Combination of Augmented Reality Based Brain-Computer Interface and Computer Vision for High-Level Control of a Robotic Arm

Xiaogang Chen, *Member, IEEE*, Xiaoshan Huang, Yijun Wang, *Member, IEEE*, and Xiaorong Gao, *Member, IEEE*

***Abstract*— Recent advances in robotics, neuroscience, and signal processing make it possible to operate a robot through electroencephalography (EEG)-based brain-computer interface (BCI). Although some successful attempts have been made in recent years, the practicality of the entire system still has much room for improvement. The present study designed and realized a robotic arm control system by combing augmented reality (AR), computer vision, and steady-state visual evoked potential (SSVEP)-BCI. AR environment was implemented by a Microsoft HoloLens. Flickering stimuli for eliciting SSVEPs were presented on the HoloLens, which allowed users to see both the robotic arm and the user interface of the BCI. Thus users did not need to switch attention between the visual stimulator and the robotic arm. A four-command SSVEP-BCI was built for users to choose the specific object to be operated by the robotic arm. Once an object was selected, the computer vision would provide the location and color of the object in the workspace. Subsequently, the object was autonomously picked up and placed by the robotic arm. According to the online results obtained from twelve participants, the mean classification accuracy of the proposed system was 93.96 ± 5.05%. Moreover, all subjects could utilize the proposed system to successfully pick and place objects in a specific order. These results demonstrated the potential of combining AR-BCI and computer vision to control robotic arms, which is expected to further promote the practicality of BCI-controlled robots.**

***Index Terms*— EEG, SSVEP, BCI, robotic arm, AR.**

I. Introduction

PARALYSIS following spinal cord injury (SCI), amyotrophic lateral sclerosis (ALS), and other injuries or neurologic disorders will make individuals unable to perform reaching and grasping movements required for various daily activities. In order to restore movement control for individuals with paralysis, several interaction techniques have attempted in recent studies. However, many existing techniques depend on the user's residual motor control. Thus the usage of these techniques is more challenging for severely motor-impaired individuals. Brain-computer interface (BCI) aims to build a new control channel between human and external devices by decoding human brain signals, and can be used to improve basic activities of daily living for severely motor-impaired individuals [1]. BCIs have been verified to be successfully used to control virtual computer cursor [2] and keyboard [3]–[6], physical objects, such as wheelchair [7], and various assistive devices [8], [9].

Recently, using BCIs to control robots have become an indispensable part of intimately blending neuroscience and robots. Invasive BCI-controlled robotic arm has shown promise in tetraplegic individuals [10], [11]. They can successfully utilize invasive BCI to control a robotic arm to carry out reach-to-grasp movements. Despite these achievements, technical difficulties and surgical risks remain major obstacles to limit invasive BCIs for daily use at home. In contrast, non-invasive BCIs are easier to be accepted due to their simplicity and safety. Currently, the most commonly used brain signal in non-invasive BCIs is electroencephalography (EEG). Several studies have demonstrated that EEG-based BCIs can provide precise control for robotic arms [12]–[16]. Although EEG-based BCI controlled robotic arms have achieved some promising results in the abovementioned studies, the control accuracy and interaction efficiency of EEG-based BCI controlled robotic arms still need to be further improved. In the existing EEG-based BCI controlled robotic arms, BCIs

Manuscript received May 15, 2020; revised October 12, 2020; accepted November 11, 2020. Date of publication November 16, 2020; date of current version January 29, 2021. This work was supported in part by the Key Research and Development Program of Guangdong Province under Grant 2018B030339001; in part by the Strategic Priority Research Program of Chinese Academy of Science under Grant XDB32040200; in part by the National Key Research and Development Program of China under Grant 2017YFB1002505; in part by the Fundamental Research Funds for the Central Universities under Grant 3332019015 and Grant 3332018191; in part by the National Natural Science Foundation of China under Grant 61431007 and Grant 61603416; in part by CAMS Innovation Fund for Medical Sciences under Grant 2016-I2M-3-023; and in part by the Young Elite Scientists Sponsorship Program, CAST, under Grant 2015QNRC001. *(Corresponding authors: Xiaogang Chen; Yijun Wang.)*

Xiaogang Chen is with the Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin 300192, China (e-mail: chenxg@bme.cams.cn).

Xiaoshan Huang and Xiaorong Gao are with the Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China (e-mail: xiaoshanhuang@gmail.com; gxr-dea@tsinghua.edu.cn).

Yijun Wang is with the State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China (e-mail: wangyj@semi.ac.cn).

Digital Object Identifier 10.1109/TNSRE.2020.3038209

usually adopt three control strategies (i.e., direct control, high-level control, and shared control) to act on robotic arms [13], [15]–[20]. In direct control, the BCI directly controls the movement of the robotic arm through simple navigation commands, such as 'move forward' or 'move up'. The user need to frequently send specific commands through the BCI to control every steps of the process. In high-level control, the user can control the robotic arm through the BCI by issuing high-level commands such as choosing red cube. The BCI-controlled robotic arm needs to be equipped with additional intelligence so that the robotic arm autonomously executes the process to achieve that goal without the user sending specific commands. In shared control, both the user (using the BCI) and the intelligent assistance (such as object recognition and localization system) share the control over the robotic arm. A feasible solution to realize shared control is that the user can flexibly switch between direct control mode and high-level control mode. To date, the majority of EEG-based BCI controlled robotic arms still adopt direct control strategy [13], [16]. Recent works have begun to attempt to investigate high-level and shared control strategies to allow users to control robotic arms to perform more complex tasks [15], [17]–[20]. Compared with the direct control strategy, the high-level and shared control strategies can effectively reduce the users' mental load and are more favored by users. Although several of these efforts employ high-level control strategy, the implementation of these systems is not the same. For example, Waytowich *et al.* (2010) used a P300-based BCI to control a robotic arm [18]. The user only needed to utilize the BCI to select the position of the object, and the robotic arm would automatically complete pick and place operations. Lampe *et al.* (2014) combined motor imagery-based BCI with camera-based visual servoing control for manipulating a robotic arm [19]. Users could select the object to be grasped through the motor imagery-based BCI, and the robotic arm autonomously performed the actual grasping actions via reinforcement learning. The work in [20] combined hybrid Gaze-BCI with image processing to control a robotic arm. The hybrid BCI were built by combing motor imagery-based BCI with eye tacking. Image processing was response for segmenting the potential objects. The user could selected the segmented objects through eye tracking. BCI was used to confirm the selected object or initiate a control command. Then the robotic arm automatically completed the reaching, delivering and releasing processes. While the user needed to manually control the grasping and lifting processes. Although some successful attempts have been made in high-level control approach [18]–[20], BCI performance of these studies was still low and needed to further improved.

