Distinguishing the roles of edge, color, and other surface information in basic and superordinate scene representation

Liansheng Yao , Qiufang Fu , Chang Hong Liu , Jianyong Wang , Zhang Yi

PII: S1053-8119(25)00102-8
DOI: https://doi.org/10.1016/j.neuroimage.2025.121100
Reference: YNIMG 121100

To appear in: *NeuroImage*

Received date: 14 October 2024
Revised date: 13 February 2025
Accepted date: 26 February 2025

Please cite this article as: Liansheng Yao , Qiufang Fu , Chang Hong Liu , Jianyong Wang , Zhang Yi , Distinguishing the roles of edge, color, and other surface information in basic and superordinate scene representation, *NeuroImage* (2025), doi: https://doi.org/10.1016/j.neuroimage.2025.121100

Highlights

- Surface information is involved exclusively in scene representation at basic level.
- Edge information is sufficient and more effective at superordinate level.
- Role of edge and surface information varies with level of abstraction.

Distinguishing the roles of edge, color, and other surface information in basic and superordinate scene representation

Liansheng Yao[1,2], Qiufang Fu[1,2*], Chang Hong Liu[3], Jianyong Wang[4], Zhang Yi[4]

[1]State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]Department of Psychology, Bournemouth University, Fern Barrow, Poole, UK

[4]Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu, China

*Correspondence author

Qiufang Fu, Ph.D.

Address: Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Chaoyang District, Beijing 100101, China

Tel: (86 10) 6484-5395

E-mail: fuqf@psych.ac.cn.

CRediT authorship contribution statement

Liansheng Yao: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Qiufang Fu:** Conceptualization, Funding acquisition, Methodology, Project administration, Writing - review & editing. **Changhong Liu:** Conceptualization, Methodology, Writing - review & editing. **Jianyong Wan:** Visualization, Writing - review & editing. **Zhang Yi:** Conceptualization, Writing - review & editing.

Abstract

The human brain possesses a remarkable ability to recognize scenes depicted in line drawings, despite that these drawings contain only edge information. It remains unclear how the brain uses this information alongside surface information in scene recognition. Here, we combined electroencephalogram (EEG) and multivariate pattern analysis (MVPA) methods to distinguish the roles of edge, color, and other surface information in scene representation at the basic category level and superordinate naturalness level over time. The time-resolved decoding results indicated that edge information in line drawings is both sufficient and more effective than in color photographs and grayscale images at the superordinate naturalness level. Meanwhile, color and other surface information are exclusively involved in neural representation at the basic category level. The time-generalization analysis further revealed that edge information is crucial for representation at both levels of abstraction. These findings highlight the distinct roles of edge, color, and other surface information in dynamic neural scene processing, shedding light on how the human brain represents scene information at different levels of abstraction.

Key words: scene representation, edge information, surface information, basic level of category, superordinate level of naturalness

1. Introduction

The human brain possesses a remarkable ability to recognize scenes and objects depicted in line drawings, despite these drawings containing only edge information, such as lines, contours, and shapes, and devoid of surface information that exists in the real world, such as color, texture, and luminance (Biederman & Ju, 1988). This ability is apparently innate, because even an infant who has no experience with line drawings can recognize objects depicted in them (Hochberg & Brooks, 1962; Yonas & Arterberry, 1994). It has also been demonstrated that the brain's response time for recognizing scenes in line drawing were comparable to those from color photographs (Lowe et al., 2018). Understanding how people recognize scenes from line drawings can help unravel the roles of edge and surface features in scene recognition and how the human brain represents scenes.

The role of edge and surface information in scene recognition remains controversial. Some researchers suggest that the edge information preserved in line drawings plays a primary role in scene categorization. They have demonstrated that the structural information retained in line drawings elicited similar neural activity in the Parahippocampal Place Area (PPA) and Retrosplenial Complex (RSC) as that elicited by color photographs (Walther et al., 2011). Moreover, it has been found that edge information receives priority processing in natural scene categorization (Fu et al., 2016). Conversely, others argue that surface information, such as texture and color, also plays a crucial role in scene recognition. For example, accuracy was reduced when scenes lacked color or had inconsistent colors (Oliva & Schyns, 2000;

Rousselet et al., 2005), and the PPA has been demonstrated to be sensitive to surface information of scenes (Lowe et al., 2017; Park & Park, 2017).

We noticed that these seemingly opposite suggestions and findings might be due to the role of edge and surface information varying with the level of abstraction. The basic level representation corresponds to the most common categories (e.g., forests, offices), whereas the superordinate level representation corresponds to more abstract categories (e.g., natural vs manmade scenes). Some authors pointed out that while horizontal and vertical lines dominate the structure of manmade scenes, natural scenes exhibit more undulating lines (Oliva & Torralba, 2001). They also observed that certain basic scene categories such as city streets and highways share similar structural shapes but differ in surface information such as city streets having a higher number of elements tend to have rougher textures and richer colors, while highways with fewer elements tend to have smoother surface and more uniform colors. Additionally, some studies have found that color plays a diagnostic role in certain basic categories, such as the overall green tones in forests and the blue waters and golden sands in beaches (Goffaux et al., 2005; Oliva & Schyns, 2000; Oliva & Torralba, 2006; Vailaya et al., 1998). Therefore, we assume that edge, color, and other surface information would play different roles in scene representation at the basic level of category and superordinate level of naturalness.

To test this, we combined EEG and MVPA methods in a one-back detection task, in which participants were asked to press a key as quickly and accurately as possible when a scene image was repeated consecutively. Compared to traditional univariate

analysis that relies on signal averages, the MVPA method analyzes the whole brain activation data and enhances the sensitivity to detect differences between different conditions (de-Wit et al., 2016; Grootswagers et al., 2017). To examine the roles of edge, color, and other surface information in scene representation at the basic category level and the superordinate naturalness level, each scene was presented in three image versions: color photographs, grayscale images, and line drawings. The significant difference in neural responses between color photographs and grayscale images would indicate the role of color information in scene representation, while the significant difference between grayscale images and line drawings would indicate the role of the other surface information such as texture and luminance in scene representation. However, as both grayscale images and line drawings share edge information, the similarity between grayscale images and line drawings would reflect the role of edge information in scene representation. Our results provided novel neural evidence for that edge information in line drawings is sufficient and more effective for neural scene representation at the superordinate naturalness level, whereas surface information is involved exclusively in initial neural scene representation at the basic category level.

2. Methods

2.1. Participants.

Twenty university students (11 female, mean age = 23.2, SD = 2.1) voluntarily participated in the study, following sample sizes from previous research (Groen et al., 2013; Harel et al., 2016; Lowe et al., 2018; Yao et al., 2023). All participants had

normal or corrected-to-normal vision. Informed consent was obtained from all participants, who were financially compensated. The experiment received approval from the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences.

2.2. Stimuli and experimental design

The study used 108 scene images (see Fig. 1a), comprising 36 different scenes each presented in color photographs, grayscale images, and line drawings. The color photographs and line drawings were adopted from a previous study (Walther et al., 2011), while the grayscale images were created by calculating the weighted average of the values from the three RGB channels of the color photographs using MATLAB (www.mathworks.com/). Images were divided into six basic level categories (beaches, city streets, forests, highways, mountains, and offices) and two superordinate level categories (natural, manmade). Stimuli were presented using Psychtoolbox 3 (Brainard, 1997; Kleiner, 2010) for MATLAB. The size of each image was 800 × 600, subtending a visual angle of 5.00 × 3.75 degrees at a viewing distance of 60 cm.

