14 Embedded Multimodal Interfaces in Robotics: Applications, Future Trends, and Societal Implications

E. A. Kirchner, S. H. Fairclough, F. Kirchner

14.1 Introduction

In the past, robots were primarily used to perform work that was either too hard, too dangerous or simply too repetitive for humans, e.g., assembly line work, or work that could be done much faster by a robotic system, such as placement work. In the future, ***human-robot interaction*** will cover a much broader range of scenarios, from working interactively with humans in the context of industrial manufacturing to robotic appliances designed to care the elderly; even in applied areas, such as autonomous robots in space or operating underwater, the demand for robots to interact or to be intuitively controlled is growing. Hence, interaction will not only involve direct control of a robot or information exchange but will include direct cooperation and physical interaction between human and robot, i.e., ***human-robot cooperation***. While direct cooperation has tremendous advantages it also presents a number of significant challenges that should not be underestimated. Advanced interfaces to enable human-robot cooperation will be required to meet these challenges and the needs of human-robot interaction in the future.

These interfaces must not only promote accurate and easy ***explicit interaction*** between humans and robots but should also enable ***implicit interaction***. Explicit interaction requires the intentional production of ***active states*** or commands, i.e., the production of speech or specific brain signals or gestures. On the other hand, implicit interaction operates by monitoring ***passive states***, i.e., emotions, mental workload, fatigue, motivation, which arise spontaneously, or requires the interpretation of spontaneous active states passively without the ***awareness*** or intention of the user to improve interaction. Examples of this mechanic would be the interpretation of the user's brain signals to prepare a robot for an upcoming movement in order to enable a greater degree of coordination between human and robot or the estimation of mental workload of the user to reduce or increase interactions task-frequency (for examples see Section 14.4). Considering enhancement of awareness for explicit interaction, it can be stated that the user is responsible to increase the awareness of the robot with respect to his or her needs and the function of the interface is to translate the active state of the human into an appropriate command, which is actioned by the robotic system. In this case, the interface between user and robot is a communicator. In case of using implicit control pathways the interface must translate the active or passive state of the human into a robotic response that is both timely and intuitive from the perspective of the user. The human does not intentionally contribute to this process and may not even be aware of the underlying mechanism. The interface serves as a monitor and translator that interprets the state of the human and possible intention (see Section 14.3 for more details). Both types of interfaces, which enable implicit or ***explicit control***, can be combined into a hybrid system. In case of this hybrid, interfaces between human and robot use multimodal input and generate multimodal output. Such advanced interfaces are often complex and of the type of embedded multimodal-multisensor interfaces.

While it is obvious that such interfaces serve human needs by enabling explicit and implicit interaction to optimize cooperation with respect to the needs of the human user, the interface provides two-way communication. For example, there are interfaces that enable robots to make use of human cognitive resources in those cases where they have reached the limits of autonomous behavior. Such interfaces are required for highly specialized robotic systems, which can be found in industrial environments. With respect to their specialized domain, these robots can outperform a human by contributing specialized solutions to a specific problem, but may not be able to cope with small deviations from a defined protocol. Furthermore, this type of two-way interface is required for the control of semi-autonomous robots, e.g., for teleoperation. One requirement of such interfaces is to provide feedback to the human with respect to the state of the robot. Therefore, a robotic system must either have a perception of itself, which can be transferred to the human, or an interface to interpret the state of the robot. In this case, the interface works by either transferring the robot's state or inferring the state of the robot.

There are robotic applications that must deliver highly demanding and sensitive modes of interaction with their human users, since they have, to an extent, replaced capabilities that would usually be performed by human being, as in the care of elderly people, or robots that extend the body, i.e., provide sensory-motor functions such as exoskeletons, surgical robots, and assistive robots that aid the human user by delivering services or information. In this case, the performance demands on the human-robot interface are even higher and often transfer or interpret states bi-directionally. In summary, a combination of the aforementioned abilities of the different types of interfaces is required to develop new and advanced multimodal-multisensor interfaces that:

- provide the human operator with a greater insight into the robot's state for better control;
- provide the robot with improved insight into the intention of the human for better support or support as needed, or allow the interaction to adapt to the human's need,
- enable the robotic device to extend the human body and senses and to be used as if it were part of the human body;
- enable the robot to learn from a human in order to imitate their behavior or to learn to understand the human behavior to become a better interaction partner; and
- inherently assure the automated detection or even avoidance of malfunction and safety of interaction.

This chapter describes some measures and approaches that can fulfill the listed requirements for advanced multimodal-multisensor interfaces. While we try to give explanations and examples for all the requirements of these systems, we will focus on approaches that make use of implicit interpretation of the human state. When using implicit interpretation of the human state, the array of measures that are required to optimize human-robot interaction depends on the type of user representation that the robot requires in order to interact or to cooperate with the person. If the robot must simply avoid hitting a person or colliding with him/her, all it needs to know is where the person is located in space. No explicit interaction is required. For this situation, only a simple awareness of the user, e.g., on her or his location in space and movements, is needed. Robots that work with the elderly and must exercise soft social skills require a much higher level of awareness of the user (and of course that means more sensors and measures). Hence, there is a relationship between the type and sophistication of the interface, its ability to interpret the human state, and the level of awareness of the user that is required for such complex interaction. Therefore, the level of awareness of the user required by the robot is hierarchical; the robot can have an awareness of the user as: (1) an object in space, (2) a co-worker or partner (what are they supposed to do, what tasks are they trying to complete), (3) an individual (gender, age, personality) and (4) a dynamic entity with respect to intentions and psychological states. Simple robots may need just (1) and (2) in order to interact or to cooperate. Robots that are designed to personalize interaction to the person would require information about stable traits of the person (3) and the means to detect dynamic changes (4). Hence, interfaces can be scaled according to the level of interaction that is required. The higher the level, the more likely it is that multimodal-multisensor interfaces are required.

The good news is that for human-robot interaction, robots are able to directly make use of a range of measures and data as part of the multimodal-multisensor interface. We will discuss the usage of ***Psychophysiological measures*** and how they can improve interaction and especially cooperation witin this type of advanced interface. However, it is very difficult or even impossible to always interpret the state or intention of the human with one 100% accuracy. This fact is the biggest challenge in any human-machine interaction and has direct implication for the subjective perception of reliability within this interaction. If a system is deemed unreliable, it will fail to win the trust of the user. The issue of trust is particularly sensitive for multimodal-multisensor interfaces. These systems are designed to respond with a degree of autonomy, hence the user must cede a degree of control to the system. In addition, these systems monitor a range of measures related to behavior and the psychological status of the person. These data are personal and sensitive, and interaction with this type of advanced robotic system may trigger a range of societal and ethical issues around data privacy and data security (see Section 14.5). Hence, trust is multifaceted during these interactions; the user must trust in the technical proficiency of the system and be confident that their personal data is secure while they interact with the interface.

Interfaces are often not stand-alone systems in robotics. They are ***embedded multimodal-multisensor interfaces*** that are deeply integrated into the system's control and into the context of interaction, requiring an automated analysis of interaction context. In the future, they will develop self-adaptive properties, which require new techniques, hardware, and algorithms as discussed in this chapter. In addition, robotics presents a huge challenge for safety, especially when humans physically interact with robots that exert high forces and accelerations together with a high net weight. Since safety of interaction is the fundamental requirement for human-robot cooperation, we will begin by discussing new and upcoming approaches from robotics to assure safety (Section 14.2). This discussion brings us directly to the definition and relevance of ***embedded multimodal interfaces*** in human robot interaction (Section 14.3), which often belong to the group of multimodal-multisensor interfaces. In Section 14.4 we give some application examples for embedded multimodal interfaces and explain how they can enable or improve human-robot interaction. Finally, in Section 14.5, future trends in embedded multimodal interfaces and societal implications are discussed.

For a detailed application scenario for an embedded multimodal interface in robotics, see the Glossary and Focus Questions.

14.2 Inherently Safe Robots—a Prerequisite for Human-Robot Cooperation

To assure safety in human-robot interaction, the most intuitive approach is to make robots inherently safe. Safety during human-robot cooperation is not only an additional benefit but often an indispensable criterion to enable sharing of common spaces or ensuring physical collaboration between fragile humans and powerful robots. Safety can be implemented on different levels during robot design to enhance reliability. Inherent safety in human-robot cooperation is achieved through a three-level process. On the lowest level (level 1), safety is ensured directly by the design of the electro-mechanical hardware. Therefore, this level is also referred to as the ***safety by design*** level: we can distinguish three parallel paths on this level (see Figure 14.1). First, the most straightforward path is the process of mechanical design itself, such as those classic lightweight designs that are standard in robotics. However, in recent years new smart materials can further enhance the lightweight design ethos. This process is enhanced by recent advances in 3D-printing technologies, which allow mechanical design to go in directions that were unthinkable using classic technologies, like embedding electronic circuitry and signals into the structure of the components. The integration of channels with complex 3D structure into the components, e.g., internal cabling, represents another approach. These developments are paralleled by advances in the design of robotic actuators. Using this approach compliant elements are embedded into actuators by serializing, e.g., a spring with a motor (and gear), which complicates the control of such an actuator on the algorithmic level, but enables built-in safety as external forces that act on the robot are absorbed by the spring - instead of the

Glossary

An **active BCI** is a brain-computer interface that derives its outputs based upon a voluntary act of explicit control from the human, e.g., generates motor imagery consistent with movement of right hand to move cursor to the right.

Active state is a psychological state associated with a volitional act or intention generated by the person, e.g., to open the door.

Application-specific safety level describes the concept to include information into the robot's behavioral control that comes from sensors that are not part of the robot itself. This concept uses the information from sensors that are placed outside of the robot to monitor the environment, e.g., a workplace in a production line and which are used in more traditional applications to create strict safety boundaries around the workplace. In more advanced approaches this information is used differently; here it helps to derive contextual information that can be used to adapt the robot's behavior instead of overwriting it. For example, in a more traditional scenario a violation of the safety boundary by a human walking by would result in a full stop of the production line. The more advanced concepts would predict the humans path and instead of stopping the production line would only reduce the speed of the moving robots. This concept therefore modifies the behavior of the robot as commanded by the ***high-level control*** module instead of overwriting it.

Awareness can refer to perceiving sensory stimuli in the environment including other actors, which may be machines or peoples.

Biocybernetic control describes a model of ***closed-loop control*** (negative or positive control) wherein measures are derived from psychophysiological or neurophysiological sources and converted into control input for an adaptive computer system.

Covert measure is a measure of human behavior or performance that cannot be detected based upon human perception, e.g., heart rate, and brain activity.

Closed-loop control is a control system that uses the concept of an open-loop system as its forward path but has one or more feedback loops (hence its name) or paths between its output and its input.

Embedded brain reading is an approach for user state detection, which is based on the online analysis of brain activity. Brain activity is used which is spontaneously evoked during human-machine interaction. The approach is deeply embedded into the system's control, the context of interaction, and makes use of multimodal data. It is applied for implicit interaction, i.e., to non-intentionally adapt or drive explicit interaction.

Embedded multimodal interface is an interface that makes use of **multimodal data** from **multimodal input** and is able to generate **multimodal output**. Its main characteristic is that it is deeply incorporated into the control of the robotic system, and may be subject to complex adaptation mechanisms such as ***reflexive adaptation***. While its function might be to gain explicit control of a system, it might be subject to implicit control to be adapted to the human's or system's needs.

Glossary

Explicit control represents a mode of input control where the user intentionally generates a specific behavior in order to achieve a specified goal, e.g., move a cursor upward.

Explicit interaction is a mode of human-computer interaction where the human user is fully cognizant of the issuing of commands and receives explicit feedback from the computer.

High-level control refers to the specification and feedback control of target positions in 3D space that must be reached by the end effector of a multi-joint robot based on information coming from sensors sampling the robot itself and its environment.

Human-robot cooperation is a subfield of human-robot interaction where a human and robot or teams of humans and robots work or act together to reach a shared goal. It often requires direct contact between human and robot or a shared workspace.

Human-robot interaction is any interaction between a human and a robot or teams of humans and robots including communication, control, feedback, direct contact, or information exchange.

A **hybrid BCI** describes a brain-computer interface that combines ***active BCI*** with either ***passive*** or ***reactive BCI*** or other measures such as eye movements or heart rate.

Imitation learning enables a robotic system to learn from demonstrations of nearly optimal policies executions given by a teacher (e.g., a human mentor). It is often used to initialize reinforcement learning to avoid time consuming learning from scratch.

Implicit interaction is a mode of human-robot interaction where the human user is not aware of the issuing of (control) commands that may be used for the control of a technical system or adaptation of an interface to the needs of the technical system or user. The user may or may not receive explicit feedback from the computer.

Internal state of a robot is computed on the basis of all sensor information directly or indirectly available to the robot. Directly available information is all information that comes from the robot's own sensors, while indirectly available information is all information that comes from sensors that are external to the robot but that the robot can access through communication pathways. The set of internal states of a robot is in most cases a finite set of eventually multi-dimensional vectors. Elements of this set are computed through means of clustering that range from simple thresholds to complex statistical methods.

Low-level control refers to the direct feedback control of movements of the motors in the joints of a multi-joint robot using sensor information coming directly from the individual motors to reach a specified position in 3D space.

Glossary *(continued)*

Neurophysiological measures represents the act of measurement based on physiological activity from cerebral sites in the human brain. These measures may be based upon electrical activity (electrocortical, electroencephalographical) or neurovascular changes (functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS)).

Overt measure is a measure of human behavior or performance that can be detected based upon human perception, e.g., voice commands, gestures. A

passive BCI is a brain-computer interface that derives its outputs from arbitrary brain activity without the purpose of voluntary control.

Passive state is a spontaneous psychological state that arises during behavior without an intention on the part of the person, e.g., fatigue, frustration.

Physiological computing refers to a field of research in human-computer interaction wherein ***Physiological measures*** derived from the human user are used as a source of input control for a computer system or interface.

Physiological measures describes the act of measurement based on processes related to human physiological functions.

Psychophysiological measures describes the act of measurement and inference wherein psychological processes and concepts are inferred on the basis of physiological measurements from the autonomic nervous system. A

reactive BCI is a brain-computer interface that derives its outputs from brain activity arising in reaction to external stimulation, e.g., a visual stimulus or sound.

Reflexive adaptation refers to a second-order process of adaptation whereby the computer makes an autonomous response and subsequently monitors the response of the human user to that response in order to inform future responses.

Safety by design refers to the fact that next-generation technical systems for human-robot cooperation will include, e.g., a compliant element in their actuators that absorbs energy. Thereby safety is an integral part of the mechanical construction of the system.

Temporal cascaded approach is an approach of using multimodal data in a timely sequenced fashion where the usage and outcome of analysis of one data type influences the analysis or choice of a second- or higher-order data type.

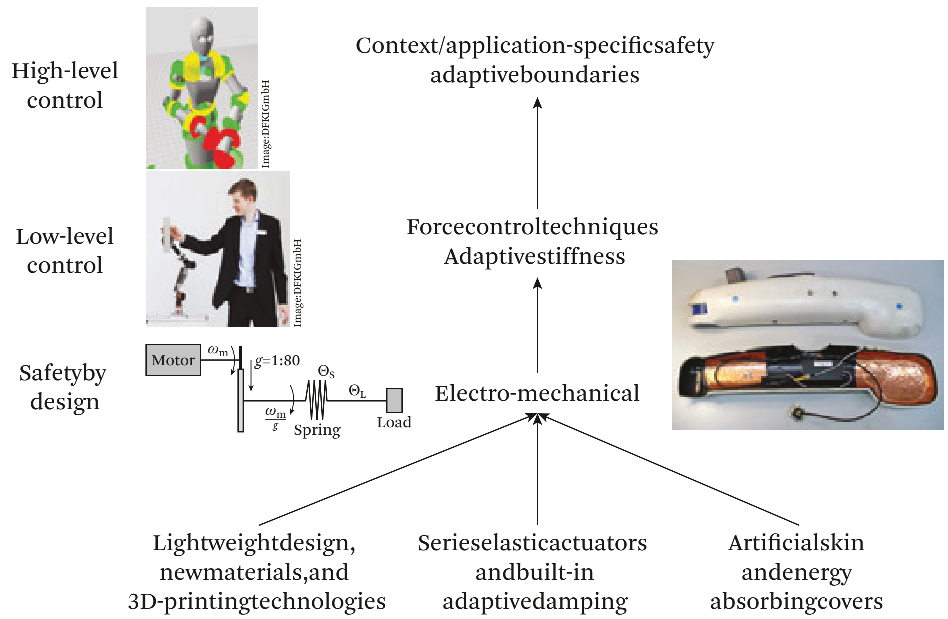

Figure 14.1 Safety levels for human-robot cooperation.

human body as in classical soft robots. Combining this approach with a high speed / high precision torque control technique on the algorithmic level [Bargsten and de Gea Fernández 2015] results in a robotic system that can be designed to be inherently safe (see video referred in Figure 14.2).

