Inferring Individualized Color-Vision Distortions from fMRI Hue-Representation Geometry

Jinil Kim[1.2.3], Albert Minkue Cho[4], Jungwoo Seo[5], Jiook Cha[2.5.6*]

[1]Dept. of Linguistics, [2]Dept. of Psychology, [3]Dept. of Computer Science and Engineering, [4]Dept. of Electrical and Computer Engineering, [5]Dept. of Brain & Cognitive Sciences, [6]Interdisciplinary Program of Artificial Intelligence, Seoul National University

Codes & Pages

KEY CLAIM

From fMRI hue geometry, we fit each person's cortical distortion model and invert it into a personalized color-correction filter.

BACKGROUND & GAP

- **Health phenotypes** (cardiac[1], GE[2]) can be represented as **structured distortions.**
- Such distortions **vary across individuals** even within a diagnostic category.
- Prior work simulates sensory loss or applies population-level fixes.
  - **Inverting a person's own distortion for intervention is under-explored**

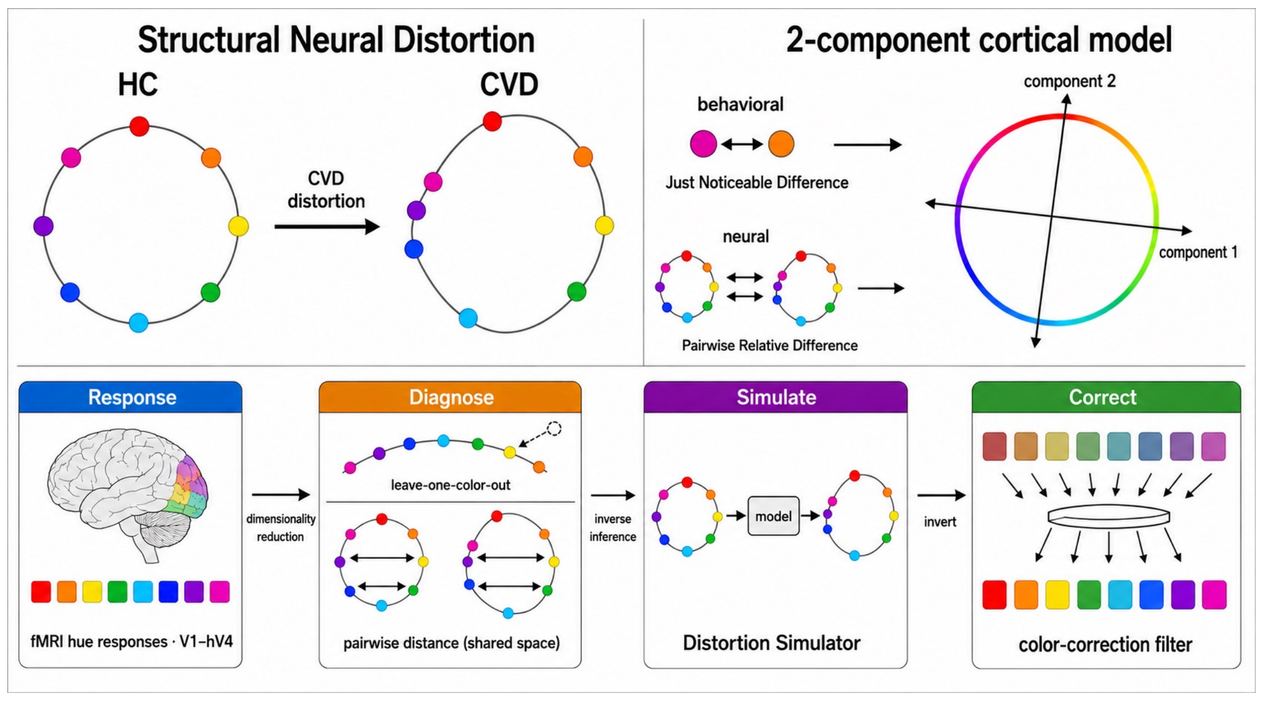
Contributions

- **We represent CVD as a distortion of the brain's hue-representation geometry**
- **We fit a cortical model per person by combining behavioral & neural losses**
- **We invert each person's fitted model into their color-correction filter**

