Improvement of the Sensory and Autonomous Capability of Robots Through Olfaction: the IRO Project

Javier Monroy$^{a,b,*}$, Jose-Raul Ruiz-Sarmiento$^{a,b}$, Francisco-Angel Moreno$^{a,b}$, Cipriano Galindo$^{a,b}$, Javier Gonzalez-Jimenez$^{a,b}$

$^a$ Machine Perception & Intelligent Robotics Group (MAPIR), System Engineering and Automation Department, University of Malaga

$^b$ Biomedical Research Institute of Malaga (IBIMA), University of Malaga

Abstract

Olfaction is a valuable source of information about the environment that has not been sufficiently exploited in mobile robotics yet. Certainly, odor information can contribute to other sensing modalities, e.g. vision, to successfully accomplish high-level robot activities, such as task planning or execution in human environments. This paper describes the developments carried out in the scope of the IRO project, which aims at making progress in this direction by investigating mechanisms that exploit odor information (usually coming in the form of the type of volatile and its concentration) in problems like object recognition and scene-activity understanding. A distinctive aspect of this research is the special attention paid to the role of semantics within the robot perception and decision-making processes. The results of the IRO project have improved the robot capabilities in terms of efficiency, autonomy and usefulness.

Keywords:
Robotics, Chemical sensors, Object recognition, Semantic networks, Machine learning

Project details:
- **Project name:** Improvement of the sensory and autonomous capability of Robots through Olfaction (IRO)
- **Reference:** 2012-TEP-530
- **Principal Investigator:** Javier Gonzalez-Jimenez
- **Project type (international, national, autonomous, transfer):** Autonomous, Junta de Andalucia
- **Funded by:** Junta de Andalucia
- **Period:** Jan 2014 / July 2019

1. Introduction

The sense of smell is not the most vital one for humans, but we certainly use it every day. When we face a cup with a dark colored liquid we can assure that it is a cup of coffee not only for what we observe, but also for what we smell. When we detect an alarming odor that might be associated to gas/butane we do not look for the possible escape in the living room but we firstly go to the kitchen, where we do not inspect randomly, but we turn our attention to those devices that use gas (e.g. hob, oven, etc.). As in the last example, the smell sense usually triggers alerts: a possible fire, a gas leak, food in poor condition, etc., but also is associated to emotionally rooted processes (Shepherd, 2004): memories, attraction or repulsion, etc. Both facets are interesting in robotics, although the latter, especially relevant in the long term for the so-called "social robots" (Leite et al., 2013; Truong y Ngo, 2018), is beyond our current interest and expectations. The IRO project focuses on the usefulness of a mobile robot able to detect and measure gases in the environment in order to identify the activities carried out in its surroundings, e.g. smoking, cooking, mopping the floor, etc. Having identified the situation, the robot should be able to act consistently, for example, locating and scolding the smoker, avoiding to pass by freshly mopped areas or, perhaps, interacting in a "social" way to help the person who is cooking.

In order to achieve these robot skills, the IRO project relied on "electronic noses" (e-noses) (Röck et al., 2008). E-noses are electronic devices composed by a set of gas sensors and different software components that provide a measure of the type and concentration level of the detected volatile substances. Despite the important advances in recent years in the development of this technology, the performance of gas sensors and algorithms for the classification of gases is still far from the olfactory capacity of humans, not to mention some other animals with much more developed olfactory capabilities.

In consequence, in spite of the limited performance of the current e-noses, the olfactory information interestingly increases the robot abilities when combined with other sensors like vision, and knowledge sources like semantics. For example, if the robot detects smoke, the utilization of vision would
be crucial for identifying an oven and inspecting it as the possible object releasing the alarming gas. Additionally, semantic information regarding the usual location of ovens, i.e. kitchens, can improve the robot actuation.

The final aim of the IRO project has been to enable a mobile robot to (i) combine olfaction and vision information and (ii) exploit semantic knowledge to smartly operate within human environments. Results of the project have been published in Monroy et al. (2016); Gongora et al. (2018); Sanchez-Garrido et al. (2014); Monroy y Gonzalez-Jimenez (2018); Ruiz-Sarmiento et al. (2016). In this paper we provide an overall and comprehensive view of the findings and results of the IRO project.

