3-D Integration of Robot Vision and Laser Data With Semiautomatic Calibration in Augmented Reality Stereoscopic Visual Interface

Salvatore Livatino, Filippo Bannò, and Giovanni Muscato, *Senior Member, IEEE*

***Abstract*—This paper proposes an augmented reality visualization interface to simultaneously present visual and laser sensors information further enhanced by stereoscopic viewing and 3-D graphics. The use of graphic elements is proposed to represent laser measurements that are aligned to video information in 3-D space. This methodology enables an operator to intuitively comprehend scene layout and proximity information and so to respond in an accurate and timely manner. The use of graphic elements to assist teleoperation, sometime discussed in the literature, is here proposed following an innovative approach that aligns virtual and real objects in 3-D space and color them suitably to facilitate comprehension of objects proximity during navigation. This paper is developed based on authors' previous experience on stereoscopic teleoperation. The approach is experimented on a real telerobotic system, where a user operates a mobile robot located several kilometers apart. The result showed simplicity and effectiveness of the proposed approach.**

***Index Terms*—Augmented reality, multiple sensors, stereo vision, telerobotics.**

I. Introduction

THERE are many applications in robotics that requires intervention in unknown, inaccessible, or dangerous environments where unpredictable situations may occur, e.g., in the case of industrial robots operating in deep waters, in planetary or volcanos exploration, in applications of safety and prevention (bomb finding and disposal) [1]–[7]. Several applications for factory automation (including industrial fabrication and manufacturing processes) also need robot teleoperation and skillful intervention [8]–[12].

These applications and many more make impossible the use of fully autonomous robotic systems and require direct human teleintervention to resolve issues. These are the cases where human cognition is irreplaceable because of the high operational accuracy that is required, as well as deep environment understanding and fast decision making, e.g., to escape deadlock situations and to perform skillful maneuvering.

Manuscript received February 21, 2011; revised May 23, 2011; accepted October 17, 2011. Date of publication October 28, 2011; date of current version January 20, 2012. Paper no. TII-11-070.

S. Livatino is with the School of Engineering and Technology, University of Hertfordshire, Hatfield, AL10 9AB, U.K. (e-mail: s.livatino@herts.ac.uk).

F. Bannò is with the Perceptual Robotics Laboratory, Scuola Superiore Sant'Anna, 56127 Pisa, Italy (e-mail: fibanno@ssc.unict.it).

G. Muscato is with the Dipartimento di Ingegneria Elettrica Elettronica e Informatica, University of Catania, 95124 Catania, Italy (e-mail: gmuscato@diees.unict.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2011.2174062

Robot teleoperation systems typically rely on 2-D displays. These systems suffer of many limitations. Among them are misjudgement of self-motion and spatial localization, limited comprehension of remote ambient layout, object size, and shape. These limitations lead to unwanted collisions during navigation, as well as long training periods for an operator. An advantageous alternative to traditional 2-D (monoscopic) visualization systems is represented by the use of stereoscopic viewing.

Previous work has been proposed in mobile robot teleguide based on stereoscopic video images [13]. The use of stereoscopic viewing increased user's sense of presence and improved understanding of the remote scene structure and distance to surrounding objects.

Stereoscopic viewing improved performance on a system relying on the visual sensor only, but there are a number of other sensors a robot can rely on, and those can well complement the visual sensor output, e.g., laser, infrared, odometry, sonars, and bumpers. The use of these sensors can significantly improve teleoperation performance because of the additional information they provide.

There is, however, a problem related to contextual visualization of different sensor measurements. This information is often shown to a user independently into the same interface (e.g., on different displays), lacking of coherence and being visualized in a nonintuitive manner. Sometimes the panel size is also limited despite several information need to be shown.

This paper proposes a method to simultaneously and coherently present both video and proximity information within an augmented reality viewing context. The approach focuses on the use of augmented objects showing (laser-based) information integrated with video objects in the visualized 3-D space. The proposed approach exploits stereoscopic visualization too.

II. Multisensor Data

Vision being the dominant human sensor modality, large attention has been paid in telerobotics to the use of visual sensors. Video images provide rich and high contrasted information and this comes presented in a way that is intuitive for humans.

There are limitations to the use of visual sensor in teleoperation, e.g., the captured images come from a constrained viewpoint being a camera typically placed very close to the robot (other limitations are, e.g., related to possible delays in image transmissions [14]–[16]).

Some of the limitations associated with the use of a visual sensor can be reduced by using additional sensors. A laser sensor can be very useful to assist robot navigation and to complement visual information. We find several contributions

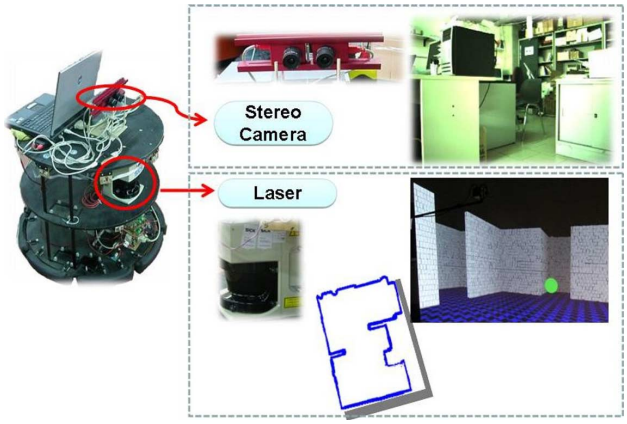

Fig. 1. Stereo camera and laser sensors with an example of the visual output they may generate. An example image of a generated 2-D floor map reconstructed by the laser is also shown.

in the literature proposing the use of camera and laser sensors, e.g., [17]–[19]. Other sensors like sonars or infrared have also been proposed, e.g., [20] and [21].

A laser range finder can be very effective in measuring position and orientation of walls and obstacles surrounding a robot, and their distance to the robot. This sensor can provide accurate estimates which has made it suitable for extracting 2-D floor maps of a robot workspace. 3-D spatial information can also be obtained by combining more sensor readings or by letting the laser device move [22]. Furthermore, it is possible to completely rely on 3-D spatial information extrapolated from 2-D laser measurements in the case of simple structured environments [15], [18], [19].

