ISBN: 978-602-5649-417

Stationary Subspace Analysis as a pre-processed Adaptive Common Spatial Pattern Data Method

Ahmad syafiq kamil[1], Febrilliyan samopa[2]

[1] Information System / Faculty of Engineering ,Universitas Nahdlatul Ulama Surabaya, East Java
ahmadsyafiqkamil@unusa.ac.id

[2] Information System / Faculty Of Information And Communication Technology, ITS Surabaya, East Java
iyan@is.its.ac.id,

Abstract

Brain activity is described by brain waves. Retrieving brain wave data can use a special recording device. The results of data taken on the record tool sometimes have bad noise. Bad noise is caused by interference received by the subject during data retrieval. The Common Spatial Pattern (CSP) method is a method for separating features from brain waves. The Common Spatial Pattern method has experienced very rapid development. Adaptive Common Spatial Pattern (ACSP) is a development method of CSP. CSP development becomes ACSP due to CSP's inability to handle data with many subjects. Previous ACSP research carried out the application of ACSP using raw data. The application of the raw data is to determine the accuracy of ACSP. In this study, we will use the stationary subspace analysis (SSA) method as a data preprocessing method. The results of this study indicate that the SSA method can increase the results of 1% accuracy.

Keywords: ***Brain Computer Interface*, *Common spatial pattern, adaptive common spatial pattern, stationary subspace analysis.***

Introduction

Brain computer interface is a technology that communicates the human brain with a system using special tools such as electrochepalogram (EEG)(Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). The last few years the development of the brain computer interface system has developed rapidly. Many BCI developments in the fields of medicine, neuroergonomic, smart environment and education and games. brain computer interfaces can be used to translate signals originating from the brain into control signals without using muscles (Wang, Gao, & Gao, 2005a).

The Common Spatial Pattern (CSP) is a method that is widely used to extract features. Ramoser introduced the common spatial pattern method for detecting hand movements (Ramoser, Müller-Gerking, & Pfurtscheller, 2000). In 1998 Johannes examined the spatial filter which was optimal for a single classification Electroencephalogram (EEG) experiment. The data used records 3 types of movements, namely the right and left movements of the index finger and the movement of the right foot. The problem in this study is the calculation of covariance matrix so that this study uses sample covariance as an estimator. This study uses the Common Spatial Pattern as a method in Feature Extraction (Müller-Gerking, Pfurtscheller, & Flyvbjerg, 1999).

Von Bunau's research demonstrates that Stationary Subspace Analysis (SSA) can be applied to BCI data. SSA can dramatically improve classification. Von Bunau concludes that SSA can improve classification and stationary and nonstationary scalp maps can be described and allow neurophysiological interpretation, due to the linearity of SSA (von Bunau et al., 2010).

In 2015 Song conducted research using the ACSP method. This method is used to examine EEG data that does not have labels from research subjects to study spatial filters. This method can be used to classify EEG data from single or multiple objects. The method developed by Song was evaluated using EEG data in multi-subject imagery motorcycles originating from BCI Competition III and IV (Song and Yoon, 2015).

The main limitation in the ACSP method proposed by Song is that this performance is influenced by EEG artifacts such as blinking of the eyes, swallowing saliva, or other activities which will later affect the EEG results used. Artifacts in training experiments in testing trials may be attenuated, where the artifacts can cause quite poor results so that it will likely result in unreliable measurements of the measurement of similarity in the ACSP method. for example, if the results of the covariance calculation dominate the Mullback-Leibler Distance

(KLD), then the similarity calculation results calculated using KLD will be unreliable (Song and Yoon, 2015). To overcome the case as above, it is necessary to know more than a few methods of preprocessing data so that the data is processed according to the needs of the research. EEG artifacts that do not need to be removed so that the classification performance and accuracy of the algorithm runs optimally.

The research aims to use SSA as a data preprocess method. This method is used to remove signal artifacts. The results data from the SSA calculation are used as input data for ACSP methods and are classified so that the accuracy of ACSP can be calculated.

Research methods

The research will be conducted in accordance with the methodology formulated as shown in Figure 1.

