Supporting Human–Robot Interaction by Projected Augmented Reality and a Brain Interface

Supporting Human–Robot Interaction by Projected Augmented Reality and a Brain Interface

Francesco De Pace, Federico Manuri, *Member, IEEE*, Matteo Bosco, Andrea Sanna, and Hannes Kaufmann

***Abstract*—This article presents a brain–computer interface (BCI) coupled with an augmented reality (AR) system to support human–robot interaction in controlling a robotic arm for pick-and-place tasks. BCIs can process steady-state visual evoked potentials (SSVEPs), which are signals generated through visual stimuli. The visual stimuli may be conveyed to the user with AR systems, expanding the range of possible applications. The proposed approach leverages the capabilities of the NextMind BCI to enable users to select objects in the range of the robotic arm. By displaying a visual anchor associated with each object in the scene with projected AR, the NextMind device can detect when users focus their eyesight on one of them, thus triggering the pick-up action of the robotic arm. The proposed system has been designed considering the needs and limitations of mobility-impaired people to support them when controlling a robotic arm for pick-and-place tasks. Two different approaches for positioning the visual anchors are proposed and analyzed. Experimental tests involving users show that both approaches are highly appreciated. The system performances are extremely robust, thus allowing the users to select objects in an easy, fast, and reliable way.**

***Index Terms*—Assistive robotics, augmented reality (AR), brain interface, NextMind, severe motor impairment, steady-state visual evoked potential (SSVEP).**

I. Introduction

HUMAN–ROBOT interaction (HRI) is an extensive and diverse research discipline, which can be divided approximately into the following four subareas [1]:

1) robots performing routine activities supervised by humans;
2) robots teleoperated by remote humans;
3) humans as passengers of automated vehicles;
4) humans assisted and/or entertained by robots.

This work falls under the last category: A novel interaction paradigm based on projected augmented reality (AR) and a brain–computer interface (BCI) to support mobility-impaired people.

Assistive-robot-based technologies might allow impaired people to restore a vital degree of independence in several daily activities. In this context, the main challenge is how users with severe motor impairments can express their intentions. Usually, people without motor impairments use physical gestures, motions, or language patterns to communicate their intentions to the robot; on the other hand, these communicative forms are often precluded to paralyzed patients who can establish communication channels only by implicit communication often based on gaze and brain activity [2].

BCIs detect and translate brain signals into commands for devices that can support users in several activities [3], [4]. For people with severe motor impairments and paralyzed patients, BCIs can be the only technology enabling them to perform daily tasks [5]. Among the different input signals BCIs can process, steady-state visual evoked potentials (SSVEPs) are signals generated by visually stimulating the visual cortex [6]. The stimulation may happen by visualizing real or virtual objects, and virtual ones can be effectively generated and displayed with AR devices. AR extends the real world with a set of computer-generated assets, which can be displayed by either wearable, hand-held, or projected devices, thus implementing different AR paradigms [7].

This article proposes a novel HRI paradigm that combines brain and projected AR technologies to support mobility-impaired people when controlling a robotic arm for pick-and-place tasks. When referring to *mobility-impaired* individuals, we are specifically addressing those who face significant impairments in their lower limbs, which may result from illnesses, injuries, disabilities, or age-related issues, leading to challenges or the inability to walk. The considered scenario (see Fig. 1-A) can be summarized as follows.

1) A robotic arm is in charge of picking up a set of objects and placing them close to the user.
2) An RGB camera frames the robot workspace to detect and recognize objects.
3) A projector augments the workspace by displaying virtual contents that trigger the visual cortex.
4) The user wears an SSVEP-BCI device.

Each virtual stimulus is uniquely associated with a specific real object, and the user can command the robot to pick it up by focusing on the virtual content. This configuration can be adapted to a solution that might help mobility-impaired people (see Fig. 1-B): the camera, the projector, and the robotic arm

Manuscript received 6 February 2024; revised 17 April 2024; accepted 13 May 2024. This work was supported by the TU Wien Bibliothek for financial support through its Open Access Funding Programme. This article was recommended by Associate Editor F. Buttussi. *(Corresponding author: Francesco De Pace.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was performed in line with the Declaration of Helsinki.

Francesco De Pace, Matteo Bosco, and Hannes Kaufmann are with Virtual and Augmented Reality Group, TU Wien, Wien 1040, Austria (e-mail: francesco.pace@tuwien.ac.at).

Federico Manuri and Andrea Sanna are with the Department of Control and Computer Engineering, Politecnico di Torino, Torino 10129, Italy.

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/THMS.2024.3414208.

Digital Object Identifier 10.1109/THMS.2024.3414208

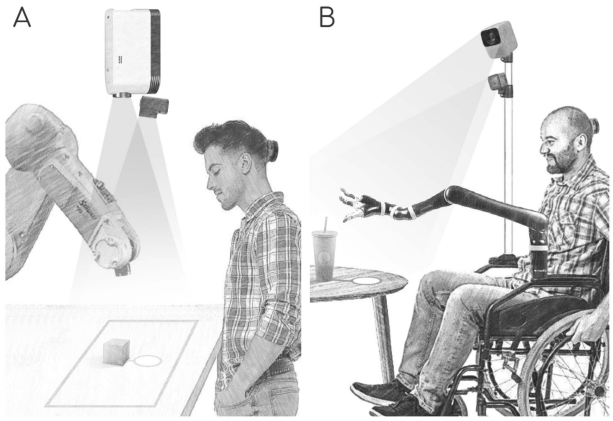

Fig. 1. Two possible scenarios. (A) has been implemented to test usability, robustness, and workload. (B) represents a possible mobile solution.

are installed on the user's wheelchair, thus providing a mobile solution.

Since the effectiveness of the SSVEP-BCI is strictly correlated with the visualization of the visual stimuli, their positioning in the real space plays an important role. Hence, two different positional approaches are proposed and compared: 1) *adaptive* (w.r.t. both the objects and the user's head position), the virtual contents are placed close to the objects in visible areas and 2) *nonadaptive*, the virtual contents are positioned linearly outside the robot workspace.

To verify whether the projected AR contents can be effectively combined with SSVEP-BCIs to assist mobility-impaired people, the proposed system and the two different approaches are first evaluated with healthy users in a user study that simulates the conditions of mobility-impaired users. Due to variability in brain signals and abilities, researchers often initiate the testing of custom BCI systems on healthy individuals to optimize parameters, assess usability, and identify potential issues. The proposed research follows this approach, beginning with healthy users but also conducting preliminary tests on healthy participants under conditions resembling those of mobility-impaired individuals to evaluate the effectiveness and feasibility of the proposed interface and robotic system. The goal is to gather insights, identify limitations and optimization opportunities, and establish a baseline for performance before testing on individuals with disabilities.

The SSVEP-BCI is implemented by the NextMind device [8]. It uses as visual stimuli the so-called *NeuroTags*, which are virtual flickering textures that elicit the visual cortex.

The effectiveness of the proposed solution is assessed by answering the following research questions.

1) *R1a*: Is the usability of the proposed solution suitable to support an effective HRI? Another secondary research question is strictly related to R1a. *R1b*: Is the adaptive solution significantly more usable than the linear one?
2) *R2*: What is the perceived workload level?
3) *R3:* Is the robustness of the proposed solution suitable to support an effective HRI?

