



# Construction of a Computer Vision Test Platform: VISART for Facial Recognition in Social Robotics

Edwin Rodríguez, Christian Gutiérrez, Cristian Ochoa, Freddy Trávez, Luis Escobar<sup>(✉)</sup>, and David Loza

Department of Energy and Mechanics, Universidad de las Fuerzas Armadas - ESPE,  
170121 Sangolquí, Pichincha, Ecuador

{rsedwin,cagutierrez6,cechoa,fstravez,lfescobar,dcloza}@espe.edu.ec  
<https://decem.espe.edu.ec/>

**Abstract.** Robotics has undoubtedly found its way deeper into every day human tasks up to the point where now they even share workspaces with people. In this context, social robotics has increased its field of action, one important part is the advances in aspects that make up the Human-Machine Interaction. This article reports the advance that an investigation team at Universidad de las Fuerzas Armadas ESPE has done to create a test platform for Computer Vision algorithms. This approach consists of a 3-DOF articulated robotic head with anthropomorphic characteristics, considering the guidelines established by different authors for social robotics, besides it provides affordable, flexible and scalable hardware resources to facilitate the implementation, testing, analysis and development of different stereo artificial vision algorithms. The system, called Visart, is intended to work in more natural situations than independent cameras, therefore, work deeply in Human-Robot Interaction. As the first report of the Visart prototype, some basic tests with the platform were performed, they consist of face detection, facial and expression recognition, object tracking, and object distance calculation. The results obtained were consistent, opening a door for further research focusing on the comparison of more computer vision algorithms to examine their performance in real scenarios and evaluate them more naturally with our prototype for Human-Robot Interaction.

**Keywords:** HRI · Computer Vision · Human-Robot Interaction · Anthropomorphic robot · Social robotics

## 1 Introduction

Since the end of the XX century, robots have found their way into the lives of humans in diverse tasks, being originally intended to partially or completely replace workers in simple repetitive duties [5, 11]. The last decade of global competition to improve and automate processes, technological advancement of robotics has greatly increased and opened up the market for new applications

[12]. Most of these include robots and humans working or interacting together in some other way, requiring a series of new challenges to be solved. Human-Robot Interaction (HRI) corresponds to a field of study dedicated to understanding, designing and evaluating robotic systems used in interaction with humans [9], which generally involves environments where conditions are unpredictable and constantly changing, unlike industrial workspaces where the different factors are under control [2, 20, 21].

Unlike HMI (Human Machine Interaction) and HCI (Human-Computer Interaction), HRI focuses on a more direct and close interaction between the whole robotic system and the user. This forces to improve robotics applications on how they can be useful in human daily life [17, 29]. This is why a complete understanding of how the user interacts with robot for long periods of time is utmost important - even more when humans and robots have interacted since 1940's [7].

Communication with robots has become more complex, resembling relationships between two humans. Consequently, the dialogue process must consider many aspects - one of them being Linguistic Competence that considers the ability to understand each participant [7]. Face to Face communication is a pillar of the dialogue action and it occurs at a time scale in the order of 40 ms, a high level of reliability in facial recognition (primitive sensory perception) by the robot is necessary considering the uncertainty present at this time scale [17, 29].

Charles Darwin acknowledged facial expression as one of the most important means for human beings communication [17, 29]. Accordingly, facial recognition makes up one of the main capabilities for Face to Face communication that a robot can be equipped to understand and adapt to changes in its environment, and thereby improve performance in its interaction, is Computer Vision [2, 14, 21]. Examples of Facial Recognition implementation can be found in prototypes such as: Kismet Robot [32], SCARINO [23], Maki Social Robot, SAM Robot, UPS, TEA Robot and NAR Robot, ESPE [22].

## 2 Social Robotics Guidelines

In recent years of robotics advancement, the areas of interest have changed their focus from the analysis in the robot environment towards the study of social and emotional aspects of interaction with humans [16]. For most people, interaction with robots is easier to carry out when the robot exhibits social behavior in some way and if it has anthropomorphic characteristics [20]. Based on the guidelines proposed by [15] for the development of robots oriented towards interaction with people, the strategies to address the following recommendations for the implementation of the robotic system are proposed as:

### 2.1 Appearance

The appearance of the robot is one of the main features, because it defines the degree of user acceptance, likelihood to interact with the system and, therefore,

the success of the robot in the desired application. [18] determined that for different applications, people prioritize the robot's appearance rather than its functionality, to which the robot's behavior is also taken into account to perform seemingly natural actions. The reason for this is that cognitive expectations that a human shows in front of a machine go hand in hand with its human-like appearance [10, 15]. To avoid venturing into the field of realism, but into the field of credibility, the robot's appearance must also be consistent with the purpose and capabilities of the robot. It may be unnecessary to provide the robot with unmotivated anthropomorphic features such as an excess of facial expressions that seek to create an intelligent social entity instead of a social robot tool [4, 30]. Managing to implement just the right amount of social traits in a machine can help to maintain user expectations closer to those that the robot may be able to fulfill.

