Topology of surface electromyogram signals: hand gesture decoding on Riemannian manifolds

Harshavardhana T. Gowda [1] **Lee M. Miller** [2 3 4 5]

Abstract

Objective. In this article, we present data and methods for decoding hand gestures using surface electromyogram (EMG) signals. EMG-based upper limb interfaces are valuable for amputee rehabilitation, artificial supernumerary limb augmentation, gestural control of computers, and virtual and augmented reality applications. *Approach.* To achieve this, we collect EMG signals from the upper limb using surface electrodes placed at key muscle sites involved in hand movements. Additionally, we design and evaluate efficient models for decoding EMG signals. *Main results.* Our findings reveal that the manifold of symmetric positive definite (SPD) matrices serves as an effective embedding space for EMG signals. Moreover, for the first time, we quantify the distribution shift of these signals across individuals. *Significance.* Overall, our approach demonstrates significant potential for developing efficient and interpretable methods for decoding EMG signals. This is particularly important as we move toward the broader adoption of EMG-based wrist interfaces.

Data and code availability

Our dataset is publicly available through the following link: DOI: 10.17605/OSF.IO/ZCR43.
Codes are available on GitHub: GitHub.

[1] Department of Electrical and Computer Engineering, University of California, Davis [2] Center for Brain and Mind, University of California, Davis [3] Department of Neurobiology, Physiology, and Behavior, University of California, Davis [4] Department of Otolaryngology/Head and Neck Surgery, University of California, Davis [5] A version of this article has been published in the Journal of Neural Engineering. *J. Neural Eng. 21 036047*. Correspondence to: Harshavardhana T. Gowda <tgharshavardhana@gmail.com>.

1. Introduction

EMG signals are recorded non-invasively by placing sensors on the skin surface to measure electrical activity generated by motor unit activation. The global characteristics of the EMG signal, such as amplitude and power spectrum, depend on numerous idiosyncratic factors. These include anatomical characteristics, such as the thickness of subcutaneous tissue, the distribution and size of motor unit (MU) territories, and the spread of endplates and tendon junctions within the MU. Physiological factors, such as the distribution of conduction velocities within motor unit fibers, the shape of intracellular action potentials (Farina et al., 2004), and muscle fatigue (Enoka & Duchateau, 2008), also influence the signal. Additionally, circumstantial factors, such as precise electrode placement (Huebner et al., 2015; Kleine et al., 2001), further contribute to signal variability. The combined effects of these factors are further complicated by interactions between signals originating from multiple neighboring muscles. As a result, signals from individual EMG electrodes tend to be highly confounded and opaque, limiting their practical utility.

We demonstrate that covariance matrices, constructed using the pairwise covariance of electrical signals measured across different electrodes, effectively capture the combined influence of various physiological and anatomical factors. Consequently, this approach provides a robust framework for quantifying differences in EMG signals across individuals. Furthermore, the spatial patterns captured by these covariance matrices exhibit rich geometric structures, which can be effectively leveraged to distinguish different hand gestures.

1.1. Prior work

Existing methods rely on constructed features such as signal root-mean-square, time-domain statistics, as described by Hudgins et al., histograms (Zardoshti-Kermani et al., 1995), marginalized discrete wavelet transform (Lucas et al., 2008), or a normalized combination of all the above. These features are typically evaluated using classifiers such as linear discriminant analysis (LDA), support vector machines (SVM), k-nearest neighbors (k-NN), and random forests.

Additionally, some studies employ deep learning models, including convolutional neural networks (CNN) (Rahimian et al., 2020a; Geng et al., 2016; Wei et al., 2019; Khushaba et al., 2022), recurrent neural networks (RNN) (Quivira et al., 2018; Sun et al., 2021), transformer-based networks (Rahimian et al., 2021; Montazerin et al., 2023), and hybrid architectures combining RNN and CNN (Rahimian et al., 2020b; Hu et al., 2018). Xiong et al. analyze EMG signals using SPD covariance matrices; however, by mapping the learned features onto a tangent plane and decoding them in Euclidean space, this approach fails to leverage the natural geometric structure of the data.

Despite their advancements, all these methods overlook the strong spatial correlations in muscle contraction patterns. Furthermore, they often involve a large number of trainable parameters - ranging in the tens of thousands - and require complex transfer learning paradigms and extensive retraining for deployment across individuals. Moreover, none of these established techniques readily adapt to signal variations caused by factors such as muscle fatigue and sensor displacement.

1.2. Our contribution

To overcome the limitations of existing methods, we propose a more principled approach by analyzing EMG signals on a Riemannian manifold, which offers a more comprehensive and natural framework for representing the data structure compared to conventional Euclidean analysis. Specifically, the SPD covariance matrices derived from multivariate EMG time series form a structured representation on a cone manifold with a Riemannian metric. This geometric framework naturally preserves the spatial correlations in muscle contraction patterns.

To ensure computational efficiency and numerical stability, we study these SPD matrices via Cholesky decomposition in the Cholesky space (Lin, 2019), which consists of lower triangular matrices with positive diagonal elements. We demonstrate that covariance matrices from different individuals occupy distinct neighborhoods within the manifold space due to the combined influence of anatomical and physiological factors on EMG signals. Furthermore, we quantify inter-individual differences in EMG signals using the geodesic distance - the length of the shortest curve between two points on a surface (between corresponding covariance matrices).

Leveraging this geometric structure, we introduce two supervised learning algorithms - manifold minimum distance to mean and manifold support vector machine - along with one unsupervised learning algorithm, manifold k-medoids clustering, to classify hand gestures directly on the Riemannian manifold. This approach not only mitigates the challenges associated with existing feature extraction techniques and

deep learning models but also offers a more interpretable and data-efficient alternative for EMG-based gesture recognition.

2. Methods and materials

We use three different datasets to demonstrate the efficacy of our methods. First, we utilize the widely used NI-NAPRO (Non-Invasive Adaptive Prosthetics) DATABASE 2-EXERCISE 1, provided by Atzori et al. Second, we employ a high-density EMG dataset from Malešević et al. Third, we use a dataset collected by us, named UCD-MYOVERSE-HAND-0. The first two datasets allow us to compare our methods with previously benchmarked approaches, while the inclusion of our own dataset ensures additional validation.

By selecting three distinct datasets, we aim to verify that our methods perform consistently well across different data collection platforms and experimental setups. Below, we provide a detailed explanation of the data acquisition protocols and EMG processing methods.

2.1. EMG data acquisition setup

A total of 30 subjects participated in our study (UCD-MYOVERSE-HAND-0). Please refer to the *ethical statement* for the subject selection criteria.

For UCD-MYOVERSE-HAND-0, we used Delsys double differential EMG electrodes (Delsys, Inc.) and an NI USB-6210 multifunction I/O device (National Instruments Corporation – 16 inputs, 16-bit, 250 kS/s) to acquire EMG data at a sampling rate of 2000 Hz. The Delsys electrodes transmitted the acquired data wirelessly to a base station, which then relayed the data to a computer via a USB connection through the NI USB-6210 data acquisition system.

A graphical user interface (GUI) was designed to display hand gestures on a screen. Subjects performed the displayed gesture with their dominant hand while seated comfortably in a chair, with their forearm resting on an elevated platform on the table. They were allowed to choose a resting position that was most comfortable for them and could adjust it throughout the experiment. Each gesture was performed for 2 s, followed by a resting period of 2 s. Specifically, the gesture was displayed on the screen for 2 s, followed by a blank screen for another 2 s. Subjects were instructed to perform the gesture for the duration the image was visible and to rest during the blank screen period.

