AUTOMATED 3D RECONSTRUCTION SYSTEM FOR AUTONOMOUS MOBILE MANIPULATOR AND VEHICLE-BORNE

David Puig, Vítor Sequeira, Emilio Ruiz, João G.M. Gonçalves, Martin Mellado*

European Commission, Joint Research Centre, TP 210, I-21020 Ispra, Italy
E-mail: {david.puig,vitor.sequeira,emilio.ruiz,joao.goncalves}@jrc.it

**Departamento de Ingeniería de Sistemas y Automática (DISA)*
Universidad Politécnica de Valencia (UPV)
Camino de Vera, s/n, 46022 – Valencia, Spain
E-mail: martin@isa.upv.es

Abstract: To construct realistic 3D models of large environments that match the physical world as accurately as possible there is the need to combine several kinds of data sources, such as: ground and aerial laser range scans, tracking data, photos, GPS data, satellite images. In this paper we concentrate on the ground level data acquisition. We present two integrated solutions for automated 3D acquisition and reconstruction of real scenes based on active range data having centimetre resolution and automatic capturing of colour: an indoor mobile robot platform prototype and an outdoor prototype mounted on roof of a car for modelling large urban areas. *Copyright © 2005 IFAC*

Keywords: data acquisition, computer vision, mobile robots, modelling, tracking systems.

1. INTRODUCTION

The use of 3D reconstruction techniques, based on data acquired from a variety of sensor (CCD cameras, laser scanners, etc) opens new horizons for environment surveillance, cultural heritage modelling, urban planning, safe vehicles navigation, safe manipulation, etc. (Boström, 2004; Sequeira 2004). 3D reconstruction is an iterative process that needs multiple views due to possible occlusion in the scene: it starts with an initial image acquisition; thereafter the geometric model of the scene is computed and analyzed to detect possible occlusions. If occlusions exist, a new image acquisition needs to be captured. The Image data from this new view is registered and merged to the previous acquired images. If new or unresolved occlusions are detected, further acquisition sessions are required in an iterative manner. For every iteration, given the already acquired range images and captured positions, the perception planning algorithm determines the next best view finding the optimal 3D position, viewing direction and others parameters like field of view, resolution, etc. The algorithm takes into consideration the already reconstructed environment and its associated constraints: topological (imposed by the objects being scanned) and operational (imposed by the reconstructed environment and by the acquisition system). The best capture point is used by a mobile robot, in the indoors prototype, to re-locate the head sensor in the position to take the next acquisition; and it is used in the outdoors prototype, to aid the driver during the process cycle. Several research groups reported scanning systems capable of scanning urban areas (Früh and Zakhor, 2001; Zhao, Shibasaki, 2003) and autonomous mobile robots equipped with relative positioning system to model indoor environments (Sequeira et al., 1999; Surmann et al., 2003).

In this paper an integrated system is described for 3D scanning using an automated vehicle or a car with a number of hardware components and a software scan management system capable of pre-processing the acquired data. The paper discusses the implemented techniques and presents experimental results from indoor buildings and large public areas, including the city centre of Verona (Italy).

2. PERCEPTION PLANNING

The objective is to provide a 3D model as complete and as accurate as possible from a real world scene for which no a priori information is available. Currently, available systems for 3D data acquisition acquire single depth images and/or have simple geometric fusion techniques producing inefficient representations of the underlying 3D surfaces. Except for trivial model-building tasks, it is necessary to locate sensors at several positions in the environment because all surfaces may not be visible from a single point of view and because data may not be acquired with enough resolution. Hence, mobility is an important asset for 3D scene reconstruction.

In the presented system, data acquisition for 3D scene reconstruction is made followed by the algorithmic process (see Fig. 1) for planning the next view using the already reconstructed environment and its associated constraints: topological and operational.

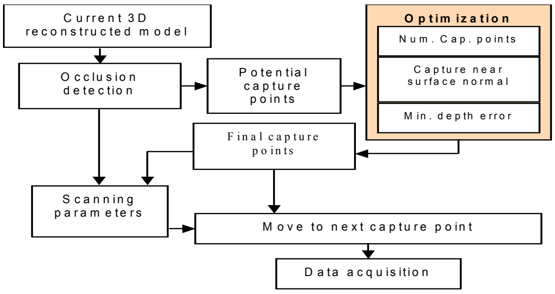

Fig. 1. Perception planning algorithm.

