Theoretical perspectives on active sensing

Scott Cheng-Hsin Yang[a,1], **Daniel M. Wolpert**[#a], and **Máté Lengyel**[#a,b]

[a]Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK

[b]Department of Cognitive Science, Central European University, Budapest H-1051, Hungary

[#] These authors contributed equally to this work.

Abstract

A key component of interacting with the world is how to direct ones' sensors so as to extract task-relevant information – a process referred to as active sensing. In this review, we present a framework for active sensing that forms a closed loop between an ideal observer, that extracts task-relevant information from a sequence of observations, and an ideal planner which specifies the actions that lead to the most informative observations. We discuss active sensing as an approximation to exploration in the wider framework of reinforcement learning, and conversely, discuss several sensory, perceptual, and motor processes as approximations to active sensing. Based on this framework, we introduce a taxonomy of sensing strategies, identify hallmarks of active sensing, and discuss recent advances in formalizing and quantifying active sensing.

Keywords

active sensing; motor planning; exploration; computational model

Introduction

Skilled performance requires the efficient gathering and processing of sensory information relevant to the given task. The quality of sensory information depends on our actions, because what we see, hear and touch is influenced by our movements. For example, the motor system controls the eyes' sensory stream by orienting the fovea to points of interest within the visual scene. Movements can therefore be used to efficiently gather information, a process termed active sensing. Active sensing involves two main processes: perception, by which we process sensory information and make inferences about the world, and action, by which we choose how to sample the world to obtain useful sensory information.

To illustrate the computational components of active sensing, we consider the task of trying to determine the time of day from a visual scene (Fig. 1). Because of the limited resolution of vision away from the fovea, sensory information at any point in time is determined by the fixation location (Fig. 1, red dot, Sensory Input). The perceptual process can be formalized in terms of an ideal observer model [1, 2] which makes task-relevant inferences. To do so,

[1]corresponding author: scott.cheng.hsin.yang@gmail.com.

the observer uses the sensory input together with a knowledge of the properties of the task (Fig. 1, Task) and the world, as well as features of our sensors, such as the acuity falloff in peripheral vision [3, 4] and processing limitations, such as limited visual memory [5, 6]. Such observers are typically formulated within the Bayesian framework. For example, the observer could use luminance information to estimate the time of the day, in this case formalized as a posterior probability distribution (Fig. 1, Observer).

The process of selecting an action can be formalized as an ideal planner which uses both the observer's inferences and knowledge of the task to determine the next movement, in this case where to orient the eyes (Fig. 1, Planner). Ultimately, the objective for the ideal planner is to improve task performance, but often it can be formalized as reducing uncertainty in task-relevant variables, such as the entropy of the distribution over the time of day. The plan is then executed, resulting in an action that leads to new sensory input (Fig. 1, Action). This closes the loop of perception and action that defines active sensing (Fig. 1, red arrow path). Although we describe these processes in discrete steps with a static stimulus and fixed task, in general, active sensing can be considered in real time with the stimulus and task changing continuously.

Active sensing as a form of exploration

As observer models have been extensively studied and reviewed [1, 2], we primarily focus here on the ideal planner which is the other key process in active sensing. In general, truly optimal planning is computationally intractable and we, therefore, need to consider approximations and heuristics. In fact, active sensing itself can be seen as emerging from such an approximation (see Box 1: Exploration, exploitation, and the value of information). The ultimate objective of behavior can be formalized as maximizing the total rewards that can be obtained in the long term [7]. This, in principle, requires considering the consequences of future actions, not only in terms of the rewards to which they lead, but also in terms of how they contribute to additional knowledge about the environment, which can be beneficial when planning actions in the more distant future. For example, when foraging for food, animals should choose actions that not only take them closer to known food sources but also yield information about potential new sources [8]. As this recursion is radically intractable, the most common approximation is to distinguish between actions that exploit current knowledge and seek to maximize future rewards, and actions that instead explore to improve knowledge of the environment [9].

Just as exploitation can be optimized to yield maximal rewards, so can exploration be optimized to yield maximal information about the environment. Exploration, thus optimized, is known as "active learning" [10, 11, 12]. In general, a huge variety of actions can be used for active learning, from turning your head towards a sound source to opening your browser to check on the meaning of an unknown phrase (such as 'epistemic disclosure' [13]). Indeed, the way participants choose queries in categorization tasks [14, 15], locate a region of interest in a variant of the game of battleship [16], or choose questions in a 20-questions-like situation [17] has been shown to be optimized for learning about task-relevant information. Active sensing, more specifically, can be regarded as the realm of active learning which involves actions that direct your sensors to gain information about quantities that change on

relatively fast time scales roughly corresponding to the time scale of single trials in laboratory-based tasks (see Box 2: Information maximization).

Common approximations in active sensing strategies

Task-related active sensing, our main focus here, makes the further approximation of breaking up "life" into discrete, known tasks. In contrast, curiosity-driven forms of information seeking [13, 18, 19] may be understood as optimized for improving an internal model for whatever task may come our way. While laboratory tasks for studying task-related active sensing are usually designed to minimize the trade off between exploration and exploitation, such that rewards only depend on task-relevant information [3, 20], curiosity-driven information seeking is often demonstrated in tasks that do have an information-independent reward structure, such that participants can be shown to actively forego these rewards for additional information [21].

Even exploration or exploitation by themselves are still intractable due to the exponential explosion of future possibilities that need to be considered. For example, maximal exploitation in the game of chess would require considering a very large number of future sequences of movement to maximize task success. Therefore, several simpler heuristics have been proposed to describe behavior. The simplest heuristic considers only the consequence of the next action, and hence is termed greedy or myopic. Thus, in the context of exploration, ideal planners are typically formalized in a way that they seek the single action that will maximize information gain or an equivalent objective (See Box 2), without considering the possibility that an action leading to suboptimal immediate information gain may allow other actions later with which total information gain would eventually become larger [3, 22, 20]. Interestingly, the strategy of greedily seeking task-relevant information has been successful in describing both human eye movements [22, 20] and the foraging trajectories of moths [23] and worms [24]. However, for some tasks it has been shown that several future actions [25, 26] are considered when planning and that there can be a trade-off between the depth of planning and the number of plans considered when there are time constraints on planning [27].

