Exploring the similarity space of visual art

Andrew Frankel (afrankel@gradcenter.cuny.edu)
Program in Cognitive Neuroscience
Graduate Center, The City University of New York
365 5th Ave
New York, NY 10016

Sunil Ale (sale000@citymail.cuny.edu)
Department of Psychology
City College, The City University of New York
160 Convent Ave.
New York, NY 10031

Colin Conwell (colinconwell@gmail.com)
Department of Cognitive Science
The Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218

Edward A. Vessel (evessel@ccny.cuny.edu)
Department of Psychology
City College, The City University of New York
160 Convent Ave.
New York, NY 10031

Abstract

What is the similarity structure of visual art? Using DreamSim (Fu et al., 2023), a variant on a CLIP neural network model optimized to produce human-like similarity scaling, we investigated the distinctiveness of artworks considering both art movements and artists, two dominant forms of category information organizing visual art. Art movements differed in their distinctiveness from other movements and also in the degree to which different artists within each movement were separable. This work highlights the promise of using linguistically-informed similarity spaces for understanding the impact of the arts on cognition.

Keywords: computational aesthetics; visual artwork; similarity

Introduction

Experiences with art can be highly impactful: they can emotionally move (Menninghaus et al., 2019), provide creative inspiration (Welke et al., 2021), and affect mood (Trupp et al., 2023). Yet the cognitive mechanisms by which art objects have these effects are still poorly understood.

Several long-standing ideas in empirical aesthetics, along with recent theoretical and experimental work, point to an intimate relationship between aesthetic judgments of artwork and the process by which people build an understanding of the world around them (Biederman & Vessel, 2006; de Cruys et al., 2024; Sarasso et al., 2020; Schmidhuber, 2009). Although our internal models of the visual environment span many hierarchical levels, it has been hypothesized that conceptual knowledge reflecting how we categorize the world around us may have an especially strong influence on aesthetic judgments (Brielmann & Pelli, 2019; Iigaya et al., 2021). In the context of visual artwork, it is likely that many forms of categorical knowledge influence how a stimulus is perceived and interpreted. For example, classification of an artwork *as an artwork* has a fundamental organizing effect on subsequent aspects of evaluation (Leder et al., 2004). Two of the most salient aspects of the categorization of artworks are by *art movement* (roughly akin to style or period) and *artist*.

We utilize recent developments in machine learning to explore the similarity space of visual artworks with respect to these two levels of categorization (art movement and artist). More specifically, we make use of DreamSim (Fu et al., 2023), a multimodal neural network that combines visual information with semantic information about the relatedness of linguistic categories and has been optimized for computing human-like similarity metrics of images.

Methods and Analyses

Stimuli. We downloaded 106,112 images of artworks from wikiart.org, spanning 15 different art movements and 1,230 artists. We restricted our analysis to art movements that had at least 8 artists with 100 unique artworks (thus excluding 4 movements). We created a balanced dataset with 100 randomly selected images for each artist.

Analysis Pipeline. We used DreamSim (Fu et al., 2023) to embed the art images into a comprehensive latent space. DreamSim is a variation on CLIP (Contrastive Language-Image Pre-training; (Radford et al., 2021)): a neural network model that is optimized for computing vector-space similarity metrics that align with human similarity ratings. For each image, DreamSim outputs a 1792-dimensional vector reflecting the embedding of each image into a human-aligned similarity space.

To evaluate the distinctiveness of different art movements (or artists within a movement) we used support-vector machines (SVM; (Vapnik, 1997)) to assess the linear separability of the classes. Within each art movement, we used a stratified k-fold (k=5) cross-validation strategy and the SVM implementation in scikit-learn (Pedregosa et al., 2011) with a linear kernel and the default regularization parameter C=1. We report the mean and standard deviation across the 5 folds in Table 1.

To visualize structure in this embedding, we first applied principal component analysis (PCA). We kept 500 dimensions, which captured 85% of the variance. We then applied t-SNE (t-distributed Stochastic Neighbor Embedding; (van der Maaten & Hinton, 2008)) a non-linear dimensionality reduction technique, to arrange items in a two-dimensional space for visualization.

Results

How distinctive are different art movements?

The similarity space embedding for the full set of artworks is illustrated in Fig. 1a, labeled by art movement. Some movements appear to be both distinctive and tightly grouped ((e.g. abstract expressionism, dark purple). Other movements such as impressionism and post-impressionism, appear highly intermixed. Romanticism is notable for consisting of well-separated groupings spread throughout the similarity space.

Within a movement, how distinctive is each artist?

