NAVIGATION OF QUADRUPED MULTIROBOTS BY GESTURE RECOGNITION USING RESTRICTED BOLTZMANN MACHINES

ARATA A. SARAIVA[1]

UTAD University,Coimbra, Portugal
ORCID: 0000-0002-3960-697X [1] E-mail: aratasaraiva@gmail.com

SANTOS, D. B. S[2] and MARQUES JUNIOR, F.C.F [3]

State University of Piaui,
Piripiri, Piaui, Bazil
[2] *ORCID: 0000-0003-4018-242X [2] E-mail: domingosbruno@prp.uespi.br*
ORCID: 0000-0002-2751-0137 [3] E-mail: dev.jrmax@prp.uespi.br

JOSE VIGNO M SOUSA[3]

State University of Piaui, Piripiri,Brazil
University Brazil, Sao Paulo, Brazil
ORCID: 0000-0002-5164-360X [3] E-mail: vigno@prp.uespi.br

N. M. FONSECA FERREIRA[5]

Department of Electrical Engineering, Institute of Engineering of Coimbra,
Coimbra, Polytechnic Institute
Rua Pedro Nunes, 3031-601 Coimbra, Portugal
Knowledge Engineering and Decision-Support Research Center (GECAD) of the
Institute of Engineering, Polytechnic Institute of Porto
INESC-TEC Technology and Science, Campus da FEUP, Rua Dr. Roberto Frias, 378,
4200 - 465 Porto, Portugal
ORCID: 0000-0002-2204-6339 [5] E-mail: nunomig@isec.pt

ANTONIO VALENTE[6]

INESC-TEC Technology and Science
Campus da FEUP, Rua Dr. Roberto Frias, 378, 4200 - 465 Porto, Portugal
School of Science and Technology, University of Trás-os-Montes and Alto Douro,
Vila Real, Portugal
ORCID: 0000-0002-5798-1298 [6] E-mail: avalente@utad.pt

This article discusses a method that performs gesture recognition, with the objective of extracting characteristics of the segmented hand, from dynamic

images captured from a webcam and identifying signal patterns. With this method it is possible to manipulate simulated multirobots that perform specific movements. The method consists of the Continuously Adaptive Mean-SHIFT algorithm, followed by the Threshold segmentation algorithm and Deep Learning through Boltzmann restricted machines. As a result, an accuracy of 82.2%.

1. Introduction

A line of research and development in robotics that has received attention in recent years is the development of bio-inspired walking robots[1].[2]

Systems with artificial intelligence are applied in pattern recognition and can be coupled to manipulable robots, being used in industrial automation, patrols and territorial surveillance,[3] have resulted in systems that can operate on structured and unstructured media,[4] through the use of advanced sensorial feedback mechanisms, making decisions based on learning algorithms and reasoning.

Based on this has inspired the design of manipulation of 4 legs spider multirobots Figure 4, quadruped simulated, by means of manual gestures. The Objects of interest that need to be detected and reviewed are coordinated as being processed in real time. The method chosen and implemented is the capture and detection of regions of interest in the scene. With a preprocessing for noises reduction in the image acquisition process, the Threshold segmentation algorithm was used. And in the last step is performed the real-time classification by an artificial intelligence, which consists of a deep learning known as Restricted Boltzmann Machines (RBM).

The method covered ensures robust coverage in image recognition,[5] under certain assumptions that will be clarified throughout the text in robotic simulation. The developed structure allows the robots to maintain or alter the formation of the specified trajectories.

The document is divided into 6 sections, in which section 2 is characterized by the model of simulated robots used, section 3 consists of kinematics, follows the formulation of the central algorithm applied in section 4 and the statistical method to verify the reliability of the used system . The results after application of the proposal are presented in section 5 and the conclusion in section 6.

2. Quadruped Robot

Locomotion of legs on natural terrain presents complex problems such as foot placement, avoiding obstacles, load distribution, general stability[6].[7]

One way to deal with these problems is to use models that describe mathematically the different situations.[8] The model adopted has 4 identical legs, as can be seen in the Figure1.

Figure 1.   Quadruped Model

3. Kinematic

Kinematics is the branch of mechanics that studies the movement of a body or a system of bodies without considering its mass or the forces acting on it.[9]Kinematics has two types, direct and inverse kinematics. In the direct kinematics it is desired to obtain the position and orientation of the actuator in the Cartesian space of the base, from the positions of the joints[10] .

