



Performance enhancement of facial electromyogram-based facial-expression recognition for social virtual reality applications using linear discriminant analysis adaptation

Ho-Seung Cha¹ · Chang-Hwan Im¹

Received: 28 July 2020 / Accepted: 20 August 2021

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Recent studies have indicated that facial electromyogram (fEMG)-based facial-expression recognition (FER) systems are promising alternatives to the conventional camera-based FER systems for virtual reality (VR) environments because they are economical, do not depend on the ambient lighting, and can be readily incorporated into existing VR headsets. In our previous study, we applied a Riemannian manifold-based feature extraction approach to fEMG signals recorded around the eyes and demonstrated that 11 facial expressions could be classified with a high accuracy of 85.01%, with only a single training session. However, the performance of the conventional fEMG-based FER system was not high enough to be applied in practical scenarios. In this study, we developed a new method for improving the FER performance by employing linear discriminant analysis (LDA) adaptation with labeled datasets of other users. Our results indicated that the mean classification accuracy could be increased to 89.40% by using the LDA adaptation method ($p < .001$, Wilcoxon signed-rank test). Additionally, we demonstrated the potential of a user-independent FER system that could classify 11 facial expressions with a classification accuracy of 82.02% without any training sessions. To the best of our knowledge, this was the first study in which the LDA adaptation approach was employed in a cross-subject manner. It is expected that the proposed LDA adaptation approach would be used as an important method to increase the usability of fEMG-based FER systems for social VR applications.

Keywords Facial-expression recognition · Facial electromyogram · Riemannian manifolds · Social virtual reality · Linear discriminant analysis adaptation

1 Introduction

With the rapid developments of virtual reality (VR) technology, the traditional social network service (SNS) has evolved into VR-based SNS (Wakeford et al. 2002; Patel et al. 2018). Various social VR applications such as Facebook Horizon,¹ vTime,² AltSpaceVR,³ VRChat,⁴ and BigScreen⁵ have been recently released to the market. Further, the outbreak of COVID-19 has accelerated the growth of VR-based communication services such as VR marketing (Wedel et al. 2020), VR church,⁶ VR conferences (Gunkel

et al. 2018), VR festivals (Kersting et al. 2020), VR education (Freina and Ott 2015), VR social science research (Pan and Hamilton 2018), and VR training (Hui and Zhang 2017).

Since humans are emotional beings, exposing human emotions in an appropriate way in a VR environment is one of the most importance factors for providing VR users with more immersive experiences (Riva et al. 2007; Mottelson and Hornbæk 2020; Rapuano et al. 2020); therefore, demand for recognizing emotional facial expressions of users wearing a head-mounted display (HMD) has been gradually increased. Emotion/facial expression can be useful not only for entertainment but also for the collaboration in a virtual meeting space or for any other application where displaying

✉ Chang-Hwan Im
ich@hanyang.ac.kr

Ho-Seung Cha
hoseungcha@gmail.com

¹ Department of Biomedical Engineering, Hanyang University, 222 Wangsimni-ro, Seoul 133-791, South Korea

¹ <https://www.oculus.com/facebook-horizon/>.

² <https://vtime.net/>.

³ <https://altvr.com/>.

⁴ <https://hello.vrchat.com/>

⁵ <https://www.bigscreenvr.com/>.

⁶ <https://www.vrchurch.org/>.

Table 1 Comparison between existing fEMG-based FER systems

References	Rezazadeh et al. (2011)	Hamedi (2011)	Hamedi (2013)	Chen et al. (2015)	Hamedi et al. (2018)	Cha (2020)
# of fEMG electrodes	6	6	6	2 (Bipolar)	6	8
# of facial expressions	5	8	10	5	11	11
# of participants	3	10	10	6	10	42
# of trials for each facial expression	10	5	5	20	5	20
# of trials for registering a single facial expression	9	4	4	14	4	1
Recognition accuracy (%)	92.60	91.80	87.10	97.10	99.31	85.01

facial-expression is relevant. Indeed, services that provide spaces for social and economic activities in a metaverse are being actively developed (Gunkel et al. 2018; Wedel et al. 2020).

The facial-expression recognition (FER) is generally based on optical cameras (Cohen et al. 2003; Agrawal et al. 2015; Chen et al. 2018; Zhang 2018; Patel and Sakadasariya 2018); however, the camera-based FER has difficulty detecting the facial movements around the eyes, because a large portion of the face is occluded by the VR-HMD (Zhang et al. 2014; Olszewski et al. 2016). To overcome this issue, researchers have attempted to incorporate additional cameras into the VR-HMD (Burgos-Artizzu et al. 2015; Thies et al. 2016; Olszewski et al. 2016). For example, Hickson et al. installed ultrathin strain gauges on the VR-HMD pad to detect the facial movements around the eyes (Hickson et al. 2015). However, these approaches made the VR HMD system bulky and increased the production cost (Hickson et al. 2015).

To address the above issues, facial electromyogram (fEMG) has been recorded around the eyes to recognize facial expressions (Yang and Yang 2011; Fatoorechi et al. 2017; Hamedi et al. 2018; Phinyomark and Scheme 2018; Lou et al. 2020; Cha et al. 2020). An fEMG indicates the electrical activity generated by facial muscle movements, which can be easily recorded using electrodes attached to the face. The fEMG-based FER is a promising alternative to the optical camera-based FER because these systems can be readily implemented using the conventional VR-HMD devices, by simply replacing the existing HMD pad with a new pad containing fEMG electrodes (Fatoorechi et al. 2017; Mavridou et al. 2017). For example, Faceteq™ developed a wearable pad⁷ with electrodes embedded which is also compatible with commercial HMDs. Additionally, the fEMG-based FER system could be fabricated at a lower cost than the optical camera-based FER system because analog front-end (e.g., ADS1298), which is widely utilized for

biosignal acquisition, does not cost as much as the image sensor (e.g., HM01B).

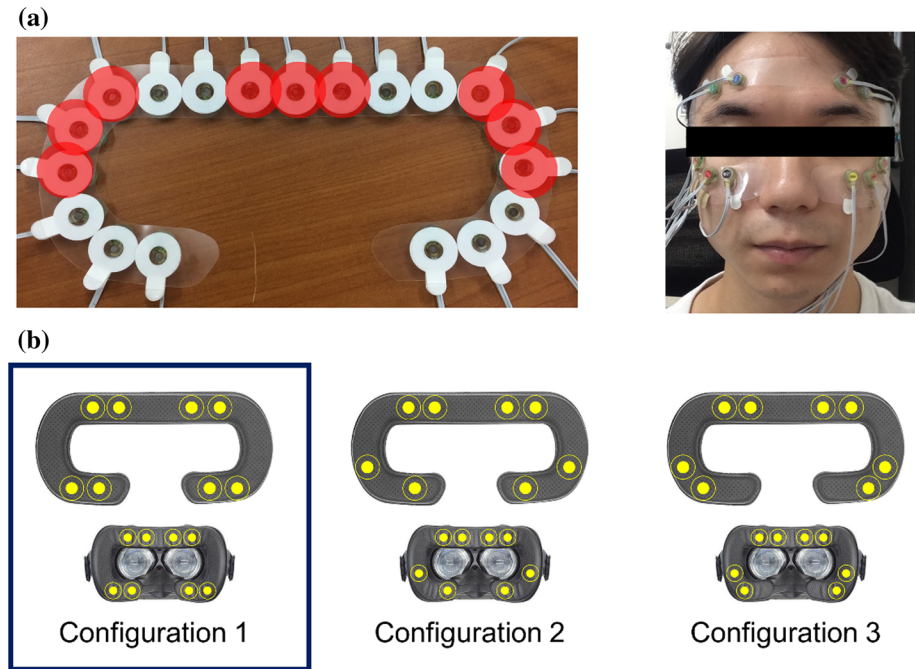
Over the past decades, various fEMG-based FER systems have been proposed as shown in Table 1 (Hamedi et al. 2011, 2013, 2018; Rezazadeh et al. 2011; Cha et al. 2020). It is to be noted that electrode locations reported in these studies were not determined considering VR applications; therefore, electrode locations varied from study to study. The highest classification accuracy reported thus far is 99.83%; in that study, 11 facial expressions were classified by attaching fEMG electrodes to the forehead and both sides of the face (Hamedi et al. 2018). However, this system required users to make facial expression four times for the registration, which does not seem to be practical enough to be used in real VR environments considering that the current FER systems require the users to repeat the registration process whenever they use the system. To address this issue, we suggested a new fEMG-based FER system in which only a single trial for each facial expression was necessary to build the classification model (Cha et al. 2020). We also implemented a real-time FER system with a processing time less than 0.05 s and succeeded in reflecting user's current facial expressions onto a virtual avatar's face in real time.