Finally, the level of safety by design can be realized by providing the new generation robots with a sense of touch. Touch or any kind of haptic information for robots has been largely ignored by robot designers over recent decades because the requisite technology was not available. Instead, vision was the dominant sensor for robotic perception and the primary means of avoiding contact with the user. Yet robots that directly interact with humans—e.g., while building or installing the windshield of an automobile on a shop floor of a car manufacturer—cannot completely avoid physical contact. Furthermore, physical contact may be unavoidable if a tool is handed over. This aspect of the interaction is where the third line of safety by design comes into action—sensor systems that sense humans via multiple modalities, not only by visual sensors but also by other sensors such as touch and distance. For example, artificial skins can be implemented because material technology and electronic circuitry have achieved a level of miniaturization and mechanical flexibility that allows us to design surfaces for the new generation of

Figure 14.2 **Video:** The COMPI arm: compliant trajectory tracking via model-based control allows safe human-robot interaction. https://youtu.be/mDODMNMC5zc.

robots that can feel [Lucarotti et al. 2013]. This is a great tool to enhance the safety of human-robot interaction, but more importantly, it allows the control programs and robotic architectures to integrate an increased range of modalities when generating a perception of the robot's environment and the interacting human or the internal state of the robot based upon pure sensor data [Kampmann and Kirchner 2015].

The next level of the hierarchy (level 2: low-level control; see Figure 14.1) can be related to the robot control regime that plays an important role in human-robot interaction. The systems within the new generation of robots are growing in complexity at a rate that is comparable to that of chip complexity growing according to Moore's Law [IntelCooperation 2005]. The low-level control (level 2 of our hierarchy) is realized in state-of-the-art robots by an approach to control actuators using torque and force rather than position [de Gea Fernández and Kirchner 2015] as the main control signal. The advantage of this control regime with respect to safety is obvious and refers to the fact that a position-controlled robot simply moves to a predefined position regardless of objects blocking its path, while a force-controlled robot also moves toward the desired position in 3D space but only as long as no external forces are encountered—forces like a collision with an object. The robots under force control regime would immediately (fractions of a second) reduce the torque on the appropriate joint resulting in an immediate stop. As soon as the external forces are gone the robot will continue on its path. In some more advanced cases the robot would in fact search for an alternative path or trajectory.

The last level in the hierarchy (level 3), also referred to as high-level control, is implemented based on sensors, internal of the robot, that are used to describe the state of the external environment of the robot [Lüth et al. 2015], e.g., cameras, lasers, or Time-of-Flight cameras. This information provides the basis for planning of manipulator arm trajectories and robot navigation paths and includes aspects of safety by avoiding to hit any obstacles (as well as humans) and describes a standard in robot control. However, these sensors can be combined with sensors that are external to the robot and that are very specific to the concrete application, e.g., a shopfloor production assembly line. There are many examples like: external overhead cameras [SafetyEye 2014] or laser range finders that are usually implemented to create safety boundaries around the robot. In traditional robot applications a violation of these borders just results in an alarm and a shutdown of the production line. In more advanced applications this information is used to adapt the behavior of the robot in case the boundaries mentioned above are violated by an object or a human [de Gea Fernández et al. 2017].

Safety boundaries are not static in modern robotics, but can be adaptive and vary with the context of the robot's task and application [Vogel et al. 2013], e.g. the robot would not go to a full stop but rather slow to a predefined speed in a human-robot cooperation scenario [de Gea Fernández et al. 2017]. Because the kind of adaptation of the robot is dependent on the context of the task, applying this adaptive type of safety level can also be seen as a context or application-specific safety level (see also Haddadin [2015]).

In summary, there are approaches to enhance safety that are inherent in the design of a robotic system. These approaches provide the robot with a good perception of the environment but do not necessarily require a concrete understanding of the human state or intention, or even recognition of the human apart from other objects in the environment. However, those internal states of robots that exist to enhance safety can also be the basis for the creation of more complex forms of awareness to support the interaction with the user. Hence, advanced interfaces that make use of multimodal data to enable explicit and implicit interaction with humans (see Section 14.3) must not only focus on establishing a representation of the user state but must also encompass a description of the status of the robot. There are simple mechanisms to present the state of the robot to a human operator, e.g., written or colored light feedback [de Gea Fernández et al. 2017], video data feedback from the point of view of the robot (i.e., its internal cameras), 3D reconstruction of the robot in its environment, or force feedback which conveys an impression of the robot's tactile perception during its interaction with the environment.

These modes of feedback are able to communicate information about the internal state of the robot. For certain tasks, such as teleoperation, continuous and clear feedback from the robot provides the user with the means to achieve easy control of the robot. In Section 14.4 we give some more examples for different applications, and in Section 14.3 we focus on the categorization of different approaches for utilizing human states for human-robot interaction and its relevance for the development of advanced embedded multimodal interfaces.

14.3 Definition and Relevance of Embedded Multimodal Interfaces

The basic goal of interfaces is to provide the robotic entity with a quantification of the state of the human user Schuller [2018] for an overview on multimodal user state recognition) and to provide the human with feedback regarding the state of a robotic system. The purpose of the interface is to generate bi-directional awareness and to enable bi-directional interaction and/or the support of the interacting human or robot. Whereas perceiving the environment via multiple modalities is very natural for more complex biological systems, the sensory capabilities of many technical systems are often limited to one modality. This modality is usually used in one way, i.e., to intentionally transfer commands or to actively perceive objects that are relevant for the system's action. In the past, this restriction of modality limited the possibilities for generating a representative level of (bi-directional) awareness. However, as pointed out in the previous section, this situation is now changing in the field of robotics. Technical systems can be equipped with different sensing modalities, which can be utilized for different purposes, with respect to both multimodal input and output [Kirchner et al. 2015]. This technical progress has been created by the increased availability and ease of usage of sensor technology, enabling robotic systems to receive a variety of data about their environment and users [Kampmann and Kirchner 2014]. This innovation requires the development of advanced multimodal-multisensor interfaces.

Let us give some examples with regard to the possible approaches that improve the interaction between human and robot using multimodal interaction. A camera can monitor the spatial position and body posture of the person and can capture any movement in space. But modern camera technology is also capable of monitoring information about facial expression and heart rate via a webcam [Monkaresi et al. 2014]. In a similar way, a conventional microphone can record sounds and utterances from the user, but is also capable of capturing those emotional responses that are inherently part of vocal expression [Bachorowksi and Owren 2010]. Both can thus be used to record psychophysiological measures of the human. Psychophysiological techniques grant technology access to signals from the autonomic nervous system via wearable sensors, which allows the robot to make inferences about the psychological state of the user. For example, changes in heart rate or galvanic skin response represent the level of psychological activation experienced by the individual; increased levels of frustration or anxiety or excitement are associated with higher psychological activation. The availability of wearable devices to measure electrocortical [Nijboer et al. 2015] and neurovascular activity [Piper et al. 2014] from the cortex would allow a robotic system to draw inferences about high-level cognitive states experienced by the user, such as intentionality, mental workload and skill acquisition [Bozinovski and Bozinovski 2015, Canning and Scheutz 2013, Kirchner et al. 2016a].

When developing a taxonomy for multimodal interfaces, it is convenient to classify techniques to monitor the status of the user into techniques that use ***overt measures*** and ***covert measures***. The former refers to methods that record and infer on the basis of what could be seen or heard by a hypothetical (human) observer. These overt methods are designed to record movements, changes in facial expression, and vocal utterances, or the same approach can be used to capture behavioral indices such as performance or task activity. Covert methods represent those measures from the user that are imperceptible to the hypothetical observer, such as psychophysiological and ***neurophysiological measures***. Certain data types, such as electrocortical activity, are invisible to the human eye, others may be perceived visually (e.g., pupil dilation) but cannot be accurately assessed in real-time by a human observer. Furthermore, a robotic system can use overt and covert measures to assess two broad categories of user state: (a) active states that represent intentionality, preparation for action and movement; and (b) passive states, such as emotions, mental workload, fatigue, and motivation, which arise spontaneously as a consequence of human-robot interaction. Table 14.1 provides two examples that capture the distinctions between overt/covert measures and active/passive state categories.

On the other hand, when classifying multimodal interfaces from a usage perspective, it becomes obvious that there is a substantial overlap between explicit or implicit interaction (see Figure 14.3). For example, when developing a speech interface for explicit control, both covert and overt measures can be used. The natural choice is to monitor the overt auditory output, but covert muscle activity, i.e., the human electromyogram (EMG), recorded by electrode arrays, can be used for such an interface [Wand et al. 2013]. The same is true for implicit control. Overt changes

Table 14.1 **Examples of active and passive states monitored using overt and covert measures.**

	Active	Passive
OVERT	user moves right hand	facial expression indicative of surprise
COVERT	increased activity in somatosensory cortex during preparation to move hand	elevation of heart rate and skin conductance level

in facial expression can be used to adapt an interface with respect to the emotional state of a user. Covert changes in skin conductance can be used to capture emotional activation. Thus, covert and overt measures are applicable for explicit and implicit control purposes. The same principle applies to the detection of those active states exhibited by the user. An active state can of course be used for explicit control. However, an active state, such as the preparation of a movement, can also be used for implicit control, i.e., to adapt a robotic system for likely upcoming movements to reduce interaction forces (see Sections 14.4 and 14.6). These examples emphasize that it is important to clearly state whether a taxonomy of interfaces is based on the techniques used to monitor the status of the user or the functionality of the multimodal interface. It is important that there is an awareness of both possibilities to avoid misunderstandings regarding the purpose of the interface.

If we consider how these techniques can be used to monitor the status of the user, a question arises concerning how various methodologies can be combined to create a dynamic and complex representation of the user state in order to enable implicit or explicit control. Brain-computer interfaces (BCI) [Wolpaw et al. 2002, Brunner et al. 2014] that often make use of the electroencephalogram (EEG) provide an interesting case for consideration. BCI technology is generally understood in its active form wherein neurophysiological correlates of voluntary control are used as an explicit control input to a robotic system (see ***active BCI*** in Figure 14.4). For example, BCI systems can use signals derived from motor imagery to intentionally direct the movements of a humanoid robot [Yongwook et al. 2012] or exoskeleton [Barsotti et al. 2015]. However, there are at least two other types of BCI that can be applied to robotic systems [Zander and Kothe 2011]. Reactive BCI describes a system where changes in brain activity in response to an external stimulus drive the output of the system. This type of BCI is activated by the neurophysiological response to a sensory event, e.g., an evoked-cortical potential (ERP) or steady-state visually evoked

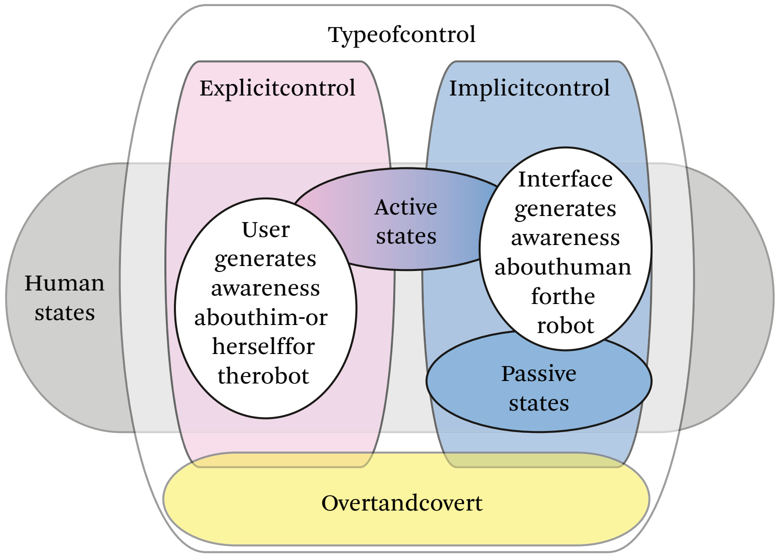

Figure 14.3 Type of control and the relation to human states and human measures.

potential (SSVEP), rather than the intention to act [Zander et al. 2014]. A reactive non-BCI interface which makes use of overt measures is, for example, a reactive eye-tracking interface (see Figure 14.4), which activates a specific action whenever the user is looking at a specific object, part of a robot, or area on a screen. The passive category of BCI (see ***passive BCI*** in Figure 14.4) describes a system where outputs are derived from changes in brain activity related to spontaneous changes in psychological states, such as: mental workload, frustration, anxiety, fatigue, etc. This 'passive' type of BCI is identical to the concept of ***biocybernetic control*** from ***Physiological computing*** [Fairclough 2009]. It is possible to create hybrid forms of BCI where active, reactive, and passive forms are used in conjunction to enhance the speed and fidelity of control [Muller-Putz et al. 2015]. For example, passive states of fatigue and frustration may affect the ability of the user to produce motor imagery, which translates into impaired control of an active BCI [Myrden and Chau 2015]. Cotrina et al. [2014] presented a ***hybrid BCI*** wherein passive state measures of electroencephalographic frontal asymmetry, a measure that is associated with emotion and motivational disposition, were used to refine the reactive response from an active BCI based upon SSVEP. Hence, active and passive state measures of the user are situated within a data space in order to improve fidelity of active con-

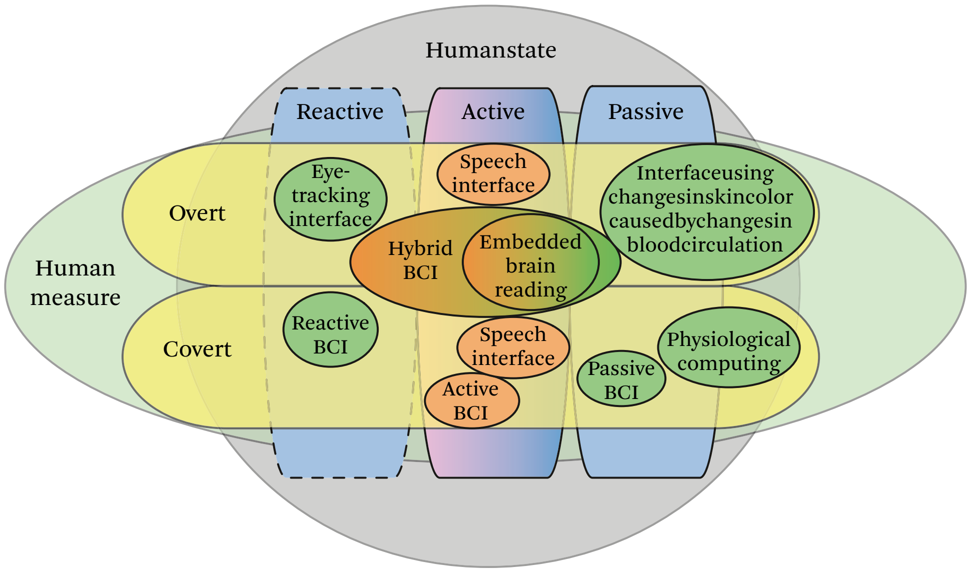

Figure 14.4 Examples of interfaces and their relation to human states and human measures.

trol over a robotic system. This data space can be further extended by adding overt measures. For example, Kim et al. [2014] designed a hybrid BCI that combined covert electroencephalographic data with overt eye movement data to navigate a quadcopter in 3D space. Hybrid BCIs can therefore be multimodal interfaces and can be used for explicit control of a technical system. They may also have integrated implicit control designed to optimize the multimodal interfaces and improve the degree of user control over the technical system. Also, if we consider the earlier example of the adaptation of an exoskeleton's control for teleoperation, this can be seen as a hybrid BCI: The user's overt arm movements while covert electroencephalographic data is used to non-intentionally and implicitly adapt and improve the explicit interaction between the human and the exoskeleton.

