

## Robot Learning by Active Imitation

Juan Pedro Bandera, Rebeca Marfil, Luis Molina-Tanco, Juan Antonio Rodríguez, Antonio Bandera and Francisco Sandoval  
*Grupo de Ingeniería de Sistemas Integrados, Universidad de Málaga  
 Spain*

### 1. Introduction

A key area of robotics research is concerned with developing social robots for assisting humans in everyday tasks. Many of the motion skills required by the robot to perform such tasks can be pre-programmed. However, it is normally agreed that a truly useful robotic companion should be equipped with some learning capabilities, in order to adapt to unknown environments, or, what is most difficult, learn to perform new tasks.

Many learning algorithms have been proposed for robotics applications. However, these learning algorithms are often task specific, and only work if the learning task is predefined in a delicate representation, and a set of pre-collected training samples is available. Besides, the distributions of training and test samples have to be identical and the world model is totally or partially given (Tan et al., 2005). In a human world, these conditions are commonly impossible to achieve. Therefore, these learning algorithms involve a process of optimization in a large search space in order to find the best behaviour fitting the observed samples, as well as some prior knowledge. If the task becomes more complicated or multiple tasks are involved, the search process is often incapable of satisfying real-time responses.

Learning by observation and imitation constitute two important mechanisms for learning behaviours socially in humans and other animal species, e.g. dolphins, chimpanzees and other apes (Dautenhahn & Nehaniv, 2002). Inspired by nature, and in order to speed up the learning process in complex motor systems, Stefan Schaal appealed for a pragmatic view of imitation (Schaal, 1999) as a tool to improve the learning process. Current work has demonstrated that learning by observation and imitation is a powerful tool to acquire new abilities, which encourages social interaction and cultural transfer. It permits robots to quickly learn new skills and tasks from natural human instructions and few demonstrations (Alissandrakis et al., 2002, Breazeal et al., 2005, Demiris & Hayes, 2002, Sauser & Billard, 2005).

In robotics, the ability to imitate relies upon the robot having many perceptual, cognitive and motor capabilities. The impressive advance of research and development in robotics over the past few years has led to the development of this type of robots, e.g. Sarcos (Ijspeert et al., 2002) or Kenta (Inaba et al., 2003). However, even if a robot has the necessary skills to imitate the human movement, most published work focus on specific components of an imitation system (Lopes & Santos-Victor, 2005). The development of a complete imitation architecture is difficult. Some of the main challenges are: how to identify which features of an action are important; how to reproduce such action; and how to evaluate the performance of the imitation process (Breazeal & Scassellati, 2002).

In order to understand and model imitation ability, psychology and brain science can provide important items and perspectives. Thus, the theory of the development of imitation in infants, starting from reflexes and sensory-motor learning, and leading to purposive and symbolic levels was proposed by Piaget (Piaget, 1945). This theory has been employed by several authors (Kuniyoshi et al., 2003, Lopes & Santos-Victor, 2005) to build robots that exhibit abilities for imitation as a way to bootstrap a learning process. Particularly, Lopes and Santos-Victor follow a previous work of Byrne and Russon (Byrne & Russon, 1998) to establish two modes of imitation defined in terms of what is shared between the model and the imitator (Lopes & Santos-Victor, 2005):

- Action level: The robot replicates the behaviours of a demonstrator, without seeking to understand them. The robot does not relate the observed behaviour with previously memorized ones. This mode is also called 'mimicking' by the authors.
- Program level: The robot recognizes the performed behaviour so it can produce its own interpretation of the action effect.

These modes can be simultaneously active, allowing for an integrated effect.

This chapter is focused on the development of a complete architecture for human upper-body behaviour imitation that integrates these two first modes of imitation (action and program levels). However, in order to simplify the described imitation architecture, and, in particular, to simplify the perception system, manipulated tools will not be taken into account.

Two main hypothesis guide the proposed work. The first is the existence of an innate mechanism which represents the gestural postures of body parts in supra-model terms, i.e. representations integrating visual and motor domains (Meltzoff & Moore, 1977). This mechanism provides the action level ability and its psychological basis will be briefly described in Section 2. The second hypothesis is that imitation and learning by imitation must be achieved by the robot itself, i.e. without employing external sensors. Thus, invasive items are not used to obtain information about the demonstrator's behaviour. This approach is exclusively based on the information obtained from the stereo vision system of a HOAP-I humanoid robot. Its motor systems will be also actively involved during the perception and recognition processes. Therefore, in the program level, the imitator generates and internally performs candidate behaviours while the demonstrator's behaviour is unfolding, rather than attempting to classify it after it is completed. Demiris and Hayes call this "active imitation", to distinguish it from passive imitation which follows a one-way perceive - recognize - act sequence (Demiris & Hayes, 2002).

The remainder of this chapter is organized as follows: Section 2 briefly discusses several related work. Section 3 presents an overview of the proposed architecture. Sections 4 and 5 describe the proposed visual perception and active imitation modules. Section 6 shows several example results. Finally, conclusions and future work are presented in Section 7.

## **2. Related work**

### **2.1 Action level imitation**

Action level imitation or mimicking consists of replicating the postures and movements of a demonstrator, without seeking to understand these behaviours or the action's goal (Lopes & Santos-Victor, 2005). This mode of imitation can be shared with the appearance and action levels of imitation proposed in (Kuniyoshi et al., 2003).

Psychology can help to develop the action level imitation mode in a robot. Thus, different theories have been proposed to justify the mimicking abilities presenting very early neonatal children. The innate release mechanism (IRM) model (Lorenz, 1966) can be briefly stated as the mechanism which predisposes an individual organism to respond to specific patterns of stimulation from its external environment. Thus, this model postulates that the behaviour of the teacher simply triggers and releases equivalent fixed-action-patterns (FAPs) by the imitator. Although IRM can be used to model the action level imitation, there is an important reason that makes it a bad candidate to inspire the general approach to this imitation mode on an autonomous robot. IRM denies any ontogenetic value to immediate imitation and emphasizes instead the developmental role of deferred imitation (Piaget, 1945). This implies the complete knowledge of the set of FAPs. The precise specification of this set is always complex and, at present, it has not been provided. In any case, it is clear that the range of imitated actions is wide and difficult to define. This claim has been also discussed from the psychology point of view. Research has shown that it is very probable that humans present some primitive capacity for behavioral matching at birth (Meltzoff & Moore, 1989). It is difficult to explain the imitation ability of a very early neonatal child based on its knowledge of a complete and previously established set of FAPs. Meltzoff and Moore pose two alternative explanations to the early presence of this mimicking ability in neonatal children (Meltzoff & Moore, 1977): i) the existence of an innate mechanism which represents the postures of body parts in terms integrating visual and motor domains; and ii) the possibility of creating such supra-modal representations through self exploratory "body babbling" during the fetus period. Although this self-learning stage will be really performed, it would not permit to imitate behaviours like facial expressions that the neonatal child has never seen before. Therefore, these authors theorize that neonatal imitation is mediated by a process of active intermodal mapping (AIM) (Meltzoff & Moore, 1989). AIM hypothesis postulates that imitation is a matching-to-target process. Infants' self-produced movements provide proprioceptive feedback that can be compared to the visually-perceived target. AIM proposes that such comparison is possible because the perception and generation of human movements are registered within a common supra-modal representational system. Thus, although infants cannot see their own bodies, these are perceived by them. They can monitor their own movements through proprioception and compare this felt activity to what they see. A similar hypothesis has been formulated by Maurer and Mondloch (Maurer & Mondloch, 2005), but while Meltzoff's AIM hypothesis appears to be activated as a choice made by the infant, they argue that, largely because of an immature cortex, the neonatal child does not differentiate stimuli from different modalities, but rather responds to the total amount of energy, summed across all modalities. The child is aware of changes in the pattern of energy and recognizes some patterns that were experienced before, but is unaware of which modality produced the pattern. As a result, the neonatal child will appear to detect cross-modal correspondences when stimuli from different modalities produce common patterns of energy change. Thus, the response of an infant is a by-product of what is termed neonatal synesthesia, i.e., the infant confuses input from the different senses.

Many mobile robot imitation approaches are closer to these hypothesis models, especially when the goal is not to recognize the behaviour performed by the demonstrator, but imitate it directly. Besides, research on imitation in robotics usually takes the approach of studying learning by imitation, assuming that the robot already possesses the skill to imitate successfully and in turn exploits this ability as a means to acquire knowledge. That is, it is

typically assumed the innate presence of an imitation ability in the robot. Thus, the robot in (Hayes & Demiris, 1994) tries to negotiate a maze by imitating the motion of another robot, and it only maintains the distance between itself and the demonstrator constant. The humanoid robot Leonardo imitates facial expressions and behaviours, in order to learn new skills and also to bootstrap his social understanding of others, by for example inferring the intention of an observable action (Breazeal et al., 2005). A computational architecture that follows more closely the AIM model was proposed in (Demiris et al., 1997). Experiments performed on a robot head in the context of imitation of human head movements show the ability of this approach to imitate any observed behaviour that the hardware of the robot system can afford.

In the AIM model, children may use imitation for subsequent learning; but they do not have to learn to imitate in the first place. Other authors support the hypothesis that the supra-modal representations that integrate visual and motor domains can be created by the robot through self-exploration. The biological basis of this approach can be found in the Asymmetric Tonic Neck reflex (Metta et al., 2000) which forces neonatal children to look at their hands, allowing them to learn the relationships between visual stimuli and the corresponding motor action. In the action-level imitation models described in (Lopes & Santos-Victor, 2005, Kuniyoshi et al., 2003), the robot learns the supra-modal representations during an initial period of self-exploration while performing movements as both visual and proprioceptive data are available. These representations can be learnt sequentially, resembling human development stages (Metta et al., 2000). Although this theory can satisfactorily explain the development of arm/hand imitation abilities, it is difficult to justify the neonatal children ability to imitate face expressions present at birth. The body babbling is therefore considered as a pre-imitation stage in which random experimentation with body movements is involved in order to learn a set of motor primitives that allow the neonatal child to achieve elementary body configurations (Breazeal et al., 2005).

