

Article

A Novel Gesture-Based Language for Underwater Human–Robot Interaction

Davide Chiarella ^{1,*†}, Marco Bibuli ¹, Gabriele Bruzzone ¹, Massimo Caccia ¹,
Andrea Ranieri ¹, Enrica Zereik ¹, Lucia Marconi ² and Paola Cutugno ²

¹ Institute of Intelligent Systems for Automation—National Research Council of Italy, Via E. De Marini 6, 16149 Genova, Italy; marco.bibuli@ge.issia.cnr.it (M.B.); gabriele.bruzzone@ge.issia.cnr.it (G.B.); massimo.caccia@ge.issia.cnr.it (M.C.); ranieri@ge.issia.cnr.it (A.R.); enrica.zereik@ge.issia.cnr.it (E.Z.)

² Institute of Computational Linguistics—National Research Council, Via E. De Marini 6, 16149 Genova, Italy; lucia.marconi@ilc.cnr.it (L.M.); paola.cutugno@ilc.cnr.it (P.C.)

* Correspondence: davide.chiarella@cnr.it; Tel.: +39-010-6475220

† Main author.

Received: 30 May 2018; Accepted: 27 July 2018; Published: 1 August 2018



Abstract: The underwater environment is characterized by hazardous conditions that make it difficult to manage and monitor even the simplest human operation. The introduction of a robot companion with the task of supporting and monitoring the divers during their activities and operations underwater can help to solve some of the problems that usually arise in this scenario. In this context, a proper communication between the diver and the robot is imperative for the success of the dive. However, the underwater environment poses a set of technical challenges which are not readily surmountable thus limiting the spectrum from which possibilities can be chosen. This paper presents the design and development of a gesture-based communication language which has been employed for the entire duration of the European project CADDY (Cognitive Autonomous Diving Buddy). This language, the *Caddian*, was built upon consolidated and standardized underwater gestures that are commonly used in recreational and professional diving. Its use and integration during field tests with a remotely operated underwater vehicle (ROV) is also shown.

Keywords: marine robotics; underwater human–robot interaction; gesture-based language; field trials

1. Introduction

Recreational and professional divers generally work in environments that are difficult to monitor and are characterized by severe conditions. In such a context, it is difficult to monitor the status of divers: any sudden episode causing them to deal with an emergency, such as technical problems or human error, may jeopardize the underwater work or even lead to dramatic consequences, which may involve the divers' safety.

In order to avoid and decrease the probability of such events, standard procedures recommend adherence to well-defined rules, for example, to pair up divers. Nevertheless, during extreme diving campaigns, divers' current best practices may not be sufficient to avoid dangerous episodes.

The EU-funded Project CADDY (Cognitive Autonomous Diving Buddy) has been developed with the aim of transferring the robotic technology into the diving world to improve the safety levels during the dives. The main objective of the project is to develop a pair of companion/buddy robots—an Autonomous Underwater Vehicle (AUV) and its counterpart, an Autonomous Surface Vehicle (ASV) (see Figure 1)—to monitor and support human operations and activities during the dive.

In this scenario, one of the major challenges consists in the development of a communication protocol, which allows the diver and the underwater robot to actively interact and cooperate for the fulfillment of the objectives of the mission.

Given the extreme attenuation of high-frequency electro-magnetic waves underwater, medium- to long-range WiFi/radio communication becomes unreliable already at low depths (i.e., 0.5 m), while optical communications are limited by the reverberation of the water and by the scattering caused by suspended debris [1]. The most used and reliable solution for underwater communication is the exploitation of acoustics, with two main disadvantages: the high prices of the devices and the very low data transmission rates [2,3].

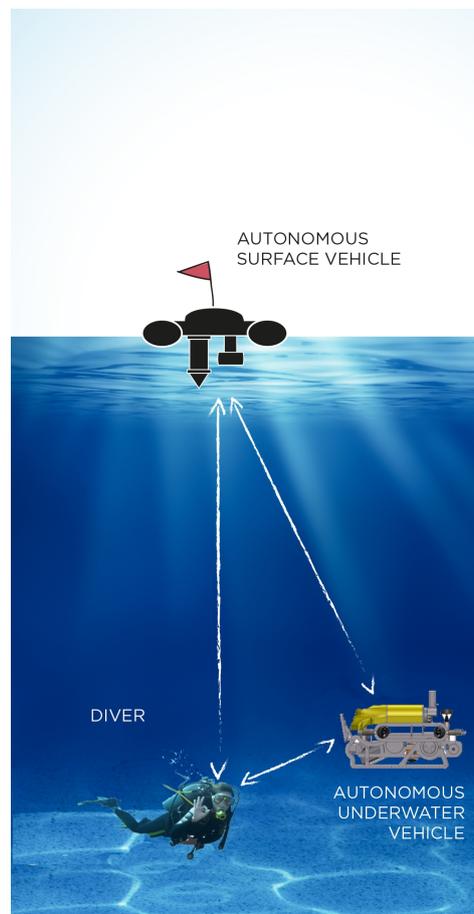


Figure 1. The CADDY (Cognitive Autonomous Diving Buddy) concept.

For all the aforementioned reasons, the solution adopted during the CADDY project has been to develop a novel communication framework with the specific purpose of letting the diver communicate through the most “natural” method available underwater: the gestures. This language, called *Caddian*, has been created as an extension to the established and universally accepted gestures employed by divers worldwide [4–7]. The choice of making the *Caddian*, for all intents and purposes, backward compatible with the current method of communication used by divers has been made in the hope of fostering a widespread adoption among communities of divers.

The field of gesture-based communication between a robotic vehicle and a human being, especially underwater, represents an open challenge in robotics. Very little work has been done.

This paper, with additional results, also extends and completes preliminary work described in [8].

2. State of the Art

The literature on human–robot interaction on dry land shows many languages based on natural language processing and gestures. For example, the authors in [9] use a finite state machine to

develop a speech interaction system. Conversely, the choice of the authors in [10] has fallen on a gesture-based human–robot interface. In this second case, however, the limited set of gestures (only five) limits the usability of the language. The work in [11] uses gestures with two alternative methods: a template-based approach and a neural network approach. As can be seen, the literature on human–robot interaction (HRI) on dry land is abundant and presents several ways to make humans and robots communicate. On the other hand, the literature on HRI languages for underwater environments is not as rich and few works present a formal language described by a formal grammar [12,13]. Among those regarding HRI in underwater environments, we cite authors in [14,15], who developed the RoboChat language providing Backus–Naur form (BNF) productions. However, the language developed was based on fiduciary markers such as ARTags [16] and Fourier tags [17], lacking the simplicity and instinctivity of gestures [18,19]. Furthermore, authors in [20] developed a programming language for AUV with essential instructions for mission control with the given grammar similar to the assembly language: in this case, the interaction between divers and robots is missing and the use of assembly language seems to be overly complex and hard to remember.

Changing perspectives and, instead of focusing on the language, focusing on the robot's ability to understand, a large amount of research (see, for example, [21,22]) has been done in order to provide robots with robust perceiving capabilities, with the aim of making them reactive to the external world and its occurring events. One of the main goals of the robotic field is to obtain a more natural interaction between men and machines without compromising the efficiency and the robustness of the "system" as a whole. To this aim, gesture recognition seems to be the most promising technique, since it is judged to be almost effortless by humans [4–7].

From the gesture recognition point of view, the literature shows many works focused on the problem of exploiting hand gesture recognition algorithms within different contexts [23], such as robotics or computer science. However, these works have almost always been developed for dry land applications, where the environment is simpler than the harsh underwater scenario (e.g., with bubbles, turbid water, etc.) or the task and the working environment are highly simplified. For example, some of them assume that the hand is the only moving object or make sure that the gestures are performed in front of a very neutral and uniform background. Furthermore, "in-air" applications can exploit cameras with extra sensors such as IR systems that are not employable in water or directly exploit RGBD (Red Green Blue Depth) cameras, like in [24].

