouch glove used in the experiments has 18 sensors for bend and abduction measurements. The FasTrack is a six degrees of freedom electro-magnetic sensor that tracks the position and orientation of a small receiver relative to a fixed transmitter. The vision system exploits a standard webcam with a CMOS image sensor with 640x480 resolution and USB 2.0 compatible connection. The camera has 30fps nominal refresh rate. In the current hardware set-up the actual measured frame rate is 12fps. For evaluation of gesture imitation, sensor devices have been configured to collect information about joint angles of the human teacher as described in section 4. In particular, the camera has been located in a fixed position to observe the human upper body. Figure 1 (right image) shows the RSV2 in its augmented configuration used for walking imitation. The webcam has been mounted on the head of the robot to minimize negative effects on its balance. This solution provides an additional advantage as the pan-tilt head motors can be exploited to rotate the camera. Camera rotation along two axes is crucial for enhancing the robot marker recognition capabilities, as it will be pointed out in section 5. The Fastrak receiver has been mounted on the back of RSV2 and is used for global localization. The USB-UIRT device has been mounted on RSV2 as well. The device is located on the torso and, being very close to the RSV2 infrared receiver, it helps in avoiding packet loss in transmission of motor commands.

3.3 Software

The application which runs the imitation system has a multithreaded architecture and has been built upon two software libraries, namely the Virtual Hand Toolkit (VHT), and the ARToolKit (ART). The main thread is devoted to sensor data acquisition from the devices, visual processing, and computation of the motor commands for the Robosapien. A second thread is in charge of sending the motor commands to the humanoid robot through the IR transmitter. The Virtual Hand Toolkit is the application development component of the VirtualHand Suite 2000 by Immersion Corporation. VHT includes a Device Configuration Utility used to initialize and calibrate the devices and a Device Manager for data acquisition. The toolkit offers high level functionalities and low level translators for specific I/O devices such as the Cybertouch and the Fastrack. ARToolKit [16] is an open-source multiplatform software library designed for augmented reality applications. The library exploits a pattern recognition algorithm and allows overlaying of computer graphics images on the video stream captured by a camera in real-time. In our experiments we used the toolkit to recognize markers on the ground acting as via points, markers attached to static obstacles and markers attached to the body of the human demonstrator for gesture imitation.

The overlay option was enabled for testing the effectiveness of the recognition algorithm. In particular, ARToolKit uses computer vision techniques to compute the real camera position and orientation relative to markers, as will be described in section 4. Markers are squares of known size (8×8cm in the proposed experiments). One snapshot of each marker must be provided to the system in advance as training pattern. The recognition algorithm is a loop which consists of different steps. First the live video image is binarized using a lighting threshold. Then the algorithm searches for square regions. The squares containing real markers are then identified and the corresponding transformation matrix is computed. OpenGL is used for setting the virtual camera coordinates and drawing the virtual images.

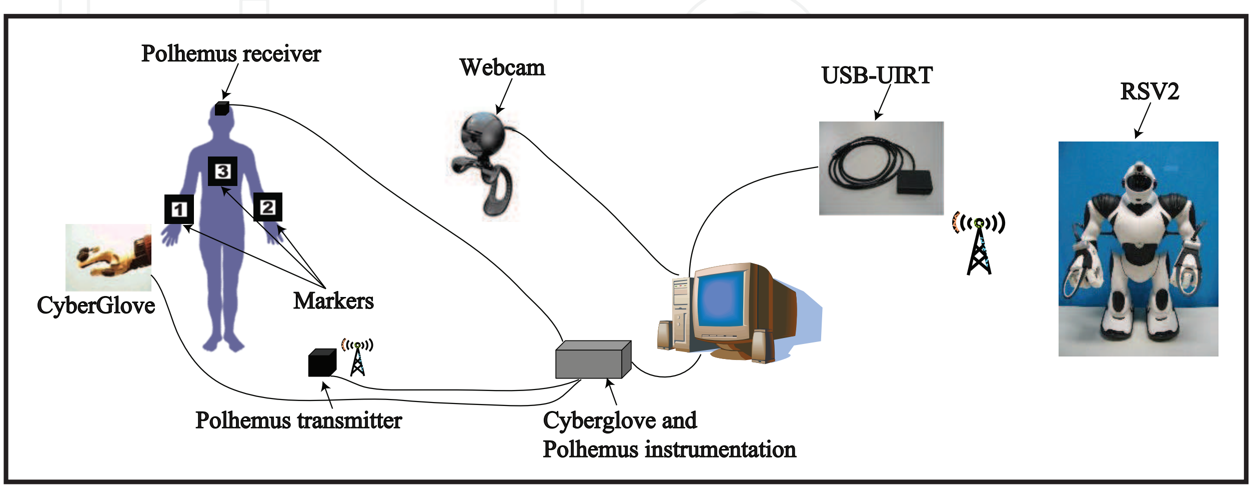

Fig. 2. System architecture for gesture imitation.

4. Imitation of human gestures

In this section we present the proposed imitation strategy for mapping the movements of the human demonstrator to RSV2. In the current setup six degrees of freedom have been considered, namely one degree of freedom for the two arms plus wrist rotation of the right arm, one degree of freedom for torso yaw and two degrees of freedom for head pitch and yaw. The associated joints can be controlled independently. The imitation strategy consists in a proportional real-time mapping between each computed sensor data and the corresponding joint. The system requires the demonstrator to perform an initial calibration routine to identify the range of motion for each movement. Each range is discretized into as many intervals as the number of feasible discrete configurations of the corresponding joint. Due to the kinematics limitations of RSV2, the resolution of the joints is quite limited and never goes beyond 8 intervals as remarked in section 4.4. After the calibration phase RSV2 starts imitating the gestures of the human demonstrator. The user can decide to play with all the six degrees of freedom concurrently or with a restricted subset by deactivating some of the sensor inputs.

Figure 2 shows the structure of the system used for gesture imitation and highlights the hardware components along with the sensors and their interconnections. Markers are attached to the body of the human demonstrator and the camera tracks the motion of the tags. The motion tracker is used to detect the motion of the head.

