Assisted Operation of a Robotic Arm Based on Stereo Vision for Positioning Near an Explosive Device

Andres Montoya Angulo *, Lizardo Pari Pinto †, Erasmo Sulla Espinoza †, Yuri Silva Vidal ‡ and Elvis Supo Colquehuanca ‡

Department of Electronic Engineering, Universidad Nacional de San Agustín de Arequipa, Arequipa 04001, Peru

* Correspondence: jmontoyaan@unsa.edu.pe
† These authors contributed equally to this work.

Abstract: This document presents an assisted operation system of a robotic arm for positioning near an explosive device selected by the user through the visualization of the cameras on the screen. Two non-converging cameras mounted on the robotic arm in camera-in-hand configuration provide the three-dimensional (3D) coordinates of the object being tracked, using a 3D reconstruction technique with the help of the continuously adaptive mean shift (CAMSHIFT) algorithm for object tracking and feature matching. The inverse kinematics of the robot is implemented to locate the end effector close to the explosive so that the operator can perform the operation of grabbing the grenade more easily. Inverse kinematics is implemented in its geometric form, thus reducing the computational load. Tests conducted with various explosive devices verified the effectiveness of the system in locating the robotic arm in the desired position.

Keywords: assisted operation; CAMSHIFT; robotic arm; explosive device

1. Introduction

Increasingly, robots are being used to carry out activities in various disciplines that are normally carried out by humans, such as assistance in surgery [1], assembly plants [2], domestic activities [3], and others. However, robots used in dangerous environments, such as rescuing people [4,5], handling radioactive elements [6], space exploration [7,8], or deactivation of explosives [9,10] have greater prominence. In addition, development of an explosive ordnance disposal (EOD) robot location system with enhanced features is being advanced as part of the ongoing project [11,12]. In order to achieve this, robots must have advanced mobility and manipulation skills that allow the operator to perform tasks very easily and quickly [13,14]. Currently, the common way to control these arms is through buttons or a joystick, and because these robots perform repetitive tasks, they generate stress for the operator [15,16]. This stress is generated because the operator tries to reach an object with the robotic arm, but by not having a clear reference of the distance at which the object is, a load of stress is generated. In addition, there is a strong pressure knowing that explosive devices are being handled [17].

The vision system provides three-dimensional information on the location of the object to be manipulated with the help of the two-dimensional location in the image that the operator provides through a touch screen, thus forming an assisted operation system. In Nadarajah’s article [18], the vision system used in Robot Soccer systems is described in a general way. First, the positioning of the cameras in parallel configuration is listed; these are used in a robotic soccer system for both federation of international robot-soccer associations (FIRA) and RoboCup. Machine vision is classified into three types: omnidirectional, binocular/stereo, and monocular. Subsequently, the image processing algorithms and the reference related to their advantages and disadvantages are explained. One of the algorithms that stands out is continuously adaptive mean shift (CAMSHIFT), an algorithm
that is used to follow a moving object. For the particular case of stereo vision, the distance is estimated with the stereo calibration of the cameras, the intrinsic and extrinsic parameters of the cameras are obtained, and then the distance of the objects captured in both cameras is calculated. In the paper by Zhao [19], a foldable manipulator applied to the five-degrees-of-freedom (5-DOF) EOD robot is presented, and the Denavit–Hartenberg parameters (D-H) method is used to introduce the virtual joint, in order to establish the direct kinematics model of the manipulator. In this way, they demonstrate that a 5-DOF robotic arm is adequate to perform this task. In [20], a system was developed that controls a robotic arm to grab an object through stereo vision in parallel configuration and with a fixed camera (the camera is not mounted on the robotic arm but placed on a turret that has a view of the arm). In addition, object tracking is achieved due to distance estimation through the triangulation method. The system is checked with some operations described in the document. In another article [21], the triangulation method is also incorporated through a stereo vision system, which grabs an object with a robotic arm that uses the CAMSHIFT algorithm that provide better tracking. A stereo vision system placed on a robotic arm in an eye-in-hand configuration [22], together with a target selection system through a touch screen [23,24], could provide an interesting solution to the problem that operators have to bring the robotic arm closer to a specific location without generating stress load.

This article presents a system that controls the movement of a robotic arm in order to grab an explosive device using non-convergent stereo vision, as part of the multimodal system developed for this project [25]. First, the police officer of the explosive disposal unit (UDEX, by its acronym in Spanish) selects the explosive device to be hit with the proposed user interface (UI). The coordinates (X,Y,Z) of the target are calculated: coordinate Z by the two cameras configuration and triangulation method, and X and Y by perspective relations. Subsequently, the CAMSHIFT algorithm maintains the tracking of the object during the movement of the arm and, at the same time, detects the corresponding characteristic (center of mass of the object) in both images. The advantage of this proposal is that the autonomous detection of some characteristic of the object to be manipulated is no longer necessary. The possibility of false positives due to disturbances such as shadows, excess or lack of lighting, among others, is eliminated due to this system, which is robust and useful in field applications [26]. Finally, the position of the target is sent to the block of the inverse kinematics of the arm, previously calculated, using geometric techniques that reduce the computational cost. The assistance system is evaluated from the point of view of usability and user experience using the NASA-TLX (NASA taskload index) [27] and the evaluation methods SUS (System usability scale) [28], to verify that this proposal reduces operator stress levels.

