Exploring the Relation Between Emotional Reception of CAV Technology and Its Acceptance: The SUaaVE* Concept

Dinesh Paudel[1], Nicolás Palomares[2], Juan-Manuel Belda-Lois[2,3], Sofía Iranzo[2], Begoña Mateo[2], Ömer Yilmaz[1], Victoria Neumann[1], and Yannick Morel[1]

[1] Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany
dinesh.paudel@tum.de,{yilmazo, victoria.neumann, yannick.morel}@in.tum.de

[2] Instituto de Biomecánica de Valencia, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
{nicolas.palomares, jmbelda, sofia.iranzo, begona.mateo}@ibv.org

[3] Grupo de Tecnología Sanitaria del IBV, CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) Valencia, Spain

Abstract. The emotional relationship users develop with a given technology has a direct impact on its acceptance. It is around this simple, though well established premise that the work pursued in SUaaVE (H2020 research project) was structured. Efforts are invested towards placing the human factor at the center of technological developments, structuring the onboard intelligence around passengers' emotions. Several complementary avenues of investigations are explored, including novel, innovative approaches to real-time emotion estimation, and at-runtime adjustments to the vehicle's road behavior in light of estimated information. This system-theoretic and algorithmic work is complemented by development of novel HMI technology, and analysis of societal factors driving acceptance of the considered technology over a representative range of stakeholder groups across Europe.

Keywords: Connected Automated Vehicles (CAVs), Human-Driven Design (HDD), Cognitive and emotional model, Affective computing.

1 Introduction

Though maturity of driving automation technology has seen major advances in recent years, its acceptance by the public has remained limited ([1]). For this technology to achieve its desired impact, societal issues pertaining to public acceptance, user awareness, and ethical considerations need necessarily be addressed. Expected benefits of the technology will fail to materialize if it is not adopted by the target user group ([2, 3]). The literature shows that, though the

* This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°814999.

public at large is fascinated by Connected and Automated Vehicle (CAV) technology, 43% of people remain afraid of traveling in an autonomous car ([4]). In addition, concerns exist in terms of compliance of the technology with expected ethical standards. The European Group on Ethics (EGE) in science and new technologies recommends that automation technology should be aligned with fundamental values adhered to by EU Treaties ([5]). Estimation of the emotional state of a car's driver has received attention in the literature, in particular relying on physiological signals ([6]). However, reliably observing the emotions of autonomous cars' passengers remains a challenge.

To directly address such issues, the SUpporting acceptance of automated Vehicle project (SUaaVE, Horizon 2020 research project) will explore the relation between prospective users' emotional reception of the technology and their acceptance of it. This will be pursued relying on a Human-Driven Design (HDD) approach, exploiting synergies between social science, human factors and automotive market research, by means of iterative assessment, co-design, and prototyping processes. The approach pursued is inclusive, it will involve users (passengers, drivers, vulnerable road users), experts, and stakeholders. In addition, developments pursued within SUaaVE will address ethical considerations related to safety-critical situations and protection of human life in the development of CAV technology. The expected outcome is the emergence of a novel automation paradigm: ALFRED (Automation Level Four and Reliable Empathic Driver), which will adjust the vehicle's road behavior in accordance with the observed emotional state of the passengers and of other actors involved in the scene. ALFRED will rely on an empathy unit (EMY) to assess the emotional and cognitive state of passengers, while taking social and ethical aspects into account. In addition, an adaptive cognitive and emotional interface will be established, including a set of services aimed at enhancing passengers' experience.

SUaaVE will explore manners in which the emotional and cognitive state of the passenger may inform the ride of an automated vehicle, to influence their emotional state positively, with consideration for societal and ethical aspects. The work conducted aims to contribute to acceptance of CAV technology.

