FORMATION CONTROL BETWEEN A HUMAN AND A MOBILE ROBOT BASED ON STEREO VISION

Flávio Garcia Pereira, Marino Frank Cypriano and Raquel Frizera Vassallo
Department of Electrical Engineering, Federal University of Espírito Santo
Av. Fernando Ferrari, 514, Goiabeiras, Vitória-ES, CEP 29075-910, Brazil

Keywords: Formation control, Stereo vision.

Abstract: This paper presents a reactive nonlinear formation control between a human and a mobile robot based on stereo vision. The human, who is considered the leader of the formation, is positioned behind the robot and moves around the environment while the robot executes movements according to human movements. The robot has a stereo vision system mounted on it and the images acquired by this system are used to detect the human face and estimate its 3D coordinates. These coordinates are used to determine the relative pose between the human and robot, which is sent to the controller in order to allow the robot to achieve and keep a desired formation with the human. Experimental results were performed with a human and a mobile robot, the leader and follower respectively, to show the performance of the control system. In this paper a load was put on a mobile robot in which the nonlinear controller was applied, allowing the transportation of this load from a initial point to a desired position. In order to verify the proposed method some experiments were performed and shown in this paper.

1 INTRODUCTION

Navigating and following are the two basic functionalities of mobile robots. Both of them are widely studied topics. Many strategies have been developed for several kinds of environments either with a single agent systems or multi-agent systems (Bicho and Monteiro, 2003; Das et al., 2002; Egerstedt and Hu, 2001; Egerstedt et al., 2001; Jadbabaie et al., 2002; Vidal et al., 2003).

In most of these studies the robots are used to navigate in order to reach a specific goal point in a given environment while avoiding obstacles (Althaus et al., 2004). However, due to limitations in sensing, most times, robots are not able to operate in complex environments, even if they are provided with accurate sensors like a 3D range finder. Therefore, we believe that a human-robot navigation system is a good solution for this kind of problem, since human can lead the robot in a cooperative task.

The motivation for the work presented in this paper comes from situations where robots are used to help people in an object transportation task where the person does not load the object, only guides the robot through the environment.

A new issue for navigation emerges from interaction tasks, which concerns the control of the platform (Althaus et al., 2004). The main idea is that the robot's behavior must be smooth and similar to the human's movements.

In human-robot interaction the actions of the robot depend on the presence of humans and their wish to interact.

In (Bowling and Olson, 2009) a human-robot team is composed by two omnidirectional robots and one human. The authors consider the case in wich a human turns abruptly a corner and the robots may keep the human in their line of sight. The strategy is accomplished by taking the center of mass of the formation and analyzing its rotation.

The work proposed in this paper consists in coordinating the movements of a mobile robot by a human in order to transport a load from a desired point to another one in a safe way. In this system, the human, which is considered the formation leader, is responsible for guiding the robot through the environment by making the load transportation from an initial point to a goal position and orientation. The robot (follower) must move according to the human maitaining the formation. This type of cooperative work needs the robot to acquire enough information about the human position to achieve the goal. This is reached through a

Garcia Pereira F., Frank Cypriano M. and Frizera Vassallo R. (2010).
FORMATION CONTROL BETWEEN A HUMAN AND A MOBILE ROBOT BASED ON STEREO VISION.
In *Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics*, pages 135-140
DOI: 10.5220/0002880601350140

vision-based nonlinear controller which estimates the linear and angular velocities that the robot must execute to follow the human. This application is useful because as the load is on the robot, the human does not need to carry the object with it. The human function is to guide the robot through the environment.

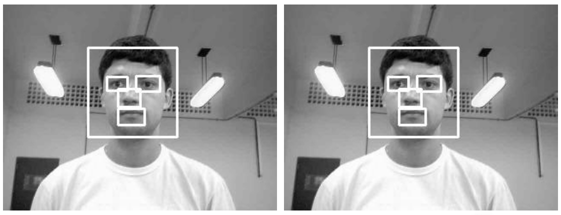
Unlike the conventional leader-follower cooperation, where commonly the formation leader is positioned in front of the follower(s), the formation control proposed in this work takes into consideration the fact that the robot navigates in front of the human, which is considered the leader of the formation. As the formation control is based on vision, a stereo vision system is mounted on top of the robot pointing backwards, which allows the robot to navigate "looking" at the human face, instead of "looking" at the back of the human. Another interesting point of this formation is the fact that, since the robot goes in front of the human (leader), the personn can observe the robot behavior during the cooperation task and, in addition, can notice the presence of obstacles and guide the robot in order to avoid them. This is something valuable because we focused on the cooperation problem and still did not include an obstacle avoidance algorithm to the robot behavior. That is one of our issues for future work.