In this branch of research (i.e., the "in-air" one) many interesting techniques are proposed; approaches based on adaptive skin detection [25–28], motion analysis [29–31], pose classification [32,33], and others are investigated.

Likewise, there are many distinctive problems essentially in relation to robustness and repeatability of the recognition procedure; for many in-air applications, a simple and structured background (e.g., uniform gray or white without other objects in view) is often considered, since many algorithms fail in more complex scenarios (e.g., in the presence of objects similar to the ones to be detected). Moreover, changes in illumination and light highly affect vision techniques, decreasing their robustness and usability in real-world situations. Conditions are even worse in underwater applications: for instance, the in-air consolidated techniques based on skin detection, or more in general relying on color detection (recall that usually divers wear gloves and masks that totally or partially cover their bodies or their faces), are not suitable in water because of hue attenuation. Due to such a phenomenon, color appearance is very different and the usually employed algorithms lose their effectiveness and robustness. Thus, a more complex approach has to be adopted; as an example, a model for light attenuation is considered in [34] and a color registration technique is tested to demonstrate the effectiveness of the overall approach.

In the specific case of divers, since their suite, gloves and all garments are usually black, difficulties can arise while segmenting their body parts: for a posture with the diver hand right in front of the chest, algorithms can be easily deceived and fail in correctly detecting the hand. To this aim, approaches based on stereocamera systems can improve the detection, exploiting the depth information: in an

image like the one above described, the hand and arm of the diver will be slightly closer to the cameras. Indeed, this difference in depth is very small, so the algorithm should be very precise and able to overcome problems badly affecting measurements, such as distortion. Furthermore, another problem strictly related to underwater perception in the presence of humans consists in occlusions due to bubbles generated by divers' breath: the object to be detected can be concealed for many frames, so some sort of prediction and tracking algorithms should be considered. These are some of the additional problems related to the underwater environment that have to be faced and solved by research on underwater perception and that can undermine robustness and repeatability of the robot behavior.

A wide number of different techniques can be exploited and has to be tailored to the specific application: geometric classifiers, Principal Component Analysis (PCA) [35–37], silhouette recognition [38], feature extraction [39], Haar classifiers [40], learning algorithms [41] and so on.

The large variety of works about gesture recognition in the relevant literature testifies to the importance of the development of a robust and effective natural HRI system and indicates that the solution to this problem is still far from being found. Moreover, most of the works presented in the literature are about in-air applications, where the operational conditions present few difficulties: the underwater environment poses many further problems, such as visibility, cloudy water, bubbles occluding the captured scene, illumination, and constraints in the range to be kept between the diver and the robot.

The hostile and harsh underwater environment and the few works addressing the problems of communication in it underline the innovation and utility of the system proposed. Moreover, in a scenario where robots will help divers in their tasks, there is a need for defining and developing a rich language to enable divers to communicate complex commands to their robotic buddies.

This article explains the first implementation and evaluation of the *Caddian* language, namely the phase following the creation of the language from alphabet, syntax and semantics, and the communication protocol that must be followed by the divers to communicate with the AUV. The work is structured as follows. Section 3 presents the definition of *Caddian* language and its communication protocol. In Section 4, the subset of the language gestures used for trials is described, while in Section 5 the robotic platform employed is outlined. Section 6 contains the description of the missions' trials and the BNF syntax of the trial language. Section 7 combines the results from the individual trials. Section 8 describes a study about the language learning curve. Section 9 presents our conclusions.

3. A Gesture-Based Language for Underwater Environments: The *Caddian* Language

3.1. Human–Robot Interaction Based on Gestures

The development of the HRI language *Caddian* is based on divers' sign language. Given the fact that a language has to be easy to learn and to be taught, *Caddian* signs have been mapped with easily writable symbols such as the letters of the Latin alphabet: this bijective mapping function translates signs to our alphabet and vice versa, as depicted in Figure 2.

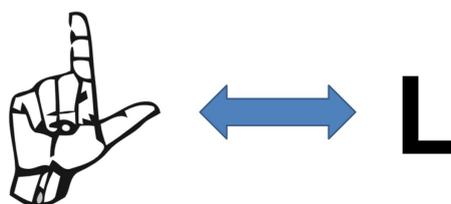


Figure 2. Bijective function mapping from gestures to letters.

Sequences of *Caddian* gestures and the corresponding sequences of characters of the written alphabet (i.e., Σ) are mapped to a semantic function that translates them into commands/messages.

A classifier encodes/decodes gestures, which should be feasible in the underwater environment and as intuitive as possible to cope both with the learning aspect of the language and divers’ acceptance. The more dimensions can be discerned by the classifier, the more gestures can be used:

- if it is able to extract and match features from both hands, the amount of recognizable gestures increases (for an example of two-handed gesture, see the “boat” signal in [5]);
- if it is able to extract features and match them in the time domain, thus being able to classify hand gestures with motion, the gestures alphabet becomes richer (for an example of motion gesture, see the “something is wrong” gesture in [5]).

In the creation of the language, the issued sentences have been sequentially defined to allow the synchronization of the recipient with the issuer, and have been defined with boundaries to ensure efficient interpretation: the “start communication” and “end of communication” gestures enclose, as the name says, the communication, while the “start communication” gesture is also used as delimiter between a message and the following one during a complex communication.

3.2. A Specialized Language

Caddian is a language for communicating between divers and underwater robots, in particular autonomous underwater vehicles (AUVs), so the list of commands/messages defined by the language in the scope of the project is strictly context-dependent. Currently, there are 40 implemented commands. The commands/messages (see Table 1 and [8]) are separated into six groups: Problems (9), Movement (5), Setting Variables (10), Interrupt (4), Feedback (3), and Works/Tasks (9).

Table 1. List of commands as seen in [8].

Group	Commands/Messages	
Problems	I have an ear problem I’m out of air [air almost over] I depleted air I’m cold I have vertigo	I’m out of breath Something is wrong [diver] Something is wrong [environment] I have a cramp
Movement	Take me to the boat Take me to the point of interest Go X Y $X \in Direction$ $Y \in \mathbb{N}$	You lead (I follow you) I lead (you follow me) Return to/come X $X \in Places$
Interrupt	Stop [interruption of action] Let’s go [continue previous action]	Abort mission General evacuation
Setting Variables	Keep this level (actions are carried out at this level) Level Off (AUV cannot fall below this level) Set point of interest Give me light (switch on the on board lights) No more light (switch off the on board lights)	Free level (“Keep this level” command does not apply any more) Slow down/Accelerate Give me air (switch on the on board oxygen cylinder) No more air (switch off the on board oxygen cylinder)
Feedback	No (answer to repetition of the list of gestures) Ok (answer to repetition of the list of gestures)	I don’t understand (repeat please)
Works	Wait n minutes $n \in \mathbb{N}$ Tell me what you’re doing Carry a tool for me Do this task or list of task n times $n \in \mathbb{N}$ Tessellation of point of interest/boat/here	Tessellation X * Y area $X, Y \in \mathbb{N}$ Photograph of X * Y area $X, Y \in \mathbb{N}$ Stop carrying the tool for me [release] Photograph of point of interest/boat/here
$Direction = \{ahead, back, left, right, Up, Down\}$		
$Places = \{pointofinterest, boat, here\}$		

3.3. Communication Protocol and Error Handling

The *Caddian* language is used inside a communication protocol that guarantees error handling. The protocol is designed having in mind a strict cooperation between the diver and the AUV: for example, we can mention the possibility for the diver to query the robot at any time about the progress of a task with the purpose of understanding whether it has been executed.

For these reasons, AUV is equipped with three light emitters (red, green, and orange):

- green = IDLE STATE — everything is ok, all tasks have been accomplished and I'm waiting for orders;
- orange = BUSY STATE — everything is ok and I'm working the last mission received;
- red = FAILURE STATE — a system failure has been detected or an emergency has been issued.

