Proceedings of the 2005 IEEE
International Workshop on Safety, Security and Rescue Robotics
Kobe, Japan, June 2005

# Sequential sensing with Biosonar for natural landmark classification

Maosen Wang, Andreas Zell
Computer Science Dept
University of Tuebingen
Sand 1, 72076 Tuebingen, Germany
Email: wangm,zell@informatik.uni-tuebingen.de

***Abstract*— Echolocating bats can make nocturnal flights in acoustically cluttered environments with the use of echolocation. Their ability to evaluate targets in complete darkness provides mobile robots an opportunity to learn target detection, classification and identification with similar biomimetic platforms. In this work, natural landmark classification with a binaural system, a sequential sensing strategy and a frequency after reconstruction algorithm were developed and tested. The aim of the work is to overcome some inherent shortcomings of airborne sonar and take advantage of bats' perceived properties for mobile robots' navigation in natural environments. Experimental results suggest considerable improvements in classification accuracy can be achieved by the use of this sequential classification method.**

## I. Introduction

Since the 19th century there is an increasing interest to study bats' marvellous ability to "see in the dark". Some of the recent work relevant to sonar target classification is reviewed here: Leonard and Durrant-Whyte used the feature of Region of Constant Depth (RCD) for navigation[5]. Kimoto and Yuta detected hedges with standard deviation of range readings[3]. McKerrow extacted features both for indoor and outdoor landmark classification[6][7]. L. Kleeman achieved accurate measurement and classification of three target types (plane, corner and edge[1]. Roman Kuc tried to recognize penny coins using binaural information of a biomimetic sonar system[4]. Walker and Peremans used a bionic sonar head successfully for single target tracking in clutter[13]. Rolf Mueller made classification of 4 trees with the features of Distant Interspike Intervals (DIIS)[8]. Although it is believed that the broadband FM echolocation calls cannot provide detailed information about the fine texture of objects [11], it still deserves study when it can provide sufficient cues for certain tasks such as mobile robot navigation with certain sensing strategies. The aim of this work is to avoid some inherent shortcomings of airborne sonar and take advantage of bats' perceived properties for mobile robot's navigation. A sequential sensing strategy and a frequency after signal reconstruction algorithm are developed and tested, aiming at biosonar's practical application in mobile robots' outdoor navigation, where complex natural landmarks are far more richer than simple geometrical forms. The organization of this paper is the followings: section 2 describes some inherent problems of the biosonar signal, section 3 introduces the method of sequential sensing. Section 4 presents experimental results, while conclusions are given in section 5.

## II. PROBLEM

The auditory system of echolocating bats assembles information over time to build a representation of the relative moving targets [9]. Successful classification of a few simple geometric targets has been achieved by special configuration of broadband ultrasonic sensors[1]. But in natural environments most of the landmarks like plants are irregular and consist of many small reflectors, which pose a special challenge for sonar systems limited to sparse sampling[8]. Eyes can image all spatial visible reflectors like leaves at the same time, but their perception mechanism is totally different to a biosonar's due to physical limitations of the airborne sonar. The first limitation of a biosonar signal is its low dimension. The 1D echoes received by ears are the result of superposition from 3D reflections. There is still no perfect means like a camera lens to decompose the signal concurrently into higher dimensions with good resolution, although some multi-pixel sonar system have reached the resolution beyond $64 \times 64$. The second limitation is the incompleteness of individual echo's representation of the observed target. All visible reflectors of a multi-faceted target shall give a reflection when ensonified, but only those that give a strong perpendicular reflection may be heard by the receiving sensors. Most features extracted from a single echo are not robust enough to be directly used for natural landmark classification. The third limitation is the signal distortion from spatial superposition. Due to the duration length $\Delta_t$ of the chirp sent, the received echo over time $A(t)$ doesn't solely represent the reflections over distance d(t). It brings trouble to our work of studying 3D geometrical difference through 1D signal probing. The second limitation gives reason for using statistical features, which can be extracted from large numbers of random echoes [8]. But it poses a hard challenge for the computing speed of mobile robots, which must make decisions quickly in real world navigation. Instead of utilizing echoes randomly, a sequential sensing strategy is used in this paper to solve this problem. By keeping the information of the original sampling sequence (section 3.1) we can extract robust cues about a target's structural geometry over sequential observations. In theory, the third limitation can be overcome with the utilization of a short enough chirp. But in reality a chirp's minimum duration is limited by the maximum sample frequency of the hardware, and a too short chirp can't generate echoes with enough energy and signal to noise ratio (SNR). In this paper a frequency after signal reconstruction algorithm is used to correct false superposition (section 3.2), it works only with FM chirps.

## III. SEQUENTIAL CLASSIFICATION METHOD

The hardware of our biosonar system consists of a robot which has two Intel Pentium 3 computer(Robin1 and Robin2), a National Instruments NI6110 analog I/O card, a Mini SSCII serial servo controller, a BNC2110 connector, and this biosonar head (figure1). The biosonar head consists of 3 Polaroid sensors in a triangular layout, similar to the layout of a bat's mouth and ears[10]: two Polaroid 600 sensors spaced 12.5$cm$ apart as ears, a Polaroid 7000 sensor as mouth in the middle between the two ears but 2.5$cm$ lower. With two turning ears, which have two degrees of freedom each, we can not only obtain a measurement of the target's distance and bearing to the robot by calculating time of flight (TOF), but also enhance the signal to noise rate (SNR) and record pairs of related echoes from 2 orientations for further study. The maximum sampling speed of the NI6110 card is 5$MHz$, we utilize 1$MHz$ in our research. The NiMH charger box provides the sensors with a 150$V$ power supply.

