Assessing video advertising engagement via Nonlinear Intersubject Correlation Analysis of EEG and eye tracking dynamics

Kostas Georgiadis
Information Technologies Inst. (ITI)
Centre for Research & Technology Hellas
Thermi-Thessaloniki, Greece
kostas.georgiadis@iti.gr

Fotis P. Kalaganis
Information Technologies Inst. (ITI)
Centre for Research & Technology Hellas
Thermi-Thessaloniki, Greece
fkalaganis@iti.gr

Nikos A. Laskaris
AIIA-lab, Dept. of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece
laskaris@csd.auth.gr

Spiros Nikolopoulos
Information Technologies Inst. (ITI)
Centre for Research & Technology Hellas
Thermi-Thessaloniki, Greece
nikolopo@iti.gr

Ioannis Kompatsiaris
Information Technologies Inst. (ITI)
Centre for Research & Technology Hellas
Thermi-Thessaloniki, Greece
ikom@iti.gr

***Abstract*— A novel framework for assessing engagement through video commercials is presented. Deviating from the current approaches, here we cast the problem as nonlinear intersubject correlation analysis and pursue the collective (i.e. across-participants) treatment of EEG signals and eye-tracking measurements. Regarding brain dynamics, two well-known descriptors, namely spatial covariance and phase locking value, undergo transcription to their 'hyperscanning' equivalents. Regarding eye-related measurements, individual paths and pupillometric measurements are summarized at population level. Using data from a recording paradigm, where each participant independently watched a cartoon video interrupted by a commercial clip, we show that our approach can provide signatures of engagement to the content being delivered and reconstruct the level of appreciation of the potential consumers.**

Keywords*— *brain dynamics, pupillometry, hybrid BCI, EEG, eye-tracking, neuromarketing

I. Introduction

The creation of a video commercial is probably the most prominent approach adopted by advertising agencies to promote products, brands, companies etc. Given the high cost of broadcasting time, the production of appealing commercials, that will eventually reach high resonance levels, has become vital. Neuromarketing exploits existing neuroimaging modalities and techniques to reinforce existing marketing practices and provides a thorough assessment of various marketing-related elements [1], including video commercials [2].

Current Neuromarketing research relies on signal processing descriptors to characterize eye and/or brain activity signals and machine learning algorithms to build models capable of predicting consumer behavior. Most frequently some features are first extracted independently from each subject and then utilized within a supervised learning pipeline, along with labels indicating a behavior outcome (e.g. 'the product observed by subject will be purchased')[3]. Deviating from this strategy, we examine here collectively the behavior of a group of potential consumers while observing a commercial clip. We study how their gaze-directions confluence and to what extend their brainwave patterns synchronize in response to the visual input and the buying attitude towards the shown product. By implementing "quasi-hyperscanning", we quantified the induced self-organization tendencies in behavioral ocular data (fixations and pupillometry measures) as well as in ongoing brainwave activity (instantaneous within-subject and across-subjects synchronization). This term is used throughout this paper to emphasize the fact that the developed methodologies work for experimental data coming from scanning multiple individuals while experiencing the exact same stimulus, but without recording their signals simultaneously.

The proposed approach is demonstrated using data from a multi-subject dataset that contains concurrent EEG and Eye Tracking (ET) data collected from 20 subjects (6 males and 14 females, aged 42.25 ± 11.95). Single-subject data were recorded, independently, while each participant was watching a cartoon video interrupted by a 30 sec commercial-clip that referred to various supermarket products. Besides neuroimaging data, questionnaire responses that evaluated the commercial's effectiveness were also provided. Our analysis commenced with the contrast between the two viewing conditions ("TV cartoons" vs "ads") and proceeded with the identification of two sub-groups within the participants: those who positively appreciated the commercial clip and those who did not. By contrasting the data from these two groups, we show that predictors of the commercial clip success can be crafted directly from EEG or ET signals. Overall, the across-subjects analysis of ET+EEG data under the close-to-real life scenario of watching a commercial clip embedded in a movies hold great promise for predicting the potential impact of a clip.