The *Caddian* protocol handles these three possible types of error:

1. the AUV does not recognize a gesture inside a sequence: the robot shows an error message and the diver repeats the gesture, but the sequence is not aborted. This allows the diver to make mistakes and, in such cases, to save time avoiding to repeat the whole sequence.
2. the AUV recognizes the sequence, but the resulting command is not semantically correct. When the whole sequence is issued, a semantical error message is shown: the sequence of gestures is aborted and must be repeated.
3. the AUV recognizes the sequence and the resulting command is semantically correct, but it is not what the diver intended. This type of error is more subtle, because it involves a semantical evaluation that only the diver is able operate with the necessary swiftness.

To deal with this last category of errors, we have to introduce the definition of "mission." As a sequence is issued to the buddy AUV, before it can turn into a real mission, the AUV repeats the sequence, writing it in plain text on its screen and waiting for the diver's confirmation. At this point, the diver simply accepts the sequence showing thumbs up, thus letting the mission start, or he refuses the sequence with thumbs down: that sequence won't turn into a mission and thus it will never be translated into a series of actions. Moreover, it would be very hard (or even impossible) for the AUV to guess where the error was in the sequence; therefore, if the diver does not confirm the sequence, the only feasible approach is to make the diver repeat it from the beginning.

Regarding the robot's mission status, two accessibility aspects were also considered crucial while creating the *Caddian* language:

1. Divers should always understand if the assigned mission has been terminated.
2. Divers should always be able to know the progress of a mission.

In the first case, the buddy AUV just turns on the idle status (i.e., green light) and remains stationary. In the second case, the proposed behavior is the following:

- the buddy is in operation executing a task;
- the diver approaches the buddy, facing it, to be clearly visible on both the camera and the sonar;
- for safety reasons, the diver always remains outside a predefined safety range (e.g., 2 m). If closer, the buddy is programmed to automatically back off from the diver.
- if all the above conditions are met, the buddy AUV suspends the current action, remaining however in the BUSY state.

In this situation, a diver can

- query the AUV on the mission's progress with the "Check" command;
- erase the current mission with the "Abort mission" command;
- report an emergency using a command belonging to the "Problems" subset;
- leave the range of safety, letting the AUV return to the assigned mission.

3.4. Language Definition

The diver-robot language has been defined as a formal language. A formal grammar can describe a formal language [12,13] and can be represented as a quadruple $\langle \Sigma, N, P, S \rangle$ as follows:

- a finite set Σ of terminal symbols (disjoint from N), the alphabet, which are assembled to form the sentences of the language;

- a finite set N of non-terminal symbols or variables or syntactic categories, which represents some collection of subphrases of the sentences;
- a finite set P of rules or productions which describe how each non-terminal is defined in terms of terminal symbols and non-terminals. Each production has the form $B \rightarrow \beta$, where B is a non-terminal and β is a string of symbols from the infinite set of strings (ΣUN);
- a differentiated non-terminal S , the start symbol, which specifies the principal category being defined, such as a sentence, a program, or a mission.

This said, the language L_G (i.e., generated by grammar G) can be formally defined as the set of strings composed of terminal symbols that can be derived through productions from the start symbol S .

$$L_G = \{s/s \in \Sigma^* \text{ and } S \rightarrow^* s\}. \quad (1)$$

For the *Caddian* language, the signs of the alphabet Σ are the set of letters belonging to the Latin alphabet mixed with some complete words, Greek letters, math symbols, and natural numbers (also used as subscripts) defined as follows:

$$\Sigma = \{A, \dots, Z, ?, \text{const}, \text{limit}, \text{check}, \dots, 1, 2, \dots\}. \quad (2)$$

By definition, the grammar of *Caddian* is a context-free grammar because on the left side of the productions only non-terminal symbols and no terminal symbols can be found [42,43]. In addition, the resulting language is an infinite language given that the first production (i.e., $\langle S \rangle$) uses recursion and the dependency graph of the non-terminal symbols contains a cycle.

3.5. Syntax

Syntax has been given through the following BNF productions:

```

⟨S⟩ ::= A ⟨α⟩ ⟨S⟩ | ∇
⟨α⟩ ::= ⟨agent⟩ ⟨m-action⟩ ⟨object⟩ ⟨place⟩ | b⟨feedback⟩ ⟨p-action⟩ ⟨problem⟩ | ⟨set-variable⟩ | ⟨feedback⟩
      | ⟨interrupt⟩ | ⟨work⟩ | ∅ | Δ
⟨agent⟩ ::= I | Y | W
⟨m-action⟩ ::= T | C | D | F | G ⟨direction⟩ ⟨num⟩
⟨direction⟩ ::= forward | back | left | right | up | down
⟨object⟩ ::= ⟨agent⟩ | Λ
⟨place⟩ ::= B | P | H | Λ
⟨problem⟩ ::= E | C1 | B3 | Pg | A1 | K | V | Λ
⟨p-action⟩ ::= H1 | B2 | D1 | Λ
⟨feedback⟩ ::= ok | no | U | Λ
⟨set-variable⟩ ::= S ⟨quantity⟩ | L ⟨level⟩ | P | L1 ⟨quantity⟩ | A1 ⟨quantity⟩
⟨quantity⟩ ::= + | -
⟨level⟩ ::= const | limit | free
⟨interrupt⟩ ::= Y ⟨feedback⟩ D
⟨work⟩ ::= Te ⟨area⟩ | Te ⟨place⟩ | Fo ⟨area⟩ | Fo ⟨place⟩ | wait ⟨num⟩ check | ⟨feedback⟩ carry | for
      ⟨num⟩ ⟨works⟩ end | Λ
⟨works⟩ ::= ⟨work⟩ ⟨works⟩ | Λ
⟨area⟩ ::= ⟨num⟩ ⟨num⟩ | ⟨num⟩
⟨num⟩ ::= ⟨digit⟩ ⟨num⟩ | Ψ

```

$\langle digit \rangle ::= 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \mid 0$

With the given syntax, we can translate the previously identified messages and commands and obtain a translation table (Table 2) [8].

Table 2. Translation table as seen in [8].

	Message/Command	Caddian	
Problems	Ear problem	$A \ bH_1 E \forall$	
	Out of breath	$A \ bB_2 B_3 \forall$	
	Out of air [air almost over]	$A \ bB_2 A_1 \forall$	
	Something is wrong [diver]	$A \ bH_1 P_g \forall$	
	Air depleted	$A \ bD_1 A_1 \forall$	
	Something is wrong [environment]	$A \ bP_g \forall$	
	I'm cold	$A \ bH_1 C_1 \forall$	
	I have a cramp	$A \ bH_1 K \forall$	
	I have vertigo	$A \ bH_1 V \forall$	
	Movement	Take me to the boat	$AYTMB \forall$ $AIFYAYCB \forall$
		You lead (I follow you)	$AIFY \forall$
Take me to the point of interest		$AYTMP \forall$ $AIFYAYCP \forall$	
I lead (you follow me)		$AYFM \forall$	
Go X Y		$AYG \ Directions \ n \forall$	
$X \in \ Directions$ and $Y \in \mathbb{N}$		$n \in \mathbb{N}$	
Return to/come X		$AYCP \forall$	
$X \in \ Places$		$AYCB \forall$ $AYCH \forall$	
Interrupt		Stop [interruption of action]	$AY \ no \ D \forall$
		Let's go [continue previous action]	$AY \ ok \ D \forall$ or $AYD \forall$
		Abort mission	$A \emptyset \forall$
	General evacuation	$A \ \Delta \ \forall$	
Setting Variables	Slow down	$AS - \forall$	
	Accelerate	$AS + \forall$	
	Set point of interest	$AP \forall$	
	Level Off	$AL \ limit \ \forall$	
	Keep this level	$AL \ const \ \forall$	
	Free level	$AL \ free \ \forall$	
	Give me air	$AA_1 + \forall$	
	No more air	$AA_1 - \forall$	
	Give me light	$AL_1 + \forall$	
	No more light	$AL_1 - \forall$	
Feedback	No	$A \ no \ \forall$	
	Ok	$A \ ok \ \forall$	
	I don't understand (repeat please)	$AU \ \forall$	
Works	Wait n minutes $n \in \mathbb{N}$	$A \ wait \ n \ \forall$	
	Tessellation X * Y area	$ATe \ n \ m \ \forall$	
	$X, Y \in \mathbb{N}$	$ATe \ n \ \forall$ [square]	
	Tessellation of point of interest/boat/here	$ATe \ P \ \forall$	
	Tell me what you're doing	$A \ check \ \forall$	
	Photograph of X * Y area $X, Y \in \mathbb{N}$	$AFo \ n \ m \ \forall$	
		$AFo \ n \ \forall$ [square]	
	Take a picture of point of interest/boat/here	$AFo \ P \ \forall$	
	Carry a tool for me	$A \ carry \ \forall$	
	Stop carrying the tool for me [release]	$A \ no \ carry \ \forall$	
	Do this task or list of task n times $n \in \mathbb{N}$	$A \ for \ n \ \dots \ end \ \forall$	
$Directions = \{ahead, back, left, right, Up, Down\}$			
$Places = \{pointofinterest, boat, here\}$			