Fig. 1 shows our robotic system equipped with a stereo camera and laser sensors, and it includes an example of the visual output these sensors may generate.

Other sensors that can be exploited to assist robot teleoperation are infrared, odometry, bumpers, and sonars.

III. Augmented Reality and Telerobotics

The different sensors information can directly be presented to a user on different displays. Alternatively, a methodology needs to be devised in order to present multisensor information more effectively. A relevant issue is to present an operator proximity measurements in a way that this information can quickly and accurately be understood during teleoperation. Furthermore, in case a visual input is also present, one wishes to combine both video and range information in an intuitive and consistent manner. In fact, an intuitive visual context would reduce the cognitive workload imposed by the interface, allowing for immediate and accurate reaction. This means to improve users' performance and make more efficient to train an operator too [23], [24].

Augmented reality represents a convenient methodology to present multisensor information. It aims at integrating the observed reality with additional virtual elements, which for example could correspond to detected sensor information. An effective and consistent augmented reality representation of multisensor information is still a challenging problem. Our focus is on visual augmented reality representation.

From when it first appeared, augmented reality has been proposed to robotic systems to assist users [25]. For example, in [26], an operator is provided with visual feedback of a virtual wireframe robot superimposed over an actual physical robot. More recently, augmented reality has been proposed on several applications, e.g., on robot programming [27], [28], on maintenance tasks [24], manual assembly [29], computer-assisted surgery [30], and telemanipulation [31].

In robot teleguide applications, a number of works have shown that augmented reality can be a very effective method for status information delivery and sensor fusion [17]–[20], [32]–[34]. Visual and range information have been combined in the same visual output, e.g., in [17]–[20]. The use of colored overlay to aid operator's comprehension has also been proposed [18], [20], [35]. However, some of the approaches proposed use bidimensional augmentations to the video image to represent range information, which represents a limit to communicate depth visually [20], [32]. Other approaches reconstruct instead a 3-D from range information and project there incoming video images (in an augmented virtuality context), which provides a more intuitive visual output [18], [19]. However, in the work of Nielsen *et al.* [19], visual information integration is limited to bidimensional projection of the video image, and in the work of Ferland *et al.* [18], augmented visual aids are not proposed.

The study presented in this paper is unique in this regard. It proposes the use of colored overlays representing laser-range information integrated three-dimensionally in the incoming video scenario (in an augmented reality context). Furthermore, the system is enhanced by 3-D stereoscopic visualization.

IV. Proposed Approach

The overall aim for the proposed approach is to represent robot sensors information visually, consistently, and three-dimensionally.

1) *Visual Scenario*. To integrate visual and proximity information into the same visual scenario, so as to avoid having different displays that compete for operator's attention.
2) *Consistent 3-D Representation*. To simultaneously present multisensor information in a coherent and intuitive manner. The display is expected to show both video and graphical objects (generated by laser measurements) within a multilayer 3-D augmented reality context.
3) *3-D Visualization*. To enhance our augmented reality interface by stereo viewing. A stereoscopic visualization of the 3-D environment allows a user to comprehend environment layout, objects size and shape, and distance relations, in a way more natural and accurate than using 2-D displays.

There are a number of issues that need to be carefully addressed in order to achieve the aforementioned objectives. A main issue is information alignment. To obtain visual consistency of the different sensor measurements, we first need to calibrate the different sensors. In the literature, we find proposals for offline calibration, e.g., [36] and [37]. This type of approach requires a number of specific actions before teleoperation takes place. A robot teleoperator needs to be trained for this calibration. An automatic calibration process would certainly be more desirable. However, automatic calibration is unreliable in practice for general application contexts.

In this paper, we propose a semiautomatic calibration process where the user follows an interactive procedure to "align" different sensor information in an easy and intuitive manner. During calibration, the system automatically processes video and laser information and it shows the result of the interaction to the operator within an augmented reality context. The operator reacts to what he/she observes by adjusting the visualization settings and selecting different options. The system then responds to operator's adjustments and reprocesses sensor measurements based on the new adjustments. The result is presented to the operator who can provide further input if needed to refine the calibration.

The process can be iterated and it is expected to take no longer than a couple of minutes. The calibration setting can typically be adopted for an entire teleoperation session and can also be stored to be reused on future sessions. The calibration can also be redone or refined at any time, also during teleoperation.

The possibility for users to immediately see the results of their actions make the proposed semiautomatic process very effective (human cognition is directly involved). An interactive semiautomatic approach to system calibration has been proposed in the past in different research areas, e.g., [38]. It has, for example, been responsible for successful applications in the area of vision-based 3-D reconstruction (an area where fully automatically processing has been sought for several years). A good example is the successful work of Debevec *et al.* [39]. We believe that a semiautomatic calibration in a multisensor-based teleoperation using augmented reality, as the one proposed, can be the key to success to bring teleoperation interfaces to a new level.

After calibration, the operator starts navigating the remote environment, which he/she observes through one display frame. The visualized scene shows both real objects captured by the onboard video cameras, and virtual objects represented through graphical elements generated from laser measurements. The virtual objects look like transparent surfaces, points, lines, or circles. Both real and virtual objects are placed in the same scene consistently and aligned in the 3-D space.

The virtual objects are named "augmented layers" and they represent a visual aid to estimate distances, environment layout, objects size and shape, etc. Therefore, they help in detecting and avoiding potential obstacles. The proposed algorithm calculates the shape and 3-D position of each augmented layer.

V. Algorithm

The stereo cameras onboard the robots observe the remote space providing frontal images of the environment, and the 2-D laser sensor horizontally scans the space in front of the robot at sensor height providing a number of distance measurements with an angular resolution of one degree, ranging from $-90°$ to $+90°$. Fig. 2 illustrates the video and laser horizontal field of views.

At the beginning of the navigation, maximum and minimum distance values are set within the laser data range. The particular values depend on the application and on the characteristics of the environment. A color lookup table is consequently generated, where the color "red" represents the smallest range value (closest distance) and the color "green" represents the largest range value (farthest distance). Each distance value lying in between these two extremes is mapped to a color ranging between red and green. Fig. 5(b) shows an example of measured laser values of the environment surrounding the robot.