Taxonomy of efficient sensing strategies

Active sensing can also be considered as part of a spectrum of strategies that organisms have developed to improve the efficiency of their sensory processing. Within this context we can consider restricted versions of the full active sensing strategy based on the extent to which the observer and planner form a full closed loop (Fig. 1, table). If the dependence on the observer is removed for all but the initial sensory input, the strategy becomes open-loop because all future actions are planned at the first time step and not updated. Examples of this sensing strategy include face recognition [28] and texture identification [29], in which the first observation is used to identify and prioritize regions of interest, and subsequent saccades follow the planned sequence without needing to update the plan.

If observer dependence is completely removed, the sensing strategy becomes entirely preplanned. Note that this may still allow the planner to depend on the sensory input, e.g.

through the bottom-up salience of different visual features [30, 31]. A preplanned strategy can be task-dependent or independent. An example of the former is the method developed for a ship to search for a submerged submarine in which a logarithmic spiral search was shown to be optimal and successfully employed [32]. An example of a task-independent preplanned strategy is a Lévy flight that mimics certain summary statistics of human eye movement under many different scenarios [33]. In a Bayesian framework, an optimal preplanned task-independent strategy could be obtained by averaging over a prior probability distribution that specifies which tasks are more probable than others (as well as the relative importance of the tasks).

Active and preplanned strategies are not mutually exclusive in that actual patterns of sensor movements seem to be influenced by some mixture of them. For example, eye movements have been found to be best predicted by a combination of several bottom-up (such as saliency) and top-down (such as reward) factors [34, 35, 36] where the contribution of these different factors can depend on the timing requirement of the task and the time course and type of eye movement [37, 38, 39, 40, 41].

Generalized efficient sensing

We can use the same taxonomy we have developed to consider active sensing beyond simply generating actions to move our sensory apparatus. This allows us to consider the more general problem of allocating our limited perceptual processing resources and, thereby, place apparently disparate aspects of sensory and perceptual processing within our unifying framework.

The design principles underlying the organization of many sensory systems both at the morphological and the neural level have often been argued to be optimized for efficient information gathering. For example, predators that need accurate depth vision for catching their prey often have forward-looking eyes and vertical pupils, while preys that need to be able to avoid predators from as wide a range of directions as possible often have laterally placed eyes with horizontal pupils [42]. Moreover, our sensors are distributed so as to increase resolution at strategic regions (e.g. fingertips and foveas). At the neural level, classical forms of efficient coding include the optimization of receptor and sensory neuron tuning curves and receptive fields based on the statistics of inputs ("natural stimulus statistics") [43]. Although these preplanned strategies are arguably optimized over an evolutionary time scale according to the tasks that animals have to achieve, we can think of these as corresponding to a task-independent strategy as they do not depend on the particular task the animal is pursuing at any one moment.

There are also aspects of perceptual processing that can be considered task-independent preplanned strategies that nevertheless depend on the stimulus. At the perceptual level, bottom-up saliency may indicate how much an image feature is informationally optimal "on average" and may thus serve as a proxy for the "real" information value that would be determined by top-down (task-dependent) processes [44]. At the sensory level, stimulus statistics can vary both spatially and temporally, and efficient coding has been successfully applied to account for the spatio-temporal context-dependence of receptive field properties

[45]. Moreover, there exists mechanisms for the early filtering of sensory information to remove its predictable components based on the current action (for example, [46, 47]), presumably to save resources for processing unpredictable sensory inputs which thus have higher information content. While these processes are still task-independent, their stimulus-dependence increases information efficiency and, in that regard, takes them closer to fully task-dependent active sensing.

Finally, attention can dynamically change receptive fields and the allocation of perceptual processing resources in a task-dependent manner, thus corresponding to a task-dependent strategy. One example of how the brain filters out task-irrelevant information in purely perceptual tasks is the phenomenon of inattentional blindness, in which people fail to notice prominent stimuli in the visual scene that are irrelevant to the task that they are performing [48]. Similarly, in motor tasks, subjects are often only aware of large sensory input changes that have a bearing on the task at the precise time of the change, and are unaware of such changes otherwise [49]. Rather than just filtering out irrelevant information, the perceptual system can also adapt more finely in a task-dependent manner how it distributes resources to processing stimuli. In an object localization task, while fixation locations did not seem to be adaptively chosen, participants' functional field of view doubled through learning [50]. Thus, participants still seem to have adopted an active sensing strategy in this more general sense, but one which did not include changing the ways in which they overtly moved their sensors (eyes). In contrast, in a face identification task allowing a single fixation, improvement was brought about by a mixture of overt (eye movement, 43%) and covert (improved processing, 57%) active sensing strategies [51]. Attention can also be updated moment-by-moment depending on sensory evidence. For example, subjects can solve a visual maze while fixating and presumably using attention to search for the exit [52]. Such continuous updating of attention is equivalent to a form of closed-loop active sensing.

Optimal stopping as active sensing

A particularly interesting aspect of active sensing, that is often treated formally but separately from other aspects, is the "optimal stopping problem" in sequential sampling. This problem involves choosing the duration (rather than the location) of sensory sampling, which is relevant as directing sensors to the same place longer usually yields more information. There is a trade-off however, as longer sampling of the same place usually entails an opportunity cost, losing out on other, potentially more informative locations, or more rewarding actions altogether. This is sometimes explicitly enforced in a task by time constraints on a final decision. While in higher level cognitive domains, "stopping decisions" are typically suboptimal (e.g. solving the "secretary problem" [53]), perceptual stopping has been shown to be near optimal in a task that was specifically designed to provide a slow accumulation of information [54]. Perceptual stopping also forms the basis of one of the key windows into infant cognition, where a standard experimental design measures looking times for different stimuli [55, 56]. Interestingly, information is rarely quantified explicitly in these experiments, and when there has been an effort made to quantify it, total entropy rather than information has been used with mixed results [57]. This may be because in active sensing, maximising total entropy rather than information (Box 2) can be greatly suboptimal in information efficiency, as has been shown in the context of

object categorisation [20]. In visual search, the duration and location of fixations has been studied in an integrated framework, using a control-theoretic approach, in which the objective function included costs on time and effort, such that there was a trade-off between the information that increases with the fixation duration and the cost of the prolongation of the task time [58].