3.6. Semantics

The communication scenarios and relating commands can be divided into six groups. In the following paragraphs, these commands sets are briefly explained.

Problems—This category of messages refers to issues happening to the diver or to the environment around the operating area. All productions contain the *b* symbol, which intrinsically denotes that there is an emergency and that any action is being executed needs to be suspended to take care of the issue.

Movement—This category of commands makes the robot move or tells the robot how to move.

Interrupt—This category of commands makes the robot stop the current task/mission. The “general evacuation” command has a special meaning: in fact, this command makes the buddy AUV abort the mission and makes it emit any possible warning signal both to the surface and to any diver in the operation area (e.g., it turns on flashing red lights and sends emergency messages through the acoustic link towards the surface vehicle).

Setting Variables—These commands set internal variables inside the robot. At this moment, there are eight of them, but only seven can be set directly by the diver (see below).

- *Speed*: the robot speed has discrete values. With the “+” or “-” signs, the diver increases or decreases this variable by a quantum.
- *Level*:
 - constant: any following command is carried out at the current level of depth;
 - off: the buddy AUV cannot move below the actual depth: if a subsequent command tries to force the buddy AUV to break this rule, the robot interrupts the mission. This behavior has been thought as a safety measure mainly for the buddy AUV (and for the diver as a direct consequence) but it may be also useful in specific scenarios, for example, during the exploration of archaeological sites;
 - free: clears previous statuses set by other commands which refer to the level of depth: the AUV is now free to move up and down underwater.
- *Point of Interest*: set a single point of interest, to be recalled later within other commands.
- *Light*: this is a binary variable which switches on or off the vehicle lights.
- *Air*: this is a binary variable which toggles on or off the vehicle’s onboard oxygen cylinder.
- *Here*: store the actual 3D coordinates of the position in memory (i.e., where the buddy is located while the command is being issued, and its yaw angle—where the buddy is facing).
- *Boat*: boat or base position, which cannot be set by the diver.

Communication Feedback—These commands refer to the communication feedback between the diver and the AUV. The diver can accept or reject a previously issued command (see Section 3.3) and can also ask the AUV to repeat the command if he did not comprehend it (by accident or by distraction).

Works—These commands refer to tasks the robot is able to do. The “tell me what you’re doing” command (i.e., check the mission progress) can be used when the diver approaches the robot (and the robot consequently pauses anything it is doing). The “wait X minutes” command instructs the robot to float and wait X minutes then proceed with the next command (useful to pause the mission or to let the seafloor dust settle down). The “carry a tool for me” command instructs the robot to carry equipment upon diver request: after the equipment has been placed into the robot compartment, the AUV waits for a physical confirmation (i.e., a button to press to give physical feedback).

4. Outline of Gestures Used during Trials

A diver’s underwater gestures are not formalized in any international standard. In fact, all organizations and diving agencies worldwide teach divers their own subset of diving hand signals, causing some gestures to vary from region to region: in this paper, the most famous and common

ones [4–7] were chosen, carefully picking both from the ones used by largest diving organizations and from gestures akin to natural or instinctive meaning. *Caddian* gestures were also chosen following the two most important requirements of the language: all gestures should be feasible and as easy as possible to perform underwater and they should be as intuitive as possible to make them easy to remember and to render the language truly effective.

In some cases, whenever possible, basic mnemonic techniques have been exploited, associating a gesture to objects related to the action it expresses: for example, in the “take a picture” gesture, the diver shows just three fingers, which can be mnemonically associated to the tripod used to stabilize and elevate a camera.

The list in Figures 3 and 4 contains a subset of gestures used during CADDY trials. As can be seen, a significant amount of task gestures and all the natural numbers are recognized, but we eventually decided not to introduce dynamic gestures, because the effort to recognize them through computer vision was too high with respect to the expected benefits.

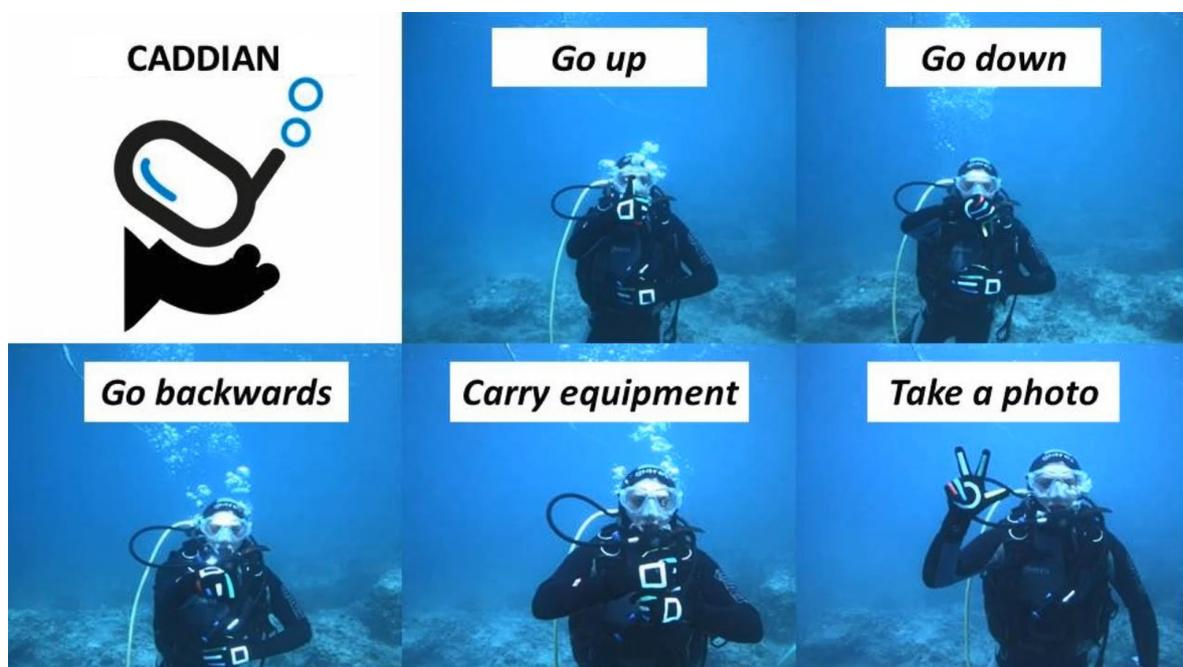


Figure 3. A subset of the *Caddian* gestures tested during 2015 validation trials.



Figure 4. Trials gestures: numbers from 1 to 5.

Mapping Gestures to Syntax; Syntax to Semantics

As already said, a bijective mapping function translates from the domain of signs to our alphabet and vice versa (Figure 5). Accordingly, one or more gestures and the corresponding characters are also mapped to a semantic function that translates them into commands/messages.