The experiment was divided into six sessions, each consisting of sixty trials - six repetitions for each of the ten gestures. The order of gestures within a session was pseudo-randomly generated to assess how variations in performing gestures affect decoding accuracy. If all six repetitions of

a gesture occurred sequentially, it could lead to repetitive, unconscious, and highly consistent movements, which we aimed to avoid. In total, each subject completed 360 trials.

To synchronize GUI instructions with data streamed from the Delsys system, we used ZeroMQ sockets (ZeroMQ) and Lab Streaming Layer (LSL) in Python. Data streams were synchronized to the master clock on the computer, which received both EMG data and event markers from the GUI. We observed a maximum clock drift of approximately 50 *ms*.

2.1.1. EMG OF DIFFERENT HAND GESTURES IN UCD-MYOVERSE-HAND-0

Forearm EMG was collected from intact subjects using twelve electrodes. Eight electrodes were placed evenly spaced around the main belly of the forearm muscles, positioned below the elbow at approximately one-third of the distance from the elbow to the wrist. The remaining four electrodes were placed evenly spaced around the wrist joint.

Each subject performed ten different hand gestures, with each gesture repeated thirty-six times. The ten gestures included wrist movements in four cardinal directions (up, down, left, and right), three types of pinches (index finger pinch, middle finger pinch, and two-finger pinch), splay, power grasp, and pointing index (figure 1). These gestures were selected to reflect commonly used interactions with computers. Wrist movements in cardinal directions can be used for navigating the screen (moving it up, down, left, and right), while the pinch gestures can be used for zooming in, zooming out, and making selections.
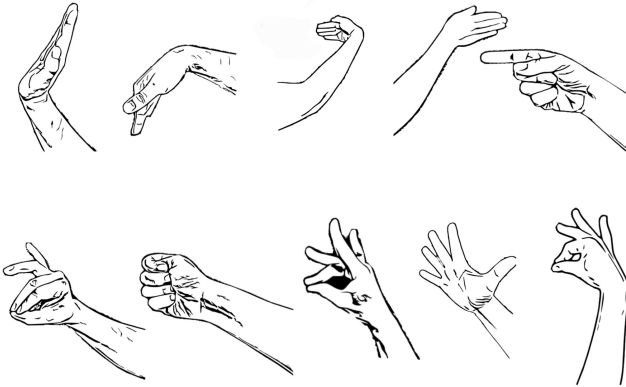

Figure 1. Ten gestures included in the UCD-MYOVERSE-HAND-0 experiment. From top-left: up, down, left, right, index point, two finger pinch, power grasp, middle finger pinch, splay, index finger pinch.

2.1.2. EMG OF DIFFERENT HAND GESTURES IN NIANPRO: DATABASE 2-EXERCISE 1

The dataset consists of EMG signals collected from forty intact subjects using twelve electrodes at 2000 Hz. Eight electrodes were placed evenly spaced around the forearm at the height of the radio-humeral joint, while two electrodes were placed on the main activity spots of the flexor digitorum superficialis and the extensor digitorum superficialis.

Subjects performed seventeen different gestures, each repeated six times. The seventeen gestures include: thumbs up, extension of the index and middle fingers with flexion of the others, flexion of the ring and little fingers with extension of the others, thumb opposing the base of the little finger, abduction of all fingers, fingers flexed together in a fist, pointing index, adduction of extended fingers, wrist supination (axis: middle finger), wrist pronation (axis: middle finger), wrist supination (axis: little finger), wrist pronation (axis: little finger), wrist flexion, wrist extension, wrist radial deviation, wrist ulnar deviation, and wrist extension with a closed hand.

For further details on data acquisition and processing, see Atzori et al.

2.1.3. HIGH DENSITY EMG OF DIFFERENT HAND GESTURES IN MALEŠEVIĆ ET AL.

The dataset consists of EMG signals collected from nineteen intact subjects using 128 electrodes placed at the forearm level at 2048 Hz. Although the experiment initially included twenty subjects, the data for *subject 5* was corrupted and therefore excluded.

Subjects performed 65 unique gestures, each formed by combining 16 basic single-degree-of-freedom movements. Each gesture was repeated five times.

For further details on data acquisition and processing, refer to Malešević et al.

2.2. EMG data preprocessing

The data collection environment was carefully controlled to eliminate AC electrical interference. EMG signals undergo minimal preprocessing. Signals are z-normalized along the time dimension for each channel. The preprocessed EMG signals are then used to construct covariance SPD matrices. We detail the EMG analysis methods in the following section.

2.3. EMG data analysis

EMG signals are collected using a set of sensors, denoted as $\mathcal{V}$, and are represented as a matrix $\mathcal{X}$ with dimensions ($\mathcal{V}$, SAMPLING FREQUENCY × GESTURE DURATION).

Signal covariance matrices are then computed from the preprocessed EMG signals using

$$\mathcal{E} = \mathcal{X}\mathcal{X}^T.$$

The resulting matrices, $\mathcal{E}$, are symmetric and positive semi-definite. To ensure positive definiteness, we apply a shrinkage estimator as follows:

$$\mathcal{E} \leftarrow (1 - \eta)\mathcal{E} + \eta\, \text{trace}(\mathcal{E})\mathcal{I}$$

where $\mathcal{I}$ is an identity matrix of the same dimensions as $\mathcal{E}$. Through empirical evaluation, we found that setting $\eta = 0.1$ is sufficient for all EMG data. The covariance matrices, $\mathcal{E}$, have dimensions $\mathcal{V} \times \mathcal{V}$.

Directly working with SPD matrices using affine-invariant or log-Euclidean metrics (Arsigny et al., 2007) involves computationally expensive operations, such as matrix exponential and matrix logarithm calculations. These operations make mappings between the manifold space and the tangent space, and vice versa, computationally intensive. To address this, Lin proposed methods to operate on SPD matrices using Cholesky decomposition. They established a diffeomorphism between the Riemannian manifold of SPD matrices and Cholesky space. In Cholesky space, the computational burden is significantly reduced: logarithmic and exponential computations are restricted to the diagonal elements of the matrix, making them element-wise operations. Additionally, the Fréchet mean (centroid) is derived in a closed form. For an SPD edge matrix $\mathcal{E}$, the corresponding Cholesky decomposition $\mathcal{L} = \text{CHOLESKY}(\mathcal{E})$ is such that $\mathcal{E} = \mathcal{L}\mathcal{L}^T$. A matrix $\lfloor \mathcal{L} \rfloor$ is the strictly lower triangular part of the matrix $\mathcal{L}$, and a matrix $\mathbb{D}(\mathcal{L})$ is the diagonal part of the matrix $\mathcal{L}$. In the following section, we explain in detail, the methods used to analyze SPD matrices.

2.3.1. DISTANCE BETWEEN TWO SPD MATRICES $\mathcal{E}_1$ AND $\mathcal{E}_2$

The geodesic distance between two SPD matrices $\mathcal{E}_1$ and $\mathcal{E}_2$ is same as the distance between the corresponding Cholesky matrices $\mathcal{L}_1$ and $\mathcal{L}_2$ and is calculated as

$$d(\mathcal{L}_1, \mathcal{L}_2) = \left\{ ||\lfloor \mathcal{L}_1 \rfloor - \lfloor \mathcal{L}_2 \rfloor ||_F^2 \right.$$
$$\left. + || \log \mathbb{D}(\mathcal{L}_1) - \log \mathbb{D}(\mathcal{L}_2)||_F^2 \right\}^{1/2}, \quad (1)$$

where $||.||_F$ denotes the Frobenius norm.

