A Model of Visually Guided Plasticity of the Auditory Spatial Map in the Barn Owl

Andrea Haessly
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712
andrea@cs.utexas.edu

Joseph Sirosh
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712
sirosh@cs.utexas.edu

Risto Miikkulainen
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712
risto@cs.utexas.edu

Abstract

In the barn owl, the self-organization of the auditory map of space in the external nucleus of the inferior colliculus (ICx) is strongly influenced by vision, but the nature of this interaction is unknown. In this paper a biologically plausible and minimalistic model of ICx self-organization is proposed where the ICx receives a *learn* signal based on the owl's visual attention. When the visual attention is focused in the same spatial location as the auditory input, the learn signal is turned on, and the map is allowed to adapt. A two-dimensional Kohonen map is used to model the ICx, and simulations were performed to evaluate how the learn signal would affect the auditory map. When primary area of visual attention was shifted at different spatial locations, the auditory map shifted to the corresponding location. The shift was complete when done early in the development and partial when done later. Similar results have been observed in the barn owl with its visual field modified with prisms. Therefore, the simulations suggest that a learn signal, based on visual attention, is a possible explanation for the auditory plasticity.

Introduction

In the brain, several computational maps process sensory information. The maps transform the sensory input into a localized activity on the map, which can be easily accessed by other neural processes. These maps self-organize so that the input space is represented topologically on the map (Knudsen et al., 1987). Typically each map is concerned with only one sensory modality. However, the barn owl is unusual in that its auditory map is strongly influenced by a different modality, vision.

Some kind of visual influence on the auditory map is known to exist because the auditory map adapts even when only vision is distorted. The auditory map in the external nucleus of the inferior colliculus (ICX) projects the auditory input to the optic tectum. In the optic tectum, the auditory information is combined with visual input to form a bimodal topographic map of space. This spatial map allows the barn owl to locate its prey using either visual or auditory cues. In order to support the bimodal organization in the optic tectum, the auditory map in the inferior colliculus must be structured appropriately; a vision-based calibration signal must be involved in its self-organizing process. In this paper, the hypothesis that a learn signal, based on coincidence of visual attention and auditory input, mediates the auditory plasticity of the ICx, is proposed and evaluated computationally.

Auditory and Visual Localization in the Barn Owl

To motivate discussion of the learn signal, let us first review the roles the inferior colliculus and the optic tectum have on localization. The barn owl primarily uses sounds to localize its prey in the dark, and it can do this with an accuracy surpassing that of most birds and mammals (Payne, 1971). The auditory maps that give the barn owl its extraordinary abilities are located in the inferior colliculus. There are three subdivisions of the inferior colliculus known as the central nucleus (ICc), external nucleus (ICx), and superficial nucleus (ICs). The ICs is a relatively small portion of the inferior colliculus, and little is known about its function other than there may be a few projections to the ICx (Knudsen, 1983). The neurons in the ICc, however, are sharply tuned to frequency and tonotopically organized, and the neurons in the ICx are broadly tuned to frequency and spatiotopically organized. The ICx receives its input from the ICc. The frequency-coded auditory input to the ICc is transformed to spatial location in the ICx by the projection from the ICc to the ICx. This transformation process, involving interaural level differences (ILD) and interaural time differences (ITD), can unambiguously determine the location, in azimuth and elevation, of the source of the auditory input (Brainard et al., 1992; Knudsen, 1987).

The spatiotopic organization that exists in the ICx is projected to the optic tectum. In the optic tectum (OT), there is a general bimodal map of space that responds to both visual and auditory input (Knudsen, 1982). This map enables the owl to determine the location of its prey either using visual cues or auditory cues. Experiments with visual stimuli have shown that neurons in the optic tectum are organized according to azimuth and elevation, and a visual stimulus in a certain area of space causes the neurons representing that area of space to fire much more rapidly (Knudsen & Konishi, 1978; Knudsen, 1982). The map responds to auditory stimuli in the same way. Most of the neurons that respond to visual stimuli also respond to auditory stimuli. In addition, the location of an auditory response is usually very close to the location of the corresponding visual response. This way, both modalities lead to similar responses, and other neural structures can process location information in the same way, irrespective of the originating modality.

