DOMAIN ADAPTATION USING RIEMANNIAN GEOMETRY OF SPD MATRICES

Gal Maman[1]*, Or Yair*[1]*, Danny Eytan*[2]*, Ronen Talmon*[1]

[1] Viterbi Faculty of Electrical Engineering
Technion, Israel Institute of Technology
Haifa, 3200003, Israel

[2] Faculty of Medicine
Technion, Israel Institute of Technology
Haifa, 3200003, Israel

ABSTRACT

In this paper, we propose a new unsupervised domain adaptation method based on the Riemannian geometry of Symmetric Positive-Definite (SPD) matrices. The proposed domain adaptation is based on parallel transport (PT) and moments alignment. We show that this method facilitates meaningful comparisons between data points from different domains, while preserving the inherent internal structure of each domain. Experimental results demonstrate the adaptation of high-dimensional noisy electrophysiological signals collected from different subjects.

Index Terms— Riemannian manifolds, parallel transport, transfer learning, high-dimensional signal analysis

1. INTRODUCTION

Recent technological progress has led to a surge in the availability of highly complex and heterogeneous data in a broad range of domains. This paper addresses two emerging problems in the analysis of such data: domain adaptation and coping with their high-dimensionality.

Notably, data in high dimension usually do not live in a Euclidean space. Therefore, applying analysis or learning techniques, which typically rely on Euclidean distances, directly to the data often leads to poor performance. By departing from Euclidean spaces, and considering instead non-Euclidean Riemannian geometry facilitates the extraction and utilization of the structure of the data, allowing for meaningful data point comparisons in high dimension. Particularly in this work, we focus on the Riemannian geometry of covariance matrices, which are Symmetric Positive-Definite (SPD) and constitute a cone manifold with a known Riemannian metric [1, 2]. The combination of covariance matrices as high-dimensional data features with their inherent Riemannian geometry has proven to be powerful in many tasks [3–5]. For example in [4, 6], a new representation in a Euclidean space was obtained by projecting each covariance matrix to the tangent plane of the Riemannian manifold at the mean of the data. This representation was shown to be highly successful in capturing the essence of complex data, leading to state of the art classification results in a large and broad variety of applications, e.g., Brain Computer Interface (BCI) [6] and medical imaging [3].

However, when considering data arising from several domains, for example, using different acquisition systems, in different sessions, or from different subjects, the covariance matrices might exhibit highly different structures, rendering the mere use of the Riemannian metric and the projection to a single tangent plane insufficient. Indeed, we show here that the covariance matrices of Electroencephalography (EEG) recordings from multiple subsets lie in different regions of the SPD manifold, and therefore the Euclidean representation proposed in [6] does not accommodate appropriate comparisons. This problem calls for domain adaptation, where a given model or representation, which is well performing in a particular domain, is adapted to a different, yet related domain [7, 8].

In this paper, we propose to use a geometric preserving transformation for domain adaptation. Specifically, we further exploit the Riemannian geometry of SPD matrices and *parallel transport* (PT) the covariance matrices along the cone manifold to a common location [9, 10]. Indeed, in [11] PT attained a joint representation of several data sets that appropriately accommodates multiple domains in BCI and sleep research problems. One shortcoming of PT is that it does not take into account the internal structure of the data from each domain, but rather depends only on the Riemannian mean of the covariance matrices. Consequently, two data sets from two domains with different structures, but with the same Riemannian mean, will undergo the same adaptation via PT. As we show in the sequel, this could significantly hamper performance.

Here, we propose to complement PT with an additional procedure based on moments alignment that takes into account the internal structure of data from each domain as well as the inter-relations between the domains. We show that this procedure facilitates improved domain adaptation. Experimental results demonstrate adequate unsupervised adaptation of high-dimensional noisy EEG recordings [12], collected from different subjects. Moreover, we show that based on our method, training a classifier on recordings from multiple subjects, and then, testing it on recordings from a new, unseen subject without any new labels, is possible and gives rise to accurate transfer learning.

This paper is organized as follows. In Section 2, we present preliminaries on the Riemannian geometry of SPD matrices. In Section 3, we formulate the problem, present the proposed domain adaptation method. Section 4 shows experimental results on high-dimensional noisy electrophysiological signals. Finally, we conclude the paper in Section 5.