When considering the use of embedded multimodal interfaces, the meaning of the word "embedded" should be explained. The measurement of psychological concepts associated with the user state is generally enhanced by consideration of task context, e.g., the type of task being performed, task criticality, difficulty, duration, etc. The monitoring capability of multimodal interfaces encompasses task models and related variables in order to monitor the psychological state of the user within a specific task context. Thus, interfaces that make use of psychological concepts must be embedded into the task context. Therefore, the automated analysis

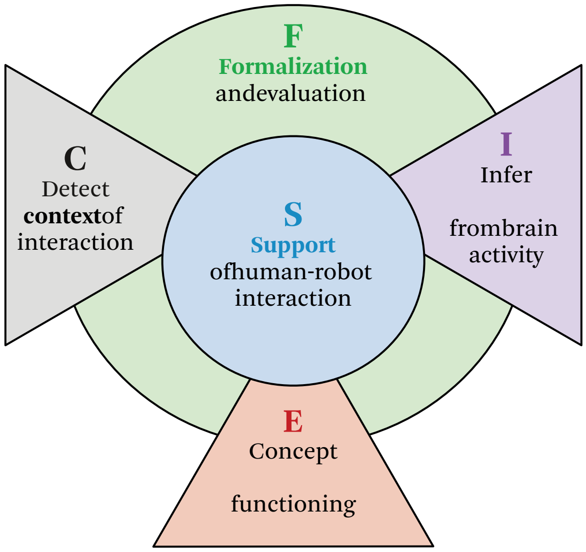

Figure 14.5 Concept for embedded brain reading for the support of human-machine interaction based on the context of interaction and inferred intentions. Figure courtesy of Kirchner et al. [2015].

of the interaction context is of tremendous relevance for embedded multimodal interfaces [Kirchner and Drechsler 2013]. On the other hand, interfaces that improve human-robot interaction must be embedded into the control of the robotic system in order to achieve and sustain safe operation [Kirchner and Drechsler 2013]. This does not necessarily mean that all data processing must be performed "on board" or that no external sources of information or processing power can be used (such as cloud-based solutions). It does imply that all processing that can only be done at the place of generation should be embedded into the system.

Figure 14.5 provides an example of embedded brain reading, a sub-type of embedded multimodal interfaces. Embedded brain reading includes the analysis of brain activity such as the EEG recorded from the surface of the head. To enable the interpretation of brain activity, embedding the analysis into the context of interaction and thus task state, human state, or spatial state is highly relevant. Without this processing module, it would nearly be impossible to interpret the user's brain activity to infer his or her intention. While context is a far-reaching term, it is used here very general. Context refers to the state of interaction in the broader sense (e.g., space, task, human state, environmental state, system state, etc.). The required interfaces are often not stand-alone systems, as a mouse or a keyboard are, but they are part of the robotic system that make use of input from internal and external sensor systems as well as sensors that are worn by the interacting human [Hung et al. 2015] designed to capture physiological measures from the body and neurophysiological measures from the brain. For example, today, exoskeletons are commonly equipped with gravity compensation [Lewis et al. 2003]. This is an algorithm which allows control of the exoskeleton so that the user does not feel the weight of the exoskeleton. This feature is most relevant for the control of distal body parts, i.e., the human arm, since the system's weight would otherwise be too high to allow extended usage, such as during teleoperation [Mallwitz et al. 2015] or for exoskeleton-assisted rehabilitation [Kirchner et al. 2013a, Kirchner et al. 2016c]. Forces can be redirected to, e.g., the hip or in case of a whole-body exoskeleton also to the ground on which the user is standing on. This control does not require knowledge about the intention of the user. However, an additional approach can be applied that enables the exoskeleton to also carry the weight of the arm. For example, gears can be locked to keep the arm in a certain position [Folgheraiter et al. 2012]. To release the locked position the system must know whether the user wants to move again. This knowledge can be gained from the analysis of the user's brain activity. However, to infer a movement intention is only relevant while the exoskeleton is keeping the arm fixed in a certain position. Thus, the exoskeleton "knows" from its own control state when to consider brain activity analysis to detect a change in the state of the human (see further explanation in Sections 14.4 and 14.6).

When using complex data such as covert brain activity to infer intentions it must be considered that the outcome of analysis can be incorrect. Therefore, embedded brain reading only considers approaches that are fault tolerant, i.e., will not lead to a malfunction of the whole embedded multimodal interface (see Folgheraiter et al. [2012], Kirchner et al. [2014]. Furthermore, methods are applied that allow formalization and evaluation of the implementations not only to estimate and measure quantitative and qualitative improvement [Folgheraiter et al. 2012] but to also verify correctness [Kirchner and Drechsler 2013].

In general, embedded brain reading describes the use of active and passive human states for explicit or implicit interaction, i.e., to non-intentionally adapt a robot or its interface with respect to an active or passive state to usually improve explicit control. The same approach can also be applied for explicit control purposes only. For example, in rehabilitation robotics, embedded brain reading can be applied to "drive" an exoskeleton to support or execute the intended movements of the patient. Spontaneously generated brain activity, which depends on the recording method (extracranial or intracranial), may not always be sufficient to control the exoskeleton in 3D. Other data such as muscle activity or eye movement data may need to be combined for an effective explicit control. Moreover, in the given example the exoskeleton's control can further be adapted to improve interaction. For example, the strength of support by the exoskeleton needed by the patient can be adapted based on an "assist-as-needed" approach [Kirchner et al. 2016b]. Such an adaptation can be achieved by directly measuring the force the patient can still exert or the strength of electromyographic signals of the supported limb or body part which can be recorded during interaction. The main goal is to combine multimodal data, such that the intended interaction or behavior can be supported best [Folgheraiter et al. 2012, Kirchner et al. 2013a, Kirchner et al. 2013b, Kirchner et al. 2014]. However, embedded brain reading can also be applied to infer the user's passive neurophysiological state, such as their current workload or task load (see also [Volume 2, Chapter 10), to adapt an interface for explicit robot control in such a way that the user is neither stressed nor bored [Kirchner et al. 2010, Kirchner et al. 2013b, Wöhrle and Kirchner 2014, Kirchner et al. 2016a] which would have negative impact on both the quality and quantity of interaction.

For embedded brain reading only brain activity is used which is spontaneously evoked (see Figure 14.6). Further, the approach is designed to interpret brain activity dynamically during interaction. Relevant data may also include spontaneous changes in brain activity in response to an external stimulus as used in reactive BCIs. To use intentionally evoked brain signals as it is often the case for many active BCIs, where, for example, the imagination of right- and left-hand movements can be used to spell a word [Blankertz et al. 2006], requires the attention of the user. Such attentional effort would require too many resources, which is one reason why classical BCIs are often considered to be inadequate for robot control in complex applications, such as, for example, in space applications. Figures 14.4 and 14.6 illustrate the interrelationship between different interfaces and embedded brain reading.

In summary, the basic goal of embedded multimodal interfaces is to provide the robotic entity with a quantification of the user state that enables easy explicit control of the robot and allows for implicit control of the robot or its interface. Such interfaces also enable the system to make use of the cognitive capabilities of the human or to enable the robot to learn from the human [Kirchner et al. 2015] (see Section 14.4). The availability of overt and covert measures to capture active and passive states (Table 14.1) allows the multimodal system to construct a dynamic representation of the user that is both sophisticated and scientifically valid as discussed in this chapter.

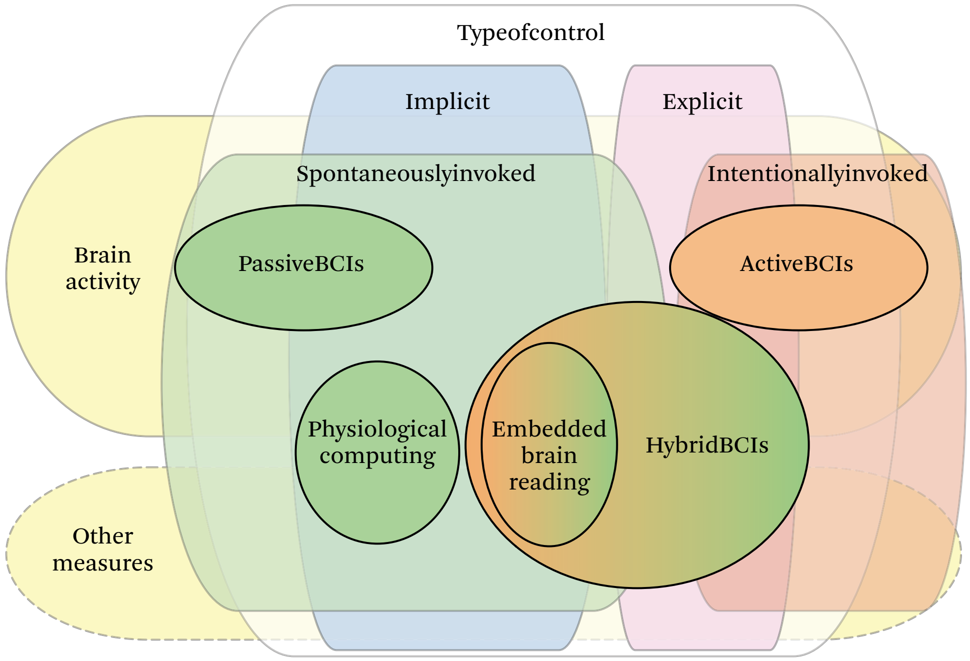

Figure 14.6 Usage of spontaneously and intentionally evoked brain activity.

14.4 Embedded Multimodal Interfaces in Robotic Applications

This section will explain the advantages of embedded multimodal interfaces and how they would function in the context of robotic applications. First, we give examples of multimodal-multisensor interfaces from two perspectives, i.e., robotics and physiological computing, and explain where and what the purposes of such interfaces usually are in both fields. Later, we give examples that focus on specific applications in robotics. We explain how both approaches, driven by the robotic control view and driven from the perspective of human state analysis, can be combined in a temporally cascaded fashion to: (1) make use of overt and covert measures to ease explicit control of robotics systems or to enable implicit control that adapts an interface or robot to improve human-robot cooperation within the same application. Further, we provide examples of how embedded multimodal interfaces will (2) improve bi-lateral awareness using covert human measures and multimodal data of the robot. Moreover, we will (3) explain how the usage of multimodal overt and covert measures allows us to increase the level of awareness.

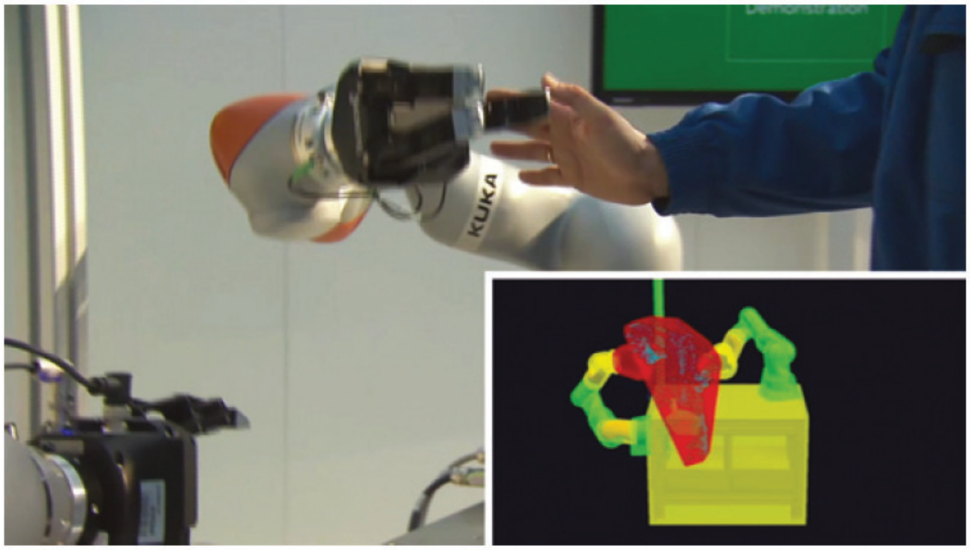

Figure 14.7 **Video:** iMRK: an embedded multimodal interface for human-robot interaction. https://www.youtube.com/watch?v=VoU3NbTyFtU.

One highly important reason for applying embedded multimodal interfaces is to enhance safety and reliability. For example, a human must be detected when he or she has penetrated the work space of an industrial robot. Different sensors on the robot or in the environment, such as lasers, 3D depth-image camera, or motion sensors (that are attached to the human or human's clothes) can be combined to detect the location of the person and predict their path [de Gea Fernández et al. 2017]. Furthermore, the weaknesses of one sensor, i.e., a restricted range, can be compensated by another to improve the recognition of a human entering the work space. A combination of sensors enables accurate predictions to interprete the situation and the user's intention. The same sensors can also be used for explicit interaction, e.g., to intentionally command the robot to stop moving or to hand over a workpiece (see the video referred to in Figure 14.7). In most cases, approaches in robotic control are designed to enhance safety and reliability, and make use of overt measures to analyze the state of the user.

There are advantages to measure a psychological concept using two or more different indices from the perspective of physiological computing. Certain psychological states may be described as many-to-one [Hettinger et al. 2003], i.e., a number of measures are required in order to represent a single psychological concept. This approach also represents a form of convergent validity wherein multiple measures are collected simultaneously in order to derive a composite score based on the degree of correlation or coherence between different measures. Therefore, the capacity of multimodal interfaces to encompass different measures allows the robotic system to monitor the user in a way that is scientifically valid. For example, Bekele and Sarker [Cacioppo et al. 2000] constructed an adaptive mode of human-robot interaction where task difficulty was dynamically adjusted in order to keep the user engaged with the task. This system combined physiological measures, i.e., covert measures, from the cardiovascular system, electromyography, and skin conductance to capture the level of task engagement exhibited by the user. The same group applied a similar approach to measuring emotional responses to create robotic interventions for children on the autistic spectrum [Liu et al. 2008] (see also Volume 3, Chapter 13.).

Making use of both covert and overt measures goes beyond classical control approaches in robotics. With the help of the application shown in the video referred to in Figure 14.8 we explain how such an approach, based on overt and covert measures, enables improvement of human-robot interaction, i.e., explicit control by non-intentionally adapting the robotic system with respect to the active human state and by applying implicit control within the same application. In this application, a user is teleoperating a robotic system by means of an active exoskeleton [Folgheraiter et al. 2012]. An active exoskeleton is a robotic system that is worn by the human and thus is in direct physical contact with the human. It is both an interface as well as a complex robotic actuator. To enhance transparency for the wearer the interaction between user and exoskeleton in this example is supported by embedded brain reading [Kirchner and Drechsler 2013, Kirchner et al. 2013b], an embedded multimodal interface that is making use of covert electroencephalographic data recorded from the scalp. Based on the analysis of the user's EEG and his or her behavior, transitions between (tele-)operation modes (see Figure 14.9) are supported (see Section 14.6 for details).

But why is covert brain signal data *not* directly used for explicit control, i.e., explicit change of operation mode? Even when using advanced signal processing and machine learning methods to infer movement intention from brain signals, state detection is inaccurate due to the complexity of the EEG signal and due to the fact that similar brain signals are generated when a human is only imaging a movement or does indeed prepare the movement to be executed. Therefore, the outcome of EEG analysis is in the given example not used to intentionally control the change between the modes. It is instead used to enhance sensitivity of the sensors that detect the movement onset, with the result that interaction forces are reduced and the operator can clearly feel the enhancement in transparency of the exoskeleton (see Section 14.6 for more details on approach and evaluation).