2 The SUaaVE Concept

The central objective of SUaaVE consists in investigating, supporting, and enhancing public acceptance of CAVs, at societal and individual level, through the integration of the human perspective within the technology. To achieve this, efforts invested aim to develop a framework explicitly accounting for the human factor (i.e. emotion, dynamic and ambient comfort, ethical aspects) within the vehicle's intelligence, and in the implemented communication channels between vehicle and passengers. More specifically, developments are structured along the following three complementary avenues of investigations (see Figure 1),

01 **ALFRED: Automation Level Four + Reliable Empathic Driver** Fundamental architecture developed to robustly assess emotions of the passenger on board (considering the emotional and cognitive state of the passenger,

and addressing ethical considerations), and adapt the vehicle behaviour and features to enhance the user's experience (through adjustment of the vehicle's road behaviour, ambient comfort factors, and adaptive communication).

02 **Immersive Virtual Human Centred Design platform (V-HCD)**
Development of an immersive CAV simulator, exploited to assess acceptance of the technology through active involvement of future users.

03 **Guidelines for the support of public authorities**
Development of policy recommendations to stimulate and promote greater levels of CAV societal acceptance, with consideration for all road users. These efforts will facilitate harmonization and alignment of implemented CAV technology across distinct national and European initiatives.

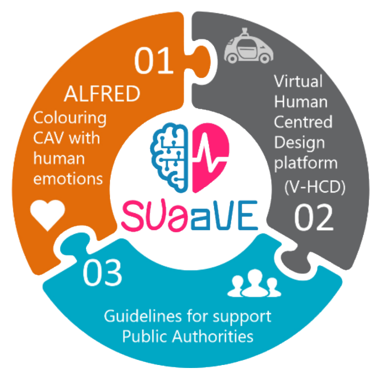

Fig. 1. Outline of the `SUaaVE` concept, sketching the path to acceptance.

Developments in `SUaaVE` lean on a Human-Driven Design (HDD) approach, a methodology shown to be of special relevance to enabling and emerging information and communication technologies. The approach revolves around the notion that the user is not only the focal point of the process, but she/he actively contributes, even leads, the definition of the concept, the technology's development, and actively participates in testing. Work conducted will focus on the human factor, directly addressing somewhat less than tangible aspects such as safety perception and, very generally, the emotional appraisal of CAVs by stakeholders. Outcome of the conducted work is aimed at all current and future technology users and stakeholders: CAV passengers, current and future drivers (with special consideration for children, senior citizens, and people with disabilities), and Vulnerable Road Users (VRUs). Efforts will be supported by the active engagement of an Advisory Board featuring public authorities, industry representatives, and a selection of concerned stakeholder associations.

`SUaaVE` will address different scenarios through the `ALFRED` concept. An example of such a scenario could be the following: "A girl is taken everyday from school to her mother's workplace. In one of the journeys, it rains heavily and the visibility becomes poor, the girl is negatively affected. `ALFRED` detects the change in emotional state through physiological signals, movements, and sound measures. To improve her state, `ALFRED` adjusts the drive to her comfort. It offers to call the girl's mother, and executes upon acceptance. The girl talks with her mother, she recovers to a normal emotional state."

3 Robust, Real-Time Emotion Estimation

There exists extensive literature on emotion recognition in speech, facial expressions, body expressions, and relevant physiological variables, as discussed in

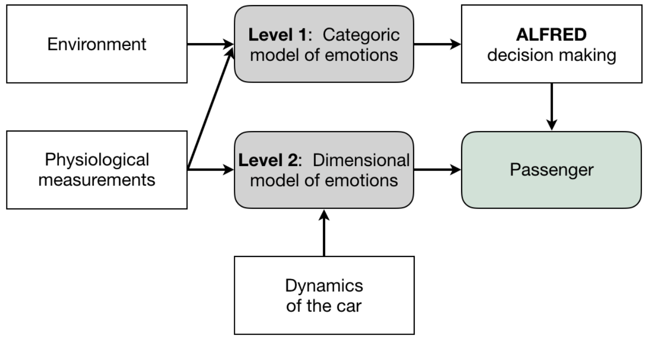

Fig. 2. The two levels of emotion analysis within SUaaVE. A categoric model is used to drive the decision making, while a dimensional model is used to convey impact of variables descriptive of the car's dynamical behavior.