The study focused on the operator assistance system, using design techniques and procedures related to vision system configuration, camera and robot calibration, and system performance analysis. The rest of the document is structured as follows: Section 2 presents the materials and methodology of the proposed system; this section is composed of the design and explanation of the interface, mathematical analysis of the stereo cameras, and control of the robotic arm. Section 3 explains the experimental results and presents discussion. Finally, conclusions and future work are presented in Section 4.

2. Materials and Methods

In Figure 1, the block diagram of the proposed assisted operation system is shown. First, the UDEX agent, through the user interface, selects the explosive device that he wishes to reach with the robotic arm. After this information is sent to the algorithm developed to calculate the angles that the robotic arm must move by calculating inverse kinematics, the stereo cameras send the captured frames to perform image processing and determine the estimated distance of the object. It is essential for the algorithm to work correctly. Finally, the values of the angles are sent to the robot and its movement is carried out. Figure 2 shows two images that show the proposed system. In the first image, the UDEX squad agent is selecting the target on the screen so that the proposed algorithm works; in the
second image, the algorithm has blurred the background with a Gaussian filter to estimate the distance to the grenades selected by the agent more quickly and accurately. The detailed development of this algorithm can be found in a previously published article [29].

![Block diagram of the proposed system.](image1)

**Figure 1.** Block diagram of the proposed system.

![Proposed visual selection system: (a) UDEX agent selecting explosive device; (b) proposed object distance estimation algorithm.](image2)

**Figure 2.** Proposed visual selection system: (a) UDEX agent selecting explosive device; (b) proposed object distance estimation algorithm.

The architecture developed in this system that moves the robotic arm through stereo vision is shown in Figure 3. It consists of five modules: user interface and assistance, stereovision analysis, tracking algorithm, manipulator kinematics, and control of the robotic arm.

![Architecture developed in this system.](image3)

**Figure 3.** Architecture developed in this system.

### 2.1. User Interface and Support (UI)

The proposed UI integrates the functions required for this system [29]. In this document, the functions related to the distance estimation of the explosive device are described; the UI is displayed in Figure 4.

#### 2.1.1. Positioning

In the button panel on the lower left side of the UI, three options are displayed that allow the operator to place the arm in the best position (in front and center of the explosive device); the distance to the object is estimated using triangulation.
2.1.2. Operator Image Adjustment

The other range of options is found in the lower right area of the UI; the operator has the freedom to manipulate the characteristics of the cameras independently (left and right) to achieve similar characteristics between the two cameras. These features are: Brightness, Hue, Contrast, Camera Gain, Saturation, Exposure, and Zoom. In case an inadequate configuration is obtained, it is possible to return to the original configuration using the “Default” button located at the top of the panel.

2.1.3. Target Selection

Finally, the operator can select the target via the “FRAME” button, which provides the option to follow the selected object and estimate its distance; see Figure 4.

![Figure 4. Target selection in the user interface.](image)

2.2. Stereo Vision Analysis

The stereo vision based distance estimation [30] is shown in Figure 5, in which the stereo camera is composed of two cameras. The points $O_1$ and $O_2$ are the optical centers of both cameras; $T$ is the baseline (distance between the centers of the cameras); and $f$ is the focal length of the lens. The point $\vec{P}$ represents the object in the real world, and $Z$ is the distance between $\vec{P}$ and the stereo cameras [29]. Using stereo vision through cameras allows the operator to have reference to the depth at which the object is located, whereas if only one camera were used, it would not be possible to obtain this [31].

![Figure 5. Schematic of the two-chamber model.](image)
To estimate the distance from the object to the base of the cameras, it is necessary to calculate the disparity between frames; see Figure 6.

**Figure 6.** Scheme of the disparity model.

The calculation of the coordinates \((X, Y, Z)\) are given by [32]:

\[
\begin{align*}
X &= \frac{x Z}{f_x} \\
Y &= \frac{y Z}{f_y} \\
Z &= \frac{f T}{d}
\end{align*}
\]  

(1)

where \(d\) is the disparity (difference of the coordinates in both images):

\[
d = |x_1 - x_2|
\]  

(2)

In the third equation of (1), it is inferred that when the distance of the object \(\overrightarrow{P}\) is greater, the disparity will be less, and vice versa: an inverse proportionality relationship. The procedure is detailed in Algorithm 1.