## 2.2 Affective Interactions and Empathy

One technique to strengthen relationships between a human and a robot is to demonstrate emotions and pro-social actions. With this approach, the distress and disagreement in the user can be greatly reduced [25]. Since the trajectory of the study of the appearance of a human-like robot is already considerably extensive, in recent years there have been a significant number of studies focused on how to express emotions and other nonverbal behaviors in social robots [15]. An area that is still poorly explored within this field is the ability to identify emotions and empathize with the user [26]. Empathy plays such an important role in human relationships that its implementation in social robots has proven effective enough to improve the user's perception of the robot as well as its interaction over extended periods [27].

## 2.3 Memory and Adaptation

Although advances in intelligence and robotic memory, consequently their direct benefits, are still under constant study, it can be anticipated that coherence in the actions of interaction of the robotic system will be improved [34]. Memory is a skill that the robot must have to establish interaction with different users. Therefore, the robot's behavior will be limited not only to predefined events, but also to learning and experiences previously obtained and stored in itself. Cloud Robotics is a technique to provide robustness in terms of memory and computing in robotic systems [13]. This perspective has benefits as Big Data, Cloud Computing, Collective Robot Learning, and Human Computation [13]. So, the system's scalability is ensured by using processing resources on cloud modules. Even simple references from users that the robot can remember increase the user's feeling of trust [15] because it is a greater step towards a more realistic personality. The analysis of behavior and social interactions has benefited from the advances in the different methods for the automatic measurement and monitoring, computer vision methods being one of the main tools [28]. The principal aspects that artificial vision usually analyzes for recognizing and sensing

the environment are shapes, sizes, location of objects, color, lighting, texture, and composition. A correct implementation of Computer Vision can help the actions and behavior of the robot easily meet the expectations of the user and thus, achieve an interactive environment where the human and robot participate in more natural and intuitive interactions [32]. The reason for this is that, by equipping the robotic system with a Computer Vision tool, it improves its environmental perception capabilities, giving it a more independent personality and functionality in addition to the potential cognitive abilities that can be achieved. Therefore, providing the robot that will be used for social interaction with a computer vision system for recognition and monitoring of objects becomes a necessity to meet the need conferring it with greater autonomy, giving it the power to decide to carry out more than its programmed activities.

### 3 Design and Construction

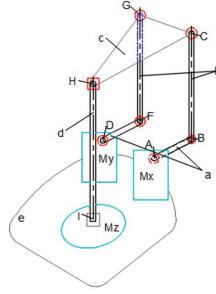
Visart design constraints were based on the social robotics guidelines and the dynamics of its Human-Robot Interaction intents. Through, the Quality Function Deployment matrix (QFD) [31] the most important technical characteristics of the system were obtained and cited as follows:

- Open source platform development
- Use of artificial vision in the final structure of the prototype
- Rapid prototyping manufacturing
- Movement using 3 DOF mechanism
- Compatibility with computer vision algorithms
- Use of HMI for results validation.

#### 3.1 Mechanical Subsystem

Taking the appearance considerations from the social robotics guidelines, the development of the platform will be limited to the construction of a robotic head. This head has to resemble a human in a minimalist way, avoiding the Mori valley. The robot will show its emotions through its eyes to take advantage of human abilities to interpret intentions based on one's actions. Additionally, the robot has to be able to move in order to operate alongside the algorithms for tracking objects, faces, and colors [4]. This is why, a 3-DOF mechanism will be designed for simulating a simplified version of the human neck (flexion, extension, rotation). The mechanical design involves the analysis of the kinematic chain of the 3-DOF mechanism (Fig. 1), according to fatigue design criteria established by [3]. In addition, the selection of actuators  $M_x$ ,  $M_y$  and  $M_z$  (Fig. 1) are also detailed. Through the use of Computer-Aided Design (CAD) and Computer-Aided Engineering (CAE) software, mechanical properties were verified. Tensile, Von Misses stress and deformation values were found through software simulation.

