Biologically Inspired Path Execution using SURF Flow in Robot Navigation

Xavier Perez-Sala*, Cecilio Angulo*, and Sergio Escalera+

* CETpD-UPC. Technical Research Centre for Dependency Care and Autonomous Living, Universitat Politècnica de Catalunya,
Neàpolis, Rambla de l'Exposició, 59-69, 08800 Vilanova i la Geltrú, Spain
+ MAiA-UB. Dept. Matemàtica Aplicada i Anàlisi, Universitat de Barcelona,
Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain,
*xavier.perez-sala@upc.edu,cecilio.angulo@upc.edu
+sergio@maia.ub.es

Abstract. An exportable and robust system using only camera images is proposed for path execution in robot navigation. Motion information is extracted in the form of optical flow from SURF robust descriptors of consecutive frames, so the method is called SURF flow. This information is used to correct robot displacement when a straight forward path command is sent to the robot, but it is not really executed due to several robot and environmental concerns. The proposed system has been successfully tested on the legged robot Aibo.

Keywords: Robot navigation, Path execution, Optical flow, SURF

1 Introduction

Navigation for autonomous mobile robots, for any kind of platform and independently to its task, implies to solve two related problems: path planning and path execution. Path planning can be defined as a high level robot guidance from a place to another place, while path execution refers to low level processes needed to fulfill path planning decisions [16]. This work is about, given a certain path plan, how to ensure path execution when the only available information for the robot is data extracted from its on-board camera. Especially, no landmarks in the environment will be considered.

Unexpected robot behaviours can be observed during path execution when a system is asked for reaching a place or set point, though it acted properly in simulated or ideal conditions. Failures in path execution, even for simple path executions like a 'go straight forward' path command, are due to several reasons: noise in the sensors, damages in the actuators, perturbations, model errors or shocks. Consequently, a feedback control would be interesting to be implemented to correct the robot from possible motion deviations.

A common approach for obtaining feedback is to consider some landmarks in the environment that help the robot to be localized in [15, 16]. However, for a general solution, no landmark should be considered, and no exact final place in the path where to arrive exist, which could act like a landmark. Another solutions focus on constrain robot motion and camera localization on the robot in order to obtain robot egomotion [2, 4, 5]. Since nor robot configuration, neither camera localization will be constrained, but be placed in the front direction, egomotion can not be considered. The general problem at hands is to ensure the execution of a 'go straight forward' path command by a general mobile robot, when frames from the on-board frontal camera is the only available information.

Our proposed approach, like those based on optical flow [2], will use consecutive frames from the on-board robot camera to extract an approximation of the displacement direction by observing 2-D displacements of brightness patterns in the image. However, unlike standard solutions, the robot direction will be computed online by extracting the so-called *SURF flow*, i.e. motion information from SURF robust descriptors of consecutive frames of image sequences provided by the robot camera. This knowledge will be the only one needed to close the control loop, and to achieve the desired straight forward movement.

Otical flow is a measure closely related with motion field [1], i.e. the projection of 3-D relative velocity vectors of the scene points onto the 2-D image plane. During a frontal displacement, motion field shows a radial configuration: vectors radiate from a common origin, the *Vanishing Point* (VP) of the translation direction. In particular, forward displacements generate vectors point away from this point, named *Focus Of Expansion* (FOE), else it is named *Focus Of Contraction* (FOC). It is proposed in this work to achieve straight forward control for mobile robots by maintaining the FOE in the center of the SURF flow.

The remaining work is organized as follows: the state of the art about robot navigation using optical flow is introduced in Section 2. Section 3 describes the solution proposed for the straight forward robot motion. In Section 4, experiments are described and results are discussed. Finally, possible improvements and further reserarch lines are listed in Section 5.

2 Related work

Biological principles of insect vision [7, 11] have inspired vision-based solutions in robot navigation for obstacle avoidance. Insects extract qualitative 3-D information using image motion to avoid obstacles. Vision-based control techniques try to balance the optical flow divergences betwen eyes/sides of the image. In [8], an approach from ecological psychology was presented to avoid obstacles based on the visual field with the lowest time to contact. As indicated in [6], qualitative measures of flow field divergence are a reliable indicator of the presence of obstacles. In the same way, it has been proposed [10] and demonstrated [9] that humans use optical flow to perceive translational direction of self-motion: radial patterns generated by the optical flow during frontal movement guide human locomotion.