Hallmarks of active sensing

We can use the taxonomy developed above to define the hallmarks of active sensing. A necessary, but not sufficient, condition to determine whether sensing is active is that the actions should be task-dependent. Yarbus, in his pioneering work in vision, showed that even when viewing the same image, humans employed distinctively different eye movements when required to make different inferences about the image [59]. Recent studies confirmed such task dependence by showing that it was possible to predict the task simply from the recording of an individual's eye movements in the Yarbus setting [60, 61], or the target they were looking for in a search task [62], or the moment-by-moment goal they were trying to achieve in a game setting [63].

Another requirement for active sensing is that different sensor and actuator properties should lead to different planning behaviors for a given task. Sensor dependence of eye movements has been demonstrated by showing that in conditions in which foveal vision is impaired (such as at low light levels, or with an artificially induced scotoma) the pattern of eye movements adapts such that it becomes fundamentally different from that in normal vision [64, 65] and near-optimal under the changed conditions [65]. Similarly, for motor dependence, patients with a cerebellar movement disorder were shown to employ eye movements that were consistent with an optimal strategy based on a higher motor cost compared to normal [66]. Indeed, several studies have suggested that motor costs may affect the active sensing strategy if participants trade-off informativeness for movement effort [58, 6].

A more stringent condition for evidence of active sensing is that actions should depend both on the task and the observations already made in the task. In their seminal work, Najemnik and Geisler [3] formalized such an active sensing strategy for a visual search task and demonstrated that several features of human eye movements are consistent with the model's behavior, including inhibition of return, the tendency not to return to a recently fixated region. More direct evidence comes from studies that quantify the informativeness of potential fixation locations and show that humans choose locations with high information value. For example, recent studies have shown that humans can direct their eyes to locations that are judged informative [67, 20] and that they make faster eye movements to such locations [68, 69].

Formalizing and quantifying active sensing

Establishing conclusive evidence for active sensing ultimately requires constructing explicit models of what a theoretically optimal active sensor would do under the same conditions that are used to test experimental participants, and having quantifiable measures of the

degree of match between the model and participants' behavior. Visual search has become a major paradigm in the study of active sensing as it typically allows a straightforward observer model which simply represents the posterior probability of the target over the potential locations. The planner then selects the fixation location that leads to the greatest probability of correctly identifying the target location – which may not be the one closest to the predicted target location [3, 70]. The use of visual search tasks is also attractive as their analysis is amenable to using simple measures of performance such as the number of eye movements to find the target [3].

However, more complex, naturalistic tasks such as categorization and object recognition, do not map directly onto visual search and require more sophisticated models. In such cases, a direct generalisation of the information maximisation objective is prohibitive, because it requires the planner to run "mental simulations" for all possible outcomes of all putative actions, and in each case compute the corresponding update to the posterior over a potentially high-dimensional and complex hypothesis space, and finally compute the entropy of each posterior. For example, when planning the next saccade, the informativeness of several putative fixation locations needs to be compared. As we don't yet know what we will see at those locations, a range of possible visual inputs need to be considered for each, and for each of these inputs the resulting posterior needs to be computed. Fortunately, it can be shown that information can be computed in a simpler (though mathematically equivalent) form that does not require simulated updates of the posterior (See Box 2), and only needs to compute entropy in the space of observations (e.g. the colour of an object seen at the fovea), which is typically much lower dimensional than the space of hypotheses (all possible object classes) [71, 20]. This not only makes modelling the information maximizing active sensor practical, but may also offer a more plausible algorithmic view on how the brain implements the active sensor. Such an approach has been successfully applied to a visual categorization task using an observer that combined information from multiple fixations to maintain a posterior over possible categories of a stimulus and a planner that selected the most informative location given that posterior [20].

Several other model objectives that the active planner is trying to optimise have also been proposed. For example, the visual search task of Najemnik and Geisler [3] used the objective of maximizing probability gain (or task performance) and in a binary categorization task in which participants could reveal one of two possible features, this objective was also better at describing their behavior than information gain [72]. Moreover, in a task in which subjects were required to perform ternary categorization, their queries were best described by a max-margin objective, where each query was optimized to resolve uncertainty between two categories at a time rather than all three simultaneously [73]. Furthermore, action objectives have been shown to adapt to task demands, with simple hypothesis-testing performed under time pressure and information maximization when temporal demands were relaxed [74]. These studies leave open the possibility that the objectives underlying different modalities of active sensing (such as eye movement or cognitive choices) may be fundamentally different (maximizing information or performance directly).

Measuring performance in naturalistic tasks also presents challenges. Classical measures of eye movements typically rely on the geometric features of performance (scan paths,

direction, number of fixations to target) [3, 22]. More recently, ideal observers have not only been used to model the task but also to obtain a fixation-by-fixation measure of performance. For example, for an information maximizing active sensor, the ideal observer model could be used to quantify the amount of information accumulated about image category with each fixation within a trial [20]. This information-based measure of performance was more robust than directly measuring distances between optimal and actual scan paths because multiple locations are often (nearly) equally informative when planning the next saccade. This means that fixation locations that are far away from the optimal fixation location, and would therefore be deemed highly suboptimal by geometric measures, can nevertheless be close to optimal in terms of information content. For example, examining the sky or the ground for cast shadows could be almost equally informative as to the time of day in the picture shown in Fig. 1, yet, these regions of the image are far away. Thus, looking at the ground may appear geometrically very suboptimal, while informationally it is near-optimal. When such information-based measures were used, the efficiency in the planning of each eye movement was shown to be around 70% [20].

