For the Workshop on Service Robotics
8th International Conference on Advanced Robotics (ICAR97)
Monterey, California, USA, July 7-9, 1997

The VIDET Project: Basic Ideas and Early Results

C. Bonivento*, L. Di Stefano*, M. Ferri†, C. Melchiorri*, G. Vassura**

* DEIS-LAR, Dip. Elettronica Informatica e Sistemistica
† DM, Dip. di Matematica
** DIEM, Dip. Ingegneria Meccanica

University of Bologna
Via Risorgimento 2, 40136 Bologna, Italy

Abstract. In this paper, the main ideas of the VIDET research project are summarized, and its current results presented and discussed. VIDET aims to increase the mobility possibilities of a visually impaired person by means of a stereo vision system and a robotic device, respectively able to reconstruct a virtual model of the environment and to provide a realistic sensation of its tactile exploration. Comments are given on the different aspects of the project (i.e. vision, definition of the virtual environment, robotic device, control aspects), along with the description of the current experimental activity.

Keywords: Haptic interfaces, virtual reality, stereo vision, wire transmission, parallel manipulators.

1 Introduction

This paper describes the VIDET (VIsual DEcoder by Touch) project, currently active at the University of Bologna. Main goal of the project is to realize a system for helping mobility of visually impaired persons by enabling them to "touch" modeled objects in a virtual representation of the surroundings. Main components of VIDET are a stereo vision system and a robotic mechanism based on force-feedback, see Fig. 1. The stereo vision system provides the depth map of a part of the environment, from which a virtual scaled model is obtained through interpolation of depth data. Then, the robotic device enables tactile interaction with the virtual model by means of force-feedback.

The idea of interacting with the environment by means of a robotic system is not new. Since the early 60's manipulators have been used to manipulate objects and interact with environments in place of human operators. Applications are found in nuclear or toxic environments, remote operations (outer space or underwater), in hazardous places (mines, industries), in medicine (devices for handicapped people).

The interaction of a mechanical device with a virtual environment is more recent, since it has required the development of more sophisticated technologies. Anyway, several devices in different applicative contexts have been developed for the interaction with virtual environments. One of the most common fields is probably the training of specialized personnel for operations such as surgery, flight, military applications, [1, 2].

With respect to medical applications, and in particular to devices for handicapped people, automated wheelchairs controlled by joystick are very common, performing many standard living activities and employing various control mechanisms, such as joysticks for hands or mouth, eye movement, head movement, and even voice commands. Many systems have been realized in the last decade also for blind people. Some significant examples are: an ultrasonic cane [3], an ultrasonic system based on the echolocation

of the bats with a stereophonic interface and two ultrasonic receivers [4], a satellite-based localization system (MoBIC) [5], the "Touch Blaster Nomad" developed at the Notthingam University, which consists in an audio-tactile system for investigating maps and graphics. Some of these devices, like the ultrasonic ones, are intended to be used for obstacle detection, while MoBIC and Touch Blaster Nomad are used as aids in the navigation but not for detection of obstacles.

Differently from the major part of other systems known in the literature, VIDET should be useful for both obstacles detection and for navigation. The only system that presents similar characteristics is VISIO, developed at the ENEA Laboratory of Frascati, Italy, [6]. VISIO, however, uses a single camera to acquire a bidimensional image of the environment and a pins matrix to display the edges of the objects.

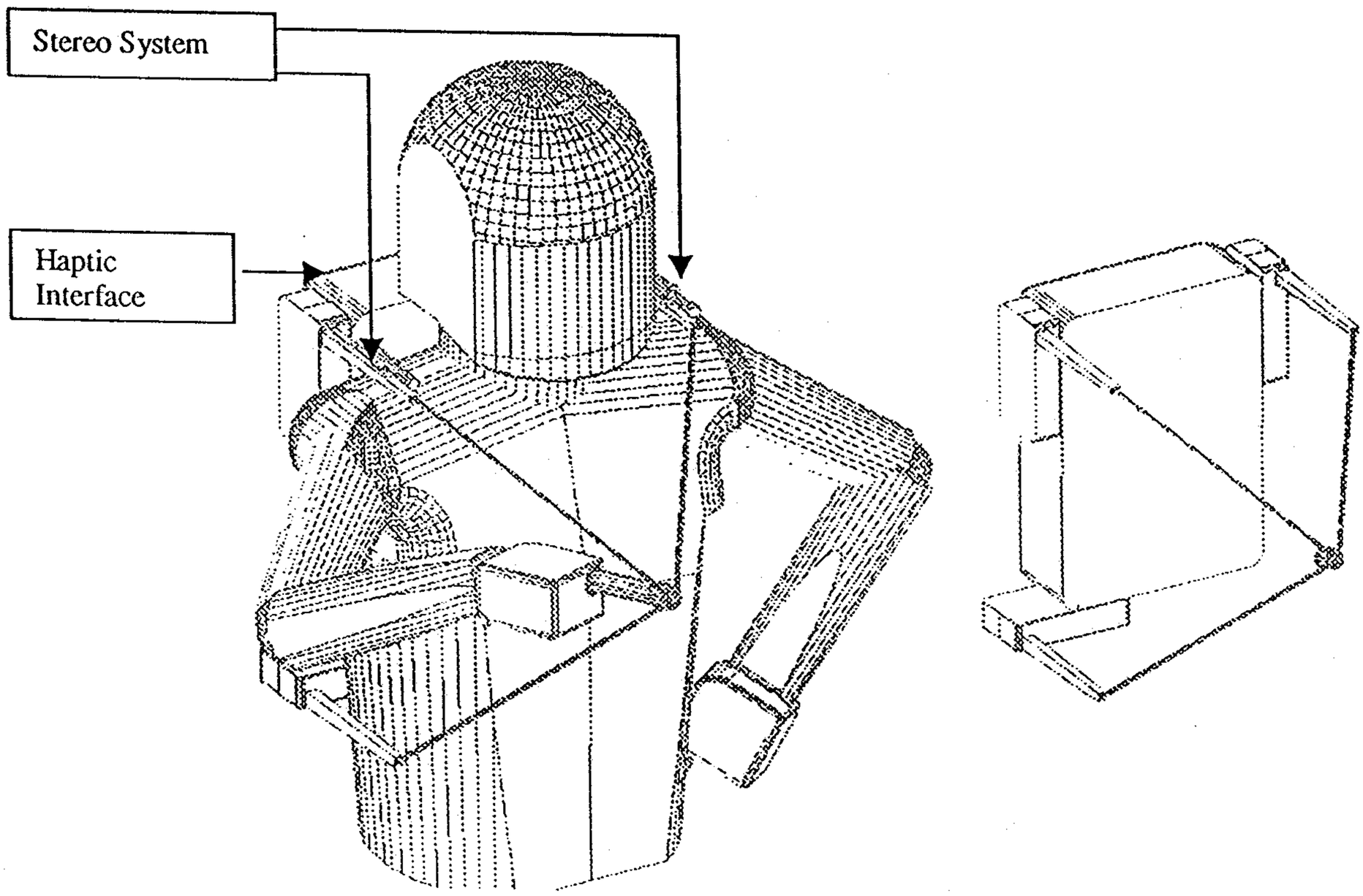

Figure 1: The VIDET system.

This paper, besides presenting the basic ideas and objectives of the overall project, [7], describes the current state of development of its main activities. In particular, the paper addresses the following main features of the device: the stereo vision system, the surface interpolation problem, the robotic device, the control aspects.

The paper concludes with the presentation of some experiments concerning the haptic interface, obtained with first laboratory prototypes. Moreover, some comments on the possibility of tactile feedback of colors are given, along with a set of preliminar tests.

2 VIDET: System Description

As already mentioned, main components of VIDET are a stereo-vision system and a robotic device (haptic interface), see Fig. 1. The general idea is to use the two cameras to recover in real time the depth map of a portion of the unknown environment. Then we build a virtual 3D model by scaling depth data and interpolating the scaled map by means of a Delaunay triangulation. In fact, as discussed in detail in [8], the information embedded into raw depth data cannot be exploited effectively while a Delaunay triangolation provides a 3D representation based on volumetric primitives suited to perform tasks such as object recognition, obstacle avoidance, navigation. Finally, the robotic system provides the interaction with the virtual surface by constraining the movements of the operator's finger along certain directions.