[7–9]. Meaningful information is contained within physiological signals (such as e.g. electromyogram, blood pressure, skin conductivity, respiration rate) which can be gainfully exploited to assess or classify the affective state of a person ([10]). Though a wide range of classifiers rely on facial expressions exclusively, supplementing such information with complementary modalities is expected to lead to improved performance. There exists a wide range of models used for classification of emotions in psychology. The most common types being referred to as categorical ([11]), dimensional ([12, 13]), and cognitive ([14]).

The emotion estimation work conducted within SUaaVE accounts for two distinct levels of interaction, as illustrated in Figure 2. The first considered level is concerned with decision making in relation to aspects of the environment that affect the emotions of the passenger (external, ambient, contextual stimuli). The second level is related to aspects of the car's road behaviour that affect the emotional state of the vehicle's passengers.

To address the first considered level, the cognitive appraisal model of emotions is particularly well suited. The approach has the benefit that it emphasises the relation between emotions and elicitors. The cognitive appraisal model pursued in the project is based on the categorical structure described in [14]. This model describes a structure of 22 emotions, described in terms of variables related to events, agents, and objects (originators of the emotions). One of the main challenges in the construction of this categorical model is to cover a wide enough variety of scenarios in the framework of the autonomous car. The collection of a robustly large range of factors is crucial. To that end, a user-centered elicitor gathering experimentation in two stages was designed. The first stage is aimed at collecting eliciting situations with questions such as: "In a journey in the autonomous car you feel anxious, fear, fright, because of the prospect of a negative event that may occur during the ride, please imagine one such event." We expect the following type of answer: "Such feelings may be elicited by a traffic jam, which may cause me to arrive late at work". A second stage is used to collect appraisal elicited by immersive experiences in a simulator. These elicitors can mostly be gathered from the analysis of the environment. Cross-referencing this information with that in a database collected in the user-centered compilation, it becomes possible to predict the emotional response of passengers in relation to given situations. In addition, a range of sensors will be used to measure the physiological response (in relevant modalities) and audiovisual information illustrative of passengers' reactions. The main considered measurements, beyond audiovisual recordings, are respiratory and heartbeat rates. Both of these may be measured in an unobtrusive manner, specifically avoiding the use of instruments which may affect (and thus color) arousal and valence of the passenger. The physiological response will help to select the most likely emotion from the emotion candidates obtained from the analysis of the environment.

The model will consist of a predictive filter of the following form,

$$\dot{x}(t) = f\left(x(t), u(t)\right) + \epsilon_1(t), \qquad x(0) = x_0, \qquad t \geqslant 0, \tag{1}$$

$$y(t) = h\left(x(t)\right) + \epsilon_2(t), \tag{2}$$

where $x(t) \in \mathbb{R}^n$ represents the state vector (descriptive of the emotional state of the passenger), $y(t) \in \mathbb{R}^m$ is the output vector (composed of the considered physiological measures informative of passengers' emotional state), $u(t) \in \mathbb{R}^p$ is the input vector (descriptive of relevant aspects of the vehicle's road behaviour, specifically those which affect emotional response of the subjects).

The second part of the filter captures the relationship between the physiological signals and the components of the emotions which may be directly controlled by the car's behaviour, these essentially entail arousal and valence. The corresponding relationship can be found in the literature ([15]). The structure of the model may be articulated as follows. A first part reflects dynamics of the emotional response. A second part is descriptive of the impact, on the emotional dynamics, of relevant variables characterising the vehicle's road behaviour. These two parts are modelled separately. Accordingly, we re-write (1) as follows,

$$\dot{x}(t) = f_1\left(x(t)\right) + f_2\left(u(t)\right), \qquad x(0) = x_0, \qquad t \geqslant 0, \tag{3}$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are estimated based on data collected through specific experiments. To identify $f_1(\cdot)$, controlled stimuli are presented to a set of subjects, dynamics of the reaction are measured and used to develop the model. The experimentation will be performed in a semi-dynamic driving simulator. While being on board, the user will be exposed to events such as risky manoeuvres and traffic jams, and be monitored to record changes in their emotional state. Similarly, $f_2(\cdot)$ is estimated by subjecting vehicle passengers to a selection of stimuli reflecting the possible range of vehicle road behaviors.

