scientific reports

Check for updates

OPEN

Decoding of image properties from single-trial visual evoked potentials recorded by ultra-high-density EEG

Sebastian Sieghartsleitner[1,2,5]✉, Leonhard Schreiner[1,3,5], Johannes Grünwald[1], Michael Jordan[1], Rossella Spataro[4], Josef Scharinger[2] & Christoph Guger[1]

Visual evoked potentials (VEPs) recorded by encephalography (EEG) allow us to study the neuronal activity non-invasively and in high temporal resolution. Traditionally, EEG analyses have relied on univariate group-level statistics and trial averaging to detect effects. However, recent advances in high-density EEG enable the investigation of brain responses at the single-subject and single-trial level. In this study, we combine ultra-high-density (uHD) EEG with cross-validated single-trial decoding to bridge both approaches, improving generalizability and reproducibility. Study participants were shown a diverse set of random images while 512 channels from the uHD system recorded their EEG over the occipital lobe. Image properties (contrast, hue, luminance, saturation and spatial frequency) were extracted for each stimuli and VEPs were used for decoding these properties in a cross-validated regression analysis. Additionally, the same data were spatially subsampled to investigate the impact of spatial resolution and electrode density on the decoding performance. Image properties could be decoded from single-trial VEPs, with contrast, saturation and spatial frequency providing the best decoding performances. Grand average decoding performance across all image properties and subjects yielded a Pearson's *r* of 0.50 between predicted and actual image property score. Greater electrode density improves decoding performance compared to standard EEG as well as subsampled configurations. Image properties robustly modulate early components of the VEP. Importantly, these modulations are pronounced enough to allow for single-trial decoding. Our analyses highlight the importance of electrode density with improvements in decoding performance extending even below 10 mm of inter-electrode distance.

Keywords Ultra-high-density EEG, Visual evoked potential, VEP, Image properties, Prediction, Regression, Decoding, Spatial resolution

The visual system enables us to perceive and interpret our environment[1–4]guides motor actions for effective interactions with it[5] and even influences our somatosensory experiences[6–9]. Thus it is no surprise that the human visual cortex occupies roughly a fifth of our cortex, spanning the entire occipital lobe and extending into portions of the temporal and parietal lobes[10]. It also surpasses the visual cortex of our primate relatives, encompassing approximately four times the area, with both early and higher-order visual regions being larger in humans, even after normalizing for brain size. This stark difference likely stems from the increased demands of visual processing, as information must be organized and relayed to regions dedicated to higher-order functions such as language and reading[11–13].

Investigating the physiological principles of the human visual cortex typically involves non-invasive and invasive imaging techniques, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), electrocorticography (ECoG) or stereo EEG (sEEG). Each method has its strengths: fMRI is renowned for its high spatial resolution across cortical and subcortical areas, while electrophysiological recordings (EEG, ECoG, sEEG) excel in temporal resolution[14]. Among these, visual evoked potentials (VEPs) recorded from EEG are particularly appealing for studying the visual cortex, offering a non-invasive, cost-effective and low-effort approach to measuring neuronal activity in response to a visual stimulus[15]. However, the relatively low spatial resolution of EEG remains a substantial limitation. In response, advancements such as very-high-density[16]

[1]g.tec medical engineering GmbH, Sierningstrasse 14, 4521 Schiedlberg, Austria. [2]Institute of Computational Perception, Johannes Kepler University, Altenberger Str. 69, 4040 Linz, Austria. [3]Institute for Integrated Circuits, Johannes Kepler University, Altenberger Str. 69, 4040 Linz, Austria. [4]Intensive Neurorehabilitation Unit, Villa delle Ginestre Hospital, ASP Palermo, 90135 Palermo, Italy. [5]These authors contributed equally: Sebastian Sieghartsleitner and Leonhard Schreiner. ✉email: sieghartsleitner@gtec.at

and ultra-high-density (uHD) EEG systems (e.g., Lee et al.[17] and Schreiner et al.[18,19]) have been developed to enhance spatial resolution.

In the current study we investigate if VEPs recorded from an uHD EEG system can decode image properties (i.e., contrast, hue, luminance, saturation and spatial frequency). The relationship between such image properties and VEPs have been investigated previously (e.g., contrast[20], luminance[21], spatial frequency[22]). Notably, these investigations typically utilize univariate analyses that assess statistical significance, which does not necessarily imply predictive power[23,24]. Predictive power, however, is crucial for improving generalizability, reproducibility and exploration of individual differences in neuroimaging research[25–29]. To overcome these shortcomings, we employ cross-validated regression analysis to evaluate both the magnitude of effect size and generalizability of observed effects to unseen data. The fact that image properties can be decoded was recently shown by Grootswagers et al.[30] and Schreiner et al.[31] using grating stimuli and classification analyses. However, moving beyond classification, we employ regression analyses to predict image properties extracted from a diverse image dataset, allowing us to show not just whether decoding is possible, but how precisely these properties can be reconstructed on a continuous scale. Finally, we also show that increased electrode density improves decoding performance.

Results

Decoding of image properties

Four subjects were presented with a variety of different visual stimuli while their EEG over the occipital lobe was recorded using a uHD EEG system (see Fig. 1a, b). For each stimuli five image properties were extracted (see Fig. 1c, d), which are known to modulate VEPs, with Fig. 1e providing an example for low vs. high contrast stimuli. Finally, subjects' VEPs were exploited to predict image property intensity utilizing a linear support vector machine (SVM) in a cross-validated regression framework.

Figure 1f shows the model performances for each image property and subject with corresponding numerical values available in Supplementary Table S1. All models performed above chance level ($P < 0.0001$, respectively). While here Pearson's r was used as the performance metric, above-chance performance was also observed with root mean square error (see Supplementary Table S2). Spatial frequency shows the best decoding performance followed by contrast and saturation (0.59, 0.52 and 0.50 on average, respectively). On the other hand, luminance showed the lowest decoding performance with 0.41 on average. In general, the Pearson's r values indicate good fit between the predicted and actual image property scores.

In additional analyses, we examined the effects of trial rejection and the use of a non-linear kernel function on decoding performance (see Supplementary Sect. 2.1 and 2.2). Neither approach led to notable changes in decoding performance, suggesting that data quality was largely unaffected by artifacts and that, given the current stimuli set, a linear kernel is sufficient to capture the relationship between visual features and neural responses. Finally, VEPs of identical stimuli (i.e., images) were averaged to improve signal-to-noise ratio and Pearson's r were recalculated, resulting in a grand average performance increase from 0.50 to 0.65 (see Supplementary Sect. 2.3).

Figure 2 shows the impact of image property intensity on VEPs at the uHD electrode closest to the Oz location (see Subsampling analyses for electrode selection method). Across image properties, VEPs consistently differ between the top and bottom quartiles, though the exact patterns vary by subject. For instance, subjects S1 and S4 show clear negative deflections around 100 ms post-stimulus, consistent with typical early VEP components, whereas this is not the case subject S3. VEPs related to high contrast and low luminance appear similar, though not identical. This similarity is due to the strong anti-correlation between these two properties in the stimulus set (see Supplementary Sect. 1). For spatial frequency, high-frequency stimuli generally evoke a stronger negative potential at ≈100 ms, except in subject S3, where low-frequency stimuli elicit a positive potential that diminishes with increasing frequency. At ≈200 ms, all subjects show a stronger positive potential in response to high spatial frequency stimuli. Supplementary Figure S3 shows an example how each uHD electrode of subject S1 is impacted by contrast intensity.

