



**HAL**  
open science

# Perceiving for Acting With Teleoperated Robots: Ecological Principles to Human-Robot Interaction Design

Bruno Mantel, Philippe Hoppenot, Etienne Colle

► **To cite this version:**

Bruno Mantel, Philippe Hoppenot, Etienne Colle. Perceiving for Acting With Teleoperated Robots: Ecological Principles to Human-Robot Interaction Design. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 2012, 42 (6), pp.1460–1475. 10.1109/TSMCA.2012.2190400 . hal-00744730

**HAL Id: hal-00744730**

**<https://hal.science/hal-00744730v1>**

Submitted on 21 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

---

# Perceiving for acting with teleoperated robots: Ecological principles to human-robot interaction design

Bruno Mantel, Philippe Hoppenot & Etienne Colle  
(IBISC, Université d'Evry, France)

*IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 2012

---



Cette version est celle des auteurs telle qu'acceptée pour publication en 2012 dans le journal *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*. En conséquence, sa pagination est susceptible de différer de celle de la version publiée.

La version publiée de l'article a pour DOI <https://doi.org/10.1109/TSMCA.2012.2190400>

Elle peut être récupérée ici (accès payant) <https://ieeexplore.ieee.org/abstract/document/6185687>

ou sur demande par email auprès de l'un des auteurs (accès gratuit).



This is the authors' accepted manuscript (AAM) version of an article published in 2012 in the journal *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*. As a consequence, the pagination of this AAM version may differ from that of the published version.

The published version of the article has DOI <https://doi.org/10.1109/TSMCA.2012.2190400>

It can be retrieve from (fee apply) <https://ieeexplore.ieee.org/abstract/document/6185687>

or by asking one of the authors by mail (for free).

Pour citer l'article (format APA) :

To cite the article (APA format) :

Mantel, B., Hoppenot, P., & Colle, E. (2012). Perceiving for acting with teleoperated robots: ecological principles to human-robot interaction design. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 42(6), 1460-1475.

# Perceiving for Acting with Teleoperated Robots: Ecological Principles to Human-Robot Interaction Design

Bruno Mantel, Philippe Hoppenot, and Etienne Colle

**Abstract**—By primarily focusing on the perceptual information available to an organism and by adopting a functional perspective, the ecological approach to perception and action provides a unique theoretical basis for addressing the remote perception problem raised by telerobotics. After clarifying some necessary concepts of this approach, we first detail some of the major implications of an ecological perspective to robot teleoperation. Based on these, we then propose a framework for coping with the alteration of the information available to the operator. While our proposal shares much with previous works that applied ecological principles to the design of man-machine interfaces (e.g., ecological interface design or EID), it puts a special emphasis on the control of action (instead of process) which is central to teleoperation but have been seldom addressed in the literature.

**Index Terms**—Affordance, ecological approach, human-robot interaction, interface design, perception-action, teleoperation.

## I. INTRODUCTION

**A** ROBOTIC teleoperation platform is a powerful tool. It enables humans to act on and explore environment from a distance. This permits, for example, an operator to reach and use remote objects and scout remote environments that otherwise may be inaccessible (as in disabled person's assistance or search and rescue) or hostile (e.g., underwater, space, rubble, radioactive area). It also enables to manipulate and examine objects at a different scale than that of the operator. For example, it can provide an operator with increased precision or strength (as in telesurgery or industrial manipulator). However, all these new opportunities come at a cost and allowing the operator to control these actions raises important human factor issues [1], [2]. A particular concern of teleoperation is the alteration of the information available to the operator and its detrimental influence on task performance [3], [4]. In this paper, we consider how the ecological approach to perception and action (e.g., [5]–[7]) might contribute to address these issues.

The ecological approach proposes that an agent's behavior is not triggered by a stimulus or a brain but emerges from the coupling between the agent and his/her environment.

Manuscript received May 3rd, 2011; revised February 8, 2012. This work was supported in part by the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n216487 (<http://www.companionable.net>).

B. Mantel, P. Hoppenot, and E. Colle are with the Computer Science, Integrative Biology and Complex Systems laboratory (IBISC), University of Evry-Val d'Essonne, 40 rue du Pelvoux - CE 1455, 91020 Evry cedex, France (e-mail: [bm@mikmuk.net](mailto:bm@mikmuk.net)).

As a consequence, it is argued that perception, action and cognition cannot be understood by focusing on the sole agent but require to consider the agent-environment (*A-E*) system as a whole. Within this perspective, perception and action are not considered as separate issues, but rather as mutually constraining each others [8], [9]. The agent uses available information about his/her relation to the environment to control his/her actions and by doing so he/she modifies the *A-E* layout and thus the information to which he/she has access (and which he/she uses to control his/her actions, and so on). This circular causality has two major implications for perception. First, to act successfully and efficiently the agent needs functional information, that is, information which is adapted to the intended action and to its context. Second, the information available to the agent depends on his/her own activity.

Several ecological concepts found an echo in telerobotics-related fields such as computer vision, e.g., [10], [11] and autonomous robots and agents, e.g., [12]–[14]. In parallel, building on shared concerns (e.g., *A-E* fit, functional perspective) and historical roots [15], [16], an ecological approach to human factors developed over the last twenty years, e.g., [17]–[22]. In the field of HCI, the ecological perspective made a breakthrough with Norman's popularization of the affordance concept [23] (but see [24], [25]) and with the advent of the so-called ecological interface design (EID) framework which developed in the context of complex work domains such as plants, e.g., [26]. The EID framework (or its first stage, work domain analysis) has seldom been applied to telerobotics. When it was, it has provided thorough analyses of the cognitive aspects (e.g., tactical) of the work domain (e.g., unmanned aerial vehicle for military use [27] or search and rescue [28]). However, we know of few works (in or outside the field of teleoperation) that used EID to address the specific issues pertaining to the control of action –as distinguished from the control of process– which are at the core of most telerobotics applications. One notable exception is the work conducted by van Paassen, Mulder and their colleagues on interface design for flight control in manned aircrafts [29]–[31]. Beside the EID framework, the ecological approach to human factors and ergonomics also led to other notable developments of displays for aircraft pilots, e.g., [32]. In telerobotics, the potential benefits of using the ecological approach to address the issues raised by the remote control of action have been acknowledged [1], but ecological principles

have little been explicitly applied to the design of interfaces for human-robot interaction (HRI)<sup>1</sup>, e.g., [33], [34].

In the present paper, we focus precisely on how the ecological approach to perception and action might be used to address the prospective and online control of action in the context of teleoperation. After clarifying essential (but sometimes misused) ecological concepts (part II) and analyzing how a teleoperation platform alters the agent-environment coupling (part III), we propose a framework for interface design which aims at providing the operator with the relevant functional information (or at compensating for not providing him/her with that information; part IV).

## II. THE ECOLOGICAL APPROACH TO PERCEPTION AND ACTION

Compared to other psychological theories, the ecological approach has the singularity to take as an entry point *what* information is available for perception, rather than *how* this information can be detected by an organism [35, p.7]. On one hand, the ‘what’ question points towards the purpose of perception. It is proposed that when considering the perceptual support for action, we should look first for information which is *functional* (as opposed to general-purpose, such as ‘3D space’). On the other hand, the ‘what’ question also points to the form in which information is available to the agent. Available information is assumed to lie in the structure of ambient energies which stimulates the perceptual systems of the agent.

In the following sections we first address the available information about opportunities for action, both at the level of the *A-E* system which constrains action (Section II-A, Affordances) and at the level of the invariant properties of stimulation which specify these *A-E* regularities (Section II-B, Invariants). We then address the available information for controlling actions (online) while they are being executed (Section II-C, Control laws). We conclude with an attempt at synthesizing the whole framework (Section II-D).

### A. Affordances of the Agent-Environment System

At a first level of description, the ecological approach proposes that what an actor needs to perceive are not the meaningless physical characteristics of the environment (e.g., this object is 43cm away, has diameter of 6 cm, weights two kilos etc.) but rather what he/she can do or not do in a given situation (e.g., is this object reachable, graspable, liftable, ‘throwable’, etc.). Such opportunities for action, which have been termed *affordances*, do not depend on mere properties of objects but on *relations* between properties of the environment and properties of the actor [5, p.129], [36], [37]. For example, a surface which is sufficiently large relative to one’s feet and sufficiently strong relative to his/her weight constitutes

a platform on which he/she can stand. An affordance can thus be defined as a relational property of the agent-environment system [38]<sup>2</sup>. Therefore, an affordance is not a mental construct. It exists whether or not it is perceived. Being grounded in the physical description of the *A-E* system, an affordance is as real as the objects and agent it characterizes.

An affordance provides information about some environmental properties scaled in terms of the perceiver’s own action system. It does not only define whether a given action is (not) possible, but also whether it is efficient, comfortable and so-on. Formally, an affordance can be characterized with a dimensionless ratio (i.e., so-called pi number). For example, Warren [40] noted that whether a stairs is climbable by an individual can be captured by the ratio  $R/L$  of riser height to leg length. If this ratio exceeds a certain value (0.88 in the case of bipedal climbing), called the *critical point* (or *absolute critical boundary* [41]; see below), the stairs cannot be climbed anymore and a new mode of action must be used (e.g., quadrupedal climbing). The  $R/L$  ratio also has an *optimal point* (0.25), at which the energy expenditure associated to climbing the stairs is minimum. When experimental participants were shown stairs with different riser heights and asked to judge whether they could climb the stairs and to rate how comfortable this would be, their responses were found to match the two previously identified values of the  $R/L$  ratio [40].

Warren’s study [40] focused on a single action mode (i.e., bipedal climbing). Mark and colleagues [41] extended this work by examining the perception of action boundaries in the context of two different modes. They investigated the perception of whether an object is within reach, and specifically the transition from reaches performed with the arm only vs. with the arm plus leaning torso forward and/or twisting shoulders (respectively termed one degree of freedom reaching, or 1-dof reaching, and multiple-dof reaching). The authors observed that when the distance to the object is increased, the participants switch from the first mode of action (1-dof) to the second (multiple-dof) before the absolute critical boundary is reached. They called this anticipated transition the *preferred critical boundary*, and showed that it coincided with a transition in felt comfort, suggesting that it could correspond to the region where the first mode of action (1-dof reaching) becomes less comfortable than the second one (multiple-dof reaching). Besides energy expenditure and comfort, preferred and optimal points also depend on other situational/task constraints such as accuracy [42] or quickness and safety [43]. The gap between preferred and absolute critical boundaries can thus be considered as a margin of safety retained by the actor, e.g., [41], [43]. Fig.1 summarizes all these notions with a schematic diagram.

Of course, the ability of an agent to perform an action does not depend on the sole geometrical characteristics of his/her body relative to the geometrical properties of the environment. Trivially, an agent wearing a heavy backpack may not be

<sup>1</sup>HRI covers a wide range of interaction schemes, that may differ according to the degree of autonomy of the machine (e.g., continuous human control, supervisory control, cooperation), according to whether the machine’s interface is the robot itself (permitting direct physical interaction) or a distinct console (as in teleoperation), according to whether the operator and the robot are co-located or distant, etc. Our present concern is with the teleoperation of robot of varying autonomy, primarily distant.

<sup>2</sup>While the principle of complementary *A-E* dispositions is largely agreed, there is an ongoing debate on the exact ontological status of affordances. The interested reader may refer for example to [39] and two special issues of *Ecological Psychology*, #12(1) and #15(2).