II. Methodology

A. Experiment and Data description

The timeline of the experiment can be seen in Fig. 1b. The design of the paradigm was such as to realistically reproduce a scenario often faced via TV viewers. Specifically, a preselected cartoon movie was presented to every participant and its presentation was interrupted, at a certain point, by a commercial clip, common to all participants. That clip lasted for 30 sec and included the advertisement of 4 different products that were in offer at a particular supermarket retail chain. Apart from being close to real-life, the selected scenario offered a natural baseline to compare against (i.e. the standard watching condition). Brain activity was recorded, at 300 Hz

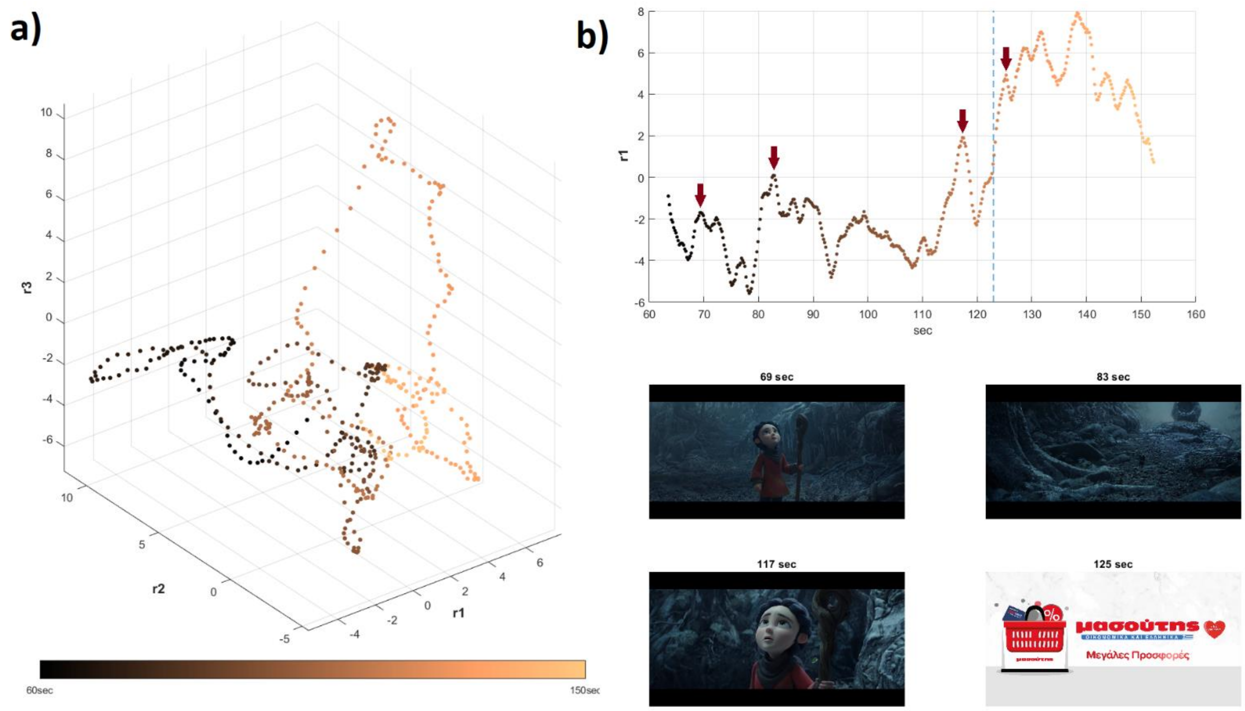

Fig. 1 a) Across-subjects, coordinated, brain dynamics during movies watching, embedded as a 3D trajectory with each dot corresponding to an "instantaneous" covariance estimate. b) The first dimension of the embedding is utilized to indicate the timeline of the recording paradigm. The vertical line indicates the start of the commercial clip and the arrows the latencies associated with the mined frames shown below.