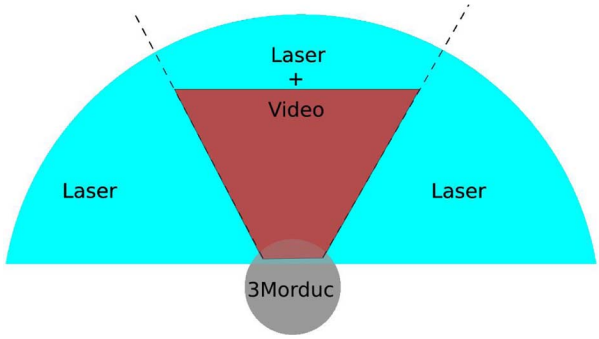

Fig. 2. Top-view representation of the laser and video sensors visibility related to robot frontal area.

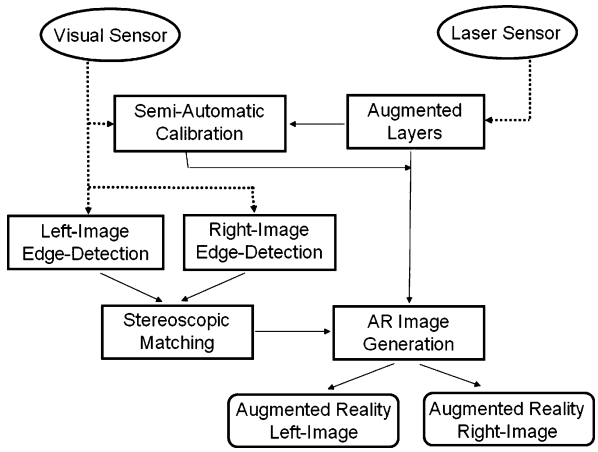

Fig. 3. Main processing steps of the proposed algorithm.

The main processing steps of the proposed algorithm are the following.

1) Augmented layers.
2) Semiautomatic calibration.
3) Edge detection.
4) Stereo matching.
5) AR-image generation.

The processing scheme is illustrated in the diagram of Fig. 3.

A. Augmented Layers

During this phase, semitransparent colored overlays are generated based on laser measurements. The degree of transparency of the overlays can be adjusted according to the operator preferences. We propose four types of augmented layers.

1) *Proximity Walls*. Lines are drawn on the horizontal plane by connecting neighbor laser values. Flat 3-D walls can then be generated by elevating those lines. The generated wall surfaces are colored, made transparent, and visualized superimposed to the video images incoming from the robot cameras. Fig. 4 illustrates the concept. The process of connecting neighbor values is carefully designed so as to detect discontinuities and discard outliers. The wall surfaces are color mapped according to the detected proximity values

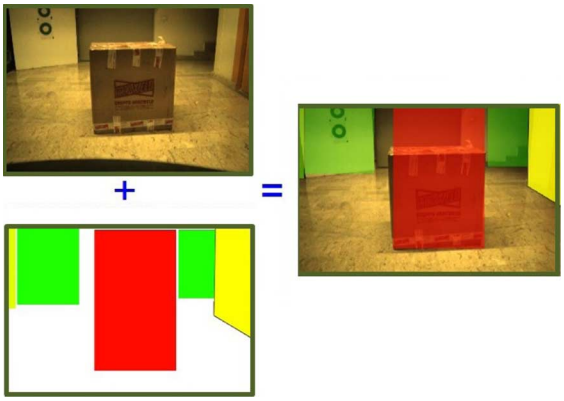

Fig. 4. Simple example illustrating the use of proximity walls. (Top left) Original video image. (Bottom left) Generated proximity walls. (Right) Resulting augmented reality view.

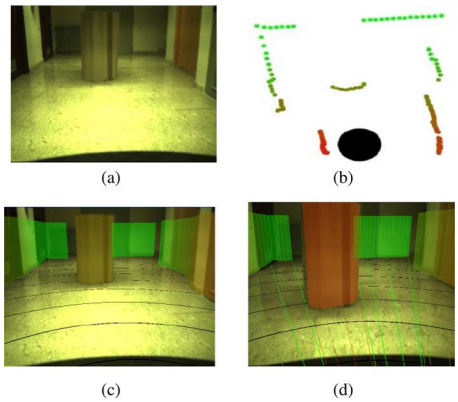

Fig. 5. Generated augmented reality views. (a) Original video image. (b) Colors associated to laser measurements. (c) and (d) Proximity walls associated with laser measurements in two different cases of distance from robot.

so that operators can visually associate a color to a distance value and get immediate understanding of possible hazards. The colors in the red–yellow–green range have a strong impact on an operator than other colors because they are conventionally associated with danger–caution–safety [40]. A different color scheme or a configurable one can nonetheless be used depending on application or cultural context. The transparent effect is obtained by manipulating the original intensity values according to the color wished for the augmented layer. The intensity values can also be adapted to current video context (e.g., illumination and background color). Fig. 5(c) and (d) shows examples for the environment in Fig. 5(a).

2) *Laser Rays*. An alternative or addition to proximity walls is a "ray effect" that can be inserted into incoming video images. The generated virtual rays depart from the robot, e.g., from the laser sensor estimated position and they reach the closest object on the specific direction (therefore, they span the laser-measured distance value). The rays are shown for a desired subset of directions, (e.g., those related to closest objects only). The rays can be colored based on the lookup table values as previously described. Fig. 6 shows example applications. In Fig. 13, both *laser rays* and *proximity walls* are used.

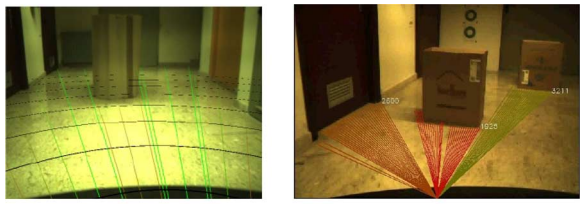

Fig. 6. Examples of generated laser rays. The rays color change according to laser measure.

3) *Concentric Circles*. A number of virtual concentric circles can be inserted into the visualized environment typically positioned at ground level to facilitate operator's space understanding. They are shown in Fig. 5.
4) *Distance Values*. The estimated distance values between robot and objects can be superimposed to the video image. This type of information can be useful when accurate quantification of the estimated distances is required, e.g., during very slow careful motion. The numbers displayed can be colored while their size and thickness can be adjusted (manually or automatically) depending on navigation context. For example, the figure size can be inversely proportional to robot speed. Fig. 6(b) shows an example where distance values are present.

