Hear–Steer–Go

The focus group „Auditory and Visual Scene Analysis“ is looking for clues amongst the behavior of bats and other specialized animals

Fellows 2008/2009

by Carl Gierstorfer

It seems like an unlikely combination. But maybe to make some progress on one of the trickiest problems in neuroscience – perception – it takes just this: Computers, bats, and the combined brainpower of four fellows at the Wissenschaftskolleg in Berlin. What the group is trying to understand is so fundamental to higher organisms that it takes some reflection to see the complexity of the problem: How animals and humans can recognize patterns in a natural scene; how we can tell a tree from a forest, marvel at a flower in the twilight, and identify the object that just flew past us as a bird. Visual scene analysis, the problem is called. And for scientists it was easier to land a robot on Mars or plunge into the strange world of quantum mechanics than to understand what each and every one of us does in every instant of our waking life: create a coherent image of the world around us.

“What makes this problem so hard is that it’s not something we can introspect on,” says Michael Lewicki, Associate Professor of Computer Science at Carnegie Mellon University, a computational neuroscientist. “We are seeing things in an abstract way already, as surfaces,

as shapes or contours. It is almost like being caught up in the famous matrix where we see the world in a consistent manner, although the objects are in reality fragments of different textures, colors, or sizes. And not only that. The problem becomes fantastically complex if we take into account that the world around us is never static; the observer himself might move, or one or more parts of what he can see moves. In other words, the characteristics of objects change constantly."

So it's all rather complicated, especially with regard to the group's goal to study the problem not in isolation, but how it occurs in the complex natural environment. "We try to understand it on different levels," says Cynthia Moss, Professor of Psychology at the University of Maryland and the convener of the group. "We want to go all the way from perception down to activity patterns of neurons in the brain." Fortunately, each problem in biology has an ideal organism for analysis. And this organism, for some members of the group, is one of the strangest animals we know – the bat.

Bats are the most diverse group of mammals with more than 1150 species worldwide. Most of them hunt insects on the wing, but many species feed on other food items like fruit, nectar, other small vertebrates, or even blood. During their evolution, which started around fifty million years ago, bats independently invented flight and developed their special form of perception: echolocation. Just like the sonar of a ship, bats send out sound waves from the mouth like other mammals or from the nose in some species to create a three-dimensional picture of the world around them by analyzing the incoming echo.

It's this active component that makes this system so attractive for studying auditory scene analysis. "The bat is sending out these signals to study and sample its environment," says Moss. "That brings us an advantage because we can study these signals and learn something about the information the animal is seeking for a particular task." It's like observing a burglar with a flashlight: by studying the movement of the flashlight and how the burglar responds to what he sees, the observer could learn a lot about his intentions and how the burglar finds his way in an unfamiliar environment.

Yet sound waves, in contrast to light waves, have another advantage: their frequency, intensity, and directionality can quite easily be studied and analyzed. So what Moss and her co-worker Annemarie Surlykke, an Associate Professor of Biology at the University of Southern Denmark, did in one of their experiments was to set up insect prey in close proximity to dense vegetation. Would the bat be able to separate the insect from the vegetation (which, in a way, represents unimportant background clutter)? And would there be a revealing pattern in the bat's vocalization?

Initially, the bat produced long signals during the search phase; it built up a general, three-dimensional picture of

the scene until something caught its attention. In this set-up it was the returning echo from the insect.
What the researchers observed was that, as the bat homed in on the insect, the characteristics of its calls began to change. The bat focused the center of the sonar beam on the point of interest and reduced the interval between successive sounds as it approached its prey. The insect, in this case without hearing the approaching bat, was sampled again and again while the bat was building up a detailed picture of the auditory scene from these snapshots. "It can filter out information on the way out by adjusting call directionality and on the way back by listening to particular echo features, and that's how the bat gets rid of what might not be so important at any moment in time," says Annemarie Surlykke. Video analysis revealed that the bat also adapted its flight path to approach the insect at the furthest distance from the vegetation. Action and perception were intimately linked – "steering by hearing" is what one of the team's papers is called.