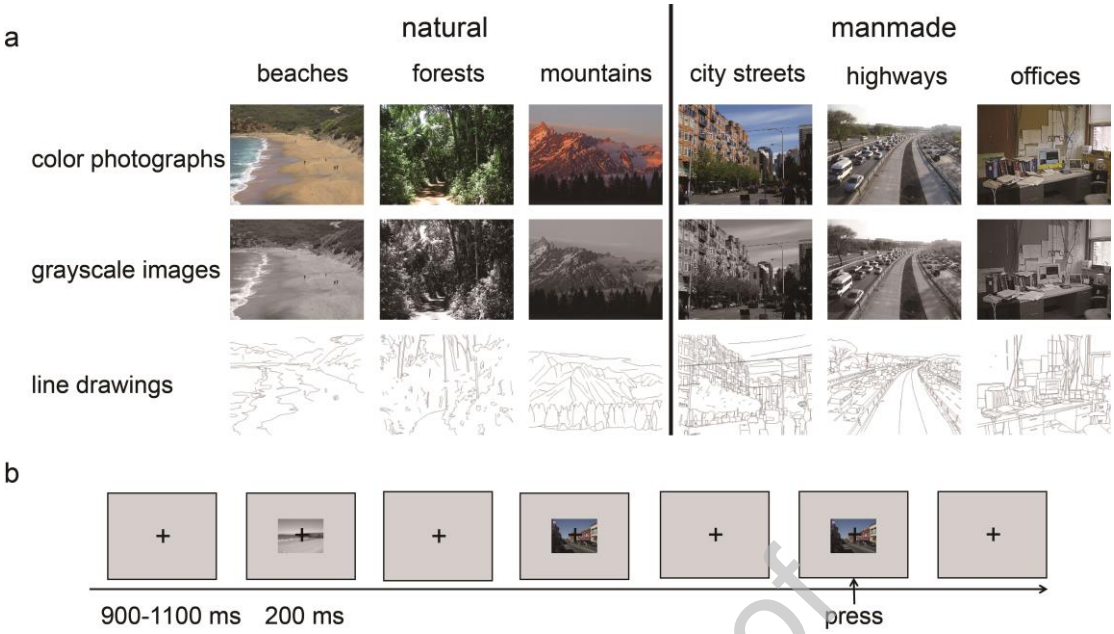

Fig. 1. Stimulus examples and experimental procedure. (a) Stimulus examples. The stimuli consisted of six basic level categories and two superordinate naturalness level categories. Each scene was presented in color photographs, grayscale images, and line drawings. (b) Experimental procedure. Each stimulus was presented for 200 ms following the central fixation cross displayed for 900-1100 ms. Participants were asked to pay attention to the central fixation cross and to press a button when an image appeared twice consecutively.

At the beginning of each block, the first stimulus was presented in the middle of the screen for 200 ms following a central fixation cross appeared for 900-1100 ms (see Fig. 1b). Participants were instructed to maintain their attention to the fixation cross. To ensure their attention on the stimuli, participants were required to perform a one-back working memory task, which required a key press as quickly and as accurately as possible when an image was repeated consecutively. Regardless of whether a response was made, the stimulus in the next trial would start. There were 135 trials in each block, in which all 108 images were presented in a randomized

order, and 27 images were randomly selected to appear consecutively. The repeated trials were not included in the EEG data analysis. The experiment consisted of 30 blocks, for a total of 4050 trials.

2.3. EEG acquisition and preprocessing.

The EEG data were recorded with a Neuroscan system using 64 electrodes. The left mastoid was served as the online reference and the right mastoid as the offline reference. The EEG signals were sampled at 1000 Hz. MATLAB and the EEGLAB toolbox (Delorme & Makeig, 2004) were used for offline preprocessing. The data were filtered between 0.1 and 30 Hz. Each trial extracted segments from 100 ms before the stimulus onset to 800 ms after the stimulus onset, and baseline corrected using the 100 ms prior to the stimulus onset. Independent Component Analysis (ICA) was conducted to identify and remove stereotypical artifacts such as eye blink artifacts. The trials with excessive artifacts (peak-to-peak deflection exceeding ± 100 μV) and incorrect responses were excluded. Data were downsampled to 200 Hz to reduce computational time and improve the signal-to-noise ratio (Grootswagers et al., 2017; Teichmann et al., 2020).

2.4. EEG analysis

Time-resolved decoding analysis. In all of the decoding analyses, patterns of brain activity from all electrodes at each time point were extracted for each participant. We used linear support vector machines (SVM; libsvm) (Chang & Lin, 2011) to train a classifier to distinguish conditions of interest, and then evaluated its ability to predict these conditions accurately in new data using independent test sets.

We conducted training and testing at each time point, with the aim of testing the classifier's ability to predict these conditions at any time point (i.e., "classification for interpretation") rather than achieving the highest decoding accuracy possible (Hebart & Baker, 2018). A 10 ms time window and a step size of 5 ms were used in the classification analysis. Decoding accuracy was significantly higher than chance level, indicating that the EEG data contained information relevant to the categories. We performed all our analyses using the CoSMoMVPA toolbox (Oosterhof et al., 2016).

For the individual image representation, there were 36 pictures for each version: color photographs, grayscale images, and line drawings. Each picture was paired with every other picture in the set, resulting in a total of 630 pairs (36 × 35/2 pairs). The classification accuracy of paired cross-validation was taken as the measure of similarity for each pair. The analysis was performed on each participant in a time-resolved manner. Initially, we divided all trials of each image into ten groups, in which nine groups were randomly selected as training sets and one group designated as the test set (i.e., ten-fold cross-validation). Subsequently, binary classification was conducted on all 630 pairs, and the classification process was repeated 100 times. The average over 100 times of decoding accuracy was taken as the value for the 36×36 decoding matrix, termed the Representational Dissimilarity Matrix (RDM). This matrix is symmetric, with the diagonal undefined. An RDM was required for each participant and each time point.

For basic category level and superordinate naturalness level scene representation, we trained classifiers to distinguish six different scene categories and between natural

versus manmade scenes for each of the three image versions, respectively. We used ten-fold cross-validation and averaged the prediction results across 100 repetitions to show how the basic and superordinate level representation of scenes evolved over time.

Spatial-resolved decoding analysis. To evaluate which electrodes contributed significantly to the neural decoding of EEG signals elicited by color photographs, grayscale images, and line drawings, we conducted decoding analyses on each electrode separately. A 50 ms time window with a 50 ms step size was used, covering 0 to 500 ms after stimulus onset, resulting in 10 time bins (Nemrodov et al., 2016; Smith & Smith, 2019). Classifiers were trained to distinguish images from one another at the individual image level, the basic level, and the superordinate level.

Time generalization analysis. To examine whether the three types of images elicit similar neural activity, the time generalization analysis was conducted following the method used in previous research (Yao et al., 2023). We trained a classifier on EEG signals at specific time points elicited by grayscale images and tested it with neural activity at all time points induced by color photographs or line drawings of the same scene. This process generated two 900 × 900 matrices (-100–800 ms to the stimulus onset), capturing the classifier generalization performance. If a classifier trained at one time point can predict the test data at other time points, it would suggest that the test and training sets exhibit similar neural activity patterns at these time points. The matrix's diagonal represented standard time-resolved decoding, while

decodable off-diagonal effects indicated temporal asynchrony of information processing between training and testing sets.

Representational similarity analysis (RSA). To explore the contribution of spatial frequency and other image attributes to scene representation, we conducted RSA by constructing RDMs for the basic level of category, superordinate level of naturalness, spatial frequency of grayscale images, and spatial frequency of line drawings. For the basic category and the superordinate naturalness RDMs, values were coded as 0 for within-category comparisons, and 1 for between-category comparisons. Spatial frequency RDMs were generated by first calculating the spatial frequency of each image using gradient-based spatial frequency analysis, followed by computing the Euclidean distances between the spatial frequency values of all image pairs. The spatial frequency RDM for grayscale images captured frequency differences both within color photographs and within grayscale images, as the spatial frequency content of color photographs was identical to that of grayscale images. In contrast, the spatial frequency RDM for line drawing captured frequency differences exclusively among line drawings. Finally, we conducted Spearman correlation analysis between each of the four RDMs and the time-resolved decoding RDM at the individual image level to assess the relationship between image attributes and neural representations.