Both the stereo-vision system and the robotic mechanism will be controlled by a elaboration unit, hosted in a 'box' carried by the operator. Main requirements of the overall system, besides the functional specifications, are weareability, flexibility, power consumption limitations, elaboration capabilities.

With respect to the mechanical device, it acts as an haptic interface with bas-relief surfaces elaborated from the data of the stereo-vision system. Given a reference coordinate system, we intend as bas-relief surface a locus of points $\mathcal{S} = \{\mathbf{p} = [x, y, z]^T\}$ with the properties that their distance from a reference plane falls inside a limited interval d_{max} (maximum depth of the bas-relief), that a unit vector $\mathbf{n}$ perpendicular to the surface can be defined $\forall \mathbf{p} \in \mathcal{S}$, and that $\mathbf{n}$ (positive in the outside direction) forms with $\mathbf{r}$, the normal to the reference plane, an angle $\alpha \in [-\pi/2, \pi/2]$, see Fig. 2.

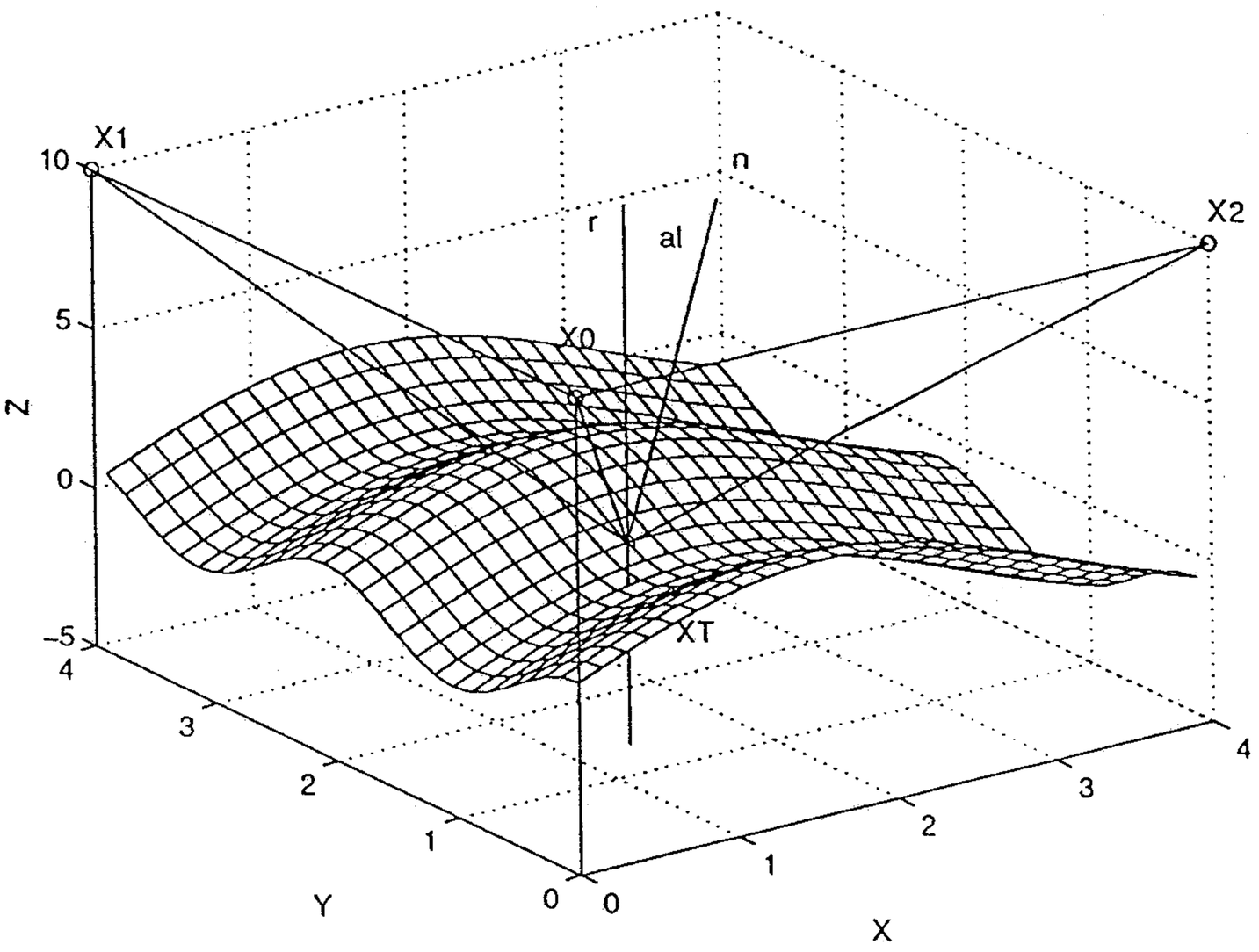

Figure 2: For a bas-relief surface, α is in $[-\pi/2, \pi/2]$.

An open problem at the basis of the VIDET project is to verify if a kinestetic perception, as the one obtained during one-finger exploration of a workspace placed in front of the operator, can allow efficient recognition of the features of a properly modelled virtual surface. At the moment, fine perception of surface details, achievable e.g. by multi-finger tactile exploration with the aid of distributed tactile perceptors, is not within the aims of this research. Then, main functional specification for the robotic device is to generate a 'good' touch perception of the virtual model with a one-finger exploration. Following these considerations, it is reasonable consider 3 as the minimum number of required degrees of freedom for the robotic interface, since only a point-contact is foreseen, and there is no necessity (neither oppor-

tunity) of interacting with the 'back' of the virtual objects/environment.

As far as 3D data acquisition is concerned, a passive stereo vision system has been preferred within the VIDET project to range sensors (such as sonar and laser) as well as to active machine vision techniques (e.g. structured light). Among VIDET's main goals and distinctive features is the ability to provide the user with a certain degree of perception of objects shape. This implies acquiring, processing and rendering to the user a (possibly) substantial amount of 3D data in real-time. Hence, though assessment of the degree of kinestetic perception actually attainable is a major issue, we have preferred stereo to active sensors due to its intrisic capability of collecting quickly a large amount of depth information. With regard to active 3D vision methods, they are far more popular and successful than stereo in industrial applications such as on-line inspection, quality control, gauging of manufactured parts. Yet, a key factor of their effectiveness in these applications is that scene's conditions (such as lighting, object-camera pose and orientation, background..) can be controlled very carefully. This is not the case of the VIDET system, which is supposed to work in a totally unstructured and highly variable scenario.

3 Stereo Vision and Surface Interpolation

The first step in the construction of the virtual model is acquisition of depth data through stereo-vision. As it is well known, the fundamental problem with stereo-vision is matching corresponding points between the two images, i.e. given a point in one image finding out which is the point in the other image associated with the same physical point. Given the inherent ambiguity of the problem, matching algorithms rely on constraints that help reducing the number of potential matches. The most powerful among these constraints is the epipolar constraint, which states that corresponding image points lie on corresponding image lines called epipolar lines [8]. The epipolar constraint can be exploited efficiently only when epipolar lines are aligned with image scanlines but, in practice, this situation is nearly impossible to obtain. Hence, a transformation called rectification is typically applied to the original image pair in order to obtain an new "aligned" pair which is actually matched [8],[9].

At the very initial stage of the project we started with the idea that VIDET's stereo subsystem should be able to provide a very dense depth map in order to build a fine model of the scene enabling recognition of objects shape by tactile exploration. Thus, we naturally addressed correlation-based matching algorithms (see for example [10]), which attempt the match at every single pixel position based on a measure of the similarity of gray-level distribution in a small window centered at the two candidates. However, in our experiments these class of algorithms has proved to be insufficiently robust due to the high number of gross matching errors which occur mainly within low-textured regions. Gross matching errors result in large errors in the depth map that significantly affect the shape of the surface reconstructed by the interpolation process. Reliability of the basic correlation algorithms may be increased by avoiding to attempt matches within low-textured regions [11] and/or by enforcing the so-called left-right consistency constraint, which consist in performing the correlation twice reversing the roles of the two-images and taking only those matches for which the two correlations agree [12]. In our experiments it turned out clearly that with both approaches obtaining robust matches implies getting far less dense depth maps.