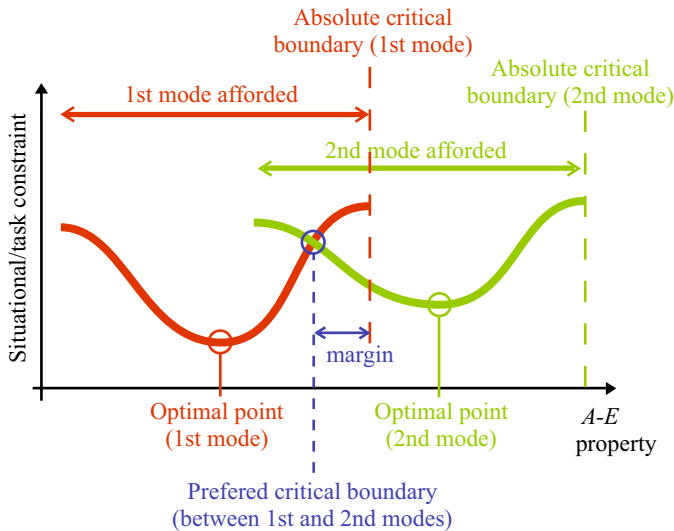


Fig. 1. Action boundaries and optimal points characterizing affordances. Given a particular action context (e.g., stairs climbing, reaching), a variable capturing a relevant situational/task constraints (e.g., energy expenditure, discomfort, danger) is plotted for two different action modes (e.g., bipedal/quadrupedal climbing, 1-dof/multiple-dof reaching) as a function of a relational property of the *A-E* system (e.g., riser height/leg length, object distance/arm length). (see text for details; the figure is abstracted from the affordance literature, in particular [40], [41]).

able to climb stairs of the same riser height as when not wearing the backpack (or not with the same efficiency, etc.). An affordance can be characterized in different ways whose relevance depends on context. Formally, this can be achieved by using different types of pi-numbers which capture the influence of different types of constraints on action [44]: geometrical (lengths combinations), kinematic (lengths and time combinations), kinetic (length, time and force combinations).

### B. Invariants in Ambient Energy Arrays

As Gibson underlined “*the central question for the theory of affordances is not whether they exist and are real, but whether information is available in ambient light for perceiving them*” [5, p.140]. Thus, at a second level of description, the ecological approach proposes that the structure of ambient energies which stimulate an agent’s perceptual systems specify the regularities of the *A-E* system. These structures which extend within and across optics, acoustics, haptics, etc. are called *invariants*. For example, light propagation is constrained by the constitutive substance of objects and the layout of their surfaces. As a consequence, the structure of light converging at a point of observation, or *optic array*, contains information about the nature of objects, their shape and their spatial layout relative to the perceiver. For example, when an object’s texture is approximately regularly distributed, the relative density of texture elements in the optic array (so-called texture gradient) specifies the slant of the surface relative to the point of observation [45, p.77]. Similarly, the specific manner in which sound waves propagating from a source are altered by the listener’s head and external ear contains information about the direction of the source relative to the head [46, p.97].

1) *Persistence over Change*: Among the most important informational structures in ambient energies, many result from changes in the agent-environment system. First, change is ubiquitous in the *A-E* system, principally in the form of movements, both on the side of the perceiver and on the side of external objects. In particular, an agent is never completely at rest (i.e., stationary). Even the most elementary aspects of its behavior, such as breathing and controlling posture, involve movements which in turn generate information. Second, change is an extremely rich source of information because it reveals the underlying structure: “*What is invariant does not emerge unequivocally except with a flux. The essentials become evident in the context of changing nonessential*” [5, p.73].

The dynamic structure of light variations that converge onto a point of observation, or *optic flow* [8], [47], [48], offers many obvious examples of these movement-induced patterns. For example, the global flow rate is directly related to the velocity of the point of observation relative to the illuminated environment [16], [49, p.805], while the relative rate of optic flow elements (i.e., so-called motion parallax) specifies the (dis-)continuity and slant of surfaces relative to the point of observation [50]. If the point of observation translates, the flow is radial in the direction of motion and laminar in the orthogonal direction. The center of the radial pattern, called the *focus of expansion*, corresponds to the direction of displacement [8]. Thus, the flow structure specifies the direction toward which one is moving. Among other well documented structures, the relative expansion rate of an object’s projection in optic flow, so-called  $\tau$  variable, specifies the first order time-to-contact, that is, the time remaining before contact if the relative speed of the point of observation and object were to be maintained [51], [52]. Of courses, similar informative structures exist in all ambient energies stimulating the senses. For example, when a source is emitting a constant sound, the relative rate of change of the sound pressure or intensity at the ears (so-called acoustic  $\tau$ ) conveys information about the source’s time-to-contact and/or distance [53], [54]. In haptic stimulation, the inertial moments of an object held or wielded conveys information about how much the tool extends beyond where it is grasped [55].

The structures of ambient energies do not only convey spatial information. For example, by simply seeing the motion of a few patches attached to the articulations of a human body (so-called point-light display), humans can perceive as diverse information as the action being performed, the morphology, age or gender of the actor, or even his/her intention [56], [57].

2) *Specification of Affordances*: One important claim of the ecological approach is that affordances are specified, that is, that the structures of ambient energies map one to one to the relational properties of the *A-E* system which define what the agent can do and not do. Given that the agent’s own body and actions also participates to the structuring of energies, it is argued that there exist higher order invariants in ambient energies arrays which provide body-scaled and action-scaled information.

For example, the height of the point of observation relative to the ground is an important constraint on optic flow

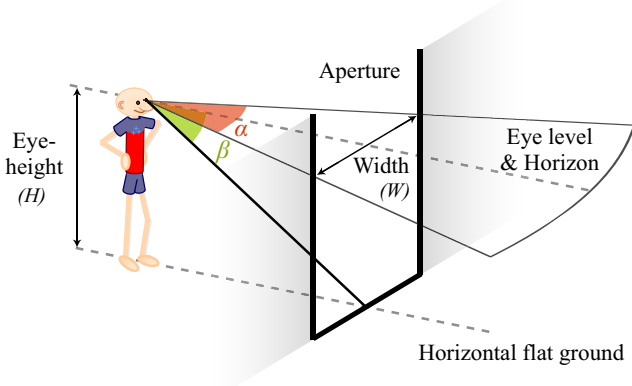


Fig. 2. Eye-height based optical specification of ‘passability’ (i.e., whether an aperture can be passed through). For an agent facing the aperture, the optic array specifies the width of the aperture in units of his/her own eye-height and thus, up to a scaling constant, in units of his/her own shoulder width (adapted from [43]).

and thus could play a major role for providing body-scaled information [5, p.164], [47], [58]. As an illustration, consider the information available about passing-ability. Fig.2 depicts an agent standing on a flat horizontal ground and facing the middle of an aperture. In that case, the relation between  $\beta$ , the angular declination below the horizon of the doorstep, and  $\alpha$ , the angular horizontal width of the object, specifies the object’s width  $W$  in units of eye-height  $H$  [43]:

$$\frac{W}{H} = \frac{\tan(\beta)}{2\tan(\alpha/2)} \quad (1)$$

If we assume that standing eye-height bears a stable relation to other anthropometrical characteristics such as shoulder width  $S$  (e.g.,  $H = c.S$ ), it follows that the optic array specifies, up to a scaling constant, the width of an aperture in units of shoulder width. In a similar fashion, the optic array under the constraint of eye-height specifies the distance of an object lying on the ground or the height of a riser in units of eye-height. Empirical studies have shown that participants exploit the constraint of eye-height on optic flow to judge whether they can pass through an aperture [43], pass under a barrier [59] or seat on a stool [60].

It is worth noting that informative structures are not only available *within* individual energies, but also extend *across* energies, either of the same type (e.g., binocular or binaural invariants) or of different types (intermodal invariants) [61]. For example, the interaural time difference and level difference contain information about the direction of a sound source relative to the listener [46, p.137]. Given that the relational properties of the *A-E* system which constrain actions often affect simultaneously multiple ambient energies, intermodal invariants could play a particularly important role for the specification of affordances. For example, humans accurately judge whether they can reach a visible object with a rod, even when they can wield the rod but not see it, suggesting that they are sensitive to information about ‘reachability’ available across optic and haptic stimulation [55], [62].

### C. Laws of Control of Movement

The specification of affordances provides an informational basis for perceiving what actions can be performed, at which cost for the organism and so on. Beside this first prospective dimension of control [39], the agent also needs information for regulating his/her actions while they are being executed. The proponents of the ecological approach argue that guiding actions with respect to environmental goals does not necessarily require to construct a 3D (internal) representation of the world and to plan movements in details. Instead, it is proposed that actions can be regulated ‘online’ by exploiting (task-) specific mappings between control variables<sup>3</sup> and informational variables, termed *control laws* [8], [37], [64].

In a seminal article [8], Gibson described a complete set of such laws that could be used to control locomotion (starting, stopping and backing-up, pursuing and fleeing, etc.). He proposed for example that steering toward a visible goal can be achieved by continuously adjusting one’s direction of motion so as to maintain the optical focus of expansion aligned with the goal’s projection in the optic array – while avoidance would consist in keeping the focus of expansion outside of the obstacle’s projection. Similarly, to steer toward a sound source, an agent can move so as to simultaneously minimize the difference in sound intensities at his/her ears and increase the overall sound intensity [35, p.83]. Among popular control laws, the so-called constant bearing and constant absolute direction strategies have been used for years by ship navigators to move out of a collision course, but also by bats to ensure a successful catch over their prey [65]. Perhaps the most striking examples of control laws which obviates the need for mental maps and action plans are the so-called balance strategy and avoid-closest strategy [66]. They show that without recovering any depth information, a small flying robot is capable of avoiding obstacles by simply (i) equating the average optic flow rate on the left and right halves of the camera image or (ii) turning away from the direction where the optically specified time-to-contact is the lowest.

Fajen and Warren [67] (see also [68], [64]) formalized a control law for steering toward goals and avoiding obstacles in which turning rate ( $\dot{\phi}$ ) is regulated as a function of the object’s distance  $d$  and of the object-relative heading specified in optic flow ( $\phi_O - \psi$ ) and across haptic/gravito-inertial and optic flows ( $\phi_{HGO} - \psi$ ). The version presented below corresponds to a situation with one goal (g) and one obstacle (o). In principle, it could be extended to more obstacles by duplicating the obstacle term (i.e., the second line):

$$\begin{aligned} \dot{\phi} = & -k_g((\phi_{HGO} - \psi_g) + wv(\phi_O - \psi_g))(e^{-c_1 d_g} + c_2) \\ & + k_o((\phi_{HGO} - \psi_o) + wv(\phi_O - \psi_o))(e^{-c_3 |\phi - \psi_o|})(e^{-c_4 d_o}) \end{aligned} \quad (2)$$

where  $k_g$  and  $k_o$  represent respectively the attractiveness of the goal and the repulsiveness of the obstacle,  $v$  is the agent’s speed and  $w$  accounts both for the amount of optical structure

<sup>3</sup>In accordance with the dynamic system theory, it is assumed that the action system self-organizes under task constraints such that its various neuro-musculo-skeletal degrees of freedom reduce to a few free parameters, or control variables, that need to be regulated [9], [63].