The aforementioned questions were used to select the assessment methodology. In particular, 22 users have been involved in tests, and different use cases (obtained by varying the number of objects in the workspace) have been considered. The system usability scale (SUS) [9] and the NASA-TLX [10] have been used to gather the users' feedback in order to assess usability and workload, whereas completion times and selection errors have been measured to assess the robustness of the proposed system. The result analysis outlines that the proposed positioning strategies are highly appreciated (high usability scores), and the adaptive placement is more usable than the linear solution. Moreover, even if the number of selection errors is always extremely low for both algorithms, the adaptive solution allows the users to interact more rapidly than the nonadaptive one, regardless of the task at hand (see Section IV-A).

The rest of the article is organized as follows. Section II presents the state of the art, with a particular focus on HRI, whereas the proposed system is presented in Section III. The user study and the system limitations can be found in Section IV. Finally, Section V concludes this article.

II. State of the Art

A. BCI Interfaces

BCIs decode brain activity and translate it into meaningful commands for "machines," thus helping humans suffering from severe motor impairments to bridge the gap with the external world [11], [12], [13]. A brain interface can be seen as a system that receives, processes, and outputs signals; the processing algorithm is in charge of mapping input signals to output signals [14]. Input signals usually belong to one of the following categories: the P300 wave of event-related potentials (ERP); SSVEP; slow cortical potentials and motor imagery (MI). The brain activity (i.e., the electroencephalogram—EEG) is usually measured noninvasively by means of electrodes (e.g., gel, water, or dry electrodes) mounted on the human scalp [15], [16]. In the proposed solution, the BCI is implemented by the NextMind device that falls under the SSVEP category.

The BCI signal processing system performs two phases: 1) calibration and 2) feedback. The calibration phase is used to train some kind of classifier (e.g., the NextMind uses a machine learning approach for the signal classification), whereas the feedback phase allows the user to control external devices/applications by brain activity; the classifier detects and decodes different mental states, thus translating them into control signals for the external world [17]. BCIs can help patients to actuate wheelchairs [18], exoskeletons [19], drones [20], robots [21], and many other types of machines/interfaces. A recent study by Peters et al. [22] investigated the usage of SSVEP BCI and eye tracking by individuals with late-stage ALS and visual impairments. The study found that the SSVEP BCI interface was more reliable and accurate in detecting user intent than the eye-tracking system.

B. Remote Interfaces for Telerobotics

While robots are advancing in autonomy and intelligence, their complete autonomy may be inadequate for tasks

involving human judgment, necessitating collaboration with human operators to teleoperate them partially or entirely. Eye-tracking interfaces have been proposed to remotely control mobile and arm robots. Minamoto et al. [23] developed an eye-tracking interface to remotely control crawler-type robots. The viewing area is divided into a 3×3 grid and the robot moves depending on the cell grid selected by using eye movements. Sharma et al. [24] exploited the movement of the eyes' pupils to control a small robotic arm for pick-and-place scenarios, by also keeping a safe distance for the entire duration of the task. For generating smooth robot trajectories, Scalera et al. [25] showed that eye fixations should be detected and removed during the detection of the eyes' movements.

Immersive virtual reality (VR) interfaces also showed their potential for remotely controlling industrial robots. For instance, in [26], a VR system for teleoperating a mobile robotic platform in pick-and-place scenarios is introduced. The primary results indicate that the system can attain a grasping accuracy of 0.5 cm and a placement accuracy of 1.5 cm. Lipton et al. [27] introduced a VR interface that separates the user and the robot by incorporating a VR control room, serving as a bridge between the human and the robot. Meanwhile, De Pace et al. [28] evaluated the efficacy of VR telerobotic interfaces for high-precision tasks.

C. BCI and AR Interfaces for HRI

The use of AR to support the interaction between humans and robots has been deeply investigated: Visual augmentation provided by AR offers many benefits when designing HRI paradigms. These benefits have been presented in [29], and they can be summarized as follows:

1) facilitated robot programming;
2) real-time support for control and navigation;
3) improved safety;
4) communication of the robot's intention;
5) increased expressiveness.

Hybrid solutions combining both BCIs and AR have also been proposed. The first attempts to deploy AR–BCI-based interfaces date back to 2010 [30]: Six cubes are tracked by using ARToolkit markers that are used to identify the real-world coordinates for the pick-and-place tasks. The AR interface augments the environment by displaying numbers close to the cubes; the user mentally counts the number associated with the desired position to activate an ERP BCI. Faller et al. [31] evaluated 3-D computer graphics for generating SSVEP stimuli, confirming their efficacy in VR and AR scenarios. In another study [32], a hybrid interface for robotic tasks combined eye-tracking and a BCI, utilizing AR for visual feedback. The integrated AR in a closed-loop system improved efficiency, reducing the number of commands and decreasing gripper height gaps during lifting. Si-Mohammed et al. [33] demonstrated the effective combination of a BCI with Microsoft HoloLens, showcasing the tolerance of small head movements during BCI use in AR contexts. Implementing a portable, closed-loop, AR-based BCI was investigated in [34] to assess the feasibility of controlling a physical device through SSVEP. The study results are promising despite the limited number of participants in the tests (three) and the reduced number of flickers used by the SSVEP system (two). A novel approach to noninvasively detect brain signals using micropatterned epitaxial graphene (MEG) sensors was investigated in [35]. The authors demonstrated a full brain–machine interface system to control a quadruped robot via the SSVEP paradigm, also discussing the importance of the EEG sensor positioning. Interested readers can refer to the work in [36] for a comprehensive review detailing the main requirements of SSVEP-AR systems.

The NextMind device has been utilized to develop hybrid BCI-AR interfaces, as seen in a comparison study [37] evaluating accuracy and mental workload in hologram selection. Additionally, the NextMind has been integrated with the Microsoft HoloLens for controlling a robotic arm in collaborative manufacturing settings [38], where Neurotags overlaid onto the real environment allow users to visually command the robot to interact with specific objects by focusing on the NeuroTags.

D. Label Placement in AR

Label placement in AR and projected AR (PAR) is a significant research area, with researchers devising techniques to position labels on virtual objects. Hirsch [39] introduced an algorithm for automatic name placement around point data in printed maps, which later influenced research on label placement in digital maps and virtual environments. Kato and Billinghurst [40] developed a PAR system using marker tracking and head-mounted display calibration to project virtual objects onto physical surfaces. They tackled challenges in label placement, considering factors such as perspective distortion and occlusion, and proposed techniques to optimize placement based on user viewpoint and surface geometry. Cotting et al. [41] proposed an adaptive PAR system that analyzes the reflection properties and the depth discontinuities of the projection area. The shape of the projected contents is continuously updated to fit the visible regions. Laplacian filters are instead used in [42] to avoid occluded areas. Riemann et al. [43] proposed a probability map that considers the smoothness and lightness of the projection surface. Nonplanar and textured surfaces are analyzed in [44] to find the most suitable projection area for label placement, considering the text legibility. Finally, Ichihashi and Fujinami [45] proposed a machine-learning approach to assess whether a reflective area is readable. The classifier changes the label properties according to the characteristics of the reflective surface.

E. Contribution

Although the state-of-the-art presents several works related to AR label positioning and SVVEP-BCIs for HRI, this work differs from it for several reasons as follows.

1) Textual labels and flickering textures provide information in different ways. Indeed, the effectiveness of textual labels strictly depends on the positioning but also on the users' ability to *comprehend* the related written text. On the contrary, the "activation" of the flickering textures relies on the users' ability to *focus* on the moving visual patterns that act as a stimulus for the visual cortex. Moreover, the recognition of the blinking frequency depends

on several factors [46] including the texture visibility and whether the positioning might affect their activation is still unclear and researched in this article.