Linear models, such as the linear SVM used here, enable the investigation of feature importance and decoding power by examining the weights assigned by the model. Specifically, features (i.e., channels and time points) which are assigned a greater absolute feature weight indicate greater importance and the sign provides information whether the polarity of the evoked potential has a positive or negative relationship with the image property. Interpreting these weights in terms of decoding power is only valid if (1) model performance is sufficiently above chance since weights from poorly performing models are not informative and (2) features were standardized prior to training[32]. Here we generated topographies using a custom software (g.tec medical engineering GmbH, Austria) according to Kubanek and Schalk[33] based on the weights averaged across the 10×10 cross-validation framework. Figure 3 shows the feature weights assigned for each image property and subject. Early time points post-stimulus onset were chosen for visualization as they reflect early VEP components as the N1. For spatial frequency, 95 ms was chosen due to its focal pattern at the occipital pole but started to disperse around 120 ms (see Supplementary Figure S4). In general, medial electrodes provide the greatest decoding power. This is especially pronounced for spatial frequency for which a clear focal increase in feature importance at the occipital pole can be observed. Finally, there is an inverse relationship between contrast and luminance in all subjects except S4.

Supplementary Figure S5 shows the average feature weight assigned to each channel across time, for the respective subjects and image properties, indicating that early time points between 100 ms to 300 ms post-stimulus onset provide the greatest feature importance. Additionally, Supplementary Figure S6 shows the topographies for the time point at which the greatest feature importance (i.e., greatest average absolute feature weight across channels) was observed for each subject and image property.

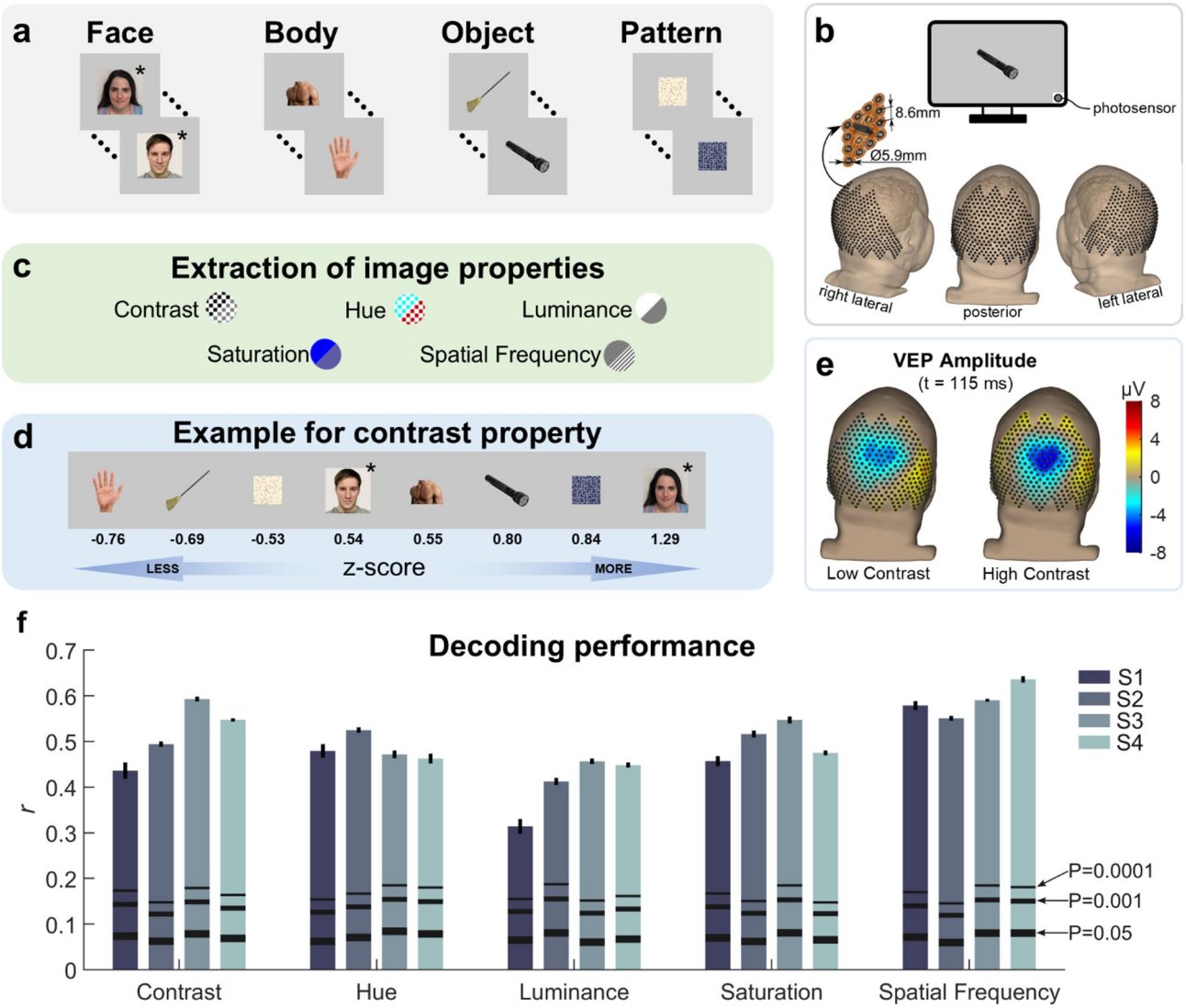


Fig. 1. (**a**) Visual stimuli of four categories (face, body, object, pattern) were presented to subjects, with the image dataset comprising 172 images in total. *Example face stimulus are not from the original dataset and were instead replaced with an image of an individual who provided written consent for illustrative use. (**b**) Stimulus presentation on a computer screen while their encephalogram was recorded with the ultra-high-density system covering the occipital lobe, recording from 16 electrode grids. (**c**) Extracted image properties of each stimulus (contrast, hue, luminance, saturation and spatial frequency) which were used for the regression analyses. (**d**) Example ranking of standardized (i.e., z-transformed) contrast scores across a subset of stimuli in increasing order. (**e**) Topography for visual evoked potential amplitude at 115 ms post-stimulus presentation for subject S1 comparing low contrast (lower 50th percentile) to high contrast (upper 50th percentile) stimuli. (**f**) Pearson's *r* between predicted and actual image property (contrast, hue, luminance, saturation, spatial frequency) for each subject (S1 to S4) obtained using single-trial visual evoked potentials. Bars reflect the mean Pearson's *r* and error bars reflect the standard deviation across the 10×10 cross-validation. Black lines for each subject and image property indicate empirical *P*-values of 0.05, 0.001 and 0.0001, respectively.

Subsampling analyses

In subsampling analyses, we investigated the impact of spatial resolution on decoding performance by subsampling electrodes from the uHD EEG system, enabling within-subject and within-recording comparisons.

Table 1 shows the regression results across all subjects for each image property and EEG system evaluated. Here, we subsampled the uHD EEG system by selecting the nearest electrode within 10 mm of each standardized position (e.g., Oz) based on Euclidean distance (see Standard EEG systems). The uHD EEG system performed best for each image property followed by the 10–05, extended 10–10, 10–10 and 10–20 system (see Supplementary Sect. 4 for statistical analysis). Hence, the decoding performance increased with increasing number of electrodes and decreasing inter-electrode distance.

In an additional subsampling analysis, we progressively increased electrode density starting from the 10–10 system until all uHD EEG electrodes were included, while ensuring even spatial coverage (see Spatial resolution).

