



Growing Shapes with a Generalised Model from Neural Correlates of Visual Discrimination

Pierre Cutellic^(✉)

Chair of Digital Architectonics, Institute of Technology in Architecture, Faculty of Architecture,
ETH Zurich, ETH Hönggerberg, Building HIB, Floor E 15, Stefano-Franscini-Platz 1,
8093 Zurich, Switzerland
cutellic@arch.ethz.ch

Abstract. This paper focuses on the application of visual Event-Related Potentials (ERP) in better generalisations for design and architectural modelling. It makes use of previously built techniques and trained models on EEG signals of a singular individual and observes the robustness of advanced classification models to initiate the development of presentation and classification techniques for enriched visual environments by developing an iterative and generative design process of growing shapes. The pursued interest is to observe if visual ERP as correlates of visual discrimination can hold in structurally similar, but semantically different, experiments and support the discrimination of meaningful design solutions. Following bayesian terms, we will coin this endeavour a *Design Belief* and elaborate a method to explore and exploit such features decoded from human visual cognition.

Keywords: Design computing and cognition · Brain-Computer interface · Machine learning · Generative design · Neurodesign

1 Introduction

Well known Event-Related Potentials (ERP) from neuropsychology [1] are widely studied and documented for reproducibility, and can serve the role of evaluating acquisition, pre-processing and classification methods. New applications seeking to involve known paradigms mean that new experiments need to be designed with these precedents in mind in order to compare meaningful results. In order to dissociate the question of acquiring, preprocessing and successfully decoding neural correlates of cognitive processes from their applications for CAAD purposes, one should refer to current EEG signal challenges and transferability of learned patterns across modalities [2].

Aims. While state-of-the-art research in cognitive science is actively dealing with that matter [3, 4], the present research focuses on the application of such potentials for better generalisations in future technologies for design and architecture. It engages with adapting known and generalised methods of acquisition, preprocessing, presentation, classification and exploitation from a P300 Visual Speller [5], for visual environments of increasing richness in information as commonly found in CAAD modelling interfaces.

It is known that, based on informational Bayesian models [6, 7], visual discrimination may occur in complex visual environments and their relevance for decision making rely on the degree of visual experience an individual may hold to construct prior beliefs upon which to infer [8, 9]. We will make use of previously built generalised techniques and trained models on a singular individual EEG signals and observe the robustness of advanced classification models to initiate the development of presentation and classification techniques for enriched visual environments by developing an iterative and generative design process of growing shapes. What is of interest is to observe if visual ERP as correlates of visual discrimination can hold in structurally similar, but semantically different, experiments and support the discrimination of meaningful design solutions. Following bayesian terms, we will coin this endeavour a Design Belief and elaborate a method to explore and exploit such features decoded from human visual cognition.

Significance. This research focuses on a generalisation and application of predefined Rapid Serial Visual Presentation of an Oddball Task (RSVP-OP) techniques and pre-trained classifiers to assess visual ERP as neural correlates of what we previously defined as design beliefs. Its goal is to advance research methods on related CAAD modelling applications.

2 Methods

The hereafter described methods are divided into sections concerning the necessary design of data flows from the generation of visual stimuli to the acquisition of EEG signals and their analysis to finally contribute in a generative process of shapes. While using the RSVP-OP as a basis, we will develop further on the generalisation of visual stimuli and their tokenisation, presentation for human visual cognition, and to which will be correlated acquired and processed aggregated EEG signals from a single person.

Tokenisation. The adapted visual stimuli use 3D metaballs rendered by a marching cubes algorithm [10, 11] in order to provide a generic and smooth visual flow in the continuous variation and presentation of generated shapes by the rendering of implicit functions of isosurface. Each flashing epoch, previously showing a row or a column in the reference case of the visual speller, is replaced by the uniform random position of a new metaball instance in spherical coordinates (Fig. 1).

The center of the spherical coordinates being either the origin of the rendered scene, or the center of one of the generated metaballs, if at least two already exist. In the case of none existing yet, a first instance will be placed at the origin for a second one to be generated from. Once the scene contains at least two instances of a metaball, the center point to generate a new one will be selected in a similar random fashion and produce the previously described relative coordinates for the new instance to be added for the rendering of the isosurface (Fig. 2).

