Object Tracking and Selective Attention in a Bat-Inspired Echolocation System

Stephanie M. Doctor, Sairina Mirchandani, and Timothy K. Horiuchi

***Abstract*— Air-coupled sonar is an important sensory system for navigation and feeding in echolocating bats and potentially in man-made flying agents operating in complete darkness. This project will illustrate how localization can be performed with a sonar system by implementing a learning approach in a neurally-inspired classifier network exposed to training data. Using this system mounted on a rotational joint, we will further demonstrate how it can be used to track objects and selectively attend to task-relevant salient points.**

***Index Terms*—sonar, echolocation, attention, neural network, radial basis function**

I. Introduction

Air-coupled sonar is an important sensory system of echolocating bats for capturing moving prey and navigating in complete darkness. These capabilities are coveted for potential applications in man-made flying agents, but current sonar technology cannot compete with a bat's echolocating abilities. Typical air-coupled sonar devices used in mobile robotics can acquire the range of the closest object and its general direction but they cannot precisely locate an object. This project aims to demonstrate how air-coupled sonar, with a more biologically-inspired approach, can be used to determine the azimuth of an object. Furthermore, two practical applications of this ability, tracking an object and creating an attentional system, will be described and implemented.

II. Setup

The laboratory setup features a sonar head mounted on a rotational servo controlled by a USB-based control board manufactured by Pololu®. The head consists of three sonar transducers (or "pingers") positioned next to each other. The two pingers on the ends are angled slightly outwards from the middle. The arrangement can be seen in Fig. 1. An important aspect of the localization portion of the project was to demonstrate a method for learning the mappings between measurements and angles, regardless of range. This allows the method to be applied to any number and/or configuration of transducers. Neither exact positioning of the sensors nor knowledge of it is required because the sensor-in-the-loop training measurements incorporate all of these transformations.

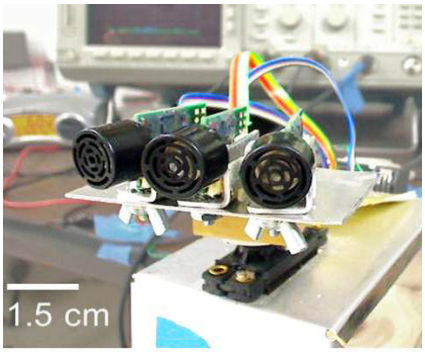

Fig. 1. Arrangement of the pingers on the rotational sonar head.

Each pinger is capable of both sending and receiving ultrasonic pulses at 40 kHz, so for a given pulse, each pinger returns the amplitude of the echo from each object in view. The microphone sensitivity patterns allow the amplitude data from the three pingers to be compared to determine the azimuth of the target.

III. Training

A neural network was chosen to train the sonar. An analytical approach would be simpler to implement, but would require detailed knowledge of both the configuration of the pingers and any nonlinearities in the signal processing. An

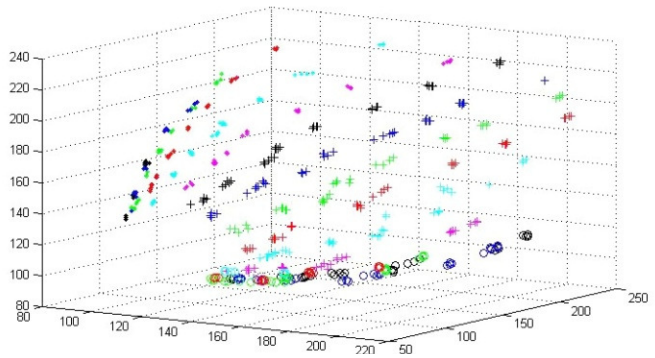

Fig. 1. Training data for 17-angle system. Different colors/markers represent different angles; clusters show different ranges. The axes are the amplitude values from the pingers. Units are arbitrary amplitude units.

Manuscript received August 1, 2011. This work was supported by the National Science Foundation (NSF) (OCI # 1063035, IOS #1010193) and the Office of Naval Research (ONR-705018).