## 2.2 Program level imitation

Robotics researchers have recognized the potential of imitation to ease the robot programming procedure. Thus, they realized that instead of going through complex programming, robots could learn how to perform new assembly tasks by imitating a human demonstrator. It must be noted that program level imitation is not always achieved from visual observation. Thus, Ogata and Takahashi (Ogata & Takahashi, 1994) use a virtual reality environment as a robot teaching interface. The movement of the demonstrator in the virtual reality space is interpreted as a series of robot task-level operations using a finite automaton model. In contrast with virtual reality, (Tung & Kak, 1995) presents a method in which a robot can learn new assembly tasks by monitoring the motion of a human hand in the real world. Their work relies on invasive sensing and can not be used easily to get accurate and complete data about assembly tasks. A more accurate method to track human hand motion is presented in (Kang & Ikeuchi, 1993). Although their work also employs a glove wired to the computer to take input from the demonstrator's hand, it uses stereo vision to improve results. One of the first examples of non-invasive teaching method is the work of Inaba and Inoue (Inaba & Inoue, 1989). This paper describes a vision-based robot programming system via a computer vision interface. Kuniyoshi et al. develop a system which can be taught reusable task plans by watching a human performing assembly tasks via a real-time stereo vision system (Kuniyoshi et al., 1994). The human instructor only

needs to perform a task in front of the system while a robot extracts task description automatically without disturbing it.

In all previously described work, the same strategy has been successfully used to allow robots to perform complex assembly tasks. This strategy can be resumed in the plan from observation (APO) paradigm proposed in (Ikeuchi & Suehiro, 1992). This passive paradigm postulates that the imitation process proceeds serially through the three stages of perception, recognition and reproduction. In a passive scheme, there is not substantial interaction between all stages, nor any relation of the perception and recognition stages to the motor systems. The motor systems are only involved in the final reproduction stage (Demiris & Hayes, 2002). Therefore, a passive paradigm implies that program level imitation should require at least an elementary level of representation, which allows for recognition of the perceived actions. The psychology basis of this passive approach can be found in the IRM model described in the previous subsection. As a learning mechanism, the IRM presents a new problem that complicates its application out of industrial assembly tasks. IRM determines that true imitation have to be novel and not already in the repertoire. Therefore, imitation is a special case of observational learning occurring without incentives, without trial and error, and requiring no reinforcement (Andry et al., 2001). Then, imitation only can provide new behaviours to the repertoire and it is not employed to improve the quality of imitated tasks or recognize known behaviours.

These claims have been discussed by neuroscientists and psychologists. While there is still some debate to define what behaviours the term imitation is exactly referring to, it is assumed that imitation is the ability to replicate and learn new skills by the simple observation of those performed by others (Billard, 2001). Thus, imitation (or program level imitation) is contracted to mimicking (or action level imitation), where imitation relies on the ability to recognize observed behaviours and not only to reproduce them by transforming sensed patterns into motor commands. In this context, experiments show that repeated imitative sessions improve imitation or recognition of being imitated (Nadel, 2004). The implication of the motor system in the imitation process defines the so-called active imitation which is biologically supported by the mirror neural system. The mirror neurons were first detected in the macaque monkey pre-motor cortex (PM), posterior parietal cortex (PPC) and superior temporal sulcus (STS) (Rizzolatti et al., 1996). Later, brain imaging studies of the human brain highlighted numerous areas, such as STS, PM and Broca (Decety et al., 2002). While the discovery of this system is certainly an important step toward a better understanding of the brain mechanisms underlying the capability of the primates to imitate, the role of the mirror neuron system as part of the general neural processes for imitation is still not completely explained.

Sausser and Billard present a model of a neural mechanism by which an imitator agent can map movements of the end effector performed by other agents onto its own frame of reference (Sausser & Billard, 2005). The model mechanism is validated in simulation and in a humanoid robot to perform a simple task, in which the robot imitates movements performed by a human demonstrator. However, this work only relies on the action level imitation (mimicking). It does not distinguish between known and novel movements, i.e. all movements are processed and imitated through the same mechanism. Therefore, there is no mechanism to improve the quality of the imitated behaviour. The passive and active paradigms are combined into a dual-route architecture in (Demiris & Hayes, 2002): known behaviours are imitated through the active route; if the behaviour is novel, evident from the

fact that all internal behaviours have failed to predict adequately well, control is passed to the passive route which is able to imitate and acquire the observed behaviour. Lopes and Santos-Victor (Lopes & Santos-Victor, 2005) propose a general architecture for action and program level visual imitation. Action level imitation involves two modules. A view-point transformation module solves the correspondence problem (Alissandrakis et al., 2002) and a visuo-motor map module maps this visual information to motor data. For program level imitation an additional module that allows the system to recognize and generate its own interpretation of observed behaviours to produce similar behaviours at a later stage is provided.

### 3. Overview of the Proposed Architecture

Fig. 1 shows an overview of the proposed architecture. The whole architecture is divided into two major modules related to visual perception and active imitation. The goal of the proposed visual perception system is the detection and tracking of the demonstrator's upper-body movements. In this work, it is assumed that in order to track the global human body motion, it is not necessary to capture with precision the motion of all its joints. Particularly, in the case of upper-body movements, it is assumed that the robot only needs to track the movement of the head and hands of the human, because they are the most significant items involved in the human-to-human interaction processes. This system works without special devices or markers, using an attention mechanism to provide the visual information. Since such system is unstable and can only acquire partial information because of self-occlusions and depth ambiguity, a model-based pose estimation method based on inverse kinematics has been employed. This method can filter noisy upper-body human postures. Running on a 850MHz PC, the visual perception system captures the human motion at 10 to 15 Hz. Finally, a retargeting process maps the observed movements of the hands onto the robot's own frame of reference. Section 4 will describe the different modules of the proposed visual perception system.

The active imitation module performs the action level and program level imitation modes. To achieve the mimicking ability, it only needs to solve the visuo-motor mapping. This mapping defines a correspondence between perception and action which is used to obtain the angle joints which move the robot's head and hands to the visually observed positions. Elbows are left free to reach different configurations. Angle joints are extracted through the use of a kinematic model of the robot body. This model includes an important set of constraints that limit the robot's movements and avoid collisions between its different body parts (these constraints are necessary, as the robot has no sensors to help in preventing collisions). The body model also determines the space that the robot's end-effectors can span. This space will be quantified to ease the memorization of behaviours. Thus, each behaviour is coded as a sequence of items of this reduced set of possible postures. In order to recognize previously memorized behaviours, the active imitation system includes a behaviour comparison module that uses a dynamic programming technique to make this comparison. Section 5 describes the proposed active imitation system.

The proposed work is inspired by the possible role that mirror neurons play in imitative behaviour. Particularly, it is related to the recent work of Demiris et al. (Demiris & Hayes, 2002, Demiris & Khadhour, 2005), Breazeal et al. (Breazeal et al., 2005) and Lopes and Santos-Victor (Lopes & Santos-Victor, 2005). Thus, the proposed system emphasizes the bidirectional interaction between perception and action and employs a dual mechanism

based on mimicking and behaviour recognition modules, as proposed in (Demiris & Hayes, 2002). However, the proposed approach does not use the adapted notion of mirror neurons to predictive forward models which match a visually perceived behaviour with the equivalent motor one. Therefore, the mirror neuron-inspired mechanism is achieved by a process where the imitator behaviours are represented as sequences of poses., Behaviours are used in the imitator's joint space as its intermodal representation (Breazeal et al., 2004). Thus, the visually perceived behaviours must be mapped from the set of three-dimensional absolute coordinates provided by the visual perception module onto the imitator's joint space. In Breazeal's proposal (Breazeal et al., 2005) this process is complicated by the fact that there is not a one-to-one correspondence between the tracked features and the imitator's joints. To solve this problem, it is proposed that the robot learns the intermodal representation from experience. In the proposed system, the imitator robot establishes this one-to-one correspondence by mapping movements of the end effectors performed by the demonstrator onto its own frame of reference (Sauter & Billard, 2005).

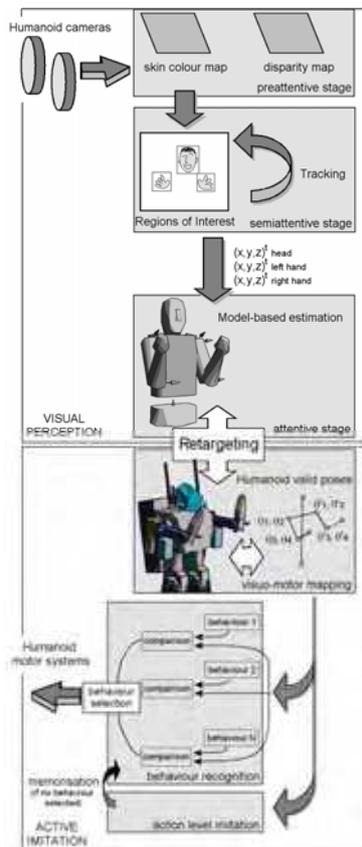


Figure 1. Overview of the proposed architecture

This transformation is achieved by a grid-based retargeting algorithm (Molina-Tanco et al., 2006). The importance given to this algorithm is influenced by the work of Lopes and Santos-Victor. These authors define an architecture based on three main modules: view-

point transformation, visuo-motor mapping and behaviour recognition (Lopes & Santos-Victor, 2005). However, their work only address postures as behaviours to imitate. In the proposed work, more complex behaviours, where the temporal chaining of elementary postures is taken into account, are addressed.