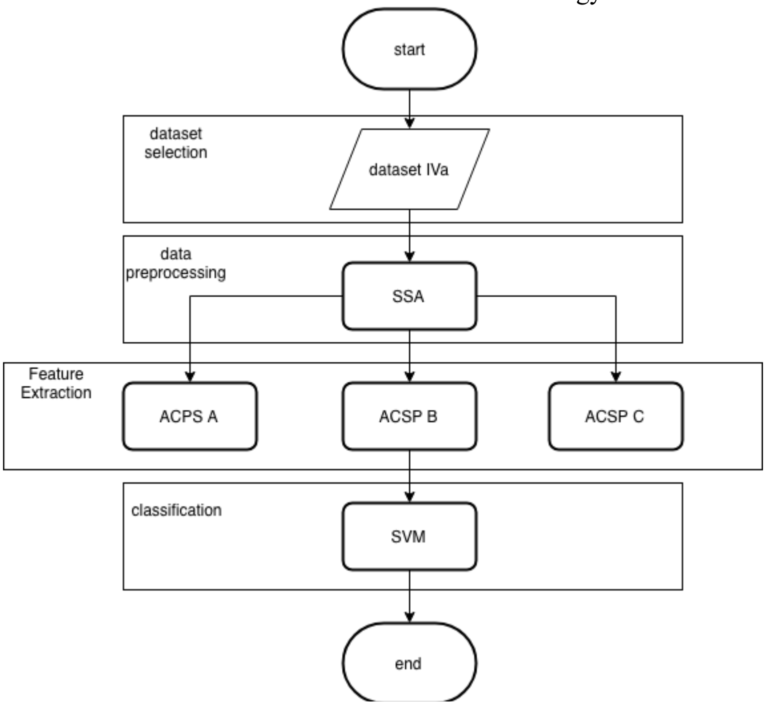

Figure 1. Research Methodology

a. Dataset selection

This research begins with searching datasets that are in accordance with previous research. In the research song used IVa dataset IVa dataset has the composition needed in this study (Song & Yoon, 2015). The IVa dataset has a description, namely the IVa dataset is a multi-subject dataset and the data has not undergone a data processing process. IVa dataset consists of 3 imagery motor classes namely left hand, right hand and right foot. This dataset is taken from 5 subjects, namely aa, al, av, aw and vv. The data of the five subjects was recorded with a predetermined scheme. The data is recorded using BrainAmp and uses an Ag / AgCi electrode cap with 128 channels from ECI. 188 EEG channels are placed in accordance with the expanded 10/20 system. The recording signal in the filter uses bandpass between 0.05 hz and 200 hz and then digitized at 1000 hz with 16 bit accuracy (0.1 uV). This dataset has a continuous signal of 118 EEG channels and a marker that shows the time at 280 signals on each of the 5 subjects (aa, al, av, aw, ay). In some markets there is a target class that has no information (NaN value). Table 1 shows the number of training trials labeled (#tr) and test trials that have no label (#te) for each subject

Table 1. Number of trial in the dataset IVa

Subjek	#tr	#te
Aa	168	112
Al	224	56
Av	84	196
Aw	56	225
Ay	28	252

b. Data preprocessing

This study uses the Stationary subspace analysis (SSA) method. The SSA method is a method that reports a high-dimensional multivariate time series into static and non-static components (von Bunau, Meinecke, Scholler, & Muller, 2010). This method assumes a signal source consists of static signals and non-static signals. In equation 1 shows that signal x (t) is a signal consisting of stationary sources $S^s(t)$ and nonstationary sources $S^n(t)$. SSA will separate the static and non-static sources so that they can be examined. Figure 2 shows an example of the results of preprocessed data. The image uses the IVa dataset on the subject aa. the image shows in line 1.