2.3.2. FRÉCHET MEAN (CENTROID) OF SPD MATRICES

Given a set of (n) SPD edge matrices $\mathcal{E}$, we first calculate their corresponding Cholesky decompositions $\mathcal{L} = \text{CHOLESKY}(\mathcal{E})$, such that $\mathcal{E} = \mathcal{L}\mathcal{L}^T$. Then, the Fréchet mean of the Cholesky decomposed matrices $\mathcal{L}$ is given by

$$\mathcal{F}_{\text{CHOLESKY}} = \frac{1}{n}\sum_{i=1}^{n} \lfloor \mathcal{L}_i \rfloor \quad +$$
$$\exp\left(\frac{1}{n}\sum_{i=1}^{n} \log(\mathbb{D}(\mathcal{L}_i))\right). \quad (2)$$

2.3.3. k-MEDOIDS CLUSTERING ALGORITHM

We implement the classic k-medoids algorithm (Kaufman & Rousseeuw, 1990) using partitioning around medoids (PAM) heuristic by replacing the Euclidean distance with the geodesic distance in equation 1.

2.3.4. MINIMUM DISTANCE TO MEAN ALGORITHM (MDM)

Given m classification classes and n training samples, SPD matrices in the training set $\{\mathcal{E}_i^j\}$, where $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., m\}$ are used to construct centroids for each of the m classes such that the centroid of class j is,

$$\mathcal{C}^j = \mathbb{E}(\{\text{CHOLESKY}(\mathcal{E}^j)\}), \quad (3)$$

where the Fréchet mean $\mathbb{E}$ is calculated according to equation 2. Given a test dataset of SPD matrices $\{\mathcal{T}\}$, $T \in \mathcal{T}$ is assigned to that class whose centroid is nearest to $\text{CHOLESKY}(T)$. That is, the class of T is

$$\arg\min_j d(\text{CHOLESKY}(T), \mathcal{C}^j), \quad (4)$$

where $d(.)$ is calculated according to equation 1.

2.3.5. SUPPORT VECTOR MACHINE (SVM)

For training the SVM, we use a kernel

$$\mathscr{K} : \text{CHOLESKY SPACE} \times \text{CHOLESKY SPACE} \rightarrow \mathbb{R},$$

such that

$$\mathscr{K} = \exp(-\gamma d^2(L_1, L_2)), \quad (5)$$

where L_1, L_2 are Cholesky factors of SPD matrice $\mathcal{E}_1$ and $\mathcal{E}_2$ and $\gamma > 0$. $d(.)$ is defined according to equation 1. In appendix A, following the arguments in Jayasumana et al., we prove that the kernel in equation 5 is a valid kernel.

2.3.6. t-SNE FOR EMG DATA VISUALIZATION

We use the t-SNE (t-Stochastic Neighbor Embedding) algorithm, adapted from Van der Maaten & Hinton, for data visualization. Unlike standard t-SNE, which takes vectors as input and uses Euclidean distance, we input edge matrices $\mathcal{E}$ and employ the distance defined in equation 1.

3. Results

Here, we demonstrate that different hand gestures can be decoded with high accuracy using simple algorithms such as *k*-medoids (*section* 2.3.3), minimum distance to mean (*section* 2.3.4), and SVM (*section* 2.3.5) on the manifold of SPD matrices. This finding highlights that EMG signal covariance matrices capture meaningful and distinguishing information, and that the manifold of SPD matrices serves as a natural embedding space for EMG signals.

3.1. Results for UCD-MYOVERSE-HAND-0

EMG data was collected at a sampling frequency of 2000 Hz using 12 electrodes, with each gesture lasting 2 *s*. For each gesture, the EMG data is represented as a matrix $\mathcal{X}$ of dimensions (12×4000), where 12 corresponds to the number of electrode channels and 4000 corresponds to the temporal dimension (i.e., $2\ s \times 2000$ Hz). The corresponding covariance SPD matrices, denoted as $\mathcal{E}$, have dimensions (12×12). We decode covariance matrices $\mathcal{E}$ using MDM, SVM, and *k*-medoids algorithms defined on the manifold of SPD matrices.

We present the average classification accuracies across subjects in table 1. Detailed subject-wise results are provided in appendix C.1.

Subject Number	Classification Methods		
	MDM	SVM ($\gamma = 1$)	*k*-medoids
Average	0.82	0.86	0.70

Table 1. Average classification accuracy for 30 subjects in UCD-MYOVERSE-HAND-0. Data from the first three sessions (i.e., the first 18 repetitions of each gesture) were used for training, while data from the last three sessions (i.e., the last 18 repetitions of each gesture) were used for testing the MDM and SVM algorithms. Chance decoding accuracy is $\frac{1}{10} = 0.1$.

3.2. Results for NINAPRO: DATABASE 2-EXERCISE 1

EMG data was collected at a sampling frequency of 2000 Hz using 12 electrodes, with each gesture lasting 5 *s*. For each gesture, the EMG data is represented as a matrix $\mathcal{X}$ with 12 rows (corresponding to the 12 electrode channels) and 10000 columns (corresponding to the temporal dimension, i.e., $5\ s \times 2000$ Hz). Therefore, the dimensions of $\mathcal{X}$ are (12×10000).

The corresponding covariance SPD matrices, denoted as $\mathcal{E}$, have dimensions (12×12). We decode these covariance matrices using the MDM, SVM, and *k*-medoids algorithms, which are defined on the manifold of SPD matrices.

To evaluate our approach, we compare the results obtained using our methods with those reported in Sun et al. and Rahimian et al. (table 2). Detailed subject-wise results for all 40 subjects using manifold-based methods are provided

in appendix C.2. For *train-test* data split, refer to Sun et al.; Rahimian et al.

	Method	Accuracy
Sun et al.	4-layer 3rd order dilation	0.824
	4-layer 3rd order dilation (LSTM)	0.797
	4-layer 2nd order dilation (LSTM)	0.796
	4-layer 1st order dilation (LSTM)	0.793
	4-layer baseline	0.753
	2-layer CNN	0.746
	2-layer LSTM	0.702
	1-layer LSTM	0.684
	2-layer MLP	0.662
	SVM (Euclidean)	0.307
Rahimian et al.	TEMGNet *200* m*s*	0.821
	TEMGNet *300* m*s*	0.829
Proposed methods	Manifold MDM	0.92
	Manifold SVM	0.93
	Manifold *k*-medoids	0.82

Table 2. Our proposed methods leverage manifold representation and outperform baselines. The unsupervised *k*-medoids algorithm is, to our knowledge, the only unsupervised method for EMG classification. Unlike neural networks with tens of thousands of parameters, our methods are computationally efficient. Chance decoding accuracy is $\frac{1}{17} = 0.059$.

3.3. Results for high density EMG in MALEŠEVIĆ ET AL.

High-density EMG was collected at 2048 Hz using 128 electrodes. The duration of each gesture varied, which we denote as the variable $\mathcal{T}$. The data corpus consisted of 65 unique hand gestures, with each gesture repeated five times. For each gesture, the EMG data is represented as a matrix $\mathcal{X}$, where $\mathcal{X}$ has $\mathcal{T}$ columns (temporal dimension) and 128 rows (corresponding to the 128 electrode channels). Thus, the dimensions of $\mathcal{X}$ are $(128 \times \mathcal{T})$.

The covariance SPD matrix $\mathcal{E}$ has dimensions (128×128). We use MDM, SVM, and *k*-medoids algorithms, as described previously, to classify the SPD matrices $\mathcal{E}$. The results of our manifold-based classification methods are presented in table 3.

Additionally, we compare our results with the methods proposed by Montazerin et al. in table 4.

3.4. EMG data visualization using *t*-SNE

We use *t*-SNE (*section* 2.3.6) to visualize the SPD covariance matrices of hand gestures. These visualizations demonstrate that the covariance matrices of EMG signals exhibit structured representations on the manifold of SPD matrices, a finding further validated through simple classification methods defined on the manifold in previous sections.