S. M. Doctor is with the Department of Biology, University of North Carolina, Chapel Hill, NC 27514 USA (e-mail: sdoctor@email.unc.edu).

S. Mirchandani is with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic University, Troy, NY 12180 USA (e-mail: mirchs@rpi.edu).

T. K. Horiuchi is with the Department of Computer and Electrical Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: timmer@umd.edu).

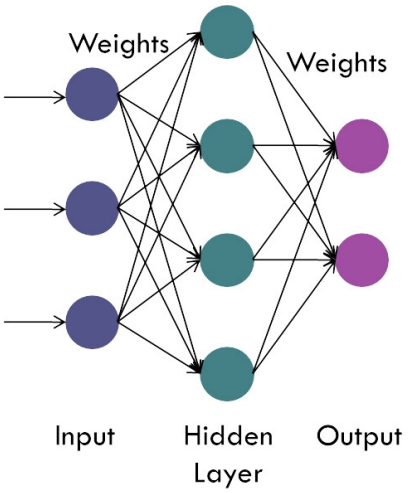

Fig. 2. Perceptron neural network diagram.

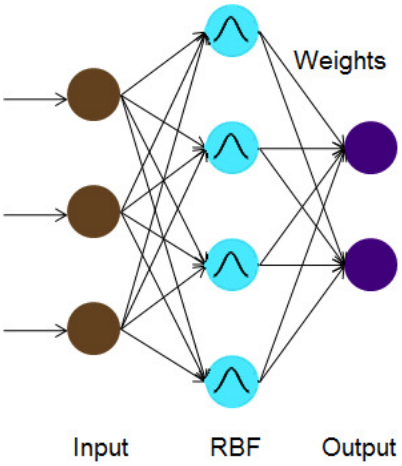

Fig. 5. Radial basis function neural network diagram.

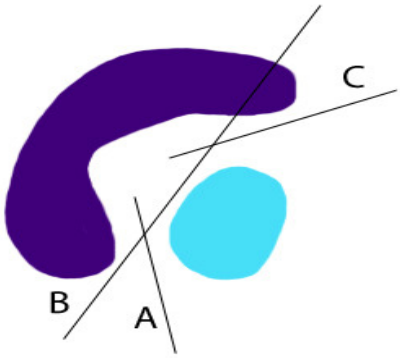

Fig. 4. Demonstration of how planes can be used to separate data. Here the relationship of a data point with the lines A, B and C determines whether the point is a member of either the blue or purple groups.

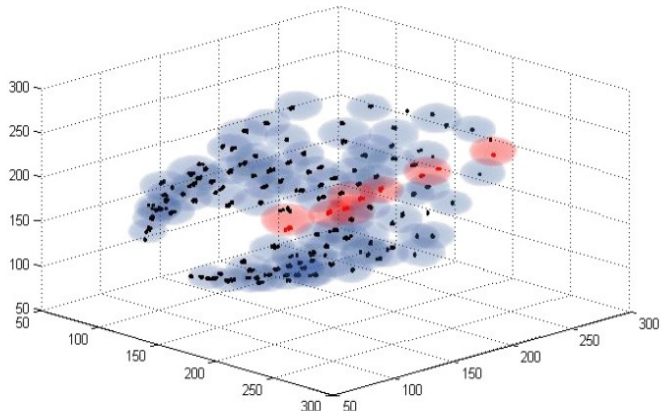

Fig. 6. RBF centers among 15-angle training data. Cloud size is 1 standard deviation. Clouds highlighted in red are most strongly associated with one example angle.

approach that associates angles with real measurements is much more robust and is more generally applicable.

A. *Data Collection*

Training data was collected for a target (1.5" dia. vertically-oriented PVC pipe) at 8 ranges for 17 sample angles. These angles ranged from -40° to 40° at 5° intervals. For each position, amplitude data was captured from each pinger, resulting in a point in 3-dimensional space. This data is shown in Fig. 1.