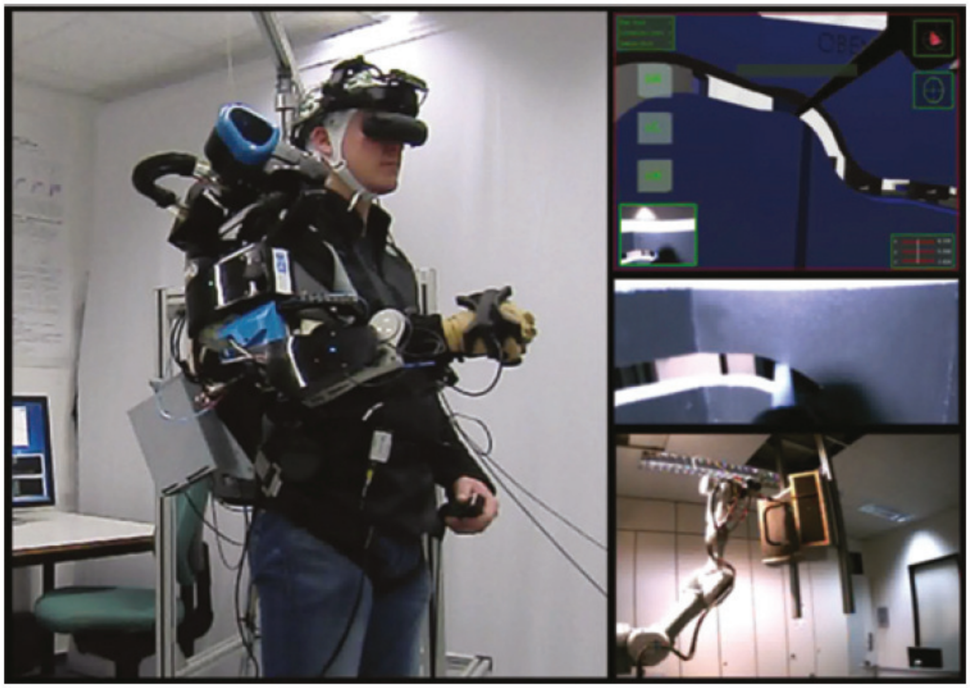

Figure 14.8 **Video:** VI-Bot: an embedded multimodal interface for robot control via teleoperation with an exoskeleton and a virtual environment. https://www.youtube.com/watch?v=3RhcgvRz_O8.

Moreover, the implemented embedded multimodal interface does not only enhance transparency but is highly reliable and safe for use. The user is always able to unlock the exoskeleton by pressing against the sensors even in a case where movement planning was missed. Furthermore, adaptation of the exoskeleton is changed every 50 ms which made it unlikely that a false positive will lead to an unwanted lock-out. During all our online tests with real online movement prediction, the user never experienced a false lock-out using this system. By combining two modalities, i.e., covert brain data and overt movement data, in a strict temporal order the interface represents an example of a temporally cascaded multimodal interfaces, which will be explained later.

Besides enhancing safety, reliability, and bilateral interaction, an embedded multimodal interfaces enhance bilateral awareness. Different modalities can be used to provide the user with a greater insight into the state of a robotic system, i.e., to improve awareness of the robot's state. For example, in the scenario which was developed for a teleoperation application (shown in the video referred to in

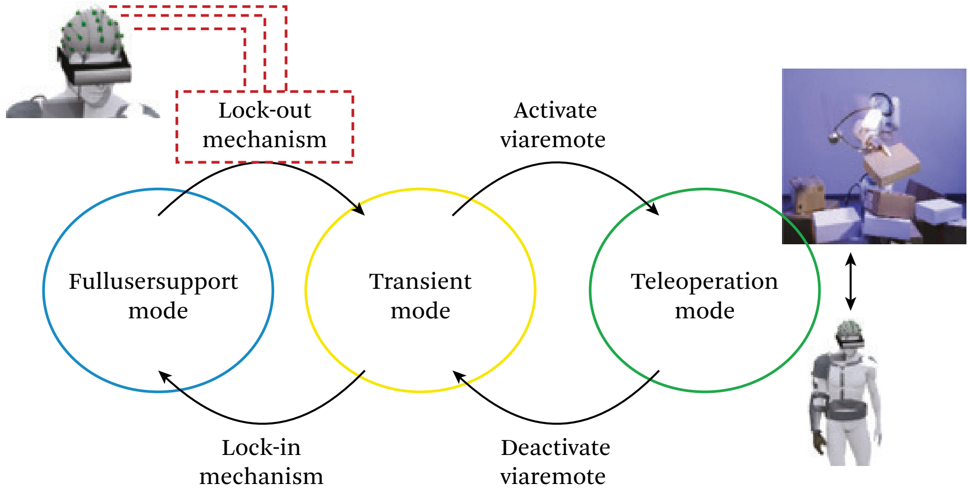

Figure 14.9 Exoskeleton states: teleoperation mode, transient mode, full user support mode. Lock-out mechanism is activated by overt measures of an active human state (arm movements detected by the exoskeleton). The activation of this mechanism is adapted by covert measures of an inferred active state change (from no movement to movement) by analysis of the human's EEG, i.e., detection of activity that correlates with movement preparation. Figure courtesy of [Folgheraiter et al. 2012].

Figure 14.8) the user is not only supported by the exoskeleton for explicit control. He or she is further virtually immersed into the situation of the robot by using 3D simulation and a head-mounted display while being able to make use of video material. Moreover, the user receives tactile force feedback from the robotic system via the exoskeleton. The combination of virtual immersion, mapping between the human and the robot's movement, and force feedback allows the user to become immersed into the situation and virtually feel what the robot feels. By means of the embedded multimodal interface, the user becomes strongly aware of the robot's states and changes in state, which eases explicit control and reduces interaction errors, such as failures in path following (see Figure 14.10).

Furthermore, the user is monitored whether he or she is indeed aware of the state of the exoskeleton and the display of important messages that are visually presented to the user. This facility is achieved by adapting the embedded multimodal interfaces based on the predictions that are made about the success of the user in recognizing these relevant messages. The adaptation of the display of messages is again based on online brain-signal analysis. Signals in the EEG are detected

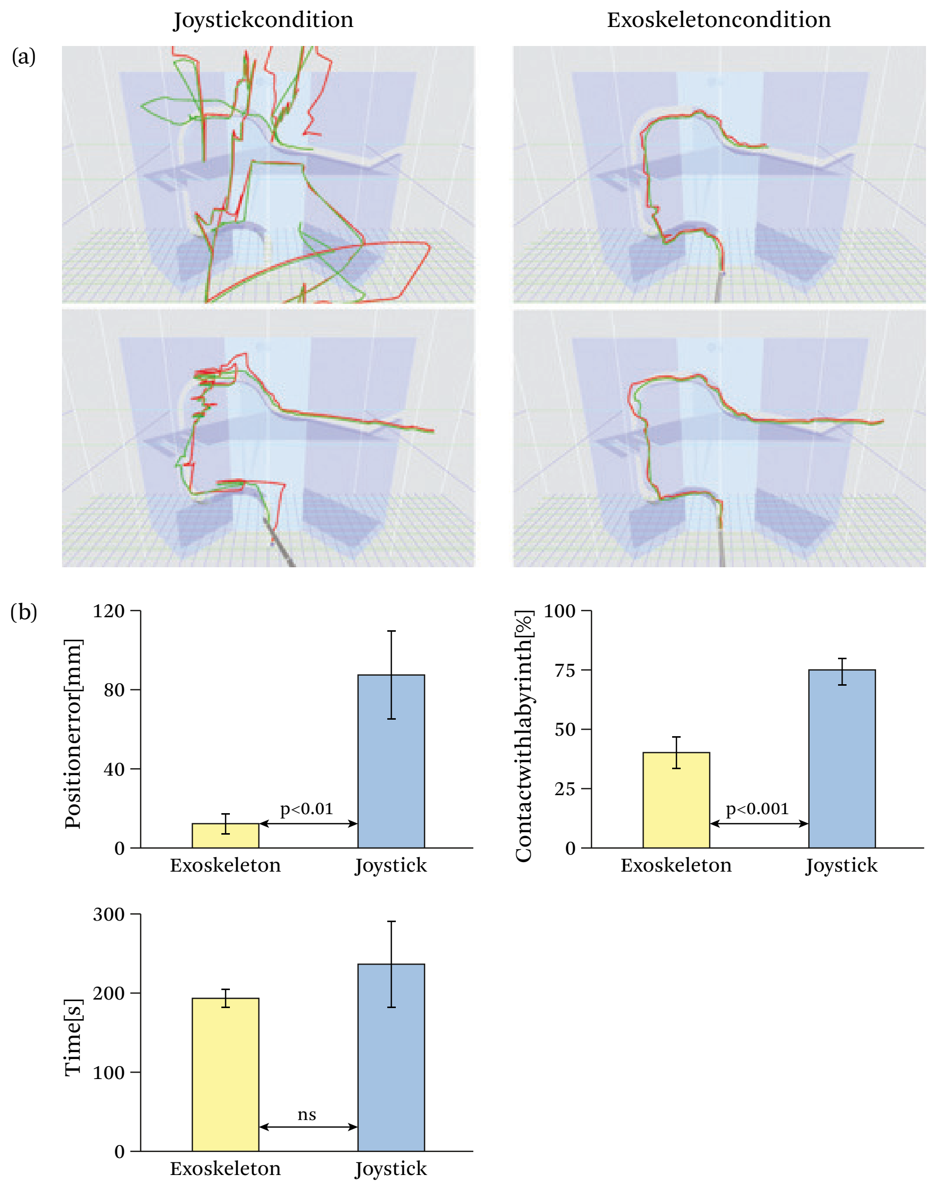

Figure 14.10 Comparison of accuracy in path following between joystick and exoskeleton control in the VI-Bot scenario. (A) Example of two subjects steering the robotic arm through a 3D maze. Green line: path corrected in the virtual environment. Red line: theoretical path without correction. (B) Behavioral analysis of nine subjects. Accumulated position errors (left), percentage where path contacted the wall of the maze (middle), and measured time for a complete sweep through the labyrinth (right). Figure courtesy of Straube et al. [2011].

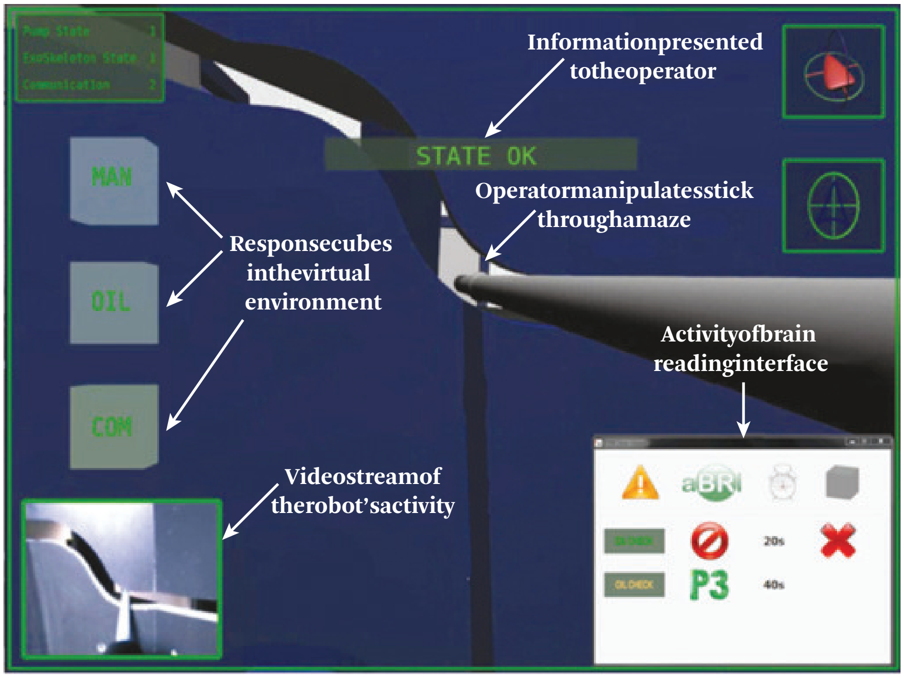

Figure 14.11 **Video:** VI-Bot - virtual immersion for holistic feedback control of semi-autonomous robots. Shown is the implicit control approach implemented to adapt the embedded multimodal interface to enhance the awareness of the user for relevant information on the robotic system based on EEG signals. https://youtu.be/8WEVZz6bpJU.

that allow the system to infer whether the user recognized the presented messages. If those signals are detected, information is not repeated for a longer time, since it is expected that the user will respond to the message. In case the relevant brain activity cannot be detected, the message is repeated instantly and at the same time highlighted to make the user more aware of the relevant information (see video referred to in Figure 14.11). The chosen approach is an implicit control of the interface [Kirchner and Drechsler 2013, Wöhrle and Kirchner 2014].

Besides enhancing ease of interaction and reducing errors, many robotic applications, especially those requiring the simultaneous control of a team of robots, have to handle the limited cognitive resources of a human that lead to cognitive overload. Embedded multimodal interfaces represent a good solution to handle this scenario by adapting the interaction with respect to the cognitive load of the human, as demonstrated in the field of physiological computing (see Zhou et al.

Figure 14.12 Immersive virtual 3D multi-robot control using a CAVE. Figure courtesy of [Kirchner and Drechsler 2013].

[2018] for indicators of cognitive load). To achieve this, the awareness by the interface of the human's state (vs. the awareness by the human of the robot's state as in the example given before) is enhanced by means of the embedded multimodal interfaces, as will be explained next.

In the video referred to in Figure 14.13 an application of multi-robot control is displayed. The control interface is a very complex one. It is based on a virtual environment using the software Machina Arte Robotum Simulans (MARS) [Rommerman et al. 2009, MARS 2015], which can be run as a 3D environment in, e.g., a CAVE (see Figure 14.12), as a 2D environment on a standard personal computer or on a multi-screen system (see the video referred to in Figure 14.13). Information about the robot is mainly presented in visual 2D or 3D mode. However, when using an exoskeleton as input device, force feedback can be applied to the operator. Thus, the exoskeleton is not only implemented to control the display but also to interface with a robotic arm, similar to the VI-Bot interface (see the video referred to in Figure 14.8). Moreover, the interface is implemented so the user can interact by using different alternative modes or different modes in combination. For example, eye tracking is implemented to enable the operator to select interaction icons instead of using a 3D mouse, a wand, or an exoskeleton to allow the operator to use the referenced input device and modality; see Figure 14.14. In the future, gestures will additionally be implemented to navigate through the scenario, to change the virtual camera position, or to select interaction modes in order to send control commands.

Figure 14.13 **Video:** Virtual multi-robot control in 2D using a multi-PC system supported by embedded brain reading. https://www.youtube.com/watch?v=zeFp_JBSBxA

While explicit control is optimized to ease the interaction in the VI-Bot scenario, the cognitive resources of the operator may be insufficient when the task load is too high. Therefore, implicit control of the interface is enabled by an embedded brain reading approach that adapts the interface with respect to the individual task-load of the user [Kirchner et al. 2016a]. This aspect is highly relevant since ease of control is one component that improves interaction quality; the skills and mental workload of the operator are other aspects that have to be considered in the context of limited resource theories (see Volume 1 Chapter 1). The embedded multimodal interfaces can be adjusted online with respect to the current task-load and general training status of the operator, resulting in different task frequencies, i.e., interstimulus intervals (ISI) for individual operators.

In Figure 14.15 it is shown that the adaptation in run 5 and 6 supported by embedded brain reading results in individual differences in total runtime, i.e., the time needed to fulfill 30 given tasks. In runs 1 and 2 and runs 3 and 4 the ISI was fixed to a long (run 1 and 2: ISI-25 condition) or shorter time (run 3 or 4: ISI-15 condition). Under both conditions the total runtime was very similar between participants. While the performance in runs 1–4 without adaptation seems very similar over the group, some participants reported that they were bored under ISI-25 or even IS-15 condition. Subjects 2 and 6 could achieve much shorter runtimes when supported

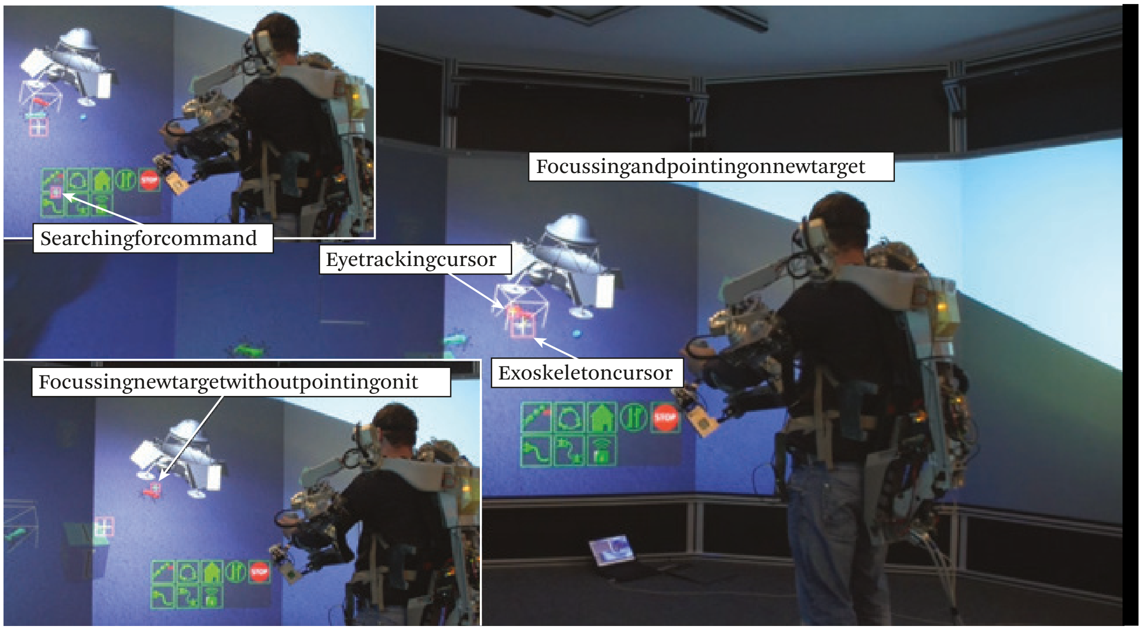

Figure 14.14 Multimodal explicit control: operator can chose to interact with interaction icons of the interface and the robotic systems via the exoskeleton or the eye tracker or a combination of both.

by embedded brain reading. To reduce runtime from the beginning would not have been a good solution, since some individuals would have become stressed, like participant 5. Thus, implicit control is required to adapt an interface to the overall needs (depending on her or his skill level) and the current needs (depending on task load) of the human operator while avoiding excessive cognitive load or boredom.