2. Project Overview

The general objective of the IRO project is to investigate mechanisms for integrating olfactory data into the robot sensing system, as well as the development of algorithms for decision making and task generation that exploit the combination of the different sensor modalities. The key idea behind the project is that the perception of gases, including both their classification and the measurement of their intensity or concentration, can improve the intelligent behavior of the mobile robot, upgrading its performance in terms of efficiency, autonomy and usefulness. Within this global target we can distinguish three partial objectives:

- **Design and fabrication of an artificial nose (e-nose) adapted to the requirements of a mobile robot.** Most of the e-noses used in mobile robotics are designed for measuring only the chemical concentration, aiming at tasks like the creation of concentration maps and/or the search of the emission sources. In the context of the present project, it is necessary that the electronic nose is designed to also provide information on the type of gas, that is, be as effective as possible in the classification of the detected chemical volatile. The objective is, therefore, to combine both facets which requires integrating different sensor technologies into a single device.

- **Gas classification and object recognition for robotics applications.** The robot, equipped with a vision system (e.g., one or multiple RGB or RGB-D cameras) and an electronic nose, could successfully improve the vision-based recognition of simple objects, exploiting the odor information gathered in the surroundings, as well as enhancing the gases classification when considering the semantic information and the probabilistic categorization of the detected object.

- **Exploiting high-level olfactory and visual semantic information in the planning and execution of tasks.** Semantics provide additional human-like information to the perceived elements. For example, a high concentration of gases related to rotten food suggest that somebody forgot about it. Semantic information can be exploited to automatically infer new robot tasks in order to maintain a set of pre-established human-like norms, in this case, rotten food should be taken out of the house. Galindo y Saffiotti (2013). Among the multiple tasks that can benefit from such inference process, we focus on the challenging task of source localization with a mobile robot in indoor environments, aiming at minimizing the necessary time to locate the object emanating the gases in the environment.

Next sections describe with more detail the work done to reach these partial objectives (see Sections 3, 4 and 6).

3. Design and Fabrication of an Artificial Nose Adapted to the Requirements of a Mobile Robot

The first step to attain the objectives identified in this project is the design and fabrication of an e-nose prototype for gas classification and concentration estimation and its integration into a mobile robot. Typically, e-noses detect volatile chemical substances by means of an array of non-selective gas sensors, that is, with sensors that react to a wide range of different gases, but provide no specific information about the chemical identity. Therefore, the output of the sensor array is usually further processed by some sort of machine learning algorithm to classify (Monroy y Gonzalez-Jimenez, 2018; Gutierrez-Osuna y Nagle, 1999) or quantify (Gunter et al., 2016; Monroy et al., 2013) the samples. E-noses offer, as a result, a relatively cheap and fast tool to assess the presence of gases, but with a substantially greater error and uncertainty margin than precise analytic methods, like gas-chromatography or mass-spectrometry (Cui et al., 2015).

Common gas sensor technologies employed to build e-noses include metal oxide (MOX), amperometric electrochemical (AEC), quartz crystal microbalance (QCM), conducting polymers (CP), and surface acoustic wave (SAW). Each of these exhibit advantages and disadvantages in terms of selectivity, sensitivity, response speed, influence by environmental conditions and drift over time, among others (Röck et al., 2008; Monroy et al., 2012). However, no single technology excels in all categories. Thus, limiting the design of an e-nose to a single sensor technology will restrict its performance and, quite often, prevent it from reaching the demanded specifications (Sanchez-Garrido et al., 2014). This motivates the combination of different gas sensor technologies into a single e-nose, which would...
result in a sensor array with better dynamic capabilities and a more informative output than any single sensor technology. Since it is unfeasible to install all possible gas sensors and technologies simultaneously on a single device, it also becomes appealing to design an e-nose in such a way that its sensor array can be reconfigured depending on the applications, keeping it cost-efficient and compact.

To this end, exploiting our previous experience in the design of e-noses (Gonzalez-Jimenez et al., 2011), the IRO project envisaged a novel e-nose architecture (Gongora et al., 2018) that combines self-contained and intelligent sensor boards (i.e., modules) with a decentralized design offering a viable solution to the problem of integrating heterogeneous gas sensors in a modular fashion. This allows us to create different and specific gas-sensing devices from inter-connectable building blocks, which not only brings versatility and reusability to the design of e-noses but also reduces development costs and ensures long-term serviceability, as new sensors can be added as needed. Fig. 1 shows a picture of the prototype built along the course of this project.