sampling rate, using Wearable Sensing's DSI-24 with 19 dry sensors placed according to the 10-20 International System. Ocular activity was simultaneously recorded, at 120 Hz, via Tobii Pro Fusion eye-tracker that provided successive measurements for eye-gaze position and pupil-diameter. Regarding preprocessing, the raw EEG signals were first bandpass filtered within [0.5–45] Hz using Butterworth filter in zero-phase mode and then FORCe [4] software was employed to remove artifactual activity. On the other hand ET-signals were interpolated for missing values. All measurements, across modalities and subjects, share a common time axis.

B. ET-data Analysis

Information from ET comes in the form of two distinct timeseries. The successive vectors r(t)=[x(t) y(t)], where x and y denote the corresponding coordinates of the gazing position, and the associated measurements of pupil diameter p(t). By collecting the gazing position timeseries $r_i(t)$ for all N participants, we expressed the time-dependent attraction of gaze based on the local-point density estimator of potential-functions (PFs), where constant d_0 is the radius of influence

$$\widehat{PF}(t) = \frac{1}{2\pi\, d_0\, N^2} \sum_{i,j=1}^{N} \exp\left(\frac{-\left\| r_i(t) - r_j(t) \right\|^2}{2\, {d_0}^2}\right) \quad (1)$$

The resulting signal takes values close to 1 whenever the gazing positions coincide.

A second signal of collective behavior was derived similarly from $p_i(t)$ timeseries. Motivated by [5], which provided empirical evidence that the temporal derivative of pupil-diameter reliably reflects attention, we first derived $\dot{p}_i(t) = \frac{dp_i}{dt}$ for each participant, then took the absolute value and finally applied a trimmed mean estimator of 10% at every latency.

C. EEG-data analysis: Spatial Covariance representation

Spatial covariance matrices (SCMs) play an important role in characterizing brain dynamics from multichannel EEG recordings containing multi-trial responses [6],[7]. They are symmetric positive definite (SPD) matrices that lie on a Riemannian manifold instead of a vector space, and hence require special treatment. With this condition fulfilled, they provide intelligible brain signal analytics and often result in fruitful data descriptions. We adopted this perspective of Riemannian geometry in two ways. First to describe participants' evolving brain dynamics as individual trajectories and correlate them to detect intersubject covariation. Second to provide a novel, collective, description of brain dynamics in a hyperscanning set up.

Let $\mathrm{X} = [\mathrm{X}(1), \mathrm{X}(2), \dots \mathrm{X}(t) \dots] \in \mathbb{R}^{\mathrm{E} \times N_t}$ be the continuous multichannel EEG signal, where E denotes the number of electrodes. We extracted, sequentially, multiple overlapping segments $\mathrm{X}_{[i]} = [\mathrm{X}(i-w), \dots, \mathrm{X}(i-1), \mathrm{X}(i), \dots, \mathrm{X}(i+w)], w \in Z^+$ and used them to derive a sequence of SCMs $\mathrm{C_i} = \frac{1}{2w-1} \mathrm{X_i X_i^T} \in \mathcal{R}^{\mathrm{E \times E}}$, where $(.)^{\mathrm{T}}$ denotes the transpose operator. The pattern of SCM was considered as reflecting the "quasi-instantaneous" brain state. By quantifying its change over time, we attempted to track its modulations induced during movies watching. To this end, we employed the geodesic distance $d(C_i, C_j) = [\![logm(C_i^{-1/2}\, C_j\, C_i^{-1/2})]\!]_F$ and applied to every pair among the set of extracted segments $\{X_{[i]}\}_{i=1:N_s}$ so as to construct the $[N_s \times N_s]$ matrix GD. Next, this matrix was fed to classical multidimensional scaling (MDS) to derive a sequence of

