Pattern Recognition Letters 68 (2015) 139–145

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Importance-weighted covariance estimation for robust common spatial pattern ☆

Alessandro Balzi [a], Florian Yger [b,*], Masashi Sugiyama [b]

[a] Department of Electronics, Information and Bioengineering, Politecnico di Milano, via Ponzio 34/5, Milan MI 20133, Italy
[b] Department of Complexity Science and Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 5 January 2015
Available online 25 September 2015

Keywords:
68T10
92C55Biomedical imaging and signal processing
Pattern recognition
Speech recognition

ABSTRACT

Non-stationarity is an important issue for practical applications of machine learning methods. This issue particularly affects Brain–Computer Interfaces (BCI) and tends to make their use difficult. In this paper, we show a practical way to make Common Spatial Pattern (CSP), a classical feature extraction that is particularly useful in BCI, robust to non-stationarity. To do so, we did not modify the CSP method itself, but rather make the covariance estimation (used as input by every CSP variant) more robust to non-stationarity. Those robust estimators are derived using a classical importance-weighting scenario. Finally, we highlight the behavior of our robust framework on a toy dataset and show gains of accuracy on a real-life BCI dataset.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Brain–Computer Interface (BCI) systems [7] have the goal to translate brain signals into a control signal that can be processed using for instance machine learning methods. Two main techniques are suitable for recording the activity of the brain: invasive and non-invasive. In this paper we will focus on a non-invasive technique called electroencephalography that had gained popularity due to its fine temporal resolution, portability and inexpensive equipment.

In recent years, approaches to BCI based on motor imagery have been developed rapidly. In this paradigm, subjects are asked to imagine the movements of their limbs or muscles. Different areas of the brain show an alteration in the regular activity according to the imagined movement performed, and this activity can be measured through an electroencephalogram (EEG). The main motivation is to establish a novel communication channel for healthy and disabled people to interact with the environment. In fact, the information of the mental state of a subject can be used for controlling a computer application or a robotic device such as a wheelchair.

A very challenging task, when dealing with a BCI system, is having a reliable representation of the brain signals. To achieve this purpose, a feature extraction method is necessary and, among all, Common Spatial Pattern (CSP) [8,14] is certainly the most popular. The idea behind CSP is to compute the most suitable spatial filters to discriminate between different types of EEG signals in a BCI protocol based on changes in oscillations (e.g. motor imagery, steady state visually evoked potential, etc.). Practically, it reduces the volume conduction effect -i.e. the spatial spread of information after the electrical signals go through the skull and skin- on the filtered signal.

However, CSP has been proven not to be robust to non-stationarity and outliers [19,20], mainly due to the difficulty of having a proper estimations of the class covariance matrices. The sources of this problem are the presence of artifacts (eye blinks, loose electrodes, etc.) that corrupt the acquired brain signals and the intrinsic non-stationarity of EEG signals. In particular between-session non-stationarity is often observed in BCI experiments and cause performance to be far from being optimal. Some of the issues inherent to BCI have been tackled through CSP variants, for example, [14] and [20].

Several approaches have been proposed to cope with non-stationarity and, among them, we can distinguish three strategies. The first strategy consists in preprocessing the signal in order to extract a stationary subspace as in [5] and [17]. Another way is to modify the CSP algorithm itself to make it robust to non-stationarity as in [9] and [21]. Finally, other approaches make the classifier robust as in [22,23] and [11].

In this paper, inspired by the approach proposed in [23], beyond making the classifier robust, we make the feature extraction step robust using the Importance-Weighting technique. More specifically, we propose a new version of the CSP algorithm based on robust estimation of the covariance matrix to cope with a type of non-stationarity called the covariate shift. This estimator is close to the one proposed in [28] but the normalization included in our estimator makes it applicable to other tasks. Moreover, in [28], the authors

☆ This paper has been recommended for acceptance by Prof. A. Heyden.

* Corresponding author. Tel.: +81 3 5841 4106.
E-mail address: florian@ms.k.u-tokyo.ac.jp, florian.yger@gmail.com (F. Yger).

http://dx.doi.org/10.1016/j.patrec.2015.09.003
0167-8655/© 2015 Elsevier B.V. All rights reserved.

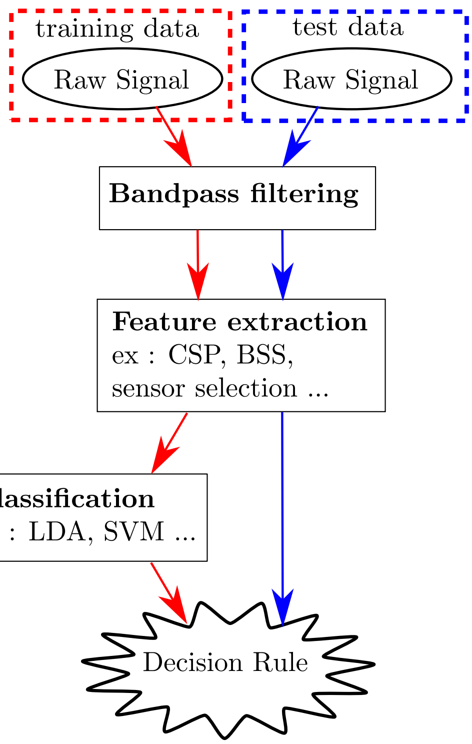

Fig. 1. Flowchart of the standard BCI framework.