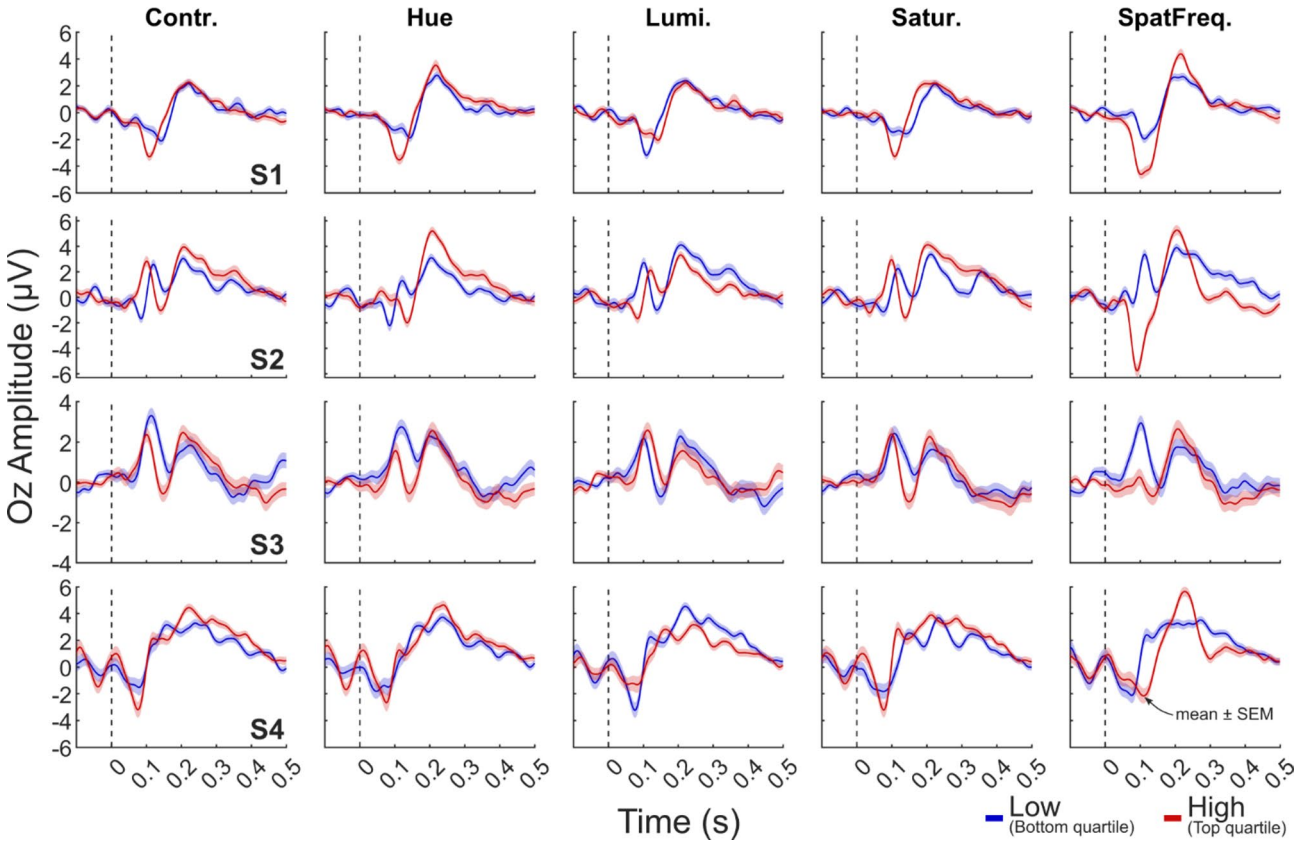


Fig. 2. Visual evoked potentials stratified into low (bottom quartile) and high (top quartile) image property intensity for each image property (contrast, hue, luminance, saturation, spatial frequency) and subject. The ultra-high density electrode closest to the Oz location was chosen for this visualization. Data were bandpass filtered from 1 to 30 Hz and baseline corrected using the pre-stimulus interval of 100 ms. The solid line and shared area represent the mean and standard error of the mean (SEM), respectively. The vertical dashed line indicates the stimulus onset.

Again, an inverse relationship between inter-electrode distance and decoding performance was observed (see Fig. 4). Here, the average Pearson's *r* across the image properties was calculated for each subject and then plotted over the inter-electrode distance for the respective subset. The non-linear behavior was investigated using a non-linear mixed-effects model according to Eq. 1.

$$f(\beta, x) = \beta_1 + \beta_2 x + \frac{\beta_3}{x} \quad (1)$$

The target variable is the obtained Pearson's *r*, x is the inter-electrode distance, and β_1, β_2, β_3 are the parameters to be estimated. Figure 4 shows the non-linear regression fit assuming β_1, β_2, β_3 to be fixed effects [β_1= 0.489 (95% CI 0.439 to 0.538), β_2= -0.0016 (95% CI -0.0028 to -0.0004), β_3= 0.211 (95% CI 0.018 to 0.405)]. Note that a mixed-effects model was necessary because the data of four subjects was fitted. Increases in decoding performance with decreasing electrode spacing (i.e., greater spatial resolution) can be observed.

Temporal generalization

The temporal generalization method (TGM)[34] was utilized to investigate when image properties became decodable and how predictive information evolved and generalized across the VEP. The TGM is based on the principle of training and testing at different (off-diagonal) and identical (main-diagonal) time points and results are typically visualized in temporal generalization maps (see Fig. 5a). Chance level performance was defined as a Pearson's *r* ranging from -0.1 to 0.1 based on the empirical *P*-values observed in Fig. 1f. Decoding performance starts increasing at roughly 75 ms post-stimulus onset and generally reaches its maximum within 200 ms post-stimulus onset, except for luminance in subject S4. Contrast, hue, and saturation TGM maps exhibit the best decoding performance with two increases in performances at roughly 100 ms and 200 ms, respectively. Additionally, the same TGM maps show off-diagonal below chance level performances (i.e., negative Pearson's *r*), which is caused by changes in polarity of the VEPs. Luminance shows the lowest decoding performances which corresponds to results shown in Fig. 1f. The highest performances can be observed along the main diagonal of the TGM maps, which reflects the temporal evolution of decoding performance where training and testing occur at identical time points. These results indicate that early VEP components carry the most predictive power with

