Real-time gas recognition and gas unmixing in robot applications

Pierre Maho, Cyril Herrier, Thierry Livache, Pierre Comon, Simon Barthelme

Real-time gas recognition and gas unmixing in a robot application

Pierre Maho[a], Cyril Herrier[b], Thierry Livache[b], Pierre Comon[a], Simon Barthelmé[a]

[a] *CNRS, GIPSA-Lab, Univ. Grenoble Alpes, F-38000 Grenoble, France*
[b] *Aryballe Technologies, 38000 Grenoble, France*

Abstract

Robot olfaction takes inspiration from animals for locating an object in the environment. This object can be a gas leak to fix or an explosive to neutralise. In these cases, the objects will emit Volatile Organic Compounds (VOCs) which can be measured with an electronic nose. This instrument has the great advantage of being able to detect a broad variety of VOCs, so the same device can then be used for a lot of different applications. In a realistic environment, several VOCs of interest can be present at the same time and mix. This creates difficulties for gas recognition, and in the literature, the problem is often ignored. In this article, we deal both with gas recognition of a large number of VOCs and gas unmixing. For that, we use a recently developed optoelectronic nose which uses peptides as sensing materials and Surface Plasmon Resonance imaging as transduction method. We present two different setups. The first setup studies the recognition of 24 gas sources of 12 VOCs disseminated in the environment. The second setup studies various realistic scenarios in which mixtures occur, due to gas sources being spatially close. We propose a real-time dictionary-based algorithm for dealing both with gas recognition and gas unmixing. We succeed in obtaining a score of 73% for the gas recognition task. For the unmixing issue, we reach at least 72% but we also show that this performance is strongly related to the VOCs used in the dictionary.

Keywords: robot olfaction, gas unmixing, gas recognition, electronic nose, Surface Plasmon Resonance imaging

1. Introduction

In nature, olfaction is a key sense used by most species of animals for locating an object in the environment. This object can be a danger to avoid, a partner for breeding or simply some food. Two main challenges are then part of any source localisation problem with olfaction. The first one is odor recognition which is essential for identifying the source of interest and for discriminating between different compounds. Odor recognition is a complicated task since an odor can be present at different levels of concentration and can also be mixed with other odors, of interest or not. The second issue is to define a movement strategy to easily and quickly locate the source. The optimal strategy can be hard to find especially due to obstacles present in environment and to variations in the wind direction.

Robot olfaction is a recent field which takes inspiration from animals (Rochel et al., 2002). To name a few examples, robot olfaction can be used for gas leak detection or demining (Ishida et al., 2012). In most cases, these applications are based on the recognition of given Volatile Organic Compounds (VOC) present in environment. This recognition can be achieved thanks to gas sensors specific to the targeted VOCs (*e.g.* Mead et al. (2013)). However, Marques et al. (2002) demonstrated that the use of an array of weakly-specific gas sensors, namely an electronic nose (see Röck et al. (2008) for a detailed definition), can also be an interesting alternative approach. In this study, we focus on this alternative class of instruments.

Electronic noses (eNoses) have been intensively studied in controlled settings in which temperature, humidity and VOC concentration are kept constant. In contrast, an olfactory robot will continuously measure VOCs with time-varying concentrations guided by diffusion and advection, and that under unstable environmental conditions (Trincavelli, 2011). In a realistic environment, several different gas sources can be present, only one of which is of interest. This makes the task of gas recognition fundamental. In addition, gases can mix, which leads to the quite complicated task of gas unmixing.

In the literature, few works have been devoted to gas recognition with multiple different VOCs in uncontrolled environments (Mcgill and Taylor, 2011) and most of them restrict their studies to 2 gas sources. Even in this case, one major assumption is that gas sources are sufficiently far away from each other to avoid interaction, meaning that there is no gas mixture (*e.g.* Fan et al. (2019)). This assumption is quite unrealistic in real-life applications.

In this work, we tackle two issues, namely gas recognition and gas unmixing, with a new optoelectronic nose (also abbreviated as eNose). This eNose is based on the use of peptides as sensing materials and on the use of Surface Plasmon Resonance imaging (SPRi) as the transduction method (Brenet et al., 2018). This instrument is quite interesting as it boards around 20 different sensing materials. This number is much higher than in other systems and is crucial for the unmixing issue. For instance, in the

linear case, the response to a mixture is just a weighted sum of the individual responses. In this case, the number of sensing materials must often be greater than the number of mixed VOCs to ensure the recovery of their concentrations (Comon and Jutten, 2010).

We have built two different open sampling systems, placed in an indoor environment with low advection. One setup is used for assessing the recognition of 12 different VOCs disseminated in 24 isolated gas sources. The other setup is used for testing the unmixing of binary and ternary VOCs mixtures. These mixtures are generated due to either the proximity of two isolated gas sources or to the succession of scented trails. We proposed a real-time processing pipeline, able to deal both with gas recognition and with gas unmixing. This pipeline proceeds in two steps. The first block deals with the baseline drift and the second one identifies the VOCs present using a dictionary.

This article starts with a short review of the main studies which deal with gas recognition and gas unmixing in an uncontrolled environment. Then, the experimental section presents the eNose and the different setups and scenarios studied here. The real-time processing pipeline is presented afterwards as well as the classification criterion used for quantifying the results. Next, results start with the recognition of the 12 VOCs and show a classification score of around 73% for the proposed method (chance level at around 8 %). Finally, we present the results for gas unmixing in various scenarios and with a dictionary based on the 12 VOCs previously studied. Classification scores show good unmixing in this complicated environment.

2. Related work

In this work, gas recognition is defined as the recognition of multiple gas sources (*i.e.* at least 2 different VOCs) present at the same time in an open environment. As a consequence, any study that only considers a single VOC source or any study with multiple VOCs but presented in separated experiments is not detailed here. Compared to studies that use controlled sampling, gas recognition in an open sampling system is a challenging task. This is due to the changes in environmental conditions (humidity, temperature) and to the fluctuations of gas concentrations which prevent the instrument from reaching a steady-state (Martinez et al., 2006). Finally, as electronic noses are likely to be weakly-specific devices, the use of highly-specific instruments with low cross-sensitivities to the target VOCs is out of the scope of this article (see for example the urban air quality study of Mead et al. (2013)).

The literature on robot olfaction describes different algorithms which are assessed with different setups which are not comparable amongst themselves (Mcgill and Taylor, 2011). In this section, we decided to split gas recognition in open sampling systems into two main types of work: first, the works assuming that the gas sources are far from each other and second those dealing with mixtures. We already emphasize that the last one is an application much more rarely considered in the literature. Although it is the more realistic, it is also the more challenging.

2.1. Recognition of isolated gas sources

One of the earliest works is probably the one of Loutfi et al. (Loutfi et al., 2005; Loutfi and Coradeschi, 2006). They studied the recognition of VOCs, up to 5, which were placed in cups present in the environment. However, their approach was based on a classical 3-phase sampling (baseline acquisition, VOC injection and recovery). This approach implies that one waits for MOX sensors to recover, which can easily take 2-5 mins.

Then, much effort has been dedicated to use only transient responses, meaning only the measurement $\boldsymbol{y}(t)$, but at the cost of a limited number of gas sources. Loutfi et al. (2009) placed two gas sources in two long corridors and proposed a supervised kernel-based approach for mapping the gas concentration of each VOC. This algorithm has been further improved by several studies from the same group (Hernandez Bennetts et al., 2014; Fan et al., 2018, 2019). Monroy et al. (2016) examined the use of time windows instead of single time points. They concluded that algorithms could be improved by taking advantage of the temporal correlation of samples in an uncontrolled environment. In the same vein, Schleif et al. (2016) proposed an algorithm to deal with short time sequences, called generative topographic mapping through time (Bishop et al., 1997). They validated their approach with a robotic arm and four different chemical substances which were presented sequentially.

Another line of research, in the absence of gas mixtures, is the examination of the effects of external parameters on classification accuracy. Palacín et al. (2019) studied the case where heating, ventilation and air conditioning were activated or not. They concluded that these parameters can help to locate gas leaks in open environments. Monroy and Gonzalez-Jimenez (2017) examined the influence of robot speed for the discrimination of 2 gas sources. By using MOX gas sensors, they particularly emphasize a loss in classification accuracy, up to 30%, when the motion speed strongly differs between the training samples and the testing. Vergara et al. (2013) made similar conclusions but regarding wind effects. They have considered the recognition of 10 VOCs with different wind speeds under tightly-controlled conditions (pressure, humidity, temperature, concentration).

An interesting conclusion of this short review is that most of the works consider only the case of a very limited number of different gas sources, which can be restrictive in practice.

2.2. Gas unmixing

The paper of Hernandez Bennetts et al. (2014) is probably one of the most realistic applications in the field. They considered 2 gas sources, separated by a small distance. They proposed an improved version of an existing

kernel-based algorithm (Loutfi et al., 2009) which can deal with mixtures with the help of a PhotoIonization Detector. However, few experiments were carried out and one of them concluded to the low probability of one VOC just near its source location, which is counterintuitive. The authors emphasized the difficulty of obtaining a ground-truth in such scenarios.

Most existing works concerning gas unmixing in uncontrolled environment rely on *mixtures-learning*. Mixtures-learning means that authors generate all possible mixtures with the VOCs of interest, measure the instrument's response, and then learn (*e.g.* with a neural network) from the measured responses to predict possible responses on novel mixtures. Marques et al. (2002) built a setup in which an Ethanol source was disturbed by a Methanol source in a turbulent regime. Their solution was then based on the training of a neural network for locating the Ethanol source. Fonollosa et al. (2014) also studied turbulent gas mixtures and proposed to use an Inhibitory Support Vector Machine (Huerta et al., 2012). MOX gas sensors were placed in a wind tunnel and the goal was then to identify the presence of Ethylene when an interfering volatile was present. In another work, Fonollosa et al. (2015) studied the composition of binary mixtures with an artificial neural network called Reservoir Computing (Maass et al., 2002). The sensors were placed in a 60 mL measurement chamber and they continuously injected random mixtures of two compounds (including "pure" mixtures) with random transition times. The approach based on Reservoir Computing performed a better estimation of the true concentration of each compound compared to simpler methods, namely a linear regression and a Support Vector Regression. An implicit drawback of a mixtures-learning approach for gas unmixing, is that it requires the generation of a lot of different mixtures for training the algorithm, ideally all the possible combinations of the studied VOCs. It can be manageable when one studies only binary mixtures but it rapidly becomes impossible when one deals with more complex mixtures. For instance, if we assume that we have R VOCs in the environment which can all mix in different proportions (including a null concentration), then we have to generate $\Omega = 2^R - 1$ mixtures to be sure that we can recognise any mixture of these R VOCs. For $R = 2$ as in the article of Fonollosa et al. (2015), we have to manage "only" 3 different kinds of mixture (2 pure, 1 real mixture) at different concentration levels. But for $R = 3$, it goes to 7 which can be more challenging regarding the experimental practicability. If we continue, for $R = 4$, we have to manage $\Omega = 15$ kinds of mixtures, for $R = 5$, $\Omega = 31$, etc...

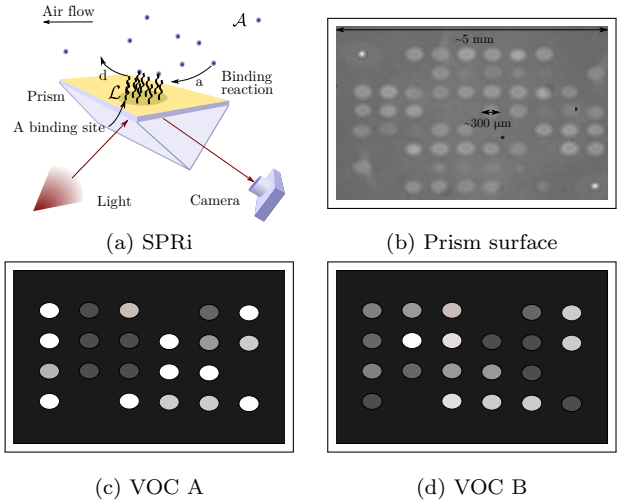
We believe that a more valuable algorithm, easier to use in practice, would deal only with the pure components to build a dictionary and try to unmix any mixture from this dictionary. In this way, only R experiments have to be generated which considerably reduces the experimental time (and human factors at the same time).

As a conclusion, little work has been done concerning the use of an electronic nose for dealing with mixtures in an uncontrolled environment.

3. Experimental

3.1. Electronic nose

3.1.1. SPR-based optoelectronic nose



(a) SPRi (b) Prism surface

(c) VOC A (d) VOC B

Figure 1: (a) Working principle of the optoelectronic nose based on SPRi. Here, only one sensing material is represented. (b) Raw image of the prism surface with some dimensions. Light areas stand for the functionalized surface. (c-d) Schematic representation of the images obtained for two different VOCs A and B.

The optoelectronic nose used in this study is the commercial version of the one described in Brenet et al. (2018). The instrument is provided by the company Aryballe Technologies. Sensing materials are mainly peptides which are fixed on the gold surface of a prism. During an acquisition, the VOC is brought above the gold surface by a flow of air (in this study, at 60 mL/min). The VOC can then interact with the sensing material through a reversible binding reaction. This reaction is both dependent on the VOC and on the sensing material. Thus, different sensing materials will lead to different chemical reactions, creating a chemical "signature" of the VOC. Since different VOCs lead to different chemical reactions, and thus different signatures, we are able to recognise VOCs. Here, the instrument uses 19 different sensing materials which are replicated 3 or 4 times on the surface, leading to an array of $P = 59$ chemical sensors, which is much higher than in most existing studies.

The binding reactions at the surface are measured using Surface Plasmon Resonance imaging (SPRi). Briefly, light is sent, reflected by the surface and caught by a simple optical camera. When a binding reaction occurs with the VOC, this changes the refraction index (more light is reflected). The changes in reflectivity are caught by the camera, which thus records in real-time the binding reactions. A representation of the working principle is presented in Figure 1a. A real image of the prism surface

is reported in Figure 1b and two synthetic images of 2 differents VOCs are shown in Figure 1c-d.

We stress that this instrument does not target specific VOCs. In fact, it can generate a signature for a broad variety of compounds. For example, Brenet et al. (2018) showed that this instrument can recognise different molecules from different chemical families (alcohols, esters, ketones, ...) but also compounds with similar molecular structures (down to only one carbon atom different).

3.1.2. Main advantages

Most existing work in robot olfaction uses MOX gas sensors (Loutfi et al., 2009). Here, we use another technology, so we deem necessary to point out some of the benefits.

First, the number of chemical sensors is large compared to other robot studies. In MOX gas sensors, this number generally reaches a maximum 5 or 6 elements, replicated or not. Here, we have four times more sensors since we have 19 different molecules on the sensing surface (replicated 3 or 4 times). It is well known that, when one deals with gas mixtures, the number of sensors is a crucial parameter when one wants to identify the components of mixtures (Comon and Jutten, 2010).

Second, this instrument demands much less power than a MOX-based system. For instance, let us consider a MOX gas sensor such as the Figaro TGS 2600 which is often used in the literature for detecting carbon monoxide or hydrogen. If only the power consumption for the heater part (which is the most power demanding by far) is considered, then according to the product's data sheet power consumption reaches 210mW (Figaro, 2013). By doing a simple multiplication and assuming that we want 60 replicates of this sensor for our robot, we reach 12.6 W. By comparison with our instrument for which the working principle only needs a camera and a LED, we obtain 1.56 W (these 2 numbers do not take into account the on-board electronic system such as the pump which creates the flow of air). Thus, this instrument reduces the power consumption by a factor $\sim$8 which is absolutely not negligible in a mobile robot application for which battery life is crucial. In addition, with this instrument, we can easily add new sensors to the prism surface without increasing system complexity (Brenet et al., 2018), meaning that we can double the number of sensors without changing power consumption.