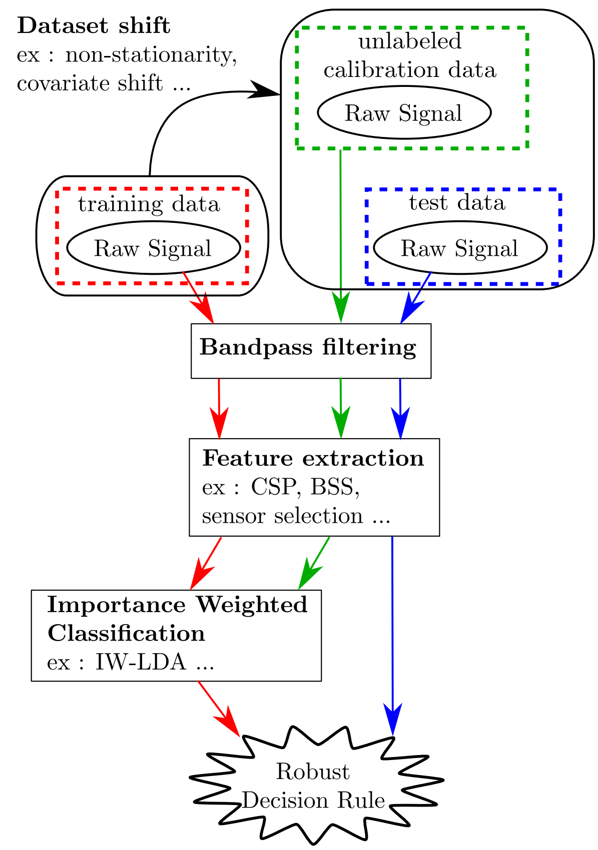

Fig. 2. Flowchart of the IWLDA framework in [23].

added to their estimator a supplementary layer of importance weighting at the epoch level. As the problem of robustness at epoch level is out of the scope of this paper, our estimator does not include such a double level of importance weighting. Indeed, we focus on the problem of robustness at sample level and leave the difficult problem of robustness at epoch level for future study.

This paper is organized as follows. First, Section 2 gives an introduction to CSP and goes through the stages needed for building a robust CSP method. Then, in Section 3 the derived CSP methods are compared to the standard method on both synthetic and real-life datasets. Finally, we conclude this work with a summary and some possible outlooks.

2. Methods

In this section, we first put the spatial filtering into perspective with the global BCI framework and then we derive our Importance-Weighted approach.

2.1. General framework

In this paper, we consider a BCI system as a pattern recognition system [8] and we try to distinguish between different motor imagery tasks by applying machine learning techniques. Over the years several different approaches have been proposed for this purpose, all of them following the standard framework summarized in Fig. 1. Three main phases can be highlighted:

Filtering: a bandpass filter is applied to both training and test raw input signals. The experiments performed in [6] showed that the filtering phase has a strong impact on the performance of feature extraction. In motor imagery applications it is common to work with frequencies on the range of 8–30 Hz [15].

Feature Extraction: attempts to create features containing important information that are used as input by a classifier. This represents a very crucial part of the framework and it is the one that this paper aims at analyzing and improving. A popular algorithm for feature extraction in BCI application is CSP [8].

Classification: a classification algorithm is used to separate different motor imagery tasks. In [13] the authors gave experimental evidence of the fact that in BCI applications linear algorithms often outperform more complex ones, hence in this paper a standard Linear Discriminant Analysis (LDA) [8] is used.

This standard framework can be slightly modified to deal with dataset shifts [16] (such as non-stationarity, covariate shift, etc.). Fig. 2 shows the approach presented in [23], in which the concept of importance is applied to the classification step obtaining covariate shift adaptation by importance-weighted LDA (IWLDA). The goal of this paper is to exploit the same concept of importance weighting but applying it in the feature extraction phase in order to improve the robustness of the model from a precedent stage. The method, called importance-weighted CSP (IWCSP), is described in Fig. 3[1]. Finally, later in the paper the two methods (IWCSP and IWLDA) are combined together and the results are discussed.

2.2. Common Spatial Pattern (CSP)

The CSP method was first introduced by Fukunaga in 1990 [8]. Nowadays it is probably the most popular algorithm for spatial filtering in motor imagery experiments. The main idea is to use a linear transformation to project the multi-channel EEG data into a low-dimensional subspace with a projection matrix. The aimed transformation maximizes the variance of signals of one class and at the same time minimize the variance of signals of the other class [4,7,14].

[1] When dealing with non-stationarity, we need to recalibrate the algorithm (either the classifier or the feature extraction) and as depicted in Figs. 2 and 3, three datasets (training, calibration and test) are needed in practice.

Download English Version:

https://daneshyari.com/en/article/536227

Download Persian Version:

https://daneshyari.com/article/536227

Daneshyari.com