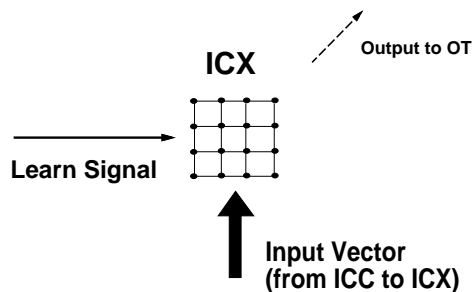Since the map in the OT is formed by merging two different

Figure 1: **The ICx model.** The two-dimensional feature map stands for the auditory spatial map in the ICx of the barn owl. The input from the ICc is a vector that is propagated to each node in the network. The learn signal is either on or off, and determines whether the map will be adapted. The neurons of the ICx project to the optic tectum.

modalities, auditory and visual input must have a compatible structure in order to assure that the bimodal map will encode the same location for both inputs. Furthermore, since the region of the bimodal map that corresponds to the area directly in front of the owl is magnified (occupying a disproportionately large portion of the map), the auditory spatial map in the ICx must have the same amount of magnification in this area in order for the visual and auditory locations to correspond. The structure of the retina causes magnification of the visual input. Since there is no corresponding mechanism that magnifies auditory input, the auditory map in the ICx must be conforming to the visual map in the optic tectum. Therefore, it seems that a visually based calibration signal must exists which guides the development of the auditory map in the ICx.

Several experiments have been performed to determine what influence vision and hearing have on the formation of the bimodal map in the optic tectum (Knudsen, 1985; 1988; Knudsen & Brainard, 1991; Knudsen & Knudsen, 1985a; 1985b; 1990). Since the OT receives its auditory input from the ICx, any changes in the representation of auditory space in the OT reflect the plasticity that is occurring in the ICx due to vision. For example, prisms or occluders were mounted over the owl's eyes to manipulate the visual information the owl received (Knudsen & Knudsen, 1985a). The adaptation usually took weeks and the prisms were left on for a period of months. While the owls still had the prisms on, their auditory localization abilities were tested by having the owl orient its head directly at the location of an auditory stimulus. However an owl, wearing right shifting prisms, localized to the right of the auditory stimulus. Even though the owl received correct auditory information, it could not accurately locate the stimulus; instead the owl chose a location that conformed to the visual distortion created by the prisms. This is an instance where a shifted visual signal causes the formation of an abnormal auditory spatial map in the ICx, and therefore in the OT, even though there were no distortions in the auditory input. Vision is used as a recalibration mechanism for the auditory spatial map in the ICx, even if the visual cues are incorrect. These experiments show an innate dominance of vision over audition.

Where does the visual recalibration signal come from? The

ICx does not respond to any visual inputs, so there are no direct visual signals available for comparison at the ICx. Anterograde labeling revealed that there was no direct feedback from the optic tectum to the ICx either (Knudsen & Knudsen, 1983). Previous computational models of visual calibration in the ICx map have relied on such connections, modeled by backpropagation of an error signal and/or a reinforcement signal (Rosen et al., 1994; Pouget et al., 1995). In addition, these models did not address how the two-dimensional maps in the ICx could self-organize from the visual input. It has been confirmed that the synaptic changes that alter the auditory maps occur in the ICx itself, and not in the lower centers or in the optic tectum (Brainard & Knudsen, 1993). To date, the nature of the recalibrating signal to the ICx is not well understood.

In the remainder of this paper, a simple biologically plausible mechanism for the self-organization and plasticity of the ICx is proposed. Simulations are performed to demonstrate the plasticity of the ICx and the effects of the proposed learn signal. The results are then discussed along with some possible future areas of research.

The Learn Signal Model

The model is based on the self-organizing feature map (Kohonen, 1981, 1989, 1990), which is an abstraction of the biological mechanisms that give rise to topographic maps. Here, a two-dimensional Kohonen map models the auditory spatial map in the ICx. The spatial location is assumed to be computed by the projection from the ICc to the ICx, and the map network receives the resulting spatial representation vector as its input (figure 1). These auditory input vectors are uniformly distributed since a sound can originate at any location in space (figure 2a).