4.1 Arms and upper-body visual tracking

The vision system was used to track the arms motion of the human demonstrator along with the upper body. One marker is attached to each arm and a third marker is attached to the human torso. The vision algorithm is able to identify and track the geometrical configuration of multiple markers concurrently. Hereafter the algorithm used by ARToolkit for estimating the transformation matrix of markers is described. The algorithm finds for each marker the transformation T_{cm} which relates the coordinates (X_m, Y_m, Z_m) in the marker reference frame to the coordinates (X_c, Y_c, Z_c) in the camera reference frame. The rotation component is estimated by projecting the two pairs of parallel lines contouring the marker. These lines in the camera screen coordinates (x, y) can be expressed as

$$\begin{aligned} a_1x + b_1y + c_1 = 0 \\ a_2x + b_2y + c_2 = 0 \end{aligned} \tag{1}$$

From camera calibration, the projection matrix which relates the camera coordinates to the screen coordinates can be computed as follows

$$\begin{bmatrix} hx \\ hy \\ h \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & 0 \\ 0 & P_{22} & P_{23} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{2}$$

where h is a constant. By substituting (x, y) of equation 2 into equation 1 the equations of two planes in space can be found as shown next:

$$\begin{aligned} a_1P_{11}X_c + (a_1P_{12} + b_1P_{22})Y_c + (a_1P_{13} + b_1P_{23} + c_1)Z_c = 0 \\ a_2P_{11}X_c + (a_2P_{12} + b_2P_{22})Y_c + (a_2P_{13} + b_2P_{23} + c_2)Z_c = 0 \end{aligned} \tag{3}$$

The outer product of the normal vectors of these planes $n_1 \times n_2$ gives the direction vector v_1 of the two parallel sides of the marker. The two direction vectors of the two pairs of parallel lines v_1 and v_2 are then compensated for errors. The outer product $v_1 \times v_2$ provides the rotation component of the transformation matrix. The translation component of T_{cm} is obtained by the correspondence between the coordinates of the four vertices of the marker in the marker coordinate frame and the coordinates of the vertices in the camera coordinate frame. The z component of the translation vector is used for tracking the movement of the arms of the human demonstrator. The rotation component R of the transformation matrix is used to compute the orientation of the marker attached to the torso. The Euler angles are extracted from R as shown in equation 4 (giving two possible solutions). The yaw angle of the marker is set equal to θ_1. The yaw angle is then mapped to the joint controlling the yaw rotation of the upper body of RSV2.

$$\begin{aligned} \theta_1 &= -\mathrm{asin}(R_{31}) \\ \theta_2 &= \pi - \theta_1 \\ \varphi_1 &= \mathrm{atan2}\left(\frac{R_{32}}{\cos\theta_1}, \frac{R_{33}}{\cos\theta_1}\right) \\ \varphi_2 &= \mathrm{atan2}\left(\frac{R_{32}}{\cos\theta_2}, \frac{R_{33}}{\cos\theta_2}\right) \end{aligned} \tag{4}$$

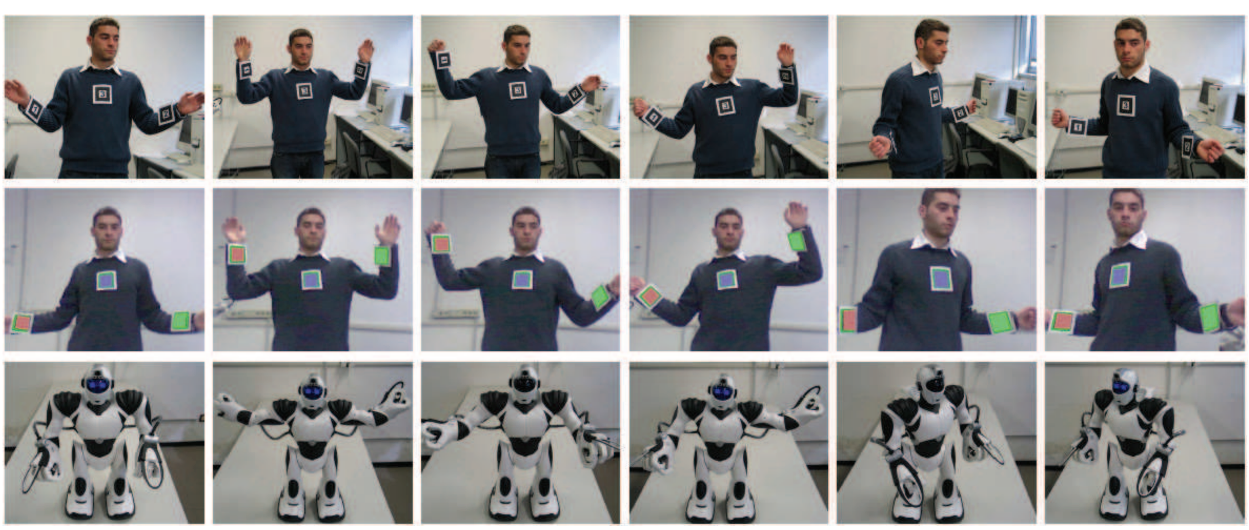

Fig. 3. Human gestures (first row), results after video processing (second row) and corresponding RSV2 configurations (third row).

4.2 Head motion tracking

The motion of the head of the human demonstrator is tracked using the FasTrack. The tracker receiver is attached to the nape of the human. The transformation matrix from the receiver frame (rx) to the world reference frame (W) is given by

$$T_{W\leftarrow rx}(t) = T_{W\leftarrow tx} T_{tx\leftarrow rx}(t) \tag{5}$$

where $T_{W\leftarrow tx}$ is the transformation between the transmitter frame to the world frame. The pitch and yaw angles of the marker are extracted from the rotation component R of $T_{W\leftarrow rx}$ and set equal to θ_2 and φ_2 respectively.

4.3 Head motion tracking

The rotational motion of the right hand wrist of the human operator is linearly mapped to the wrist of the right hand of RSV2. The motion is measured by the wrist sensor of the CyberGlove. The wrist angle γ of the CyberGlove is computed using the relation

$$\gamma = G * (\text{Digital\_Value} - O) \tag{6}$$

where G and O are the gain and offset calibration values.

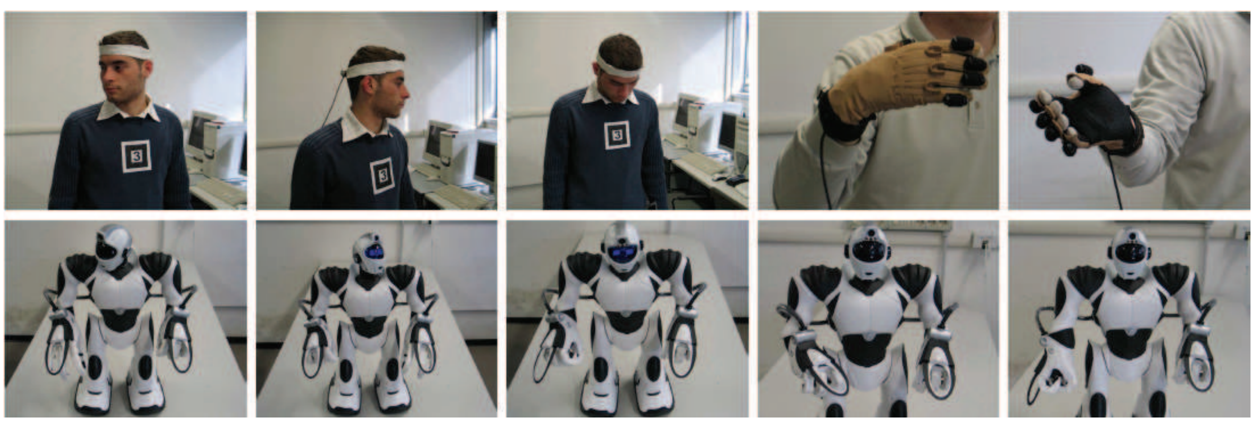

Fig. 4. Imitation of head and right wrist movements.

