20th IMEKO TC4 International Symposium and
18th International Workshop on ADC Modelling and Testing
Research on Electric and Electronic Measurement for the Economic Upturn
Benevento, Italy, September 15-17, 2014

Autonomous navigation of mobile robots: from basic sensing to problem solving

Paolo Tripicchio[1], Massimo Satler[1], Carlo Alberto Avizzano[2], Massimo Bergamasco[2]

[1]*Gustavo Stefanini advanced robotics research center, Scuola Superiore Sant'Anna, La Spezia (Italy), p.tripicchio@sssup.it,m.satler@sssup.it*
[2]*PERCeptual RObotics laboratory, Scuola Superiore Sant'Anna, Pisa (Italy), c.avizzano@sssup.it,m.bergamasco@sssup.it*

Abstract – **Autonomous navigation is a complex task that requires both sensing capabilities to react to sudden environmental changes or map the environment and reasoning to schedule the next action to perform. Starting from basic sensing technology used in the majority of mobile robotic systems, the introduction of sensor fusion techniques allows to obtain useful information to solve the localization, mapping and navigation problems. Applications of these methods to achieve specific robot capabilities will be presented starting from object detection and recognition, passing to scene classification and ending with an industrial related application: the visual inspection of industrial facilities by means of a flying vehicle.**

I. INTRODUCTION

In the last years the field of mobile robotics is transitioning from user-teleoperated systems to completely or semi-autonomous systems. An autonomous robot is a robot that performs behaviors or tasks with a high degree of autonomy, which is particularly desirable in cases like environments exploration, in activities that are potentially dangerous for humans, or again in applications where the robot performs better or in a shorter period of time.

To achieve autonomous navigation and intelligent behavior in robots, it is necessary to solve specific problems: starting from the well-known mobile robotics questions "Where am I?" and "How can I reach a specific location?". Such questions are related to the localization, mapping, relocation and path planning problems that require reactive behaviors of the robots like collision avoidance, location reaching, following a target and so on and robot capabilities like moving objects detection, people recognition, place classification, etc.

In general a fully autonomous robot should be able to gain information about the surrounding environment, work for an extended period of time without human intervention and possibly avoid situations that are harmful to people, property, or itself. An autonomous robot should also be able to learn or gain new knowledge to adapt to surroundings changes. To fulfill all these capabilities, the robot requires several proprioceptive and exteroceptive sensors.

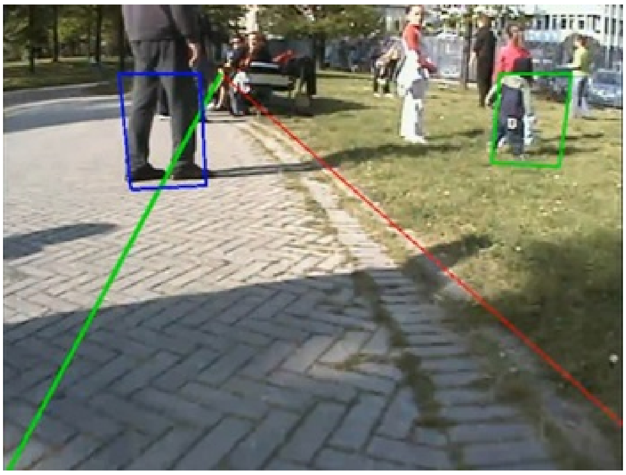


Fig. 1. Robot navigating in outdoor environment recognizing driving lane and moving people.

Proprioception, or sensing one's own internal status, is required for robots to work autonomously near people and in harsh environments. Common proprioceptive sensors include thermal, optical, electrical and haptic sensing. Exteroception is sensing things about the environment. Autonomous robots must have a range of environmental sensors to perform their task while keeping an adequate level of knowledge and safety. Common exteroceptive sensors include the electromagnetic, sound, touch, chemical, temperature, range, pressure, and altitude sensors. All these sensors cannot work indipendently but their measures should be fused to obtain good estimation of environmental and internal variables.

The paper presents possible approaches to solve the autonomous navigation problems and the specific solutions to obtain the required robotic capabilities. Section ii. will introduce a short state of the art on localization and mapping techniques introducing both indoor and outdoor navigation. In section iii. a possible approach to obtain autonomous navigation will be introduced. Section iv. will present specific inplementations to achieve the required sensing capabilities. The last section will present the final

ISBN-14: 978-92-990073-2-7

considerations and conclusions.

