NeuroImage 111 (2015) 167–178

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Learning a common dictionary for subject-transfer decoding with resting calibration

Hiroshi Morioka [a,b], Atsunori Kanemura [c], Jun-ichiro Hirayama [a], Manabu Shikauchi [a], Takeshi Ogawa [a], Shigeyuki Ikeda [a], Motoaki Kawanabe [a], Shin Ishii [a,b,*]

[a] ATR Cognitive Mechanisms Laboratories, Kyoto 619-0288, Japan
[b] Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
[c] National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8568, Japan

ARTICLE INFO

Article history:
Accepted 8 February 2015
Available online 13 February 2015

Keywords:
Brain–machine interface (BMI)
Subject-transfer decoding
Electroencephalography (EEG)
Dictionary learning and sparse coding
Multi-subject–session analysis
Spatial attention

ABSTRACT

Brain signals measured over a series of experiments have inherent variability because of different physical and mental conditions among multiple subjects and sessions. Such variability complicates the analysis of data from multiple subjects and sessions in a consistent way, and degrades the performance of subject-transfer decoding in a brain–machine interface (BMI). To accommodate the variability in brain signals, we propose 1) a method for extracting spatial bases (or a dictionary) shared by multiple subjects, by employing a signal-processing technique of dictionary learning modified to compensate for variations between subjects and sessions, and 2) an approach to subject-transfer decoding that uses the resting-state activity of a previously unseen target subject as calibration data for compensating for variations, eliminating the need for a standard calibration based on task sessions. Applying our methodology to a dataset of electroencephalography (EEG) recordings during a selective visual–spatial attention task from multiple subjects and sessions, where the variability compensation was essential for reducing the redundancy of the dictionary, we found that the extracted common brain activities were reasonable in the light of neuroscience knowledge. The applicability to subject-transfer decoding was confirmed by improved performance over existing decoding methods. These results suggest that analyzing multisubject brain activities on common bases by the proposed method enables information sharing across subjects with low-burden resting calibration, and is effective for practical use of BMI in variable environments.

Introduction

One major issue in neural decoding for brain–machine interfaces (BMI) (Dornhege et al., 2007; Graimann et al., 2011; Tan and Anton, 2010; Wolpaw and Wolpaw, 2012), as well as neuroscience studies (Haxby et al., 2001; Haynes and Rees, 2006; Horikawa et al., 2013; Kamitani and Tong, 2005; Shibata et al., 2011) is how to deal with undesired variability among different subjects or different recording sessions from a single subject. For instance, electroencephalography (EEG), a widely-used neuroimaging modality in real environments, often suffers from physical and mental drifts. Physical variations include misalignment of sensors (electrodes) over days or recording sessions, different shapes of the head or skull across subjects, and changes in sensor impedance over time. Even when subjects perform exactly the same task, brain activity patterns also vary substantially across subjects (Garrett et al., 2011; McIntosh et al., 2014). Brain signals vary even in the same subject because of different physical and mental conditions, or disturbance by task-irrelevant brain activity. Such unavoidable variability is an obstacle to highly successful applications of BMI in daily life, and to neuroscientific group or longitudinal analyses using large-scale databases.

Several researchers (Devlaminck et al., 2011; Fazli et al., 2009; Kang and Choi, 2014; Lotte and Guan, 2010; Samek et al., 2014) tackled the subject-to-subject variability of EEG to achieve *subject-transfer decoding*, the goal of which is to classify the mental state of a previously unseen *target subject* based on the data or pre-trained decoders of other *training subjects*, so that BMI is instantly usable. There are two main approaches. First, the subject-invariant approach (Fazli et al., 2009; Samek et al., 2014) builds a universal decoder that is constructed only with data from training subjects, ignoring the specificity of target subjects. It thus reconciles to suboptimality if the target subject is dissimilar to any of the training subjects. Second, the task-calibration approach (Devlaminck et al., 2011; Kang and Choi, 2014; Lotte and Guan, 2010) conducts experiments with the target subject to obtain a task-based calibration dataset, which is used for tuning the decoder. This approach

* Corresponding author at: 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan. Fax: +81 75 753 4907.
E-mail address: ishii@i.kyoto-u.ac.jp (S. Ishii).

http://dx.doi.org/10.1016/j.neuroimage.2015.02.015
1053-8119/

can accommodate subject-specific variation, but acquiring the task-based calibration data is often too costly in practice, especially in daily-life applications of BMI. A third approach is thus needed, which explicitly considers the variability of target subjects but is applicable based on small efforts by them.