This paper is organized as follows. The technique for estimating the face pose is briefly discussed in Section 2 and the formation controller as its stability analysis are shown in Section 3. Some experimental results are given in Section 4 and, finally, the conclusions and future work are commented in Section 5.

2 HUMAN DETECTION

The nonlinear controller that will be presented in Section 3 needs the distance from the human to the robot (d) and the angle between the vector which links the human position with the center of the robot (θ). In order to estimate these two variables, it is used a calibrated stereo vision system attached to the robot. The human pose is estimated by detecting the face in both images captured by the vision system and, afterwards, recovering its 3D coordinates.

The face detection is performed with the algorithm proposed in (Viola and Jones, 2001), which utilizes several images to train a detection algorithm based on Adaboost. After detecting the face in both images, an algorithm to detect some facial characteristics is applied and the coordinates of these points in the image plane are determined. Knowing the positions of these points in both images, their 3D coordinates can be recovered. These facial features are, for instance, the eyes, nose and mouth. Figures 1 (a) and 1 (b) show the images captured by the stereo vision system and the detected facial features.



(a) (b)

Figure 1: Images captured by the stereo vision system.

The face coordinates is considered as being the average of the facial features coordinates, i. e.,

$$[X,Y,Z] = \frac{1}{N}\sum_{i=1}^{N}[x_i, y_i, z_i], \quad (1)$$

where N is the number of detected features and $[x_i, y_i, z_i]$, are the 3D coordinates of each detected facial feature in the vision system reference frame attached to the robot. Figure 2 illustrates the robot and the stereo vision system reference frames.

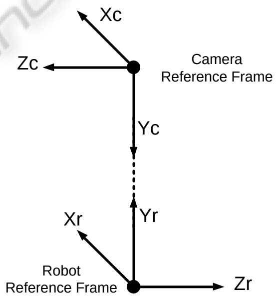


Figure 2: Robot and vision system reference frames.

The relation between these coordinate frames is given by a rotation of 180° around the x-axis and a translation of $1.2m$ in the y-axis. Thus,

$$R_f = T_y(1.2)\,R_x(180)\,C_f \quad (2)$$

where R_f and C_f are respectively the coordinates in the robot reference frame and in the camera reference frame.

After estimating the 3D coordinates of the face, the variables d and θ are calculated on the vision system's XZ-plane (ground plane). Figure 3 shows the human and the vision system positions in the XZ-plane.

The values of d and θ, that define the distance and the angle between the human and the robot, are deter-

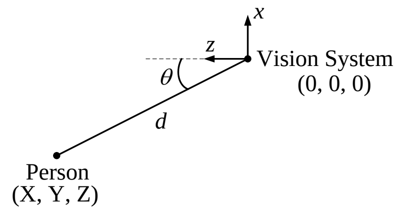


Figure 3: Human and vision system position in the XZ-plane.

mined as

$$\begin{cases} d &= \sqrt{X^2+Z^2} \\ \theta &= \text{atan}\left(\dfrac{X}{Z}\right), \end{cases} \tag{3}$$

where X and Z are, respectively the x and z coordinates of the face. Notice that, however d and θ where calculated in the camera frame, they are valid in the robot frame, due to the aligment between the two frames.

3 CONTROLLER

The aim of the proposed controller is to keep the robot with a desired distance and orientation to the human partner in order to accomplish a specific task. That is done by using the information obtained from the pair of images provided by the stereo vision system assembled on the robot.

Figure 4 illustrates a mobile robot with a stereo vision system attached to it and one person behind it. Considering this case, the kinematic equations that describe the robot movement with respect to the person are given by

$$\begin{cases} \dot{d} &= v_h - v_r\cos(\theta) \\ \dot{\theta} &= \omega + \dfrac{v_r\sin(\theta)}{d} + \dfrac{v_h\sin(\tilde{\theta})}{d}, \end{cases} \tag{4}$$

where v_r and ω are, respectively, the robot linear and angular velocities, v_h is the human velocity, d is the distance between the person and the vision system while θ is the angle of the human face according to the stereo vision system frame and $\tilde{\theta}$ is the angular error.