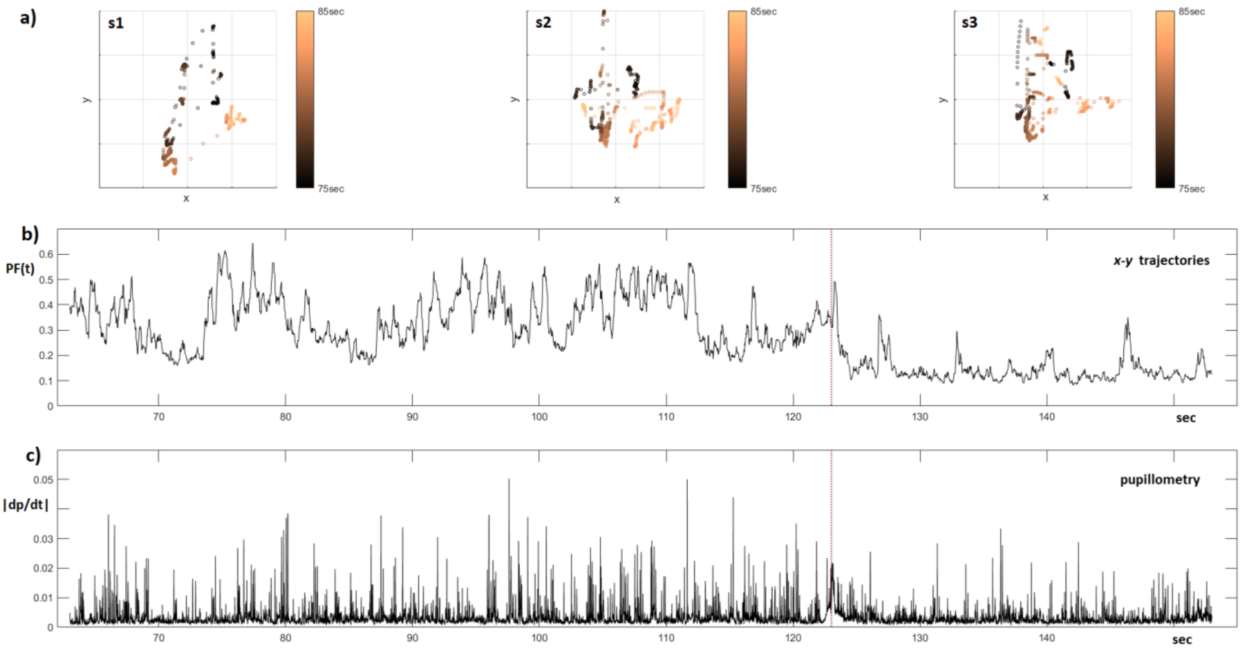

Fig. 2 a) The ET-trajectories of 3 random subjects during the [75-85] sec time interval. b) The time-dependent local point density across all subjects. c) The across-subjects estimate of temporal derivative of the pupillometry signal. Vertical lines indicate the commercial clip onset.

vectors $\{Y_{[i]}\}_{i=1:N_s}$, that constituted a low-dimensional embedding (3D) of induced modulations in brain's covariation pattern. The resulting sequence, derived independently for each participant, was plotted as a trajectory with the color indicating the time evolution (see Fig. 3a) and can be thought of as a parsimonious reconstruction of the "true" underlying brain dynamics.

Although these trajectories were aligned in time, they represented embeddings that had been carried out separately and, hence, could not been directly compared. To tackle this issue, we resorted to *distance correlation* (distcorr) [8], a nonlinear measure of dependence between distinct but paired vectorial observations that is known to overcome distribution-shift issues like scale differences or translations. We adopted the distcorr function so as to derive a time-dependent measure, by correlating (nonlinearly) time-aligned subsequences between trajectories from any pair of participants, i.e. $distcorr^{i,j}(t') = distcorr\left(\{Y^i_{[t']}\}, \{Y^j_{[t']}\}\right), i,j = 1, \dots, N$. These computations resulted in $\frac{N(N-1)}{2}$ timeseries. By computing the mean value (m) and standard deviation (std) for each t', an aggregate latency-dependent profile of dependence was finally derived that peaked at instances where SCMs were modulated more similarly (see Fig. 3b).