Many of the examples and figures in this report use the 17-angle system; however, late in the project the system was recalibrated after one of the pingers was displaced from its original position. The new configuration had a narrower field of view and only 15 angles at 5 degree separation were possible. The remaining figures are based on the 15-angle system.

B. *Learning*

Two types of neural network learning approaches were compared to determine which one would create a more accurate system. In each, a three-neuron input layer represents the amplitude data from the three receivers and 17 output neurons represent the training angles. For a given input, the output neuron population activity, when normalized, represents a probability distribution of the target location.

The first approach for classification used a three-layer perceptron network, which is represented in Fig. 3. In this network, there is a "hidden layer" of neurons between the input and output layers, which represent planes in 3-dimensional space that separate the measurements of targets at different angles. The algorithm uses the amplitude values from the 3 pingers as inputs, and then activates output neurons using the location of the data with respect to the planes defined by the hidden layer. For the example in Fig 4., a given data point can be classified as being in the purple group if it is to the left of plane B or above plane C. During training, if the output is incorrect, the positions of the planes and the manner in which they influence the output are adjusted according to the perceptron learning rule (see [1]).

The second approach for classification was a radial basis function (RBF) network, which is represented in Fig. 5. In this network, the distances between an input data point and a set of RBFs, which are Gaussian distributions, are calculated. The outputs from the RBF layer are multiplied by weights and summed to activate the corresponding output neurons; the overall function is defined by

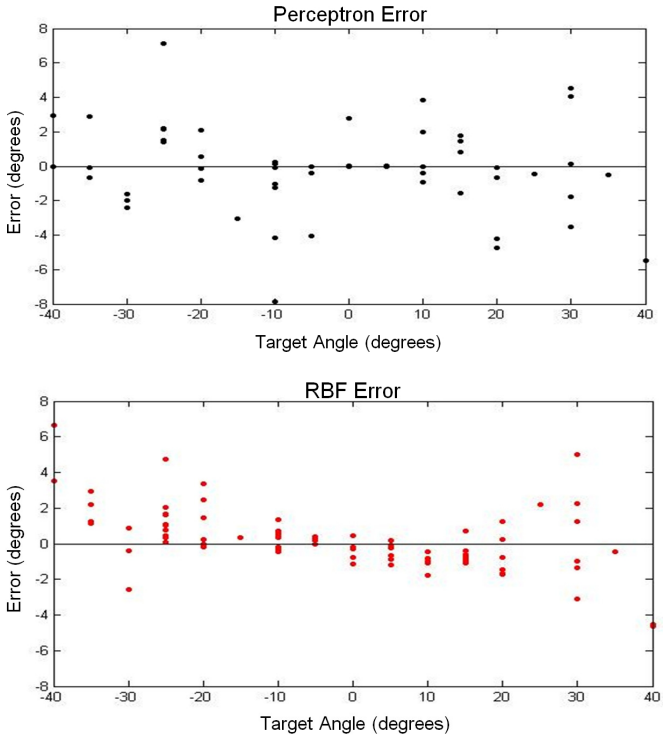

Fig. 7. Error of Perceptron learning algorithm (top) and RBF learning algorithm (bottom) for a subset of the 17-angle training data. Error is the difference between the calculated angle and the actual target angle.

$$f_r(x) = \lambda_0 + \sum_{i=1}^{n_r} \lambda_i \phi\left(\|x - c_i\|\right) \quad (1)$$

where x is the input vector, $\Phi=e^{-v^2/B^2}$, B is a constant, λ_i are the weights on the RBF layer, c_i are the RBF centers, and n_r is the number of centers [2]. During the initial phase of the training process, the RBFs are scattered among the data points and adjusted by position and size to reflect clusters and subspaces in the training data (as seen in Fig. 6); the weights between the RBFs and the output layer are then modified to generate the desired angle output pattern.

IV. System Performance

To compare the two learning algorithms, they were tested using a subset of the training data. For both approaches, the weighted average of the output neurons (i.e., expectation value for the angle) was calculated for each test point. This was compared to the actual angles, and the error was plotted in Fig. 7. The RBF algorithm overall showed fewer errors than the perceptron algorithm, so the sonar utilized the RBF algorithm for subsequent projects.