Moreover, the proposed e-nose architecture also enables the integration of other electronic components like GPS for georeferenced measurements, or wireless communications for remote readings, a feature which, despite not being a technological contribution, provides an improvement over most commercial e-noses and facilitates applications of mobile robot olfaction.

4. Gas Recognition and Classification for Robotic Applications

The task of odor recognition deals with the problem of identifying a volatile sample among a set of possible categories (Trincavelli et al., 2009). This process plays an important role in the development of many applications, such as city odor mapping (Onkal-Engin et al., 2005; Monroy et al., 2014), pollution monitoring (Hasenfratz et al., 2015), breath analysis in clinical environments (Guo et al., 2010), or the nowadays common estimation of blood alcohol content for drivers (Gibb et al., 1984; Hlastala, 1998). Among them, there are some applications like pollution monitoring or leak detection that require to measure the environment continuously and/or at different locations. For such scenarios, the use of a mobile robot with the capability of identifying and measuring the volatiles concentration is of great help, as already reported in Marques et al. (2002).

4.1. Gas Classification

The classification of volatile substances is, possibly, the most studied application of e-noses. Traditionally, this has been performed by analyzing the response of an array of gas sensors when exposed to pulse-like gas excitation under well-controlled measurement conditions (i.e. temperature, humidity, exposure time, etc.). Unsurprisingly, dozens of works report less than 10% classification error rate under these specific circumstances. However, when the classification is to be performed on a real, uncontrolled scenario, and particularly for the case where the e-nose is collecting samples onboard a moving platform, assumptions such as a perfect alignment or equally length of patterns do not hold (Vergara et al., 2013). This, which is due to the dynamic and chaotic nature of gas dispersal, together with the strong dynamics shown by most gas sensor technologies, notably increases the complexity of the classification problem (Monroy et al., 2016).

4.2. Continuous Chemical Classification

The discrimination of gases performed with a robot equipped with an array of gas sensors presents a number of additional challenges when compared to standard identification applications. While standard classification tasks usually host gas sensors inside a chamber with controlled humidity, temperature and airflow conditions, in robotics olfaction there is no control over the sensing conditions. This entails that the sensor signals to be processed are noisy and dominated by the signal transient behavior (Trincavelli, 2011). Under these challenging conditions, chemical recognition can be seen as a particular case of time series classification, characterized by working on sub-sequences of the main data stream (see Fu (2011) for a complete review). Nevertheless, most of these approaches are proposed for univariate time series, while e-nose data is fundamentally multivariate (i.e based on an array of gas sensors with different dynamic responses). This, together with the aforementioned challenges of real data, make most segmentation approaches difficult to apply to e-nose data, which, in turn, affect negatively to the classification rate.

A novel approach was published in Schleif et al. (2016) as a partial result of the IRO project to address the aforementioned issues. This approach is based on generative topographic mapping through time (GTM-TT) and integrates supervised classification and relevance learning (SGTM-TT) to the problem of volatile identification in mobile robotics. By exploiting the strong temporal correlation of the e-nose data, the method is capable to classify gases with high accuracy employing short data sequences (1s, 10s and 20s). Given the ephemeral nature of gas dispersion, it is also analyzed the impact of the data sequence length on the classification performance, trying to push the limits towards a fast-response chemical recognition system. Furthermore, another remarkable advantage for robotics applications is the introduction of a relevance value, by studying the relevance of the different sensors composing the e-nose, and the time points in the data sequence, for predicting the class label (see Fig. 2).

Later, in Monroy et al. (2016) we advocated the use of the well known sliding window approach to avoid feature based segmentation and to study up to which extent considering delayed samples contributes to exploit the temporal correlation of e-nose data. This technique is attractive because it is simple, intuitive, and, moreover, because it is amenable to online applications, which is a primary focus of the IRO project. We analyzed the impact of the window length on the classification accuracy (see Fig. 3) for three state of the art classifiers, a variety of experimental scenarios, e-nose configurations and gas classes (employing three different olfaction datasets). The main conclusion of such work is that, for online chemical classification in uncontrolled environments, feeding the classifiers with additional delayed samples leads to a small, yet important, improvement (up to 6% units) on the classification accuracy.