#### 4. Visual Perception System

To interact meaningfully with humans, it is interesting that robots will be able to sense and interpret the same phenomena that humans observe (Dautenhahn & Nehaniv, 2002). This means that, in addition to the perception required for conventional functions (localization, navigation or obstacle avoidance), a robot that interacts with humans must possess perceptual capabilities similar to humans.

Biological-plausible attention mechanisms are general approaches to imitate the human attention system and its facility to extract only relevant information from the huge amount of input data. In this work, an attention mechanism based on the Feature Integration Theory (Treisman & Gelade, 1980) is proposed. The aim of this attention mechanism is to extract the human head and hands from the scene. The proposed system integrates bottom-up (data-driven) and top-down (model-driven) processing. The bottom-up component determines and selects salient image regions by computing a number of different features (preattentive stage). In order to select the demonstrator's head and hands as relevant objects, skin colour has been included as input feature. Disparity has been also employed as input feature. It permits to take into account the relative depth of the objects from the observer. Similar features have been used in (Breazeal et al., 2003). The top-down component uses object templates to filter out data and track only relevant objects (semiattentive stage). The tracking algorithm can handle moving hands and head in changing environments, where occlusions can occur. To support the tracking process, the model includes weighted templates associated to the appearance and motion of head and hands. Then, the proposed system has three steps: parallel computation of feature maps, feature integration and simultaneous tracking of the most salient regions. The motivation of integrating an attention mechanism in this architecture to reduce the amount of input data is twofold: i) the computational load of the whole system is reduced, and ii) distracting information is suppressed. Besides, although in the current version of the proposed architecture, the attention mechanism only employs skin colour and depth information to extract the relevant objects from the scene, new features like colour and intensity contrasts could be easily included in subsequent versions. Thus, the mechanism could be used to determine where the attention of the observer should be focused when a demonstrator performs an action (Demiris & Khadhouri, 2005).

The outputs of the semiattentive stage are the inputs of a third module that performs the attention stage. In this work, a model of human appearance is used in the attention stage with the main purpose of filtering fast, non-rigid motion of head and hands. Besides, it can provide the whole range of motion information required for the robot to achieve the transformation from human to robot motion. To estimate articulated motion, the human model includes a 3D geometric structure composed of rigid body parts.

##### 4.1 Preattentive stage

In this work, the visual perception system is applied to track simultaneously the movements of the hands and the head of a human demonstrator in a stereo sequence. The depth of the

tracked objects is calculated in each frame by taking into account the position differences between the left and right images. The preattentive stage employs skin colour and disparity information computed from the available input image in order to determine how interesting a region is in relation to others. Attractivity maps are computed from these features, containing high values for interesting regions and lower values for other regions. The integration of these feature maps into a single saliency map provides to the semiattentive stage the interesting regions of the input video sequence.

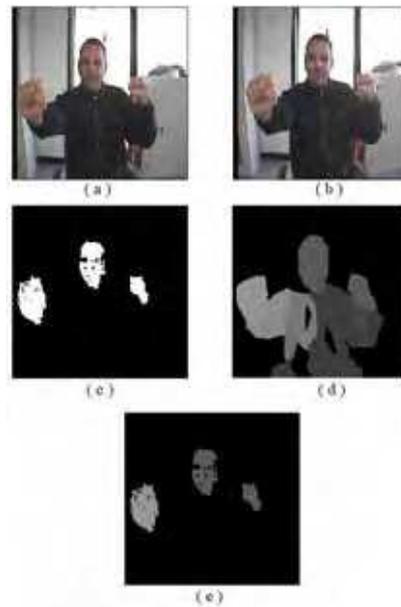


Figure 2. a-b) Input stereo pair; c) skin colour; d) disparity map; and e) saliency map

Fig. 2 shows an example of saliency map obtained from a stereo pair. In order to extract skin colour regions from the input image, an accurate skin chrominance model using a colour space can be computed and then, the Mahalanobis distance from each pixel to the mean vector is obtained. If this distance is less than a given threshold  $T_s$  then the pixel of the skin feature map is set to 255. In any other case, it is set to 0. The skin chrominance model used in the proposed work has been built over the TSL colour space (Terrillon & Akamatsu, 1999). Fig. 2b shows the skin colour regions obtained from the left image of the stereo pair (Fig. 2a). On the other hand, the system obtains the relative depth information from a dense disparity map. Closed regions, with high disparity values associated, are considered more important. The zero-mean normalized cross-correlation measure is employed as disparity descriptor. It is implemented using the box filtering technique that allows to achieve fast computation speed (Sun, 2002). Thus, the stereo correlation engine compares the two images for stereo correspondence, computing the disparity map at about 15 frames per second. Fig. 2c shows the disparity map associated to the stereo pair at Fig. 2a. Finally, and similarly to other models (Itti & Koch, 2001), the saliency map is computed by combining the feature maps into a single representation (Fig. 2d). The disparity map and the skin probability map are

then filtered and combined. A simple normalized summation has been used as feature combination strategy, which is sufficient for systems with a small number of feature maps.

#### 4.2 Semiattentive stage

The semiattentive stage tracks the head and hands of the human demonstrator, which are selected from the input saliency map. Firstly, a Viola-Jones face detector (Viola & Jones, 2001) runs on each significant region to determine whether it corresponds to a face. The closest face to the vision system is considered as the demonstrator's face. Connected components on the disparity map are examined to match the hands which correspond with this selected face (Breazeal et al., 2003). Finally, a binary image including the head and hands of the demonstrator is built. It must be noted that this process is run only as an initialization step, i.e. to search for a human demonstrator. Once the demonstrator has been found, hands and head are tracked over time.

The proposed method uses a weighted template for each object to track which follows its viewpoint and appearance changes. These weighted templates and the way they are updated allow the algorithm to successfully handle partial occlusions. To reduce the computational cost, templates and targets are hierarchically modeled using Bounded Irregular Pyramids (BIP) that have been modified to deal with binary images (Marfil et al., 2004, Molina-Tanco et al., 2005).

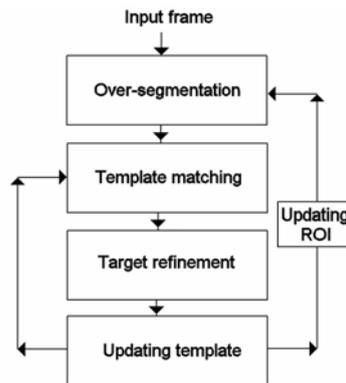


Figure 3. Data flow of the tracking algorithm

The tracking process is initialized as follows: once the demonstrator's head and hands are found, the algorithm builds their hierarchical representations using binary BIPs. These hierarchical structures are the first templates and their spatial positions are the first regions of interest (ROIs), i.e. the portions of the current frame where each target is more likely located. Once initialized, the proposed tracking algorithm follows the data flow shown in Fig. 3. It consists of four main steps which are briefly described below (see Appendix A for further details):

- Over-segmentation: in the first step of the tracking process a BIP representation is obtained for each ROI.
- Template matching and target refinement: once the hierarchical representation of the ROIs are obtained, each target is searched using a hierarchical template matching process. Then, the appearance of each target is refined following a top-down scheme.

- Template updating: as targets can represent severe viewpoint changes over a sequence, templates must be updated constantly to follow up varying appearances. Therefore, each template node includes a weight which places more importance to more recent data and allows to forget older data smoothly.
- Region of interest updating: once the targets have been found in the current frame, the new ROIs for the next frame are obtained.

### 4.3 Attentive stage

In the proposed system, our robot performs imitation and learning by imitation by itself, i.e. without employing external sensors. No markers or other invasive elements are used to obtain information about the demonstrator's behaviour. Therefore, this approach is exclusively based on the information obtained from the stereo vision system of the imitator. In order to filter the movements of all tracked items, the attentive stage employs an internal model of the human.

This work is restricted to upper-body movements. Therefore, the geometric model contains parts that represent hips, head, torso, arms and forearms of the human to be tracked. Each of these parts is represented by a fixed mesh of few triangles, as depicted in Fig. 4. This representation has the advantage of allowing fast computation of collisions between parts of the model, which helps in preventing the model from adopting erroneous poses due to tracking errors.

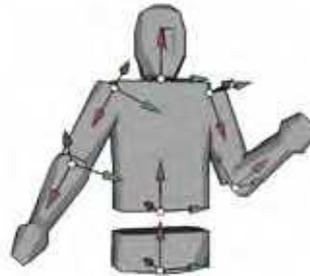


Figure 4. Illustration of the human upper-body kinematic model

Each mesh is rigidly attached to a coordinate frame, and the set of coordinate frames is organized hierarchically in a tree. The root of the tree is the coordinate frame attached to the hips, and represents the global translation and orientation of the model. Each subsequent vertex in the tree represents the three-dimensional rigid transformation between the vertex and its parent. This representation is normally called a skeleton or kinematic chain (Nakamura & Yamane, 2000) (Fig. 4). Each vertex, together with its corresponding body part attached is called a bone. Each bone is allowed to rotate --but not translate-- with respect to its parent around one or more axes. Thus, at a particular time instant  $t$ , the pose of the skeleton can be described by  $\Phi^{(t)} = (R^{(t)}, \vec{s}^{(t)}, \phi^{(t)})$  where  $R^{(t)}$  and  $\vec{s}^{(t)}$  are the global orientation and translation of the root vertex, and  $\phi^{(t)}$  is the set of relative rotations between successive children. For upper-body motion tracking, it is assumed that only  $\phi$  needs to be updated --this can be seen intuitively as assuming that the tracked human is seated on a chair.

The special kinematic structure of the model can be exploited to apply a simple and fast analytic inverse kinematics method which will provide the required joint angles from the Cartesian coordinates of the tracked end-points (see Appendix B).