It is also interesting to note that in Fig. 7 the RBF shows smaller error in the center of the field of view, suggesting that it can resolve targets toward the center better than those at the edges.

Fig. 8 shows tuning curves, or neural activation, of three adjacent output neurons in the system, for 36 angles between approximately -13° and 12°. The angles on which the neurons were trained are highlighted in red. These tuning curves

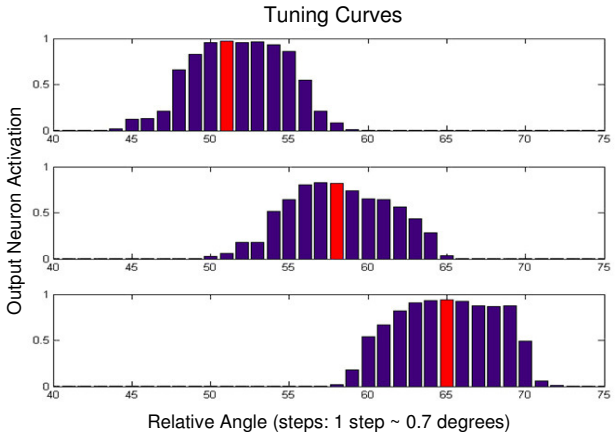

Fig. 8. Tuning curves for three adjacent neurons of the 15-angle system over 36 angles between approximately -13° and 12°. The angles on which the neurons were trained are highlighted in red.

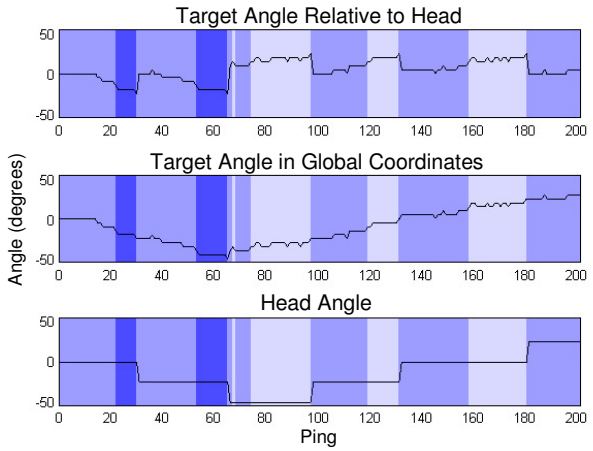

Fig. 9. Behavior of the 15-angle system when tracking an object for 200 pings. Background color indicates which pinger is firing: light blue represents left pinger, medium blue represents middle pinger, dark blue represents right pinger.

display overlap of neural activity, which shows the system's potential for hyperacuity, or ability to estimate angles between the trained ones.

V. Tracking

The higher reliability of estimates in the center of the field of view and ability to differentiate angles motivated the design of an algorithm that tracks the closest target in its field of view. The algorithm reacts to the target's relative azimuth according to the following rules: 1) if the target angle is within 10° of where the head is facing, the sonar stays in the same place and fires the middle pinger; 2) if the target angle is between 10° and 20° away, the sonar does not turn but fires the side pinger closest to the target. 3) If the target angle is more than 20° off-center, the sonar turns to center on the target. Fig. 9 displays the behavior of this algorithm in real time.

VI. Attention

From both a biological and an engineering perspective, selective attention is highly advantageous for the efficient use

of limited resources. Selective attention typically refers to the phenomenon of preferentially allocating resources to a subset of available sensory inputs (e.g., different locations in space). In bats, selective attention is overtly expressed when the head turns to spotlight different objects or when changes are made to their vocalizations. Covertly (i.e., internally), however, bats may be focused on the signals associated with only one of many objects in the field of view, such as an insect. Applying some of the established ideas of attention to a sonar system might allow it to handle some of the complex sensorimotor tasks associated with navigation in cluttered three-dimensional environments. Such capabilities would be valuable for man-made flying agents.

