Designing Artificial Vision Policies: Rods, Cones, and Latent Filtering

Introduction
In biological vision, rods and cones serve different purposes: rods are sensitive to light and motion, while c
In artificial agents, we can replicate these functional roles by deciding how frequently to process grayscale

Rod-Dominant Vision
This policy favors grayscale input for efficiency and speed
The agent primarily perceives the world in black and white, with occasional RGB frames providing color an
This is analogous to having many rods and fewer cones
It reduces compute cost while still allowing color differentiation.

Cone-Heavy Vision
This mode emphasizes color detail, with RGB inputs processed every frame
Grayscale can be added as a parallel channel to stabilize luminance information
This policy is closer to camera-like vision, with color treated as primary information.

Balanced Camera-Like Vision
Here, grayscale and RGB are treated equally, with every frame processed in full RGB
This is the least efficient mode, but ensures maximum fidelity
It is appropriate for environments where fine color discrimination is essential.

Optimization Strategies
To manage compute costs, batching multiple teacher frames (RGB, grayscale, R, G, B) into one encode st
Alternatively, color processing can be scheduled at intervals (e.g., every 4th frame) to reduce load while m
This creates a hybrid policy, blending rod-dominant efficiency with cone-based richness.

Design Consequences
The chosen policy defines the agent's perceptual world
Rod-heavy systems will deprioritize color and may exhibit behavior similar to partial color blindness, while
Adding a foveal ROI (mouse-centered region with sharp detail and blurred periphery) could further bias the

Conclusion
By tuning the ratio of grayscale to RGB inputs, developers control whether an agent evolves rod-like efficie
These design decisions fundamentally shape how intelligence emerges from perception.