4.3. Gas Classification in Motion

Having demonstrated that online chemical classification is feasible with a mobile robot, IRO also investigated the impact
of carrying such task while the robot is navigating. We analyzed the induced changes in the gas sensors response and determined that the movement of the robot has an important impact on the classification accuracy if not properly considered, resulting in a decrease of up to 30% in some configurations (Monroy y Gonzalez-Jimenez, 2017). We supported our conclusions with an extensive experimental evaluation consisting of a mobile robot inspecting a long indoor corridor with two chemical volatile sources (ethanol and acetone) more than 240 times, at different motion speeds.

To analyze to which extent the motion of the gas sensing device may affect the classification accuracy, we trained multiple classifiers with samples of each chemical volatile collected in a traditional static setup (i.e. both robot and gas source standing still), and then, analyzed the classification performance for a set of increasing motion velocities. Fig. 4 shows the results of the experiments from which a noticeable reduction in the classification accuracy is observed when increasing the motion speed. This confirms our suspicions about the negative impact that the motion speed of the robot has over classification rate.

To overcome, to a certain degree, the aforementioned effect, we also analyzed the classification accuracy when the classifier is also trained with in-motion data samples, proposing different training schemes. We showed that training a classifier with data collected in motion yields, on average, more accurate outcomes (see Fig. 4 (right)) than using a static setup (Fig. 4 (left)). Moreover, we found that it is not necessary to train the classifiers with data gathered at the same speed than the testing data to remove this negative correlation, but it suffices to capture the underlying dynamics. As a general conclusion, the absolute speed is not a determinant parameter, but the gap between the speeds used to collect the training and testing datasets is an aspect to be taken into consideration when deploying real olfaction applications with a mobile robot.

5. Object recognition and Semantic Knowledge for Robotic Applications

From the object recognition side, the peculiarities of the acquisition process of visual data by a mobile robot permits the inspection of larger portions of the robot workspace, gathering rich semantic information. In this case, semantic information comes in the form of contextual relations, i.e. objects that are found according to certain configurations: keyboards are usually in front of computer screens, microwaves are in the same room as refrigerators, tables are typically surrounded by chairs, etc (Galleguillos y Belongie, 2010). Thereby, during the object recognition process, the presence of a refrigerator in a room

Figure 2: Illustration of the sensor relevance (upper row) estimated for four gas classes and time points relevance profile (bottom row) averaged over all classes (left) and mean prediction accuracy over time for window length’s of = 1, 10, 20 sec (right). These results corresponds to an e-nose dataset collected under semi-controlled measurement conditions.

Figure 3: Average classification accuracy of a naive Bayes classifier for different lengths and positions of the sliding window within the time-series e-nose data.
helps to disambiguate the classification of a white, box-shaped object as a microwave and not as a night stand (Ruiz-Sarmiento et al., 2016; Oliva y Torralba, 2007).

To exploit these contextual relations in the IRO project we make use of Conditional Random Fields (CRFs), a model from the Probabilistic Graphical Models (PGMs) family (Koller y Friedman, 2009), and combine them with Ontologies (Uschold y Gruninger, 1996) to achieve a more robust performance. CRFs represent the objects in the environment as nodes in a graph, where edges are used to link contextually related objects (see Fig. 5). In (Ruiz-Sarmiento et al., 2017c) a survey on different learning approaches for these models is presented, performing a comparative analysis focusing on the time needed for training and the achieved recognition accuracy. This analysis is especially targeted at finding the most suitable one for scene object recognition, providing Loopy Belief Propagation (LBP) the best results (Murphy et al., 1999). These comparisons were done with two state-of-the-art datasets, including a particular one, called Robot@Home one (Ruiz-Sarmiento et al., 2017b), specifically conceived to serve as a testbed for the evaluation of semantic mapping algorithms, mainly those exploiting contextual information.