An additional way to incorporate SCM representation in the analysis of brain dynamics was attempted by offline simulation of "brainsourcing" [9]. We took advantage of the fact that the timeline during movies presentation was common for all subjects and pooled the signals from N=20 of them. The signal became $\boldsymbol{X}^{group} = \left[\dots X^1_{(t)}, \dots \vdots \;\; \dots X^2_{(t)}, \dots \vdots \;\; \dots X^N_{(t)}, \dots\right] \in \mathbb{R}^{N \cdot E \times N_t}$ and the procedure for deriving SCMs proceeded accordingly so as to reconstruct the trajectory of a "collective brain state" (see Fig. 1a).

D. EEG-data analysis: Phase synchrony

While SCM representation encapsulates the signal power of brain activations, it overlooks important information conveyed by the phase of brainwaves. In such a paradigm of dynamic nature like ours, we considered important to investigate this aspect as well. Treating the multiple recordings as performed simultaneously ("quasi-hyperscanning"), we investigated for possible, film-induced, alignments of the brainwaves emanating from each brain region.

The employed algorithmic procedure started by band-pass filtering the signals within the range of canonical rhythms (δ,θ,α,β1,β2,γ1) and deriving the instantaneous phases by means of Hilbert transform. Then, separately for each rhythm and sensor (denoted $c_i = 1,2 \dots, E$), the phases $\theta^j_{c_i}(t)$ from all N subjects were utilized repeatedly in pairwise cross-subjects computations of phase synchrony. Specifically, the estimator of Phase Locking Value (PLV) [10] was employed in time-resolved manner $PLV^{jl}_{c_i}(t) = \frac{1}{(2w+1)} \left|\sum_{t'=t-w}^{t+w} e^{i\,\Delta\theta(t')}\right|$ where, $\Delta\theta(t) = \theta^j_{c_i}(t) - \theta^l_{c_i}(t)$, $j, l = 1,2 \dots, N$, and w a resolution parameter that controls the temporal window. The computations resulted in a set of $\frac{N(N-1)}{2}$ timeseries, based on which the levels of intersubject synchrony were statistical compared between cartoon film and commercial clip viewing. Multiple statistical tests (Wilcoxon rank sum) were performed for the distinct brainwave rhythms and sensors and results were reported after Bonferroni adjustment.