In order to select the next view, the perception planning process takes into account the constraints to have sufficient overlap between views and estimates the view direction of missing data that gives the maximal amount of new information regardless the final number of views. The technique is occlusion driven (occlusions being determined as step discontinuities in the range data), and selects a next-best view by choosing a location that:

- Maximise the visible volume by resolving occluded volumes.
- Maximise the area of the surface measured with sufficient sampling density.
- Minimise the resources needed to reconstruct the scene with a predefined quality (number of acquisitions, amount of data).
- Terminating the reconstruction at any point in time should yield an optimal result.

These goals are incorporated in an objective function G which evaluates a parameter set for a given scene description based on the previously taken images. The view planning problem is then addressed by solving a constrained continuous global optimisation problem for the search space $D \subset \Re^8$ (Klein, 2000). Besides the first range image, the algorithm requires the desired sampling density as an input parameter.

2.1 Iterative surface construction

We use a binary partition of space into the visible volume, which is located between any of the capture positions and the surface measured from that position, and the invisible volume, which is unseen in the images already acquired because it is located outside all fields of view or because it is occluded by measured surfaces (see Fig. 2).

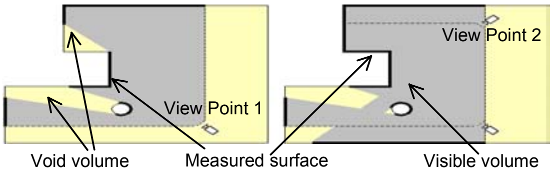

Fig. 2. Visible volume and measured surfaces.

The invisible volume consists of the object volume and additional void volume. Accordingly, the surface S of the visible volume is partitioned into a measured part and the unmeasured void surface.

2.2 Objective function

For the incorporation of a quality criterion we define a function $\beta : S \to \Re$ which yields the sampling density for a given point on the surface S. Points on void surfaces yield a function value of 0. Additionally, we associate each point p on S with a desired sampling quality $\beta_{\max}(p)$. Accordingly, we define a function $F : S \times \Re^8 \to \Re$, which expresses the expected sampling density yielded at a point on the surface S when viewed with the evaluated set of parameters. With the solid angle A_{patch} covered by one pixel at the position in the image grid in standard resolution onto which the point $P \in S$ is projected, the function F is formulated by:

$$F(p,x,\varphi_a,\varphi_e,\vartheta_h,\vartheta_v,m) := m \cdot \frac{\vec{n}(p-x)}{d^3 \cdot A_{patch}} \qquad (1)$$

Where $\vec{n}$ is the normal of the surface S in point p and $d = |p-x|$ is the distance between the capture point x and P.

3. SPATIAL DATA ACQUISITION

In large environments for which many capture locations may be required, it is not viable to involve human operators in the positioning of the Data Acquisition System (DAS). In hazardous environments as nuclear power plants and storages of strategic material, human operators access is even restricted for safety and security reasons. In order to cover a large variety of real scenarios and applications, we designed two different kinds of acquisition systems in terms of hardware and software components. These DAS are mounted on:

- A mobile platform for automated 3D reconstruction. It consists of an indoor vehicle platform with a number of sensors for its navigation, and of a manipulator arm used for positioning the data acquisition system.
- A vehicle for acquiring 3D data from urban areas.

3.1 Data acquisition system on the mobile robot

Although there is a vast amount of research effort on mobile robots in the literature, the research on mobile manipulator platforms is an area which have not been discussed in depth. Yamamoto and Yun (1993), considered a mobile manipulator consisting of a LABMATE non-holonomic mobile platform and a PUMA 250 with six DOF that uses the measured joint position of the arm for the mobile platform motion planning. Dario et al. (1999) propose a mobile robot composed of a four-wheel mobile base and a robotic arm, with a pan-tilt head supporting two cameras, and a laser scanner to reconstruct the localization of the robot. In order to outline a possible preliminary solution, we investigated the use of multi-purpose robotics vehicles with manipulator arms, and we used a prototype for Nuclear Storages operations, named the Robotics Inspection Vehicle – RIV (see Fig. 3a). This section is focused on the main design features of the RIV mobile robotics platform prototype that are suitable for the purpose of 3D data scanning.