Third, the instrument has very good selectivity and sensitivity. Sensitivity varies depending on the compound (much like a biological nose), but Brenet et al. (2018) found a sensitivity in the range of ppb for some VOCs. Selectivity is also very high, as evidenced by data from Brenet et al. (2018), who managed to discriminate compounds that differ by a single carbon atom. We note however that these experiments were carried out under controlled conditions (fixed temperature and humidity, controlled three-phase sampling), and performance in more difficult settings may not be as high.

VOC	Formula	Molar mass ($g.mol^{-1}$)	P_{sat} (mbar)	Chemical family
(S)-Limonene	$C_{10}H_{16}$	136.2	2.1	Alcene
β-pinene	$C_{10}H_{16}$	136.2	3.2	Alcene
Allyl-hexanoate	$C_9H_{16}O_2$	156.2	0.91	Ester
Geranyl acetate	$C_{12}H_{20}O_2$	196.3	0.035	Ester
Butanol	$C_4H_{10}O$	74.1	8.9	Alcohol
Cis-3-hexenol	$C_6H_{12}O$	100.2	1.4	Alcohol
Linalool	C_7H_6O	154.2	0.021	Alcohol
Benzaldehyde	C_7H_6O	106.1	1.7	Aldehyde
Trans-2-octenal	$C_8H_{14}O$	126.2	0.73	Aldehyde
Citral	$C_{10}H_{16}O$	152.2	0.12	Aldehyde
Acetic acid	$C_2H_4O_2$	60.0	21	Acid
Guaïacol	$C_7H_8O_2$	124.1	0.24	Phenol

Table 1: List of the VOCs used and some of their properties (the vapor pressure P_{sat} has been estimated at 25°C).

Finally, a common concern with eNoses is stability over the medium and long term: are signatures stable enough that VOCs can be reliably recognise over weeks or months? Using a mobile robot, Maho et al. (2019) ran repeated measurements of three different VOCs over several months and under different environmental conditions. The signatures did indeed drift over time, but a standard correction method (Artursson et al., 2000) is enough to stabilise them. The results show that the signatures used for recognition can be used without re-training over a period of several months.

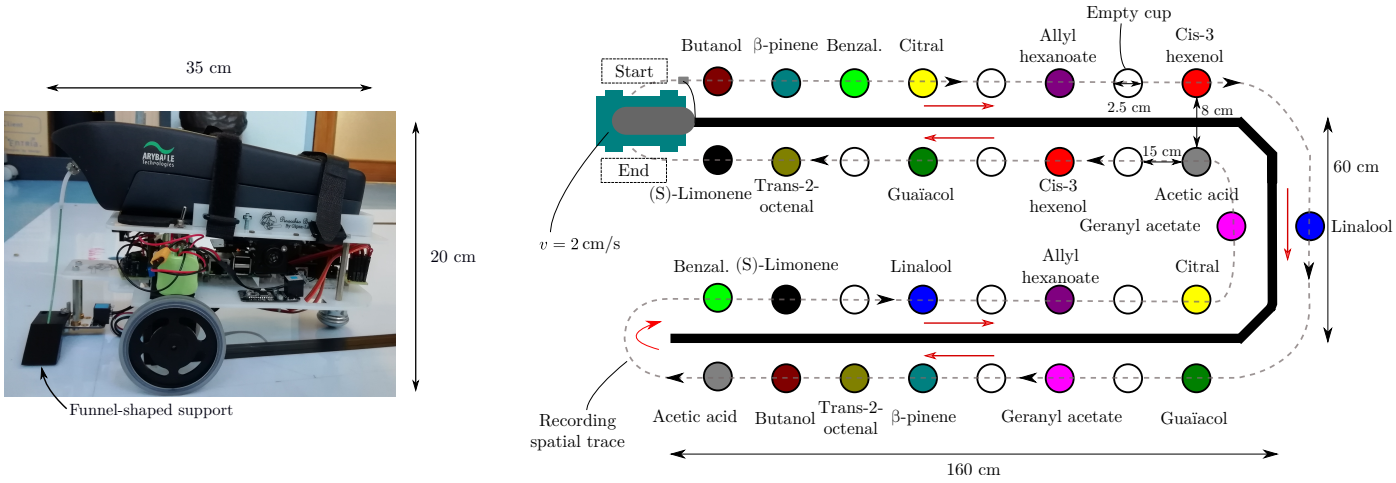
All these characteristics match the desirable attributes needed for a robot application described by Russell (2001).

3.2. Experimental setups

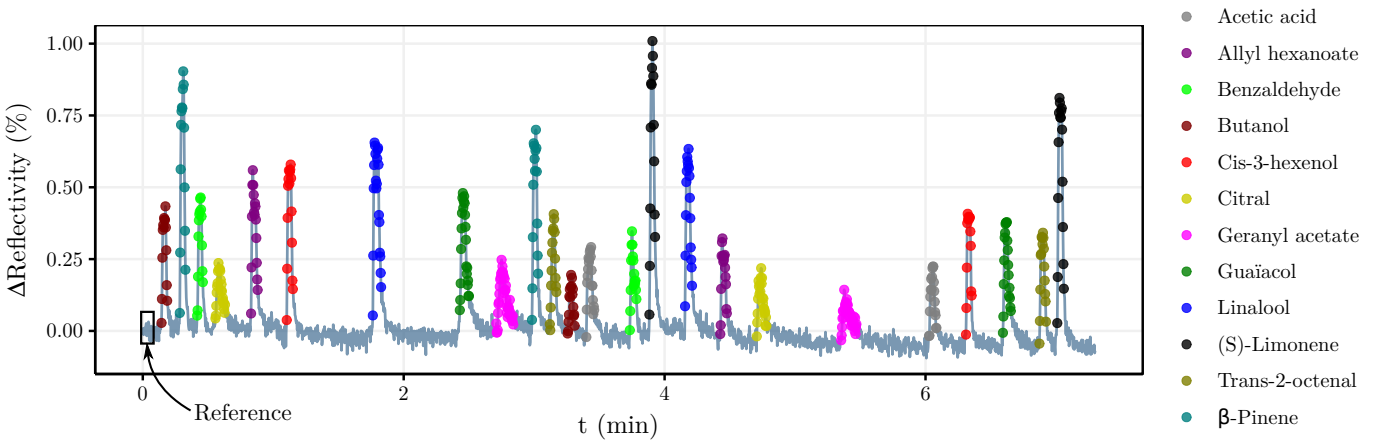
In this article, we present two different open sampling systems. The first one lets us perform gas recognition of a large number of VOCs, up to 12, in an uncontrolled environment with low advection. The second one is more flexible and can generate different scenarios in which gas mixtures occur. A desirable property shared by these 2 setups is that they can produce large datasets without requiring much human intervention. The many repeated measurements ensure good statistical reliability in this study. We add that in both cases the path of the robot is always predefined since navigation is not the purpose of this paper.

3.2.1. Setup 1: Sniffer robot

The first open sampling system was already presented in a previous publication (Maho et al., 2019). The platform (see Figure 2a) consists of a robot that carries the previously introduced eNose, moving over a flat surface where gas sources are placed. A funnel-shaped support has been made with a 3D printer and the PEEK (PolyEtherEtherKetone) injection tube of the eNose is inserted in this sup-



(a)



(b)

Figure 2: Sniffer robot platform (a video is available here: http://bit.ly/snifferRobot) (a) The platform and some dimensions. (b) A time series of one chemical sensor (with reference subtracted, meaning the first points) during one lap. The segmentation used in this study is shown by the colored points (each color stands for a VOC).

port in order to increase the suction area. The ground is a 2m × 1m × 2.5mm polycrystal plate which is lifted by 1.5cm. On this plate, a black path is drawn for the robot to follow. Along this path, the plate is pierced at 34 different locations thus enabling to slide below the plate up to 34 gas sources at the same time. Here, gas sources are small cups in which liquid solutions of VOCs are placed. This setup is placed in an indoor environment,basically a normal office with low natural advection (no ventilation system). The liquid solutions (∼250μL) are put in the cups just before the experiment. The speed of the robot is set to 2 cm/s, the frame rate of the camera to 5 Hz and the airflow to 60 mL/min. The frame rate can be increased and the airflow can be decreased but these values are sufficient in practice to measure the chemical reactions and their dynamics.

The 12 VOCs we selected are listed in table 1. This selection has been only based on product availability and on safety, but *not* on whether their signatures are easy to tell apart. Each VOC is repeated twice along the track. The position of each VOC is optimized manually according two criteria: first, we have to limit the desorption of one compound on the next compound and second, two cups of a VOC must have different neighbours. In this way, each gas source is sufficiently distant from its neighbours which limits the mixing but does not completely prevent it. This also explains why some cups are left empty in Figure 2a and why "*only*" 24 cups out of 34 are really occupied by a VOC. The path is repeated 28 times (the total duration of the experiment is then 3h and 44 min), so for each VOC we have at most 56 peaks (such the ones represented in Figure 2b).

Finally, a raw time series (with reference subtracted) of one chemical sensor is represented in Figure 2b. A segmentation has been done to highlight each VOC (different color) and to make the comparison easier with the setup

in Figure 2a.

3.2.2. Setup 2: Sniffer arm

The second setup is used to evaluate different realistic scenarios that can be encountered in a robot olfaction application.

The main part is an aluminium trapezoidal shaft that moves the eNose along one dimension. The total length of this path is 36.5cm. With this setup, we can perform sweeps multiple times, from left to right and vice versa. To save space, we detail only the data from one direction of sweep, but the rest is presented in the Supplementary Materials (the two directions are highly similar). Again, a funnel-shaped support is added to increase suction area. Gas sources are now scent strips, directly placed along the 1D-path, on which few drops are deposited. The speed of arm movement is set to 1 cm.s^{-1}.

In the literature on gas recognition in an uncontrolled environment, authors generally work only with pure compounds which are far from each other in order to avoid mixing. We believe that these restrictions are not realistic in the field. That is why we focus in this study on binary, or ternary, mixtures. To this end, we propose different realistic scenarios with increasing complexity:

Scenario ①, Figure 3b: isolated gas sources are spatially close which leads to gas mixtures. We consider two scent strips which are separated either by 3 cm or by 1 cm. The two gas sources are Citral and (S)-Limonene. One drop of their liquid solutions ($\sim$50 μL) is deposited on the scent strips. The averaged ΔReflectivities of 20 sweeps on the Figure 3b clearly indicates that the 2 gas sources mix. Unsurprisingly, the amount of mixing depends on the distance separating the two gas sources.

Scenario ②, Figure 3c: gas sources are no longer isolated but trails. We use three scent strips (5 cm long) which are placed one after the other. The first trail contains Citral, the second contains (S)-Limonene and the last one contains Guaïacol. Three drops of their liquid solutions ($\sim$150 μL) are deposited on the scent strips. The averaged ΔReflectivities on the Figure 3c of 20 sweeps clearly demonstrate the complexity of the task in which transitions between VOCs are unclear.

Scenario ③, Figure 3d: gas sources are again trails but no longer pure compounds. We consider three scent strips (5 cm long) which are placed one after the other. The first trail contains Citral and Cis-3-hexenol, the second contains Citral and (S)-Limonene and the last one contains (S)-Limonene and Cis-3-hexenol. For a given scent strip, three drops of each liquid solution ($\sim$150 μL) of the 2 compounds are deposited on the scent strip. Thus, the mixtures operate both in the gas phase and in the liquid phase. Again, the averaged ΔReflectivities on the Figure 3d of 20 sweeps clearly demonstrate the complexity of the task which requires unmixing algorithms that can deal with ternary mixtures.

In each one of these scenarios, the experiments involve placing the scent strips, depositing the compounds and doing several sweeps (N=20). At the end of each sweep, the arm stays in place for 20 sec. The response of the eNose during the sweeps is recorded continuously (without any interruption between each sweep).

The spatial scale of this setup is limited compared to other experimental setups that can be found in the literature. However, in large-scale setups it is very difficult to obtain a ground-truth (Hernandez Bennetts et al., 2014). Indeed, nobody can tell if the results of the proposed algorithm far from the gas sources are valid or not since no one can tell exactly the proportion of a VOC at these distances in a realistic environment. Thus, we consider that a small-scale setup is completely appropriate since we would not have more information with a larger one.

4. Data analysis

In this section, we describe the proposed real-time algorithm and the quantitative criteria we used to assess the performance of our algorithm.

4.1. A real-time unmixing algorithm

We propose a single processing pipeline which we use in all the scenarios previously introduced. It can be used for the recognition of isolated gas sources (as in Setup 1, described in Section 3.2.1) and for the unmixing of binary and ternary mixtures (as in Setup 2, see Section 3.2.2).

Recently, Fan et al. (2019) highlighted the need for real-time algorithms in a robot olfaction application such as emergency response scenarios. Real-time algorithms must be of low algorithmic complexity to avoid any processing lag, and the processing pipeline described here is designed with that constraint in mind.

Beyond the constraint of low algorithmic complexity, real-time data-processing algorithms must be causal: at the current time point t, we can only use the data acquired up to time t. All the results shown include this constraint [1]. Another direct implication of the real-time constraint is that the whole processing pipeline for the current time point t must have a computation time lower than the sampling period (here, 200 ms). We report computation time below, but need to point out that it depends on the data, on the programming language (here, `R`) and on computer performance (Dell Inc. Latitude XT3, CPU i7-2640M 2.8GHz, Ubuntu 16.04). The timings we report could be greatly reduced, as `R` in particular is an interpreted language that we use for convenience. An order-of-magnitude improvement is to be expected from implementing the algorithms in a compiled language.