The *t*-SNE embedding reveals that SPD matrices from different subjects occupy distinct neighborhoods on the mani-

Fold number	Classification methods		
	MDM	SVM ($\gamma = 0.1$)	k-medoids
1	0.83 ± 0.07	0.84 ± 0.08	
2	0.95 ± 0.04	0.96 ± 0.03	
3	0.96 ± 0.03	0.97 ± 0.03	0.87 ± 0.05
4	0.95 ± 0.04	0.96 ± 0.04	
5	0.91 ± 0.05	0.92 ± 0.04	
Average	$\mathbf{0.92 \pm 0.07}$	$\mathbf{0.93 \pm 0.07}$	$\mathbf{0.87 \pm 0.05}$

Table 3. Classification accuracy averaged across *19* subjects in Malešević et al. Following the work in Montazerin et al., we perform 5-fold cross validation analysis. Chance decoding accuracy is $\frac{1}{65}$.

	Method	Accuracy
Montazerin et al.	*512* sample window and *128* channels (*their best*)	0.92
Proposed methods	Manifold MDM	0.92
	Manifold SVM	0.93
	Manifold k-medoids	0.87

Table 4. The proposed methods achieve performance comparable to the transformer-based approach in Montazerin et al., while being computationally efficient. Moreover, they are better suited for deployment across individuals and real-time adaptation due to their transparent nature. For instance, drift in EMG signals caused by sensor position shifts or muscle fatigue can be addressed simply by updating the centroids of the SPD matrices in the manifold MDM algorithm, rather than repeatedly fine-tuning or retraining a transformer. Chance decoding accuracy is $\frac{1}{65}$.

fold, as shown in figure 2. This variation in EMG signals arises from a combination of anatomical, physiological, and circumstantial factors and can be quantified using a single geodesic distance metric across all individuals, as depicted in figure 3. Due to these differences in spatial muscle contraction patterns among individuals, generalizing deep learning algorithms across the entire population remains challenging.

However, with this appropriate non-Euclidean representation of the data, spatial muscle contraction patterns within an individual are sufficiently distinct to enable unsupervised classification of various gestures, as demonstrated in figure 4, thereby eliminating the need for complex transfer learning paradigms or model retraining when deploying across individuals.

Furthermore, since patterns from different individuals occupy separate neighborhoods on the manifold, directly comparing gestures across all subjects leads to extremely poor gesture differentiation (figure 5).

Nevertheless, using parallel transport (see appendix B), these distinct gesture patterns can be aligned across individuals, with an accuracy bounded by the accuracy of unsupervised classification (figure 6). Figure 7 illustrates why *parallel* transport (see appendix B), as opposed to Euclidean transport, is the appropriate approach in the manifold space.

4. Observations and conclusions

① EMG signals are multivariate and capture muscle activation patterns across spatially separated locations using multiple sensor nodes. This multivariate nature of EMG, which reflects a distributed yet concerted effort from multiple muscles to execute a given gesture, suggests that the data structure is not inherently Euclidean. In this article, we empirically demonstrate that this is indeed the case.

Signal covariance matrices, which represent the covariances of EMG signals across all pairs of sensor nodes, form structured representations on the manifold of SPD matrices, naturally distinguishing different hand gestures. This suggests that covariance features, rather than convolutional filters, are more suitable for EMG signal feature extraction.

② We showed that EMG signals exhibit discriminative representations on the manifold, allowing classification even with unsupervised clustering algorithms. However, EMG signals from different individuals occupy distinct neighborhoods within the manifold space.

We need to verify whether EMG signals from certain individuals or groups - such as those with a high BMI, who typically have more subcutaneous fat - lie in the same neighborhood or whether each individual is inherently different. If the latter is true, we must further explore its implications for generalizable machine learning models.

③ Following the previous point, we should explore whether zero-shot learning algorithms can be developed, allowing a model trained on a set of individuals to generalize to unseen individuals.

As demonstrated here, few-shot learning is straightforward. If we have a few labeled samples for each gesture, we can perform unsupervised classification and use parallel transport to align EMG signals across individuals.
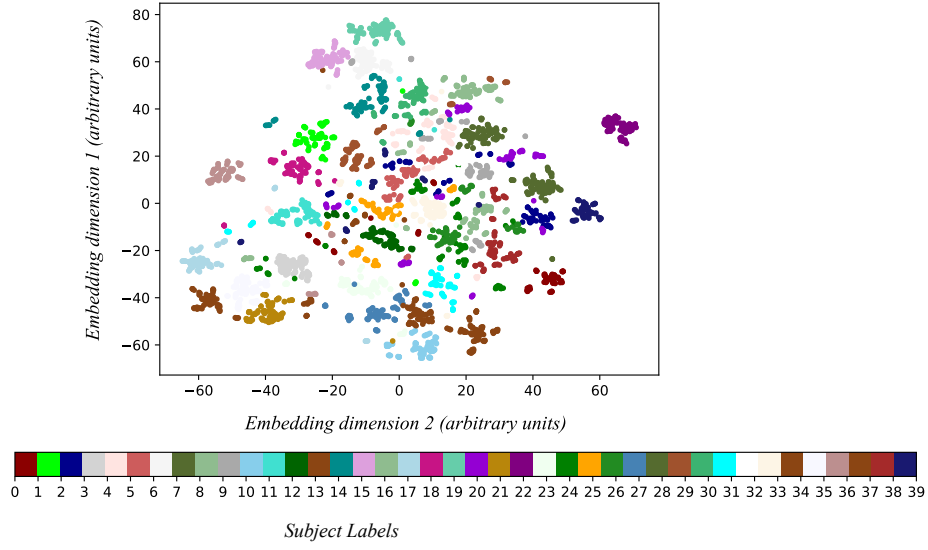
Figure 2. *t*-SNE of SPD covariance matrices using Riemannian distance indicates that the SPD matrices from different subjects lie in different neighborhoods of the manifold. This is due to shift in EMG signals owing to the combined effect of various anatomical, physiological, and circumstantial factors. Embedding is for NINAPRO: DATABASE 2-EXERCISE 1. Embedding is colored according to subjects. Each of the 40 subjects performed 102 trials (17 gestures, each repeated 6 times).
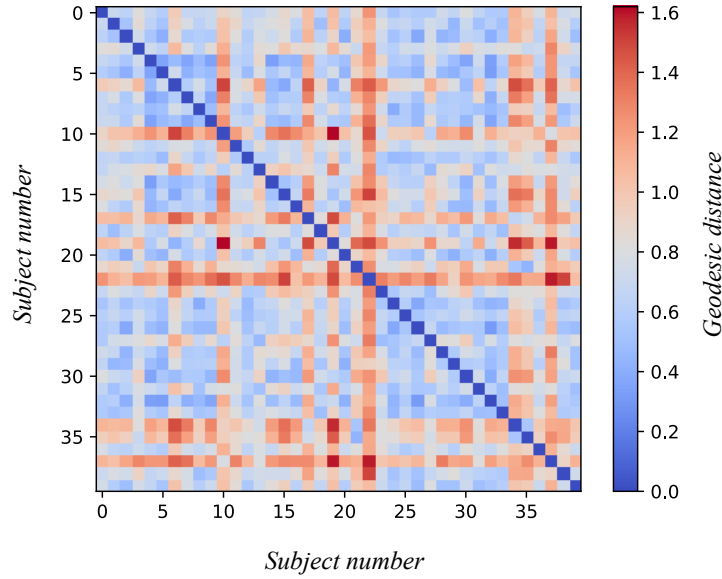