Besides qualitative information, motion field can provide more accurate measurements. It is possible to estimate the relative motion between camera and scene, i.e. egomotion, by considering some hard assumptions. In [2], constraints are met and optical flow is used as an approximation of motion field to compute translational and angular velocities of a car. Egomotion can also be used to localize the robot in the environment. In [4, 5], the navigation task is divided in three phases: localization, path finding, and path execution. Optical flow is used to correct localization. In [4], odometry computed from wheel encoders is improved with an inaccurate egomotion, computing vehicle speed from optical flow. In [5], better results are presented from visual odometry, and localization is made only using egomotion. However, for path execution, our goal, global localization is a hard task to be avoided. Hence, a system is described in [3] allowing a wheeled robot to drive through the center of a corridor by controlling the steering angle. Robot navigates aligning the camera to the wall, at a certain distance, only using a rigidly mounted camera.

Using steering angle as control signal, a novel method will be proposed to detect translational direction without global localization (egomotion) or relative references (landmarks or a wall). Mimicking the human use of optical flow, steering angle will be calculated from radial patterns around the vanishing point (FOE in our case) that optical flow generates during translational movements. Several works exist where FOE is located from optical flow, but none of them use it as a feedback signal to correct robot navigation. For pure translation displacements, FOE calculation is completely described in [1]. Else, when the rotational component is non-zero, optical flow vectors will not intersect on FOE. However, it is the most trivial method to compute FOE, as it was pointed out in [14], where FOE is computed for locomotion control using an Artificial Neural Network, but it was never implemented for this goal. A simple method to solve rotations was introduced in [13] by discounting arbitrary rotations and applying the method for pure translation. However, it is claimed in [12] that navigation methods using optical flow are usually based on unrealistic assumptions about the scene, and unrealistic expectations about the capabilities of motion estimation techniques. Better results could be obtained by directly determining general qualitative properties of the motion structure (FOE computation), instead of a precise analysis of rotational parameters.

3 Robot Navigation Control

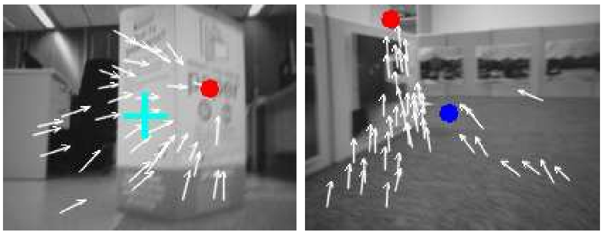
A method to control the path execution during the navigation of mobile robots is introduced. A closed loop is implemented to control straight forward displacements, with feedback signal extracted from robot camera images. Proposed procedure is composed by three steps: firstly, motion information is extracted from consecutive frames through SURF flow computation. Next, instantaneous direction of translation is computed by finding the Focus Of Expansion (FOE) from SURF flow vectors. Finally, control loop is closed, maintaining constant the direction of translation. Hence, straight forward displacements are ensured without the use of egomotion, odometry information is omitted, robot localizationt is avoided, and computational resources are dedicated to achieve reliable orientation measurements for the control module.

Procedure 1 Vision-based navigation control at instant k

Input: Current image I_k from the camera (Fig. 1(b)), number of frames taken during a robot step h, horizontal camera resolution res_x, horizontal opening angle oa_x and set point in pixels $sp_x^p = res_x/2$

Output: Steering angle: $e_{x_k}^o$

loop
  Compute SURF descriptors and keypoint locations: P_k
  Find correspondences between P_k and P_{k-1}: M_k
  Compute intersections of motion vectors M_k: C_k
  Estimate Vanishing Point from highest density region in C_k: (VP_{x_k}, VP_{y_k})
  Apply temporal filter using h last VP: $\mathrm{AVP}_k = \frac{1}{h+1}\sum_{i=k-h}^{k} \mathrm{VP}_i$
  Compute horizontal error in pixels: $e_{x_k}^p = AVP_{x_k} - sp_x^p$
  Transform error $e_{x_k}^p$ to angles: $e_{x_k}^o = e_{x_k}^p \,(oa_x/res_x)$
end loop



(a) (b)

Fig. 1. (a) Error signal depends on distance from the center of the image to the VP (b) Average Vanishing Point computation. Red-top point represents the current vanishing point (VP_k), and blue-centred point is the averaged one (AVP_k).