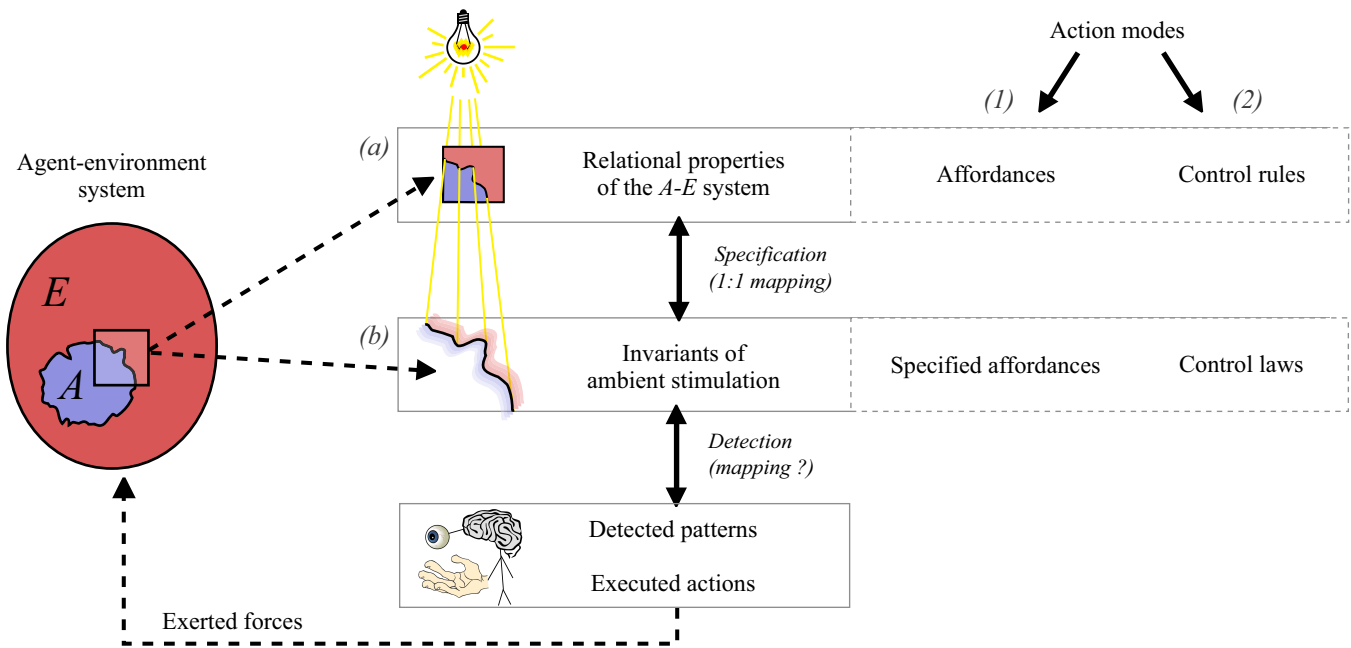


Fig. 3. Perception for action and action for perception in the agent-environment system (see text for details).

and for the size of the field of view, and  $c_i$ 's constants define the exponential rate of decay and minimal value of the distance's influence on steering.

A great number of studies also investigated how the optical variable  $\tau$  could contribute to the control of actions such as braking or intercepting a mobile object. For example, Lee (1976) showed that an agent can achieve a smooth contact with an object by regulating its braking in such a way that the first temporal derivative of  $\tau$  is kept at a certain value (-0.5). Later, Yilmaz and Warren (1995) observed that instead of maintaining a steady state, experimental participants appeared to be adjusting the brake position  $x$  as a function of the deviation of  $\dot{\tau}$  from a reference value (-0.52):

$$\Delta x = b(-0.52 - \dot{\tau}) + \epsilon \quad (3)$$

where  $b$  is a scaling constant and  $\epsilon$  a noise term.

**Control Rules:** We propose to complement this formalism by adding a new level of description, called *control rules*. In control rules, the control variables are expressed as a function of the relational properties of the A-E system which constrain action, in place of the invariants of ambient stimulation which specify these constraints (as it is the case for control laws).

Within this formalism,  $\dot{\phi}$  in (2) would be expressed as a function of distance and object-relative heading and  $\Delta x$  in (3) as a function of the rate of change of first order time-to-contact, regardless of how these properties are specified in ambient stimulation (e.g., in optics, acoustics or across different ambient energies).

The reason for adding this new level of description is twofold. First, it provides a unified (two levels) framework for characterizing the available information about action opportunities (i.e., affordances and specifying invariants) and the information for regulating actions while they are executed (i.e., control rules and control laws). Second, it could prove

particularly useful in the context of HCI and of telerobotics in particular, because these induce a decoupling between the properties constraining actions and the ambient energy structures specifying these constraints (a point on which we will come back later).

#### D. Synthesis

Fig. 3 summarizes the information available to an agent for controlling his/her action. On the right side of the figure, the two aspects of control are represented in two different columns. Specifically, it is proposed (see also [37]) that for each action mode there exists information about the opportunity to perform the action, its efficiency, etc. (column 1) and information for regulating the free parameters of the action system while the action is executed (column 2). For example, the action mode 'walking through' requires information about whether the aperture can be passed through (1) and information for controlling locomotion on-line relative to the (edges of the) opening (2). Then, it is proposed that for both aspects of control, the characterization of available information involves two distinct levels of description, represented with the two horizontal boxes in the upper part of the figure: first, the level of the relational properties of the A-E system at which action is constrained (box *a*); second, the level of the invariant properties of ambient stimulation which specify the control-relevant A-E properties of the first level (box *b*).

One central claim of the ecological approach is that the relation between the two levels is one of *specification*. Specification refers to a lawful 1:1 mapping between invariants of ambient stimulation and the properties of the A-E system that give rise to them. Accordingly, information available to an agent is assumed to be unambiguous. However, this does not mean that all affordances and control rules are specified all

the time. For example, some objects might occlude others, a room might be dark and a place filled with fog. More generally, some properties of an object won't affect the optic, acoustic or haptic flows stimulating the agent's perceptual systems unless he/she rotates the object, knocks it, wields it or scrapes its surface. It follows that perception shall not be conceived as a linear process (going from the top to the bottom box of Fig. 3) but as a circular process in which the agent adjusts his/her actions until the invariants of ambient stimulation specify the information sought: on Fig. 3, perception is spread over the whole loop which goes from box (b) down to the bottom, and then follows the dotted arrows leftward to the A-E system and then rightward, back to box (b). The two rightward dotted arrows emphasize the dual role played by the agent's action: performatory and exploratory [7, p.80]. Altogether, the notions of affordance, invariant, control rule and control law emphasize the mutual dependency of perception and action, or so-called perception-action cycle [9].

### III. TELEOPERATION AS TOOL-MEDIATED INTERACTION

Teleoperation falls in the category of instrumented interactions. A teleoperation platform is a tool. Tools vary greatly in complexity, ranging from projectiles, hammer or magnifying glass to cars, sonars or computers, but they all share a same function. As noted by Shaw and colleagues [44, p.306]: “Tools enhance, extend, or restore the action or perception capabilities of humans and animals”.

By mediating the relation of an agent to his/her environment, a tool alters simultaneously the A-E properties constraining actions (Fig. 3, box a) and the invariants in ambient stimulation specifying these properties (Fig. 3, box b). In Sections III-A and III-B below we first consider the implications of each of these levels for teleoperation. Then we emphasize the two different perspectives they represent for design (Section III-C).

#### A. Alteration of Affordances and Control Rules

A tool changes the layout of affordances. It provides new opportunities for action and withdraws other opportunities that were previously afforded. An agent holding a hammer gains new opportunities for hitting things, for reaching farther objects, etc. At the same time, he/she may not be able to grasp a glass of water anymore or to explore a surface with his/her hand in the search of a light switch. Similarly, by teleoperating a robotic arm, a technician gains the opportunity to assemble pipes underwater or a disabled person gains the opportunity to reach and grasp distant objects. However, while doing so, these operators may not be able to simultaneously sign a register, drive a car or a wheelchair, read a book, etc.

When considering the affordances of an actor-tool-environment system (A-T-E system), a first distinction can be made as a function of whether the tool is manipulated or not. The opportunities for action granted by a tool considered as a detached object of the environment, A-(T-E) are not the same as the opportunities for action granted by the same tool considered as a functional extension of the actor's action system, (A-T)-E [44]. For example, a hammer lying on a table

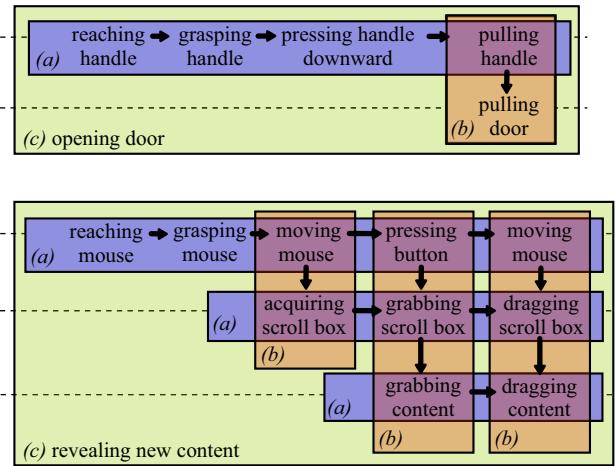


Fig. 4. Different types of action modes nesting illustrated in the context of opening a door (top) and revealing window content on a GUI (bottom). Nested action modes can be either sequential (a), concomitant (b), or they can embrace each other and form a hierarchy (c) (see text for details).

can afford picking up, while when firmly held in the hand it can afford hammering. This dual function of a tool can be reformulated within the more general framework of action modes *nesting*, e.g., [22], [37]. We propose to distinguish three different kinds of nesting: sequence, concomitance and hierarchy. They are illustrated on Fig. 4 with the now classic examples of a door with a pivoting handle and of a scrollbar on a GUI [23], [69]. *Sequence* refers to the nesting of action modes in time (Fig. 4, boxes a) [69]. For example, a horizontal door handle first affords grasping to the actor, then once grasped it affords pressing downward, and once pressed it affords pulling. Similarly, in a scroll-bar, the moveable box first affords grabbing to the user, and when grasped it affords dragging<sup>4</sup>. *Concomitance* designates the relation between action modes which are simultaneous but nevertheless depend on each others (Fig. 4, boxes b). For example, if a door is not locked, a door handle affords both pulling the handle and pulling the door. Similarly, if appropriately programmed, a grasped scroll-bar box affords both dragging the box and dragging the content of the linked window. Lastly, *hierarchy* refers to the relation by which an action mode embraces other modes (Fig. 4, boxes c) [22]. For example, the action mode ‘picking up a hammer’ can be decomposed into the lower order modes of reaching, grasping and lifting the hammer. Similarly, grasping a door handle, pressing it downward, and pulling it are components of the higher order action mode of opening a door, while grabbing and dragging a scroll-bar are part of the higher order action mode of revealing new content [69].

Importantly, a door handle can be pressed only if it can be reached and grasped first, and at the same time, the handle won't be grasped in the same way if for pressing downward (to

<sup>4</sup>Remind that we are discussing affordances (i.e., what the agent can do or not do) and not the availability of information about these affordances. Hence, the scroll box refers here to a functional area (i.e., an area programmed such that it can be clicked, grabbed and dragged) and not to something displayed on a screen that would make this area visible. The opportunity to drag a box does not depend on whether there is information about such opportunity. The box within a scroll-bar can be dragged even when the computer screen is off.



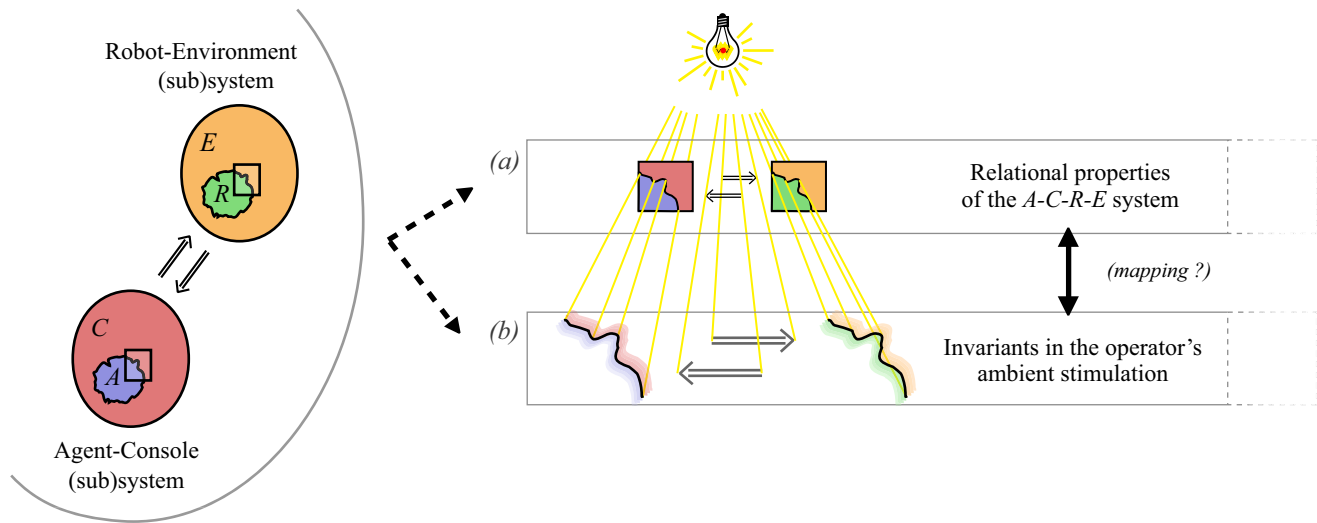


Fig. 5. Redefining the two levels for describing the information available to the agent in the context of robot teleoperation (see text for details).

open the door) or for blocking it upward (to prevent someone from opening the door). These two examples underline that sequential action modes do not merely follow each others in time: within a sequence each mode is contingent on both its predecessor and its successor. The transition from one mode to the other does not occur at a predetermined moment in time but (if and) when the second mode becomes afforded, or more efficient than the first, etc. (see Section II-A; [37]). Similar causal dependencies exist between concomitant modes and between modes within a hierarchy and they also have important consequences on both affordances and control rules. For example, when revealing content on a GUI, the action mode ‘moving the mouse’ is constrained by the concomitant mode ‘dragging content’ in such a way that hand movement can be regulated online on the basis of the induced scrolling of the content on the screen.