**Algorithm 1 Triangulation Method**

1: procedure TRIANGULATION(center\textsubscript{L}, center\textsubscript{R}, f, T) \(\triangleright\) T is baseline
2: \(x_L \leftarrow \text{center}_L[0]\)
3: \(x_R \leftarrow \text{center}_R[0]\)
4: \(\text{disparity} \leftarrow x_L - x_R\)
5: if \(\text{disparity} = 0\) then
6: \(\text{disparity} \leftarrow 1\)
7: end if
8: \(Z \leftarrow (f \times T) / \text{disparity}\)
9: \(Z \leftarrow |Z|\)
10: return \(Z\)
11: end procedure

2.3. Object Tracking Algorithm

The CAMSHIFT algorithm is based on the MeanShift algorithm; the disadvantage of MeanShift is that its region of interest (ROI) is a fixed value. Therefore, when the target object gets closer to the lens, the object in the image becomes larger and the effect of the fixed ROI is small. However, when the target object is far from the lens, the object in the image becomes smaller. The proportion of smaller objects in the ROI makes tracking unstable and causes errors in judgment [25]. The CAMSHIFT tracking algorithm has the ability to adjust the seek box on every frame. It uses the position of the centroid and the zero-order moment of the search window in the top frame to set the location and dimensions of the
search window for the next frame [33]. Figure 7 shows the flow diagram of the CAMSHIFT algorithm [34,35].

![Flowchart of the CAMSHIFT algorithm](image)

**Figure 7.** Flowchart of the CAMSHIFT algorithm.

### 2.4. Robotic Arm Control

The study of the direct kinematic problem of a robot can be carried out by different methods; a commonly used method is based on the Denavit–Hartenberg parameters [36]. It is a systematic method and best suited for modeling manipulators in series. The D-H method has been used for the development of the kinematic model of this robot due to its versatility and acceptability to model any number of joints and links of a serial manipulator. Figure 8 shows the schematic design of the 5-DOF robotic arm from which the D-H parameters in Table 1 were extracted. The most common manipulators are 3, 4, or 6 degrees of freedom. The higher the degrees of freedom, the more flexible the manipulator, but the higher the number of degrees of freedom, the more difficult the manipulator will be to control [19]. For that reason, a 5-DOF robotic arm was chosen. Table 2 shows the lengths of the links of the robot used in testing this system.
Table 1. D-H parameters.

| Joint | Joint Name | $\theta$ | $d$ | $a$ | $\alpha$ |
|-------|------------|----------|-----|-----|----------|
| 1     | Waist      | $-175^\circ \leq q_1 \leq 175^\circ$ | $d_1$ | 0   | $\pi/2$ |
| 2     | Shoulder   | $-80^\circ \leq q_2 \leq 60^\circ$ | 0   | $a_2$ | 0        |
| 3     | Elbow      | $0^\circ \leq q_3 \leq 90^\circ$ | 0   | $a_3$ | 0        |
| 4     | Forearm    | $-45^\circ \leq q_4 \leq 45^\circ$ | 0   | 0   | $\pi/2$ |
| 5     | Wrist      | $-180^\circ \leq q_5 \leq 180^\circ$ | $d_4$ | 0   | 0        |

Table 2. Link length.

| $d_1$ | 0.30 m | $a_2$ | 0.40 m | $a_3$ | 0.35 m | $d_4$ | 0.10 m |

The calculations of the arm motion matrices are shown below, where $T$ represents the position and orientation of the end effector.

$$ T = A_1^0A_2^1A_3^2A_4^3A_5^4 = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

The geometric solution of the inverse kinematics is an intuitive method that requires figures that model the Euclidean space of the robot. Given this need, the top view of the robotic arm is presented in Figure 9.

Looking at Figure 9, we see:

$$ \theta_1 = \arctan \left( \frac{p_y}{p_z} \right) $$

To solve joint $\theta_3$, the kinematic decoupling method is used; the position point $\gamma_m$ of the robot wrist is calculated from the orientation constant of the final effector $\gamma$, as shown in Figure 10.
After applying the decoupling method and solving Equation [37], the final equation of motion for $\theta_3$ can be obtained:

$$\theta_3 = \arctan \left( \frac{a_2 a_3 \sqrt{1 - \frac{R^2 - a_2^2 - a_3^2}{R^2 - a_2^2 - a_3^2}}}{R^2 - a_2^2 - a_3^2} \right)$$

(5)

To determine the value of joint $\theta_2$, first the values of the angles $\phi$ and $\psi$ must be determined. Looking at Figure 10, the following equation for the angle $\phi$ is obtained:

$$\phi = \arctan \left( \frac{d_{2m} - d_1}{\gamma m} \right)$$

(6)

The equation for the angle $\psi$ is:

$$\psi = \arctan \left( \frac{a_3 \sin \theta_3}{a_2 + a_3 \cos \theta_3} \right)$$

(7)
Due to the restrictions of $\theta_2$, the robot only has the elbow down configuration. Looking at Figure 10, $\theta_2$ can be obtained as:

$$\theta_2 = \phi + \phi$$

(8)

