Towards an EEG Search Engine

Nima Bigdely-Shamlo, Ken Kreutz-Delgado, Christian Kothe and Scott Makeig

Swartz Center for Computational Neuroscience,
University of California San Diego,
La Jolla, CA, U.S.A

Abstract—**The current EEG analysis→publication workflow mostly documents qualitative descriptions of event-related EEG dynamics. This makes it difficult to look for comparable results in the literature since search options are limited to textual descriptions and/or similar-appearing results depicted in the paper figures. We demonstrate a method for quantitative comparison of source-resolved results (e.g., ERPs, ERSPs) across different EEG studies. Our proposed *source-resolved EEG measure search engine* receives search queries composed of event-related EEG measures, each associated with an estimated brain source location to be compared using Measure Projection Analysis (MPA) to all records in the search engine database accumulated by automated data analysis workflows applied to data of multiple studies. A similarity-ranked list of events from other studies that have elicited similar EEG dynamics in nearby source-locations is then returned to the user along with their experiment and event metadata. We performed a search query using this method through 52 event-related ERSP measures from four studies for Left Precentral Gyrus source measures similar to an input ERSP pattern associated with RSVP target detection, obtaining results consistent with the EEG literature and giving information about the interpretation of the phenomenon of interest not possible to infer from the single input study.**

I. Introduction

The current EEG analysis→publication workflow is mostly based on qualitative descriptions of EEG dynamic events (e.g., "a 3-8 µV positivity at 350 ms or theta burst near 800 ms over frontocentral scalp"). Such descriptions make it difficult to look for comparable results in the literature. For example, if a researcher finds some event-related feature in their data, s/he must perform a tedious text search through an ever-increasing number of published papers for comparable text descriptions and/or similar-appearing measures depicted in the paper figures. It is not yet possible to quantitatively compare source-resolved results from different EEG studies.

Also, no single study is ever large enough to account for all the variability originating from subject group differences (age, gender, etc.), and each experimental event-type explores only a subset of associated circumstances. For example, every study of EEG dynamics involving subject 'error' awareness includes at most a few types of 'error' conditions. Quantitative comparison of EEG dynamics associated with a much larger collection of 'error'-recognition conditions across many EEG studies could give much useful information for scientific interpretation of new single-study results.

Here we propose a method for building a *Source-resolved EEG Measure Search* engine built on an EEG measure database that would receive search queries composed of event-related EEG measures each associated with a brain source location estimate, and optionally a machine-readable description of the time-locked events (e.g. HED tags [1, 2]) and/or, an anatomical brain region of interest (ROI). This query information would then be compared to all the same-measure-type records in the search engine database accumulated from previous automated data analysis workflows applied to other study data and potentially other search queries. A similarity-ranked list of the most related-measure and source-location pairs (events that have elicited similar EEG dynamics in similar source-locations) from other studies will then be returned to the user along with their metadata (including pointers to any relevant publications). Fig. 1 shows a flowchart of the proposed search operation.

II. Methods

The first step of the search operation is to transform the query data into the (time/space/frequency/…) coordinate systems of the database. For instance, we use here the spatial coordinate system of the MNI (Montreal Neurological Institute [3]) template brain. Next, a mapping that assigns a similarity index (or equivalently a distance measure) to each database entry for some given query data is invoked.

In our implementation we achieve this by first applying the Measure Projection Analysis (MPA) method [4] to interpolate all source measures associated with each database item (subject, session or study data associated with a single event), originally assigned to discrete and sparse spatial locations, over a dense standardized 3-D grid of spatial brain-space voxels. The result is a common spatial representation of source-resolved EEG measures. MPA amounts to first convolving a 3-D spatial Gaussian kernel (whose standard deviation represents expected uncertainties in localization error and subject head model error) with the entered source locations to produce a source density value at each 3-D brain grid location and then calculating source density weighted average of nearby source measures at each brain voxel. Finally, the similarity of between the query EEG measures and stored measures is calculated by aggregating their similarities across all voxels in the region of interest (ROI) specified in the query.