4 Beyond Static Classifiers: Dynamic Observers

In Section 3, we have discussed the prospect of exploiting measures of external stimuli, relevant to the passenger's emotions (such as the vehicle's dynamic

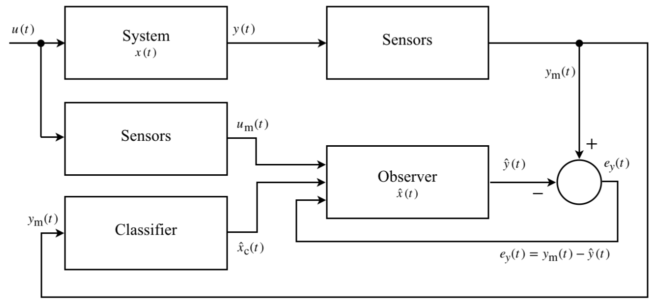

Fig. 3. Schematic representation of the considered observation problem; the considered system is descriptive of the subject's emotional response, its output $y(t)$, $t \geqslant 0$, is a set of physiological variables, measures $u_{\mathrm{m}}(t)$ of relevant stimuli, output measures $y_{\mathrm{m}}(t)$, and the outcome $\hat{x}_{\mathrm{c}}(t)$ of a classification procedure (as discussed in Section 3), are used together to support observation of the system state $x(t)$.

behavior, for instance), together with physiological measures, to estimate and possibly classify the passenger's emotional state. The notion of exploiting a dynamical model (such as that represented by (1)–(2)) in such a respect is not particularly widespread in the literature, although one may find a number of results, such as that in [16], which consider temporal behavior of input data. In [17], the authors do propose the use of one (simple, linear) such dynamical system to reflect the dynamical nature of the subjects' emotional reactions. However, only a limited number of results exploring this avenue of investigation can be found. Those that do remain typically limited to rather simple linear forms, such as that in [17].

Simple notions of cognitive psychology are sufficient to appreciate that our emotional response typically behaves in a dynamic manner; our emotional state is a direct function of influencing factors (external stimuli, or system input in a system-theory formulation, in particular when considering a state-space representation as that in (1)–(2), see [18]). It is not a static, instantaneous relationship however, hence the *dynamic* label. In addition, availability of a model of this system is of special import in the perspective of estimating a person's emotional state. If the relationship from emotional state to correlated physiological measure clearly also is of a dynamic nature, time-scales involved are significantly faster, such that it is oftentimes treated as a direct, static relationship as described by (2) (for simplicity). However the time-scale from external stimuli $u(t)$, $t \geqslant 0$, to emotional state $x(t)$ is significantly slower, motivating the consideration of the time differential in (1).

External stimuli ($u(t)$ in Figure 3) playing a significant determining role in the evolution of our emotions, it stands to reason to account for such factors when determining a passenger's emotional state. The relation from stimulus to state not being direct, but described by a dynamic map of the form of (1), motivation to develop a model of such a map is straightforward. Such a model constitutes a tool that allows to gainfully exploit measures of relevant external stimuli to inform an emotional state estimate. In addition, there also exists information within the form of (1). In particular, specifics of $f(\cdot)$, directly descriptive of the considered subject's emotional dynamics, provide information on the manner in which the person may be expected to react to considered stimuli. This presents a challenge as, if the system model information proves faithful enough (i.e. proficient in accurately describing the subject's reactions), the model allows to appropriately account for the impact of external stimuli. In the opposite eventuality however, relying on such a model may not only prevent us from exploiting measures of relevant stimuli, but actively deteriorate performance when combining the outcome of this process with output-based estimation (i.e. estimation of the emotional state based on measured physiological signals).