Fig. 1b shows the same embeddings, labeled by artist. A comparison of the movement-labeled embedding to the artist-labeled embedding makes it clear that many of the individual groupings within Romanticism comprise paintings from individual artists, each of which occupies a distinctive region of similarity space despite sharing the common movement label. On the other hand, the impressionist painters appear far less tightly grouped, with artworks from individual painters being more similar to those of other artists. Abstract-expressionism shows an intermediate pattern: all painters appear very close together, but still show distinctiveness from painter to painter.

Are there differences across movements in the distinctiveness of individual artists?

To formally test differences across genres, we measured the separability of the 8 most prolific artists in each movement (Table 1). Separability scores ranged from 94.8% correct classification for Surrealism to 79.1% correct classification for Impressionism. Separability for Romanticism was 92.1%, confirming that individual Romantic artists are more separated than Impressionist artists. Separate visualizations of the embeddings for these two movements (Fig. 1c,e) clearly show that

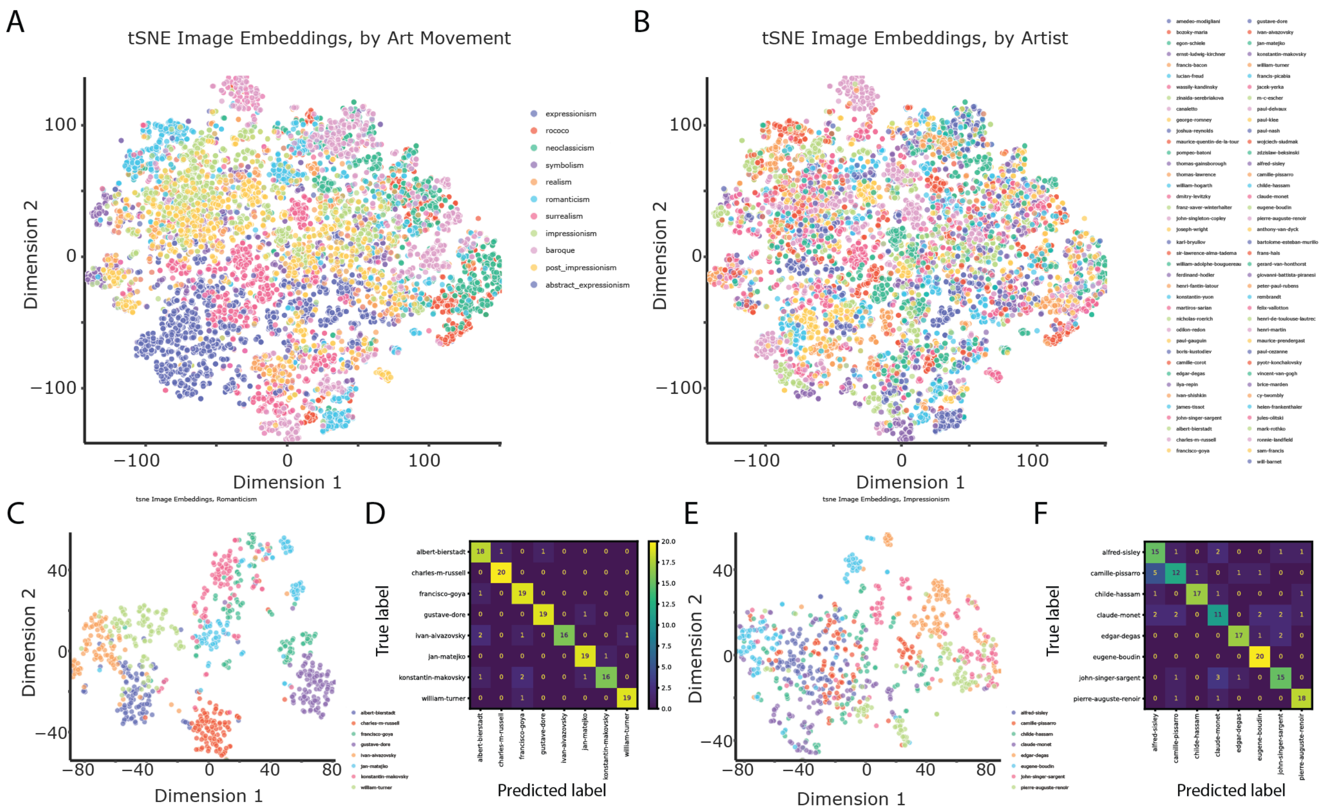

Figure 1: a) Visualization of 8800 artworks in a 2-dimensional similarity space, labeled by art movement (8 artists by 11 art movements, 100 items/artist). The embedding was computed using DreamSim (Fu et al., 2023), and visualized in 2-dimensions using tSNE after dimensionality reduction (see Analysis). b) Visualization of the same 8800 artworks labeled by artist. c) Similarity space for Romanticism labeled by artist and d) the confusion matrix for the Romanticism test set (160 images), illustrating a very high degree of distinctiveness. e) Similarity space for Impressionism labeled by artist and f) the confusion matrix for the Impressionism test set (160 images), illustrating a lower degree of distinctiveness.

individual artists form much more distinct clusters for Romanticism than Impressionism. Confusion matrices for the test trials (Fig 1d,f) reveal that works by the Impressionist artist Monet in particular are highly confusable with other artists.