These considerations, together with the understanding that, even if we succeeded in building a very fine model of the scene, this large amount of information could hardly be rendered to the user by the force-feedback device, have led us to redirect the depth acquisition and interpolation activity towards the recovering of a reliable, though coarse, representation. Hence, we are currently working on feature-based stereo algorithms, that can provide reliable, but sparse, depth maps by matching high-level tokens such as edges, corners, curves.

In particular, our matching program is based on the binocular algorithm for matching line segments proposed by Ayache [9]. As it is the case of all feature-based algorithms, Ayache's algorithm requires a number of preliminary steps dedicated to extraction of the features to be matched from the original gray-level images. First, we extract edge curves using Canny's algorithm [13], then we approximate each curve with a set of straight line segments by starting from the segment through the curve's endpoints and recursively splitting the current segment at the edge point where the displacement is maximum [14].

Once extracted from both images, line segments are matched using a rather complex process consisting

of three independent stages called hypotheses prediction, hypotheses propagation and validation. In the prediction stage a set of potential matches (hypotheses) is determined on the basis of the epipolar constraint and of tight constraints on the allowable length and orientation differences between corresponding segments. In the propagation stage, starting from hypotheses, matches are propagated recursively so as to try to match segments that are located in a neighborhood of those already matched; the set of constraints exploited in this stage includes the epipolar constraint, weaker constraints on the allowable lengths and orientations differences and the continuity constraint which, on the basis of the assumption that object surfaces are sufficiently regular, states that neighbor segments have similar disparities. Eventually, in the validation stage every new match is compared against those already obtained to ensure that each segment has only one homolog (uniqueness constraint). After the matching step, we obtain 3D structure by projecting back into the world coordinate system the endpoints of matched segments; as it is well known, the stereo rig must be calibrated in order to perform this final step.

So far we have tested our line segments matching program on a set of about 20 image pairs taken from indoor scenes. Results show that the program is capable of matching roughly 50% of the extracted segments (depending on the image, from slightly less that one hundred to about three hundreds matches), with a very low number of errors (3-4% of the matches). Besides, structure is reproduced properly when image segments are projected back into the 3D space, since localisation of segments belonging to the same object is correct and also the relative distances between objects are recovered correctly; currently, we do not project back segments which are nearly horizontal since, due to the nature of epipolar geometry, 3D reconstruction of segments lying along epipolar lines requires very accurate localisation of segment's endpoints.

In the following, some results are reported. Figure 3 contain the rectified image pair of an object, consisting of two perpendicular planes with a regular pattern of squares, used to calibrate the stereo rig. The left image of Fig. 4 shows all the segments extracted from the left image of the calibration object, while the right one those actually matched by the algorithm; in this case, 422 and 446 segments have been extracted respectively from the left and right images, with the number of matched segments being 211 and that of matching errors just 6 (i.e. less than 3%). Figure 5 shows two views of the set of 3D segments obtained by back-projection of matched image segments: the left one is a frontal view while in the right one the viewpoint is roughly aligned with the left plane of the calibration object. The frontal view allows us to verify that 3D segments are localised correctly, while we can perceive structure by the change of the viewpoint; in the lateral view of Fig. 5 the bunch of segments associated with the cables beneath the chair appears correctly behind the left plane of the calibration object, which in turn is localised behind the chair's leg; the right plane of the object appears in the correct position and orientation too.

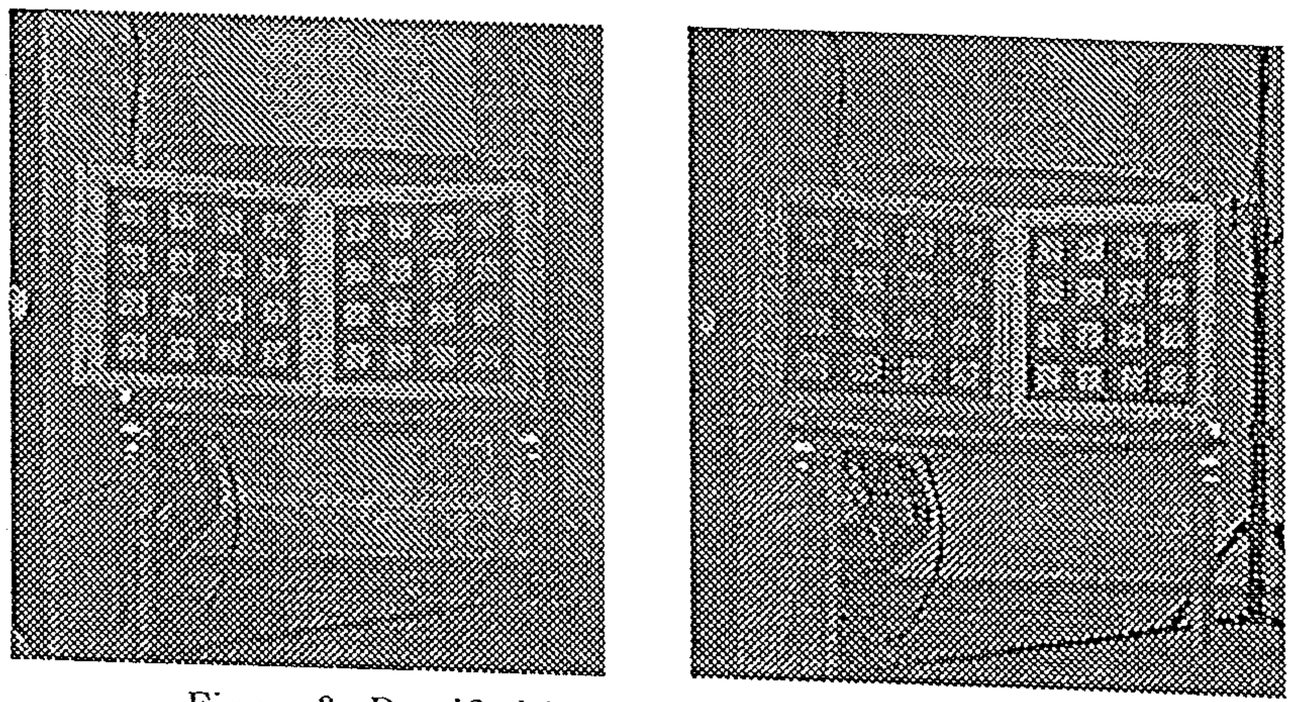

Figure 3: Rectified images of the calibration object

Given the goal for the stereo system, we consider these results satisfactory, since the matching program is very reliable and allows the recovering of a coarse model of the scene consisting of 3D line segments.