2) In addition to providing several details regarding the performances of the proposed solution (e.g., computational time, number of failures, etc.), this article places a strong emphasis on the system usability and workload.
3) The user study simulates the conditions of a mobility-impaired person. Specifically, the users cannot change their body position and can only rotate their heads to interact with the NeuroTags. These limitations may affect the capability of the users to focus on the visual stimuli, and their effects on the activation of SSVEP are still to be investigated.

The presented solution is different from the work presented in Sanna et al. [38] as they do not consider projected AR interfaces nor do they investigate the effectiveness of different positional algorithms. Moreover, their user study does not consider the limitations of mobility-impaired people. On the contrary, the proposed work improves the current AR-BCI literature by highlighting the importance of the SSVEP stimuli positional approach and also evaluating the usability, workload, and robustness of the proposed system.

III. Proposed System

This section describes 1) the BCI-AR interface, 2) the NeuroTag positional approaches, and 3) the performance of the proposed system.

A. Brain-AR Interface

The proposed interface allows users to pick up objects by controlling a robotic arm with a brain interface and a projected AR visualization system. Specifically, by wearing the NextMind device, users can select the NeuroTags projected on a plane. Each NeuroTag (i.e., the visual stimuli) blinks at a specific frequency; thus, it can be uniquely identified, generating a visual signature in the users' brains. A single NeuroTag is associated with a single object. Therefore, the users can easily command the robot to pick up the object by selecting the corresponding NeuroTag using the NextMind device. The NeuroTags are enhanced with graphical feedback to help users understand the level of focus they are achieving while performing a selection. Since the device is not open source and it is only possible to detect whether a NeuroTag has been recognized or not, the raw data as well as the algorithm used for their processing are not publicly available. However, a recent study showed that the median activation time is around 2.35 s and the device uses stimulus frequencies of 3 Hz [47]. To clearly detect different object-NeuroTag pairs, a color palette characterized by a considerable spatial distance has been used to color each pair. The color palette has been ordered, trying to juxtapose complementary colors. The coloring system takes inspiration from the color strategies used to represent the subway maps of the world's largest cities [48].

The AR interface consists of a projector pointing downward toward a working table. The projected user interface (UI) displays the following:

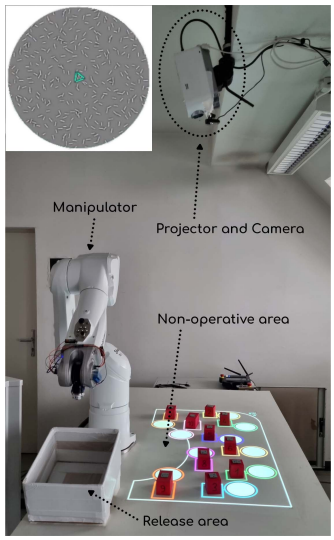

Fig. 2. NeuroTag (top-left) and the projected interface. The NeuroTags are positioned close to the objects to be picked up by the manipulator.

1) 2-D bounding boxes (BBs) that surround the real objects;
2) the NeuroTags;
3) the operative area (i.e., the area that the robotic manipulator can reach);
4) the nonoperative area (i.e., the area that the robotic manipulator cannot reach);
5) some icons that indicate the connection and battery status of the NextMind.

The projector (full HD, 1920 $\times$ 1080p) is positioned 143 cm from the table and it covers an area of 1000 $\times$ 560 mm. Thus, the resolution is of $\simeq$ 2 mm/px. Fig. 2 shows the projected AR interface.

B. Hardware and Software Architecture

The hardware architecture is composed of a Windows 10 i9-11900K CPU @ 3.50 GHz PC (W-PC), an Ubuntu i7-6700 CPU @ 3.40 GHz PC (U-PC), and a Staubli TX2-60 6-DOF robot arm [49], connected to a shared local area network. An LG HF80LA laser projector [50] is connected to the U-PC, whereas the NextMind sends data to W-PC through Bluetooth. Finally, W-PC controls a custom-made electromagnetic gripper designed with Arduino through Bluetooth. The W-PC application has been developed using the Unity3D [51] game engine, whereas U-PC handles several ROS nodes developed using ROS Melodic [52].

Referring to Fig. 3, the camera sends a snapshot of the current objects' configuration to U-PC (1). After a calibration phase, the acquired snapshot is analyzed, identifying the objects' 2-D BBs that are sent to W-PC as ROS messages (2). Then, W-PC computes the NeuroTag positions using one of the two available approaches, either adaptive or nonadaptive (3). The NeuroTags are displayed using the laser projector (4). Finally, the user selects one of the NeuroTags using the NextMind (5), and the robot picks up the corresponding object using the custom-made gripper (6).

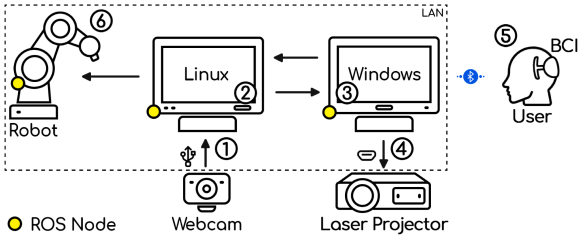

Fig. 3. High-level overview of the proposed architecture.

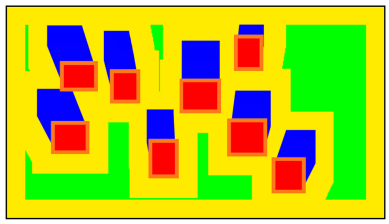

Fig. 4. Operative area, divided into available and occupied cells. Available cells are in green and correspond to the search space subdivided with the quadtree data structure. For the occupied cells, the red color is for the BBs, the orange color is for the BB borders, the yellow color is for the borders, and the blue color is for the occlusions.

The ROS infrastructure is composed of several nodes that handle the object detection module, the robot path planning, and the synchronization between the real manipulator and its virtual counterpart instantiated in the Unity3D application. The Unity3D application can read and write data to the ROS network by using the ROS# [53] library. In order to compute a feasible robot trajectory, an inverse kinematic (IK) algorithm runs at the ROS side using the object position as a goal position. The IK is provided by the MoveIt ROS package [54] and it takes into account the limits of the robot joints, described in the corresponding URDF file.

Since the main goal of this research concerns the NeuroTag positioning, the pickable objects are represented by 10 colored wooden blocks (see Fig. 2) equipped with a metal plate (used by the electromagnetic gripper). Hence, the object detection module uses a color thresholding approach robust enough to detect the objects' BBs (2-D BBs). Once detected, the BBs are sent to W-PC and used as input for the two positioning approaches called adaptive positioning approach (APA) and non-APA (NAPA). The former tries to place the NeuroTags in positions that are always visible to the users and close to the objects. The latter places the NeuroTags in a predefined area close to the users. It is worth noticing that both approaches compute the NeuroTag positions by determining the pixel to be centered on. Hence, both algorithms reach pixel accuracy that depends on the height at which the projector is placed. The two approaches are described in the next sections.