To set notation, we note $y_p(t)$ the time series of the chemical sensor p. We note respectively N_t, P and R, the duration of the recording, the number of chemical sensors

[1] Performance could be improved by processing the data in a batch, but then the algorithms would not be real-time anymore

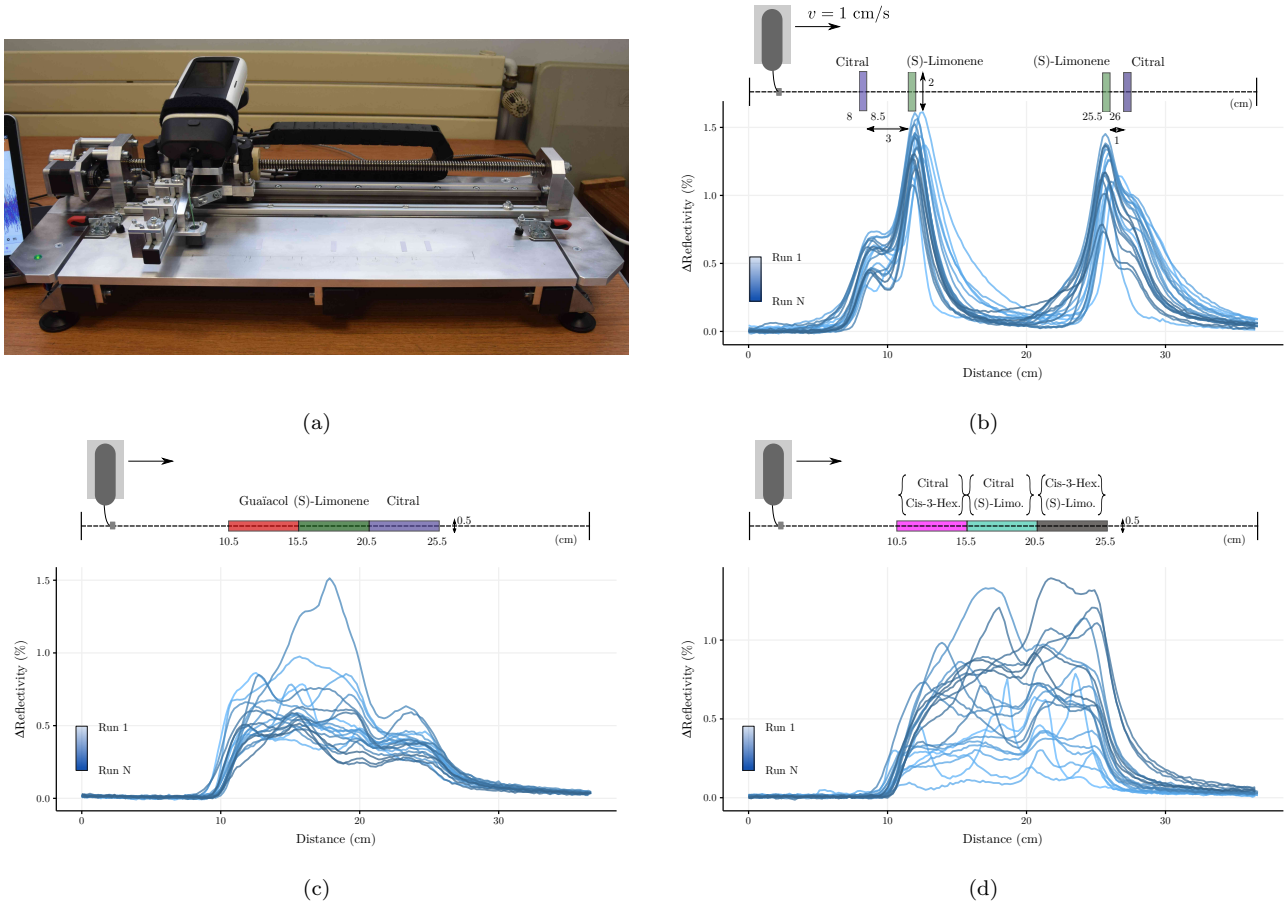


Figure 3: Sniffer arm setup. (a) The setup (a video is available here: http://bit.ly/snifferArm). (b) Scenario ①: isolated gas sources are spatially close. (c) Scenario ②: successive trails of gas sources (pure VOCs). (d) Scenario ③: successive trails of gas sources (binary VOC mixtures). Below each scenario, we represent the averaged ΔReflectivity for each sweep (first sweep is the lighter). More details are given in the text.

Algorithm 1 Real-time estimation of the baseline $b_p(t)$

Require: q, k
  if $t \leq k$ **then**
    $s_p(t) = \{y_p(1), ..., y_p(t)\}$
  else
    $s_p(t) = \{y_p(t-k+1), ..., y_p(t)\}$
  end if
  $\hat{b}_p(t) = \inf\{x \in s_p(t) : q \leq \hat{F}(x)\}$ $\#\hat{F}(x)$ *is the empirical cumulative distribution function of* $\mathcal{Y}_p$ *sampled by* $s_p(t)$

and the number of VOCs. As explained in the introduction, the instrument generates a signature $\boldsymbol{k}_r \in \mathbb{R}^P$ for each VOC r (the extraction process is detailed in Section 5.1). We stack all these signatures into a matrix $\boldsymbol{K} = [\boldsymbol{k}_1, ..., \boldsymbol{k}_R] \in \mathbb{R}^{P \times R}$. The intensities $\boldsymbol{c}(t) \in \mathbb{R}^R$ at time t are stacked into a matrix $\boldsymbol{C} = [\boldsymbol{c}(1), ..., \boldsymbol{c}(N_t)] \in \mathbb{R}^{R \times N_t}$ and measurements are stacked into a matrix $\boldsymbol{Y} = [\boldsymbol{y}(1), ..., \boldsymbol{y}(N_t)] \in \mathbb{R}^{P \times N_t}$.

4.1.1. Baseline-drift correction

Instrumental drift is a long-standing issue in the field of electronic noses, whatever the technology used (Vergara et al., 2012). Here, the drift particularly affects the baseline (reference measurement) in a time range of several minutes as can be seen in Figure 4 for the data from Setup 1 (but the same observation holds for Setup 2). This short-term drift can be explained by several changes over time: environmental conditions (temperature, humidity), heating of the electronic system, reference gas (ambient air is more and more polluted by the evaporation of the gas sources), etc. Here, the difficulty is that VOC injections and baseline drift appear at the same time, even if VOC injections are much shorter over time (∼1 s).

The problem of estimating the baseline drift is similar to the estimation of a smoothly-varying trend in time series. This problem can be encountered in several domains such as economy, chemistry or medicine. The approach proposed here is based on quantile filtering (Hyndman and Fan, 1996) which enables to estimate the trend by avoiding the peaks. Considering an integer k and a scalar $q \in [0, 1]$, the estimation of the baseline $b_p(t) \in \mathbb{R}$ corresponds to the q-quantile of $\{y_p(t-k+1), ..., y_p(t)\}$ (*cf* Algorithm 1).

This non-linear filtering technique is quite fast as it requires only ∼10 ms for a vector $\boldsymbol{y}(t) \in \mathbb{R}^{59}$ ($k = 100$ sec and $q = 0.1$ for the two setups for all the results). For the

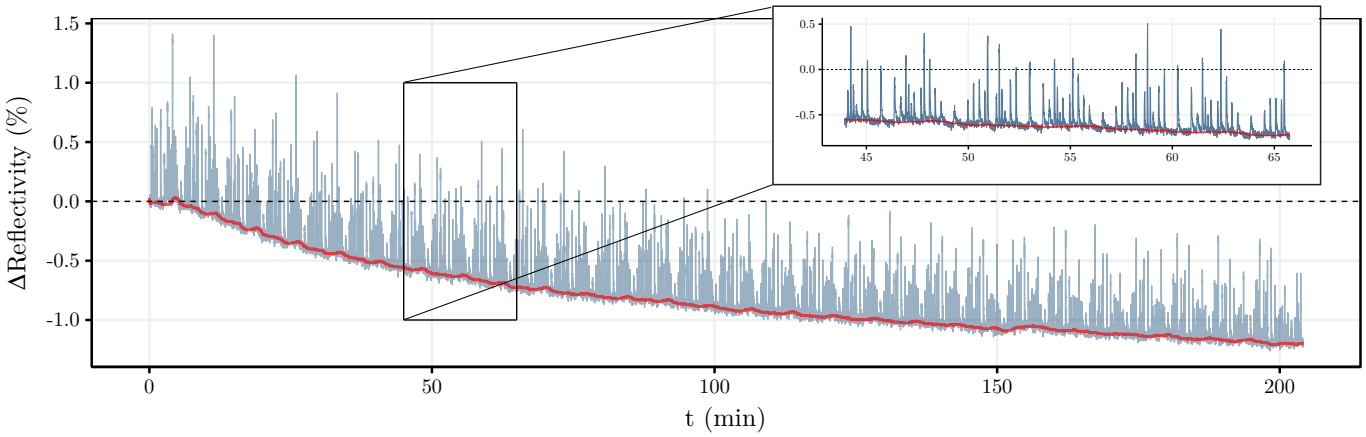


Figure 4: Time series of one chemical sensor (blue) for the entire experiment from Setup 1. In red, we represent the estimation of the baseline drift with Algorithm 1 ($k = 100$ s and $q = 0.1$). The zoom corresponds to 3 laps.

k first time points, we simply reduce the window size to t. The estimation of the baseline for Setup 1 is presented in red in Figure 4.

The size k of the time window and the scalar q can be important, here they have been manually selected for the experiments. Real-time tuning of these parameters is an interesting topic for further research.

In the following, we assume that Algorithm 1 has been applied to estimate $b_p(t)$ and that $\hat{b}_p(t)$ has been subtracted to $y_p(t)$. The notation remains unchanged.

4.1.2. Sparse linear unmixing

In this section, we detail the formulation of an real-time algorithm able to deal both with gas recognition of a large number of pure VOCs and gas unmixing in an uncontrolled environment.

As described in Section 2.2, a solution proposed by some authors is based on mixtures-learning. We already discussed the main drawback of this strategy. In this paper, we propose to use a known dictionary $K \in \mathbb{R}^{P\times R}$ containing all the signatures of the R studied VOCs. This method requires little experimental time (in the training phase), as only pure VOCs have to be measured. The goal is then to estimate a vector $\boldsymbol{c}(t) \in \mathbb{R}^R$ which gives the intensity of each VOC at time t. This implies the formulation of a model relating $\boldsymbol{K}$ and $\boldsymbol{c}(t)$ to the measurement $\boldsymbol{y}(t)$.

The model. The model we formulate is linear, because linearity greatly simplifies the computations. We know from theoretical work (Maho et al., 2018) that linearity cannot hold in general, but holds approximately in a low concentration or low affinity regime. In addition, we used a linear model successfully in an other work (Maho et al., 2019) to normalise signatures for concentration. Another piece of evidence in favour of using linear models is provided in Section 5.1 in which we represent the linear fitting of data of a single VOC.

Given a vector of sensor responses at time t, $\boldsymbol{y}(t) \in \mathbb{R}^P$, we express the linear model both in presence of a single VOC r and of a mixture of VOCs:

$$\begin{array}{cc} \text{Single VOC} & \text{Mixture of VOCs} \\ \boldsymbol{y}(t) = \boldsymbol{k}_r c_r(t) & \boldsymbol{y}(t) = \boldsymbol{K}\boldsymbol{c}(t) \end{array} \quad (1)$$

The model for a single VOC is used in Section 5.1 for the estimation of the signature $\boldsymbol{k}_r \in \mathbb{R}^P$. In the following, we focus on the linear mixing model which simply assumes that the measurement $\boldsymbol{y}(t)$ is a linear combination of the signatures of pure VOCs. Again, we emphasize that this method requires little experimental time since we only need to generate data with pure VOCs for the estimation of $\boldsymbol{K}$.

The optimisation problem. Despite its simplicity, linear unmixing has been successfully applied in many fields such as remote sensing (Bioucas-Dias, 2009) or fluorescense microscopy (Dickinson et al., 2001). A simple way to estimate $\boldsymbol{c}(t)$ is to minimize a least-squares cost function:

$$\hat{\boldsymbol{c}}(t) = \arg\min_{\boldsymbol{c}} \|\boldsymbol{y}(t) - \boldsymbol{K}\boldsymbol{c}\|_2^2 \quad (2)$$

In this case, the solution is easily obtained from the pseudo-inverse of $\boldsymbol{K}$. However, the latter analytic solution may produce negative intensity values, which is not physically possible. To avoid this, we add a non-negativity constraint on $\boldsymbol{c}(t)$. In addition, we do not expect that all the VOCs from the dictionary are present in the mixture (recall that $R = 12$ and that we study ternary mixtures in the worst case). This implies that we may add a sparsity constraint on $\boldsymbol{c}(t)$. Sparsity can help the estimation since $\boldsymbol{K}$ can suffer from an ill-conditioning due to correlated signatures. The sparsity of $\boldsymbol{c}(t)$ can be naturally imposed by the ℓ_0-"norm" which counts the number of non-null elements in $\boldsymbol{c}(t)$:

$$\hat{\boldsymbol{c}}(t) = \arg\min_{\boldsymbol{c}\geq 0} \|\boldsymbol{y}(t) - \boldsymbol{K}\boldsymbol{c}\|_2^2 \text{ s.t. } \|\boldsymbol{c}\|_0 \leq \gamma \quad (3)$$

Here, it implies that at most γ intensity values of $\boldsymbol{c}(t)$

will be non-null (we call *support* $\mathcal{S}(t)$ of $\boldsymbol{c}(t)$ the position of non-null coefficients). However, this optimization problem is NP-hard and the global solution can be reached only in small-scale applications. Indeed, worst-case computation time increases exponentially with the size R of the dictionary. Another approach, which we choose here, is to relax this constraint with the ℓ_1-norm, $\|\boldsymbol{c}\|_1 = \sum_{r=1}^{R} |c_r|$, through a penalty term weighted by $\lambda \geq 0$:

$$\hat{\boldsymbol{c}}(t) = \underset{\boldsymbol{c}\geq 0}{\arg\min} \frac{1}{2P}\|\boldsymbol{y}(t) - \boldsymbol{K}\boldsymbol{c}\|_2^2 + \lambda\|\boldsymbol{c}\|_1 \qquad (4)$$

Finally, we add a last constraint to improve the solution, which we call *support continuity*. Given a time t and a small integer ω, the supports $\{\mathcal{S}(t-\omega), ..., \mathcal{S}(t)\}$ of $\{\boldsymbol{c}(t-\omega), ..., \boldsymbol{c}(t)\}$ must be similar. This constraint implies a weak assumption which is that during a small time frame (defined by ω) the composition of the mixture does not change much (here, $\omega = 1$ sec). This constraint could be formulated as an additional penalty term such as $\mu\|\boldsymbol{c} - \boldsymbol{c}(t-1)\|_2^2$ which would smooth the estimation of $\boldsymbol{c}(t)$ in relation to $\boldsymbol{c}(t-1)$. However, this would require tuning an additional hyperparameter μ. Instead, we propose an heuristic that we detail in the next paragraph. For the rest, we define $\boldsymbol{c}^{\mathcal{S}(t)}(t) \in \mathbb{R}^{\gamma}$ the restricted version of $\boldsymbol{c}(t)$ to the support $\mathcal{S}(t)$, $\boldsymbol{C}_{\omega}^{\mathcal{S}(t)}(t) = [\boldsymbol{c}^{\mathcal{S}(t)}(t-\omega), ..., \boldsymbol{c}^{\mathcal{S}(t)}(t)] \in \mathbb{R}^{\gamma\times(\omega+1)}$, $\boldsymbol{Y}_{\omega}(t) = [\boldsymbol{y}(t-\omega), ..., \boldsymbol{y}(t)] \in \mathbb{R}^{P\times(\omega+1)}$ and $\boldsymbol{K}^{\mathcal{S}(t)} \in \mathbb{R}^{P\times\gamma}$ the restricted version of $\boldsymbol{K}$ to the support $\mathcal{S}(t)$. We define the subproblem:

$$\hat{\boldsymbol{C}}_{\omega}^{\mathcal{S}(t)}(t) = \underset{\boldsymbol{C}\geq 0}{\arg\min} \|\boldsymbol{Y}_{\omega}(t) - \boldsymbol{K}^{\mathcal{S}(t)}\boldsymbol{C}\|_F^2 \qquad (5)$$

where $\boldsymbol{C} \geq 0$ means that all the elements of $\boldsymbol{C}$ are non-negative and $\|\cdot\|_F$ stands for the Frobenius norm.

The optimisation method. The proposed method is divided into two main blocks: first, estimate which VOCs have a non-null intensity in $\boldsymbol{c}(t)$, *i.e.* solve Problem (4), and, second, integrate the support continuity constraint.