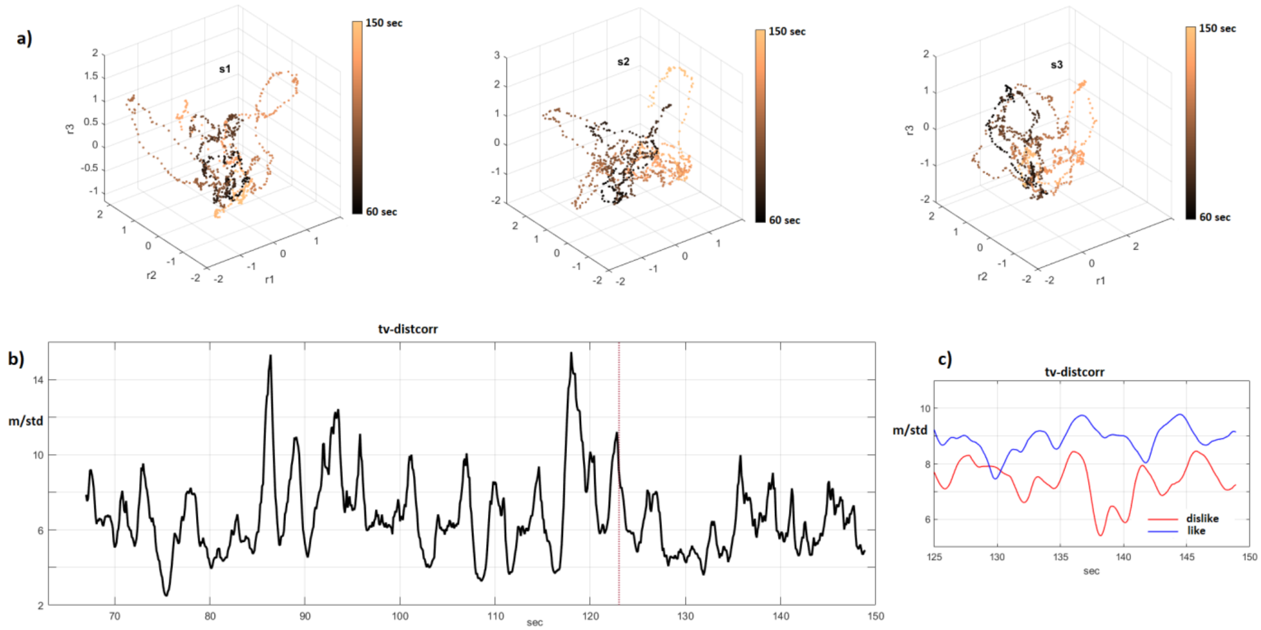

Fig. 3 a) Trajectories from random subjects formed while viewing the same movies. b) Aggregate measure of distance correlation as a function of time; peaks correspond to instances at which trajectories co-evolve across subjects and vertical line indicates the onset of commercial clip. c) The aggregate measure of distance correlation conditioned on the subjective assessment of clip.

III. Results

A. A gross picture of induced brain dynamics

We commence by providing a collective picture of brain dynamics in response to watching the film and the commercial clip in immediate succession. Using wide-band filtered signals from N=20 participants we reconstructed the group-based trajectory followed (by the across-subjects covariance) during the [60 153] sec interval. Covariance estimates were derived every 0.5 sec based on segments of 2 sec. By inspecting the corresponding Fig. 1a, and considering the distance-preserving character of the mapping, it becomes clear that the beginning of commercial clip marks an excursion of brain dynamics towards a more diverse sequence of covariance patterns. On the contrary, the group-covariance appears to alter slowly during the cartoon-film presentation, and this smooth motion indicates slowly-varying visual inputs, as can be seen in Fig. 1b, where we have mined input frames according to selected turning points of the trajectory.

B. Across-participants alignment of ET dynamics

Next, we discuss the trajectories formed by eye-gaze during movies/clip presentation and the extent at which they overlap between participants due to common visual input. Considering that ET-measurements were carried out after calibration, their time-indexed coincidence (as estimated via the measure of PFs in eq.(1)) can be considered as directly reflecting engagement. Three random trajectories are depicted in Fig. 2a. The time varying PF-index derived from N=20 participants is shown in Fig. 2b, where it becomes evident the difference in population-level dynamics between "baseline" (i.e. cartoon movies) and advertisement (i.e. commercial clip). Movies capture more, and sporadically, the participants' attention. It is noteworthy that the peaks seen during the commercial clip correspond to the different sequentially presented products.

Apart from gaze positioning that may be considered a somehow reflexive behavioral characteristic, we also present the results from pupillometric analysis which is considered as reliably reflecting subjective evaluation. The across-subjects estimate of temporal derivative of the pupillometry signal is provided in Fig. 2c. Considering that all participants reported the cartoon film as more interesting than the commercial clip, it becomes evident that the group-aggregated $\dot{p}(t)$-profile is suggestive of engagement with the visual content.