The usage of multimodal and multisensor data in embedded multimodal interfaces have many positive effects on human-robot interaction. Whereas we have already discussed that a temporal cascaded approach allows a smoother interaction, the combination of multimodal data can further increase the dimension of insight possible into the intention of the user more than a single modality could. Some examples should be given next that accord with the Gestalt theory of perception (see Volume 1, Chpter 1), that the whole can be more than the sum of the individual parts. An example is the combination of different covert physiological measures and overt measures from the human, like eye movement data, EEG, and EMG data for the control of a robotic rehabilitation system [Kirchner et al. 2013a]

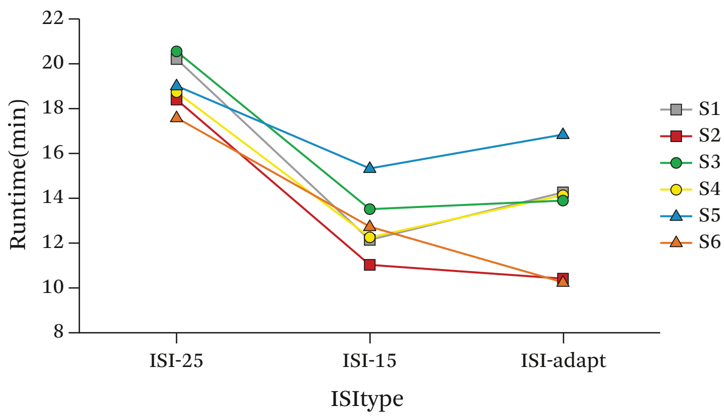

Figure 14.15 The means of runtime for a fixed ISI of 25 S (ISI-25), a fixed ISI of 15 S (ISI-15), and ISIs adapted by embedded brain reading (ISI-adapt) are depicted. Figure courtesy of Kirchner et al. [2016a].

(for more information on multimodal behavioral signal processing systems see Volume 2, Chapter 10 and Volume 2, Chapter 12).

Using each individual signal not only reduces the reliability but would further generate less ”meaning”, e.g., to look at an object does not tell us whether interaction is wanted. EMG activity alone does not tell us whether interaction is desired to begin, e.g., to move the arm to the object, or whether the subject was just bumped by someone and used the arm to balance or has experienced a spasm. In addition, EEG cannot clearly tell us whether the subject is planning to actually execute the interaction or is only imagining the interaction, since internal visualization and planning of motor execution result in very similar brain patterns. If these individual signals are combined, the outcome becomes more reliable, especially if the expected temporal order of signals is consistent, i.e., a temporal cascaded approach is followed. Moreover, only if at the same time or afterward brain activity related to movement planning is detected in the EEG, the user is probably not only looking at the object but at least is thinking about starting an interaction. Finally, by adding the EMG signal, it becomes quite clear that at the moment when the subject looks at the object, plans a movement and tries to activate the muscles, that he or she indeed wants to start a movement. Therefore, the combination of all three signals tells us more than the individual signals alone and results in very reliable interpretations of the human’s intention.

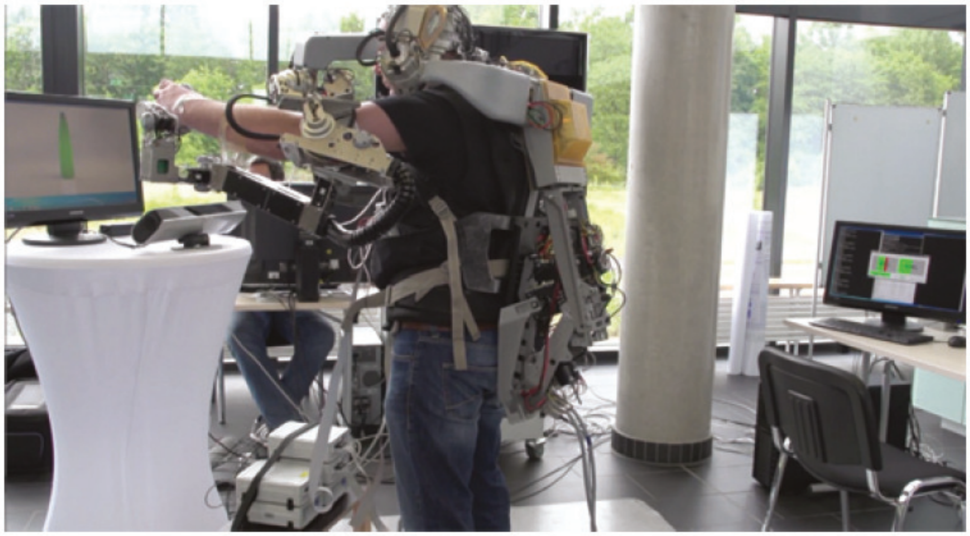

Figure 14.16 **Video:** Demonstration of biosignal usage to control a robotic system: the exoskeleton is moving the right or left arm in case that the bottle is in focus and biosignal analysis detects the intention of the operator to move from EEG and EMG signals. https://www.youtube.com/watch?v=BRpbZFOXdRk

Deciphering the intention of the human user is highly relevant for rehabilitation robotics. In the video referred to in Figure 14.16 it is shown how the combination of EEG, EMG, and eye tracking data can be used to control an assistive device, i.e., an exoskeleton. While the most straightforward approach is to combine all signals in an "and" fashion, i.e., to only drive the exoskeleton in case that all three signals are detected, EEG and EMG data can, to some degree, substitute for each other (see decision-level fusion in Volume 2 Chapter 12, Section 12.8). Therefore an "or" combination is also possible. Depending on the application or state of rehabilitation it might be more useful to use the "and" or the "or" combination, i.e., different temporal cascaded approaches.

For example, during early rehabilitation, the patient should strongly be motivated to train her or his impaired limb, however muscle and brain signals might be weak. It might be better to use the "or" combination here, i.e., to drive the exoskeleton in case there is either the expected EEG or EMG signal detected. Later in rehabilitation the patient might be more annoyed by unwanted movements or the patient should be more strongly involved in the interaction. Here, the "and" combination might be more suitable. The kind of combination of different signals has therefore an impact on the behavior of the interface and on the reliability of intention recognition (see Figure 14.17 Kirchner et al. [2014]). Hence, the temporally cascaded approach that is followed does strongly determine the quality of awareness of the human's state that a robotic system can derive.

In summary, embedded multimodal interfaces enable bilateral interaction. They make use of overt and covert measures to detect active and passive human states to make the robotic system aware of the state of the human. Furthermore, they use multimodal- multisensor data to make the human aware of the robot's state. This bilateral awareness strongly improves interaction, reducing cognitive load, interaction forces, and interaction errors. Based on a carefully chosen temporally cascaded approach, they allow that multimodal data can be combined, such that the whole can be more than the sum of the individual parts (see Panagakis et al. [2018]). The later requires a deep integration or embedding of the multimodal interface into the system control.

14.5 Future Trends: Self-Adapting Embedded Multimodal Interfaces and Societal Implications

In this chapter, we first introduce three main trends in embedded multimodal interfaces which will enable future applications in human-robot interaction.

1. The functionality of embedded multimodal interfaces must be highly adaptable in an online fashion. They should adapt to changes in the context of interaction, changes in the state of the robot or the human, or changes in availability of certain sensors or sensor data.
2. To enable this level of adaptability future interfaces must be deeply embedded into the control of a robot in order to develop specific hardware and software solutions.
3. There is a trend to embed the interface into applications and systems, in order to make them smaller and more energy-efficient and to allow both long-term usage and a high degree of flexibility.

While many developments are new research territory for human-robot interaction, we will show with examples that research from other areas can be used and solutions can be adapted. Using these approaches, a more direct and efficient interaction between human and robot will be enabled. However, these changes and future developments will have societal implications.

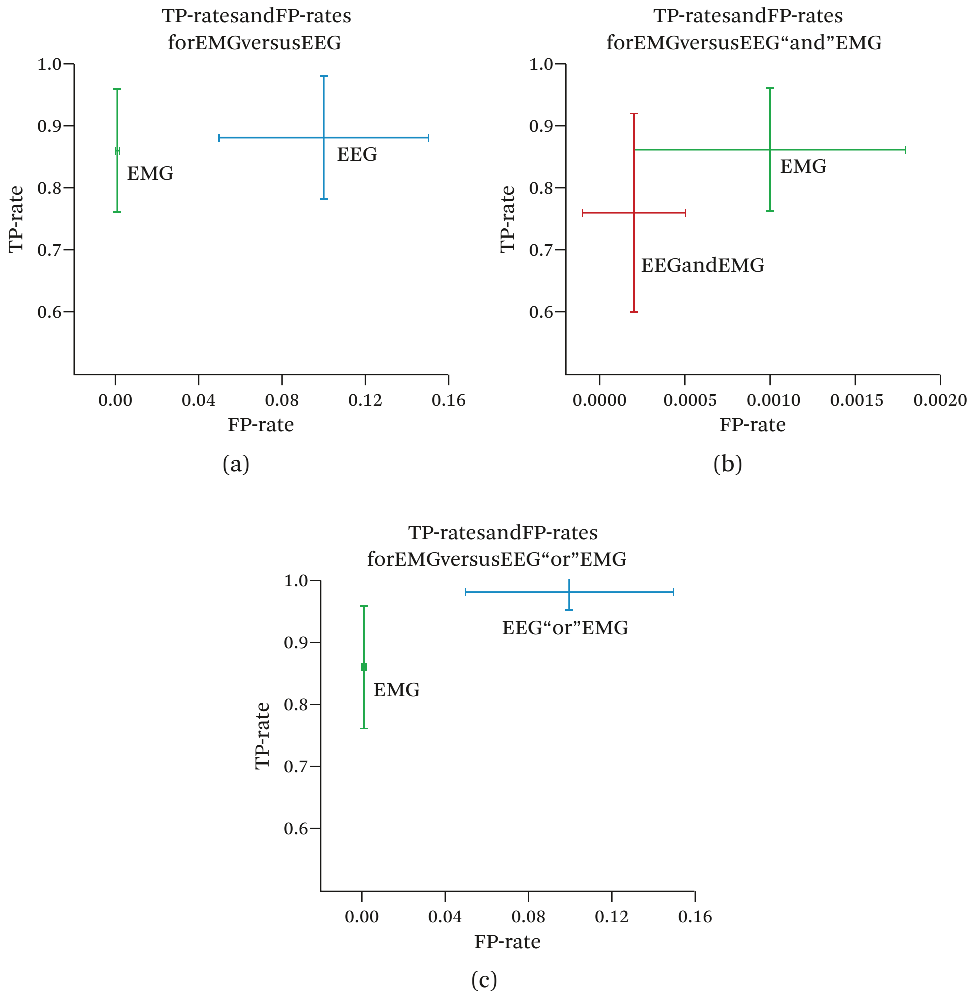

Figure 14.17 Effect of different combinations of EEG and EMG data on true and false positive rates during classification. (A) Prediction results in TP- and FP-rate for EEG (blue) and EMG (green) analysis. (B) Prediction results in TP- and FP-rate for EEG “AND” EMG (red) and EMG (green) analysis. (C) Prediction results in TP- and FP-rate for EEG “OR” EMG (black) and EMG (green) analysis. Figure adapted from Kirchner et al. [2014].

14.5.1 Inherent Self-Adaptation and Deep Integration

Multimodal-multisensor interfaces that are embedded into application procedures, robotic systems, application environments, into the users’ clothes, or that are implemented as wearable devices will be perceived as normal in the future. Considering continuously changing interaction environments and interaction tasks, it is obvious that multimodal-multisensor interfaces must be able to adapt to changes by themselves in the future. Today's robotic systems and their interfaces are already of such complexity that any hard-coded change will require a high number of specialists to perform the required software and hardware adaptations. For the common user, it will become impossible to handle this kind of technical detail. However, the need for adaptation not only for robotic systems but also for their interfaces will increase.

14.5.1.1 Closed-Loop Design for Self-Adaptation to the State of the Human

An embedded multimodal interface allows the robotic system to adapt to the movement intentions and states of human operators. This system works via ***closed-loop control*** logic whereby multimodal signals from the user are monitored, analyzed, classified and converted into appropriate outputs or adaptive responses from the robot. This closed-loop design requires a set of rules whereby a target state triggers an adaptive response. However, this may not be an exclusive relationship and a range of potential responses are available once a specific state has been identified. For example, if the robotic system detects high mental workload during a process control task, it could slow the pace of operations (see Section 14.4) or suggest a break. The rules that translate multimodal detection into an adaptive response at the interface draw from a repertoire of possibilities, all of which are equally likely to create the desired effect on user behavior. This scenario poses the question: How does the multimodal interface select the most appropriate response from an existing repertoire of possible responses?

A closed-loop system for either positive or negative control is often characterized with reference to a single discrete cycle of monitoring and adaptation. In this case, a single cycle may describe how the detection of high mental workload is translated into a reduction of task pacing in order to aid the human operator. This is a first-order process of adaptation wherein the loop detects and responds to a target state in the short term. Once this adaptation is activated, it is possible for the multimodal system to detect whether its own adaptive response had the desired effect on the human operator. If slowing the pace of the task has reduced the mental workload of the operator (assessed via neurophysiological monitoring), this adaptive response is deemed to be successful. The other possibility is that the adaptive response failed to reduce the mental workload of the human operator, in which case the multimodal interface must enter a second cycle of monitoring and adaptation in order to select another response, e.g., suggest a break. This latter process is called second-order adaptation or ***reflexive adaptation*** [Serbedzija and Fairclough 2012] because an adaptive response is based upon the closed-loop monitoring of the consequences of its own intervention on the state of the user.

A second-order process of reflexive adaptation can facilitate machine learning over a sustained period of interaction with an individual operator. In order for the multimodal system to adjust to the individual, it must accumulate a database that identifies those adaptive responses found to be effective for a particular user. Second-order adaptation describes a generative process where the repertoire of adaptive responses is 'pruned' or customized on the basis of repeated interaction with a specific user. This evolving cycle has been described as a process of mutual adaptation with three main phases [Fairclough 2015]. The initial encounter between the multimodal system and user is characterized by improvisation. The system responds to the user in a generic fashion using default adaptations with no prior knowledge of individual preferences. Adaptation may be perceived by the user to be less than optimal during this early phase. As the user spends more time interacting with the system, second-order adaptation improves the acceptability of responses from the perspective of the user. This second phase of reciprocal coupling is characterized by tailoring the adaptive repertoire of the robotic system to the individual. If we look further ahead in time, in terms of months and years, it is reasonable to expect that any stable model of preferences acquired during reciprocal coupling will have limited longevity—as the user acquires greater skill or habituates to popular adaptive responses or experiences cognitive changes due to aging. The third phase of co-evolution describes a process of updating a stable model of user preferences over a longer period of time.