For any limited-view sensor, movements provide an expanded field of view, but they may also improve perception when the sensor's performance varies across its field of view. To direct resources, attentional systems are commonly designed around a "saliency map" that represents the desirability of focusing on a given direction. Saliency is considered to be partly innate and partly task dependent.

The task chosen for this project was to map and monitor the positions of objects in the near-field environment. For this task, the factors of importance taken into consideration were the localization uncertainty of detected objects and the time elapsed since a particular direction had been looked at.

A. *Certainty*

The first factor of importance, the certainty of an object's location, was quantified by calculating the entropy of the output distribution from the RBF network (as in [3]) with the formula

$$Entropy = -\sum p(x) \log p(x) \tag{2}$$

By calculating entropy as in (2), its value provides information regarding the overall shape of the distribution. High entropy suggests a "flatter" distribution, or less certainty; low entropy indicates certainty of the target's location. The entropy could then be used to create a saliency map for that object by multiplying the normalized RBF output by the entropy. This way, high entropy, or less certainty, would lead to high saliency, or more interest in looking in that direction to clarify the object it has seen. In Fig. 10, six example distributions are shown with their associated entropies and

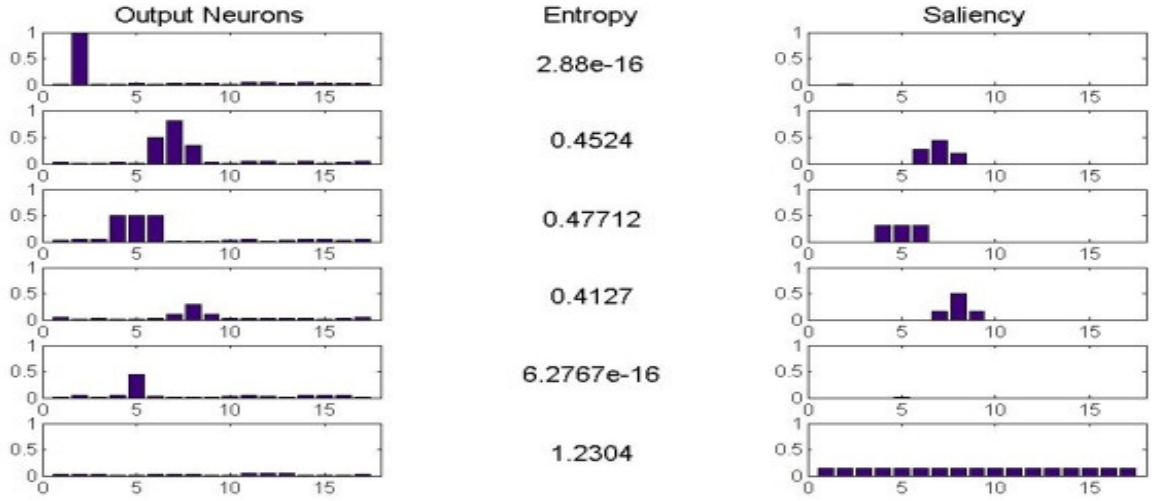

Fig. 10. Six sample output distributions of the 17-angle system and their associated entropies and saliency maps.

