

Received March 5, 2020, accepted March 21, 2020, date of publication March 26, 2020, date of current version April 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983608

# Real-Time Recognition of Facial Expressions Using Facial Electromyograms Recorded Around the Eyes for Social Virtual Reality Applications

HO-SEUNG CHA<sup>ID</sup>, SEONG-JUN CHOI, AND CHANG-HWAN IM<sup>ID</sup>, (Member, IEEE)

Department of Biomedical Engineering, Hanyang University, Seoul 04763, South Korea

Corresponding author: Chang-Hwan Im (ich@hanyang.ac.kr)

This work was supported by the Samsung Science & Technology Foundation [SRFC-TB1703-05, fEMG-based facial expression recognition for interactive VR applications].

**ABSTRACT** Recent developments of social virtual reality (VR) services using avatars have increased the need for facial expression recognition (FER) technology. FER systems are generally implemented using optical cameras; however, the performance of these systems can be limited when users are wearing head-mounted displays (HMDs) as users' faces are largely covered by the HMDs. Facial electromyograms (fEMGs) that can be recorded around users' eyes can be potentially used for implementing FER systems for VR applications. However, this technology lacks practicality owing to the need for large-scale training datasets; furthermore, it is hampered by a relatively low performance. In this study, we proposed an fEMG-based FER system based on the Riemannian manifold-based approach to reduce the number of training datasets needed and enhance FER performance. Our experiments with 42 participants showed an average classification accuracy as high as 85.01% for recognizing 11 facial expressions with only a single training dataset for each expression. We further developed an online FER system that could animate a virtual avatar's expression reflecting a user's facial expression in real time, thus demonstrating that our FER system can be potentially used for practical interactive VR applications, such as social VR networks, smart education, and virtual training.

**INDEX TERMS** Facial expression recognition, facial electromyography, riemannian manifolds, human-machine interface, virtual reality.

## I. INTRODUCTION

With rapid developments in virtual reality (VR) technologies, there is an increasing interest in social network VR applications using human or human-like avatars [1]–[4]. For example, Facebook recently released a new VR social network application called Facebook Space [5] and started to evolve their mobile-based social network service (SNS) platform into a VR-based SNS. In addition, various VR applications, such as Vtime, VRchat, AltSpaceVR, and High Fidelity VR have been released into the market to keep pace with rapidly changing social interaction environments [6], [7].

To allow users a more immersive experience during VR-based social interaction, it is important to precisely recognize their facial expressions and visualize them on an avatar in VR. Facial expression recognition (FER) has been studied widely using optical cameras in the field of computer vision [8]–[13]. In camera-based FER systems,

multiple facial components, such as eyes, nose, cheeks, mouth, eyebrows, ears, and forehead, need to be detected from a camera image to recognize a user's facial expressions [13]. However, these systems cannot effectively detect facial expressions when a user wears a head-mounted display (HMD) as a large portion of the face is covered by the HMD. For example, Olszewsk *et al.* [14] attached an RGB camera onto an HMD and continuously captured images of a user's lips to estimate the user's facial expressions. Similarly, Burgos-Artizzu *et al.* [15] used a webcam placed at a distance from users to recognize their facial expressions when wearing HMDs. These camera-based systems could precisely detect changes in the users' facial expressions around the lips; however, they failed to estimate motions related to other parts of the face that were covered by the HMD, such as eyes and eyebrows.

To tackle this issue, infrared cameras were attached on the inner side of HMDs to track user's eye and eyebrow movement [16], [17]. To implement this system, at least three wide angle cameras are needed, which not only makes the

The associate editor coordinating the review of this manuscript and approving it for publication was Zhihan Lv<sup>ID</sup>.

**TABLE 1.** Studies on facial expression recognition based on fEMG.

Year/Reference	2004 [27]	2009 [22]	2010 [23]	2011 [24]	2013 [25]	2015 [26]
Number of electrodes for fEMG recording	6	16	6	6	6	2 (Bipolar)
Facial muscles from which fEMG activity was recorded	C, L, M	C, L, M, F, O, Z, D, Men	F, T	F, T	F, T	F
Number of facial expressions	3	6	5	8	10	5
Number of participants who were recruited for the experiment	6	1	3	10	10	6
Number of trials for registering a single facial expression	n.r.	n.r.	9	4	4	14
Recognition accuracy (%)	85	92	92.60	91.80	87.10	97.10
Information transfer rate in bits/trial	0.74	1.98	1.77	2.34	2.34	2.07

Note that “C”, “L”, “M”, “F”, “O”, “Z”, “D”, “Men”, and “T” represent corrugator, levator labii superioris, masseter, frontalis, orbicularis oculi, zygomaticus major, depressor anguli oris, mentalis, and temporalis, respectively. Information transfer rate in bits/trial was computed as  $\log_2 N + P \log_2 P + (1-P) \log_2 [(1-P)/(N-1)]$ , where  $N$  and  $P$  represent the number of possible selections and recognition rate, respectively. “n.r.” represents neither used nor mentioned in the referred study.

system expensive but also increases the total amount of data to be analyzed and transmitted. Furthermore, Li *et al.* [18] estimated muscular movement around the eyes using strain recorded from eight thin film strain gage sensors attached to an HMD; however, the use of thin and flexible strain sensors also increases the total cost of the system [19]. Moreover, an additional camera is still needed to capture muscular movements around the lips.

Recently, an FER system estimating facial expressions using facial electromyograms (fEMGs) recorded from electrodes attached to an HMD is being developed [20], [21]; however, its performance has not yet been publicized. This system seems promising because fEMG-based FER systems can be implemented at a much lower cost than camera- and strain-based FER systems. Table 1 lists a series of representative studies on fEMG-based FER [22]–[27]. Note that the studies listed in Table 1 determined electrode positions without considering potential VR applications and thus electrode locations were scattered all over the face. Moreover, the FER studies referred to in Table 1 might not be appropriate for practical social VR applications owing to the following limitations. Firstly, large training datasets were required to achieve a fairly high accuracy in classifying discrete facial expressions. For example, a recent study [26] required 14 trials to register just a single expression, which implies that

the users of this particular system have to go through long and tedious calibration sessions ( $14 \times 5 = 70$  repeated trials in this study) to recognize just five facial expressions. In the FER studies listed in Table 1, the least number of trials required to register a single expression was four [24], [25], which is still too large for use in practical VR applications. We think that in practical scenarios, an FER system should show a high classification performance after only a single registration of a user’s facial expressions. Moreover, previously reported fEMG-based FER systems were not validated with many participants. The maximum number of participants in previous studies was just ten [24], [25]. Lastly, to the best of our knowledge, online fEMG-based FER systems have never been reported.

The main aim of the present study is to implement a new fEMG-based FER system to address the limitations of conventional methods; the new system does not require a large training dataset but yields high classification performance. To this end, a Riemannian manifold-based classifier [28]–[30] was employed for the first time in the field of myoelectric interfaces. The performance of the implemented FER system was compared with that of conventional systems with a large number of participants ( $N = 42$ ), in terms of recognition accuracy and information transfer rate (ITR) [31], when only a single training trial was used to build a classifier.

We further implemented an online FER system, in which a virtual avatar mimicked a user’s facial expressions in real time.