II. STATE OF THE ART

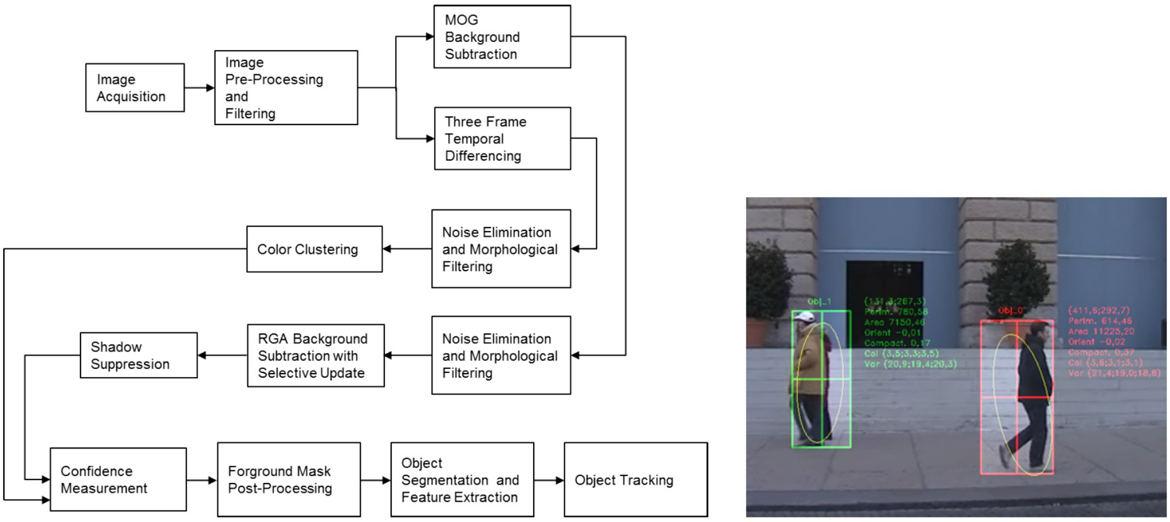
To obtain autonomous navigation, the robot requires to solve the localization problem and to have at least the capability to navigate point-to-point. Starting in the 1970s with wire-guidance and progressed in the early 2000s to beacon-based triangulation, current state of the art robots autonomously navigate based on sensing natural features. At first, autonomous navigation was based on planar sensors, such as laser range-finders, that can only sense at one level. Nowadays, most advanced systems fuse information gained from various sensors for both localization (position) and navigation purposes. Sensor fusion is the key technique to obtain precise positioning and good mapping capabilities as well as exact scene understanding. Many techniques are used in literature for the position estimation. Basically they fuse the odometric information, the inertial measurements and the range sensor readings by means of a state observer or Kalman filter. Algorithms based on scan matching [18] employ all the data coming from range scanners to generate a local map and then measure the motion occurred between consecutive scanned maps computing the minimum error coordinate transformation. SLAM (Self Localization And Mapping) algorithms based on grid mapping [8] make uses of a small number of particles which contain each a map represented as an occupancy matrix. These maps represent free and occupied space on a probabilistic base. Feature based SLAM algorithms [7] work on high level information (i.e. features) that are few in number with respect to sensors acquisitions and thus the computational load is reduced. The map representation is reduced as well. More recently, visual odometry algorithms [17] are based on feature extraction and on the association of these features between successive frames. The identification of such features is performed with vision algorithms like SURF [5] or SIFT [6] that identifies points or points clusters with an associated descriptor. Analyzing the motion of these features it is possible to estimate the motion of the camera and at the same time of the robot rigid body.

Moving to outdoor exploration other issues arises. In fact outdoor autonomy is more difficult for ground vehicles, due to the presence of three-dimensional uneven terrains, disparities in surface density and different and changing wheater conditions. In the contrary outdoor autonomy is most easily achieved in the air, since obstacles are rare. In the case of outdoor navigation, the use of sensor like GPS allows to obtain a good positional accuracy with ease during free space movements. For precise positioning however sensor fusion techniques have still to be used.

III. POSSIBLE APPROACH

In many applications, the robot should be able to move independently in a semi-structured environment such as an office, store or house. In most cases it is useful that the robot moves through a series of waypoints or to a specific point of the environment without knowing the obstacles present in the environment. To achieve this goal we propose a hybrid software architecture that consists of a reactive and a deliberative part. The reactive part controls the local navigation responding to obstacles that are found in the environment and to the commands of the deliberative part that knows all the waypoints and navigation controls globally.