Nevertheless, the performance of our previous fEMG-based FER system was still inadequate for applications in practical scenarios. In the present study, we developed a new method for improving the FER performance without increasing the size of training datasets. We attempted to use labeled datasets acquired from other users to improve the FER performance. To implement this idea, we adjusted a specific user's linear discriminant analysis (LDA) classifier through the adaptation of additional LDA classifier constructed from other users' labeled datasets, which has never been proposed to the best of our knowledge.

We organized the remainder of this paper as follows. Materials for experiments including electrode placement, reference photographs of emotional faces, and experimental paradigms are introduced in Sect. 2. Methods for data analyses including preprocessing, feature extraction, classification, and LDA adaptation technique are provided in

⁷ <https://www.emteqlabs.com/about-us/>.

Fig. 1 **a** The left figure of the panel shows a plastic film pad on which 19 sticker electrodes are densely attached. The electrodes shaded in red circles are those frequently detached from the facial surface. The right figure of the panel shows a user wearing the electrode pad. **b** Three candidate electrode configurations with eight electrodes tested in our previous study (Cha et al. 2020). The electrode configuration 1 was employed in this study



Sect. 3. Detailed analysis results are reported in Sect. 4. Finally, discussions and conclusions are presented in Sects. 5 and 6, respectively.

2 Materials

2.1 Electrode placement

To determine the optimal electrode placement, a preliminary experiment was conducted. First, we cut a polypropylene plastic clear file in a shape of VR frame; hereafter we call this plastic film. Nineteen sticker electrodes were attached to the plastic film so that the electrodes were densely arranged as shown in Fig. 1a. Next, three male adults were asked to freely move their facial muscles with the plastic film attached on their faces. From this preliminary experiment, it was found that electrodes above specific facial muscles such as the temporalis and corrugator frequently detached from the skin, which are marked with nine red circles in Fig. 1a. Eventually, fEMG was recorded from ten remaining electrodes. Among the ten electrodes, only eight electrodes were selected based on the classification performance evaluated for three different electrode configurations shown in Fig. 1b. According to our previous study (Cha et al. 2020), the highest recognition accuracy was achieved when employing the electrode configuration 1; therefore, we decided to use the first configuration in this study.

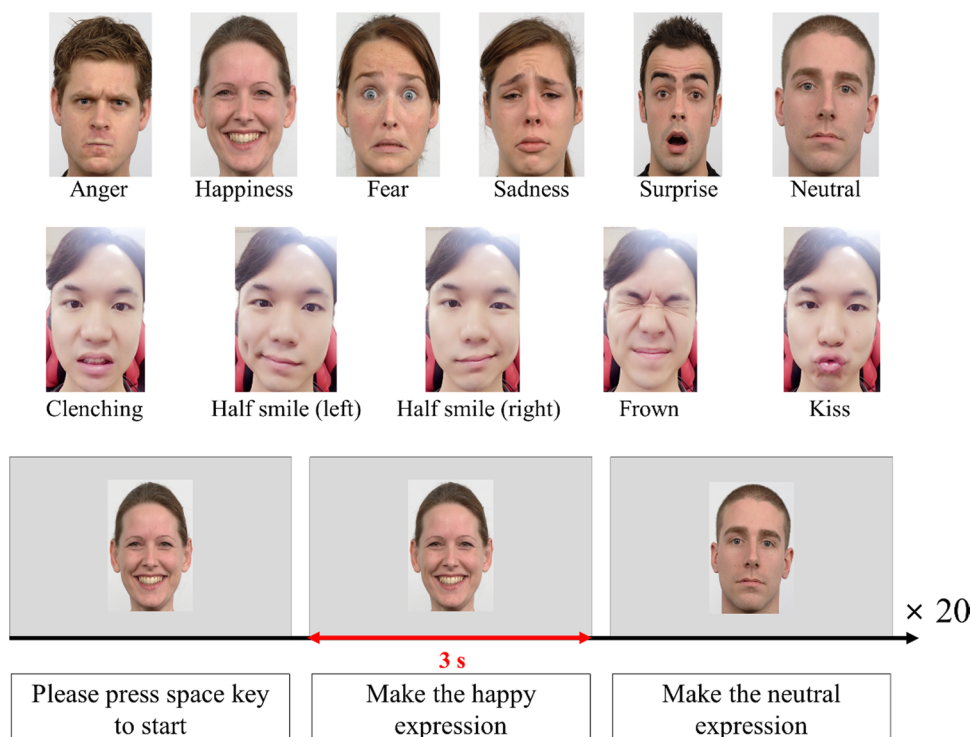
2.2 Photographs of emotional faces

We tried to include as many emotional facial expressions as possible in our FER system based on the previous studies summarized in Table 1; therefore, we decided to employ 11 emotional-face pictures as the reference pictures that participants mimicked during the experiments. Six emotional-face pictures were obtained from the Radboud database (Langner et al. 2010). The Radboud database contained a facial picture set of 67 models displaying emotional expressions based on a facial action coding system (Ekman 1993; Ekman and Rosenberg 2005; Sato and Yoshikawa 2007). The emotions represented in the selected pictures included anger, fear, happiness, neutrality, sadness, and surprise. The six pictures are presented in the first row of Fig. 2. We also took facial pictures of the first author of this paper, while he was making five facial expressions: clenching, half smile (left and right), frown, and kiss. These five pictures are presented in the second row of Fig. 2.

2.3 Participants

Forty-two healthy native Korean participants (17 males and 25 females) volunteered to participate in this study. Their ages ranged from 21 to 29 years (mean = 24.07, standard deviation = 1.89). No participants reported severe health problems that could have affected the experiment, e.g., Bell's palsy, stroke, or Parkinson's disease. They all were instructed to not to drink alcohol and sleep enough to avoid the psychological health problem during experiments

Fig. 2 Eleven facial-expression picture stimuli and the experimental procedure



before the day of the experiment. All the participants were provided with a detailed explanation of the experimental protocols and signed a written informed consent form. The study protocol was approved by the Institutional Review Board (IRB) of Hanyang University, South Korea (IRB No. HYI-14–167-11).

2.4 Experimental procedure

fEMG data were collected at a sampling rate of 2048 Hz using a Biosemi Active Two system (Biosemi, B.V., Amsterdam, The Netherlands). The recording system included two additional electrodes—common mode sense (CMS) and driven right leg (DRL)—which were used as reference and ground channels, respectively. We attached the CMS and DRL electrodes to the left and right mastoids, respectively.

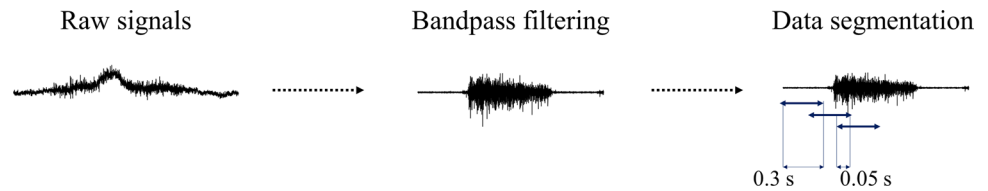
Before the main experiment, a short training period was provided for the participants to become accustomed to mimicking the 11 emotional faces shown in Fig. 2. The selected emotional-face pictures were presented on a computer monitor using E-prime 2.0 (Psychology Software Tools, Sharpsburg, PA, USA). During the experiment, each participant mimicked the 11 emotional faces presented on the monitor repeatedly 20 times. Note that we repeated 20 times based on the maximum repetitions reported in the previous studies (see Table 1). The overall experimental procedure for a single trial is presented in the bottom panel of Fig. 2. First, a reference emotional picture for the participant to mimic (e.g., happy face) was presented on the monitor. The

participant pressed the space bar when he/she was ready to move to the next step. After the space bar was pressed, a short “beep” sound was generated, and the participant mimicked the emotional-face picture for 3 s. After the 3 s, the participant made a neutral facial expression and waited for the next trial. The 11 emotional-face pictures were randomly presented, to reduce the possibility of an order effect. This procedure would be needed to be done for every user to generate a user-customized classifier model in the application of the proposed FER system to practical VR applications. It is to be noted that only a single training trial per each facial expression is needed for the generation of the user-customized classifier model in our study. Each participant completed a total of 220 trials (11 facial expressions \times 20 repetitions). The corresponding dataset (.bdf format) is available at <https://doi.org/10.6084/m9.figshare.9685478.v1>. It is expected that this dataset can be utilized to develop new algorithms to enhance the overall performance of the fEMG-based FER system in a VR-HMD environment.

3 Methods

The fEMG-based FER system is a pattern-recognition-based myoelectric interface, similar to a multifunction prosthesis. (Asghari Oskoei and Hu 2007; Hakonen et al. 2015; Geethanjali 2016; Phinyomark and Scheme 2018). The multifunction prosthesis which provides multiple control options allows patients to manipulate prosthesis in more flexible

Fig. 3 Signal preprocessing stages



manner. To enable the multiple options, various pattern recognition techniques have been developed in many literatures. As for other myoelectric interfaces, the data-analysis procedure of the fEMG-based FER system includes preprocessing, feature extraction, and classification (Hakonen et al. 2015). In this section, we introduce the three data-analysis procedures and then describe the concept of LDA adaptation with labeled datasets of other users in a detailed manner.