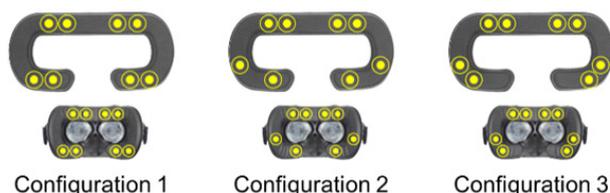
**II. EXPERIMENT ENVIRONMENT**

**A. DETERMINATION OF ELECTRODE CONFIGURATION**

Firstly, we fabricated a transparent plastic film identical in shape and size to an actual HMD face pad sold in the market (hereafter we call this plastic film an ‘HMD pad’). Nineteen sticker electrodes were attached to the plastic film; the electrodes were densely arranged as shown in the left part of Fig. 1. Such HMD pads were attached to the faces of three male adults, who were asked to freely move their facial muscles. In this preliminary test, it was observed that electrodes above specific facial muscles, such as the temporalis and corrugator, detached frequently from the skin. Therefore, we decided to exclude nine electrodes within the shaded rectangles in the left part of Fig. 1 from the list of electrode candidates. Eventually, fEMG data were recorded using ten electrodes as shown in Fig. 1. As the maximum number of channels supported by widely used commercial analog front ends for biomedical signal acquisition (e.g., ADS1298, Texas Instruments™) is eight, we decided to select eight electrodes from the ten electrodes. We evaluated the performance of three different electrode configurations with eight electrodes shown in Fig. 2, in terms of recognition accuracy. After determining the optimal electrode configuration, further analyses including real-time FER were conducted using the fEMG data acquired from the optimal eight electrodes.



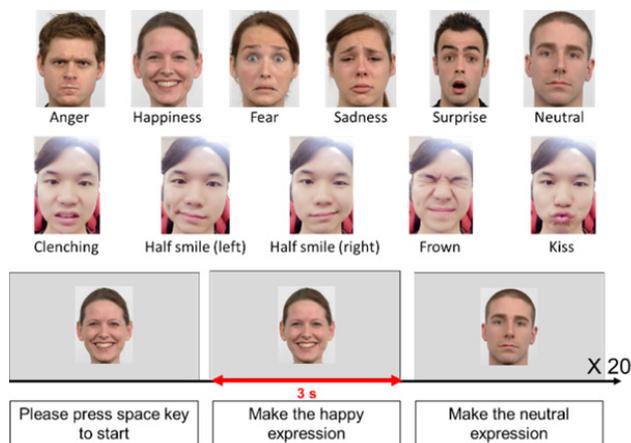
**FIGURE 1.** HMD pad for recording fEMGs and a user wearing an HMD pad with ten electrodes. Nine electrodes within the shaded rectangles were excluded in this study.



**FIGURE 2.** Three candidate electrode configurations for recording facial EMG.

**B. PICTURES OF EMOTIONAL FACES**

Anger, fear, happiness, neutrality, sadness, and surprise were sourced from the Radboud database [32], which contains a facial picture set of 67 models displaying emotional expressions based on a facial action coding system (FACS). The six selected pictures are presented in the first row of Fig. 3.



**FIGURE 3.** Eleven facial expression picture stimuli and the experimental procedure of performing a facial expression.

We selected the Radboud database as the facial pictures provided by the Radboud database were relatively easier to be used as reference pictures than those provided by other databased such as SEWA [33] and BP4D+ [34]. Five pictures of the first author’s face displaying different expressions including clenching, half smile (left and right), frown, and kiss were additionally used as reference pictures (see the second row in Fig. 3). In total, 11 pictures were employed as the reference facial expressions that the study participants mimicked during experiments.

**C. PARTICIPANTS**

Forty-two participants (17 males and 25 females) volunteered to participate in our study (age:  $24.07 \pm 1.89$  years ranging from 21 to 29 years), and all the participants were native Korean. None of the participants reported any serious health problems, such as Bell’s palsy, stroke, or Parkinson’s disease, that might affect the study. Before conducting the experiments, all the participants were given a detailed explanation of the experimental protocols and in addition, they signed a written consent. The participants received monetary compensation for their participation in the experiments. The study protocol was approved by the Institutional Review Board (IRB) of Hanyang University, South Korea (IRB No. HYI-14-167-11).

**D. EXPERIMENTAL PROCEDURE**

fEMG data were acquired using HMD pads with ten electrodes (see Fig. 1). We used a commercial biosignal recording system (Active Two; Biosemi B.V., Amsterdam, The Netherlands) and the sampling rate was set at 2,048 Hz. Each electrode was referenced and grounded by a common mode sense (CMS) electrode and a driven right leg (DRL) electrode attached at the left and right mastoids, respectively.

Before conducting the experiment, the participants were allowed a short training period to familiarize themselves with the 11 selected facial expressions as shown in Fig. 3. Reference facial pictures were presented on a computer

monitor using E-prime 2.0 (Psychology Software Tools, Sharpsburg, PA, USA), which also recorded the time points at which the facial pictures were presented.

To construct a database, each participant was asked to make 11 designated facial expressions repeatedly for 20 times. Fig. 3 shows the experimental procedure of a single trial. First, a facial expression that a participant needs to make (e.g., happy expression in Fig. 3) is presented on the monitor. The participant is asked to press the space bar when he/she is ready to move on to the next step. After the participant presses the space bar, a short beep sound is heard, at which time he/she needs to start mimicking the emotional face presented on the monitor for three seconds. Subsequently, the participant can make a neutral expression before the next designated facial expression is presented.

Pictures with different facial expressions were randomly presented. Each participant underwent a total of 220 trials (11 facial expressions  $\times$  20 repetitions). This dataset was uploaded to Figshare and is now freely available at <https://figshare.com/s/c2866bb4285af6d8b612>.

### III. CONVENTIONAL APPROACH

Typical pattern-recognition-based myoelectric interfaces, such as multifunction prosthesis, include a series of processes consisting of preprocessing, data segmentation, feature extraction, and classification, each of which has been well established [35]–[41]. We implemented an fEMG-based FER system using conventional methods widely used in pattern recognition-based myoelectric interfaces. All the analyses were conducted using Matlab 2018b (MathWorks, Inc., Natick, MA, USA).

#### A. PREPROCESSING

Traditionally, raw EMG signals are acquired in a bipolar configuration, which is known to alleviate movement artifacts and noises from stray potentials [42]. As each of the three different electrode configurations we tested (Fig. 2) contain eight electrodes, four electrode pairs were selected to acquire bipolar EMG signals. Two neighboring electrodes constituted a single channel, resulting in four electrode pairs – (1) two electrodes on the left forehead, (2) two electrodes on the right forehead, (3) two electrodes on the left cheekbone, and (4) two electrodes on the right cheekbone. Later, the raw fEMG signal was notch-filtered at 60 Hz to eliminate AC noise and band-pass filtered using a 4<sup>th</sup> order Butterworth filter with cutoff frequencies of 20 and 450 Hz.

#### B. DATA SEGMENTATION

Filtered fEMG signals were segmented into a series of short segments using a sliding window. To determine an optimal sliding window length, various window lengths (from 50 ms to 1,500 ms with a step size of 50 ms, i.e., 50, 100, ..., 1,450, 1,500) were tested. The sliding window was slid from 0 ms to the end of the signal with a fixed time interval of 50 ms, which implies that the facial expression of a participant was recognized at every 50 ms.

