Augmented Reality User Interface for Reconnaissance Robotic Missions

L. Zalud[1]

[1]L. Zalud, Department of Control and Instrumentation, Brno University of Technology, Kolejni 4, Brno, Czech Republic,
e-mail: zalud@feec.vutbr.cz

Abstract— **The problem of visual telepresence and augmented reality control of reconnaissance mobile robots is described. ARGOS (Advanced Robotic Graphical Operation System) for teleoperation of various mobile robots through sensory supported visual telepresence is presented. Two robots - Orpheus and Hermes, and one embedded system Orpheus EB - made on Department of Control and Instrumentation (DCI) are described as examples of systems with different features and capabilities that may be controlled through ARGOS. Data fusion of CCD color camera data, thermovision data and 3D proximity data through extended 3D evidence grids is described.**

I. Introduction

Reconnaissance of dangerous areas is one of the most challenging tasks for today's robotics. According to many indications, e.g. from Robocup Rescue League community, where the DCI team is involved (see [10]), it seems that nowadays the development of practically usable reconnaissance robots is aimed to the following tasks:

- More robots controlled by one operator, where the operator should aim to crucial tasks, like victim identification or solving problems, while the robots should do the basic tasks, like mapping, autonomously.
- Easy and intuitive human-to-robot interface should be done, since the real operators will be rescuers rather than robotic specialists.

The author tries to describe possible solution of the named problems through advanced user interface program based on visual telepresence called ARGOS and shows its application on three hardware platforms.

It has to be pointed out that the remote robotic reconnaissance of dangerous areas is a very complex and interdisciplinary task, and only well-tuned robotic systems, with good software, hardware, communication and sensory subsystem may succeed. Mobility and ability to work reliably in hard conditions are very important. It also induces that mechanical construction and hardware of the robots play very important role in this complex task.

Regarding the sensory subsystem, the most important part for successful mission is visual, since it is the most important and intuitive for the operator. Here the main problem is to allow robot control under any visibility conditions that may happen during rescue mission – e.g. smog, fog, complete darkness, etc. Author tries to indicate possible solution through data-fusion from three matrix sensors – CCD color camera, thermovision imager and 3D optical proximity scanner.

Since the robots described later are primarily teleoperated, the whole robotic system typically consists of two main parts:

- One or more Operator stations.
- One or more robots.

The configuration of an operator station may vary significantly according to mission necessities. The basic part of the operator station is a computer – it may be not only desktop or notebook computer, but also PDA with ARGOS-PDA client program.

They are two main "classes" of robots that may be controlled with ARGOS:

- Bigger and more complex robots with PC onboard.
- Smaller robots without PC.

The idea is each robot has its profile in ARGOS, so the program sends and gets appropriate data to/from the robots, but the style of control and user interface is similar.

Two different robots made at DCI – Orpheus and Hermes, as well as embedded system Orpheus-EB are roughly described in the article showing an example of remotely controlled robotic systems that may be controlled by ARGOS.

II. ARGOS-Client

If laptop or personal computer is used on base station, the robots may be controlled through ARGOS-client program.

The main, although not the only, control technique of the robots in ARGOS is so called visual telepresence. The operator has a head mounted display with inertial head movement sensor. The movements of the operator are measured, transformed and transmitted to robot. The camera copies movements of the operator's head and since the operator can see the picture from it, he/she feels to be in the robot's environment (see more in [9] and [10]).

The movements of the whole robot and all of the other functions are controlled by advanced two-hand joystick system similarly to airplanes. The used concept is known as Hands on Throttle and Stick (H.O.T.A.S.) and the basic idea is that the operator does not need to use any other control devices during the whole mission.

978-1-4244-1635-6/07/$25.00 ©2007 IEEE.