Visart's mechanical design also includes kinematic analysis for the ball type joint, according to Denavit Hartenberg criteria [6]. Kinematic chains were



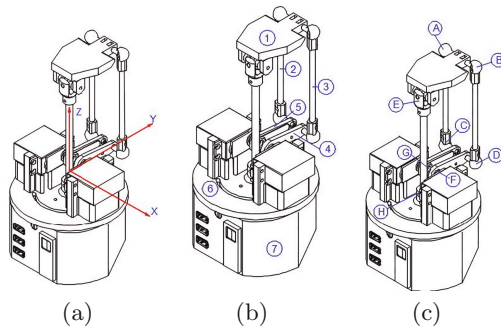
**Fig. 1.** Visart 3-DOF mechanism. General scheme.

described according to Fig. 2, a total of 7 links, 3 rotational joints, 3 universal joints, and 2 spherical joints were defined. The result is a 3-DOF mechanism. The head’s rotational and translational matrix were obtained and shown in Eqs. 3 and 1 respectively. Subsequently, the dynamic analysis of the system was carried out, starting from a kinematic analysis and using the Euler-Lagrange method to obtain the Lagrangian of the system, resulting in Eq. 2.

$$P_c = \begin{bmatrix} -19.5c(\phi) - 10s(\phi) - 40c(q_2)s(\phi) \\ 10c(\phi) - 19.5s(\phi) + 40cos(q_2)c(\phi) \\ 40s(q_2) \end{bmatrix} \tag{1}$$

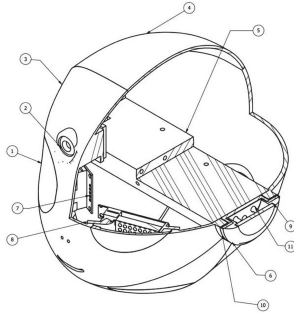
$$\begin{bmatrix} 0.024372 & -0.0806804 \\ -0.0806804 & 0.024372 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_i \\ \ddot{\theta}_d \end{bmatrix} = \begin{bmatrix} \tau_i \\ \tau_d \end{bmatrix} \tag{2}$$

$$R_c = \begin{bmatrix} c(\phi) & -c(q_2)s(\phi) & s(q_2)s(\phi) \\ s(\phi) & c(\phi)c(q_2) & -s(q_2)c(\phi) \\ 0 & s(q_2) & c(q_2) \end{bmatrix} \tag{3}$$



**Fig. 2.** Visart 3-DOF kinematic analysis. Links and joints distribution. (a) Coordinate system. (b) Links (c) Rotational and universal joints.

Since the robotic system will be used for computer vision algorithms, the visual sensory system will be specially highlighted, which will be achieved through the shape and position of the eyes and cameras. The final system structure is shown in Fig. 2 for the neck and Fig. 3 for the head. A total of 32 components are included on Visart’s mechanical design. Most elements, including body and face, were manufactured by additive manufacturing using FDM (Fused Deposition Modeling) with PLA (Polylactic acid or polylactide) as prime material.



**Fig. 3.** Visart’s head assembly. 1: Glasses, 2: Camera, 3: Front Head Case, 4: Back Head Case, 5: Head Bracket, 6: Ear Base, 7: Led Matrix, 8: Ear Protection, 10: Ears, 11: Ear LED ring.

### 3.2 Electronic Subsystem

In order to show empathy, one of the most effective tools are the robot’s eyes [20]. With this consideration, Visart was built with an  $8 \times 8$  LED matrix as an interactive component of the robot to show its emotions and transmit expressions through its gaze. This simple method combined with the head’s movement and face tracking capabilities take advantage of human abilities to interpret intentions based on one’s actions [4]. Additionally, the robotic platform has a modular design using a minicomputer to control the prototype, this computer can communicate through internet for later developments.

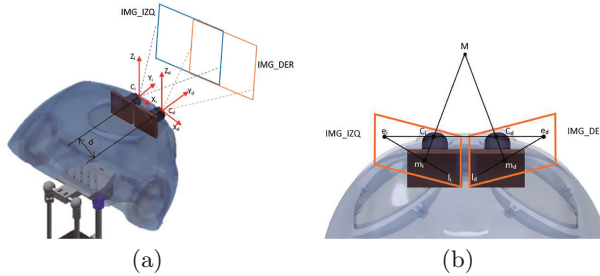
### 3.3 Computer Vision Subsystem

Computer Vision subsystem focuses on the robot’s ability to extract visual information from the physical space. The algorithm for image processing was programmed with Python language alongside OpenCv (Open Source Computer Vision Library) module, on Ubuntu 16.04 OS. Following the stereo vision guidelines two parallel web cams located on the top of the head were used (Fig. 3).