.

$$x(t) = As(t) = [A^s \ \ A^n]\begin{bmatrix} S^s(t) \\ S^n(t) \end{bmatrix} \qquad (1)$$

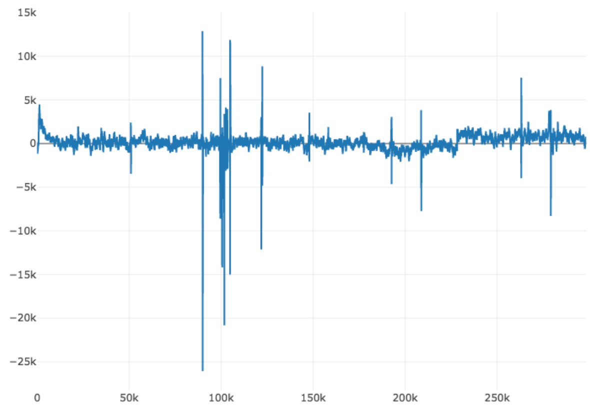

Figure 2 SSA Results

```
[Psaa{i}, Pnaa{i}, As, An, al_results] = ssa(aa, 117, 'reps', 1,
'equal_epochs', 61);
```

Figure 3 Source Code of SSA

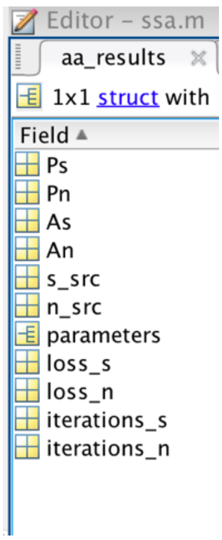

Figure 4 SSA Result Component

In this study each subject will be treated equally. Figure 3 has the source code of SSA with the attributes of the source code. Data generated from the source code is a data with static dimensions 117. The data is processed with 1 repetition. The results of this SSA will only be taken Ps and Ps as a result of calculating the toolbox. Ps and Pn are static and non-static projections from the test results. Static and non-static sources are obtained by multiplying the projection of static and non-static sources to signal sources, namely the dataset IVa.

c. Feature extraction

This study uses the ACSP method as a method to extract features from the IVa dataset. ACSP is an improved CSP method so that it can be used to analyze data with multi subjects. The multi-subject calculation uses the variance-based (FV), kullback-leibler distance (KLD), and FN norms. all three methods are used because they can measure the similarity of data. Implementation of ACSP is carried out by the procedure as follows:

- Stage 1: training signal data is calculated using the CSP method.
- Stage 2: enter test data from the target subject.
- Stage 3: calculate the projection of the data test feature using each method to measure the data similarity so that a new covariance matrix is obtained.
- Stage 4: estimate $\emptyset_1$ and $\emptyset_2$ with each method.
- Stage 5: estimate $\overline{C_1}$ and $\overline{C_2}$ and update the training data
- return to stage 2 for the next experiment.

d. classification

The result of feature extraction processed by the ACSP method will be an x_r feature that is relevant to this research. The next stage after all these features are ready, all these features will be processed using the SVM linear classification method. The linear SVM will be tested again when a new feature is updated by the ACSP for each target. new trial (Song & Yoon, 2015). The SVM method is chosen as a classifier method because according to Abdulkader SVM it is known that it can generate properties that exist in the data (Abdulkader, Atia, & Mostafa, 2015).

Research Results and Discussion

a. Plain ACSP

Table 2 shows the results of ACSP calculations without using SSA as a data preprocess. In aa subject, the average classification result is 67.9%. Subjects experienced an average increase in classification accuracy by 23%. Subject al got the average classification result of 75.6%. Subject al experienced an increase in the accuracy of the classification results by 51%. Av subject gets an average classification result of 78.9%. Subject al experienced an increase in the accuracy of the classification results by 39%. Aw subjects get an average classification result of 87.1%. Subject al experienced an increase in the accuracy of the classification results by 45%. The subject of ay gets an average classification result of 91.1%. Subject al experienced an increase in accuracy of classification results by 36%.