3.1 Preprocessing

Figure 3 shows stages of the fEMG signal preprocessing. The fEMG signals recorded from eight electrodes were notch-filtered at 60 Hz and bandpass-filtered at 20–450 Hz using a fourth-order Butterworth filter. The filtered fEMG signals were split into a series of short segments using a sliding window. The sliding window is one of the digital signal processing techniques; once a virtual window is set with its size, the window is sled with a fixed size until the window reached the end of a signal. In this study, we simply truncated the signal before and after of the window at each window. We set the sliding-window length to 300 ms and moved the sliding window from 0 ms to the end of the signal with a fixed time interval of 50 ms. According to the average fEMG onset time of 1.02 ± 0.34 s after the presentation of the beep sound (Cha et al. 2020), the first 1 s of the fEMG signals was excluded from the analysis; only the last 2 s of the fEMG signals was used.

3.2 Feature extraction in Riemannian manifold

The Riemannian manifold is a real, smooth (differentiable) manifold in which a finite-dimensional Euclidean space is defined on a tangent space at a point (Förstner and Moonen 2003; Wang et al. 2012; Morerio and Murino 2017). The space of a symmetric and positive-definite (SPD) matrix becomes a Riemannian manifold (Förstner and Moonen 2003; Wang et al. 2012; Morerio and Murino 2017); therefore, an SPD matrix can be considered as a point on a Riemannian manifold. This property allows a covariance matrix to be used in the Riemannian manifold, because the covariance matrix has the properties of the SPD matrix. Unfortunately, in the Riemannian manifold, mathematical operations defined in the Euclidean space cannot be employed. To deal with the SPD matrix in the Euclidean manner, Arsigny et al. (Arsigny et al. 2007) proposed a logarithmic map defined as

$$S = C_r^{\frac{1}{2}} \log m \left(C_r^{-\frac{1}{2}} C C_r^{-\frac{1}{2}} \right) C_r^{\frac{1}{2}}, \quad (1)$$

where $\log m(\cdot)$ represents the logarithm of a matrix and C represents an SPD matrix. This equation allows C on a Riemannian manifold to be mapped to S on a tangent space generated by a reference point C_r . A tangent space on a Riemannian manifold is locally isometric to a Euclidean space. Barachant et al. (Barachant et al. 2010, 2013) employed this approach and utilized the upper triangular elements of S as features in an electroencephalography-based brain–computer interface.

For each fEMG segment $D \in R^{E \times S}$, a sample covariance matrix (SCM) C can be computed as $1/(S-1)DD^T \in R^{E \times E}$, where S and E represent the number of samples and fEMG channels, respectively. Before the SCM is projected onto a tangent space, the reference point C_r for forming the tangent space should be determined. We followed the approach of Barachant et al., who employed the reference point as a geometric mean of the SCMs in the training dataset (Barachant et al. 2010, 2013). The geometric mean is the mean of the SCMs in the Riemannian sense, and the algorithm for computing it is presented in Appendix 1. After the reference point C_r was computed, an SCM C was mapped onto S in a tangent space using (1). Finally, the upper triangular elements of S were used as the features, which constituted the vector x . The number of feature dimensions was 36 ($= 8 \times 9 / 2$).

3.3 Classification

Our preliminary test showed that the average classification accuracies achieved by using LDA, support-vector machine, tree, and k-nearest neighbors were 85.01, 79.14, 81.06, and 81.06%, respectively. Based on these results, we chose LDA as the classification algorithm. LDA is one of the most frequently used algorithms for myoelectric interfaces (Hakonen et al. 2015). The LDA model can be statistically derived by assuming that the data within each class follow a multivariate normal distribution (Morrison 1969). Let the fEMG feature vector and a facial-expression class label be x and k , respectively; then, the feature vector x can be predicted as follows:

$$\hat{y} = \operatorname{argmax}_k \varphi_k(x) \quad (2)$$

where \hat{y} represents the predicted label and $\varphi_k(\mathbf{x})$ represents the decision function. The decision function $\varphi_k(\mathbf{x})$ is defined as

$$\varphi_k(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_k - \frac{1}{2} \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_k + \log(\pi_k) \quad (3)$$

where $\boldsymbol{\mu}_k \in \mathbb{R}^{36}$ is a mean vector of features corresponding to label k and $\boldsymbol{\Sigma} \in \mathbb{R}^{36 \times 36}$ is a pooled covariance matrix (PCM). The $\boldsymbol{\mu}_k$ for every class label ($k = 1, 2, \dots, 11$) and $\boldsymbol{\Sigma}$ can be estimated using the training dataset. The estimation of $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}$, as well as the derivation of the decision function, is presented in detail in Appendix 2.

The first trials for each facial expression were used as the training datasets, and the remaining 19 trials were used as the test datasets to evaluate the performance of our FER system. No samples were excluded from the original dataset. We defined the classification accuracy as the number of correctly classified samples divided by the total number of samples.

3.4 LDA model adaptation with labeled datasets

We employed only a single trial (first trial) as the training dataset; thus, the user's LDA model could easily be overfitted, degrading the FER performance. To enhance the FER performance, we attempted to generalize the user's LDA model by adapting it with another LDA model constructed using datasets of other users. We assumed that these datasets were already collected; therefore, no additional training datasets of the user were required. Hereinafter, the dataset collected from other users is denoted as DB (representing "database").

Let $\boldsymbol{\mu}_{tr_k}$ and $\boldsymbol{\Sigma}_{tr}$ be the mean vector and the PCM of a user's training dataset, respectively. Similarly, let $\boldsymbol{\mu}_{DB_k}$ and $\boldsymbol{\Sigma}_{DB}$ be the mean vector and the PCM of the dataset of other users (DB), respectively. We applied two shrinkage parameters (α and β) to the two mean vectors ($\boldsymbol{\mu}_{tr_k}$ and $\boldsymbol{\mu}_{DB_k}$) and the two PCMs ($\boldsymbol{\Sigma}_{tr}$ and $\boldsymbol{\Sigma}_{DB}$), as follows:

$$\tilde{\boldsymbol{\mu}}_k = (1 - \alpha) \boldsymbol{\mu}_{tr_k} + \alpha \boldsymbol{\mu}_{DB_k} \quad (4)$$

$$\tilde{\boldsymbol{\Sigma}} = (1 - \beta) \boldsymbol{\Sigma}_{tr} + \beta \boldsymbol{\Sigma}_{DB} \quad (5)$$

where $0 \leq \alpha, \beta \leq 1$; $\alpha, \beta \in \mathbb{R}$; and $\tilde{\boldsymbol{\mu}}_k$ and $\tilde{\boldsymbol{\Sigma}}$ are the newly adapted mean vector and PCM, respectively. This adaptation strategy was adopted from previous studies (Zhang et al. 2013; Vidovic et al. 2014, 2016); however, our approach differed from the previous ones in that we performed the LDA adaptation among different users (i.e., cross-subject settings), whereas in the previous studies (Zhang et al. 2013; Vidovic et al. 2014, 2016), LDA adaptation was performed for the same user and different sessions (cross-session settings).

To investigate the effect of the size of DB on the FER performance, we prepared various DBs that included different numbers of participants. We increased the number of participants from 0 to 41 ($n = 0, 1, \dots, 41$). Then, we conducted the LDA adaptation using (4) and (5). The maximum number of participants that could be included in DB was 41, because 42 participants were recruited for this study. Here, $n = 0$ indicates that no adaptation was performed.

Two different strategies were used for selecting n participants for constructing DB: (1) *rnDB*, i.e., randomly selecting n participants among 41 participants, and (2) *riDB*, i.e., selecting n participants in the order of closest distance between the user's training dataset and other user's dataset. We measured the Riemannian distances. Specifically, we first computed the geometric mean of a user's training dataset (\mathbf{C}_r^{tr}) and the geometric mean for another participant p (\mathbf{C}_r^p). Next, we computed all the distances in the Riemannian manifold between \mathbf{C}_r^{tr} and \mathbf{C}_r^p and selected n participants in the ascending order of the Riemannian distances. The distance between \mathbf{C}_1 and \mathbf{C}_2 on a Riemannian manifold is defined as

$$\delta_R(\mathbf{C}_1, \mathbf{C}_2) = \log(\mathbf{C}_1^{-1} \mathbf{C}_2)_F = \left[\sum_{i=1}^n \log^2 \lambda_i \right]^{\frac{1}{2}} \quad (6)$$

where λ_i represents the real positive eigenvalues of $\mathbf{C}_1^{-1} \mathbf{C}_2$.