As far as computation time is concerned, the matching algorithm itself is very fast (a few hundreds of msec.) but the overall process is dramatically slowed down by segments extraction, which can last tens

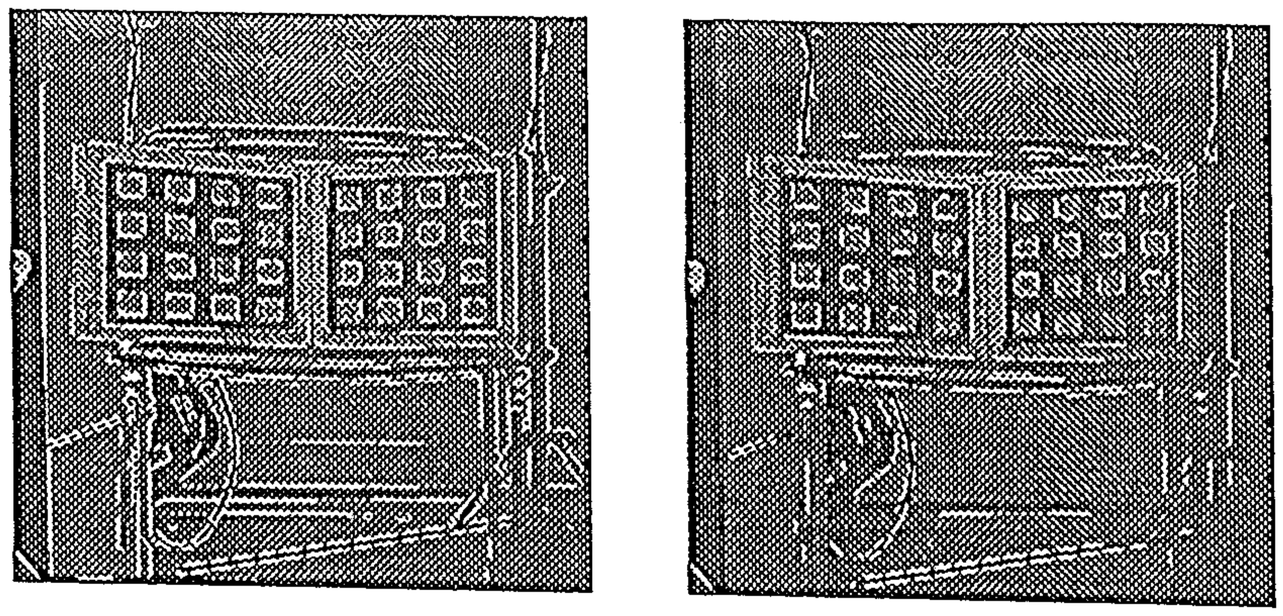

Figure 4: Left image of the calibration object: extracted segments and matched segments

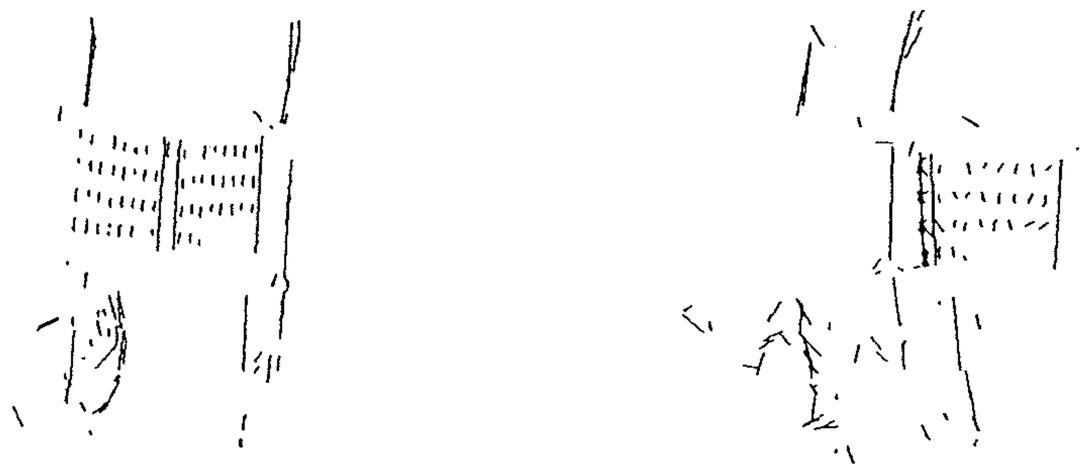

Figure 5: Frontal and lateral views of 3D segments

of seconds. Thus, future stereo activity will address a "local" implementation of the matching algorithm aimed at reducing substantially computation time. Indeed, a local approach seems to be quite natural in the context of the VIDET project. In fact, since the robotic device enables the user to explore just a small portion of the scene within a given time frame, what is really needed is a local 3D model to be updated in real-time. This could be achieved by forwarding to the stereo subsystem the position and velocity data of the tip of the robotic device. This information, together with the camera calibration parameters and the known mapping function which relates the robotic device work-space to the cameras field-of-view, would allow the determination of the two small regions of interest in the image pair where matching must be actually carried out in order to reconstruct the portion of the scene currently under exploration.

We have conducted a few experiments, finding out that if we simply reduce the area of interest in the two images to 100×100 pixels (original images were typically 512×512) the whole stereo process takes about 700-800 msec. and the matching algorithm still behaves correctly, i.e., with respect to the global approach, we do not miss matches nor obtain new errors within the regions of interest.

After recovering depth through stereo vision, the second step is to obtain a scaled surface representing the environment. Interpolation of depth data can be seen as the *trait d'union* between sensors (optics, image processing, depth computation) and actuators (control and mechanics). At the very initial stage of the project, before actual data were available as output of the stereo system and correlation based algorithms were under investigation as the matching approach, the very strong simplification that dense depth data located on a uniform Cartesian grid could be available had been assumed. This hypothesis was of great help for testing control algorithms, but, as discussed previously, recovering a reliable 3D model implies obtaining depth data which have very different densities throughout the image. Hence, in the current interpolation approach we fit triangles into the set of points and segments obtained by the

stereo process. We do it by requesting that the projection of this triangulation on a plane at constant depth be a Dealunay triangulation [15]. This has been obtained by implementing a standard Green and Sibson algorithm, modified to accept the constraint of including some prescribed segments [16].

Once the triangulation is established, the main task is determining the impact point between the reconstructed surface and the robotic device. The last two sampled positions of the exploring finger yield a vector; by extending it, we find the intersection point of the positive half–line in that direction with the triangulated surface; from this we compute the distance to impact, which is essential for the control part. The overall computation from data input to building the triangulation, to the final computation of distance to impact, keeps a standard i486 microprocessor busy for well under one second on real data. Consistently with the local approach under investigation within the stereo activity, we are currently working at a reduced computation interpolation, driven by the position of the finger; i.e., we try to perform the triangulation of the virtual surface just around the trajectory of the finger.

Table 1 shows the configuration of the stereo vision system currently used for developement of matching algorithms.

Cameras:	Sony AVC-D5CE	500x582 Pixels, CCIR output, C-Mount Lens, Int/Ext Sync, 380 TV lines, 3 Lux at F1.43, S/N Ratio:50db, Power Consumption: 2.9 Watts, Dimensions: 55x50x128 mm, Weight: 290 grams
Frame Grabber:	Data Translation DT2867-LC	4 Input channels, SW selectable gain,offset and range, 2 8-bit buffers, 8 input LUT, overlay output LUT, Zoom, Pan and Scroll circuitry, 2 asyncronous high-speed data ports (DT-connect)
Processing:	Intel Premiere PCI II	Pentium 90 Mhz, 32MB Ram, 512MB HD

Table 1: Stereo System's current configuration.

4 The Robotic Device: WireMan

Since the expected interactions of the robotic devices are only with bas-relief surfaces, and since the interaction is only by a single point, a three dof mechanism is sufficient for the goals of the VIDET project. With this respect, a robotic device based on classical open kinematic chains was initially considered as haptic interface, and the spherical (RRP) and articulated (RRR) configurations have been analyzed. Then, considerations related to the weight and the precision requested to the haptic interface led to a parallel-type configuration, which has developed in the design based on three parallel wires shown in Fig. 1.

Note that this device, considered as a 3D haptic interface, is defective since 4 wires are necessary to fully interact with virtual 3D environments, [17, 18]. Nevertheless, the particular nature of the virtual surfaces considered in VIDET (i.e. bas-reliefs), a proper control strategy, and a suitable training of the operator allow for the use of this simple and light device, as also partially demontrated in laboratory experiments. With this respect, some preliminar experiments have been carried out, besides on the 3 wires device, also on a 1-wire prototype showing some interesting results, as discussed in the following Sections.

Note that very few haptic interfaces have been designed with a parallel wire configuration. In fact, some limitations characterize this type of devices, for example the possibility of twisting the wires, the large space occupied for obtaining a reasonable workspace, the necessity of having $n+1$ actuators for rendering a full virtual reality interaction with an n-dimensional object, the intrinsic properties of rendering only a 'single-point' interaction. On the other hand, this configuration has very interesting properties: it allows very ligth structures, it has a parallel actuation, presents reduced problems of transmission nonlinearities, and has a better overall weight/payload ratio.