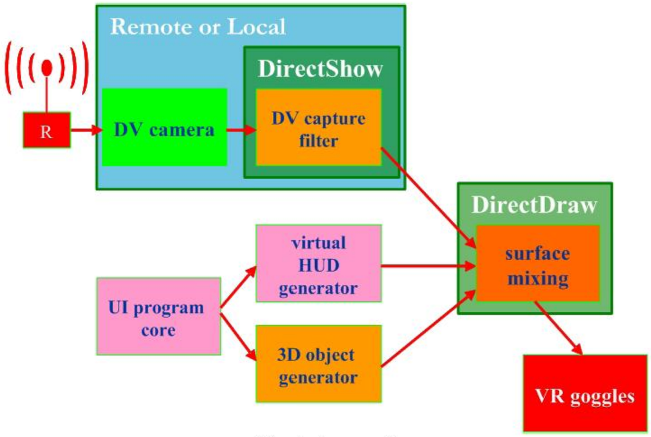

Fig. 1 Argos scheme.

The scheme of ARGOS system is depicted on Fig. 1. The program is developed in Microsoft .NET 2005 C++ language and uses DirectX 9.0c library.

At the moment, following DirectX components are used:

- **Direct3D -** to draw 3D objects.
- **DirectShow -** to perform the video-grabbing. The advantage of DirectShow is that virtually any Microsoft Windows compatible video-digitizing device may be used with it.
- **DirectInput -** to process the data from human machine interface devices like various joysticks and two-hands joystick systems, head movement sensors, etc.
- **DirectSound -** is used to play the voice messages that inform the operator about vital events that happen during operation.

A very important issue is that video is rendered to DirectX surfaces. It means virtually any (even 3D) object may be covered by this surface what gives us a big potential for video data fusion tests. The video rendering is also under full programmer's control in this case.

Another important feature is that any combination of 2D and 3D objects may be displayed by ARGOS. It means 3D data may be displayed not only in small virtual head-up displays (VHUD's – see below) but also over the main screen. The advantage is also that DirectX has good support for aplha-blending, so transparency may be easily defined for each object (in 3D) and even each pixel (in 2D). It is also possible to combine more video-images in the form of layers, again with transparency defined for each pixel. This feature is used for CCD camera and thermovision camera image data fusion (see below).

Although ARGOS is very flexible in its displaying and control capabilities of user interface, the most often used method that comes from author's experience (see [10]) is sensory supported visual telepresence (SSVT). SSVT uses previously described visual telepresence as a comfortable and intuitive method of camera movement control, while it adds semi-transparent displays (VHUD's) with other important data to the operator's view. The transparency of them may be easily and quickly controlled by the operator through e.g. analogue-like potentiometer on joystick according to actual mission necessities.

The typical look of user interface used most time by the author's team is depicted on Fig. 2 and will be shortly described in the following text.

Central VHUD - the central part of the operator's view is called Central VHUD and contains the most important data for the operator's instant orientation.

Note this is the only VHUD that is not in dark box – it is both because if its importance and placement to the central (and most important) part of camera view.

The central green reference cross forms base point to which all the other data are related.

Head rotation cross is painted by blue ink and represents operator's head relative rotation according to the central point. Although it may seem to be useless (since the operator normally knows where he is looking to…) it appears to be a very useful tool during mission because of head movement sensor temperature drift, etc.

2D scan displays - laser scanner data from actual 2D measurement may be easily displayed to the user interface display in the form of small circular nets representing the polar distances from the sensor (see the central upper and lower circular parts of the user interface display on Fig. 6. Although this display is very informative, it only shows one 2D scan from the scanner at a time.

3D scan display - to increase the information value, 3D representation of the proximity data is suitable. The right-lower display is used for 3D data representation. In the actual version of ARGOS it may be used to display so called instant data or 3D evidence grids data.

If the instant data are displayed, the data are recalculated from the original coordinates by and rendered to a surface by Direct3D, while each distance measurement is represented by a box or by a point-sprite.

Since the whole scene is on a Direct3D surface, it may be easily integrated to the operator's view (see right-lower part of ARGOS user interface on Fig. 2). The same data may be also displayed over the main surface, it means they may appear to be on the camera image.

