

Automated Data Cleaning for the Muse EEG

Arnaud Delorme

CerCo CNRS, Paul Sabatier
University, Toulouse, France
SCCN, INC, UCSD, La Jolla
CA, USA

<https://orcid.org/0000-0002-0799-3557>

Jeffery A. Martin

Center for the Study of
Non-Symbolic Consciousness,
Newport, KY, USA
Lab. for Consciousness Science,
CIHS, Carlsbad, CA, USA
Transformative Technology Lab,
Stanford, CA, USA
School of Medicine, Stanford
University, Stanford, CA, USA

<https://orcid.org/0000-0003-4158-9571>

Abstract—Wearable EEG headsets have transformed the landscape of EEG research. It is no longer necessary to use expensive equipment and over 30 minutes of preparation time to collect EEG data. Instead, participants may do it themselves in a few minutes from the comfort of their home. Confronted with processing thousands of such recordings, manual data cleaning has become a bottleneck, so we tested automated methods for cleaning data. To validate these methods, we asked three trained EEG human raters to clean 100 files of 12 minutes each, and use these manual rejections to assess the performance of automated cleaning methods for both channel rejections and continuous data rejections. We showed that rejecting channels based on abnormal spectrum yielded the best results. We also showed that the Artifact Subspace Reconstruction rejection method was the best method to reject continuous portions of data. Inter-rater consistency is the gold standard to assess the quality of automated data rejection methods, and we showed that our best rejection methods were not significantly different and might even outperform human raters. We provide a simple recipe and plugin for the popular EEGLAB software for automated data cleaning of Muse data. We hope this new tool will allow more widespread use of wearable EEG in clinical and research settings where large quantities of wearable EEG data need to be processed.

Keywords— Muse, EEG, data cleaning, artifact subspace reconstruction, inter-rater consistency

I. INTRODUCTION

Wearable EEG, ranging in a couple of hundred dollars have transformed the landscape of EEG research. The data quality of such devices has proven of sufficient quality to collect EEG in various conditions, including continuous EEG [1] and event-related potentials [2]. Some of these wearable systems, like the Muse, use active electrodes similar to the one available on research-grade high-density EEG systems [3]. While the number of channels of wearable EEG systems remains limited (4 channels for the Muse), because of volume conduction, most EEG channels record activity from the whole brain. Therefore, for some use cases, a few channels might be sufficient to draw conclusions regarding the frequency content of the EEG and by inference some cognitive mechanisms. While wearable EEG will not replace the use of professional-grade EEG systems anytime soon, it allows

collecting data repetitively on a large quantity of participants, in our case several thousand recordings. This feat is not possible with professional-grade EEG systems; because of long preparation time and expenses associated with collecting data, most experimental studies include less than 32 participants.

Thousands of recordings bring new challenges to process data. With a few dozen recordings, it is usually possible to manually reject artifacts. However, this is no longer possible with thousands of recordings. Methods have been developed to automatically reject artifacts [4,5,11,12,13]. However, first these methods are not compared to human inter-rater reliability so it is difficult to assess their actual performance compared to humans. Second, they are designed for high density high quality EEG recording: they are not properly validated for wearable EEG where the number of channels is reduced and the data quality is usually lower than research-grade systems, especially when participants are asked to fit the wearable headset by themselves.

Here, we validate automated data rejection methods applied to wearable EEG, based on the manual rejection of three EEG human raters. Inter-rater consistency remains the gold standard for assessing automated rejection methods, so we compared how various rejection methods matched manual rejections.

II. METHODS

A. Data collection

123 participants self-enrolled in a 4-month long meditation course, for which they were asked to collect their EEG brainwave on a daily basis with the Muse headset [9]. Participants signed consent forms and the study was approved by the Center for the Study of Non-Symbolic Consciousness Ethical Review Board.

Data collection was performed with the 4-channel Muse 1 headset at 220 Hz sampling rate and a custom smartphone app developed by NeuroMore Inc. saved the raw time series (this private app was designed for the meditation course and is not available for public download). Participants went through a