In telerobotics, it is the robot which ultimately acts in and on the remote environment (e.g., manipulating objects, revealing their characteristics). Thus, it may seem at first that what the operator needs to know are the affordances of the robot-environment system ( $R-E$  system). Can the mobile robot pass through that opening? Is that object within reach of the robot’s arm? Such  $R-E$  functional fits can definitely not be ignored. However, they are not sufficient for teleoperation<sup>5</sup>. If the operator is to control the robot’s actions, he/she will not only need to know what the robot can do but what he/she can *make* the robot do [70]. Can I drive the mobile robot through that opening? Can I reach that object with the robotic arm?

The nesting of action modes is further increased in telerobotics because the tool is distributed among a command console and a robot. Fig. 5 takes up the top left part of Fig. 3 and applies it to the case of teleoperation. As illustrated in box (a), the operator’s opportunities to act on the remote environment (through the platform) are constrained by the

properties of the agent-console-robot-environment system ( $A-C-R-E$  system). An affordance (or a control rule) of the  $A-C-R-E$  system depends simultaneously on relational properties of the robot-environment subsystem ( $R-E$ ), on relational properties of the agent-console subsystem ( $A-C$ )<sup>6</sup>, and on the existence of an adequate mapping between the actions granted by these two sets of constraints. For example, the opportunity to drive a robot uphill is constrained by: 1) the relation between the power of the robot and the slant of the path; 2) the size and stiffness of the console’s joystick relative to the operator’s size and strength; and 3) the characteristics of the mapping (if any) between the operator’s opportunities for manipulating the joystick and the robot’s opportunities for climbing the hill. Each type of constraints does not only affect which actions can or cannot be performed, but also how efficient and comfortable these actions are (for the operator; cf. II-A).

An important concern with such a complex tool is that the mapping between the actions performed by the operator and their consequences in the remote environment can change across teleoperation sessions or even within a session. This will be the case for example when an obstacle avoidance system (or any other assistance) is turned on, when the console’s inputs (keyboard, joystick) are changed, when transmission delay changes, when a tire is under-inflated, etc. In some cases (e.g., when the batteries are low) these alterations can be informative, to the extent that they are specified (see III-B below). However, in all cases these alterations also change the control rules and affordances of the  $A-C-R-E$  system, forcing the operator to rediscover them.

### B. Alteration of Specifying Invariants in Ambient Stimulation

In addition to changing the opportunities for action, a tool also alters the structure of ambient energies which stimulates the perceptual systems of the agent, and thus the information

<sup>5</sup>This could be enough if the operator were to team up with an autonomous robot in order to achieve a collaborative task, or if the operator were to delegate some actions to the robot, a point on which we will come back later.

<sup>6</sup>By ‘console’ we refer here to both the console itself and the operator’s surroundings.

available to the agent. On one hand, this alteration is partly functional as it allows the new (in-)opportunities for action and the new control rules to be specified. On the other hand, this alteration also restrains available perceptual information.

In the case of teleoperated robotics, the alteration of available perceptual information is all the more important that – again – the tool is split up into a command console and a robot. As shown on Fig. 5 (box *b*), the availability of information specifying what the operator can (make the robot) do depends on the availability of patterns, within ambient stimulation, which specify the relevant properties of the *A-C-R-E* system. Within the *A-C* subsystem, the physical characteristics of both the console and the operator participate to the structuring of the ambient energies which stimulate the operator’s perceptual systems. It follows that the information about what the operator can do with the console’s inputs (reaching, grabbing and moving a mouse or a joystick, reaching a screen, etc.) can, in principle, be directly available in the structure of ambient stimulation. As concerns the *R-E* subsystem, the physical characteristics of the robot and of its environment also participate to the structuring of ambient energies. However, as the operator and the robot are often not physically located at the same place, these structures may not reach the operator’s perceptual system (e.g., if the robot is not in direct line of sight, not within hearing distance, etc.). In that case, availability of information about what the robot can do depends on the ability of the robot (or any other entity) to gather the relevant information (e.g., by sampling the specifying energy structures), on the capability of the communication system to transmit this information, and on the capability of the console to render it appropriately to the operator. Similarly, whether and how the mapping between *A-C*’s and *R-E*’s functions will influence the structure of ambient energies stimulating the operator is also contingent on the robot’s sensors (camera, laser, sonar, etc.), the console’s displays (screen, speaker, force-feedback joystick, etc.), and the communication between the two.

*Exploratory activity, multimodality, and specification:* Besides the aforementioned aspects, there are at least three transverse aspects that are worth mentioning. First, the information available to the operator also depends on the opportunities for action granted by the teleoperation platform. As we argued earlier, movement is central to perception (see Sections II-B, II-D) and thus, any limitation on the operator’s ability to control the robot’s body and sensors is also a restriction on his/her ability to create information [62], [71]–[73]. For example, a displayed video is all the more flat and lacking information about spatial layout that it is hermetic to the operator scrutinizing. Typically, moving the head will not reveal the objects that are occluded; looking successively at objects lying at different depth will not change the distance of convergence of lines of sight; nor will it sharpen the contour of the focused object (or blur the contour of other objects). Second, the teleoperation platform does not only affect the structure of individual ambient energies (optics, haptics, acoustics, etc.), but also alters the information available across energies. Even if an operator is provided with coordinated patterns in optic and in haptic/gravito-inertial flows (e.g., by wearing a HMD

and driving the robot camera with his/her head), the availability of information about the scale of the scene will depend on the gain between the two patterns of stimulation [62]. Third, the teleoperation platform also questions the principle of specification, that is, that there exists a lawful 1:1 mapping between *A-C-R-E* properties and the structure of stimulation. In complex tools such as teleoperation platforms, the mapping is mediated by complex mechanical and electronic devices<sup>7</sup>. As a consequence, the mapping between *A-C-R-E* properties and the resulting structure in ambient energies can change from one platform to the other or even from one session to the next (e.g., if the camera orientation has changed, if one sensor is off or not calibrated, etc.). This is a major concern because in absence of specification (i.e., if the mapping is 1:many, many:1 or many:many), the structure of ambient stimulation is ambiguous relative to the opportunities for action and to how actions can be regulated.

### C. Different Entry Points for Designers

The designer of a teleoperation platform can step in at different levels that shall not be confused. At a first level, the level of affordances and control rules, he/she can work on laying out an appropriate functional fit within the agent-console-robot-environment system. As noted by McGrenere and Ho [25], a designer working on this first level addresses both usefulness and usability. He/she first addresses usefulness because the relevance of having the system afford a particular action depends on the appropriateness of the action relative to the goals of the platform. He/she also addresses usability, because the relevance of affording an action also depends on the cost (e.g., energy expenditure) of performing this action for the operator.

At a second level, the level of invariants of ambient stimulation, the designer can work on the information available to the operator for controlling the robot. In other words, he/she can work on the means by which the structure of ambient stimulation can be made to specify the control-relevant properties of the *A-C-R-E* system of the first level. When working on the specification of affordances and control rules, a designer thus addresses another aspect of usability, the availability and salience of the information which is necessary for controlling the robot [25]. He/she further addresses usability through the cost associated to the exploratory activity which generates the information.

The two levels are of course not independent of one another. However, as we noted at the end of Section III-B, in the case of telerobotics there is no guarantee that the relation between control-relevant *A-C-R-E* property and the invariants in ambient stimulation will be one of specification (i.e., 1:1 mapping). The existence and consistency of this mapping, on which depend its tractability and the operator’s ability to exploit it, are in large part up to the designer (thence the importance of standards [74]).

<sup>7</sup>Interestingly, several GUI design choices can be viewed as attempts to restore artificially a seemingly unmediated mapping between ambient energies structure and afforded actions. For example, in a scroll-bar the empty areas extending above and below the cursor but not on its left and right indicate that the box can be dragged upward and downward, but not side to side [69].

#### IV. DESIGNING INTERFACES FOR TELEOPERATED ROBOTICS

##### A. Data vs. Information

An important concern in research on HRI for robot teleoperation is the severe alteration of the information available to the operator and its detrimental consequences on task performance and felt comfort [1]–[3]. As a consequence, HRI designers are faced with the question of determining what type of sensors and displays are needed, and how the data should be presented.

Traditional interfaces for robot teleoperation have classically used one different display to represent the data provided by each sensor (i.e., termed conventional 2D interface in [33]). As researchers worked on improving teleoperation performance, this so-called philosophy of one-measurement-one-display (or single-sensor-single-indicator) led to a proliferation of displays and alarms (e.g., video, laser, sonar, robot pose, energy gauge, compass, etc. [33]; the argument was initially made in the context of process control, see [20], [22]). Recently, however, there has been an important trend in HRI design community toward combining data from several sensors within a single display (e.g., overlaying heat and sound data on video [75], fusing laser, sonar and video data [76], integrating video with a 3D reconstruction of laser data and/or known map [33], [77]–[80], integrating images from different cameras [34]). Still, in most cases, the content of the interface did not change (i.e., the same set of data was displayed), only the form in which data were presented was modified (typically, by representing them within a common reference frame –or at least within close frames<sup>8</sup>).

1) *A Functional Perspective*: We propose to shift the primary focus from *how* data should be presented to *what* data should be presented [73] (see also [33]), and hence to move from designing displays of data to designing displays of information [20], [22]. Given our ecological perspective, information has here to be understood as ‘behaviorally-relevant’ (i.e., actor-, task-, and environment-specific). We argue that the primary objective of console displays is to provide the operator with the information that is relevant to achieve the actions for which the teleoperation platform is intended. We further argue that this objective is the cornerstone of the whole interface design process, the point on which all other decisions depend. As Flach and colleagues [81, p.297] noted: *The limitations of technologies, humans, and control systems are important considerations – but the significance of these limitations can only be appreciated relative to the functional demands of a work domain*. To underline this central aspect and to contrast it with other approaches to design whose primary focus is on structure rather than on function (e.g., technology-centered, user-centered and control-centered), these authors labeled the ecological perspective *use-centered* [81].