BCI performance determines performance of BCI-controlled robotic arms to a large extent. Thus high-performance BCIs are more suitable for building BCI-controlled robotic arms. Among these existing EEG-based BCIs, BCI based on steady-state visual evoked potential (SSVEP) is gaining more and more attention due to its less training and high performance [4], [21], [22]. Therefore, SSVEP-based BCIs promise to build high-performance BCI-controlled robotic arms. Recently, SSVEP-BCIs have been successfully applied for the control of robotic arms [13], [15], [16]. However, most of the existing studies adopted computer monitors as visual stimulators for inducing SSVEPs. Users required to shift attention between the computer monitor and the robotic arm, which might lead to a poor user experience. Augmented reality (AR) technology provides a possible solution to this issue. Since AR technology involves in real-time integration of virtual objects into a real environment, which makes virtual and real content coexist in the same space [23]–[25]. Currently, AR environment can be realized through a portable and wearable glass (i.e., Microsoft HoloLens). Combing AR and BCI can improve the flexibility and portability of SSVEP-based BCIs. The field of AR-based SSVEP-BCI is still in its infancy, only a few studies have been conducted to date [16], [26] and further research is needed to improve our understanding of AR-based SSVEP-BCI.

The present study implemented a novel robotic arm control system by combing computer vision, AR technology, and SSVEP-BCI. To the best of our knowledge, this study was the first time to combine SSVEP-BCI, AR technology, and computer vision for realizing the control of a robotic arm. In this study, a Microsoft HoloLens was used to build AR environment and present flickering stimuli for eliciting SSVEPs, in which users could see both the robotic arm and the user interface of the BCI. A four-command SSVEP-BCI was utilized for users to choose the specific object to be operated by the robotic arm. Once an object was selected, the computer vision would provide the position and color of the object in the workspace. Subsequently, the robotic arm autonomously picked up and placed the object based on information obtained from the BCI and the computer vision. The feasibility of the proposed system was verified by offline and online experiments.

II. Methods

A. Subjects

Thirteen healthy volunteers (7 males and 6 females), who were aged between 19 to 31 years old, took part in the present study. All individuals reported normal vision or corrected-to-normal vision. The offline experiment was carried out on ten subjects, while the online experiment was performed on twelve subjects. Nine subjects participated in these two experiments. The Institution Review Board of Tsinghua University approved this study (No. 20180041). All subjects signed informed consent. During the experiments, subjects wore an EEG cap and a HoloLens, and sat on a chair, facing the robotic arm.

B. EEG Data Acquisition

EEG data were acquired through a Neuracle EEG amplifier with a sampling rate of 1000 Hz. The reference electrode and the ground electrode were placed on the vertex and the forehead respectively. Only 9-channel (Pz, POz, Oz, PO5, PO3, O1, PO4, PO6, and O2) EEG signals were recorded in this study. The impedance of each electrode is lower than 10 kΩ. The experiments were conducted in a normal office room without any electromagnetic shielding.

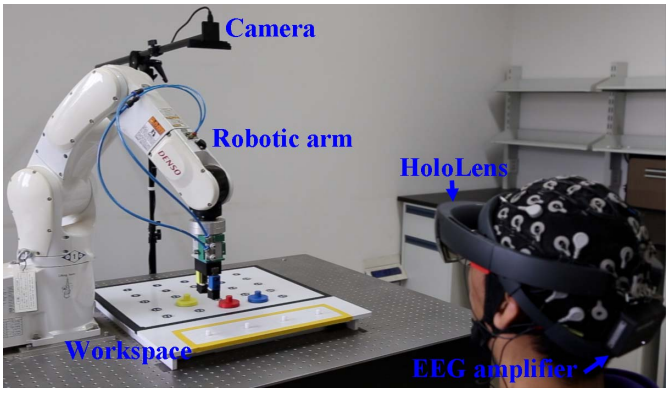

Fig. 1. Experimental setup of the proposed AR-BCI controlled robotic arm.

C. System Description

The proposed AR-BCI controlled robotic arm mainly consisted of robotic arm subsystem, AR display subsystem, SSVEP-BCI subsystem, and computer vision subsystem. Three plastic objects with different color (i.e., red, blue, and yellow) were placed on the workspace. Fig. 1 illustrates the experimental setup of the proposed AR-BCI controlled robotic arm. The AR subsystem was implemented by a HoloLens (Microsoft, US), which was responsible for eliciting SSVEPs. The SSVEP-BCI subsystem was utilized to choose the object to be manipulated by the robotic arm. The computer vision subsystem provided the color and location information for the three plastic objects in the workspace. The data processing of the SSVEP-BCI subsystem and computer vision subsystem were done on one computer. The two subsystems worked independently of each other. The outputs of the computer vision subsystem and the SSVEP-BCI subsystem would form a robot command through the computer and be transmitted to the robot controller through TCP/IP. The robotic arm would automatically operate according to the received robot command. Fig. 2 shows the system architecture of the proposed system. As shown in Fig. 2, the proposed AR- BCI controlled robotic arm mainly included four subsystems: the robotic arm, the AR display, the SSVEP- BCI, and the computer vision.