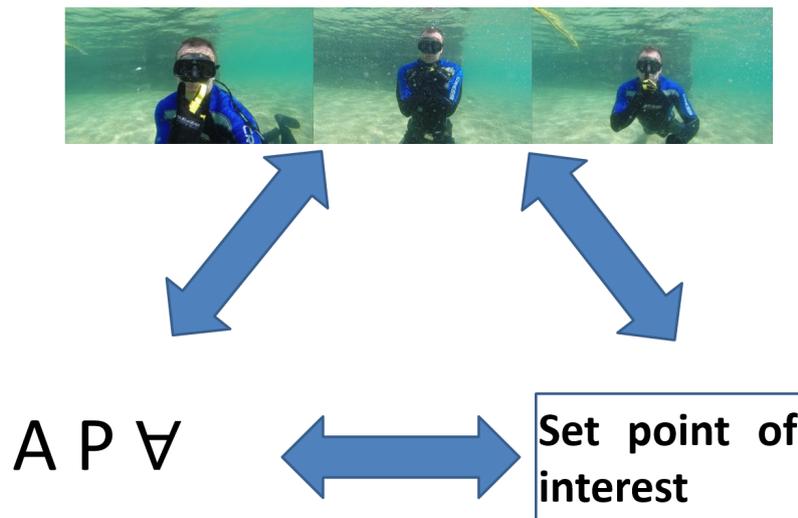


Figure 5. Gestures, written alphabet, and semantics.

5. Outline of Trial Vehicle: R2 ROV

At-field trials have been carried out employing the robotic platform R2 ROV/AUV design and developed by CNR-ISSIA (Figure 6); the R2 underwater vehicle is the product of a retrofit process of the former Romeo ROV, built and developed during the 1990s. R2 is characterized by an open-frame structure and a full-actuated motion capability. Thanks to its compact size, 1.3 m long, 0.9 m wide, and 1.0 m tall, it is still considered a small-/medium-class ROV/AUV. Depending on the specific payload for each mission (dedicated sensor package, manipulation systems, etc.), the total weight can vary from 350 to 500 Kg in-air. The overall motion control is provided by a redundant and fault-tolerant thruster allocation with four vertical thrusters for vertical motion and four horizontal thrusters for the 2D horizontal positioning. In ROV mode, a fiber-optic based link provides real-time data transfer for both direct piloting/control/supervision of the robot, as well as on-line data gathering and analysis.

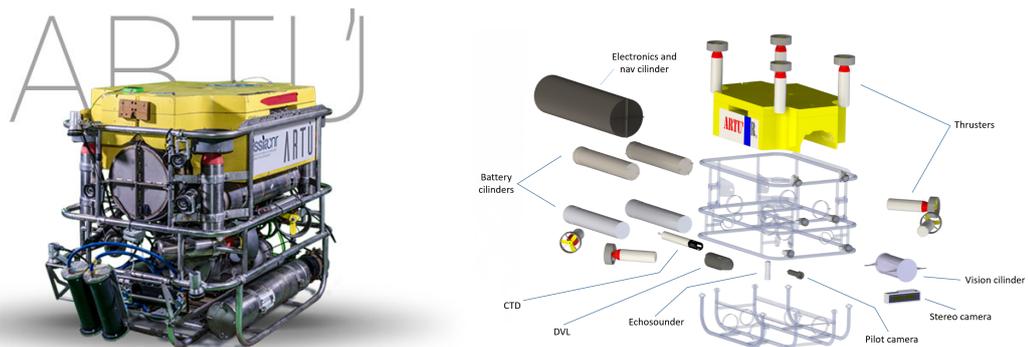


Figure 6. R2 (Artù) ROV. (left) a picture; (right) CAD design with sensors and actuators.

The R2 ROV/AUV during the trials was equipped with:

- IMU (Inertial Measurement Unit): Quadrans 3-axis Fiber Optic Gyro and 3DM-GX3-35 MicroStrain AHRS (Attitude Heading Reference System);
- GPS (Global Positioning System): 3DM-GX3-35 MicroStrain AHRS (Attitude Heading Reference System);
- Stereo camera: BumbleBee XB3 13SC-38 3 sensor multi-baseline color camera;
- CTD (Conductivity, Temperature, and Depth sensors) : OceanSeven 304 Plus;

- Sonars: 2 Tritech PA500 echosounders (1 for seafloor detection and 1 for front obstacle detection);
- Lights: 6 front-mounted LED-based (Light-emitting diode) high intensity spot lights.

6. Trials at Sea and at Pool

Two experimental campaigns were carried out in 2015 [44,45], one in Biograd Na Moru (Croatia), at sea, in October and the second one in Genova (Italy), at pool, in November. Both campaigns were focused on the validation of the interaction capabilities of the robot, mostly related to the gesture recognition and compliant robot reactions to the desired commands. Five professional divers were involved, respectively four for the Biograd Na Moru campaign and one for the Genova one. All the divers were trained for about half an hour before the first dive, as the set of gestures used was minimal (i.e., 22 gestures) and it was mostly a subset of the “common gestures set” already used in the diving world (i.e., 15 out of 22). The Biograd Na Moru’s trials were made in open sea at a 4 m depth over one week. Depending on the day, trials were carried out with the presence of currents or at calm sea. During the Genova campaign, trials were made at a 3-m-deep outdoor pool. Given that these trials were only focused on preliminary functional tests of the computer vision algorithms and aimed at gathering as much data as possible for their further refinement, the R2 ROV (see Figure 6), described in Section 5, was employed as the buddy AUV. According to the envisioned use-case scenarios of the CADDY project, trials were made up of four kinds of mission.

- **Movement Missions**—In this kind of mission, the diver issues a movement command with a number (i.e., “Go up 1 m”).
- **“Take a photo” Missions**—In this kind of mission, the diver commands the AUV to take a picture from the point where it is stationing.
- **“Do a mosaic” Missions**—In this kind of mission, the diver commands the AUV to do a mosaic/tessellation of an area $n \times m$ of the seabed (see Section 3 under Works).
- **Complex Missions**—In this kind of mission, the diver commands the AUV to go to the boat and bring back a tool.

During the trials, a minimal subset of a modified version of *Caddian* has been used. Here, the syntax of the minimal language is followed:

$$\begin{aligned} \langle S \rangle &::= A \langle \alpha \rangle \langle S \rangle \mid \forall \\ \langle \alpha \rangle &::= \langle direction \rangle \langle num \rangle \mid \langle place \rangle \mid \langle work \rangle \\ \langle direction \rangle &::= forward \mid back \mid up \mid down \\ \langle place \rangle &::= B \mid H \mid \Lambda \\ \langle work \rangle &::= Te \langle area \rangle \mid Fo \mid carry \mid \Lambda \\ \langle area \rangle &::= \langle num \rangle \langle num \rangle \mid \langle num \rangle \\ \langle num \rangle &::= \langle digit \rangle \langle num \rangle \mid \Psi \\ \langle digit \rangle &::= 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9 \mid 0 \end{aligned}$$

Applying the syntax rules, the commands for each mission were as follows:

- **Movement missions:**
 - “Go up 1 m”: $A \text{ up } 1 \Psi \forall$
 - “Go down 1 m”: $A \text{ down } 1 \Psi \forall$
 - “Go back 1 m”: $A \text{ back } 1 \Psi \forall$
 - “Go forward 1 m”: $A \text{ forward } 1 \Psi \forall$
- **“Take a photo” mission:** $A \text{ Fo } \forall$
- **“Do a mosaic” mission:** $A \text{ Te } 2 \Psi 4 \Psi \forall$

- **Complex mission:** A B A carry A H ∇

The complex mission deserves a special mention because in its syntax we can observe three concatenated commands between the delimiters (i.e., “A” and “∇”). The first one is “go to the boat” (“B”), followed by “carry a tool” (“carry”) and then “return here” (“H”). Consequently, the robot moves to the boat, opens the compartment to hold the tool requested, and returns to the point where the entire mission has been started.