Statistical analysis. To evaluate whether the analysis results of the EEG data were significantly above chance, we utilized threshold-free cluster enhancement (TFCE) (Smith & Nichols, 2009) implemented in the CosMoMVPA toolbox. First, we shuffled the trial labels for permutation testing, reanalyzing the data using these

shuffled labels. This process was repeated 100 times for each participant, thereby creating a null distribution for each participant. Subsequently, a group-level null distribution was constructed using Monte Carlo sampling, comprising 1000 permutations with shuffled labels (Stelzer et al., 2013). This method utilized clustering for multiple comparison correction. TFCE created an empirical distribution of maximum cluster sizes without applying a specific threshold, and the 95th percentile of this distribution was used as the significance threshold (equivalent to $p < 0.05$, one-tailed).

Bootstrap tests were conducted to assess whether there were significant differences on the onset latency (i.e., the first time point following stimulus onset where decoding accuracy significantly surpassed the chance level), peak latency (i.e., the time point at which the maximum accuracy was reached), and peak value (i.e., the highest accuracy) of scene representation for different versions of images, as well as the training / testing times for maximum decoding accuracy in time generalization. Decoding accuracy from participants' time-resolved decoding analysis and time generalization analysis was bootstrapped 10,000 times, resulting in empirical distributions for onset latency, peak latency, and peak value to establish 95% confidence intervals (CIs). To compare the differences between different image types in onset latency, peak latency, and peak value, as well as the differences in training and testing times corresponding to the maximum decoding accuracy in time generalization analysis, we computed the differences between these 10,000 bootstrapped samples for each dataset. This process produced an empirical

distribution of differences. A p-value was defined as the proportion of samples with a difference greater than or less than 0 divided by the number of permutations (i.e., a two-tailed test), corrected with false discovery rate (FDR). A significant difference was indicated if the 95% CI did not include 0.

To evaluate the statistical significance of the correlations between different RDMs, we performed a permutation test. First, we calculated the Spearman correlation coefficient (*R*) for the original image attribute RDMs, then randomly shuffled the image labels and recalculated the Spearman's *R*. This procedure was repeated 10,000 times to obtain an empirical null distribution. The p-value was defined as the number of values in the null distribution with an absolute value greater than the absolute value of the true correlation, divided by the total number of permutations (i.e., a two-tailed test), corrected with false discovery rate (FDR).

3. Results

3.2. Behaviroal results

The mean proportion accuracy of the one-back repetition detection task was 0.956 (*SD* = 0.032) for color photographs, 0.918 (*SD* = 0.048) for grayscale images, and 0.933 (*SD* = 0.042) for line drawings. Repetition trials were excluded from the subsequent EEG analysis.

3.2. EEG decoding results

We first decoded the EEG signals elicited by color photographs, grayscale images, and line drawings at the individual image level as in previous studies (Cichy et al., 2014; Dobs et al., 2019; Singer et al., 2023; Yao et al., 2023). Since both basic

category level and superordinate naturalness level scene representations can contribute to the decoding at the individual image level, we then separately decoded the EEG signals elicited by the three types of images at the basic and superordinate levels. The time-resolved decoding analysis, the spatial decoding analysis, and the time generalization analysis were conducted at each level of analysis.

3.2.1. Decoding results at the individual image level

To investigate the role of edge, color, and other surface information in scene representation at the individual image level, we trained a classifier to distinguish each individual scene image from all other images based on neural activity elicited by color photographs, grayscale images, and line drawings.

Fig. 2a shows the results of the time-resolved decoding. As can be seen from the figure, all types of individual scene images were decoded significantly above the chance level (0.50). The decoding accuracy for color photographs reached significance at 75 ms, and peaked at 130 ms with a value of 0.597; grayscale images reached significance at 75 ms, and peaked at 120 ms with a value of 0.598; line drawings reached significance at 105 ms, and peaked at 145 ms with a value of 0.56.

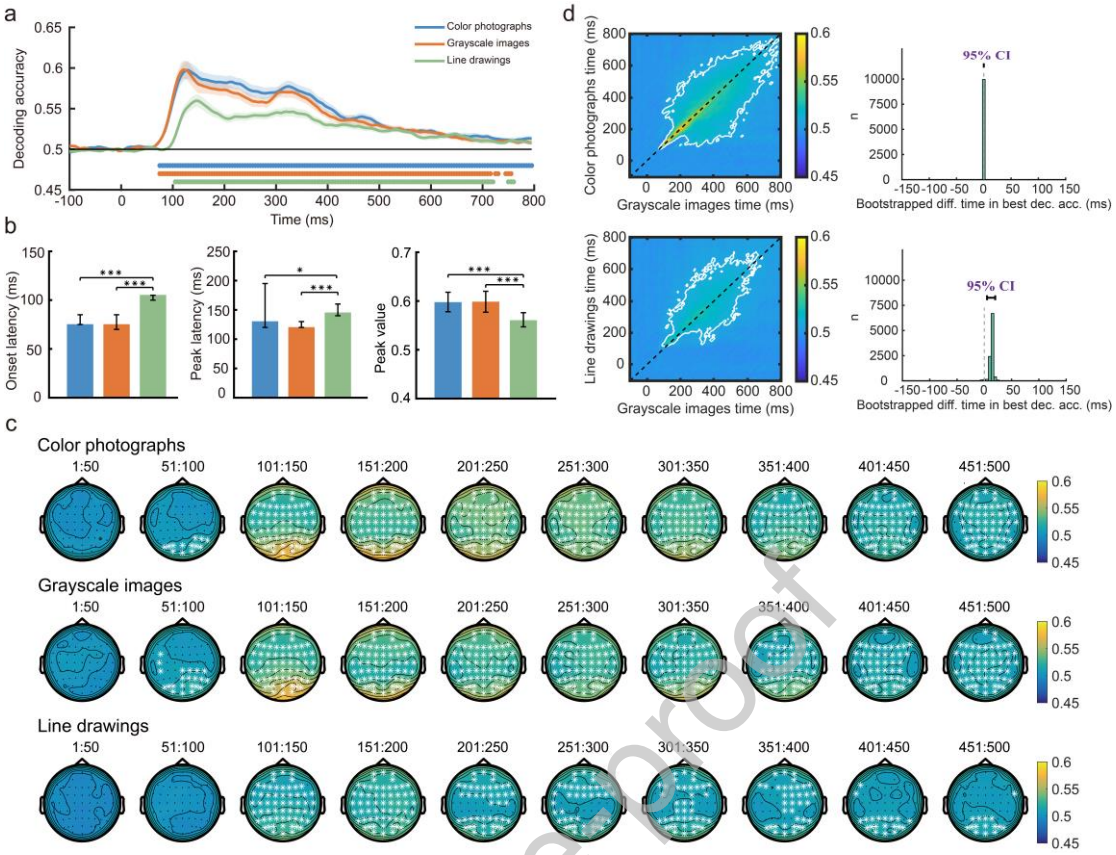

Fig. 2. Results of the time-resolved decoding for three versions of images at the individual image level. (a) Time course of decoding accuracy for color photographs, grayscale images, and line drawings, with points below indicating time points where decoding accuracy was significantly above chance ($p < 0.05$, threshold-free cluster enhancement (TFCE) corrected). The shaded areas represent one standard error (SE) about the means. (b) Onset latency, peak latency, and peak value of the time-course decoding for three versions of images. Error bars depict bootstrapped 95 % confidence intervals (CIs). *$p < 0.05$, ***$p < 0.001$ (false discovery rate (FDR) corrected). (c) Spatial resolved decoding of color photographs, grayscale images, and line drawings. Each map depicts whether decoding accuracy at each electrode exceeded chance level within a given time window. White stars indicate electrodes with decoding accuracy significantly above chance ($p < .05$, TFCE corrected). (d) Results of the time generalization analysis for the individual image. Grayscale images were used as the training set, color photographs or line drawings as the testing

sets. The left figures show the decoding accuracy for each time point combination, the white outline indicates significant clusters ($p < 0.05$, TFCE corrected), and the right figures show the bootstrap differences between the training set time and testing set time corresponding to the best decoding accuracy.