A 6-axis Denso VS-060 robotic arm was adopted in this study. Meanwhile, we also equipped the robotic arm with a parallel gripper DHPS-25-A from Festo. The workspace was divided into two areas, one was a black area and the other was a yellow area. Twenty-five blue circles were in the black area, and these 25 circles represented 25 different pre-taught locations. Three plastic objects with different color (i.e., red, blue, and yellow) were randomly arranged in three of the 25 different pre-taught locations by the robotic arm. Participants were instructed to use the proposed system to pick up the three objects from the black area and move them to desired locations (i.e., 3 different locations) in the yellow area. The robotic arm placed the three objects in three different positions in the yellow area from left to right in the order of picking objects. The three-dimensional coordinates of the 25 different positions in the black area and the 3 different positions in the yellow area were pre-determined by the robotic arm.

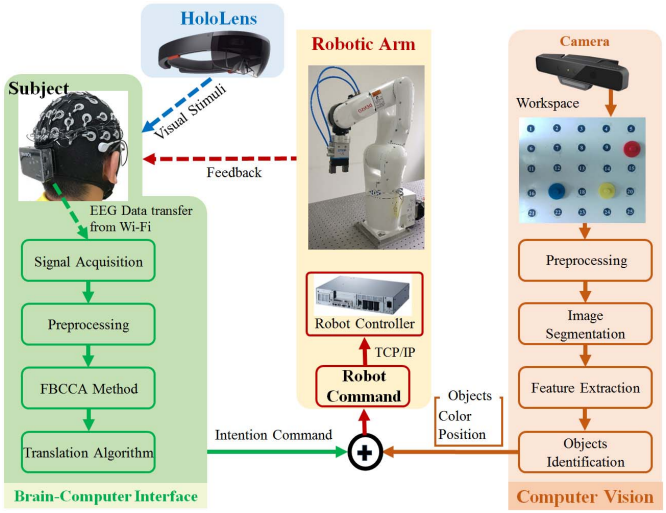

Fig. 2. System architecture of the proposed AR-BCI controlled robotic arm.

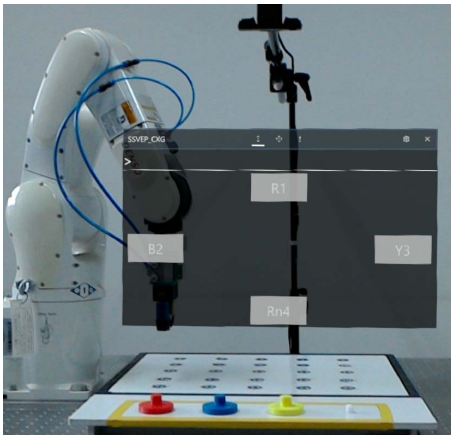

Fig. 3. First-person perspective seen from HoloLens.

In the SSVEP-BCI subsystem, a Neuracle EEG amplifier was adopted to acquire EEG signals. A four-command SSVEP-BCI was designed for subjects to send intention commands. A synchronous mode was adopted in the proposed BCI for sending commands. Flickering stimuli for eliciting SSVEPs were presented on the Microsoft HoloLens. The HoloLens's refresh rate was 60 Hz. Visual stimuli were displayed on-screen as a flickering white transparent mask. The robotic arm was situated behind the four targets. The visual stimuli in the HoloLens were developed with Unity 3D Engine (C#). Fig. 3 shows the first-person view as displayed in the HoloLens. As shown in Fig. 3, the proposed BCI user interface included four commands. The four commands were flickering at 9 Hz, 11 Hz, 10 Hz, and 12 Hz (see Fig. 4), which represented choosing red cube (i.e., "R1"), choosing blue cube (i.e., "B2"), choosing yellow cube (i.e., "Y3"), and undoing the last operation (i.e., "Rn4") (see Fig. 3). As shown in Fig. 2, the output of the FBCCA method was the recognition result of user's gazed stimulation frequency using the recorded EEG signals. According to the command corresponding to each stimulation frequency (see Fig. 3 and

Fig. 4. Stimulation frequency for each target.

Fig. 4), the translation algorithm converted the output of the FBCCA into the command executed by the robotic arm.

The computer vision subsystem used an Intel RealSense SR300 camera to acquire images of the workspace. The typical filter for homomorphic filtering process was first used to correct illumination of the original image and enhance the original image's details [27]. Second, the homomorphic filtered image in RGB was converted to HSV color space. In order to recognize the color information of these objects, we defined hue, saturation and value ranges for the three colors (i.e., red, blue, and yellow) respectively. For red color, the HSV components values were defined in the following ranges: $-250°$ $\sim 50°$ for hue, 0.05$\sim$1 for saturation, and 0$\sim$1 for value. The HSV components values for blue color were quantized to the following ranges: $180° \sim 280°$ for hue, 0.555$\sim$1 for saturation, and 0.2$\sim$1 for value. For yellow color, the HSV components values were limited in the following ranges: $30° \sim 90°$ for hue, 0.1$\sim$1 for saturation, and 0.2$\sim$1 for value. In this stage we could obtain three intermediary images corresponding to red, blue, and yellow colors. In each intermediary image, all pixels classified as specific color were fixed to value 255, and the other pixels were set to 0. Third, a median filter was applied to each intermediary image. The median filtered image was then converted to a binary image. Fourth, we removed from the binary image all connected objects that had fewer than 3000 pixels. Last, the image was evenly divided into 25 equal sub-images. For each sub-image, the boundary of hole was calculated for determining the object location. The outputs of the computer vision subsystem contained the color and location of the three objects. The detailed three-dimensional coordinates of the object could be determined by matching the position information recognized by the computer vision with the 25 different pre-taught locations. According to the detailed three-dimensional coordinates of the object, the robotic arm would automatically pick up the object and place it in the desired location.