Later, only those segments of the fEMG signals acquired while a participant was making a designated facial expression were selected. To exclude fEMG signals recorded during transition intervals, segments including the first 1 second signal were not selected. The average fEMG onset time evaluated using a simple thresholding method was  $1.02 \pm 0.34$  seconds. To evaluate the fEMG onset time, the average RMS value of fEMG signals in the time interval of 2 – 3 s after the beep sound was set as the threshold value for each trial. Then, the RMS values at every time point in a trial were compared with the threshold value, and the first time point at which the RMS value exceeded the threshold value was determined as the fEMG onset time of the trial. Indeed, it has been reported that the use of transient EMG signals could degrade the performance of myoelectric interfaces [35], [39].

#### C. FEATURE EXTRACTION

Root-mean-square (RMS), wavelength (WL), sample entropy (SE), and cepstral coefficient (CC) were used as the features for the pattern classification. We selected these features as they showed the best performance in myoelectric pattern recognition studies [39]. Indeed, the combination of all four features not only showed the highest averaged classification accuracy but also exhibited robustness with respect to different segment lengths [39].

The equations to compute the features are presented in Table 2. The numbers of RMS, WL, SE, and CC features were 4, 4, 4, and 16 ( $= 4 \times 4$ ), respectively, as there were four bipolar fEMG channels. Therefore, the total number of features was 28, which constitute the feature vector denoted by  $FV_w(f, t)$ .

#### D. CLASSIFICATION

As the main aim of our study was to develop a minimum training fEMG interface, only one trial among 20 trials was selected and used to train the linear discriminant analysis (LDA) classifier [35], [36]. The first trials for each of the 11 facial expressions were used as the training dataset and the remaining 19 trials were used as the test dataset to evaluate the performance of our FER system. Note that we did not exclude any samples in the original dataset. The accuracy for each facial expression was defined as the number of correctly classified segments divided by the total number of samples. The accuracy for each participant was then calculated by averaging the accuracies of 11 facial expressions. In this study, the total number of samples was 760 (19 trials  $\times$  40 segments) for each participant and each facial expression.

### IV. RIEMANNIAN MANIFOLD-BASED APPROACH

Recently, Riemannian manifold-based pattern classification has attracted much interest for brain computer interface (BCI) applications as it has been reported that BCI performance can be significantly improved using this approach [28]–[30]. We hypothesized that the Riemannian approach could also be successfully applied to our problem because this approach utilizes spatial channel covariance unlike the conventional

**TABLE 2.** Features used for fEMG pattern classification.

Feature	Equation
Root-mean-square (RMS)	$\sqrt{\frac{1}{S} \sum_{s=1}^S x^2(s)}$
Wave length (WL)	$\sqrt{\sum_{s=1}^{S-1}  x(s+1) - x(s) }$
Sample entropy (SE)	$-\ln\left(\frac{A^2(r)}{B^2(r)}\right)$ , where $r = 0.2 \times \text{std}(x(s))$
Cepstral coefficient (CC)	$CC_1 = -a_1$ $CC_p = -a_p - \sum_{l=1}^{p-1} \left(1 - \frac{1}{p}\right) a_p CC_{(p-l)}$ where $1 < p \leq 4$ and $a_p$ is the coefficient of the autoregressive model.

$x(s)$  is the EMG signal of a single channel. When computing SE, A indicates the number of template vector pairs satisfying  $d[X_{m+1}(s), X_{m+1}(s)] < r$  while B indicates the number of template vector pairs satisfying  $d[X_m(s), X_m(s)] < r$ , where  $X_m(i) = \{x_s, \dots, x_{s+m-1}\}$  and  $d[X_m(i), X_m(j)]$  ( $i \neq j$ ) represents the Chebyshev distance. *std* indicates standard deviation. When computing CC, coefficients of an autoregressive model can be computed as  $x(s) = c + \sum_{p=1}^4 a_p x(s-p) + \varepsilon$ , where  $c$  is a constant and  $\varepsilon$  is the white noise of  $x(s)$ .

features, such as RMS, WL, SE, and CC, which are evaluated independently from each channel signal. It was expected that the Riemannian approach might effectively consider the muscle synergy during the expressive movements of the face as this phenomenon is similar to the synchronous brain activations during specific mental tasks in BCI. However, to the best of our knowledge, this approach has not been employed in myoelectric interfaces. In this study, we investigated whether Riemannian manifold-based pattern recognition can enhance the overall performance of our FER system.

### A. PREPROCESSING

Contrary to the conventional approach that uses bipolar montage, raw fEMG signals recorded from eight channels (see Fig. 2) in a unipolar setup were used in the Riemannian manifold-based approach. Raw fEMG signals were notch-filtered at 60 Hz and band-pass filtered using a 4<sup>th</sup> order Butterworth filter with cutoff frequencies of 20 and 450 Hz, similar to the conventional approach.

### B. DATA SEGMENTATION

The data segmentation process was identical to that of the conventional process (see section 2.5.2) except that the number of channels was different. Hereafter, a single segment of fEMG signals will be denoted by  $x_w \in \mathbb{R}^{C \times S}$ , where  $w = 1, 2, 3, \dots, W$ , with  $W$  being the number of windows,  $C$  the number of channels, and  $S$  the number of samples in a single segment. In this study,  $W$  and  $C$  were 40 ( $= 2 \text{ s} / 0.05 \text{ s}$ ) and 8, respectively, and  $S$  was determined by the analysis window length (various lengths were tested from 50 to 1,500 ms with a step size of 50 ms).

### C. FEATURE EXTRACTION

For each fEMG segment  $x_w$ , a  $C \times C$  sample covariance matrix (SCM) can be computed as  $C_w = 1/(S-1)x_w x_w^T$ . The SCM is a symmetric and positive-definite (SPD) matrix

and it is known that the space of the SPD matrices become a Riemannian manifold [43], which implies that it can be regarded as a point on Riemannian manifolds. In Riemannian manifolds, a finite-dimensional Euclidean space can be defined on a tangent space at an SCM. Therefore, it is important to determine a reference SCM, which will be used to form a tangent space [29]. In other words, mapping the  $C_w$  onto a tangent space can be regarded as new fEMG feature extraction in the Euclidean space, and thus machine learning-based classification algorithms can also be applied.

The reference SCM,  $C_{ref}$ , can be computed using Algorithm 1 in a previous literature [28] (Please see the Appendix for the details of the algorithm). Once the  $C_{ref}$  was computed,  $C_w$  computed with  $x_w$  was mapped onto the tangent space formed by  $C_{ref}$ . Eventually, the upper triangular elements of matrices in the tangent space were used as features. The tangent space mapping process is well described in Algorithm 2 in section 4 of [30]. Note that the number of features was 36 ( $= (8 \times 9) / 2$ ) in this study. More detailed theoretical and mathematical descriptions on the Riemannian geometry-based feature extraction can be found in literatures [28]–[30].

### D. CLASSIFICATION

A classification method identical to that of the conventional approach was employed, except that the training and test dataset features could be evaluated using the tangent space of  $C_{ref}$  once  $C_{ref}$  was computed using the  $C_w$ 's computed from the first trial of each of the 11 selected facial expressions.