B. Semiautomatic Calibration

The self-calibration phase is to align visual and laser sensors information. The result is a function that maps each laser measurement to an image pixel. This allows for determining where augmented layers should be located in the 3-D space. The layers appearance can, therefore, be projected consistently in the visualized scene.

Initially, we know the camera and laser field of views, and we have a rough estimate of the relative position between camera and laser. We assume each camera is oriented such that its axes is parallel to that of the laser. Based on the aforementioned, the horizontal image coordinate corresponding to each laser measurement is estimated from the input given by the operator.

The operator observes laser measurements projected onto the video image, which gives him/her a clear visual feedback of the level of sensors alignment. The operator observes possible misalignments between the transparent layers and the visualized environment objects, and he/she consequently resolves them by manually adjusting the calibration parameters present in the visual interface. The operator may start with adjusting the focal length and proceeds with x-, y-, and z-axis. Fig. 7 shows example images for this step.

The input provided by the operator allows the system to estimate camera focal length and the horizontal offset between laser and camera position. Assuming less priori knowledge, we could have a more general model. This would, however, increase the number of configurable parameters; therefore, this option has not been considered. We want to keep the calibration simple and we think the proposed assumption is sufficient for several applications.

The left and right camera images are also aligned during the calibration phase. This avoids that possible inaccuracies in

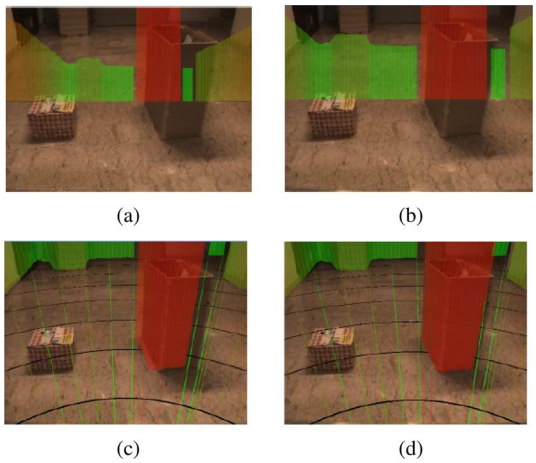

Fig. 7. Images from the calibration phase. The different transparent colored layers should superimpose on corresponding video objects. (a) Graphic overlay at the beginning of the calibration. (b) After the adjustment of the focal length. (c) After the adjustment of the y-coordinate offset. (d) After the adjustment of the xx-coordinate offset.

camera positions or differences in the location of their principal points may cause vertical disparities which reduces depth sensation. The horizontal disparity offset can on the other hand be easily tuned by hand to make stereoscopic viewing more comfortable and suitable to current scenario and operator's vision. The entire calibration process is fast because the operator gets immediate feedback on the consequence of his/her actions.

C. Edge Detection

The objective of this phase is to detect reliable edges in camera images. These edges allow augmented layers to be correctly placed within the image plane. Furthermore, the edge detection procedure allows for detecting obstacles unseen by the laser sensor. The proposed image processing runs independently on both left and right camera images. The main computational steps are the following.

1) *Contrast Stretching*. This step is to discard edges in image regions that are not relevant to aid navigation, e.g., those representing floors and ceilings, and to enhance edges in otherwise relevant areas, e.g., those representing obstacles. To achieve this result, the input image is converted to gray scale and a contrast stretching procedure is applied. The result is a new image where intensity contrast among pixels is reduced in regions with very low or high intensity, and increased in the other regions.
2) *Canny's Algorithm*. This popular edge detector is implemented using the OpenCV framework [41], [42]. An example result in our scenario is shown in Fig. 8.
3) *Invisible Obstacle Enhancement*. The obstacles outside laser visible range can be discovered and their lower edges highlighted in the incoming images, so as to get operator's attention on them. This result is possible by comparing the position of each estimated edge pixel and the position of the pixel corresponding to each laser measurement. If an edge pixel is located in the image plane below the pixel corresponding to laser measurement, an obstacle may have been detected, which is highlighted in the image plane. Fig. 9 illustrates the process through an example.
4) *Calibration Refinement*. If the comparison performed in the previous step results in an edge pixel corresponding to laser measurement (within a certain error tolerance), the detected edge position is used to refine the previously performed self-calibration. The edge position becomes then the correct projection of laser measurements and the calibration parameters are adjusted accordingly. Fig. 10 shows an example.

Fig. 8. (Right) Result of image processing applied to (left) a camera image.

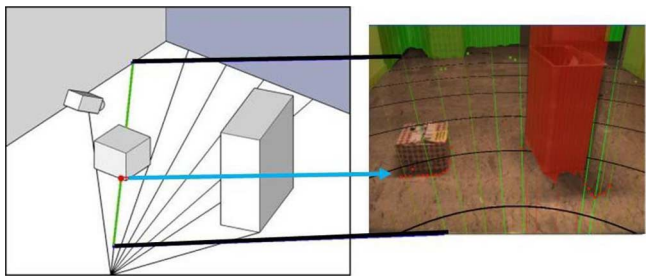

Fig. 9. Example of object invisible to our laser highlighted by the proposed *edge detection*. The left image represents the environment. The right image shows the camera view with the laser invisible object highlighted.

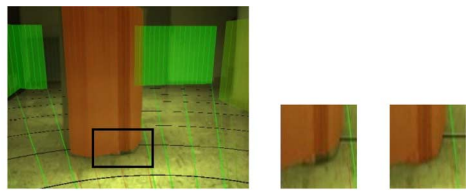

Fig. 10. (a) Alignment using *calibration refinement*. (b) Margin of the graphic overlay is initially slightly detached from the base of the box. (c) Alignment is improved after the refinement step.

D. Stereo Matching

False edges generated by our detector or image artifacts could challenge our system. A technique is then proposed to cope with this issue, which exploits correspondences between stereo images.