Table 2 ACSP without SSA (in percentage)

	CSP	ACSP A	ACSP B	ACSP C	Average
Aa	48,8	75	71,4	76,2	67,9
Al	42,4	87,5	86,6	85,7	75,6
Av	53,6	85,7	88,1	88,1	78,9
Aw	53,6	98,2	98,2	98,2	87,1
Ay	64,3	100,0	100,0	100,0	91,1

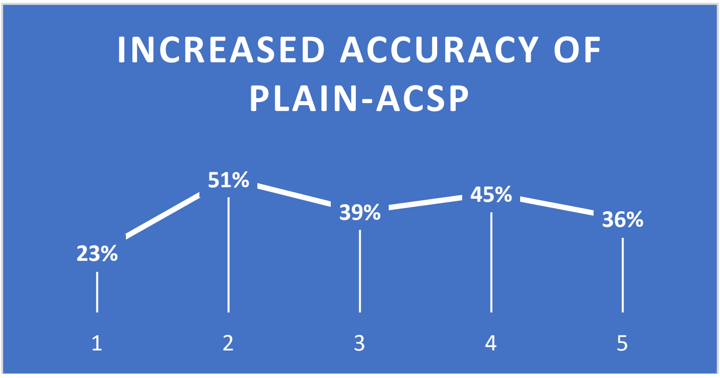

Figure 5 increased accuracy of plain-ACSP

b. ACSP-SSA

Table 3 shows the results of ACSP calculations without using SSA as a data preprocess. In aa subject, the average classification result is 70.5%. Subjects experienced an average increase in the accuracy of classification results by 35%. Subject al got the average classification result of 76.7%. Subject al experienced an increase in the accuracy of the classification results by 32%. Subject av gets an average classification result of 87.1%. Subject al experienced an increase in the accuracy of the classification results by 32%. Aw subjects get an average classification result of 87.1%. Subject al experienced an increase in the accuracy of the classification results by 45%. The subject of ay gets an average classification result of 90.8%. Subject al experienced an increase in accuracy of classification results by 37%.

Table 3 ACSP With SSA (in percentage)

	CSP	ACSP A	ACSP B	ACSP C	average
Aa	58,9	73,8	75,6	73,8	70,5
Al	56,7	87,5	87,9	86,6	79,7
Av	56,0	82,1	84,5	82,1	76,2
Aw	53,6	98,2	98,2	98,2	87,1
Ay	63,3	100,0	100,0	100,0	90,8

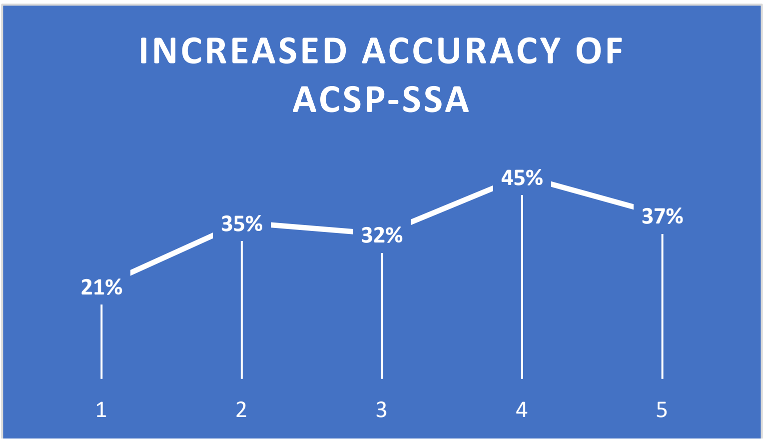

Figure 6 increased accuracy of ACSP-SSA

Figure 7 shows a graph of the comparison of the results of Plain-ACSP and ACSP-SSA. The figure shows an increase in the accuracy results shown by the ACSP-SSA at an average of 1%. These results indicate that ACSP-SSA can increase the accuracy of Plain-ACSP.