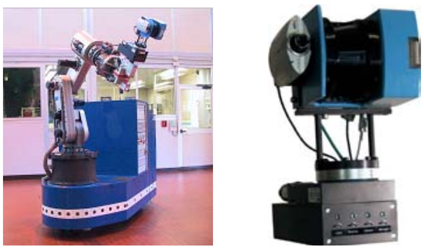

Fig. 3. a) RIV mobile manipulator with mounted DAS. b) DAS prototype I.

An important constraint on the use of robotics for indoor 3D reconstruction is that these environments must not be adapted or specifically designed to accommodate these technologies. For this reason RIV implements (Ruiz, et al., 2003) an holonomic design to carry out motion in any direction.

3.1.1 RIV design aspects

An important design feature is that RIV's sensors and electro-mechanical equipment are based on industrial components off-the-shelf smartly integrated in the GENERIS software architecture for robotics (Ruiz, et al., 1999). In addition, the vehicle design is very compact - only 2.3 x 1.15 meters - considering that it carries a heavy payload manipulator and a high-capacity battery for up to 8 hours of continuous operation. Every steering-driving wheel unit is equipped with an incremental encoder for the driving wheel and with an absolute encoder for the steering axis. Many different mechanisms were studied to achieve holonomic motion (Holmberg and Khatib, 1999). These include various configurations of single or double universal wheels, Mecanum wheels, chains of spherical or cylindrical wheels, orthogonal wheels, and ball wheels. All of these mechanisms, except for some using ball wheels, have discontinuous wheel contact points which are a great source of vibration. To overcome these issues, RIV was designed with three contact points given by three steering and driving wheels units located at the vertices of an isosceles triangle. Each contact point coincides with a steering axis, and their location ensures the vehicle stability for any configuration of the manipulator arm.

The forward kinematics method of the wheeled mobile platform is based on a particular case of a common approach found throughout literature: the instantaneous centre of rotation (ICR). The main idea used to solve the forward kinematics problem stands on commanding all motions with respect to the instantaneous centre of rotation along the trajectory (see Fig. 4). We suppose that:

- The linear speed measured by the encoder is the one of the origin of the mobile frame.
- The trajectory is always locally circular with respect to the ICR.

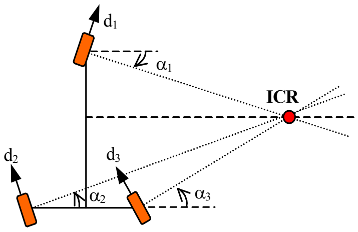

Fig. 4. Vehicle kinematics model.

3.1.2 DAS system prototype I

This system is based on a commercial 2D laser scanner and rotary stage (shows in Fig. 3b). In addition, the DAS mounts a colour CCD camera for automated colouring of the model. The system is controlled by a tablet PC that can configure the devices, acquire the 3D data and command the mobile manipulator using a TCP/IP wireless connection. The DAS has an accuracy of proximally 2.5cm, and a maximum range of 8m. the resolution of the system is programmable with a scanning space of 180° (V) and 360° (H). The DAS is mounted in the end-effector of the robot arm (see Fig. 3a).

3.2 Data acquisition system mounted on a car

3D reconstruction of large areas requires data acquisition from different captured locations. Two methods can be used for doing this: Stop Scan and Go (SSG) or Scan While Driving (SWD). The first method aided by the perception planning algorithm performs the acquisition process in an iterative way taking scans from different static positions. In the second one, the data acquisition is achieved by driving the vehicle equipped with the system along the road. The final resolution of the 3D model depends not only on the resolution of the scanners sensors but also on the driving speed. In this paper we present an outdoor DAS capable to use both methods.

3.2.1 DAS system prototype II

We designed and developed a compact portable DAS for 3D modelling of large areas with a compact mechanical design mountable on the roof of any personal car (see Fig. 5). An important feature of the system is that the DAS's sensors and equipment are based on commercial off-the-shelf components. For the conceptual design, a graphics computer simulation was achieved on the Virtual Robot Simulator VRS® (Mellado et al., 2003) to verify its suitability in a typical urban environment.