C. Adaptive Positioning Approach

The main goal of APA is to place the NeuroTags in positions that are always visible to the users and close to the objects. Before receiving the BB data, the operative area is represented by an M matrix composed of 500 $\times$ 280 free cells (i.e., M is composed of 500 $\times$ 280 pixels, given a projection area of 1000 $\times$ 560 mm and a resolution of 2px/mm). Since a NeuroTag is represented by a 2-D circular image of radius r, the M matrix is first modified, marking all cells lying in a border of width r as occupied. Then, after having received the BBs, M is further modified, marking as occupied all cells that are 1) inside the BBs, 2) used to display a colored border that highlights the objects, and 3) inside a border of width equal to r that surrounds the BB and the projected UI. Finally, the occlusions derived from the user's point of view are also marked as occupied cells. Specifically, considering the wooden blocks shown in Fig. 2, their four upper vertices are projected on the plane using four rays generated from the position P_u of the user's head. Then, the scan-line algorithm [55] has been used to identify the cells that are inside the 4-side polygon defined by the four projected vertices. These cells are marked as occupied cells, also considering the polygon's borders. Fig. 4 shows the available cells (in green color) and the occupied cells (in red, orange, yellow, and blue colors).

Once all the available cells have been identified, the following strategy is adopted to find the NeuroTag positions. First, the BBs are ordered depending on the number of cells contiguous to a BB border, defining a list $L_{\mathrm{BB}i}$, with i representing the ith BB. Therefore, free cells that are contiguous to borders generated by occlusions are discarded. By sorting the BBs in ascending order, the most critical BBs are prioritized with respect to the others. BBs are sequentially processed, starting from the BB with the lowest number of available cells. If $\mathrm{len}(L_{\mathrm{BB}i}) > 0$ (i.e., there is at least one solution), it is selected the cell $c \in L_{\mathrm{BB}i}$ that minimizes the distance $d = \mathrm{dist}(c, P_u)$ (the $\mathrm{len}(L_{\mathrm{BB}i}) = 1$ case is straightforward). On the contrary, if $\mathrm{len}(L_{\mathrm{BB}i}) = 0$, the following approach is proposed. The search space (i.e., the M matrix) is recursively subdivided using a quadtree data structure [56], with a resolution of four free cells per leaf node. Then, starting from the BB_i center position, the space surrounding the BB_i border is searched for free cells using a searching area with initial dimensions equal to the BB_i border. Its dimensions are increased by 1 pixel until one or multiple free cells are found, generating the $L_{\mathrm{BB}i}$ list. Finally, the cell closest to the user is selected using $d = \mathrm{dist}(c, P_u)$. Once the final cell is selected, M is modified, marking as occupied the cells covered by the NeuroTag, and the algorithm starts over with the next BB.

When all the NeuroTags have been placed, their positioning is analyzed to verify whether the BB_i-$\mathrm{NeuroTag}_i$ path is obstacle-free. Specifically, a ray is generated between the BB_i and $\mathrm{NeuroTag}_i$ centers, checking that no collision occurs and a free obstacle path is found. If all paths are obstacle-free [see Fig. 5(a)], the solution is considered to be valid. On the contrary, if one or multiple paths have collisions [see Fig. 5(b)], the solution is considered to be invalid, and the algorithm starts over from the definition of M. However, the NeuroTag diameter is reduced by 5 mm to improve the chances of finding free cells (the initial NeuroTag diameter is equal to 10 cm). If a valid solution is not found after 10 trials, the NeuroTags that do not have a free collision path are placed in the nonoperative area of

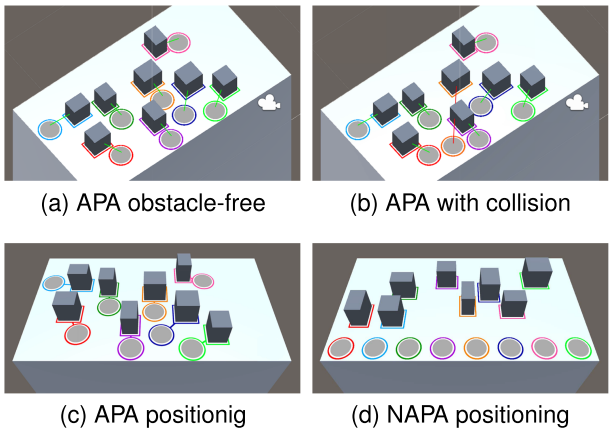

Fig. 5. (a) All NeuroTags are in the "line of sight" with their corresponding objects. (b) The orange NeuroTag is not in the "line of sight" with its object. (c) and (d) The final APA and NAPA positioning, respectively.

the projected UI (see Section III-A). Fig. 5(c) shows the final APA positioning.

To summarize, APA takes as input the position of the user's head and derives the objects' occlusions. Then, it computes the best NeuroTag positioning by considering the available free space. Hence, APA is optimized for each user and it is not affected by the viewing angle.

D. Nonadaptive Positioning Approach

NAPA places the NeuroTags in a predefined area, at the bottom of the projected UI. When W-PC receives the BBs, the positions P_i of the corresponding NeuroTags are determined using the following formula:

$$P_i = l \times (1/2n + i/n), \quad 0 \leq i < n \tag{1}$$

where l is the UI width, n the number of BBs, and P_i the fraction of l, starting from the UI left-border, respectively. The BBs are processed starting from the left-side border of the projected UI. The NeuroTag dimension does not vary, and it is equal to 7.5 cm. Fig. 5(d) shows the final positioning of eight NeuroTags.

E. Evaluation of NeuroTag Positioning and Robot Movement

Several tests have been carried out to assess 1) the minimum height that allows APA to successfully position 10 NeuroTags, 2) the APA-NAPA performances in terms of time and failures, and 3) the time required by the robot to move an object to the designated position. To assess the minimum height and the APA-NAPA performances, a virtual camera, representing the user's head, was placed at $P_u = (78, y, 0)$ cm, with P_{u_x} being the horizontal distance from the center of projection and P_{u_y} the vertical distance from the floor that varied depending on the test. For the APA minimum height test, P_{u_y} could assume values between 130 and 180 cm, whereas, for the APA-NAPA performance test, it was set equal to 180 cm.

Referring to the minimum height, the results showed that APA was able to successfully display 10 NeuroTags when the height was greater than 140 cm. This result defines the minimum height threshold to display correctly 10 NeuroTags.

TABLE I
APA AND NAPA PERFORMANCE

	Avg. time (ms)		Avg. failures (%)	
Objects	**2**	**10**	**2**	**10**
APA	43.63	648.5	0	3.9
NAPA	0.0006	0.01	0	0

TABLE II
ROBOT EXECUTION TIME CONSIDERING CLOSE AND FAR LOCATIONS

	T_{ik_s}	T_{mv_s}	T_p	T_{ik_e}	T_r	T_{mv_g}	T_{robot}
Close	<0.01	3.4	3	<0.01	2.5	3	11.9 + $\simeq$ 0.01
Far	<0.01	9.6	3	<0.01	9.4	3	25 + $\simeq$ 0.01

For the APA-NAPA performances, both approaches were tested considering 2 and 10 objects, running for 1000 iterations, and evaluating the time required to compute a solution and the percentage of failure. The number of *failures* represents the number of times APA was not able to compute a NeuroTag position and the NeuroTag was placed in the nonoperative area. The results are depicted in Table I and they show that both approaches are fast and reliable when considering two objects. When computing the NeuroTag positioning for 10 objects, the APA performance decreases as the algorithm has to partition the projection area more precisely. However, the percentage of failure remains very low and the computation time is low enough to use the system in low mobility conditions. It is worth noting that, due to its design, NAPA is inherently failure-resistant, ensuring a consistent zero-failure rate. This characteristic has been included in the table to facilitate a meaningful comparison with APA.