For the mathematical modeling of stereo vision camera arrangement, the cameras were considered parallel to  $b$  distance (Fig. 4: a). The origin of the

coordinate system was set on the left camera. Using homogeneous matrices, any point projection within the robot's visual field can be calculated. For a projected point  $M$  (Fig. 4: b) on the stereo system, at a focal length  $f$ , according to the pin-hole model, the matrices shown in Eq. 4 were obtained.

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}_{left} = \begin{bmatrix} fk_x & 0 & fC_x & hfC_x \\ 0 & 1 & 0 & 0 \\ 0 & k_z & fC_z & hfC_z \end{bmatrix} \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} \quad (4)$$



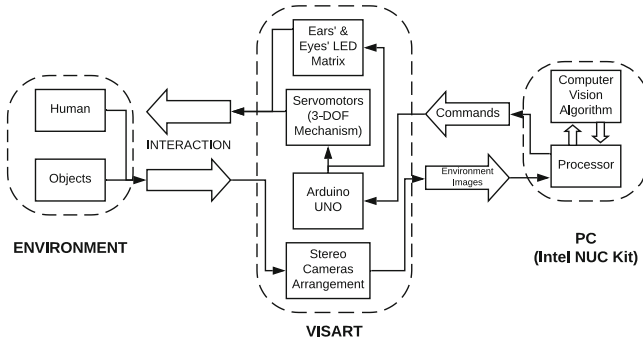
**Fig. 4.** Visart stereo vision field space for mathematical modeling.

## 4 Validation Methodology

Once all Visart robot subsystems have been implemented, the test platform is set-up as indicated in the block diagram shown in Fig. 5. The computer vision algorithms will be loaded on the Intel NUC Personal Computer kit, equipped with an Intel Core i5-4250U processor with a 2.6 GHz speed, integrated with an Intel HD Graphics 5000 2GB graphics processor and 6 GB of RAM. The 2 cameras of the robot will send the captured images directly to the computer for processing. After the analysis, the program will send the respective movement and expression commands to the control board. This platform will be in charge of writing the commands received by the computer in the Visart actuators.

With the complete platform set-up, the cameras are first calibrated. For this task, the ChessBoard calibration method will be used to align the lines captured in the images between both cameras and with the real world. The algorithm code is based on the work of [1]. Once the image calibration of both cameras is achieved, the image is trimmed to repair the resulting distortion, so that the actual coordinates of the detected objects can be obtained.

For platform validation tests, a test protocol is established to verify the operation and capabilities of the system to use computer vision algorithms. Based on [8, 33] consideration of ISO/IEC 19794-5:2005, the following protocol also takes



**Fig. 5.** Visart function block diagram

into account the parameters and formats of the scenes, properties, attributes, and best practices to obtain good images and video of faces for analysis in biometric applications. The established protocol was then developed in 4 stages detailed in Fig. 6:

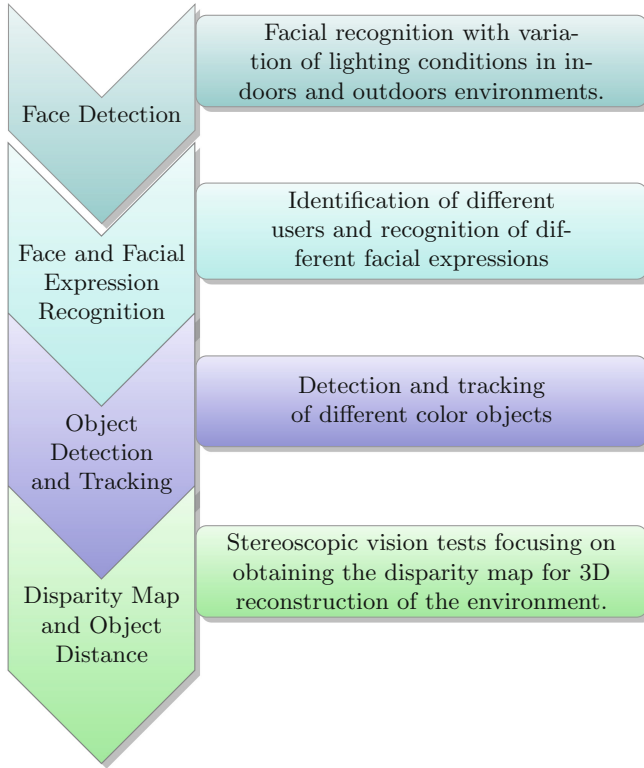
In the first 2 stages, the platform tests were performed with 5 subjects to analyze functionality. In the first stage of Face Detection, images were taken while varying the human robot distance, as well as the capture planes (front, right and left side) and light incidence in order to check the ability to recognize human faces in different ambient. Fluorescent lights were used to control the surrounding light in indoors. For the registration process of a user, 15 images of his face in different orientations are stored. Likewise, the facial expression recognition algorithm -limited to joy, surprise, sadness, and anger- was trained with a database of 100 images per expression using the Python Tensorflow tool. The tests were performed with all the individuals in the same scene.