To calculate the similarity between two search items, each associated with a set of n equivalent dipoles $D_i(y), i = 1, .., n$ located at brain

locations $y_i \in ROI \subset R^3$ and their associated measure vectors (e.g. vectorized ERSP) $\overrightarrow{M_i}, i = 1, .., n$, we first calculate the expected, or *projected*, value for the measure vector for each item at location $y_i \in ROI$ from

$$E\{\overrightarrow{M}(y)\} = \left\langle \overrightarrow{M}(y) \right\rangle = \frac{\sum_{i=1}^{n} P_i(y)\overrightarrow{M_i}}{\sum_{i=1}^{n} P_i(y)}. \quad (1)$$

Here $P_j(y) = TN(y; \hat{x}_j, \sigma^2 \cdot I, t)$ where TN is a normalized, truncated (at $t\sigma$, here t=3) Gaussian distribution with standard deviation σ and $\hat{x}_j$ is the estimated location of dipole D_j. We then compare projected measure vectors $\overrightarrow{M_i^1}$ and $\overrightarrow{M_i^2}$ of the two items at each voxel $y_k \in ROI, k = 1, .., m$ and average these values over query-specified ROI locations (m voxels) to obtain a final similarity value between the two search items with

$$Similarity = \frac{\sum_{k=1}^{m} F(Correlation(\overrightarrow{M_k^1}, \overrightarrow{M_k^2}))}{m}, \quad (2)$$

where

$$F(c) = \text{Fisher's Z}(c) = \frac{1}{2}\ln\frac{1+c}{1-c}. \quad (3)$$

Fig. 2 gives a flow chart of the measure comparison process between two conditions measures, and potentially from different studies, using Measure Projection.

III. RESULTS

We implemented the EEG measure search algorithm described above in MATLAB (Mathworks, Inc.) and applied it to grand mean (across-subjects) event-related time/frequency ERSP (event-related spectral perturbation) measures for 52 study event types within four EEG studies:

1. Target detection task during rapid serial visual presentation (RSVP) without immediate button press feedback [5].

2. A two-back Continuous Performance task with auditory feedback in which subjects have to indicate, by button press, whether the current letter presented on the screen is the same as the letter presented two letters earlier [6].