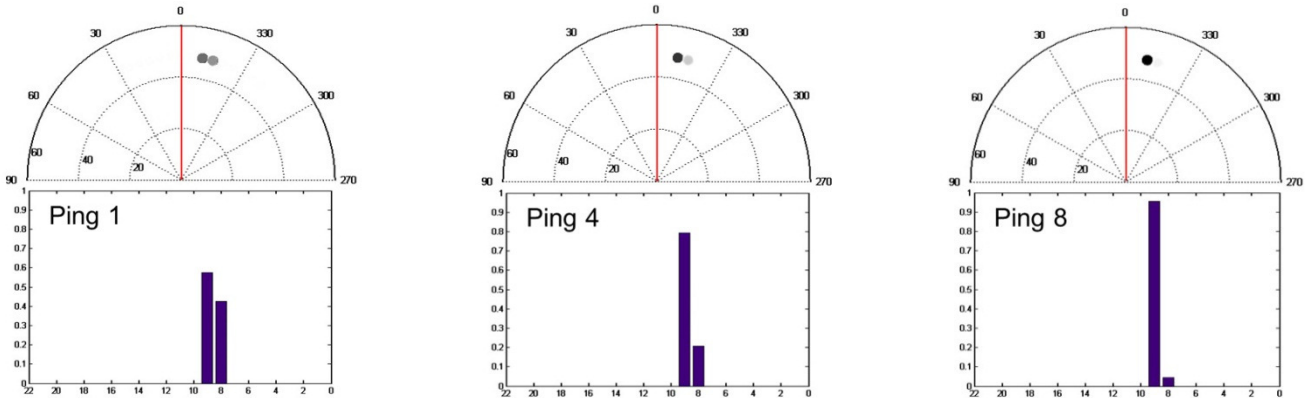

Fig. 11. Demonstration of how merging looks improves the certainty of observed objects. The top plots map the position of an object. A darker point indicates higher certainty. The lower plots are bars of the normalized output for the object. Data was collected from sonar using the 15-angle system.

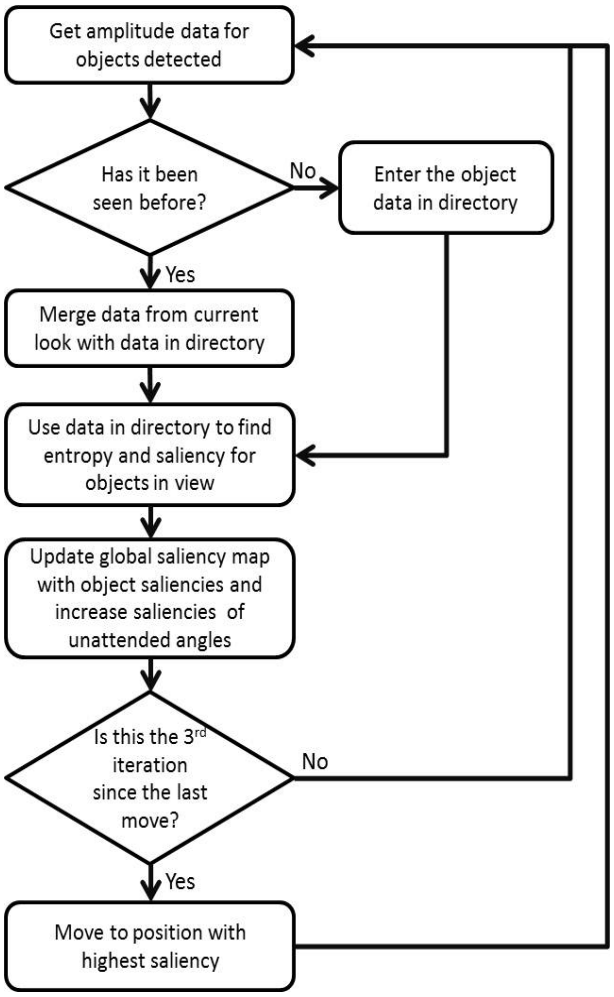

Fig. 12. Attention algorithm flow chart.