2. RIEMANNIAN GEOMETRY OF SPD MATRICES

A symmetric matrix $\boldsymbol{P} \in \mathbb{R}^{d\times d}$ is positive-definite if all its eigenvalues are strictly positive, or equivalently, if $\boldsymbol{v}^T\boldsymbol{P}\boldsymbol{v} > 0$ for every nonzero vector $\boldsymbol{v}$. The set of all SPD matrices is an open convex cone, constituting a differential Riemannian manifold $\mathcal{M}$. Let $\mathcal{T}_{\boldsymbol{P}}\mathcal{M}$ be the tangent space at the point $\boldsymbol{P} \in \mathcal{M}$ equipped with the following inner product

$$\left\langle \boldsymbol{S}_1, \boldsymbol{S}_2 \right\rangle_{\mathcal{T}_{\boldsymbol{P}}\mathcal{M}} = \left\langle \boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{S}_1\boldsymbol{P}^{-\frac{1}{2}}, \boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{S}_2\boldsymbol{P}^{-\frac{1}{2}} \right\rangle, \qquad (1)$$

where $\boldsymbol{S}_1, \boldsymbol{S}_2 \in \mathcal{T}_{\boldsymbol{P}}\mathcal{M}$ are two symmetric matrices. This Riemannian manifold has a unique geodesic between any two SPD matrices $\boldsymbol{P}_1, \boldsymbol{P}_2 \in \mathcal{M}$, whose length defines a distance given by:

$$\delta_R^2(\boldsymbol{P}_1,\boldsymbol{P}_2) \triangleq \left\|\log\left(\boldsymbol{P}_1^{-\frac{1}{2}}\boldsymbol{P}_2\boldsymbol{P}_1^{-\frac{1}{2}}\right)\right\|_F^2 = \sum_{i=1}^{d}\log^2(\lambda_i), \tag{2}$$

where $\|\cdot\|_F$ is the Frobenius norm, $\log(\cdot)$ is the matrix logarithm and λ_i is the ith eigenvalue of the matrix $\boldsymbol{P}_1^{-\frac{1}{2}}\boldsymbol{P}_2\boldsymbol{P}_1^{-\frac{1}{2}}$. For an SPD matrix $\boldsymbol{P}\in\mathcal{M}$, the Logarithm chart mapping an SPD matrix $\boldsymbol{P}_i\in\mathcal{M}$ to $\boldsymbol{S}_i\in\mathcal{T}_{\boldsymbol{P}}\mathcal{M}$ is defined by

$$\boldsymbol{S}_i = \mathrm{Log}_{\boldsymbol{P}}(\boldsymbol{P}_i) \triangleq \boldsymbol{P}^{\frac{1}{2}}\log\left(\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{P}_i\boldsymbol{P}^{-\frac{1}{2}}\right)\boldsymbol{P}^{\frac{1}{2}}. \tag{3}$$

The inverse map is the Exponential chart, given by

$$\boldsymbol{P}_i = \mathrm{Exp}_{\boldsymbol{P}}(\boldsymbol{S}_i) \triangleq \boldsymbol{P}^{\frac{1}{2}}\exp\left(\boldsymbol{P}^{-\frac{1}{2}}\boldsymbol{S}_i\boldsymbol{P}^{-\frac{1}{2}}\right)\boldsymbol{P}^{\frac{1}{2}}. \tag{4}$$

The Riemannian mean of a set $\left\{\boldsymbol{P}_i\in\mathcal{M}\right\}_{i=1}^N$ is defined by the Fréchet mean:

$$\overline{\boldsymbol{P}} \triangleq \underset{\boldsymbol{P}\in\mathcal{M}}{\arg\min}\sum_{i=1}^{N}\delta_R^2(\boldsymbol{P},\boldsymbol{P}_i) \tag{5}$$