3.1 Feedback Control

To achieving a straight forward displacement, the robot motion target will be to hold the same orientation during all the path execution. From the camera point of view, this target is similar to hold the vanishing point in the center of the image (Fig. 1(a)). The error signal to close the loop will be calculated from video signal feedback, by computing distance between VP and the actual center of the image. Since the control variable will be the steering angle, only horizontal component of distance will be used to define it.

3.2 Vanishing Point

During frontal displacements, motion field displays a radial vector configuration around a common origin, the vanishing point of the translation direction. Motion field is not a directly accessible measure, but it is closely related with optical flow, under certain circumstances [2]: (1) robot moves on a flat ground, with (2) on-board camera translating in parallel to the ground and (3) its angular velocity is perpendicular to the ground plane. For general robots like that used in this work, nevertheless, constraints do not meet. The Sony Aibo robot is a quadruped robot with a camera on its "nose". Thus, image sequence are more instable than those provided by a wheeled vehicle, with a camera mounted rigidly on its structure. Image instability is due to neck joints, causing head vibrations transmitted to the camera, and specially, for robot walking. Legged robot steps produce very different movements than wheeled robot displacements, usually smoother than Sony Aibo gait. Walk behaviour in our experiments generates vertical and left-right pendular movements, i.e. camera suffers simultaneous roll and pitch rotations. Only the first assumption could be fulfilled in this case.

The hardest assumption of our approach is made at this point. Since Aibo robot gait is symmetric and periodic, restrictions two and three can be assumed as satisfied 'in average' and they will be extrapoled, during robot displacements, for instantaneous translation. Therefore, Sony Aibo gait deviations will be considered like shocks and vibrations which the controller will correct. As shown in Section 4, our qualitative approach is enough to control the desired legged robot navigation. A temporal filter is performed to compute VPs as averaged during robot gait. The Averaged Vanishing Point (AVP), described in Algorithm 1, is the point from which is computed the steering control.

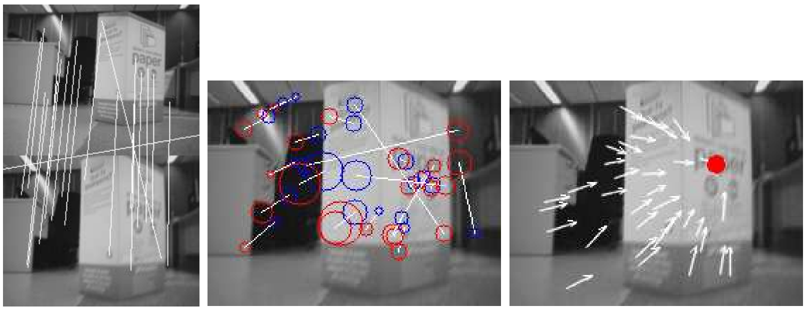
As it was pointed out, calculated optical flow vectors do not converge to an unique point (FOE), even when assumptions are met. Hence, VP has been extracted by clustering intersections, since they form a cloud around VP.

3.3 SURF Flow

SURF flow is defined as 2-D displacements of SURF patterns in the image, where SURF is referred to Speeded Up Robust Features [17]. It is the field resulting from correspondences between SURF keypoints from consecutive frames in a video sequence. Unlike optical flow or the more similar SIFT flow [19], SURF flow is not a dense flow. It is only performed between high confidence keypoints in the image, selected by using a multi-scale Hessian detector to find image corners. SURF flow computation is faster than SIFT flow, since correspondences are only searched for a few hundreds of keypoints in each image (depending on the image texture), and corner detection and SURF description are computed using Haar wavelets on the integral image representation. Result of this correspondence is shown in Fig. 2(a) and Fig. 2(b).