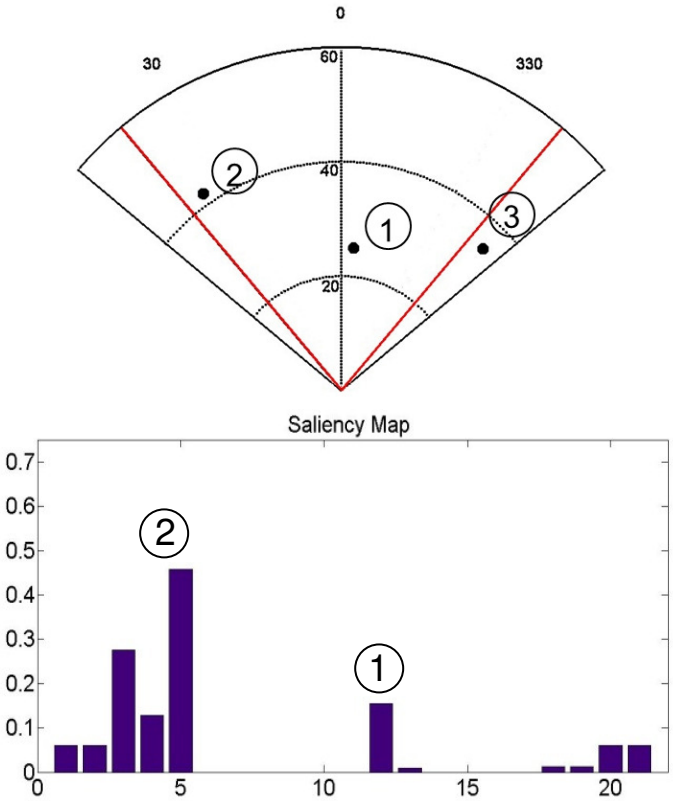

Fig. 13. Example of the function of the attentional system using the 17-angle configuration. The top plot displays the positions of sample targets 1, 2 and 3. Targets 1 and 2, which are within the field of view, appear in the saliency map. Target 3 is undetected. Unobserved angles (edges of the saliency map) rise slowly. With this data, the sonar would turn toward target 2, since it has the highest saliency, which should then drop. The saliency at the right would continue to rise until the sonar turned toward it, at which point it would locate target 3.

saliency maps.

The entropy of an object could also be affected by the number of times the object had been seen. Seeing an object for a second time at the same location should make the sonar more confident of the object's position, regardless of the value of its entropy. To address the problem of how to merge information from data taken at different times and different angles, the target distributions were aligned to a global coordinate system. If the sonar saw a target and determined that it had been seen before, its old and new distributions were multiplied and then renormalized; then its entropy was recalculated. Often this led to lower saliency, allowing confidence to grow without necessarily moving the sonar head. An example of how certainty grows using this method is shown in Fig. 11.

B. *Time Since Last Look*

The other factor that was used to create the saliency map was the time since a particular direction had been in the sonar's field of view. When monitoring an environment, even if there is absolute certainty of what has been seen, it is not necessary that this information will always be correct. An environment changes with time and the system must take that into account. This was done by slowly increasing the saliency of undetected angles in the global saliency map. This way, if the sonar continued to look away from an angle for an extended period of time, its saliency would eventually become the highest and the sonar would turn in its direction.

C. *Algorithm Design*

The algorithm structure is outlined in Fig. 12. An example of how the system works is shown in Fig. 13.

VII. Conclusion

Although sonar sensors are used extensively in mobile robotics, they are not used as a primary sensory modality due to their coarse angular localization. Given the capabilities of echolocating bats, however, it is clear that bats have much to teach us. This system was trained to (and achieved) an azimuthal resolution of 5° (error = ±6°). Future work should build off of this foundation. The recent approach in the computational neuroscience community of generating and using probability distributions for signaling opens many interesting possibilities for neural processing as well as for deeper mathematical insights into neural computation. Using biological systems, such as that in bats, as inspiration for sonar

technology will allow steps to be taken toward a man-made system that can perform as well as biological ones.

ACKNOWLEDGMENT

We thank Tarek Massoud for advice during the project and Professor Cynthia Moss for discussions of the behavior of the big brown bat.

REFERENCES

[1] A. K. Jain and J. Mao, "Artificial neural networks: A tutorial," *IEEE Computer*, vol. 29, pp. 31–44, Mar. 1996.

[2] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Networks*, vol. 2, pp. 302–309, Mar. 1991.

[3] L. W. Renninger, P. Verghese, and J. Coughlan, "Where to look next? Eye movements reduce local uncertainty," *Journal of Vision*, vol. 7(3):6, pp. 1–17, Feb. 2007.