4.4 Results

Experiments involving gesture imitation were carried out to test the effectiveness of the system. Figure 3 shows RSV2 imitating arms and upper-body movements by visual tracking of trained patterns attached to the user's body. The same figure shows also the results of the video processing algorithm. Three squares with different colors (or gray shadings) are superimposed to the recognized markers. Figure 4 shows an experiment involving the imitation of head and wrist movements.

The evolution of the command signals sent to the humanoid robot compared to the sensed data for each degree of freedom has been reported in figure 5. The curves representing the command signals are step functions obtained by sampling the input data signals. The cardinality of the set of steps of each curve equals the resolution of the corresponding degree of freedom of the robot. For example the arms and the head yaw movements have the highest resolution with 8 and 12 steps respectively. It can be seen from figure 5 that the robot degrees of freedom track with acceptable fidelity commanded values, up to the limited resolution available.

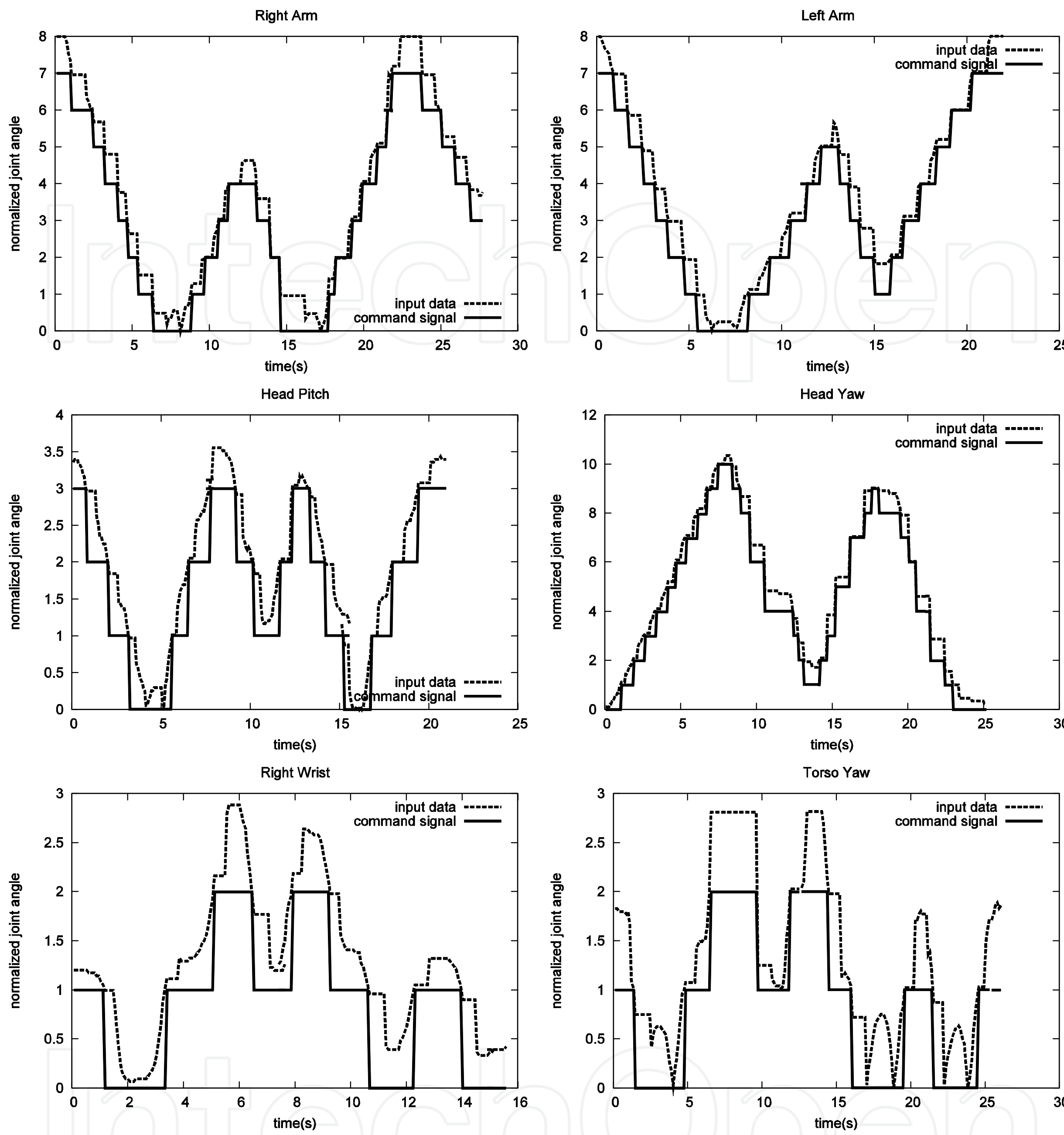

Fig. 5. Evolution of the command signal compared to the sensed data for each degree of freedom.

5. Walking imitation

Figure 6 shows the experimental setup used for walking imitation tasks. The system comprises a demonstration phase, where the user demonstrates a walking path, and an imitation phase, where the robot follows the provided trajectory. The setup also includes a fixed ceiling camera for initial recognition of the topology of the markers on the ground. In the first experiment a simple walking imitation task has been tested where the robot has to follow a taught route. In the demonstration phase a human performs a walking path while his movements are tracked by the Fastrak receiver, which is attached to one of the legs of the demonstrator. After the demonstration phase RSV2 imitates the demonstrated path relying only on the tracker sensor without visual guidance. The transformation matrix from the Fastrak receiver frame (rx) to the world reference frame (W) is given by equation 5. Figure 7 shows a sequence of images taken from both the demonstration phase and the imitation phase. The demonstrated path is a smooth U-turn around a static obstacle. Figure 8 reports the evolution of RSV2 walking trajectory compared to the demonstrated trajectory. The demonstrated trajectory is approximated as a NURBS (Non Uniform Rational B-Spline) curve (Piegl, 1991). The systems computes a piecewise linear approximation of the NURBS generating a set of viapoints and RSV2 follows the approximated path through a sequence of rotations and translations. The computed mean error of the path following task is about 15cm. The demonstrated path, shown in figure 8, appears deformed on the right side due to sensor inaccuracies of the tracker, which increase as the distance between the transmitter and the receiver increases. The same inaccuracies are the main cause of the worsening of path following performance in the last stage of the task after the U-turn.