How could visual input calibrate the auditory map formed by the Kohonen algorithm? Because the ICx does not respond to any visual inputs directly, the calibration signal, while visually based, must be of a different form. A simple *learn* signal that turns the synaptic learning on or off is proposed in this paper. When the visual attention and the location of the sound source coincide, the learn signal is turned on and allows the map to adapt. Thus, the signal forces the map to learn the portion of the input space currently attended to. Since the owl at-

2

(a) Auditory input space   (b) Visual attention centers   (c) Auditory spatial map
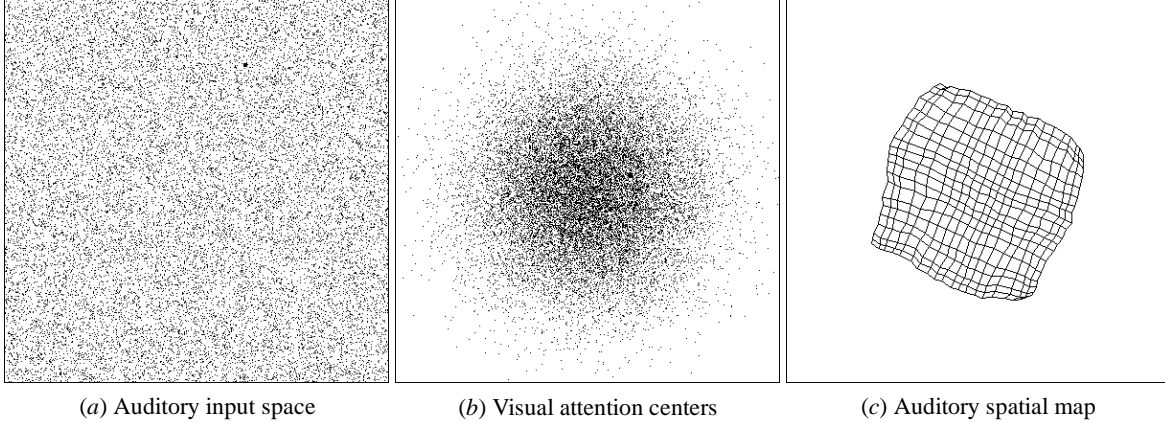
Figure 2: **Self-organization of the normal topographic ICx map.** The square region in (a) is a two-dimensional representation of the auditory space, and the inputs are uniformly distributed in this space. Figure (b) shows the distribution of visual attention centers that were used to determine the on/off value of the visual instructive signal. The attention is distributed about the center of the input space in a gaussian fashion, so that the center is attended to more frequently than the periphery. When trained with these signals, a topographic map of the input space develops, as shown in figure (c). The width of the map corresponds to the spread of the attention signal.

tends more often to the center of its visual field than to the periphery, a gaussian distribution around the center of the input space is used to generate the visual attention (figure 2b).

In the Kohonen map the neuron that is most similar to the input vector is known as the excitation center. The excitation center of the auditory map, for input v is defined as the neuron r' for which

$$\forall r \,:\, \| v - w_{r'} \| \leq \| v - w_r \|, \tag{1}$$

where r are the nodes in the network and w_r is the vector of weights. The excitation center is the image of the auditory input on the map.

During training the learn signal must be computed. If the Euclidean distance between the auditory input and the visual attention is within a certain threshold θ, the learn signal will be on; otherwise the signal is off. When the signal is on, the synaptic strengths of the neighborhood around the excitation center are modified according to the standard feature map learning algorithm

$$w_r^{new} = w_r^{old} + \alpha \left(v - w_r^{old} \right), \tag{2}$$

where α is the learning rate. Adaptation occurs only when the owl is attending to the area of space where the sound originates, that is, when the auditory input and visual attention coincide.
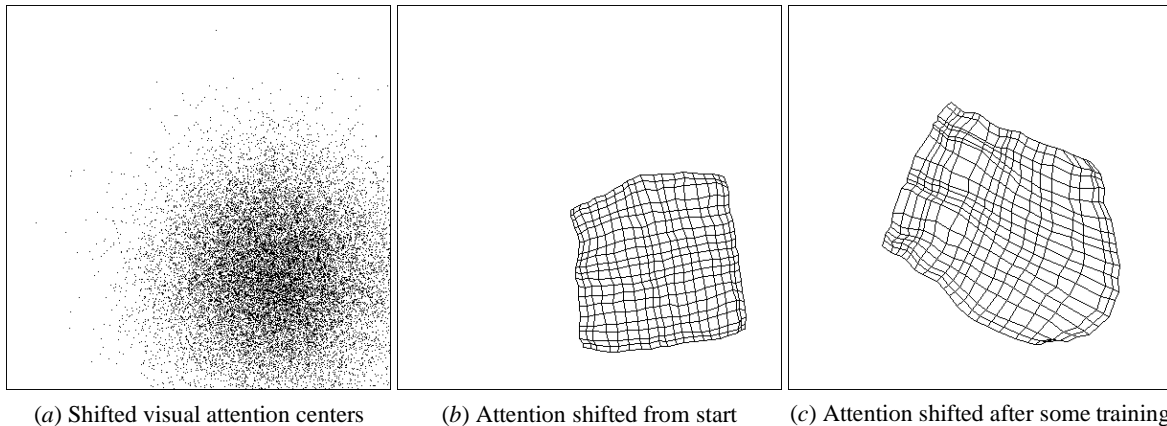
Simulations