Conclusion

The use of the ideal observer-planner framework has allowed both qualitative and quantitative description of human sensing behaviors and thus offers insights into the computational principles behind actions and sensing. The next big challenge is to construct a flexible representation that connects sensory inputs to large classes of high-level natural tasks, such as estimating the wealthiness, age or intent of a person [59], and feed this construction into an ideal observer-planner model. Although constructing such algorithms is far from trivial, as Bayesian inference algorithms using structured probabilistic representations become increasingly sophisticated and powerful in matching human-level cognition [75], their integration into the active sensing framework will provide practical solutions for modeling such tasks. This advance will not only allow comparison with experiments on more natural tasks for a richer understanding of the active sensing process, but will also be a step towards applying such theories to improve learning in real-world situations.

Acknowledgements

References

[1]. Geisler WS. Ideal Observer Analysis. The visual neurosciences. 2003:825–837.

[2]. Geisler WS. Contributions of ideal observer theory to vision research. Vision Res. 2011; 51:771–781. [PubMed: 20920517]

[3]. Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. Nature. 2005; 434:387–391. [PubMed: 15772663]

[4]. Geisler WS, Perry JS, Najemnik J. Visual search: the role of peripheral information measured using gaze-contingent displays. J Vis. 2006; 6:858–873. [PubMed: 17083280]

[5]. Horowitz TS, Wolfe JM. Visual search has no memory. Nature. 1998; 394:575–577. [PubMed: 9707117]

[6]. Le Meur O, Liu Z. Saccadic model of eye movements for free-viewing condition. Vision Res. 2015; 116:152–164. [PubMed: 25724662] [• This study presents a model of free viewing that incorporated bottom-up saliency, visual memory, occulomotor bias, and a sampling-based action objective. For several datasets of images, the distributions of saccades predicted by the model are very similar to those made by human.]

[7]. Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT Press; Cambridge, Massachusetts, USA: 1998.

[8]. Averbeck BB. Theory of Choice in Bandit, Information Sampling and Foraging Tasks. PLOS Comput Biol. 2015; 11:e1004164. [PubMed: 25815510]

[9]. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006; 441:876–879. [PubMed: 16778890]

[10]. MacKay DJC. Information-based objective functions for active data selection. Neural Comput. 1992; 4:590–604.

[11]. Settles B. Active Learning Literature Survey. Technical Report. 2010

[12]. Borji A, Itti L. Bayesian optimization explains human active search. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, editorsAdvances in Neural Information Processing Systems 26; NIPS; 2013. 55–63.

[13]. Gottlieb J, Oudeyer P-Y, Lopes M, Baranes A. Information-seeking, curiosity, and attention: computational and neural mechanisms. Trends Cogn Sci. 2013; 17:585–593. [PubMed: 24126129]

[14]. Markant DB, Gureckis TM. Is it better to select or to receive? Learning via active and passive hypothesis testing. J Exp Psychol Gen. 2014; 143:94–122. [PubMed: 23527948] [• This study shows that the queries we choose for ourself when performing a task are more informative, on average, than when we are provided with queries chosen by others when performing the same task. This result highlights the personalized nature of active sensing in cognition.]

[15]. Castro RM, Kalish C, Nowak R, Qian R, Rogers T, Zhu X. Human Active Learning. In: Koller D, Schuurmans D, Bengio Y, Bottou L, editorsAdvances in Neural Information Processing Systems 21; Cambridge, MA: NIPS, MIT Press; 2009. 241–248.

[16]. Gureckis TM, Markant DB. Active Learning Strategies in a Spatial Concept Learning Game. Proceedings of the 31st Annual Conference of the Cognitive Science Society; 2009. 3145–3150.

[17]. Nelson JD, Divjak B, Gudmundsdottir G, Martignon LF, Meder B. Children's sequential information search is sensitive to environmental probabilities. Cognition. 2014; 130:74–80. [PubMed: 24184396]

[18]. Barto AG. Intrinsically Motivated Learning in Natural and Artificial Systems. Springer Berlin Heidelberg; Berlin, Heidelberg: 2013.

[19]. Gordon G, Ahissar E. Hierarchical curiosity loops and active sensing. Neural Networks. 2012; 32:119–129. [PubMed: 22386787]

[20]. Yang SC-H, Lengyel M, Wolpert DM. Active sensing in the categorization of visual patterns. eLife. 2016 [•• This study focused on a higher-level visual task that requires the integration of visual information from multiple observations to categorise a visual image. In addition to showing that human eye movements depended on the stimulus and are hence active, by using an ideal observer-actor model, the authors were able to estimate, for the first time, that in a high-level categorization task humans are about 70% efficient in planning each saccade.]

[21]. Bromberg-Martin ES, Hikosaka O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron. 2009; 63:119–126. [PubMed: 19607797]

[22]. Renninger LW, Verghese P, Coughlan J. Where to look next? Eye movements reduce local uncertainty. J Vis. 2007; 7:6.

[23]. Vergassola M, Villermaux E, Shraiman BI. 'Infotaxis' as a strategy for searching without gradients. Nature. 2007; 445:406–409. [PubMed: 17251974]

[24]. Calhoun AJ, Chalasani SH, Sharpee TO. Maximally informative foraging by Caenorhabditis elegans. eLife. 2014; 3:e04220. [• This study shows that the foraging trajectories of worms is governed by active sensing. The mathematical model that generated these trajectories is the same one used to describe the foraging trajectories of moths [23] and human eye movements [20], suggesting that active sensing is a general principle employed across species.]

[25]. Vries JPD, Hooge ITC, Verstraten FAJ. Saccades Toward the Target Are Planned as Sequences Rather Than as Single Steps. Psychol Sci. 2014; 25:215–223. [PubMed: 24166857] [•• This study show active sensing strategies can plan saccades up to three steps ahead, but given enough time, they can be modified according to how the visual scene changes configuration.]