Let us assume that we have a measurement $\boldsymbol{y}(t)$ from which we want to identify $\boldsymbol{c}(t)$. The first problem is to find the support $\mathcal{S}(t)$ of $\boldsymbol{c}(t)$ (the position of non-null coefficients). This can be done by solving the optimisation problem in eq. (4), in which the parameter λ influences the degree of sparsity of $\hat{\boldsymbol{c}}(t)$, meaning the estimated number of VOCs in the mixture $\boldsymbol{y}(t)$; the larger λ, the sparser $\hat{\boldsymbol{c}}(t)$. On the other hand, if $\lambda = 0$ then all coefficients of $\hat{\boldsymbol{c}}(t)$ can be non-null and only the data misfit will matter. λ is therefore a crucial parameter in the estimation of the support $\mathcal{S}(t)$. In order to estimate $\lambda_{\min}$ and solve the problem (4) for $\lambda = \lambda_{\min}$, we use the algorithm proposed by Friedman et al. (2010) and the associated `R` package `glmnet`. The algorithm relies on a grid search and a β-fold cross-validation procedure. Given a λ, $\beta \in \mathbb{N}$ folds are generated from $\boldsymbol{y}(t)$, then β-1 folds are used for estimating $\hat{\boldsymbol{c}}(t)$ with a coordinate descent, and finally the remaining fold is used for prediction from which a mean squared error is estimated. This procedure is repeated for each fold, which gives an estimation of the mean squared error for the given λ. This procedure is repeated with the same folds for the sequence of λ. The sequence of values of λ is picked using the method proposed by Friedman et al. (2010). Cross-validation is a common technique to estimate the prediction performance of a model. In our case, it has the advantage of being sensitive to the number of non-null coefficients. Indeed, since the error is computed on data which have not been used for the optimisation, more parameters in the model does not imply less error (in fact, these parameters will learn the noise in the training data which will negatively impact the performance on new data). Finally, the value of λ corresponding to the minimum mean squared error is chosen as $\lambda_{\min}$. With $\lambda = \lambda_{\min}$, the algorithm then estimates $\hat{\boldsymbol{c}}(t)$ from the whole $\boldsymbol{y}(t)$ via a coordinate descent. The support $\mathcal{S}(t)$ is finally identified from $\hat{\boldsymbol{c}}(t)$ as $\mathcal{S}(t) = \{r \in [\![1, R]\!] : \hat{c}_r(t) \neq 0\}$. We emphasize that this method has the great advantage of not requiring the knowledge of the number of VOCs in the mixture $\boldsymbol{y}(t)$ (this number is estimated through $\lambda_{\min}$). In Algorithm 2, we don't represent the cross-validation procedure for the estimation of λ in the minimization process (4) for lack of space.

At this stage, we have an estimation of $\mathcal{S}(t)$ from $\hat{\boldsymbol{c}}(t)$ considering only the time t and we have already estimated the previous ω supports $\{\mathcal{S}(t-\omega), ..., \mathcal{S}(t-1)\}$. In order to integrate the support continuity constraint, we just

Algorithm 2 Real-time estimation of the intensities $\boldsymbol{c}(t)$

Require: ω, τ, β
  # Estimate the support of $\boldsymbol{c}(t)$
  $\hat{\boldsymbol{c}}(t) = \arg\min_{\boldsymbol{c}\geq 0} \frac{1}{2P}\|\boldsymbol{y}(t) - \boldsymbol{K}\boldsymbol{c}\|_2^2 + \lambda\|\boldsymbol{c}\|_1$
  $\mathcal{S}(t) = \{r \in [\![1, R]\!] : \hat{c}_r(t) \neq 0\}$
  # Homogenise the support in relation to the window
  Let $\mathcal{S}_M(t)$ be the most frequent support in $\{\mathcal{S}(t-\omega), ..., \mathcal{S}(t)\}$
  if $\mathcal{S}(t) \neq \mathcal{S}_M(t)$ **then**
    # Compute the cross-validation error for each support
    Split the P rows of $\boldsymbol{Y}_{\omega}(t)$ into β folds f_i, $i \in [\![1, \beta]\!]$. f_i contains the indices of the test fold i. $\boldsymbol{A}_{f_i}(t)$ contains the rows $\in f_i$ and $\boldsymbol{A}_{-f_i}(t)$ contains the rows $\notin f_i$.
    $J, J_M = 0$
    for $i = 1$ **to** β **do**
      $J = J + \min_{\boldsymbol{C}\geq 0} \|\boldsymbol{Y}_{\omega,-f_i}(t) - \boldsymbol{K}_{-f_i}^{\mathcal{S}(t)}\boldsymbol{C}\|_F^2$
      $J_M = J_M + \min_{\boldsymbol{C}\geq 0} \|\boldsymbol{Y}_{\omega,-f_i}(t) - \boldsymbol{K}_{-f_i}^{\mathcal{S}_M(t)}\boldsymbol{C}\|_F^2$
    end for
    if $J_M < J$ **then**
      $\mathcal{S}(t) = \mathcal{S}_M(t)$
    end if
  end if
  # Estimate $\boldsymbol{c}(t)$ *on the final support*
  $\hat{\boldsymbol{c}}(t) = 0_R$
  $\hat{\boldsymbol{c}}^{\mathcal{S}(t)}(t) = \arg\min_{\boldsymbol{c}\geq 0} \|\boldsymbol{y}(t) - \boldsymbol{K}^{\mathcal{S}(t)}\boldsymbol{c}\|_2^2$
  # Smooth $\hat{\boldsymbol{c}}(t-\tau)$
  $\hat{\boldsymbol{c}}(t-\tau) = \frac{1}{2\tau+1}\sum_{i=-\tau}^{\tau} \hat{\boldsymbol{c}}(t-\tau-i)$

find the most frequent support $\mathcal{S}_M(t)$ of the $\omega+1$ supports. If the support $\mathcal{S}_M(t)$ equals $\mathcal{S}(t)$, then we go to the next step. Otherwise, we need to identify which support better explains the whole time window. To that end, we perfom again a β-fold cross-validation. We carry out the cross-validation on the rows of $\boldsymbol{Y}_\omega$ instead of its columns, since the two supports are known. We split the P rows of $\boldsymbol{Y}_\omega(t) \in \mathbb{R}^{P\times(\omega+1)}$ into β folds f_i, $i \in [\![1,\beta]\!]$ (f_i contains the indices of each fold). For each support and each fold, we solve the subproblem (5) considering $\boldsymbol{Y}_{\omega,-f_i}(t)$ (it contains the rows of $\boldsymbol{Y}_\omega(t)$ which are not in f_i), with a quasi-Newton algorithm called L-BFGS (Byrd et al., 1995). We then predict $\boldsymbol{Y}_{\omega,f_i}(t)$ (it contains the rows of $\boldsymbol{Y}_\omega(t)$ which are in f_i) and compute the sum of squared errors. We repeat the procedure for each fold, which gives an estimation of the sum of squared errors for each support. Finally, we identify $\mathcal{S}(t)$ as the support with the minimal error. For the first ω measurements, we just set ω to $t-1$.

Eventually, the final step is then just to solve the problem (5) for the final support $\mathcal{S}(t)$. Furthermore, a smoothing filter is then applied to $\hat{\boldsymbol{c}}(t-\tau)$, where τ is an additional parameter (here, $\tau = 0.4$ sec). We simply replace $\hat{\boldsymbol{c}}(t-\tau)$ by the mean value of $\{\hat{\boldsymbol{c}}(t-2\tau), ..., \hat{\boldsymbol{c}}(t)\}$.

The algorithm is summarized in Algorithm 2. It is important to note that this method does not assume that we know the number of VOCs in the mixture $\boldsymbol{y}(t)$ and does not normalize the data. So, the vector $\hat{\boldsymbol{c}}(t)$ which contains the intensity of each signature in the measurement $\boldsymbol{y}(t)$ should be related to the VOC concentration in some way. This claim must be confirmed by further experiments measuring the true VOC concentration with an additional instrument that can determine ground-truth concentration.

Computation time. In Section 5.4, we perform cross-validation in order to estimate a score for the data from Setup 1 ($R = 12$ VOCs). Out of ∼600 000 measurements $\boldsymbol{y}(t)$, the computation time of Algorithm 2 was then 75 ±17 ms. Eventually, the computation time needed by the two algorithms is lower than the sampling period (200 ms), as required.

4.2. Quality assessment

In this study, we have no way of measuring the *actual* concentration of each VOC, which limits the assessment of our method. However, we can use the spatial position of each gas source in order to compute a score. We can associate this spatial position with a set of time points, say $\{t_1^r, ..., t_{N_r}^r\}$ for the VOC r. Then, we can associate to each time t in this temporal range, a label $\ell(t)$ according to the maximum intensity:

$$\forall t \in \{t_1^r, ..., t_{N_r}^r\}, \quad \ell(t) = \arg\max_i c_i(t) \tag{6}$$

From this label $\ell(t)$, we can then compute a score based on the position of the gas sources. However, this score can only work for pure compounds. Since we have also binary mixtures at a time t (Scenario ③ of Setup 2), we extend the previous criterion to the mixtures of γ VOCs by simply taking the γ first maxima as the classification result:

$$\forall t \in \{t_1^{r_1,...,r_\gamma}, ..., t_{N_r}^{r_1,...,r_\gamma}\}, \forall j \in [\![1,\gamma]\!] \quad \ell_j(t) = \arg\max_{i\setminus\{\ell_1(t),...,\ell_{j-1}(t)\}} c_i(t) \tag{7}$$

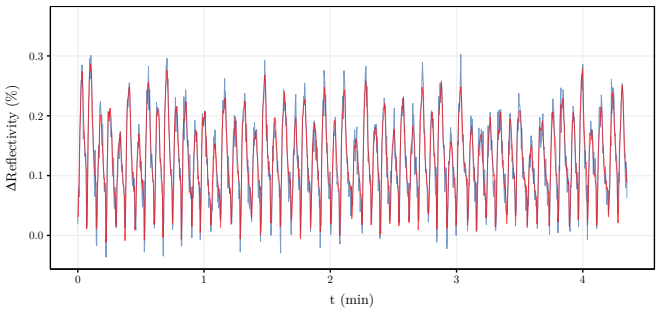
From the labels $\{\ell_1(t), ..., \ell_j(t)\}$, we consider the output of the algorithm a success if it correctly identifies all the VOCs present in the mixture. The output is considered false if at least one VOC has not been correctly identified.

We warn the reader that the spatial area of the gas sources (diameter of the cups, length of the scent strips) does not really correspond to the beginning and to the end of interactions (meaning, the rise and the baseline return). Indeed, there exists some inherent lags, such as: a chemical lag due to the dynamics of the chemical reactions (adsorption, desorption), a lag due to the airflow (the delay in transporting one molecule from the floor to the prism surface), a lag due to diffusion (the spatial range is in fact greater than the spatial area of the gas source) and another lag due to the funnel. Consequently, the response can often spread over a spatial range larger than the spatial area of the gas source. Future developments can integrate some deconvolution tools to compensate for all these lags (see for example the recent study of Martinez et al. (2019)).

Quality assessment for data from Setup 1. In this setup, the spatial area of the cups and the corresponding temporal range are quite small. As a reminder, the cups have a diameter of 2.5 cm and the speed of the robot is 2 cm/s, leading to only 6 measurements if we consider only the spatial area of a cup. To avoid an overoptimistic score (due to a too small temporal range), we perform a segmentation of the signal and use this segmentation for the definition of the temporal ranges from which we extract labels. This segmentation assumes that there is no mixture between two successive cups (which may not be true) and is detailed in the Supplementary Materials. An example of the segmentation obtained is shown in Figure 2b, in which each color corresponds to the temporal range of each one of the 24 gas sources. This approach increases the spatial range to 8 ± 3.2 cm (*i.e.* the temporal range to 4 ± 1.6 sec).

In each temporal range, we then perform a vote between all the labels extracted and associate with the gas source the majority label. This label is then compared to the experimental plan in Figure 2a.

Quality assessment for data from Setup 2. In this setup, mixtures occur in different scenarios and it is hard to predict what is occurring outside the gas sources (*i.e.* scent strips). So the spatial range is not increased, and is defined as the spatial area occupied by the scent strips. However, instead of taking a majority label for the whole area as for Setup 1, we keep each label $\ell(t)$ for the score. For instance, for **Scenario ②**, the Guaïacol source starts from 10.5cm to 15.5cm which corresponds to



(a)

VOC	R^2(%)
Butanol	98.0
β-Pinene	99.7
Benzaldehyde	98.0
Citral	92.0
Allyl hexanoate	98.7
Cis-3-hexenol	99.0
Linalool	99.1
Guaïacol	97.9
Geranyl	91.1
Trans-2-octenal	97.9
Acetic acid	94.8
(S)-Limonene	99.7

(b)

Figure 5: Justification of the rank-1 approximation. (a) On the left, the blue curve corresponds to the original pure time series of Citral for one chemical sensor and red curve corresponds to its rank-1 approximation. On the right, another representation (data vs prediction): the more the point cloud is aligned with $y = x$, the better. (b) The goodness of fit R_r^2 (Eq. (8)) for each VOC: the closer to 100% R_r^2 is, the better.

25 measurements (speed = 1cm/s and frame rate = 1Hz). For each measurement $\boldsymbol{y}(t)$, we predict the label $\ell(t)$ based on the criterion (6). If the label $\ell(t)$ is Guaïacol then the classification is correct, otherwise the prediction is false and so forth for the other gas sources. For **Scenario ③** (binary mixtures), we consider a success the identification of the full mixture (i.e. not only one VOC among the two).

5. Gas recognition of isolated sources

In this section, we tackle the issue of gas recognition with the data from Setup 1 (see Section 3.2.1). As a reminder, these data correspond to 28 laps of a path along which 24 gas sources of 12 VOCs (repeated twice) have been disseminated. The goal is then to apply Algorithm 2 to recognise each one of these gas sources.

In the following, by considering a measurement $\boldsymbol{y}(t)$, we assume that the baseline drift has been subtracted from each chemical sensor p using Algorithm 1.

5.1. Signature extraction and linear model justification

The method used in this article is a supervised one, meaning that we need the signatures of the 12 VOCs. To this end, we describe a simple method for estimating them.

According to the model we assume here (eq. (1)), the response $\boldsymbol{y}_r(t) \in \mathbb{R}^P$ to a VOC r is simply proportional to its "concentration" $c_r(t) \in \mathbb{R}$. Let Let us assume that we measure this VOC r during a long period of time, say N_r, at a variable concentration and that we stack these measurements into $\boldsymbol{Y}_r \in \mathbb{R}^{P\times N_r}$. Then, model (1) implies that the rank of $\boldsymbol{Y}_r$ is 1. In other words, we can simply write $\boldsymbol{Y}_r = \boldsymbol{u}_r\boldsymbol{v}_r^T$ (in the absence of noise) with $\boldsymbol{u}_r \in \mathbb{R}^P$ and $\boldsymbol{v}_r \in \mathbb{R}^{N_r}$. $\boldsymbol{v}_r$ contains the variations over time and $\boldsymbol{u}_r$ contains the variations across the chemical sensors, in other words $\boldsymbol{u}_r$ is the signature of the VOC r.

Here, we do not measure the VOCs with individual experiments, so we do not have directly a matrix $\boldsymbol{Y}_r$ containing the pure measurements. We extract this matrix for each VOC r by segmenting the signals with a method detailed in the Supplementary Materials. With this segmentation, we stack all the peaks corresponding to the VOC r to generate P time series containing "only" the VOC r (assuming that there is no mixture between two cups). Let us assume that these time series are of length N_r with $N_r \geq P$, so we have the matrix $\boldsymbol{Y}_r \in \mathbb{R}^{P\times N_r}$. Then, we perform a Singular Value Decomposition (SVD) of this matrix, say $\boldsymbol{Y}_r = \boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{V}_r^T$ with $\boldsymbol{U}_r \in \mathbb{R}^{P\times P}$, $\boldsymbol{\Sigma}_r \in \mathbb{R}^{P\times P}$ a diagonal matrix and $\boldsymbol{V}_r \in \mathbb{R}^{N_r\times P}$. Finally, we identify the first column of $\boldsymbol{U}_r$ as the signature $\boldsymbol{k}_r$ of the VOC r. Another way could be to perform a Non-negative Matrix Factorisation (NMF) instead of the SVD in order to ensure that the coefficients of $\boldsymbol{k}_r$ are non-negative. However, in practice, the coefficients of the extracted $\boldsymbol{k}_r$ are all of the same sign (sometimes negative, so we just flip the sign). All these signatures $\boldsymbol{k}_r$ are then stacked into a matrix $\boldsymbol{K} \in \mathbb{R}^{P\times R}$, which we call the dictionary. Each column of this dictionary is a unit-norm vector (*i.e.* $\|\boldsymbol{k}_r\|_2 = 1$).