2) *Behavioral Fidelity, not Presence*: In mediated perceptual experiences, the concept of presence (or telepresence) refers to the impression or subjective feeling of being in the remote environment, to an illusion of reality and of non-mediation [82]–[84]. An important corollary of the functional

perspective promoted here is that presence does not appear to be a relevant criterion for designing teleoperation interfaces (see also [1], [85], [86]). Building on the literature on virtual reality, we argue that there are at least three reasons for that. First, there exists no evidence that the subjective feeling of presence is a prerequisite for behavioral performance [86]. On the contrary, it has been showed for example that using wire-frame vs. photo-realistic computer graphics (CG) had no significant influence on humans’ ability to judge distance in virtual environments [87]. Similarly, humans can accurately judge whether a virtual object is within reach or not, even though they are perfectly aware that this object is immaterial and thus cannot, properly speaking, be ‘reached’ [62]. Therefore, if one is interested in behavioral performance (task success, efficiency, comfort, etc.) achieving a subjective feeling of presence does not appear to be necessary. Second, it is very unlikely that operators will ever be fooled and take the world interfaced as being their real (un-interfaced) surroundings [86]. The reason is that any alteration of the structure of ambient stimulation caused by the teleoperation platform does not only potentially alter the information about the world interfaced: it also specifies the interface itself (i.e., the teleoperation platform). Thus, eliciting a sensation of presence is not a realistic goal to pursue, at least within the next decades. Third, any research on interface design that restricts itself to trying to improve stimulation correspondence (instead of information correspondence) will, in all likelihood, also restrict itself to providing the operator with the same means for controlling actions that he/she could have used in unaided situations [88]. Thus, achieving a subjective feeling of presence restrains the repertoire of solutions and thereby potentially underutilizes technology.

It follows that, from an ecological point of view, the challenge is not to design an interface that would be invisible, but rather to create an interface that would be *functionally transparent*, that is, transparent to the informative patterns that are relevant to achieve the task at hand [22], [86].

##### B. Proposed Framework

Determining the information the operator needs to control the robot is a complex task, but it is only then that we can devise how this information can be sampled and rendered to the operator –or any other means to allow the operator aided with the teleoperation platform to achieve the intended actions [20], [22], [88]. We propose that the overall process can be divided into three steps: a first step of identifying all the action modes that the *A-C-R-E* system should afford; a second step of characterizing these modes with the control relevant properties of the *A-C-R-E* system; and a third step of implementing these action modes, which will principally consist in modifying the platform such that the ambient energies stimulating the operator specify these *A-C-R-E* properties. In the terminology of Vicente and Rasmussen (1990), the first two steps aim at identifying the *content* and *structure* of information while the third step is concerned with the *form* in which this information is given to the operator. In essence, steps 1 and 2 are similar to the work domain analysis used to identify constraints on

<sup>8</sup>In [33], [75], [79], [80], the video and laser/map data on the interface still belong to two different frames and thus their contents remain incoherent.

process control in cognitive systems engineering (for example as part of the EID framework, e.g., [22], [26], [27]). However, in the present case our concern is with the control of action, not process. This implies notably to take into account the constraints arising from the operator (Section III-A). Below we go back to each of these steps.

1) *Identifying Action Modes*: The first step deals with identifying all the affordances of the *A-C-R-E* system that need to be realized (i.e., the actions that need to be performed) to fulfill the goals for which the teleoperation platform is intended. This first step thus addresses the following questions: what is the purpose of the teleoperation platform? What is it designed for? What should it allow the operator to (make the robot) do? Given the nested structure of action modes and its consequences on control (see Section III-A), the purpose of this first step is not just to build a list of all the action modes that the system should afford, but also to identify and keep track of the nesting relations between actions [22]. An abstraction hierarchy (AH) could be used to model the hierarchical relations between actions [22], [89]. However, the two other types of nesting (sequence, concomitance) also need to be modeled (III-A). In addition, such a use of the AH, although it would comply with the principles of the AH (e.g., [22]) would differ from the way it is traditionally used in at least two respects. First, there is no reason why the number of levels in a hierarchy of action modes should be known in advance: in all likelihood, this number will depend on the purpose of the platform being studied (see also [37]). Second, this AH would only capture the relations between action modes. As an alternative, an AH could also be used to model how (individual) actions relate to the underlying physical laws and properties of the system which constrain them. This second perspective, which seems to be the way the AH is primarily used in the literature (e.g., in EID for flight control, [29]–[31]), addresses the second step of our framework: the characterization of action modes (Section IV-B2 below). The identification of action modes can be handled as a recursive process in which higher order modes (e.g., searching victims in rubble) are decomposed into lower order modes (e.g., locomoting, scanning), which are themselves decomposed into lower modes (e.g., steering toward a goal, avoiding obstacles, passing a doorway, scrutinizing, un-occluding, etc.), and so on.

2) *Characterizing Action Modes*: Using the action modes and the nesting structure previously identified (IV-B1 above), the second step then consists in characterizing the associated affordances and control rules. In other words, this step is concerned with identifying the relational properties of the *A-C-R-E* system which define (i) whether the actions can be performed or not (and at which cost, etc.) and (ii) how the corresponding action variables can be regulated when actions are executed (see Section II).

Being relational properties of the *A-C-R-E* system, the affordances are not only action-specific but also contingent on the characteristics of the actor (operator, robot) and of the environment within which the action takes place (command console, robot environment). Affordance characterization thus requires to take into account constraints such as those arising from the

special populations that will operate the system (operators with particular disabilities, children, elderly, engineers, medical staff, etc.) or from the particular ecological niche in which the robot is to operate (home, underwater, forest, rubbles, etc.). Typically, an affordance will be characterized with one or several dimensionless ratios (see Section II-A). The relevance of using a particular type of ratio (e.g., anthropometrical, biomechanical, bio-energetical) and of using a given ratio in place of, or in conjunction with others depends on context. For example, whether a given path affords locomotion or not depends principally on its width (relative to the robot's) in a flat home environment, while if in rubbles or outdoor it may further depend on whether the floor is sufficiently strong (relative to the robot's weight), on the slant and material of the surface (relative to the robot's power, grip and mass distribution) and so on. Importantly, finding the description that best captures a functional fit does not necessarily imply to reach completeness or the highest precision, but rather to fulfill the precision requirements of the action considered.

A similar analysis also applies to control rules. Like affordances, control rules are defined across properties of the *A-C-R-E* system and thus the degrees of freedom that have to be actively regulated when an action is being performed also depend on the interplay of task, actor and environmental constraints (see Sections II-C, III-A). For example, steering toward a goal requires to adjust the direction of self-motion (i.e., heading) so as to minimize the difference between heading and the direction of the goal. However, if the goal is the entrance of a tunnel further ahead on a road, steering can also be defined as adjusting heading so as to stay in the lane (or, for certain combinations of vehicle and roads, as adjusting heading and speed, etc.).

3) *Implementing Action Modes*: Once the functional information the operator needs to control the robot has been identified, the final step consists in adapting the teleoperation platform in such a way that the operator aided by the console-robot interface can successfully perform all the actions identified during step 1 (IV-B1 above). We propose that the means for adapting the interface can be divided in four categories: anthropomorphism, information augmentation, action augmentation, and delegation. The first three approaches aim at granting the operator access to the functional information identified during step 2 (IV-B2 above). The fourth approach is concerned with compensating for not granting him/her this access. The four approaches are not mutually exclusive, and in fact it is likely that they will be used in combination rather than in isolation. In the next sections, we describe each of these techniques and illustrate them with concrete realizations.

a) *Anthropomorphism*: This first approach aims at taking advantage of two facts. First, the operator has considerable prior experience with interacting through his/her own body. Second, even when operating the robot, the operator continues to co-perceive his/her own body (and presumably in a far more richer way than any display could ever permit). Anthropomorphism proceeds by improving the robot's resemblance to its human operator (cf. also the *bionic approach* [90]). It rests upon the idea that by mapping certain properties of the robot to those of the operator, the operator will perceive (or be

attuned to) these robot's characteristics through the perception of his/her own body (and/or his/her pre-attunement to these).

The properties matched can be of two different types. First, a mapping can be established between properties of the operator and of the robot which constrain what they can do (e.g., arm length for reaching, grasp size for grasping, body width for passing through). More generally, a relational property of the  $R-E_R$  subsystem which constrains action can be made to reflect an equivalent property of the  $A-E_A$  subsystem (where  $E_R$  and  $E_A$  denote the environment where the robot and agent evolve, respectively). For example, the ratio of the robot's frontal width to the size of remote apertures can be mapped to the ratio of the operator's shoulders width to the size of his/her every-day doors.

Second, a mapping can also be established between properties of the operator and of the robot (or relational properties of  $R-E_R$  and  $A-E_A$ ) which constrain how opportunities for action and control rules are *specified* within the structure of ambient stimulation. A typical example is that of eye-height. For example, under the constraint of an agent's eye-height, the optic flow can specify, up to a scaling factor, the width of an aperture in units of the agent's shoulders width (cf. Fig. 2). In mobile telerobotics, it has been shown that raising the robot's camera height toward the operator's own eye-height had a beneficial influence on his/her ability to judge whether the robot can pass through an aperture [91]. Eye-height is also known to influence the way other agent-environment properties map to particular features of optic flow, such as the relation between ground speed and global optic flow rate. With practice, humans become attuned to these relations and subsequently they may be fooled (at least for a certain period) when this relation is changed. For example, Owen and Warren [92] proposed that the higher height of 747 cockpits (compared to the plane on which pilots were taught) could explain the excessive taxiing speed of the first 747 pilots. Similarly, drivers switching between car and truck may undergo a recalibration phase [93]. This suggests that in mobile telerobotics placing the robot's camera at the operator's eye-height (e.g., current, walking, or car driving [1]) could prevent operators from misperceiving the robot's speed and allow for shorter adaptation times.

*b) Information Augmentation:* Information augmentation is the most intuitive solution. As anthropomorphism it consists in providing the operator the functional information he/she needs to control the actions identified during step 1 (Section IV-B2). However, this time the idea is to proceed by adding new feedbacks to the console. Information augmentation thus addresses the question of how to equip the console's displays such that the resulting structure of ambient stimulation (optics, acoustics, etc.) specifies the functional  $A-C-R-E$  relationships that have been identified during step 2 (Section IV-B1).

Before this functional information can be rendered to the operator, it must first be available to the platform. Some  $A-C-R-E$  system properties characterizing affordances and control rules are known a priori, either whole or in part, and thus can be directly stored on the teleoperation platform. For example, this is often the case for the robot's size and shape (e.g., which constrain its ability to pass through apertures). In all

other situations, the relevant variables have to be captured by the platform and thus the designer has to determine how they might be picked up. Typically, this preliminary phase will involve (i) identifying the ambient energy invariants that specify the affordances and control rules gathered during step 2 and (ii) identifying the technologies with which these invariants could be sampled (i.e., the sensors, possibly supplemented with data processing techniques).

The core phase of information augmentation then consists in rendering this functional information to the operator. This is achieved by laying out the console's displays so as to establish new 1:1 mappings between the control-relevant variables from step 2 and the invariants in ambient energies that stimulate the operator's perceptual systems (see Fig. 5 and Section III-B). This formulation underlines that the information available to the operator does not merely depend on what is rendered on the console. Rather, it is also contingent on the characteristics of the  $A-C$  system (e.g., spatial layout). As a consequence, the  $A-C$  relationship must be taken into account (and possibly adapted) when augmenting information.

A possible means for providing the operator with robot-scaled information about the remote environment is to include parts of the robot's body in the depiction of its environment on the console. In the context of mobile robotics for example, the camera position, orientation or size of the field of view can be adapted in such a way that the visual field comprises the robot's body [91]. Similarly, when the display includes a CG reconstruction of the environment in 3D (e.g., built from laser or map data), a scale robot avatar can be incorporated into it [33]. In both cases, if the robot stands sufficiently close to an aperture, the optical structure induced by the display will specify whether or not the robot can pass through the aperture. However, given the well known scale ambiguity of video images [94] and virtual environments [95], the relation between the width of an aperture and the width of the robot may be not specified in the optic array for apertures farther away. For example, the robot scale will not propagate to visible distant paths and objects if the area in-between is not filled with enough objects and/or texture, grid, or random tiles on the ground that would materialize the spatial continuity between the robot and the distant elements [45, p.77], [88]. One possibility is to display a perspective presentation of video feeds from several slanted cameras, or so-called perspective folding [34] (note however that the authors did not include the robot in the display). For ground vehicles, another solution would be to present a bird's eye view of the scene (i.e., with the line of sight orthogonal to the ground).