Particularly, the Riemannian mean of two SPD matrices $\boldsymbol{P}_1,\boldsymbol{P}_2\in\mathcal{M}$ is located at the midpoint of the connecting geodesic and is explicitly given by

$$\overline{\boldsymbol{P}} = \boldsymbol{P}_1^{\frac{1}{2}}\left(\boldsymbol{P}_1^{-\frac{1}{2}}\boldsymbol{P}_2\boldsymbol{P}_1^{-\frac{1}{2}}\right)^{\frac{1}{2}}\boldsymbol{P}_1^{\frac{1}{2}}$$

The Riemannian mean of a general number of matrices can be computed using an iterative procedure [4, 11]. Given a set $\left\{\boldsymbol{P}_i\in\mathcal{M}\right\}_{i=1}^N$, the pairwise Riemannian distances can be approximated by

$$\delta_R(\boldsymbol{P}_i,\boldsymbol{P}_j) \approx \left\|\hat{\boldsymbol{S}}_i - \hat{\boldsymbol{S}}_j\right\|_F$$

where $\hat{\boldsymbol{S}}_i = \overline{\boldsymbol{P}}^{-\frac{1}{2}}\mathrm{Log}_{\overline{\boldsymbol{P}}}(\boldsymbol{P}_i)\,\overline{\boldsymbol{P}}^{-\frac{1}{2}}$ and $\overline{\boldsymbol{P}}$ is the Riemannian mean of $\{\boldsymbol{P}_i\}$. Lastly, we denote the vector representation of any symmetric matrix $\boldsymbol{S}$ by

$$\boldsymbol{s} = \mathrm{vec}(\boldsymbol{S})$$

where $\mathrm{vec}(\cdot)$ is a vector consisting of the elements of the upper triangular of a symmetric matrix $\boldsymbol{S}$, with $\sqrt{2}$ weights on its off-diagonal elements such that

$$\|\boldsymbol{s}_i - \boldsymbol{s}_j\|_2 = \|\boldsymbol{S}_i - \boldsymbol{S}_j\|_F$$

for any two symmetric matrices $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$, where $\boldsymbol{s}_i = \mathrm{vec}(\boldsymbol{S}_i)$ and $\boldsymbol{s}_j = \mathrm{vec}(\boldsymbol{S}_j)$.

3. PROPOSED METHOD

Let $\left\{\boldsymbol{X}_i^{(1)}\right\}_{i=1}^{N_1},\left\{\boldsymbol{X}_i^{(2)}\right\}_{i=1}^{N_2},\ldots,\left\{\boldsymbol{X}_i^{(K)}\right\}_{i=1}^{N_K}$ be K subsets, where each subset k contains N_k matrices of data observations $\boldsymbol{X}_i^{(k)}\in\mathbb{R}^{d\times T_i^{(k)}}, i=1,\ldots,N_k$, such that d is the observation (features) dimension and $T_i^{(k)}$ is the observation length. Let $\boldsymbol{P}_i^{(k)}\in\mathbb{R}^{d\times d}$ denote the sample covariance of the data matrix $\boldsymbol{X}_i^{(k)}\in\mathbb{R}^{d\times T_i^{(k)}}$, given by $\boldsymbol{P}_i^{(k)} \triangleq \frac{1}{T_i^{(k)}}\boldsymbol{X}_i^{(k)}\left(\boldsymbol{X}_i^{(k)}\right)^T$, where for simplicity we assume observations with zero mean.

We focus here on a body of previous work [1, 4, 5], where covariance matrices, which are particular SPD matrices, are used as features, exploiting their Riemannian geometry. In their seminal work [4], Barachant *et al.* proposed to project each covariance matrix $\boldsymbol{P}_i^{(k)}$ to the tangent space $\mathcal{T}_{\overline{\boldsymbol{P}}}\mathcal{M}$, where $\overline{\boldsymbol{P}}$ is the Riemannian mean of all the matrices. This representation was shown to be highly successful in capturing the essence of complex data. However, we show here that when the data sets live in various domains this representation highly depends on the domain.

To facilitate comparisons between data from various domains, we propose to further exploit the Riemannian geometry of SPD matrices and to PT all the matrices along the cone manifold to a common location. This method was first introduced in [11], and here we briefly review its main principles and steps.