The deliberative architectures have historically been the first to be used for the control of autonomous robots. They are based on a functional approach, which expects to acquire an ideally complete knowledge of the environment in order to build an accurate model on which to plan a series of actions that lead to the target. The goal of our robot is to reach a point on the map, while the actions that we have to plan are a sequence of movements of the robot. The part of the system that generates the trajectory starting from the internal representation of the environment is called planner. The algorithms used to implement the planner generally explore a tree of trajectories, excluding those who violate the physical constraints (e.g. obstacles), mechanical (e.g., radii of curvature too tight) or performance (e.g., maximum time, distance covered). The trajectory generated can also be optimum with respect to a cost function, which is generally the minimum time or minimum covered space. If the environment in which the robot moves is dynamic, the path must be recalculated whenever a change occurs, which happens at each cycle in the case of moving robot in dynamic unknown environments. The computing power needed to explore a tree of possible trajectories is very large. The computers will be big and therefore expensive in terms of energy consumed. Generally an entire computer is dedicated only to trajectory planning [9], [13]. To reduce the dimension of the solution two planners can be employed, one long and one short-term: the first deals with creating a series of waypoints up to the final goal, the second generates the path to reach the next waypoint. When changes occur in the world only the second planner must be invoked [12]. The main advantage of using a planner with respect to any other architecture is the security of a good result, a reactive architecture instead can indeed vanish completely in balanced situations such as when a target is behind to a dead end [11]. The main disadvantage is the high computing power required that increases with the raise of the details of a map. The risk of relying on just a planner however is to let the robot out of control in the case of collisions with new barriers along the old planned route [14]. At the end of the 80s it was proposed a new paradigm based on the concept of behavior. The consider-



(a) (b)

Fig. 2. Automated Moving Object Detection and Tracking Algorithm(a). Example of detecting people that move in the scene(b).

ations that led to this new kind of architecture were the inability to model a complex world and the slow response of the robot to changes in the environment. A reactive robot has only laws that define how to react to certain inputs of the sensors, without the use of abstract representations or memory of past events and states. The intelligence of a reactive robot is created by many behavior, coordinated with each other by some competitive or collaborative function. The disadvantages of this type of architecture is the lack of planning, which prevents an optimized navigation, and a map, without which it is possible to get into dead ends or to wander endlessly without ever achieve the goal.

Choosing an hybrid solution allows to take advantage of the good characteristics of both the methods.

For the reactive part there are various approaches,a possible solution can be of using nonlinear control based on attractors and repulsors implemented with differential equations, following the theory of dynamical systems[19].

The deliberative part instead will use modern machine learning and AI techniques to perform reasoning on the action to accomplish and plan tasks esecution.

Both the parts require specific knowledge acquired from the environment and the robot internal state variables

IV. SENSING COMPONENTS

As discussed above, to perform autonomous navigation in an indoor or outdoor environment, some sensing components are necessary in order to be able to execute specific behaviors of the robot or to construct a correct representation of the environment map. This section proposes some solution to acquire moving object and people detection capability useful for collision avoidance, place recognition[16] and object recognition[10] that are useful for scene understanding and reasoning and autonomous navigation capabilities in unstructured environments[15].

A. *Moving objects and People detection*

For object and people detection we use the software architecture shown in Figure 2 a.

In the MOG (Mixture Of K Gaussians) Background Subtraction module the pixel color values are modeled as a Mixture Of K Gaussians $(\mu_i, \sigma_i, \omega_i)$ [2]. Multiple adaptive Gaussians are necessary to cope with acquisition noise and lighting changes. Pixel values that do not fit the background distributions (Mahalanobis distance) are considered foreground. The model copes with multimodal background distributions. The RGA Background Subtraction with Selective Update module is based on the approach proposed in [3] where the background at each (i; j) pixel location is modelled indipendently. The Running Average technique is used for performing a more accurate Background Subtraction (BS). The Confidence Measurement combines the results from BS and Color Clustering (CC) [4]. It reduces false segmentation and provides more accurate object boundary. The Object Tracking tracks objects in successive frames to detect trajectories performing these tasks: finding correspondence, motion estimation, corrective feedback and occlusion detection. The tracking is robust by integrating feature vector template matching and Kalman filtering. An example is reported in Figure 2 b.

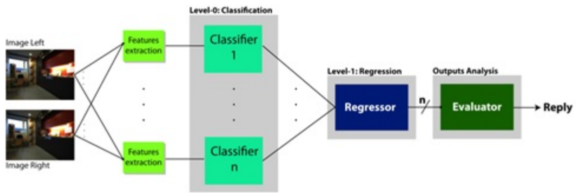


Fig. 3. Combination of classifiers in a two stage stack.