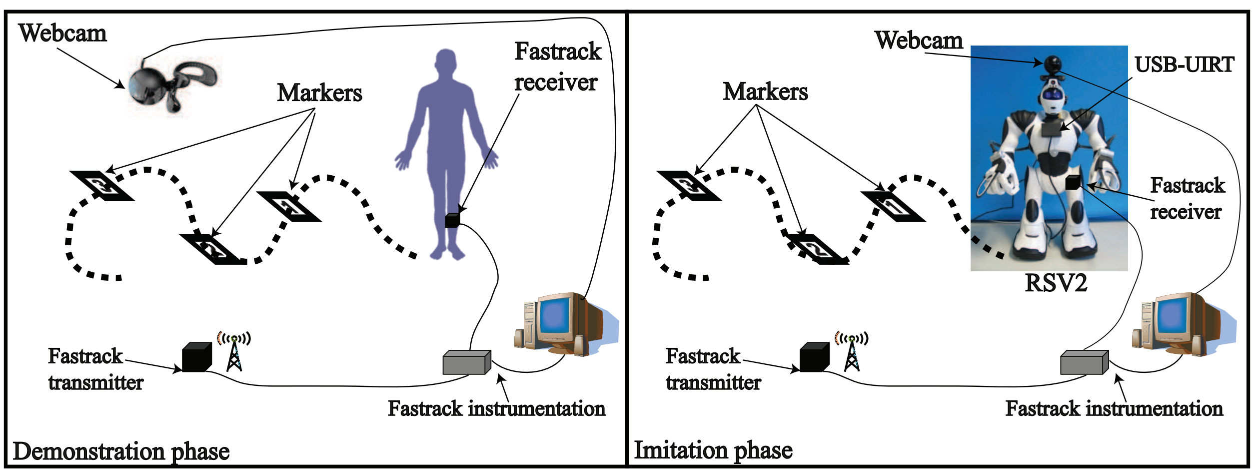

Fig. 6. System architecture for walking imitation.

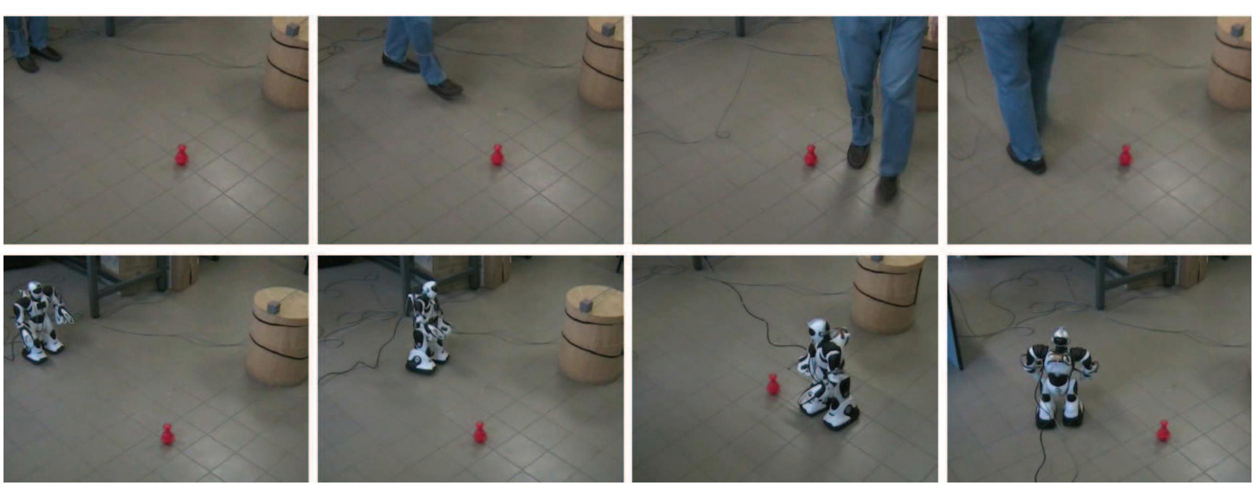

Fig. 7: Walking experiment 1: demonstration phase (top row) and imitation phase (bottom row).

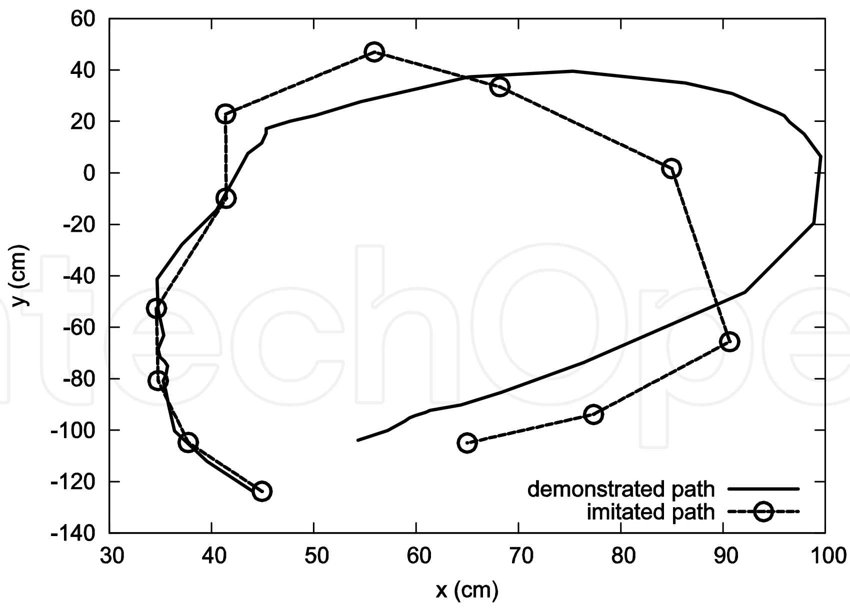

Figure 8: Walking experiment 1: demonstrated path and imitated path.