First, the Riemannian mean of each subset k, denoted by $\overline{\boldsymbol{P}}^{(k)}$, is computed. Second, a common point $\overline{\boldsymbol{Q}}\in\mathcal{M}$ is set, for example, the Riemannian mean of $\left\{\overline{\boldsymbol{P}}^{(k)}\right\}_{k=1}^K$. Third, the PT from $\overline{\boldsymbol{P}}^{(k)}$ to $\overline{\boldsymbol{Q}}$ is applied to each SPD matrix $\boldsymbol{P}_i^{(k)}$, which is defined by

$$\boldsymbol{\Gamma}_i^{(k)} = \Gamma_{\overline{\boldsymbol{P}}^{(k)}\to\overline{\boldsymbol{Q}}}\left(\boldsymbol{P}_i^{(k)}\right) = \boldsymbol{E}\boldsymbol{P}_i^{(k)}\boldsymbol{E}^T, \tag{6}$$

where $\boldsymbol{E} = \left(\overline{\boldsymbol{Q}}\left(\overline{\boldsymbol{P}}^{(k)}\right)^{-1}\right)^{\frac{1}{2}}$. Finally, the transported matrices $\left\{\boldsymbol{\Gamma}_i^{(k)}\right\}_{i,k}$ are projected to the tangent plane at $\overline{\boldsymbol{Q}}$ and represented as vectors in a Euclidean space

$$\boldsymbol{z}_i^{(k)} = \mathrm{vec}\left(\overline{\boldsymbol{Q}}^{-\frac{1}{2}}\mathrm{Log}_{\overline{\boldsymbol{Q}}}(\boldsymbol{\Gamma}_i^{(k)})\overline{\boldsymbol{Q}}^{-\frac{1}{2}}\right)\in\mathbb{R}^{D\times 1}. \tag{7}$$

where $D = d(d+1)/2$. For details and rigorous analysis, we refer the readers to [11].

One shortcoming of PT is its sole dependence on the Riemannian mean of each subset, and that it does not take into account the structure of the subset. Therefore, we propose to complement PT with an additional refinement step, particularly, a unitary rotation in order to align the primary sources of variability of the subsets.

Formally, let $\left\{\boldsymbol{z}_i^{(1)}\right\}_{i=1}^{N_1},\left\{\boldsymbol{z}_i^{(2)}\right\}_{i=1}^{N_2},\ldots,\left\{\boldsymbol{z}_i^{(K)}\right\}_{i=1}^{N_K}$ be the vectors in $\mathbb{R}^D$ resulted from PT (7). For each $k\in\{1,2,\ldots,K\}$, let

$$\boldsymbol{Z}^{(k)} = \left[\boldsymbol{z}_1^{(k)},\boldsymbol{z}_2^{(k)},\ldots,\boldsymbol{z}_{N_k}^{(k)}\right]\in\mathbb{R}^{D\times N_k} \tag{8}$$

be a matrix consisting of the new representation of the data in each data set. In addition, let $\boldsymbol{U}^{(k)}\in\mathbb{R}^{D\times D}$ be a matrix whose columns are the left-singular vectors of $\boldsymbol{Z}^{(k)}$, ordered by their respective singular-values, from the largest to the smallest. Note that by our previous assumption, the sample mean of the vectors is zero. The proposed adaptation refinement is carried out by the following rotation:

$$\boldsymbol{Y}^{(k)} = \left(\boldsymbol{U}^{(k)}\right)^T\boldsymbol{Z}^{(k)}, \quad \forall k.$$

Several remarks due at this point. First, we assume that the singular-values are simple so that each singular-vector is unique. Second, to circumvent possible singular-vector sign ambiguity, under the assumption that the angles between the corresponding singular vectors are acute, we flip the orientation of the singular-vectors to satisfy this assumption. Namely, we set a certain subset to be a reference set. Without loss of generality, let subset $k=1$ be the