3. A visual response task in an EEG study on the domain-generality of visual self-representation. Domain-general self-representation was operationally defined as the self-relevance common to one's own Face and Cup [7].
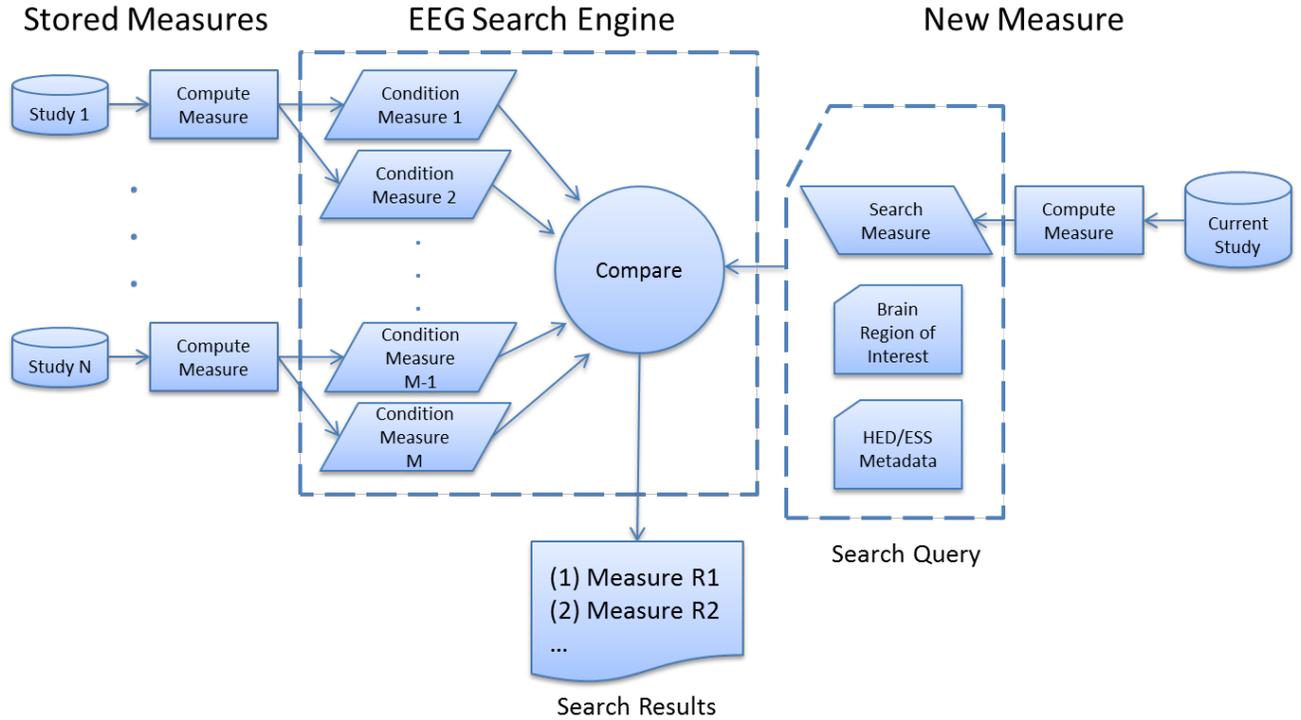
Figure 1. Flowchart of the proposed HeadIT Measure Search Engine.
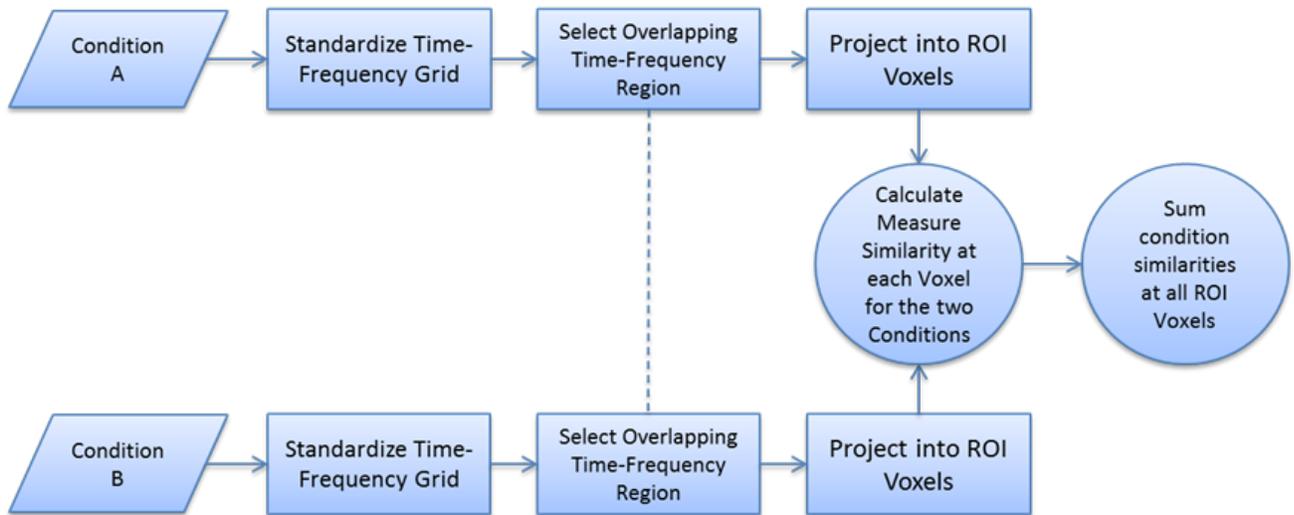
Figure 2. Flowchart of event-related measure comparison using measure projection.

4. An inter-modal audiovisual task in which subjects either switched attention upon receiving audiovisual cues, or ignored the cues and continued to focus on the interleaved audio or visual stimulus stream, responding to occasional target stimuli with button presses [8].

Input Search Condition (Target in an RSVP task):

Region of Interest

Target Condition ERSP
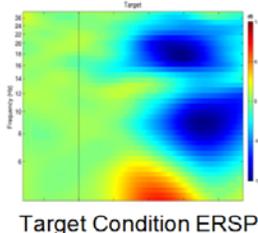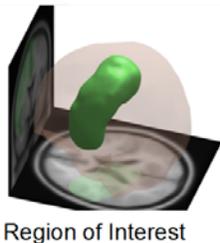
Top 3 Search Result Conditions (from the total 51):

1. Auditory Target Response during the Auditory Focus condition

2. Visual Target Response during the Visual Focus condition

3. Visual Target Response during the switch from Auditory to Vision Focus condition
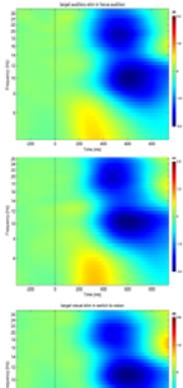
Figure 3. (Top) input query, and (bottom) top-ranked ERSP measure search results.