[26]. Shen K, Pare M. Predictive Saccade Target Selection in Superior Colliculus during Visual Search. J Neurosci. 2014; 34:5640–5648. [PubMed: 24741054]

[27]. Snider J, Lee D, Poizner H, Gepshtein S, Jolla L. Prospective optimization with limited resources. PLoS Comput Biol. 2015; 14:e1004501.

[28]. Peterson MF, Eckstein MP. Looking just below the eyes is optimal across face recognition tasks. Proc Natl Acad Sci USA. 2012; 109:E3314–E3323. [PubMed: 23150543]

[29]. Toscani M, Valsecchi M, Gegenfurtner KR. Optimal sampling of visual information for lightness judgments. Proc Natl Acad Sci USA. 2013; 110:11163–11168. [PubMed: 23776251]

[30]. Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res. 2000; 40:1489–1506. [PubMed: 10788654]

[31]. Borji A, Sihite DN, Itti L. Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study. IEEE Trans Image Process. 2013; 22:55–69. [PubMed: 22868572]

[32]. Langetepe E. On the optimality of spiral search. SODA '10 Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms; Society for Industrial and Applied Mathematics; 2010. 1–12.

[33]. Brockmann D, Geisel T. The ecology of gaze shifts. Neurocomputing. 2000; 32–33:643–650.

[34]. Navalpakkam V, Koch C, Rangel A, Perona P. Optimal reward harvesting in complex perceptual environments. Proc Natl Acad Sci USA. 2010; 107:5232–5237. [PubMed: 20194768]

[35]. Clavelli aKaratzas D, Llados J, Ferraro M, Boccignone G. Modelling Task-Dependent Eye Guidance to Objects in Pictures. Cog Comp. 2014; 6:558–584.

[36]. Miconi T, Groomes L, Kreiman G. There's Waldo! A Normalization Model of Visual Search Predicts Single-Trial Human Fixations in an Object Search Task. Cereb Cortex. 2015:1–19. [PubMed: 23926113]

[37]. Schütz AC, Trommershäuser J, Gegenfurtner KR. Dynamic integration of information about salience and value for saccadic eye movements. Proc Natl Acad Sci USA. 2012; 109:7547–52. [PubMed: 22529390]

[38]. Schütz AC, Lossin F, Gegenfurtner KR. Dynamic integration of information about salience and value for smooth pursuit eye movements. Vision Res. 2015; 113:169–178. [PubMed: 25175113]

[39]. Paoletti D, Weaver MD, Braun C, van Zoest W. Trading off stimulus salience for identity: A cueing approach to disentangle visual selection strategies. Vision Res. 2015; 113:116–124. [PubMed: 25152318]

[40]. Kovach CK, Adolphs R. Investigating attention in complex visual search. Vision Res. 2014; 116:127–141. [PubMed: 25499190]

[41]. Ramkumar P, Fernandes H, Kording K, Segraves M. Modeling peripheral visual acuity enables discovery of gaze strategies at multiple time scales during natural scene search. J Vis. 2015; 15:19.

[42]. Banks MS, Sprague WW, Schmoll J, Parnell JA, Love GD. Why do Animal Eyes have Pupils of Different Shapes? Sci Adv. 2015; 1:e1500391. [PubMed: 26601232]

[43]. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annual review of neuroscience. 2001; 24:1193–1216.

[44]. Bruce NDB, Tsotsos JK. Saliency, attention, and visual search: An information theoretic approach. Journal of Vision. 2009; 9:1–24.

[45]. Schwartz O, Hsu A, Dayan P. Space and time in visual context. Nature reviews. Neuroscience. 2007; 8:522–535. [PubMed: 17585305]

[46]. Seki K, Perlmutter SI, Fetz EE. Sensory input to primate spinal cord is presynaptically inhibited during voluntary movement. Nature neuroscience. 2003; 6:1309–1316. [PubMed: 14625555]

[47]. Bays PM, Wolpert DM. Computational principles of sensorimotor control that minimize uncertainty and variability. The Journal of physiology. 2007; 578:387–396. [PubMed: 17008369]

[48]. Most SB, Scholl BJ, Clifford ER, Simons DJ. What You See Is What You Set: Sustained Inattentional Blindness and the Capture of Awareness. Psychological Review. 2005; 112:217–242. [PubMed: 15631594]

[49]. Triesch J, Ballard DH, Hayhoe MM, Sullivan BT. What you see is what you need. Journal of Vision. 2003; 3:86–94. [PubMed: 12678628]

[50]. Holm L, Engel S, Schrater P. Object learning improves feature extraction but does not improve feature selection. PLoS ONE. 2012; 7:e51325. [PubMed: 23251499]

[51]. Peterson MF, Eckstein MP. Learning optimal eye movements to unusual faces. Vision Res. 2014; 99:57–68. [PubMed: 24291712]

[52]. Crowe DA. Neural Activity in Primate Parietal Area 7a Related to Spatial Analysis of Visual Mazes. Cerebral Cortex. 2004; 14:23–34. [PubMed: 14654454]

[53]. Seale DA, Rapoport A. Sequential decision making with relative ranks: An experimental investigation of the "secretary problem". Organ Behav Hum Dec. 1997; 69:221–236.

[54]. Drugowitsch J, DeAngelis GC, Angelaki DE, Pouget A. Tuning the speed-accuracy trade-off to maximize reward rate in multisensory decisionmaking. Elife. 2015; 4:e06678. [PubMed: 26090907] [• By analyzing the speed-accuracy trade-off in a multisensory integration task, this study shows that humans shown to choose optimally when to stop integrating evidence and make a decision so as to maximise overall reward rate.]