Considering the robot task time, the manipulator should first move from its rest pose to the object selected by the user, then to the area designated to release the objects. Hence, the robot execution time T_{robot} can be expressed as

$$T_{\text{robot}} = T_{\text{ik}_s} + T_{\text{mv}_s} + T_p + T_{\text{ik}_e} + T_{\text{mv}_g} + T_r \tag{2}$$

where T_{mv_s} represents the time required by the robot to move from the idle position to the selected object (20 cm above, vertically positioned), T_p is the picking time, T_{mv_g} the time to move from the object location to the release area, and T_r the time to release the object. T_{ik_s} and T_{ik_e} are the time required by the IK to compute a suitable path, from the rest pose to the object position and from the object position to the release area, respectively. Since $T_{\text{ik}_s}, T_{\text{mv}_s}, T_{\text{ik}_e}$, and T_{mv_g} depend on where the objects are positioned, several tests have been carried out, each involving a single object positioned both nearby and at a distance from the release area. The main outcomes are shown in Table II, in terms of average execution times expressed in seconds. As expected, the most expensive tasks are the robot movements, which highly depend on the start and goal positions. T_{ik_s} and T_{ik_e} are instead negligible, whereas the picking and releasing times are constant. Finally, the execution time doubles when the object is far from the release area. To summarize, when

considering a projector positioned at 143 cm from the table and a virtual camera placed at 78 cm from the center of projection, the minimum height required to correctly display 10 NeuroTags is equal to 141 cm. Both APA and NAPA demonstrate effective and swift positioning of up to 10 NeuroTags. Additionally, the task time of the robot is significantly influenced by movement time, with execution times ranging from 12 to 25 s. Considering the time required 1) by the algorithms to find a suitable position for the NeuroTags, 2) by the NextMind to detect the NeuroTags ($\simeq 2.5$ s), and 3) by the robot to move the selected objects, the users have to stand still for the entire duration of the interaction, without moving their bodies. Hence, the system is well suited for situations in which the user is restricted from movement but desires interaction with surrounding objects. Examples include a worker engaged in manual maintenance operations or individuals with mobility impairments.

IV. User Study

Tests have been carried out at TU Wien University to verify the effectiveness of the proposed interface by comparing the APA-NAPA approaches. Twenty-two users were asked to perform three selection tasks, of increasing difficulty. The users had to select 3 out of 3 total objects (T_1), then 4 NeuroTags out of 7 total objects (T_2), and, finally, 5 NeuroTags out of 10 total objects (T_3). The real objects (i.e., the red wooden blocks) were enumerated from 1 to 10, and, for each task, the users had to select the objects following a given ordered sequence. The sequence was written on a piece of paper placed on the working table, in front of the users. To ensure that the proposed interface and robotic system can effectively support individuals with mobility impairments, all tests were conducted with users in a specific and fixed position, emulating the constraints experienced by individuals with severe motor disabilities. The users positioned themselves at a specific distance from the center of projection (the same distance used in Section III-E) and their height was used to calibrate the APA algorithm. They were asked to stand still for the entire duration of the experiment, thus, providing a configuration similar to the one presented in Fig. 1-B.

Each user performed the tasks using both approaches; the order of the heuristics was changed after each test to avoid learnability effects. The overall procedure can be summarized as follows.

1) The user is introduced to the proposed research and the test procedure is explained in detail.
2) The user fills out a form for general user data.
3) The NextMind standard calibration is performed to assure that the system is properly calibrated based on the specific user; then, the user is able to practice with the NextMind to get accustomed to it.
4) The user performs the main task sequence (T_1, T_2, and T_3) with one of the approaches (APA or NAPA).
5) The user fills out the SUS, the NASA-TLX, and the single ease question [57] (SEQ).
6) The user repeats steps 3)–5) using the other approach.
7) The user is interviewed regarding the overall experience.

After the first calibration, the users were able to practice with the NextMind (around 10 min), selecting some NeuroTags and, thus, picking up the objects using the robotic manipulator. The positions of the real objects for T_1, T_2, and T_3 were the same for all the users (and for APA and NAPA), allowing for a fair result comparison. In addition to collecting subjective data through questionnaires, the time required to select a single NeuroTag, and the number of errors were also monitored and collected. The errors were represented by the number of times the users selected the wrong NeuroTag. With the term *errors*, false positives are considered: they represent wrong activations due to the user selecting a NeuroTag despite focusing on another one. False positives were correctly computed by comparing the given pick-up sequence and the test outcome, eventually asking the user for confirmation.

A. Results

Twenty-two students and researchers from TU Wien participated in the user tests. Before the experiment, they signed a consent form that explained the goals of the study and the anonymity of the subjects' data. Participants gave also written informed consent in accordance with the Declaration of Helsinki [58]. All collected data have been anonymized, and it is not possible to access the brain's raw data from the device. There were 15 males and 7 females, with ages between 21 and 35 years old (average $\mu = 28$, dev. std. $\sigma = 4.4$). All users were volunteers and did not receive any reward for participating in the tests. The average calibration scores for APA and NAPA were equal to 3.63 ($\sigma = 0.85$) and 3.5 ($\sigma = 0.91$), respectively. According to the NextMind standard calibration procedure, they can be both considered "Good-Medium" calibrations. The users' knowledge of BCI-AR systems and their expertise with the NextMind device was evaluated on a $(1-5)$ scale, with $1 =$ never used and $5 =$ every day usage: The users showed very little knowledge of BCIs ($\mu = 1.32$, dev. std. $\sigma = 0.76$) or of the NextMind device ($\mu = 1.41$, $\sigma = 0.94$); they had no previous experience with the NextMind at all ($\mu = 1.00$, $\sigma = 0.00$), and almost no experience with AR projected systems ($\mu = 1.23$, $\sigma = 0.42$).

To assess the SUS, NASA-TLX, and SEQ data, their distributions were initially examined through the Shapiro–Wilk test. Subsequently, for normally distributed data, the dependent *t*-test was employed as a pairwise test. In the case of nonnormal distributions, either the Wilcoxon signed-rank test or the sign test was utilized, depending on whether the difference of the distributions (DDs) exhibited a symmetrical or nonsymmetrical shape, respectively. The APA and NAPA usability results present nonuniform distributions ($p_{\text{APA}} = 0.049$, $p_{\text{NAPA}} = 0.039$) and DDs are not symmetrical. The median values for APA and NAPA are $m_{\text{APA}} = 86.25$ and $m_{\text{NAPA}} = 82.5$, respectively, and the sign test shows statistically significant differences between the two groups ($p_{\text{sign}} = 0.012$). Fig. 6-left shows the SUS results. The APA and NAPA NASA-TLX outcomes (the weighted dimensions) have been analyzed by comparing the six dimensions. More than one dimension presents nonuniform distributions, and all the DDs are not symmetrical; thus, the sign test has been used. However, it was not possible to determine statistically significant differences among the NASA-TLX dimensions. The resulting average task index loads are, respectively, $\mu_{\text{APA}} = 25.49$, $\sigma = 16.81$ and $\mu_{\text{NAPA}} = 30.64$, $\sigma = 19.213$. Finally, the SEQ data

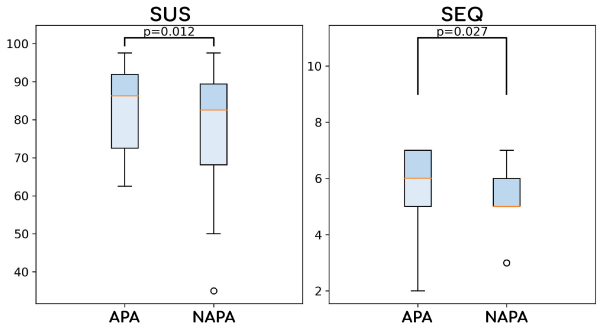

Fig. 6. SUS and SEQ results for APA and NAPA. The median values are represented with orange lines.