For the last two test stages, objects of different colors (green, red, blue and yellow) were used. The objective was to recognize and track the object along Visart visual field. For this task  $6 \times 6$  cm, 3D-printed PLA cards were used. Finally, in order to analyze the bifocal setup of computer vision of the robot, Kolmogorov and Zabih-GC's graph-cutting algorithm [24] was tested for the generation of disparity maps, showing a grayscale pattern in several planes of the image allowing 3D reconstruction of its environment. The test consisted of an object placed at different distances from the platform (5, 10, 20, 40, 80, 120, 160, 200 cm) and then estimated through the computer vision algorithm.

## 5 Results

Indoor tests were carried out in a controlled environment. A controlled beam of light directed to the user's face was emitted by fluorescent lamps. Under these conditions, all the tests detailed in Sect. 4 were performed. Outdoor tests, on the other hand, brought out unpredictable outcomes due to the environmental conditions. Each stage test resulted as follows:





**Fig. 6.** Validation methodology stages.

### 5.1 Face Detection

Figure 7 shows the data obtained by varying the distance from the user to Visart, and trying several face positions. The values shown below correspond to the means of the experiments. The data were taken from 6 ranges, starting at 80 cm up to 350 cm with a 54 cm increase. For frontal face plane faces are detected 80% of the times, while right face planes 56.67% and left face plane 53.33% of the times. For distances greater than 296 cm the detection percentage is zero.

### 5.2 Face and Facial Expression Recognition

The experiment started with the 15 images database for each user, and was performed in a dark room at 100% illumination -with the fluorescent lamp pointed towards the user- and 0% illumination -without the fluorescent lamp-. In addition, recognition for several subjects in the same image frame was tested. In all cases for registered users, the system could recognize them 100% of the times with 100% illumination at close range. As a result, the system was able to recognize expressions such as happiness, sadness, and surprise.

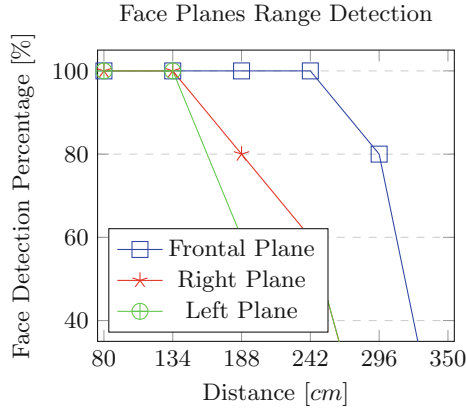


Fig. 7. Face planes range detection.

### 5.3 Object Detection and Tracking

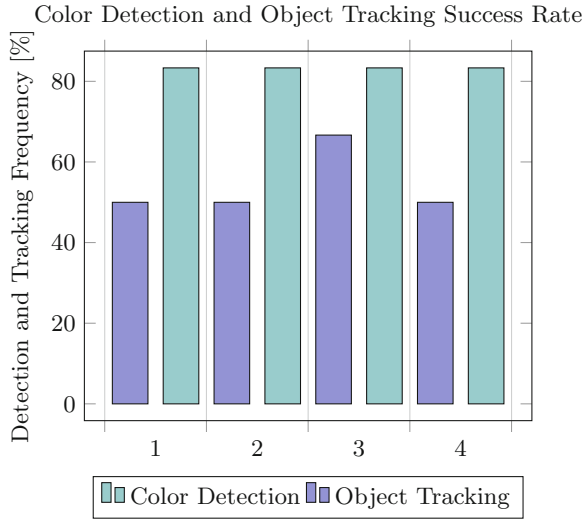
Figure 8 shows the portion of events in which the system was able to detect the test figures, as well as the frequency when the algorithm could successfully track the colored cards. The tests were performed on distances from 80 to 350 cm, with a total of 6 tests for each color. Consequently, on 83.33% of the events the system could detect any color, and on 54.16% it could successfully achieve object tracking. The blue testing probe is the color with a higher tracking accuracy (66.66%).