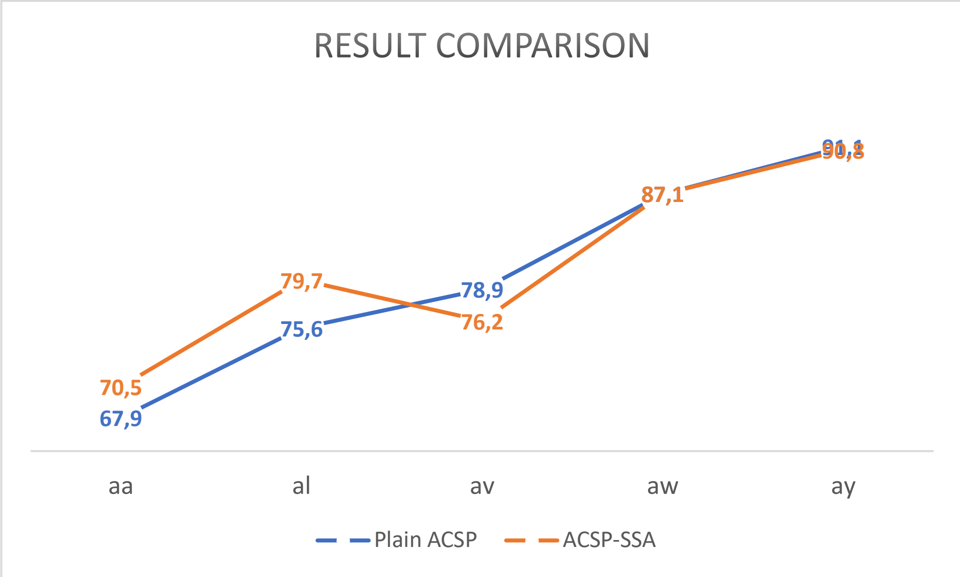

Figure 7 Results of Comparison between Plain-ACSP and ACSP-SSA

With a signal consisting of static and non-static sources, it will produce a different level of accuracy. The results of Plain ACSP use raw signals where the signal still has a source of static and non-static signals. The results obtained indicate that SSA can increase the accuracy of ACSP. Accuracy improvement because data used in ACSP calculations only has static signals. The SSA method has separated the signal data into a source of static and non-static signals. According to Prado, the time series in the EEG is always assumed to be static where the signal does not depend on when we start the observation (Prado, 1998). Which means that in the time series it will be seen the same in time intervals. In the SSA method the signal used is a static signal as a result of the separation of the sources of static and non-static signals so that the data preprocess results do not see the time coefficient. The use of static signals affects the results of the accuracy of the SSA method because there is a potential for static signal sources that are considered false (von Bunau et al., 2010). This research only uses statiss sources because the toolbox from von Bunau's research optimizes static sources rather than non-static sources.

Conclusions and recommendations

With SSA's ability to separate signal artifacts in the form of static and non-static signals, the results of the accuracy of the research that has been carried out can increase. The results of this accuracy increase because ACSP only uses static data obtained by SSA calculations. The accuracy of the results of this study increased only 1% so that the results of this study cannot be said to be significant. Research related to SSA must be improved so that it can improve the results of the accuracy of SSA.

References

Abdulkader, S. N., Atia, A., & Mostafa, M.-S. M. 2015. Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal*, *16*(2), 213–230. http://dx.doi.org/10.1016/j.eij.2015.06.002

Müller-Gerking, J., Pfurtscheller, G., & Flyvbjerg, H. 1999. Designing optimal spatial filters for single-trial {EEG} classification in a movement task. *Clinical Neurophysiology*, *110*(5), 787–798. http://dx.doi.org/10.1016/S1388-2457(98)00038-8

Prado, R. 1998. *Latent structure in non-stationary time series*. Duke University Durham, NC.

Ramoser, H., Müller-Gerking, J., & Pfurtscheller, G. 2000. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *8*(4), 441–446.

Song, X., & Yoon, S.-C. 2015. Improving brain-computer interface classification using adaptive common spatial patterns. *Computers in Biology and Medicine*, *61*, 150–160. https://doi.org/10.1016/j.compbiomed.2015.03.023

von Bunau, P., Meinecke, F. C., Scholler, S., & Muller, K.-R. 2010. Finding stationary brain sources in EEG data. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2010*, 2810–2813. https://doi.org/10.1109/IEMBS.2010.5626537

Wang, Y., Gao, S., & Gao, X. 2005a. Common Spatial Pattern Method for Channel Selelction in Motor Imagery Based Brain-computer Interface. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 5392–5395). https://doi.org/10.1109/IEMBS.2005.1615701

Wang, Y., Gao, S., & Gao, X. 2005b. Common Spatial Pattern Method for Channel Selelction in Motor Imagery Based Brain-computer Interface. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 5392–5395). https://doi.org/10.1109/IEMBS.2005.1615701

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *113*(6), 767–791.