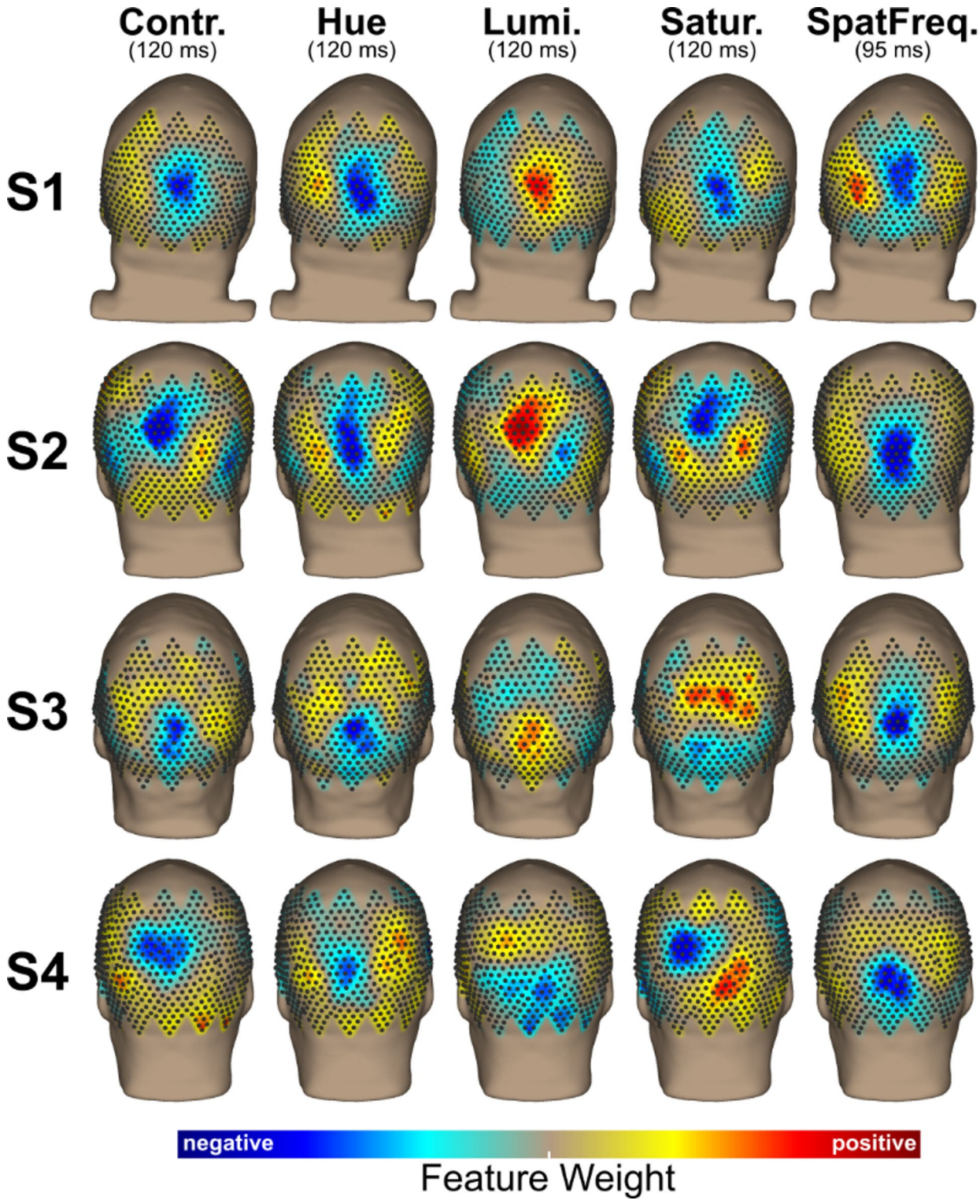


Fig. 3. Feature weights assigned by the linear support vector machine regression model for each subject (S1 to S4) and image property (contrast, hue, luminance, saturation, spatial frequency). The feature weights reflect the average across the 10 repetitions of 10-fold cross-validation. Note that for spatial frequency an earlier time point was chosen (95 ms post-stimulus onset) than for all other properties (120 ms post-stimulus onset).

	uHD	10–05	ext. 10–10	10–10	10–20
Contrast	0.52 (0.07)	0.50 (0.06)	0.46 (0.07)	0.45 (0.07)	0.39 (0.04)
Hue	0.49 (0.03)	0.47 (0.03)	0.44 (0.04)	0.44 (0.05)	0.36 (0.04)
Luminance	0.41 (0.07)	0.39 (0.06)	0.34 (0.10)	0.34 (0.08)	0.29 (0.06)
Saturation	0.50 (0.04)	0.47 (0.03)	0.43 (0.06)	0.41 (0.08)	0.36 (0.07)
SpatFreq.	0.59 (0.04)	0.57 (0.05)	0.52 (0.06)	0.51 (0.08)	0.46 (0.06)
Average	**0.50 (0.07)**	**0.48 (0.06)**	**0.44 (0.07)**	**0.43 (0.06)**	**0.37 (0.07)**

Table 1. Regression results obtained in the 10 × 10 cross-validation showing the pearson's *r* in mean (standard deviation) between predicted and actual image property (contrast, hue, luminance, saturation, Spatial frequency) for the ultra-high density, 10–05, extended 10–10, 10–10 and 10–20 EEG system.

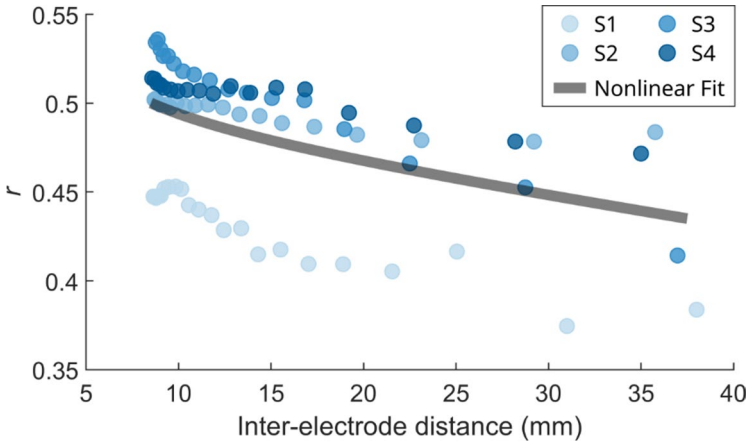


Fig. 4. Pearson's *r* obtained by the linear support-vector machine regression model and averaged across image properties (contrast, hue, luminance, saturation, spatial frequency) for each subject (S1 to S4) over inter-electrode distance.

decoding performance sharply rising at ≈75 ms and again declining at ≈250 ms (see Fig. 5b). The main-diagonal decoding performance is additionally shown in Supplementary Figure S7 for each subject and image property.

Discussion

Decoding of image properties

Visual stimuli comprising a diverse set of images including faces, bodies, objects, and patterns, evoke distinct VEPs that vary with the visual properties of each image. Our results show that these properties can be reliably predicted from VEPs recorded over the occipital cortex.

The investigation of feature importance, based on the weights assigned by the linear SVM regression model, revealed that early components of the VEP near the occipital pole show the strong predictive power (see Fig. 3 and Supplementary Figure S5). These findings correspond to our current understanding of visual processing in the human brain[1,3,20–22,35].

Visualization of VEPs (see Fig. 2) and feature weights (see Fig. 3) offers several insights. First, VEPs exhibit considerable inter-subject variability. While this may partly be due to the use of naturalistic images, which could introduce greater variability than classical stimuli such as grating, the precise reasons are difficult to determine. One possibility could be individual differences that can occur even in early components and are often overlooked when studies focus on group averages[36,37]. For instance, although all people fundamentally have the same VEP components, their shape may be different due to variations in cortical folding[36,38]. Supporting this notion, the feature weight topographies were more consistent across subjects, suggesting that similar VEP components are modulated in comparable ways. Notably, medial electrodes near the occipital pole, tended to provide the greatest decoding power. For spatial frequency, all subjects had large negative weights assigned to a focal area over their medial occipital lobe, likely reflecting the location of V1[39]. Thus, our model inferred an inverse relationship between the VEP's polarity around 100 ms post-stimulus onset and spatial frequency. This aligns with previous findings that VEPs tend to be more negative for high spatial frequency stimuli[22,40]. A similar, albeit less focal, pattern was observed for contrast with greater negative weights over the occipital pole aligning with previous reports of larger N1 amplitudes for higher contrast stimuli[20] (see Fig. 2). Hue, like spatial frequency and contrast, was also associated with more negative feature weights for medial occipital electrodes. Moreover, an inverse relationship between feature weights for contrast and luminance was observed in all subjects except S4. This goes in hand with observations made by Johannes et al.[21] that high stimulus luminance provokes larger occipital P1. Finally, saturation shows varying feature weight topographies between subjects with subjects S1 and S2

