clean_rawdata suite

Makoto Miyakoshi

Automated artifact rejection in The 30th EEGLAB workshop

June 18, 2021 10:50-11:05 pm
Did you know that?

• A historical fact is that the offline version of artifact subspace reconstruction (ASR) implemented in `clean_rawdata()` plugin (…) was specifically developed for this project upon our request by the main developer of BCILAB (Kothe and Makeig 2013). The original solution was called Christian-Nima Combo after the developers, but formally changed into `clean_rawdata()` on June 26, 2013 to be implemented as an plugin for EEGLAB…

Miyakoshi et al. (2020)
Christian-Nima combo available outside SCCN?

Makoto Miyakoshi
Dear Christian, How can we use clean_artifact() outside SCCN? Is it already distributed as a part of BCILAB package? Thank you.

Christian Kothe
Hi Makoto, it's not yet in BCILAB, but will be on the next release. We should also think about making an EEGLAB plugin. (If one of us finds the development time).

Makoto Miyakoshi
Dear Christian, Is it just taking several files from your library and write a wrapper with GUI? I'd be happy to do that for Sandy.

Christian Kothe
Yes, that's what it boils down to. Thanks for considering it!

Makoto Miyakoshi
Why don't you tell me the minimum necessary files to run it? Do you also agree to distributing as a EEGLAB plugin?

Christian Kothe
Yes, I do. Which files you need depends on what cleaning workflow you would like to offer. The function clean_artifacts is basically the core function that call

Makoto Miyakoshi
OK, I'll wait for your return. Thank you.

Christian Kothe
I think you can already start wrapping it -- I believe that the function interface is not actually different in the new version.

Lea, Sandra
Thanks to all of you (especially Makoto) for your help with this. It's very much appreciated! Sandy 2013/6/6 Christian Kothe <christiankothe@gmail.com> <mailto:christiankothe@gmail.com>

Makoto Miyakoshi
Dear Christian, I'm thinking about writing a wrapper for Sandy. I wonder which version of the script should I include. Could you show me the path to the file?

Christian Kothe
I believe the version of the script is 2013/06/04.
Christian Kotho

Hi Makoto. I actually don't know where it is (Nima wrote it and I haven't yet tested it myself) -- please ask him. Best, Christian

Christian Kotho

Oh, wait a second -- I was assuming that you'd like to wrap the combo method (from the email title). To get it coded up, please start with the files in /data/co

Makoto Miyakoshi <mmiyakoshi@ucsd.edu>

to Christian

Dear Christian,

Thank you. To respect your work, I also would like you to name the plugin clean_artifacts() sounds just too generic. I mean, for most of ERP researchers cleaning artifact means epoch rejection. To avoid confusion, I would suggest something like clean_rawdata(). Please tell me what you think.

Makoto

***

[Message clipped] View entire message

Christian Kotho

Good question ... maybe Artifact Removal Tools? You might also give Scott the option to weigh in (since he has a good sense of naming). By the way, I could imag

Makoto Miyakoshi

All right, I'll give it a temporarily name for working. We can change it later. EEGLAB provides detailed backward compatibility support. You may want to ask Arno

Christian Kotho <christiankotho@gmail.com>

to me

Sure, go ahead. If you need any low-level routines, they are in code/misc next to dependencies.

Christian

***
Flow of the processes
1. High-pass filter

Kaiser-window with $\beta=7.8573$

Attenuation: -80dB

Specify transition band edges, not cutoff frequencies!
2. Channel rejection

• If a continuous flat signal is detected.
• If high-freq noise is detected.

\[
\text{noisiness} = \frac{\text{MAD}(\text{high-pass filtered data})}{\text{MAD}(\text{low-pass filtered data})}
\]

[MAD: median absolute deviation]

\[
z\text{noise} = \frac{(\text{noisiness} - \text{median(noisiness)})}{\sigma}
\]

\[
\sigma = \text{MAD}(\text{noisiness}, 1) \times 1.4826
\]

\[
n \text{oise\_mask} = z\text{noise} > \text{noise\_threshold}
\]

• If correlation with the surrounding channels is poor. If elec locations are available, it uses RANSAC. Slow!
3. ASR

- Can handle a non-stationary large artifact
  - An ideal preprocessing for ICA!
- Standard deviation of 20-30 is suggested.
- Riemannian option may have an issue. Needs more validation.

