Counting the Features a Colormap Invents: A Persistent-Homology Audit of Scientific Color Maps Across Observer Models

Satyajeeth Suresh Kannan
Innovation Academy · Mary's Lab
github.com/JeetuSK0808/maryslab

Preprint · August 2026 · Mary's Lab technical report TR-02

ABSTRACT

The rainbow colormap is widely criticised for introducing visual structure that the underlying data does not contain. The criticism is normally argued perceptually and asserted qualitatively. We make it a measurement. Treating the displayed image as a scalar field of perceived lightness, we compute zero-dimensional persistent homology on both the data field and its displayed rendering, and report the difference: features displayed but not present are invented by the colormap. Applied to a synthetic field containing exactly one true maximum, viridis and greyscale display exactly 1 feature and are faithful for every observer model tested, while jet displays 24 and rainbow 52 — respectively 23 and 51 phantom features. Running the analysis per observer model reveals an effect not visible under normal vision alone: simulated dichromacy substantially changes the phantom count, in one case reducing it, which we argue makes per-observer auditing a necessary rather than optional part of the procedure.

Keywords: scientific visualization; colormaps; persistent homology; color vision deficiency; perceptual uniformity

1. Introduction

That rainbow colormaps mislead is among the most repeated claims in scientific visualization. The argument is normally made in terms of perceptual non-uniformity: the map's lightness is not monotonic in the data value, so equal steps in the data do not produce equal steps in perceived brightness, and viewers consequently perceive boundaries where the data is smooth.

The argument is correct and it is not a measurement. It provides no number to report, no threshold to enforce in review, and no way to rank two candidate maps. Practitioners are asked to accept an aesthetic judgment on authority.

We propose that the claim has an exact form. If a colormap causes a viewer to see a ridge, that ridge is a topological feature of the displayed image that is absent from the data. Topological features admit counting, and persistent homology supplies both the count and a magnitude per feature — allowing noise-scale wobbles to be separated from structure a reader will actually perceive.

2. Method

The displayed field

Let the data be a scalar field `f` on a grid. A colormap is a path `c(t)` through colour space; the displayed image is `c(f_normalised)`. We define the *displayed field* as the perceived lightness of each displayed cell — the Oklab L coordinate of the rendered colour, optionally after passing through an observer model.

A faithful colormap is one whose displayed field is topologically equivalent to the data field. This is a stronger and more useful statement than "the map is perceptually uniform", and it is checkable.

Persistence

We compute zero-dimensional persistent homology by sublevel-set filtration, implemented as a union–find over cells sorted by value with the elder rule: when two components merge, the younger dies and the older survives. Each feature receives a birth and a death, and the difference is its persistence. Low-persistence features are quantisation and interpolation noise; high-persistence features are structure a reader perceives. A single threshold separates them, and we report it with every result.

The audit computes the persistence diagram of the data field and of the displayed field, counts significant features in each, and reports the surplus as *spurious* and the deficit as *lost*.

Observers

The displayed field depends on who is looking. We therefore repeat the analysis under simulated protanopia, deuteranopia and tritanopia using the Machado model, in addition to normal trichromatic vision. A colormap may be faithful for one observer and not another; reporting a single number hides that.

3. Evaluation

Inputs are synthetic, and deliberately so. The test field is a single isotropic Gaussian peak on a 64×64 grid. It has exactly one maximum. This is the property that makes the experiment interpretable: any feature beyond the first that a colormap displays is necessarily invented, because the ground truth is known by construction. A measured dataset would confound the measurement with uncertainty about its own topology.

Table 1. Features displayed versus features present, on a field with exactly one true maximum. Persistence threshold 0.02. "Spurious" is the count of displayed features with no counterpart in the data.

colormap	observer	data	displayed	spurious	monotonic
greyscale	normal	1	1	0	yes
viridis	normal	1	1	0	yes
viridis	deutan	1	1	0	yes
jet	normal	1	24	23	no
jet	protan	1	1	0	no
jet	deutan	1	8	7	no
jet	tritan	1	20	19	no
rainbow	normal	1	52	51	no
rainbow	deutan	1	44	43	no
coolwarm	normal	1	5	4	no

The separation is stark and it is not a matter of degree. Greyscale and viridis display exactly the one feature the data contains, for every observer model. Jet displays 24 features to a normal trichromat, of which 23 do not exist. Rainbow displays 52. These are not marginal excursions above a threshold; they are dozens of distinct perceived maxima imposed on a single smooth hill.

The per-observer results contain a finding we did not anticipate. Under simulated protanopia, jet's phantom count falls to 1 — the dichromatic projection collapses precisely the red–green excursions that generate jet's phantom ridges for a trichromat. The map remains classified as distorted because its lightness profile is still non-monotonic, but the *kind* of damage differs by observer. A practitioner auditing only under normal vision would draw incomplete conclusions, and a practitioner auditing only under simulated CVD could conclude jet was acceptable. Both single-observer procedures are inadequate.

Coolwarm, a diverging map often recommended as a safe default, displays 4 spurious features on this field. This is expected rather than damning: a diverging map is designed for data with a meaningful midpoint, and applying one to a single-peaked field is a misuse. The audit measures the pairing of map and data, not the map in isolation, and that is the correct unit of analysis.

Reproducing these numbers

Clone `github.com/JeetuSK0808/maryslab` and run `node papers/build/experiments.mjs`. Every figure in this report is written by that script into `papers/build/results.json`; the instrument itself is exercised by `audit_colormap_topology`. The engines carry a 401-vector conformance suite shared by a TypeScript reference implementation, a C++20 core, and a WebAssembly build, so a reported number is a property of the specification rather than of one implementation.