## V. RESULTS

### A. INFLUENCE OF WINDOW LENGTH ON PROCESSING DELAY

In real-time FER systems, a short recognition time is important. Therefore, we first investigated the influence of window length on the processing delay time. Recog-

tion time includes the time taken for signal preprocessing, feature extraction, and classification. Recognition times corresponding to conventional and Riemannian approaches were measured using a series of windows of different lengths (50–1,500 ms with a step size of 50 ms; i.e., 50, 100, ..., 1,450, 1,500) on a desktop PC (Windows 10, 16 GB RAM, Intel Core i7 7700 3.60 GHz).

Fig. 4 shows the changes occurring in recognition time as a function of analysis window length in the conventional and Riemannian approaches. While the recognition time for conventional approach increased significantly at higher window lengths, the time required for the Riemannian approach did not increase much. To investigate the main cause for this difference, we measured the time taken to calculate each feature of conventional and Riemannian approaches. The average times taken to calculate RMS, CC, WL, SE, and the Riemannian feature from a 300 ms segment were 0.17 ms, 0.28 ms, 0.16 ms, 2.80 ms, and 0.57 ms, respectively. It took much longer time to calculate SE than other features, which is thought to be the main reason why the conventional approach spent longer recognition time than the Riemannian approach. To implement an online FER system with a speed of 20 frames per second, the classification decision should be made within 50 ms. Therefore, an analysis window length of 1,100 ms was determined as the maximum window length (see Fig. 4) as the signal processing time exceeds 50 ms for an analysis window length longer than 1,100 ms.

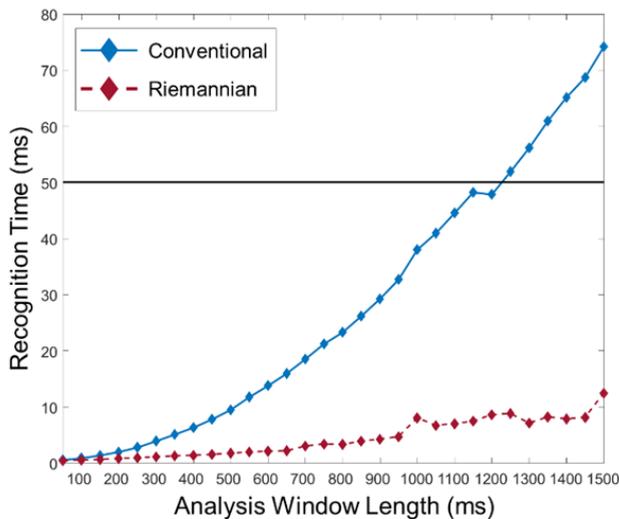


FIGURE 4. Time taken for recognizing a facial expression with respect to different analysis window lengths in conventional and Riemannian approaches.

**B. DETERMINATION OF OPTIMAL ELECTRODE CONFIGURATION AND WINDOW LENGTH**

The optimal electrode configuration and analysis window length for conventional and Riemannian approaches were determined based on recognition accuracy. Recognition accuracies with respect to different window lengths and candidate electrode configurations are shown in Fig. 5.

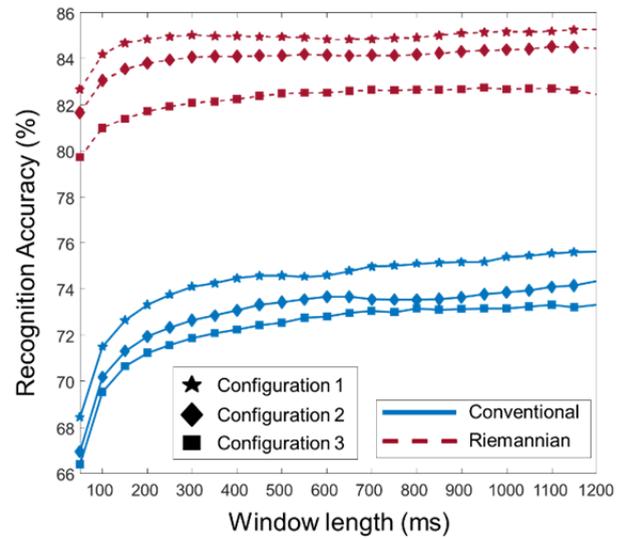


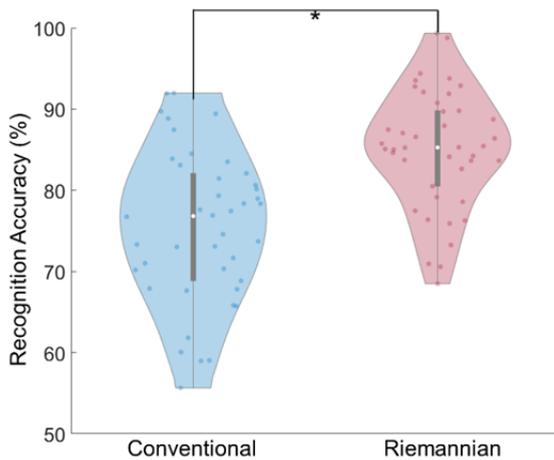
FIGURE 5. Recognition accuracies with respect to different window lengths and candidate electrode configurations.

As for the electrode configuration, configuration 1 yielded the highest recognition accuracy for both conventional and Riemannian approaches and hence was selected as the optimal electrode configuration.

With respect to window length, recognition accuracies of both approaches increased as the window length increased. These results are in line with those reported in the previous literature on myoelectric interfaces, where larger window sizes were recommended to avoid the increase in the bias and variance of features [36]. The maximum recognition accuracy was achieved at a window length of 300 ms in the Riemannian approach, while the maximum accuracy was achieved at a window length of 1,100 ms in the conventional approach. Accordingly, 300 ms and 1,100 ms were determined as the optimal window lengths for the Riemannian and conventional approaches, respectively.

When the optimal electrode configuration (Configuration 1) and optimal window lengths (300 ms and 1,100 ms) were used, the average recognition accuracies for the conventional and Riemannian approaches were  $74.58\% \pm 9.44\%$  and  $85.01\% \pm 7.30\%$  (mean  $\pm$  standard deviation), respectively. Fig. 6 represents a violin plot [44] of the average recognition accuracies of conventional and Riemannian approaches. It could be observed that the average recognition accuracy of the Riemannian approach was statistically significantly higher than that of the conventional approach (paired t-test:  $p = 3.66e-16$ ). The standard deviation of the accuracies was lower in the Riemannian approach than in the conventional approach, implying that the inter-individual variability in the recognition of facial expressions was also reduced by employing the Riemannian approach.

Fig. 7 shows the recognition accuracies of conventional and Riemannian approaches for each participant; error bars indicate standard deviation. The Riemannian approach



**FIGURE 6.** Violin plots of average recognition accuracies evaluated using the optimal window lengths and electrode configurations for conventional and Riemannian approaches. \* indicates a statistically significant difference ( $p < 0.05$ ).

always yielded a higher recognition accuracy than the conventional approach. The largest and second-largest differences in accuracy between the conventional and Riemannian approaches were 20.73% and 17.82%, respectively. The increases in the classification accuracy were much higher in the participants who showed low classification accuracies in the conventional approach (e.g., participants 16 and 7) than in the participants who showed high classification accuracies (e.g., participants 11 and 1), which mainly contributed to the significant increase in the average classification accuracy.