There were two methods for selecting the reference points for the tangent space when the Riemannian features were extracted from DB: 1) using the geometric mean of a user's training dataset (\mathbf{C}_r^{tr}) and 2) using the geometric mean of DB (\mathbf{C}_r^{DB}).

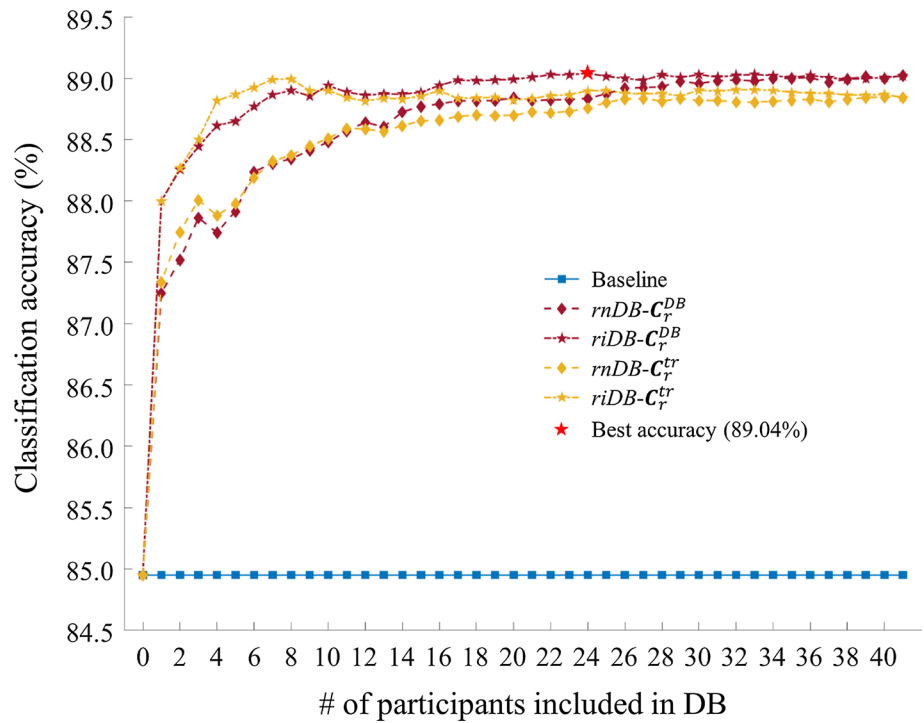
With the combination of participant selection strategies (*rnDB* and *riDB*) and reference-point selection strategies to include in DB (\mathbf{C}_r^{tr} and \mathbf{C}_r^{DB}), four adaptation approaches were available, which were denoted as *rnDB-C_r^{DB}*, *rnDB-C_r^{tr}*, *riDB-C_r^{DB}*, and *riDB-C_r^{tr}*. For each approach, the common α and β for all the participants were optimized with regard to the classification accuracy via a grid search. Specifically, we computed the average classification accuracies while varying the α and β values from 0 to 1 with a fixed step size of 0.1 (i.e., 0, 0.1, 0.2, 0.3, ..., 0.9, 1). Next, we set α and β to the values yielding the highest classification accuracy.

4 Results

4.1 Determination of optimal LDA adaptation approach

We determined the optimal LDA adaptation approach according to the average classification accuracy. Figure 4 shows the classification accuracy as a function of the number

Fig. 4 Classification accuracy as a function of the number of participants included in DB for the baseline and four LDA adaptation approaches. The highest accuracy is marked in red star



of participants included in DB for four different LDA adaptation approaches. The baseline represents the condition where no LDA adaptation was applied. Except for the baseline, the classification accuracy tended to increase with the number of participants included in DB. When C_r^{DB} was used as the reference point, a higher accuracy could be achieved regardless of the DB selection strategy ($rnDB$ or $riDB$). When the $rnDB$ strategy was employed, a larger number of participants was needed to achieve a similar accuracy level, compared with the case where the $riDB$ strategy was employed. Among the four LDA adaptation approaches, $riDB-C_r^{DB}$ yielded the highest accuracy (89.04%) when 24 participants were included in DB (as indicated by the red star in Fig. 4).

4.2 Analysis of LDA shrinkage parameters

Figure 5 presents the classification accuracy for different values of the parameters α and β , in the case where $riDB-C_r^{DB}$ was employed. As shown in Fig. 5, the classification accuracy reached 89.04% at $\alpha = 0.5$ and $\beta = 0.1$. This accuracy was 4.09 pp (percentage point) higher than that for the no-adaptation condition ($\alpha = 0$ and $\beta = 0$), which was 85.04%. A Wilcoxon signed-rank test indicated that the difference in classification accuracy between the baseline (no adaptation) and the optimal LDA adaptation condition was statistically significant ($p < 0.001$). Interestingly, a high accuracy of 82.97% was achieved using an LDA model constructed solely with DB (i.e., $\alpha = 1$ and $\beta = 1$), indicating the potential of the user-independent

FER system. The lowest classification accuracy (78.01%) was observed when the mean vectors of DB and the PCM of the user’s training data were used ($\alpha = 1$ and $\beta = 0$).

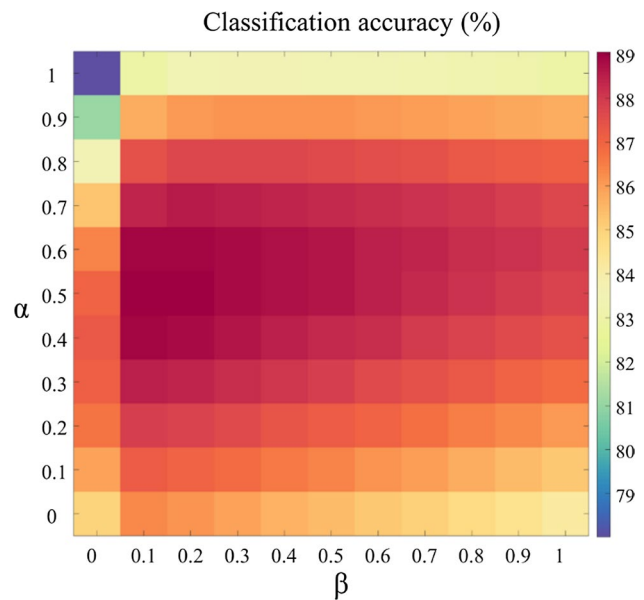


Fig. 5 Classification accuracy for the optimal LDA adaptation condition with respect to the LDA shrinkage parameter α and β . Each color is mapped to a specific accuracy

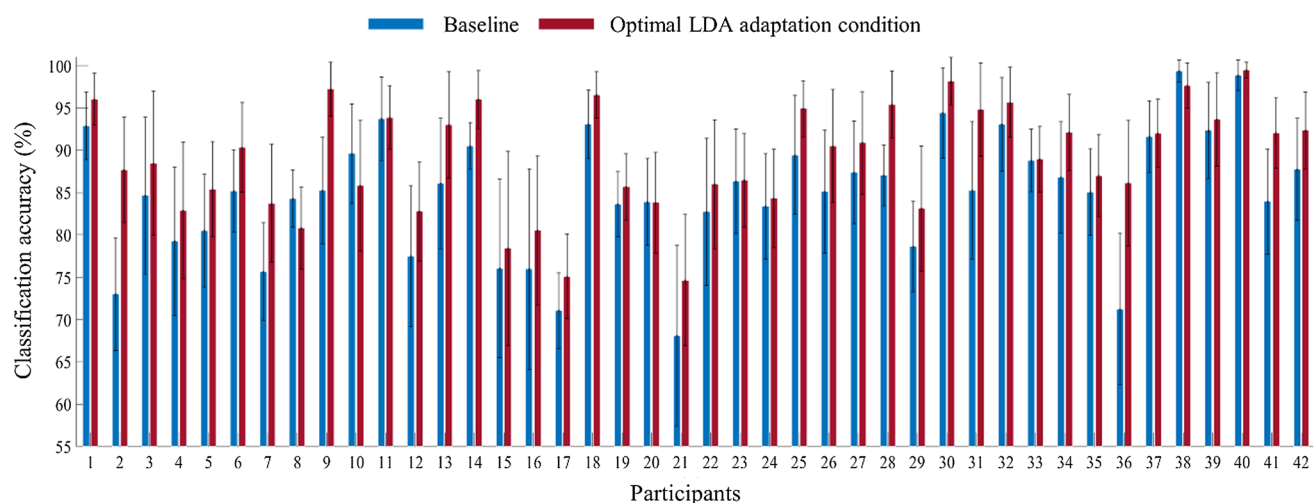


Fig. 6 Classification accuracy for each of the 42 participants for the baseline and optimal LDA adaptation conditions. The error bars indicate the standard deviations

4.3 Further analysis with optimal LDA adaptation condition

Figure 6 shows the classification accuracies for each of the 42 participants, relative to the baseline (no adaptation) and optimal LDA adaptation conditions. The error bars indicate the standard deviations across 19 test trials. The classification accuracies for all the participants except four were increased by employing the LDA adaptation approach. The three largest accuracy increments between the baseline and optimal LDA adaptation conditions were observed for participants No. 36, 2, and 9; the increments were $14.88 \text{ pp} \pm 10.15 \text{ pp}$, $14.67 \text{ pp} \pm 5.60 \text{ pp}$, and $11.99 \text{ pp} \pm 6.46 \text{ pp}$, respectively.