Angle $\theta_4$ is obtained from the relation:

$$\gamma = \theta_2 + \theta_3 + \theta_4$$

(9)

where by clearing $\theta_4$, this is obtained:

$$\theta_4 = \gamma - \theta_2 - \theta_3$$

(10)

To determine the value of joint $\theta_5$, the following equality may be used, in order to keep the end effector coordinated with the robot base:

$$\theta_5 = \theta_1$$

(11)

2.5. Evaluation Methodology

In order to validate this proposal, tests were carried out with police officers from the UDEX. At the beginning of the tests, the participants were informed about the general description of the assistance system in addition to the correct manipulation of the robotic arm and the user interface. Before starting each test, the participants had 5 min to familiarize themselves with the robot, the buttons of the classic system, the developed interface, and the developed assistance system. After the participants completed their training, tests of arm approach to the target, in this case explosive devices, were performed. Each of the agents had 5 min to achieve the greatest number of correct answers in the task of bringing the arm to the target region. A successfully completed task was for the fathom to get close to the explosive. The test scenario conditions were normal, with average room brightness. After the participants performed the robot handling test, they were provided with a sheet with the NASA-TLX test [27] and the SUS questionnaire [28]. To end the tests, each participant was interviewed to confirm that they successfully completed each questionnaire. Figure 11 shows the three steps described above.

![Figure 11](image)

**Figure 11.** Tests carried out in the work environment: (a) information on the methodology and explanation of the tests to the UDEX agents; (b) development of tests; (c) responses of the participants in the NASA-TLX and SUS tests.

3. Results and Discussion

3.1. Performance of the Proposed Algorithm

In this article, the Dobot Magician training robotic arm was used, and the two webcams were Xioami CMSXJ22A, whose resolution was $1280 \times 720$ pixels. Instead of using the clamp of the arm, it was decided to use a support for the stereo cameras, of our own manufacture, which is shown in Figure 12. This replacement was decided due to the limitation of the space in the end effector of the Dobot Magician. Prior to testing, the
cameras were calibrated using a $9 \times 8$ calibration pattern with a grid size of $30 \times 30$ mm. The Matlab Calibration Toolbox [38] was used, using the Zhang method [39].

Once the intrinsic parameters of the camera are known through a previous calibration, it was possible to start tracking the object. First, the operator selected the object to follow in the user interface. Once this was done, the algorithm obtained the coordinates of the object in the plane of each of the two images (coordinates in pixels) in order to be able to calculate the target depth and tracking frame. The coordinate of the object in the left camera was converted to the position of the object with respect to the center of the camera (mounted on the arm end effector) by applying the mathematics developed in Section 2.2. Subsequently, the values obtained from the object in 3D and in real measurements served as input data to the inverse kinematics of the arm so that it can move to said position. When the arm reached the position, the tracking stopped and the operator was notified of the successful movement; if it failed to arrive, the movement continued while the object continued to be tracked. Verification of the estimated distance of the tracked object was achieved by comparing the estimated measurement with the actual measurement of the object. Figure 13 shows a graph of the accuracy achieved by progressively placing the object at different distances, with an average accuracy of 99.18% achieved when comparing the estimated distance versus the real distance. The method proposed in this article has been compared with other distance estimation methods [40,41]. A more detailed explanation of this process can be found in [29].
Figure 13. Accuracy graph of the estimated distances. The data of the extraction method of the author Yu [34] are considered, while the data of the other method are of the author Xiu [35].

The objectives that were followed in this experiment were real explosives, provided by the UDEX squad of the Peruvian National Police. Specifically, a military hand grenade and a type 322 mortar grenade were tracked; these explosives have had the most presence in attacks in Arequipa city [42]. They are shown in Figures 14 and 15, respectively.

Figure 14. Military hand grenade: (a) tracking in the left camera; (b) follow-up in the right camera.

Figure 15. Type 322 mortar grenade: (a) left camera tracking; (b) follow-up in the right camera.

Figure 16 shows a sequence of images of the mortar tracking process.
Table 3 shows the values obtained by bringing the end effector to different desired positions, together with its corresponding error value for 15 follow-ups. It was observed that the differences between the values of the real coordinates and the coordinates reached were very small, less than 2.64% average error, with this error being smaller in Z (depth).