The vehicle-mounted system is composed of three laser range scanners: two of them are mounted in vertical direction to collect the 3D data used to reconstruct the surface model of the scanned objects. The other scanner is mounted horizontally and is used for registering the vertical range profiles taken by the vertical scanners. In addition the system is composed of two calibrated colour CCD cameras for instant colouring of the model and of a GPS receiver for global positioning and internal synchronization. The system is controlled by two touch-screen tablet-PCs. All the equipment is portable and battery powered with 4 batteries of 60Ah – 12V in the vehicle giving autonomy of approximately 9 hours.

Fig. 5. Vehicle-borne and GPS reference station.

In order to reduce the occlusions produced by the pedestrians, lamppost, etc, and to acquire the data in only one step, the two vertical lasers are mounted with different angle among each others (see Fig 6).

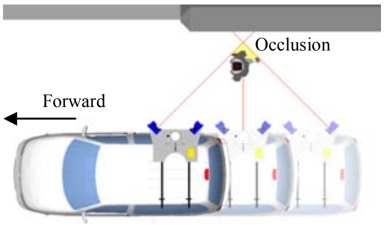

Fig. 6. The red lines represent the scanning planes.

3.2.2 Modelling

The acquisition hardware collects data from four different sources (see Fig. 7). To form a model, these input sources need to be managed in 2 consecutive steps as shown in Fig. 8.

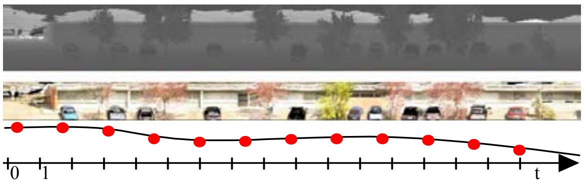

Fig. 7. Data sources: range, colour and tracking data.

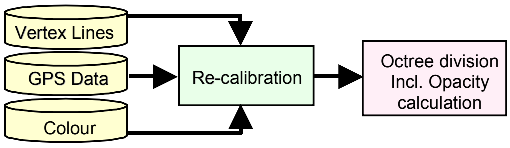

Fig.8. Modelling steps.

For the reconstruction of outdoor environments and their alignment with the physical world, the absolute position and orientation of the DAS is needed. A positioning sensor is a device that can measure the position and/or orientation relative to a reference frame. The positioning data is then passed to a pre-processing module witch converts the acquired points from a local to a global reference frame (Re-calibration step), in order to produce a 3D model of the physical world as realistic and accurate as possible. Details for our approach can be found in the next section. A second pre-processing step, called Octree-division and opacity calculations, is required to facilitate quick and efficient rendering for large models. In this stage the scan is converted to an octree and the opacity for each leaf is calculated.

3.2.2.1 Recalibration

The vertical profiles re-calibration problem consists of estimating the car position and orientation for each vertical scan. As said before, the vertical laser takes scan profiles in a 2D plane; the movement of the vehicle provides the 3rd dimension necessary to obtain the 3D model.

Working outdoors presents a number of problems that are not noticed when dealing with indoor tracking systems. The use of positioning devices in an indoor environment is simplified due to known limitations of the working environment. On the other hand, when working outdoors the environment is virtually unlimited in size and setting up infrastructure may be difficult. One of the problems that involve the use of a GPS receiver as positioning system is that it elaborates the position from the signal received from several satellites that are not geostationary. For that reason, the number and position of the available satellites can change in the time, influencing the system precision. In order to increase the precision of the positioning system, we used GPS RTK technique, thus a network of GPS reference stations distributed on the territory, even 100 km apart, from which GPS correction data is retrieved and interpolated to create a virtual reference station located nearby the rover. With this kind of configuration, it is possible to have centimetre-level positioning accuracy. For the vertical profiles re-calibration we suppose that the car moves on a flat environment. The car localization is given by (X, Y, ϕ), were X and Y are the cartesian coordinates of the vehicle and ϕ is the vehicle orientation with respect the x-axis.

The GPS receiver provides positions at 1Hz, P_0 to P_n, acquired during the acquisition. The lasers take profiles at 20Hz and 15Hz. Due to the different rate between the laser scanners and the GPS we need an interpolation of the position and the orientation. Our approach for interpolation is based on the Catmull-Rom local interpolating spline (Catmull and Rom, 1974) which enables us to estimate the car position and orientation at any position along the trajectory for every vertical scan profile (Fig. 9 shows the trajectory interpolation). The points are interpolated by the spline, which passes through each GPS point in a direction parallel to the line between the adjacent points.