To resolve these difficulties, we develop a novel framework for analyzing multisubject EEG data using the unsupervised signal processing technique of *dictionary learning* after compensating for variations between subjects and sessions, and design low-cost calibration for subject-transfer decoding using resting-state data. The proposed framework decomposes multichannel EEG data into a subject- (and session-) invariant dictionary of spatial pattern bases, subject- (and session-) specific linear transforms to adjust the dictionary to each subject (and session), and sparse codes. The subject–session-specific transforms are newly introduced to the dictionary learning framework to modulate the dictionary to allow compensating for variability across subjects and sessions. Dictionary learning is useful in its own right, e.g., for clearer visualizations of spatial patterns to disambiguate their neurophysiological interpretations (Barthélemy et al., 2013; Chevallier et al., 2014), and recently used for neural decoding (Hammer et al., 2011; Zhou et al., 2012). Kang and Choi (2014) has recently proposed the idea of using a latent subspace shared across subjects and tuning it with subject-specific transforms based on a Bayesian probabilistic model. We then develop a novel subject/session-transfer scheme, which uses resting-state brain activities as calibration data. The resting-state data are easy to collect, and the proposed scheme does not require expensive task-based calibration, which would be beneficial for subjects to easily use BMI. The underlying idea is that resting-state brain activity reflects the subject-specific nature of brain activity that is consistent over subsequent task sessions. Recent studies have revealed that brain activity during resting states exhibits quite organized and stable patterns of functional connectivity, such as the default mode and the dorsal attention networks (Brookes et al., 2011; de Pasquale et al., 2012; Fox et al., 2005), and is likely intrinsic to individual brains (Baldassarre et al., 2012; Massar et al., 2014; Mueller et al., 2013; Wu et al., 2014). We thus make use of resting-state data to extract subject-specific characteristics that are intrinsic and specific to individual brains, and supposed to vary more between different brains than in the same brain between different sessions. To our knowledge, the present study is the first to apply resting-state EEG data to BMI applications.

The contribution of this study is threefold.

1. We develop a new dictionary learning technique for extracting common spatial bases while compensating for variability across subjects/sessions.
2. We propose the use of resting-state data for calibration in subject-transfer decoding, which is made possible with the proposed dictionary learning technique.
3. By using real EEG recordings from more than forty subjects performing a selective visual–spatial attention task (Morioka et al., 2014), we validate the proposed dictionary learning technique, estimating interpretable spatial patterns that are consistent with existing neuroscience knowledge, and also show that the proposed subject-transfer decoding framework performs better than existing decoding methods.

Method and material

Three core assumptions

The proposed method is based on the following three core assumptions:

A1) At each time point, brain activities as a spatial pattern can be expressed as a combination of a small number of spatial bases common across subjects and sessions.

A2) Actual signals measuring the brain activities are deformed by subject–session-specific spatial transforms.

A3) For the same subject, spatial transforms are consistent over task sessions and preceding resting-state sessions.

Fig. 1 depicts the outline of the proposed method built based on these assumptions.

Requirements for data

To accomplish our goal, we require that a dataset satisfies the following three properties, which correspond to the three assumptions above, respectively:

R1) All data samples share common underlying activities; that is, all the subjects perform the same mental task such as selective spatial attention or sensorimotor rhythm modulation.

R2) Yet, the data generation process may vary over subjects and sessions due to, e.g., variations in the brain structure, differences in channel positions, changes of the conductance of skin and gel, and slight differences of brain regions activated by the same task between subjects and even sessions of the same subject.

R3) Recordings of resting-state activities of the target subjects just before performing BMI task sessions are available.

Basic dictionary learning

Dictionary learning is a data analysis method that estimates overcomplete bases for sparsely representing measurable signals. It has its origin in neuroscience under the name of sparse coding (Olshausen and Field, 1997), which is still a current research topic (Hunt et al., 2013), and has been applied to signal processing (Elad, 2010; Mallat, 2008; Patel and Chellappa, 2013) for denoising, compression, and so on. Sparse representation has also been shown to improve classification performance in pattern recognition (Gao et al., 2010; Mairal et al., 2008; Zhang and Li, 2010). Instead of using predefined bases like discrete cosine transform (DCT) bases or wavelets, dictionary learning adaptively constructs a set of bases, or a *dictionary*, from the given data with sparseness constraints, so that the dictionary is best suited for representing the data at hand. Principal component analysis (PCA) also estimates orthogonal bases from the given data; however, the sparse method is more flexible as its overcomplete bases can cover some dynamic characteristics that may be possessed by many real-world signals.

We satisfy the first assumption A1) by dictionary learning. According to the basic formulation of dictionary learning, a vector of measured signals[1] $\mathbf{x}_t \in \mathbb{R}^M$ at time t is represented by

$$\mathbf{x}_t \approx \mathbf{D}\alpha_t, \tag{1}$$

where $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_K] \in \mathbb{R}^{M \times K}$ is a dictionary matrix whose column vectors $\mathbf{d}_k$ are called atoms and $\boldsymbol{\alpha}_t \in \mathbb{R}^K$ is called a sparse code. This equation can be seen as a conversion from a signal $\mathbf{x}_t$ to a sparse code $\boldsymbol{\alpha}_t$. If $\mathbf{D}$ were fixed at a set of DCT bases, then $\boldsymbol{\alpha}$ would be a frequency domain variable. Dictionary learning estimates $\mathbf{D}$ adaptively based on the given data. In general, we take $K > M$ (more bases than the signal dimensionality), that is, we use an "overcomplete" dictionary $\mathbf{D}$. The

[1] The vector $\mathbf{x}$ may be direct measurements, or signals preprocessed depending on the characteristics of the data. We will revisit the preprocessing in the Data-dependent considerations: Experimental procedure and feature extraction section.