D. Offline Experiment

The aim of the offline experiment was to mainly evaluate and optimize the performance of the proposed computer vision and SSVEP-BCI. During the offline experiment, the robotic arm did not start and was motionless. The offline BCI experiment carried out five blocks and in each block there were 32 trials, for a total of 160 trials in the offline BCI experiment. According to Fig. 3 and Fig. 4, there were four targets in the proposed BCI. For each block, there were 8 trials for each target and the order of targets was randomized. Therefore, a total of 40 trials were obtained for each target in the offline BCI experiment. Each trial lasted 9.2 s, of which 5 s was used for visual stimulation and 4.2 s was utilized for gaze shifting. Participants were instructed to focus their gaze on a target stimulus according to an auditory cue. No feedback was provided in the offline BCI experiment. In addition, we utilized the Intel RealSense SR300 camera to capture 400 images, in which three objects with different colors were randomly placed in the workspace. The 400 images were used for evaluating the computer vision performance.

E. Online Experiment

The purpose of the online experiment was to assess the performance of the proposed AR-BCI controlled robotic arm. Two tasks were contained in the online experiment, one was a cued-guided robot movement task and the other is a freely controlled robot movement task. According to the optimized results from the offline BCI experiment, the online experiment adopted a 2.6 s visual stimulation time. Thus each trial lasted 6.8 s, of which 2.6 s was used for visual stimulation and 4.2 s was utilized for gaze shifting. Participants were instructed to focus their gaze on a target stimulus according to an auditory cue. The cued-guided robot movement task contained 10 blocks. Each block included 4 trials, each of which corresponded to 4 targets. The presentation order of the 4 trials was pseudo-random. The auditory cue was provided right after the end of the 2.6 visual stimulation time. Real-time visual feedback was provided. Namely, the robotic arm autonomously picked up the selected object from the black area and move it to desired location in the yellow area. For the freely controlled robot movement task, participants were asked to utilize the proposed AR-BCI controlled robotic arm to pick up and place the objects in a specific order (i.e., red, blue, and yellow). For each participant, the freely controlled robot movement task was executed five times.

F. EEG Data Analysis

EEG epochs were segmented in [0.14 s 0.14+d s] based on the stimulation triggers produced by the stimulus program, where d represented data length utilized in subsequent data analysis. This study first removed the linear trend in each epoch and then each segment was notch-filtered at 50 Hz to reduce the power-line frequency noise. Subsequently, all epochs were down-sampled from 1000 Hz to 250 Hz. In the offline analysis, we analyzed the amplitude spectrum and signal-to-noise ratio (SNR) of SSVEPs. Fourier transform was used to calculate amplitude spectrum $y(f)$. The data length used in the offline analysis was 5 s. Thus the frequency resolution of Fourier transform was 0.2 Hz. The SNR was calculated by the following equation:

$$\mathrm{SNR} = 20\log_{10}\frac{8 \times y(f)}{\sum_{k=1}^{4}\left[y(f-0.2\times k)+y(f+0.2\times k)\right]} \tag{1}$$

In the present study, we utilized the filter bank canonical correlation analysis (FBCCA) algorithm to detect

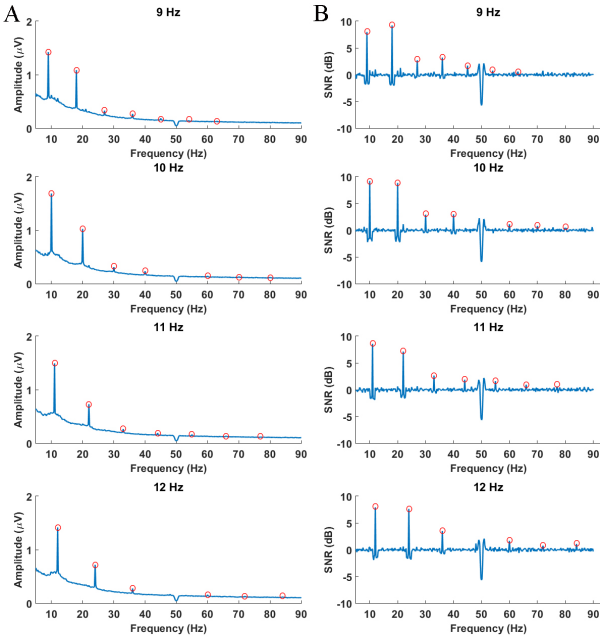

Fig. 5. Amplitude spectrum (A) and SNR (B) of SSVEPs averaged across all subjects for each target. The red circles represent the stimulus frequency and its harmonics.