METHODS – EXPERIMENT & ANALYSIS

- **Participants (N = 10)**: 7 Healthy Controls (HC) + 3 CVD (two deutans, one protan)
- **fMRI Experiment**: **8 isoluminant hues, 6 runs, ROIs V1-hV4**
- **LOCO (leave-one-color-out)**: can a held-out hue be rebuilt from the other seven?
  - **Forward-Encoding model** maps hue → 6 channels → neural responses
  - Tests **continuous hue structure** in neural hue representations
- **RDM (relative distances)**: how far apart are hue pairs in the brain representation?
  - **Shared Response Model (SRM)** projects HC & CVD into a **shared low(4)-dim space**
  - ΔRDM = CVD − HC = which pairs are **relatively merged or split**

FRAMEWORK & MODEL - LOSS



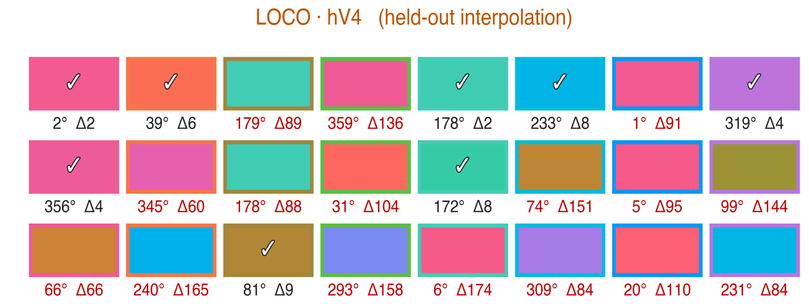
Diagnose each person's hue distortion**, fit** a cortical model as **a deficiency simulator,** then **invert** it into a **color-correction filter.**

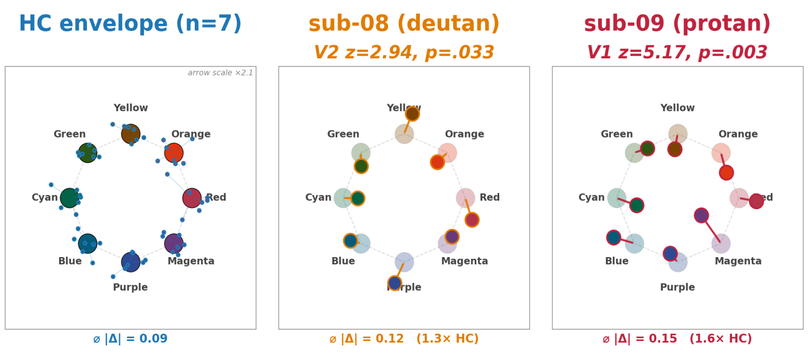
Model mechanism	Parameters	Form
Retinal Cone-Shift[6]	1	$\Delta\lambda$
Retinal + Cortical compensation	1	$\delta\theta = (2-g)\cdot\delta\theta\_Machado$
Cortical two-component	2	$\theta' = \theta + \beta\_s \cos(\theta-90°) + \beta\_c \cos(\theta-\theta\_conf)$

Loss = behavioral JND (Lγ) + V1,V2 ΔRDM (L_RDM) + hV4 LOCO (L_LOCO);
selected by held-out generalization

RESULTS

A. Represent · A structured distortion





Category discrimination preserved

- LORO accuracy difference: p=0.668
- Every hue stays decodable.

But continuous interpolation is impaired under LOCO

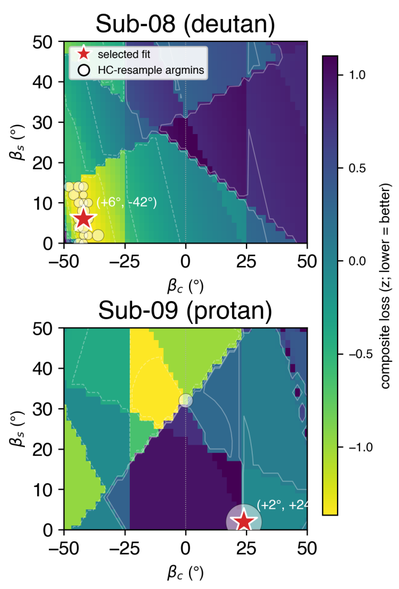
- LOCO Accuracy: HC 0.47

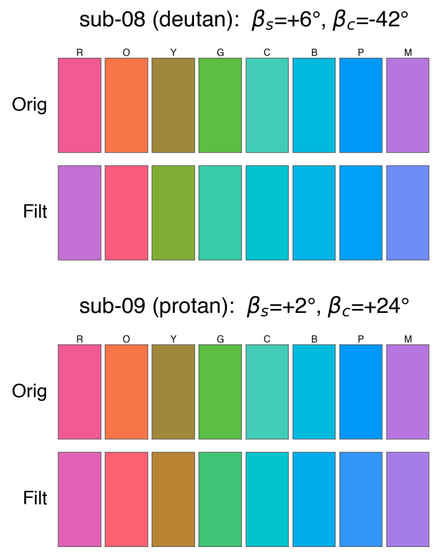
→ Deutan 0.25 (p=.082) / Protan 0.13 (p=.024)