In the approaches mentioned above, the affordance was rendered to the operator by presenting on the console's display the constitutive elements of the  $A-C-R-E$  relationship which define the affordance (e.g., the robot and the aperture width). An alternative approach for providing body- and action-scaled information is to directly display the  $A-C-R-E$  relationship itself and/or the associated action boundaries and optimal point (see Section II-A). For example, as the robot's width is known and the width of paths can be retrieved (e.g., online from laser data), it should be possible to directly mark on the virtual ground of a laser-based 3D reconstruction, the paths that afford

passage and those that do not. More generally, the virtual ground could be painted with gradients of colors (e.g., from green to red) to reflect the degree to which safe, easy and efficient passage is afforded, given all that is known a priori and being gathered by the sensors: the size of the passage relative to the robot, the strength of the surface of support relative to its weight, the ambient heat relative to what the robot can handle, the path's slope and material relative to the power, mass distribution, and grip of the robot, etc.

Predictive displays (e.g., [80], [96], [97]) could also constitute an interesting tool for augmenting information. Indeed, by readily providing the operator with information about the consequence of his/her actions, this technique could allow to preserve the perception-action loop [98]. However, to qualify as augmentation of information in the ecological sense, the overlaid computer simulation must be designed in such a way that it, and more generally the induced structures within stimulation, provides the operator with the relevant functional information (e.g., the particular informational variables by which he/she can control the action at stakes). This brings us back to the three steps procedure we are describing.

During the last decades, the related research field of interface design for aircraft control has offered several benchmark examples of how displays can directly provide the pilot with informational variables that are relevant for flight control. A first example is the WrightCAD [32]. Unlike traditional primary flight displays which show state variables (altitude, compass, attitude, airspeed, vertical speed, etc.), the WrightCAD presents the pilot with functional information. For example, instead of materializing the plane's airspeed, it shows horizontal 'depression lines' whose rate of flow represents the difference between current speed and stall speed. A second example is the so-called perspective flight path display or tunnel-in-the-sky display, e.g., [99]–[101]. It presents the pilot with the flight trajectory to be followed by overlaying a perspective view of a wire frame tunnel on the display monitor. A central characteristic of the resulting display is that the tunnel's geometrical shape (and the way it changes) specifies the error in aircraft's position and orientation relative to the flight path to be followed (together with error dynamics). As a consequence, it provides the pilot with functional information that he/she can directly use to regulate his/her action on the aircraft's controls. Another example is the energy augmented version of the tunnel-in-the-sky display [29]. It provides to the pilot information about the aircraft's energy states and energy rates of change (e.g., the deviation of current total energy from the reference profile) which thus maps to the precise aspects over which the pilot's has control over (i.e., total energy through engine thrust and exchange between kinetic and potential through elevator). Continuing this line of work, new interface elements providing functional information about the spatiotemporal separation from other aircrafts [31] and from the ground [30] have also been developed.

c) *Action augmentation*: As we underlined in Sections II-B1 and II-D, the availability of information in the structure of ambient energies strongly depends on the exploratory actions in which the agent might engage to create information. It follows that, on some points, an action can play the same

role as a new display. An action can change the mapping between the properties of the *A-E* system and the structure of ambient stimulation, such that properties that were previously not specified become specified (and vice versa). Thus, an action can reveal agent-environment properties that would otherwise remain hidden to the perceiver.

As for anthropomorphism and information augmentation, the objective of action augmentation is to establish new mappings between functional variables identified during step 2 and patterns in the ambient energies stimulating the operator (see Fig. 5 and Section III-B). However, contra the two former approaches, the latter proceeds by granting the operator the opportunity to perform information-generating actions. Action augmentation is thus concerned with “*enhancing the operator's resources for action*” [102, p.16].

Of course action augmentation will fail to provide functional information to the operator if there is no display on the console. Action augmentation requires some information augmentation. Nonetheless, the approach is worth using because its original perspective can lead to consider different solutions. A simple example is the depth of objects which can be rendered either through stereo images or by providing the operator with the opportunity to generate motion parallax (e.g., [103]; note that depth, per se, is not information in the ecological sense). To count as action augmentation, the camera motion, or any other actions, have to be actively controlled by the operator as opposed to being passively experienced (e.g., driven by a computer program). Besides a difference in terms of stimulation content, the two versions also differ as a function of whether the informative patterns are obtained by the operator or imposed to him/her. This latter aspect is a central feature of action augmentation. This technique gives the operator the opportunity to generate information when he/she needs it and about the properties he/she is interested in. For example, instead of continuously overlaying a specific color on all the objects of a video display which are within reach of the robot, the mouse pointer could simply take a specific color (e.g., green) when passing over a reachable object.

By mapping control-relevant *A-C-R-E* properties on invariants of ambient stimulation, information augmentation provides the operator with an external model of task dynamics, and thus obviates the need for the operator to build an internal model of this dynamics [26]. In a similar fashion, by granting the operator with the opportunity to reveal hidden *A-C-R-E* constraints on action, action augmentation allows the operator to create his/her own external model of task dynamics [102]<sup>9</sup>.

d) *Delegation*: The last approach consists in delegating to the robot the action modes for which the information available to the operator is insufficient. Instead of trying to augment the operator in one way or another, the robot is left with the responsibility to achieve the actions whose control require information that the operator is missing. As a result, the operator remains in charge of the sole actions for which he/she has access to the necessary functional information.

<sup>9</sup>In his paper, Kirlik illustrates that point very neatly with the example of a cook preparing steaks on a grill in a restaurant. Even if the domain is not directly related to robotics, the example is really worth reading.



### 1. Identifying action modes

What is the purpose of the platform? What actions should it allow to do? How are they nested?

### 2. Characterizing action modes

What *A-C-R-E* properties define the corresponding affordances and control rules?

### 3. Implementing action modes

Anthropomorphism	Information augmentation	Action augmentation	Delegation
What operator's properties constrain these actions and their specification in equivalent unmediated situations?	How to make these <i>A-C-R-E</i> properties available to the platform? 1. by storing those which are known a priori 2. by capturing the others	How are they specified? Which sensors+processing can sample them?	
How to make the operator's stimulation <i>specify</i> these control-relevant properties of the <i>A-C-R-E</i> system...			How to make the platform control the action modes for which the information available to the operator is insufficient?
...by laying out the robot (and console) such that its characteristics match the operator's	...by laying out the console's display(s)	...by granting the operator opportunities to perform information-generating actions	

Fig. 6. Synthesis of the proposed framework for compensating for the alteration of available perceptual information in teleoperation.

Even if this technique does not imply to render information to the operator, it nevertheless necessitates that this information be available to the robot. Accordingly, delegation requires a preliminary phase of finding the means by which to gather the control-relevant *A-C-R-E* properties identified during step 2 (see IV-B3b: Information augmentation).

The delegation of control is a technique widely used in telerobotics (e.g., obstacle avoidance, safe guarding) and in related field of manned vehicle control (e.g., driving: ABS, power steering, cruise control, turn suggestion, etc.). An important concern of delegation is that unless the control of actions is completely entrusted to the robot, it necessary implies that control be shared between the operator and the robot. This cohabitation of human and autonomous control raises important issues, e.g., [104], [105]. From an ecological point of view, the presence of closed control loops (e.g., obstacle avoidance) changes the mapping between the operator's actions and their remote consequences (cf. Section III-A) and thus can impair the operator's ability to discover the invariants of stimulation which maps to the control-relevant *A-C-R-E* properties (cf. Section III-B). For example, when an obstacle avoidance system is activated, a same command issued by the operator can result in completely different robot behaviors depending on the spatial surroundings of the robot.

The nesting structure of actions modes identified during step 1 (Section IV-B1) could provide a useful framework for sharing out the work between the operator and the robot on a functional basis. In particular, the hierarchical relations between modes could be used to delegate the control of subsidiary actions to the robot while the operator remains in charge of controlling actions that exist at a higher level of abstraction. For example, to (make the robot) reach an object, the operator can be responsible for controlling the position of the robotic arm's end point (e.g., the robot's grasp) relative to the targeted environmental object, while the subsequent control of individual joints of the robotic arm and of the

mobile basis are delegated to the robot [106]. In complement, the distinction between the prospective vs. online control of action (which require knowledge of affordances and control rules, respectively) could also provide a useful delimitation for splitting responsibilities between the operator and the robot. For example, experimental participants seem to be more precise in judging whether a robot can pass through an opening than in being actually able to drive it through [70].

### C. Concluding Remarks

The proposed framework for coping with the alteration of available information is summarized on Fig. 6. One important characteristic of this framework (which is also common to EID and WDA) is that it requires knowledge of the domain being considered (steps 1 and 2). As a corollary, it will not be suited to design interfaces for controlling all-purpose robots –in the sense of platforms with no defined purpose– because this would require to identify and characterize every single possible action. As we already noted, the four techniques proposed for implementation (step 3) shall not be considered as independent solutions but more likely as complementary facets of a solution. These four approaches have different scopes and come with different advantages and disadvantages. For example, action augmentation could prove particularly useful for revealing the dynamical properties of the *A-C-R-E* system which constrain actions but are not specified in the absence of movement (e.g., maximum turning rate or maximum deceleration in mobile robotics). For their part, anthropomorphism and delegation could facilitate and accelerate the operator's mastery of the actions he/she has to perform with the teleoperation platform. However, the scope of action modes that can be handled with these three methods (e.g., the *A-C-R-E* properties that can be made available through action augmentation and anthropomorphism) is much narrower than for information augmentation. On the other hand, information augmentation proceeds by adding new feedbacks on the con-

sole and thus comes at the risk of overloading the display and of drowning relevant information of the moment, whereas the three techniques of action augmentation, anthropomorphism and delegation offer alternatives to imposing more information on the console's display.

From a methodological point of view, the work to be carried on in each step (identification, characterization and implementation) requires both analytical work and empirical investigation. For example, analysis can be used to determine what needs to be displayed to make the operator's ambient stimulation specify control-relevant *A-C-R-E* properties (what Flach called *inverse ecological optics* in case of a visual display [20]). However, the objective is not merely to make the structure of ambient stimulation specific but also to ensure that these structures can be picked up by the operator and successfully exploited to achieve the task at hand. Thus, experimental tests remain of great importance. Although this may appear a colossal work, it can capitalize on a large existing literature on human perception-action (but also in animals and autonomous agents) in which themes such as locomotion (walking, driving, flying, etc.) and manipulation (reaching, grasping, intercepting moving objects, etc) have extensively been studied.

## V. CONCLUSION

In this paper we attempted to illustrate how the ecological approach to perception and action could provide a useful basis for addressing the issues raised by the perceptual control of remote actions. It is worth noting that, in return, teleoperated robotics provides a unique opportunity for psychologists to put their models and hypotheses to the test. In particular, by lowering the number of degrees of freedom involved in the agent-environment interaction, instruments such as teleoperation platforms should allow for a better tractability of the processes at work [44]. Our approach differs from the EID framework in that it focuses on the control of action, not process. Throughout the paper, we pointed out some implications of this difference in scope for domain analysis (in particular that constraints on action spread over the entire *A-C-R-E* system). Our analysis of the means by which the interface can be made to support for action modes also provides a complementary perspective on the problem of the so-called semantic mapping [26]. Besides the four techniques we described, the proposal that, in fine, it is the structure of the operator's stimulation –and not the structure of the display per se– that should map to the domain constraints, could provide a common ground to address the use of 2D visual displays, of other displays (e.g., stereo, head-driven, see-through, auditory, haptic, multimodal) and thus of alternative interaction paradigms (enaction, mixed-reality).

After clarifying important ecological concepts and analyzing the impact of a teleoperation platform on perception and action from an ecological perspective, we proposed a framework for interface design which aims at providing the operator with the relevant functional information (or at compensating for not providing him/her that information). Our goal was not to provide a fully operational toolbox but rather to give some guidelines and general principles of an ecological approach

to designing HRI for teleoperated robotics. We hope that this contribution will stimulate debate, applications, and that it will be completed and improved. In a forthcoming paper, we will present a concrete example of how we used these guidelines to develop an interface for navigating a robot in a home environment.