Figure 7 presents the recall, precision, and F1 score (percentage) for each expression, relative to the baseline and optimal LDA adaptation conditions. The F1 score was computed using the harmonic mean of the recall and precision. The facial expressions on the three bar graphs were arranged in the order of decreasing accuracy relative to the optimal LDA adaptation. The recall, precision, and F1-score values were increased for all the facial expressions when the optimal LDA adaptation was utilized, except for the recall for happiness. The recall for the happiness expression was slightly reduced (by 0.75 pp) from 96.01% but still remained high (95.26%). The three facial expressions with the largest increases in the F1 score were fear, kiss, and anger, with increments of 7.58 pp, 6.67 pp, and 6.19 pp, respectively.

4.4 Confusion analysis

Figure 8 presents the confusion matrices of the classification results for the baseline and optimal LDA adaptation

conditions. The diagonals of the confusion matrices indicate the recalls. The facial-expression labels in the two confusion matrices were arranged in the order of decreasing recall for the facial expressions of the baseline. The average decrease for all the confusions was 0.41 pp. The top five largest decreases in the confusion were observed when (1) fear was misclassified as surprise, (2) surprise was misclassified as fear, (3) anger was misclassified as a frown, (4) sadness was misclassified as anger, and (5) clenching was misclassified as fear. The decreases in these five cases were 3.42 pp, 3.37 pp, 3.23 pp, 2.95 pp, and 2.94 pp, respectively. Although the average confusions were reduced, confusions for angry and surprise were increased for some participants (participant 8 and 38), leading to the deterioration of overall FER performance of those participants. Introduction of an improved technique to further elevate the FER performance after the LDA adaptation might be necessary in future studies.

4.5 Online demonstration

Figure 9 shows a snapshot of the online experiment taken when a participant was making a happy facial expression. It can be seen that a virtual avatar is mimicking the facial expression of the participant. Note that the electrodes for acquiring the fEMG signals were directly attached to the commercial HMD pad in this online demonstration. Classification decision was made at every 0.05 s (20 frames per second). The demonstration video can be found at https://youtu.be/9_VFJrZ-0Gk.

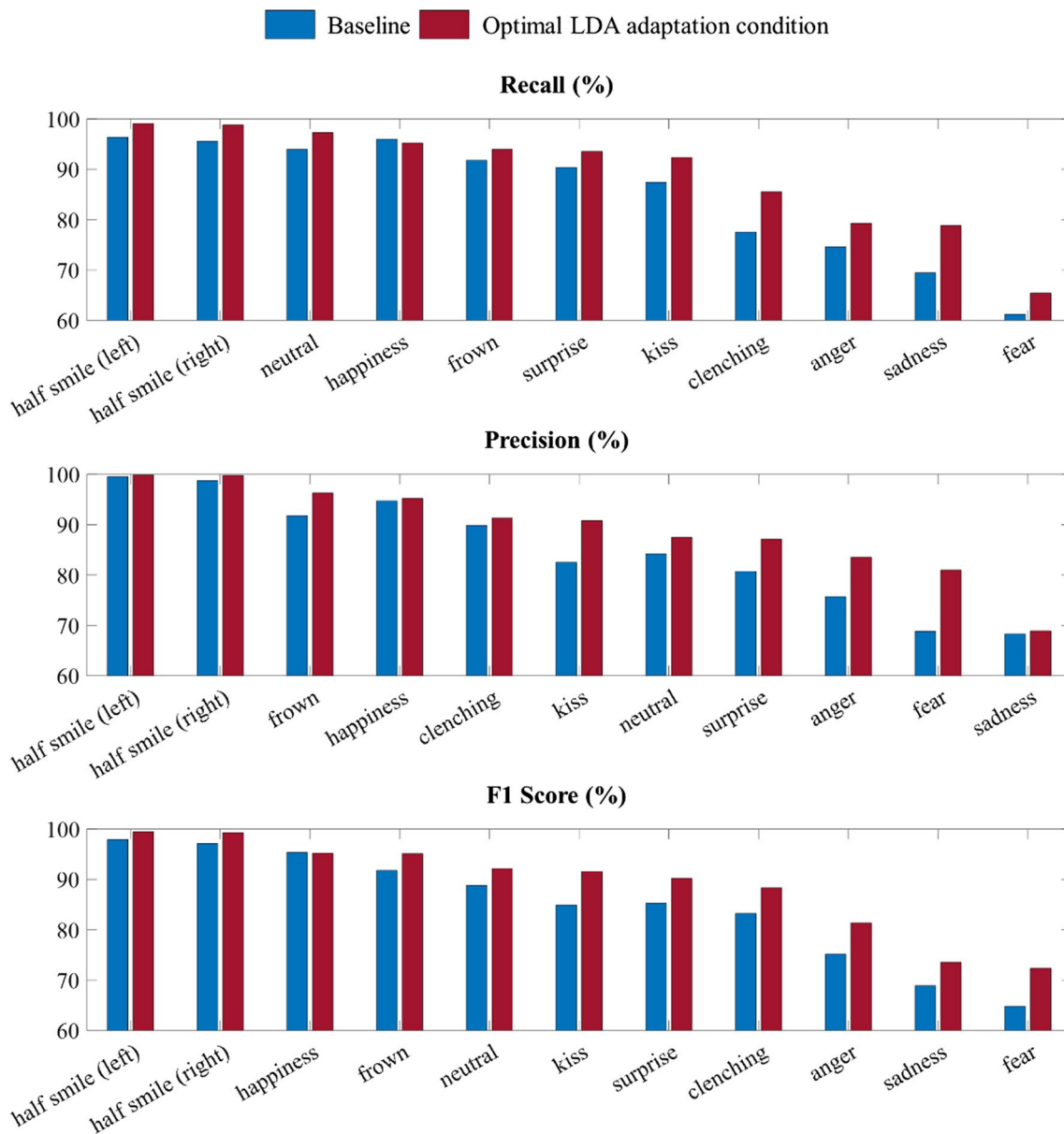


Fig. 7 Recall, precision, and F1 score for each facial expression. The F1 score was the harmonic mean of the recall and precision. The facial expressions on the three bar graphs were arranged in the order of decreasing accuracy for the optimal LDA adaptation condition

5 Discussion

We improved the performance of fEMG-based FER using LDA model adaptations with datasets of other users. In our previous study, we implemented an fEMG-based FER system that requires only a single training dataset, but performance degradation was inevitable owing to the limited training dataset (Cha et al. 2020). The objective of the present study was to enhance the FER performance without collecting an additional training dataset from the user. To this end, we adjusted the LDA shrinkage parameters of the user according to those of other users. To the best of our

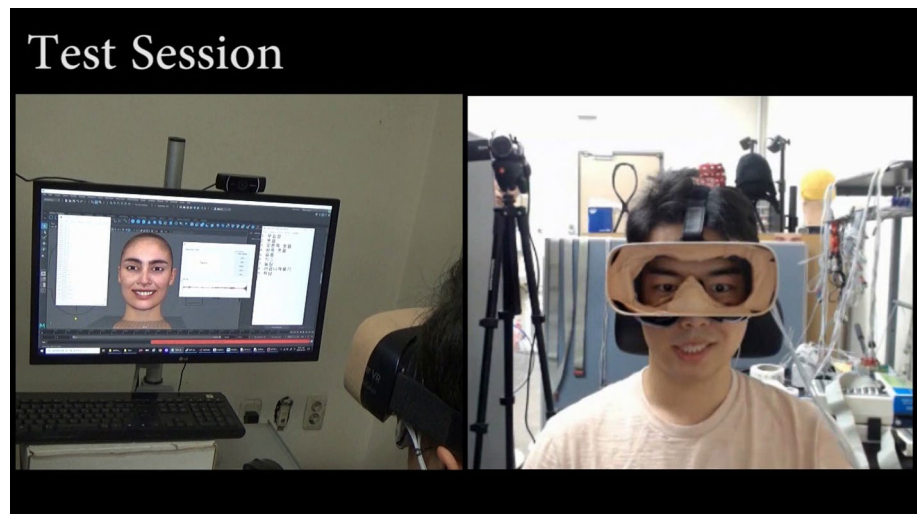
knowledge, this was the first study in which the LDA adaptation approach was employed in a cross-subject manner. We believe that our technique being able to mirror the user's face onto their avatars' faces will be practically utilized in social VR or any other applications requiring personal virtual avatar.