Some examples of robotic devices (both used as haptic interfaces and more general applications) based on wires are: Touch Feely, a four-wire device for 3D interactions, [19]; Spidar, a similar device presented in [17]; the master robot for teleoperations discussed in [20]; the crane-type manipulator of [21].

4.1 Aspects of the mechanical design

The following issues have been considered as general and important points on which develop the design of the mechanical structure:

- The system must be wearable. The wearability depends on a number of factors, including size and bulk, adaptability to different body features (including male/female differences) and requires that the presence of the equipment does not interfere with the free movements of the limbs.
- The system must exhibit a good level of flexibility, with the possibility of easly change the configuration of the wire guide-points. In fact, the position of each guide-point contributes to the definition of the workspace with respect to the arm and the body of the operator. The search for optimal positioning of the wire guide-points has to be performed on the field, as CAD simulation cannot give a realistic idea of how the operator really interacts with the system.
- Actuation and control must provide for reasonably accurate performance, even if the requested dynamic response of the system is not too high.
- The system must offer good reliability and robustness with respect to several types of disturbances. In particular, the relative position of the wire guide-points must not be influenced by changes in the body posture or by variations in the wire tension. This means that a fairly rigid system must connect the wire guide-points without influencing on the other hand the body's mobility.
- Easy maintenance and inspection are required, including fast and easy replacement in case of wire breakage.
- Intrinsic safety features must be adopted, in particular introducing wire tension limitations in order to avoid possible damages to the operators body in case of uncorrect routing of the wires (e.g. winding of the tensed wires over the finger due to unproper movements).
- The use of off-the-shelf solutions both for the actuation and for the sensory and control equipment is recommendable, as the global cost and the set-up time of the system should be low, considered the possible use of the device as personal aid for impaired persons.

According to the above mentioned issues, the actual overall design looks like in Fig. 1. The system basically consists of a sort of backpack, with a rigid frame connected to stuffed adaptable belts. The rigid frame is manufactured in light weight magnesium alloy, and has the function of fixed reference frame for both the robotic interface and the two video cameras. Moreover, it hosts some electronic boards for data acquisition and low-level control of the actuators. This frame will host also boards for signal processing. At present, the elaborations of videa data and the control procedures are not performed "on board", but on a PC-based set-up, connected to the mobile frame by means of a parallel line.

On the sides of the rigid frame three actuation modules are rigidly connected, directly or by interposition of purposely shaped spacers. In this way, any planar arrangement of the three wire guide-points can be obtained, in order to test different operating configurations. Some examples of possible alternative configurations are shown in Fig. 6.

Each actuation module contains a wire guide-point and the relative winding-unwinding system under sensorized control. In practice, as shown in Fig. 7, the wire guide point is a purposely shaped ring made in low-friction anti-wear material placed at the top of a guide spacer 2 coming out of the actuation box 3. The wire is made of twisted Kevlar yarns with a resultant title of about 3000 m/kg: it is very rigid (about 3% strain at breakage, with an ultimate strength of 250 N). Its very low diameter allows considering negligible the errors in position control due to overwinding of the thread on the collecting pulley.

Before winding on the main collecting pulley 6, the wire is routed over an idle pulley 4 connected to a small flexure beam 5 instrumented with a full strain-gauge bridge for the measurement of the wire tension. The main pulley 6 is driven by a worm gear speed reducer with a transmission ratio of 1:25. The worm is connected to the shaft of a Minimotor 2444 Brushless motor.

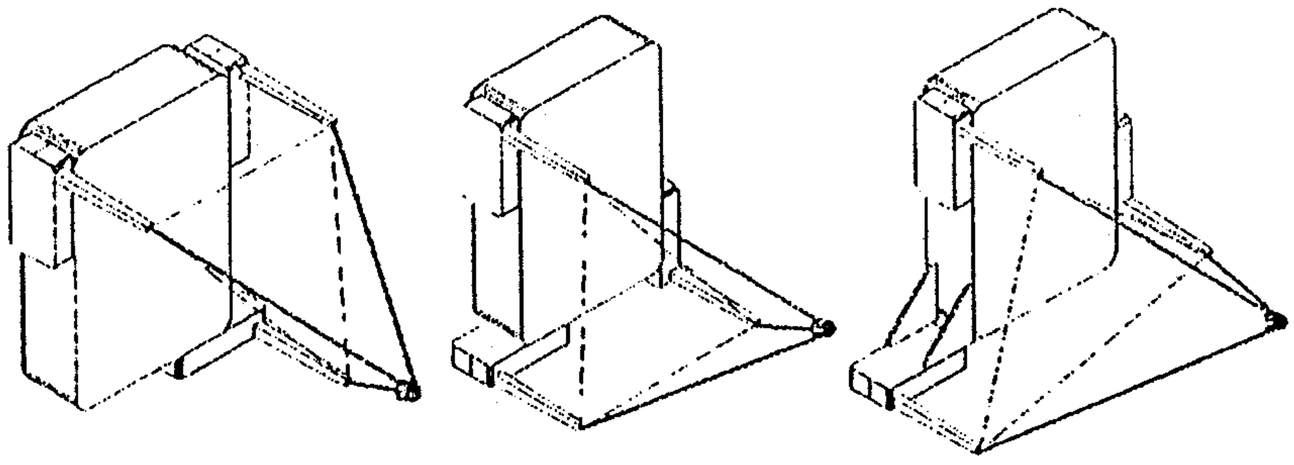

Figure 6: Different configurations of the wire modules.

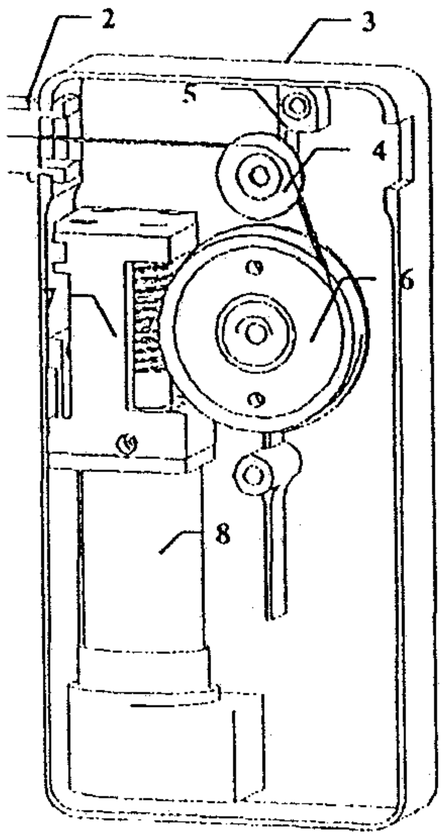

Figure 7: The wire actuation module.

An encoder with 500 pulses per turn generates information on the shaft angular displacement and, consequently, on the winded-unwinded stretch of the yarn. The control of backlash between the two meshing gears is obtained by center distance adjustment, being the motor fixed to the external box frame by means of an elastic structure, whose deformation can be easily controlled by means of an adjusting screw.

Some design data relative to the actuation box and to the backpack are summarized in Table 2.

4.2 The thimble

The design of the thimble is a very critical activity for the practical success of the proposed wire-operated system. At the moment several design solutions are being considered, but only extended on-field experimentation will allow determining the most suitable solution.

Different kind of problems arise, the major of them concerning how and where to connect the three wires to the thimble itself: the optimal solution should allow the respect of the theoretical tetrahedron

Maximum allowed speed of each wire	800 mm/s
Weight of the actuation box	0.48 Kg
Maximum tension allowed on each wire	15 N
Size of the actuation box	130x75x33 mm
Linear displacement resolution	0.01 mm
Approx weight of the whole bag	4 Kg
Size of the bag frame	370x300x80 mm

Table 2: Main mechanical design data.

configuration, with the three wires always converging into one point, whatever the position of this vertex and the consequent shape of the tetrahedron. At the same time full mobility of the operator's hand should be allowed, without any interference with the wires. It must be considered that the wires cannot be permanently fixed to the thimble, as this would seriously obstacle the wear-on operations of the backpack.