In order to check whether the rank-1 approximation is correct or not, we compare the matrix $\boldsymbol{Y}_r$ to its best rank-1 approximation $\hat{\boldsymbol{Y}}_r = \sigma_r\boldsymbol{u}_r\boldsymbol{v}_r^T \in \mathbb{R}^{P\times N_r}$. For easier visualisation, we just represent in Figure 5a the time series $\boldsymbol{y}_{p,r} \in \mathbb{R}^{N_r}$ and its approximation $\hat{\boldsymbol{y}}_{p,r} = \sigma_r u_{p,r}\boldsymbol{v}_r$ for a given chemical sensor p and a given VOC r (here, Citral). We also compute the goodness-of-fit R_r^2:

$$R_r^2 = 1 - \frac{\sum_{p,n}(Y_{pn,r} - \hat{Y}_{pn,r})^2}{\sum_{p,n}(Y_{pn,r} - \bar{y}_r)^2} \tag{8}$$

where $Y_{pn,r}$ denote the (p,n) entries of $\boldsymbol{Y}_r$ and $\bar{y}_r \in \mathbb{R}$ their average.

	(S)-Limonene	Citral	Guaïcaol	Cis-3-hexenol
VIF	5.3	10.1	19.7	31.8

Table 2: Variance Inflation Factor for 4 VOCs which are used in Setup 2.

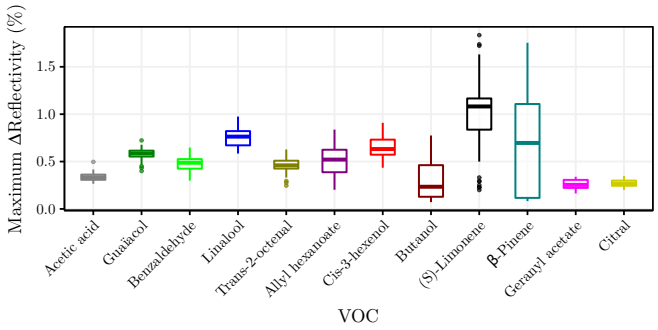


Figure 6: Distribution of the maximum intensities over the laps for each VOC.

R_r^2 equals the proportion of variance explained by the model compared to the total variance in $\boldsymbol{Y}_r$. The values of R_r^2 for each VOC are reported in Figure 5b and show a generally good fit for all of the VOCs.

5.2. Analysis of the dictionary

The method proposed in this paper is based on a penalised linear regression. A linear regression can suffer from collinearities or multicollinearites which may exist between the signatures of the matrix $\boldsymbol{K} \in \mathbb{R}^{P\times R}$. A multicollinearity is present when one column of $\boldsymbol{K}$ is equal to a linear combination of the other colums. For non-penalised linear regression, it is easy to show that multicollinearity will cause problems. Indeed, in this case, the classical least-squares solution of $\boldsymbol{y} = \boldsymbol{K}\boldsymbol{c}$ is $\hat{\boldsymbol{c}} = (\boldsymbol{K}^{\mathrm{T}}\boldsymbol{K})^{-1}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{y}$. The inversion of $\boldsymbol{K}^{\mathrm{T}}\boldsymbol{K}$ requires that $\mathrm{rank}(\boldsymbol{K}^{\mathrm{T}}\boldsymbol{K}) = \mathrm{rank}(\boldsymbol{K}) = R$ (assuming $P > R$). This condition is then violated if there exists at least one multicollinearity.

To check for multicollinearity, we use a well-known indicator, namely the Variance Inflation Factor (VIF) (James et al., 2013). Considering a signature $\boldsymbol{k}_r \in \mathbb{R}^P$:

$$\mathrm{VIF}_r = \frac{\sum_p (k_{p,r} - \bar{k}_r)^2}{\sum_p (k_{p,r} - \hat{k}_{p,r})^2} \tag{9}$$

with $\hat{\boldsymbol{k}}_r = \alpha_0 - \sum_{i\neq r} \alpha_i \boldsymbol{k}_i$ and $\bar{k}_r$ the mean of $\boldsymbol{k}_r$.

Like the R^2 criterion (eq. (8)), the VIF is dependent on the notion of variance explained. Here, we regress $\boldsymbol{k}_r$ against the other signatures and the greater the VIF is, the more $\boldsymbol{k}_r$ is linearly dependent on the other signatures. A VIF equals 1 if and only if $\boldsymbol{k}_r$ is linearly independent from the other signatures. A classical rule is that a VIF which is greater than 5 or 10 indicates collinearity problems (James et al., 2013).

To illustrate the collinearity problem in our case, we focus on a smaller dictionary which will be used for Setup 2. The results are reported in Table 2. From the Table 2, the factors indicate strong collinearities between the signatures even with only 4 VOCs in the dictionary (for the complete dictionary, the results are even worse). These results motivate the use of Algorithm 2 and its ℓ_1-penalty which can help to combat these multicollinearities.

5.3. Cross-validation

In order to avoid an overestimation of the score (with the criterion defined in Section 4.2), we perform a cross-validation. In fact, if we extract the signatures with all 28 laps and then predict the labels for these same laps, we may introduce a bias as the training set (extraction of the signatures) is then the same as the testing set.

To perform cross-validation, we divide the 28 laps into β folds (here, $\beta = 5$). The laps corresponding to the first $\beta - 1$ folds are taken for extracting the dictionary and the β^{th} fold is used for testing. Concretely, for the testing laps, we apply Algorithm (2) which returns an intensity vector $\boldsymbol{c}(t)$ for a measurement vector $\boldsymbol{y}(t)$. From this intensity vector, we extract a label for each region of interest which has been previously identified with a segmentation step (see Supplementary Materials). We then repeat the procedure for all the β folds and compute the classification rate and the confusion matrix. Finally, we repeat the entire cross-validation 10 times with new folds. From these 10 cross-validations, the classification rates and the confusion matrices are averaged.

5.4. Results

The confusion matrix is reported in Table 3 and the average classification score is 73.7 ± 21.9 % which is much larger than the chance level (8.33%). Despite the difficulty of the task, some VOCs are even perfectly or almost perfectly identified. It means that almost all the gas sources containing these VOCs have been correctly recognised (*e.g.* Acetic acid or Benzaldehyde). However, it happens that Algorithm 2 does not find any VOC (*i.e.* all the VOCs have a null intensity) even if the segmentation does. These estimations have been classified as "*None*" in the confusion matrix and correspond to regions of interest with low signal to noise ratio. For these regions, the null solution must have the lower cost compared to solutions with one or more VOCs.

Figure 6 represents the distribution of the maximum intensities for each VOC over the time. The diversity of intensities is clear, showing the exhaustion of some VOCs (*e.g.* β-Pinene) and the stability of others (*e.g.* Guaïacol). It demonstrates that the good classification score cannot be attributed to a single factor such as a simple difference of intensity between VOCs. Recognition performance can only be explained by the variability in affinities between the sensing materials and the VOCs that produces recognisable signatures.

Prediction \ Truth	Acetic acid	Guaïacol	Benzaldehyde	Linalool	Trans-2-octenal	Allyl hexanoate	Cis-3-hexenol	Butanol	(S)-Limonene	β-Pinene	Geranyl acetate	Citral
Acetic acid	**100**	0	0	0	2	1	0	5	0	3	0	0
Guaïacol	0	**100**	0	0	0	0	0	0	0	0	0	0
Benzaldehyde	0	0	**99**	1	1	1	0	6	0	1	0	4
Linalool	0	0	0	**92**	2	0	1	0	0	0	0	0
Trans-2-octenal	0	0	0	1	**63**	6	0	3	0	4	0	17
Allyl hexanoate	0	0	0	0	18	**72**	0	0	0	1	3	0
Cis-3-hexenol	0	0	0	0	0	1	**59**	16	0	1	0	0
Butanol	0	0	1	1	4	3	40	**59**	2	4	0	16
(S)-Limonene	0	0	0	0	1	5	0	7	**72**	24	9	0
β-Pinene	0	0	0	0	1	6	0	0	20	**48**	3	0
Geranyl acetate	0	0	0	5	1	3	0	1	4	4	**84**	35
Citral	0	0	0	0	7	2	0	1	1	1	0	**28**
None	0	0	0	0	0	0	0	2	0	9	0	0

Table 3: Confusion matrix for the data from Setup 1. The colored cells correspond to pairs of VOCs which are hard to differentiate. A *None* class has been added when Algorithm (2) did not find any VOC. In fact, it happens that for some regions of interest, Algorithm (2) does not find any VOC (all the intensities are null) whereas the segmentation method did not discard them (especially areas with low signal to noise ratio).

Misclassifications are sometimes due to the fact that some pairs of VOCs are quite hard to differentiate (colored cells in Table 3). It is interesting to note that these pairs are sometimes from the same chemical family. For instance (Butanol, Cis-3-Hexenol) are both Alcohols and (β-Pinene, (S)-Limonene) are both Alcenes and they even share the same molar mass. Of course, chemical similarity is not the whole story, since e.g. Linalool is not confused with other Alcohols. The misclassifications between Citral and Geranyl acetate can be attributed to the lower signal-to-noise ratio for these two VOCs. This low signal-to-noise ratio can be explained by both their low volatility and their low affinity with the sensing materials.

6. Gas unmixing

In this section, we tackle the issue of gas unmixing with the data from Setup 2 (see Section 3.2.2). We remind the reader that we have generated various realistic scenarios with increasing complexity: ① the gas sources are isolated and spatially close, ② the gas sources are successive trails of pure compounds and ③ the gas sources are successive trails of binary compounds. In each scenario, the data correspond to 20 sweeps (from left to right, the results for the other direction is in the Supplementary Materials). Each one of these 20 sweeps is then processed separately for the unmixing with the previously generated dictionary and with Algorithm 2.

In the following, when considering a measurement $\boldsymbol{y}(t)$, we assume that baseline drift has been subtracted using Algorithm (1).

6.1. Building and pruning of the dictionary

The dictionary $\boldsymbol{K} \in \mathbb{R}^{P\times R}$ is built with the method detailed in Section 5.1 and from the whole data set from Setup 1. This last point is an important aspect of the following results: it implies that the training setup is not the same as the testing setup. This characteristic is quite important in practice. Indeed, the dictionary will be always generated in a separated setup in order to be used in the field afterwards. So, training with Setup 1 and testing with Setup 2 will assess the robustness of the proposed method.

From the dictionary, we first extract the set of the 4 VOCs which are present in the scenarios ① to ③, namely: Citral, (S)-Limonene, Guaïacol and Cis-3-hexenol. At least one VOC is always used as a control, meaning that at least one VOC is present in the dictionary but not in the experiment. We call the VOCs actually present the *target VOCs*. We expect that the estimated intensity of the control VOC will be close to or equal to 0. This is a key point of our results; indeed, an eNose is a non-specific device that can generate signatures for a broad variety of VOCs, contrary to specific sensors which are designed for one or two VOCs. In practice, an eNose may be less effective than specialised sensors if we only target 2 specific VOCs. So there is no reason to favour an eNose except if we want to use it for a large amount of applications with a lot of different VOCs. That is why, in practice, the dictionary will be as large as possible and only an unknown subset of VOCs from the dictionary will be relevant for a given application. To our knowledge, by considering a larger dictionary than the number of VOCs actually present, we go further than most of the studies in the field and es-

	Citral		(S)-Limonene		Distance	
	Source 1	Source 2	Source 1	Source 2	Source 1	Source 2
Mean ± Standard deviation (cm)	8.6 ± 0.35	27.9 ± 0.50	11.9 ± 0.31	25.5 ± 0.28	3.3 ± 0.41	2.3 ± 0.47
Ground-truth (cm)	[8,8.5]	[27,27.5]	[11.5,12]	[25.5,26]	3	1

Table 4: Estimation of the position of the isolated gas sources from Scenario ① with the default dictionary (Citral, (S)-Limonene, Guaïacol, Cis-3-hexenol), based on the maximum intensity. The column distance refers to the spatial distance between Citral and (S)-Limonene.

pecially studies which perform mixtures-learning. What's more, these 4 VOCs are clearly not the best subset of the 12 VOCs from Setup 1. Indeed, by looking at the confusion matrix in Table 3, the subset containing Acetic acid, Guaïacol, Benzaldehyde and Linalool would probably lead to much better results than the ones presented here.

The results with the dictionary of 4 VOCs are described in the next section. Afterwards, we progressively extend the dictionary by adding one VOC at a time. The order of the VOCs is defined with the help of the confusion matrix (Table 3). Given the current set of VOCs in the sub-dictionary, we simply add all the confusions made with these VOCs (*i.e.* we add amongst themselves the columns of the confusion matrix containing the set). Then, we take as new member of the set the VOC which has the lowest confusion coefficient. With this method, the order of the VOCs (after the 4 VOCs already chosen by default) is: Acetic Acid, Allyl-hexanoate, Linalool, Benzaldehyde, Trans-2-octenal, β-Pinene, Geranyl acetate and Butanol. The task is then harder and harder and, at the end, the whole dictionary is used in the unmixing.

6.2. Results with a dictionary of 4 VOCs

The intensities estimated for the 4 VOCs in the 3 scenarios are represented in Figure 7. Each color stands for a VOC and each line corresponds to a sweep (a color gradation indicates the sweep index).

For **Scenario** ①, the intensities are reported in the top panel of Figure 7. The intensities of Citral and (S)-Limonene clearly indicate good unmixing from the signals previously shown in Figure 3b. We can notice that the intensities of the controls (namely Cis-3-hexenol and Guaïacol) are not strictly null. However, these intensities correspond mainly to transition areas (from one gas source to another) or to areas with low signal to noise ratio (SNR). What's more, the intensities of the controls are much lower than the intensities of the two VOCs actually present. By computing the classification criterion defined in Section 4.2, we reach a noteworthy score of 90% (98% for the other direction) for the spatial areas defined by the scent strips (corresponding to a total of 240 measurements). In this scenario, we even go further by estimating the spatial position of each isolated gas source. For that, we simply take as the position of a gas source, the position of its maximum intensity. The average estimated intensities for the 20 sweeps are reported in Table 4. These results highlight a good estimation of the position of the gas sources when they are far enough from each other (separated by 3cm). The same task is harder for a smaller distance (1cm) for which the distance is overestimated. In addition, (S)-Limonene presents better estimation results compared to Citral. This can be explained by the possible lags introduced in Section 4.2, especially the chemical lag and the lag related to the airflow. Indeed, Citral is a heavier molecule than (S)-Limonene (see Table 1) so its transport and its interaction with the chemical sensors can take more time and thus delay the measurement.

For **Scenario** ②, the results are represented in the middle panel of Figure 7. Again, the estimated intensities correspond to the spatial areas of the gas sources. However, the estimated intensity of the control (here, Cis-3-hexenol) is no more negligible and no more restricted to low-SNR regions. We notice that the intensity of the control VOC depends mainly on the VOC present. Indeed, the gas source containing Guaïacol is well estimated whereas the estimations for the other gas sources ((S)-Limonene and Citral) are more affected by Cis-3-hexenol. This observation is a consequence of the existence of linear dependencies between the signatures. Despite these correlations, we find a classification score of 83% (92% for the other direction), out of 1 520 measurements. This score is quite good in view of the difficulty of the task. Even if the intensity of the control VOC is high, it is still lower than those of the target VOCs.