SSVEPs [28]. The FBCCA method mainly involves three major aspects: filter bank analysis, standard CCA analysis, and target identification. For extracting and integrating information from fundamental and harmonic SSVEP components, the FBCCA method first performs a filter bank analysis to decompose EEG data into multiple sub-bands. In this study, SSVEP harmonics were still clearly visible at 84 Hz in the SNR image (see Fig. 5). According to the stimulation frequencies (i.e., 9 Hz, 11 Hz, 10 Hz, and 12 Hz) and the SNR analysis of SSVEPs, the frequency range between 8 and 88 Hz was used for the filter bank. Our previous study has validated that sub-bands containing multiple harmonic bands with a high cut-off frequency is the effective method for the design of sub-bands in the filter bank [28]. Thus the frequency range of the n-th sub-band was from $n \times 8$ Hz to 88 Hz in this study. According to the previous study [28], the number of sub-bands was set to 7. Subsequently, the standard CCA was performed on each sub-band. The canonical correlation values $(\rho_k^n, n = 1, 2, \cdots, N)$ between sub-bands and reference signals of sine-cosine waves are obtained, where N is the number of sub-bands. CCA was implemented using MATLAB's *canoncorr* function. The number of harmonics in the CCA process was set to 4. A weighted sum of squares of the correlation coefficients corresponding to all sub-bands is used for target recognition:

$$\tilde{\rho}_k = \sum_{n=1}^{N} w(n) \cdot \left(\rho_k^n\right)^2, \quad k = 1, 2, 3, 4 \tag{2}$$

where n is the sub-band index, and $w(n)$ is the weight for each sub-band. In this study, $w(n) = n^{-1.25} + 0.25$ based on the previous work [28]. The target is recognized as

$$F_{target} = \max_k \tilde{\rho}_k \tag{3}$$

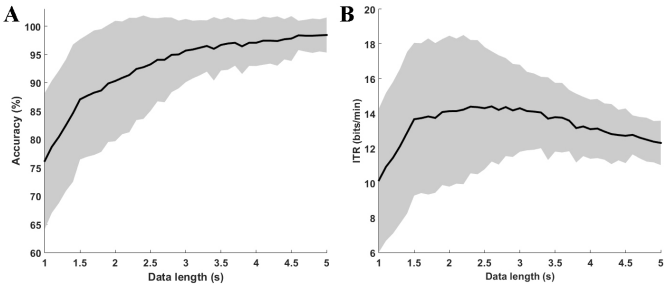

Fig. 6. Accuracy (A) and ITR (B) of the proposed SSVEP-based BCI with different data lengths (from 1 s to 5 s with an interval of 0.1 s). The shaded area indicates the standard deviation.

III. Results

A. Offline Performance

To assess the signal quality of SSVEPs induced by the HoloLens, amplitude spectrum and SNR were calculated for the SSVEPs. SSVEP amplitude of each epoch was calculated using Fourier transform, and then the averaged SSVEP amplitudes were obtained by frequency domain averaging. Fig. 5 shows the SNR and amplitude spectrum of the SSVEPs induced by the HoloLens. As shown in Fig. 5, the SSVEPs exhibit peak amplitudes and SNRs at the stimulus frequency and its harmonics. The amplitudes of the SSVEP harmonics decrease obviously when increasing the response frequency. While the SNRs of the SSVEP harmonics declines relatively slowly when increasing the response frequency. These results demonstrated that the HoloLens could serve as a visual stimulator for eliciting SSVEPs.

In the SSVEP-BCI subsystem, the FBCCA method was adopted to classify the four-target SSVEPs. Information transfer rate (ITR), i.e., the amount of information transferred per unit time, is the most popular evaluation metric used in BCIs [29]. Therefore, we also calculated the ITR of the SSVEP-BCI subsystem. A 4.2 s gaze shifting time was added to the ITR estimation. Fig. 6 shows the accuracy and ITR with different data lengths. As depicted in Fig. 6, the classification accuracy increases as the data length increases. When the data length is 2.6 s, the maximum ITR (i.e., 14.41 bits/min) is obtained. The classification accuracy corresponding to the maximum ITR is 94.06%. Therefore, the online experiment adopted a 2.6 s visual stimulation time. Fig. 7 illustrates the confusion matrix of each command at the 2.6 s visual stimulation time. The mean accuracy of each command is higher than 90%. A one-way repeated measures analysis of variance (ANOVA) was utilized to evaluate the difference of recognition accuracy among different commands, showing that there was no obvious difference in the accuracy of the four commands [$F(3, 27) = 1.46$, $p > 0.05$].

For computer vision performance evaluation, this study collected 400 images containing three objects with different colors. The proposed computer vision method could accurately recognize the three objects from these 400 images, thus achieving 100% recognition accuracy. Fig. 8 shows an example of the implementation of the proposed computer vision approach. As shown in Fig. 8, the proposed computer vision method could correctly detect the colors and positions of the three objects. In the computer vision subsystem,

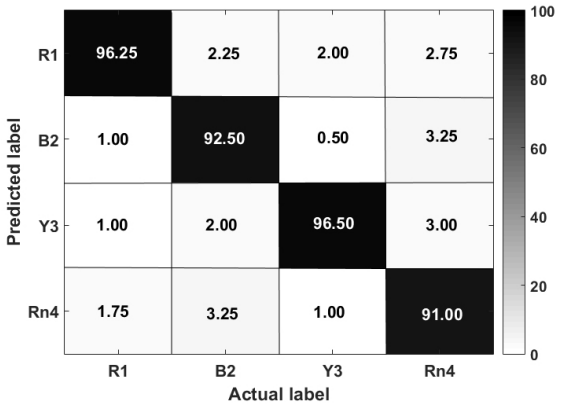

Fig. 7. Confusion matrix of each command at the 2.6 s visual stimulation time. The color scale shows the classification accuracy, and the diagonal reveals the correct classification accuracy of each command.

the three objects' colors were determined according to the defined hue, saturation and value ranges for the three colors. Since the 25 possible locations of the objects were accurately determined via the teach pendant of the robotic arm, the three objects' positions could be determined through mapping the 25 segmented sub-images with the 25 possible locations.