7. Results

The first phase of the trials has been focused on simple gesture recognition, namely the recognition of single gestures by the robot and the execution of the associated command. Divers performing the gestures are shown in the following figures. Figure 3 shows a sample of the single-gestured commands used during the trials, while Figures 7–9 show gesture recognition in different kinds of environment to emphasize the different challenges that had to be solved to have correctly classified gestures. For example, Figure 7 shows the diver in different kinds of water at varying distance from the camera: the green and yellow circles shows that the recognition process has been successful for all four images.

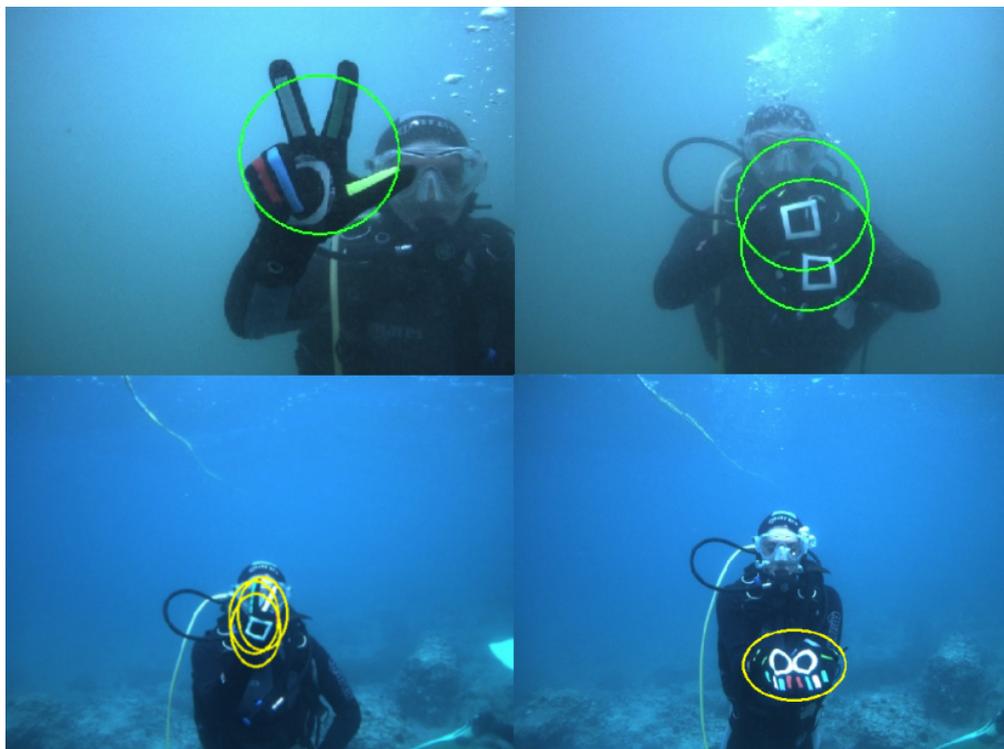


Figure 7. Detected gestures, from the top left, clockwise: take a photo, carry equipment, start communication, and go to the boat.



Figure 8. “Go down 2 m” gesture sequence executed during trials.



Figure 9. CADDY simple gesture recognition in a harsh underwater environment.

Extensive tests focused on the complex gesture sequence recognition were carried out during the second phase of the trial. Complex sequences are enriched sets of gestures that represent dialogues or sentences containing more information that the robot has to recognize, separate, and interpret in order to achieve the desired goal requested by the diver. Examples of complex sequences are “go to boat and carry equipment here” or “execute a mosaic of $N \times M$ meters.” Complex sequences have been tested 15 times each in order to evaluate the reliability and success rate of the system, obtaining good results in terms of recognition capabilities (i.e., all complex sequences had a success rate which varies from 87 to 99.9% except for the “do a mosaic” mission, which achieved 60%). An example of a complex sequence executed by a diver is the “Go down 2 m”, depicted in Figure 8. From Figure 9, it is possible to appraise the harsh conditions that the CADDY system has to face while working in the underwater environment. The water visibility level can very badly affect the gesture interaction, as well as the illumination condition and gesture occlusions due to bubbles. The CADDY system was successful in the interaction through gestures even during days with a low visibility condition (2.5 m with suspended sediments) and with a strong current and strong waves; the robot lit up its lights to signal to the diver that his gestures were recognized and the command executed.

8. Cognition and Ease of Language Learning: Evaluation on Dry Land

The same gestures and relative missions performed during the trials were tested on dry land in order to evaluate the language learning curve and the language cognitive load in divers. The research team tested 22 volunteers: all of them completed the trials. Seven out of 22 volunteers were female and 4 out of 22 had previous diver experience. No replication of the experiment with the same volunteer has been made and missions have been provided sequentially with no order randomization. The evaluation of the language took place in the following stages: first, the language has been explained to the volunteers (the explanation had a duration of about five minutes); afterward, the volunteers were asked to repeat the six trials missions, which were as follows:

- **Mission 1:** “Go up 1 m”;
- **Mission 2:** “Go down 1 m”;
- **Mission 3:** “Go back 1 m”;
- **Mission 4:** “Take a photo”;
- **Mission 5:** “Do a mosaic”;
- **Mission 6:** “Go to boat, bring me something (carry equipment), come back here.”

Figure 10 describes the results of this evaluation. As can be observed, the error value decreases with the progress of the missions, proving that the volunteers learnt the language very quickly, given the fact that only five minutes were dedicated to the explanation of the whole language used in trials, made up of 22 gestures.

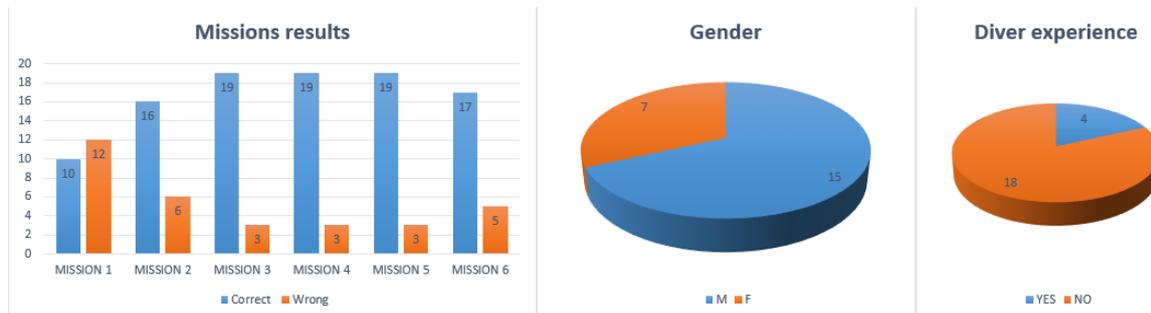


Figure 10. Evaluation of the six missions and description of volunteers data set.

The observed errors in the mission enunciation were as follows:

- **CLOSE_NUM_1:** forgetting “close number” when issuing the mission.
- **NUMBER_ONE:** confusing/swapping the “number one” with “close communication.”
- **UP:** forgetting the “UP” gesture and issuing only the meters to go. The error of forgetfulness of a gesture appears only in the first and third mission. This error is supposed to be due to the initial nervousness of the candidate.
- **CLOSE_COMM:** confusing/swapping the “close communication” with “number one.”
- **DOWN:** wrong orientation of the hand.
- **CLOSE_NUM_2:** confusing/swapping the “close number” with “close communication.”
- **BACKWARDS:** forgetting the “BACKWARDS” gesture and issuing only the meters to go. The error of forgetfulness of a gesture appears only in the first and third missions. This error is supposed to be due to the nervousness of the candidate.
- **TAKE_PHOTO:** wrong orientation of the hand.
- **MOSAIC:** wrong orientation of the hands.
- **NUMBERS:** wrong orientation of the hand.
- **BOAT_CARRY:** the candidate did not remember the gestures for “boat” or “carry.”
- **HERE_1:** the candidate did not remember the gestures for “come here.”
- **START_MSG:** candidate uses, in the complex mission, the “start message” gesture to close the communication. This may also be an indication for a further improvement of the language where a single gesture is used to both open and close the communication. In this case, the interarrival time could be used as information to separate one communication from the following one.
- **HERE_2:** confusing/swapping the “here” gesture with “go backwards.”