As a result, each new metaball instance P is parametrised with its coordinates xyz , and two parameters of field strength St and subtract Su related to the isosurface calculations. Ideally, the radius R of the sphere to be rendered as a metaball is $R = (St/Su)^{0.5}$, such that an instance of P can be parametrised as

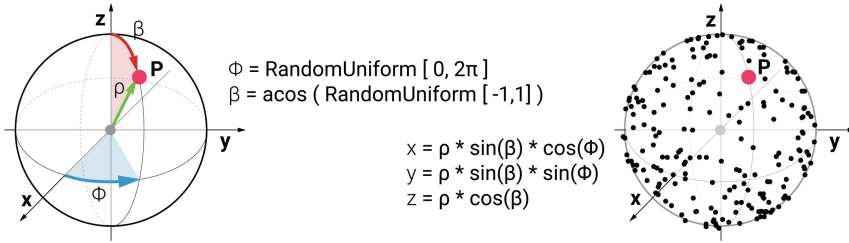


Fig. 1. Random uniform spherical distribution of solutions for new possible coordinates of a metaball instance being later presented as a new stimulus. From a given center point O (grey) of a sphere of radius ρ (green), one generates the new coordinates x, y, z , of a point P (pink) and relative to O .

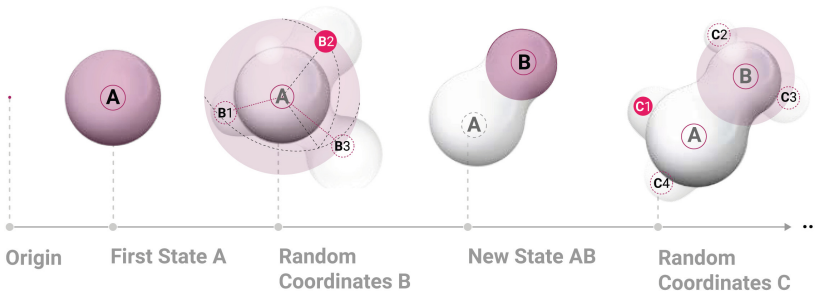


Fig. 2. Sequence for updating the rendered isosurface. From Left to Right: a first instance is placed at the *Origin* of the scene: *First State A*. New coordinates (i.e. $B1, B2, B3, \dots$, as examples of possible generations) are generated, given the center of A and its radius: *Random Coordinates B*. From $B2$, a new instance B is added to the scene: *New State AB*. Similarly, new coordinates are generated (i.e. $C1, C2, C3, C4, \dots$, as examples) from a randomly selected preexisting instance A or B : *Random Coordinates C*.

$P : (Px, Py, Pz, Pst, Psu)$ and for an entire token T constituted of nP such that $T : [(P_0x, P_0y, P_0z, P_0st, P_0su), \dots, (P_{n-1}x, P_{n-1}y, P_{n-1}z, P_{n-1}st, P_{n-1}su)]$. Eventually the final distance added to the coordinates of a new instance from the center of a previous one is equal to the radius of the later and the new resulting radius R . Each token can possibly have different distances between connected metaballs and each metaball can possibly have different radii (Fig. 3). One will consider these two configurations as two distinct classes $C1$ (same distances and radii) and $C2$ (random distances and radii).

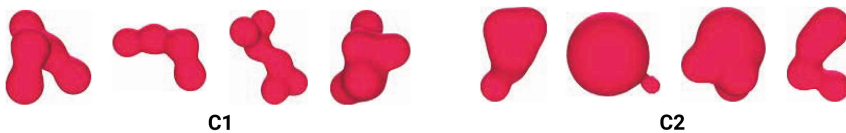


Fig. 3. Left: class $C1$ samples with same radii and distances. Right: class $C2$ samples with random radii and distances.

In addition, three main kind of shaders are applied to each tokens: $S1$ - a plain white shader with no depth or shadow, $S2$ - a Phong material shader with specularity and reflectance, $S3$ - a black and white dot-patterned shader with no depth or shadow but applied on the uv coordinates of the shape (Fig. 4). These three shaders allow for three different kinds of visual distinction of the complex geometry, depth, silhouette and curvature being rendered. They all relate to a certain kind of basic information sent to the visual system for early processing and known as information of shape from texture and motion [12, 13]. The three applied shaders will be considered as three unrelated categories $Q1$, $Q2$, $Q3$ for comparison of results, as providing different degrees of shape information.