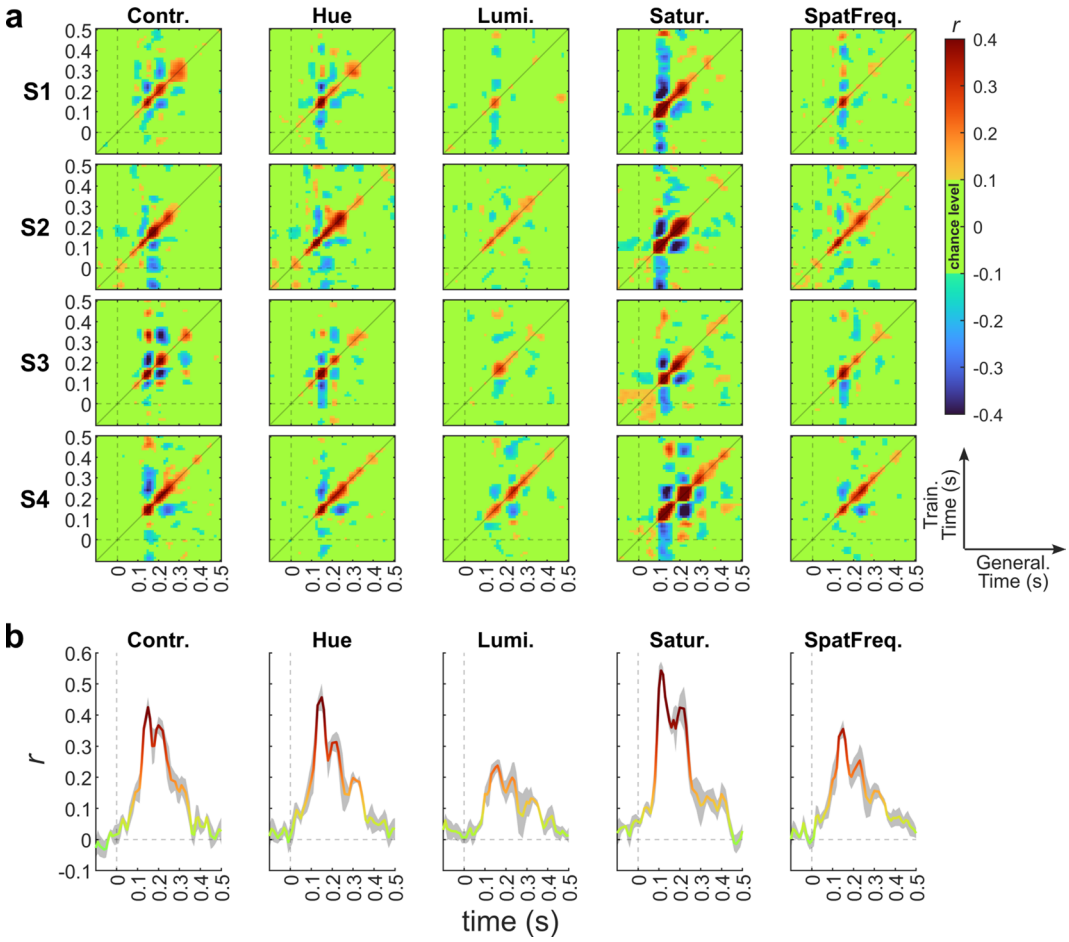


Fig. 5. (**a**) Temporal generalization method results showing the model performance (Pearson's *r*) for a given training time (y-axis) and generalization time (x-axis) for subjects S1 to S4 and image property. (**b**) Model performance for the main-diagonal, when training and generalization time are identical, showing the decoding performance across the trial. The horizontal and vertical dashed lines indicate model performance of zero and stimulus onset, respectively. The solid line and shaded area represent the mean and standard error of the mean across subjects, respectively.

again exhibiting more medial occipital patterns and S3 and S4 lateralized patterns. The importance of medial structures such as V1 was expected as V1 is strongly involved in the processing of color features[35].

Taken together, here we showed that single-trial VEPs were able to decode image properties extracted from more naturalistic images, in comparison to abstract stimuli (e.g., gratings and shapes). The latter is typically chosen to investigate image properties and their impact on the visual system as they can be artificially generated to isolate a given image property. However, our results show that these image properties can nonetheless be extracted and decoded from VEPs even when more naturalistic images are used. Furthermore, feature-weight topographies largely align with expectations from literature while also emphasizing the importance of individual differences.

Subsampling analysis

The usage of the uHD EEG system allowed us to investigate how spatial resolution and electrode density impact decoding performance. Note that this system has been shown to improve motor decoding performance for individual finger movements[17] and hand gestures[18] compared to standard EEG configurations. However, these studies utilized spectral features based on event-related desynchronization, whereas here we investigate if these findings carry over to evoked-potentials recorded over the occipital lobe. To this end, the uHD EEG electrodes were subsampled to obtain the respective EEG systems and thus the same subjects with the same data could act as their own controls providing a highly controlled setting.

The uHD EEG system consistently yielded the best performance in comparison to subsets based on the 10–05, extended 10–10, 10–10 and 10–20 system (see Table 1). Furthermore, the average performance across subjects and image properties decreased with greater inter-electrode distance of the EEG systems. While this is in line with previous studies[16–18] one may reasonably expect the opposite especially in machine learning contexts. Specifically, greater electrode density over the same scalp area increases electrode count and thus feature count (in our case: 22 additional features per additional electrode). This may lead to poorly performing models, as they

could overfit on training data but perform poorly on unseen data, especially if features are redundant or irrelevant. However, our results indicate that the additional EEG channels and resulting features are neither irrelevant nor redundant but informative as improvements in decoding performance were observed. Finally, improvements in said decoding performance extended even below 10 mm of inter-electrode distance (see Fig. 4)[41]. This may seem at odds with earlier theoretical reports that distances of 20 to 30 mm yield maximum possible resolution from EEG signals[42,43] which have since been questioned[44]. However, it is important to distinguish between decoding performance gains arising from (1) access to additional neural information and (2) improved estimation of the underlying neural signal due to increased signal-to-noise ratio. While these two mechanisms are not mutually exclusive, the latter appears to play a more prominent role based on analyses presented in Supplementary Sect. 7.

Temporal generalization

Temporal generalization maps indicate that early components of the VEP can decode image properties (see Fig. 5b). While this is to be expected based on statistical analyses[20–22,40] it is important to note that this also upholds in cross-validated predictive analyses[24].

As noted in Temporal generalization, contrast, hue and saturation maps exhibit negative Pearson's r in the off-diagonal. Changes in polarity across the classified VEP can lead to such below chance level performance. Here, we can take a closer look at subject S1's contrast temporal generalization map (see top left map in Fig. 5a) and corresponding VEP time course (see top left VEP in Fig. 2). Figure 2 shows that electrodes near the occipital pole are strongly modulated by contrast and experience a polarity change from negative to positive. Additionally, greater contrast leads to both a greater negative (110 ms) and positive peak (220 ms) peak, providing the inverse relationship which goes in hand with the below chance level performance in Fig. 5a.

While feature weight topographies for the saturation property show some variation across subjects (S1/S2 vs. S3/S4 in Fig. 3), temporal generalization maps do not. Instead, two distinct increases in decoding performance at ≈100 ms and ≈200 ms post-stimulus can be identified across all subjects reflecting the early visual components[45]. Note that the feature weights are closely linked to the VEP morphology (e.g., polarity and amplitude) and their topography, whereas the temporal generalization maps simply show the decoding performance of the VEP components. Taken together, these findings further bolster the notion that while people likely have the same VEP components, their morphology may vary due to differences in cortical folding[36,38].

Comparing results utilizing the whole VEP time course (see Fig. 1) versus single time points (see Fig. 5), shows that grand average performance is greater for the former (Pearson's $r = 0.50$ vs. 0.44). Notably, the grand average performance of 0.44 is even based on the peak performance observed for each temporal generalization map in Fig. 5a, which is not attainable for brain-computer interface applications as the exact time point of peak performance is not known a priori. These results indicate, that considering the whole VEP time course, rather than single time points, allows models to capture more complex relationships, such as interactions between VEP components and temporal patterns such as small drifts in the VEPs.

Limitations and future work

One key strength of this study is the use of more naturalistic images, rather than abstract stimuli. However, the uHD EEG system could also be applied to abstract stimuli, as they enable precise isolation of low-level image properties, something which is not possible in the current image dataset. Using abstract stimuli such as grating patterns may enhance the information provided by decoding models and source reconstruction algorithms. Furthermore, these stimuli could sample a larger space of a given image property allowing for better characterization of population level neural responses. Additionally, future work could explore real-time decoding of image properties and image reconstruction.