### 5.4 Disparity Map and Object Distance

Through the disparity map generation algorithm, a grayscale-image was obtained. This allowed 3D reconstruction of the robot environment though using both camera images. Therefore, it was possible to compute the estimated distance from any object to the robot and then calculate measurement errors. Deviation values are shown in Fig. 9. The best results were obtained from distances from 20 to 120 cm with an error lesser than 5%.

## 6 Discussion

The results obtained show that facial recognition is effective on face to face interaction among the user and Visart. However, the platform face recognition ability is greatly reduced when Visart captures only right or left face profile. This issue may be due to the vision algorithm used or the image quality that is influenced by the environmental brightness or cameras' resolution. Besides, Visart can successfully recognize 3 basic facial expressions [29] used during people interaction. As well as the ability to track color testing probes. In the latest instance, the repetitiveness is determined by the corresponding color. Experiments with the



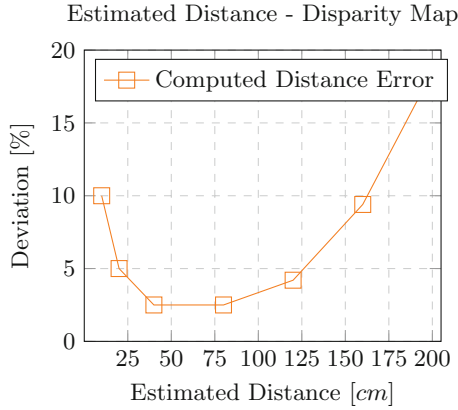
**Fig. 8.** Visart success rate for color detection and object tracking. Column 1: Green card, 2: Red card, 3: Blue card and, 4: Yellow card. (Color figure online)

blue testing probe got the highest tracking accuracy rate. This may be related to the probe color opacity and the type of material used to manufacture the cards and the 3D, including the printing parameters such as infill and layer height. The contrast of the color and the environment is also a concern considering that the color detection algorithm interprets the information as grayscale.

## 7 Future Work

A robust validation process is necessary, in order to obtain a greater amount of data to identify the platform behavior. For this purpose works like [23] and [19] will be taken as a reference. In order to improve the platform's functionality, tests in more diverse environments are planned to be done including the integration an analysis of several computer vision algorithms. This will lead to making the best choice for every test stage before taking the platform into real life scenarios.

Since educational robotics has also advanced at the same pace as other areas of social robotics [30], more tools such as speech recognition can be implemented into VISART to fulfill different functions that involve interaction with young students either as a tutor or class assistant. Because of this, it is important to carry out a validation process of social interaction with students of different ages and identify the strengths and weaknesses of the platform. This will also promote the use of new educational tools and methodologies that encourage curiosity and technology development by students; and even the identification of possible obstacles and risks in the use of robots for teaching.



**Fig. 9.** Distance estimation using Stereo Vision.

Another great area to develop is the use of VISART as a interaction tool for management of Cyber Physical Systems in elderly care, home automation and as office assistant.

## 8 Conclusion

The Visart platform was proven useful for face detection in any of the planes up to 134 cm - between user and robot. However, for longer distances only frontal plane facial recognition is guaranteed. Considering that 242 cm is the maximum reliable detection distance, the system validates itself for use in social interaction as most social actions take place within the 100 cm range. In this length, facial expressions such as happiness, sadness and surprise could also be recognized successfully.

Object tracking capabilities of the system are significantly encouraging with its 80% accuracy in recognition obtained. The object in some cases could not be tracked due to several factors such as the reflection index of the test cards, the different light conditions consequence of the robot’s movement, the image processing algorithm and even the cameras used.

The platform is based on open-source software and has the capability for local data processing. These features provide the operating conditions to mitigate network connection problems - such as low-latency responses, quality of service and downtime.