Simulations were performed using a 20×20 neuron network with random initial values for all weights. Each training trial consisted of four steps: (1) An input vector was generated and the excitation center was determined using equation 1. (2) An attention center was generated and (3) compared with the excitation center. If the signals were relatively close (within the threshold θ), the learn signal was turned on; otherwise the learn signal was off. (4) If the learn signal was

on, synapses were modified according to equation 2. A total of 20,000 training trials were required for the map to organize.

A series of experiments were performed to simulate the different experimental conditions on owls with and without prisms. The first experiment simulated the control case where the owl is allowed normal vision. The visual attention is generated from a gaussian that is centered over the input space (figure 2b). The resulting topographic map is shown in figure 2c. The map is centered in the input space, and the extent to which the map covers the space is determined by the spread of the gaussian attention signal. Thus, the learn signal focuses the map to the most attended portion of the input space.

To simulate the development of the ICx map with prisms, the center of the gaussian distribution of attention was shifted relative to the input space (figure 3a). The second experiment simulated the owl wearing the prisms before its eyes had opened. In this case, the gaussian was shifted before any training steps. The resulting network had a similar shape as in the control case, but the entire network had shifted in the direction of the learn signal (figure 3b). Here, the map was forced to learn the inputs in the shifted region.

In the third experiment, the center of the gaussian was shifted after 10,000 training trials, simulating a period of normal development, after which the prisms were placed over the owl's eyes. Initially the map was forming in the center of the input space. After the shift occurred, the map slowly moved towards the new attention center. The area of the map furthest away from the signal was slower to adapt and looked similar to the map where the learn signal is centered. In conclusion, if the shift was introduced right from the start, the network learned only the attended region (figure 3b); however, if the shift was introduced in the middle of training, the map shifted only partially (figure 3c). These results are in agreement with those observed experimentally in the barn owl (Knudsen & Knudsen, 1990).

(*a*) Shifted visual attention centers   (*b*) Attention shifted from start   (*c*) Attention shifted after some training

Figure 3: **Self-organization of the ICx map with shifted inputs.** Figure (*a*) shows the shifted distribution of visual attention centers. Figure (*b*) displays the map trained from the start with the shifted attention signal. The entire map has shifted to coincide with the position of the attention signal. However, if the attention signal is shifted midway during training, only the portion of the map close to the new attention center shifts, as shown in figure (*c*).

Discussion

The proposed model of the visually guided plasticity in the ICx is minimalistic in its assumptions and biologically well-motivated. The single Kohonen map represents the auditory spatial map that exists in the ICx of the barn owl. It is not necessary to model the OT because the plasticity of the auditory map occurs at the level of the ICx. The model shows that a simple on/off learn signal, based on the coincidence of visual attention and auditory input, is a sufficient explanation for auditory map plasticity. The learn signal does not need connections from the optic tectum to the ICx, which several other models rely on, because the learn signal is not based on a comparison of the visual and auditory inputs in the OT. The learn signal is also much simpler than an error signal, which would include information such as the magnitude and location and even direction of error. The learn signal could originate from the higher cortical areas of the brain where visual input has already been processed, and the location for visual attention formed.

In humans and other animals, sensory modalities are combined to give a single cohesive view of the world. In certain cases the perception of the world can be distorted, as in an illusion, because of conflicting information from different sensory inputs. The visual dominance in the formation of the multimodal map in the barn owl gives insight into the mechanisms used for the integration of different sensory modalities and how one modality can distort the perceptions of other modalities.

In future work, we plan to extend the model with more realistic neurons with lateral connections where the weight modification process is completely unsupervised (e.g. Sirosh & Miikkulainen, 1994). Furthermore, we plan to include the ICc to ICx connections, which are responsible for computing the spatial input representation from the frequency-specific interaural level differences (ILD) and interaural time differences (ITD). This way the ICx would be organized according to di-

rect inputs from the ICc instead of intermediate spatial representations as in the current model. The bimodal map that exists in the OT could also be included. This map would not only represent the visual input, but it would also incorporate the auditory input from the Kohonen map in the existing model. Such a comprehensive model would be a major step toward verifying that the learn signal is still sufficient for the plasticity of the ICx on this large scale, and that it causes a similarity in structure of both the auditory and visual spatial maps so that the merging done in the OT is possible.