To exploit measures of external relevant stimuli together with measures of physiological signals, we construct an observer of the following form,

$$\dot{\hat{x}}(t) = \varphi\left(\hat{x}(t), u_{\mathrm{m}}(t)\right) + \gamma(e_y(t), \hat{x}_{\mathrm{c}}(t)), \qquad \hat{x}(0) = \hat{x}_0, \qquad t \geqslant 0, \tag{4}$$

$$\hat{y}(t) = \chi\left(\hat{x}(t)\right), \tag{5}$$

where $\hat{x}(t) \in \mathbb{R}^n$, $t \geqslant 0$, is the observed estimate of the subject's emotional state $x(t)$, $u_{\mathrm{m}}(t) \in \mathbb{R}^p$ represents the vector of measured relevant stimuli, $\hat{y}(t) \in \mathbb{R}^m$ is the predicted output, $e_y(t) \triangleq y_{\mathrm{m}}(t) - \hat{y}(t)$ corresponds to an output error, $y_{\mathrm{m}}(t)$ is the measured output (vector of physiological measures), $\hat{x}_{\mathrm{c}}(t) \in \mathbb{R}^n$ reflects the outcome of a classification procedure (as discussed in Section 3), $\varphi(\cdot)$ is an estimate of $f(\cdot)$ in (1), $\chi(\cdot)$ an estimate of $h(\cdot)$ in (2), $\gamma(\cdot)$ is to be defined by the designer, and n, m, $p \in \mathbb{N}$.

The general form of (4)–(5) is typical of Kalman filters or Luenberger-type observers ([19, 20]), wherein $\varphi(\cdot)$ plays the role of predictor (predicting expected change in state for the measured stimuli), and $\gamma(\cdot)$ that of corrector. The predicted state is used in (5) to produce a predicted output, which is then compared to the measured output. The resulting error $e_y(t)$, $t \geqslant 0$, is used to correct the prediction. A non-trivial difference in the approach considered here is the inclusion of $\hat{x}_{\mathrm{c}}(t)$ in the corrective term, which allows to account for the outcome of the classification procedure within the proposed observation framework.

As alluded to previously, accuracy of the dynamics model (that is, how closely $\varphi(\cdot)$ captures $f(\cdot)$) is of crucial importance, and viability of the overall approach hinges on it. More specifically, development of observation tools commonly rely on strong model knowledge assumption. To such an extent that, relaxing such assumptions typically undermines achievement of any type of observation guarantees (see the discussion in [21, 22]). As pertains to the considered problem however, the existence of some measure of uncertainty on the system dynamics (1) is likely inescapable. Processes involved (of emotional response) are commonly described in cognitive psychology, but in qualitative terms. Conversely, quantitative representations or closed-form models which may be of direct use in (1) are not common in the literature. Accordingly, we will follow a modeling/identification process to develop sufficiently faithful models of the considered dynamics. Specifically, qualitative considerations developed in the cognitive psychology literature will be relied upon to define a general frame for the considered state-space model. A range of practical experiments will be conducted to collect representative data, considering a sufficiently wide range of subjects. This data will be relied upon to adjust the aforementioned general model in such a manner that it reflects its contents. A wide range of techniques may be considered to that end, including for instance indirect adaptive approaches ([23]) supported by a Q–modification scheme ([24]). Alternatively, one may choose to rely on Gaussian regression ([25]) to adjust form of the model to reflect collected data, though such an approach requires transformation of the time-differential relation in (1) to an algebraic one. This is routinely achieved using, for instance, nonlinear swapping (as discussed in [26]). Such techniques typically rely on a nonlinear parametrisation of uncertainty, a common assumption that often proves straightforward to satisfy. In particular, should one approximate for instance the right-hand-side of (1) using a linear combination of unknown (uncertain) parameters and known nonlinear functions of the state and the input, the residual error, corresponding to the difference between this estimate (assuming ideal parameters) and the actual value of $f(\cdot)$, may also be represented (up to arbitrary accuracy) using such a linear parameterisation, provided that $f(\cdot)$ is a continuous function of its arguments. This directly follows from Weierstrass' approximation theorem ([27]), see for illustration the discussion in [28].

5 Conclusion

The approach explored in the `SUaaVE` project aims at promoting acceptance of vehicular automation by placing the human factor at the center of technological developments. This is pursued by affording special attention to CAV users' emotional response to the technology. In this perspective, a number of approaches are followed to develop robustly reliable real-time estimators, allowing to assess passengers' emotional response at run-time. This is achieved by using a number of classifiers, but also dynamic observers. Building upon these developments, the work within the project is currently oriented towards exploiting the assessed information to, in real-time, adjust the vehicle's on-road behavior in such a manner that passengers' emotional state is positively affected.