As shown in Fig. 4, classification accuracy was increased as the number of participants included in DB was increased. This might be explained as follows: the original LDA model, which was overfitted owing to the limited training dataset (only a single training dataset), became more generalized via the LDA adaptation with large datasets from other users.

		Baseline										Optimal LDA adaptation condition													
True Labels	Half smile (left)	96.4%	0.6%	0.1%	0.5%	0.2%		0.4%	0.7%	1.0%	0.1%	0.0%	99.1%	0.3%	0.1%	0.2%			0.1%	0.1%	0.1%	0.0%			
	Happiness	0.0%	96.0%	0.1%	0.6%	0.1%	0.0%	0.1%	1.4%	0.7%	0.0%	0.8%	0.0%	95.3%	0.0%	0.8%	0.1%		0.1%	1.9%	1.0%	0.3%	0.7%		
	Half smile (right)		0.7%	95.6%	0.8%	0.0%		1.6%	0.4%	0.2%	0.0%	0.7%		0.1%	98.8%	0.1%			0.0%	0.4%	0.3%	0.0%	0.2%		
	Neutral		0.0%	0.0%	94.0%		0.6%	1.6%	0.4%	0.2%	2.1%	1.1%	0.0%		0.0%	97.3%		0.3%	0.4%	0.2%	0.2%	1.0%	0.5%		
	Frown		0.2%		0.6%	91.8%	0.1%	1.2%	0.8%	3.9%	0.5%	1.0%		1.1%		0.3%	94.0%		0.2%	0.8%	2.7%	0.2%	0.7%		
	Surprise		0.0%		1.2%	0.1%	90.4%	0.1%	0.0%	0.2%	1.5%	6.4%		0.0%		1.5%	0.1%	93.6%	0.0%		0.1%	1.7%	3.1%		
	Kiss	0.0%	0.2%	0.1%	5.1%	0.1%	1.1%	87.4%	0.6%	0.4%	3.8%	1.2%	0.0%		2.7%		0.6%	92.4%	0.8%	0.2%	2.8%	0.6%			
	Clenching	0.2%	1.7%	0.6%	2.5%	1.0%	0.1%	7.8%	77.6%	2.5%	0.8%	5.2%	0.0%	1.9%	0.0%	2.8%	0.9%	0.1%	5.2%	85.5%	0.5%	0.8%	2.2%		
	Anger	0.1%	0.4%	0.0%	1.4%	4.8%	2.0%	2.7%	0.7%	74.6%	10.1%	3.2%	0.1%	0.2%		0.7%	1.6%	1.0%	1.7%	0.8%	79.3%	13.5%	1.2%		
	Sadness	0.0%			3.6%	1.1%	5.4%	1.4%	0.5%	10.2%	69.5%	8.2%				2.9%	0.4%	3.0%	0.9%	0.5%	7.3%	78.8%	6.2%		
	Fear	0.0%	1.5%	0.2%	1.4%	0.8%	12.3%	1.5%	3.2%	4.5%	13.4%	61.2%		1.2%	0.0%	2.0%	0.6%	8.8%	0.7%	2.5%	3.3%	15.3%	65.4%		
			Half smile (left)	Happiness	Half smile (right)	Neutral	Frown	Surprise	Kiss	Clenching	Anger	Sadness	Fear	Half smile (left)	Happiness	Half smile (right)	Neutral	Frown	Surprise	Kiss	Clenching	Anger	Sadness	Fear	
			Predicted Labels												Predicted Labels										

Fig. 8 Confusion matrices of the classification results for the baseline and optimal LDA adaptations. The facial-expression labels on the two confusion matrices were arranged in the order of decreasing recall of the baseline (the diagonals of the confusion matrices indicate the recalls)

Fig. 9 A snapshot of the online experiment taken when a participant was making a happy facial expression (the demonstration video can be found at https://youtu.be/9_VFJrZ-0Gk)



However, increasing patterns of classification accuracy differed among the four LDA adaptation strategies. The reason why the classification accuracy increased more rapidly when selecting the data in terms of the Riemannian distance ($riDB$) might be that the data with similar distributions with the user's training dataset were chosen first. Therefore, this strategy would be useful when DB is not sufficiently collected. On the other hand, the reason why the classification accuracies when a full DB was used were different depending on the selection of the reference covariances (C_r^r and C_r^{DB}) might be explained by the generalization of LDA parameters. When features were extracted from the user's

domain (C_r^r), features that had similar distribution with the user's features could be extracted. This might lead to overfitting of LDA parameters, and thus the classification accuracy would not be increased much. Based on this result, selection of C_r^{DB} as the reference covariance is highly recommended to improve the overall FER performance.

We found the optimal LDA parameters α and β , which can be universally applied for all users. Our analysis results for the variation of the classification accuracy with respect to α and β indicated that mean vector μ had a significantly larger effect on the performance than the PCM Σ . A similar effect of the mean vector in LDA adaptation was observed

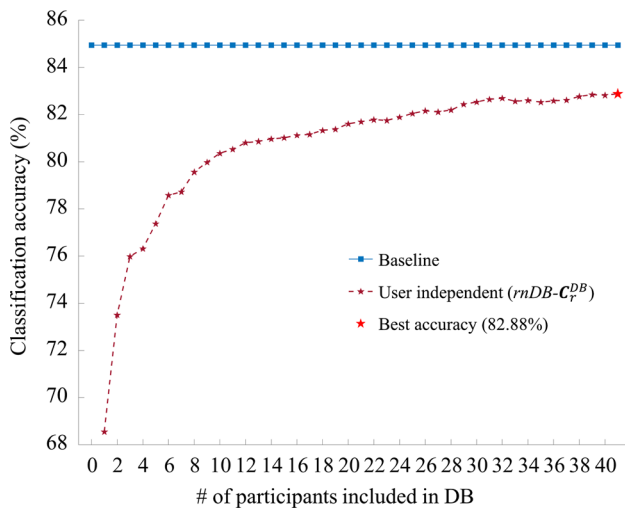


Fig. 10 Classification accuracy as a function of the number of participants included in DB for the baseline and user-independent conditions ($rnDB-C_r^{DB}$). The highest accuracy for the user-independent condition is marked in red

in previous studies (Vidovic et al. 2014, 2016), although the LDA adaptation was conducted using datasets of the same participants. Nevertheless, the adaptation for the PCM was still necessary for enhancing the overall performance of the FER system. Our results indicated that the highest classification accuracies for each value of α were always achieved with $\beta \neq 0$ or $\beta \neq 1$.

A user-independent FER system is a system that users can employ without a training session. Thus, the development of a user-independent practical FER system is an important goal (Matsubara and Morimoto 2013; Khushaba 2014; Xiong et al. 2015; Kim et al. 2017). In this study, our FER system became user independent under a specific condition, i.e., when the $rnDB-C_r^{DB}$ approach was employed with $\alpha = \beta = 1$. To confirm the feasibility of the user-independent system, we computed the classification accuracy in this condition. The results are presented in Fig. 10. Interestingly, the classification accuracy increased with the number of participants. The highest accuracy (82.88%) was achieved when all 41 participants' data were employed for the training. Although this accuracy was lower than that of the baseline system (85.04%) trained with the user's own dataset, the result is promising in that no training dataset from the user was required. The high accuracy is explained as follows: the geometric mean of the large DB yielded a large tangent space, which was helpful for making the feature distributions of the specific user and the other users similar. In the future study, we plan to develop an online user-independent FER system with a better performance, which is expected to increase the practicality of the FER system in many VR applications as the VR users can use the FER function without a need for a cumbersome registration session.

Our study has the following ripple effects: (1) This study can accelerate and expand the metaverse world by adding facial-expression recognition to VR avatars. The biggest drawback of the avatars in current social VR services is that they fail to convey users' emotional expressions. This greatly reduces VR users' level of immersion and acts as an obstacle to natural communication between users. The proposed method that can maximize FER performance with only a single training dataset can greatly contribute to building a huge metaverse world of the future. (2) The datasets available from this study can contribute to meaningful research exchanges with interested researchers on how to analyze data in VR environments. Unlike the data available in general environments, the data analyzed in this study are based on VR environments. It can be of great value to several researchers and companies interested in analyzing biosignal data in VR environments. (3) This study can contribute to the expansion of VR convergence research by increasing interest in applying biosignal analysis in VR environments. Recent advances in VR-based digital healthcare (Buettner et al. 2020) have made it increasingly important to monitor biosignals in VR environments. We hope this study help expand various research areas in VR environments.