## REFERENCES

- [1] D. D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human-robot coordination in future operations," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, no. 2, pp. 210–218, 2004.
- [2] R. R. Murphy, "Human-robot interaction in rescue robotics," *IEEE Trans. Syst., Man, Cybern. C*, vol. 34, no. 2, pp. 138–153, 2004.
- [3] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Trans. Syst., Man, Cybern. C*, vol. 37, no. 6, pp. 1231–1245, 2007.
- [4] J. S. Tittle, A. Roesler, and D. D. Woods, "The remote perception problem," in *Proc. Hum. Fac. Erg. Soc. 46th Annual Meeting*, 2002, pp. 260–264.
- [5] J. J. Gibson, *The Ecological Approach to Visual Perception*. Boston, MA, USA: Houghton-Mifflin, 1979.
- [6] C. F. Michaels and C. Carello, *Direct perception*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1981.
- [7] E. S. Reed, *Encountering the world: Toward an ecological psychology*. New York, USA: Oxford University Press, 1996.
- [8] J. J. Gibson, "Visually controlled locomotion and visual orientation in animals," *Brit. J. Psychol.*, vol. 49, pp. 182–194, 1958.
- [9] P. N. Kugler and M. T. Turvey, *Information, natural law, and the self-assembly of rhythmic movement*. Mahwah, NJ, USA: LEA, 1987.
- [10] R. Bajcsy, "Active perception," *Proc. IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.
- [11] J. Y. Aloimonos and A. Badyopadhyay, "Active vision," in *Proc. First IEEE Int. Conf. Comput. Vis.*, 1987, pp. 35–54.
- [12] A. P. Duchon, L. P. Kaelbling, and W. H. Warren, "Ecological robotics," *Adapt. Behav.*, vol. 6, no. 3–4, pp. 473–507, 1998.
- [13] R. R. Murphy, "Case studies of applying Gibson's ecological approach to mobile robots," *IEEE Trans. Syst., Man, Cybern. A*, vol. 29, no. 1, pp. 105–111, 1999.
- [14] A. C. Slocum, D. C. Downey, and R. D. Beer, "Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention," in *From animals to animats: Proc. 6th Int. Conf. Simul. Adapt. Behav.* Cambridge, MA, USA: MIT Press, 2000, pp. 430–439.
- [15] J. J. Gibson, "A theoretical field-analysis of automobile-driving," *Am. J. Psychol.*, vol. 51, no. 3, pp. 453–471, 1938.
- [16] J. J. Gibson, P. Olum, and F. Rosenblatt, "Parallax and perspective during aircraft landings," *Am. J. Psychol.*, vol. 68, no. 3, pp. 372–385, 1955.
- [17] J. M. Flach, P. A. Hancock, J. Caird, and K. J. Vicente, Eds., *Global Perspectives on the Ecology of Human-Machine Systems*. Hillsdale, NJ, USA: LEA, 1995.
- [18] P. A. Hancock, J. M. Flach, J. Caird, and K. J. Vicente, Eds., *Local Applications of the Ecological Approach To Human-Machine Systems*. Hillsdale, NJ, USA: LEA, 1995.
- [19] G. Smets, "Industrial design engineering and the theory of direct perception and action," *Ecol. Psychol.*, vol. 7, no. 4, pp. 329–374, 1995.
- [20] J. M. Flach, "The ecology of human-machine systems I: Introduction," *Ecol. Psychol.*, vol. 2, no. 3, pp. 191–205, 1990.
- [21] J. M. Flach and P. A. Hancock, "An ecological approach to human-machine systems," in *Proc. Hum. Fac. Erg. Soc. 36th Annual Meeting*, Chicago, IL, USA, 1992, pp. 1056–1058.
- [22] K. J. Vicente and J. Rasmussen, "The ecology of human-machine systems II: Mediating 'direct perception' in complex work domains," *Ecol. Psychol.*, vol. 2, no. 3, pp. 207–250, 1990.
- [23] D. A. Norman, *The psychology of everyday things*. New York, NY, USA: Basic Books, 1988.
- [24] —, "Affordance, conventions, and design," *interactions*, vol. 6, no. 3, pp. 38–43, 1985.
- [25] J. McGrenere and W. Ho, "Affordances: Clarifying and evolving a concept," in *Proc. Graphic Interface*, Montreal, Canada, 2000, pp. 179–186.

- [26] J. Rasmussen and K. J. Vicente, "Coping with human errors through system design: Implications for ecological interface design," *Int. J. Man Mach. Stud.*, vol. 31, no. 5, pp. 517–534, 1989.
- [27] J. Rasmussen, "Ecological interface design for complex systems: An example: Sead-uav systems," HURECON, Denmark, Tech. Rep., 1998.
- [28] J. A. Adams, C. M. Humphrey, M. A. Goodrich, J. L. Cooper, B. S. Morse, C. Engh, and N. Rasmussen, "Cognitive task analysis for developing unmanned aerial vehicle wilderness search support," *J. Cogn. Eng. Decis. Mak.*, vol. 3, no. 1, pp. 1–26, 2009.
- [29] M. H. J. Amelink, M. Mulder, M. M. van Paassen, and J. M. Flach, "Theoretical foundations for a total energy-based perspective flight-path display," *Int. J. Aviat. Psychol.*, vol. 15, no. 3, pp. 205–231, 2005.
- [30] C. Borst, "Ecological approach to pilot terrain awareness," Ph.D. dissertation, Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands, 2009.
- [31] S. B. van Dam, M. Mulder, and M. M. van Paassen, "Ecological interface design of a tactical airborne separation assistance tool," *IEEE Trans. Syst., Man, Cybern. A*, vol. 38, no. 6, pp. 1221–1233, 2008.
- [32] J. M. Flach, "Ready, fire, aim: A "meaning-processing" approach to display design," in *Attention and Performance XVII*, D. Gopher and A. Koriat, Eds. Cambridge, MA, USA: MIT Press, 1999, ch. 7, pp. 197–222.
- [33] C. W. Nielsen, M. A. Goodrich, and R. W. Ricks, "Ecological interfaces for improving mobile robot teleoperation," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 927–941, 2007.
- [34] M. Voshell, D. D. Woods, and F. Phillips, "Overcoming the keyhole in human-robot interaction: Simulation and evaluation," in *Proc. Hum. Fac. Erg. Soc. 49th Annual Meeting*, Orlando, FL, IL, USA, 2005, pp. 442–446.
- [35] J. J. Gibson, *The Senses Considered as Perceptual Systems*. Boston, MA, USA: Houghton-Mifflin, 1966.
- [36] —, "The theory of affordances," in *Perceiving, acting, and knowing: Toward an ecological psychology*, R. E. Shaw and J. Bransford, Eds. Hillsdale, NJ, USA: LEA, 1977, pp. 67–82.
- [37] W. H. Warren, "Action modes and laws of control for the visual guidance of action," in *Complex movement behaviour: 'The' motor-action controversy*, O. G. Meijer and K. Roth, Eds. Amsterdam, Netherlands: Elsevier, 1988, ch. 14, pp. 339–380.
- [38] T. A. Stoffregen, "Affordances as properties of the animal-environment system," *Ecol. Psychol.*, vol. 15, no. 2, pp. 115–134, 2003.
- [39] M. T. Turvey, "Affordances and prospective control: An outline of the ontology," *Ecol. Psychol.*, vol. 4, no. 3, pp. 173–187, 1992.
- [40] W. H. Warren, "Perceiving affordances: visual guidance of stair climbing," *J. Exp. Psychol. Human*, vol. 10, no. 5, pp. 683–703, 1984.
- [41] L. S. Mark, K. Nemeth, D. Gardner, M. J. Dainoff, J. Paasche, M. Duffy, and K. Grandt, "Postural dynamics and the preferred critical boundary for visually guided reaching," *J. Exp. Psychol. Human*, vol. 23, no. 5, pp. 1365–1379, 1997.
- [42] D. L. Gardner, L. S. Mark, J. A. Ward, and H. Edkins, "How do task characteristics affect the transitions between seated and standing reaches?" *Ecol. Psychol.*, vol. 13, no. 4, pp. 245–274, 2001.
- [43] W. H. Warren and S. Whang, "Visual guidance of walking through apertures: body-scaled information for affordances," *J. Exp. Psychol. Human*, vol. 13, no. 3, pp. 371–383, 1987.
- [44] R. E. Shaw, O. M. Flascher, and E. E. Kadar, "Dimensionless invariants for intentional systems: Measuring the fit of vehicular activities to environmental layout," in *Global Perspectives on the Ecology of Human-Machine Systems*, J. M. Flach, P. A. Hancock, J. Caird, and K. J. Vicente, Eds. Hillsdale, NJ, USA: LEA, 1995, pp. 293–358.
- [45] J. J. Gibson, *The perception of the visual world*. Boston, MA, USA: Houghton-Mifflin, 1950.
- [46] J. Blauert, *Spatial hearing. The psychophysics of human sound localization*. Cambridge, MA, USA: MIT Press, 1997.
- [47] D. N. Lee, "The optic flow field: The foundation of vision," *Philos. T. Roy. Soc. B*, vol. 290, no. 1038, pp. 169–179, 1980.
- [48] J. J. Koenderink, "Optic flow," *Vision Res.*, vol. 26, no. 1, pp. 161–179, 1986.
- [49] H. von Helmholtz, *Optique physiologique, volumes I et II*. Paris, France: Masson, 1867.
- [50] E. J. Gibson, J. J. Gibson, O. W. Smith, and H. Flock, "Motion parallax as a determinant of perceived depth," *J. Exp. Psychol.*, vol. 58, no. 1, pp. 40–51, 1959.
- [51] D. N. Lee, "Visual information during locomotion," in *Perception: Essays in honour of J. J. Gibson*, R. B. MacLeod and J. Herbert L. Pick, Eds. Ithaca, NY, USA: Cornell University Press, 1974, ch. 14, pp. 250–267.
- [52] —, "A theory of visual control of braking based on information about time-to-collision," *Perception*, vol. 5, no. 4, pp. 437–459, 1976.
- [53] B. K. Shaw, R. S. McGowan, and M. T. Turvey, "An acoustic variable specifying time-to-contact," *Ecol. Psychol.*, vol. 3, no. 3, pp. 253–261, 1991.
- [54] D. H. Ashmead, D. L. Davis, and A. Northington, "Contribution of listeners' approaching motion to auditory distance perception," *J. Exp. Psychol. Human*, vol. 21, no. 2, pp. 239–256, 1995.
- [55] H. Y. Solomon and M. T. Turvey, "Haptically perceiving the distances reachable with hand-held objects," *J. Exp. Psychol. Human*, vol. 14, no. 3, pp. 404–427, 1988.
- [56] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Percept. Psychophys.*, vol. 14, no. 2, pp. 201–211, 1973.
- [57] S. Runeson and G. Frykholm, "Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention," *J. Exp. Psychol. Gen.*, vol. 112, no. 4, pp. 585–615, 1983.
- [58] H. A. Sedgwick, "Space perception," in *Handbook of Perception and Human Performance, volume I: Sensory Processes and Perception*, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds. New York, NY, USA: Wiley, 1986, ch. 21, pp. 21.1–21.57.
- [59] R. Marcilly and M. Luyat, "The role of eye height in judgment of an affordance of passage under a barrier," *Curr. Psychol. Lett.*, vol. 24, no. 1, 2008.
- [60] L. S. Mark, "Eyeheight-scaled information about affordances: a study of sitting and stair climbing," *J. Exp. Psychol. Human*, vol. 13, no. 3, pp. 361–370, 1987.
- [61] T. A. Stoffregen and B. G. Bardy, "On specification and the senses," *Behav. Brain Sci.*, vol. 24, no. 2, pp. 195–261, 2001.
- [62] B. Mantel, B. G. Bardy, and T. A. Stoffregen, "Multimodal perception of reachability expressed through locomotion," *Ecol. Psychol.*, vol. 22, no. 3, pp. 192–211, 2010.
- [63] J. A. S. Kelso, *Dynamic Patterns: The Self-organization of Brain and Behavior*. Cambridge, MA, USA: MIT Press, 1995.
- [64] W. H. Warren, "The dynamics of perception and action," *Psychol. Rev.*, vol. 113, no. 2, pp. 358–389, 2006.
- [65] K. Ghose, T. K. Horiuchi, P. S. Krishnaprasad, and C. F. Moss, "Echolocating bats use a nearly time-optimal strategy to intercept prey," *PLoS Biol.*, vol. 4, no. 5, p. e108, 2006.
- [66] A. P. Duchon and W. H. Warren, "Robot navigation from a gibsonian viewpoint," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, San Antonio, TX, USA, 1994, pp. 2272–2277.
- [67] B. R. Fajen and W. H. Warren, "Behavioral dynamics of steering, obstacle avoidance, and route selection," *J. Exp. Psychol. Human*, vol. 29, no. 2, pp. 343–362, 2003.
- [68] G. Schöner, M. Dose, and C. Engels, "Dynamics of behavior: Theory and applications for autonomous robot architectures," *Robot. Auton. Syst.*, vol. 16, no. 2–4, pp. 213–245, 1995.
- [69] W. W. Gaver, "Technology affordances," in *Proc. CHI'91*, New Orleans, LA, USA, 1991, pp. 79–84.
- [70] K. S. Jones, E. A. Schmidlin, and L. B. R. Johnson, "Tele-operators' judgments of their ability to drive through apertures," in *Proc. 4th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI'09)*, La Jolla, CA, USA, 2009, pp. 249–250.
- [71] C. Lenay, "Enaction, externalisme et suppléance perceptive," *Intellectica*, vol. 43, pp. 27–52, 2006.
- [72] B. G. Bardy and B. Mantel, "Ask not what's inside your head, but what your head is inside of (Mace, 1977), a commentary on Lenay (2006) and Jacob (2006)," *Intellectica*, vol. 43, pp. 53–58, 2006.
- [73] T. A. Stoffregen, B. G. Bardy, and B. Mantel, "Affordances in the design of enactive systems," *Virtual Reality*, vol. 10, no. 1, pp. 4–10, 2006.
- [74] D. A. Norman and J. Nielsen, "Gestural interfaces: A step backward in usability," *interactions*, vol. 17, pp. 46–49, September 2010.
- [75] M. Baker, R. Casey, B. Keyes, and H. A. Yanco, "Improved interfaces for human-robot interaction in urban search and rescue," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, vol. 3, 2004, pp. 2960–2965.
- [76] G. Terrien, T. Fong, C. Thorpe, and C. Baur, "Remote driving with a multisensor user interface," in *Proc. 30th SAE Int. Conf. Environ. Syst., Toulouse*, France, July 2000.
- [77] F. Ferland, F. Pomerleau, C. T. L. Dinh, and F. Michaud, "Egocentric and exocentric teleoperation interface using real-time, 3d video projection," in *Proc. 4th ACM/IEEE Int. Conf. Hum.-Robot Interact. (HRI'09)*, La Jolla, CA, USA, 2009, pp. 37–44.
- [78] H. K. Keskinpala and J. A. Adams, "Objective data analysis for a PDA-based human robotic interface," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2004, pp. 2809–2814.