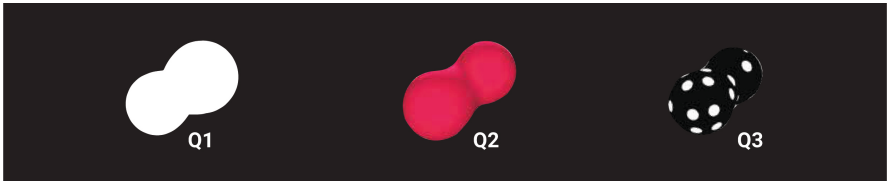


Fig. 4. Rendering of the three distinct shaders as visual categories. A black background is used during the recording sessions.

Visual Presentation. From the previous study of visual spelling with an ERP-BCI [5], the Rapid Serial Visual Presentation of the Oddball Paradigm Task (RSVP-OP) is preserved with a similar time and tokenisation structure. Each presentation contains a sequence of 12 tokens shuffled and shown 15 times so that each token would be viewed 15 times in a random order of appearance. An initial period, to ease-in the user's attention into the visual scene and show how the tokenisation will be presented, is set to 2.5 s. Similarly, a minimum of 2.5 s of a break period is set between presentation periods to avoid rapid fatigue and disengagement. Since the temporal method used for classification is *offline* learning and the next presentation period is dependent to the processing and the returned discriminated token by a pre-trained classifier (i.e. a new tokenisation can happen only if there is a new state returned), the break period is also extended until a value is returned (i.e. the index of one of the presented tokens or none in case of no discrimination found). Each token is presented on screen for a duration of 100 ms and followed by a blank screen for a duration of 75 ms while the standard refresh rate of the visual presentation is approximately 60 Hz. Each recording session has been kept under a maximum time of 18 s (excluding the break periods) and 6 discriminated tokens forming the overall shape. The main adaptation from the generalised RSVP-OP consists in augmenting its temporal structure (Fig. 5).

While the RSVP-OP occurs, data is acquired accordingly. And while the data is being processed during break periods, the current state of the shape is kept visible until a new value is returned and the new state of the shape is shown for a second before starting the new RSVP-OP and in order to generate the new tokens. Additionally, and since the complexity of visual scenes presented is more important than in the case of a word speller, the RSVP is adapted at every token flashed so that its silhouette appearing

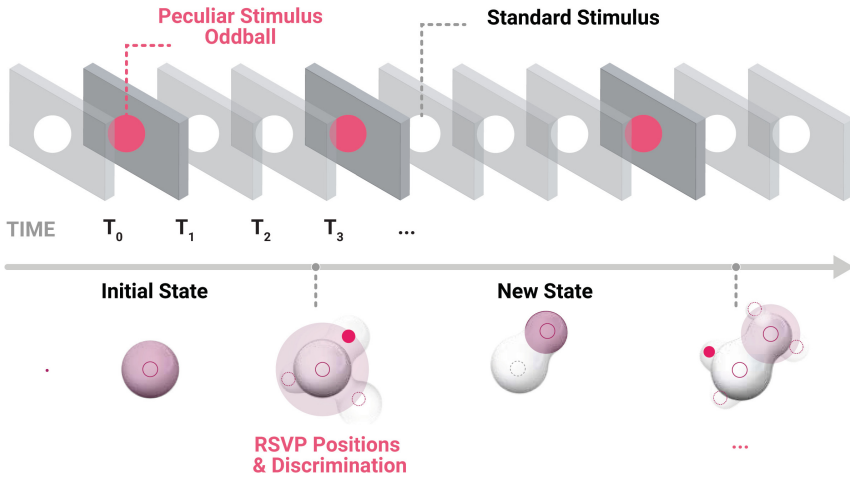


Fig. 5. Adaptation of the typical RSVP-OP task. Between stimulus presentation, another presentation period is introduced to show the new state of the overall shape formed by discriminated tokens and from which new tokens will be generated for stimulus presentation. This new period also plays the role of helping to reduce prolonged cognitive workload and await for discrimination and generation of new stimulus for the next RSVP-OP.

on screen is maximised. This effect is achieved by measuring the angle between: a - the line formed by the centroid of the shape and and the center of the presented token; b - the X -axis of the scene always horizontal and parallel to the camera X -axis. A rotation is then applied to the shape as in Fig. 6.

