Efficiently searching through large tACS parameter spaces using closed-loop Bayesian optimization

Romy Lorenz a, b, *, Laura E. Simmons c, 1, Ricardo P. Monti d, 1, Joy L. Arthur c, Severin Limal e, Ilkka Laakso f, Robert Leech g, Ines R. Violante h, **

a MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, CB2 7EF, UK
b Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig, 04303, Germany
c Computational, Cognitive and Clinical Neuroscience Laboratory, Department of Medicine, Imperial College London, London, W12 0NN, UK
d Gatsby Computational Neuroscience Unit, University College London, London, W1T 4JG, UK
e Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, OX1 3PT, UK
f Department of Electrical Engineering and Automation, Aalto University, Espoo, 02150, Finland
g Centre for Neuroimaging Science, King’s College London, London, SE5 8AF, UK
h School of Psychology, Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7XH, UK

A R T I C L E I N F O

Article history:
Received 9 March 2019
Received in revised form 26 June 2019
Accepted 1 July 2019
Available online xxx

A B S T R A C T

Background: Selecting optimal stimulation parameters from numerous possibilities is a major obstacle for assessing the efficacy of non-invasive brain stimulation.
Objective: We demonstrate that Bayesian optimization can rapidly search through large parameter spaces and identify subject-level stimulation parameters in real-time.
Methods: To validate the method, Bayesian optimization was employed using participants' binary judgements about the intensity of phosphenes elicited through tACS.
Results: We demonstrate the efficiency of Bayesian optimization in identifying parameters that maximize phosphenes in a short timeframe (5 min for >190 possibilities). Our results replicate frequency-dependent effects across three montages and show phase-dependent effects of phosphenes. Computational modelling explains that these phase effects result from constructive/destructive interference of the current reaching the retinas. Simulation analyses demonstrate the method's versatility for complex response functions, even when accounting for noisy observations.
Conclusion: Alongside subjective ratings, this method can be used to optimize tACS parameters based on behavioral and neural measures and has the potential to be used for tailoring stimulation protocols to individuals.

© 2019 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

A challenge when selecting optimal parameters for studies involving transcranial alternating current stimulation (tACS) is that it rapidly results in a combinatorial explosion of experimental conditions, as physiologically plausible frequencies (0.1–100 Hz) and relative phase differences between electrodes (0°–359°) span a wide range of possibilities; thus, exhaustively testing all possibilities results in impractical experiment durations [1]. To overcome this challenge, active sampling approaches have been developed that automatically choose samples from which they progressively learn in real-time. Active sampling is useful when exploring a large space of possible experimental conditions and when the acquisition of appropriate data comes at a cost, either in terms of time, financial costs, data quality or subject comfort. In psychometrics, the potential of Bayesian active sampling approaches has long been recognized. Typically, parametric Bayesian methods are used [2–5]; however they rely on strong assumptions about the underlying objective function which are difficult to justify given limited a priori knowledge. By contrast, nonparametric Bayesian active sampling provides far more flexibility and allows
accommodating many different types of functions. Additionally, they are often faster and conceptually easier to implement [6].

To validate nonparametric Bayesian optimization in the context of non-invasive brain stimulation, we looked at phosphenes: illusory flash-like visual percepts that can be reliably induced by tACS. The current understanding is that phosphenes are generated in the retina by current spreading from the stimulation electrodes [7–11]. To date, two tACS-related stimulation parameters have been shown to affect phosphene perception: frequency and intensity. The strongest perception occurs for frequencies in the lower beta range, i.e., 14–22 Hz [12,13], and increases linearly with stimulation intensity [9,12,14]. However, the effect of relative phase has not yet been investigated. Depending on the montage employed surface electrodes can generate superimposition of currents injected at different relative phases, which the retinas might be sensitive to, providing a compelling model to test our nonparametric Bayesian optimization approach.

Thus, for this proof-of-concept, tACS frequency and phase parameters were explored conjointly. Conventionally assessing phosphene perception during tACS involves subjects judging parameters to be applied in the next iteration (Fig. 1a). In Study 1, choice, the algorithm automatically proposed a new pair of tACS parameters to the subject and were selected by the acquisition function (see Supplementary Methods C). This is problematic as it relies on judgments of absolute magnitudes; however, humans are better at making relative judgments [18–22], expressing preference for one option over another. Thus, we applied nonparametric Bayesian optimization [1,23,24] based on relative judgements [25], to search through a large tACS parameter space, with the aim of identifying frequency-phase combinations that elicit the strongest phosphene perception in individual subjects.