4. Real-data evaluation: a measured field with structure nobody planted

The field above contains exactly one feature because we put exactly one there. That makes the ground truth unarguable and the test easy: a map has to invent a feature out of a smooth background, which is the mildest version of the failure. The harder question is what these maps do to a field with real structure, where features sit at many scales and some of them are genuinely marginal.

Snapshot `usgs-splib07a-1` supplies one directly. A matrix of 1752 measured reflectances against the 38 grid wavelengths is not a contrived test image — it is the heatmap a spectroscopy paper prints. We took 256 of those materials, min–max normalised the matrix, and audited the same five maps against it. The field carries 99 features above threshold in the data itself.

Table 2. Displayed feature count on a real 256 × 38 reflectance matrix carrying 99 true features. Every cell is what a reader would count off the rendered image; the data count is identical in all cases because the field never changes.

colormap	normal	protan	deutan	tritan
grayscale	96	96	96	96
viridis	97	97	97	97
jet	136	151	136	132
rainbow	194	188	177	184
coolwarm	158	103	123	154

The ordering survives the move to real data, and the magnitudes get worse. Rainbow renders 194 features where the data has 99 — it very nearly doubles the apparent structure of the field. Jet adds 37. A reader counting bands off either image and reasoning about how many distinct material classes the library contains would be reasoning about the colormap.

Two things here that the synthetic field could not show. Coolwarm is the most observer-unstable map in the set: it displays 158 features to a normal trichromat and 103 under simulated protanopia. The same figure, the same data, and a difference of dozens of apparent features depending on who is reading it — a divergence the single-peak field had no room to express.

And no map is faithful here, including the two that were faithful on the synthetic field. Grayscale displays 96 and viridis 97, both short of 99: they *lose* features rather than invent them. That is the better failure — a reader misses something rather than seeing something that is not there — but it is still a failure, and the honest conclusion is not that viridis is safe. It is that on a real field every map in this set distorts, and they differ in which direction and by how much.

Reproducing these numbers

Clone `github.com/JeetuSK0808/maryslab` and run `node papers/build/experiments.mjs`. Every figure in this report is written by that script into `papers/build/results.json`; the instrument itself is exercised by `audit_colormap_topology`. The engines carry a 401-vector conformance suite shared by a TypeScript reference implementation, a C++20 core, and a WebAssembly build, so a reported number is a property of the specification rather than of one implementation.

5. Limitations

This section states what the result does not establish. It is written to be read before the abstract is quoted.

- **The measured field's ground truth is our own computation.** The 99-feature count comes from the same persistence calculation being evaluated, applied to the data before rendering. It is the right reference for asking what a colormap does to a field, and it is not an independent count of how many mineral classes the library contains.
- **The threshold is a judgment.** Persistence assigns every feature a magnitude; deciding which magnitudes a reader will perceive requires a cutoff, and that cutoff is an engineering choice. We report it (0.02) with every result so it can be varied.
- **Lightness is not all of perception.** The displayed field is built from perceived lightness. Chroma and hue transitions also drive perceived boundaries, and this instrument does not model them. The counts are therefore a lower bound on perceived structure, not a complete account.
- **No human subjects were involved.** We measure a property of the displayed image under a model of the visual system. We do not measure what viewers report seeing. The inference from phantom topological features to reader misinterpretation is argued, not demonstrated, and a psychophysical study would be required to close it.
- **Observer models are models.** The per-observer analysis inherits every limitation of the Machado simulation, which parameterises dichromacy but does not capture the full variation of real colour vision deficiency.
- **A clean audit is not an endorsement.** Zero spurious features means the map does not invent structure

on this field. It says nothing about legibility in print, at small sizes, or against a given background.

- **Prior art was surveyed informally.** Persistent homology is standard in topological data analysis and colormap criticism is a large literature; a search found no tool applying the former to quantify the latter per observer model. That is a statement about a search, not a proof of novelty.

6. Acknowledgements and disclosure

The instruments described here were implemented in Mary's Lab, an open-source computational color science project. Software design, implementation, and the preparation of this report were carried out with substantial assistance from a large language model (Anthropic Claude), under the author's direction and review. All numerical results were produced by executing the released code; none were generated by a language model. The author is responsible for the claims made here.

References

1. Borland, D. and Taylor, R. M. (2007). Rainbow color map (still) considered harmful. *IEEE Computer Graphics and Applications*, 27(2), 14–17.
2. Moreland, K. (2009). Diverging color maps for scientific visualization. *Proceedings of the International Symposium on Visual Computing*, 92–103.
3. Crameri, F., Shephard, G. E., and Heron, P. J. (2020). The misuse of colour in science communication. *Nature Communications*, 11, 5444.
4. Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2002). Topological persistence and simplification. *Discrete & Computational Geometry*, 28, 511–533.
5. Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: An Introduction*. American Mathematical Society.
6. Machado, G. M., Oliveira, M. M., and Fernandes, L. A. F. (2009). A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1291–1298.
7. Ottosson, B. (2020). *A perceptual color space for image processing (Oklab)*. bottosson.github.io.
8. Nuñez, J. R., Anderton, C. R., and Renslow, R. S. (2018). Optimizing colormaps with consideration for color vision deficiency. *PLOS ONE*, 13(7), e0199239.

Source, engines, and the script that produced every number in this report: `github.com/JeetuSK0808/maryslab`. Results regenerate with `node papers/build/experiments.mjs`. Inputs: synthetic where §3 says so, and measured USGS splib07a spectra (snapshot `usgs-splib07a-1`) where a real-data section says so. Prepared with AI assistance (see Acknowledgements).