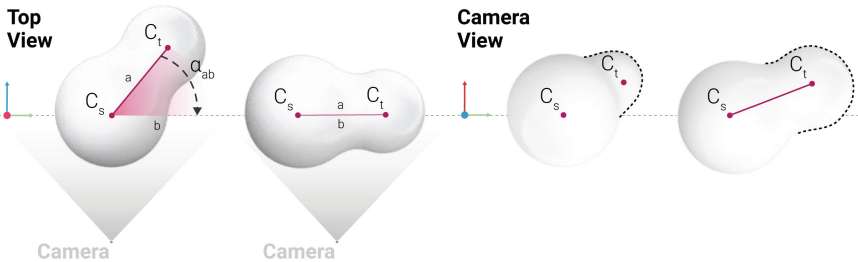


Fig. 6. Rotation applied to the shape around its X-axis during each token presentation of the RSVP-OP period in order to increase each presented token's visibility. C_s is the centroid of the shape in its current state and C_t is the centroid of the token being presented as a stimuli. Both are forming a line a from which is measured its angle ∂_{ab} formed with a vector parallel to the camera's Y-axis and passing by C_s , noted b . The rotation is applied on the whole shape (including the token) around the vector parallel to the camera's X-axis and passing by C_s , so that $\partial_{ab} = 0^\circ$.

This method provides a new view angle of the shape at each RSVP and allows for novel information of the overall shape from motion and texture [12, 13], while the visibility of the token is emphasised to ease the discrimination. Additionally, a random

rotation is constantly applied during break periods to show more information of the overall shape after and before RSVP. A re-centering and re-scaling of the camera occurs before every new RSVP to ensure that the centroid of the shape remains at the center of the scene and the whole shape is being contained and visible on the screen.

Signal Acquisition and Processing. The EEG data is acquired through a *Lab Streaming Layer protocol* [14], synchronised, and by a 16 channels *OpenBCI* [15] (i.e. *Daisy + Cyton* configuration) at a sampling frequency of 125 Hz with electrodes placement at *FC5, FC6, C3, Cz, C4, CP1, CP2, P3, Pz, P4, PO3, POz, PO4, O1, O2* and *Oz* positions of the *Modified Combinatorial Nomenclature* (MCN) of the *International 1020* placement [16]. Signals are digitised at their device’s sampling frequency and then filtered with an eight-order bandpass filter with low and high cut-off frequencies of 0.1 and 20 Hz, to finally build epochs -0.100 to 0.700 ms onset visual stimulus and downsample each signal to the high pass limit since most ERP components can be found below 20 Hz [17]. No particular artifact rejection method is applied except for amplitudes superior at $75 \mu\text{v}$ to reject outliers from muscular movements. This allows for a minimisation of data points to process within the range of ERP detection.

For similar reasons explained in previous experiments [5] concerning challenging EEG signals features for stable classification (mainly signal-to-noise ratio and non-stationarity), the capacity for a given classifier to learn across different modes (different sessions, experiments and users) without calibration is a question of research on *Transfer Learning* itself [2] and can be approached by either *Information Geometry* [3] or *Deep Learning* [4] methods. Given the low amount of data and the user-based approach of the experiment, information geometry classifiers have been chosen and trained for a single person on multiple recording sessions of a P300 word speller, so that the assumed learning across-modalities would concern only the cross-experiments mode (i.e. From spelling words to growing shapes). The pre-trained classifier is a *riemannian* classification pipeline constituted of ERP covariance matrices and projection on the tangent space [18–20] with an AUC accuracy of 97.5% after 12 training sessions. Given previously mentioned EEG features and an increasing variance in the data when applying new experiments, the robustness of such method is evaluated by observing the difference of averaged discriminated samples recorded during the new experiment (Fig. 7). Though observed on less data amounts than for training, one can see that despite changes in the morphologies of signals and presence of noise, the classification accuracy across experiments for a single user can be maintained to a certain degree, although it may not provide for a continuous and fully robust adaptive classification across all mentioned modalities.