#### 4.4 Retargeting

In any form of imitation, a correspondence has to be established between demonstrator and imitator. When the imitator body is very similar to that of the demonstrator, this correspondence can be achieved by mapping the corresponding body parts (Nehaniv & Dautenhahn, 2005). Thus, Lopes and Santos-Victor propose two different view-point transformation algorithms to solve this problem when the imitator can visually perceive both the demonstrator's and its own behaviour (Lopes & Santos-Victor, 2005). However, the similarity between the two bodies is not always sufficient to adopt this approach. Often the imitator's body will be similar to the demonstrator's, but the number of degrees of freedom (DOFs) will be very different. In these cases, it is not possible to establish a simple one-to-one correspondence between the coordinates of their corresponding body parts. Thus, more complex relations and many-to-one correspondences are needed to perform imitation correctly (Molina-Tanco et al., 2006). Sauser and Billard (Sauser & Billard, 2005) describe a model of a neural mechanism by which an imitator can map movements of the end-effector performed by other agents onto its own frame of reference. Their work is based on the mapping between observed and achieved subgoals, where a subgoal is defined as to reach a similar relative position of the arm end-effectors or hands. Our work is based on the same assumption.

In this work, the mapping between observed and achieved subgoals is defined by using three-dimensional grids, associated to each demonstrator and imitator hand. Fig. 5b shows the grid associated to the left hand of the demonstrator. This grid is internally stored by the robot and can be autonomously generated from the human body model. It provides a quantization of the demonstrator's reachable space. The demonstrator's reachable space cells can be related to the cells of the imitator's grids (Fig. 5c). This allows defining a behaviour as a sequence of imitator's grid elements. This relation is not a one-to-one mapping because the robot's end-effector is not able to reach to all the positions that the human's hand can reach. Thus, the proposed retargeting process involves a many-to-one correspondence that has to solve two main problems: i) how to perform re-scaling to the space reachable by the imitator's end-effectors and ii) how to obtain the function that determines the egocentric (imitator) cell associated to an observed allocentric (demonstrator) cell. The presented system solves these problems using look-up tables that establish a suitable many-to-one correspondence. Briefly, two techniques can be used to relate demonstrator and imitator cells. See (Molina-Tanco et al., 2006) for further details:

- Uniform scale mapping (USM). The length of a stretched arm for both the demonstrator and the imitator gives the maximum diameters of the corresponding grids. The relation between these lengths provides a re-scaling factor applied to demonstrator's end-effector position to obtain a point in the imitator grid. The nearest cell to this point is selected as imitator's end-effector position. Although it is a valid option, USM may distort quality of the imitated behaviour if a large part of the motion is performed in an area that the imitator cannot reach.
- Non-uniform scale mapping (NUSM). This approach is applied when it is possible to set a relation between the shapes of demonstrator and imitator grids. This relation

allows to apply a geometric transformation to all demonstrator poses that roughly translate the movements to the imitator's reachable area. NUSM improves USM results but it needs a better *a priori* knowledge about both demonstrator's and imitator's bodies. As in this case we are able to determine both human and HOAP-I reachable spaces, and it is possible to set a rough relation between them (Molina-Tanco et al., 2006), the NUSM approach is chosen.

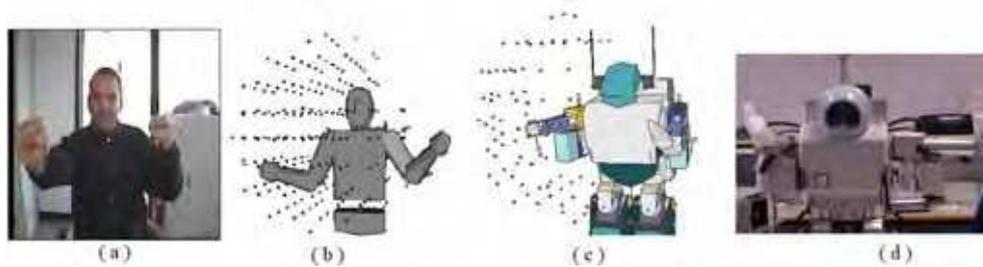


Figure 5. Overview of the retargeting process: a) human demonstrator; b) internal human model (only a sub-region of the grid is shown for clarity); c) internal humanoid model (only a sub-region of the grid is shown for clarity); and d) real humanoid imitator

## 5. Active Imitation Module

The main characteristic of an active imitation architecture is that the same motor representations that are responsible for the generation of a movement are recruited in order to perform behaviour recognition (Demiris & Hayes, 2002). Thus, the active imitator does not execute the passive perception -- recognition -- action cycle, but actively generates possible behaviours concurrently with the perception of the demonstrator's behaviour. The most similar behaviour to the perceived one is selected. If the observed behaviour is not recognized, it is added to the memorized repertoire. The original idea of active imitation was proposed by Demiris and Hayes (Demiris & Hayes, 2002). The same concept has been subsequently employed by other authors, e.g. (Breazeal et al., 2005) or (Lopes & Santos-Victor, 2005). The proposed approach is similar to the one presented in (Demiris & Hayes, 2002). The following section introduces the three components of the active imitation module: visuo-motor mapping, action level imitation and behaviour recognition (Fig. 1).

### 5.1 Visuo-motor mapping

The visuo-motor mapping (VMM) defines a correspondence between an observed behaviour and executed action. In the proposed system, the demonstrator hands are tracked and their coordinates are used to extract a coherent upper-body human pose using a human model. The hand poses provided by this model are translated to an egocentric image by the retargeting method described in Section 4.4. The VMM will relate the egocentric image coordinates of these hands to the actual joint angles, in terms of forward/inverse kinematics. Then, the VMM can be used to infer the motor commands used to achieve the perceived pose. This ability is subsequently used to make recognition in motor space and to imitate (Demiris & Hayes, 2002, Lopes & Santos-Victor, 2005).

Several authors propose a VMM algorithm that defines a direct translation from the imitator's end-effector coordinates to the imitator's joint angles which must be sent to the

robot motors to achieve a pose similar to the observed one (Lopes & Santos-Victor, 2005, Molina-Tanco et al., 2005). In the proposed approach, from the coordinates of the hands on the imitator's frame of reference, the imitator obtains the rest of joint coordinates using a model of its own body (Figs. 5b-d). The mechanism to obtain the joint angles from the hand positions is explained in Appendix B. The robot model (Fig. 5c) is used not only to extract the pose of the joints, but also to impose a set of constraints that limit the robot's movements and avoid collisions between its different body parts. The joint limits can be taken into account by the motor control software, but there is no internal mechanism to avoid the collision between two robot parts. Although the implementations are very different, the proposed approach can be related to the free-elbow method proposed in (Lopes & Santos-Victor, 2005).

### 5.2 Action level imitation

The action level imitation or mimicking can be directly achieved by sending to the robot motor controllers the joint angles provided by the VMM. However, in the proposed imitation architecture, the action level imitation module is also the responsible of generating the representation of every observed behaviour. This representation will be memorized if it is not present in the actual repertoire of behaviours.

The behaviour is not exactly memorized as it was observed. All the different modules of the architecture introduce some errors, resulting in noisy observed behaviour. Therefore, the action level imitation module performs outlier removal during the acquisition process. Outliers are removed using local windows centered at each element of the observed sequence. They are detected as small size clusters on the sequence. Fig. 6a shows the trajectory for an observed behaviour when outlier removal has not been performed. It can be noted that the movement trajectory is noisy. On the contrary, Fig. 6b illustrates the same behaviour after outlier removal. This filtered representation will be memorized and included into the behaviour repertoire.

Once outliers have been removed, the behaviour is stored as a sequence of transitions between different elements of the grid. Self-transitions are discarded.

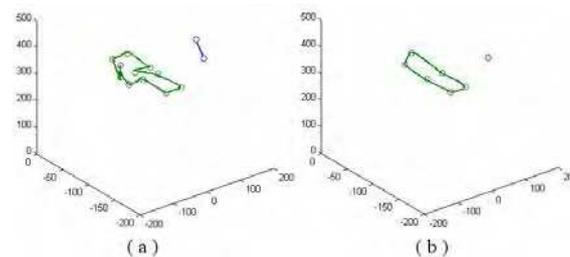


Figure 6. a) Trajectories of the end-effectors for an observed behaviour; and b) trajectories for the same behaviour in a) after outlier removal.

### 5.3 Behaviour recognition

When comparing an observed behaviour of arbitrary duration to the repertoire of memorized behaviours, time must be normalized out. In this work, an online version of a classical time normalization algorithm based on dynamic programming, which is commonly

known as Dynamic Time Warping (DTW) (Sakoe & Chiba, 1978), is employed. DTW simultaneously calculates the optimal time warping function and the resulting minimum distance between two patterns.

Demiris and Hayes introduced the concept of confidence value, as an indicator of how confident an imitator's behaviour is that it can match an observed behaviour (Demiris & Hayes, 2002). In the proposed system, the confidence value for an observed behaviour is inversely proportional to the distance between the observed and memorized behaviours.

#### 5.4 Behaviour learning

When the human demonstrator performs a behaviour that the imitator does not know, none of the memorized behaviours' confidence value stands out. This behaviour is therefore considered as novel, and subsequently added to the behaviour repertoire, i.e. learnt. In this work, it is assumed that learning is the process of acquiring a novel behaviour, either its trajectory specifications or the motor commands needed to achieve it (Demiris & Hayes, 2002). Learning as used here does not imply generalization or adaptation to different circumstances or any other processes as used in the field of machine learning. Once the novel behaviour is learnt, it can be recognized in subsequent performances of the future demonstrators.