2D map display - the left-lower display of ARGOS user interface is intended as 2D map display.

In the current version it displays 2D evidence grids.

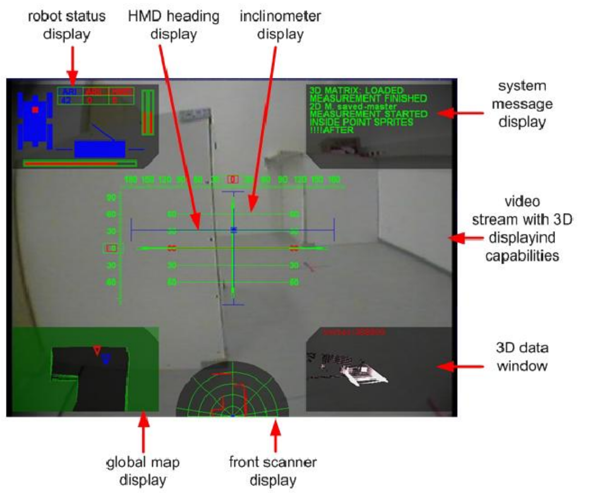

Fig. 2 ARGOS – User interface.

System message display - the right-upper VHUD serves as a message display. It works as a log window and displays events that happen only once (e.g. warning messages, etc.). It appears as an infinite rolling paper with messages, and it is not suitable to display varying variables.

Robot status display - the status of the robot is displayed on upper-left display. It shows the most important status variables related to robot's mechanics and electronics. It informs operator if the robot is moving and what is the direction and speed of the motors. It also shows how the 2DOF arm is extended, which camera is active, etc.

III. ARGOS-PDA

The very novel part of the Orpheus and Hermes robotic systems is PDA version of the operator's station. Although only the basic functions of the robots may be nowadays controlled by it, and the device itself may be only seen as a demonstration of the capabilities, it represents an important tool to make the robots much more versatile.

The hardware used is Pocket Look 720 PDA by Siemens company and the main features of it are:

- Processor ARM 620MHz.
- Display resolution up to 640x480.
- Wi-fi capabilities.

Since the PDA is equipped with different operation system named Microsoft Windows Mobile 2003, new version of the ARGOS system called ARGOS-PDA had to be developed.

IV. Data Fusion

The goal of the research described in this chapter is to improve the ARGOS to a system that:

- Makes robot control possible virtually under any visibility conditions, like fog, smoke, darkness, ...
- Displays the most appropriate data or fused data in a convenient and intuitive way.
- Visually emphasizes alive victims.
- Permits to use same data for digital map building and self-localization.

This is done through data-fusion from three sensors with somewhat complementary features:

- CCD color camera.
- Thermovision camera.
- 3D proximity optical scanner.

A. Visible Spectrum and Thermovision Data Fusion

The basic problem of data fusion of CCD color camera and thermovision camera with micro-bolometer is that the goal is to mix the data from visible spectrum sensed by color CCD camera and LWIR (long wave infrared) spectrum measured by the thermovision imager to the visible spectrum displayed by display (either TFT or CRT monitor or head mounted display with LCDs). Since the spectrum to be displayed is evidently wider than the one we have at our disposal, it seems this task may not be done without some kind of compression of the color data. But it has to be firstly pointed out that the data sensed by CCD camera and displayed by the monitors are not perfect anyway and do not correspond to reality.

They are many ways how a color may be represented. The most commonly used representation for additive color mixing is RGB model. It nicely corresponds to the way we biologically percept colors in principal, but does not correspond well to the way we "feel" the colors. There is another possibility how to describe colors by three parameters. If we look at the standard RGB color cube along the black/white diagonal, we will see the top of the so called HSV hexcone.

The HSV color model describes all possible colors by following parameters:

- H - HUE, $H \in \langle 0^o;360^o)$.
- S - SATURATION, $S \in \langle 0,1 \rangle$.
- V - BRIGHTNESS (luminancy), $V \in \langle 0,1 \rangle$.