Figure 3. Riemannian geodesic distance between the centroids of SPD covariance matrices of 40 subjects in NINAPRO: DATABASE 2-EXERCISE 1. Geodesic distance quantifies the differences in EMG signals between subjects due to the combined effect of various physiological and anatomical factors. The centroid of a given subject is calculated as the *Log-Cholesky* average of SPD covariance matrices of all 102 trials (17 gestures, each repeated 6 times). *X* and *Y* axes are numbered according to subjects.
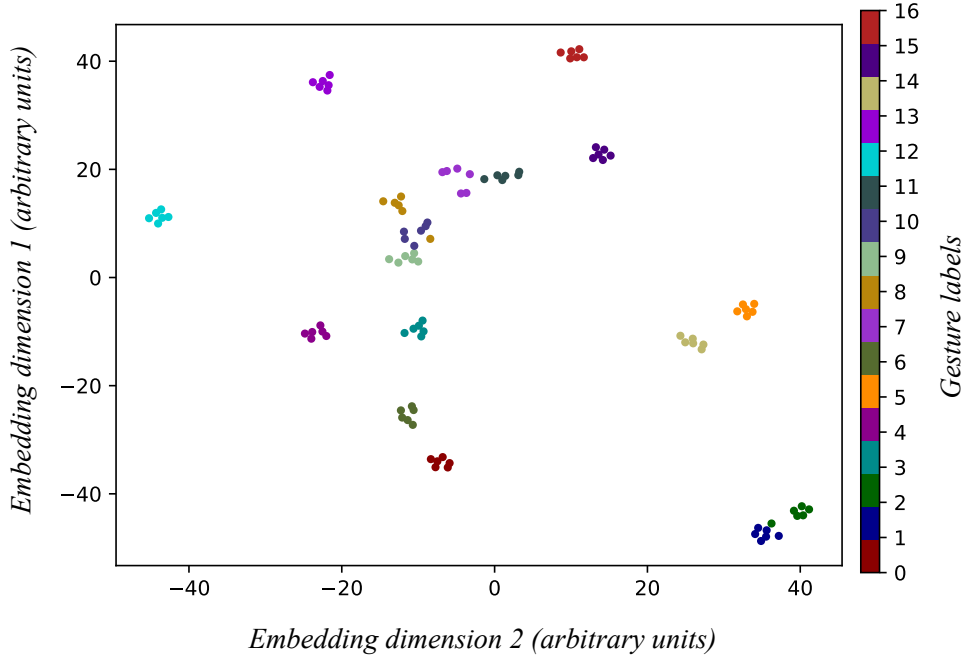
Figure 4. t-SNE of SPD matrices of *subject 0* from NINAPRO: DATABASE 2-EXERCISE 1 shows that different gestures within a subject have contrasting spatial patterns. We can classify these distinct gestures with supervised algorithms such as minimum distance to mean (MDM) and support vector machine (SVM) or unsupervised algorithms such as *k*-medoids clustering using Riemannian distance. Classification accuracy using the above methods is presented in appendix C.2 for all 40 subjects. Embedding is colored according to gestures. The subject performed 17 gestures with each gesture repeated six times.

The gestures are: 0: thumb up, 1: extension of index and middle - flexion of the others, 2: flexion of ring and little finger - extension of the others, 3: thumb opposing base of little finger, 4: abduction of all fingers, 5: fingers flexed together in fist, 6: pointing index, 7: adduction of extended fingers, 8: wrist supination (axis: middle finger), 9: wrist pronation (axis: middle finger), 10: wrist supination (axis: little finger), 11: wrist pronation (axis: little finger), 12: wrist flexion, 13: wrist extension, 14: wrist radial deviation, 15: wrist ulnar deviation, and 16: wrist extension with closed hand.
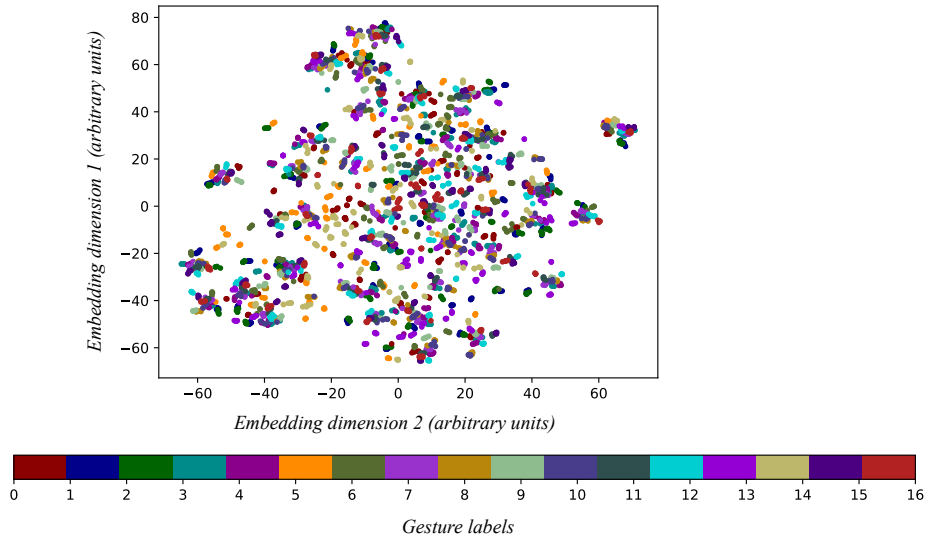
Figure 5. *t*-SNE of SPD covariance matrices using Riemannian distance without accounting for intersubject differences reveals that the SPD matrices for the same gesture from different subjects do not cluster together. This is the same embedding in figure 2 colored according to gestures (instead of subjects). Each of the 40 subjects performed 102 trials (17 gestures, each repeated 6 times).

The gestures are: 0: thumb up, 1: extension of index and middle - flexion of the others, 2: flexion of ring and little finger - extension of the others, 3: thumb opposing base of little finger, 4: abduction of all fingers, 5: fingers flexed together in fist, 6: pointing index, 7: adduction of extended fingers, 8: wrist supination (axis: middle finger), 9: wrist pronation (axis: middle finger), 10: wrist supination (axis: little finger), 11: wrist pronation (axis: little finger), 12: wrist flexion, 13: wrist extension, 14: wrist radial deviation, 15: wrist ulnar deviation, and 16: wrist extension with closed hand.
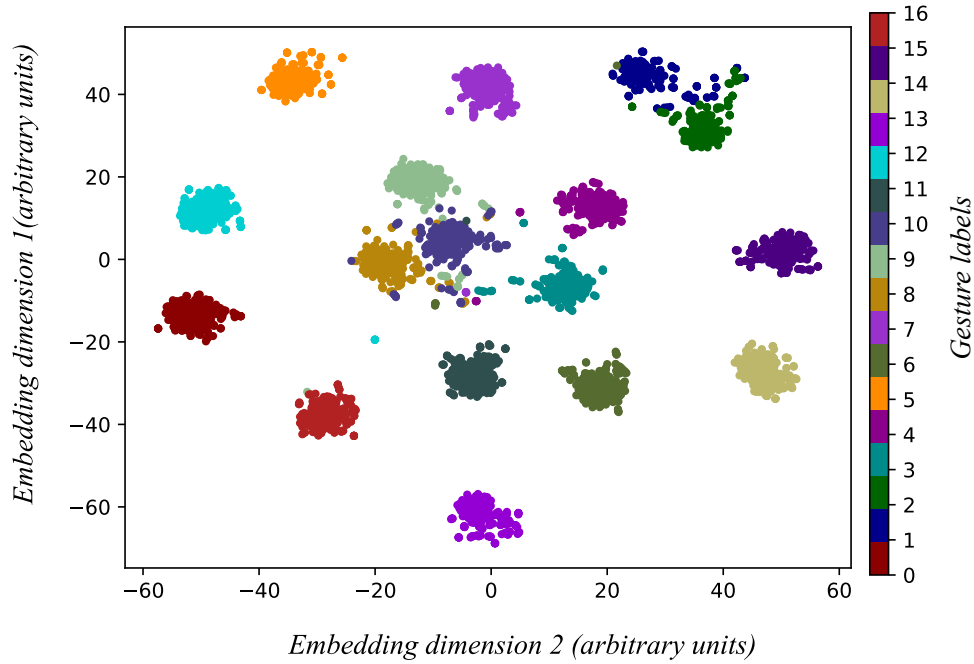
Figure 6. t-SNE of SPD covariance matrices using Riemannian distance after aligning the gestures using *parallel transport* as described in appendix B and illustrated in figure 7. Data is aligned assuming ground truth labels for visualization purposes. Otherwise, alignment accuracy is bounded by the accuracy of unsupervised classification. This demonstrates how data from different individuals in figure 5 can be aligned according to gestures so that the same gestures from different individuals are transported to the same neighborhood in the manifold space. Embedding is colored according to gestures. Each of the 40 subjects performed 102 trials (17 gestures, each repeated 6 times).