Fig. 2b shows the onset latency, peak latency, and peak value of the time course decoding at the individual image level. The bootstrap test was used to compare the onset latency, peak latency, and peak value among three versions of images. It revealed that, for both onset latency and peak latency, color photographs (onset latency $p < 0.001$, peak latency $p = 0.034$, FDR corrected) and grayscale images (all $ps < 0.001$, FDR corrected) were earlier than line drawings. For the peak value, color photographs ($p < 0.001$, FDR corrected) and grayscale images ($p < 0.001$, FDR corrected) were higher than line drawings, while there was no significant difference between color and grayscale images. These results suggested that the other surface information such as texture, rather than color, played an important role in the speed and decoding accuracy of scene representation at the individual image level.

To examine which electrodes significantly contributed to decoding the individual image scene for the three image types, we conducted independent analyses for each electrode using a time window of 50 ms (see Fig. 2c). The results showed that for color photographs and grayscale images, significant decoding electrodes appeared from 51-100 ms, and primarily distributed in the posterior brain regions. In the subsequent time window of 101-150 ms, electrodes across the entire brain showed significant contributions to this decoding. For line drawings, significant decoding

electrodes appeared from 101-150 ms throughout the entire brain, and from 201-250 ms, the number of significant electrodes began to decrease.

To examine whether similar neural activity was elicited by the three image versions, the time generalization analysis was conducted. Fig. 2d shows the time generalization results for the individual image. As in the previous research (Yao et al., 2023), when a classifier was trained with neural activity by grayscale images and tested with those by color photographs or line drawings, the significant decoding accuracy at each time point combination indicated that the training and testing sets elicited similar neural activity at that time point. The results revealed that when the classifier was trained with grayscale images and tested with color photographs, significant decoding accuracy was observed over a large time window, and no significant shift from the diagonal of the matrix was observed in the best decoding accuracy. However, when the classifier was trained with grayscale images and tested with line drawings, significant decoding accuracy was still observed across a wide time window, but the best decoding accuracy showed a 15 ms shift from the matrix's diagonal. These results indicated that at the individual image level, color photographs, grayscale images, and line drawings elicited similar neural activity, but the lack of surface information such as texture other than color resulted in a slower neural representation speed.

3.2.2. Decoding results at the basic category level

To investigate the role of edge, color, and other surface information in basic level scene representation, we trained a classifier to distinguish six basic categories based on neural activity elicited by color photographs, grayscale images, and line drawings.

Fig. 3a shows the results of the time-resolved decoding. Decoding accuracy for color photographs reached significance at 95 ms, and peaked at 210 ms with a value of 0.244; grayscale images reached this at 85 ms, and peaked at 130 ms with a value of 0.234; and line drawings reached this at 105 ms, and peaked at 145 ms with a value of 0.230.

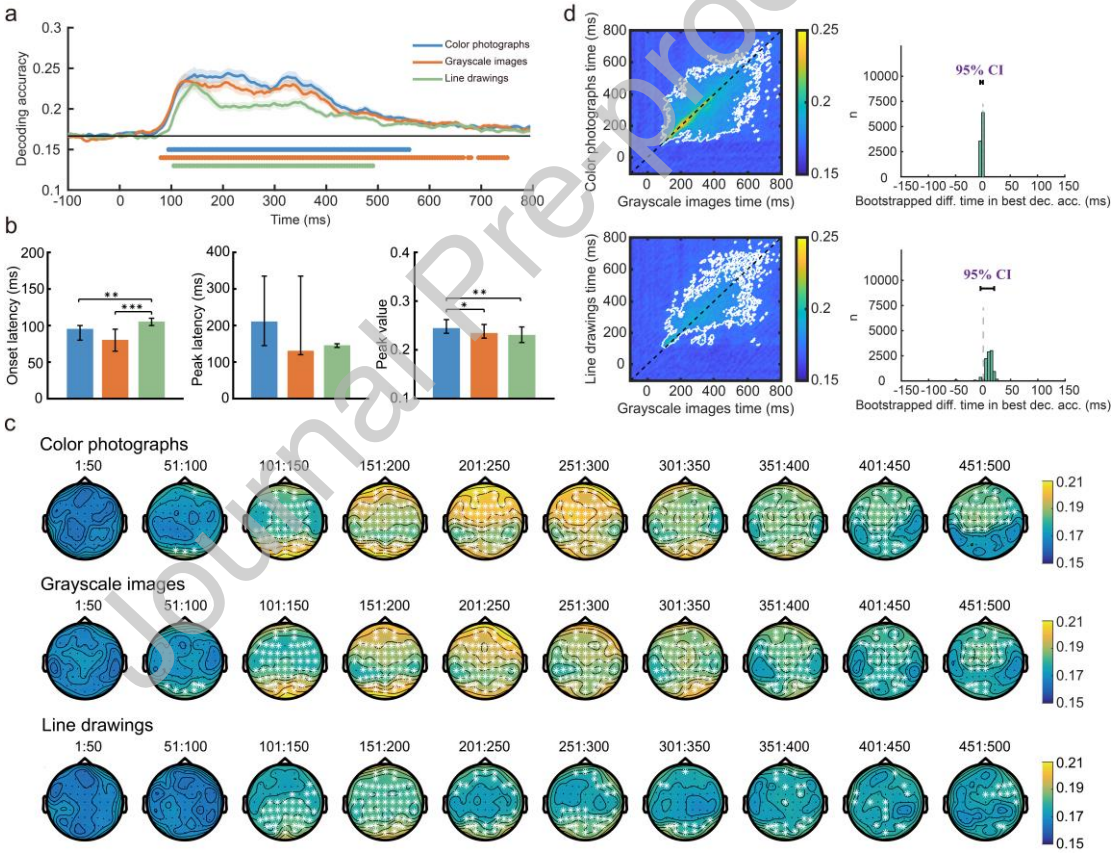

Fig. 3. Results of the time-resolved decoding for three versions of images at the basic level representations. (a) Time course of decoding accuracy for color photographs, grayscale images, and line drawings, with points below indicating time points where decoding accuracy was significantly above chance ($p < 0.05$, TFCE, corrected). The shaded areas represent one SE about

the means. (b) Onset latency, peak latency, and peak value of the time course decoding for three versions of images. Error bars depict bootstrapped 95 % CIs. *p < 0.05, **p < 0.01,***p < 0.001 (FDR corrected). (c) Spatial-resolved decoding of color photographs, grayscale images, and line drawings at the basic level representations. Each map depicts whether decoding at each electrode exceeded chance level within a given time window. White stars indicate electrodes with decoding accuracy significantly above chance (p < .05, TFCE corrected). (d) Results of the time generalization analysis for the basic level. Grayscale images were used as the training set, color photographs or line drawings were used as the testing sets, respectively. The left figures show the decoding accuracy for each time point combination, the white outline indicates significant clusters (p < 0.05, TFCE corrected), and the right figures show the bootstrap differences between the training set time and testing set time corresponding to the best decoding accuracy.

