Smart, general, and situation specific sensors

Tjeerd Andringa[1,2]

[1] INCAS[3], Dr Nassaulaan 9, Assen, The Netherlands
[2] Artificial Intelligence, University of Groningen, Nijenborg 9, Groningen, The Netherlands
tjeerd@ai.rug.nl, tjeerdandringa@incas3.eu

Abstract. The need for central processing in animals and animats can be reduced with smart sensors that reliably indicate the presence of behaviorally relevant events and environmental affordances. Another, equally important, reduction can be realized through a smart choice of what needs to be sensed and how much detail the task requires. The combination of smart processing and 'expectation management' accounts for some of the core properties of cognition and might be used to understand animal perception and to design improved animat sensors. In particular the combination couples gist, sensorimotor, and global workspace theories of perception and behavior selection to smart sensors in animals and animats.

Keywords: Smart sensors, cognitive models, perception, audition, vision, attention, salience, meaning, gist, affordances, global workspace

1 Introduction

Animals perceive to act [1, 2]. For animals like insects with a small behavioral repertoire, detailed perception is a burden because it forces central processing to sift through an abundance of data to detect and combine behaviorally relevant cues that indicate what the environment affords. For these animals it is desirable that the perceptual system is exclusively sensitive to behaviorally relevant cues. In animals like humans, with a rich and precise behavioral repertoire, the perceptual system must perceive and combine a huge variety of behaviorally relevant details from a number of sensory modalities. Crucially, the system does not need to be sensitive to every detail all the time. On the contrary: most of the time the system does not need detailed perception at all. Moreover when behavior depends on detailed perception, the perceptual system needs only to be sensitive to the level of detail relevant for the behavior it is initiating or engaged in. Normally this is a small subset of all the details that can be perceived. This paper addresses the question of how to reconcile the seemingly conflicting demands of being able to perceive a huge range of specific details while processing (and representing) no more detail than necessary.

2 The orienting and the directed search mode of perception

We start with human perception as an example of a sensor system that reconciles the ability of exquisitely precise perception with the capacity not to be bothered with irrelevant details. In fact, humans even have words for the processes that address these demands [3]. For the auditory domain we have the words 'hearing' and 'listening' and for the visual domain 'seeing' and 'looking'. "I hear music" typically entails that I am are able to detect the presence of music. Likewise "I see a picture" acknowledges the visual existence of a picture. In contrast "I listen to music" and "I look at a picture" entails typically a more careful in depth analysis of the target. In addition the linguist forms 'listen *to*' and 'look *at*' suggest a directedness that is not required for 'hearing' and 'seeing'. 'Hearing' and 'seeing' can be described as an orienting process to detect potential relevance and – justifiably – ignore irrelevance. This process does not require attention – even while asleep we hear. 'Listening to' and 'looking at' can be thought of as a knowledge-guided directed search for task-relevant evidence on which intentional (non-automated or novel) behavior is based.

In vision research the visual gist phenomenon [4, 5] was established as a rapidly build-up, but meaningful, representation of a visual scene based on global scene statistics. A review of audition research [6] concluded that evidence for a similarly meaningful auditory gist was abundant and introduced the terms 'hearing' stage and 'listening' stage. According to [6] audition and vision share five properties:

1. only the gist of the scene or object is initially processed;
2. processing of the gist is rapid, based on simple statistics, but leads to the activation of interpretation hypotheses (i.e., possibly meaningful interpretations);
3. the focus of attention is deployed strategically according to prior knowledge and the perception of the gist;
4. conscious detailed analysis is possible on the part of the scene within the focus of attention;
5. only limited processing of the unattended parts of the scene occurs.

The properties 1 and 2 refer to the orienting process (hearing/seeing), while 4 and 5 refer to the knowledge guided directed search for evidence (listening/looking). Property 3 corresponds to a process that determines the salience of the stimulus. Stimuli of which the low-level properties are such that they may lead to a change in intentional behavior – salient stimuli – will be analyzed attentively, while stimuli deemed irrelevant allow the system to continue its current tasks. Salient, task-irrelevant stimuli like loud passing aircraft that (persistently) disrupt central processing are typically interpreted as annoying.

Figure 1 shows a functional model of the interplay of these two processes, updated from [3], where audition is in the service of behavioral option selection of sound producing events conform the listed conclusions. The left describes the orienting mode. In this mode auditory input is processed up to the point of estimated irrelevance [3]. Irrelevance is defined as having no consequences for the behavioral option selection process depicted in the upper-right of figure 1. Stimulus irrelevance occurs typically when the input complies with coarse expectations of the sound producing events in the environment. In this case the expectations of the sonic environment are automatically – without attention and conscious control – updated

with matching task-irrelevant sonic events. This ensures that the global world-model (or context) remains in sync with the environment. This process may have been observed directly as Stimulus Specific Adaptation (SSA) as a reduction of the response to a common sound [7]. The small blue circles in the layer that matches statistical expectations with signal input signifies this. During hearing we might be able to have an implicit gist-like awareness of basic or super-ordinate categories [4], such as passing cars or traffic, but we are unable to access signal details. However when part of the input does not comply with the expectations, a mismatch occurs and the hearing process is able to disrupt (non-auditory) task performance and, involuntarily, directed attention is aimed at the mismatching signal evidence. Behavioral option selection will now include whatever the stimulus interpretation affords.