6 Conclusion

In this study, we succeeded in improving the performance of fEMG-based FER by using LDA adaptation in the Riemannian manifold without any additional training dataset of the user. However, for the system to be used in realistic scenarios, its limitations must be considered. First, the test/retest reliability should be tested to determine whether the LDA adaptation method is still feasible in cross-session environments. It is well known that the user's data domain can be affected by several factors, e.g., electrode shifts, humidity changes, and impedance changes (Young et al. 2012; Muceli et al. 2014; Li et al. 2016; Vidovic et al. 2016). Second, new domain adaption technique based on deep learning should be researched. One sample Kolmogorov–Smirnov test for the EMG features rejected the hypothesis that the features are not normally distributed, which is opposite to the assumption of the LDA that data are normally distributed. This indicated that the LDA might not be the best option for our system. It will be interesting to develop new deep learning-based domain adaption technique which is applied well with the EMG data. Third, our method must be validated using a dataset collected from a dry electrode-based EMG recording system. The portable biosignal acquisition system is generally susceptible to external noise and artifacts. Thus, an additional signal-processing method for denoising would be helpful for the development of a robust fEMG-based FER system. Fourth, our adaptation method resulted in better

performance when the stimuli were static pictures of facial expressions, but it has not yet been tested in more realistic settings (e.g., presentation of video stimuli or natural interaction with others). Further investigation needs to be conducted under more realistic environments so that the proposed method can be utilized in practical VR applications. Fifth, the current electrode systems need to be further developed. Further studies are needed to enhance the attachability to the curved facial surface, increase robustness against temperature changes or sweats of the skin, and improve the recorded signal quality. Recently, ultra-thin, flexible, and breathable electrodes are actively developed as the substitute of the current rigid electrodes (Fan et al. 2020), which is expected to be incorporated with the VR-HMD system in the near future. Lastly, directly capturing facial motions in a regression manner rather than a classification manner could be effective for developing a practical FER system. The pattern-classification method does not provide a solution to deal with unregistered facial expressions, which were not present in the training dataset. Thus, a regression-based FER approach should be investigated, which is an interesting research topic.

Appendix 1

Algorithm: Geometric mean of N SCM matrices

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

Output: The estimated geometric mean C_r .

Initialize: $C_r^{(t)} = \frac{1}{N} \sum_{n=1}^N C_n$, $t = 1$, and $\epsilon = 10e - 5$

While

$$\bar{S} = \frac{1}{N} \sum_{n=1}^N \log_m C_n^{(t)}$$

$$C_r^{(t+1)} = \exp_{C_r^{(t)}}(\bar{S})$$

$$t = t + 1$$

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

Return $C_r^{(t+1)}$

$\log_C(C_w)$ is defined as $C^{\frac{1}{2}} \log_m(C^{-\frac{1}{2}} C_w C^{-\frac{1}{2}}) C^{\frac{1}{2}}$, where $\log_m(\cdot)$ denotes the logarithm of a matrix. $\exp_C(\bar{S})$ is defined as $C^{\frac{1}{2}} \exp_m(C^{-\frac{1}{2}} \bar{S} C^{-\frac{1}{2}}) C^{\frac{1}{2}}$, where $\exp_m(\cdot)$ denotes the exponential of matrix. $\|\cdot\|_F$ denotes Frobenius Norm.

Appendix 2

The LDA model can be statistically derived by assuming that distribution of data within each class follows a multivariate normal distribution. Let the random variables $X \in \{x_1, x_2, x_3, \dots\}$ and $L \in \{1, 2, \dots, k\}$ represent the feature vector and label, respectively. According to the assumption of LDA, the probability density function of the d -dimensional feature vector $x \in \mathbb{R}^d$ within the class label k can be given as

$$p(X = x|L = k) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma^{-1}(x-\mu_k)} \quad (7)$$

where $\mu_k \in \mathbb{R}^d$ is the mean vector of the features corresponding to label k , and $\Sigma \in \mathbb{R}^{d \times d}$ is a PCM. μ_k and Σ can be estimated as follows:

$$\mu_k = \sum_{\forall i.s.t.L=k} x / N_k \quad (8)$$

$$\Sigma = \frac{1}{N - K} \sum_{k=1}^K \sum_{\forall i.s.t.L=k} (x - \mu_k)(x - \mu_k)^T \quad (9)$$

where N , N_k , and K represent the total number of samples, the number of samples corresponding to label k , and the total number of labels, respectively.

The posterior probability that the label is k given $X = x$ can be expressed according to Bayes' rules:

$$p(L = k|X = x) = \frac{p(L = k)p(X = x|L = k)}{p(X = x)} \quad (10)$$

Let $p(L = k)$ and $p(X = x|L = k)$ be π_k and $f_k(x)$, respectively. Then, the numerator can be rewritten as $\pi_k f_k(x)$. Considering that $\pi_k f_k(x)$ is a monotonic increment function, $\pi_k f_k(x)$ is proportional to $\log(\pi_k f_k(x))$. We can employ $\log(\pi_k f_k(x))$ as the decision function $\varphi_k(x)$, which is expressed as

$$\varphi_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log(\pi_k) \quad (11)$$

Finally, a class label k is predicted in the test stage using the following equation:

$$\hat{y} = \underset{k}{\operatorname{argmax}} \varphi_k(x) \quad (12)$$

In summary, μ_k and Σ are estimated using the training dataset via (2) and (3) in the training stage, and a new feature vector x is predicted via (6) in the test stage.

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

Author contributions HC wrote a major part of the paper (Introduction, Methods, Results, Discussion, and Conclusion) and conducted main analyses. CI provided important insight for the design of the paper and revised the manuscript. All authors listed have made considerable contribution to this paper and approved the submitted version.

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

Declarations

Conflicts of interest The authors declare that there are no conflicts of interest regarding the publication of this paper.

Availability of data and material The raw dataset was uploaded to Figshare, which is available at <https://doi.org/10.6084/m9.figshare.9685478.v1>.

References

- Agrawal S, Khatri P, Gupta S et al (2015) Facial expression recognition techniques : a survey. *Int J Adv Electron Comput Sci* 2:61–66. <https://doi.org/10.1016/j.procs.2015.08.011>
- Arsigny V, Fillard P, Pennec X, Ayache N (2007) Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM J Matrix Anal Appl* 29:328–347. <https://doi.org/10.1137/050637996>
- Asghari Oskoei M, Hu H (2007) Myoelectric control systems-A survey. *Biomed Signal Process Control* 2:275–294
- Barachant A, Bonnet S, Congedo M, Jutten C (2010) Riemannian Geometry Applied to BCI Classification. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp 629–636
- Barachant A, Bonnet S, Congedo M, Jutten C (2013) Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Neurocomputing* 112:172–178. <https://doi.org/10.1016/j.neucom.2012.12.039>
- Buettner R, Baumgartl H, Konle T, Haag P (2020) A Review of Virtual Reality and Augmented Reality Literature in Healthcare. 2020 IEEE Symp Ind Electron Appl ISIEA 2020. <https://doi.org/10.1109/ISIEA49364.2020.9188211>
- Burgos-Artizzu XP, Fleureau J, Dumas O, et al (2015) Real-time expression-sensitive HMD face reconstruction. In: *SIGGRAPH Asia 2015 Technical Briefs, SA 2015*. ACM Press, New York, USA, pp 1–4
- Cha H-S, Choi S-J, Im C-H (2020) Real-time recognition of facial expressions using facial electromyograms recorded around the eyes for social virtual reality applications. *IEEE Access* 8:62065–62075. <https://doi.org/10.1109/access.2020.2983608>
- Chen J, Chen Z, Chi Z, Fu H (2018) Facial expression recognition in video with multiple feature fusion. *IEEE Trans Affect Comput* 9:38–50. <https://doi.org/10.1109/TAFFC.2016.2593719>
- Chen Y, Yang Z, Wang J (2015) Eyebrow emotional expression recognition using surface EMG signals. *Neurocomputing* 168:871–879. <https://doi.org/10.1016/j.neucom.2015.05.037>
- Cohen I, Sebe N, Garg A et al (2003) Facial expression recognition from video sequences: temporal and static modeling. *Comput vis Image Underst* 91:160–187. [https://doi.org/10.1016/S1077-3142\(03\)00081-X](https://doi.org/10.1016/S1077-3142(03)00081-X)
- Ekman P (1993) Facial expression and emotion. *Am Psychol* 48:384–392. <https://doi.org/10.1037/0003-066X.48.4.384>
- Ekman P, Rosenberg EL (2005) *What the face reveals basic and applied studies of spontaneous expression using the facial action coding system (FACS)*. Oxford University Press, Oxford
- Fatoorechi M, Archer J, Nduka C, et al (2017) Using facial gestures to drive narrative in VR. In: *SUI 2017 - Proceedings of the 2017 Symposium on Spatial User Interaction*. ACM Press, New York, USA, p 152
- Fan YJ, Yu PT, Liang F, Li X, Li HY, Liu L, Cao JW, Zhao XJ, Wang ZL, Zhu G (2020) Highly conductive, stretchable, and breathable epidermal electrode based on hierarchically interactive nano-network. *Nanoscale* 12:16053–16062. <https://doi.org/10.1039/D0NR03189E>
- Förstner W, Moonen B (2003) A Metric for Covariance Matrices. In: Grafarend EW, Krumm FW, Schwarze VS (eds) *Geodesy-The challenge of the 3rd millennium*. Springer, Berlin Heidelberg, pp 299–309
- Freina L, Ott M (2015) A literature review on immersive virtual reality in education: State of the art and perspectives. *Proc eLearning Softw Educ (eLSE)(Bucharest, Rom April 23--24, 2015)* 8
- Geethanjali P (2016) Myoelectric control of prosthetic hands: state-of-the-art review. *Med Devices Evid Res* 9:247–255
- Gunkel SNB, Stokking HM, Prins MJ, et al (2018) Virtual reality conferencing: Multi-user immersive VR experiences on the web. In: *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*. Association for Computing Machinery, Inc, New York, NY, USA, pp 498–501
- Hakonen M, Piitulainen H, Visala A (2015) Current state of digital signal processing in myoelectric interfaces and related applications. *Biomed Signal Process Control* 18:334–359. <https://doi.org/10.1016/j.bspc.2015.02.009>
- Hamed M, Salleh S-H, Astaraki M et al (2013) EMG-based facial gesture recognition through versatile elliptic basis function neural network. *Biomed Eng Online* 12:73. <https://doi.org/10.1186/1475-925X-12-73>
- Hamed M, Salleh S-H, Swee TT et al (2011) Surface electromyography-based facial expression recognition in Bi-polar configuration. *J Comput Sci* 7:1407
- Hamed M, Salleh SH, Ting CM et al (2018) Robust facial expression recognition for MuCI: a comprehensive neuromuscular signal analysis. *IEEE Trans Affect Comput* 9:102–115. <https://doi.org/10.1109/TAFFC.2016.2569098>
- Hickson S, Kwatra V, Dufour N et al (2015) Facial performance sensing head-mounted display. *ACM Trans Graph* 34(4):1–9. <https://doi.org/10.1145/2766939>
- Hui Z, Zhang H (2017) Head-mounted display-based intuitive virtual reality training system for the mining industry. *Int J Min Sci Technol* 27:717–722. <https://doi.org/10.1016/j.ijmst.2017.05.005>
- Kersting M, Steier R, Venville G (2020) Exploring participant engagement during an astrophysics virtual reality experience at a science festival. *Int J Sci Educ Part B Commun Public Engagem*. <https://doi.org/10.1080/21548455.2020.1857458>
- Khushaba RN (2014) Correlation analysis of electromyogram signals for multiuser myoelectric interfaces. *IEEE Trans Neural Syst Rehabil Eng* 22:745–755. <https://doi.org/10.1109/TNSRE.2014.2304470>
- Kim K-T, Park K-H, Lee S-W (2017) An Adaptive Convolutional Neural Network Framework for Multi-user Myoelectric Interfaces. In: *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*. IEEE, pp 788–792
- Langner O, Dotsch R, Bijlstra G et al (2010) Presentation and validation of the radboud faces database. *Cogn Emot* 24:1377–1388. <https://doi.org/10.1080/02699930903485076>
- Li QX, Chan PPK, Zhou D, et al (2016) Improving robustness against electrode shift of sEMG based hand gesture recognition using online semi-supervised learning. In: *2016 International*