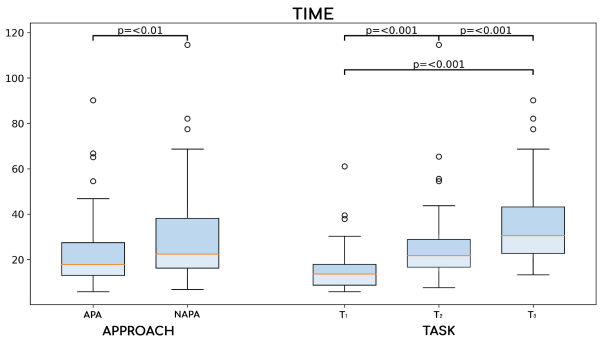

Fig. 7. Time results. The orange lines represent the mean values.

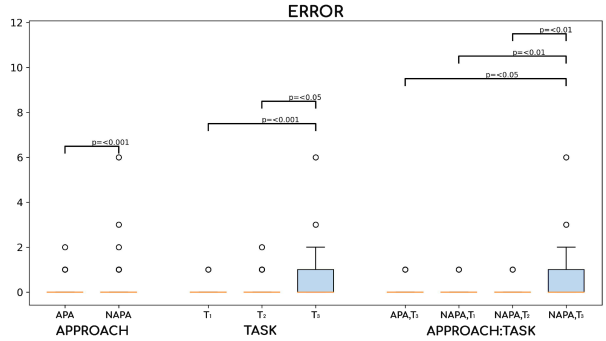

Fig. 8. Error results. The orange lines represent the mean values.

(Fig. 6-right) present nonuniform distributions ($p_{\text{APA}} = 0.001$, $p_{\text{NAPA}} = 0.005$) and nonsymmetrical DDs. The APA and NAPA median values are $m_{\text{APA}} = 6$ and $m_{\text{NAPA}} = 5$, respectively, showing statistically significant differences between the two algorithms ($p_{\text{sign}} = 0.027$).

Concerning the time and error results, given an ith task T_i (with $1 \leq i \leq 3$), the data related to each single NeuroTag have been summed up, computing the total time and the total number of errors of T_i. Time outcomes present a nonuniform distribution across all levels of the independent variables. The nonparametric aligned rank transform (ART) test [59] was performed to evaluate the effect of *Task* and *Approach* on the time required to complete the tasks. The results revealed that there was a statistically significant effect of *Approach* ($F(1, 105) = 10.172, p < 0.01, \eta_p^2 = 0.088$) and *Task* ($F(2, 105) = 77.951, p < 0.001, \eta_p^2 = 0.597$), but no interaction. Following the work in [60], multifactor contrast tests have been carried out as post-hoc analysis, showing statistically significant differences between APA-NAPA and among the three different tasks $T_1 - T_2$, $T_1 - T_3$, and $T_2 - T_3$ (Fig. 7 shows the time results). Similarly to the time data, the error outcomes show nonuniform distributions and the ART test revealed that there was a statistically significant interaction between the effects of *Approach* and *Task* ($F(2, 105) = 13.93, p < 0.001, \eta_p^2 = 0.209$) as well as a statistically significant effect of *Task* ($F(2, 105) = 8.98, p < 0.001, \eta_p^2 = 0.146$) and *Approach* ($F(1, 105) = 34.31, p < 0.001, \eta_p^2 = 0.246$). Contrast tests indicate there were statistically significant differences between APA-NAPA as well as between T_1–T_3 and T_2–T_3. Concerning the interaction, significant differences were found between T_3^{APA}–T_3^{NAPA}, T_1^{NAPA}–T_3^{NAPA}, and T_2^{NAPA}–T_3^{NAPA}. Fig. 8 shows the error results.

Since the users did not complain about being unable to perform a selection while performing the tasks, the number of false negatives has been computed equal to zero. The precision P for the ith task was computed according to the formula $P_{\text{alg}}^i = \text{True\_Positives}/(\text{True\_Positives} + \text{False\_Positives})$. The final APA precision outcomes were $P_{\text{APA}}^1 = 1$, $P_{\text{APA}}^2 = 0.95$, and $P_{\text{APA}}^3 = 0.97$, whereas the NAPA configuration obtained $P_{\text{NAPA}}^1 = 0.94$, $P_{\text{NAPA}}^2 = 0.96$, and $P_{\text{NAPA}}^3 = 0.75$.

B. Discussion

Although the users were not allowed to move their bodies, both positioning approaches were deemed usable and effective in interacting with the flickering textures. The high levels of usability (>80) indicate that motion-restricted situations do not affect the activation of the NeuroTags. This is an important result because it suggests that they might be effectively employed for helping people with severe motor impairments. Moreover, the APA usability outcomes are slightly higher than the NAPA ones, suggesting that the distance object-NeuroTag should be kept as small as possible. This result is in line with the state of the art related to AR label positioning. Similar to the usability outcomes, the SEQ data indicate that both approaches allowed the users to easily interact with the NeuroTags. However, the users perceived APA as simpler and more straightforward than NAPA. These results support both *R1a* and *R1b*, confirming the great effectiveness and usability of the positioning solutions, with APA being significantly more usable than the nonadaptive one.

Referring to *R2*, the NASA-TLX results show that there are no significant differences between the two approaches. This result might be due to the relatively small sample size and, thus, to the lack of statistical power. Generally, the workload (i.e., the task index) is low independently of the chosen positional strategy, with the mental workload having more influence than the other dimensions, confirming the results presented in [37].

Concerning the time outcomes, two interesting insights can be derived. First, it is unsurprising that the time needed to complete tasks rises with an increasing number of NeuroTags. Second,

APA consistently enables users to select the NeuroTags faster than NAPA across all tasks. Consequently, the positional approach remains independent of the specific task, and the adaptive approach consistently reduces interaction time, regardless of the number of NeuroTags considered. The error outcomes present a similar trend. APA is generally less prone to errors than the NAPA and the number of errors tends to increase as the number of objects increases, especially when using NAPA for T_3. Since in T_3, the NeuroTags are positioned quite close to each other, some false positive errors were detected, indicating that the users selected the wrong NeuroTags (when using NAPA). This result confirms the importance of the adopted positioning approach when interacting with the flickering textures.

Referring to *R3*, the proposed hardware/software architecture and the positioning approaches can be effectively used to interact with the NeuroTags for pick-and-place scenarios. APA and NAPA were really precise (worst precision $P^3_{\text{NAPA}} = 0.75$) with a few numbers of false positives. In addition, the users were fast in selecting the NeuroTags. The most complex task (T_3) required on average 40 s (NAPA) and 20 s (APA) to be completed, indicating 3 s and 6 s per NeuroTag, respectively. Finally, the exceptionally low percentage of failures and the APA fall-back mechanism, which positions NeuroTags in the nonoperative area in case of failures, ensure robustness and reliability.