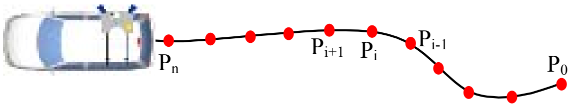

Fig. 9. Catmull-Rom trajectory interpolation.

Where (P_0, P_1,…, P_n) are the GPS control points which require Catmull-Rom spline to pass through, $t \in [0, 1]$.

3.2.2.2 *Trajectory correction*

The decrease of GPS precision can affect the final model. In order to improve the results, we use a Kalman filter (Kalman, 1960) to estimate the vehicle location when the quality of the position read from the GPS receiver is too low (see Fig. 10).

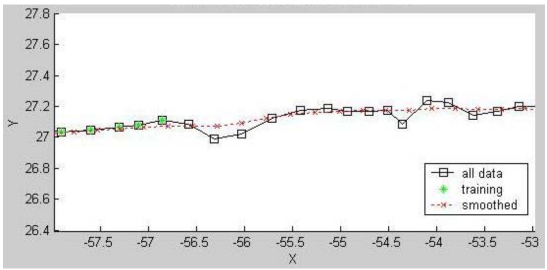

Fig. 10. Trajectory filtered.

The travelling of the vehicle around the road can be approximated by a linear system and can be described by the following equations, where $x(t)$ is the hidden state at time t, and $y(t)$ is the observation.

$$x_t = Ax_{t-1} + \omega_{t-1} \tag{2}$$

$$y_t = Cx_t + z_t \tag{3}$$

ω_{t-1} represents the process and z_t represents noise. They are assumed to be independent and with normal probability distribution. The time update projects the current state estimate ahead in time. The measurement update adjusts the projected estimate by an actual measurement at that time. The new position $(x1, x2)$ will be the previous position plus the velocity $(dx1, dx2)$ plus noise ω.

$$\begin{bmatrix} x1_t \\ x2_t \\ dx1_t \\ dx2_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x1_{t-1} \\ x2_{t-1} \\ dx1_{t-1} \\ dx2_{t-1} \end{bmatrix} + \begin{bmatrix} \varpi_{x1} \\ \varpi_{x2} \\ \varpi_{dx1} \\ \varpi_{dx2} \end{bmatrix} \tag{4}$$

We assume that only observe the position of the vehicle.

$$\begin{bmatrix} y1_t \\ y2_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x1_t \\ x2_t \\ dx1_t \\ dx2_t \end{bmatrix} + \begin{bmatrix} z_{x1} \\ z_{x2} \end{bmatrix} \tag{5}$$

3.2.3 *Experimental results.*

We carried out experiments in two areas: the city centre of Verona (Italy) and a large site containing office buildings, warehouses and general outdoor surrounding. The final models have the following characteristics:

	Site I	Site II
Vertex	16.1 million	59.6 million
Acq. time	37 min.	1 h 31 min.
Recalibration	15 sec.	60 sec.

The above results where obtained with an Intel P4/Hyperthreading CPU of 3.2 GHz, 2 GB RAM, SCSI 15000 rpm HDD of size 68.3 GB.

4. CONCLUSIONS

The paper illustrated a new automated technique for the creation of realistic 3D models of real world scenes (see some examples in Fig 11, 12, 13 and 14). The main advantage is that it takes reality as the base of modelling.

A main aspect of the overall system is its integrated approach, producing automatic modelling procedures, including detection of non-modelled parts and computation of the next best view, to progress from range and visual data acquisition to final high quality models of interiors and exteriors. Our current work concentrates in increase the tracking accuracy of the positioning system by using a hybrid positioning system combining two or more sensor technologies.

5. ACKNOWLEDGMENT

We would like to thank the city of Verona and the University of Brescia, Italy, for the kind co-operation and support when performing the data acquisitions.

The work presented was developed within VISNET (http://www.visnet-noe.org), a European Network of Excellence funded under the European Commission IST FP6 programme.

6. MODEL SNAPSHOTS

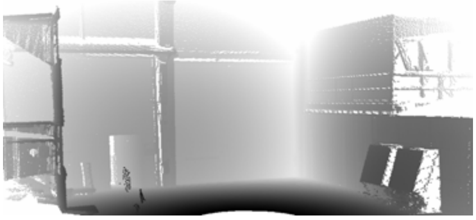

Fig. 11. Indoor environment scanned by the system prototype I.