14.5.1.2 Self-Adaptation to the Context of Interaction

Besides adapting an interface to the user, adapting robotic systems and their interfaces to the context of interaction will become an important direction for the future of robotics. To allow this development, the context of interaction must be recognizable from the perspective of a robotic entity. In an industrial context where operational sequences are prescribed and must accurately be fulfilled, it seems that the recognition of the context of interaction should be relatively straightforward and may even be predefined. However, flexible support must consider deviations from the procedure and personal preferences of the human user. For example, personal preferences can be handled by systems that learn during interaction from physiological data (such as error-related potentials in the human EEG) what the user means for example by certain individually chosen gestures, i.e., it learns the mapping between gesture and action [Kim et al. 2017]. In non-predefined interaction scenarios, the recognition of context becomes even more relevant. To recognize the context of interaction, different data sources can be considered if available data about procedures and preferences of the human can be used. For example, knowledge about preferences of patients might even be more relevant than the analysis of physiological measures to optimally support them by a robotic system [Novak et al. 2013]. However, physiological measures as explained in the example in the video referred to in Figure 14.16 can also be used to detect the context of interaction. Moreover, the supporting system itself is also able to detect the context of interaction, as explained in the example presented in the video referred to in Figure 14.8. On the other hand, interaction usually requires physical activity of the human. Thus, complex movement data is a highly relevant source of information to deduce the context of interaction. While today human movement behavior is often analyzed to develop approaches that enable robotic systems to learn to imitate human behavior [Metzen et al. 2013, Mülling et al. 2013, Pastor et al. 2009], movement data can also be used to recognize the context of interaction [Senger and Kirchner 2016], especially body posture can be very informative about the behavioral context of interaction. In a simple case, the direction of movement with respect to a robotic system can tell the robot whether a human wants to interact or not [de Gea Fernández et al. 2017].

14.5.1.3 Software Frameworks and Hardware Solutions for Deep System Integration

To use the power of embedded approaches, future multimodal interfaces must potentially make use of any available data, like sensor data of the robotic system or from environmental supervision, data about procedures, and even (physiological) data about the interacting human. This brings along some challenges in data storage, processing, selection, and handling. Thus, new software and hardware solutions, such as the open source software framework pySPACE [Krell et al. 2013] or the specialized software framework reSPACE [Wöhrle and Kirchner 2015] which can run on embedded systems, must be developed that allow for flexible usage of multimodal data. Optimization during runtime will require reconfiguration of hardware at runtime to allow optimal, i.e., time, resource, and energy-efficient interpretation of potentially changing data sets. These approaches have to consider which data is most relevant to interpret the current situation, to infer the intention of a human in a certain situation, or to estimate the best possible support of a human by a robotic system at hand.

Supervised and unsupervised learning methods must be applied on top of sophisticated signal processing that reduces data and filters relevant information. All this must be achieved on embedded hardware that allows fast but also resource-saving analysis. Different approaches must be combined. For example, the usage of processing units based on Field Programmable Gate Array (FPGA) was shown to support powerful and fast processing with low energy consumption on small-sized, embeddable devices [Wöhrle et al. 2014]. However, these approaches are still limited. Thus, a combination of small embedded devices and powerful central processing units must be considered and promoted. While the embedded system will perform data analysis within the interface, the robot or the wearable device, the central processing unit or units will be able to perform more complex calculations required to optimize data selection or combination, processing, and, hence, adaptation of the current processing flows. Thus, hybrid hardware/software solutions will enable self-adapting embedded multimodal interfaces.

In summary, future robotic systems must not only behave autonomously but must deeply understand humans to better support them and to allow flexible interaction and cooperation. While this seems to require extra effort at first glance, on closer examination two things become obvious. (1) Approaches that are currently developed to enable robots to behave better and to perform complex task, such as approaches that enable imitation learning, i.e., learning from human demonstrations [Schaal 1997, Argall et al. 2009], are also relevant to improve interaction. (2) The amount of multimodal data from multiple sources will increase. Thus, for both approaches new software and hardware solutions must be developed in order to profit from each other.

In order for next-generation robotic systems and their embedded multimodal interfaces to accommodate the requirements dictated by the human-robot cooperation it will be necessary to develop standardized robot control frameworks and architectures that not only allow easy integration of internal parts of the robot such as motors, cameras etc., but also adapt the systems control to human interaction. The frameworks must be designed such that they are flexible toward changes in multimodal input and multimodal output during runtime, e.g., changes in sensor input or changes in interface input. Examples are the Robot Construction Kit (ROCK) and DROCK [DRock 2015]) software frameworks. These model-based approaches allow the designer of a robotic system to define the system from a library of well-defined and mathematically modeled components. This approach works from the hardware as well as from the software perspective. Therefore, software-based concepts like adaptation and learning can be integrated with standard planning and control approaches via machine learning to enable usage by non-specialists.

It is obvious that both main research directions, i.e., autonomous artificial intelligent robotic systems and human-machine interaction become inseparable to develop future robotic systems that are able to optimally support humans and to interact with them intuitively. Robotics is a very good example that shows how different fields of research must not only work together to their mutual advantage.

14.5.2 Current and Arising Societal Implications

As stated before an embedded multimodal interface is capable of enhancing human-robot cooperation by: (a) increasing machine awareness of the user via monitoring and (b) personalizing the behavior of the robot to the preferences of the individual via the closed-loop process of reflexive adaptation. This emergent approach is designed to evolve the technical sophistication of how people interact with robots. However, these technological advancements are associated with a number of societal implications.

It is important to understand that closed-loop systems are driven by goal-directed logic. The closed-loop within the multimodal interface is programmed with a specific directive, e.g., to prevent mental overload, to improve performance efficiency, to preserve the safety of the operator. Its repertoire of adaptive responses are simply the means by which the system achieves its specified goal. Unlike the inert and passive technology of today, this symmetrical interaction is characterized by a degree of agency on the part of the machine and the requirement for a human operator to cede a degree of control to the system. Given this, it is important to define the agenda of the machine to be effective, reliable, and not to lead to unforeseen circumstances [Kirchner and Drechsler 2013].

14.5.2.1 Establishing Trust

The challenge for multimodal interfaces is how to make the robotic system an effective "team-player" from the perspective of a human user [Klein et al. 2004]. In order for a robotic system to work with human users, it is important for embedded multimodal interfaces to establish a degree of 'trust' with their users. Miller Miller [2005] argued that technology could earn the trust of the user by transparency, i.e., the laws of cause and effect encapsulated within the closed-loop are clearly understood by the user. This transparency can be enhanced by clear feedback to the user during interaction and predictable behavior on the part of the robotic system. The development of trust between robot and user requires time and can only be achieved through repeated interaction over a long period.

Multimodal interfaces utilize increased data processing capacity to: (1) monitor the behavior/physiology of the user; (2) make inferences about the psychological status of the user based on monitoring; and (3) translate those inferences into timely and intuitive responses at the interface. In order to monitor-infer-adapt within a working control loop, multimodal interfaces must operate as surveillance systems, gathering data on individual users in order to respond proactively in an intelligent fashion. In addition, the multimodal interface requires a degree of autonomy to adapt to changes in user state without any requirement for explicit commands. This combination of intensive user monitoring with autonomous function is the price to be paid for the advanced level of functionality characterized by multimodal interfaces.

Societal issues of trust and system autonomy are both significant and inherently interconnected for the introduction of multimodal interfaces, particularly those designed to capture non-intentional responses as part of passive or reactive systems. The first concern is the degree of confidence that the user has in the technical prowess of the system. In other words, can the system collect data with sufficient fidelity? Is it capable to make a sensitive and accurate discrimination between different psychological states? Can the system successfully translate these multimodal data into sensitive and intelligent adaptation at the interface? If the user can answer those questions in the positive, he is likely to trust the system and will be comfortable ceding control to autonomous functions. If not, the user will either desire a return to manual control (if that is possible) or work unhappily and suspiciously with a technology that he views as erratic and unpredictable; in either case, the proposed advantages of multimodal interfaces will be lost.

A second issue concerns the 'values' that are inherent in the control directives of a 'machine with an agenda' [Fairclough 2015]. In order for the multimodal interface to respond to the detection of specific psychological states, it must translate the detection of a target state into an appropriate response at the interface. This process of translation can be straightforward. If the user is working on a safety-critical task, the adaptive logic of the multimodal interface should promote sustained engagement with the task in order to achieve error-free performance and maximize safety. If the user becomes bored or complacent, the multimodal interface will evoke a strategy to restore task engagement, e.g., to transfer tasks from autonomous to manual control in order to re-engage the user with task requirements. This enhanced autonomy, which is characteristic of multimodal interfaces, permits the system to adapt in order to operate upon the user—to effectively manage the psychological state of the person. Naturally it is important for the user to trust the system if he is to be completely comfortable with this type of advanced interaction. It is also possible that the goals and desires of the human user may diverge from the control directives of the multimodal interface: the user may desire to take a rest break, may feel unwell, or may resent working with a system that seems to expect him to do all the work. There is potential for an interaction with the multimodal interface to descend into a battle of "wills" where the human is forced to subjugate his wishes or desires in the face of a technical entity, which is both implacable and incapable of behaving with sufficient flexibility. This example demonstrates how implicit control based on monitoring combined with autonomous function can 'snowball' into a subversion of human goals and desires. This is why trust is such an important societal factor for acceptance of multimodal interfaces, humans must: (1) have faith in the technical proficiency of the system in order to comfortably relinquish some control over the interaction; and (2) interact with the system in the knowledge that autonomous decision-making will not subvert their autonomy and rights as human beings.

14.5.2.2 The Relevance of Consent and Data Privacy

Consent is a second important societal factor to be considered when the behavior and physiology of the person is monitored by technology. The data that streams from the user to the multimodal interface must be considered to be personal data. Issues surrounding data privacy and data ownership may be crucial influences on the extent to which users will accept the introduction of multimodal interfaces. It has been argued that openness with respect to data acquisition, storage and sharing is fundamental to the relationship between the user and the system, i.e., reciprocal accountability [Brin 1999]. In this case, the user allows personal data to be collected in full knowledge on how it will be used, stored ,and protected by the system. There is evidence that users would prefer to have a contractual arrangement whereby data is only obtained, stored and shared with full written consent [Reynolds and Picard 2005]. At the time of writing, research governance when performing experiments with human participants often requires written consent and compliance with data protection laws before personal data can be collected. But the degree of control that users can legally exercise over the ways in which personal data is stored and used is fundamentally determined by the extent to which users are deemed to own their own data [Fairclough 2014]. If the user is granted full ownership, they can control what is stored and who can access these data. Full ownership would allow a user to remove data from the system if they wished to do so, the user could even charge for access to their data.

Personal ownership of data and informed consent are important steps to protect the individual equipped with multimodal sensors. However, it is equally important that personal data is stored and managed in a way that continues to safeguard the rights of the individual. The management of personal data by external agencies sits at the heart of the General Data Protection Regulation (GDPR) that came into force throughout Europe in May 2018. This legislation grant greater protection and rights to the individual whose personal data are held by 'controllers' or 'processors', whether they be individuals, companies, or organizations. The issue of consent is central to GDPR and entities that hold personal data are subject to fines if a breach of confidentiality is detrimental to the individual. While GDPR is a positive development with respect to accountability, the primary risk to privacy from multimodal monitoring is a process of inference fueled by data aggregation from multiple sources [Friedland and Tschantz 2018]. The process of data aggregation is an important technique for accurate assessment of the operator state, for example, the detection of high mental workload is improved by cross-referencing task monitoring (e.g., activity or phase of operation) with neurophysiological data, such as EEG. The former provides a meaningful context for the latter. Similarly, emotional responses can be characterized by a combination of facial expression, autonomic psychophysiology, body posture, and vocal expression. It is important for users of multimodal sensor systems to understand which data sources are active and how they aggregate in order to deliver an inference about operator state, especially if aggregation occurs across databases that are controlled by different 'controllers' or 'processors.' In the previous example, these types of data can also serve secondary purposes, such as identifying individuals via unique features such as facial expression or voice, which leads to a scenario where individuals can effectively 'profiled' and compared with respect to specific operator states, e.g., John displayed high mental workload twelve times during the task compared to Jane who only entered a state of high workload twice.

The issue of data ownership points to another dimension of trust in multimodal interfaces—the issue of data privacy. Certain types of multimodal interfaces rely on the monitoring and measurement of psychological states, such as mental workload or frustration or fatigue (see also Volume 3, Chapter 13), but which other parties are allowed to access these data? And perhaps more importantly, can the individual user be identified on the basis of these data? A company may wish to record data from all user sessions, particularly during safety-critical activities, for purposes of accident investigation. In this case, part of the conditions of employment would require a user to share personal data and to be identified with that data. The societal issues associated with system use are more profound when a company wishes to access data from multimodal interfaces for purposes of performance management. For example, to assess the level of concentration exhibited by an employee during their duties or to capture episodes of frustration or to gauge alertness at the beginning of the work session. In this case, personal data is collected and interpreted from the individual in a way that could actively disadvantage that person. It is unlikely that the user would trust the system in this scenario, not because of what the technology is designed to do, but because of the way data is harvested and used by the system administration. It is also possible to imagine a more positive scenario where employee monitoring is performed to identify instances of high occupational stress in order to bring about changes to working conditions that ultimately benefit the employee. If we extend this scenario to service industries, such as multimodal interfaces in the context of internet search, a company can argue that harvesting data tied to an individual can be used to improve the quality of service offered to that user. The key issue in this case is whether the user must surrender their ownership of personal data in order to use that service, especially a service as pervasive and indispensable as internet search. A secondary problem that relates to the recent GDPR legislation concerns the sharing of personal data by a service provider with other entities, i.e., if the user surrenders ownership to use a service, do they also surrender the right to control the distribution of their personal data?

The issue of data privacy, anonymization, and sharing is particularly pertinent for systems that collect data from the brain and body as part of the human-computer interaction (see Friedland and Tschantz [2018] for further discussion). It is possible to extract information about the health of the person based on these data. As two examples, the presence of epilepsy can be detected from an EEG record collected on a long-term basis and it would also be possible to identify markers of cardiovascular disease from the regular acquisition of electrocardiographic data. It would be controversial for an employer to collect data from a specific individual for the purposes of mental workload monitoring, for example, and use the same data set as part of a health assessment. The sharing of sensitive, personal data of this type with other legal entities, such as health insurance companies, is also problematic from the perspective of the user. If the individual cannot own their data, they cannot control how it will be stored, shared, and analyzed, hence, some kind of informed consent possibly in the form of a contractual agreement along the lines of GDPR is likely to be a prerequisite for users of multimodal interfaces.

The ethical issues around data ownership have a number of practical implications for the design of multimodal interfaces. An individual could log on to the system anonymously and wipe the data record after use. This level of data protection is not possible when the system must store data that is linked to an individual user, for example, when the multimodal interface is designed to personalize adaptation to that person during sustained and repeated system use (see above). In this case, data must be stored and storage should be secure, e.g., password protected and encrypted. One solution is local storage, using a physical media such as a USB pen drive, that is owned by the user and used as a repository for all data collection during the interaction. However, due to the size of the data file over a period of time and the need for a back-up, it is unlikely that a physical device would serve as a practical solution in isolation and remote storage elsewhere would require protection and anonymization. There are a number of studies (e.g., Barra et al. [2016]) where data from the brain and body has been used to identify the individual and the use of a biometric key to unlock data from the same source offers an intriguing solution for the future.

Embedded multimodal interfaces offer the possibility of enormous technical advances but progress in this direction does introduce novel problems of user trust, system autonomy, and data ownership.

14.6 Supplementary Digital Materials: Exoskeleton's Mode Change Supported by Embedded Brain Reading—Approach and Evaluation

The VI-Bot exoskeleton (see the video referred to in Figure 14.8) has three operation modes: full user support mode, teleoperation mode, and transient mode (see Figure 14.9). During teleoperation mode, the movements of the human are mapped to the robotic system and the user receives force feedback from the robotic system. During full user support mode, the exoskeleton keeps the user's arm in a fixed position. During transient mode, the user can move freely without controlling the robotic system.