- [79] B. Keyes, M. Micire, J. L. Drury, and H. A. Yanco, "Improving human-robot interaction through interface evolution," in *Human-Robot Interaction*, D. Chugo, Ed. Croatia: INTECH, 2010, pp. 183–202.
- [80] R. W. Ricks, C. W. Nielsen, and M. A. Goodrich, "Ecological displays for robot interaction: A new perspective," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, vol. 3, Sendai, Japan, 2004, pp. 2855–2860.
- [81] J. M. Flach, K. J. Vicente, F. Tanabe, K. Monta, and J. Rasmussen, "An ecological approach to interface design," in *Proc. Hum. Fac. Erg. Soc. 42nd Annual Meeting*, Chicago, IL, USA, 1998, pp. 295–299.
- [82] E. Pasquinelli, "Presence (theories of)," in *Enaction and enactive interfaces: A handbook of terms*, A. Luciani and C. Cadoz, Eds. Grenoble, France: Enactive System Books, 2007.
- [83] T. B. Sheridan, "Teleoperation, telerobotics and telepresence: A progress report," *Control Eng. Pract.*, vol. 3, no. 2, pp. 205–214, 1995.
- [84] J. M. Loomis, "Distal attribution and presence," *Presence-Teleop. Virt.*, vol. 1, pp. 113–119, 1992.
- [85] R. Casati and E. Pasquinelli, "Is the subjective feel of "presence" an uninteresting goal?" *J. Vis. Lang. Comput.*, vol. 16, no. 5, pp. 428–441, 2005.
- [86] T. A. Stoffregen, B. G. Bardy, L. J. Smart, and R. Pagulayan, "On the nature and evaluation of fidelity in virtual environments," in *Virtual and adaptive environments*, L. J. Hettinger and M. Haas, Eds. Mahwah, NJ, USA: LEA, 2003, pp. 111–128.
- [87] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall, "Does the quality of the computer graphics matter when judging distances in visually immersive environments?" *Presence-Teleop. Virt.*, vol. 13, no. 5, pp. 560–571, 2004.
- [88] P. J. Stappers, W. W. Gaver, and C. J. Overbeeke, "Beyond the limits of real-time realism: Moving from stimulation correspondence to information correspondence," in *Virtual and adaptive environments*, L. J. Hettinger and M. Haas, Eds. Mahwah, NJ, USA: LEA, 2003, pp. 91–110.
- [89] T. A. Stoffregen and T. Nonaka, "Nested affordances," presented at the 14th International Conference on Perception and Action, Yokohama, Japan, July 2007.
- [90] Y. Rybarczyk, D. Mestre, P. Hoppenot, and E. Colle, "Implémentation télérobotique de l'anticipation sensorimotrice pour optimiser la coopération homme-machine," *Trav. Humain*, vol. 67, no. 3, pp. 209–233, 2004.
- [91] K. S. Moore, J. A. Gomer, C. C. Pagano, and D. W. Moore, "Perception of robot passability with direct line of sight and teleoperation," *Hum. Factors*, vol. 51, no. 4, pp. 557–570, 2009.
- [92] D. H. Owen and R. Warren, "Perceptually relevant metrics for the margin of safety: A consideration of global optical flow and density variables," Ohio State University Research Foundation, Columbus, OH, USA, Tech. Rep., 1982.
- [93] W. Schiff and W. Arnone, "Perceiving and driving: Where parallel roads meet," in *Local Applications of the Ecological Approach To Human-Machine Systems*, P. A. Hancock, J. M. Flach, J. Caird, and K. J. Vicente, Eds. Hillsdale, NJ, USA: LEA, 1995, vol. 2, ch. 1, pp. 1–35.
- [94] P. J. Stappers, "Scaling the visual consequences of active head movements: A study of active perceivers and spatial technology," Ph.D. dissertation, Delft University, Delft, Netherlands, 1992.
- [95] J. M. Loomis and J. M. Knapp, "Visual perception of egocentric distance in real and virtual environments," in *Virtual and adaptive environments*, L. J. Hettinger and M. Haas, Eds. Mahwah, NJ, USA: LEA, 2003, pp. 21–46.
- [96] M. V. Noyes, "Superposition of graphics on low bit rate video as an aid in teleoperation," Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, June 1984.
- [97] T. B. Sheridan, "Space teleoperation through time delay: Review and prognosis," *IEEE J. Robot. Autom.*, vol. 9, no. 5, pp. 592–606, 1993.
- [98] H. G. Stassen and G. J. F. Smets, "Telemanipulation and telepresence," *Control Eng. Pract.*, vol. 5, no. 3, pp. 363–374, 1997.
- [99] A. J. Grunwald, J. B. Robertson, and J. J. Hatfield, "Evaluation of a computer-generated perspective tunnel display for flight-path following," NASA, Washington (D.C.), U.S.A., Tech. Rep. 1736, 1980.
- [100] M. Mulder, "Cybernetics of a tunnel-in-the-sky display," Ph.D. dissertation, Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands, 1999.
- [101] E. Theunissen, "Integration of human factors aspects in the design of spatial navigation displays," in *Virtual and adaptive environments*, L. J. Hettinger and M. Haas, Eds. Mahwah, NJ, USA: LEA, 2003, pp. 369–389.
- [102] A. Kirlik, "The ecological expert: Acting to create information to guide action," *Proceedings of the Fourth Symposium on Human Interaction with Complex Systems*, pp. 15–27, 1998.
- [103] G. J. F. Smets, C. J. Overbeeke, and M. H. Stratmann, "Depth on a flat screen," *Percept. Motor Skill*, vol. 64, no. 3c, pp. 1023–1034, 1987.
- [104] J.-M. Hoc, F. Mars, I. Milleville-Pennel, E. Jolly, M. Netto, and J.-M. Blosseville, "Evaluation of human-machine cooperation modes in car driving for safe lateral control in bends: Function delegation and mutual control modes," *Trav. Humain*, vol. 69, pp. 153–182, 2006.
- [105] D. J. Bruemmer, D. A. Few, R. L. Boring, J. L. Marble, M. C. Walton, and C. W. Nielsen, "Shared understanding for collaborative control," *IEEE Trans. Syst., Man, Cybern. A*, vol. 35, no. 4, pp. 494–504, 2005.
- [106] K. Nait-Chabane, S. Delarue, P. Hoppenot, and E. Colle, "Strategy of approach for seizure of an assistive mobile manipulator," *Robot. Auton. Syst.*, vol. 57, no. 2, pp. 222–235, 2009.



**Bruno Mantel** received the Dipl.Ing. degree in computer science (spec. cognitive science) from the Graduate School of Computer Science and Advanced Technologies (EPITA), Kremlin-Bicêtre, France in 2002, the D.E.A. degree in human motor skills from Université Paris-Sud, Orsay, France in 2004 and the Ph.D. degree in human movement sciences from Montpellier I University, Montpellier, France in 2009.

He is currently a post-doctoral fellow in robotics at the Computer science, integrative biology & complex systems (IBISC) laboratory, University of Evry Val d'Essonne, France. His research focuses on the intrinsically multimodal, active and functional nature of perception and on its implications for the design and evaluation of man-machine interfaces.

Dr. Mantel is a member of the International Society for Ecological Psychology, of the French Association for Cognitive Research (ARCo), and of the Association for Computing Machinery (SIGCHI).



**Philippe Hoppenot** received the Dipl.Ing. degree from the National Institute of Telecommunication (INT), Evry, France in 1991, the master degree in applied physics from Paris XII University, Créteil, France, in 1992 and the PhD degree in robotics from the University of Evry, Evry, France, in 1997.

He is presently professor at the University of Evry Val d'Essonne, France. His research interests include man-machine cooperation, rehabilitation engineering, robotics and remote control. He has worked on disabled people assistance by the way of giving them the possibility to remotely control a mobile arm to perform daily life activities. He is currently involved in European and French projects on helping elderly people, especially with cognitive impairments, to stay at home.

Prof. Hoppenot is a member of the IFRATH (Federative Institute for Research on Technical Aids to Disabled People) and a member of the editorial committee of the Handicap conference, organised by IFRATH every two years.



**Etienne Colle** received the P.G.Dip. and the PhD in assistive robotics from Paris XII University, Créteil, France, in 1979 and 1983 respectively.

He is a first-class professor at the University of Evry Val d'Essonne, France. He is the former director (until 2009) of the Complex Systems Laboratory (LSC) and IBISC laboratory. He is currently head of the HANDS group (HANDicap et Santé). His field of expertise includes robotics and telerobotics for disabled and elderly people, and human-robot interaction (or cooperation). His current work focus on coordinated control of a mobile arm, man-machine cooperation, control modes, user-centered design and more recently ambient robotics.

Prof. Colle is a co-founder (and co-leader) of the IFRATH (Federative Institute for Research on Technical Aids to Disabled People), an institution created in 1996 which aims at federating rehabilitation researches in France.