We used the ERSP associated with target presentation in the RSVP study as the search event and compared it to event-related source-resolved dynamics for the remaining 51 event types within a Left Precentral Gyrus region of interest. A 3-D Gaussian with 12-mm standard deviation was used in measure projection of the other 51 ERSPs. The brain volume was segmented into 3,908 voxels in a cubic grid with 8-mm voxel spacing. Fig. 3 (top) shows the ERSP pattern and ROI used as the input search query and Fig. 3 (bottom) three most-similar results based on the computed event-related measure similarities.

IV. DISCUSSION

The first three results, shown in Fig. 3 (bottom), have a quite similar ERSP pattern in Left Precentral Gyrus and are all associated with target detection, although derived from another study (Attention Shift [8]). In both the input condition and the top three matching results the participant was instructed to press a button with their right hand after detecting the target, which associated as expected from the literature with the prominent mu rhythm blocking beginning at 400 ms, in or near the right hand somatomotor area in the Left Precentral Gyrus. Identifying this factor common to the matching condition strengthens the implication that the EEG dynamics here represent dynamics relating to purposive hand movements. Given a much larger database of source-resolved study measures, further study or more formal analysis of commonalities in the matching condition descriptors might give further insights into the functional role of this cortical area and of the captured EEG dynamics.

To our knowledge this is the first demonstration of EEG source pattern search across different studies. A future direction is to create an EEG search engine with much larger database of source-resolved EEG dynamics, each tagged with a machine readable description of associated cognitive events or state, e.g. using Hierarchical Event Descriptor (HED) tags [1, 2] and subject information (age, gender…), e.g. using EEG

Study Schema (ESS) [9]. This information can be added to the search query and used to refine the results. Such a large database of source EEG patterns, accompanied with task and subject information, could be built over an existing raw EEG archive such as HeadIT [10] and have applications beyond EEG search. For example it may enable the creation of more accurate BCI models by providing neuroscience-informed priors to be used in BCI machine learning algorithms.

Another future direction is to integrate the abovementioned search engine into widely-used EEG analysis software, e.g. EEGLAB [11, 12], and to present relevant search results to researchers during their analysis so as to facilitate better interpretation of patterns present in their data.

REFERENCES

[1] *Hierarchical Event Descriptor (HED) Tags.* Available: http://sccn.ucsd.edu/wiki/HED

[2] N. Bigdely-Shamlo, K. Kreutz-Delgado, K. Robbins, M. Miyakoshi, M. Westerfield, T. Bel-Bahar, C. Kothe, J. Hsi, and S. Makeig, "Hierarchical Event Descriptor (HED) Tags for Analysis of Event-Related EEG Studies," presented at the IEEE GlobalSIP, Austin, TX (submitted), 2013.

[3] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, 1993, pp. 1813-1817 vol.3.

[4] N. Bigdely-Shamlo, T. Mullen, K. Kreutz-Delgado, and S. Makeig, "Measure projection analysis: a probabilistic approach to EEG source comparison and multi-subject inference," *NeuroImage,* vol. 72, pp. 287-303, May 15 2013.

[5] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain Activity-Based Image Classification From Rapid Serial Visual Presentation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 16, pp. 432-441, Oct 2008.

[6] T. Mullen, J. Onton, A. Delorme, and S. Makeig, "Analysis and visualization of theta-band information flow dynamics in an ERN-producing task," presented at the Human Brain Mapping, Barcelona, Spain, June 2010.

[7] M. Miyakoshi, N. Kanayama, T. Iidaka, and H. Ohira, "EEG evidence of face-specific visual self-representation," *NeuroImage,* vol. 50, pp. 1666-75, May 1 2010.

[8] R. Ceponiene, M. Westerfield, M. Torki, and J. Townsend, "Modality-specificity of sensory aging in vision and audition: evidence from event-related potentials," *Brain research,* vol. 1215, pp. 53-68, Jun 18 2008.

[9] *EEG Study Schema (ESS).* Available: http://sccn.ucsd.edu/wiki/ESS

[10] *Human Electrophysiology, Anatomic Data, and Integrated Tools (HeadIT).* Available: http://headit.org/

[11] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods,* vol. 134, pp. 9-21, Mar 15 2004.

[12] A. Delorme, T. Mullen, C. Kothe, N. Bigdely-Shamlo, Z. Akalin, A. V. Acar, and S. Makeig, "EEGLAB, MPT, NetSIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG/MEG processing," *Computational intelligence and Neuroscience,* 2011.