Shape Generation. During the developed RSVP-OP sessions and parameters, two types of data are recorded into a user’s database: *a* - aggregated and processed time series used as input for the classifying pipeline, segmented by presentation period (i.e. one for each discriminated token). *b* - generated shapes directories, containing their *Q-C* shape labels (see section: *Tokenisation*), the mesh file and its associated material (Fig. 8.) computed by the programmed shaders (in **.obj* and **.mtl* formats), and a **.json* data file containing all parameters used to procedurally generate the given shape (Table 1.). The later is used for further understanding on the extent of the produced solution space

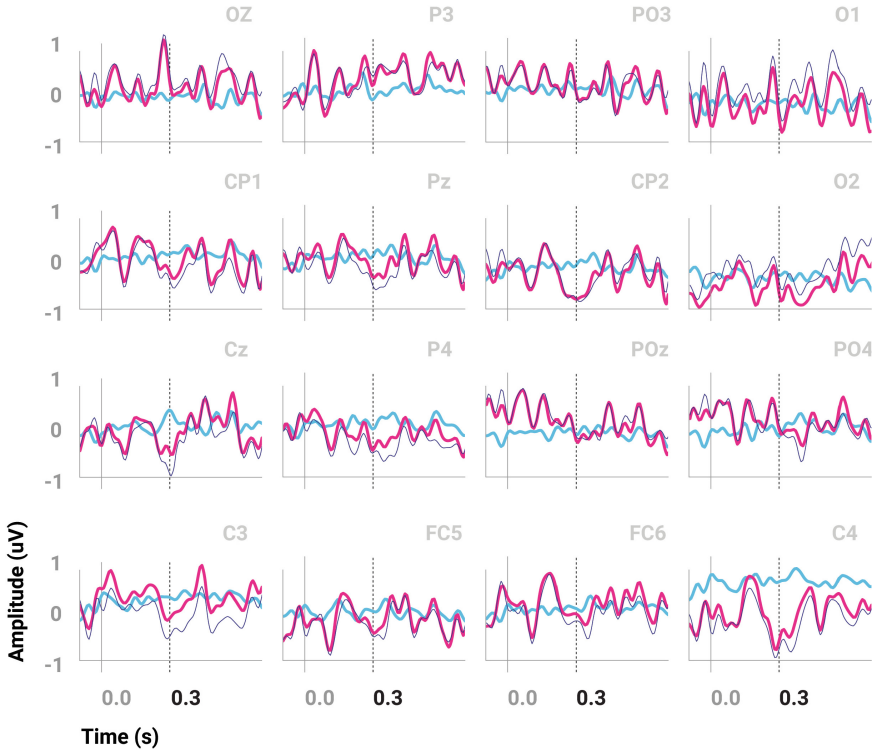


Fig. 7. ERP binary conditions of discriminated data with the shape generation experiment. Red plots show positive Target. Blue plots show negative ones. A difference waveform is plotted in dark blue.

and its features. Eventually, a similar method can be used to proceed from an inverse modelling fashion to generate such shapes given an adequate artificial generator.

3 Results

From all data files of Q - C shapes generated, a dimensionality reduction is applied from the initial 36 shape dimensions (6 instances X 6 params) to a 2d mapping using T - SNE [21] and $UMAP$ [22] to evaluate the topology of the aggregated data and account for possible manifolds. In order to observe a differentiation between possibly random shape generations and otherwise meaningful ones, they are compared with randomly generated data using similar procedural methods and parameters for both $C1$ and $C2$ classes (Fig. 9).

Both T - SNE and $UMAP$ methods show similar clusters and suggest that discriminated data correlate only in part with random data points. As some clusters appear outside the random ones in more compact topologies, they suggest a meaningful convergence for some generated shapes. Since visual ERP is clearly correlated with visual



Fig. 8. Grid of all the generated shapes by Q - C order. Original shaders have been replaced here by colors for better overall visualisation. Bottom to top, Left to Right: Q_2 (pink) with 50 C_1 shapes and 50 C_2 shapes; similarly followed by Q_3 (black) and Q_1 (grey).