References

1. F. Ekman, M. Johansson, and J. Sochor. Creating appropriate trust for autonomous vehicle systems: A framework for hmi design. Technical report, 2016.
2. M. Cunningham and M. A. Regan. Autonomous vehicles: human factors issues and future research. In *Proc. of the 2015 Australasian Road Safety Conf.*, 2015.
3. T. Horberry, A. Stevens, and M. A. Regan. *Driver acceptance of new technology: Theory, measurement and optimisation.* CRC Press, 2014.
4. M. Schmidt. Unfälle mit selbstfahrenden autos: Für viele ist der hersteller verantwortlich. 2016.
5. E.S.I. Group. E.S.I. Group.
6. C. Dobbins and S. H. Fairclough. Detecting negative emotions during real-life driving via dynamically labelled physiological data. *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 830–835, 2018.

7. A. Kleinsmith and N. Bianchi-Berthouze. Affective body expression perception and recognition: A survey. *IEEE Trans. on Affective Comp.*, 4(1):15–33, Jan 2013.
8. Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, Jan 2009.
9. R. W. Picard, E. Vyzas, and J. Healey. Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191, Oct 2001.
10. E. Vyzas and R. W. Picard. Affective pattern classification. In *Emotional and Intelligent: The Tangled Knot of Cognition*, pages 176–182, 1998.
11. P. Ekman. *Basic Emotions*, chapter 3, pages 45–60. John Wiley & Sons, 2005.
12. A.R. Mehrabian and A.James. *An approach to environmental psychology.* the MIT Press, 1974.
13. J.A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.
14. A.C. Ortony, L. Gerald, and A. Collins. *The Cognitive Structure of Emotions.* Cambridge University Press, 1988.
15. A. Haag, S. Goronzy, P. Schaich, and J. Williams. Emotion recognition using biosensors: First steps towards an automatic system. In *Affective Dialogue Systems*, pages 36–48, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
16. T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li. Spatial–temporal recurrent neural network for emotion recognition. *IEEE Trans. on Cybernetics*, 49(3):839–847, 2019.
17. M. Karg, S. Haug, K. Kuhnlenz, and M. Buss. A dynamic model and system-theoretic analysis of affect based on a piecewise linear system. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 238–244, Sep. 2009.
18. H. K. Khalil. *Nonlinear Systems.* Prentice Hall, third edition, 2002.
19. R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. of the ASME Journal of Basic Engineering*, 82(Series D):35–45, 1960.
20. D. G. Luenberger. Observer for multivariable systems. *IEEE Trans. on Automatic Control*, AC-11(2):190–197, 1966.
21. Y. Morel and A. Leonessa. Prediction-based observation of nonlinear systems non-affine in the unmeasured states. In *Proc. 48th IEEE Conf. Dec. Contr.*, 2009.
22. Y. Morel and A. Leonessa. Nonlinear predictor-based output feedback control for a class of uncertain nonlinear systems. In *Proc. 2010 ASME Dynamic Systems and Control Conference*, Boston, MA, 2010.
23. K. S. Narendra and A. M. Annaswamy. *Stable Adaptive Systems.* Pr. Hall, 1989.
24. K. Y. Volyanskyy, A. J. Calise, and B.-J. Yang. A novel q-modification term for adaptive control. In *Proc. 2006 IEEE Am. Contr. Conf.*, 2006.
25. C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning.* MIT Press, 2006.
26. M. Krstić, I. Kanellakopoullos, and P. Kokotović. *Nonlinear and Adaptive Control Design.* John Wiley and Sons, Inc, New York, NY, 1995.
27. H. Jeffreys and B. S. Jeffreys. *Methods of Mathematical Physics,* 3rd Ed. Cambridge University Press, Cambridge, England, 2000.
28. Y. Morel, M. Porez, and A. J. Ijspeert. Estimation of relative position and coordination of mobile underwater robotic platforms through electric sensing. In *Proc. 2012 IEEE Int. Conf. on Rob. and Aut.*, pages 1131–1136, St. Paul, MN, 2012.