If we express a CCD camera images in HSV mode and make histograms of it, we clearly see that the distribution of S and V is not balanced and they are almost none (or even none) pixels with S=1 and V=1. This comes out from the restrictions of CCD (and CMOS) sensors and their data processing.

From this we can conclude that if we add pixels with colors that have arbitrary H parameter and S=1 and V=1, they will be clearly perceptible by the operator. In other words, we can also say we are adding rainbow colors to the ordinary image and since they are not common (or even present) in image, they may be well used to emphasize parts of image.

Regarding the previous text, the whole procedure of data mixing that may be seen on Fig. 3 with the result on Fig. 4 may be described as follows:

- The image from Flir A20 camera is digitized. More exactly, since we are using FireWire output, it is already digital, but author is using a mode where brightness of a pixel is used to code its temperature (or rather LWIR frequency with maximum intensity) - since this interpretation is the only linear one that the camera provides. So digitalization here means temperature assignment and scaling.

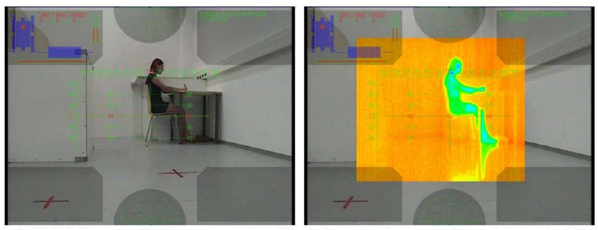

Fig. 3 CCD image (left), corresponding thermo-image placed over it (right).

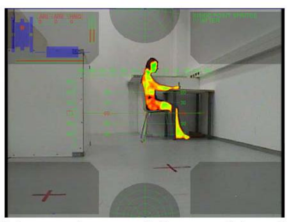

Fig. 4 Visible spectrum and thermovision data fusion.

- The temperatures that are not near human body temperatures are filtered off. More exactly the alpha of the corresponding pixels is set to 0, so the pixels keep the temperature information, what may be useful for further data processing or in certain display modes.
- The pixels that are to be displayed are recalculated to "rainbow" colors, so H corresponds to frequency, S=1 and V=1.
- This thermovision image is placed over the CCD camera image.

What has not been mentioned, but must be done, is CCD camera and thermovision camera images spatial correspondence scaling. This comes out from the fact the field-of-view of the cameras is different, they have different position, they are not aligned to be perfectly parallel, and each of the sensors has different resolution.

The only problem that may not be solved completely when combining thermovision and color camera image without 3D spatial data is the one that comes from their spatial arrangement. Since the cameras cannot occupy the same place, the optical paths they measure are different and it is impossible to calibrate the system to have well aligned image for all distances from the cameras.

To conclude this subsection, we can say there is almost philosophical question how to display frequencies of electromagnetic spectra that humans cannot see. The problem is much more complex in our case, where the image made by these "imperceptible colors" is mixed with the one we know (visible spectra). The technique described hereinbefore takes advantage of the imperfection of CCD imagers and commonly used displays together with the ability of our brain to accommodate to something that does not represent the reality perfectly, but to clearly distinguish something uncommon. In other words, the somewhat limited image from camera is taken as sufficient and appears to describe our world faithfully and thermovision image in the form of "pure rainbow" colors works as an exclamation for the operator, that there is something that should attract the attention.

It also has to be emphasized that if using ARGOS system, the operator has the possibility to change the display parameters in real-time during operation.

B. Visible Spectrum and Proximity 3D Data Fusion

The second sub-problem of data fusion is to mix data from CCD camera and 3D proximity scanner.

The first step, correspondence search - or calibration, may be described with help of Fig. 5 and 6. The 3D data have to be adjusted to fit perfectly to the CCD camera image. The process is quite complicated, since the arrangement has to be done in all six degrees of freedom. The 3D data projection to the CCD image is done by perspective projection. The camera position related to the one of the lidar is used to calculate and display the data.