The analysis of the users' comments supports these findings. Specifically, all users appreciated both approaches, with more positive comments for APA. As an example, the users stated "I prefer APA because the tag is connected to the object. NAPA stresses a bit because of fear of looking at the tag next to the right one" or "I felt more confident with APA because I had an easier time working with it, and it seemed more intuitive with the focus point being near the objects to be moved," showing that reducing the NeuroTag-object distance greatly improves interface usability. However, a few users reported that they preferred NAPA because the NeuroTags were always positioned in the same area, close to the user. It is important to recognize that while SSVEP-BCIs typically necessitate minimal training, users may encounter challenges in obtaining optimal performances. Thus, a training phase is advisable. The raw data and a video are available online at the link[1] and in the supplementary materials.

C. Limitations

First, despite these promising results, it is uncertain whether the proposed system will be equally effective for motor-impaired individuals who rely on wheelchairs for mobility. To address the needs of this user group, the system should be made portable and adaptable to accommodate a different robotic manipulator, such as the Kinova Jaco [61]. Second, although the users were explicitly instructed to remain stationary throughout the experiment, with no reported collisions with the manipulator, it is advisable to implement additional safety measures to prevent any potential human–robot collision. For example, this could involve tracking the upper limbs, such as hand movements, or considering the use of a collaborative robot instead of an industrial one. Finally, the current implementation lacks the head's position tracking, which means that the position of the projected AR contents does not dynamically update as the user's head changes position.

[1][Online]. Available: https://doi.org/10.48436/s881n-vzs78

V. Conclusion and Future Works

This article presents a projected AR system combined with a BCI for robotic pick-and-place tasks in motion-restricted conditions. By using the NextMind device, the users can command a Staubli TX2-60 manipulator to pick up objects positioned on a flat surface. Two different approaches to position the NeuroTags have been proposed and assessed in a user study under conditions similar to those of mobility-impaired people. The experimental tests involving users showed that the proposed interface and its two NeuroTag positioning approaches are highly appreciated and effective in allowing users to easily and reliably select objects for pick-and-place tasks. The main outcomes show that APA and NAPA can be effectively used to place the NeuroTags, with APA being more usable, less time-demanding, and error-prone than NAPA. The system's robust performance, as demonstrated by the user tests, underscores the potential of the proposed system to enhance mobility and autonomy for individuals with motor impairments. Future improvements will focus on replacing the current object detection module with a relevant and recent object recognition framework; this upgrade will help in considering alternative use cases, such as picking objects of the user's choice. In addition, further evaluations will be conducted to determine whether the BCI device can achieve information transfer rates comparable to the work in [47] when the visual stimuli are projected on real surfaces. Finally, in order to further confirm these preliminary results, people with severe motor disabilities should be involved in the user tests, verifying the usability and easiness of use of the NextMind device and assessing their preferences toward the APA or NAPA approaches.

References

[1] T. B. Sheridan, "Human–robot interaction: Status and challenges," *Hum. Factors*, vol. 58, no. 4, pp. 525–532, 2016.
[2] S. Li and X. Zhang, "Implicit intention communication in human–robot interaction through visual behavior studies," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 4, pp. 437–448, Aug. 2017.
[3] A. Bablani, D. R. Edla, D. Tripathi, and R. Cheruku, "Survey on brain-computer interface: An emerging computational intelligence paradigm," *ACM Comput. Surv.*, vol. 52, no. 1, pp. 1–32, 2019.
[4] D. Yadav, S. Yadav, and K. Veer, "A comprehensive assessment of brain computer interfaces: Recent trends and challenges," *J. Neurosci. Methods*, vol. 346, 2020, Art. no. 108918.
[5] F. Cincotti et al., "Non-invasive brain–computer interface system: Towards its application as assistive technology," *Brain Res. Bull.*, vol. 75, no. 6, pp. 796–803, 2008.
[6] F. Beverina et al., "User adaptive BCIs: SSVEP and P300 based interfaces," *PsychNol. J.*, vol. 1, no. 4, pp. 331–354, 2003.
[7] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 355–385, 1997.
[8] Snapchat, "NextMind." Accessed: Jan. 22, 2024. [Online]. Available: https://github.com/Snapchat/NextMind
[9] J. Brooke et al., "SUS—A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.
[10] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Advances in Psychology*, vol. 52. Amsterdam, The Netherlands: Elsevier, 1988, pp. 139–183.

[11] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation," *Physiol. Rev.*, vol. 97, no. 2, pp. 767–837, 2017.

[12] X. Chen, B. Zhao, Y. Wang, and X. Gao, "Combination of high-frequency SSVEP-based BCI and computer vision for controlling a robotic arm," *J. Neural Eng.*, vol. 16, no. 2, 2019, Art. no. 026012.

[13] M. Alimardani and K. Hiraki, "Passive brain-computer interfaces for enhanced human-robot interaction," *Front. Robot. AI*, vol. 7, pp. 125–137, 2020.

[14] P. Stegman, C. S. Crawford, M. Andujar, A. Nijholt, and J. E. Gilbert, "Brain–computer interface software: A review and discussion," *IEEE Trans. Human-Mach. Syst.*, vol. 50, no. 2, pp. 101–115, Apr. 2020.

[15] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 145–147, Jun. 2003.

[16] K. J. Miller, D. Hermes, and N. P. Staff, "The current state of electrocorticography-based brain–computer interfaces," *Neurosurgical Focus*, vol. 49, no. 1, p. E2, 2020.

[17] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, pp. 41–56, 2008.

[18] K.-T. Kim, H.-I. Suk, and S.-W. Lee, "Commanding a brain-controlled wheelchair using steady-state somatosensory evoked potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 3, pp. 654–665, Mar. 2018.

[19] A. L. Benabid et al., "An exoskeleton controlled by an epidural wireless brain–machine interface in a tetraplegic patient: A proof-of-concept demonstration," *Lancet Neurol.*, vol. 18, no. 12, pp. 1112–1122, 2019.

[20] S. Kim, S. Lee, H. Kang, S. Kim, and M. Ahn, "P300 brain–computer interface-based drone control in virtual and augmented reality," *Sensors*, vol. 21, no. 17, 2021, Art. no. 5765.

[21] D. Irimia, R. Ortner, G. Krausz, C. Guger, and M. Poboroniuc, "BCI application in robotics control," *IFAC Proc. Volumes*, vol. 45, no. 6, pp. 1869–1874, 2012.

[22] B. Peters et al., "SSVEP BCI and eye tracking use by individuals with late-stage ALS and visual impairments," *Front. Hum. Neurosci.*, vol. 14, 2020, Art. no. 595890.

[23] M. Minamoto, Y. Suzuki, T. Kanno, and K. Kawashima, "Effect of robot operation by a camera with the eye tracking control," in *Proc. IEEE Int. Conf. Mechatronics Automat.*, 2017, pp. 1983–1988.

[24] V. K. Sharma, L. Murthy, and P. Biswas, "Comparing two safe distance maintenance algorithms for a gaze-controlled HRI involving users with SSMI," *ACM Trans. Accessible Comput.*, vol. 15, no. 3, pp. 1–23, 2022.

[25] L. Scalera, S. Seriani, P. Gallina, M. Lentini, and A. Gasparetto, "Human–robot interaction through eye tracking for artistic drawing," *Robotics*, vol. 10, no. 2, pp. 54–58, 2021.

[26] C.-Y. Kuo, C.-C. Huang, C.-H. Tsai, Y.-S. Shi, and S. Smith, "Development of an immersive SLAM-based VR system for teleoperation of a mobile manipulator in an unknown environment," *Comput. Ind.*, vol. 132, 2021, Art. no. 103502.