Conclusion

In the current study, we investigated if image properties (contrast, hue, luminance, saturation, spatial frequency), extracted from a diverse dataset of images, can be decoded based on VEPs recorded over the occipital cortex. To this end, single-trial VEPs were used to predict the image properties in a cross-validated regression analysis. Additionally, we explored the impact of spatial resolution and electrode density on decoding performance.

Results revealed that image properties robustly modulate early components of the VEP allowing decoding of image properties with contrast, spatial frequency and saturation yielding greatest decoding performance. Higher electrode density provides superior decoding performance in comparison to both standard EEG systems and subsampled subsets.

Materials and methods

Subjects

Four healthy male subjects [S1 to S4, age: 35 (10) years in mean (SD)] with normal or corrected to normal vision participated in the study. Subjects received no prior training regarding the experiment.

The study was conducted in accordance with the principles embodied in the Declaration of Helsinki and approved by the ethics committee IRCCS Sicilia under approval number U148/19. All subjects provided written informed consent before taking part in the experiment.

Experimental protocol

Subjects sat in a comfortable chair and were asked to focus on the visual stimuli (i.e., images) presented on a computer screen (24 inches, 1920 × 1080, 60 Hz, 2 ms reaction time). The visual stimuli comprised four categories: face, body, object and pattern (see Fig. 1a). The whole dataset consisted of 46 face, 41 body, 57 object and 28 pattern images[46,47] which were presented in a pseudorandom manner to balance the number of images

across categories. Two example images per category are shown in Fig. 1a. The images were set against a gray background and a small white square was added in the bottom right corner to obtain the stimulus onset timing using a photosensor (see Fig. 1b). The respective images were shown for a random duration of 0.75, 1.0–1.25 s followed by a random inter-stimulus (ISI) of 0.6, 0.8 and 1.0 s. During this ISI, the same gray background was shown, and the white square was replaced with a black one. The experiment was split into 10 runs (i.e., blocks) during which 20 images per image category were randomly shown. In total, this yielded 800 stimuli (i.e., trials) (10 runs × 4 image categories × 20 images) per subject.

The experiments were performed using MATLAB/Simulink (The MathWorks, Inc., MA, USA) and g.HIsys Professional (g.tec medical engineering GmbH, Austria). Stimulus presentation was implemented using ParadigmPresenter as part of g.HIsys Professional. The luminance change in the bottom right corner during the presentation of an image (i.e., switch from black to white) was registered by the photosensor, converted to a TTL signal by the g.TRIGbox (g.tec) and finally input to the digital channel of the g.HIamp biosignal amplifier (g.tec).

Ultra-high-density EEG system

For this study, we used the g.Panoglin uHD EEG system (g.tec), whose electrode grids feature 16 electrodes each with an inter-electrode distance of 8.6 mm and an exposed sensor diameter of 5.9 mm (see Fig. 1b).

Before mounting g.Pangolin, the subjects' heads were shaved and cleaned with medical alcohol to achieve optimal conductivity between the electrode grids and the scalp. Additionally, a standard and appropriately sized EEG cap (g.GAMMAcap, g.tec) was used to identify the Oz electrode location, which was defined to coincide with first channel of the first grid. Subsequent grid placements were pre-determined by electrode grid geometry.

We used 512 channels per subject, which required two electrode connector boxes and two 256-channel g.HIamp biosignal amplifiers. Both connector boxes and amplifiers shared a common ground electrode which was located on the right mastoid. The acquired data was sampled at 512 Hz. Note that covering the whole scalp would require 1024 channels in total (see Fig. 3a, in Schreiner et al.[19]).

Magnetic resonance imaging scans

Anatomical MRI scans were performed for all subjects and the resulting T1-weighted images were then utilized to reconstruct brain and skull structures in FreeSurfer (Martinos Center for Biomedical Imaging, Cambridge, MA, USA)[48]. During the MRI scans, subjects wore a standard and appropriately sized EEG cap equipped with MRI markers at standard electrode positions. These markers caused MRI artifacts during the scans, which were used for co-registration in a custom montage software (g.tec). Specifically, this software was used for generating subject-specific montages.

Data processing

All data processing was performed in MATLAB.

Image properties

The following five image properties were extracted for each image of the dataset: contrast, hue, luminance, saturation and spatial frequency. These image properties were used as the target variable (i.e., dependent variable) in the regression analyses (see Cross-validated regression).

Contrast describes differences in luminance and was estimated by first converting the image to grayscale, then calculating the root mean square (RMS) of the grayscale intensity values across all pixels. This results in a single value known as the RMS contrast and was chosen as it is independent of the spatial frequency and spatial distribution of the image[49].

Hue was calculated by converting the images from red, green, blue (RGB) color space to the hue, saturation, value (HSV) color space followed by calculating the mean hue value across all pixels of the image[50].

Luminance (i.e., brightness) was estimated by converting the images from RGB to YCbCr color space with Y being luminance and Cb and Cr being the chrominance (i.e., color information). Finally, the mean luminance across all pixels of the image was calculated.

Saturation describes the color intensity and was calculated by converting the images from RGB to HSV color space followed by calculating the mean saturation across all pixels of the image[50].

Spatial frequency reflects the level of clarity and image activity with lower spatial frequency corresponding to greater clarity and smoother, less intricate content. Here, we estimated the spatial frequency using the 2D Fourier Transform. First the images were converted to grayscale, followed by the application of the 2D Fourier Transform to obtain the images' frequency domain representation. The magnitude of the Fourier coefficients was then weighted by their distance from their Euclidean distance from the center of the frequency spectrum (i.e., the zero-frequency component). This ensures that high-frequency components have more weight than low-frequency components. Finally, the sum of these weighted Fourier coefficients is divided by the total sum of Fourier coefficients, making the measure independent of the overall image brightness (i.e., scale-invariant) and thus comparable between different images.

Finally, the Yeo-Johnson transformation was used to improve Gaussianity of each of the five obtained image property populations[51] and a z-transform was used for standardization. Figure 1d shows example images including their contrast score in increasing order. Supplementary Sect. 1 describes the relationship between the image properties, as well as image categories.

EEG preprocessing

The raw EEG data were initially multiplied by a factor of 0.1, to remove the pre-amplification factor of 10, and then notch-filtered at 50 Hz and its harmonics using a 3rd -order Butterworth filter. Note that all filters were implemented as zero-phase digital filters (*filtfilt* in MATLAB) to preserve the phase of the underlying signal.

Bad channels were identified and removed according to Schreiner et al.[19] and the remaining channels were referenced to the common average.

EEG feature extraction and epoching
A bandpass filter (1 to 20 Hz) was applied for the cross-validated regression analysis described in Cross-validated regression, whereas a bandpass filter from 1 to 40 Hz was applied for temporal generalization method (TGM) described in Temporal generalization. The relatively small upper cut-off value of 20 Hz was chosen to keep the number of features used in the regression analysis reasonably small, while still capturing most of the energy in the VEP (see Cross-validated regression for further information). Finally, the data were epoched into trials ranging from 0.1 s to 0.7 s pre- and post-stimulus onset, respectively.

Cross-validated regression
The epoched EEG data of size $N \times T \times C$, with N being trials, T being time samples and C being channels, were then downsampled according to the Nyquist theorem. Specifically, a downsampling factor of 12 was applied as the EEG data were previously low pass filtered at 20 Hz, resulting in a new sampling rate of approximately 42.67 Hz. A time window of 0.0 to 0.5 s post-stimulus onset was chosen as it is expected to contain the relevant VEP. Finally, the last two dimensions of the epoched and downsampled EEG data $N \times T_d \times C$ were serialized (i.e., flattened), resulting in $T_d \times C$ features per trial, with T_d being the time samples after downsampling. Brigell et al.[52] stated in their guideline that most of the VEP energy is within 3 to 30 Hz. However, increasing the upper cut-off value from 20 Hz to 30 Hz would have increased the number of features by nearly 50% without increasing decoding performance (see Supplementary Sect. 2.4).