Peak distortion in a subject-specific ROI

- Deutan V2 p=0.040, Protan V1 p=0.007

B. Fit · Deficiency modeling





Retina models failed:

- Overcompensation, non-invertible

Result (β_s, β_c; final loss):

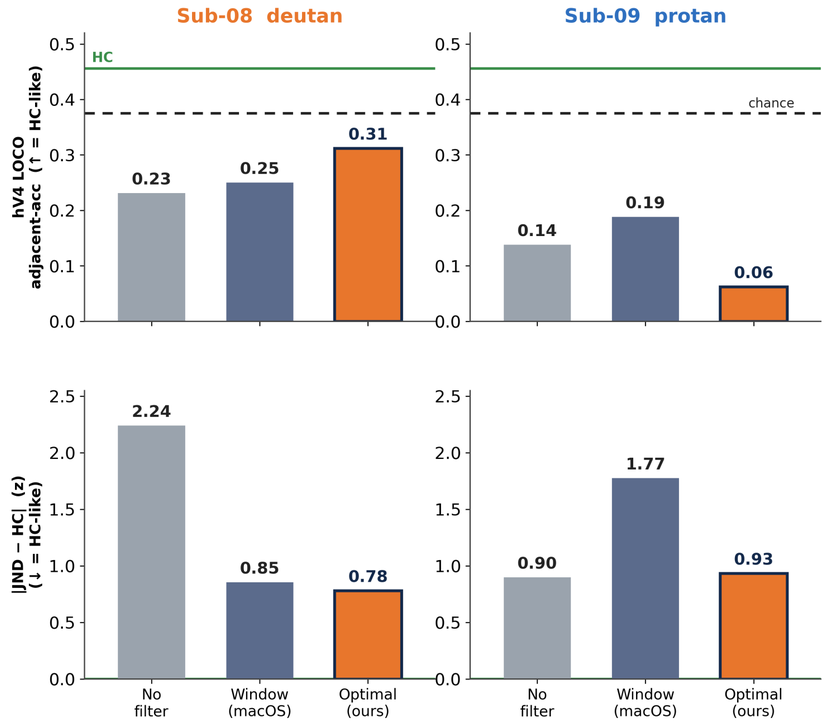
- Deutan: (+6°, −42°); L = γ_OY + ΔRDM_V2
- Protan: (+2°, +24°); L = γ_all + ΔRDM_V1

Neural loss stabilizes parameters

- Deutan: β_c sign stable, no zero-crossing
- Protan: decisive only with neural

(3 → 7 of 7 held-out folds)

C. Validation (2nd-MRI, Exploratory)



Behavior (JND HC-disparity, ↓ = HC-like):

- **Optimal ≥ Window in HC-likeness**
- Protan: Window harms, Optimal preserves

Neural (hV4 LOCO):

- **Deutan rises at Optimal but below chance**
- Protan remains floored.

Neural (V1/V2 SRM-RDM):

- Deutan shifts away from HC
- **Protan recovers toward HC under both filters**

CONCLUSIONS

End-to-end framework:

- **Structured distortion → model fitting → inversion → personalized filter**

Behavioral HC-likeness favors the personalized filter; exploratory **neural** effects

- **Limits:** N=2; heterogeneous results across metrics/subjects
- **Next:** a unified geometry target (integration across metrics, ROIs), larger N

REFERENCES

1. Biffi C et al. IEEE Trans Med Imaging 39, 2088–2099 (2020).
2. Lotfollahi M et al. Nat Methods 16, 715–721 (2019).
3. Brouwer G, Heeger D. J Neurosci 29, 13992–14003 (2009).
4. Kriegeskorte N et al. Front Syst Neurosci 2, 4 (2008).
5. Bannert M, Bartels A. J Neurosci 38, 3657–3668 (2018).
6. Machado G et al. IEEE Trans Vis Comput Graph 15, 1291–1298 (2009).
7. Somers L et al. Vision Res 218, 108390 (2024).
8. Tregillus K et al. Curr Biol 31, 936–942 (2021).