In order to combine different sources of contextual information, novel environment representations can be used such as the so-called Multiversal Semantic Map (Ruiz-Sarmiento et al., 2017a). This map is an extension of traditional semantic maps for robotics (Galindo et al., 2005), with the ability to coherently manage uncertain information coming from, for example, object recognition or gas classification processes, and reference them to the location where they were acquired into a metric map. Additionally, it also comprises semantic information codified by means of an Ontology, enabling the execution of high-level reasoning tasks (Kostavelis y Gasteratos, 2015), which are of special interest in this project.

6. Exploiting High-level Olfactory and Visual Semantic Information in the Planning and Execution of Tasks

Mobile robots operating in human environments like offices, hospitals, or factories benefit from the fusion of different sensing modalities to efficiently accomplish tasks that are hard or even unfeasible to address if only one sensor is employed (Kam et al., 1997). As mentioned, the IRO project we focus on two of these modalities, namely vision and artificial olfaction, and study their application to a challenging problem: the localization of gas emission sources within real-world indoor environments, commonly referred as gas source localization (GSL) (Kowadlo y Russell, 2008). For that, the robot would need not only to detect the volatile chemical substance that is being release, but also pinpoint the location of its release source.

Figure 4: Average classification accuracy for different motion speeds when: (left) the classifiers have been trained with static data samples, and (right) the classifiers have been trained with data collected in motion.

Figure 5: Left, Scene from the NYUv2 dataset with segmented patches and their ids ($s_1, \ldots, s_{12}$). Right, Conditional Random Field (CRF) graph built according to the patches in the NYUv2 scene (the node and relations of the wall, $x_{i0}$, have been omitted for clarity). The orange area illustrates the scope of a pairwise factor modeling the relations between two objects, while the blue one stands for the scope of an unary factor classifying an object according to its features. Black boxes represent the expected results from a probabilistic inference process over such CRF.
As stated, enriching the search process with visual sensory information and considering semantic relationships through an inference process will enhance the current state of art of GSL algorithms.

To demonstrate this claim two parallel approaches were considered: on the one hand, we relied on human intervention by means of a teleoperated mobile platform (Monroy et al., 2017b), delegating the inference of the most likely source location to the human tele-operator, and, on the other hand, we developed a fully autonomous system able to infer the most likely source location based on the sensory data available on the robot and high-level semantic reasoning Monroy et al. (2018a). Both approaches are detailed in the following sections.

6.1. Olfactory telerobotics

Since inferring the type of object (and the location in the environment) of the gas source which is releasing the gases that have been detected by the robot is not straightforward, we simplified the problem by introducing the human factor and its powerful reasoning capabilities to solve this challenging problem (Gongora et al., 2017). In this context, olfactory telerobotics can be seen as the augmentation of the sensing capabilities of a conventional teleoperated mobile robot to acquire information about the surrounding air (i.e., gases, wind-speed, etc.) in addition to the usual audio and video streams (see Fig. 7).

To evaluate whether the human reasoning can be exploited through a teleoperated robot to efficiently locate the gas source, we collected a dataset comprised of 60 GSL experiments with a teleoperated mobile robot (Gongora y Gonzalez-Jimenez, 2019). The goal of the human operators was to identify and locate the gas source among several visually-identical candidate objects (see Fig. 8). Results demonstrate that humans had over 75% success rate for search times between three to four minutes, supporting our hypothesis that semantic reasoning is indeed used by humans when locating the gas source with this configuration.

6.2. Semantic-based Autonomous Gas Source Localization

The use of visual information when locating a gas source is not a novel approach, yet, it has been only superficially explored in the literature with very simple problem domains where the robot exploited prior knowledge about the source physical characteristics to reduce the locations to search (Ishida et al., 2006). Moreover, a formal way to define and exploit the relationships among gases and objects (i.e., their semantics) it is still missing, aspect which could assist the GSL process in a more flexible way. In Monroy et al. (2018b), as a partial result of the project, presented a novel GSL system that pursues both efficiency by exploiting the semantics between the detected gases and the objects in the environment, and coherence through the consideration of the uncertainty in the identification of gases and objects. To encode these semantic relationships (e.g. that heaters can release smoke), we rely on an ontology (Uschold y Gruninger, 1996). These factors makes this approach particularly suitable for structured-indoor environments containing multiple objects likely to release gases where semantic relationships can be exploited.