Moreover, an image correspondence post-processing is applied in order to achieve better VP computation. This refinement, shown in Fig. 2(c), takes place once SURF flow is extracted and an estimation of VP is computed (see Section 3.2). It consists on search for better correspondences for each keypoint in current image, looking for similar SURF descriptors in a restricted area of previous image. This search area is defined by the triangle ABC, where vertex A is the keypoint in current image, the middle point of edge BC is the estimated VP and angle $\widehat{BAC}$ defines the search range. Once correspondences are refined, VP is computed again, using the same process described above.

Method effectiveness depends, as usual, on assuming that keypoints are found in images, i.e. a textured environment exists. In fact, typical human-hand scenes



(a) (b) (c)

Fig. 2. (a) Keypoint correspondences between consecutive images (b) Motion **vectors in** the newest image (c) Refined motion vectors with the correspondent vanishing point.

have enough corners for achieve SURF flow performance. On the other hand, SURF flow is robust to optical flow methods' limitations [20]: brightness constancy, temporal persistence or "small movements", and spatial coherence.

4 Results and Discussion

Results presented in this work are obtained using a Sony Aibo ERS-7 robot wirelessly communicated with a standard dual-core PC. Experiments are performed using the robot for environment interaction and the computer for hard computation processing. Path execution has been divided in reactive collision avoidance and straight forward control. Obstacle avoidance procedure is performed on-board, as a reactive behaviour using the robot infrared sensor, and computation to go straight forward is executed in the external computer. Sony Aibo camera captures the image and it is sent to the PC every 100ms, through wireless connection. Application running on the computer, first of all, extracts SURF flow from consecutive frames; then, the VP of the translation direction and the steering angle are computed; and finally, walking direction is sent to Sony Aibo. Gait behaviour for the robot is based on the Tekkotsu software[1].

Experiments are performed in an artificial grass surface of about $4m^2$, containing two crossing corridors. It is a natural scenario without artificial landmarks and small variability of the light level. To allow a future development in unstructured environments, corridor walls are wallpapered with pictures of real halls and corridor walls; providing enough textures to the system to ensure the correct performance of image processing algorithms. Used image resolution is 208×159 pixels.

In order to achieve qualitative results of the system performance in different relative positions between the robot and walls, 8 representative starting positions and orientations are chosen around the scenario, equally distributed, and 5 trials

[1] `http://www.tekkotsu.org/`

Fig. 3. Navigation sequence in open loop control.

Fig. 4. Navigation sequence, with straight forward control.

are launched for each one. Results show the difference between non-controlled straight forward behaviour and the controlled one. In open loop control, due to their mechanical characteristics, robot walks drawing a curve (Fig. 3). When feedback control is applied, Sony Aibo robot goes successfully straight forward (Fig. 4), correcting faulty displacements and performing the desired behaviour.

Some problems with wireless connection are observed, and sometimes image is not sent at time from robot to computer. When it occurs in consecutive images, it produces large oscillations, which can be corrected or not, depending on the number of frames lost. If problem persist, it can produce uncontrolled behaviours. A precise study about the maximum number of lost images supported should be completed depending on the last order sent by the computer, which will be repeated during all the non-informed period.

5 Conclusions and Future Work

In this work it is proposed a biological inspired vision-based navigation control to walk straight forward in a reliable way. Moreover, implementation is exportable to other robotic platforms with different configurations. Results shown that objectives introduced in this work have been accomplished without the use of artificial landmarks, taking into account some assumptions about the robot movement. Since Aibo's camera suffers simultaneous roll and pitch rotations during the robot gait, future work will avoid the hardest assumption proposed. The robot will correct its trajectory using motor information.