The gestures are: 0: thumb up, 1: extension of index and middle - flexion of the others, 2: flexion of ring and little finger - extension of the others, 3: thumb opposing base of little finger, 4: abduction of all fingers, 5: fingers flexed together in fist, 6: pointing index, 7: adduction of extended fingers, 8: wrist supination (axis: middle finger), 9: wrist pronation (axis: middle finger), 10: wrist supination (axis: little finger), 11: wrist pronation (axis: little finger), 12: wrist flexion, 13: wrist extension, 14: wrist radial deviation, 15: wrist ulnar deviation, and 16: wrist extension with closed hand.
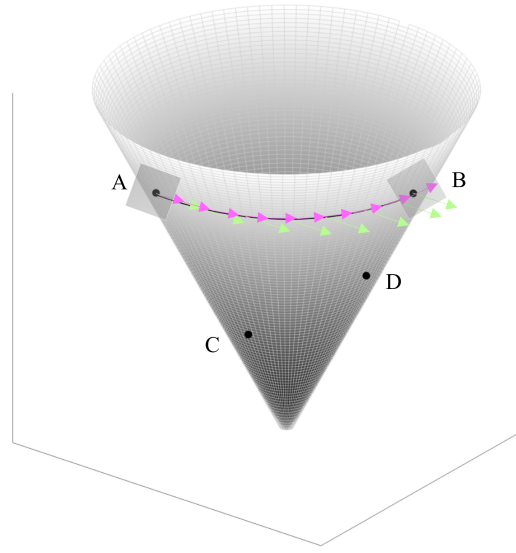
Figure 7. Euclidean transport is inappropriate in the manifold space. Point A is an SPD matrix that represents the EMG signal covariance for a given gesture, say a pinch, for a given subject. Point B is another SPD matrix representing the reference location for that gesture on the manifold: say, a particular reference subject's centroid for the pinch gesture. The length of the curve joining A and B is the geodesic distance between the two points. Points C and D, like A and B are SPD matrices that represent different gestures; for instance, the reference locations for a power grasp and a wrist flexion. *Green* arrows illustrate the Euclidean transport of a tangent vector from A to B. *Pink* arrows represent the *parallel* transport of a tangent vector from A to B. Parallel transport rotates the vector along the path from A to B so that the vector in the tangent plane of A remains in the tangent plane of B when transported. It can be seen that the Euclidean transport of a vector in the tangent plane of A results in vectors that are no longer in the tangent plane along the path to B. The Euclidean transport of vectors therefore cannot appropriately traverse a geodesic. The *square patches* at A and B represent the tangent planes.
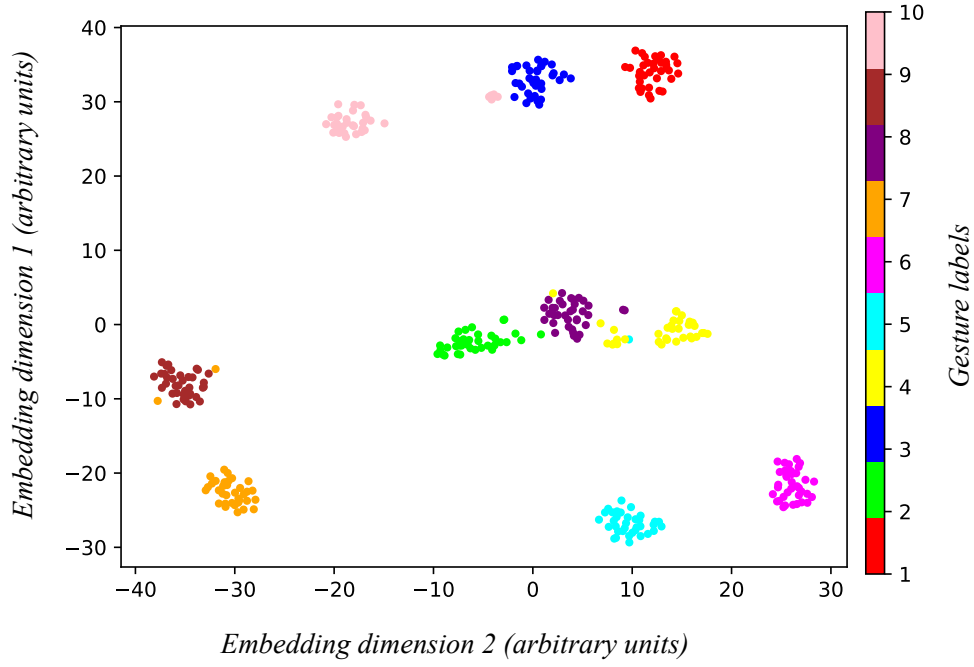
Figure 8. t-SNE of SPD matrices of *subject 0* from UCD-MYOVERSE-HAND-0 shows that different gestures within a subject have contrasting spatial patterns. We can classify these distinct gestures with supervised algorithms such as minimum distance to mean (MDM) and support vector machine (SVM) or unsupervised algorithms such as *k*-medoids clustering using Riemannian distance. Average classification accuracy across subjects using the above methods is presented in table 1. Embedding is colored according to gestures. The subject performed 10 gestures with each gesture repeated 36 times.

Ten gestures are: 1 - Down, 2 - Index finger pinch, 3 - Left, 4 - Middle finger pinch, 5 - Index point, 6 - Power grasp, 7 - Right, 8 - Two finger pinch, 9 - Up, 10 - Splay.

Author Contributions L. M. M. conceptualized the study and framed the problem statements. H. T. G. formulated the mathematical framework, implemented the codes, and collected the data. H. T. G. and L. M. M. authored and reviewed the manuscript.

Competing Interests

The authors declare no competing interests.

References

Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007. doi: 10.1137/050637996.

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A.-G. M., Elsig, S., Giatsidis, G., Bassetto, F., and Müller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific Data*, 1(1):1–13, 2014. doi: 10.1038/sdata.2014.53.

Enoka, R. M. and Duchateau, J. Muscle fatigue: what, why and how it influences muscle function. *Journal of Physiology*, 586(1):11–23, January 2008. doi: 10.1113/jphysiol.2007.139477.

Farina, D., Merletti, R., and Enoka, R. M. The extraction of neural strategies from the surface EMG. *Journal of Applied Physiology*, 96(4):1486–1495, 2004. doi: 10.1152/japplphysiol.01070.2003.

Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., and Li, J. Gesture recognition by instantaneous surface EMG images. *Scientific Reports*, 6:36571, 2016. doi: 10.1038/srep36571.

Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M., and Geng, W. A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. *PLOS ONE*, 13(10):e0206049, 2018. doi: 10.1371/journal.pone.0206049.

Hudgins, B., Parker, P., and Scott, R. N. A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 40(1):82–94, 1993. doi: 10.1109/10.204774.

Huebner, A., Faenger, B., Schenk, P., Scholle, H.-C., and Anders, C. Alteration of surface EMG amplitude levels of five major trunk muscles by defined electrode location displacement. *Journal of Electromyography and Kinesiology*, 25(2):214–223, 2015. doi: 10.1016/j.jelekin.2015.01.004.

Jayasumana, S., Hartley, R., Salzmann, M., Li, H., and Harandi, M. Kernel methods on Riemannian manifolds with Gaussian RBF kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(12):2464–2477, 2015. doi: 10.1109/TPAMI.2015.2408352.

Kaufman, L. and Rousseeuw, P. J. Partitioning around medoids (program pam). In *Wiley Series in Probability and Statistics*, pp. 68–125. John Wiley & Sons, Inc., Hoboken, NJ, USA, March 8 1990. ISBN 978-0-470-31680-1. doi: 10.1002/9780470316801.ch2. URL https://doi.org/10.1002/9780470316801.ch2. Retrieved 2021-06-13.

Khushaba, R. N., Al-Timemy, A. H., Samuel, O. W., and Scheme, E. J. Myoelectric control with fixed convolution-based time-domain feature extraction: Exploring the spatio–temporal interaction. *IEEE Transactions on Human-Machine Systems*, 52(6):1247–1257, 2022. doi: 10.1109/THMS.2022.3184572.

Kleine, B. U., Stegeman, D. F., Mund, D., and Anders, C. Influence of motoneuron firing synchronization on sEMG characteristics in dependence of electrode position. *Journal of Applied Physiology (1985)*, 91(4):1588–1599, October 2001. doi: 10.1152/jappl.2001.91.4.1588.

Lin, Z. Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition. *SIAM Journal on Matrix Analysis and Applications*, 40(4):1353–1370, 2019. doi: 10.1137/18M1195746.

Lucas, M.-F., Gaufriau, A., Pascual, S., Doncarli, C., and Farina, D. Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. *Biomedical Signal Processing and Control*, 3(2):169–174, 2008. doi: 10.1016/j.bspc.2007.09.003. Surface Electromyography.

Malešević, N., Olsson, A., Sager, P., Andersson, E., Cipriani, C., Controzzi, M., Björkman, A., and Antfolk, C. A database of high-density surface electromyogram signals comprising 65 isometric hand gestures. *Scientific Data*, 8(1):63, February 2021. doi: 10.1038/s41597-021-00861-9.

Montazerin, M., Rahimian, E., Naderkhani, F., Atashzar, S. F., Yanushkevich, S., and Mohammadi, A. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals. *Scientific Reports*, 13(1):11000, July 2023. doi: 10.1038/s41598-023-38161-4.

Quivira, F., Koike-Akino, T., Wang, Y., and Erdogmus, D. Translating sEMG signals to continuous hand poses using recurrent neural networks. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 166–169, 2018. doi: 10.1109/BHI.2018.8333395.

Rahimian, E., Zabihi, S., Atashzar, S. F., Asif, A., and Mohammadi, A. XceptionTime: Independent time-window XceptionTime architecture for hand gesture classification. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1304–1308, 2020a. doi: 10.1109/ICASSP40776.2020.9054586.

Rahimian, E., Zabihi, S., Atashzar, S. F., Asif, A., and Mohammadi, A. Surface EMG-based hand gesture recognition via hybrid and dilated deep neural network architectures for neurorobotic prostheses. *Journal of Medical Robotics Research*, 5(01n02):2041001, 2020b. doi: 10.1142/S2424905X20410014.

Rahimian, E., Zabihi, S., Asif, A., Farina, D., Atashzar, S. F., and Mohammadi, A. TEMGNet: Deep transformer-based decoding of upperlimb sEMG for hand gestures recognition. *arXiv preprint arXiv:2109.12379*, 2021.

Sun, T., Hu, Q., Gulati, P., and Atashzar, S. F. Temporal dilation of deep LSTM for agile decoding of sEMG: Application in prediction of upper-limb motor intention in neurorobotics. *IEEE Robotics and Automation Letters*, 6(4):6212–6219, 2021. doi: 10.1109/LRA.2021.3077167.

Van der Maaten, L. and Hinton, G. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanhalli, M., and Geng, W. A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface. *Pattern Recognition Letters*, 119:131–138, 2019. doi: 10.1016/j.patrec.2018.09.003.

Xiong, D., Zhang, D., Zhao, X., Chu, Y., and Zhao, Y. Learning non-euclidean representations with SPD manifold for myoelectric pattern recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:1514–1524, 2022. doi: 10.1109/TNSRE.2022.3182030.

Zardoshti-Kermani, M., Wheeler, B., Badie, K., and Hashemi, R. EMG feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333, 1995. doi: 10.1109/86.481972.

A. Manifold SVM kernel

Here, we prove that the kernel used in *section 2.3.5* is a valid kernel.

We shall prove that the kernel $\mathscr{K} = \exp(-\gamma d^2(L_1, L_2))$ for $L_1, L_2 \in$ CHOLESKY SPACE is a valid kernel for all $\gamma > 0$.

For a kernel to be valid, it must be positive definite (Jayasumana et al., 2015).

From *theorem 5.2* in Jayasumana et al., a kernel

$$\mathscr{K} : \text{CHOLESKY SPACE} \times \text{CHOLESKY SPACE} \to \mathbb{R}$$
$$:= \exp(-\gamma d^2(L_1, L_2))$$

is positive definite for all $\gamma > 0$ if and only if $d^2(L_1, L_2)$ is negative definite.

From *lemma 5.5* in Jayasumana et al., if

$$\psi : \text{CHOLEASKY SPACE} \to F,$$

where F is the Frobenius inner product space is a function, then,

$$\mathscr{O} : \text{CHOLESKY SPACE} \times \text{CHOLESKY SPACE} \to \mathbb{R}$$

defined by

$$\mathscr{O}(L_1, L_2) = ||\psi(L_1) - \psi(L_2)||_F^2$$

is negative definite.

Following *theorem 6.1* (Jayasumana et al., 2015),

$$\mathscr{K} : \text{CHOLESKY SPACE} \times \text{CHOLESKY SPACE} \to \mathbb{R}$$

given by

$$\mathscr{K}(L_1, L_2) = \exp(-\gamma d^2(L_1, L_2))$$

is a positive definite kernel for all $\gamma > 0$ if and only if there exists a function

$$\psi : \text{CHOLESKY SPACE} \to F$$

such that

$$d(L_1, L_2) = ||\psi(L_1) - \psi(L_2)||_F.$$

We can now define

$$\psi = \lfloor L \rfloor + \log \mathbb{D}(L)$$

for all $L \in \text{CHOLESKY SPACE}$.

We have

$$d(L_1, L_2) = \{|| \lfloor L_1 \rfloor - \lfloor L_2 \rfloor ||_F^2 + \\ || \log \mathbb{D}(L_1) - \log \mathbb{D}(L_2) ||_F^2 \}^{1/2}.$$

$$||\psi(L_1) - \psi(L_2)||_F = \\ || \lfloor L_1 \rfloor + \log \mathbb{D}(L_1) - \lfloor L_2 \rfloor - \log \mathbb{D}(L_2) ||_F \\ = \{|| \lfloor L_1 \rfloor - \lfloor L_2 \rfloor ||_F^2 + || \log \mathbb{D}(L_1) - \log \mathbb{D}(L_2) ||_F^2 \}^{1/2}$$

Therefore,

$$d(L_1, L_2) = ||\psi(L_1) - \psi(L_2)||_F.$$

Hence, the kernel

$$\mathscr{K} = \exp\left(-\gamma d^2(L_1, L_2)\right)$$

is a valid kernel.