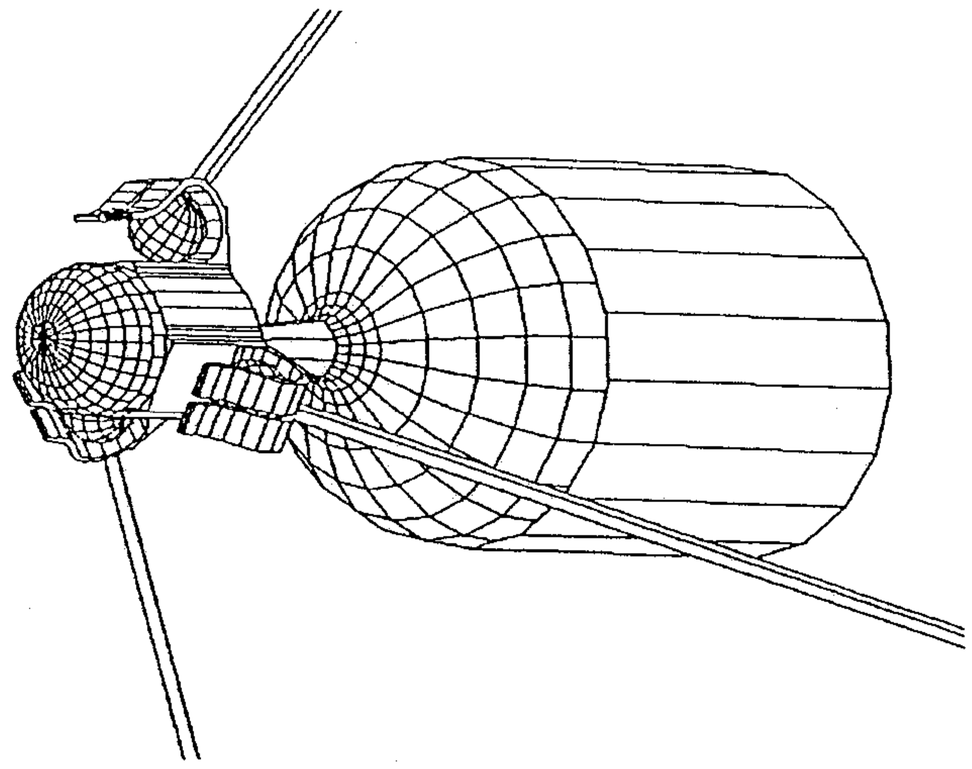

Figure 8: Details of the thimble.

Another problem is related to the need of disjoining the operator's finger orientation from that of the resultant force vector applied by the wires: it would result very uncomfortable to have to continuosly adapt the finger posture according to the variation of the applied force (moreover, this operation could be sometimes impossible). To this purpose, the optimal solution should allow a spherical joint between the thimble structure and the wire system, with its center coinciding with the vertex of the tetrahedron.

The position of the vertex (or ball-joint center) should be as close as possible to the fingertip: a position outside or inside the fingertip along the finger axis could generate bending effects that could alter the perception by the operator (this matter should be investigated in practice by testing the different configurations).

The solution that is now under development places the vertex outside the fingertip, but close to it.

Each wire is terminated with a small rigid sphere, to wich it is simply secured by a knot. These end spheres are connected to a small joining element by means of built-in snap fasteners that allow easy connection-disconnection. The distance between the spheres is very small, and a good approximation of the required one-point connection can be obtained. The thimble is connected to the terminal element through a ball joint, so that limited angular motion between the finger and the joining element is permitted. In Fig. 8 a sketch of the solution under development is presented.

The possibility to connect the three wires to a handle instead of a thimble is also being considered. Besides reducing the possibility of interference of the thimble with the wires, due to remote position of the handle with respect to the tetrahedron vertex, this configuration can suggest testing different

procedures of exploration of the virtual surface. For example, future work could study how to reproduce the sensations generated on the operator's hand-wrist by the use of a virtual stick acting against virtual obstacles. Fig. 9 shows a sketch of a possible design of this solution.

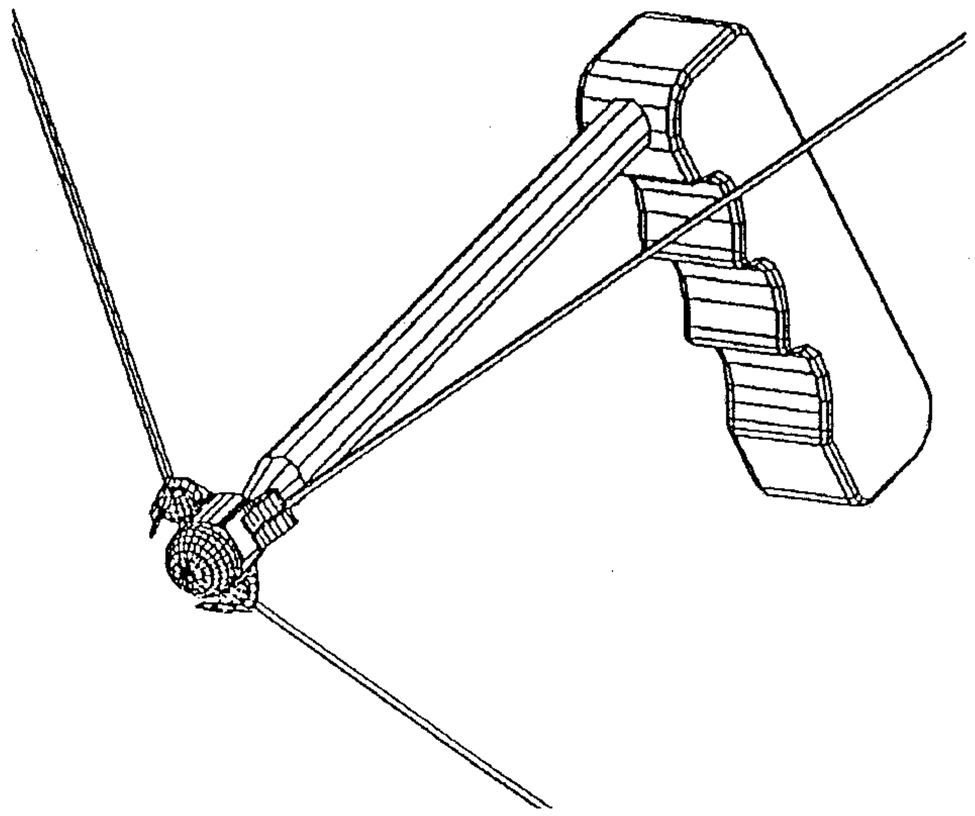

Figure 9: The handle.

4.3 Some wearability issues

A second design problem concerns the procedure by which the backpack can be weared-on by the operator and the wire-system initialized.

In order to make wear-on operations easier, the wires and the thimble are not permanently connected. In the rest position, after the control activation, each wire is completely wound with its end pulled against the guide-ring under proper tension. The operator can connect the wires to the thimble simply pulling out each wire and connecting its end to the terminal element by means of the above described fast connectors.

Then the system can start an initialization procedure, that can automatically recognize the relative positions of the base vertex (separately winding the three wires after the connection, it is possible to measure the length of the three sides of the base triangle, therefore to compute all the successive tetrahedron features).

This auto-initialization procedure is very practical in case of wire replacement due to breakage (in spite of their high tensile strength, bare Kevlar yarns are very sensitive to wear). It is not necessary to make a precise control of the length of the replaced yarn, but simply to be sure its length exceeds the minimum required.

After this initialization procedure, the operator can insert the thimble over his fingertipand insert the ball tip of the thimble inside the hole of the terminal, thus being ready for operation.

A final comment about the motor sizing. The actuators have been oversized, in order to have the possibility to implement control procedures that are based on the superposition of an arbitrary "force field" (see Section 5). For safety reasons, a torque limiting procedure has to be introduced in order to prevent injury to operator's fingers.