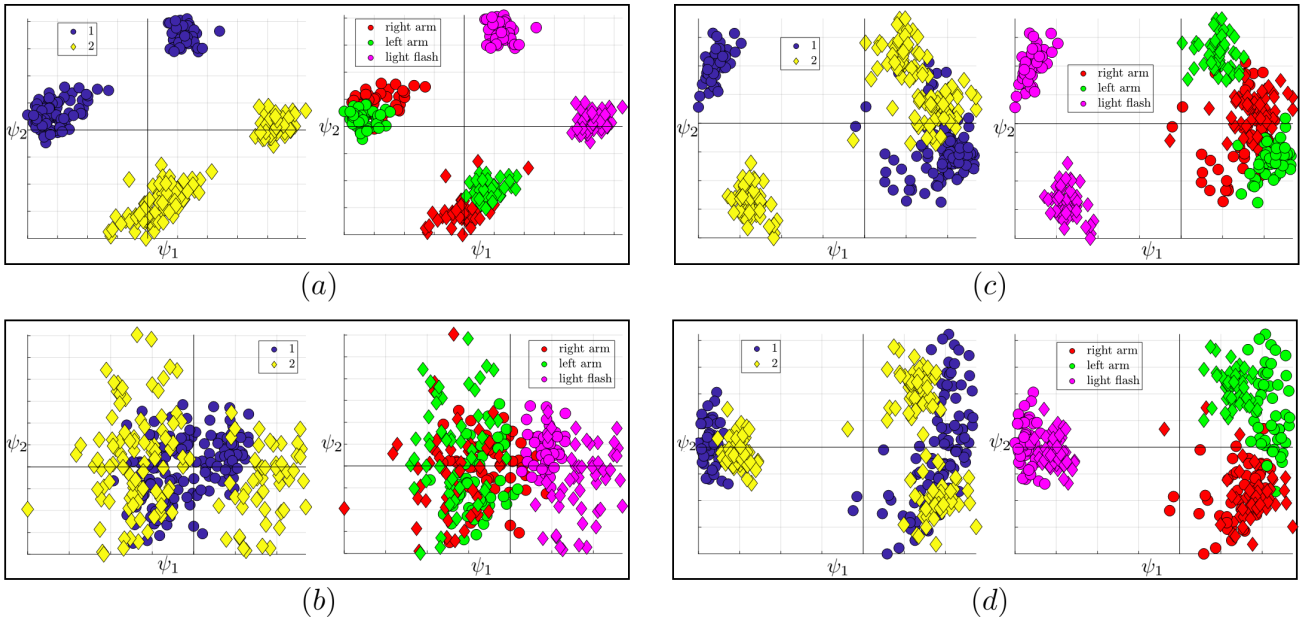

Fig. 1. The first two principal components of the representation of the trials obtain by the different algorithms: (a) "Baseline", (b), "Mean Transport", (c) steps (1) - (3a) of Algorithm 1 and (d) Algorithm 1. Each point represents a single trial. Trials from two subjects and the response to two somatosensory stimuli and one visual stimulus are presented. The plots on the left are colored by the subject and the plots on the right are colored by the stimulus.

reference set. For each of the other subsets $k \neq 1$, we flip the orientation of the singular-vector $\boldsymbol{u}_j^{(k)}$, given as the jth column of $\boldsymbol{U}^{(k)}$, according to:

$$\boldsymbol{u}_j^{(k)} \leftarrow \text{sign}\left(\left\langle \boldsymbol{u}_j^{(1)}, \boldsymbol{u}_j^{(k)} \right\rangle\right)\boldsymbol{u}_j^{(k)}, \quad \forall j.$$

The domain adaptation algorithm is summarized in Algorithm 1.

4. EXPERIMENTAL RESULTS

We applied the proposed domain adaptation to EEG recordings from 64 channels organized in subsets corresponding to 11 subjects in ages ranging from 7 to 16 years, where each subject is considered as a different domain. The recordings were conducted in trials of about 1 second. In each trial, the EEG response to one of several stimuli was recorded. More details about the data appear in [13]. There are many reports that study the electrical response patterns to stimulation, named evoked potentials [14–17]. However, they all rely on specific feature extraction that are heavily tailored and adapted to the specific study and paradigm, in contrast to the currently proposed method.

Three pre-processing steps were applied: (i) down sampling to 1KHz, (ii) exclusion of malfunctioning electrodes and highly noisy trials, and (iii) applying the Fourier transform to each channel recording in each trial and taking the absolute value. After pre-processing, the entire data set included 37 electrodes (channels) and contained $80 - 500$ repeated trials per stimulus for each subject. Let $\boldsymbol{X}_i^{(k)} \in \mathbb{R}^{37\times T_{i,k}}$ be the absolute value of the Fourier transform of the EEG recordings from the k-th subject at the i-th trial.