Figure 16. Tracking sequence of the type 322 mortar grenade: (a) left camera images; (b) images from the right camera.
Table 3. Values of different end effector points.

| #  | Real Coordinate (cm) | Reached Coordinate (cm) | Error (%) |
|----|----------------------|--------------------------|-----------|
|    | X        | Y        | Z        | X        | Y        | Z        | X | Y | Z |
| 1  | 20.7     | 12.9     | 73.2     | 20.1     | 12.8     | 71.7     | 2.9 | 0.8 | 2.0 |
| 2  | 20.1     | 11.6     | 71.8     | 19.7     | 11.5     | 70.7     | 2.0 | 0.9 | 1.5 |
| 3  | 18.9     | 11.2     | 69.2     | 18.1     | 10.9     | 68.7     | 4.2 | 2.7 | 0.7 |
| 4  | 18.2     | 10.7     | 67.8     | 17.4     | 10.6     | 66.2     | 4.4 | 0.9 | 2.4 |
| 5  | 17.3     | 10.1     | 65.3     | 16.8     | 10.0     | 64.9     | 2.9 | 1.0 | 0.6 |
| 6  | 16.5     | 9.6      | 63.1     | 16.0     | 9.2      | 62.7     | 3.0 | 4.2 | 0.6 |
| 7  | 14.8     | 9.0      | 59.9     | 14.2     | 8.7      | 59.4     | 4.0 | 3.3 | 0.8 |
| 8  | 13.9     | 8.4      | 55.4     | 13.1     | 8.2      | 55.0     | 5.8 | 2.4 | 0.6 |
| 9  | 13.1     | 7.8      | 51.1     | 12.7     | 7.6      | 50.8     | 3.0 | 2.6 | 0.6 |
| 10 | 12.6     | 7.2      | 48.2     | 11.8     | 7.0      | 47.5     | 6.3 | 2.8 | 1.5 |
| 11 | 11.4     | 6.5      | 46.6     | 11.1     | 6.1      | 45.9     | 2.6 | 6.2 | 1.5 |
| 12 | 10.3     | 6.1      | 43.2     | 9.9      | 6.0      | 42.8     | 3.9 | 1.6 | 0.5 |
| 13 | 9.5      | 5.7      | 42.1     | 9.2      | 5.4      | 40.9     | 3.2 | 5.3 | 2.9 |
| 14 | 8.7      | 5.3      | 40.7     | 8.4      | 5.0      | 39.9     | 3.4 | 5.7 | 2.0 |
| 15 | 8.0      | 5.0      | 39.0     | 7.6      | 4.9      | 38.4     | 5.0 | 2.0 | 1.5 |

The tracking sequence of the explosive device by the Dobot is shown in Figure 17.

Figure 17. Cont.
3.2. Experimental Evaluation Results

Figures 18 and 19 show the results obtained from the tests carried out using the NASA-TLX and SUS methods. These graphs show the scores obtained by the 15 participants for each of the two robot control systems: traditional control system (robot control by means of joysticks and buttons) and proposed assistance system (robot control by the graphical interface). In Figure 18, the average of each of the six categories evaluated is presented: mental demand, physical demand, temporary demand, performance, effort, and degree of frustration. In general, the assistance system developed presented a lower workload in the six categories evaluated, being more notable in the categories of frustration and mental and temporary demands. In this evaluation, a value near 20 means that it is not pleasant for the operator.

Figure 19 shows the degree of usability and workload (stress generation) of each of the two systems for each participant during the experiment. The color background of this graph shows three different scoring areas: light red for poor usability ($SUS_{score} < 50$), light yellow for good usability ($85 > SUS_{score} \geq 50$), and light green for excellent usability ($SUS_{score} \geq 85$).

Table 4 shows a summary of the results of both evaluations for the two compared systems. Statistical parameters are used to obtain more reliable values such as the mean, standard deviation, and standard error. In the NASA-TLX test column, the average value of the traditional method of buttons and joysticks is $X_{WT} = 15.13$, a considerably high value, whereas the value of the proposed method is $X_{WP} = 8.25$, making it clear that it is comfortable for operators. In the SUS tests column, it can be seen that the average score of the proposed system is $X_{SP} = 82.51$, much higher than the score obtained for the joysticks and buttons system, which is $X_{ST} = 46.65$. The proposed assistance system is considered to be a good interface, that is, it has a user-friendly interface.
Figure 18. Results of the tests carried out using the NASA–TLX test.

Figure 19. Results of the tests carried out using the SUS test.

Table 4. Experimental results of the NASA Task Load Index (NASA–TLX) and system usability scale (SUS) tests.

| Evaluation Categories | NASA-TLX | SUS |
|-----------------------|----------|-----|
|                       | Traditional Method | Proposed Method | Traditional Method | Proposed Method |
| Average               | 15.13     | 8.25 | 46.65 | 82.51 |
| Standard deviation    | 2.78      | 2.96 | 8.45  | 8.07  |
| Standard error        | 1.13      | 1.21 | 2.18  | 2.08  |
3.3. Discussion

In general, the assistance system proposed in this document is a good method to be applied in EOD handling. The system evaluated with the NASA-TLX method presents very good results in all categories. The stress load is clearly reduced compared to the traditional method of robotic arm control, which includes a lower execution time to perform the task of reaching the objective in the proposed system compared with the reference system. These data are related to the number of successes in the tests carried out, as the proposed system reduces human error to a minimum. The operator found it easier to manipulate the robot through our user interface. According to the data obtained from the SUS evaluation, the proposed assistance system has better results in usability, that is, it is very easy to use and understand its operation; this confirms that the developed system is a good interface.