Fig. 8 shows the recall, precision, and F1 score for each facial expression. Note that facial expression types in the three bar graphs were sorted by recognition accuracy for the Riemannian approach in a descending order. The Riemannian approach outperformed the conventional approach for all the tested facial expression types, in terms of recall, precision, and F1 score. The Riemannian approaches were particularly

effective to increase F1 scores for three facial expressions of kiss, sadness, and anger. The increments of F1 scores for kiss, sadness, and anger were 14.24%p, 13.73%p, and 13.14%p, respectively. For example, the kiss expression was often misclassified as the sadness and anger expressions when the conventional approach was used, but the misclassification rates were considerably dropped from 7.2% to 3.7% and from 4.7% to 0.4%, respectively, by using the Riemannian approach.

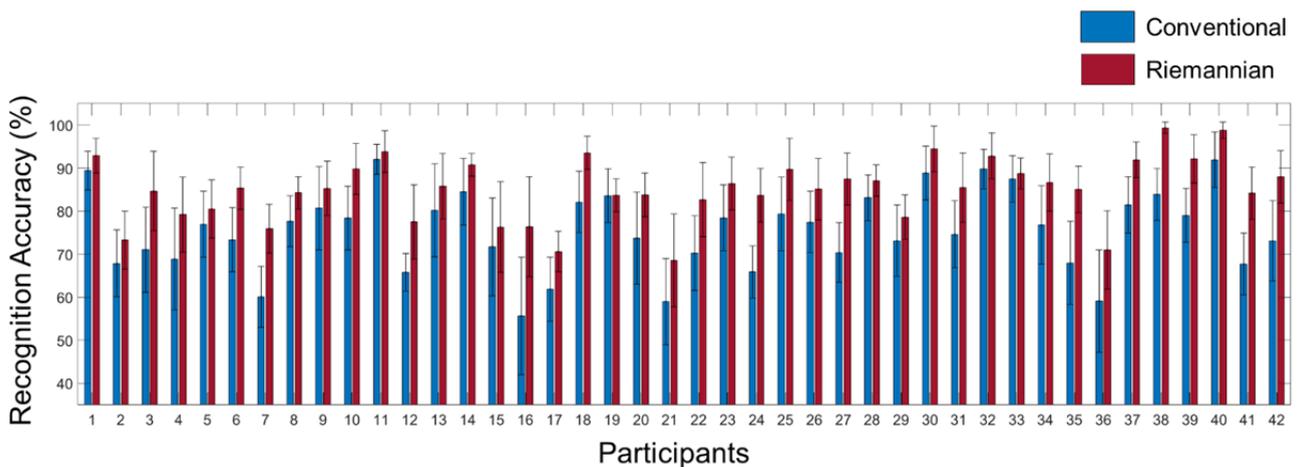
**C. FURTHER ANALYSIS WITH OPTIMAL CONDITIONS**

ITRs for the conventional and Riemannian approaches were 1.76 and 2.33 bits/trial, respectively. An ITR of 2.33 bits/trial was comparable to the highest reported ITR (2.34 bits/trial) among six studies on fEMG-based FER listed in Table 1. It should be noted that the ITR of 2.33 bits/trials reported in this study was achieved using only a single training dataset, while four repeated training sessions were required to achieve similar ITRs in previous studies [24], [25].

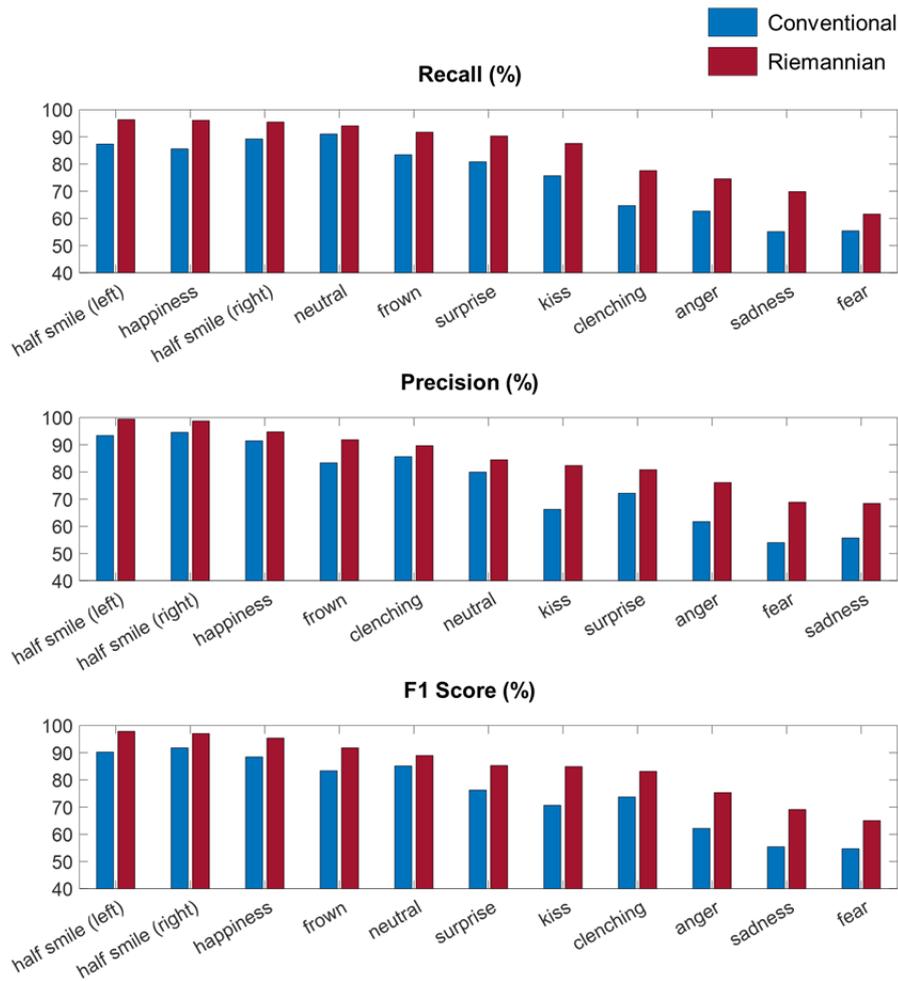
**D. REAL-TIME FACIAL ANIMATION WITH AN AVATAR**

An online fEMG-based FER program was implemented using the optimal conditions derived from offline results using Matlab. Louise’s 3D head model (EISKO, Paris, France) developed using commercial 3D-CG software Maya (Autodesk, Inc., San Rafael, CA, USA) was utilized as an avatar. The Avatar’s 11 facial animations were embodied based on the 11 facial expression picture stimuli shown in Fig. 3. The designated facial expressions were animated in real time based on the recognition results transported via user datagram protocol (UDP) communication. A snapshot of the online experiments is included in Fig. 9 and the movie can be found at <https://youtu.be/KLhR6EVkGOM>.

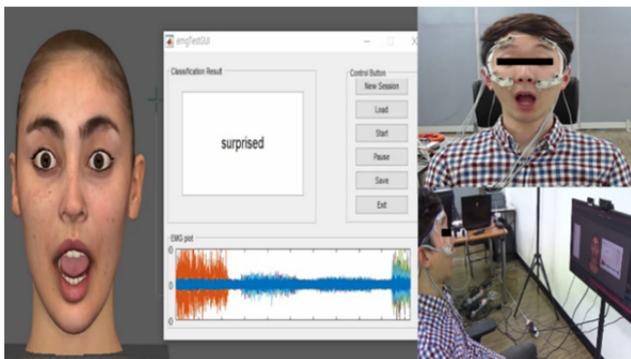
In this demonstration experiment, a male participant (age: 26 years old) was asked to make 11 facial expressions twice, during which his expressions were reflected in the avatar’s face in real time. Prior to the online experiment,



**FIGURE 7.** Recognition accuracies of conventional and Riemannian approaches for each participant. Error bars indicate standard deviation.