The inverse kinematics consists of the process of determining the angles in terms of the coordinates for the desired position of the leg in the Cartesian system.[9]

The procedure for obtaining the equations is strongly dependent on the configuration of the robot.[11] The kinematics chosen by the present work is the inverse, in this case, is obtained through geometric definitions based on the shape of the leg Figure2, the equations obtained are 1, 2, 3.
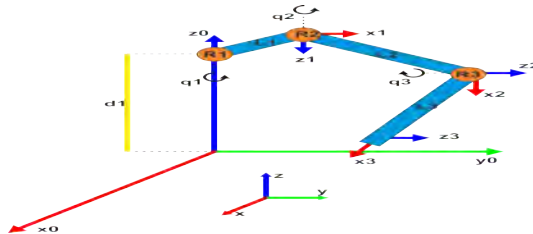

Figure 2.   Kinematic model of the legs.

$$\theta_1 = \arccos\left(\frac{y_1}{x_1}\right) \tag{1}$$

$$\theta_2 = \arccos\left(\frac{L_2^2 + x_3^2 + y_3^2 - L_3^2}{2.L_2.\sqrt{x_2^3 + y_3^2}}\right) + \arctan\left(\frac{y_3}{y_3}\right) \tag{2}$$

$$\theta_3 = \pi - \arccos\left(\frac{L_2^2 + L_3^2 - \left(x_3^2 + y_3^2\right)}{2.L_2.L_3}\right) \tag{3}$$

Each leg of the Quadruped Robots is composed of a chain of three links of freedom Figure 2.The assignment model of the link frame of the links is based on Denavit-Hartenberg direct geometric modeling algorithm.

The Figure 2 illustrates the leg used for the Robots, in which it has three joints R1, R2 and R3 respectively. R1 is the point at which the leg is attached to the body of the robot, L1 represents the "thigh", L2 is called the "femur" and L3 is the "tibia".

The Table 1 represents the parameters obtained through the application of the Danavit-Hartenberg method.

Table 1. Denavit-Hartenberg parameters.

Link	ai	αi	di	θi
1	$L1$	$\pi/2$	0	$\theta 1$
2	$L1$	0	0	$\theta 2$
3	$L1$	0	0	$\theta 3$

4. Methodology

In this section we will discuss the structure of the adopted system, the manual gestures chosen, the structure of the algorithm and finally the method of the classifier analysis.

The structure of the system corresponds to a sequence of steps that will allow the tracking of gestures in real time. Each step of recognition begins with algorithms to establish a region of interest, segmenter to reduce noise and improve classifier processing and lastly the use of the classifier. From the recognized gesture the robots will execute a certain movement. The operation of the used algorithm can be analyzed in the Figure 3.
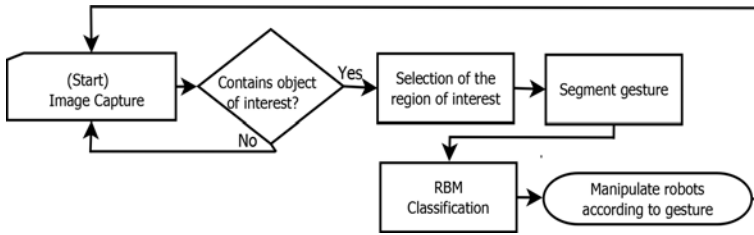
Figure 3. Flowchart of the proposed method

For the real-time tracking of the object of interest was used the continuously adaptive algorithm CamShift which is an adaptation of the algorithm of MeanShift. For each captured frame, the original image is converted to a probabilistic color distribution of the image, which using a histogram model of the color to be traced.[12] In this case, the characteristic color to be tracked is that of the hand, after having recognized this pattern, the desired location is made in the scene, and then the segmentation of this as shown in Figure 4.b.
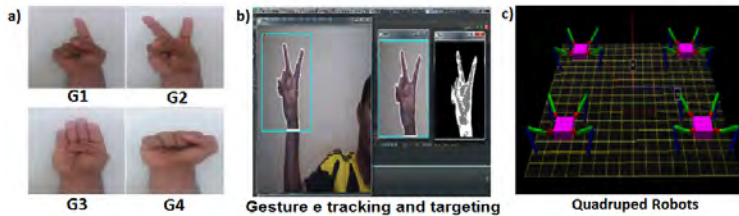

Figure 4. a) Gestures used. b) Tracking and segmentation the object of interest. c)Used robots.

Four gestures were used to control the robots Figure 4.a, each gesture transmits an action to the robots being able to be of two categories, collective or individual. The gesture "G1" conveys a collective action for the robots which consists of the same ones executing a forward movement movement, "G2" just like the previous gesture transmits a collective action causing the robots to return to the initial position, "G3" by its Instead it consists of an individual action that causes the robot from the lower right corner to move to the center, "G4" transmits an individual action that causes the robot from the upper left corner to move up to the scepter.

The next step is segmentation. The algorithm chosen to perform this

stage of the system is Theshold algorithm, in order to reduce the noises in the image to obtain a greater precision in the classifier. This type of algorithm is based on the principle that an object of interest in the scene has distinct quantifiable characteristics, such characteristics as intensity or magnitude of the gradient.[13]

This segmentation process is based on the search for pixels whose values are within the ranges defined as thresholds. From a threshold established with the characteristics of the objects you want to isolate, the image can be segmented into two groups: the group of pixels with gray levels below the threshold and the group of pixels with gray levels above the threshold. If the value of the pixel is greater than a value stipulated as a limit, it is assigned a value, in this case it is white, however otherwise, another value is assigned that is black, the relevance of this method can be seen in Figure 4.b.