Finally, for **Scenario** ③, the intensities are reported on the bottom panel of Figure 7. At first sight, the results seem better than for **Scenario** ② which is simpler, especially if we focus on the control VOC. In fact, the control VOC is no more Cis-3-hexenol but now Guaïacol. Guaïacol seems to be a much better control VOC than Cis-3-hexenol. Indeed, the estimated intensities of Guaïacol are close to 0, as expected. This can be explained by a VIF which is lower for Guaïacol (see Table 2) and by the Table 3 which shows that Guaïcaol is perfectly identified despite the 11 other VOCs. For the present VOCs, the estimated intensities match with the spatial position of the gas sources, especially for the gas sources containing Citral and (S)-Limonene. The intensities of Cis-3-hexenol are less simple to analyse since the intensities exceed the spatial areas of its gas sources which could be due to ternary mixtures. However, a comforting fact is that these intensities clearly decrease when the eNose goes over a gas source which does not contain Cis-3-hexenol. The score reaches

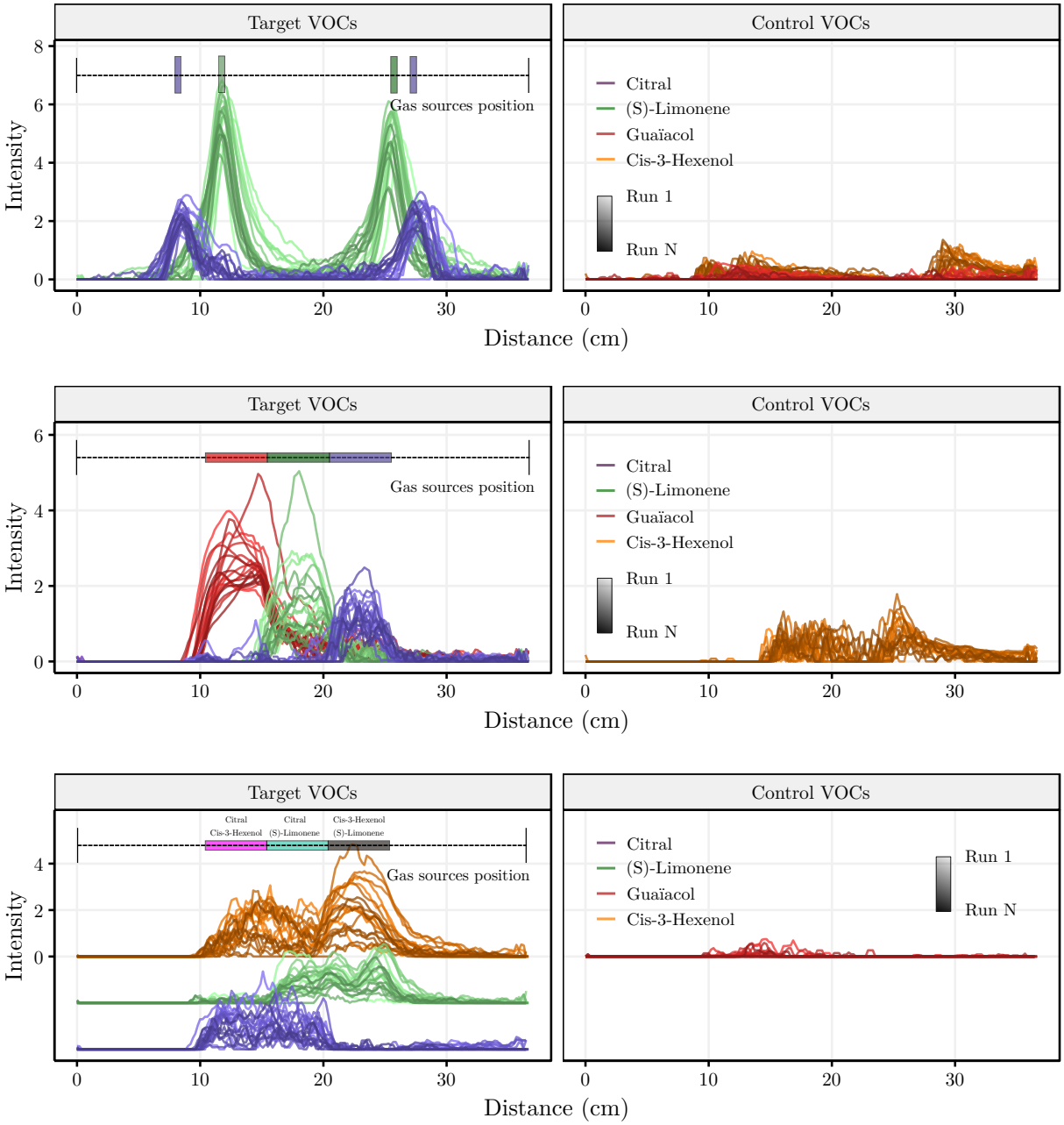


Figure 7: Results of the proposed algorithm for the different scenarios introduced in Section 3.2.2. Each line corresponds to one sweep (first sweep is the lighter). Each color corresponds to the estimated intensity of the given VOC at the distance d. The results have been generated with the default dictionary ((S)-Limonene, Citral, Guaïacol, Cis-3-hexenol). For Scenario ③, the intensities of Citral and (S)-Limonene have been vertically shifted for easier visualisation.

72% (71% in the other sweep direction) which means that the identity of 72% of all binary mixtures (corresponding to 1,520 measurements, i.e. 1,520 mixtures) has been well predicted.

6.3. Results with a dictionary of increasing size

The previous results emphasize that we can achieve good classification performance even with one or two VOCs which are not present in the experiment but present in the dictionary. Here, we go even further by adding one-by-one each VOC from the full dictionary. As a reminder (see Section 6.1 for details), the order according which the VOCs are added, is the following: Acetic acid, Allyl-hexanoate, Linalool, Benzaldehyde, Trans-2-octenal, β-Pinene, Geranyl acetate and Butanol.

Concretely, we start by adding the signature of Acetic acid to the previous dictionary of size $R = 4$ ((S)-Limonene, Citral, Guaïacol, Cis-3-hexenol). Then, we apply Algo-

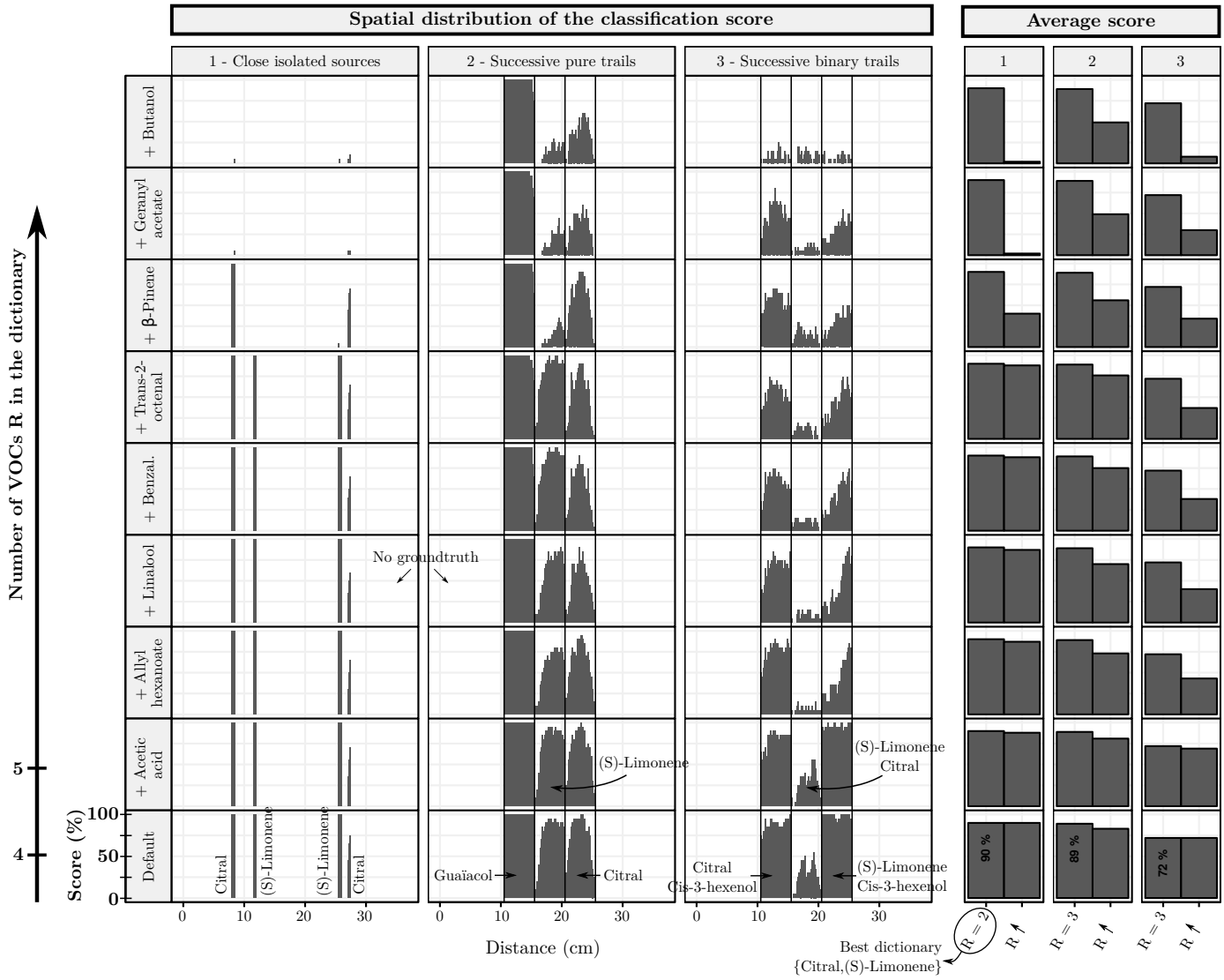


Figure 8: Influence of the size of the dictionary on unmixing performance. From the bottom up, the number of VOCs in the dictionary is increasing, starting from the default dictionary ((S)-Limonene, Citral, Guaïacol, Cis-3-hexenol). Left to right, the three first left panels correspond to the spatial distribution of the score for each scenario (for a position d, the score is the average across the 20 sweeps). The spatial distribution highlights the misclassifications (mainly in transition areas) in each scenario and for each dictionary. For **Scenarios** ② and ③, a black line indicates the transitions between one gas source to another. Finally, the right hand panels correspond to the average score in each scenario and for each dictionary (averaging over space and across the two sweeps). For each individual panel, the left bar corresponds to the score of the best dictionary (*i.e.* the smallest) and the right bar correspond to the score of the current dictionary.

rithm (2) to unmix the signals and estimate the intensity of each VOC of this new sub-dictionary. Ideally, the estimated intensity of Acetic acid is close to 0. However, in practice the new VOC could considerably disrupt the unmixing, especially due to correlations which may exist between this new signature and the signatures already present. To assess the influence of this new VOC, we simply generate the label $\ell(t)$ based on the maximum intensity of $\boldsymbol{c}(t)$ (see Section 4.2 for details). This predicted label is then compared to the position of the gas sources. In Figure 8, we report both the spatial distribution of the errors averaged across the sweeps (left panel) and the average score (right panel). Afterwards, we repeat the procedure by adding Allyl-hexanoate and so forth for the others. Until we reach the upper limit of $R = 12$ VOCs (the full dictionary), the task becomes harder and harder with each new VOC we add.

From Figure 8, for **Scenarios** ① and ②, we see that the scores start to considerably decrease only at $R = 10$ VOCs. It is noteworthy that even with 9 VOCs in the dictionary, we can reach a classification score of 88% for **Scenario** ① and 76% for **Scenario** ②. In fact, $R = 10$ corresponds to the addition of β-Pinene which is highly-correlated (99.8 %) to (S)-Limonene (for the two scenarios, the gas sources of (S)-Limonene are no longer identified). Afterwards, at $R = 11$, the addition of Geranyl acetate produces another large decrease of the scores (again, for the two scenarios, the gas sources of Citral are no longer

identified). In fact, the confusion matrix (see Table 3) already showed that Geranyl acetate can cause a lot of misclassifications for Citral. At $R = 11$ or $R = 12$ (the full dictionary), the final classification scores reach 2% for **Scenario** ① and 49% for **Scenario** ②. This shows that the addition of Butanol has no effect on the classification. The difference between **Scenario** ① and ② is explained by the gas source containing Guaïacol which is present in **Scenario** ② but not in **Scenario** ①. The correct identification of Guaïacol agrees with Table 3, in which Guaïacol is also reliably identified. Finally, the hardest areas to identify are the transition areas. This observation is really clear in **Scenario** ② and reflects a drawback of our assessment method. Indeed, the score is based on the spatial position of the gas sources but it does not take into account the diffusion of the VOC or some interactions which can occur between different VOCs. For example, there is no reason that the gas source of (S)-Limonene for **Scenario** ② starts exactly at the beginning of its scent strip and not a little time after or before due to diffusion or due to interaction with Guaïacol. So our assessment method may be biased to some extent.

In **Scenario** ③, things are even more complicated. If we only refer to the scores, they "continuously" decrease starting with the introduction of Allyl-hexanoate at $R = 6$. It means that when we add two control VOCs (Guaïacol and Acetic acid), there is no strong effect on classification performance compared to the default dictionary. In this case, most of the misclassifications are due to the middle gas source in which the intensity of Cis-3-hexenol seems to be over-estimated (see an example in Figure 7). However, again, the score computed with the spatial position of the gas sources does not take into account all the possible variations which can occur. Here, it is likely that ternary mixtures occur around the middle gas source (Cis-3-hexenol is present in the two extreme binary trails). Ternary mixtures are not taken into account by the criterion and a single misclassification leads to a global misclassification of the mixture according to our criterion. This explains why the scores for Scenario ③ are lower and decrease faster than in Scenarios ① and ②.

6.4. Discussion

The unmixing results seem quite acceptable in all scenarios. To our knowledge, it is the first demonstration of such unmixing results with an electronic nose in an uncontrolled environment with uncontrolled sampling. What's more, these results have four additional key aspects: first, the mixtures have not been previously trained (only the dictionary of pure compounds and, on top of that, with a separated setup), second, the number of VOCs in the mixtures is at no time assumed during the unmixing process, third, the dictionary is much larger than the number of VOCs actually present, and finally, all the results have been generated in a real-time fashion.

However, the results are obviously perfectible. First, they can be affected by the correlations or the linear dependencies which exist between the signatures of the dictionary. The ℓ_1-penalty of Algorithm 2 improves the solution in presence of collinearity compared to classical least-squares solutions. But it improves the solution in a way which can be unacceptable in our case. Indeed, if two signatures are strongly correlated then the Lasso will tend to randomly select one of the two signatures, putting the other to zero (Tibshirani, 2011). In other words, in the field, if only the VOC A is present and strongly correlated to a VOC B present in the dictionary, then the Lasso solution may indicate the presence of sometimes A, sometimes B. If A is a harmful VOC and B is a safe one, it is easy to understand the consequences of such an error. Improvements can be made by decorrelating A from B or by simply discarding the signature of B from the dictionary. Another line of research to differentiate A and B, is also the design of additional sensing materials. Indeed, if A and B have currently correlated signatures, it means that the actual sensing materials are not enough discriminative regarding these two compounds.

Second, even if the linear model seems a good approximation here, a non-linear relationship exist certainly and may improve the unmixing. A non-linear model motivated by theory would be hardly achievable due to the variations in multiple crucial parameters during an experiment, such as humidity, temperature or concentration of the VOC. An empirical improvement could be to add an interaction term such the one proposed by Llobet et al. (1998) for the MOX sensors. In our notation, the model would correspond to $\boldsymbol{y} = \sum_r (\boldsymbol{k}_r c_r - \sum_{i \neq r} \boldsymbol{a}_{ir} c_i c_r)$ where $\boldsymbol{a}_{ir} \in \mathbb{R}^P$ is the interaction vector between the VOC i and the VOC r. This interaction term could then capture some variations which are not taken into account by the linear model. However, some trade-offs are involved since a more complicated model can require more experiments (for calibrating some parameters for instance) and can also be much more complicated to fit due to the existence of local minima, indeterminations, or numerical unstability.