### 6. Experimental Results

The first experiments reported here show that the proposed architecture can provide mimicking abilities to our HOAP-I humanoid robot. HOAP-I is a mini-humanoid robot built by Fujitsu, provided with 20 degrees of freedom, including 4 on each arm. Fig. 7 illustrates the trajectories of the demonstrator's and imitator's hands for several examples. It can be observed that the imitation is qualitatively good although it suffers a non-uniform scale mapping to bring observed motion into the reachable space of the robot. Using this mimicking ability, we have generated a set of behaviours which will constitute the behavioural repertoire of the robot. Behaviours in Fig. 7 constitute the set of memorized trajectories. For this experiment we have chosen the trajectories to correspond to different diving signals. Each signal conveys a message to fellow divers: '(It is) Cold', '(I am) OK', etc. The meaning of each signal is illustrated Fig. 7. In the rest of this Section, the abbreviated terms are used: [Cold], [OK], etc.

In the second experiment, different demonstrators perform behaviours of a reduced repertoire that the robot has memorized. Table 1 shows the confidence values associated to these demonstrations. Note that this is a much more challenging task for the robot, and thus we chose a reduced signal repertoire. No feedback was provided to the demonstrators as to how 'well' the diving signals were performed. The movements were described verbally to the performers, and they had not seen the motion of previous demonstrators nor had information about their results during the experiment. The main consequence of these restrictions is that some movements were performed in quite different ways by different demonstrators. Particularly, the [Help] behaviour in which the extended right arm moves up and down generated some confusion, as users did not know relative position of the hand nor amplitude of movements. Table 1 shows that this situation makes the system consider the unrecognized motion as a new behaviour, and consequently it is stored again in the data base. The rest of the movements were correctly recognized as Table 1 depicts.

		1. Performer #1				Performer #2				Performer #3			
		2. [Cold]	3. [OK]	4. [Help]	5. [Well]	[Cold]	[OK]	[Help]	[Well]	[Cold]	[OK]	[Help]	[Well]
Stored behaviours	[Cold]	<b>1.00</b>	0.38	0.16	0.26	<b>0.69</b>	0.55	0.31	0.24	<b>0.89</b>	0.56	0.35	0.26
	[OK]	0.39	<b>1.00</b>	0.20	0.47	0.44	<b>0.75</b>	0.46	0.39	0.52	<b>0.79</b>	0.51	0.39
	[Help]	0.16	0.20	<b>1.00</b>	0.53	0.17	0.19	0.11	0.38	0.17	0.22	0.24	0.40
	[Well]	0.26	0.45	0.54	<b>1.00</b>	0.31	0.32	0.23	<b>0.98</b>	0.27	0.36	0.33	<b>0.97</b>
		6. Training data				7. Test data							

Table 1. Confidence values

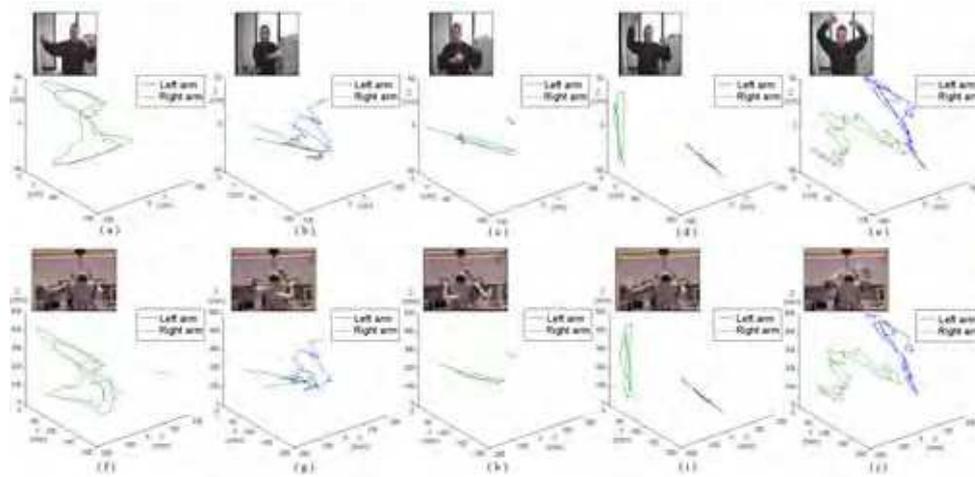


Figure 7. a-e) Trajectories followed by the demonstrator’s hands; and f-j) trajectories followed by the imitator’s hands. These behaviours also constitute the memorized repertoire: a) [Help]; b) [Cold]; c) [NoAir]; d) [GoUp] and e) [Well]

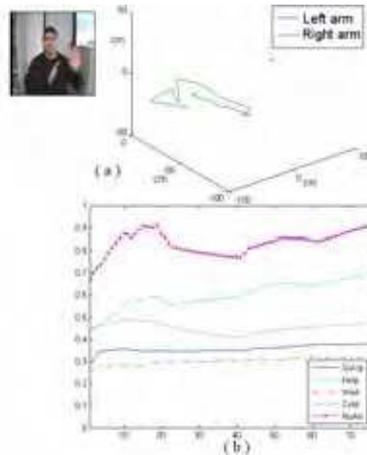


Figure 8. a) Trajectory followed by the demonstrator’s hands –behaviour [NoAir]–; and b) confidence values of memorized behaviours {[Help], [Cold], [NoAir], [GoUp], [Well]} when demonstrator executes behaviour [NoAir]

Recognition can be achieved even when only a part of the stored motion is detected by the system, as the online version of the DTW algorithm updates confidence values for each received frame. Fig. 8a shows the evolution of these confidence values while a demonstrator is executing the behaviour [NoAir]. It can be noted that the memorized behaviour [NoAir] not only consistently gets the highest confidence value at the end of the motion, but it is also the most similar behaviour during all the sequence. Thus, it could have been correctly recognized even if the communication would have been interrupted during the test. Although a result based on partial observation should be always considered with caution, it can be very useful in real interaction scenarios where the data flow is usually disturbed by external factors.

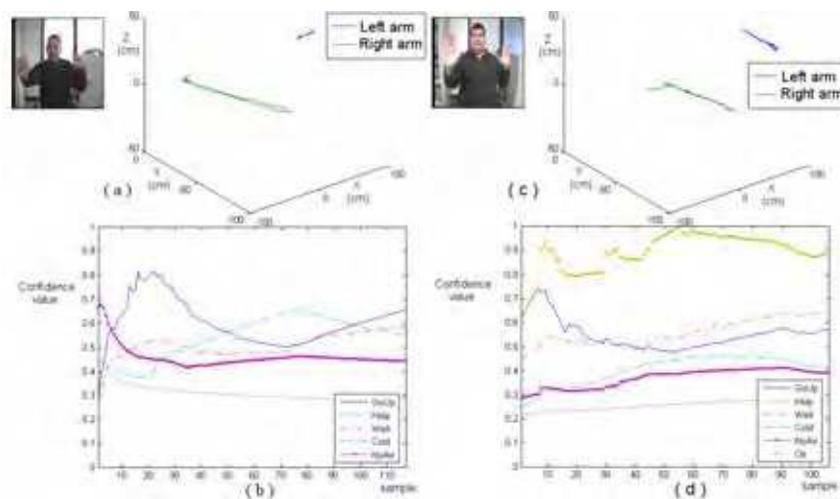


Figure 9. a) Trajectory followed by the demonstrator's hands -behaviour [OK]-; b) confidence values of memorized behaviours {[Help], [Cold], [NoAir], [GoUp], [Well]} when demonstrator executes behaviour [OK]; c) trajectory followed by the demonstrator's hands - behaviour [OK]-; and d) confidence values of memorized behaviours {[Help], [Cold], [NoAir], [GoUp], [Well], [OK]} when demonstrator executes behaviour [OK]

Finally, Fig. 9a illustrates an observed behaviour which is not in the memorized repertoire. Fig. 9b shows the confidence values for this behaviour. It can be appreciated that none of the behaviours' confidence value stands out. This triggers the learning module, which adds the novel behaviour which to the imitator's repertoire (in the experiment illustrated by Fig. 9a-b, the novel behaviour corresponds to the signal [Well]). If a different demonstrator performs the behaviour [Well] (Fig. 9c), the imitator has now this behaviour into the memorized set. Fig. 9d illustrates the confidence values. It can be appreciated that, in this occasion, the confidence value of the recently learn behaviour stands out, resulting in the behaviour being correctly recognized.

## 7. Conclusions and Future Work

In this chapter, an architecture that endows a robot with the ability to imitate has been described. This architecture has modules that provide action level and program level

capabilities. The program level imitation is achieved by a behaviour recognition module which compares previously memorized and observed behaviours. If there are no behaviours that can match sufficiently well, the action level imitation module provides a representation acquired from the observed behaviour. This representation is added to the repertoire of memorized behaviours. Experiments have been performed using a real humanoid robot HOAP-I and different human demonstrators. These experiments have showed that the architecture is able to imitate known behaviours, as well as acquiring new ones which are successfully employed later.

As behaviour complexity increases, other communication channels between demonstrator and imitator such as verbal instruction or attentional cues are required (Nicolescu & Mataric, 2005). In the proposed work, imitation learning is not augmented by allowing the demonstrator to employ additional instructive activities. Neither the learned representations are refined through generalization from multiple learning experiences, nor through direct feedback from the teacher. These items will be taken into account in future work. Current efforts focus on improving the proposed attention mechanism and including verbal communication modules in the architecture.