Methods

Empirical studies

We applied Bayesian optimization (Fig. 1a) in two studies (Study 1: N = 10, 6 females, mean age±SD: 26.3±5.59 years; Study 2: N = 10, 5 females, mean age±SD: 26.6±4.69 years). All participants reported normal or corrected-to-normal vision. Exclusion criteria were self-reported metal implants in the head or implanted electronic devices, history of neurological problems or head injury, skin sensitivity, pregnancy, use of psychoactive medication. Subjects gave written informed consent for their participation. The study conformed to the Declaration of Helsinki and ethical approval was granted through the local ethics board (NRES Committee London – West London & GTAC).

At each of the 20 iterations, subjects were exposed to two successive blocks of tACS (peak-to-peak amplitude of 1 mA) lasting 5 s each (with 1 s rest between blocks, Fig. 1b). These two blocks differed in the tACS frequency-phase-combination applied to the subject and were selected by the acquisition function (see Bayesian Optimization). Optimization was driven based on the subject’s rating of which of the two blocks evoked the strongest phosphene perception. If subjects perceived no clear difference between blocks, or did not perceive phosphenses, they were instructed to select a block at random. Based on the subject’s choice, the algorithm automatically proposed a new pair of tACS parameters to be applied in the next iteration (Fig. 1a). In Study 1, three montages were tested in separate runs (within-subject design): Cz-Oz, F4–P4 and O1–O2 (the latter two montages connected to separate returns on the shoulders, see Fig. 1c). Montage Cz-Oz was selected to validate the method against previous studies showing frequency dependency of phosphenes [12]; here phosphene perception was only optimized across frequencies. Montages O1–O2 and F4–P4 were chosen as they generate superimposition of currents that allowed the investigation of the effect of frequency and phase on phosphene perception across different electrode scalp distributions. Results of Study 1 were used to obtain more accurate priors for Study 2. In Study 2, phosphene perception was assessed for montages F4–P4 and O1–O2 at a higher level of detail by narrowing the frequency-range considered (Fig. 1d, grey space).

Bayesian optimization

Bayesian optimization is a two-stage procedure that repeats iteratively. In the data modelling stage, a probabilistic surrogate model is used to estimate the objective function. While commonly, Bayesian optimization requires scalar responses, Brochu et al. [25] have proposed an approach based on preferences, which was employed here (Brochu’s Python implementation is available on https://github.com/misterwindupbird/IBOQ). By using a Thurstone-ModeLLer model with a Gaussian process (GP), it is possible to relate binary observations to a continuous function (Supplementary Fig. 4). In the guided search stage, an acquisition function is used to propose two points in the parameter space from which to sample next (i.e., chooses the two blocks with different tACS parameters the subject will receive in the next iteration). As such, the acquisition function balances the trade-off between exploring the parameter space and exploiting the parameters for which measurements have already been collected; this allows for an efficient and reliable search over an exhaustive parameter space. Here, we applied the expected improvement (EI) acquisition function (see Supplementary Methods A).

Computational modelling

Current density on the retinas for montage F4–P4 and O1–O2 was approximated using the finite element method and a realistic head model (see Supplementary Methods B).

Simulation analyses

Simulations were carried out to demonstrate the method’s applicability to variously complex objective functions (featuring multiple local optima), as well as, how the choice of the acquisition function impacts the results (comparing the EI with the upper-confidence bound (GP-UCB) acquisition function). Further, we wanted to understand how instructing participants to give random judgements in cases when they were not able to notice any differences in phosphene intensity, may have affected our results. For each possible combination of human rater sensitivity level, acquisition function and objective function, we ran 50 simulations with the number of iterations set to 39. Identical to the empirical studies, at each iteration, the acquisition function proposes two points in the experiment space to be sampled from. Optimization was driven by the binary information about which of the two points corresponds to a larger “ground truth” function value. For details please refer to the Supplementary Methods C.
Replication of frequency-dependent phosphene perception

Group-level Bayesian models were obtained for each study separately by inferring the GP based on all available preference ratings of all subjects for a given montage. Group-level Bayesian mean model for montage Cz-Oz (Fig. 1g; for group-level variance see Supplementary Fig. 1; subject-level results are available in Supplementary Fig. 2) implicate that the strongest phosphene perception was predicted at 19 Hz, in close agreement with previous studies [12].

Table 1

<table>
<thead>
<tr>
<th>Montage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>F4-P4</td>
<td>-3.95</td>
<td>-0.02</td>
<td>0.64</td>
<td>-1.55</td>
<td>-3.75</td>
<td>-9.65</td>
<td>-33.40</td>
<td>-93.40</td>
<td>-127.88</td>
<td></td>
</tr>
<tr>
<td>Study 1</td>
<td>-50.89</td>
<td>-52.06</td>
<td>-37.36</td>
<td>-25.92</td>
<td>-20.87</td>
<td>-20.69</td>
<td>-24.00</td>
<td>-29.49</td>
<td>-31.71</td>
<td></td>
</tr>
<tr>
<td>Montage</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>Cz-Oz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Phase-dependent phosphene perception