5.1 Dynamic obstacle avoidance

In a second experiment, a vision-guided imitation task was performed which requires dynamic obstacle avoidance. After the demonstration phase an obstacle is positioned on the original path. RSV2 is programmed to follow the path and to scan the environment for possible obstacles by turning its head whenever it reaches a via-point. A marker is attached to the obstacle for detection. Figure 9 shows a sequence of images taken from the imitation phase of the task. Figure 10 reports the demonstrated path together with the replanned path and the actual imitated path (the obstacle is detected when the robot reaches the second via-point). Figure 11 shows the output of the onboard camera when the marker is recognized, with a colored square superimposed to the recognized marker. The algorithm used for replanning the trajectory performs a local deformation of the original NURBS around the obstacle. Initially the NURBS is resampled and a deformation is applied to each sample point $(x(u_i), y(u_i))$ which falls within a certain distance from the obstacle according to the following equation

$$\begin{cases} \tilde{x} = x(u_i) + \alpha e^{-\frac{d^2}{\beta}}(x(u_i) - x_m) \\ \tilde{y} = y(u_i) + \alpha e^{-\frac{d^2}{\beta}}(y(u_i) - y_m) \end{cases} \qquad (7)$$

where α and β are constants. The deformation stretches the samples with a gaussian modulation factor which depends on the distance d between the robot and the obstacle. The value of d is approximated by the measured distance d_{cm} between the camera and the marker which are approximately orthogonal. The obstacle coordinates (x_m, y_m) are set equal to the coordinates of the marker in the world reference frame and are given by

$$\begin{cases} x_m = x_r + d\cos\theta_m \\ y_m = y_r + d\sin\theta_m \end{cases} \qquad (8)$$

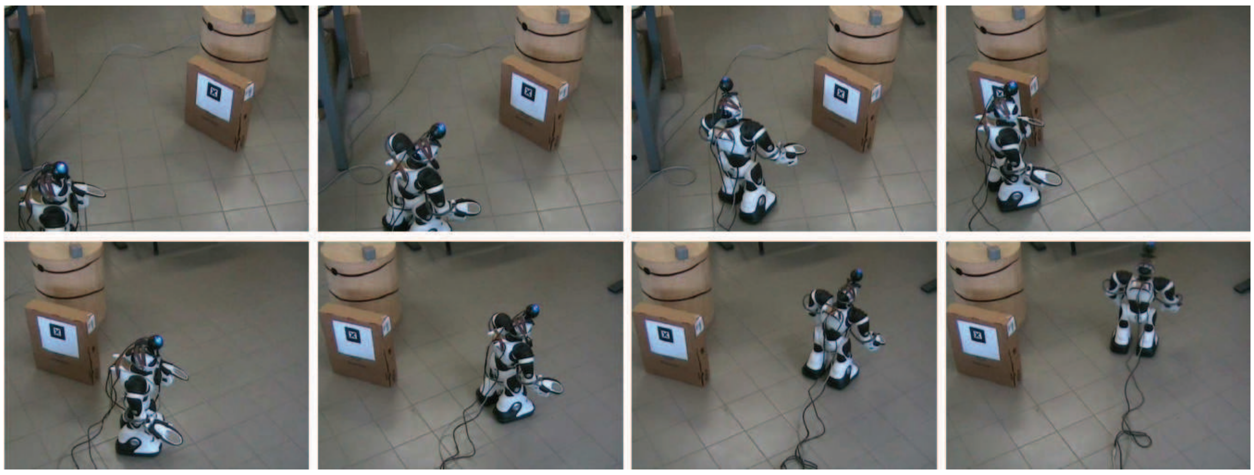

Figure 9: Walking experiment 2: imitation phase with dynamic vision-guided obstacle avoidance.

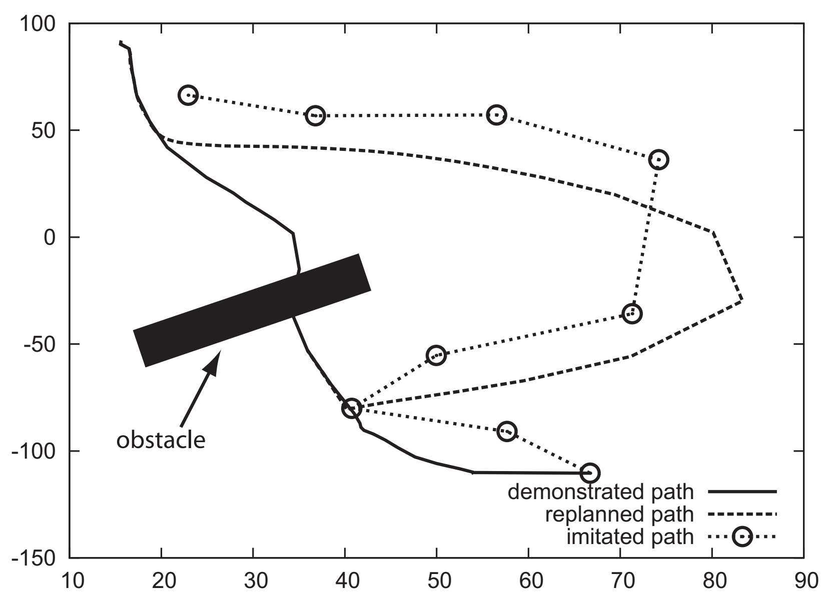

Figure 10: Walking experiment 2: demonstrated path, replanned path and imitated path.

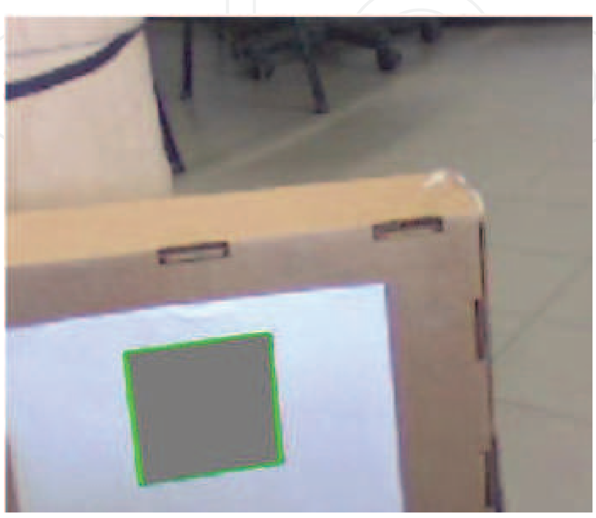

Figure 11: Walking experiment 2: obstacle detection.