## 8. Appendix A. Tracking using BIPs

In the binary BIP, each node  $n$  is identified by  $(i,j,l)$  where  $l$  represents the level and  $(i,j)$  are the co-ordinates within the level. To build the different levels of the pyramid, each node has two associated parameters:

- Homogeneity,  $Hom(i,j,l)$ . This is set to 1 if the four nodes immediately underneath have homogeneity values equal to 1. Otherwise, it is set to 0. It must be noted that in the base of the structure (level 0) only nodes which correspond to image pixels of the interest regions have homogeneity values equal to 1.
- Parent link,  $(X,Y)_{(i,j,l)}$ . If  $Hom(i,j,l)$  is equal to 1, the values of the parent link of the four nodes immediately underneath are set to  $(i,j)$ . Otherwise, these four parent links are set to a null value.

Each template  $M$  and target  $T$  is represented by using binary BIP structures:

$$M^{(t)}(l) = \bigcup_{ij} m^{(t)}(i,j,l) \quad T^{(t)}(l) = \bigcup_{ij} t^{(t)}(i,j,l) \quad (1)$$

$M^{(t)}(l)$  and  $T^{(t)}(l)$  being the level  $l$  of the pyramid structure corresponding to the template and target in frame  $t$ , respectively. Each level of the template and the target are made up of a set of homogeneous nodes.

### A.1 Oversegmentation

The hierarchical representation of a region of interest  $ROI^{(t)}$  in the current frame  $t$  depends on the target position in the previous frame, and is updated as described in Subsection 8.4. The hierarchical structure can be represented in each level as:

$$ROI^{(t)}(l) = \bigcup_{ij} p^{(t)}(i,j,l) \quad (2)$$

being  $p$  a node of the bounded irregular pyramid built over the  $ROI$ .

### A.2 Template matching and target refinement

The process to localize the target in the current frame  $t$  is a top-down process which starts at a working level  $l^{(t)}_w$  and stops in the level where the target is found. In each level  $l$ , the template  $M^{(t)}(l)$  is placed and shifted in  $ROI^{(t)}(l)$  until the target is found or until  $ROI^{(t)}(l)$  is completely covered. If  $ROI^{(t)}(l)$  was completely covered and the target was not found, the target localization would continue in the level below. The displacement of the template can be represented as  $d_k^{(t)}=(d_k^{(t)}(i),d_k^{(t)}(j))$ , being  $d_0^{(t)}$  the first displacement and  $d_f^{(t)}$  the final displacement.  $d_f^{(t)}$  is the displacement that situates the template in the position where the target is placed in the current frame. The algorithm chooses as initial displacement in the current frame  $d_0^{(t)}=d_f^{(t-1)}$ . In order to localize the target and obtain  $d_f^{(t)}$ , the overlap  $O_{d_k^{(t)}}^{(t)}$  between  $M^{(t)}(l)$  and  $ROI^{(t)}(l)$  in each template displacement  $k$  is calculated:

$$O_{d_k^{(t)}}^{(t)} = \sum_{ij \in \xi} w^{(t)}(m(i, j, l_w^{(t)})) \quad (3)$$

being  $w^{(t)}(m(i, j, l))$  a weight associated to  $m^{(t)}(i, j, l)$  in the current frame  $t$ , as explained in Section 8.3.  $\xi$  is the subset of nodes that satisfy the following two conditions:

$$\begin{aligned} Hom(f(m^{(t)}(i, j, l_w^{(t)}), a(t))) &= 1 \\ Hom(p^{(t)}(i + d_k^{(t)}(i), j + d_k^{(t)}(j), l_w^{(t)})) &= 1 \end{aligned} \quad (4)$$

being  $f(m^{(t)}(i, j, l_w^{(t)}), a(t))$  a coordinate transformation of  $m^{(t)}(i, j, l_w^{(t)})$  that establishes the right correspondence between  $m^{(t)}(i, j, l_w^{(t)})$  and  $p^{(t)}(i + d_k^{(t)}(i), j + d_k^{(t)}(j), l_w^{(t)})$ .  $a(t)$  denotes the parameter vector of the transformation, which is specific for the current frame. Equation 4 is satisfied when a match occurs. All the ROI nodes that match with nodes of the template are marked as nodes of the target. Thus, the hierarchical representation of the target  $T^{(t)}$  is obtained.

In order to refine the target appearance, its hierarchical representation is rearranged level by level following a top-down scheme (Marfil et al., 2004). This process is applied to all homogeneous nodes of ROI which have not been marked as target nodes in the template matching process. If one of these nodes has a homogeneous neighbor node that belongs to the target, it is also marked as a target node.

### A.3 Template updating

In order to update the template, a new parameter is included in the template model:

- $w^{(t)}(m(i, j, l))$ . It is a weight associated to each node  $m^{(t)}(i, j, l)$  of the template  $M^{(t)}$  in the current frame  $t$ .

The whole template is updated at each sequence frame:

$$m^{(t+1)}(i, j, l) = \begin{cases} m^{(t)}(i, j, l) & \text{if no match} \\ f^{-1}(q^{(t)}(i, j, l), a^{(t)}) & \text{if match} \end{cases} \quad (5)$$

$$w^{(t+1)}(m(i, j, l)) = \begin{cases} w^{(t)}(m(i, j, l)) - \alpha & \text{if no match} \\ 1 & \text{if match} \end{cases} \quad (6)$$

where the superscript  $(t)$  denotes the current frame and the forgetting constant,  $a$ , is a predefined coefficient that belongs to the interval  $[0,1]$ . This constant dictates how fast the forgetting action will be.

#### A.4 Region of interest updating

This process has two main steps:

1. *ROI<sup>(t+1)</sup>(0) selection*: Level 0 of the new region of interest is obtained by taking into account the position where the target is placed in the original image of the frame  $t$ . Firstly, the bounding-box of  $T^{(t)}(0)$  ( $BB(T^{(t)}(0))$ ) is computed. Then,  $ROI^{(t+1)}(0)$  will be made up of the pixels of the next frame  $p^{(t+1)}(i,j,l)$  which are included in the bounding box plus the pixels included in an extra border  $\varepsilon$  of the bounding box. This extra border tries that the target in the next frame will be placed in the new ROI.

$$ROI^{(t+1)}(0) = \bigcup_{ij \in \{BB(T^{(t)}(0)) + \varepsilon\}} p^{(t+1)}(i,j,0) \quad (7)$$

2. *Over-segmentation of ROI<sup>(t+1)</sup>(0)*: The hierarchical structure  $ROI^{(t+1)}$  is built. This step is performed at the beginning of the tracking process  $t+1$  (subsection 8.1).

## 9. Appendix B. Model pose estimation

As shown in Fig. 11, each arm is modelled with a two-bone kinematic chain. The parent bone corresponds to the upper arm and is allowed to rotate around three perpendicular axes. This provides a simplified model of the shoulder joint.  $T_1^w R$  is the local transformation between the upper-arm reference frame  $O_1$  and a coordinate frame attached to the torso and centered at the shoulder joint  $w$ . The bone representing the lower arm is allowed to rotate around a single axis, corresponding to the elbow joint.  $T_1^2 R, {}^1\bar{l}_1$  denotes the local transformation between the upper-arm reference frame  $O_1$  and the lower-arm reference frame  $O_2$ , where  ${}^1\bar{l}_1 = (0,0,l_1)^T$ , being  $l_1$  the length of the upper-arm, and  ${}^2_1 R$  corresponds to the rotation  $\theta_e$  about the elbow axis.

Given a desired position for the end-point of the arm at time instant  $t+1$ ,  ${}^w\bar{p}_d^{(t+1)}$ , and given the rotation matrices  ${}^w_1 R^{(t)}$  and  ${}^1_2 R^{(t)}$  at the previous time instant  $t$ , the problem is then to find the updated matrices  ${}^w_1 R^{(t+1)}$  and  ${}^1_2 R^{(t+1)}$ . A simple geometric method is summarized here that can solve such problem. See (Mitchelson, 2003) for further details.

1. Bring  ${}^w\bar{p}_d^{(t+1)}$  within reach of the arm:

$$\text{if } \left| {}^w\bar{p}_d^{(t+1)} \right| > (l_1 + l_2) \text{ then } {}^w\bar{p}_d^{(t+1)} \leftarrow {}^w\bar{p}_d^{(t+1)} \frac{l_1 + l_2}{\left| {}^w\bar{p}_d^{(t+1)} \right|}$$

2. *Compute elbow circle*: Posing the model arms is an under-constrained problem, as four degrees of freedom must be specified from only three constraints, corresponding to the

co-ordinates of the desired end-point position  ${}^w\vec{p}_d^{(t+1)}$ . The elbow circle is defined as the set of positions that the elbow is free to adopt when the end-point of the arm reaches  ${}^w\vec{p}_d^{(t+1)}$ . It has a radius  $r$  and it is contained in a plane perpendicular to the vector  ${}^w\vec{p}_d^{(t+1)}$  at a distance  $b$  to the shoulder joint.

$$r^2 = \frac{(d+l_1+l_2)(-d+l_1+l_2)(d-l_1+l_2)(d+l_1-l_2)}{2d}$$

$$b = \sqrt{l_1^2 - r^2}$$

where  $d = |{}^w\vec{p}_d^{(t+1)}|$

3. Choose updated elbow axis  ${}^w\vec{x}_2^{(t+1)}$  and location  ${}^w\vec{l}_2^{(t+1)}$ : We chose the elbow axis at time instant  $t+1$  to be the closest to the one at the previous time instant,  ${}^w\vec{x}_2^{(t)}$ :

$${}^w\vec{x}_2^{(t+1)} = ({}^w\vec{p}_d^{(t+1)} \wedge {}^w\vec{x}_2^{(t)}) \wedge {}^w\vec{p}_d^{(t+1)}$$

$${}^w\vec{l}_1 = b \frac{{}^w\vec{p}_d^{(t+1)}}{|{}^w\vec{p}_d^{(t+1)}|} + r \frac{{}^w\vec{x}_2^{(t+1)} \wedge {}^w\vec{p}_d^{(t+1)}}{|{}^w\vec{x}_2^{(t+1)} \wedge {}^w\vec{p}_d^{(t+1)}|}$$

4. Fill updated rotation matrices  ${}^w_1R^{(t+1)} = ({}^w\vec{x}_1 \ {}^w\vec{y}_1 \ {}^w\vec{z}_1)$  and  ${}^1_2R^{(t+1)} = ({}^1\vec{x}_2 \ {}^1\vec{y}_2 \ {}^1\vec{z}_2)$  with:

$$\begin{aligned} {}^w\vec{x}_1 &= {}^w\vec{x}_2 & {}^1\vec{x}_2 &= (1,0,0) \\ {}^w\vec{z}_1 &= {}^w\vec{l}_1 / |{}^w\vec{l}_1| & {}^1\vec{z}_2 &= {}^wR_1({}^w\vec{p}_d - {}^w\vec{l}_1) \\ {}^w\vec{y}_1 &= {}^w\vec{z}_1 \wedge {}^w\vec{x}_1 & {}^1\vec{y}_2 &= {}^1\vec{z}_2 \wedge {}^1\vec{x}_2 \end{aligned}$$

The proposed inverse kinematics method can obtain an arm pose that will put the hand of the model in the required position. But this pose must be analyzed in order to determine if it respects model joint limits and does not produce a collision between different links.