In general, through the experiences of the users evaluated, the developed interface allows the manipulation of the robot to be easier and the assistance system considerably reduces the stress load of the operator. The handling of explosive devices is more efficient and safer when using this system, improving the experience of UDEX agents.

4. Conclusions and Future Work

In this document, a robotic arm control system was presented, and using two non-converging stereo cameras, the position of an explosive device selected by the operator in 3D coordinates was obtained. The triangulation method was used to determine the depth of the explosive device in relation to the location of the cameras, and the CAMSHIFT algorithm was used to track it. After the coordinates were sent to the robotic arm’s inverse kinematics block, the arm was able to reach the location of the explosive, successfully completing the test. The tests show that the system works correctly in estimating the distance of the explosive device with an accuracy of 99.18%, a higher percentage than in other related work. The results of both evaluations prove that the proposed assistance system is better than the traditional robot handling system for EOD tasks. The proposed method reduces the stress load by 18% and, moreover, the success rate is very high.

Future work will focus on improving detection accuracy and implementing a robust detection method that can overcome illumination intensity variations present in a outdoor environment.

Author Contributions: Conceptualization, A.M.A. and L.P.P.; methodology, A.M.A.; software, A.M.A. and E.S.C.; validation, E.S.E.; formal analysis, L.P.P. and Y.S.V.; investigation, A.M.A.; resources, E.S.E.; writing—original draft preparation, A.M.A.; writing—review and editing, L.P.P. and E.S.C.; supervision, E.S.E.; project administration, L.P.P.; funding acquisition, Y.S.V. and E.S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad Nacional de San Agustín de Arequipa with contract number IBA-IB-27-2020-UNSA.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: With the support of the Universidad Nacional de San Agustín de Arequipa with the following contract No IBA-IB-27-2020-UNSA and UDEX-AQP; for the development of the project through the information collected and for their invaluable guidelines to explore the facets of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vittoria, S.; Lahlou, G.; Torres, R.; Daoudi, H.; Mosnier, I.; Mazalaigue, S.; Sterkers, O. Robot-based assistance in middle ear surgery and cochlear implantation: First clinical report. *Eur. Arch. Otorhinolaryngol.* 2021, 278, 77–85. https://doi.org/10.1007/s00405-020-06070-z.
2. Baskaran, S.; Niaki, F.A.; Tomaszewski, M.; Gill, J.S.; Chen, Y.; Jia, Y.; Mears, L.; Krovi, V. Digital Human and Robot Simulation in Automotive Assembly using Siemens Process Simulate: A Feasibility Study. *Procedia Manuf.* 2019, 34, 986–994. https://doi.org/10.1016/j.promfg.2019.06.097.

3. Yamamoto, T.; Takagi, Y.; Ochiai, A.; Iwamoto, K.; Itozawa, Y.; Asahara, Y.; Ikeda, K. Human Support Robot as Research Platform of Domestic Mobile Manipulator. In *RoboCup 2019: Robot World Cup XXIII*. RoboCup 2019; Chalup, S., Niemueller, T., Suthakorn, J., Williams, M.A., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2019; Volume 11531. https://doi.org/10.1007/s00468-019-01323-3.

4. Jahanshahi, H.; Jafarzadeh, M.; Sari, N.N.; Pham, V.-T.; Huynh, V.V.; Nguyen, X.Q. Robot Motion Planning in an Unknown Environment with Danger Space. *Electronics* 2019, 8, 201. https://doi.org/10.3390/electronics8020201.

5. Tuba, E.; Strumberger, I.; Zivkovic, D.; Bacanin, N.; Tuba, M. Mobile Robot Path Planning by Improved Brain Storm Optimization Algorithm. In Proceedings of the 2018 IEEE Congress on Evolutionary Computation (CEC), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–8. https://doi.org/10.1109/CEC.2018.8477928.

6. Bandala, M.; West, C.; Monk, S.; Montazeri, A.; Taylor, C.J. Vision-Based Assisted Tele-Operation of a Dual-Arm Hydraulically Actuated Robot for Pipe Cutting and Grasping in Nuclear Environments. *Robotics* 2019, 8, 42. https://doi.org/10.3390/robotics8020042.

7. Ichter, B.; Pavone, M. Robot Motion Planning in Learned Latent Spaces. *IEEE Robot. Autom. Lett.* 2019, 4, 2407–2414. https://doi.org/10.1109/LRA.2019.2901898.

8. Arm, P.; Zenkl, R.; Barton, P.; Beglinger, L.; Dietsche, A.; Ferrazzini, L.; Hutter, M. SpaceBok: A Dynamic Legged Robot for Space Exploration. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 6288–6294. https://doi.org/10.1109/ICRA.2019.8794136.