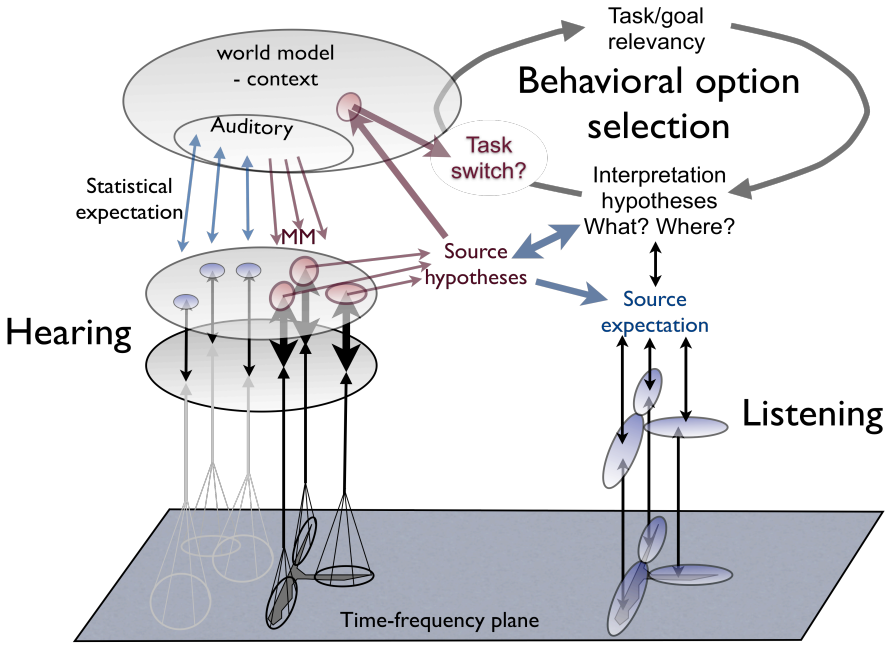

Fig. 1. Functional model of the orienting (*Hearing*, left) and the directed search (*Listening*, right) modes of audition. The switch from orienting – based on coarse signal statistics – to a directed search of detailed patters occurs when the *Behavioral Option Selection* process needs detailed perceptive input. This might be initiated by a mismatch (*MM*) between context dependent signal constituent expectations and the actual input. Red indicates a mismatch between signal and expectation and blue indicates a match.

While the orienting mode on the left is always active – albeit with high thresholds and a minimal content of the world-model during sleep – the right part is only active when we are conscious and a global workspace [8] of all task and situation relevant information (a correlate of consciousness) exists. In fact a salient sound during sleep can lead to the activation of a global workspace for intentional behavioral option selection. When awake, a pattern of mismatches between expectancies and signal input might indicate the presence of a source that is inconsistent with the current

mental tasks. In this case, mental tasks are suspended and the system involuntarily attends its perceptual modalities with preactivated object or event specific expectations. This corresponds to the knowledge-guided directed search for task-relevant evidence, which can be implemented through preactivating the tuning properties of processes closer to the input levels. This conforms with the reverse hierarchy theory of visual learning [9] that states that perceptual learning progresses from 'the top' backwards to the input levels, which have a better signal-to-noise ratio. Practice leads to rapid (direct [1]) and specific detectors in perceptual task experts that connect detailed specialized analyses knowledge, which influences behavioral option selection. These more specific detectors become available for gist-based [5] processing as in the left side of figure 1. This allows agents to include (new) environmental affordances [1] in the gist. Note that the presented model is consistent with the interplay of consciousness and attention in [10].

3 Conclusion

The use of smart sensors to reduce central processing demands constitutes half the perception-for-action optimization process. The other half is to adapt perception to environmental and task demands. This can be in the form of a knowledge-guided directed search for task-relevant signal details. Animals and animats with this property are, at minimal processing costs, open to the unexpected, while being able to process the expected to a tunable degree.

References

1. Gibson, J., The Ecological Approach to Visual Perception. Psychology Press (1986)
2. Bargh, J.A., Chartrand, T.L., The unbearable automaticity of being. American Psychologist, 54 (7), 462 -- 479 (1999)
3. Andringa, T.C., Audition: from sound to sounds. Machine Audition: Principles, Algorithms and Systems, W.W. Wang (ed), IGI Global Publisher (2010)
4. Potter, M.C., Short-term conceptual memory for pictures. Journal of Experimental Psychology: Human Learning and Memory, 2 (5), 509 -- 522 (1976)
5. Oliva, A., Gist of a scene, Neurobiology of Attention, 251--256 (2005)
6. Harding et al. Auditory gist perception: an alternative to attentional selection of auditory streams? Lecture Notes in Computer Science: Attention in Cognitive Systems. Theories and Systems from an Interdisciplinary Viewpoint, Springer-Verlag Berlin Heidelberg, 4840 pp. 399 -- 416 (2007)
7. Winkler, I., Denham, S.L., Nelken, I., et al. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends in Cognitive Sciences (2009) vol. 13 (12) pp. 532 -- 540
8. Dehaene, S., Naccache, L., Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition, 79 (1-2), 1 -- 37 (2001)
9. Ahissar, M., Hochstein, S., The reverse hierarchy theory of visual perceptual learning. Trends in Cognitive Sciences , vol. 8 (10), 457 -- 464 (2004)
10. Koch, C. and Tsuchiya, N., Attention and consciousness: two distinct brain processes. Trends in Cognitive Sciences, 11 (1), 16 -- 22, (2007)