Changes between full user support mode and transient mode are intentionally controlled by overt arm movements, i.e., an active human state is detected by the exoskeleton. Changes between transient mode and teleoperation mode and back are intentionally commanded by overt hand gestures. Again, an active human state is detected by the interface. Being in the transient mode, there is a mode change possible back to full user support mode. This change is non-intentionally controlled whenever the user is *not* moving his or her arm for a certain time period. The user is usually not pausing to intentionally control the mode change but is stopping during robot control for different reason, e.g., to change position to get a better view of the robot's arm or to think about a solution for a difficult situation or to handle additional requests like communication with a second person, while the exoskeleton is supporting the user by keeping the arm in a fixed position such that the operator does not have to hold his arm by himself. The latter becomes relevant in case of very fine manipulations that are interrupted by longer breaks. In such situations, the user wants to avoid any big movements and wants to keep his arm in a specific position.

Thus, while mode changes between transient mode and teleoperation mode are intentionally controlled by the human, mode changes between transient mode and full user support mode are non-intentionally elicited (see above). Further, changes between full user support mode back to transient mode which are intentionally elicited by the human are additionally non-intentionally adapted by the interface itself by detecting specific active states of the human from covert measures, i.e., brain signals recorded from the surface of the human's head Figure 14.18. As mentioned above, when the system is in the full user support mode, the human can intentionally control the interface to change back to transient mode simply by starting to move his arm again. This overt behavior is measured by the exoskeleton to detect the active state of the human, i.e., arm movement.

To intentionally activate a mode change from full user support mode to the transient mode the user must press against force sensors of the exoskeleton for a certain time that is long enough to avoid false commands by, e.g., muscle twitches or movement of the upper body. However, even if the strength of the forces and the duration of pressure application to the exoskeleton are individually optimized for each user to cope with different body measures and muscle strength, the user will feel that the exoskeleton will not directly respond to her or his movements. Therefore, the explicit control from the full user support mode back to the transient mode is adapted by another implicit control loop that makes use of covert measures, i.e., brain activity, that enable the interface to detect another active state of the human, i.e., preparation of an arm movement. This cascaded approach of combining explicit control and implicit control for adaptation enhances transparency, since the movement planning phase can be detected from EEG before the movement is executed. This information can be used to prepare the exoskeleton for the explicit control by a later arm movement by enhancing the sensitivity of the sensors that detect the movement onset. As for the implicit control example, the adaptation requires the system to infer a human state or upcoming human state. No explicit command is required from the human.

To optimize the general sensitivity of the exoskeleton for each subject, a time threshold was defined individually, i.e., a minimum time he or she had to press against the exoskeleton with a minimum force to release it from full user support mode. Both minimum time threshold and force threshold were chosen by performing a calibration procedure in which the subjects had to keep their forearm completely extended and the shoulder flexed forward in order to bring the third joint of the exoskeleton in an angular range between $15 - 20°$. The user was then asked to repeat an oscillatory movement of about 90° with a regular speed (see Figure 14.19A). The requested type of movement and speed was shown to the user beforehand in a movie that showed the exact movement sequence and timing. At the beginning of each session, the force threshold was set to a maximum value

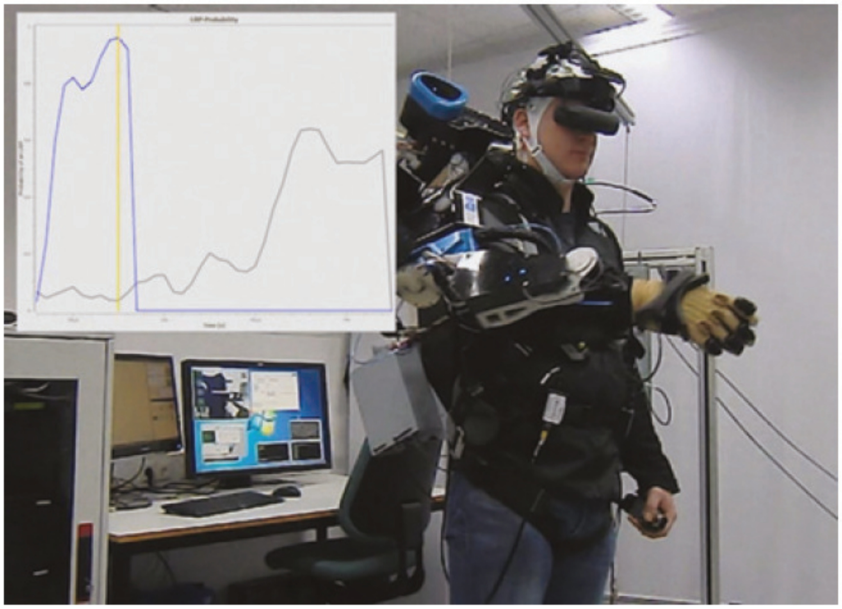

Figure 14.18 **Video:** VI-Bot implicit control to adapt the exoskeleton's control to the user's inferred needs: support of interaction by movement prediction based on EEG signals. https://youtu.be/fiI4MKPTFg0.

F_{th}^{Max} and the time threshold to a minimum value (T_{th}^{Min}) in order to assure that no movement would cause the system to lock-out. The user was asked to perform the rotational movement ten times. If the lock-out did not occur, F_{th} was decreased for 10% and the experiment was repeated. This went on until the user locks out due to the rotational movement. At this point, the force threshold was kept constant while the time threshold was increased until the perturbation movement was no longer able to cause the lock-out event. The discovered thresholds F_{th}^{Min} and T_{th}^{Max} were the values that were individually estimated for each subject. These values represented the time and force of pressure that must have been exceeded to lock out the system from full user support mode to transient mode. Thus, the time T_{th}^{Max} was chosen which was the shortest one that did not cause an unwanted lock-out at a minimum interaction force F_{th}^{Min}. This individually defined time threshold (T_{th}^{Max}, see T_max in Figure 14.19b) could further be reduced based on the outcome of the embedded brain reading interface that predicted movement preparation, while the required force (F_{th}^{Min}) was not adapted. The user had to press shorter (shorter than T_{th}^{Max}, see T max in Figure 14.19b) against the sensors the more likely movement planning was predicted based on EEG analysis. This adaptation of the exoskeleton required less force over time from the human for explicit control of the exoskeleton, i.e., to make the exoskeleton change from full user support mode to transient mode (see Figure 14.19). To calculate the effort required to pass from full user support mode to transient mode, we integrated the force according to Equation 14.1:

$$I = \int_{T_0}^{T_{\mathrm{Lout}}} F_{\mathrm{int}}(t)dt. \tag{14.1}$$

To analyze how much the required minimum interaction force I could be reduced by the approach, an experiment was performed in which a prediction score equivalent to 75% correct movement planning detection and a prediction score equivalent to 100% correct movement planning detection was randomly chosen to simulate the adaptation by embedded brain reading. Both conditions were interleaved with a no adaptation condition ($T_{\mathrm{th}}^{\mathrm{Max}}$; see $T\_\max$ in Figure 14.19b) chosen to adapt the exoskeleton control. In case of 100% correct movement planning the individually estimated time threshold $T_{\mathrm{th}}^{\mathrm{Max}}$ was reduced to $T_{\mathrm{th}}^{\mathrm{Min}}$, i.e., 10 ms (see $T\_\min$ in Figure 14.19b) which is the minimum time that the exoskeleton needs to react to a signal. This minimum response time is caused by the exoskeleton's $100Hz$ control cycle. In case of 75% correct movement planning, the individually estimated time threshold $T_{\mathrm{th}}^{\mathrm{Max}}$ was reduced to a time value between $T_{\mathrm{th}}^{\mathrm{Max}}$ and $T_{\mathrm{th}}^{\mathrm{Min}}$, i.e., $T_{\mathrm{th}}^{\mathrm{Mid}}$ (see T_mid in Figure 14.19b). Thus, $T_{\mathrm{th}}^{\mathrm{Min}}$ was equivalent with maximal adaptation by embedded brain reading, $T_{\mathrm{th}}^{\mathrm{Max}}$ with no adaptation by embedded brain reading, and $T_{\mathrm{th}}^{\mathrm{Mid}}$ with a medium adaptation by embedded brain reading. Interaction forces under all three conditions ($T_{\mathrm{th}}^{\mathrm{Max}}$, $T_{\mathrm{th}}^{\mathrm{Mid}}$, and $T_{\mathrm{th}}^{\mathrm{Min}}$) were measured by the force sensors which were embedded into the exoskeleton. Mean values were calculated for each prediction score value for ten measured movements across five subjects. It could clearly be shown that the interaction force applied over time was reduced (see value for force integral under all three condition), i.e., for more than one third under maximum adaptation compared to no adaptation. Subjects reported that they could clearly feel the differences in transparency of the exoskeleton even in case of medium adaptation, i.e. in case of $T_{\mathrm{th}}^{\mathrm{Mid}}$ (T_mid in Figure 14.19b).

Focus questions

14.1. What is the difference between explicit and implicit control approaches?

14.2. Which three safety levels can be defined for robots?

14.3. How can Moore's Law be related to robotics?

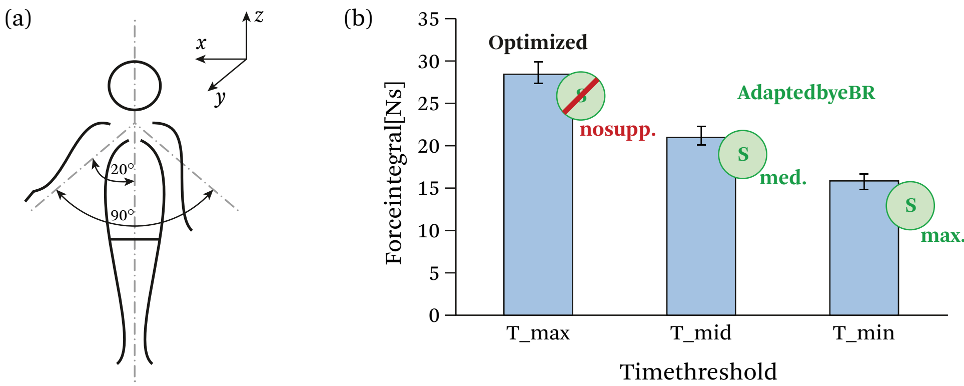

Figure 14.19 Reduction of interaction force between exoskeleton and human by adapting the exoskeleton's control by means of embedded brain reading. (A) positioning of forearm and movements of the upper body during the calibration session are depicted. (B) Mean values for applied force over time (force integral) calculated for ten measured movements across five subjects under each time threshold condition are depicted. T: time threshold, i.e., time to press against the sensors that must be exceeded to unlock the exoskeleton from full user support mode. T_max: individually estimated for each subjects based on calibration measurements with no support by embedded brain reading (no supp.); $T\_\text{min} = 10$ ms at S_{max}: maximum adaptation (prediction score equivalent with 100% correct movement planning detection); T_mid: medium time threshold value (between T_max and T_min) at medium adaptation S_{med} by embedded brain reading (prediction score equivalent with 75% correct movement planning detection). Figure modified after Folgheraiter et al. [2012].

14.4. When classifying interfaces which two types of behaviors and two types of methodologies can be used?

14.5. What is the difference between overt and covert measures of the user?

14.6. How does a passive BCI differ from an active BCI?

14.7. What is a hybrid BCI?

14.8. What is used in physiological computing and to what can it be compared?

14.9. How can a speech interface be implemented from the point of view of human measures?

14.10. What does "embedded" in "embedded multimodal interface" mean?

14.11. What is embedded brain reading?

14.12. How is the VI-Bot application non-intentionally adapted based on covert measures? Explain both approaches.

14.13. Why can the whole be more than the sum of the individual parts when combining different measures in embedded multimodal interfaces?

14.14. What are the two main trends for future embedded multimodal interfaces?

14.15. What is the relevance of the closed-loop design for personalization and machine intelligence?

14.16. How can we design an embedded multimodal interface while preserving the privacy of the individual?

References

B. D. Argall, S. Chernova, M. Veloso, and B. Browning. 2009. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5): 469–483. DOI: 10.1016/j.robot.2008.10.024. 253

J. Bachorowksi and M. J. Owren. 2010. Vocal expression of emotion. In M. Lewis et al., editor, *Handbook of Emotions*, pp. 196–210. Guilford Press. 229

V. Bargsten and J. de Gea Fernández. 2015. COMPI: Development of a 6-DOF Compliant Robot Arm for Human-Robot Cooperation. In *Proceedings of the 8th International Workshop on Human-Friendly Robotics, (HFR-2015)*, Technische Universität München (TUM), Munich, Germany. 225

S. Barra, A. Casanova, M. Fraschini, and M. Nappi. 2016. Fusion of physiological measures for multimodal biometric systems. *Multimedia Tools and Applications*, pp. 1–13. 10.1007/s11042-016-3796-1. 259

M. Barsotti, D. Leonardis, C. Loconsole, M. Solazzi, E. Sotgiu, C. Procopio, C. Chisari, M. Bergamasco, and A. Frisoli. 2015. A full upper limb robotic exoskeleton for reaching and grasping rehabilitation triggered by mi-bci. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pp. 49–56. DOI: 10.1109/ICORR.2015.7281174. 230

B. Blankertz, G. Dornhege, M. Krauledat, M. Schröder, J. Williamson, R. Murray-Smith, and K.-R. Müller. 2006. The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell. In *Verlag der Technischen Universität Graz*, pp. 108–109. 235

S. Bozinovski and A. Bozinovski. 2015. Mental states, eeg manifestations, and mentally emulated digital circuits for brain-robot interaction. *IEEE Transactions on Autonomous Mental Development*, 7: 39–51. DOI: 10.1109/TAMD.2014.2387271. 229

D. Brin. 1999. *The Transparent Society: Will Technology Force Us to Choose between Privacy and Freedom*. Perseus Books. 256

C. Brunner, B. Blankertz, F. Cincotti, A. Kubler, D. Mattia, F. Miralles, A. Nijholt, B. Otal, P. Salomon, and G. R. Muller-Putz. 2014. BNCI horizon 2020—towards a roadmap for brain/neural computer interaction. In *Brain-Computer Interfaces 2*, pp. 1–10. DOI: 10.1007/978-3-319-07437-5_45. 230

J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson. 2000. Psychophysiological science. In J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, editors, *Handbook of Psychophysiology*, pp. 3–26. Cambridge University Press, Cambridge UK. 238

C. Canning and M. Scheutz. 2013. Functional near-infrared spectroscopy in human-robot interaction. *Journal of Human-Robot Interaction*, 2(3): 62–84. DOI: 10.5898/JHRI.2.3.Canning. 229

A. Cotrina, A. Benevides, A. Ferreira, T. Bastos, J. Castillo, M. L. Menezes, and C. Pereira. 2014. Towards an architecture of a hybrid BCI based on SSVEP-BCI and passive-BCI. In *2014 36th Annual International Conference of the IEEE*, pp. 1342–1345. Engineering in Medicine and Biology Society (EMBC). DOI: 10.1109/EMBC.2014.6943847. 231

J. de Gea Fernández and F. Kirchner. 2015. Predictive compliance for interaction control of robot manipulators. In *Proceedings of the International Conference on Intelligent Robots and Systems, (IROS-2011)*, pp. 4134–4140. Technische Universität München (TUM), San Francisco, CA. DOI: 10.1109/IROS.2011.6094476. 226

J. de Gea Fernández, D. Mronga, M. Günther, T. Knobloch, M. Wirkus, M. Schröer, M. Trampler, S. Stiene, E. A. Kirchner, V. Bargsten, T. Bänziger, J. Teiwes, T. Krüger, and F. Kirchner. 2017. Multimodal sensor-based whole-body control for human–robot collaboration in industrial settings. *Robotics and Autonomous Systems*, 94: 102–119. http://www.sciencedirect.com/science/article/pii/S0921889016305127. DOI: 10.1016/j.robot.2017.04.007. 227, 237, 252

D-rock. 2015. http://robotik.dfki-bremen.de/en/research/projects/d-rock.html. 253

S. H. Fairclough. 2009. Fundamentals of physiological computing. *Interacting With Computers*, 21: 133–145. DOI: 10.1016/j.intcom.2008.10.011. 231

S. H. Fairclough. 2014. Physiological data should remain confidential. *Nature*, 505:263. DOI: 10.1038/505263a. 256

S. H. Fairclough. October 2015. A closed-loop perspective on symbiotic human-computer interaction. In B. Blankertz, G. Jacucci, L. amberini, A. Spagnolli, and J. Freeman, editors, *Symbiotic Interaction: 4th International Workshop, Symbiotic 2015*, Berlin, Germany, pp. 57–67. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-24917-9_6. 251, 255