All these regards contribute to Visart being a platform that brings together the necessary characteristics in both appearance and functionality, that ensure an effective HRI in social robotics [9, 15, 23]. Therefore, it is possible to take the next step towards testing its effectiveness with more real people interactions against the use of various algorithms for face detection and facial recognition, object detection and monitoring, and proximity estimation in the interaction environment.

## References

1. Armea, A.: Calculating a depth map from a stereo camera with OpenCV (2017). <https://albertarmeacom/post/opencv-stereo-camera/>
2. Azevedo, H., Belo, J.P.R., Romero, R.A.F.: Cognitive and robotic systems: speeding up integration and results. In: 2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), Curitiba, pp. 1–6. IEEE (2017). <https://doi.org/10.1109/SBR-LARS-R.2017.8215337>
3. Budynas, R.G., Keith, N.: Diseño en ingeniería mecánica de Shigley, 8th edn. McGrawHill, Mexico (2008)
4. Duffy, B.R.: Anthropomorphism and the social robot. *Robot. Auton. Syst.* **42**(3–4), 177–190 (2003). [https://doi.org/10.1016/S0921-8890\(02\)00374-3](https://doi.org/10.1016/S0921-8890(02)00374-3). <https://linkinghub.elsevier.com/retrieve/pii/S0921889002003743>
5. Faust, R.A.: Robotics in Surgery: History, Current and Future Applications. Nova Publishers, New York (2007). Google-Books-ID: p70afWYqcrMC
6. Fernández, L., Sotomayor, L.: Análisis cinemático inverso y directo del robot paralelo (2016)
7. Fong, T., Thorpe, C., Baur, C.: Collaboration, dialogue, human-robot interaction. In: Jarvis, R.A., Zelinsky, A. (eds.) *Robotics Research*. Springer Tracts in Advanced Robotics, vol. 6, p. 10. Springer, Heidelberg (2003). [https://doi.org/10.1007/3-540-36460-9\\_17](https://doi.org/10.1007/3-540-36460-9_17)
8. Gomez Camacho, M., Barrera, Y.V.: Prototipo de un sistema biométrico integrado para la identificación de personas - fase I (2016). <https://doi.org/10.13140/rg.2.1.3267.2881>
9. Goodrich, M.A., Schultz, A.C.: Human-robot interaction: a survey. *Found. Trends Hum. Comput. Interact.* **1**(3), 203–275 (2007). <https://doi.org/10.1561/1100000005>
10. Hayashi, K., Shiomi, M., Kanda, T., Hagita, N.: Who is appropriate? A robot, human and mascot perform three troublesome tasks. In: 19th International Symposium in Robot and Human Interactive Communication, Viareggio, Italy, pp. 348–354. IEEE (2010). <https://doi.org/10.1109/ROMAN.2010.5598661>
11. Hockstein, N.G., Gourin, C.G., Faust, R.A., Terris, D.J.: A history of robots: from science fiction to surgical robotics. *J. Robot. Surg.* **1**(2), 113–118 (2007). <https://doi.org/10.1007/s11701-007-0021-2>
12. IFR: World Robotics Survey: Industrial robots are conquering the world (2015). <https://ifr.org/news/world-robotics-survey-industrial-robots-are-conquering-the-world-/>
13. Kehoe, B., Patil, S., Abbeel, P., Goldberg, K.: A survey of research on cloud robotics and automation. *IEEE Trans. Autom. Sci. Eng.* **12**(2), 398–409 (2015). <https://doi.org/10.1109/TASE.2014.2376492>
14. Krishnan, A., Hs, A.: Design and implementation of robotic vision for face recognition. In: 2017 International Conference on Circuits, Controls, and Communications (CCUBE), Bangalore, pp. 135–138. IEEE (2017). <https://doi.org/10.1109/CCUBE.2017.8394148>
15. Leite, I., Martinho, C., Paiva, A.: Social robots for long-term interaction: a survey. *Int. J. Social Robot.* **5**(2), 291–308 (2013). <https://doi.org/10.1007/s12369-013-0178-y>