[55]. Aslin RN, Fiser J. Methodological challenges for understanding cognitive development in infants. Trends Cogn Sci. 2005; 9:92–98. [PubMed: 15737816]

[56]. Csibra G, Hernik M, Mascaro O, Tatone D, Lengyel M. Statistical treatment of looking-time data. Dev Psychol. 2016; 52:521–36. [PubMed: 26845505]

[57]. Kidd C, Piantadosi ST, Aslin RN. The Goldilocks effect: human infants allocate attention to visual sequences that are neither too simple nor too complex. PLoS ONE. 2012; 7:e36399. [PubMed: 22649492]

[58]. Ahmad S, Huang H, Yu AJ. Cost-sensitive Bayesian control policy in human active sensing. Front Hum Neurosci. 2014; 8:955. [PubMed: 25520640]

[59]. Yarbus AL. Eye movements and Vision. Plenum Press, New York; New York, USA: 1967.

[60]. Borji A, Itti L. Defending Yarbus : Eye movements reveal observers' task. J Vis. 2014; 14:29.

[61]. Haji-Abolhassani A, Clark JJ. An inverse Yarbus process: Predicting observers' task from eye movement patterns. Vision Res. 2014; 103:127–142. [PubMed: 25175112]

[62]. Borji A, Lennartz A, Pomplun M. What do eyes reveal about the mind? Neurocomputing. 2015; 149:788–799.

[63]. Rothkopf CA, Ballard DH, Hayhoe MM. Task and context determine where you look. J Vis. 2007; 7:16.

[64]. Walsh DV, Liu L. Adaptation to a simulated central scotoma during visual search training. Vision Res. 2014; 96:75–86. [PubMed: 24456805]

[65]. Paulun VC, Schütz AC, Michel MM, Geisler WS, Gegenfurtner KR. Visual search under scotopic lighting conditions. Vision Res. 2015; 113:155–168. [PubMed: 25988753] [•• Following the approach of [3, 70], this study shows that human eye movements are similar to those predicted by an ideal observer-actor model under both photopic and scotopic conditions during visual search. The clear difference in eye movements in the two conditions shows that humans can adapt their saccade planning to account for changes in sensor properties in a near-optimal way, which is a hallmark of active sensing.]

[66]. Veneri G, Federico A, Rufa A. Evaluating the influence of motor control on selective attention through a stochastic model: the paradigm of motor control dysfunction in cerebellar patient. BioMed Res Int. 2014; 2014 162423.

[67]. Janssen CP, Verghese P. Stop before you saccade: Looking into an artificial peripheral scotoma. J Vis. 2015; 15:7.

[68]. Bray TJP, Carpenter RHS. Saccadic foraging: reduced reaction time to informative targets. Eur J Neurosci. 2015; 41:908–913. [PubMed: 25659260]

[69]. Kowler E, Aitkin CD, Ross NM, Santos EM, Zhao M. Davida Teller Award Lecture 2013 : The importance of prediction and anticipation in the control of smooth pursuit eye movements. J Vis. 2014; 14:10.

[70]. Najemnik J, Geisler WS. Eye movement statistics in humans are consistent with an optimal search strategy. J Vis. 2008; 8:4.

[71]. Houlsby N, Huszár F, Ghahramani Z, Lengyel M. Bayesian active learning for classification and preference learning. arXiv. 2011 1112.5745.

[72]. Nelson JD, McKenzie CRM, Cottrell GW, Sejnowski TJ. Experience matters: Information acquisition optimizes probability gain. Psychol Sci. 2010; 21:960–969. [PubMed: 20525915]

[73]. Markant DB, Settles B, Gureckis TM. Self-Directed Learning Favors Local, Rather Than Global, Uncertainty. Cognitive Sci. 2015:1–21. [• This study addresses the question of what action objective do people use by designing a three-category categorization task that can empirically distinguish between two heuristics strategies, namely, one that simultaneously reduces uncertain about all three categories and one that disambiguates the top two. The results show that humans prefer the latter strategy. The study is a nice example of how one can narrow in to the computational principles behind human actions using an ideal observer-actor framework.]

[74]. Coenen A, Rehder B, Gureckis T. Strategies to intervene on causal systems are adaptively selected. Cognitive Psychol. 2015; 79:102–133.

[75]. Lake BM, Salakhutdinov R, Tenenbaum JB. Human-level concept learning through probabilistic program induction. Science. 2015; 350:1332–1338. [PubMed: 26659050]

[76]. Bellman R. Dynamic Programming. Princeton University Press; 1957.

[77]. Kaelbling LP, Littman ML, Cassandra AR. Planning and acting in partially observable stochastic domains. Artificial Intelligence. 1998; 101:99–134.

[78]. Berger JO. Statistical decision theory and Bayesian analysis. Springer Science & Business Media; 2013.

Box 1

Exploration, exploitation, and the value of information

The ideal observer performs inference, using Bayes' rule, about several variables characterizing the state of the environment simultaneously, $\mathbf{x}$ (e.g. what objects are present in the scene, their configuration, features, etc.), given the sensory inputs up to the current moment, $z_{0:t}$, and an internal model of the environment and its sensory apparatus, M:

$$\mathbb{P}(\mathbf{x}|z_{0:t}, M) \propto \mathbb{P}(z_{0:t}|\mathbf{x}, M)\, \mathbb{P}(\mathbf{x}|M) \quad (1)$$

A task defines a reward function over actions, a, that depends on the true state of the environment, $\mathcal{R}(a, \mathbf{x})$. There are two aspects of the reward function that make it task-dependent. First, it typically depends only on a subset of state variables, $\mathbf{x}_T$ (e.g. defining whether it is the time of day, or the age of the people in the picture of Fig. 1 that you want to estimate), and second, it has a particular functional form (e.g. determining how much under- or overestimating the time of day matters).

In general, an agent navigating the environment cannot use the reward function directly to select actions for two reasons. First, it does not directly observe the true state of the environment, so it must base its decisions on its *beliefs* about it as given by the ideal observer (Eq. 1). Second, its objective is to maximize total reward in the long run, and so the consequences of its actions in terms of how they change environmental states (or, more precisely, the agent's beliefs about them) must also be taken into account. Thus, we can write the value of an action, $\mathcal{Q}$, as the sum of its immediate and future values, each depending on the agent's current beliefs:

$$\mathcal{Q}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, M)) = \mathcal{Q}_{\text{now}}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, M)) + \mathcal{Q}_{\text{future}}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, M)) \quad (2)$$

This equation is the well-known Bellman optimality equation [76] but, rather than expressing values directly for the states of the environment $\mathbf{x}$, as typically done, here it is applied to the *belief* "states" of the agent, $\mathbb{P}(\mathbf{x}/z_{0:t}, M)$, i.e. the beliefs it holds about those states [77].