[27] J. I. Lipton, A. J. Fay, and D. Rus, "Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing," *IEEE Robot. Automat. Lett.*, vol. 3, no. 1, pp. 179–186, Jan. 2018.

[28] F. De Pace, G. Gorjup, H. Bai, A. Sanna, M. Liarokapis, and M. Billinghurst, "Leveraging enhanced virtual reality methods and environments for efficient, intuitive, and immersive teleoperation of robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 12967–12973.

[29] R. Suzuki, A. Karim, T. Xia, H. Hedayati, and N. Marquardt, "Augmented reality and robotics: A survey and taxonomy for AR-enhanced human-robot interaction and robotic interfaces," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2022, pp. 1–33.

[30] A. Lenhardt and H. Ritter, "An augmented-reality based brain-computer interface for robot control," in *Proc. 17th Int. Conf. Neural Inf. Process. Models Appl.*, 2010, pp. 58–65.

[31] J. Faller, R. Leeb, G. Pfurtscheller, and R. Scherer, "Avatar navigation in virtual and augmented reality environments using an SSVEP BCI," in *Proc. Int. Conf. Appl. Bionics Biomech.*, 2010, pp. 1–4.

[32] H. Zeng et al., "Closed-loop hybrid gaze brain-machine interface based robotic arm control with augmented reality feedback," *Front. Neurorobot.*, vol. 11, pp. 60–73, 2017.

[33] H. Si-Mohammed et al., "Towards BCI-based interfaces for augmented reality: Feasibility, design and evaluation," *IEEE Trans. Visualization Comput. Graph.*, vol. 26, no. 3, pp. 1608–1621, Mar. 2020.

[34] C. Gorman and Y.-K. Wang, "A closed-loop AR-based BCI for real-world system control," in *Proc. IEEE Symp. Ser. Comput. Intell.*, 2021, pp. 1–7.

[35] S. N. Faisal et al., "Noninvasive sensors for brain–machine interfaces based on micropatterned epitaxial graphene," *ACS Appl. Nano Mater.*, vol. 6, no. 7, pp. 5440–5447, 2023.

[36] L. Angrisani, P. Arpaia, E. De Benedetto, L. Duraccio, F. L. Regio, and A. Tedesco, "Wearable brain-computer interfaces based on steady-state visually evoked potentials and augmented reality: A review," *IEEE Sensors J.*, vol. 23, no. 15, pp. 16501–16514, Aug. 2023.

[37] S. Da Col, E. Kim, and A. Sanna, "Human performance and mental workload in augmented reality: Brain computer interface advantages over gestures," *Brain- Comput. Interfaces*, vol. 9, no. 4, pp. 211–225, 2022.

[38] A. Sanna, F. Manuri, J. Fiorenza, and F. De Pace, "BARI: An affordable brain-augmented reality interface to support human–robot collaboration in assembly tasks," *Information*, vol. 13, no. 10, pp. 460–474, 2022.

[39] S. A. Hirsch, "An algorithm for automatic name placement around point data," *Amer. Cartographer*, vol. 9, no. 1, pp. 5–17, 1982.

[40] H. Kato and M. Billinghurst, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," in *Proc. 2nd IEEE ACM Int. Workshop Augmented Reality*, 1999, pp. 85–94.

[41] D. Cotting and M. Gross, "Interactive environment-aware display bubbles," in *Proc. 19th Annu. ACM Symp. User Interface Softw. Technol.*, 2006, pp. 245–254.

[42] T. Siriborvornratanakul and M. Sugimoto, "Clutter-aware adaptive projection inside a dynamic environment," in *Proc. ACM Symp. Virtual Reality Softw. Technol.*, 2008, pp. 241–242.

[43] J. Riemann, M. Khalilbeigi, M. Schmitz, S. Döweling, F. Müller, and M. Mühlhäuser, "FreeTop: Finding free spots for projective augmentation," in *Proc. CHI Conf. Extended Abstr. Hum. Factors Comput. Syst.*, 2016, pp. 1598–1606.

[44] D. Iwai, T. Yabiki, and K. Sato, "View management of projected labels on nonplanar and textured surfaces," *IEEE Trans. Visualization Comput. Graph.*, vol. 19, no. 8, pp. 1415–1424, Aug. 2013.

[45] K. Ichihashi and K. Fujinami, "Estimating visibility of annotations for view management in spatial augmented reality based on machine-learning techniques," *Sensors*, vol. 19, no. 4, pp. 939–967, 2019.

[46] F. Manuri, A. Sanna, M. Bosco, and F. De Pace, "A comparison of three different NeuroTag visualization media: Brain visual stimuli by monitor, augmented and virtual reality devices," in *Proc. Hum. Interact. Emerg. Technol.: Artif. Intell. Future Appl. AHFE Int.*, 2020, pp. 150–158.

[47] B. J. Hou et al., "Feasibility of a device for gaze interaction by visually-evoked brain signals," in *Proc. Symp. Eye Tracking Res. Appl.*, 2022, pp. 1–7.

[48] S. Griffioen, "Maximizing color differences in metro maps," Ph.D. dissertation, Faculty Sci. Eng., Univ. Groningen, 2014. [Online]. Available: https://fse.studenttheses.ub.rug.nl/11812/

[49] Staubli, "TX2-60." Accessed: Jan. 22, 2024. [Online]. Available: https://www.staubli.com/tw/en/robotics/products/industrial-robots/tx2-60.html

[50] LG, "HF80LA." Accessed: Jan. 22, 2024. [Online]. Available: https://www.lg.com/us/support/product/lg-HF80LA.AUS

[51] Unity, "Unity3D." Accessed: Jan. 22, 2024. [Online]. Available: https://unity.com/pages/unity-pro-buy-now?gclsrc=ds&gclsrc=ds

[52] ROS, "ROS melodic morenia." Accessed: Jan. 22, 2024. [Online]. Available: http://wiki.ros.org/melodic

[53] Siemens, "ROS#." Accessed: Jan. 22, 2024. [Online]. Available: https://github.com/siemens/ros-sharp

[54] MoveIt, Accessed: Jan. 22, 2024. [Online]. Available: https://moveit.ros.org

[55] K. Lin, Y. Chen, M. Qiu, M. Zeng, and W. Huang, "SLGC: A fast point-in-area algorithm based on scan-line algorithm and grid compression," in *Proc. IEEE 11th Int. Conf. Comput. Sci. Educ.*, 2016, pp. 352–356.

[56] R. A. Finkel and J. L. Bentley, "Quad trees a data structure for retrieval on composite keys," *Acta Informatica*, vol. 4, pp. 1–9, 1974.

[57] T. Rotolo, "The single ease question." Accessed: Jan. 22, 2024. [Online]. Available: https://trymata.com/blog/2015/03/04/measuring-task-usability-the-single-ease-question/

[58] "Declaration of Helsinki." Accessed: Jan. 22, 2024. [Online]. Available: https://tinyurl.com/7s6djxpn

[59] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2011, pp. 143–146.

[60] L. A. Elkin, M. Kay, J. J. Higgins, and J. O. Wobbrock, "An aligned rank transform procedure for multifactor contrast tests," in *Proc. 34th Annu. ACM Symp. User Interface Softw. Technol.*, 2021, pp. 754–768.

[61] "Kinova Jaco." Accessed: May 8, 2023. [Online]. Available: https://assistive.kinovarobotics.com/product/jaco-robotic-arm