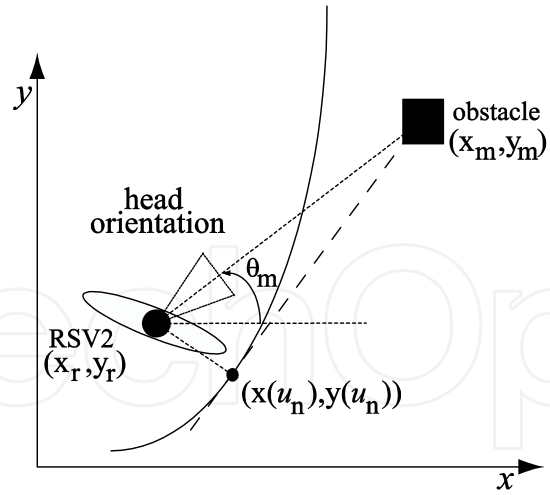

Figure 12: Walking experiment 2: approximation of obstacle coordinates.

where (x_r, y_r) is the current position of the robot and θ_m is the angle between the x-axis and the obstacle, which depends on both RSV2 body and head rotation. The angle θ_m can not be reliably computed, as the actual head rotation is hard to estimate due to the non ideal alignment between the head and the body when the robot stops. Therefore, the obstacle coordinates are approximated as follows

$$\begin{cases} x_m \approx x(u_n) + d_{cm} \dfrac{\dot{x}(u_n)}{\sqrt{\dot{x}^2(u_n) + \dot{y}^2(u_n)}} \\ y_m \approx y(u_n) + d_{cm} \dfrac{\dot{y}(u_n)}{\sqrt{\dot{x}^2(u_n) + \dot{y}^2(u_n)}} \end{cases} \tag{9}$$

where $(x(u_n), y(u_n))$ is the point on the original NURBS nearest to the current position of the robot. This means that the obstacle is assumed to lie on the tangent to the curve at the closest point to the robot.

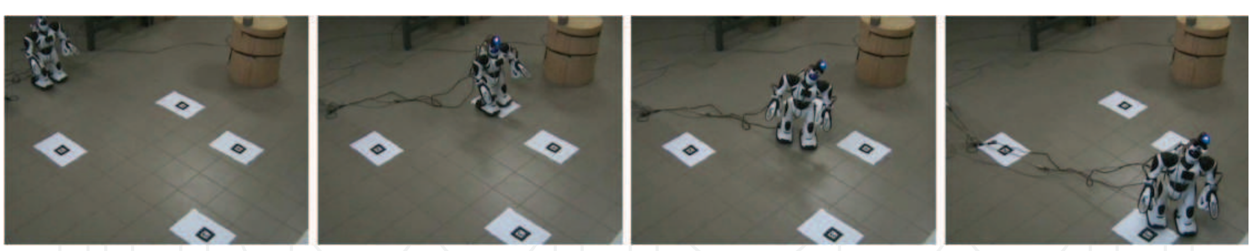

Figure 13: Walking experiment 3: imitation phase with vision guided marker following.

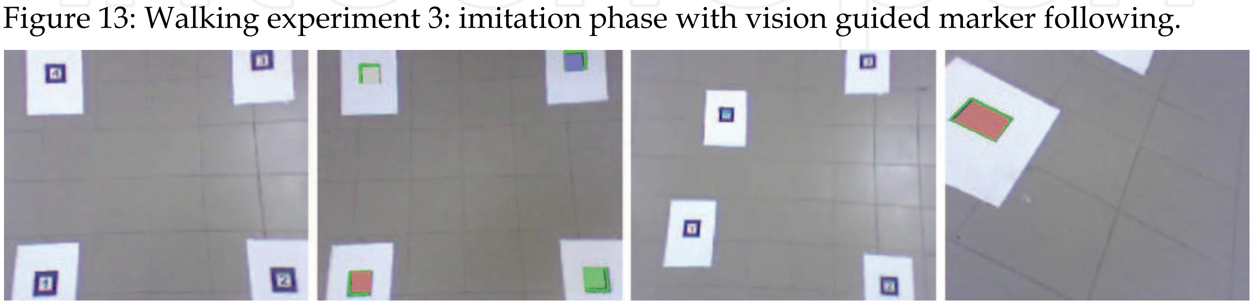

Figure 14: Walking experiment 3 form left to right: initial marker configuration, marker detection from ceiling camera, modified marker configuration and marker detection from onboard camera.

5.2 Vision-guided landmark following

In a third experiment another vision-guided imitation task has been investigated. A set of markers is positioned on the ground of the workspace and a fixed ceiling camera detects the initial configuration of the markers (the estimation error is approximately 7cm). The demonstration consists of a walking task where a human reaches a sequence of markers. The correct sequence of reached markers is computed from the measured trajectory and from the topology of the markers. After the demonstration phase, the markers are slightly moved into a new configuration and RSV2 is programmed to reach the demonstrated sequence of markers. In the proposed experiment the task is to reach marker 1 and then marker 3. Figure 13 reports images of the imitation phase. Figure 14 shows the initial marker configuration, the modified marker configuration, and the marker detection from onboard camera. Figure 15 shows the demonstrated path, the configuration of the markers and the imitated path. RSV2 imitates the reaching task by following the demonstrated path. Once the robot detects a marker on the floor it tries to approach the marker through a sequence of small movements by estimating the position and the distance of the target with the onboard camera with equation 10. The estimation error is less critical than in experiment 2 since the robot performs small movements while approaching the marker as shown in figure 15.