The corresponding pairs of stereo images (left and right) captured by the onboard cameras and sent through the network are very similar but not identical. They differ by the stereoscopic viewing difference (images are taken slightly apart from each other), This image difference is what causes an observer to appreciate the 3-D effect [43].

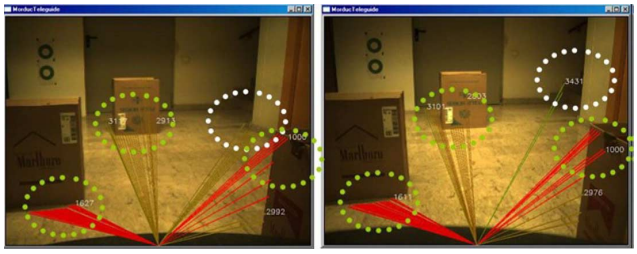

Fig. 11. Example of correspondent and noncorrespondent features in stereo images. The green circles show correspondent features, while the white circles show a noncorrespondent one.

The consequence of having not identical images is that the same image processing may lead to (slightly) different results. An algorithm is then proposed to select only those edges that can be matched in a stereo-image couple. This algorithm classifies edges as "corresponding" and "noncorresponding." In particular:

1) *Corresponding Edges*. The edges detected represent the same environment feature in corresponding stereo images. In other words, the same part of the same object is detected on both left and right images. The detected edges may not be located in the same image pixel on corresponding left and right images.
2) *Noncorresponding Edges*. The edges detected represent different environment features in corresponding stereo images. This is especially the case for features appearing close to image borders, but it also happens for some other features, e.g., when they are "weakly" represented.

The proposed algorithm is applied to each pair of stereo-images (left and right), and it follows the procedure in the following.

1) For each image within a stereo pair, each detected edge is assigned the corresponding laser measure. This will depend on edge position within the image and the result of the self-calibration.
2) If the two assigned laser measures for a stereo pair have different values, then that edge is classified as "noncorresponding edges." In this case, at least one of the two edges is likely an image artifact.
3) If the two assigned laser measures for a stereo pair have the same value, then that edges are classified as "corresponding edges." To be classified as "corresponding," we also require that the edges' highest point (in the image vertical axis) is lower than a set threshold.

Fig. 11 illustrates examples of corresponding (green circles) and noncorresponding edges (white circles).

E. AR Image Generation

The previously produced augmented layers are now aligned in the final images according to the detected edges to generate the final augmented reality image. In case a pair of corresponding edges has a disparity value above a predetermined threshold, the augmented layer position is estimated as the average of the two edge positions. This is to allow for comfortable stereoscopic visualization.

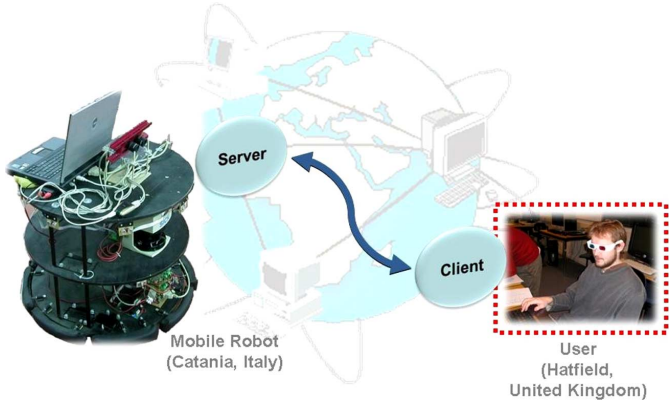

Fig. 12. Schematization of the system setup.

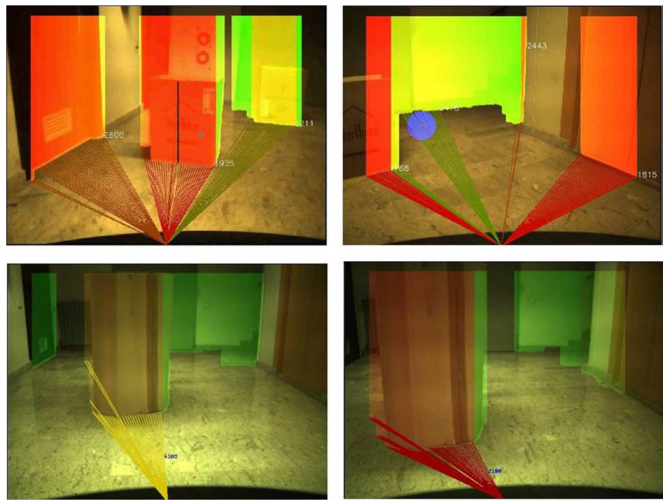

Fig. 13. Snapshots of different outputs as they appear to our operators during test trials.

The *noncorresponding edges* are considered outliers and will not generate augmented layers. This is relevant to visualize a correct output and also to avoid generation of isolated augmented layers which would cause uncomfortable stereo viewing to an operator. For example, in Fig. 11, we have noncorresponding features which would generate uncomfortable viewing if observed in stereo.

VI. Experimentation

A. System Setup

The proposed experimentation aimed at assessing feasibility and performance of the proposed augmented reality visual interface on a real setting. The proposed telerobotic system allowed users to operate a robot located approximately 2500 km away from them. The robot was located at the Robotics Lab at the University of Catania, Italy, while the operator sat at the 3D Visualization and Robotics Lab, University of Hertfordshire, Hatfield, U.K. Fig. 12 shows a schematization of the system setup.

A simple working environment was proposed which can be observed in Fig. 13. The environment was suitable for laser detection and it can represent a typical working scenario. The proposed laser system has nonetheless been proven reliable also on more articulated setups [15].

The robot was equipped with a stereoscopic video camera (Videre Design STH-MDCS2-VAR-C). The baseline of the onboard video cameras was suitably set according to the expected scenario [45]. This was 7 cm, which is very close to the average human interocular distance (6.5 cm), therefore, providing a realistic viewing experience. The robot also mounted a bidimensional laser rangefinder (Sick LMS series).

The telerobotic system was an evolution of the system adopted in [13], but the system features a new operational setup and processing scheme because of the different sensors information being processed and displayed simultaneously. The processing steps illustrated in Fig. 3 took place at operator's site. We used the C++ programming language, the OpenCV libraries for image processing and the OpenGL libraries for the graphical rendering.