C. Inter-subject correlation as trajectory covariation

In the third step of our analysis, we investigate whether SCM-modulations reveal any dependence among subjects, that can be attributed to the way film and clip influence ongoing brain dynamics, and hence can serve for building biomarkers of engagement to the visual content. Three embeddings of SCM-modulations in the form of dynamic trajectories have been included in Fig. 3a for the case of three random participants. Initial covariance estimates were derived every 0.1 sec, based on 1sec segments from the wide-band EEG signal. The common color-code, associated with the underlying timeline, is suggestive of some analogies but not identical dynamics. We systematically pursued this trend by estimating the distance correlation between all possible pairs among N=20 participants and at the level of subsequences of 5sec in duration. The obtained results have been aggregated as shown in Fig. 3b. The depicted profile indicates that, at the group level, brain dynamics covariate more strongly during the film presentation. Going one step further, we have included in direct contrast (Fig. 3c) the corresponding correlation profiles from the participants that liked the commercial and those who did not.

D. Inter-subject phase synchrony

Finally, we present the results from analyzing inter-subject phase synchrony as a function of sensor placement and brainwave rhythm. By statistically comparing the levels

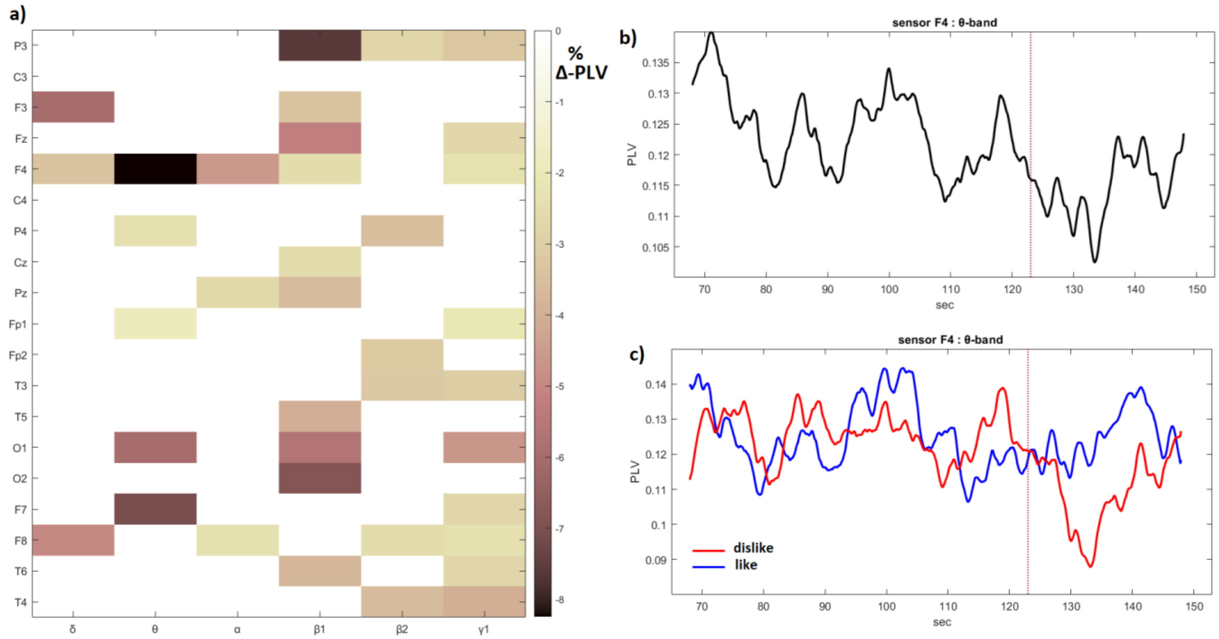

Fig. 4 a) Statistical comparison of across-subject phase-synchrony as a function of sensor and frequency band. Color coded in the relative difference in PLV-level between watching commercial-clip and movies. A mask has been applied to those differences less significant than P-value of 0.001; Bonferroni corrected. b) The time course of across-subject phase synchrony corresponding to sensor-band pair of F4- θ band. c) Across-subject phase synchrony, of the same sensor-band pair, conditioned on the subjective assessment of commercial clip.

of phase synchrony between the two conditions (i.e. movies and commercial film viewing), we detected statistically significant drops in inter-subject phase synchrony during the television ads (Fig. 4a). In particular, right frontal activity in θ-band (4-8 Hz) and left parietal activity in β1-band (13-20 Hz) seem to reflect mostly the disengagement of the viewers, reaching a percentage reduction in PLV level ($\Delta PLV = \frac{"ads\text{−}cartoon"}{"carton"}$) of 8%. The group-level profile of the most sensitive biomarker of inter-subject phase synchrony change is provided in Fig. 4b (via trimmed mean estimator of 10%). In addition, the corresponding profiles are shown in Fig. 4c., derived after confining the involved computations only among those who liked the commercial (blue trace) and those who did not (red trace).

