Learning to Perceive, Perceiving to Act: An Information-Theoretic View of Robot Decision Making

Shivaram Kumar Jagannathan
shivaram.jagannathan [at] outlook [dot] com

August 12, 2026

Abstract

Robot decision making is only as reliable as the information supplied by perception. This paper develops a complementary view of learning-based robotics in which deep visual models, information theory, and physical inductive biases work together. A perception system should not merely predict what is visible; it should estimate which uncertainties matter for the next decision and choose observations that reduce them. We formalise this idea with mutual information, Bayesian belief updates, and expected information value. We then connect these objectives to physics-informed representation learning, where geometry, dynamics, and conservation laws regularise learned models. Optimisation-based controllers, search-based planners, and learned policies are treated as downstream decision mechanisms that can all benefit from a better belief state. The result is a view of perception as an active, decision-aware inference process rather than a passive front end.

1 Introduction

Robotic systems must infer hidden state from incomplete, noisy, and often adversarial observations before they can choose a useful action. A camera does not reveal mass, friction, contact mode, or whether an occluded object will move when pushed. A learned policy can be expressive, while a symbolic planner can be interpretable and compositional, but both still depend on the quality of the state or belief supplied to them.

The choice of decision mechanism does not remove this dependence. Optimisation, search, sampling, and learned policies provide different ways to convert beliefs into actions, but each inherits the blind spots of its perceptual interface. Our claim is therefore narrow: perception should be trained and evaluated by how it changes downstream decisions, not only by pixel-level or object-level accuracy.

The central question is therefore:

> Which observation should a robot seek next, and how much will that observation improve the decisions it can make?

Figure **??** gives the proposed loop. A deep model extracts a representation from sensor data, an information-theoretic objective identifies decision-relevant uncertainty, and a planner or controller chooses an action that both pursues the task and improves the next belief update.

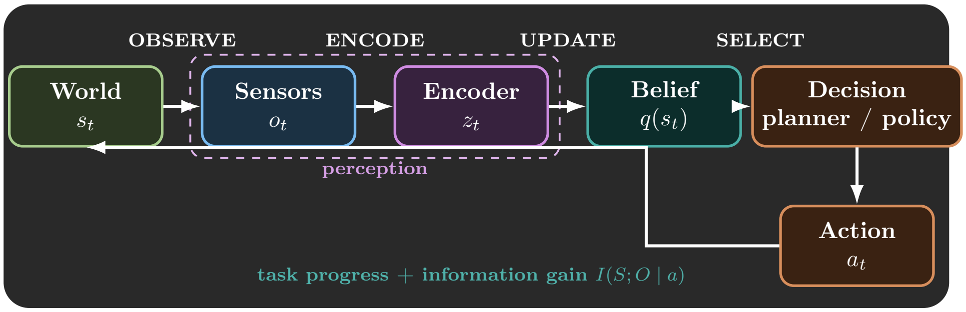

Figure 1: Perception and decision making as one loop. The robot can choose actions for task progress and for information gain, while the downstream planner, controller, or policy operates on an explicitly uncertain belief.

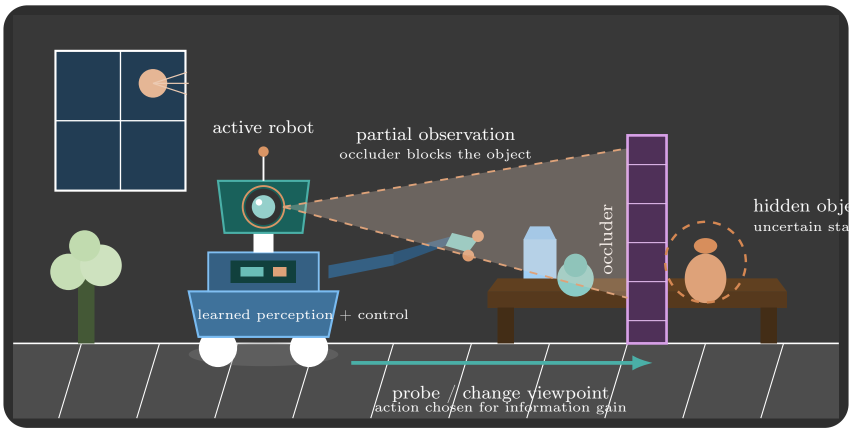

Figure 2: A pictorial view of active perception. The robot does not treat the occluded object as a nuisance outside the planner: it moves its viewpoint or probes the scene to reduce uncertainty before committing to a decision.

2 Information as a Decision Resource

Let S denote latent task state, O an observation, A an action, and G the task objective. A belief $q(s)$ summarises what the robot currently knows. The uncertainty of that belief can be measured by entropy,

$$H_q(S) = -\mathbb{E}_{s\sim q}\left[\log q(s)\right]. \tag{1}$$

An observation is useful when it reduces uncertainty about variables that affect future decisions. Conditional mutual information captures this value:

$$\mathcal{I}(A) = I(S; O \mid A = a) = H(S \mid A = a) - H(S \mid O, A = a). \tag{2}$$

The action need not be a dedicated sensing motion. A grasp, camera pan, or contact-rich push can simultaneously make progress and reveal geometry or material properties. We can express this trade-off as

$$a_t^* = \arg\max_{a \in \mathcal{A}} \underbrace{\mathbb{E}[R(G) \mid a]}_{\text{task value}} + \beta \underbrace{I(S; O \mid a)}_{\text{information value}} . \tag{3}$$

The coefficient β does not declare planning or control unimportant. It controls when the robot should spend effort resolving uncertainty rather than optimising under an unreliable belief. In a well-observed state, the first term dominates; near an ambiguous contact or occlusion, the second can be decisive.