4.4 Kinematic models

In the following, two different configurations with 3 and 1 actuated wires are considered. In order to compute the forward and inverse kinematics for the two devices, the following assumptions are made: a) all the wires are always in tension; b) the origin of the workspace reference frame is located in one of the

wire actuation modules; c) the x-y plane of the reference frame coincides with the plane of the actuation modules. Therefore, the coordinates of the base points of the wires (in the general case of 3 wires) are $\mathbf{x}_i = [x_i, y_i, z_i]^T, i = 0, 1, 2$, with $x_0 = y_0 = z_0 = z_1 = z_2 = 0$, see Fig. 10.

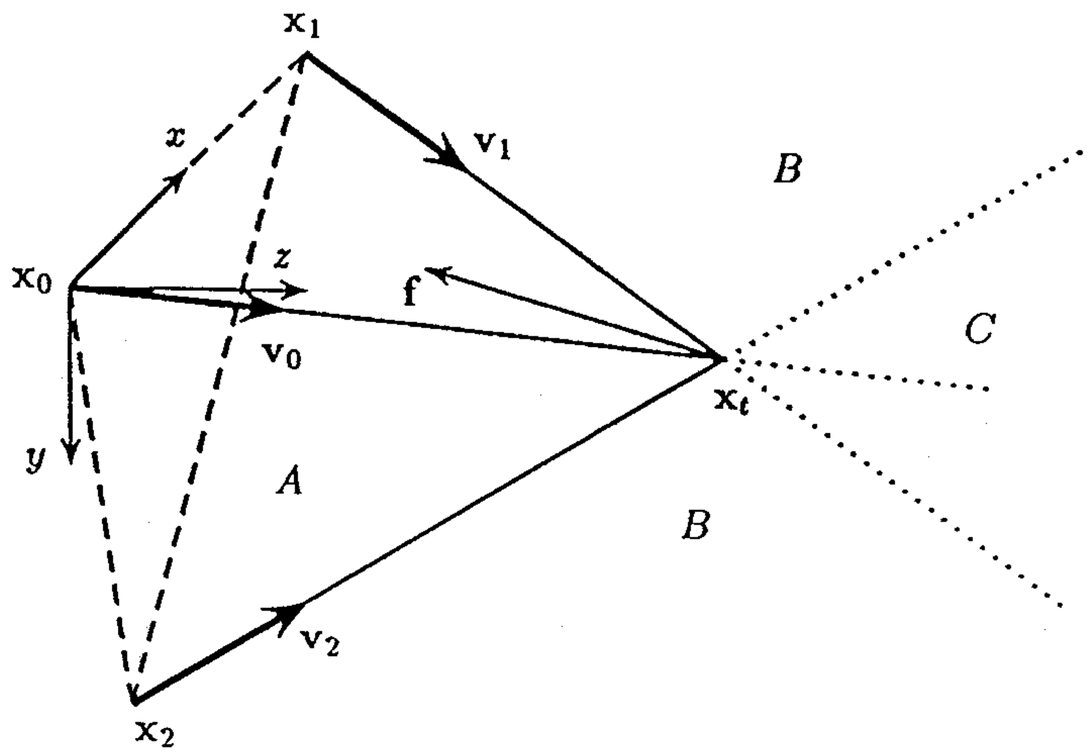

Figure 10: Sub-regions of the workspace for the WireMan.

In the following, the position of the thimble in the workspace is indicated as $\mathbf{x}_t = [x, y, z]^T$, while l_0, l_1, l_2 denote the lenghts of the cables (assumed always non-negative), which can be obtained from the measurement of the encoders.

Three-wire configuration. In this case, the solution of the inverse kinematic problem is simply given as:

$$l_i = ||\mathbf{x}_t - \mathbf{x}_i|| = \sqrt{(x - x_i)^2 + (y - y_i)^2 + z^2} \qquad i = 0, 1, 2 \tag{1}$$

The forward kinematic problem can be solved as:

$$\begin{cases} x &= \dfrac{(l_0^2 - l_1^2 + x_1^2 + y_1^2)y_2 - (l_0^2 - l_2^2 + x_2^2 + y_2^2)y_1}{2(x_1 y_2 - x_2 y_1)} \\ y &= \dfrac{(l_0^2 - l_2^2 + x_2^2 + y_2^2)x_1 - (l_0^2 - l_1^2 + x_1^2 + y_1^2)x_2}{2(x_1 y_2 - x_2 y_1)} \\ z &= \sqrt{l_0^2 - x^2 - y^2} \end{cases} \tag{2}$$

Because of the rather simple kinematic configuration, also the computation of the Jacobian matrix is straighforward. Given the position $\mathbf{x}_t = [x, y, z]^T$ of the thimble, the unit vectors associated to the wires, see Fig. 12, are:

$$\mathbf{v}_i = \frac{\mathbf{x}_t - \mathbf{x}_i}{||\mathbf{x}_t - \mathbf{x}_i||} = \frac{\mathbf{x}_t - \mathbf{x}_i}{l_i} \qquad i = 0, 1, 2 \tag{3}$$

Then, the tensions along the wires necessary to originate a force $\mathbf{f} = [f_x, f_y, f_z]^T$ to the thimble, with the sign defined as in Fig. 12, can be computed from:

$$\mathbf{f} = -\mathbf{v}_0 t_0 + -\mathbf{v}_1 t_1 + -\mathbf{v}_2 t_2 = \mathbf{V}\mathbf{t} = \mathbf{J}^{-T}\mathbf{t} \tag{4}$$

where $\mathbf{J} = [-\mathbf{v}_0, -\mathbf{v}_1, -\mathbf{v}_2]^{-T}$ can be considered as the Jacobian matrix of the device, and $\mathbf{t} = [t_0, t_1, t_2]^T$ the vector of the wire tensions. Therefore, given a force vector $\mathbf{f}$, the corresponding $\mathbf{t}$ is obtained as

$$\mathbf{t} = \mathbf{J}^T \mathbf{f} \tag{5}$$

Note that, since it is impossible to push the wires, only traction forces are allowed, i.e. forces corresponding to $t_i > 0$. We define $\mathcal{F} = \{\mathbf{f} : \mathbf{f} = \mathbf{J}^{-T}\mathbf{t};\ t_i > 0\}$ as the subset of forces $\mathbf{f}$ that can be actuated by the wires.

Similarly, in the velocity domain one obtains:

$$\dot{\mathbf{x}}_t = \mathbf{J}\dot{\mathbf{l}} \tag{6}$$

where $\dot{\mathbf{l}}$ collects the linear velocities of the winding/unwinding wires, and $\dot{\mathbf{x}}_t$ is the cartesian (linear) velocity vector of the thimble.

One-wire configuration. As mentioned above, some experiments have been executed also on a 1-wire device, schematically represented in Fig. 11. In this case, the forward and inverse kinematics are simply obtained by considering a polar reference system with the three variables l, α, β (only l is actuated, i.e. the tension can be controlled, while α, β are passive and moved by the operator while exploring the environment):

$$\begin{cases} x &= l\cos\beta\cos\alpha \\ y &= l\cos\beta\sin\alpha \\ z &= l\sin\beta \end{cases} \qquad \begin{cases} l &= \sqrt{x^2+y^2+z^2} \\ \alpha &= \text{atan2}(y,x) \\ \beta &= \text{atan2}(z,\sqrt{x^2+y^2}) \end{cases} \tag{7}$$

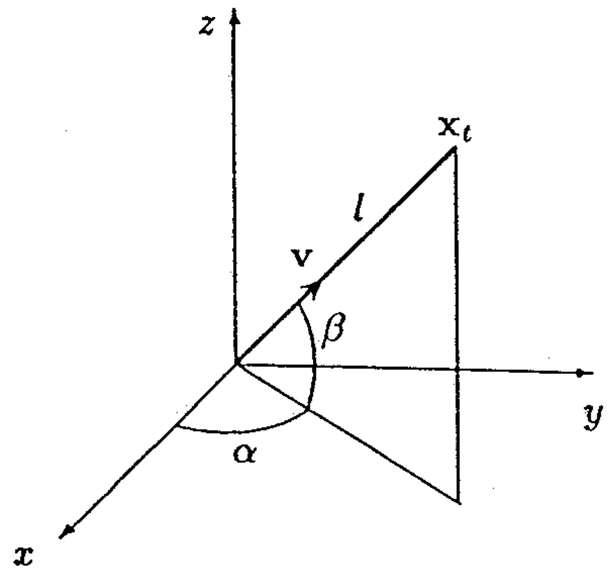

Figure 11: Polar configuration of a 1-wire device.