Fig. 3b shows the onset latency, peak latency, and peak value of the time course decoding for different image versions. For onset latency, color photographs (p = 0.002, FDR corrected) and grayscale images (p < 0.001, FDR corrected) were significantly earlier than line drawings, indicating that surface information other than color influenced the basic level representation. For peak value, color photographs were significantly higher than both grayscale images (p = 0.014, FDR corrected) and line drawings (p = 0.002, FDR corrected), suggesting that color information affected the decoding accuracy of basic level representations. There were no differences on peak latency among the three image versions.

To further investigate the role of color information in representing each of six basic categories, we trained a classifier to differentiate neural activities induced by

color photographs and grayscale images within the same category (See Supplementary Fig. 1). The results revealed significant decoding points for beaches, forests, city streets, and offices, but not for highways and mountains. Additionally, a bootstrap test on the peak values of these four scene categories showed that the peak value for beaches was significantly higher than those for forests, city streets, and offices. These results suggest that the role of color in neural representation varies across categories, with the most pronounced effect observed for beaches.

Fig. 3c shows the electrodes significantly contributed to the decoding of basic level representation. For color photographs and grayscale images, significant decoding electrodes appeared from 51-100 ms, with widespread significant electrodes across the entire brain observed in the 101-150 ms time window. However, for line drawings, significant decoding electrodes appeared from 101-150 ms over the posterior brain regions, became widespread across the entire brain in the 151-200 ms window before decreasing.

The time-generalization analysis revealed that when a classifier used neural activity elicited by grayscale images as the training set, significant decoding accuracy was observed over a large time window regardless of whether the neural activity by color photographs or line drawings were used as the testing set. Additionally, the best decoding accuracy did not shift from the diagonal of the matrix (see Fig. 3d). The results suggested that the three types of images elicited similar neural activity at the basic level, and other surface information did not affect the speed of similar neural representation.

3.2.3. Decoding results at the superordinate level of naturalness

To investigate the role of edge, color, and other surface information in the superordinate level representation of naturalness in scenes, we trained a classifier to distinguish between natural and manmade scenes based on neural activity elicited by color photographs, grayscale images, and line drawings.

Fig. 4a shows the results of the time-resolved decoding. Color photographs were significantly decoded above chance level at 110 ms, and peaked at 275 ms with a value of 0.564. Grayscale images could be significantly decoded at 105 ms, and peaked at 305 ms with a value of 0.56. Line drawings were significantly decoded at 100 ms, and reached a peak at 135 ms with a value of 0.582.

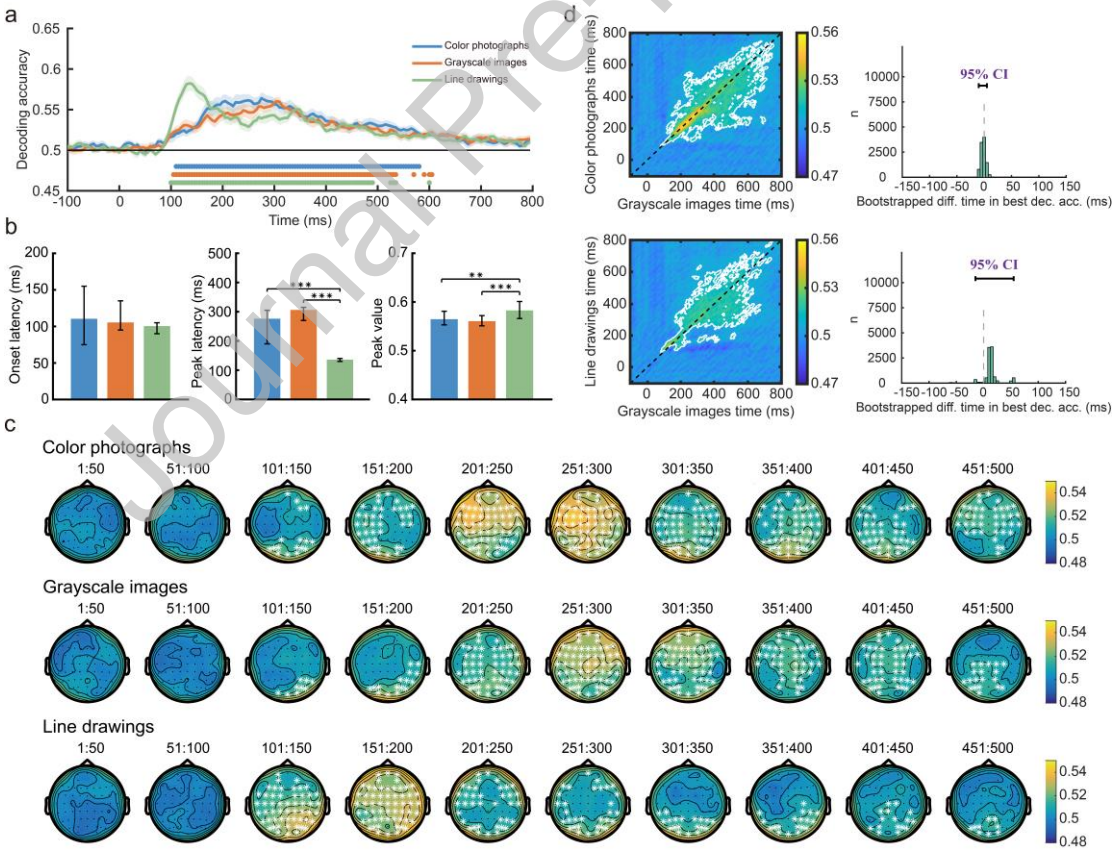

Fig. 4. Results of the time-resolved decoding for three versions of images at superordinate level naturalness representations. (a) The time course decoding accuracy of color photographs,

grayscale images, and line drawings with points below indicating time points where decoding accuracy was significantly above chance (*ps* < 0.05, TFCE corrected). The shaded areas represent one SE about the means. (b) Onset latency, peak latency, and peak value of the time course decoding for the three versions of images. Error bars depict bootstrapped 95 % CIs. *p < 0.05, **p < 0.01,***p < 0.001 (FDR corrected). (c) Spatial-resolved decoding of color photographs, grayscale images, and line drawings at the superordinate naturalness level representations. Each map depicts whether decoding at each electrode exceeded chance level within a given time window. White stars indicate electrodes with decoding accuracy significantly above chance (*ps* < .05, TFCE corrected). (d) Results of the time generalization analysis for the superordinate naturalness level. Grayscale images were used as the training set, and color photographs and line drawings were used as the testing sets, respectively. The left figures show the decoding accuracy for each time point combination, the white outline indicates significant clusters (*ps* < 0.05, TFCE corrected), and the right figures show the bootstrap differences between the training set time and testing set time corresponding to the best decoding accuracy.

Fig. 4b shows the onset latency, peak latency, and peak value of the time course decoding for three image versions. There were no significant differences on onset latency among the three image versions. The peak latency was significantly earlier for line drawings than for color photographs (p < 0.001, FDR corrected), and grayscale images (p < 0.001, FDR corrected). The peak value was also higher for line drawings than for color photographs (p = 0.008, FDR corrected) and for grayscale images (p < 0.001, FDR corrected).

Fig. 4c shows the electrodes significantly contributed to the decoding of superordinate naturalness level representation. For line drawings, significant decoding electrodes appeared from 101-150 ms across the entire brain, while for color photographs and grayscale images, only limited number of electrodes were significantly decodable during this time window. Moreover, the decoding peak for line drawings corresponded to the time window of 101-150 ms, with higher decoding accuracy at the posterior electrodes. However, the decoding peaks for color photographs and grayscale images corresponded to the time window of 251-300 ms and 301-350 ms separately, with higher decoding accuracy at the anterior electrodes.

The results of time-generalization analysis showed that when a classifier used neural activity elicited by grayscale images as the training set, significant decoding accuracy was observed over a large time window, regardless of whether neural activity by color photographs or line drawings were used as the testing set. Additionally, the best decoding accuracy did not shift from the diagonal of the matrix (see Fig. 4d). The results suggested that the three types of images elicited similar neural activity at the superordinate naturalness level, and surface information did not affect the neural representation speed.