For illustration purposes, we first considered recordings only from two subjects and their EEG responses to 2 somatosensory stimuli (right arm and left arm nerve stimulation) and 1 visual stimulus (light projection with a flash). From each stimulus, we collected the responses from 50 repeated trials. Consequently, we organized the recordings in two subsets $\left\{\boldsymbol{X}_i^{(1)}\right\}_{i=1}^{150}$ and $\left\{\boldsymbol{X}_i^{(2)}\right\}_{i=1}^{150}$.

We applied Algorithm 1 to $\left\{\boldsymbol{X}_i^{(1)}\right\}_{i=1}^{150}$ and $\left\{\boldsymbol{X}_i^{(2)}\right\}_{i=1}^{150}$ and compared the result to two other methods. In the first method, which we term "Baseline", the covariance matrices are used as features with their Riemannian distance, without additional processing. In the second method, which we term "Mean Transport" (MT), the Riemannian mean of each subset is subtracted from the covariance matrices as proposed in [4]. To emphasize the contribution of the different algorithmic steps, and particularly, the refinement by moments alignment, we also report the result of steps (1)-(3a) of Algorithm 1.

For visualization purposes, we applied PCA to the obtained representation of the trials by the different algorithms and depict in Figure 1 the two principal components. Consequently, each point in the figure is associated with the obtained representation of one trial. Circles mark trials of subject 1 and diamonds mark trials of subject 2. The left and right scatter plots only differ by color – on the left, the trials are colored by the subjects, and on the right, the trials are colored by the different stimuli.

Figure 1(a) presents the PCA of the representation obtained by the "Baseline" method. We observe that trials with the same stimulus applied to different subjects are embedded in different locations. Figure 1(b) presents the PCA of the representation obtained by the "Mean Transport" method. On the left, we observe that the trials are not clustered by the subjects, implying on some degree of domain adaptation. However, on the right, we observe that the internal structure is lost, i.e., responses to the somatosensory stimuli are mixed. Figure 1(c) presents the PCA of the representation obtained by steps (1)-(3a) of Algorithm 1 and Figure 1(d) presents the representation

Algorithm 1 Domain adaptation using Riemannian geometry

Input: K subsets $\left\{\boldsymbol{X}_i^{(k)} \in \mathbb{R}^{d\times T_i^{(k)}}\right\}_{i=1}^{N_k}, k=1,\ldots,K.$

Output: K matrices $\left\{\boldsymbol{Y}^{(k)} \in \mathbb{R}^{D\times N_k}\right\}_{k=1}^{K}, D = d(d+1)/2$ whose columns are the aligned vector representations of the input data matrices.

1. For each subset $\left\{\boldsymbol{X}_i^{(k)}\right\}_{i=1}^{N_k}$, compute the covariance matrices $\left\{\boldsymbol{P}_i^{(k)}\right\}_{i=1}^{N_k}$ and their Riemannian mean $\overline{\boldsymbol{P}}^{(k)}$.
2. Compute $\overline{\boldsymbol{Q}}$, the Riemannian mean of $\left\{\overline{\boldsymbol{P}}^{(k)}\right\}_{k=1}^{K}$.
3. For $k=1\ldots,K$:
   (a) For $i=1,\ldots,N_k$:
      i. Apply PT (6): $\boldsymbol{\Gamma}_i^{(k)} = \Gamma_{\overline{\boldsymbol{P}}^{(k)}\to\overline{\boldsymbol{Q}}}\left(\boldsymbol{P}_i^{(k)}\right)$
      ii. Project $\boldsymbol{\Gamma}_i^{(k)}$ to a Euclidean space (7):

$$\boldsymbol{z}_i^{(k)} = \mathrm{vec}\left(\overline{\boldsymbol{Q}}^{-\frac{1}{2}}\mathrm{Log}_{\overline{\boldsymbol{Q}}}\left(\boldsymbol{\Gamma}_i^{(k)}\right)\overline{\boldsymbol{Q}}^{-\frac{1}{2}}\right)$$