Fusing the classification results (from both the detected gases and the recognized objects in the environment) together with the semantic information, a probabilistic Bayesian framework is proposed to assign to each detected object a probability of being the gas source. Finally, a path planning algorithm based on Markov Decision Processes (MDP) merges these probabilities with the navigation distances from the current robot location to the different objects (i.e., a cost value related to the time the robot would spend to reach the candidate object), to produce a plan that minimizes the search time. Both simulated (using computational fluid dynamic tools and GADEN gas dispersion
simulate. Monroy et al. (2017a) and real experiments demonstrate the feasibility of this novel approach by considerably reducing the search times and producing more coherent gas source searches.

7. Conclusions

In this paper we have described and reviewed the goal and main contributions of the IRO project, focused on the improvement of the sensory and autonomous capability of mobile robots through olfaction.

We have first reviewed the concept of electronic nose, rising some specific issues when used on-board a mobile robot, and described a design of a modular e-nose suited for mobile robotics applications. Then, having in mind the final goal of fusing different sensing modalities, we have focused on the intermediate tasks of visual object recognition and gas classification. Here, the project contribution consists of different algorithms and experimental evaluations towards improving the recognition rates when these tasks are carried out with a mobile robot while navigating.

Finally, we have introduced semantic reasoning to successfully fuse multiple sensing modalities when solving the challenging problem of gas source localization with a mobile robot. In this point, the project contributes with a novel architecture able to exploit the information provided by the vision and olfaction sensory sub-systems, as well as handling their respective uncertainties. For each detected object in the environment, a probability of being the gas source is estimated and afterward fed to a probabilistic framework that outputs the optimal path the robot should follow when inspecting the different objects in the environment, minimizing the search time.

Acknowledgements

This work has been funded by the Andalusia Regional Government and the European Regional Development Fund (ERDF) under project TEP2012-530.

References


DOI: 10.1016/j.jpba.2014.08.030


DOI: 10.1016/j.engappai.2010.09.007


DOI: 10.1016/j.robot.2012.12.007


DOI: 10.1109/IROS.2005.1545511


DOI: 10.1016/j.cviu.2010.02.004


DOI: 10.1016/S0196-6464(84)80517-X


DOI: 10.1016/j.robot.2018.02.008


DOI: 10.1109/ISOEN.2017.7968999


DOI: 10.3390/s110606145


DOI: 10.1021/acssensors.6b00008


DOI: 10.1109/TBME.2010.2055864


DOI: 10.1109/3477.790446


DOI: 10.1016/j.pmcj.2014.11.008


DOI: 10.1152/jappl.1998.84.2.401


DOI: 10.1007/s10462-004-9037-2


DOI: 10.1109/TRO.1997.554212


DOI: 10.1016/j.robot.2014.12.006

DOI: 10.1177/027836490809518
DOI: 10.1007/s12369-013-0178-y
DOI: 10.1016/S0040-6090(02)00951-X
DOI: 10.1016/j.snb.2016.09.013
DOI: 10.4018/978-1-5225-3862-2.ch012
DOI: 10.3390/s121013664
DOI: 10.1109/ICSENS.2014.6985265
DOI: 10.3390/s17071479
DOI: 10.1016/j.snb.2013.06.053
DOI: 10.1109/ROMAN.2017.8172427
DOI: 10.1016/j.chemolab.2016.08.011
DOI: 10.1007/978-3-319-91479-4_42
DOI: 10.3390/s18121474
DOI: https://doi.org/10.1016/j.tics.2007.09.009
DOI: 10.1016/j.envsoft.2004.04.012
DOI: 10.1021/cr068121j
DOI: 10.1007/978-3-319-48799-1_1
DOI: 10.1016/j.knosys.2016.12.016
DOI: 10.1177/0278364916695640
DOI: 10.1016/j.ijar.2016.10.009
DOI: 10.1109/ICSENS.2014.6985418
DOI: 10.1007/s10244-014-0442-2
DOI: 10.1371/journal.pbio.0020146
DOI: 10.1515/kien.0020146
DOI: 10.1016/j.snb.2009.03.018
DOI: 10.1109/TCDS.2017.2751963
DOI: 10.1017/s0269888900007797
DOI: 10.1016/j.snb.2013.05.027