3.2.4. RSA results of the effect of spatial frequency in scene representation

To explore the contribution of spatial frequency and other image attributes to scene representation, we performed the RSA. The results, shown in Fig. 5, revealed that the spatial frequency RDM for both grayscale images (Spearman's $R = 0.195$, $p < 0.001$) and line drawings (Spearman's $R = 0.206$, $p < 0.001$) were significantly

correlated with the basic category RDM. These results suggest that spatial frequency differences in both grayscale images and line drawings are associated with different categories at the basic level. Interestingly, only the spatial frequency RDM for line drawings (Spearman's $R = 0.128$, $p < 0.01$) showed a significant correlation with the superordinate naturalness level RDM. This indicates that variations in naturalness at the superordinate level are primarily driven by spatial frequency differences in line drawings.

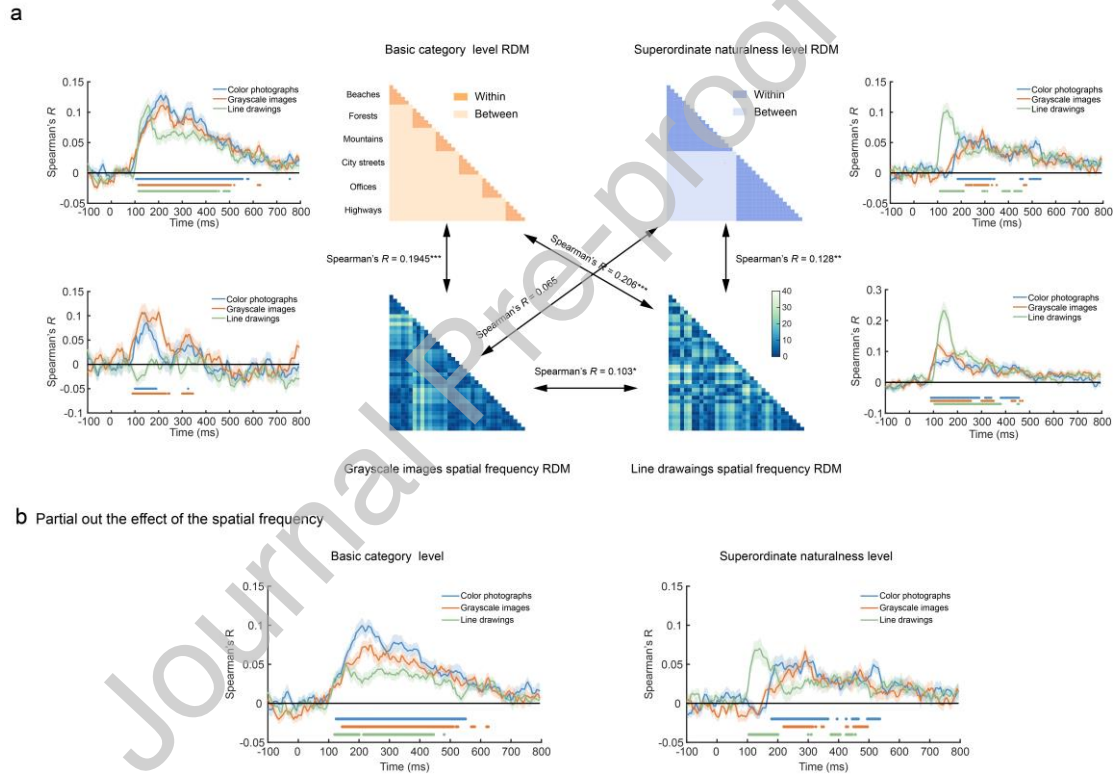

Fig. 5. Results of the representational similarity analysis for different image attributes. (a) The central panel displays the RDMs for the basic category level, superordinate naturalness level, spatial frequency of grayscale images, and spatial frequency of line drawings (with 1 corresponding to between-category and 0 corresponding to within-category comparisons). *$p <$ 0.05, **$p < 0.01$, ***$p < 0.001$. The panels on the left and right depict the temporal dynamics of the Spearman correlation coefficients between each of the four RDMs and the decoding RDM at

the individual level. (b) The temporal dynamics of the Spearman correlation coefficients at the basic category level and superordinate naturalness level, after controlling for the spatial frequency RDM. The markers below the curves indicate the time points where decoding accuracy was significantly above chance ($p < 0.05$, threshold-free cluster enhancement corrected). Shaded areas represent one SE about the mean.

To further explore the contribution of various image attributes to scene representation, we conducted Spearman correlation analysis between each of the four RDMs and the time-resolved decoding RDMs at the individual image level. For the basic category level RDM, significant correlations were observed between 105-580 ms for color photographs, 115-520 ms for grayscale images, and 115-500 ms for line drawings. In contrast, for the superordinate naturalness level RDM, significant correlations were observed between 185-340 ms and 450-535 ms for color photographs, 220-330 ms for grayscale images, 110-210 ms and 375-455 ms for line drawings. These results suggest that the correlation onset latency occurred earlier for color photographs than for grayscale images or line drawings at the basic category level, but onset latency was earlier for line drawings than for grayscale images and color photographs at the superordinate naturalness level, indicating potential differences in processing hierarchies.

Additionally, for the spatial frequency RDM of grayscale images, significant correlations were observed between 100-190 ms for color photographs and 90-240 ms and 300-345 ms for grayscale images, but no significant correlation was found for line-drawings. In contrast, the spatial frequency RDM of line drawings exhibited

significant correlations between 90-455 ms for color photographs, 90-470 ms for grayscale images, and 90-380 ms for line drawings. These results suggest that the spatial frequency of grayscale images is not associated with neural activity induced by line drawings, while the spatial frequency of line drawings is related to neural activity induced by both color photographs and grayscale images.

Finally, to investigate whether the higher decoding accuracy for line drawings at the superordinate naturalness level representations is driven by abstract scene information, or by low-level features such as spatial frequency, we conducted partial correlation analyses. Specifically, we controlled for spatial frequency RDM by partialling it out from both the basic category and superordinate naturalness level RDMs, as well as from the time-resolved decoding RDM at the individual image level. The results (Fig. 5b) showed that significant correlations between decoding RDM and both the basic category and the superordinate naturalness level RDMs persisted even after controlling for the influence of spatial frequency. Specifically, at the basic category level, significant correlations were observed between 125-550 ms for color photographs, 145-525 ms for grayscale images, and 120-445 ms for line drawings. At the superordinate naturalness level, significant correlations were found between 180-365 ms, 445-465 ms, and 495-535 ms for color photographs; 220-315 ms and 450-495 ms for grayscale images; and 105-200 ms and 375-445 ms for line drawings. These results suggest that the brain's higher decoding of line drawings at the superordinate naturalness level is not solely attributable to low-level features such

as spatial frequency, but genuinely reflects the representation of abstract scene information.

4. Discussion

The current study combined the EEG technique and the MVPA methods to examine the role of edge, color, and other surface information in scene representation with different levels of abstraction. The decoding results showed that the peak accuracy was significantly higher for color photographs than for grayscale images and line drawings at the basic category level, but significantly lower for color photographs and grayscale images than for line drawings at the superordinate naturalness level. Consistently, the onset latency was significantly faster for color photographs and grayscale images than for line drawings at the basic category level, but not at the superordinate naturalness level. These results provide neural evidence for how edge, color, and other surface information play distinct roles in basic category level and superordinate naturalness level scene representations and provide insights into how the human brain represents scenes over time.