16. Lindblom, J., Andreasson, R.: Current challenges for UX evaluation of human-robot interaction. In: Schlick, C., Trzcieliński, S. (eds.) *Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future. Advances in Intelligent Systems and Computing*, vol. 490, pp. 267–277. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-41697-7\\_24](https://doi.org/10.1007/978-3-319-41697-7_24)
17. Littlewort, G., et al.: Towards social robots: automatic evaluation of human-robot interaction by facial expression classification. In: Thrun, S., Saul, L.K., Schölkopf, B. (eds.) *Advances in Neural Information Processing Systems 16*, pp. 1563–1570. MIT Press, Cambridge (2004)
18. Lohse, M., Hegel, F., Wrede, B.: Domestic applications for social robots: an online survey on the influence of appearance and capabilities. *J. Phys. Agents (JoPha)* **2**(2), 21–32 (2008). <https://doi.org/10.14198/JoPha.2008.2.2.04>
19. Martinez, B., Valstar, M.F., Jiang, B., Pantic, M.: Automatic analysis of facial actions: a survey. *IEEE Trans. Affect. Comput.* **10**(3), 325–347 (2019). <https://doi.org/10.1109/TAFFC.2017.2731763>. <https://ieeexplore.ieee.org/document/7990582/>
20. McGinn, C.: Why do robots need a head? The role of social interfaces on service robots. *Int. J. Social Robot.* (2019). <https://doi.org/10.1007/s12369-019-00564-5>
21. Nagai, K., Sakabe, H., Ohka, M.: Finger direction recognition toward human-and-robot cooperative tasks. In: *2017 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, Nagoya, pp. 1–3. IEEE (2017). <https://doi.org/10.1109/MHS.2017.8305242>
22. Ochoa, C., Trávez, F.: *Diseno y construcción de una plataforma robótica para pruebas técnicas sobre visión artificial que permita el aprendizaje y desarrollo del software en robots interactivos*. Ph.D. thesis, Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador (2018)
23. Pinillos, R., Marcos, S., Feliz, R., Zalama, E., Gómez-García-Bermejo, J.: Long-term assessment of a service robot in a hotel environment. *Robot. Auton. Syst.* **79**, 40–57 (2016). <https://doi.org/10.1016/j.robot.2016.01.014>. <https://linkinghub.elsevier.com/retrieve/pii/S0921889015300440>
24. Pinto Sánchez-Matamoros, L.F., Carrasco, J.A.P.: Análisis de la aplicación de algoritmos de K-means y Continuous Max-Flow a la segmentación de imágenes en color. *Dep. Teoría de la Señal y Comunicaciones Escuela Técnica Superior de Ingeniería Universidad de Sevilla*, p. 80 (2015)
25. Preston, S.D., de Waal, F.B.M.: Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* **25**(1), 1–20 (2002). <https://doi.org/10.1017/S0140525X02000018>
26. Reyes Fernández, B.: Empathy and moral development: implications for caring and justice. *Actualidades en Psicología* **20**(107), 141 (2011). <https://doi.org/10.15517/ap.v20i107.40>
27. Riek, L.D., Robinson, P.: When my robot smiles at me: enabling human-robot rapport via real-time head gesture mimicry. *J. Multimodal User Interfaces* **3**(1–2), 99–108 (2010)
28. Robie, A.A., Seagraves, K.M., Egnor, S.E.R., Branson, K.: Machine vision methods for analyzing social interactions. *J. Exp. Biol.* **220**(1), 25–34 (2017). <https://doi.org/10.1242/jeb.142281>
29. Stewart, M., Gwen, B., Ian, L., Javier, F., Movellan, R.: *Real time face detection and facial expression recognition: development and applications to human computer interaction*, Madison, Wisconsin, USA, p. 53. IEEE (2003)
30. Tapus, A., Maja, M., Scassellatti, B.: The grand challenges in socially assistive robotics. *IEEE Robot. Autom. Mag.* **14**(1) (2007)

31. Tontini, G.: Integrating the Kano model and QFD for designing new products. *Total Qual. Manag. Bus. Excell.* **18**(6), 599–612 (2007). <https://doi.org/10.1080/14783360701349351>
32. Wiese, E., Metta, G., Wykowska, A.: Robots as intentional agents: using neuroscientific methods to make robots appear more social. *Front. Psychol.* **8**, 1663 (2017). <https://doi.org/10.3389/fpsyg.2017.01663>. <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.01663/full>
33. Yi, D., Liu, R., Chu, R.F., Lei, Z., Li, S.Z.: Face matching between near infrared and visible light images. In: Lee, S.-W., Li, S.Z. (eds.) *ICB 2007*. LNCS, vol. 4642, pp. 523–530. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-74549-5\\_55](https://doi.org/10.1007/978-3-540-74549-5_55)
34. Yii Lim, M., Aylett, R.: Human-like memory retrieval mechanisms for social companions. In: *AAMAS*, Taipei (2011)