M. Folgheraiter, M. Jordan, S. Straube, A. Seeland, S. K. Kim, and E. A. Kirchner. 2012. Measuring the improvement of the interaction comfort of a wearable exoskeleton. *International Journal of Social Robotics*, 4(3): 285–302. DOI: 10.1007/s12369-012-0147-x. 234, 235, 238, 240, 263

G. Friedland and M. Tschantz. 2018. Privacy concerns of multimodal sensor systems. In S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *The Handbook of Multimodal-Multisensor Interfaces, Volume 3: Multimodal Language Processing, Software Tools, Commercial Applications, and Emerging Directions*. Morgan & Claypool Publishers, San Rafael, CA, 257, 258

S. Haddadin. 2015. *Intuitive Interaction with Robots–Technical Approaches and Challenges*, pp. 249–271. Springer Verlag GmbH Heidelberg.227

L. J. Hettinger, P. Branco, L. M. Encarnaco, and P. Bonato. 2003. Neuroadaptive technologies: applying neuroergonomics to the design of advanced interfaces. *Theoretical Issues in Ergonomic Science*, 4: 220–237. DOI: 10.1080/1463922021000020918. 237

K. Hung, C. C. Lee, and S.-O. Choy. 2015. Ubiquitous health monitoring: Integration of wearable sensors, novel sensing techniques and body sensor networks. In S. Adibi, editor, *Mobile Health: A Technology Road Map*, pp. 319–342. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-12817-7_15. 234

IntelCooperation. 2005 Excerpts from a conversation with Gordon Moore: Moore's law. http://large.stanford.edu/courses/2012/ph250/lee1/docs/Excepts_A_Conversation_with_Gordon_Moore.pdf. [Online; accessed September 15, 2016]. 226

P. Kampmann and F. Kirchner. 2014. Integration of fiber-optic sensor arrays into a multi-modal tactile sensor processing system for robotic end-effectors. *Sensors–Open Access Journal*, Special Issue Tactile Sensors and Sensing Systems, 14(4): 6854–6876. DOI: 10.3390/s140406854. 228

P. Kampmann and F. Kirchner. May 2015. Towards a fine-manipulation system with tactile feedback for deep-sea environments. *Robotics and Autonomous Systems*, 67(C): 115–121. 10.1016/j.robot.2014.09.033. 226

B. H. Kim, M. Kim, and S. Jo. 2014. Quadcopter flight control using a low-cost hybrid interface with eeg-based classification and eye tracking. *Computers in Biology and Medicine*, 51: 82–92. 10.1016/j.compbiomed.2014.04.020. 232

S. K. Kim, E. A. Kirchner, A. Stefes, and F. Kirchner. 2017. Intrinsic interactive reinforcement learning—using error-related potentials for real world human-robot interaction. *Scientific Reports*, 7: 17562, 2017. https://doi.org/10.1038/s41598-017-17682-7. 251

E. A. Kirchner and R. Drechsler. 2013. A formal model for embedded brain reading. *Industrial Robot: An International Journal*, 40(6): 530–540. 10.1108/IR-01-2013-318. 233, 234, 238, 242, 243, 254

E. A. Kirchner, H. Wöhrle, C. Bergatt, S. K. Kim, J. H. Metzen, D. Feess, and F. Kirchner. 2010. Towards operator monitoring via brain reading—an EEG-based approach for space applications. In *Proceedings of the 10th International Symposium Artificial Intelligence, Robotics and Automation in Space*, pp. 448–455, Sapporo. 235

E. A. Kirchner, J. C. Albiez, A. Seeland, M. Jordan, and F. Kirchner. February 2013a. Towards assistive robotics for home rehabilitation. In M. F. Chimeno, J. Solé-Casals, A. Fred, and H. Gamboa, editors, *Proceedings of the 6th International Conference on Biomedical Electronics and Devices (BIODEVICES-13)*, pp. 168–177, Barcelona, Spain. SciTePress. 234, 235, 245

E. A. Kirchner, S. K. Kim, S. Straube, A. Seeland, H. Wöhrle, M. M. Krell, M. Tabie, and M. Fahle. 2013b. On the applicability of brain reading for predictive human-machine interfaces in robotics. *PLoS ONE*, 8(12): e81732, 12. 10.1371/journal.pone.0081732. 235, 238

E. A. Kirchner, M. Tabie, and A. Seeland. 2014. Multimodal movement prediction - towards an individual assistance of patients. *PLoS ONE*, 9(1): e85060, 01. 10.1371/journal.pone.0085060. 234, 235, 248, 249

E A. Kirchner, J. de Gea Fernandéz, P. Kampmann, M. Schröer, J. H. Metzen, and F. Kirchner. 2015. *Intuitive Interaction with Robots - Technical Approaches and Challenges*, pp. 224–248. Springer Verlag GmbH Heidelberg. DOI: 10.1007/978-3-658-09994-7_8. 228, 233, 235

E. A. Kirchner, S: K. Kim, M. Tabie, H. Wöhrle, M. Maurus, and F. Kirchner. 2016a. An intelligent man-machine interface - multi-robot control adapted for task engagement based on single-trial detectability of p300. *Frontiers in Human Neuroscience*, 10:291. ISSN 1662-5161. 10.3389/fnhum.2016.00291. 229, 235, 244, 246

E. A. Kirchner, N. Will, M. Simnofske, L. M. Vaca Benitez, B. Bongardt, M. M. Krell, S. Kumar, M. Mallwitz, A. Seeland, M. Tabie, H. Wöhrle, M. Yüksel, A. Heß, R. Buschfort, and F. Kirchner. 2016b. Recupera-Reha exoskeleton technology with integrated biosignal analysis for sensorimotor rehabilitation. In *Proceedings of the Second Conference on smartASSIST 2016*. 235

E. A. Kirchner, N. Will, M. Simnofske, L. M. Vaca Benitez, B. Bongardt, M. M. Krell, S. Kumar, M. Mallwitz, A. Seeland, M. Tabie, H. Wöhrle, M. Yüksel, A. Heß, R. Buschfort, and F. Kirchner. December 2016c. Recupera-reha: Exoskeleton technology with integrated biosignal analysis for sensorimotor rehabilitation. In *2. Transdisziplinäre Konferenz "Technische Unterstützungssysteme, die Menschen wirklich wollen." Transdisziplinäre Konferenz "Technische Unterstützungssysteme, die Menschen wirklich wollen,"*, pp. 504–517. Elsevier, Hamburg, Germany. https://www.dfki.de/web/forschung/publikationen/renameFileForDownload?filename=20161121_Recupera-Reha_ExoskeletonTechnology with Integrated.pdf&file_id=uploads_2992. 234

G. Klein, D. D. Woods, J. M. Bradshaw, R. R. Hoffman, and P. J. Feltovich. 2004. Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems*, 19: 91–95. 254

M. M. Krell, S. Straube, A. Seeland, H. Wöhrle, J. Teiwes, J. H. Metzen, E. A. Kirchner, and F. Kirchner. December 2013. pySPACE - a signal processing and classification environment in Python. *Frontiers in Neuroinformatics*, 7(40). https://github.com/pyspace. 10.3389/fninf.2013.00040. 252

F. L. Lewis, D. M. Dawson, and C. T. Abdallah. 2003. *Robot Manipulator Control: Theory and Practice*. CRC Press. 234

C. Liu, K. Conn, N. Sarkar, and W. Stone. 2008. Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder. *International Journal of Human-Computer Studies*, 66: 662–677. 238

C. Lucarotti, C. M. Oddo, N. Vitiello, and M. C. Carrozza.2013. Synthetic and bio-artificial tactile sensing: A review. *Sensors*, 13(2): 1435. URL http://www.mdpi.com/1424-8220/13/2/1435. 10.3390/s130201435. 226

C. Lüth, U. Frese, H. Täubig, D. Walter, and D. Hausmann. 2015. Sams sicherheitskomponente für autonome mobile serviceroboter. In *Proceedings ROBOTIK 2008*, Munich, Germany. VDI-Verlag. 227

M. Mallwitz, N. Will, J. Teiwes, and E. A. Kirchner. 2015. The capio active upper body exoskeleton and its application for teleoperation. In *Proceedings of the 13th Symposium on Advanced Space Technologies in Robotics and Automation*. ESA. 234

MARS. 2015. Mars - a cross-platform simulation and visualization tool. http://rock-simulation.github.io/mars/. Last accessed on April 11, 2015. 243

J. H. Metzen, A. Fabisch, L. Senger, J. Gea Fernández, and E. A. Kirchner. 2013. Towards learning of generic skills for robotic manipulation. *KI - Künstliche Intelligenz*, 28(1): 15–20. 10.1007/s13218-013-0280-1. 252

C. A. Miller. 2005. Trust in adaptive automation: the role of etiquette in tuning trust via analogic and affective methods. In *First International Conference on Augmented Cognition*, Las Vegas, NV. 254

H. Monkaresi, S. Hussain, and R. Calvo. 2014. Using Remote Heart Rate Measurement for Affect Detection. *Florida Artificial Intelligence Research Society Conference*. 228

G. Muller-Putz, R. Leeb, M. Tangermann, J. Hohne, A. Kubler, F. Cincotti, D. Mattia, R. Rupp, K. R. Muller, and J. Del R Millan. 2015. Towards noninvasive hybrid brain-computer interfaces: Framework, practice, clinical application, and beyond. *Proceedings of the IEEE*, 103: 926–943. DOI: 10.1109/JPROC.2015.2411333. 231

K. Mülling, J. Kober, O. Koemer, and J. Peters. 2013. Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, 32: 263–279. DOI: 10.1177/0278364912472380. 252

A. Myrden and T. Chau. 2015. Effects of user mental state on EEG-BCI performance. *Frontiers in Human Neuroscience*, 9:308. DOI: 10.3389/fnhum.2015.00308. 231

F. Nijboer, B. van de Laar, S. Gerritsen, A. Nijholt, and M. Poel. 2015. Usability of three electroencephalogram headsets for brain–computer interfaces: A within subject comparison. *Interacting with Computers*, 27: 500–511. DOI: 10.1093/iwc/iwv023. 229

D. Novak, X. Omlin, R. Leins, and R. Riener. 2013. Predicting targets of human reaching motions using different sensing technologies. *IEEE Transactions on Biomedical Engineering*, 60(9): 2645–2654. http://www.biomedsearch.com/nih/Predicting-targets-human-reaching-motions/23674417.html. DOI: 10.1109/TBME.2013.2262455. 252

Y. Panagakis, O. Rudovic, and M. Pantic. 2018. Learning for multi-modal and context-sensitive interfaces. In S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *The Handbook of Multimodal-Multisensor Interfaces*, Volume 2: Signal Processing, Architectures, and Detection of Emotion and Cognition. Morgan & Claypool Publishers, San Rafael, CA. 248

P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal. 2009. Learning and generalization of motor skills by learning from demonstration. In *2009 IEEE International Conference on Robotics and Automation*, pp. 763–768. IEEE. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5152385. 10.1109/ROBOT.2009.5152385. 252

S. K. Piper, A. Krueger, S. P. Koch, J. Mehnert, C. Habermehl, J. Steinbrink, H. Obrig, and C. H. Schmitz. 2014. A wearable multi-channel fnirs system for brain imaging in freely moving subjects. *Neuroimage*, 85(1): 64–71. DOI: 10.1016/j.neuroimage.2013.06.062. 229

C. Reynolds and R. W. Picard. 2005. Affective sensors, privacy and ethical contracts. In *CHI'2005*, Vienna, Austria. DOI: 10.1145/985921.985999.256

ROCK. 2015. Rock - robot construction kit. http://robotik.dfki-bremen.de/en/research/softwaretools/rock-1.html. Last visited on 28.02.2016.

M. Rommerman, D. Kühn, and F. Kirchner. 2009. Robot design for space missions using evolutionary computation. In *Evolutionary Computation, 2009. CEC '09. IEEE Congress on*, pp. 2098–2105. 10.1109/CEC.2009.4983200. 243

SafetyEye. 2014. SafetyEye from PILZ. http://brochures.pilz.nl/bro_pdf/SafetyEYE_2014.pdf. Accessed August 16, 2016. 227

S. Schaal. 1997. Learning from demonstration. In *Advances in Neural Information Processing Systems* 9, pp. 12–20. MIT Press. 253

B. Schuller. 2018. Multimodal user state and trait recognition: An overview. In S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *The Handbook of Multimodal-Multisensor Interfaces*, Volume 2: Signal Processing, Architectures, and Detection of Emotion and Cognition, chapter 5. Morgan & Claypool Publishers, San Rafael, CA. 228

L. Senger and E. A. Kirchner. 2016. Automatic Detection and Recognition of Human Movement Patterns in Manipulation Tasks. In *Proceedings of the International Conference on Physiological Computing Systems (PhyCS 2016)*, Lisbon, Portugal. SCITEPRESS Digital Library. 252

N. Serbedzija and S. H. Fairclough. 2012. Reflective pervasive systems. *ACM Transactions on Autonomous and Adaptive Systems*, 7. 250

S. Straube, M. Rohn, M. Roemmermann, C. Bergatt, M. Jordan, and E. Kirchner. 2011. On the closure of perceptual gaps in man-machine interaction: virtual immersion, psychophysics and electrophysiology. *Perception*, 40 ECVP Abstract Supplement:177. 241

C. Vogel, C. Walter, and N. Elkmann. 2013. A projection-based sensor system for safe physical human-robot collaboration. In *IROS*, pp. 5359–5364. IEEE. 227

M. Wand, C. Schulte, M. Janke, and T. Schultz. 2013. Array-based electromyographic silent speech interface. In *6th International Conference on Bio-inspired Systems and Signal Processing*. http://www.csl.uni-bremen.de/cms/images/documents/publications/BS13_WandSchulteJankeSchultz_ArrayBasedEMGSSI.pdf. BIOSIGNALS 2013. 229

H. Wöhrle and E. A. Kirchner. 2014. Online classifier adaptation for the detection of p300 target recognition processes in a complex teleoperation scenario. In H. P. da Silva, A. Holzinger, S. Fairclough, and D. Majoe, editors, *Physiological Computing Systems*, volume 8908 of *Lecture Notes in Computer Science*, pp. 105–118. Springer Berlin Heidelberg. 235, 242

H. Wöhrle and F. Kirchner. 2015. Reconfigurable hardware-based acceleration for machine learning and signal processing. In R. Drechsler and U. Kühne, editors, *Formal Modeling and Verification of Cyber-Physical Systems: 1st International Summer School on Methods and Tools for the Design of Digital Systems*, Lecture Notes in Computer Science, chapter Springer Fachmedien Wiesbaden. Springer. 252

H. Wöhrle, J. Teiwes, M. Tabie, A. Seeland, E. A. Kirchner, and F. Kirchner. 2014. Prediction of Movements by Online Analysis of Electroencephalogram with Dataflow Accelerators. In *Proceedings of the International Congress on Neurotechnology, Electronics and Informatics (NEUROTECHNIX 2014)*, pp. 31–37, Rome, Italy. ScitePress. 253

J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6): 767–791. 230

C. Yongwook, J. Jaeseung, and J. Sungho. 2012. Toward brain-actuated humanoid robots: Asynchronous direct control using an EEG-based BCI. *IEEE Transactions on Robotics*, 28: 1131–1144. 230

T. O. Zander and C. Kothe. 2011. Towards passive brain-computer interfaces: applying brain- computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, 8:1–5. 230

T. O. Zander, J. Bronstruo, R. Lorenz, and L. R. Krol. 2014. Towards BCI-based implicit control in human-computer interaction. In S. H. Fairclough and K. Gilleade, editors, *Advances in Physiological Computing*, pp. 67–90. Springer-Verlag. 231

J. Zhou, K. Yu, F. Chen, Y. Wang, and S. Z. Arshad. 2018. Multimodal behavioural and physiological signals as indicators of cognitive load. In S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, and A. Krüger, editors, *The Handbook of Multimodal-Multisensor Interfaces*, Volume 2: Signal Processing, Architectures, and Detection of Emotion and Cognition. Morgan & Claypool Publishers, San Rafael, CA. 242, 243