   (b) Compute $\boldsymbol{Z}^{(k)}$ (8) and apply SVD to obtain $\boldsymbol{U}^{(k)}$.
   (c) If $k \neq 1$, for each $j \in \{1,2,\ldots,D\}$ update the columns of $\boldsymbol{U}^{(k)}$:

$$\boldsymbol{u}_j^{(k)} \leftarrow \mathrm{sign}\left(\left\langle \boldsymbol{u}_j^{(1)}, \boldsymbol{u}_j^{(k)}\right\rangle\right)\boldsymbol{u}_j^{(k)}$$

   (d) Compute $\boldsymbol{Y}^{(k)} = \left(\boldsymbol{U}^{(k)}\right)^T\boldsymbol{Z}^{(k)}$.

obtained by the full algorithm. In Figure 1(d), we detect 3 distinct clusters corresponding to the 3 stimuli without dependence on the particular subject, indicating successful domain adaptation. In addition, Figure 1(c) entails that the refinement step in Algorithm 1 is essential.

Next, we extended the examination to the entire set of 11 subjects, considering their EEG responses to 3 different types of stimuli: somatosensory, visual and auditory. Figure 2(a) presents the results obtained by the "Baseline" method and Figure 2(b) presents the results obtained by Algorithm 1. We observe that Algorithm 1 attains three distinct clusters corresponding to the three different stimuli without dependence on the subjects identity. To provide quantitative assessment, we applied classification with leave one subject out cross-validation based on the representation obtained by the 3 completing methods. Figure 3 presents the classification results when using a cubic SVM. The obtained average classification rates are 63.2% ("Baseline"), 72.2% ("Mean Transport"), 84.9% (steps (1)-(3a) of Algorithm 1), and 94.9% (the full Algorithm 1). We remark that the sign of the second principal components of subjects 5 and 6 were manually flipped (note that their respective angles were 89° and 95°). In practice, this could be achieved with an extra knowledge of only few labeled points. Importantly, besides the above note, the domain adaptation reported here was achieved in an unsupervised manner. Namely, without any knowledge of the stimuli, the high-dimensional noisy EEG recordings from different subjects were adequately adapted, so that a classifier can be trained on data from one subject and applied to data from another test subject, without any labels of the test subject.

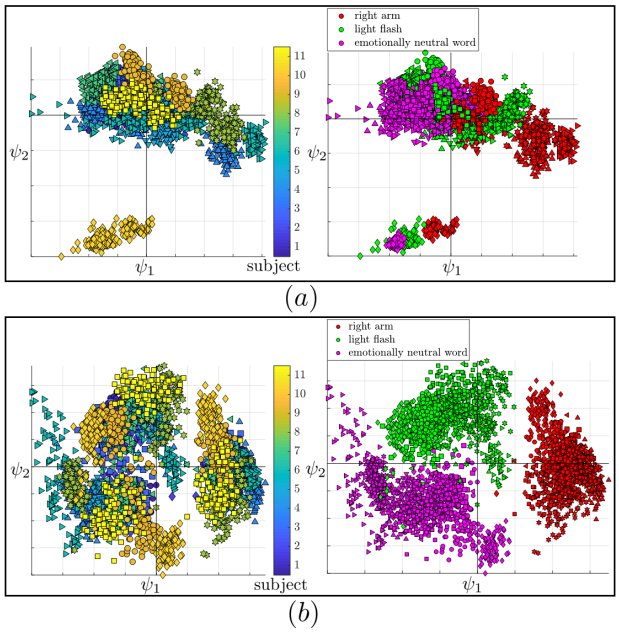

Fig. 2. The first two principal components of the representation of the trials obtain by: (a) "Baseline", (b) Algorithm 1. Each point represents a single trial. Trials from 11 subjects and the response to three different types of stimuli are presented. The plots the left are colored by the subject and the plots on the right are colored by the stimulus.