9. Grigore, L.S.; Oncioiu, I.; Priescu, I.; Joita, D. Development and Evaluation of the Traction Characteristics of a Crawler EOD Robot. *Appl. Sci.* 2021, 11, 3757. https://doi.org/10.3390/app11093757.

10. Jiang, J.; Luo, X.; Xu, S.; Luo, Q.; Li, M. Hand-Eye Calibration of EOD Robot by Solving the \( AXB = YCZD \) Problem. *IEEE Access* 2022, 10, 3415–3429. https://doi.org/10.1109/ACCESS.2021.3136850.

11. Postigo-Malaga, M.; Supo-Colquehuaquena, E.; Matta-Hernandez, J.; Pari, L.; Mayhua-Lopez, E. Vehicle location system and monitoring as a tool for citizen safety using wireless sensor network. In Proceedings of the 2016 IEEE ANDESCON, Arequipa, Peru, 19–21 October 2016, pp. 1–4. https://doi.org/10.1109/ANDESCON.2016.7836272.

12. Vidal, Y.S.; Supo, C.E.; Cacllata, C.M.; Mamani, G.J.; Pino, C.B.; Pinto, P.L.; Espinoza, E.S. Analysis and Evaluation of a EOD Robot Prototype. In Proceedings of the 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTROCNICS), Toronto, ON, Canada, 1–4 June 2022; pp. 1–6. https://doi.org/10.1109/IEMTROCNICS55184.2022.9795740.

13. Song, X.; Yang, Y.; Choromanski, K.; Caluwaerts, K.; Gao, W.; Finn, C.; Tan, J. Rapidly Adaptable Legged Robots via Evolutionary Meta-Learning. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25–29 October 2020; pp. 3769–3776. https://doi.org/10.48550/arXiv.2003.01239.

14. Paxton, C.; Ratliff, N.; Eppner, C.; Fox, D. Representing Robot Task Plans as Robust Logical-Dynamical Systems. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 4–8 November 2019; pp. 5588–5595. https://doi.org/10.1109/ROSTROCNICS.2019.8867861.

15. Cio, Y.S.L.K.; Raison, M.; Ménard, C.L.; Achiche, S. Proof of Concept of an Assistive Robotic Arm Control Using Artificial Stereovision and Eye-Tracking. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2019, 27, 2344–2352. https://doi.org/10.1109/TNSRE.2019.2950619.

16. Rafael Verano, M.; Jose Caceres, S.; Abel Arenas, H.; Andres Montoya, A.; Joseph Guevara, M.; Jarelh Galdos, B.; Jesus Talavera, D., Suthakorn, J., Williams, M.A., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; Volume 1498. https://doi.org/10.1007/978-3-030-90176-9_37

17. Vilcapaza Goyzueta, D.; Guevara Mamani, J.; Sulla Espinoza, E.; Supo Colquehuanca, E.; Silva Vidal, Y.; Pinto, P.P. Evaluation of a NUI Interface for an Explosives Deactivator Robotic Arm to Improve the User Experience. In *HCI International 2021—Late Breaking Posters. HCII 2021*, Stephanidis, C., Antona, M., Ntoa, S., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2021; Volume 1498. https://doi.org/10.1007/978-3-030-91076-9_37.

18. Nadarajah, S.; Sundaraj, K. A survey on team strategies in robot soccer: Team strategies and role description. *Artif. Intell. Rev.* 2013, 40, 271–304. https://doi.org/10.1007/s10462-011-9284-0.

19. Zhao, J.; Han, T.; Ma, X.; Ma, W.; Liu, C.; Li, J.; Liu, Y. Research on Kinematics Analysis and Trajectory Planning of Novel EOD Manipulator. *Appl. Sci.* 2021, 11, 9438. https://doi.org/10.3390/app1109438.

20. Du, Y.C.; Taryudi, T.; Tsai, C.T.; Wang, M.S. Eye-to-hand robotic tracking and grabbing based on binocular vision. *Microsyst. Technol.* 2021, 27, 1699–1710. https://doi.org/10.1007/s00542-019-04475-3.

21. Wang, M.S. Eye to hand calibration using ANFIS for stereovision-based object manipulation system. *Microsyst. Technol.* 2018, 24, 305–317. https://doi.org/10.1007/s00542-017-3315-y.

22. Esteves, J.S.; Carvalho, A.; Couto, C. Generalized geometric triangulation algorithm for mobile robot absolute self-localization. In Proceedings of the 2003 IEEE International Symposium on Industrial Electronics (Cat. No.03TH8692), Rio de Janeiro, Brazil, 9–11 June 2003; Volume 1; pp. 346–351. https://doi.org/10.1109/ISIE.2003.126722.