B. Experimental Design

The aim of the evaluation was to assess the effectiveness of the algorithms proposed for information alignment, image processing, and stereoscopic viewing. The research question involved the following three aspects.

- *Semiautomatic calibration*. How simple, intuitive, and reliable is the proposed semiautomatic calibration process?
- *Image Processing*. How reliable and suitable to our application is the proposed edge detection and stereo matching processing?
- *Stereoscopic Viewing*. How effective and comfortable is to have stereoscopic viewing in the proposed mixed reality context?

The usability evaluation was designed following recommendations gathered from the literature on evaluation of virtual reality applications, e.g., [44]. The study is a within-subject evaluation with 15 participants chosen among students and staff of the university. The target population was composed of participants with varying backgrounds having mostly no experience with teleoperation, but a half of them were frequent video games players. The age of the participants ranged between 22 and 44, with an average of 28.9. Each participant was asked to teleguide the remote robot system for approximately 30 min performing typical robot movements such as forward/backward translation and rotations parallel to floor.

Each run produced log files containing quantitative information related to robot positions, sensor data, and timings. Images from the navigation trials were also stored. Users' observations were noted during and after test trials. The users were asked to answer a questionnaire designed according to the research question. Fig. 13 shows snapshots of different outputs as they appear to our operators during test trials.

C. Test Outcome

1) Telecommunication and Delays: We have estimated time delays related to response to user's actions. As observed in previous studies, the delay was nearly constant and had an average of 1 s. Such a delay, expected not to cause significant decrease of performance [14], allowed for interactive teleoperation. The robot could be well managed (at the price of a slower drive). We considered the aforementioned delay as a realistic setting for many applications requiring visual feedback in teleoperation. The delay was mainly due to the transmission of the two video-images, with resolution of 640×480 pixels, which were sent jpeg compressed through the network. The time to transmit the laser-sensor information as well as to compress/decompress the video images were negligible. The laser data sent through the network were only those taken at the same time of the captured images (to keep time consistency in the augmented reality scenes). The time delay may vary accordingly to connection speed. To analyze teleoperation performance under different time delays was, however, not among the objectives of our study. We were interested in testing our system under a typical teleoperation delay.

2) Result Analysis: The results of the quantitative and qualitative data analysis were very encouraging proving feasibility and reliability of the proposed methodology. In particular:

- *Semiautomatic calibration*. All users agreed that the semiautomatic calibration was an intuitive process easy to get acquainted with, which confirmed this is a strong aspect of the proposed method. The typical error in object misalignment was in average 5 pixels. We experienced a maximum misalignment of 31 pixels (represented in bottom right of Fig. 13). With our method, it is anyway always possible to stop during navigation to reexecute the semiautomatic calibration process (e.g., when the error becomes unsuitable for the task). We have learnt that the effectiveness of the calibration process depends on the initial view. If this represented a very simple environment area with objects at the same distance from the robot, the calibration may turn imprecise leading to (slight) misalignments between environment objects and augmented layers. Conversely, a better choice for the initial view, e.g., having three objects at different distances from the robot, did not require to reexecute the calibration on long runs.
- *Image Processing*. The edge detection was successful on all the input images. It allowed reliable edges to be detected on most relevant objects and it provided correct visualization of augmented layers. The procedure for invisible objects enhancement worked very well. This happened because as typical true edges appeared grouped along a line while false edges were typically isolated, therefore, not attracting operator's attention. The general limit of the proposed procedure may be the complexity of the robot operating environment (our test environment was relatively simple). Nonetheless, a simple factory-like environment as in our experiments can be found in many real applications. The stereoscopic matching procedure worked well but we should notice that it needs careful tuning on the type of the expected environment to effectively recognize outliers and to allow for comfortable stereo viewing. Situations like that represented in Fig. 11 would be resolved by the proposed algorithm.
- *Stereoscopic Viewing*. It had already been demonstrated in previous work the great advantage of using stereo viewing in separate experiments using either video images [13] or laser-based images [15]. This time both modalities were put together, which represented a challenge being the two outputs of different nature (video and graphic). Nonetheless, this did not lead to problems with the proposed algorithm and all users appreciated the augmented reality stereoscopic effect. Possible misalignments between video

and graphic images could be spotted occasionally and they were emphasized when watching the augmented reality scenario in stereo viewing. Nonetheless, the support to navigation provided by stereoscopic viewing was judged very high. In order to obtain comfortable viewing, it was important that the stereo-matching process was effective in discarding false matches and outliers, and that the semiautomatic calibration process was capable of removing vertical disparities (easy to be spotted).

The performed test succeeded in demonstrating feasibility of the proposed approach and it provided useful insight about its use and reliability in indoor factory-like scenarios.

VII. Conclusion

This paper presented a new methodology for robot teleoperation based on augmented reality visualization further enhanced by 3-D stereoscopic viewing and 3-D graphics. The aim was a users observation accurate and timely.

A method was described which included a strategy for semiautomatic calibration and image processing that led to a consistent way to visually represent video and laser information. An algorithm was proposed and implemented on a real telerobotic systems. The results proved feasibility and simplicity of the proposed method. The calibration worked very well and it was judged intuitive and versatile. The system made no relevant errors on a simple working environment which represents many applications. The edge detection and stereo matching were robust which resulted in augmented layers correctly placed into the generated views. The stereoscopic visualization was comfortable (and relevant outliers were removed). The stereo-matching process needed to be tuned to the specific application contexts.

The performed experimentation was a fundamental phase before proceeding with further developments of the proposed methodology, which will include a more articulated design of the augmented layers (including configurable color schemes) and a higher number of robot sensors represented inside one visual output. The use of augmented reality visualization in telerobotic applications is considered by the authors the natural step forward and it will certainly become very popular in the near future.

References

[1] R. Murphy, "Human-robot interaction in rescue robotics," *IEEE Trans. Syst., Man, Cybern., C, Appl. Rev.*, vol. 34, no. 2, pp. 138–153, May 2004.

[2] Z. Zheng, M. Shugen, L. Zhenli, and C. Binggang, "Communication mechanism study of a multi-robot planetary exploration system," in *IEEE Int. Conf. Robot. Biomimet.*, Dec. 2006, pp. 49–54.

