ROBUST MODELING OF A BODY
FOR HUMAN-ROBOT INTERACTION

Manel Frigola, Alberto Rodriguez, Josep Amat and Alicia Casals
Automatic Control Dep. UPC
Centre of Research of Biomedical Engineering
Technical University of Catalonia
manel.frigola@upc.edu

1. INTRODUCTION

The introduction of robots in environments close to humans, as for instance in medical applications and in the services sector in general, demands for new performances and robustness. Robots need to be endowed with higher adaptability to changing conditions in order to guarantee safety operation, specially when they work close to humans. In such applications, where robots operate in “natural”, no pre-structured environments, human-robot collaboration is highly desirable, specially when the programmed task can produce damages, or even when changes of robot strategy are needed. This collaboration carries with it new interaction requirements, easiness of operation, sociability, etc. Socially interactive robots requirements as communication, understanding human orders and human intention (face and gestures recognition), learning or imitation are analysed from different points of view in [1].

With the aim of achieving higher human-robot interaction capabilities and for a human being able to have a bearing in the execution of the robot task, either modifying on-line the robot trajectory or the position of working points, or even changing robot strategies, different interaction devices are available.

Apart from using solutions such as oral orders, robot teach pendant or other manual devices, research is going on towards the increase of performances of gesture based systems. The measure of forces applied over the robot in conjunction with vision systems can provide very important information for understanding human gestures [2, 3, 4].

This work is oriented to obtain a more reliable and precise model of the body of a person, operating close to the robot, in order to provide information about human intention by the interpretation of their movements and gestures. Gesture orders complement the information provided by the measure of the forces applied on the robot. This intuitive way of controlling the robot and the processing of the data in real time allows a human to modify “on line” the task going on.

2. ON LINE ROBOT CONTROL FROM GESTURES

The direct interaction with robots during the execution of a task, either to modify the task or to avoid risky situations, can be performed using contact techniques based on the tactile capacity of the robot cover, on a sensitized handle, or on force sensors located at the end-effector.

Alternative non contact techniques, based on vision, are more flexible and anticipative, but they present reliability problems. So as to increase the reliability of detection and modelling systems based on vision, we have worked to advance in the techniques of modelling the human figure of persons that appear in an area around the robot by fusing data obtained from different techniques and sensors. In the developed system, the 3D model of the body is based on the fusion of a cluster of points obtained from a set of TV cameras which detect the human body from movement, and the volume obtained through the carving of the 2D images obtained from each one of the cameras, from background occlusion.

The 3D model thus obtained is validated verifying the compatibility of each part of the model constituted by a set of cylinders, with a volumetric model of the volume covered by the infrared sensors, the presence sensors used in the scene.

The problem of extracting the human figure from the background has been tackled for many authors. Some works rely basically on carving, in [5] the segmented figure, from colour analysis, is reconstructed by carving for the precise location of a passenger to activate a smart airbag. A process for the 3D reconstruction process, in a multicamera system is described in [6], and the estimation of the human posture from the 3D reconstruction is depicted in [7]. A study of the human body kinematics, from a single camera is described in [8]. A main problem is the previous image segmentation. In [9] a regions based segmentation procedure, oriented to multimedia systems, is based on a graph structure from which different objects can be selected. A system more similar to ours, described in [10], tries to
avoid quick illumination changes and relocation of background objects, using colour and gradient information for background subtraction, operating first at pixel level and in a second step operate at region level.

3. FUSION METHOD

The method proposed aims to obtain better and more reliable results in human modelling from different views by fusing different techniques. The construction of the approximate 3D model of a body, which is tracked within the robot environment, is carried out by fusing the data corresponding to the images of the perceived movement in the scene with the voxels obtained from carving. From the former data a good spatial resolution can be obtained through the triangulation of singular points. Nevertheless, its reliability is low because of misinterpretation due mainly to potential occlusions and/or lack of movement.

The voxels that constitute the volume resulting from the process of carving are obtained from the intersection of volumes generated by 2D to 3D projection of the segmented images from each camera. The “carving” volume results in dense cloud of voxels as it is compared with those extracted by image difference, but a high precision construction in real time is computationally costly.

Fig. 1 shows the proposed procedure followed for obtaining the model of the human body that fuses both techniques. The images acquired from the multiple cameras around the working area are segmented via two different modules, one is background subtraction and the second technique is the image difference between successive frames taken with an adaptive time interval. After this initial step, a rough body adjustment is applied over the “carving” volume. This primary body adjustment is performed by estimating the main body axis and possible locations of the extracted upper-tips.

Simultaneously, a process of singular points detection in the 2D difference images is carried out. The singular point extraction, that consist in locating tips in the extracted silhouette, is assisted by the information provided by the above mentioned 3D human model fitted with the carving image. Therefore, the singular points searching area can be limited within a volume located around the extreme points of the human figure (assumed to correspond to the head, arm tips and elbows). Afterwards, the 3D singular points obtained from triangulation are filtered by a tracking process to improve stability. In a further process of fitting them with the constraints that characterize a human figure a polycylindrical model of the human body is generated, making a connectivity hypothesis along the arms that has to be compatible with the volumetric data obtained from carving.

![Diagram](image)

Fig 1. Body model adjustment procedure
4. RESULTS

Image comparison at a variable rate is used here for tracking of human motion in image sequences. Fig 2 shows the experimental scenario with two views of the scene and the results of movement detection. Triangulation of the singular points of the two stereo images provides their 3D location, necessary to interpret gestures. By using several computational improvements like binary decomposition of the 3D space, precomputing the camera model associated with every voxel of the carving matrix and preprocessing of the images an efficient implementation of carving has been achieved. Thanks to these improvements it is possible to attain near real-time 3D image generation, approximately 16 carving images per second with low-resolution (128x128x128 voxels). Therefore carving techniques can provide a real time method for obtaining an alternative 3D representation of objects. The results of carving are shown in fig. 3.

Fig. 2 Image segmentation obtained from movement. Two simultaneous views of the scene (a) and image difference results (b)

Fig. 3 Carving of the scene from images acquired from four cameras. Simultaneous views of the scene and the extracted volume by carving reconstruction

The simplified polyclindrical 3D model obtained from fusion is shown in fig. 4

This interaction has been experimented in the laboratory, working in different situations:
- Interrupting the execution of a task, both at the level of impeding its execution or modifying it using the perception of the “obstructionist force”, with the force sensors located at the wrist, [3].
- Obstructing the movement through gesture orders, using the polyclindrical model used, [4].

This process is to be used in rehabilitation. In operations such as rehabilitation exercises, massages, etc, it is usually possible to define a kind of repetitive movement. This is a kind of operation that can be programmed to be carried out by a robot. Operations such as moving an arm up and down, sweeping a surface with a rotating movement are some examples of actions that require slight changes from time to time, for instance a larger swing of the arm or displacing the rotating movement over the working surface, respectively.
Fig 4. Polycylindrical model of the human body

Fig. 5 shows an example of rehabilitation by a therapist that steers the patient’s arm up and down repetitively. This process has been tracked by the system of detection and modeling of the human figure, fig 5 b), and the process is to be reproduced in an experimental cell composed with two 6 d.o.f robots. During its operation, the physiotherapist can interact with the robot to modify the exercises previously learned, that is, defining new extreme points of the movement, different inclination, etc., according to the progressive therapy or patient response to it.

With this procedure the 3D polycylindrical models obtained initially working uniquely from movement is improved. Consequently, a better interpretation of gestures and a more effective human robot interaction has been achieved.

Fig. 5 a) Scenario for rehabilitation exercises, b) Polycylindrical model over the therapist

References