Our results revealed that it was other surface information, rather than color information, that plays an important role in scene representation at the individual image level and at the basic category level. The results showed that the decoding onset latency was around 100 ms for all three versions of images, consistent with previous findings that scene-related neural activity can be detected in EEG around 100 ms (Cichy et al., 2017; Lowe et al., 2018; Orima & Motoyoshi, 2023). Importantly, we found that the decoding onset latency was significantly earlier for

grayscale images than for line drawings at both the individual image level and the basic level. This finding indicates that other surface information receives early processing and plays a role in initial scene representation at these levels. Previous studies found that early basic level scene classification is more similar to texture models (Renninger & Malik, 2004). Our findings extend this by showing that other surface information contributes to initial scene representation at both the individual image level and the basic category level.

Consistently, we found that for the decoding peak latency, grayscale images were significantly earlier than line drawings at the individual image level, but there were no significant differences at the basic category level. The difference on the decoding peak latency among the three image versions at the individual image level might be due to the additional processing of line drawings, which involves filling-in of surface information that is not required for the other two versions of images. However, this salient fill-in process might not occur at the basic category level, as the representation at this level is much more abstract than the individual image level.

Moreover, our results revealed that the decoding peak value was significantly higher for color photographs than for grayscale images at the basic category level. Specifically, we found that color influences the neural representation of beaches, forests, city streets, and offices, with its effect being most pronounced for beaches. These results indicate that color information contributes to the basic category level representation. This might be because color can serve as a diagnostic feature for basic level scene categories, such as blue for beaches and green for forests (Oliva &

Torralba, 2006), but not at the individual image level, where scenes within the same category usually share similar colors.

Importantly, our decoding results showed that edge information depicted in line drawings plays a crucial role in scene representation at the superordinate naturalness level. At this level, line drawings not only exhibited an earlier peak latency but also a higher peak value than color photographs and grayscale images. The results indicate that edge information in line drawings is both sufficient and more effective than in color photographs and grayscale images at the superordinate naturalness level. The peak latency for line drawings was at 135 ms, with higher decoding accuracy at posterior electrodes, corresponding to the feature extraction and perception stages of scene recognition. Prior ERP research also showed that the differences between manmade and natural scenes were detected earlier by ERP responses to line drawings than to color photographs (Lowe et al., 2018).

However, manmade scenes tend to contain more lines and, consequently, exhibit much higher spatial frequency compared to natural scenes. This raises concerns about whether the early peak observed in the time-resolved decoding reflects the abstract scene information necessary for distinguishing between manmade and natural environments. To address this, we conducted a representational similarity analysis. The results revealed that spatial frequency differences in both grayscale images and line-drawings were associated with categories distinctions at the basic level. In contrast, variations in naturalness at the superordinate level were primarily linked to the spatial frequency differences in line drawings. Moreover, the spatial frequency of

grayscale images was not associated with neural activity induced by line drawings, whereas the spatial frequency of line drawings was related to neural activity induced by color photographs and grayscale image, suggesting distinct neural mechanisms.

To further investigate the role of abstract scene information, we compared the time-resolved decoding RDM at the individual image level with different model RDMs representing the basic category level or superordinate naturalness level, while controlling for the influence of spatial frequency. The results revealed that line drawings continued to exhibit earlier and stronger correlations with superordinate naturalness representations, even after partialling out the influence of spatial frequency. These findings suggest that the brain's more accurate and robust decoding of line drawings at the superordinate naturalness level is not solely driven by low-level visual features but genuinely reflects the representation of abstract scene information.

Our time generalization analysis results revealed that the neural activity patterns elicited by line drawings were similar to those evoked by color photographs or grayscale images. In this analysis, when grayscale images were used as the training set and color images or line drawings as the testing sets, the neural activity patterns elicited by grayscale images could be generalized to those elicited by either color photographs or line drawings at each level of abstraction. These results provided direct neural evidence that the neural activities elicited by the three versions were similar. Consistently, previous research found that color photographs and line drawings of scenes at the basic category level elicited similar patterns of neural

activation, especially in the Parahippocampal Place Area (PPA) (Morgan et al., 2019; O'Connell et al., 2018; Walther et al., 2011), which is primarily associated with the spatial layout of scenes (Bilalic et al., 2019; Chaisilprungraung & Park, 2021; Park et al., 2011). Our previous research also found that color photographs, grayscale images, and line drawings of objects at the individual image level elicited similar neural activity (Yao et al., 2023). These findings indicate that line drawings preserve important spatial structure information of both scenes and objects, which plays a crucial role in scene and object recognition (Choo & Walther, 2016; O'Connell et al., 2018; Walther & Shen, 2014).

Interestingly, however, the time corresponding to the best decoding accuracy at the individual image level was significantly later (approximately 15 ms) for line drawings than for grayscale images when grayscale images were used as the training set and line drawings as the testing set, but not when the color photographs were used as the testing set. These findings suggest that the surface information such as textures other than colors can affect scene processing speed at the individual image level. However, there were no such differences at the basic category level and superordinate naturalness level. There may be two reasons for this. Firstly, neural representation at the individual image level is more specific than that at the basic category level or superordinate naturalness level, as it requires more detailed information to differentiate between two similar scenes, like two different forest scenes. That is, surface information is likely to be more important in the scene representation at a more specific level rather than at a more abstract level. Secondly, when visual scenes

lack surface information, top-down processing may fill the missing information by comparing it with stored memory of scenes (Morgan et al., 2019; Mudrik et al., 2014; Rahman & Sommer, 2008). This process could have slowed down the processing of line drawings in the current experiment.

5. Conclusion

To sum up, our findings provide electrophysiological evidence for the distinct roles of edge, color, and other surface information in scene representation across different levels of abstraction. First, our time decoding results demonstrate that edge information is sufficient and more effective at the superordinate naturalness level, whereas color and other surface information are exclusively involved in scene representation at the basic category level. Second, our time generalization results reveal that line drawings evoke neural activity patterns similar to those elicited by grayscale images across all levels of scene abstraction. These results highlight the crucial role of edge information in scene representation at all levels of abstraction and provide insights into the distinct contributions of edge, color and surface information to scene recognition. This work advances our understanding of how the human brain represents and processes scenes.

Acknowledgements

The research was funded by the National Key Research and Development Program of China (No. 2021ZD0204202) and the National Natural Science Foundation (32471112).

References

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive psychology*, *20*(1), 38-64. https://doi.org/10.1016/0010-0285(88)90024-2

Bilalic, M., Lindig, T., & Turella, L. (2019). Parsing rooms: the role of the PPA and RSC in perceiving object relations and spatial layout. *Brain Structure & Function*, *224*(7), 2505-2524. https://doi.org/10.1007/s00429-019-01901-0

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436. https://doi.org/10.1163/156856897x00357

Chaisilprungraung, T., & Park, S. (2021). "Scene" from inside: The representation of Observer's space in high-level visual cortex. *Neuropsychologia*, *161*, Article 108010. https://doi.org/10.1016/j.neuropsychologia.2021.108010

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for Support Vector Machines. *Acm Transactions on Intelligent Systems and Technology*, *2*(3). https://doi.org/10.1145/1961189.1961199

Choo, H., & Walther, D. B. (2016). Contour junctions underlie neural representations of scene categories in high-level human visual cortex. *Neuroimage*, *135*, 32-44. https://doi.org/10.1016/j.neuroimage.2016.04.021

Cichy, R. M., Khosla, A., Pantazis, D., & Oliva, A. (2017). Dynamics of scene representations in the human brain revealed by magnetoencephalography and

deep neural networks. *Neuroimage*, *153*, 346-358. https://doi.org/10.1016/j.neuroimage.2016.03.063

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455-462. https://doi.org/10.1038/nn.3635

de-Wit, L., Alexander, D., Ekroll, V., & Wagemans, J. (2016). Is neuroimaging measuring information in the brain? *Psychonomic Bulletin & Review*, *23*(5), 1415-1428. https://doi.org/10.3758/s13423-016-1002-0