[3] G. Muscato, D. Caltabiano, S. Guccione, D. Longo, M. Coltelli, A. Cristaldi, E. Pecora, V. Sacco, P. Sim, G. S. Virk, P. Briole, A. Semerano, and T. White, "ROBOVOLC: A robot for volcano exploration result of first test campaign," *J. Ind. Robot.*, vol. 30, no. 3, pp. 231–242, 2003.

[4] R. Safaric, M. Debevc, R. M. Parkin, and S. Uran, "Telerobotics experiments via internet," *IEEE Trans. Ind. Electron.*, vol. 48, no. 2, pp. 424–431, Apr. 2001.

[5] A. Kouba, R. Severino, M. Alves, and E. Tovar, "Improving quality-of-service in wireless sensor networks by mitigating hidden-node collisions," *IEEE Trans. Ind. Informat.*, vol. 56, no. 3, pp. 299–313, Aug. 2009.

[6] X. Shao and D. Sun, "Development of a new robot controller architecture with FPGA-based IC design for improved high-speed performance," *IEEE Trans. Ind. Informat.*, vol. 54, no. 4, pp. 312–321, Nov. 2007.

[7] R. Wirz, R. Marin, M. Ferre, J. Barrio, J. M. Claver, and J. Ortego, "Bidirectional transport protocol for teleoperated robots," *IEEE Trans. Ind. Electron.*, vol. 56, no. 9, pp. 3772–3781, Sep. 2009.

[8] P. Arena, P. Di Giamberardino, L. Fortuna, F. La Gala, S. Monaco, G. Muscato, A. Rizzo, and R. Ronchini, "Toward a mobile autonomous robotic system for Mars exploration," *J. Planet. Space Sci.*, vol. 52, pp. 23–30, 2004.

[9] T. Chang, P. Jaroonsiriphan, M. Bernhardt, and P. Ludden, "Web-based command shaping of cobra 600 robot with a swinging load," *IEEE Trans. Ind. Informat.*, vol. 2, no. 1, pp. 59–69, Feb. 2006.

[10] S. Bernardi and J. Campos, "Computation of performance bounds for real-time systems using time Petri Nets," *IEEE Trans. Ind. Informat.*, vol. 5, no. 2, pp. 168–180, May 2009.

[11] R. Bautista-Quintero and M. J. Pont, "Implementation of H-infinity control algorithms for sensor-constrained mechatronic systems using low-cost microcontrollers," *IEEE Trans. Ind. Informat.*, vol. 4, no. 3, pp. 175–184, Aug. 2008.

[12] G. Astuti, G. Giudice, D. Longo, C. Melita, and G. Muscato, "An overview of the "volcan project": An UAS for exploration of volcanic environments," *J. Intell. Robot. Syst.*, vol. 54, pp. 471–494, 2009.

[13] S. Livatino, G. Muscato, C. Koeffel, S. Sessa, C. Arena, A. Pennisi, D. Di Mauro, and E. Malkondu, "Mobile robotic teleguide based on video images," *IEEE Robot. Autom. Mag.*, vol. 14, no. 4, pp. 58–67, Dec. 2008.

[14] L. J. Corde, C. R. Caringnan, B. R. Sullivan, D. L. Akin, T. Hunt, and R. Cohen, "Effects of time delay on telerobotic control of neural buoyancy," in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington, DC, 2002, pp. 2874–2879.

[15] S. Livatino, G. Muscato, S. Sessa, and V. Neri, "Depth-enhanced mobile robot teleguide based on laser images," *Spec. Issue Design Control Methodol. Telerobot., Elsevier Mechatron. J.*, vol. 20, no. 7, pp. 739–750, 2010.

[16] F. Zeiger, M. Schmidt, and K. Schilling, "Remote experiments with mobile-robot hardware via internet at limited link capacity," *IEEE Trans. Ind. Electron.*, vol. 56, no. 12, pp. 4798–4805, Dec. 2009.

[17] H. Baltzakis, A. Argyros, and P. Trahanias, "Fusion of Laser and Visual Data for Robot Motion Planning and Collision Avoidance," in *Machine Vision and Applications*. New York: Springer-Verlag, 2003.

[18] F. Ferland, F. Pomerleau, C. T. Le Dinh, and F. Michaud, "Egocentric and exocentric teleoperation interface using real-time, 3D video projection," in *Proc. 4th ACM/IEEE Int. Conf. Human Robot Interact.*, 2009, pp. 37–44.

[19] C. W. Nielsen, M. A. Goodrich, and R. W. Ricks, "Ecological interfaces for improving mobile robot teleoperation," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 927–941, Oct. 2007.

[20] R. Meier, T. Fong, C. Thorpe, and C. Baur, "A Sensor Fusion Based User Interface for Vehicle Teleoperation," Field and Service Robots, Pittsburgh, PA, Aug. 1999.

[21] G. Terrien, T. Fong, C. Thorpe, and C. Baur, "Remote driving with a multisensor user interface," presented at the presented at the Int. Conf. Environmental Syst., Toulouse, France, Jul. 2000.

[22] Fraunhofer IAIS Institute 3DLS: 3D Laser Scanner Dec. 2009 [Online]. Available: http://www.3d-scanner.net

[23] J. W. S. Chong, S. K. Ong, A. Y. C. Nee, and K. Youcef-Youmi, "Robot programming using augmented reality: An interactive method for planning collision-free paths," *Robot. Comput.-Integr. Manuf.*, vol. 25, pp. 689–701, 2009.

[24] U. Neumann and A. Majoros, "Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance," in *Proc. IEEE Virtual Reality Annu. Int. Symp. (VRAIS)*, Atlanta, GA, Mar. 1998, pp. 4–11.

[25] P. Milgram, S. Zhai, D. Drasic, and J. Grodski, "Applications of augmented reality for human robot communication," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Yokohama, Japan, Jul. 1993, pp. 1467–1472.

[26] A. Rastogi and P. Milgram, "Augmented telerobotic control: A visual interface for unstructured environments," in *Proc. KBS/Robot. Conf.*, Montreal, PQ, Canada, Oct. 1995, pp. 16–18.