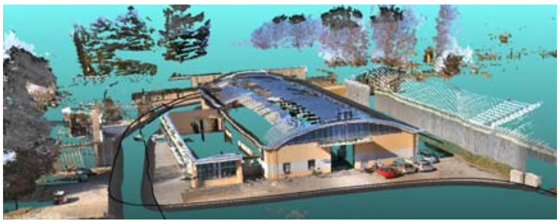

Fig. 12. SWD. Warehouse area.

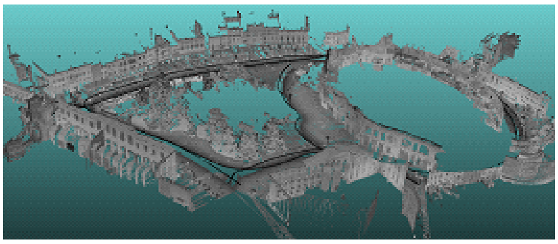

Fig. 13. SWD. City centre of Verona (Italy).

Fig. 14. SSG. Office buildings.

REFERENCES

Früh, C., and A. Zakhor (2001). Fast 3D Model Generation in Urban Environments. *Int. Conf. Multisensor Fusion and Integration for Intelligent Systems, Baden-Baden, Germany.*

Surmann, H., A. Nüchter, J. Hertzberg (2003). An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments. *Robotics and autonomous systems 45, pp. 181 – 198.*

Zhao, H., R. Shibasaki (2003). A Vehicle-Borne Urban 3-D Acquisition System Using Single-Row Laser Range Scanners. *IEEE Transactions on systems, man, and cybernetics.*

Sequeira, V, .J.G.M. Gonçalves, G. Boström, M. Fiocco, and D. Puig (2004). Automatic Scene Change Analysis of Large Areas, *Proc. INMM 45th Annual Meeting, Orlando, FL USA, 18-22.*

Klein, K., V. Sequeira (2000). View planning for the 3D modelling of real world scenes. *IEEE/RSJ International conference on intelligent robots and systems. Takamatsu, Japan.*

Boström, G., M. Fiocco, D. Puig, A. Rossini, J.G.M. Gonçalves, V. Sequeira (2004). Acquisition, Modelling and Rendering of Very Large Urban Environments. *2nd International Symposium on 3D Data Processing Visualization & Transmission, Thessaloniki, Greece.*

Ruiz, E., M. Ricci, A. Dosio (1999). A generic and opened software control system for robotics. *International Workshop on Parallel Kinematic Machines, Milan, Italy.*

Ruiz E., G. Renaldi, D. Puig, M. Franzetti, C. Correcher (2003). Robotics Inspection Vehicle for Advanced Storages. *Proc. 25th ESARDA Symposium on Safeguards and Nuclear Material Management, Stockholm, Sweden.*

Catmull, E., R. Rom (1974). A class of Interpolating Splines in Computer-Aided Geometric Design. *Academic Press.*

Mellado M., C. Correcher, J.V. Catret, D. Puig (2003). VirtualRobot: An open general-purpose simulation tool for Robotics. *European Simulation and Modelling Conference, Naples, Italy.*

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transaction of the ASME-Journal of Basic Engineering, pp. 35-45.*

Sequeira, V., K. Ng, E. Wolfart, J. Goncalves, and D. Hogg (1999). Automated 3d reconstruction of interiors with multiple scan-views. *In Proceedings of Videometrics VI (SPIE vol. 3641), pp. 106–17. SPIE.*

Yamamoto, Y. and X. Yun (1993). Coordinating Locomotion and Manipulation of a Mobile Manipulator. *Recent trends in mobile robots, World Scientific publishing Co. Pte. Ltd.*

Dario, P., E. Guglielmelli, C. Laschi, G. Teti. (1999) MOVAID: a personal robot in everyday life of disabled and elderly people. *Technology and Disability Journal, No.10, IOS Press, The Netherlands, pp.77-93.*

Holmberg R. and Khatib O. (1999). Development of a Holonomic Mobile Robot for Mobile Manipulation Tasks. *International Conference on Field and Service Robotics, PA.*