Cross-validated regression models were employed to investigate if the single-trial VEPs can predict image properties. Specifically, an SVM regression model was employed and evaluated in a 10-fold cross-validation (CV) framework, which was repeated 10 times[53]. A linear SVM was chosen as it is robust to outliers and can handle datasets which have more features than observations (i.e., trials in the current context). SVM parameters were set as follows: The scale of the linear kernel was automatically selected by MATLAB using a heuristic approach, features were standardized and a box constraint of 1.0 was used to prevent overfitting. As performance metric, the linear Pearson correlation coefficient was used, which will be referred to as Pearson's *r* subsequently. Pearson's *r* was employed as performance metric instead of the root mean square error as the Pearson's *r* is easier to interpret in the current context.

Empirical *P*-values were calculated to assess the likelihood of the observed model performance occurring by chance. To this end, a permutation approach was employed, where the target variable was shuffled, generating null models based on the assumption of no relationship between the target variable and the features[17]. This permutation approach was performed for each image property and subject, resulting 20 *P*-values. Finally, *P*-values were corrected for multiplicity (i.e., multiple hypothesis testing) using the false discovery rate (i.e., FDR, Benjamini-Hochberg) correction[54]. Model performance was considered better than chance for $P < 0.05$.

Subsampling analyses

Two analyses were performed to investigate the impact of spatial resolution and electrode density on decoding performance. Notably, these analyses are based on the same subjects, data and processing as described in Data processing.

Standard EEG systems
The performance obtained by the uHD EEG system was compared to those obtained from conventional EEG systems such as the: 10–20, 10–10, extended 10–10 and 10–05 system. The standard electrode positions were extracted for each subject using the FieldTrip toolbox[55] and four reference points (nasion, inion, left and right pre-auricular points). Finally, for each standardized electrode position (e.g., Oz), the nearest uHD electrode was associated with the respective system. Finding the nearest uHD electrode was based on the Euclidean distance, where only standard positions within 10 mm of their uHD counterpart were considered. As a result, electrodes such as AFz, Fz, and others were not approximated, as they were not covered by the uHD system. Finally, the EEG data from these selected electrodes were processed according to EEG preprocessing, EEG feature extraction and epoching and Cross-validated regression.

Spatial resolution
Spatial resolution, quantified by inter-electrode distance, and its impact on decoding performance was investigated by first selecting electrodes according to the 10–10 system. The number of electrodes was then progressively increased (doubling, tripling, quadrupling, etc.), until all uHD EEG electrodes were included. At each stage of this densification process, newly added electrodes were selected to maintain an even distribution and ensure comparable inter-electrode distance. Inter-electrode distance was calculated as the mean Euclidean distance of each electrode and its four nearest neighbors. Again, EEG data from selected electrodes were processed according to EEG preprocessing, EEG feature extraction and epoching and Cross-validated regression.

Temporal generalization

The TGM[34] was utilized to investigate the temporal dynamics of VEPs and how they relate to their predictive power. A broader time window (-100 to 500 ms relative to stimulus onset) was used in the TGM analysis to confirm that decoding performance starts at chance level prior to stimulus presentation. Furthermore, a larger upper cut-off value of 40 Hz for the bandpass filter was chosen to provide higher temporal resolution. Consequently, the downsampling factor of 6 was chosen according to the Nyquist Theorem, resulting in a new sampling rate of 85.3 Hz and 69 time samples after downsampling and epoching.

The TGM is based on training and testing (i.e., evaluating) the model on all combinations of these 44 time points allowing us to gain insights into when image properties become decodable and how this generalizes across time. Thus, the number of features reflects the number of channels in this analysis. The evaluation framework chosen was 10 repetitions of a 10-fold CV. Importantly, the performance metric used for TGM analyses must be criterion-free, meaning the metric does not depend on an arbitrary decision threshold and is thus robust to systematic biases[34]. These systematic biases can arise in TGM analyses when training and testing time points are at different time points. The Pearson's *r* satisfies this requirement and was thus chosen for this analysis.

This analysis was performed for each of the five image properties, four subjects, and 69 time samples, resulting in 1380 different models.

Data availability

A sample dataset of subject S1, including an appropriate description of the dataset, is available in an online repository at: https://osf.io/h7s9r/?view_only=1a3444d18cee4f47a7c7717672fdee09. Requests to access complete datasets should be directed to CG, guger@gtec.at.

Received: 1 April 2025; Accepted: 1 September 2025
Published online: 25 September 2025

References

1. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
2. Li, F. F., VanRullen, R., Koch, C. & Perona, P. Rapid natural scene categorization in the near absence of attention. *Proc. Natl. Acad. Sci. U S A.* **99**, 9596–9601 (2002).
3. Johnson, J. S. & Olshausen, B. A. Timecourse of neural signatures of object recognition. *J. Vis.* **3**, 4 (2003).
4. Johnson, J. S. & Olshausen, B. A. The earliest EEG signatures of object recognition in a cued-target task are postsensory. *J. Vis.* **5**, 2 (2005).
5. Goodale, M. A. Visuomotor control: where does vision end and action begin? *Curr. Biol.* **8**, R489–R491 (1998).
6. Kennett, S., Taylor-Clarke, M. & Haggard, P. Noninformative vision improves the Spatial resolution of touch in humans. *Curr. Biol.* **11**, 1188–1191 (2001).
7. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
8. Taylor-Clarke, M., Kennett, S. & Haggard, P. Vision modulates somatosensory cortical processing. *Curr. Biol.* **12**, 233–236 (2002).
9. Schaefer, M., Heinze, H. J. & Rotte, M. Seeing the hand being touched modulates the primary somatosensory cortex. *NeuroReport* **16**, 1101–1105 (2005).
10. Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. Visual field maps in human cortex. *Neuron* **56**, 366–383 (2007).
11. Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. Visual Cortex in Humans. In *Encyclopedia of Neuroscience* 251–257Elsevier, (2009). https://doi.org/10.1016/B978-008045046-9.00241-2.
12. Meyer, E. & Arcaro, M. Larger area size, not increased number, better explains expansion of human visual cortex. *J. Vis.* **23**, 5557 (2023).
13. Meyer, E. E., Martynek, M., Kastner, S., Livingstone, M. S. & Arcaro, M. J. Expansion of a conserved architecture drives the evolution of the primate visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2421585122 (2025).
14. Haufe, S. et al. Elucidating relations between fMRI, ecog, and EEG through a common natural stimulus. *NeuroImage* **179**, 79–91 (2018).
15. Carter, J. L. Visual evoked potentials. in Clinical Neurophysiology 311–322 (Oxford University Press) https://doi.org/10.1093/med/9780195385113.003.0022. (2009).
16. Robinson, A. K. et al. Very high density EEG elucidates Spatiotemporal aspects of early visual processing. *Sci. Rep.* **7**, 16248 (2017).
17. Lee, H. S. et al. Individual finger movement decoding using a novel ultra-high-density electroencephalography-based brain-computer interface system. *Front. Neurosci.* **16**, 1009878 (2022).
18. Schreiner, L., Sieghartsleitner, S., Mayr, K., Pretl, H. & Guger, C. Hand gesture decoding using ultra-high-density EEG. in *11th International IEEE/EMBS Conference on Neural Engineering (NER)* 01–04 (IEEE, Baltimore, MD, USA, 2023). 01–04 (IEEE, Baltimore, MD, USA, 2023). https://doi.org/10.1109/NER52421.2023.10123901 (2023).
19. Schreiner, L. et al. Mapping of the central sulcus using non-invasive ultra-high-density brain recordings. *Sci. Rep.* **14**, 6527 (2024).
20. Martinovic, J., Mordal, J. & Wuerger, S. M. Event-related potentials reveal an early advantage for luminance contours in the processing of objects. *J. Vis.* **11**, 1–1 (2011).
21. Johannes, S., Münte, T. F., Heinze, H. J. & Mangun, G. R. Luminance and Spatial attention effects on early visual processing. *Cogn. Brain. Res.* **2**, 189–205 (1995).
22. Kenemans, J. L., Baas, J. M. P., Mangun, G. R., Lijffijt, M. & Verbaten, M. N. On the processing of Spatial frequencies as revealed by evoked-potential source modeling. *Clin. Neurophysiol.* **111**, 1113–1123 (2000).
23. Lo, A., Chernoff, H., Zheng, T. & Lo, S. H. Why significant variables aren't automatically good predictors. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13892–13897 (2015).
24. Gordillo, D., Ramos Da Cruz, J., Moreno, D., Garobbio, S. & Herzog, M. H. Do we really measure what we think we are measuring? *iScience* **26**, 106017 (2023).
25. Poldrack, R. A. Inferring mental States from neuroimaging data: from reverse inference to Large-Scale decoding. *Neuron* **72**, 692–697 (2011).
26. Gabrieli, J. D. E., Ghosh, S. S. & Whitfield-Gabrieli, S. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron* **85**, 11–26 (2015).
27. Dubois, J. & Adolphs, R. Building a science of individual differences from fMRI. *Trends Cogn. Sci.* **20**, 425–443 (2016).
28. Jollans, L. & Whelan, R. The clinical added value of imaging: A perspective from outcome prediction. *Biol. Psychiatry: Cogn. Neurosci. Neuroimaging.* **1**, 423–432 (2016).
29. Westfall, J. & Yarkoni, T. Statistically controlling for confounding constructs is harder than you think. *PLoS ONE.* **11**, e0152719 (2016).
30. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. Mapping the dynamics of visual feature coding: insights into perception and integration. *PLoS Comput. Biol.* **20**, e1011760 (2024).
31. Schreiner, L. et al. Advancing Visual Decoding in EEG: Enhancing Spatial Density in Surface EEG for Decoding Color Perception. in. *IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)* 952–957 (IEEE, St Albans, United Kingdom, 2024). (2024). https://doi.org/10.1109/MetroXRAINE62247.2024.10796840.