The immediate value $\mathcal{Q}_{\text{now}}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, M))$ can be computed as the reward expected under the current posterior distribution provided by the ideal observer ("expected" gain in Bayesian decision theory [78]):

$$\mathcal{Q}_{\text{now}}(a) = \sum_{\mathbf{x}} \mathcal{R}(a, \mathbf{x})\, \mathbb{P}(\mathbf{x}|z_{0:t}, M) \quad (3)$$

The future value $\mathcal{Q}_{\text{future}}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, M))$ has a more complicated form that is generally computationally intractable, but can be shown to depend on two factors: 1. the way the

action leads to future rewards by steering the agent into future environmental states, and 2. the way it leads to new observations based on which the ideal observer can update its beliefs so that its uncertainty decreases, allowing better informed decisions and therefore higher rewards in the future. These two factors are commonly referred to as "exploitation" and "exploration", respectively, and are often treated separately due to the intractability of $\mathcal{Q}_{\text{future}}(a, \mathbb{P}(\mathbf{x}|z_{0:t}, \mathbf{M}))$, but as we see from Eq. 2 they both factor into the same greater objective of maximizing value (i.e. "exploitation") in terms of belief states. Importantly, evaluating both factors requires recursion into the future so they are each intractable even when treated separately.

Box 2

Information maximization

A useful proxy for the explorational value of an action, *a* (see Box 1), is the information that one expects to gain about environmental state, **x**, by taking that action. The (Shannon) information exactly expresses the expected reduction in uncertainty about **x**:

$$\mathscr{I}(\mathbf{x}, z_{t+1} | a) = \mathscr{H}(\mathbf{x} | z_{0:t}, \mathrm{M}) - \left\langle \left\langle \mathscr{H}(\mathbf{x} | z_{0:t+1}, \mathrm{M}) \right\rangle_{\mathbb{P}(z_{t+1} | a, \mathbf{x}, \mathrm{M})} \right\rangle_{\mathbb{P}(\mathbf{x} | z_{0:t}, \mathrm{M})} \quad (4)$$

where $\langle \cdot \rangle$ denotes an average according to the specific distribution given by the subscript, and

$$\mathscr{H}(\mathbf{x}) = - \left\langle \log \mathbb{P}(\mathbf{x}) \right\rangle_{\mathbb{P}(\mathbf{x})} = - \sum_{\mathbf{x}} \mathbb{P}(\mathbf{x}) \log \mathbb{P}(\mathbf{x}) \quad (5)$$

is the entropy of $\mathbb{P}(\mathbf{x})$ quantifying uncertainty about **x**. The first term of Eq. 4 expresses the uncertainty about the state of the environment, **x**, under the current posterior, $\mathbb{P}(\mathbf{x}/z_{0:t}, \mathrm{M})$, while the second term expresses the average expected uncertainty for an updated posterior, $\mathbb{P}(\mathbf{x}/z_{0:t+1}, \mathrm{M})$, once we make a new observation, z_{t+1}, upon executing action *a*. The information we gain by this action is thus the reduction of uncertainty as we go from the current posterior to a fictitious new posterior, and an information maximizing ideal planner simply chooses the action *a* which maximizes this information:

$$a_{t+1} = \underset{a}{\mathrm{argmax}} \ \mathscr{I}(\mathbf{x}, z_{t+1} | a) \quad (6)$$

There are two ways in which Eq. 4 is merely an approximation to explorational value. First, it only considers the next action without performing the full recursion for future belief states and actions (hence is "greedy" or "myopic"). Second, explorational value typically depends on a task-specific reward function (because it is derived from Eq. 2) and thus it might favor gathering information about particular aspects (dimensions or regions) of **x** to which the reward function is particularly sensitive, while Eq. 4 treats all aspects of **x** equal and is thus agnostic as to the reward function. Nevertheless, when the reward function only depends on a subset of the variables in **x**, $\mathbf{x}_T$ (Box 1), as it usually does in most everyday and laboratory tasks, this aspect of the reward function can be taken into account by simply computing entropies over $\mathbf{x}_T$ rather than the full **x**. In this case, the resulting active sensing strategy is still *task-dependent*. However, when even this aspect of the current task is ignored, and so information is computed for a predetermined set of variables, or the average information for different sets of variables (corresponding to different potential tasks) is computed, the strategy becomes *task-independent*. This kind of task-independent strategy can also be thought of as *curiosity-driven*. When the

posterior is updated based on observations one expects *a priori*, rather than on actual observations, the strategy becomes *preplanned*. (See Figure 1 for the different strategies.)

Even in its approximate form, Eq. 4 is computationally prohibitive. This is because for each action, *a*, it requires computing (the entropy of) an updated posterior, $\mathcal{H}(\mathbf{x}/z_{0:t+1}, \mathbf{M})$, for all possible fictitious observations, z_{t+1}, for each state of the world, $\mathbf{x}$, that is thought possible according to the current posterior, $\mathbb{P}(\mathbf{x}/z_{0:t}, \mathbf{M})$. Fortunately, using the property that Shannon information is symmetric, it can be shown that Eq. 4 can be rewritten in a different, though mathematically equivalent form (for full derivation, see Ref. 71), which does not require fictitious posterior updates:

$$\mathcal{I}(\mathbf{x}, z_{t+1}|a) = \mathcal{H}(z_{t+1}|a, z_{0:t}, \mathbf{M}) - \left\langle \mathcal{H}(z_{t+1}|a, \mathbf{x}) \right\rangle_{\mathbb{P}(\mathbf{x}|z_{0:t}, \mathbf{M})} \qquad (7)$$