IV. Discussion

Intersubject correlation is a well-established metric of engagement even in dynamic, naturalistic scenarios where an explicit response model may be unavailable [11],[12]. Deviating from its original linear flavor, we have introduced here three novel non-linear variants (based on ET-dynamics, spatial covariance dynamics and brainwave synchrony) and applied to the task of evaluating commercial advertisement clips.

Acknowledgment

This work was a part of project NeuroMkt, co-financed by EU's European Regional Development Fund and Greek National Funds via the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH CREATE INNOVATE (Project No T2EDK-03661) & project BINGO, financed by EU – NextGenerationEU, in the H.F.R.I framework call, "Basic research Financing (Horizontal support of all Sciences)" under the National Recovery and Resilience Plan "Greece 2.0" (H.F.R.I. Project No 15986).

References

[1] K. Georgiadis et al. "NeuMa-the absolute Neuromarketing dataset en route to an holistic understanding of consumer behaviour." Scientific Data, 10(1), 508 (2023).

[2] F. Kalaganis et al. "Unlocking the subconscious consumer bias: a survey on the past, present, and future of hybrid EEG schemes in neuromarketing". Frontiers in Neuroergonomics, 2, 11 (2021). https://doi.org/10.3389/fnrgo.2021.672982

[3] Hakim, A. et al.. "Machines learn neuromarketing: Improving preference prediction from self-reports using multiple EEG measures and machine learning." IJRM, 38(3), 770-791. 2021

[4] I. Daly, R. Scherer, M. Billinger, G. Müller-Putz. "FORCe: Fully online and automated artifact removal for brain-computer interfacing". IEEE transactions on neural systems and rehabilitation engineering, 23(5), 725-736 (2014).

[5] R.Brink., P. Murphy, S. Nieuwenhuis. "Pupil Diameter Tracks Lapses of Attention". Plos One. https://doi.org/10.1371/journal.pone.0165274

[6] F. Kalaganis, N. Laskaris, E. Chatzilari, S. Nikolopoulos, I Kompatsiaris. "A Riemannian geometry approach to reduced and discriminative covariance estimation in brain computer interfaces". IEEE Transactions on Biomedical Engineering 67 (1), 245-255 (2019)

[7] K Georgiadis et al., "^{R}Neumark: A Riemannian EEG analysis framework for neuromarketing". Brain Inform., 9 (1) (2022), p. 22. https://doi.org/10.1186/s40708-022-00171-7

[8] G. Székely, M. Rizzo, N. Bakirov. "Measuring and testing dependence by correlation of distances." Ann. Statist. 35 (6) 2769 - 2794, 2007.

[9] K. Davis et al. "Brainsourcing: Crowdsourcing Recognition Tasks via Collaborative Brain-Computer Interfacing." CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

[10] N. Laskaris, D. Adamos, and A. Bezerianos. "Graph Theory for Brain Signal Processing". In:Thakor (eds) Handbook of Neuroengineering. Springer, Singapore. 2021. https://doi.org/10.1007/978-981-15-2848-4_66-2

[11] J. Dmochowski, P. Sajda, J. Dias, L. Parra. "Correlated components of ongoing EEG point of emotionally laden attention – a possible marker of engagement?" Frontiers in Human Neuroscience, 6(112), 1–9. 2012.

[12] S. Cohen et al., "Neural engagement with online educational videos predicts learning performance for individual students",6,60-64, 2018.