**FIGURE 8.** Recall, precision, and F1 score of each facial expression for conventional and Riemannian approaches. These parameters were calculated using the following equations:  $\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$ ,  $\text{precision} = \text{TP} / (\text{TP} + \text{FP})$ ;  $\text{F1 score} = 2 \times (\text{recall} \times \text{precision}) / (\text{recall} + \text{precision})$ , where TP and FP represent true positive and false positive, respectively.



**FIGURE 9.** A snapshot of online experiments taken when a participant was making a surprised facial expression (a YouTube video can be found at <https://youtu.be/KLhR6EVkGOM>).

the participant was asked to make 11 facial expressions to train a classifier. During the online experiment, fEMG data were stored in a circular buffer, when a series of signal processing steps including preprocessing and Riemann

feature extraction were conducted simultaneously. The extracted fEMG features were then fed into the pre-built classifier in real time with a rate of 20 times per second. The online FER system could correctly identify all designated facial expressions (22 trials); however, during the transient period between two facial expressions, unstable fluctuations of classification results were sometimes observed, which needs to be enhanced through further studies.

## VI. DISCUSSION

In this study, we implemented an FER system based on fEMG recorded around the eyes to reflect a user's facial expressions on the VR avatar for realizing a highly interactive social VR environment. Our FER system could classify 11 facial expressions with a fairly high recognition accuracy of over 85% with just a single registration for each facial expression. An increment of more than 10% could be achieved in recognition accuracy using the Riemannian manifold-based classifier when compared to the conventional feature extraction method. It is noteworthy that we could

True Labels	Half Smile (left)	96.3%	0.6%	0.1%	0.5%	0.2%		0.4%	0.8%	0.9%	0.1%	0.0%	
	Happiness	0.1%	96.1%	0.2%	0.6%	0.1%	0.1%	0.1%	1.4%	0.7%	0.0%	0.7%	
	Half Smile (right)		0.7%	95.5%	0.9%	0.0%		1.7%	0.4%	0.2%	0.0%	0.6%	
	Neutral		0.0%	0.0%	94.0%		0.6%	1.7%	0.4%	0.2%	2.0%	1.0%	
	Frown		0.2%		0.4%	91.7%	0.1%	1.4%	0.9%	3.8%	0.5%	1.0%	
	Surprise		0.0%		1.3%	0.1%	90.3%	0.1%	0.0%	0.1%	1.6%	6.5%	
	Kiss		0.0%	0.2%	0.1%	5.0%	0.1%	1.0%	87.6%	0.6%	0.4%	3.7%	1.2%
	Clenching		0.3%	1.7%	0.6%	2.5%	1.0%	0.1%	7.8%	77.6%	2.5%	0.8%	5.2%
	Anger		0.1%	0.4%	0.0%	1.4%	4.8%	2.0%	2.6%	0.7%	74.6%	10.1%	3.3%
	Sadness		0.0%			3.4%	1.0%	5.4%	1.4%	0.5%	10.2%	69.9%	8.1%
	Fear		0.0%	1.5%	0.2%	1.4%	0.8%	12.1%	1.6%	3.3%	4.2%	13.4%	61.6%
			Half Smile (left)	Happiness	Half Smile (right)	Neutral	Frown	Surprise	Kiss	Clenching	Anger	Sadness	Fear
		Predicted Labels											

FIGURE 10. A confusion matrix. Each row and column of the confusion matrix represent true labels and predicted labels, respectively.

achieve a high ITR of 2.33 with only a single training session; this value is similar to ITRs reported in previous studies that used four repeated training sessions [24], [25]. Although there has been extensive research to realize fEMG-based FER systems, no previous study tried to reduce the number of training sessions nor did they optimize the parameters affecting the performance of fEMG-based FER systems. In this study, we not only found an optimal fEMG electrode configuration on the HMD, but also investigated the influence of window length on system performance. Most importantly, we implemented an online FER system in which a virtual avatar could mimic a user’s facial expressions in real time at a high refreshing rate of 20 frames per seconds.

To identify frequently occurring errors in the implemented FER system, a confusion matrix was constructed (see Fig. 10). The rows and columns in the confusion matrix represent true labels and predicted labels, respectively. In the confusion matrix, values in the diagonal elements represent recalls while off-diagonal elements represent false negative rates (FNR). It can be seen from the figure that there was high confusion among fear, surprise, and sadness expressions.

For example, sadness was frequently misclassified as anger and fear with high FNRs of 10.2% and 8.1%, respectively. As seen in Fig. 3, the sadness expression does not show distinct facial gestures compared to other expressions such as surprise and happiness, making it difficult for the participants to mimic its unique facial gesture. Indeed, according to our post-experimental surveys, many participants answered that they had difficulty in mimicking the sadness expression. Fear was also frequently misclassified as surprise with a high FNR of 12.1%. This high confusion rate might originate from the similar facial gestures around the eyes between the fear and surprise expressions. It can be readily seen from Fig. 3 that both facial expressions include sharp raising of both eyebrows. On the other hands, the anger and fear expressions showed a low average FNR of 3.79% even though both expressions are associated with negative emotions. This low confusion rate might originate from the distinct differences in both eyebrow and lip gestures between them. As seen in Fig. 3, both eyebrows are lowered, and the lips are firmly closed in the anger expression; whereas both eyebrows are raised, and both labial commissures are lowered in the fear expression.

It is noteworthy that Riemannian approach yielded high classification accuracy with a relatively shorter analysis window length (300 ms) compared to the conventional approach (optimal window length = 1,100 ms). This implies that the maximum time delay of Riemannian approach is 300 ms, assuming that the FER processing time is negligible (less than 2 ms as shown in Fig. 4). It is well-reported that a threshold of detectable visual feedback delay of one's own body movement is about 200 ms [42]. Even when the analysis window length is reduced to 200 ms, the Riemannian approach can still yield a high classification accuracy of over 84% (see Fig. 5), suggesting that the proposed fEMG-based FER system might be used in practical scenarios without making the users of the system feel uncomfortable due to the delayed visual feedback. We believe that the main reason why the Riemannian approach outperformed the conventional approach in terms of classification accuracy is that the Riemannian approach evaluated the interrelationship among EMG signals recorded from multiple channels, allowing for the consideration of spatial information contained in the fEMG datasets, which could not be considered in the conventional approaches.