The detection of limits violation and collisions is merely used to correct tracking errors and produce a more natural motion in the human model. But for HOAP-1 model, these two features become a crucial part of the motion generation, as an incorrect pose could damage the real robot if it is not previously detected and avoided.

- *Joint limits detection.* Given the updated shoulder and elbow rotation matrices, it is necessary to extract joint angles from these matrices that correspond to the real DOFs of the model.

This process is made by applying a parameterization change to rotation matrices. There is a direct correspondence between Denavith-Hartenberg (DH) (Craig, 1986) parameters and model joint angles, so the local axes referred angles are converted to DH parameters. The shoulder conversion can be done applying the following parameterization to the rotation matrix  ${}^w_1R$ :

$${}^w_1R = \begin{pmatrix} c\theta_2c\theta_3 & -c\theta_2s\theta_3 & s\theta_2 \\ s\theta_1s\theta_2c\theta_3 + c\theta_1s\theta_3 & -s\theta_1s\theta_2s\theta_3 + c\theta_1c\theta_3 & -s\theta_1c\theta_2 \\ -c\theta_1s\theta_2c\theta_3 + s\theta_1s\theta_3 & c\theta_1s\theta_2s\theta_3 + s\theta_1c\theta_3 & c\theta_1c\theta_2 \end{pmatrix} \quad (8)$$

where  $\theta_1$ ,  $\theta_2$  and  $\theta_3$  are the real DOFs of the model arm,  $c\theta_i$  is  $\cos(\theta_i)$  and  $s\theta_i$  is  $\sin(\theta_i)$ .

The elbow angle is much easier to obtain: as there is only one DOF in the elbow, the local rotation angle is equal to model  $\theta_4$  angle.

Both human and humanoid robot models distribute the DOFs in the same way: two located in the shoulder, and two in the elbow. But the inverse kinematics method provides a solution in which the shoulder contains three DOFs. It is required to move this DOF from shoulder to elbow. This operation can be easily done, given the chosen parameterization, as the third DOF of the shoulder corresponds with the rotation along the segment axis and so it can be directly translated as the first elbow DOF.

Once the model DOFs are computed, the system can directly check if any of them lies beyond its limits.

- *Collision detection.* This process is much more complex than previous one. Our method uses RAPID (Gottschalk et al., 1996) as the base of the collision detection module. This library provides functions that can quickly and efficiently check collisions between meshes composed by triangles, and attached to model links.

Once the system detects an incorrect position (i.e. joint limit or collision), it follows these steps:

1. The system looks for alternative poses (i.e. different arm configurations). Imitation requires to place hands in certain coordinates, but the elbow is free to move in the circle presented in Fig. 10a. Thus, alternative poses will preserve hand positions, but will move the elbow in this circle.
2. The motion of the arm should be as smooth as possible. Thus, alternatives should be more densely searched near the current elbow location. This is implemented by exponentially distributing the alternatives around the initial incorrect elbow position, as shown below:

$$\theta_{2i} = \pi \frac{1}{100 \frac{(n-i)^n}{n}} \quad (9)$$

where  $\theta_{2i}$  and  $\theta_{(2i+1)}$  correspond to two symmetric alternatives on the elbow circle with respect to the current pose, and  $n = N/2$ , being  $N$  the number of alternative poses that are tested when the current pose is erroneous.

Fig. 10b shows alternatives given a certain pose. As required, alternative poses are placed on the elbow circle (Fig. 10a) and are more deeply distributed near the current elbow position.

3. The system chooses the nearest valid alternative.
4. If there is no valid alternative, the arm remains in the last valid position.

The speed of the process depends on the number of alternatives it needs to check. A system using a correct number of alternatives should produce smooth movements and work in real-time even if all of them are to be tested.

The alternative evaluation module has been also used when the system is in a valid pose: in these cases, the two nearest alternatives to current pose are checked. If one of them locates

the elbow in a lower vertical position, and do not produce limits violation nor collisions, then the elbow is moved to that position. This allows the model to adopt more natural poses when possible.

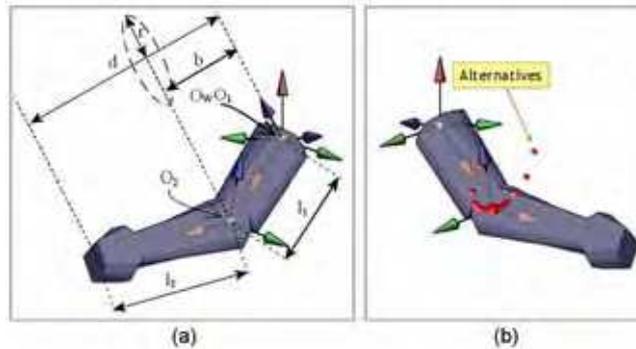


Figure 10. a) Kinematic model of the human arm showing local coordinate frames and elbow circle; and b) alternative poses (red spheres) for a given elbow position

## 10. Acknowledgement

This work has been partially granted by the Spanish Ministerio de Educación y Ciencia (MEC) and FEDER funds, Project n. TIN2005-01359, and by the European Robotics Research Network (EURON), Project VISOR.

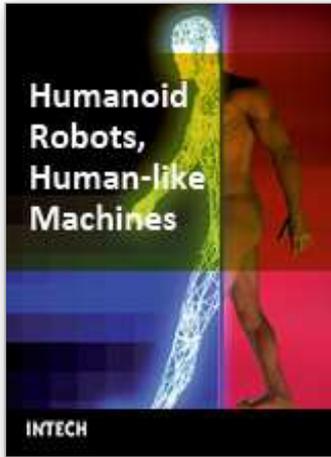
## 11. References

- Alissandrakis, A.; Nehaniv, C. & Dautenhahn, K. (2002). Imitation With ALICE : Learning to Imitate Corresponding Actions Across Dissimilar Embodiments, *IEEE Trans. on Systems, Man and Cybernetics - Part A: Systems and Humans*, Vol. 32, No. 4, July 2002, pp. 482-496.
- Andry, P.; Gaussier, P.; Moga, S.; Banquet, J. P. & Nadel, J. (2001). Learning and communication via imitation: an autonomous robot perspective, *IEEE Trans. on Systems, Man and Cybernetics - Part A: Systems and Humans*, Vol. 31, No. 5, September 2001, pp. 431-442.
- Billard, A. (2001). Learning motor skills by imitation: a biologically inspired robotic model, *Cybernetics and Systems Journal*, Vol. 32, 2001, pp. 155-193.
- Breazeal, C. & Scassellati, B. (2002). Challenges in building robots that imitate people. In: *Imitation in Animals and Artifacts*, Dautenhahn, K. & Nehaniv, C. (Ed.), pp. 363-390, MIT Press, Cambridge, MA, USA.
- Breazeal, C.; Brooks, A.; Gray, J.; Hancher, M.; McBean, J.; Stiehl, D. & Strickon, J. (2003). Interactive robot theatre, *Communications of the ACM*, Vol. 46, No. 7, July 2003, pp. 76-84.
- Breazeal, C.; Buchsbaum, D.; Gray, J.; Gatenby, D. & Blumberg, B. (2005). Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others, *Artificial Life*, Vol. 11, No. (1-2), January 2005, pp. 31-62.