7. Conclusion and future works

This study reports two main results regarding the use of a SPR-based optoelectronic nose (abbreviated as eNose) for a robot application. First, we succeeded in recognising up to ∼73% of 24 isolated gas sources containing 12 different VOCs. Second, we succeeded in unmixing binary and ternary mixtures occurring in various realistic scenarios. When isolated gas sources are close enough for significant mixing to take place, we reach a classification score of 90%. When the robot goes over several successive trails of pure compounds, we reach a score of 89%. Finally, when the robot goes over several successive trails of binary compounds, we reach a score of 72%.

Our methodology has several noteworthy features. First, the results have been obtained with a single, unified, processing pipeline. Second, this pipeline is based on the use of a dictionary of pure VOCs, which is quick to generate.

Third, the unmixing algorithm is based on a real-time implementation, and does not require any information about the number of VOCs present in the mixtures. Finally, the dictionary was generated in a different setup and was always larger (and sometimes much larger) than the number of VOCs currently present.

In the following, we detail possible directions for future work.

The experiments we have carried out can be extended in many ways. First, besides the convection induced by the funnel, there is not so much advection in our experiments because we are in a closed room. To be even more realistic, further studies can incorporate some wind, but wind can decrease recognition performance (see the study of Vergara et al. (2013)). Second, the actual concentration of each VOC is unknown in our experiments. This was a restrictive factor for the assessment of our results and it should be removed in further studies. The concentration would be especially highly valuable in order to assess if the intensity matrix that we estimate is related in some way to the true concentration of the VOCs. For instance, the use of a PhotoIonization Detector can provide such information (see Hernandez Bennetts et al. (2014)) but at the cost of a high airflow ($\sim$500 mL/min) which could easily impoverish the gas mixture. Third, the scale of the setups that we have built can be considered small compared to other setups in the literature (see Fan et al. (2019)). Even if the small scale used here was appropriate for assessing our algorithm, the next setup must be larger to gain even more realism. Finally, the experiments were carried out on the same day or with a one-day interval. In practice, the dictionary is generated once and has to be used several days, weeks or even months after its generation. It could then be interesting and highly valuable to reproduce the results of Maho et al. (2019) in the more complicated setting used here.

Regarding the algorithms that we have proposed, the development of methods for decorrelating the signatures and the enrichment of the mixing model have been already discussed in Section 6.4. However, other improvements can be imagined. First, if we try to extend the results of Maho et al. (2019) on drift correction, we may need to design a different correction method. Second, measurement is affected by temporal lag, related for instance to the chemical reactions between the VOC and the sensing materials (see Section 4.2 for details). In source localisation, these lags can impair the estimation of the position of the gas source. Deconvolution techniques can then be deployed to mitigate the effects of the lags (see Martinez et al. (2019)). For gas mixtures, these techniques can be particularly difficult to develop since the chemical lag depends strongly on the VOCs in the mixture, which are unknown. Finally, an interesting extension of our results would be to consider the blind case, in other words the Blind Source Separation (BSS) framework. The main difference with our study is that we would no longer assume knowledge of the signatures matrix $\boldsymbol{K}$ for the unmixing. Only very few recent studies have dealt with this issue (*e.g.* Maho et al. (2018), Madrolle et al. (2018)) but a BSS method could considerably reduce the experimental time required for learning.

Acknowledgments

This work was supported by the FUI-WISE AAP21 minalogic project, BPI. The authors would like to acknowledge Hiroshi Ishida from the University of Tokyo for dicussions and comments on this work.

References

Artursson, T., Eklöv, T., Lundström, I., Martensson, P., Sjöström, M., Holmberg, M., Sep. 2000. Drift correction for gas sensors using multivariate methods. Journal of Chemometrics 14 (5-6), 711–723.

Bioucas-Dias, J. M., Aug. 2009. A variable splitting augmented Lagrangian approach to linear spectral unmixing. In: 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. pp. 1–4.

Bishop, C. M., Hinton, G. E., Strachan, I. G. D., Jul. 1997. GTM through time. In: Fifth International Conference on Artificial Neural Networks. Cambridge, UK, pp. 111–116.

Brenet, S., John-Herpin, A., Gallat, F.-X., Musnier, B., Buhot, A., Herrier, C., Rousselle, T., Livache, T., Hou, Y., Jul. 2018. Highly-Selective Optoelectronic Nose Based on Surface Plasmon Resonance Imaging for Sensing Volatile Organic Compounds. Analytical Chemistry.

Byrd, R., Lu, P., Nocedal, J., Zhu, C., Sep. 1995. A Limited Memory Algorithm for Bound Constrained Optimization. SIAM Journal on Scientific Computing 16 (5), 1190–1208.

Comon, P., Jutten, C., Feb. 2010. Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press.

Dickinson, M. E., Bearman, G., Tille, S., Lansford, R., Fraser, S. E., Dec. 2001. Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy. BioTechniques 31 (6).

Fan, H., Bennetts, V. H., Schaffernicht, E., Lilienthal, A. J., Apr. 2018. A cluster analysis approach based on exploiting density peaks for gas discrimination with electronic noses in open environments. Sensors and Actuators B: Chemical 259, 183–203.

Fan, H., Hernandez Bennetts, V., Schaffernicht, E., Lilienthal, A. J., Jan. 2019. Towards Gas Discrimination and Mapping in Emergency Response Scenarios Using a Mobile Robot with an Electronic Nose. Sensors 19 (3), 685.

Figaro, 2013. TGS 2600 - for the detection of Air Contaminants.

Fonollosa, J., Rodríguez-Luján, I., Trincavelli, M., Vergara, A., Huerta, R., Oct. 2014. Chemical Discrimination in Turbulent Gas Mixtures with MOX Sensors Validated by Gas Chromatography-Mass Spectrometry. Sensors 14 (10), 19336–19353.

Fonollosa, J., Sheik, S., Huerta, R., Marco, S., Aug. 2015. Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring. Sensors and Actuators B: Chemical 215, 618–629.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of statistical software 33 (1), 1–22.

Hernandez Bennetts, V., Schaffernicht, E., Pomareda, V., Lilienthal, A., Marco, S., Trincavelli, M., Sep. 2014. Combining Non Selective Gas Sensors on a Mobile Robot for Identification and Mapping of Multiple Chemical Compounds. Sensors 14 (9), 17331–17352.

Huerta, R., Vembu, S., Amigó, J. M., Nowotny, T., Elkan, C., Sep. 2012. Inhibition in multiclass classification. Neural Computation 24 (9), 2473–2507.

Hyndman, R. J., Fan, Y., 1996. Sample Quantiles in Statistical Packages. The American Statistician 50 (4), 361–365.

Ishida, H., Wada, Y., Matsukura, H., Nov. 2012. Chemical Sensing in Robotic Applications: A Review. IEEE Sensors Journal 12 (11), 3163–3173.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. Vol. 6. Springer.

Llobet, E., Vilanova, X., Brezmes, J., Sueiras, J. E., Alcubilla, R., Correig, X., Jan. 1998. Steady-state and transient behavior of thick-film tin oxide sensors in the presence of gas mixtures. Journal of The Electrochemical Society 145 (5), 1772–1779.

Loutfi, A., Broxvall, M., Coradeschi, S., Karlsson, L., Sep. 2005. Object recognition: A new application for smelling robots. Robotics and Autonomous Systems 52 (4), 272–289.

Loutfi, A., Coradeschi, S., Jun. 2006. Smell, think and act: A cognitive robot discriminating odours. Autonomous Robots 20 (3), 239–249.

Loutfi, A., Coradeschi, S., Lilienthal, A. J., Gonzalez, J., 2009. Gas distribution mapping of multiple odour sources using a mobile robot. Robotica 27 (2), 311–319.

Maass, W., Natschläger, T., Markram, H., Nov. 2002. Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Computation 14 (11), 2531–2560.

Madrolle, S., Duarte, L. T., Grangeat, P., Jutten, C., Sep. 2018. A Bayesian Blind Source Separation Method for a Linear-quadratic Model. In: 2018 26th European Signal Processing Conference (EUSIPCO). pp. 1242–1246.

Maho, P., Barthelmé, S., Comon, P., Jul. 2018. Non-linear source separation under the Langmuir model for chemical sensors. In: 10th IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM 2018). IEEE, Sheffield, United Kingdom, pp. 380–384.

Maho, P., Dolcinotti, C. L., Livache, T., Herrier, C., Andreev, A., Comon, P., Barthelmé, S., May 2019. Olfactive robot for gas discrimination over several months using a new optoelectronic nose. In: 2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN). pp. 1–3.

Marques, L., Nunes, U., de Almeida, A. T., Oct. 2002. Olfaction-based mobile robot navigation. Thin Solid Films 418 (1), 51–58.

Martinez, D., Burgués, J., Marco, S., Jan. 2019. Fast Measurements with MOX Sensors: A Least-Squares Approach to Blind Deconvolution. Sensors 19 (18), 4029.

Martinez, D., Rochel, O., Hugues, E., Jun. 2006. A biomimetic robot for tracking specific odors in turbulent plumes. Autonomous Robots 20 (3), 185–195.

Mcgill, K., Taylor, S., Apr. 2011. Robot Algorithms for Localization of Multiple Emission Sources. ACM Comput. Surv. 43 (3), 15:1–15:25.

Mead, M. I., Popoola, O. A. M., Stewart, G. B., Landshoff, P., Calleja, M., Hayes, M., Baldovi, J. J., McLeod, M. W., Hodgson, T. F., Dicks, J., Lewis, A., Cohen, J., Baron, R., Saffell, J. R., Jones, R. L., May 2013. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. Atmospheric Environment 70, 186–203.

Monroy, J. G., Gonzalez-Jimenez, J., Mar. 2017. Gas classification in motion: An experimental analysis. Sensors and Actuators B: Chemical 240, 1205–1215.

Monroy, J. G., Palomo, E. J., López-Rubio, E., Gonzalez-Jimenez, J., Nov. 2016. Continuous chemical classification in uncontrolled environments with sliding windows. Chemometrics and Intelligent Laboratory Systems 158, 117–129.

Palacín, J., Martínez, D., Clotet, E., Pallejà, T., Burgués, J., Fonollosa, J., Pardo, A., Marco, S., Apr. 2019. Application of an Array of Metal-Oxide Semiconductor Gas Sensors in an Assistant Personal Robot for Early Gas Leak Detection. Sensors 19 (9).

Rochel, O., Martinez, D., Hugues, E., Sarry, F., Sep. 2002. Stereo-olfaction with a sniffing neuromorphic robot using spiking neurons. In: 6th European Conference on Solid-State Transducers.

Röck, F., Barsan, N., Weimar, U., 2008. Electronic nose: current status and future trends. Chemical reviews 108 (2), 705–725.

Russell, R. A., Feb. 2001. Survey of Robotic Applications for Odor-Sensing Technology. The International Journal of Robotics Research 20 (2), 144–162.

Schleif, F.-M., Hammer, B., Monroy, J. G., Jimenez, J. G., Blanco-Claraco, J.-L., Biehl, M., Petkov, N., Feb. 2016. Odor recognition in robotics applications by discriminative time-series modeling. Pattern Analysis and Applications 19 (1), 207–220.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society. Series B (Statistical Methodology) 73 (3), 273–282.

Trincavelli, M., 2011. Gas Discrimination for Mobile Robots. Ph.D. thesis, Örebro University.

Vergara, A., Fonollosa, J., Mahiques, J., Trincavelli, M., Rulkov, N., Huerta, R., Aug. 2013. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. Sensors and Actuators B: Chemical 185, 462–477.

Vergara, A., Vembu, S., Ayhan, T., Ryan, M. A., Homer, M. L., Huerta, R., May 2012. Chemical gas sensor drift compensation using classifier ensembles. Sensors and Actuators B: Chemical 166-167, 320–329.

Supplementary materials
Online gas recognition and gas unmixing in robot application

Pierre Maho, Cyril Herrier, Thierry Livache,
Pierre Comon, Simon Barthelme

1. Segmentation process

For the data from Setup 1, a segmentation process is required to identify regions of interest (ROIs). A ROI is defined as the temporal range during which a VOC is picked up by the device. The ROIs are only used for evaluating the performance of our algorithm but are not required for applying it. Here, we detail step-by-step the segmentation that we carry out. This section is partly translated from a previous conference article from our group (Maho et al. (2019), in French).

In the following, the segmentation is always performed on the average signals across the chemical sensors.

1.1. Synchronisation of the laps

As a reminder, the data from Setup 1 is the repetition of a predefined path. Given N laps of this path, we need to find N signals of duration N_t (the duration of a lap) which have the same ROIs. In order to improve the signal to noise ratio and to make easier the detection of these ROIs, we work with the average lap $\bar{y}(t)$.

Currently, the robot is not able to report its spatial position, so we have to find the beginning of each lap n in order to synchronize them. A solution could be to simply take a multiple of N_t. However, in practice, the duration of a lap can vary a bit and this variation can correspond to a ROI (~ 1 sec). To have a better estimation of the beginning, we assume that we know the first lap, denoted by $s(t)$. In practice, this lap is manually extracted from the signals. Then, we define the sliding correlation $m(t) \in [-1, 1]$ as:

$$m(t) = \frac{\sum_{n=1}^{N_t} s(n)y(n+t)}{(\sum_{n=1}^{N_t} s(n)^2 \times \sum_{n=1}^{N_t} y(n+t)^2)^{\frac{1}{2}}} \tag{1}$$

$m(t)$ is then equal to 1 if there is a perfect match with the template $s(t)$. So we identify the beginning of each lap as the position of the maxima of $m(t)$, greater than a threshold (0.7).

1.2. Detection of the ROIs

The detection of the ROIs from the signals is a three-step procedure. The first two steps correspond to a rough detection from $\bar{y}(t)$ based on a thresholding and on the derivative of the signal. The third step refines the ROI for each lap using a model.

First, from $\bar{y}(t)$, we detect the maxima in a given temporal range (here, $\pm$5s due to the distance between two cups and due to the robot speed) which are greater than a threshold ($5\sigma_n$ with σ_n the standard deviation of the noise).

Second, from these maxima, we apply a heuristic procedure to approximate the ROIs. Starting from a local maximum, we increase the area around this position until there is a significant change in the sign of the derivative. This change signals the beginning of a new ROI. Thus, we identify $2R = 24$ ROIs which are represented by colored lines in Figure 1.

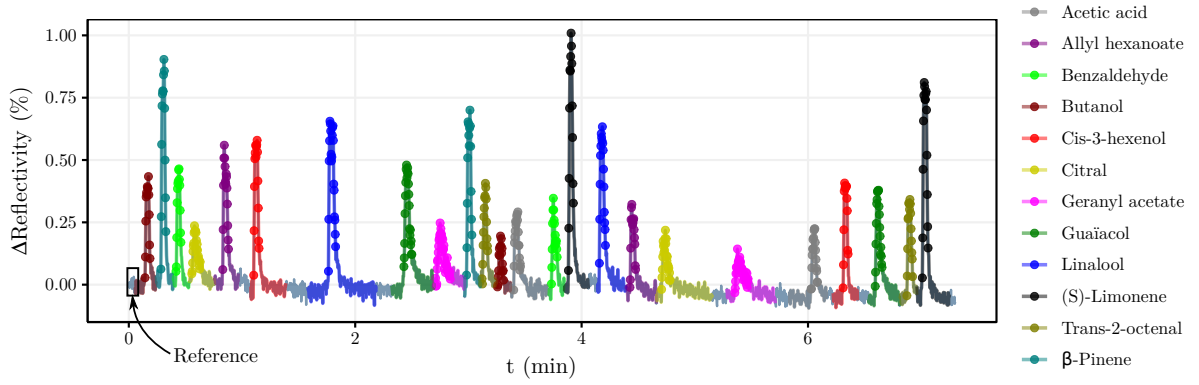


Figure 1: Results of the segmentation for one lap and one chemical sensor. Each color stands for a different VOC. Colored lines correspond to rough first segmentation and colored poinds correspond to the final segmentation which is used in the article (after refining with the model).