Once the mutual spatial correspondences of the sensors are known, homogeneous transformation is used to position the 3D data correctly to the image, and inverse perspective projection is used to get the color from image corresponding to the actual 3D measurement point. The result is colored 3D spatial map of the robot's environment (see Fig. 6).

Fig. 5 CCD image (left) and 3D proximity data integrated to it (right).

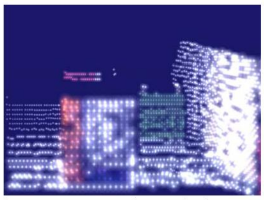
Fig. 6 Visible spectrum and proximity 3D data fusion.

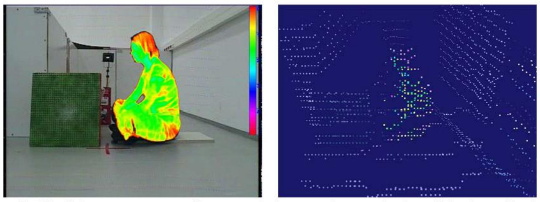
Fig. 7 Visible spectrum, thermovision and proximity 3D data fusion.

The method the author wants to "focus on" for the future development on ARGOS system and Orpheus and Hermes robots is data-fusion of all three named sensors – CCD camera, thermovision and 3D proximity scanner. This method has not been finished to be practically usable yet but the preliminary results are very promising (see Fig. 7).

V. Teleoperated Robot Examples

A. Orpheus-X2

The Orpheus robot (see Fig. 1) has been developed in our department from the beginning of year 2003. The project is a natural continuation of "mainly research" U.T.A.R. project [1], [2], [3]. The robot is intended as a teleoperated system for various reconnaissance missions – it may serve as a rescue robot, pyrotechnical robot or a robot for firemen. So its primary mission is to search objects (mostly live humans) in harsh terrain such as building ruins.

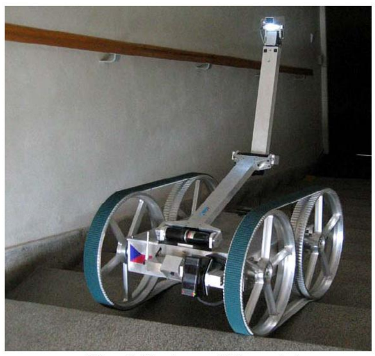
Fig. 8 Orpheus-X2 prototype.

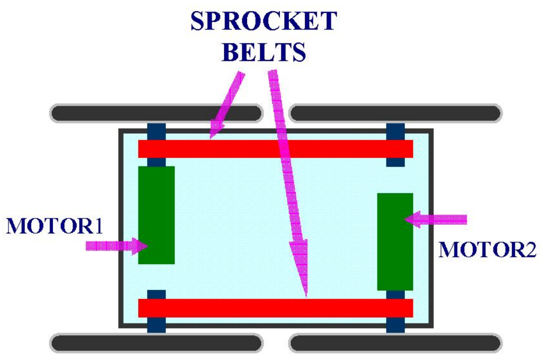

Fig. 9 Simplified scheme of the locomotor.

Locomotion Subsystem

Our department has developed a new Skid-steered Mobile Platform (SSMP) for the Orpheus mobile robot. The SSMP is intended to be both indoor and outdoor device, so its design was set up for this purpose. The scheme of the platform is shown on Fig. 9.

Electronics

The electronics of the Orpheus robot consists of two main parts:

- Atmel AVR based microcontroller subsystem.
- NEXCOM processor board for on-board calculations and Ethernet.

Communication

Communication over the Orpheus robotic system is divided into several parts (see Fig. 10).

The processor board is equipped with four serial lines. One of them is used for communication with the microcontroller subsystem, others communicate with the 3D proximity scanners because of their need for quite high communication bandwidth.