- Byrne, R. & Russon, A. (1998). Learning by imitation: a hierarchical approach, *Behavioral and Brain Sciences*, Vol. 21, No. 5, October 1998, pp. 667-721.
- Craig, J. J. (1986). *Introduction to Robotics*, Addison-Wesley, Boston, MA, USA.
- Dautenhahn, K. & Nehaniv, C. (2002). *Imitation in animals and artifacts*, MIT Press, Cambridge, MA, USA.
- Decety, J.; Chaminade, T.; Grezes, J. & Meltzoff, A. (2002). A PET exploration of the neural mechanisms involved in reciprocal imitation, *Neuroimage*, Vol. 15, January 2002, pp. 265-272.
- Demiris, J.; Rougeaux, S.; Hayes, G. M.; Berthouze, L. & Kuniyoshi, Y. (1997). Deferred imitation of human head movements by an active stereo vision head, *Proceedings of the 6th IEEE Int. Workshop on Robot Human Communication*, pp. 88-93, Sendai-Japan, September-October 1997.
- Demiris, J. & Hayes, G. M. (2002). Imitation as a dual-route process featuring predictive and learning components: A biologically plausible computational model, In: *Imitation in Animals and Artifacts*, Dautenhahn, K. & Nehaniv, C. (Ed.), pp. 327-361, MIT Press, Cambridge, MA, USA.
- Demiris, Y. & Khadhour, B. (2005). Hierarchical, Attentive Multiple Models for Execution and Recognition (HAMMER), *Proceedings of the Int. Conf. on Robotics and Automation, Workshop on Robot Programming by Demonstration*, pp. 38-41, Barcelona-Spain, April 2005.
- Gottschalk, S.; Lin, M. C. & Manocha, D. (1996). OBB-Tree: A Hierarchical Structure for Rapid Interference Detection, *Computer Science*, Vol. 30, August 1996, pp. 171-180.
- Hayes, G. M. & Demiris, J. (1994). A robot controller using learning by imitation, *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*, pp. 198-204, Grenoble-France, July 1994.
- Ijspeert, A.; Nakanishi, J. & Schaal, S. (2002). Movement imitation with nonlinear dynamical systems in humanoid robots, *Proceedings of the IEEE Int. Conf. on Robotics and Automation*, pp. 1398-1403, Washington-USA, May 2002.
- Ikeuchi, K. & Suehiro, T. (1992). Towards an assembly plan from observation - part I: assembly task recognition using face-contact relations (polyhedral objects), *Proceedings of the IEEE Int. Conf. on Robotics and Automation*, pp. 2171-2177, Nice-France, May 1992.
- Inaba, M. & Inoue, H. (1989). Visual-based robot programming, *Proceedings of the 5th Int. Symp. on Robotics Research*, pp.129-136, Tokyo, August 1989.
- Inaba, M.; Mizuuchi, I.; Tajima, R.; Yoshikai, T.; Sato, D.; Nagashima, K. & Inoue, H. (2003). Building spined muscle-tendon humanoid, In: *Robotics Research: The Tenth International Symposium*, Jarvis, R. A. & Zelinsky, A. (Ed.), pp. 113-130, Springer-Verlag, Berlin.
- Itti, L. & Koch, C. (2001). Feature combination strategies for saliency-based visual attention systems, *Journal of Electronic Imaging*, Vol. 10, No. 1, January 2001, pp. 161-169.
- Kang, S. B. & Ikeuchi, K. (1993). Toward automatic robot instruction from perception - Recognizing a grasp from observation, *IEEE Trans. on Robotics Robots and Automation*, Vol. 9, No. 4, 1993, pp. 432-443.
- Kuniyoshi, Y.; Inaba, M. & Inoue, H. (1994). Learning by watching: extracting reusable task knowledge from visual observation of human performance, *IEEE Trans. on Robotics and Automation*, Vol. 10, No. 6, December 1994, pp.799-822.

- Kuniyoshi, Y.; Yorozu, Y.; Inaba, M. & Inoue, H. (2003). From visual-motor self learning to early imitation - a neural architecture for humanoid learning, *Proceedings of the 2003 IEEE Int. Conf. on Robotics and Automation*, pp. 3132-3139, September 2003.
- Lopes, M. & Santos-Victor, J. (2005). Visual learning by imitation with motor representations, *IEEE Trans. on Systems, Man and Cybernetics - Part B: Cybernetics*, Vol. 35, No. 3, June 2005, pp. 438-449.
- Lorenz, K. (1966). *On aggression*, Harcourt, Brace and World, Orlando.
- Marfil, R.; Bandera, A.; Rodríguez, J. A. & Sandoval, F. (2004). Real-time Template-based Tracking of Non-rigid Objects using Bounded Irregular Pyramids, *Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robotics and Systems*, Vol. 1, pp. 301-306, Sendai-Japan, September-October 2004.
- Maurer, D. & Mondloch, C. (2005). The infant as synaesthete, In: *Attention and Performance XXI: Processes of Change in Brain and Cognitive Development*, Munakata, Y. & Johnson, M. (Ed.).
- Meltzoff, A. & Moore, M. (1977). Imitation of facial and manual gestures by human neonates, *Science*, Vol. 198, October 1977, pp. 75-78.
- Meltzoff, A. & Moore, M. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms, *Developmental Psychology*, Vol. 25, 1989, pp. 954-962.
- Metta, G.; Manzotti, R.; Panerai, F. & Sandini, G. (2000). Development: Is it the right way toward humanoid robotics?, *Proceedings of the IAS*, pp. 249-253, Venice-Italy, July 2000.
- Mitchelson, J. R. (2003). *Multiple-Camera Studio Methods for Automated Measurement of Human Motion*, Ph.D. dissertation, CVSSP, School of Electronics and Physical Sciences, Univ. of Surrey, UK.
- Molina-Tanco, L.; Bandera, J. P.; Marfil, R. & Sandoval, F. (2005). Real-time human motion analysis for human-robot interaction, *Proceedings of the IEEE/RSJ Int. Conf. on Robotics and Intell. Systems*, pp. 1808-1813, Alberta-Canada, August 2005.
- Molina-Tanco, L.; Bandera, J.P.; Rodríguez, J.A.; Marfil, R.; Bandera, A. & Sandoval, F. (2006). A Grid-based Approach to the Body Correspondence Problem in Robot Learning by Imitation, *Proceedings of the European Robotic Symposium (EUROS 2006), Workshop on Vision Based Human-Robot Interaction*, Palermo-Italy, March 2006.
- Nadel, J. (2004). Early imitation and the emergence of a sense of agency, *Proceedings of the 4th Int. Workshop on Epigenetic Robotics*, pp. 15-16, Genoa-Italy, August 2004.
- Nakamura, Y. & Yamane, K. (2000). Dynamics computation of structure-varying kinematic chains and its application to human figures, *IEEE Trans. on Robotics and Automation*, Vol. 16, No. 2, April 2000, pp. 124-134.
- Nehaniv, C. & Dautenhahn, K. (2005). The correspondence problem in social learning: What does it mean for behaviors to "match" anyway?, In: *Perspectives on Imitation: From Cognitive Neuroscience to social science*, Hurley, S. (Ed.), MIT Press, Cambridge, MA, USA.
- Nicolescu, M. & Mataric, M. (2005). Task learning through imitation and human-robot interaction, In: *Models and mechanisms of imitation and social learning in robots, humans and animals*, Dautenhahn, K. & Nehaniv, C. (Ed.), Cambridge University Press, Cambridge, UK.

- Ogata, H. & Takahashi, T. (1994). Robotic assembly operation teaching in a virtual environment, *IEEE Trans. on Robotics and Automation*, Vol. 10, No. 3, 1994, pp. 391-399.
- Piaget, J. (1945). *La formation du symbole chez l'enfant*, Delachaux et Niestle, Paris.
- Rizzolatti, G.; Fadiga, L.; Gallese, V. & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions, *Cognitive Brain Research*, Vol. 3, No. 2, March 1996, pp. 131-141.
- Sakoe, H. & Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-26, No. 1, February 1978, pp. 43-49.
- Sausser, E. & Billard, A. (2005). View sensitive cells as a neural basis for the representation of others in a self-centered frame of reference, *Proceedings of the Third International Symposium on Imitation in Animals and Artifacts*, pp. 119-127, Hatfield-UK, April 2005.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots?, *Trends in Cognitive Sciences*, Vol. 3, No. 6, June 1999, pp. 233-242.
- Sun, C. (2002). Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques, *Int. Journal of Computer Vision*, Vol. 47, No. 1/2/3, April-June 2002, pp. 99-117.
- Tan, K. C.; Chen, Y. J.; Tan, K. K. & Lee, T. H. (2005). Task-Oriented Developmental Learning for Humanoid Robots, *IEEE Trans. on Industrial Electronics*, Vol. 52, No. 3, June 2005, pp. 906-914.
- Terrillon, J. C. & Akamatsu, S. (1999). Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images, *Proceedings of the 12th Conf. on Vision Interface*, Vol. 2, pp. 180-187, May 1999.
- Treisman, A. & Gelade, G. (1980). A feature integration theory of attention, *Cognitive Psychology*, Vol. 12, No. 1, January 1980, pp. 97-136.
- Tung, C. P. & Kak, A. C. (1995). Automatic learning of assembly tasks using a dataglove system, *Proceedings of the Int. Conf. on Intelligent Robots and Systems*, pp. 1-8, 1995.
- Viola, P. & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features, *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 511-518, Kauai-USA, December 2001.



## **Humanoid Robots, Human-like Machines**

Edited by Matthias Hackel

ISBN 978-3-902613-07-3

Hard cover, 642 pages

**Publisher** I-Tech Education and Publishing

**Published online** 01, June, 2007

**Published in print edition** June, 2007

In this book the variety of humanoid robotic research can be obtained. This book is divided in four parts: Hardware Development: Components and Systems, Biped Motion: Walking, Running and Self-orientation, Sensing the Environment: Acquisition, Data Processing and Control and Mind Organisation: Learning and Interaction. The first part of the book deals with remarkable hardware developments, whereby complete humanoid robotic systems are as well described as partial solutions. In the second part diverse results around the biped motion of humanoid robots are presented. The autonomous, efficient and adaptive two-legged walking is one of the main challenge in humanoid robotics. The two-legged walking will enable humanoid robots to enter our environment without rearrangement. Developments in the field of visual sensors, data acquisition, processing and control are to be observed in third part of the book. In the fourth part some "mind building" and communication technologies are presented.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Juan Pedro Bandera, Rebeca Marfil, Luis Molina Tanco, Juan Antonio Rodriguez, Antonio Bandera and Francisco Sandoval (2007). Robot Learning by Active Imitation, Humanoid Robots, Human-like Machines, Matthias Hackel (Ed.), ISBN: 978-3-902613-07-3, InTech, Available from:  
[http://www.intechopen.com/books/humanoid\\_robots\\_human\\_like\\_machines/robot\\_learning\\_by\\_active\\_imitation](http://www.intechopen.com/books/humanoid_robots_human_like_machines/robot_learning_by_active_imitation)

**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](http://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

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.