Humans, of course, do not have such an active system. They are visual animals. However, during their studies, Annemarie Surlykke and Cynthia Moss have discovered many similarities to how we gather information from a visual scene. For example, the pattern by which bats move their head to direct the center of their sonar beam to objects of interest is comparable to our eye movements when we encounter an unfamiliar visual scene. So maybe there is something to be learned about visual scene analysis by listening to the bat. The problems, after all, remain the same: How do we separate background noise from what is interesting? How do we recognize the form of an object?

In a recent paper in the prestigious journal *Nature*, Lewicki and his co-workers asked exactly this question: How do you start to organize the information in a scene? On close inspection, the problem becomes tricky because the individual light receptors, comparable to pixels in a digital camera, capture only a tiny part of the image. This information is then analyzed by neurons in the visual cortex of the brain. But just as when closely inspecting an oversized print, it is impossible to recognize distinct objects by the individual pixels alone. In addition, the texture or color of the object might vary. So how does the brain assimilate all this information and recognize distinct objects?

"The visual system transforms all this variability into something that is less variable," Lewicki believes. It's analogous to extracting common denominators in all this variation and then grouping them into classes, just as the letter "L" is always recognized as such regardless of color or size. Lewicki and his group at Carnegie Mellon University have built a computer model that learns to handle this variability automatically. The model also predicts how neurons should behave, how they should be connected, and how they should respond. "It matches very closely with what people have measured in the visual system," Lewicki explains.

This is one of the areas where the group members' expertise overlaps: Lewicki now wants to adapt his model to the problem of organizing information in auditory scenes. Whether you are flying around catching insects or listening to a friend at a cocktail party, your auditory system has to pull out of a complex acoustic background numerous tiny fragments of sound to recognize distinct auditory objects. He does not expect this to be an easy problem but still believes that the fundamental principles remain the same.

Which is where the fourth member of the group comes in, Bruno Olshausen, Associate Professor of Neuroscience and Optometry at the University of California, Berkeley. Olshausen, like Lewicki, works on computational models, and has been focusing on the principles governing networks of neurons in the brain.

“Another challenge we are up against is that we don't fully understand the computational primitives that biology uses,” Olshausen says. It's hard to overestimate the immense complexity of even simple nervous systems. In any mammalian brain there are billions of neurons, any one of which is connected to thousands of others. To make matters worse, the flow of information is reciprocal: any given network of neurons processing a scene fragment will feed its information to a higher-level network. And, at the same time, this higher-level network feeds back on the lower-level network. It's a complicated exchange of information unlike anything in current computational devices.

Such complex systems are a nightmare to decipher and maybe neuroscience is decades away from fully mapping out the brain. Nevertheless, Olshausen is convinced that hidden within these feedback networks, nature has important lessons for engineers and computer scientists attempting to build artificial vision systems. Precisely because it is exactly what is needed in order to analyse a complex scene: a system that uses context from the “bird's eye view” of the overall scene to interpret all the loose fragments. But it also raises a chicken-or-egg question: “The higher levels can not be sure of how all the fragments fit together unless they are certain about what the fragments are,” Olshausen says. “And at the same time the networks processing the scene fragments cannot be certain of what information to pass up to higher levels without the proper context.”

Part of the answer may lie in the very fact that information flies in both directions. Higher-level networks can hypothesise on the information they receive from the lower levels. These hypotheses are then sent back and the lower-level networks can process this information. Developing concrete, computational models of this process is what Olshausen, in collaboration with Lewicki, tries to do during his time at the Wissenschaftskolleg.

So, after all, is there a universal information-processing algorithm common to all nervous systems? Might the bat experience its world just as we do? This idea harbours the thought that the mammalian brain is comparable to the central processing unit of a computer; where you can

Braunes Langohr
Plecotus auritus

plug in, via USB-connection so-to-speak, different sensory organs while the experience of the world stays the same: It's a question that has perplexed philosophers for centuries.

The group is sceptical because, as Lewicki points out, with each new sensory organ you open up another experience of the world. "You can make distinctions you could not make before," he says. The snake sensing infrared heat or the dog sniffing out each ingredient you put in your pasta is an experience that will be forever hidden from us. Just as the bat will never know from hearing alone what colour is.

But there is all reason to believe that underlying mechanisms that analyse and make sense of all this information may share many similarities. After all, who would have thought that bats can teach us what it really means to see?

v.l.n.r.:
Bruno Olshausen, Annemarie Surlykke,
Michael Lewicki, Cynthia Moss