A more detailed view of the errors can be seen in Table 3.

Table 3. Summary of the observed errors.

Gesture	Semantic	Type of Error	Example	Occurrences
Ψ	Close number	CLOSE_NUM_1	A up 1 ∇	7
1	Number one	NUMBER_ONE	A up ∇ Ψ ∇	4
up	Go up	UP	A 1 Ψ ∇	2
∇	Close communication	CLOSE_COMM	A up 1 Ψ 1	5
down	down	DOWN	Wrong orientation	1
Ψ	Close number	CLOSE_NUM_2	A up 1 ∇ ∇	1
backwards	Go backwards	BACKWARDS	A 1 Ψ ∇	1
Fo	Take a photo	TAKE_PHOTO	Wrong orientation	3
Te	Do a mosaic	MOSAIC	Wrong orientation	2
2,4	Number two and four	NUMBERS	Wrong orientation	1
B	Go to the boat	BOAT_CARRY	A A A H ∇	1
H	Come back here	HERE_1	A B A carry A ∇	2
A	I’m starting a message	START_MSG	A B A carry A H A	1
H	Come back here	HERE_2	A B A carry A backwards ∇	1

9. Conclusions

In this paper, a novel gesture-based language for underwater human–robot interaction (UHRI) has been proposed: the description of the language, called *Caddian*, is provided with alphabet, syntax, semantics, and a communication protocol. The presented work has mostly focused on the definition of the language and on showing its potential and likely acceptance by the diving community. A description of the performed trials using a minimal modified subset of the language has been reported and is preliminary, but encouraging results are provided.

The classifier performed quite well both in normal and in stormy weather, thus testifying that gestures outside the standardized ones have been chosen correctly. Moreover, divers learnt the language very quickly, showing that the associated cognitive load on the divers is acceptable.

However, the trials also showed that the success rate of the classification can still be improved for some gestures (e.g., “do a mosaic”), and the syntax of natural numbers for humans is not as intuitive as preliminarily designed. Tests with the whole set of gestures have to be made because dynamic gestures, which were not involved during these trials, did not have the same rate of success as the static ones.

Regarding the characteristics of the current version of the language, it would be worth considering adding the ability of changing a mission with an updated parameter (for instance, a new “here” position), or setting macros and then using parameterized missions (i.e., functions).

A deeper study of the use of the whole language will be the next steps of the research, which might involve new scenarios with a consequent creation of new commands; more results about the performance of the classifier must be collected to prove the CADDY framework robust. On the other hand, acceptance by divers of the language will be taken into account. Through the study of their feedback, the language might be changed accordingly.

Author Contributions: Conceptualization, D.C., M.C., L.M. and P.C.; data curation, D.C., M.B. and A.R.; formal analysis, D.C.; funding acquisition, G.B., M.C., L.M. and P.C.; investigation, D.C., M.B., A.R. and E.Z.; methodology, D.C. and M.B.; project administration, M.C. and L.M.; resources, G.B. and M.C.; software, D.C., M.B., A.R. and E.Z.; supervision, G.B., M.C., L.M. and P.C.; validation, D.C., M.B., A.R. and E.Z.; visualization, D.C., M.B., A.R. and E.Z.; writing—original draft, D.C.; writing—review & editing, D.C., M.B., A.R. and E.Z.

Funding: The research leading to these results received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 611373.

Acknowledgments: The authors would like to thank Giorgio Bruzzone and Edoardo Spirandelli for their invaluable assistance during trials and Mauro Giacomelli for the *Caddian* language pictures. We would also like to show our gratitude to M.Sc. Angelo Odetti and M.Sc. Roberta Ferretti for their support. We are also immensely grateful to M.Sc. Arturo Gomez Chavez who developed the vision system of CADDY and Dr Luca Caviglione who provided insight and expertise that greatly assisted the writing of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Cu, J.H.; Kong, J.; Gerla, M.; Zhou, S. The challenges of building mobile underwater wireless networks for aquatic applications. *IEEE Netw.* **2006**, *20*, 12–18.
2. Kilfoyle, D.; Baggeeroer, A. The state of the art in underwater acoustic telemetry. *IEEE J. Ocean Eng.* **2000**, *25*, 4–27. [[CrossRef](#)]
3. Neasham, J.; Hinton, O. Underwater acoustic communications—How far have we progressed and what challenges remain. In Proceedings of the 7th European Conference on Underwater Acoustics, Delft, The Netherlands, 5–8 July 2004.
4. Confédération Mondiale des Activités Subaquatiques. Segni Convenzionali CMAS. Online pdf, 2003. Available online: <https://www.cmas.ch/docs/it/downloads/codici-comunicazione-cmas/it-Codici-di-comunicazione-CMAS.pdf> (accessed on 25 May 2018).

5. Recreational Scuba Training Council. Common Hand Signals for Recreational Scuba Diving. Online Pdf, 2005. Available online: <http://www.neadc.org/CommonHandSignalsforScubaDiving.pdf> (accessed on 25 May 2018).
6. Scuba Diving Fan Club. Most Common Diving Signals. HTML Page. 2016. Available online: http://www.scubadivingfanclub.com/Diving_Signals.html (accessed on 25 May 2018).
7. Jorge, M. Diving Signs You Need to Know. HTML Page. 2012. Available online: <http://www.fordivers.com/en/blog/2013/09/12/senales-de-buceo-que-tienes-que-conocer/> (accessed on 25 May 2018).
8. Chiarella, D.; Bibuli, M.; Bruzzone, G.; Caccia, M.; Ranieri, A.; Zereik, E.; Marconi, L.; Cutugno, P. Gesture-based language for diver-robot underwater interaction. In Proceedings of the OCEANS 2015—Genova, Genoa, Italy, 18–21 May 2015; pp. 1–9. [CrossRef]
9. Tao, Y.; Wei, H.; Wang, T. A Speech Interaction System Based on Finite State Machine for Service Robot. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12–14 December 2008; Volume 1, pp. 1111–1114. [CrossRef]
10. Xu, Y.; Guillemot, M.; Nishida, T. An experiment study of gesture-based human-robot interface. In Proceedings of the 2007 IEEE/ICME International Conference on Complex Medical Engineering, Beijing, China, 23–27 May 2007; pp. 457–463. [CrossRef]
11. Waldherr, S.; Romero, R.; Thrun, S. A Gesture Based Interface for Human-Robot Interaction. *Auton. Rob.* **2000**, *9*, 151–173. [CrossRef]
12. Chomsky, N. Three models for the description of language. *IRE Trans. Inf. Theory* **1956**, *2*, 113–124. [CrossRef]
13. Backus, J.W. The Syntax and Semantics of the Proposed International Algebraic Language of the Zurich ACM-GAMM Conference. In Proceedings of the International Conference on Information Processing, UNESCO, Paris, France, 15–20 June 1959.
14. Dudek, G.; Sattar, J.; Xu, A. A Visual Language for Robot Control and Programming: A Human-Interface Study. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007; pp. 2507–2513.
15. Xu, A.; Dudek, G.; Sattar, J. A natural gesture interface for operating robotic systems. In Proceedings of the 2008 IEEE International Conference on Robotics and Automation, ICRA, Pasadena, CA, USA, 19–23 May 2008; pp. 3557–3563.
16. Fiala, M. ARTag, a Fiducial Marker System Using Digital Techniques. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; IEEE Computer Society: Washington, DC, USA, 2005; Volume 5, pp. 590–596. [CrossRef]
17. Sattar, J.; Bourque, E.; Giguere, P.; Dudek, G. Fourier tags: Smoothly degradable fiducial markers for use in human-robot interaction. In Proceedings of the Fourth Canadian Conference on Computer and Robot Vision (CRV '07), Montreal, QC, Canada, 28–30 May 2007; pp. 165–174. [CrossRef]
18. Islam, M.J.; Ho, M.; Sattar, J. Dynamic Reconfiguration of Mission Parameters in Underwater Human-Robot Collaboration. *arXiv* **2017**, arXiv:1709.08772.
19. Xu, P. Gesture-based Human-robot Interaction for Field Programmable Autonomous Underwater Robots. *arXiv* **2017**, arXiv:1709.08945.
20. Kim, B.; Jun, B.H.; Sim, H.W.; Lee, F.O.; Lee, P.M. The development of Tiny Mission Language for the ISiMI100 Autonomous Underwater Vehicle. In Proceedings of the OCEANS 2010 MTS/IEEE SEATTLE, Seattle, WA, USA, 20–23 September 2010; pp. 1–6. [CrossRef]
21. Garg, P.; Aggarwal, N.; Sofat, S. Vision based hand gesture recognition. *World Acad. Sci. Eng. Technol.* **2009**, *49*, 972–977.
22. Manresa, C.; Varona, J.; Mas, R.; Perales, F. Hand tracking and gesture recognition for human-computer interaction. *Electr. Lett. Comput. Vis. Image Anal.* **2005**, *5*, 96–104. [CrossRef]
23. Rautaray, S.S.; Agrawal, A. Vision Based Hand Gesture Recognition for Human Computer Interaction: A Survey. *Artif. Intell. Rev.* **2015**, *43*, 1–54. [CrossRef]
24. Biswas, K.; Basu, S.K. Gesture Recognition using Microsoft Kinect[®]. In Proceedings of the 5th International Conference on Automation, Robotics and Applications, Wellington, New Zealand, 6–8 December 2011; pp. 100–103.