B. Online Performance

For the performance evaluation of the proposed AR-BCI controlled robotic arm, this study executed the online experiment. Two tasks (i.e., a cued-guided robot movement task and a freely controlled robot movement task) were included in the online experiment. TABLE I lists the online results of the cued-guided robot movement task. As listed in table I, the proposed system obtains a mean accuracy of 93.96 ± 5.05 % with a mean ITR of 14.21 ± 2.53 bits/min. The minimum and maximum accuracies for the proposed system is 82.5% and 100%, respectively. TABLE I lists the online results of the freely controlled robot movement task. All participants could complete the freely controlled robot movement. In average, they were able to finish the task in 107.67 ± 12.24 s. Nine of the twelve subjects could successfully complete the task without error. If the BCI command is identified incorrectly, the subject can use the undo command to undo the last operation. These online results verified that the proposed AR-BCI controlled robotic arm was feasible.

IV. Discussion

The present study showed the proof of concept for a high-level robot control system based on computer vision and AR-BCI to control a robotic arm. In the offline analysis, the proposed computer vision method could accurately determine the color and location information of the objects in the workspace. The HoloLens could elicit stable SSVEPs, and the proposed SSVEP-based BCI could decode the human intent with high accuracy. The classification accuracy corresponding to the maximum ITR is 94.06%. Additionally, the recognition accuracy of the online experiments was 93.96% across twelve subjects. All participants could utilize the proposed system to complete the pick-and-place task. These offline and online results verified the feasibility of healthy subjects manipulating a high-level robot control system.

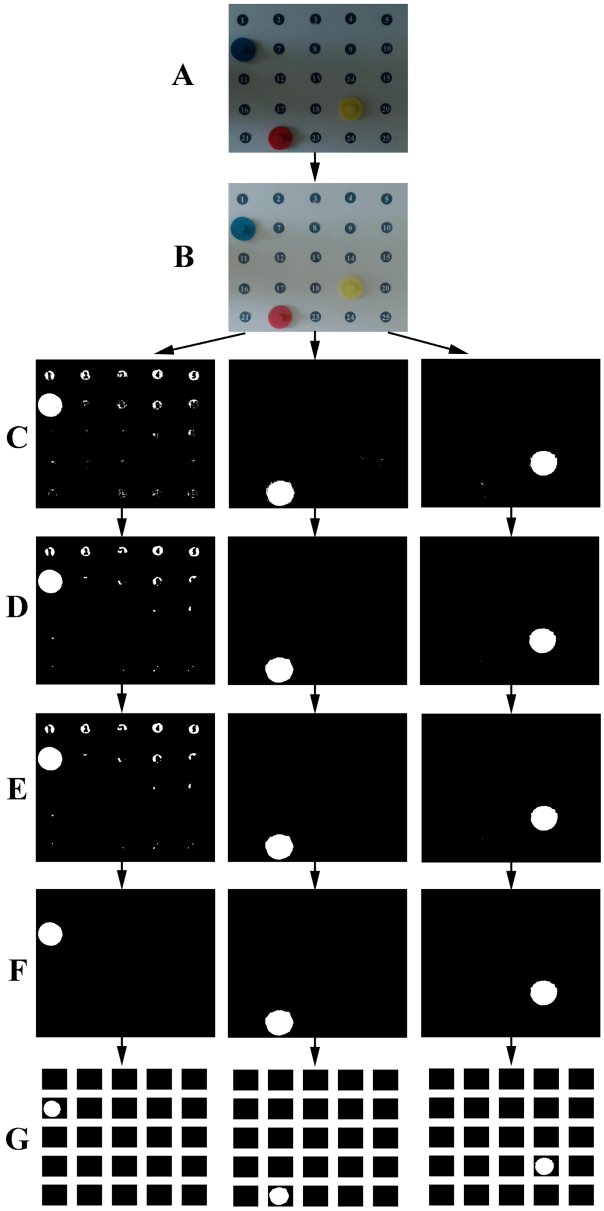

Fig. 8. An example of the implementation of the proposed computer vision approach. (A) Original image, (B) Homomorphic filtered image, (C) Detecting colors using HSV color space, (D) Median filtered image, (E) Binary image, (F) Removing small objects from binary image, and (G) Segmented image.

Currently, most SSVEP-BCIs adopt a "bulky" computer screen as a visual stimulator. This limits the portability of SSVEP-BCIs to some extent. AR is able to achieve enhanced perception of the real world by real-time integrating virtual objects into a real environment. Recently, with growing maturity of AR technology, it is possible to adopt AR device for eliciting SSVEP [16], [24], [26]. Portable and wearable AR devices further increase the portability of SSVEP-BCIs. Currently, combining AR technology with SSVEP-BCI is still at an early stage, only few studies are available on this issue [16], [24], [26]. As far as we know, the present study is the first work that realizes a high-level robot control system via combining AR, SSVEP-BCI, and computer vision. In this study, a wearable AR device was used to elicit SSVEPs. The AR environment was realized through the HoloLens worn by users. In comparison with computer

TABLE I
RESULTS OF THE CUED-GUIDED ROBOT MOVEMENT TASK

Subject	Accuracy (%)	ITR (bits/min)
S1	92.50	13.21
S2	95.00	14.42
S3	100.00	17.65
S4	100.00	17.65
S5	100.00	17.65
S6	95.00	14.42
S7	90.00	12.11
S8	92.50	13.21
S9	95.00	14.42
S10	82.50	9.30
S11	90.00	12.11
S12	95.00	14.42
Mean±SD	93.96±5.05	14.21±2.53