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. *Nature Communications*, *10*, Article 1258. https://doi.org/10.1038/s41467-019-09239-1

Fu, Q. F., Liu, Y. J., Dienes, Z., Wu, J. H., Chen, W. F., & Fu, X. L. (2016). The role of edge-based and surface-based information in natural scene categorization: Evidence from behavior and event-related potentials. *Consciousness and Cognition*, *43*, 152-166. https://doi.org/10.1016/j.concog.2016.06.008

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition*, *12*(6), 878-

892. https://doi.org/10.1080/13506280444000562

Groen, I. I. A., Ghebreab, S., Prins, H., Lamme, V. A. F., & Scholte, H. S. (2013). From Image Statistics to Scene Gist: Evoked Neural Activity Reveals Transition from Low-Level Natural Image Structure to Scene Category. *Journal of Neuroscience*, *33*(48), 18814-18824. https://doi.org/10.1523/jneurosci.3128-13.2013

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677-697. https://doi.org/10.1162/jocn_a_01068

Harel, A., Groen, I. I. A., Kravitz, D. J., Deouell, L. Y., & Baker, C. I. (2016). The Temporal Dynamics of Scene Processing: A Multifaceted EEG Investigation. *eNeuro*, *3*(5). https://doi.org/10.1523/ENEURO.0139-16.2016

Hebart, M. N., & Baker, C. I. (2018). Deconstructing multivariate decoding for the study of brain function. *Neuroimage*, *180*, 4-18. https://doi.org/10.1016/j.neuroimage.2017.08.005

Hochberg, J., & Brooks, V. (1962). PICTORIAL RECOGNITION AS AN UNLEARNED ABILITY - A STUDY OF 1 CHILDS PERFORMANCE. *American Journal of Psychology*, *75*(4), 624-&. https://doi.org/10.2307/1420286

Kleiner, M. (2010). Visual stimulus timing precision in Psychtoolbox-3: Tests, pitfalls and solutions. *Perception*, *39*, 189-189.

https://doi.org/10.1177/03010066100390S101

Lowe, M. X., Rajsic, J., Ferber, S., & Walther, D. B. (2018). Discriminating scene categories from brain activity within 100 milliseconds. *Cortex*, *106*, 275-287. https://doi.org/10.1016/j.cortex.2018.06.006

Lowe, M. X., Rajsic, J., Gallivan, J. P., Ferber, S., & Cant, J. S. (2017). Neural representation of geometry and surface properties in object and scene perception. *Neuroimage*, *157*, 586-597. https://doi.org/10.1016/j.neuroimage.2017.06.043

Morgan, A. T., Petro, L. S., & Muckli, L. (2019). Scene Representations Conveyed by Cortical Feedback to Early Visual Cortex Can Be Described by Line Drawings. *Journal of Neuroscience*, *39*(47), 9410-9423. https://doi.org/10.1523/jneurosci.0852-19.2019

Mudrik, L., Shalgi, S., Lamy, D., & Deouell, L. Y. (2014). Synchronous contextual irregularities affect early scene processing: Replication and extension. *Neuropsychologia*, *56*, 447-458. https://doi.org/10.1016/j.neuropsychologia.2014.02.020

Nemrodov, D., Niemeier, M., Mok, J. N. Y., & Nestor, A. (2016). The time course of individual face recognition: A pattern analysis of ERP signals. *Neuroimage*, *132*, 469-476. https://doi.org/10.1016/j.neuroimage.2016.03.006

O'Connell, T. P., Sederberg, P. B., & Walther, D. B. (2018). Representational differences between line drawings and photographs of natural scenes: A dissociation between multi-voxel pattern analysis and repetition suppression.

Neuropsychologia, *117*, 513-519. https://doi.org/10.1016/j.neuropsychologia.2018.06.013

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive psychology*, *41*(2), 176-210. https://doi.org/10.1006/cogp.1999.0728

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*(3), 145-175. https://doi.org/10.1023/a:1011139631724

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: the role of global image features in recognition. In S. MartinezConde, S. L. Macknik, L. M. Martinez, J. M. Alonso, & P. U. Tse (Eds.), *Visual Perception, Pt 2: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception* (Vol. 155, pp. 23-36). https://doi.org/10.1016/s0079-6123(06)55002-2

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in neuroinformatics*, *10*, 27. https://doi.org/10.3389/fninf.2016.00027

Orima, T., & Motoyoshi, I. (2023). Spatiotemporal cortical dynamics for visual scene processing as revealed by EEG decoding. *Frontiers in Neuroscience*, *17*, Article 1167719. https://doi.org/10.3389/fnins.2023.1167719

Park, J., & Park, S. (2017). Conjoint representation of texture ensemble and location in the parahippocampal place area. *Journal of Neurophysiology*, *117*(4), 1595-1607. https://doi.org/10.1152/jn.00338.2016

Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling Scene Content from Spatial Boundary: Complementary Roles for the Parahippocampal Place Area and Lateral Occipital Complex in Representing Real-World Scenes. *Journal of Neuroscience*, *31*(4), 1333-1340. https://doi.org/10.1523/jneurosci.3885-10.2011

Rahman, R. A., & Sommer, W. (2008). Seeing what we know and understand: How knowledge shapes perception. *Psychonomic Bulletin & Review*, *15*(6), 1055-1063. https://doi.org/10.3758/pbr.15.6.1055

Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research*, *44*(19), 2301-2311. https://doi.org/10.1016/j.visres.2004.04.006

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition*, *12*(6), 852-877. https://doi.org/10.1080/13506280444000553

Singer, J. J. D., Cichy, R. M., & Hebart, M. N. (2023). The Spatiotemporal Neural Dynamics of Object Recognition for Natural Images and Line Drawings. *Journal of Neuroscience*, *43*(3), 484-500. https://doi.org/10.1523/jneurosci.1546-22.2022

Smith, F. W., & Smith, M. L. (2019). Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks [Article]. *Neuroimage*, *195*, 261-271. https://doi.org/10.1016/j.neuroimage.2019.03.065

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement:

Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, *44*(1), 83-98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *Neuroimage*, *65*, 69-82. https://doi.org/10.1016/j.neuroimage.2012.09.063

Teichmann, L., Quek, G. L., Robinson, A. K., Grootswagers, T., Carlson, T. A., & Rich, A. N. (2020). The Influence of Object-Color Knowledge on Emerging Object Representations in the Brain. *Journal of Neuroscience*, *40*(35), 6779-6789. https://doi.org/10.1523/jneurosci.0158-20.2020

Vailaya, A., Jain, A., & Zhang, H. J. (1998). On image classification: City images vs. landscapes. *Pattern recognition*, *31*(12), 1921-1935.

Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(23), 9661-9666. https://doi.org/10.1073/pnas.1015666108

Walther, D. B., & Shen, D. D. (2014). Nonaccidental Properties Underlie Human Categorization of Complex Natural Scenes. *Psychological science*, *25*(4), 851-860. https://doi.org/10.1177/0956797613512662

Yao, L. S., Fu, Q. F., & Liu, C. H. (2023). The roles of edge-based and surface-based information in the dynamic neural representation of objects. *Neuroimage*, *283*,

Article 120425. https://doi.org/10.1016/j.neuroimage.2023.120425

Yonas, A., & Arterberry, M. E. (1994). Infants perceive spatial structure specified by line junctions. *Perception*, *23*(12), 1427-1435. https://doi.org/10.1068/p231427

Data and code availability

The available data and code can be found at https://doi.org/10.57760/sciencedb.psych.00262.

Declaration of Interest Statement

We declare that we have no financial and personal relationship with other people or organization that can inappropriately influence our work, and there is no professional or other personal interest of any nature or kind in any service or company that could be construed as influencing the review of the manuscript entitled.