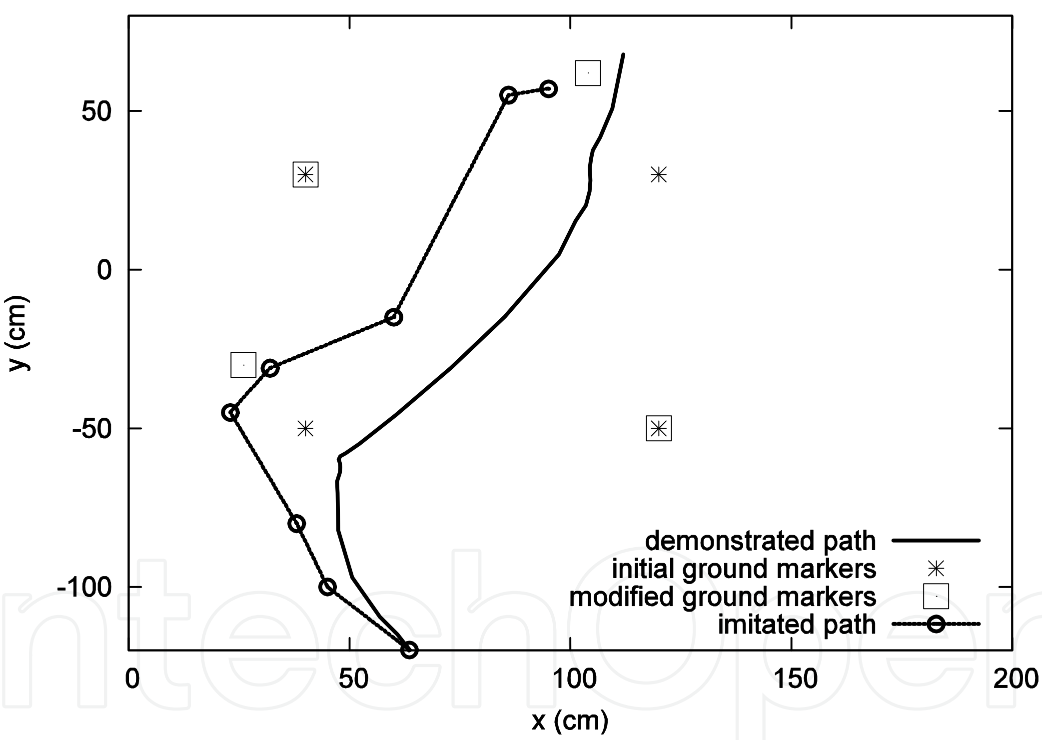

Figure 15: Walking experiment 3: demonstrated path, marker positions and imitated path.

5.3 Vision-guided grasping

RSV2 autonomous grasping skills have been evaluated in a further experiment where the task is to reach a landmark on the floor indicating the position of a set of graspable objects. The robot starts approximately 80cm away from the landmark and in its initial configuration the head and the body are properly aligned. Hence RSV2 is able to detect the landmark using the onboard camera with a tolerable error and moves towards the objects. Once the robot gets closer to the landmark it tries to reach a proper alignment with the target with small rotations and translations. Figure 16 shows the execution of the grasping task while figure 17 shows the marker detection phase at two different configurations. Due to the limitations of RSV2 the grasping procedure is not controllable and allows only to lower the upper body and one arm of the robot trying to pick up objects which are located near the foot of the robot. To overcome these kinematics limitations of the toy and the sensor errors, multiple objects have been placed near the landmark for redundancy. However, the rate of successful grasping operations is still low (approximately 40%). This result suggests that improved grasping capabilities might be required to exploit a low-cost humanoid for experimental research involving grasping.

Figure 16: Walking experiment 4: vision guided grasping task.

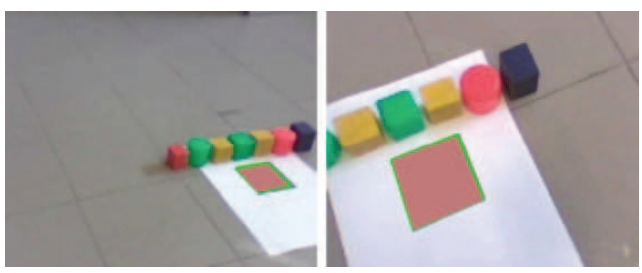

Figure 17: Walking experiment 4: marker detection.

6. Towards a complete low-cost platform

In the previous sections we have presented a system for humanoid task programming based on imitation which is focused on the use of a low-cost robot. The cost of the robot is only few hundreds of dollars, which is far less than the cost of more advanced humanoids. However, it is still necessary to discuss whether the proposed platform can be considered as a fully low-cost platform as a whole. The software developed is mostly open-source, except for low-level drivers connecting with glove and tracker devices which could be replaced by functionally equivalent open-source drivers. Additional costs arise from the sensor devices used for imitating human gestures and walking path. The USB camera and the USB-UIRT device are also very cheap (about 50$ each). By contrast, the Polhemus FasTrack and Immersion CyberTouch are high quality measurement devices which come at high price. Therefore, the developed platform cannot be considered as a fully low-cost system.
We are therefore investigating alternative tracking devices that can be exploited for motion capture. In this section we report our initial investigation on the exploitation of a Nintendo Wii remote controller for gesture recognition. This device costs less than 100$ and hence would keep the entire system low-cost.
A peculiar feature of the Wii console is its main wireless controller, the Wiimote, and the secondary controller, the Nunchuk, which can be used in conjuction to detect acceleration

and orientation in three dimensions (figure 18, left image). The Wii controllers use a combination of accelerometers for tracking human motion (the Wiimote also supports infra-red sensors to detect pointing tasks). The controllers support Bluetooth connection which can be used to connect the device to a normal PC. Figure 18 (right image) shows RSV2 imitating upper arm gestures in a similar experiment to the one presented in figure 3. The user is handling the Wii controllers in her hands and the system detects the posture of the arms (one degree of freedom for each arm) by sensing accelerations. Although the accuracy of the Nintendo controller is not comparable to the high accuracy of the FasTrack, we believe that the use of such a simple wearable device can lead to interesting results, especially if employed for entertainment or education.

Figure 18: Nintendo Wii controllers (left image) and imitation experiment (right image).

7. Conclusion

In this chapter, a robot programming by demonstration system oriented to imitation of human gestures for a humanoid robot has been presented, along with a path specification system based on imitation of human walking paths. The experimental system investigated comprises a Robosapien V2 humanoid and includes multiple sensor devices such as an electromagnetic tracker and a monocular vision system. The novelty of the approach is the investigation of a high level programming paradigm such as imitation with a low cost humanoid toy robot. Indeed, the ability to teach elementary motions is a prerequisite step toward programming more structured tasks. Vision guided trajectory imitation and dynamic obstacle avoidance have also been successfully experimented, along with autonomous grasping skills. The experiments show that the ability to teach motion paths enables a toy robot to achieve rather complex navigation tasks. The proposed imitation approach is quite general, even though its implementation is constrained by some limitations of RSV2 and by sensor inaccuracies. The usable workspace is currently restricted to a square of approximately 2m^2 due to the range resolution of the tracking device. Moreover, the tilting motion of the upper body while the robot is walking prevents the use of a continuous visual feedback. Therefore the vision system can be exploited only when the robot is not moving. Finally, the non ideal alignment between the head and the body of the robot when it stops strongly affects the estimation of landmarks pose used for visual guidance. The kinematics limitations of RSV2 also affect the imitation accuracy. Such limitations include the difficulty of performing sharp turns, the absence of side steps commands and the simplified grasping abilities. Nonetheless, we believe that low-cost humanoid platforms such as RSV2 provide an exciting and affordable opportunity for research in humanoid programming based on imitation.

8. Acknowledgment

This research is partially supported by Laboratory LARER of Regione Emilia-Romagna, Italy.

9. References

A. Arakawa, M. Hiyama, T. Emura, and Y Kagami. Trajectory Generation for Wheeled Mobile Robot Based on Landmarks. In Proceedings of the International Conference on Systems, Man and Cybernetics, Vancouver, Canada, Oct. 1995.