Conclusion

The simulations reported in this paper demonstrate that a simple visually-based learn signal is a sufficient explanation of the auditory plasticity observed in the ICx of the barn owl. Unlike in previous models, an error signal is not necessary to calibrate the auditory map. Rather, the simple coincidence of visual attention and spatial location of auditory input may alone drive the plasticity of the ICx. The coincidence signal may be generated in the cortical area that is responsible for attention. Direct feedback projections from the optic tectum or close coupling of the OT and ICx are not necessary. In the future, biological experiments should be performed to verify whether such a learn signal exists, and also to determine the signal pathway to the ICx from the higher cortical areas, possibly via the ICs.

References

Brainard, M.S. & Knudsen E.I. (1993). Experience-dependent plasticity in the inferior colliculus: A site for visual calibration of the neural representation of auditory space in the barn owl. *The Journal of Neuroscience* 13, 4589–4608.

Brainard, M.S., Knudsen, E.I. & Esterly S.D. (1992). Neural derivation of sound source location: Resolution of spatial ambiguities in binaural cues. *The Journal of the Acoustical Society of America* 91, 1015–1027.

Knudsen, E.I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *The Journal of Neuroscience* 2, 1177–1194.

Knudsen, E.I. (1983). Subdivisions of the inferior colliculus in the barn owl (*Tyto Alba*). *The Journal of Comparative Neurology* 218, 174–186.

Knudsen, E.I. (1985). Experience alters the spatial tuning of auditory units in the optic tectum during a sensitive period in the barn owl. *The Journal of Neuroscience* 5, 3094–3109.

Knudsen, E.I. (1987). Neural derivation of sound source location in the barn owl. An example of a computational map. *The Annals of the New York Academy of Science* 5, 3094–3109.

Knudsen, E.I. (1988). Early blindness results in a degraded auditory map of space in the optic tectum of the barn owl. *The Proceedings of the National Academy of Science* 85, 6211–6214.

Knudsen, E.I. & Brainard, M.S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science* 253, 85–87.

Knudsen, E.I., du Lac, S. & Esterly, S.D. (1987). Computational maps in the brain. *Annual Review of Neuroscience* 10, 41–65.

Knudsen E.I. & Knudsen P.F. (1983). Space-mapped auditory projections from the inferior colliculus to the optic tectum in the barn owl (*Tyto Alba*). *The Journal of Comparative Neurology* 218, 187–196.

Knudsen, E.I. & Knudsen P.F. (1985a). Vision calibrates sound localization in developing barn owls. *The Journal of Neuroscience* 9(9), 3306–3313.

Knudsen, E.I. & Knudsen P.F. (1985b). Vision guides the adjustment of auditory localization in young barn owls. *Science* 230, 545–548.

Knudsen, E.I. & Knudsen P.F. (1990). Sensitive and critical periods for visual calibration of sound localization by barn owls. *The Journal of Neuroscience* 10, 222–232.

Knudsen, E.I. & Konishi, M. (1978). A neural map of auditory space in the owl. *Science* 200, 795-797.

Kohonen, T. (1981). Automatic formation of topological maps of patterns in a self-organizing system. In *Proceedings of the 2nd Scandinavian Conference on Image Analysis* (pp. 214–222). Espoo, Finland: Pattern Recognition Society of Finland.

Kohonen, T. (1989). *Self-Organization and Associative Memory*, chapter 5. Berlin; Heidelberg; New York: Springer.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE* 78, 1464–1480.

Payne, R.S. (1971) Acoustic location of prey by barn owls (*Tyto Alba*). *The Journal of Experimental Biology* 54, 535–573.

Pouget, A., Deffayet, C. & Sejnowski, T.J. (1995). Reinforcement learning predicts the site of plasticity for auditory remapping in the barn owl. In *Advances in Neural Information Processing Systems 7* San Mateo, CA: Morgan Kaufman Publishers.

Rosen, D.J., Rumelhart, D.E. & Knudsen, E.I. (1994). A connectionist model of the owl's sound localization system. *Advances in Neural Information Processing Systems, 6.*

Sirosh, J. & Miikkulainen, R. (1994). Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics* 71(1), 66–78.