The values of d and θ are obtained from the 3D coordinates of the middle of the face after 3D reconstruction. The way to obtain these values was explained in Section 2.

After determining d and θ, the distance and orientation errors ($\tilde{d}$, $\tilde{\theta}$) are calculated as

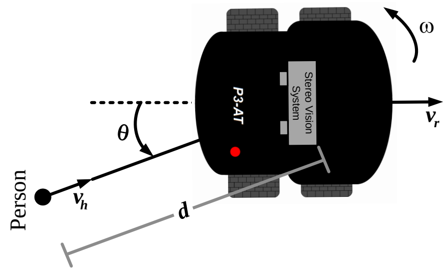


Figure 4: Relative position between the robot and the person.

$$\begin{cases} \tilde{d} &= d_d - d \\ \tilde{\theta} &= \theta_d - \theta, \end{cases} \tag{5}$$

where d_d and θ_d are respectively the desired distance and orientation that the robot has to keep from the human.

For the presented system, it is proposed a nonlinear controller in order to accomplish the objective of control, i. e, to keep the robot following the human with specified distance and orientation. For this controller, the linear and angular velocities are

$$\begin{cases} v_r &= \dfrac{1}{\cos(\theta)}\left(v_h - k_d\tanh(\tilde{d})\right) \\ \omega &= k_\theta\tanh(\tilde{\theta}) - v_r\dfrac{\sin(\theta)}{d} - v_h\dfrac{\sin(\tilde{\theta})}{d}, \end{cases} \tag{6}$$

with $k_d,\ k_\theta > 0$.

The control structure used in this work appears in Figure 5. The operations of detecting faces and facial features, recovering the 3D coordinates and determining the values of d and θ, are executed by the "Image Processing" block.

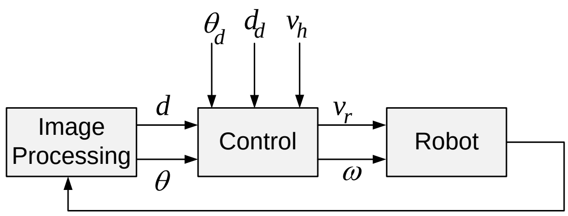


Figure 5: Control structure.

3.1 Stability Analysis

For the nonlinear controller proposed in (6) the following candidate Lyapunov function is taken to check the system stability

$$V(\tilde{d},\tilde{\theta}) = \frac{1}{2}\tilde{d}^2 + \frac{1}{2}\tilde{\theta}^2, \tag{7}$$

which is positive semi-defined and has the following time-derivative function

$$\dot{V}(\tilde{d},\tilde{\theta}) = \tilde{d}\dot{\tilde{d}} + \tilde{\theta}\dot{\tilde{\theta}}. \qquad (8)$$

Thus, differentiating (5) with respect to time and replacing the result in (8) we will have

$$\dot{V}(\tilde{d},\tilde{\theta}) = -\tilde{d}\dot{d} - \tilde{\theta}\dot{\theta}. \qquad (9)$$

Now, the values of $\dot{d}$ and $\dot{\theta}$ in (9) are substituted by the respective values from (4). Then we have

$$\dot{V}(\tilde{d},\tilde{\theta}) = -\tilde{d}[v_h - v_r\cos(\theta)] - \tilde{\theta}\left[\omega + \frac{v_r\sin(\theta)}{d} + \frac{v_h\sin(\tilde{\theta})}{d}\right]. \qquad (10)$$

Replacing the control actions proposed (6) in (10), the time-derivative of the Lyapunov function is rewritten as

$$\dot{V}(\tilde{d},\tilde{\theta}) = -k_d\tanh(\tilde{d})\tilde{d} - k_\theta\tanh(\tilde{\theta})\tilde{\theta}. \qquad (11)$$

As mentioned k_d and k_θ are positive, and functions like $x.\tanh(x) \geq 0\,\forall\,x$. So, (8) is negative semi-defined, which means the asymptotic stability of the proposed nonlinear controller, i. e., $\tilde{d},\tilde{\theta} \rightarrow 0$ when $t \rightarrow \infty$.

Notice that to estimate the angular velocity it is necessary to know the human velocity, v_h. This velocity can be estimated as

$$v_h \approx \frac{d(t) - d(t-1)}{\Delta t} - v_r\cos(\tilde{\theta}). \qquad (12)$$

4 EXPERIMENTAL RESULTS

In order to prove the reliability of the proposed method, some experiments were performed and two of them will be shown in this section. These experiments were performed with a mobile robot Pioneer 3AT from ActivMedia Robotics, a stereo vision system built with two USB web cameras, which provide images with 320x240 pixels. The control system is written in C++ and each control loop is executed in 250ms. The robot with the stereo vision system used to perform the experiments can be seen in Figure 6.