25. Kawulok, M. Adaptive skin detector enhanced with blob analysis for gesture recognition. In Proceedings of the 2009 International Symposium ELMAR, Zadar, Croatia, 28–30 September 2009; pp. 37–40.
26. Elsayed, R.A.; Sayed, M.S.; Abdalla, M.I. Skin-based adaptive background subtraction for hand gesture segmentation. In Proceedings of the 2015 IEEE International Conference on Electronics, Circuits, and Systems (ICECS), Cairo, Egypt, 6–9 December 2015; pp. 33–36.
27. Chang, C.W.; Chang, C.H. A two-hand multi-point gesture recognition system based on adaptive skin color model. In Proceedings of the 2011 International Conference on Consumer Electronics, Communications and Networks (CECNet), XianNing, China, 16–18 April 2011; pp. 2901–2904. [[CrossRef](#)]
28. Ghaziasgar, M.; Connan, J.; Bagula, A.B. Enhanced adaptive skin detection with contextual tracking feedback. In Proceedings of the 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech), Stellenbosch, South Africa, 30 November–2 December 2016; pp. 1–6. [[CrossRef](#)]
29. Barneva, R.P.; Brimkov, V.E.; Hung, P.; Kanev, K. Motion tracking for gesture analysis in sports. In Proceedings of the 2016 IEEE Western New York Image and Signal Processing Workshop (WNYISPW), Rochester, NY, USA, 18 November 2016; pp. 1–5. [[CrossRef](#)]
30. Jiang, Y.; Hayashi, I.; Hara, M.; Wang, S. Three-dimensional motion analysis for gesture recognition using singular value decomposition. In Proceedings of the 2010 IEEE International Conference on Information and Automation, Harbin, China, 20–23 June 2010; pp. 805–810. [[CrossRef](#)]
31. Jost, C.; Loor, P.D.; Nédélec, L.; Bevacqua, E.; Stanković, I. Real-time gesture recognition based on motion quality analysis. In Proceedings of the 2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN), Turin, Italy, 10–12 June 2015; pp. 47–56.
32. Czuszynski, K.; Ruminski, J.; Wtorek, J. Pose classification in the gesture recognition using the linear optical sensor. In Proceedings of the 2017 10th International Conference on Human System Interactions (HSI), Ulsan, Korea, 17–19 July 2017; pp. 18–24. [[CrossRef](#)]
33. Ng, C.W.; Ranganath, S. Gesture recognition via pose classification. In Proceedings of the 15th International Conference on Pattern Recognition, ICPR-2000, Barcelona, Spain, 3–7 September 2000; Volume 3, pp. 699–704. [[CrossRef](#)]
34. Yamashita, A.; Fujii, M.; Kaneko, T. Color registration of underwater images for underwater sensing with consideration of light attenuation. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007; pp. 4570–4575.
35. Oliveira, M.; Sutherland, A.; Farouk, M. Two-stage PCA with interpolated data for hand shape recognition in sign language. In Proceedings of the 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, USA, 18–20 October 2016; pp. 1–4. [[CrossRef](#)]
36. Birk, H.; Moeslund, T.B.; Madsen, C.B. Real-Time Recognition of Hand Alphabet Gestures Using Principal Component Analysis. In Proceedings of the 10th Scandinavian Conference on Image Analysis, Lappeenranta, Finland, 9–11 June 1997.
37. Saxena, A.; Jain, D.K.; Singhal, A. Sign Language Recognition Using Principal Component Analysis. In Proceedings of the 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, India, 7–9 April 2014; pp. 810–813. [[CrossRef](#)]
38. Masurelle, A.; Essid, S.; Richard, G. Gesture recognition using a NMF-based representation of motion-traces extracted from depth silhouettes. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 1275–1279. [[CrossRef](#)]
39. Edirisinghe, E.M.P.S.; Shaminda, P.W.G.D.; Prabash, I.D.T.; Hettiarachchige, N.S.; Seneviratne, L.; Niroshika, U.A.A. Enhanced feature extraction method for hand gesture recognition using support vector machine. In Proceedings of the 2013 IEEE 8th International Conference on Industrial and Information Systems, Peradeniya, Sri Lanka, 17–20 December 2013; pp. 139–143. [[CrossRef](#)]
40. Chavez, A.G.; Pfingsthorn, M.; Birk, A.; Rendulić, I.; Misković, N. Visual diver detection using multi-descriptor nearest-class-mean random forests in the context of underwater Human Robot Interaction (HRI). In Proceedings of the OCEANS 2015—Genova, Genoa, Italy, 8–21 May 2015; pp. 1–7. [[CrossRef](#)]
41. Saha, H.N.; Tapadar, S.; Ray, S.; Chatterjee, S.K.; Saha, S. A Machine Learning Based Approach for Hand Gesture Recognition using Distinctive Feature Extraction. In Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 91–98. [[CrossRef](#)]

42. Hopcroft, J.E.; Motwani, R.; Ullman, J.D. chapter Context-Free Grammars and Languages. In *Introduction to Automata Theory, Languages, and Computation*, 2nd ed.; Addison-Wesley: Boston, MA, USA, 2001; pp. 169–217.
43. Jurafsky, D.; Martin, J.H. chapter Context-Free Grammars. In *Speech and Language Processing*; Prentice Hall: Upper Saddle River, NJ, USA, 2014; pp. 395–435.
44. Mišković, N.; Bibuli, M.; Birk, A.; Caccia, M.; Egi, M.; Grammer, K.; Marroni, A.; Neasham, J.; Pascoal, A.; Vasilijević, A.; Vukić, Z. CADDY—Cognitive Autonomous Diving Buddy: Two Years of Underwater Human-Robot Interaction. *Mar. Technol. Soc. J.* **2016**, *50*, 54–66. [[CrossRef](#)]
45. Mišković, N.; Pascoal, A.; Bibuli, M.; Caccia, M.; Neasham, J.A.; Birk, A.; Egi, M.; Grammer, K.; Marroni, A.; Vasilijević, A.; et al. CADDY Project, Year 2: The First Validation Trials. In *Proceedings of the 10th IFAC Conference on Control Applications in Marine Systems CAMS*, Trondheim, Norway, 13–16 September 2016; Volume 49, pp. 420–425. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).