In the above, $d(.)$ is defined according to equation 1 and $||.||_F$ is the Frobenius norm.

B. Parallel transport

Here, we explain parallel transport in detail.

As given in Lin, the logarithm map from the manifold to its tangent space at L is given by

$$\widetilde{\text{Log}}_L K = \lfloor K \rfloor - \lfloor L \rfloor + \mathbb{D}(L) \log\{\mathbb{D}(L)^{-1}\mathbb{D}(K)\}, \quad (6)$$

where L, $K \in \text{CHOLESKY SPACE}$ and $\widetilde{\text{Log}}_L K \in \text{TANGENT SPACE}$

The exponential map from the tangent space to the manifold space is given by,

$$\widetilde{\text{Exp}}_L X = \lfloor L \rfloor + \lfloor X \rfloor + \mathbb{D}(L) \exp\{\mathbb{D}(X)\mathbb{D}(L)^{-1}\}, \quad (7)$$

where L, $\widetilde{\text{Exp}}_L X \in \text{CHOLESKY SPACE}$ and $X \in \text{TANGENT SPACE}$.

A tangent vector $X \in \text{TANGENT SPACE}$ is parallelly transported to the tangent vector

$$\lfloor X \rfloor + \mathbb{D}(K)\mathbb{D}(L)^{-1}\mathbb{D}(X) \quad (8)$$

at K (Lin, 2019).

Given a set of subjects (S) and a set of gestures (G), let us denote gestures of a particular hand movement (Cholesky factorized covariance SPD matrices) belonging to a particular subject as $\mathcal{G}_s^g$, $s \in S$ and $g \in G$. Parallel transport algorithm is as defined below.

Algorithm 1 Parallel transport
___
Input: Cholesky matrices $\mathcal{G}_s^g$
Output: Parallelly transported Cholesky matrices.

Compute the Riemannian mean
$\bar{\mathcal{G}}_s^g = \mathbb{E}(\text{all elements in } \mathcal{G}_s^g)$ using equation 2.

Map all elements in $\mathcal{G}_s^g$ to the tangent space at $\bar{\mathcal{G}}_s^g$ using equation 6.
Denote this mapped set as $\text{TANGENT MAPPED}(\mathcal{G}_s^g)$.

Parallelly transport the tangent vectors in $\text{TANGENT MAPPED}(\mathcal{G}_s^g)$ (using equation 8) to the tangent space at $\bar{\mathcal{G}}_{s\,=\,\text{Reference}}^g$. We choose Subject 0 as the reference.

Map all the parallelly transported vectors in the tangent space of $\bar{\mathcal{G}}_{s\,=\,\text{Reference}}^g$ back to the manifold space using equation 7.

For visualization purposes in figure 6, we assume ground truth to identify clusters of a particular gesture g in G. In practice, EMG signals corresponding to different hand gestures can be classified using unsupervised methods such as k-medoids clustering. In this case, the accuracy of parallel transport is bounded by the accuracy of the unsupervised classifier. Algorithm 1 presents a method for aligning multivariate EMG time series from different individuals to the same neighborhood in the manifold space.

C. Detailed results

Here, we present detailed subject-wise results for UCD-MYOVERSE-HAND-0 and NINAPRO: DATABASE 2-EXERCISE 1 datasets .

C.1. Results for UCD-MYOVERSE-HAND-0

The UCD-MYOVERSE-HAND-0 database consists of 30 subjects. Each subject performed 10 different hand gestures (a total of 360 trials: 10 gestures, each repeated 36 times). Therefore, the chance-level decoding accuracy is $\frac{1}{10} = 0.1$. We present the detailed subject-wise results in table 5.

Subject number	Classification methods		
	MDM	SVM $(\gamma = 1)$	k-medoids
0	0.97	0.99	0.94
1	0.61	0.69	0.48
2	0.62	0.76	0.61
3	0.76	0.83	0.60
4	0.92	0.93	0.74
5	0.82	0.86	0.63
6	0.94	0.97	0.88
7	0.94	0.95	0.74
8	0.97	0.97	0.93
9	0.92	0.94	0.74
10	0.84	0.86	0.62
11	0.79	0.81	0.62
12	0.95	0.97	0.93
13	0.84	0.84	0.73
14	0.82	0.84	0.64
15	0.57	0.66	0.50
16	0.92	0.93	0.76
17	0.85	0.89	0.85
18	0.53	0.57	0.35
19	0.82	0.84	0.61
20	0.83	0.87	0.72
21	0.82	0.89	0.72
22	0.93	0.94	0.70
23	0.88	0.92	0.91
24	0.55	0.55	0.26
25	0.78	0.82	0.55
26	0.83	0.89	0.83
27	0.78	0.86	0.68
28	0.99	1.0	0.99
29	0.89	0.91	0.72
Average	**0.82**	**0.86**	**0.70**

Table 5. Classification accuracy for *30* subjects in UCD-MYOVERSE-HAND-0. Data from the first 3 sessions are used for training (first 18 repetitions of each gesture), and data from the last 3 sessions (last 18 repetitions of each gesture) are used for testing the MDM and SVM algorithms.

C.2. Results for NINAPRO: DATABASE 2-EXERCISE 1

The NINAPRO: DATABASE 2-EXERCISE 1 consists of 40 subjects. Each subject performed 17 different hand gestures (a total of 102 trials: 17 gestures, each repeated 6 times). Therefore, the chance-level decoding accuracy is $\frac{1}{17} = 0.059$. We present the detailed subject-wise results in table 6.

Subject number	Classification methods		
	MDM	SVM ($\gamma = 8$)	k-medoids
0	1.0	1.0	0.90
1	0.94	0.97	0.86
2	0.97	0.97	0.78
3	0.94	0.94	0.79
4	0.97	1.0	0.85
5	0.85	0.85	0.79
6	0.88	0.88	0.69
7	0.85	0.91	0.80
8	0.97	1.0	0.85
9	0.88	0.91	0.87
10	0.97	0.91	0.84
11	0.85	0.91	0.79
12	0.97	0.94	0.89
13	1.0	0.97	0.93
14	0.91	0.97	0.81
15	0.91	0.88	0.66
16	0.97	0.97	0.78
17	0.88	0.91	0.83
18	0.91	0.97	0.85
19	0.88	0.88	0.83
20	0.88	0.94	0.84
21	0.91	0.91	0.85
22	0.88	0.88	0.83
23	0.91	0.94	0.78
24	0.91	0.94	0.84
25	0.97	0.97	0.85
26	0.97	0.94	0.87
27	0.91	0.94	0.87
28	0.91	0.88	0.75
29	0.82	0.91	0.76
30	0.88	0.88	0.68
31	0.91	0.88	0.84
32	0.94	0.97	0.85
33	0.76	0.82	0.73
34	0.97	0.97	0.84
35	0.97	0.88	0.80
36	0.94	1.0	0.89
37	0.97	0.97	0.96
38	0.97	0.94	0.84
39	0.97	0.97	0.86
Average	**0.92**	**0.93**	**0.82**

Table 6. Classification accuracy for 40 subjects in NINAPRO: DATABASE 2-EXERCISE 1. Following the works in Sun et al. and Rahimian et al., for each gesture, we use repetitions *1, 3, 4,* and *6* for training and repetitions *2* and *5* for testing of MDM and SVM algorithms.