The last step is to start the classifier training that will be responsible for performing the signal recognition. The RBM algorithm was used. In addition, RBMs can be used as non-linear classifiers and can achieve better performance compared to conventional neural networks and SVMs.[14]

A RBM is a stochastic neural network composed of two layers,[14] visible and hidden. The visible units layer represents the observed data and is connected to the hidden layer, which in turn must learn to extract characteristics of this data. In a RBM neurons of the same layer are not connected between themselves. Therefore, there is only a connection between neurons of different layers, so the machine is restricted.

For the RBM classifier, 5685 images were used, which are an adaptation of the datasets provided by,[15] 3411 used for training consisting of 60 % of the total and 2274 images for testing consisting of 40 %. As a statistical tool we have the confusion matrix,[16] which provides the basis for describing the accuracy of the classification and characterizing the errors, helping to refine the classification.[17]

5. Results

Satisfactory results were obtained for each gesture evaluated, worst case, "G1", with 61% as shown in Table 2. The overall accuracy of the approach is 82.2%

Table 2.   Confusion matrix

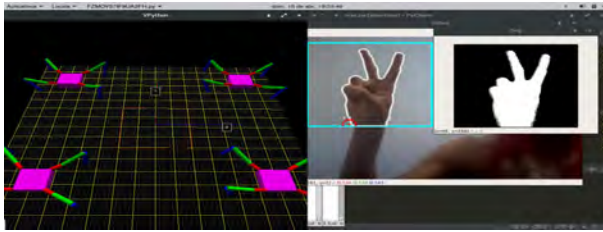| Known	G1	G2	G3	G4
G1	61%	-	-	-
G2	-	95%	-	-
G3	-	-	76%	-
G4	-	-	-	92%

Note: Predicted %

Figure 5.   Manipulating the robots.

In the Figure 5 is displayed to the real-time evaluation of the method used in the project, it can be observed that the gesture used is the "G4", that means that the command is transmitting a collective action, causing the robots to return to their initial position.

6. Conclusion

This article presents a real-time multirobot control system from robust and reliable hand signals. In addition, a good accuracy was obtained in the application of the algorithm chosen to classify the gestures Table 2.

The system presented satisfactory results with regard to real-time video processing, obtaining analysis and classification of the gestures, in order to transmit information to the robots, and with this to be performed actions instantly.

As future work it is suggested to increase the number of gestures, to test with other classifiers and to implement in real robots.

Acknowledgments

Bibliography

1. M. F. Silva, J. A. T. Machado and A. M. Lopes, *Robotica* **23**, p. 595–606 (2005).
2. N. M. Fonseca Ferreira, F. Moita, V. Santos, J. Ferreira, J. Santos, F. Santos and M. Silva, Education with robots inspired in biological systems(04 2018).
3. Y. Girdhar and G. Dudek, *Autonomous Robots* **40**, 1267 (2016).
4. M. Li, M. Su, R. Xie, Y. Zhang, H. Zhu, T. Zhang and Y. Guan, Development of a bio-inspired soft gripper with claws, in *Robotics and Biomimetics (ROBIO), 2017 IEEE International Conference on*, 2017.
5. *transações IEEE em imagens médicas* **35**, 1262.
6. M. Luneckas, T. Luneckas, D. Udris and N. M. Fonseca Ferreira, *Elektronika ir Elektrotechnika* **20**, 7(09 2014).
7. X. Xiong, F. Wörgötter and P. Manoonpong, *IEEE transactions on cybernetics* **46**, 2521 (2016).
8. H. Yu, H. Gao, L. Ding, M. Li, Z. Deng and G. Liu, *IEEE Transactions on Industrial Electronics* **63**, 5488 (2016).
9. P. Corke, *Robotics, Vision and Control: Fundamental Algorithms In MATLAB® Second, Completely Revised* (Springer, 2017).
10. A. Zeiaee, R. Soltani-Zarrin, R. Langari and R. Tafreshi, Design and kinematic analysis of a novel upper limb exoskeleton for rehabilitation of stroke patients, in *Rehabilitation Robotics (ICORR), 2017 International Conference on*, 2017.
11. M. S. Couceiro, N. M. Fonseca Ferreira and J. A. T. M. Dainius, *Journal of Communications in Nonlinear Science and Numerical Simulation* **17**, 893(02 2012).
12. C.-H. Hsia, Y.-J. Liou and J.-S. Chiang, *Journal of Real-Time Image Processing* **12**, 183 (2016).
13. M. Yazdi and M. Fani, Shot boundary detection with effective prediction of transitions' positions and spans by use of classifiers and adaptive thresholds, in *Electrical Engineering (ICEE), 2016 24th Iranian Conference on*, 2016.
14. C. Gou, K. Wang, Y. Yao and Z. Li, *IEEE Transactions on Intelligent Transportation Systems* **17**, 1096 (2016).
15. P. K. Pisharady, P. Vadakkepat and A. P. Loh, *International Journal of Computer Vision* **101**, 403 (2013).
16. X. Deng, Q. Liu, Y. Deng and S. Mahadevan, *Information Sciences* **340**, 250 (2016).
17. A. Saraiva, R. T. Melo, V. Filipe, J. V. M. Sousa, N. M. Fonseca Ferreira and A. Valente, *international journal of systems applications, engineering and development* **12**, 63 (2018).