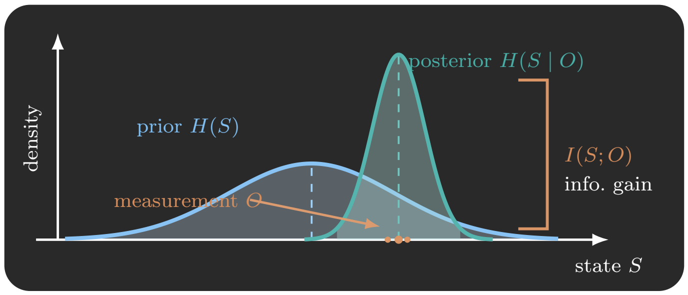

Figure 3: An informative observation transforms a broad prior into a more useful posterior. The value is not abstract certainty; it is the reduction in uncertainty relevant to the next action.

3 Deep Representations with Physical Structure

Deep learning provides flexible representations for pixels, point clouds, proprioception, and contact histories. Yet unconstrained representations can use shortcuts that fail under viewpoint changes, new materials, or altered dynamics. Physical structure offers inductive biases that improve data efficiency and improve the meaning of latent variables.

For a learned dynamics model f_θ, a physics-informed objective can combine data fit with a residual from a known differential equation:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{data}}(\theta) + \lambda_{\text{phys}} \, \mathbb{E}_t \left[\left\| \frac{\partial \hat{x}_\theta}{\partial t} - F_{\text{phys}}(\hat{x}_\theta, u_t) \right\|^2 \right] . \tag{4}$$

The same principle applies beyond explicit dynamics. SE(3)-equivariant features respect rigid-body transformations; geometric encoders preserve relations between points; differentiable simulators expose contact and constraint structure; and neural operators can learn corrections to PDE or rigid-body models. These priors do not replace learning. They restrict the hypothesis space to functions that are easier for a robot to identify.

4 A Unified Learning Objective

The preceding pieces suggest a training objective that scores representations by both predictive quality and their usefulness for action. Let $Z = f_\theta(O)$ be a learned representation and π_ψ a decision mechanism. A schematic objective is

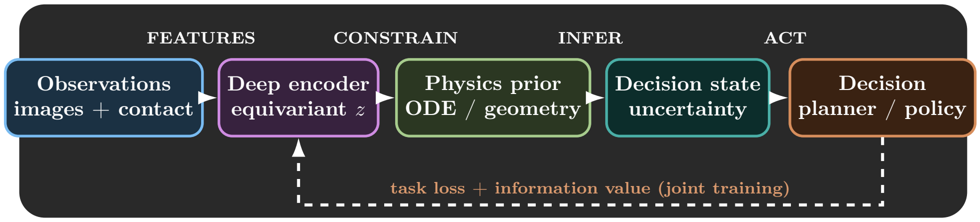

Figure 4: A hybrid perception model. Deep features capture appearance and contact cues, while physical structure regularises the latent state passed to the planner or controller. The decision objective can train the whole system.

$$\min_{\theta,\psi} \mathbb{E}\left[\mathcal{L}_{\text{task}}(\pi_\psi(Z), G) + \alpha\, \mathcal{L}_{\text{phys}}(Z) - \beta\, I(Z; S \mid G)\right]. \tag{5}$$

The final term is deliberately conditional on the goal. Maximising information about every possible property is wasteful; the robot should preserve information that changes the decision. In practice, $I(Z; S \mid G)$ can be approximated with variational bounds, contrastive objectives, ensembles, or uncertainty-aware world models. The physics term can be a differentiable residual, an equivariance penalty, or a simulator consistency loss.

This view also clarifies the role of existing decision systems. A structured planner can use a belief whose state variables carry calibrated uncertainty. An optimisation-based controller can select trajectories that expose informative contact modes. A learned policy can receive a latent state that separates task-relevant ambiguity from nuisance variation. The decision mechanism changes; the information interface remains.

5 Discussion

Several challenges remain. Mutual information is difficult to estimate in high dimensions, and a poor variational bound can reward representation artifacts. Physics models are incomplete, so an overly strong residual can make the model less accurate in regimes where the assumed equations fail. Finally, active information gathering has a real cost: time, energy, collision risk, and hardware wear must enter the objective alongside information value.

These limitations motivate a pragmatic combination of methods rather than a single replacement for existing robotics stacks. Deep networks provide rich perceptual features, physical priors improve extrapolation, information theory identifies what uncertainty matters, and planners or controllers turn beliefs into structured action. The most useful systems will likely compose all four.

6 Conclusion

Perception and decision making should not be separated by an interface that forgets uncertainty. A robot should learn representations that preserve the information needed to act, use physical structure to make those representations stable, and choose observations when they have high expected decision value. This perspective is compatible with search, optimisation, and learned policies: it gives each of them a better, more informative belief on which to operate.

References

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, 2006.

[2] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021.

[3] N. Tishby and N. Zaslavsky. Deep learning and the information bottleneck principle. In *IEEE Information Theory Workshop*, 2015.