Group-level Bayesian mean models for montages F4–P4 (Fig. 1e; for group-level variance see Supplementary Fig. 1) show that the strongest phosphene perception was predicted at 22 Hz with 30° (Study 1) and at 19 Hz with 0° (Study 2). For montage O1–O2 (Fig. 1f; for group-level variance see Supplementary Fig. 1), both studies predicted the optimum at 16 Hz, either with 30° (Study 1) or 0° (Study 2) phase difference. Subject-level results are available in Supplementary Fig. 2/3. The log marginal likelihood (lml) of group-level Bayesian models was computed as a function of varying length-scale parameter for “Phase”, while keeping the length-scale for “Frequency” fixed (for details see Supplementary Methods A). Lml automatically incorporates a trade-off between model fit and model complexity, and higher values are desirable. Directly comparing lml for varied “Phase” length-scale parameters, thus serves as a good indicator for assessing phase-dependent effects of phosphene perception. While high lml values for small to medium length-scales (i.e., 1–10) would indicate a phase effect, high lml values for large length-scale parameters (i.e., 20–50) would contradict a phase-dependent effect. Results are listed in Table 1 and indicate a clear effect of phase for both montages as medium-sized length-scale parameters resulted in higher lml values.

Computational model explains phase-dependence of phosphenes resulting from wave interference

According to computational modelling, both F4–P4 (Fig. 2a) and O1–O2 (Fig. 2b) montages yield the highest retinal current density...
when tACS is applied in phase (i.e., 0°), and the lowest current density when the stimulation is applied at opposite phase (i.e., 180°). In the 0° condition, the currents produced by the two stimulators pass through the eyes with the same phase and approximately in the same direction, producing constructive interference. In the 180° condition, the retinal current densities interfere destructively, and the currents are mostly restricted in between stimulation electrodes.

**Simulation results confirm the versatility of method**

When simulating “perfect” human rater sensitivity (i.e., 100%), the algorithm identifies the optimum for all three objective functions, illustrated by the small Euclidean Distance between predicted and true optimum. Equally, the algorithm captures the whole distribution of the functions, reflected in moderate (rho > .5) to high spatial correlations (rho > .8) between predicted and true functions. For simple (Sum Squares, Fig. 2c) and highly complex functions (Camelback, Fig. 2e), results remain robust even for low human rater sensitivity levels. This is not the case for the moderately complex function (Brain, Fig. 2d), for which the algorithm seems to settle for the local optimum instead. In this case but also more generally, we observe the explorative GP-UCB acquisition function resulting in better results than the EL acquisition function especially for low human rater sensitivity levels.

**Discussion**

We demonstrate that Bayesian optimization based on binary preference ratings provides a feasible and efficient method for searching through large tACS parameter spaces (312/192 possible combinations in Study 1/Study 2). While hypothetically testing each possible parameter combination would have taken up to 83/51 min in Study 1/Study 2, the method took 5 min per participant. We not only replicate well-known frequency-dependent effects, but we also demonstrate phase-dependent interference effects of phosphene perception. Phase dependency is consistent with the changes of the current density on the retinas, which supports the hypothesis of retinal origin of TACS-induced phosphene [7–11]. This effect is relatively subtle when compared to frequency-dependent effects; therefore, more conventional rating scales may have lacked sufficient sensitivity to assess these differences in human perception.

While we show optimization based on subjective ratings, importantly this approach can be used to find the optimal TACS parameters for behavioral (e.g., reaction time) or neural (e.g., fMRI or EEG) target measures [23,24,26–28]); in those cases it can also optimize scalar values instead of binary observations [25]. Our simulation analyses showed that the method can model variously complex and noisy objective functions. This is critical as we hypothesize that identifying the optimal TACS frequency-phase combination for effectively modulating cognitive processes might be more challenging: the relationship between TACS parameters and the subject’s neural or behavioral response could be less uniform, feature multiple optima and may be affected by a lower contrast-to-noise ratio. Interestingly, the approach can also be extended to search through multi-dimensional parameter spaces (e.g., including stimulation intensity as an additional dimension). While this approach has originally been proposed in robotics, artificial intelligence [25] and computer graphics [25], this study may pave the way to an application in the field of brain stimulation. As the method works on a per subject-basis, it has the potential to be used for tailoring stimulation protocols to individuals.

**Acknowledgements**

This work was supported by the NIH Imperial BRC (P60478), the Leverhulme Trust (P45930), EPSRC (P70597) and BBSRC (BB/S008314/1). Wellcome Trust funding for Romy Lorenz (209139/Z/17/Z) and Ines Violante (103045/Z/13/Z).

**Conflicts of interest**

The authors report no conflicts of interest.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brs.2019.07.003.

Brochu E, Cora VM, de Freitas N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. 2010 Dec 12.