4.1 Experiment 1

In this experiment, the human leads the robot from the initial position, which has coordinates $(X,\ Z) = (0m,\ 0m)$ to the coordinates $(X,\ Z) = (3m,\ -5m)$ through a "S-like" path. The trajectory described by the robot can be seen in Figure 7.

The linear and angular velocities performed by the robot are shown, respectively, in Figures 8 (a) and (b).

Figure 6: Robot used to perform the experiments.

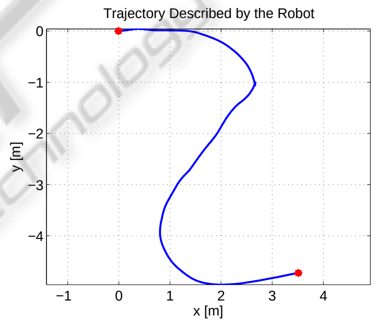
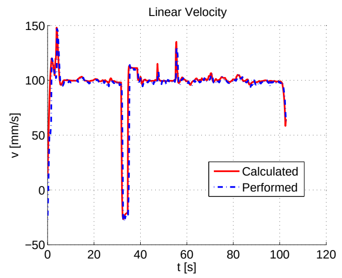


Figure 7: Experiment 01: Trajectory described by the robot.

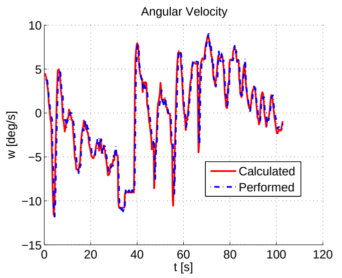
The continuous line represents the velocities calculated by the controller and the velocities performed by the robot are plotted as dashed line.

In Section 3.1 the control system stability was demonstrated, i. e., it was proved that the system errors tend to zero when the time grows and if the human stops at some place. In order to illustrate this, the graphics shown in Figure 9 present the errors during the experiment. Figure 9 (a) shows the distance error during the experiment while the angular error is presented in Figure 9 (b).

By observing Figure 7 it is possible to see that the robot reached the goal describing desired trajectory. Moreover, as can be seen in Figure 9, both the distance error and the orientation error tend to zero, which means the robot accomplished the task. For the next experiments, the errors will not be shown.

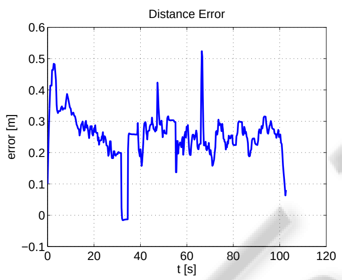


(a)

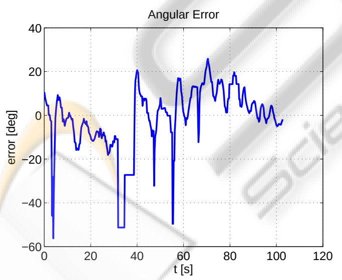


(b)

Figure 8: Experiment 01: Velocities performed by the robot. Linear (a) and Angular (b).



(a)



(b)

Figure 9: Experiment 01: Distance error (a) and Angular error (b).

4.2 Experiment 2

In this experiment the human should guide the robot from the initial position $(X, Y) = (0m, 0m)$ to the final coordinates $(X, Y) = (5m, 0m)$. However, there was a circular obstacle with diameter equal to $0.4m$. The trajectory described by the robot during the execution of this task is shown in Figure 10.

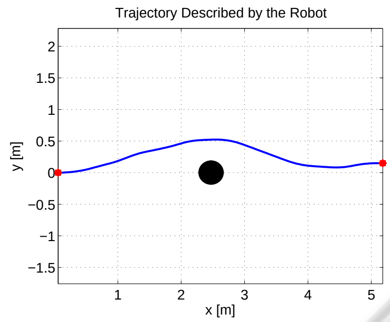
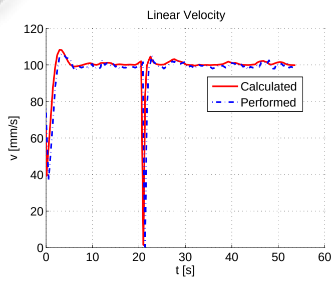


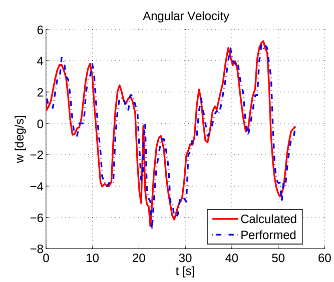
Figure 10: Experiment 02: Trajectory described by the robot.