Art Movement	Artist Separability Score Mean	Score standard dev. 5-fold CV
Surrealism	0.94750	0.008478
Expressionism	0.92875	0.021506
Abstract-expressionism	0.92375	0.014470
Romanticism	0.92125	0.009354
Post-impressionism	0.86500	0.018371
Rococo	0.86000	0.025495
Symbolism	0.85500	0.009186
Baroque	0.84750	0.021139
Neoclassicism	0.83375	0.034596
Realism	0.82875	0.032500
Impressionism	0.79125	0.015104

Table 1: Separability of artists in each of 11 art movements, computed as a proportion of correct classifications in the test set (100 images per movement) across 5 cross-validation folds.

Conclusion

Using a deep neural network trained to produce human-like similarity ratings, we analyzed the distinctiveness of two dominant forms of categorical structure, art movements and artists, across a large database of visual artwork. We found notable differences across different art movements, both in their degree of distinctiveness between different art movements and in the overall similarity of individual paintings within a movement.

We note that the degree to which the observed similarity structure would reflect the judgments of individual human observers remains to be seen. In addition, it is unclear how low-level features reflecting style and medium, versus high-level semantic information informed by the inclusion of language in DreamSim training, shape the resultant similarity space.

Despite these shortcomings, this approach shows great promise. These results form a basis for understanding how the category structure of visual artwork, a human-made, highly-varied class of objects, relates to their underlying visual and semantic similarity. This information will enable the development and testing of models for how the the impact of an art object relates the conceptual and categorical structure of internal representations.

References

Biederman, I., & Vessel, E. A. (2006). Perceptual pleasure and the brain. *American Scientist*, *94*, 247-253. doi: 10.1511/2006.3.247

Brielmann, A. A., & Pelli, D. G. (2019). Intense beauty requires intense pleasure. *Frontiers in Psychology*, *10*, 1-17. doi: 10.3389/fpsyg.2019.02420

de Cruys, S. V., Frascaroli, J., & Friston, K. (2024, 1). Order and change in art: towards an active inference account of aesthetic experience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*. Retrieved from https://royalsocietypublishing.org/doi/10.1098/rstb.2022.0411 doi: 10.1098/rstb.2022.0411

Fu, S., Tamir, N. Y., Sundaram, S., Chai, L., Zhang, R., Dekel, T., & Isola, P. (2023). *Dreamsim: Learning new dimensions of human visual similarity using synthetic data.* Retrieved from https://dreamsim-nights.github.io/

Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., & O'Doherty, J. P. (2021). *Nature Human Behaviour*. doi: 10.1038/s41562-021-01124-6

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*, 489-508.

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*, *126*, 171-195. Retrieved from https://www.aesthetics.mpg.de/fileadmin/user_upload/Dateien_LIT/Menninghaus_et_al._What_are_aesthetic_emotions.pdf doi: 10.1037/rev0000135

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... Sutskever, I. (2021). *Learning transferable visual models from natural language supervision.*

Sarasso, P., Neppi-Modona, M., Sacco, K., & Ronga, I. (2020). "stopping for knowledge": The sense of beauty in the perception-action cycle. *Neuroscience and Biobehavioral Reviews*, *118*, 723-738. Retrieved from https://doi.org/10.1016/j.neubiorev.2020.09.004 doi: 10.1016/j.neubiorev.2020.09.004

Schmidhuber, J. (2009). Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. *Anticipatory Behavior in Adaptive Learning Systems*, *5499*, 48-76.

Trupp, M. D., Bignardi, G., Specker, E., Vessel, E. A., & Pelowski, M. (2023, 8). Who benefits from online art viewing, and how: The role of pleasure, meaningfulness, and trait aesthetic responsiveness in computer-based art interventions for well-being. *Computers in Human Behavior*, *145*. doi: 10.1016/j.chb.2023.107764

van der Maaten, L., & Hinton, G. E. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, *9*, 2579-2605. Retrieved from https://api.semanticscholar.org/CorpusID:5855042

Vapnik, V. N. (1997). The support vector method. In W. Gerstner, A. Germond, M. Hasler, & J.-D. Nicoud (Eds.), *Artificial neural networks — icann'97* (pp. 261–271). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/BFb0020166

Welke, D., Purton, I., & Vessel, E. A. (2021). Inspired by art: Higher aesthetic appeal elicits increased felt inspiration in a creative writing task. *Psychology of Aesthetics, Creativity, and the Arts*. doi: 10.1037/aca0000393