J.P. Bandera, L. Molina-Tanco, R. Marfil, and F. Sandoval. A model-based humanoid perception system for real-time human motion imitation. In Proceedings of the IEEE Conference on Robotics, Automation and Mechatronics, volume 1, pages 324–329, 1-3 Dec. 2004.

S. Behnke, J. Müller, and M. Schreiber. Playing soccer with robosapien. In Proceedings of The 9th RoboCup International Symposium, Osaka, Japan, Jul. 2005.

A. Billard. Robota: Clever toy and educational tool. Robotics and Autonomous Systems, 42(3-4):259–269, 2003.

C. Breazeal and B. Scassellati. Challenges in building robots that imitate people. In K. Dautenhahn and C. Nehaniv, editors, Imitation in Animals and Artifacts. MIT Press, 2001.

S. Calinon, F. Guenter, and A. Billard. Goal-directed imitation in a humanoid robot. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Barcelona, Spain, Apr. 18-22, 2005.

Hon Nin Chow, Yangsheng Xu, and S.K. Tso. Learning human navigational skill for smart wheelchair. In Proceedings of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems, (IROS), Sept. 2002.

K. R. Dixon and P. K. Khosla. Learning by Observation with Mobile Robots: A Computational Approach. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), New Orleans, LA, Apr., 2004.

J. H. Hwang, R. C. Arkin, and D. S. Kwon. Mobile robots at your fingertip: Bezier curve on-line trajectory generation for supervisory control. In Proceedings of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems, (IROS), Las Vegas, Nevada, Oct., 2003.

K. Ikeuchi and T. Suehiro. Toward an assembly plan from observation, Part I: Task recognition with polyhedral objects. 10(3):368–385, 1994.

T. Inamura, N. Kojo, T. Sonoda, K. Sakamoto, K. Okada, and M. Inaba. Intent imitation using wearable motion capturing system with on-line teaching of task attention. In Proceedings of the IEEE-RAS Int'l Conference on Humanoid Robots, Tsukuba, Japan, Dec., 2005.

T. Inamura, Y. Nakamura, H. Ezaki, and I. Toshima. Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Seoul, Korea, May 2001.

M. Ito and J. Tani. On-line imitative interaction with a humanoid robot using a mirror neuron model. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), New Orleans, LA, Apr., 2004.

M. J. Mataric. Getting humanoids to move and imitate. IEEE Intelligent Systems, 15(4):18–24, 2000.

S. Kagami, K. Nishiwaki, J.J. Kuffner, K. Okada, M. Inaba, and H. Inoue. Vision-Based 2.5D Terrain Modeling for Humanoid Locomotion. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Taipei, Taiwan, Sept., 2003.

H. Kato and M. Billinghurst. Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. In Proceedings of the 2nd International Workshop on Augmented Reality (IWAR), Oct., 1999.

Y. Kuniyoshi, Y. Yorozu, M. Inaba, and H. Inoue. From visuo-motor self learning to early imitation a neural architecture for humanoid learning. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Taipei, Taiwan, Sept. 2003.

D. Matsui, T. Minato, K. MacDorman, and H. Ishiguro. Generating natural motion in a android by mapping human motion. In IEEE/RSJ Int'l Conference on Intelligent Robots and Systems (IROS), Edmonton, Canada, Aug., 2005.

P. Menezes, F. Lerasle, J. Dias, and R. Chatila. A single camera motion capture system dedicated to gestures imitation. In Proceedings of the IEEE-RAS Int'l Conference on Humanoid Robots, Tsukuba, Japan, Dec., 2005.

P. Michel, J. Chestnutt, S. Kagami, K. Nishiwaki, J. Kuffner, and T. Kanade. Online Environment Reconstruction for Biped Navigation. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Orlando, Florida, May, 2006.

P. Michel, J. Chestnutt, J. Kuffner, and T. Kanade. Vision-Guided Humanoid Footstep Planning for Dynamic Environments. In Proceedings of 5th IEEE-RAS International Conference on Humanoid Robots, 2005.

K. Morioka, Joo-Ho Lee, and H. Hashimoto. Human-Following Mobile Robot in a Distributed Intelligent Sensor Network. 51(1):229–237, 2004.

S. Nakaoka, A. Nakazawa, F. Kanehiro, K. Kaneko, M. Morisawa, and K. Ikeuchi. Task model of lower body motion for a biped humanoid robot to imitate human dances. In Proceedings of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems (IROS), Edmonton, Canada, Aug., 2005.

A. Nakazawa, S. Nakaoka, K. Ikeuchi, and K. Yokoi. Imitating human dance motions through motion structure analysis. In Proceedings of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems (IROS), Lausanne, Switzerland, Oct., 2002.

L. Piegl. On NURBS: A Survey. IEEE Computer Graphics and Applications, 11(1):55–71, Jan., 1991.

N.S. Pollard, J.K. Hodgins, M.J. Riley, and C.G. Atkeson. Adapting human motion for the control of a humanoid robot. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Washington, DC, May, 2002.

M. Riley, A. Ude, K. Wade, and C.G. Atkeson. Enabling real-time full-body imitation: a natural way of transferring human movement to humanoids. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Taipei, Taiwan, Sept., 2003.

S. Schaal. Is imitation learning the route to humanoid robots? Trends in Cognitive Sciences, 3(6):233–242, Jun., 1999.

A.P. Shon, K. Grochow, and R.P.N. Rao. Robotic imitation from human motion capture using gaussian processes. In Proceedings of the IEEE-RAS Int'l Conference on Humanoid Robots, Tsukuba, Japan, Dec., 2005.

L. Tang and S. Yuta. Vision Based Navigation for Mobile Robots in Indoor Environment. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Seoul, Korea, May, 2001.

S. Wanitchaikit, P. Tangamchit, and T. Maneewarn. Self-Organizing Approach for Robots Behavior Imitation. In Proceedings of the IEEE Int'l Conference on Robotics and Automation (ICRA), Orlando, Florida, May, 2006.

R. Zöllner, O. Rogalla, R. Dillmann, and M. Zöllner. Understanding Users Intention: Programming Fine Manipulation Tasks by Demonstration. In Proceedings of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems, (IROS), Sept., 2002.

Service Robot Applications
Edited by Yoshihiko Takahashi

ISBN 978-953-7619-00-8
Hard cover, 400 pages
Publisher InTech
Published online 01, August, 2008
Published in print edition August, 2008

The aim of this book is to provide new ideas, original results and practical experiences regarding service robotics. This book provides only a small example of this research activity, but it covers a great deal of what has been done in the field recently. Furthermore, it works as a valuable resource for researchers interested in this field.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821