Although the present study demonstrated the possibility of practical fEMG-based FER systems, some issues still remain to be addressed. Firstly, it should be examined whether our FER system can be used repeatedly for a long period of time. It is well known that the performance of myoelectric interfaces degrades gradually due to several factors including electrode shift, humidity, and impedance changes [45]. Therefore, test-retest reliability and reusability of the FER system need to be verified in future studies. Secondly, although an fEMG-based FER system could be implemented with a single training dataset in this study, implementation of a user-independent FER system is also a promising future research topic. Our offline datasets can be employed to investigate the feasibility of the user-independent FER system as our datasets were collected by presenting a set of designated facial expressions to the participants. Thirdly, feasibility of voluntary emotional FER needs to be validated. In this study, participants were asked to mimic several designated emotional faces, which might be somewhat impractical because people have their own facial expressions reflecting their emotional states. Fourthly, the classification performance of our FER system might be further enhanced by employing more numbers of EMG electrodes embedded in the HMD face pad. We are now implementing an FER system with 16 electrodes by reducing the size of each electrode and designing an arrayed electrode configuration. Lastly, a method for tracking natural facial muscle movements needs to be developed. Our system could recognize only 11 designated facial expressions but could not recognize facial movements not included in the training set. A regression-based approach would be more appropriate to implement such a system rather than the present classification-based approach. This is an interesting topic that we would like to pursue in our future studies.

## APPENDIX

### Algorithm 1 Geometric Mean of SCM Matrices

**Input:** SCM matrices  $\{C_1, C_2, \dots, C_W\}$  and tolerance  $\epsilon > 0$ .

**Output:** The estimated geometric mean  $C_m$ .

**Initialize:**  $C_m^{(t)} = \frac{1}{W} \sum_{w=1}^W C_w$ ,  $t = 1$ ,  $\epsilon = 10e-5$

**While**

$$\bar{S} = \frac{1}{W} \sum_{w=1}^W \text{Log}_{C_m^{(t)}}(C_w)$$

$$C_m^{(t+1)} = \text{Exp}_{C_m^{(t)}}(\bar{S})$$

$$t = t + 1$$

**Until**  $\|C_m^{(t+1)}\|_F < \epsilon$

**Return**  $C_m^{(t+1)}$

$\text{Log}_C(C_w) = C^{\frac{1}{2}} \log_m \left( C^{(-\frac{1}{2})} C_w C^{(-\frac{1}{2})} \right) C^{\frac{1}{2}}$ , where  $\log_m(\cdot)$  denotes the logarithm of a matrix. Please note that the logarithm of a diagonalizable matrix  $A = \mathbf{VDV}^{-1}$  is defined as  $\log_m(A) = \mathbf{VDV}^{-1}$ , where the elements of  $\mathbf{D}$  are given by  $d_{(i,j)} = \log(d_{(i,j)})$ .  $\text{Exp}_C(\bar{S}) = C^{\frac{1}{2}} \exp \left( C^{(-\frac{1}{2})} \bar{S} C^{(-\frac{1}{2})} \right) C^{\frac{1}{2}}$ , where  $\exp(\cdot)$  denotes the exponential of matrix. Please note that the exponential of a diagonalizable matrix  $A = \mathbf{VDV}^{-1}$  is defined as  $\exp(A) = \mathbf{VDV}^{-1}$ , where the elements of  $\mathbf{D}$  are given by  $d_{(i,j)} = \exp(d_{(i,j)})$ .  $\|\cdot\|_F$  denotes Frobenius Norm.

## REFERENCES

- [1] R. Schroeder, A. Huxor, and A. Smith, "Activeworlds: Geography and social interaction in virtual reality," *Futures*, vol. 33, no. 7, pp. 569–587, Sep. 2001.
- [2] J. N. Bailenson, A. C. Beall, J. Loomis, J. Blascovich, and M. Turk, "Transformed social interaction, augmented gaze, and social influence in immersive virtual environments," *Human Commun. Res.*, vol. 31, no. 4, pp. 511–537, Oct. 2005.
- [3] R. Schroeder, "Social interaction in virtual environments: Key issues, common themes, and a framework for research," in *The Social Life of Avatar*. London, U.K.: Springer, 2011, ch. 1, pp. 1–18.
- [4] X. Pan and A. F. D. C. Hamilton, "Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape," *Brit. J. Psychol.*, vol. 109, no. 3, pp. 395–417, Aug. 2018.
- [5] *Facebook Spaces*. [Online]. Available: <https://www.facebook.com/spaces>
- [6] A. Bonasio, "Social VR: Who Is Going to Get It Right in This Industry First?" UploadVR, San Francisco, CA, USA. Accessed: Nov. 22, 2016. [Online]. Available: <https://uploadvr.com/social-vr-whos-going-to-get-it-first>
- [7] J. K. Patel and A. Sakadasariya, "Survey on virtual reality in social network," in *Proc. 2nd Int. Conf. Inventive Syst. Control (ICISC)*, Seoul, South Korea, Jan. 2018, pp. 1341–1344.
- [8] W. Sun and B. Wang, "A survey of facial expression recognition," *Comput. Knowl. Technol.*, vol. 2, no. 9, pp. 33–38, 2012.
- [9] S. K. Shah, "A survey of facial expression recognition methods," *IOSR J. Eng.*, vol. 4, no. 4, pp. 01–05, Apr. 2014.
- [10] T. Zhang, "Facial expression recognition based on deep learning: A survey," in *Proc. Adv. Intell. Syst. Comput.*, vol. 686, 2017, pp. 345–352.
- [11] S. Agrawal, P. Khatri, and S. Gupta, "Facial expression recognition techniques: A survey," *Int. J. Adv. Electron. Comput. Sci.*, vol. 2, no. 1, pp. 61–66, 2015.
- [12] T. Wu, S. Fu, and G. Yang, "Survey of the facial expression recognition research," in *Proc. Int. Conf. Brain Insp. Cog. Sys.*, 2012, pp. 392–402.
- [13] J. Kumari, R. Rajesh, and K. M. Pooja, "Facial expression recognition: A survey," *Procedia Comput. Sci.*, vol. 58, pp. 486–491, Jan. 2015.
- [14] K. Olszewski, J. J. Lim, S. Saito, and H. Li, "High-fidelity facial and speech animation for VR HMDs," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1–14, Nov. 2016.
- [15] X. P. Burgos-Artizzu, J. Fleureau, O. Dumas, T. Tapie, F. LeClerc, and N. Mollet, "Real-time expression-sensitive HMD face reconstruction," in *Proc. SIGGRAPH ASIA Tech. Briefs (SA)*, 2015, pp. 1–4.