TABLE II
RESULTS OF THE FREELY CONTROLLED ROBOT MOVEMENT

Subject	Trial length (s)	Total number of trials	Completion time (s)
S1	6.8 (2.6+4.2)	17	115.60
S2	6.8 (2.6+4.2)	15	102.00
S3	6.8 (2.6+4.2)	15	102.00
S4	6.8 (2.6+4.2)	15	102.00
S5	6.8 (2.6+4.2)	15	102.00
S6	6.8 (2.6+4.2)	15	102.00
S7	6.8 (2.6+4.2)	15	102.00
S8	6.8 (2.6+4.2)	15	102.00
S9	6.8 (2.6+4.2)	15	102.00
S10	6.8 (2.6+4.2)	21	142.80
S11	6.8 (2.6+4.2)	17	115.60
S12	6.8 (2.6+4.2)	15	102.00
Mean±SD	-	15.83±1.80	107.67±12.24

monitors, the HoloLens is a lightweight headset with good portability. When adopting a computer monitor as a visual stimulator, the visual stimuli and the robotic arm are not presented in the same field of view [13], [15] or only limited components of the robotic arm can be presented on the computer monitor through cameras [16]. Users usually need to shift attention between the robotic arm and the computer monitor to enable better control the movement of the robotic arm. Unlike the computer monitor, the AR display can mitigate the attention shift problem since the see-through characteristic of AR display allows the robotic arm and the visual stimuli to be within the users' field of vision. With the help of AR, the interaction between user and robotic arm becomes more intuitive and natural. In this study, SSVEPs elicited by the HoloLens exhibited obvious peaks at the stimulus frequency and its harmonics (see Fig. 5). Moreover, the recognition accuracy of the proposed SSVEP-BCI was higher than 90% in both online and offline experiments. The results demonstrated that the HoloLens could present visual stimuli with stable frequencies for eliciting SSVEPs, and it is feasible to use SSVEP-based BCIs in the HoloLens-based AR environment. Although AR-based SSVEP-BCI could achieve high recognition accuracy in this study, using an AR display seem to slightly degrade SSVEP-based BCI performance compared with our previous studies using conventional computer monitor as visual stimulator [13], [15]. A similar degradation in BCI performance was also observed in two recent HoloLens-based AR studies [16], [26]. A possible reason for this phenomenon is that the inherent properties of AR display and conventional computer monitor are different. Unlike the computer monitor, the see-through characteristic of AR display allow users to simultaneously see visual stimuli and surrounding scenes, which may distract users' attention from visual stimuli [16], [26]. Moreover, the luminance and contrast of the visual stimuli on an AR display are lower than those presented on a computer monitor [26]. The high recognition accuracy of the present study demonstrated that this degradation would not affect the practical application of AR-based SSVEP-BCI. Compared with recent HoloLens-based AR studies [16], [24], [26], the proposed AR-BCI had a high recognition accuracy.

In terms of EEG signal acquisition, this study adopted a wireless, portable EEG amplifier (i.e., Neuracle). The high accuracy reported by the proposed SSVEP-BCI demonstrated that the wireless, portable EEG amplifier was suitable to build SSVEP-based BCIs. In SSVEP-based BCIs, visual stimulator and EEG amplifier are the main hardware components. The portability of the visual stimulator and EEG amplifier in this study facilitated the portability of the proposed SSVEP-BCI and allowed users to operate the robotic arm without staying in a fixed place. EEG decoding algorithms also play a vital role in SSVEP-based BCIs. The present study utilized the FBCCA approach for recognizing SSVEPs. The FBCCA method, which is a training-free method, does not require calibration data to train a calculation model. Thus even new users can use the proposed SSVEP-based BCI directly without recording training data. Additionally, the SSVEP-based BCI system designed for robotic arm control in the present study is also suitable for other applications of SSVEP-based BCIs.

Objects in the workspace were captured by an Intel RealSense SR300 camera. The computer vision recognized the objects in the workspace based on the images obtained by the camera. Then the color and location of objects in the workspace could be obtained. In this study, these three objects were randomly placed by the robotic arm in three of the 25 pre-taught locations. This facilitated the identification of the objects' locations to some extent. Furthermore, the three colors (i.e., red, blue, and yellow) used in the present study were very distinguishable in the HSV color space. These facilitated accurately identification of the color and location information of these objects in the workspace. Both offline and online results verified that the proposed computer vision method could accurately recognize these objects. According to the results obtained from the SSVEP-BCI and the computer vision, the robotic arm would automatically perform the picking and placing operations. The BCI high-level control strategy with computer vision assistance allowed users to send the mental intention by using a simple command. On the one hand, this high-level control strategy can reduce the users' mental load. On the other hand, high-level decisions are determined by users, which helps the total system to deal with uncertainty, complexity and ambiguity in decision making. Although numerous EEG-BCI controlled robotic arms have been developed in recent years, studies on high-level or shared control strategies are still few [15], [17]. Moreover, studies of combing high-level control strategy and AR technology to

control robotic arms are even rarer. The high-level robotic arm control system based on computer vision and AR-BCI developed in this study would provide new insights into BCI-controlled robots.

In order to promote the daily usability of the proposed system in everyday life environments, further improvements of the proposed system can be achieved through the following aspects. First, the usability of the proposed system will be tested on elderly individuals or paralyzed patients in the near future. Second, the proposed SSVEP-BCI performance can be improved by adopting training-based approaches for SSVEP detection [4], [22]. Meanwhile, we will try to use transfer-learning techniques to shorten the calibration time without affecting recognition accuracy [30], [31]. Last, each successful trial of the proposed system needed the computer vision to accurately provide the color and position information of the desired object. We will focus on the addition of cameras to provide more robust recognition of various objects [32].

V. Conclusion

A high-level robotic arm control system based on computer vision and AR-BCI was designed and realized in the present study. The offline results demonstrated that both AR-BCI and computer vision had high accuracy. The feasibility of the proposed system in online applications was also verified through a cued-guided robot movement task and a freely controlled robot movement task. The proposed system obtained a mean accuracy of 93.96 $\pm$ 5.05 % in the cued-guided robot movement task. Furthermore, all participants could complete the freely controlled robot movement task. The offline and online results verified the potential of combining computer vision and AR-BCI to operate a robotic arm.