[27] Intelligent Robot Painting Dec. 2009 [Online]. Available: http://www.inropa.com, Inropa Company

[28] T. Pettersen, J. Pretlove, C. Skourup, T. Engedal, and T. Lokstad, "Augmented reality for programming industrial robots," in *Proc. Int. Symp. Mixed Augmented Reality*, Tokyo, Japan, Oct. 2003, pp. 319–320.

[29] J. Molineros and R. Sharma, "Computer vision for guiding manual assembly," in *Proc. IEEE Int. Symp. Assembly Task Plann.*, Fukuoka, Japan, May 2001, pp. 362–368.
[30] R. T. Azuma, "A survey of augmented reality," *Presence*, vol. 6, no. 4, pp. 355–385, 1997.
[31] R. Marin, P. J. Sanz, P. Nebot, and R. Wirz, "A multimodal interface to control a robot arm via the web: A case study on remote programming," *IEEE Trans. Ind. Electron.*, vol. 52, no. 6, pp. 1506–1520, Dec. 2005.
[32] M. Baker, R. Casey, B. Keyes, and H. A. Yanco, "Improved interfaces for human-robot interaction in urban search and rescue," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2004, vol. 3, pp. 2960–2965.
[33] J. Scholtz, J. Young, J. Drury, and H. Yanco, "Evaluation of human-robot interaction awareness in search and rescue," in *IEEE Int. Conf. Robot. Autom.*, 2004, vol. 3, pp. 2327–2332.
[34] H. A. Yanco and J. Drury, "Where am I? Acquiring situation awareness using a remote robot platform," in *IEEE Int. Conf. Syst., Man Cybern.*, 2004, pp. 2835–2840.
[35] B. Giesler, T. Salb, P. Steinhaus, and R. Dillmann, "Uning augmented reality to interact with an autonomous mobile platform," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2004, vol. 1, pp. 1009–1014.
[36] Y. Bok, Y. Hwang, and I. So Kweon, "Accurate motion estimation and high-precision 3D reconstruction by sensor fusion," in *IEEE Int. Conf. Robot. Autom.*, Roma, Italy, Apr. 2007, pp. 4721–4726.
[37] Q. Zhang and R. Pless, "Extrinsic calibration of camera and laser range finder," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sendai, Japan, Sep. 2004, pp. 2301–2306.
[38] M. A. Al-Mouhamed, O. Toker, and A. Al-Harthy, "A 3-D vision-based man-machine interface for hand-controlled telerobot," *IEEE Trans. Ind. Electron.*, vol. 52, no. 1, pp. 306–319, Feb. 2005.
[39] P. E. Debevec, C. J. Taylor, and J. Malik, "Modeling and rendering architectures from photographs: A hybrid geometry- and image-based approach," in *Proc. SIGGRAPH*, Aug. 1996, pp. 11–20.
[40] E. Williams and B. Andrews, *The Non-Designer's Design Book*. Berkeley, CA: Peachipt Press, 1994.
[41] Open Source Computer Vision: OpenCV [Online]. Available: http://opencv.willowgarage.com 2009, OpenCV Software
[42] J. Canny, "A computational approach to edge detection," *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms*, vol. 184, pp. 87–116, 1987.
[43] G. J. Kim, *Designing Virtual Reality Systems. The Structured Approach*. New York: Springer-Verlag, 2005, .
[44] S. Livatino and C. Koeffel, "Handbook for evaluation studies in virtual reality," in *Proc. IEEE Symp. Virtual Environ., Human-Comput. Interface Meas. Syst.*, Ostuni, Italy, 2007, pp. 1–6.
[45] M. Ferre, R. Aracil, and M. A. Sanchez-Uran, "Stereoscopic human interfaces," *IEEE Robot. Autom. Mag.*, vol. 14, no. 4, pp. 50–57, Dec. 2008.

Salvatore Livatino received the M.Sc. degree from the University of Pisa, Pisa, Italy, and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 1993 and 2003, respectively.

He is currently an Associated Professor in the School of Engineering and Technology, University of Hertfordshire, Hatfield, U.K. He was a Researcher at the Scuola Superiore Sant´Anna, Pisa (from 1993 to 1997), the University of Leeds, U.K. (1995), INRIA Grenoble, France (1996), and the University of Edinburgh, U.K., in 2001. He was at Aalborg University, Denmark, for 12 years, first as a Research Fellow, then as an Assistant and an Associate Professor. He is the author of several journal and conference papers, and contributed to several European projects and U.K. grants. His teaching experience has mostly been within problem-based learning and multidisciplinary educations. His research interests include 3-D visualization, virtual/augmented reality, telerobotics, telemedicine, computer graphics and vision, and computer games, with research combining those fields, e.g., 3-D stereoscopic vision and augmented reality visual interfaces for robot telecontrol, depth-enhanced medical endoscopic teleintervention, and immersive computer games.

Filippo Bannò received the B.Sc. and M.Sc. degrees in computer engineering from the University of Catania, Catania, Italy, in 2007 and 2010, respectively. He received an additional specialization degree from Scuola Superiore di Catania (high excellence education center), Catania, in 2010. He is currently working towards the Ph.D. degree at Perceptual Robotics Laboratory, Scuola Superiore Sant´Anna, Pisa, Italy.

He was a one-year Visiting Researcher at the University of Hertfordshire, U.K. His current research interests include virtual reality systems, 3-D graphics, computer vision, and robotics.

Giovanni Muscato (M'91–SM'01) received the Degree in electrical engineering from the University of Catania, Catania, Italy, in 1988.

After completing graduation, he was with the Centro di Studi sui Sistemi, Turin, Italy. In 1990, he joined the Dipartimento di Ingegneria Elettrica Elettronica e Informatica, University of Catania, where he is currently a Full-Time Professor of Robotics and Automatic Control. He was the Coordinator of the EC project Robovolc and is the local Coordinator of several national and European projects in robotics. He is the author of more than 250 papers in scientific journals and conference proceedings and three books in the fields of control and robotics. His current research interests include service robotics and the use of soft-computing techniques in the modeling and control of dynamical systems.

Prof. Muscato is the Co-Chair of the Service Robotics Technical Committee. He is with the Board of Trustees of the Climbing and Walking Robots Association.