B. *Visual place recognition*

We introduce a machine learning architecture to address the problem of scene or visual place classification in order to localize the robot and in the same time focusing on generalization of the classification technique. Such architecture provides the ability of understanding the topological relations existing in an indoor environment. In particular, this method associates semantic terms (i.e., "Corridor", "Kitchen") to each specific place, giving an "intuitive" idea of the position of the robot. To accomplish this goal, four different methodologies have been integrated. All these methodologies are based on the same multi-level machine learning architecture (Figure 3). The developed architecture is made of a first level of "weak" classifiers based on ”visual” features extracted from stereo images couple, and a second level of classifiers performing fusion and regression of the first level outputs. Four different configurations of the regressor stage have been compared in order to obtain the best results to problem solving. These configurations employ different methods that we call: "Committee of Experts", "Stacked Regression with Support Vector Machines (SVM) stage", "Stacked Regression with Artificial Neural Network stage", "Weighted Linear Combination". The experimental tests show that "Stacked Regression with Support Vector Machines stage" result to be the best approach providing a high generalization capability [16].

Fig. 4. Visual place classification fusing different classification methods.

C. *Object recognition*

A similar approach has been used to solve the problem of object recognition. The approach uses a stacked generalization classification method to perform a multi-layer object recognition fusing heterogeneous spatial and color data acquired with an RGB-D camera. The method assumes a specific a-priori knowledge of the objects that may be present in a scene. This tool exploits the accuracy of the SVM introducing a multilevel architecture with a stacked generalization approach. The main idea here consists in the introduction of a second stage of generalization to deduce the biases of the original generalizers. The second layer input is the output of the first layer of generalizers, and its output is the generalization guess. The raw sensor data are fed to a three stages pipeline, consisting on preprocessing, object detection and classification steps. The output is the identifier of every detected object. This technique has a low computational cost and is suitable for on-line applications, such as robotic manipulation or automated logistic systems[10]. An example of multiple object recognition in presence of occlusions is reported in Figure 5.

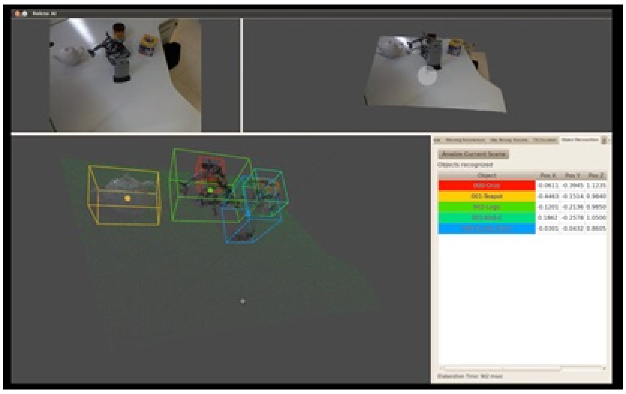


Fig. 5. Detection and classification of multiple objects with occlusions.

D. *Atonomous navigation*

Fusing techniques have also been employed in a flying robot to optimally merge all information gained from the on-board sensors in order to perform simultaneous localization and mapping (SLAM). The flying robot is a Micro Aerial Vehicle capable of navigating in outdoor, indoor and confined environments. The aim of the work was to replace the human operator during the inspection of failures or procedures correctness inside dangerous areas where human presence should be avoided for long or short time periods. To be able to accomplish this specific task the system presents a semi-autonomous behavior. The human operator instruct the system about the trajectory to be performed and the system is able to process all the navigation data onboard so that even in presence of comunication losses the system remains stable and safe.

Given the limited amount of computational power of the embedded computer, an optimized SLAM algorithm that allows correct navigation of the robot has been implemented. The designed system has been oriented towards the specific context of visual inspection of industrial area and it has been tested firstly in a laboratory test-bed [15] and later in a real power plant, see Fig. 6.

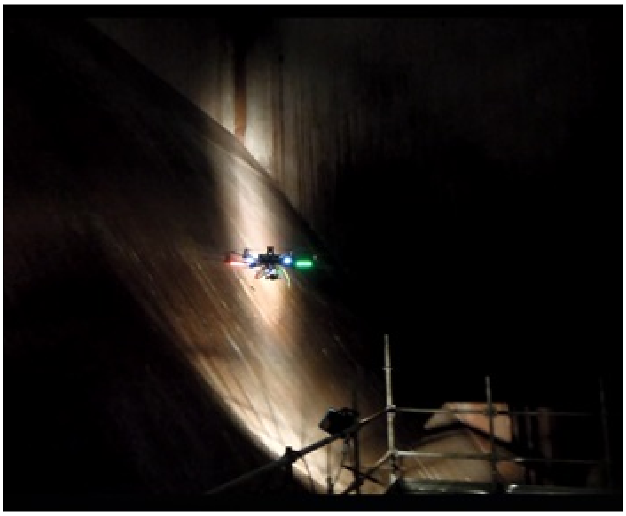


Fig. 6. Flying robot inspectioning an industrial boiler.