The terms in this form also have intuitive meaning: the first term formalizes the planner's total uncertainty about what the next observation, z_{t+1}, might be, including uncertainty due to the fact that we are also uncertain about the state of the environment, $\mathbf{x}$, whereas the second term expresses the average uncertainty we would have about z_{t+1} if we knew what $\mathbf{x}$ was. In other words, the informational value is high for actions for which the main source of uncertainty about their consequent observation is the uncertainty about $\mathbf{x}$ (ie. their uncertainty is potentially reducible) and not because these observations are just more inherently noisy (which implies irreducible uncertainty). In contrast, an alternative approach, called maximum entropy, would only care about total uncertainty, $\mathcal{H}(z_{t+1}|a, z_{0:t}, \mathbf{M})$, and therefore ignores this distinction between reducible and irreducible uncertainty.
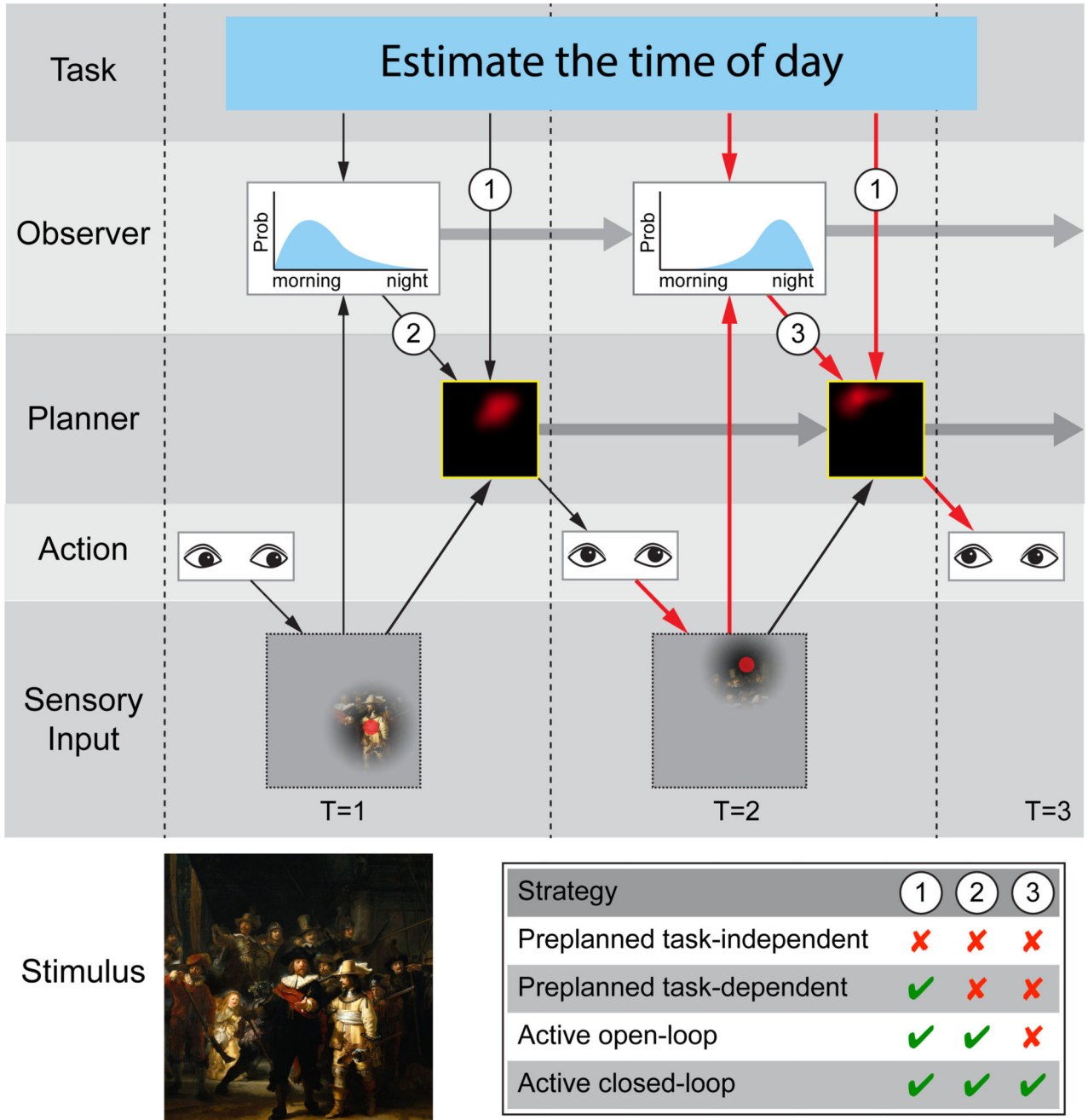
Figure 1. Active sensing framework and taxonomy.
An example of the temporal evolution of an active sensing strategy in which the task is to the estimate the time of day from an image (Stimulus: in this case the *The Night Watch* by Rembrandt). The gaze direction of the eye (Action) determines the fixation location (red dot in Sensory Input), and the Sensory Input is then limited by the typical fall-off of acuity with eccentricity (illustrated in Sensory Input). Given the Task and the Sensory Input, the Observer computes a probability distribution over the time of day. In this case the bright area fixated may suggest morning. Given the Task and the Observer's inference, the Planner

determines the expected value of moving the eyes to fixate different locations in the image (red intensity indicates value in the Planner). For example, the value could be the expected reduction in entropy in the Observer's inference distribution. The eyes can then be moved to a location with high or maximum value (such as examining the sky). This leads to new Sensory Input which updates the Observer's inference (to correctly suggest night). The larger gray arrows that link the Observer's inference and the Planner's action-objective map across time indicates that all the information from previous time steps are passed onto the current observer and planner. The red arrows in the figure highlights the components involved in a single loop of an active closed-loop strategy. The removal of specific sets of interactions (1, 2 & 3) leads to different sensing strategies (Table inset).