The Jacobian operator can be derived similarly to the three-wire case once the direction of the unit vector associated to the wire is known.

5 Control Strategies

For the sake of simplicity, reference is made in the following to a planar 2D case. Anyhow, the presented considerations and concepts apply with very minor changes also to the 3D case, as partially described in [22]. The control strategy of the device must consider two main working cases: the free space motion of the thimble and its interaction with a surface.

Free space. During the free space movements, it is necessary to maintain the wires at a proper level of tension, guaranteeing at the same time that the movements of the operator are not restricted and that the position of the thimble can be correctly computed. For this purpose, a fictitious force field can be generated at the thimble by the wires. In this manner, besides having the tendons at a positive tension value, the user perceives a force in the free workspace, which can be considered similar to the gravity force of our everyday experience.

Contact with a virtual surface. When the thimble is 'in contact' with the virtual surface, an interaction force $\mathbf{f}_e$ can be computed from the model of the environment. Different models of the environment can be considered, leading to different behaviours of the system and sensations experienced by the operator. In the following, the model is assumed as a spring with direction locally perpendicular to the virtual surface. Therefore, the desired environment reacting force $\mathbf{f}_e$, to be generated by the wires and opposing to the force $\mathbf{f}_o$ applied by the operator, is directed along $\mathbf{n}$, the direction locally perpendicular to the surface, i.e. $\mathbf{f}_e = -K\Delta x\,\mathbf{n}$, where K is the stiffness of the virtual surface and Δx the 'penetration' in the surface.

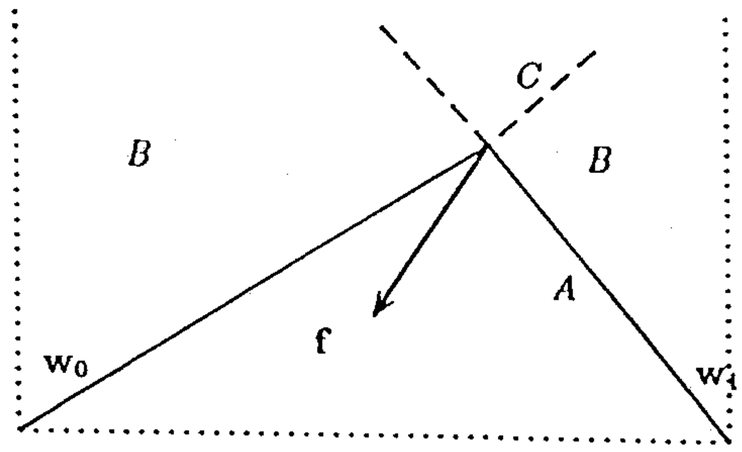

Figure 12: Sub-regions of the workspace in the planar, 2D case.

The possibility of actually generating the desired force $\mathbf{f}_e$ depends on the kinematic configurations of the device (number of actuated wires) and on the direction of $\mathbf{f}_e$ with respect on the subdivision of the workspace (regions A, B, C for the planar device shown in Fig. 12).

The contact with the surface can be reproduced correctly, i.e. $\mathbf{f}_e = \mathbf{f}$, whenever $\mathbf{n}$ belongs to the region A in Fig. 12 (delimited by the wires). In this case, two values $t_0, t_1 \geq 0$ are given by (5) and the desired surface reaction $\mathbf{f}_e$ is generated.

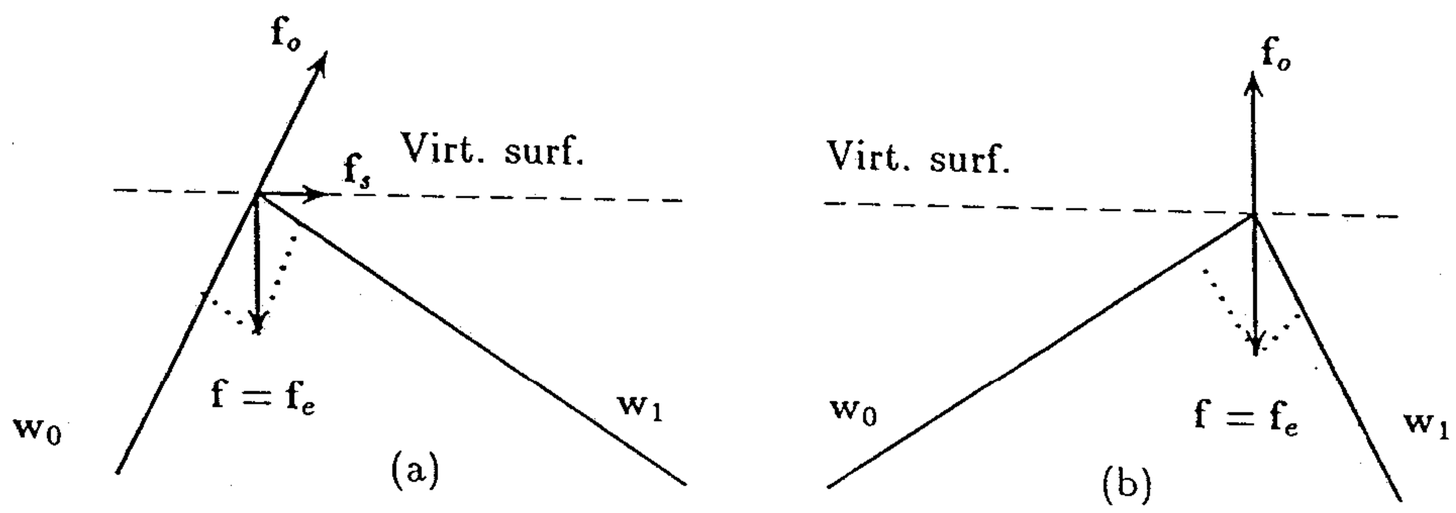

Figure 13: Motion (a) and equilibrium (b) on a surface; $\mathbf{f}_s$ is an unbalanced force component causing the motion along the surface.

Whenever the resultant force $\mathbf{f}$ generated by the wires has the direction of the driving force $\mathbf{f}_o$ applied by the finger to the thimble, an equilibrium condition is possible, and the thimble stops on the virtual surface, pushing on it. As soon as the finger tends to add a tangential force component, the thimble starts sliding over the surface, see Fig. 13 or Fig. 14. This behaviour is very useful for the exploration of the surface.

If $\mathbf{n}$ lies outside region A, for example in region B of Fig. 12, it is not possible with the wires to generate the resultant force $\mathbf{f}_e$, since this could only be obtained with a negative tension of at least one wire ($\mathbf{f}_e \notin \mathcal{F}$).

Anyhow, the device can still apply a reacting force, and therefore let the operator experience an interaction with the virtual surface. To this purpose, different algorithms can be developed for the computation of the wire tensions. These algorithms basically differ for the adopted method of projecting the desired force $\mathbf{f}_e$ in $\mathcal{F}$, the subset of forces $\mathbf{f}$ which can be generated by the wires. In the following, the algorithms applied in the experimental activity are briefly described.

Algorithm P1. If $\mathbf{t} = \mathbf{J}^T\mathbf{f}_e$ gives a vector with some negative components t_i, these are put to a minimum positive value t_{min} necessary to maintain the wire in tension, while the others are actuated. This corresponds to decompose the force $\mathbf{f}_e$ along the wires, actuating in practice only the positive tensions. The result is that the user can still perceive the surface, and explore its border, but the experienced stiffness is different (minor) than the programmed one. This drawback is shown in Fig. 14, since in steady state (case b) $|\mathbf{f}_e| = K|\Delta x| \neq |\mathbf{f}_{on}| = |\mathbf{f}_n|$.