V. CONCLUSIONS

Autonomous navigation requires complex robotic abilities. The robot should be able to react to changes in the environments still reaching the planned goal.This paper presented a possible hybrid approach that makes use of both a reactive and a deliberative part for reasoning. To be able to react to sudden changes in the environment, specific vision and machine learning algorithms has been presented for the detection and recognition, in first instance, of objects and people moving in the environment, and in second instance, to classify and recognize the environment itself (and take advantage of the robot knowledge). The presented sensor fusion techniques are also applied to obtain simultaneous localization and mapping in embedded computing hardware and allow small robots to safely operate in harsh environments.

REFERENCES

[1] S.Mallat, “A wavelet tour of Signal Processing. The Sparse Way”, 3rd edition, Academic Press, New York, USA, 2009.

[2] Stauffer, C., Grimson,“W.E.L.: Adaptive Background Mixture Models for Real-Time Tracking”. IEEE CVPR (1999)

[3] Wren, C., Azarbayejani, A., et al.:“Pfinder: real-time tracking of the human body". IEEE Transaction on Pattern Analysis and Machine Intelligence (1997)

[4] Zhu, Y.:“Moving Object Detection and Segmentation Based on Background Subtraction Image Over-Segmentation". J. Software, 6 (2011)

[5] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, 2008

[6] Lowe, David G., ”Object recognition from local scale-invariant features”, Proceedings of the International Conference on Computer Vision, vol. 2, 1999, pp. 1150-1157

[7] Young H. Choi, Tae K. Lee, Se Y. Oh, ”A line feature based SLAM with low grade range sensors using geometric constraints and active exploration for mobile robot”. Auton. Robots, Vol. 24 (January 2008), pp. 13-27

[8] Grisetti, G.; Stachniss, C.; Burgard, W., ”Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters,” Robotics, IEEE Transactions on , vol.23, no.1, pp.34,46, Feb. 2007

[9] Monteiro, S., Bicho, E.: ”A dynamical systems approach to behavior-based formation control". IEEE ICRA, pp. 2606-2611 (2002)

[10] Peppoloni, L., Satler, M., Luchetti, E., Avizzano, C.A., Tripicchio, P.:" Stacked Generalization for scene analysis and object recognition". IEEE International Conference on Intelligent Engineering Systems (2014)

[11] Saffiotti, A., Konolige, K., Ruspini, E.H.:” A multi-valued logic approach to integrating planning and control”. Artificial Intelligence, pp. 481-526 (1995)

[12] Ajzen, I.: "From Intentions to Actions: A Theory of Planned Behavior. Action-control: From cognition to behavior", pp. 11-39. Springer, Heidelberg (1985)

[13] Barraquand, J., Langlois, B., Latombe, J.C.: "Numerical potential field techniques for robot path planning". IEEE Transactions on Systems, Man, and Cybernetics, pp. 224-241 (1992)

[14] Hagras, HA: "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots". IEEE Transactions on Fuzzy Systems, pp. 524-539 (2004)

[15] Satler, M.; Unetti, M.; Giordani, N.; Avizzano, C.A.; Tripicchio, P., ”Towards an autonomous flying robot for inspections in open and constrained spaces”, Multi-Conference on Systems, Signals & Devices (SSD), 2014 11th International , vol., no., pp.1,6, 11-14 Feb. 2014

[16] Walter Lucetti, Emanuel Luchetti, ” Combination of Classifiers fot Indoor Room Recognition”, In proceeding of: CLEF 2010

[17] Scaramuzza, D.; Siegwart, R., "Appearance-Guided Monocular Omnidirectional Visual Odometry for Outdoor Ground Vehicles," Robotics, IEEE Transactions on , vol.24, no.5, pp.1015,1026, Oct. 2008

[18] A. Censi, "An ICP variant using a point-to-line metric" Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2008
[19] S. Behnke, B. Frotschl, R. Rojas, P. Ackers, W. Lindstrot, M. de Melo, A. Schebesch, M. Simon, M. Sprengel, O. Tenchio."Using Hierarchical Dynamical Systems to Control Reactive Behavior". In Lecture Notes in Computer Science Volume 1856, 2000, pp 186-195