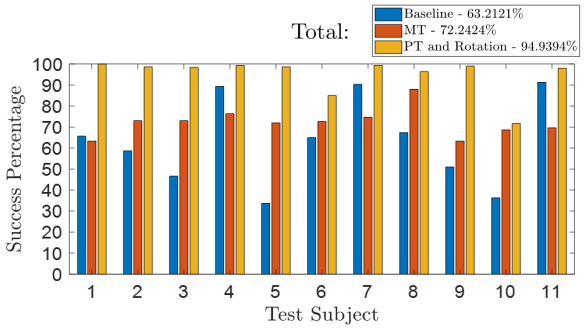

Fig. 3. Classification results of the response to three different stimuli from 11 subjects, using a cubic SVM, based on the representation obtained by the different algorithms.

5. CONCLUSIONS

Analyzing related data sets living in different domain is a long-standing problem, which calls for domain adaptation techniques. In this paper, we exploit the Riemannian geometry of SPD matrices and present an unsupervised approach for domain adaptation based on PT and moments alignment. The use of Riemannian geometry facilitates natural incorporation of both the intrinsic structure of each data set as well as the relations between the different data sets. Experimental results demonstrate the applicability of the presented domain adaptation method to a challenging problem involving high-dimensional noisy electrophysiological signals.

6. REFERENCES

[1] X. Pennec, P. Fillard, and N. Ayache, "A riemannian framework for tensor computing," *International Journal of computer vision*, vol. 66, no. 1, pp. 41–66, 2006.

[2] R. Bhatia, *Positive definite matrices*, Princeton university press, 2009.

[3] M. Lorenzi, N. Ayache, and X. Pennec, "Schilds ladder for the parallel transport of deformations in time series of images," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 2011, pp. 463–474.

[4] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain–computer interface classification by riemannian geometry," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920–928, 2012.

[5] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on riemannian manifolds," *IEEE transactions on pattern analysis and machine intelligence*, vol. 30, no. 10, pp. 1713–1727, 2008.

[6] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Classification of covariance matrices using a riemannian-based kernel for bci applications," *Neurocomputing*, vol. 112, pp. 172–178, 2013.

[7] S. Ben-David, J. Blitzer, K. Crammer, and F. Pereira, "Analysis of representations for domain adaptation," in *Advances in neural information processing systems*, 2007, pp. 137–144.

[8] S. Jialin Pan, I. W Tsang, J. T Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2011.

[9] O. Freifeld, S. Hauberg, and M. J Black, "Model transport: Towards scalable transfer learning on manifolds," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1378–1385.

[10] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu, "Transfer learning: a riemannian geometry framework with applications to brain-computer interfaces," *IEEE Transactions on Biomedical Engineering*, 2017.

[11] O. Yair, M. Ben-Chen, and R. Talmon, "Parallel transport on the cone manifold of spd matrices for domain adaptation," *arXiv preprint arXiv:1807.10479*, 2018.

[12] E. Niedermeyer and FH L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*, Lippincott Williams & Wilkins, 2005.

[13] D. Eytan, E. W Pang, S. M Doesburg, V. Nenadovic, B. Gavrilovic, P. Laussen, and A. Guerguerian, "Bedside functional brain imaging in critically-ill children using high-density eeg source modeling and multi-modal sensory stimulation," *NeuroImage: Clinical*, vol. 12, pp. 198–211, 2016.

[14] A. KC Lee, E. Larson, R. K Maddox, and B. G Shinn-Cunningham, "Using neuroimaging to understand the cortical mechanisms of auditory selective attention," *Hearing research*, vol. 307, pp. 111–120, 2014.

[15] M. M Murray, C. M Michel, R. Grave De Peralta, S. Ortigue, D. Brunet, S. Gonzalez Andino, and A. Schnider, "Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging," *Neuroimage*, vol. 21, no. 1, pp. 125–135, 2004.

[16] A. H Harrison and J. F Connolly, "Finding a way in: a review and practical evaluation of fmri and eeg for detection and assessment in disorders of consciousness," *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 8, pp. 1403–1419, 2013.

[17] C. M Michel, M. M Murray, G. Lantz, S. Gonzalez, L. Spinelli, and R. Grave de Peralta, "Eeg source imaging," *Clinical neurophysiology*, vol. 115, no. 10, pp. 2195–2222, 2004.