Finally, we refine these ROIs using a model. We note $y_{rn}(t)$ the ROI of the average chemical sensor corresponding to the VOC r during the lap n and $y_{prn}(t)$ the signal of the chemical sensor p.

We can model partly the response of the chemical sensor p by considering a chemical model, namely the Langmuir model (Maho et al., 2018). It allows to model the chemical interactions between the VOC r and the chemical sensor p during the lap n using 3 parameters: the adsorption constant a_{pr}, the desorption constant d_{pr} and the concentration $c_{rn}(t)$ of the VOC. By noting t^s_{rn} and t^e_{rn}, the beginning and the end of the VOC injection, the model is the following:

$$\theta_{prn}(t) = \begin{cases} 0 & \text{si } t < t^s_{rn} \\ \frac{a_{pr}c_{rn}(t)}{a_{pr}c_{rn}(t)+d_{pr}}(1 - e^{-(a_{pr}c_{rn}(t)+d_{pr})(t-t^s_{rn})}) & \text{si } t \in [t^s_{rn}, t^e_{rn}] \\ \theta_{prn}(t^e_{rn})e^{-d_{pr}(t-t^e_{rn})} & \text{si } t > t^e_{rn} \end{cases} \tag{2}$$

Unfortunately, the model (2) makes too many assumptions to be valid here (*e.g.* temperature and humidity are constant). So we use a simpler form to model the average response which does not assume any relationship between the parameters:

$$\theta_{rn}(t) = \begin{cases} 0 & \text{si } t < t^s_{rn} \\ \theta^{\text{eq}}_{rn}(1 - e^{-\alpha_{rn}(t-t^s_{rn})}) & \text{si } t \in [t^s_{rn}, t^e_{rn}] \\ \theta_{rn}(t^e_{rn})e^{-\beta_{rn}(t-t^e_{rn})} & \text{si } t > t^e_{rn} \end{cases} \tag{3}$$

We use the model (3) only for the identification of the beginning t^s_{rn} and the end t^e_{rn} of the VOC respons r for each lap n. For t^s_{rn}, we extract the adsorption signal (rise) by approximating t^e_{rn} as the time point with max amplitude. Then, for each possible value of t^s_{rn}, we estimate the parameters (θ^{eq}_{rn}, α_{rn}) which minimize the quadratic cost, using the L-BFGS algorithm (Byrd et al., 1995). Finally, we choose t^s_{rn} as the one with the minimal cost. For t^e_{rn}, the procedure is similar: assuming that t^s_{rn} is now known, we estimate the parameters (θ^{eq}_{rn}, α_{rn}, β_{rn}) from the entire signal $y_{rn}(t)$ for values of t^e_{rn} close to the max, then we associate t^e_{rn} with the lowest cost. We this method, there is no need for an extra hyperparameter, such as a threshold value, for identifying the injection range in the ROI. Some results are represented in Figure 2.

Finally, we refine each ROI by taking the points in $[t^s_{rn}, t^e_{rn} + \frac{1}{\beta_{rn}}]$. The final ROIs for one lap and one biosensor are represented by colored points in Figure 1.

1.3. Outliers detection

Among all the ROIs extracted, some have too little signal to be useful, which is mainly due to VOC exhaustion. We detect and remove all the ROIs for which the average $y_{rn}(t)$ around

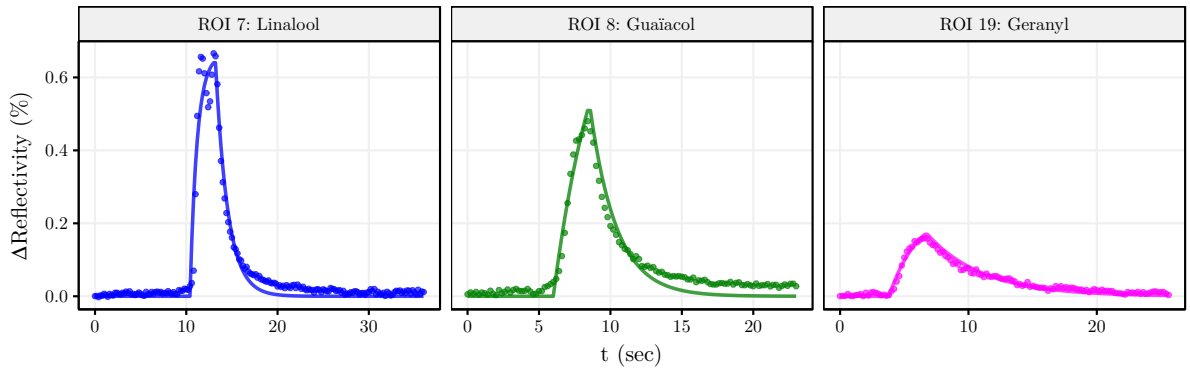


Figure 2: Results of the fit for 3 different ROIs. The points correspond to the average chemical sensor and the lines correspond to the fit.

the max is lower than $5\sigma_n$. The final number of ROIs is then 663 (out of $28 \times 24 = 672$ possible ROIs), corresponding to 663 labels of the 12 VOCs used.

2. Setup 2 - Results for the sweep from right to left

	Citral		(S)-Limonene		Distance	
	Source 1	Source 2	Source 1	Source 2	Source 1	Source 2
Mean ± Standard deviation (cm)	9.8 ± 0.32	29.1 ± 0.21	11.1 ± 0.21	24.9 ± 0.17	1.2 ± 0.28	4.2 ± 0.31
Ground-truth (cm)	[9,9.5]	[28,28.5]	[10.5,11]	[24.5,25]	1	3

Table 1: Estimation of the position of the isolated gas sources from Scenario ① with the default dictionary (Citral, (S)-Limonene, Guaïacol, Cis-3-hexenol), based on the maximum intensity. The column distance refers to the spatial distance between Citral and (S)-Limonene.

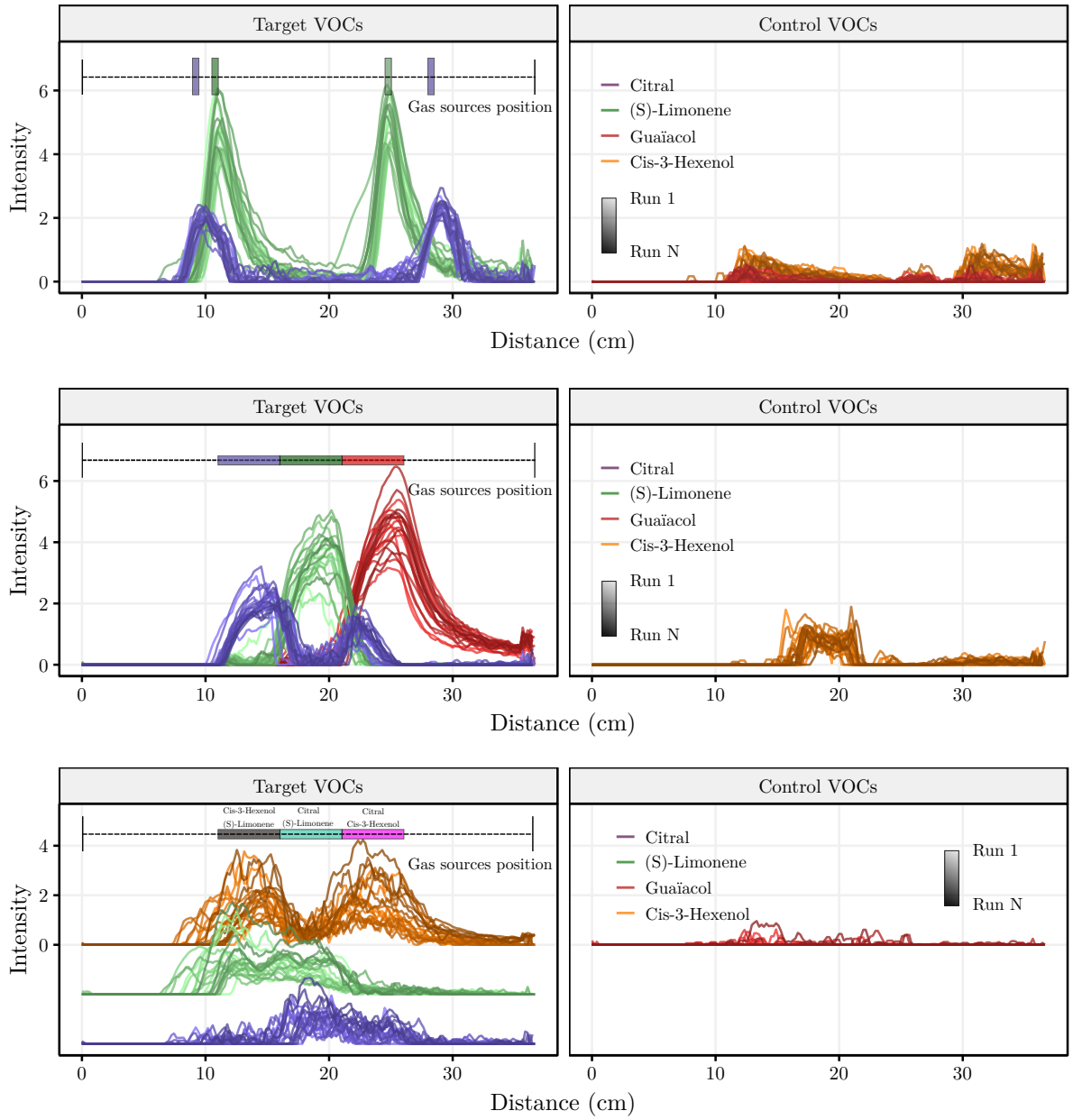


Figure 3: Right to left direction. Results of the proposed algorithm for the different scenarios introduced. Each line corresponds to one sweep (first sweep is the lighter). Each color corresponds to the estimated intensity of the given VOC at the distance d. The results have been generated with the default dictionary ((S)-Limonene, Citral, Guaïacol, Cis-3-hexenol). For Scenario ③, the intensities of Citral and (S)-Limonene have been vertically shifted for easier visualisation.

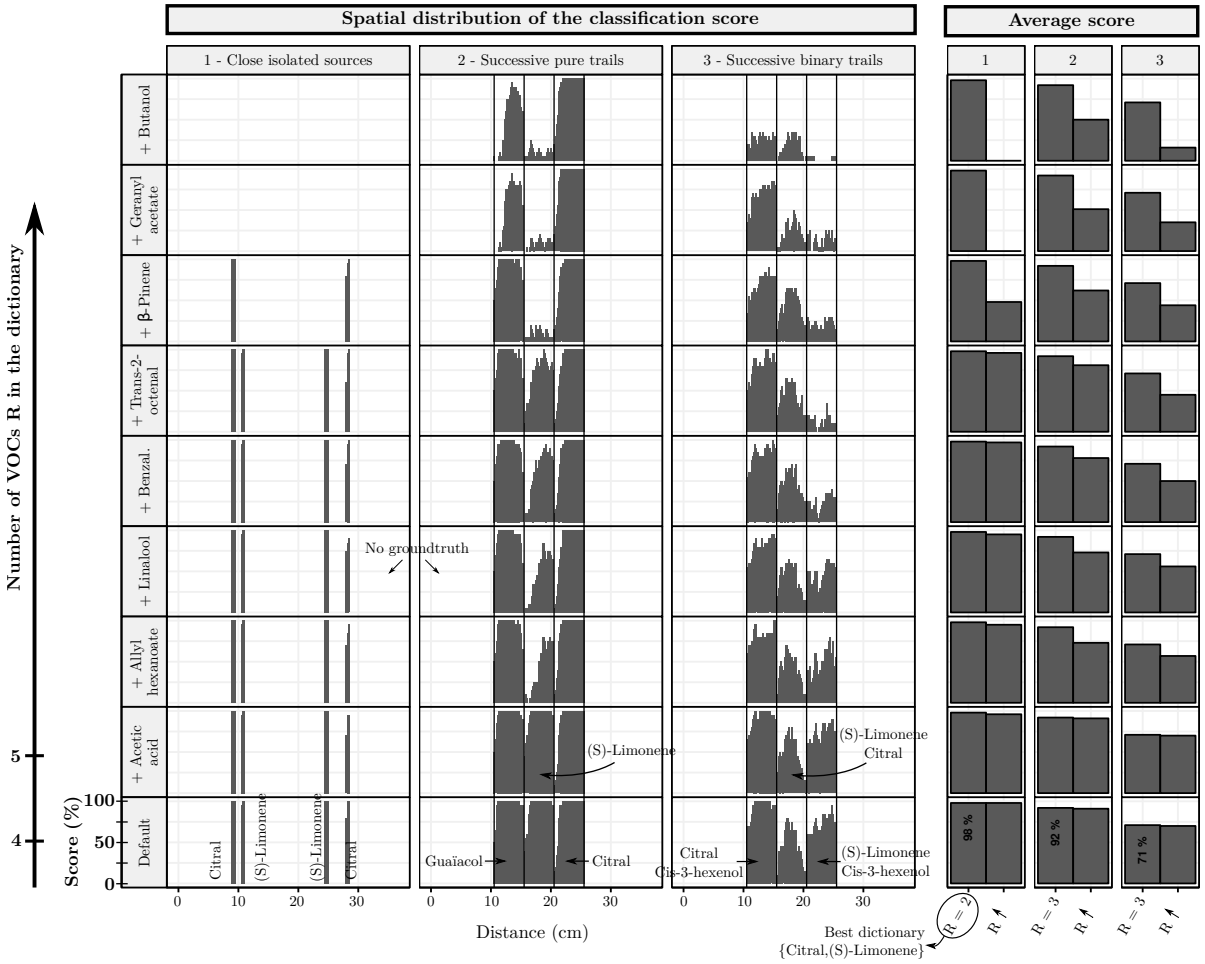


Figure 4: Influence of the size of the dictionary on the demixing performance. From the bottom up, the number of VOCs in the dictionary is increasing, starting from the default dictionary ((S)-Limonene, Citral, Guaïacol, Cis-3-hexenol). Left to right, the three first left panels correspond to the spatial distribution of the score for each scenario (for a position d, the score is the average across the 20 sweeps). The spatial distribution highlights the misclassifications (mainly in transition areas) for each scenario and each dictionary. For **Scenarios** ② and ③, a black line indicates the transitions between one gas source to another. Finally, the right hand panels correspond to the average score for each scenario and each dictionary (so average over space and across the sweeps). For each individual panel, the left bar corresponds to the score of the best dictionary (*i.e.* the smallest) and the right bar correspond to the score of the current dictionary.

References

Byrd, R., Lu, P., Nocedal, J., Zhu, C., Sep. 1995. A Limited Memory Algorithm for Bound Constrained Optimization. SIAM Journal on Scientific Computing 16 (5), 1190–1208.

Maho, P., Barthelmé, S., Comon, P., Jul. 2018. Non-linear source separation under the Langmuir model for chemical sensors. In: 10th IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM 2018). IEEE, Sheffield, United Kingdom, pp. 380–384.

Maho, P., Dolcinotti, C. L., Livache, T., Herrier, C., Andreev, A., Comon, P., Barthelmé, S., Aug. 2019. Reconnaissance de plusieurs composés chimiques à l'aide d'un robot équipé d'un nez électronique. In: XXVIIème colloque GRETSI (GRETSI 2019). Lille, France.