Removing bad data period is Artifact Subspace Non-reconstruction.
4. Window rejection

- Not power but sqrt power (i.e. abs amplitude). -inf means even if a signal is too quiet, do not reject it.
- 25 means if more than 25% of channels are bad, reject that window.
5. Before-after comparison

- Remove channel drift (data not already high-pass filtered)
  - Linear filter (FIR) transition band [lo hi] in Hz: 0.25 0.75

- Remove bad channels
  - Remove channel if it is flat for more than (seconds): 5
  - Max acceptable high-frequency noise std dev: 4
  - Min acceptable correlation with nearby chan [0-1]: 0.8

- Perform Artifact Subspace Reconstruction bad burst correction
  - Max acceptable 0.5 second window std dev: 20
  - Use Riemannian distance metric (not Euclidean) - beta

- Additional removal of bad data periods
  - Acceptable [min max] channel power range (+/- std dev): -Inf 7
  - Maximum out-of-bound channels (%): 25

- Pop up scrolling data window with rejected data highlighted

Keyboard shortcuts: [r] read data, [o] old data, [b] both data, [d] difference, [+ ] increase amp scale, [- ] decrease amp scale, [*] shrink time scale, [/] expand time scale, [h] show/hide slider.
Tips for performance 1

**asr_calibrate()** line 120  \(2^{21} \rightarrow 2^{20}\)

```matlab
if ~exist('maxmoms', 'var')
    maxmoms = hlp_memfree/(2^21); end
blocksize = max(blocksize,ceil((C*C*S*8*3*2)/(maxmoms*(2^21))));
```

**clean_artifacts()** line 180  64 -> 4096 (as much as your RAM allows)

```matlab
hlp_vararginstruct varargin,
    ('chancorr_crit','ChannelCorrelationCriterion','ChannelCriterion'), 0.8, ...
    ('line_crit','LineNoiseCriterion'), 4, ...
    ('burst_crit','BurstCriterion'), 5, ...
    ('window_crit','WindowCriterion'), 0.25, ...
    ('highpass_band','Highpass'), [0.25 0.75], ...
    ('channel_crit_maxbad_time','ChannelCriterionMaxBadTime'), 0.5, ...
    ('burst_crit_refmaxbadchns','BurstCriterionRefMaxBadChns'), 0.075, ...
    ('burst_crit_reftolerances','BurstCriterionRefTolerances'), [-inf 5.5], ...
    ('distance2','Distance'), 'euclidian', ...
    ('window_crit_tolerances','WindowCriterionTolerances'),[-inf 7], ...
    ('burst_rejection','BurstRejection'), 'off', ...
    ('nolocs5_channel_crit','NoLocsChannelCriterion'), 0.45, ...
    ('nolocs_channel_crit_excluded','NoLocsChannelCriterionExcluded'), 0.1, ...
    ('max_mom','MaxMom'), 4, ...
    ('flatline_crit','FlatlineCriterion'), 5);
```
Fine tuning supported!

- There are as many as 31 tuning points!
To avoid trouble

• The input data must be full-ranked and double precision!

• To check if your data is full-ranked, run the following check:
  
  \[ \text{sum(eig(cov(double(EEG.data'))))} > 1E-7 \] == EEG.nbchan
General problem: both ASR and ICA increases gamma power

- Raw - ASR
- Raw – (ASR+ICA rej1)
- Raw – (ASR+ICA rej2)

https://sccn.ucsd.edu/wiki/Makoto%27s_preprocessing_pipeline#Why_does_IC_rejection_increase_gamma_power.2C_or_why_is_an_IC_not_broadband-independent.3F_.28For_160.2C000_page_views.2C_05.2F10.2F2021_added.29
Conclusion

• Originally requested for a tic study in UCLA.

• `clean_rawdata()` is a suite of solutions including high-pass filter, channel rejection, ASR, and window rejection.

• ASR with SD==20 is a recommended value for a standard laboratory data.

• Input data must be in double precision and full-ranked.