Table 1. A sample *.json file containing parameters P_x , P_y , P_z , P_{st} and P_{su} for each instance (0 to 5) necessary to procedurally generate its associated shape.

	P_x	P_y	P_z	P_{st}	P_{su}
0	0.5	0.5	0.5	0.61	15.6204993518133
1	0.556925825892438	0.495865511380644	0.645404113938775	0.61	15.6204993518133
2	0.454769577708112	0.495483133279051	0.763573292424071	0.61	15.6204993518133
3	0.511977763440335	0.644353703985817	0.663595549798839	0.61	15.6204993518133

attention [1], another index of engagement is added to help in visualising the relation of engagement with discriminated shapes (Fig. 10.). The index used for this is a commonly used *Beta to (Alpha + Theta)* index [23], where the mean relative band power of *Theta* (4–8 Hz), *Alpha* (8–12 Hz) and *Beta* (12–30 Hz) frequency bands are computed for each aggregated time series of a shape. Since the index E is computed on pre-processed EEG data which has been filtered and resampled to a maximum of 20 Hz, the *Beta* band is being cut by approx. 55%, a naive factor k is applied on the mean value of the *Beta* band such that $k = 0.55$ and: $index = (\beta + k\beta)/(\alpha + \theta)$.

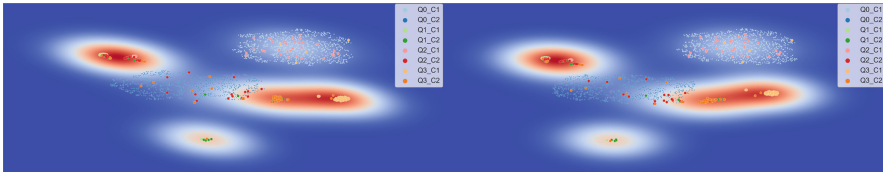


Fig. 9. T-SNE and UMAP 2D dimension reduction of the generated data. Both are run with several perplexity/proximity parameters (euclidean, 15 to 45 with step of 10) to observe the persistence of clusters topology across global/local structures. The map contains 20000 random data points and 300 discriminated ones. A gradient plot in the background shows the proximity of discriminated data points in clusters. Label $Q0$ is a random generation with $C1$ and $C2$ parameters found in $Q1$, $Q2$ and $Q3$. Sample 1–2: T-SNE (Left) and UMAP (Right) clusters with perplexity/proximity parameters 35. Similar results are found across all parameter sets.

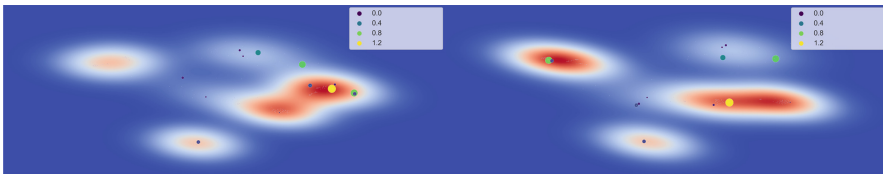


Fig. 10. Mapping of the Engagement index for each discriminated data points already projected on the 2-dimensional plane with T-SNE with perplexity parameter of 15, 25, 35 and 45. From Left to Right: sample plot with perplexity parameter of 15 and 35.

4 Conclusion

The mapping of engagement index on clustered discriminated data shows peaks of engagement both in specific clusters and random ones. It also shows that very few low peaks are present on the specific clusters. One can interpret such topology by summing that some meaningful clusters are formed but some data points outside of them might also be of interest and that such index would be helpful to adjust their meaningfulness. The robustness of generalising the acquisition and classification methods across experiments for a single user can be maintained to some extent and would greatly benefit from further adaptive research in stimulus presentation and transfer learning. We have engaged into modifying typical RSVP methods to the end of easing the rendering of complexified stimulus presentations towards design and architectural modelling purposes. Through the accumulation of generated shapes, we have shown that some meaningful clusters emerge to form what we can now call a Design Belief in the way they aggregate around regions in the latent space for certain design solutions and parameter ranges over time and based on typical informational bayesian prior beliefs. In addition, engagement indices of visual attention such as the one used in the present experiment can be purposed to value and ponder both formed design beliefs and episodic discrimination outside such regions but with high engagement index in order to notice other possible regions of interest. This should allow to further devise for a method to generate design solutions based on the discrimination of such design belief together with an exploitation/exploration ratio of the design space, in order to maintain variance over time in the

generation of design solutions. Further experiments will develop this combined discriminative/generative method together with a better granularity of ERP classifications and stimulus presentations moving from the generation of shapes to the spatial articulation of parts for architectural modelling implementations.