Moreover, shocks and vibrations suffered by the camera will be compensated by tacking in account robot configuration. Future work will be an improving of the system presented in this work, to be used in legged robots. In [2], motion field is formulated supposing an error component due to shocks and vibrations. Nevertheless, motion field error in x and y axis are roughly estimated. At this point, we are in an advantageous position, because it is assumed that our shocks and vibrations are movements resulting to the quadruped robot gait, and these movements are possible to be modelled through direct kinematics. Other improvements include decreassing sampling rate and the duration of actions.

References

1. Trucco, E. Verri, A.: Introductory Techniques for 3-D Computer Vision. Prentice Hall PTR, Upper Saddle River, NJ, USA (1998)
2. Giachetti, A., Campani, M., Torre, V., CRS, C.: The Use of Optical Flow for Road Navigation. In: IEEE Trans. on Robotics and Automation, vol. 14, pp. 34–48 (1998)
3. Dev, A., Krose, B., Groen, F.: Navigation of a mobile robot on the temporal development of the optic flow. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 558–563 (1997)
4. Nagatani, K., Tachibana, S., Sofne, M., Tanaka, Y.: Improvement of odometry for omnidirectional vehicle using optical flow information. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 1, pp. 468–473 (2000)
5. Sorensen, D.K., Smukala, V., Ovinis, M., Lee, S.: On-line optical flow feedback for mobile robot localization/navigation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1246–1251 (2003)
6. Nelson, R.C., Aloimonos, J.: Obstacle avoidance using flow field divergence. In: Trans. on Pattern Analysis and Machine Intelligence, vol. 11, pp. 1102–1106 (2002)
7. Srinivasan, MV., Chahl, JS., Weber, K., Venkatesh, S., Nagle, MG., Zhang, SW.: Robot navigation inspired by principles of insect vision. Robotics and Autonomous Systems 26, 203–216 (1999)
8. Duchon, A.P., Warren, W.H.: Robot navigation from a Gibsonian viewpoint. In: IEEE International Conference on Systems, Man, and Cybernetics, 'Humans, Information and Technology', vol. 3, pp. 2272–2277 (1994)
9. Warren, W.H., Kay, B.A., Zosh, W.D., Duchon, A.P., Sahuc, S.: Optic flow is used to control human walking. Nature Neuroscience 4, 213–216 (2001)
10. Warren, W.H., Hannon, D.J.: Direction of self-motion is perceived from optical flow. Nature 336, 162–163 (1988)
11. Santos-Victor, J., Sandini, G., Curotto, F., Garibaldi, S.: Divergent stereo in autonomous navigation: From bees to robots. International Journal of Computer Vision 14, 159–177 (1995)
12. Thompson, W.B., Kearney, J.K.: Inexact Vision. In: Proc. Workshop on Motion: Representation and Analysis, pp. 15–22 (1986)
13. Negahdaripour, S., Horn, B.K.P.: A Direct Method for Locating the Focus of Expansion. Computer Vision, Graphics, and Image Processing 46, 303–326 (1989)
14. Branca, A., Stella, E., Attolico, G., Distante, A.: Focus of Expansion estimation by an error backpropagation neural network. Neural Computing & Applications 6, 142–147 (1997)
15. Yoon, K., Jang, G., Kim, S., Kweon, I.: Color landmark based self-localization for indoor mobile robots. J. Control Autom. Syst. Eng. 7, 9, 749–757 (2001)
16. Deng, X. and Milios, E. and Mirzaian, A.: Landmark selection strategies for path execution. Robotics and Autonomous Systems 17, 171–185 (1996)
17. Herbert, B., Tinne, T., Luc, VG.: Surf: Speeded up robust features. Lecture Notes in Computer Science 3951, 404–417, Computer Vision–ECCV 2006 (2006)
18. Lowe, D.G.: Object recognition from local scale-invariant features. In: IEEE International Conference on Computer Vision, vol. 2, pp. 1150, USA (1999)
19. Liu, C., Yuen, J., Torralba, A., Sivic, J., Freeman, W.: SIFT flow: dense correspondence across different scenes. Lecture Notes in Computer Science 5304, 28–42, Computer Vision–ECCV 2008 (2008)
20. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: IJCAI, vol. 3, pp. 674–679, Citeseer (1981)