32. Sieghartsleitner, S. et al. Analysis of Cortical Excitability During Brain-Computer Interface Stroke Rehabilitation of Upper and Lower Extremity. in. *IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)* 1106–1111 (IEEE, St Albans, United Kingdom, 2024). (2024). https://doi.org/10.1109/MetroXRAINE62247.2024.10796097.
33. Kubanek, J., Schalk, G. & NeuralAct A tool to visualize electrocortical (ECoG) activity on a Three-Dimensional model of the cortex. *Neuroinform* **13**, 167–174 (2015).
34. King, J. R. & Dehaene, S. Characterizing the dynamics of mental representations: the Temporal generalization method. *Trends Cogn. Sci.* **18**, 203–210 (2014).
35. Xing, D. et al. Brightness–Color interactions in human early visual cortex. *J. Neurosci.* **35**, 2226–2232 (2015).
36. Woodman, G. F. A brief introduction to the use of event-related potentials in studies of perception and attention. *Atten. Percept. Psychophys.* **72**, 2031–2046 (2010).
37. Luck, S. J. *An Introduction To the Event-Related Potential Technique* (MIT Press, 2014).
38. Nunez, P. L. & Srinivasan, R. *Electric Fields of the Brain* (Oxford University Press, 2006). https://doi.org/10.1093/acprof:oso/9780195050387.001.0001
39. De Valois, R. L. & De Valois, K. K. Spatial vision. *Annu. Rev. Psychol.* **31**, 309–341 (1980).
40. Boeschoten, M. A., Kemner, C., Kenemans, J. L. & Van Engeland, H. Time-varying differences in evoked potentials elicited by high versus low Spatial frequencies: a topographical and source analysis. *Clin. Neurophysiol.* **116**, 1956–1966 (2005).
41. Freeman, W. J., Holmes, M. D., Burke, B. C. & Vanhatalo, S. Spatial spectra of scalp EEG and EMG from awake humans. *Clin. Neurophysiol.* **114**, 1053–1068 (2003).
42. Srinivasan, R., Nunez, P. L. & Silberstein, R. B. Spatial filtering and neocortical dynamics: estimates of EEG coherence. *IEEE Trans. Biomed. Eng.* **45**, 814–826 (1998).
43. Srinivasan, R., Tucker, D. M. & Murias, M. Estimating the Spatial Nyquist of the human EEG. *Behav. Res. Methods Instruments Computers.* **30**, 8–19 (1998).
44. Grover, P. & Venkatesh, P. An Information-Theoretic view of EEG sensing. *Proc. IEEE.* **105**, 367–384 (2017).
45. Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S. & Hillyard, S. A. Cortical sources of the early components of the visual evoked potential. *Hum. Brain. Mapp.* **15**, 95–111 (2002).
46. Minear, M. & Park, D. C. A lifespan database of adult facial stimuli. *Behav. Res. Methods Instruments Computers.* **36**, 630–633 (2004).
47. Khuvis, S. et al. Face-Selective units in human ventral Temporal cortex reactivate during free recall. *J. Neurosci.* **41**, 3386–3399 (2021).
48. Dale, A. M., Fischl, B. & Sereno, M. I. *Cortical Surface-Based Anal. NeuroImage* **9**, 179–194 (1999).
49. Peli, E. Contrast in complex images. *J. Opt. Soc. Am. A.* **7**, 2032 (1990).
50. Smith, A. R. Color gamut transform pairs. *SIGGRAPH Comput. Graph.* **12**, 12–19 (1978).
51. Yeo, I. K. A new family of power transformations to improve normality or symmetry. *Biometrika* **87**, 954–959 (2000).
52. Brigell, M., Bach, M., Barber, C., Moskowitz, A. & Robson, J. Guidelines for calibration of stimulus and recording parameters used in clinical electrophysiology of vision. *Doc. Ophthalmol.* **107**, 185–193 (2003).
53. Bouckaert, R. R. & Frank, E. Evaluating the replicability of significance tests for comparing learning algorithms. in Advances in Knowledge Discovery and Data Mining (eds Dai, H., Srikant, R. & Zhang, C.) vol. 3056 3–12 (Springer Berlin Heidelberg, Berlin, Heidelberg, (2004).
54. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.: Ser. B (Methodol.).* **57**, 289–300 (1995).
55. Oostenveld, R., Fries, P., Maris, E., Schoffelen, J. M. & FieldTrip Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* 1–9 (2011). (2011).

Acknowledgements

We extend our gratitude to Leon Y. Deouell and Gal Chen for generously providing the image dataset. We also thank Pauline Schomaker and Matteo La Rosa for their contributions to the experiments. Finally, we are grateful to the subjects for their valuable time and to Christoph Kapeller for his methodological input regarding the spatial subsampling analysis.

Author contributions

SS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. LS: Conceptualization, Data curation, Investigation, Software, Visualization, Writing – original draft. JG: Methodology, Validation, Visualization, Writing – review & editing. MJ: Formal analysis, Software, Visualization, Writing – review & editing. RS: Data curation, Investigation, Resources, Writing – review & editing. JS: Supervision, Writing – review & editing. CG: Conceptualization, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Declarations

Competing interests

SS, LS, MJ and JG are employed by g.tec medical engineering GmbH. CG is the CEO of g.tec medical engineering GmbH. The remaining authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-18275-5.

Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025