References

1. Kutas, M., Kiang, M., Sweeney, K.: Potentials and paradigms: event-related brain potentials and neuropsychology. In: Faust, M. (ed.) *The Handbook of the Neuropsychology of Language*, pp. 543–564. Wiley, Oxford (2012)
2. Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F.: A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *J. Neural Eng.* **15**(3), 031005 (2018)
3. Rodrigues, C., Luiz, P., Jutten, C., Congedo, M.: Riemannian procrustes analysis: transfer learning for brain-computer interfaces. *IEEE Trans. Biomed. Eng.* **66**(8), 2390–2401 (2018)
4. Tuleuov, A., Abibullaev, B.: Deep learning models for subject-independent ERP-based brain-computer interfaces. In: 9th International IEEE/EMBS on Neural Engineering, pp. 945–48 (2019)
5. Cutellic, P.: Towards encoding shape features with visual event-related potential based brain–computer interface for generative design. *IJAC* **17**(1), 88–102 (2019)
6. Bayes, T., Price, R.: *Essai en vue de résoudre un problème de la doctrine des chances*, vol. 18. Cahiers d’histoire et de philosophie des sciences. Paris (1763)
7. Pierce, J.R.: *An Introduction to Information Theory: Symbols, Signals & Noise*. 2nd, revised edn. Dover Publications, New York (1980)
8. Lindsay, P.H.: *Human Information Processing: An Introduction to Psychology*, 2nd edn. Academic Press, New York (1977)
9. Goldstein, E.: Bruce, and Thomson Learning (Firm). *Sensation and Perception*. Thomson Wadsworth, Belmont (2007)
10. Blinn, J.F.: A generalization of algebraic surface drawing. *ACM Trans. Graph.* **1**(3), 235–256 (1982)
11. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. In: *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*, New York, pp. 163–169
12. Palmer, S.E.: *Vision Science: Photons to Phenomenology* 3rd printing. MIT Press, Cambridge (2002)
13. Stone, J.V.: *Vision and Brain: How We Perceive the World*. MIT Press, Cambridge (2012)
14. LSL is developed and hosted by The Swartz Center for Computation Neuroscience at UCSD San Diego. https://github.com/scn/lsl_archived
15. OpenBCI specifications online
16. Sharbrough, F., Chatrian, G.-E., Lesser, R.P., Lüders, H., Nuwer, M., Picton, T.W.: American electroencephalographic society guidelines for standard electrode position nomenclature. *J. Clin. Neurophysiol.* **8**(2), 200 (1991)
17. Luck, S.J.: *An Introduction to the Event-Related Potential Technique*. A Bradford Book, 2nd edn. The MIT Press, Cambridge (2014)
18. Barachant, A., Congedo, M.: A Plug&Play P300 BCI Using Information Geometry. [arXiv: 1409.0107](https://arxiv.org/abs/1409.0107) [Cs, Stat]. 30 August 2014
19. Congedo, M., Barachant, A., Bhatia, R.: Riemannian geometry for EEG-based BCI; a primer and a review. *Brain-Comput. Inter.* **4**(3), 155–174 (2017)

20. Barachant, A.: Python Package for Covariance Matrices Manipulation and Biosignal Classification with Application in BCI. see Alexandrebarachant/PyRiemann. Python (2019)
21. van der Maaten, L., Hinton, G.: Visualizing data using T-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008)
22. McInnes, L., Healy, J.: UMAP: uniform manifold approximation and projection for dimension reduction. [arXiv:1802.03426](https://arxiv.org/abs/1802.03426) [Cs, Stat]. 9 February 2018
23. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* **40**(1–2), 187–195 (1995)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