It is important to mention that the robot does not have an obstacle avoidance algorithm. So, the deviation that can be seen in Figure 10 occurred due to the human action, i. e., the person who leads the robot noticed an object in the middle of the desired trajectory and guided the robot in order to deviate from it. Even though it seems that the person has to walk in an unnatural way in order to prevent the robot from crashing into obstacles or corner in a hallway, this problem will be solved by adding an obstacle avoiding strategy such as in (Chuy et al., 2007; Pereira, 2006).

The linear velocity is shown in Figure 11 (a) while the angular velocity appears in Figure 11 (b).



(a)



(b)

Figure 11: Experiment 02: Velocities performed by the robot. Linear (a) and Angular (b).

Notice that around the 20th second of the experiment, both linear and angular velocities calculated by the controller are equal to zero. This happened because in this instant, the robot lost the information about the person for, at least, one of the cameras of the stereo vision system. This behavior was adopted in order to avoid having the robot be out of control every time the vision system loses the information about the human. Thus, whenever the robot loses the information about its leader, it stops and waits for a new face detection.

Although an obstacle in the middle of the desired trajectory, the human managed to guide the robot to avoid this object and reach the desired point.

5 CONCLUSIONS AND FUTURE WORK

This paper presented an approach to a formation control between a human and a mobile robot using stereo vision. The strategy uses the recognition method presented by (Viola and Jones, 2001) in order to find the facial features needed to estimate the human position. A stable nonlinear controller is proposed to allow the robot to perform the task in cooperation with a human. The effectiveness of the proposed method is verified through experiments where a human guides the robot from an initial position to a desired point, when the human moves either forward or backwards.

Our future work is concerned in improving the features detection in order to better estimate the human position and orientation. Besides, we also intend to introduce a second robot to the formation. Thus, the robots will be able to carry a bigger and heavier load.

REFERENCES

Althaus, P., Ishiguro, H., Kanda, T., Miyashita, T., and Christensen, H. (2004). Navigation for human-robot interaction tasks. In *ICRA '04. 2004 IEEE International Conference on Robotics and Automation, 2004. Proceedings.*, volume 2, pages 1894–1900 Vol.2.

Bicho, E. and Monteiro, S. (2003). Formation control for multiple mobile robots: a non-linear attractor dynamics approach. In *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 2, pages 2016–2022 vol.2.

Bowling, A. and Olson, E. (2009). Human-robot team dynamic performance in assisted living environments. In *PETRA '09: Proceedings of the 2nd International Conference on PErvsive Technologies Related to Assistive Environments*, pages 1–6, New York, NY, USA. ACM.

Chuy, O., Hirata, Y., and Kosuge, K. (2007). Environment feedback for robotic walking support system control. In *ICRA*, pages 3633–3638.

Das, A., Fierro, R., Kumar, V., Ostrowski, J., Spletzer, J., and Taylor, C. (2002). A vision-based formation control framework. *Robotics and Automation, IEEE Transactions on*, 18(5):813–825.

Egerstedt, M. and Hu, X. (2001). Formation constrained multi-agent control. In *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, volume 4, pages 3961–3966 vol.4.

Egerstedt, M., Hu, X., and Stotsky, A. (2001). Control of mobile platforms using a virtual vehicle approach. *Automatic Control, IEEE Transactions on*, 46(11):1777–1782.

Jadbabaie, A., Lin, J., and Morse, A. (2002). Coordination of groups of mobile autonomous agents using nearest neighbor rules. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 3, pages 2953–2958 vol.3.

Pereira, F. G. (2006). Navegação e desvio de obstáculos usando um robô móvel dotado de sensor de varredura laser. Master's thesis, Universidade Federal do Espírito Santo - UFES.

Vidal, R., Shakernia, O., and Sastry, S. (2003). Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation. In *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, volume 1, pages 584–589 vol.1.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001.*, 1:I–511–I–518 vol.1.