References

[1] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain-computer interfaces for communication and rehabilitation," *Nature Rev. Neurol.*, vol. 12, no. 9, pp. 513–525, 2016.

[2] J. R. Wolpaw and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 51, pp. 17849–17854, Dec. 2004.

[3] R. Fazel-Rezai, B. Z. Allison, C. Guger, E. W. Sellers, S. C. Kleih, and A. Kübler, "P300 brain computer interface: Current challenges and emerging trends," *Frontiers Neuroeng.*, vol. 5, p. 14, Jul. 2012.

[4] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T. P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Natl. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, 2015.

[5] M. J. Vansteensel, E. G. M. Pels, M. G. Bleichner, M. P. Branco, T. Denison, Z. V. Freudenburg, P. Gosselaar, S. Leinders, T. H. Ottens, M. A. Van Den Boom, P. C. Van Rijen, E. J. Aarnoutse, and N. F. Ramsey, "Fully implanted brain–computer interface in a locked-in patient with ALS," *New England J. Med.*, vol. 375, no. 21, pp. 2060–2066, 2016.

[6] C. Pandarinath *et al.*, "High performance communication by people with paralysis using an intracortical brain-computer interface," *eLife*, vol. 6, Feb. 2017, Art. no. e18554.

[7] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3156–3166, Nov. 2013.

[8] K. K. Ang *et al.*, "A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clin. EEG Neurosci.*, vol. 42, no. 4, pp. 253–258, Oct. 2011.

[9] M. Al-Quraishi, I. Elamvazuthi, S. Daud, S. Parasuraman, and A. Borboni, "EEG-based control for upper and lower limb exoskeletons and prostheses: A systematic review," *Sensors*, vol. 18, no. 10, p. 3342, Oct. 2018.

[10] L. R. Hochberg *et al.*, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012.

[11] J. L. Collinger *et al.*, "High-performance neuroprosthetic control by an individual with tetraplegia," *Lancet*, vol. 381, no. 9866, pp. 557–564, Feb. 2013.

[12] Z. Zhang *et al.*, "An intention-driven semi-autonomous intelligent robotic system for drinking," *Frontiers Neurorobotics*, vol. 11, p. 48, Sep. 2017.

[13] X. Chen, B. Zhao, Y. Wang, S. Xu, and X. Gao, "Control of a 7-DOF robotic arm system with an SSVEP-based BCI," *Int. J. Neural Syst.*, vol. 28, no. 08, Oct. 2018, Art. no. 1850018.

[14] B. J. Edelman *et al.*, "Noninvasive neuroimaging enhances continuous neural tracking for robotic device control," *Sci. Robot.*, vol. 4, no. 31, Jun. 2019, Art. no. eaaw6844.

[15] 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, Apr. 2019, Art. no. 026012.

[16] Y. Ke, P. Liu, X. An, X. Song, and D. Ming, "An online SSVEP-BCI system in an optical see-through augmented reality environment," *J. Neural Eng.*, vol. 17, no. 1, Feb. 2020, Art. no. 016066.

[17] A. Úbeda, E. Iáñez, and J. M. Azorín, "Shared control architecture based on RFID to control a robot arm using a spontaneous brain–machine interface," *Robot. Auto. Syst.*, vol. 61, no. 8, pp. 768–774, Aug. 2013.

[18] N. Waytowich, A. Henderson, D. Krusienski, and D. Cox, "Robot application of a brain computer interface to staubli Tx40 robots-early stages," in *Proc. World Automat. Congr. (WAC)*, 2010, pp. 1–6.

[19] T. Lampe, L. D. J. Fiederer, M. Voelker, A. Knorr, M. Riedmiller, and T. Ball, "A brain-computer interface for high-level remote control of an autonomous, reinforcement-learning-based robotic system for reaching and grasping," in *Proc. 19th Int. Conf. Intell. User Interface (IUI)*, 2014, pp. 83–88.

[20] H. Zeng *et al.*, "Closed-loop hybrid gaze brain-machine interface based robotic arm control with augmented reality feedback," *Frontiers Neurorobotics*, vol. 11, p. 60, Oct. 2017.

[21] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory Brain–Computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.

[22] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[23] K. Takano, N. Hata, and K. Kansaku, "Towards intelligent environments: An augmented reality–brain–machine interface operated with a see-through head-mount display," *Frontiers Neurosci.*, vol. 5, p. 60, Apr. 2011.

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

[25] Z. Makhataeva and H. Varol, "Augmented reality for robotics: A review," *Robotics*, vol. 9, no. 2, p. 21, Apr. 2020.

[26] S. Park, H.-S. Cha, and C.-H. Im, "Development of an online home appliance control system using augmented reality and an SSVEP-based brain–computer interface," *IEEE Access*, vol. 7, pp. 163604–163614, 2019.

[27] S. A. M. Saleh and H. Ibrahim, "Mathematical equations for Homomorphic filtering in frequency domain: A literature survey," in *Proc. Int. Conf. Inf. Knowl. Manag.*, Singapore, 2012, pp. 74–77.

[28] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008.

[29] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, 2002.

[30] P. Yuan, X. Chen, Y. Wang, X. Gao, and S. Gao, "Enhancing performances of SSVEP-based brain–computer interfaces via exploiting inter-subject information," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046006.

[31] M. Nakanishi, Y.-T. Wang, C.-S. Wei, K.-J. Chiang, and T.-P. Jung, "Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1105–1113, Apr. 2020.

[32] D. P. Mcmullen *et al.*, "Demonstration of a semi-autonomous hybrid brain–machine interface using human intracranial EEG, eye tracking, and computer vision to control a robotic upper limb prosthetic," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 784–796, Jul. 2014.