- [16] S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa, "Eyemotion: Classifying facial expressions in VR using eye-tracking cameras," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Hawaii, HI, USA, Jan. 2019, pp. 1626–1635.
- [17] J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt, and M. Nießner, "FaceVR: Real-time gaze-aware facial reenactment in virtual reality," *ACM Trans. Graph.*, Mar. 2018.
- [18] H. Li, L. Trutoiu, K. Olszewski, L. Wei, T. Trutna, P.-L. Hsieh, A. Nicholls, and C. Ma, "Facial performance sensing head-mounted display," *ACM Trans. Graph.*, vol. 34, no. 4, Jul. 2015, Art. no. 47.
- [19] H. J. Xian, C. R. Cao, J. A. Shi, X. S. Zhu, Y. C. Hu, Y. F. Huang, S. Meng, L. Gu, Y. H. Liu, H. Y. Bai, and W. H. Wang, "Flexible strain sensors with high performance based on metallic glass thin film," *Appl. Phys. Lett.*, vol. 111, no. 12, Sep. 2017, Art. no. 121906.
- [20] *Emteq*. Accessed: Apr. 2, 2020. [Online]. Available: <https://emteq.net>
- [21] I. Mavridou, M. Hamed, M. Fatoorechi, J. Archer, A. Cleal, E. Balaguer-Ballester, E. Seiss, and C. Nduka, "Using facial gestures to drive narrative in VR," in *Proc. 5th Symp. Spatial User Interact. (SUI)*, Brighton, U.K., Oct. 2017, p. 152.
- [22] G. Gibert, M. Pruzinec, T. Schultz, and C. Stevens, "Enhancement of human computer interaction with facial electromyographic sensors," in *Proc. 21st Annu. Conf. Austral. Comput.-Hum. Interact. Special Interest Group Design, Open 24/7 (OZCHI)*, 2009, pp. 421–424.
- [23] I. M. Reza zadeh, X. Wang, M. Firoozabadi, and M. R. H. Golpayegani, "Using affective human-machine interface to increase the operation performance in virtual construction crane training system: A novel approach," *Autom. Constr.*, vol. 20, no. 3, pp. 289–298, May 2011.
- [24] M. Hamed, S.-H. Salleh, and T. T. Swee, "Surface electromyography-based facial expression recognition in bi-polar configuration," *J. Comput. Sci.*, vol. 7, no. 9, pp. 1407–1415, Sep. 2011.
- [25] M. Hamed, S.-H. Salleh, M. Astaraki, and A. Noor, "EMG-based facial gesture recognition through versatile elliptic basis function neural network," *Biomed. Eng. OnLine*, vol. 12, no. 1, p. 73, Jul. 2013.
- [26] Y. Chen, Z. Yang, and J. Wang, "Eyebrow emotional expression recognition using surface EMG signals," *Neurocomputing*, vol. 168, pp. 871–879, Nov. 2015.
- [27] L. Benedict P. Ang, E. F. Belen, R. A. Bernardo, E. R. Boongaling, G. H. Briones, and J. B. Coronel, "Facial expression recognition through pattern analysis of facial muscle movements utilizing electromyogram sensors," in *Proc. IEEE Region Conf. (TENCON)*, vol. 100, 2004, pp. 600–603.
- [28] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Classification of covariance matrices using a Riemannian-based kernel for BCI applications," *Neurocomputing*, vol. 112, pp. 172–178, Jul. 2013.
- [29] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to BCI classification," in *Proc. Int. Conf. Latent Variable Anal. Signal Separat.* Berlin, Germany: Springer, Sep. 2010, pp. 629–636.
- [30] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain-computer interface classification by Riemannian geometry," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, Oct. 2012.
- [31] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.
- [32] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg, "Presentation and validation of the Radboud Faces Database," *Cognition Emotion*, vol. 24, no. 8, pp. 1377–1388, Dec. 2010.
- [33] J. Kossaiji, B. W. Schuller, K. Star, E. Hajiyev, M. Pantic, R. Walecki, Y. Panagakis, J. Shen, M. Schmitt, F. Ringeval, J. Han, V. Pandit, and A. Toisoul, "SEWA DB: A rich database for audio-visual emotion and sentiment research in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Oct. 1, 2019, doi: [10.1109/TPAMI.2019.2944808](https://doi.org/10.1109/TPAMI.2019.2944808).
- [34] X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu, and J. M. Girard, "BP4D-spontaneous: A high-resolution spontaneous 3D dynamic facial expression database," *Image Vis. Comput.*, vol. 32, no. 10, pp. 692–706, Oct. 2014.
- [35] P. Geethanjali, "Myoelectric control of prosthetic hands: State-of-the-art review," *Med. Devices, Evidence Res.*, vol. 9, pp. 247–255, Jul. 2016.
- [36] M. Asghari Oskoei and H. Hu, "Myoelectric control systems—A survey," *Biomed. Signal Process. Control*, vol. 2, no. 4, pp. 275–294, Oct. 2007.
- [37] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.
- [38] H. Huang, P. Zhou, G. Li, and T. Kuiken, "Spatial filtering improves EMG classification accuracy following targeted muscle reinnervation," *Ann. Biomed. Eng.*, vol. 37, no. 9, pp. 1849–1857, Sep. 2009.
- [39] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, "EMG feature evaluation for improving myoelectric pattern recognition robustness," *Expert Syst. Appl.*, vol. 40, no. 12, pp. 4832–4840, Sep. 2013.
- [40] J. M. Hahne, B. Graimann, and K. R. Müller, "Spatial filtering for robust myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1436–1443, Jun. 2012.
- [41] T. W. Beck, T. J. Housh, J. T. Cramer, M. H. Malek, M. Mielke, R. Hendrix, and J. P. Weir, "A comparison of monopolar and bipolar recording techniques for examining the patterns of responses for electromyographic amplitude and mean power frequency versus isometric torque for the vastus lateralis muscle," *J. Neurosci. Methods*, vol. 166, no. 2, pp. 159–167, Nov. 2007.
- [42] M. Mohr, T. Schön, V. von Tscharnner, and B. Nigg, "Intermuscular coherence between surface EMG signals is higher for monopolar compared to bipolar electrode configurations," *Frontiers Physiol.*, vol. 9, p. 566, May 2018, doi: [10.3389/fphys.2018.00566](https://doi.org/10.3389/fphys.2018.00566).
- [43] W. Förstner and M. Boudewijn, "A metric for covariance matrices," in *Geodesy—The Challenge of the 3rd Millennium*. Berlin, Germany: Springer, 2003, pp. 299–309.
- [44] J. L. Hintze and R. D. Nelson, "Violin plots: A box plot-density trace synergism," *Amer. Statistician*, vol. 52, no. 2, pp. 181–184, May 1998.
- [45] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: Implications for EMG pattern recognition in prosthesis control," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046005.



**HO-SEUNG CHA** was born in Seoul, South Korea, in 1989. He received the B.S. degree from the Department of Biomedical Engineering, Yonsei University, South Korea, in 2013, and the M.S. degree from the Department of Biomedical Engineering, Hanyang University, South Korea, in 2015, where he is currently pursuing the Ph.D. degree. His research interests include biosignal processing, biosignal-based human computer interfaces, and brain-computer interfaces. He was a recipient of the Best Oral Presentation Award at the International Biomedical Engineering Conference (IBEC), in 2016, and Travel Award and Young Investigator Award at the international Society for Medical Innovation and Technology (SMIT), in 2018.



**SEONG-JUN CHOI** received the B.S. degree from the Department of Biomedical Engineering, Hanyang University, South Korea, in 2015, where he is currently pursuing the M.S. degree with the Department of Biomedical Engineering. His research interest includes biosignal analysis based on deep learning.



**CHANG-HWAN IM** (Member, IEEE) received the B.S. degree from the School of Electrical Engineering, Seoul National University, South Korea, in 1999, and the M.S. and Ph.D. degrees from Seoul National University, in 2001 and 2005, respectively. He was a Postdoctoral Associate with the Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN, USA, from 2005 to 2006. From 2006 to 2011, he was with the Department of Biomedical Engineering, Yonsei University, South Korea, as an Assistant/Associate Professor. Since 2011, he has been with the Department of Biomedical Engineering, Hanyang University, South Korea, as an Associate/Full Professor. He is currently the Director of the Computational Neuroengineering Laboratory, Hanyang University. He has authored over 150 articles in peer-reviewed international journals. His research interests include various fields of neuroelectromagnetic and computational neuroengineering, especially brain-computer interfaces, the diagnosis of neuropsychiatric diseases, noninvasive brain stimulation, bioelectromagnetic source imaging, and dynamic neuroimaging.