- Conference on Machine Learning and Cybernetics (ICMLC). IEEE, pp 344–349
- Lou J, Wang Y, Nduka C et al (2020) Realistic facial expression reconstruction for VR HMD Users. *IEEE Trans Multimed* 22:730–743. <https://doi.org/10.1109/TMM.2019.2933338>
- Matsubara T, Morimoto J (2013) Bilinear modeling of EMG signals to extract user-independent features for multiuser myoelectric interface. *IEEE Trans Biomed Eng* 60:2205–2213. <https://doi.org/10.1109/TBME.2013.2250502>
- Mavridou I, McGhee JT, Hamed M, et al (2017) FACETEQ interface demo for emotion expression in VR. In: *IEEE Virtual Reality*. pp 441–442
- Morerio P, Murino V (2017) Correlation Alignment by Riemannian Metric for Domain Adaptation. arXiv
- Morrison DG (1969) On the Interpretation of Discriminant Analysis. *J Mark Res* 6:156. <https://doi.org/10.2307/3149666>
- Mottelson A, Hornbæk K (2020) Emotional avatars: The interplay between affect and ownership of a virtual body. arXiv
- Muceli S, Jiang N, Farina D (2014) Extracting signals robust to electrode number and shift for online simultaneous and proportional myoelectric control by factorization algorithms. *IEEE Trans Neural Syst Rehabil Eng* 22:623–633. <https://doi.org/10.1109/TNSRE.2013.2282898>
- Olszewski K, Lim JJ, Saito S, Li H (2016) High-fidelity facial and speech animation for VR HMDs. *ACM Trans Graph* 35:1–14. <https://doi.org/10.1145/2980179.2980252>
- Pan X, de Hamilton AF (2018) Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. *Br J Psychol* 109:395–417. <https://doi.org/10.1111/bjop.12290>
- Patel AN, Howard MD, Roach SM et al (2018) Mental state assessment and validation using personalized physiological biometrics. *Front Hum Neurosci* 12:1–13. <https://doi.org/10.3389/fnhum.2018.00221>
- Patel JK, Sakadasariya A (2018) Survey on virtual reality in social network. In: *Proceedings of the 2nd International Conference on Inventive Systems and Control, ICISC 2018*. IEEE, pp 1341–1344
- Phinyomark A, Scheme E (2018) EMG pattern recognition in the era of big data and deep learning. *Big Data Cogn Comput* 2:21. <https://doi.org/10.3390/bdcc2030021>
- Rapuano M, Ferrara A, Sbordone FL et al (2020) The appearance of the avatar can enhance the sense of co-presence during virtual interactions with users. *CEUR Workshop Proc* 2730:1–10
- Rezazadeh IM, Firoozabadi M, Hashemi Golpayegani MR et al (2011) Using affective human-machine interface to increase the operation performance in virtual construction crane training system: a novel approach. *Autom Constr* 20:289–298. <https://doi.org/10.1016/j.autcon.2010.10.005>
- Riva G, Mantovani F, Capideville CS et al (2007) Affective interactions using virtual reality: the link between presence and emotions. *Cyberpsychol Behav* 10:45–56. <https://doi.org/10.1089/cpb.2006.9993>
- Sato W, Yoshikawa S (2007) Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition* 104:1–18. <https://doi.org/10.1016/j.cognition.2006.05.001>
- Thies J, Zollhöfer M, Stamminger M et al (2016) FaceVR: real-time facial reenactment and eye gaze control in virtual reality. *ACM Trans Graph*. <https://doi.org/10.1145/3182644>
- Vidovic MMC, Hwang HJ, Amsuss S et al (2016) Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation. *IEEE Trans Neural Syst Rehabil Eng* 24:961–970. <https://doi.org/10.1109/TNSRE.2015.2492619>
- Vidovic MMC, Paredes LP, Hwang HJ, et al (2014) Covariate shift adaptation in EMG pattern recognition for prosthetic device control. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. pp 4370–4373
- Wakeford N, Hong S (2002) The social life of avatars presence and interaction in shared virtual environments reviewer. *Sociol Res Online* 7:137–138. <https://doi.org/10.1177/136078040200700211>
- Wang R, Guo H, Davis LS, Dai Q (2012) Covariance discriminative learning: A natural and efficient approach to image set classification. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp 2496–2503
- Wedel M, Bigné E, Zhang J (2020) Virtual and augmented reality: advancing research in consumer marketing. *Int J Res Mark* 37:443–465. <https://doi.org/10.1016/j.ijresmar.2020.04.004>
- Xiong A, Zhao X, Han J, et al (2015) An user-independent gesture recognition method based on sEMG decomposition. In: *IEEE International Conference on Intelligent Robots and Systems*. pp 4185–4190
- Yang S, Yang G (2011) Emotion recognition of EMG based on improved L-M BP neural network and SVM. *J Softw* 6:1529–1536. <https://doi.org/10.4304/jsw.6.8.1529-1536>
- Young AJ, Hargrove LJ, Kuiken TA (2012) Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration. *IEEE Trans Biomed Eng* 59:645–652. <https://doi.org/10.1109/TBME.2011.2177662>
- Zhang H, Zhao Y, Yao F, et al (2013) An adaptation strategy of using LDA classifier for EMG pattern recognition. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. IEEE, pp 4267–4270
- Zhang L, Tjondronegoro D, Chandran V (2014) Random Gabor based templates for facial expression recognition in images with facial occlusion. *Neurocomputing* 145:451–464. <https://doi.org/10.1016/j.neucom.2014.05.008>
- Zhang T (2018) Facial expression recognition based on deep learning: a survey. *Adv Intell Syst Comput* 686:345–352. https://doi.org/10.1007/978-3-319-69096-4_48

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.