The communication between the Nexcom processor board and operator's station computer is done through Ethernet. The advantage is that the connection may be accomplished through Ethernet cable or wirelessly using Wi-fi, etc.

Sensory Subsystem

The robot contains four cameras. Two main cameras are on a sensory head. They have two degrees of freedom – may move left to right and up to down. The movements limits are similar to the ones of a human head. The cameras are sensitive, high resolution color cameras with Sony chips. The other two, so called side or secondary cameras, are black&white highly sensitive cameras with one degree of freedom.

An infrared thermosensor is used for object temperature measurement. The thermosensor is placed beside the main camera and rotates with it. It causes the temperature of the object in the centre of the camera picture is measured.

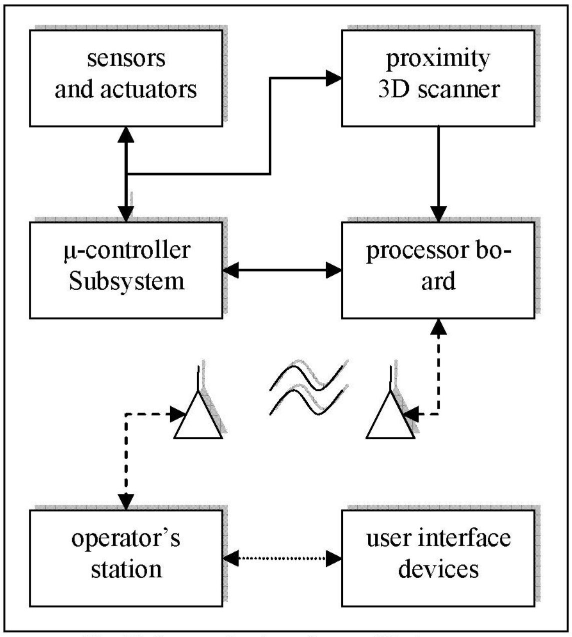

Fig. 10 Communication scheme of Orpheus.

The robot is equipped with other sensors, like accelerometers working as inclinometers, magnetometer, microphones, etc.

B. Orpheus-EB

The Orpheus EB represents a new approach in survey systems on DCI. It is not a complete robot, but rather embedded system that may grab video from up to four sources and control 2-DOF high-speed camera positioning device. It has 100Mbit/s Ethernet as well as Wi-Fi connection. It is battery-powered and may operate independently for about 80 minutes.

Physically Orpheus EB is a weather- and dust-proof box, which has all necessary electronics inside, plus camera with positioning device.

Orpheus EB is very good platform for testing on new robotic platforms as well as it may be practically used, e.g. on airships (see Fig. 12), model-sized RC airplanes, etc.

Fig. 11 Orpheus EB.

Fig. 12 Orpheus EB mounted on 9-meters airship.

C. Hermes

Hermes is a mobile robot built on omnidirectional platform (see Fig. 13). The robot is being developed since summer 2005 and it's purpose is completely different from the one of Orpheus. Hermes was built as an experimental robot and is meant to work mainly indoors rather than in ruins.

The platform was not built at DCI, but it was equipped here with all the necessary electronics and sensors. It uses four DC motors with integrated gearboxes directly linked with omni-wheels. The motors are aligned to the shape of corners of a square, so there must be used wheels with rollers oriented 45 degrees to the circumference of the wheel.

The advantage of omnidirectional drive is that it may move in any direction without any repositioning, which is very valuable when testing, e.g. mapping algorithms or virtually any position-based data processing. The advantage of the drive used on Hermes is that it is rugged, big and strong enough, so up to 70kg of payload may be placed on it. It also has quite low centre of gravity, since the lead batteries are in the bottom part of Hermes, so the risk of turn over is decreased.