23. Dune, C.; Leroux, C.; March, E. Intuitive human interaction with an arm robot for severely handicapped people—A One Click Approach. In Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics, Noordwijk, The Netherlands, 13–15 June 2007; pp. 582–589. https://doi.org/10.1109/ICORR.2007.4428484.
24. Kim, D.; Lovelett, R.; Behal, A. An empirical study with simulated ADL tasks using a vision-guided assistive robot arm. In Proceedings of the 2009 IEEE International Conference on Rehabilitation Robotics, Kyoto, Japan, 23–26 June 2009; pp. 504–509. https://doi.org/10.1109/ICORR.2009.5209527.

25. Goyzueta, D.V.; Guevara, M.J.; Montoya, A.A.; Sulla, E.E.; Lester, S.Y. Analysis of a user interface based on multimodal interaction to control a robotic arm for EOD applications. Electronics 2022, 11, 1690. https://doi.org/10.3390/electronics1111690.

26. Bradski, G.R. Computer vision face tracking for use in a perceptual user interface. Intel Technol. J. 1998. https://doi.org/10.1109/11.775417.

27. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in Psychology; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

28. Bangor, A.; Kortum, P.T.; Miller, J.T. An empirical evaluation of the system usability scale. Int. J. Hum.-Comput. Interact. 2008, 24, 574–594. https://doi.org/10.1080/10447310802205776.

29. Montoya, A.; Pari, L.; Elvis, S.C. Design of a User Interface to Estimate Distance of Moving Explosive Devices with Stereo Cameras. In Proceedings of the 2021 6th International Conference on Image, Vision and Computing (ICIVC), Qingdao, China, 23–25 July 2021; pp. 362–366. https://doi.org/10.1109/ICIVC52351.2021.9526934.

30. Cheng, Y. Mean shift, mode seeking, and clustering. IEEE Trans. Pattern Anal. Mach. Intell. 1995, 17, 790–799. https://doi.org/10.1109/34.400568.

31. Murray, D.; Jennings, C. Stereo vision based mapping and navigation for mobile robots. In Proceedings of International Conference on Robotics and Automation, Albuquerque, NM, USA, 20–25 April 1997; Volume 2, pp. 1694–1699 https://doi.org/10.1109/ROBOT.1997.614387.

32. Corke, P. Robotics, Vision and Control, Fundamental Algorithms in Matlab, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2017. https://doi.org/10.1007/978-3-642-37371-8_33.

33. Sooksatra, S.; Kondo, T. CAMSHIFT-Based Algorithm for Multiple Object Tracking; Springer: Berlin/Heidelberg, Germany, 2013. https://doi.org/10.1007/978-3-642-37371-8_33.

34. Yu, Y.; Bi, S.; Mo, Y.; Qiu, W. Real-time gesture recognition system based on Camshift algorithm and Haar-like feature. In Proceedings of the 2016 IEEE International Conference on Cyber Technology in Automation, CONTROL, and Intelligent Systems (CYBER), Chengdu, China, 19–22 June 2016; pp. 337–342. https://doi.org/10.1109/CYBER.2016.7574846.

35. Xiu, C.; Ba, F. Target Tracking Based on the Improved Camshift Method; IEEE: Piscataway, NJ, USA, 2016.

36. Corke, P.I. A Simple and Systematic Approach to Assigning Denavit–Hartenberg Parameters. IEEE Trans. Robot. 2007, 23, 590–594. https://doi.org/10.1109/TRO.2007.896765.

37. Trucco, E.; Verri, A. Introductory Techniques for 3-D Computer Vision; Prentice Hall: Hoboken, NJ, USA, 1998.

38. Fetić, A.; Jurić, D.; Osmanović, D. The procedure of a camera calibration using Camera Calibration Toolbox for MATLAB. In Proceedings of the 2012 Proceedings of the 35th International Convention MIPRO, Opatija, Croatia, 21–25 May 2012; pp. 1752–1757.

39. Lee, R.; Wu, T.e.; Guo, J. An Adaptive Cross-Window stereo camera Distance Estimation technology and its system implementation for multiple applications. In Proceedings of the 2017 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), Hsinchu, Taiwan, 24–27 April 2017; pp. 1–4. https://doi.org/0.1109/VLSI-DAT.2017.7939669.

40. Zhang, J.; Chen, J.; Lin, Q.; Cheng, L. Moving Object Distance Estimation Method Based on Target Extraction with a Stereo Camera. In Proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), Xiamen, China, 5–7 July 2019; pp. 572–577. https://doi.org/10.1109/CIVC47709.2019.8980940.

41. Guevara, J.; Pari, P.; Vilcapaza, D.; Supo, E.; Sulla, E.; Silva, Y. Compilation and Analysis of Requirements for the Design of an Explosive Ordnance Disposal Robot Prototype Applied in UDEX-Arequipa. In Proceedings of the HCI International 2021 23rd International Conference on Human-Computer Interaction, Virtual, 24–29 July 2021. https://doi.org/10.1007/978-3-030-78642-7_18.