Hermes is equipped with:

- High resolution Sony CCD camera.
- FLIR A20 thermovision camera.
- SICK laser proximity scanner with additional rotational degree of freedom

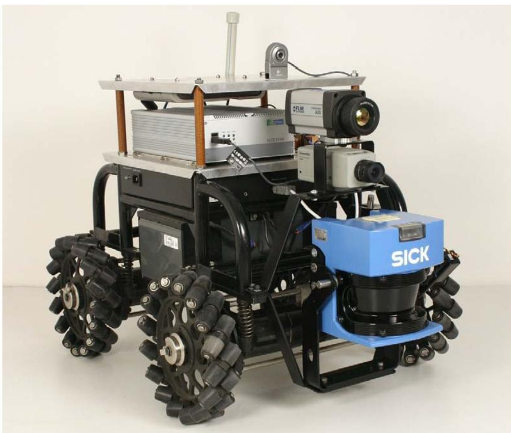

Fig. 13 Hermes robot.

All these sensors are aligned in such a way that they all point to the same direction, so their fields of view penetrate to each other, which is the prime condition for data fusion described in Chap. (VII).

VI. Conclusion

Possible solutions of the most demanding problems in reconnaissance robotics, described in introduction, have been shown. It has to be pointed out that although the presented methods do work in real-time and may be practically used, they do not represent final stage of the research, but rather state-of-the-art of the author's research.

ARGOS system presented in this paper may be used to remotely control different mobile robots. It has to be said that ARGOS is not meant as completely universal mobile robot user interface tool, but is oriented towards mainly teleoperated mobile machines for reconnaissance of dangerous environments. It supposes some basic set of features on the controlled robot, but once it is fulfilled, different robotic systems may be controlled easily and in similar way, what is demonstrated on Orpheus and Hermes robots and Orpheus EB.

The method the author wants to "focus on" for the future development is data-fusion of all three named sensors – CCD camera, thermovision and 3D proximity scanner. This method has not been finished to be practically usable yet but the preliminary results are very promising (see Fig. 7).

VII. Acknowledgements

This project was supported by the Ministry of Education of the Czech Republic under Project 1M6840770004.

This work was supported by the Academy of Sciences of the Czech Republic under project no 1ET100750408.

VIII. References

[1] L.Zalud and L. Kopecny and T. Neuzil, "Laser Proximity Scanner Correlation Based Method for Cooperative Localization and Map Building", in *Proc. 7th International Workshop on Advanced Motion Control.* Maribor, University of Maribor, Slovenia, 2002, pp 480-486.

[2] H.REverett, Sensors for Mobile Robots, Theory and Applications, AK Peters, Ltd., USA, ISBN: 1-56881-048-2, 1995.

[3] C.M.Martin and H.P.Moravec, Robot Evidence Grids, The Robotics Institute Carnegie Melon University, Pittsburgh, Pennsylvania 15213, 1996

[4] T.Neuzil and L.Kopecny and L.Zalud, "Vision System of Universal Telepresence and Autonomous Robot," in *The 12th International Danube Adria Association for Automation & Manufacturing Symposium*, pp 323-324, Jena, Germany.

[5] E.Oyama and N.Tsunemoto and S.Tachi and S. Inoue, "Experimental Study on Remote Manipulation using Virtual Reality," Presence, Volume 2 Number 2, 1993, pp 112-124.

[6] T.B.Sheridan, "Telerobotics, Automation, and Human Supervisory Control", MIT Press, Cambridge, USA, 1992.

[7] E.Wise, "Applied Robotics", Prompt Publications, USA, ISBN: 0-7906-1184-8, 1999.

[8] L.Zalud, "Universal Autonomous and Telepresence Mobile Robot Navigation," in *32nd International Symposium on Robotics - ISR 2001*, Seoul, Korea, 2001, pp 1010-1015.

[9] L.Zalud, "ORPHEUS – Reconniaissance Teleoperated Robotic System" in *16th IFAC World Congress*, Prague, Czech Republic, 2005, pp. 1-6

[10] L.Zalud "Rescue Robot League - 1st Place Award Winner," RoboCup 2003: Robot Soccer World Cup VII, Springer, Germany, ISBN 3-540-22443-2, 2004