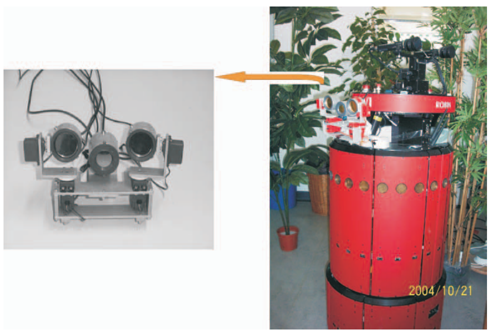


Fig. 1. Biosonar head

The software consists of four layers in Red Hat Linux (figure2). The NI6110 driver software is required for the AI/AO card to achieve 1$MHz$ input sampling speed.

### A. Sequential sensing

Since an individual echo is only a partial and incomplete representation of the observed target, and bats typically make some inspections of a potential target through hovering before a decision is made[2], we believe that the sequence of echoes can provide decisive information for complex natural landmark classification. The dynamic sensing strategy of bats is simulated by looking around the target with certain intervals and certain observing ranges. The sequence of continuous probing chirps along the navigation track is kept as the sensing sequence for the following processing. The primary movement strategy used in this paper is the biosonar's relative turning around 3 artificial landmarks. In this way an additional dimension of sequential movement is added into the observation space. The inter-echo's variations can also yield additional cues for classification.

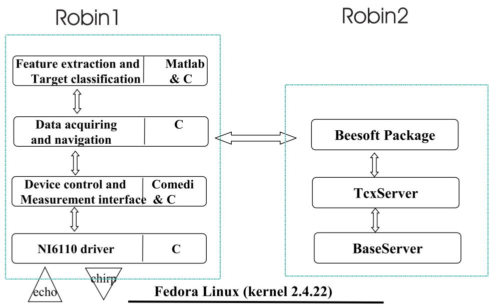


Fig. 2. Software structure

### B. Sequential reconstruction

Before extracting features, we used a reconstruction algorithm that we invented to overcome the superposition problem in sequentially sensed echoes (figure3). Through reconstruction of different frequency components in echo $A(t)$, we got the resulting signal $A'(t)$. $A'(t)$ corresponds directly with reflection along distance d(t).

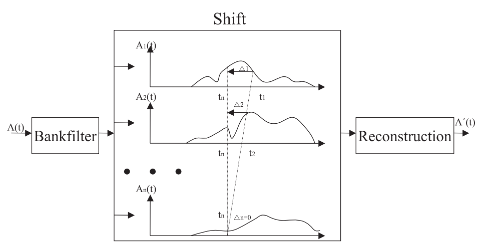


Fig. 3. Echo shifting and reconstruction according to frequencies

$$A^{'}(t) = A_1(t) + A_2(t + \Delta_2) + \cdots + A_n(t + \Delta_n) \quad (1)$$

$$\Delta_i = \Delta_t \times \frac{(f_{high} - f_i)}{(f_{high} - f_{low})} \quad (2)$$

$$\Delta_t = N/S \quad (3)$$

Where $Ai(t)$ is the result of band pass gamma tone filtering, whose centre frequency is $f_i$, $f_{high}$ and $f_{low}$ are respectively the maximum and minimum frequency of the sent FM chirp. Shifting time $\Delta_i$ is decided by the position of $f_i$ in the chirp, whose duration $\Delta_t$ is certain under certain chirp length ($N = 256$) and sampling rate($S = 1Mhz$).

### *C. Feature extraction and matching*

After signal reconstruction, Seven features are extracted (see table I ) from an individual echo signal $A'(t)$. They are selected from twelve tested features due to their performances. Their original sampling sequences are kept by storing them in feature arrays $f_i(\theta, d)$, where $\theta$ is the horizontal orientation, $d$ the distance between biosonar head and target, $i$ the feature code number (1-7). Those features' performance varies under different circumstances. In order to reduce their dependence on distance and keep the amplitude difference between trees, a normalization is performed before feature extraction (figure4), $A^{"}(t) = A^{'}(t)/\sum A^{'}(t)$.

Following extraction of seven 1D features (in table 1) from all individual $A"(t)$ in the observation range, three new 2D features are then calculated (S1, S2, S3) out of them, where the second dimension is the original sampling sequence and s3 is a combined feature from 3 to 7 features. In the following sections we only discuss three 2D signatures generated from those 7 features:

- S1, amplitude envelope signature (feature 1 in table)
- S2, local maximum's position signature (feature 2)
- S3, combined signature $\{f(i, j)\}$, $i \in [3, 7]$

Feature S1 and S3 are then processed with a moving window 2D correlation algorithm.

$$r = \frac{\sum_m \sum_n (A_{mn} - A)(B_{mn} - B)}{\sqrt{\left(\sum_m \sum_n (A_{mn} - A)^2\right)\left(\sum_m \sum_n (B_{mn} - B)^2\right)}} \quad (4)$$

where the sampled feature matrix $B$ is correlated with the stored matrix templates $A$ like a moving window. But feature S2 is so salient that a single randomly emerged perpendicular leaf may change it radically. It can only be used through statistical calculation. Here a Kolmogorov-Smirnov algorithm is used to calculate the individual signal similarity in sequence.

$$K^2 = \sum_{i=1}^{k} \frac{(EXP_i - OBS_i)^2}{EXP_i} \quad (5)$$

Three matching results from these 3 features are the input of the k-nearest neighbour algorithm to make the final classification. The whole process of this method is shown in figure 4.

Three matching results from these 3 features are the input of the k-nearest neighbour algorithm to make the final classification. The whole process of this method is shown in figure 4.

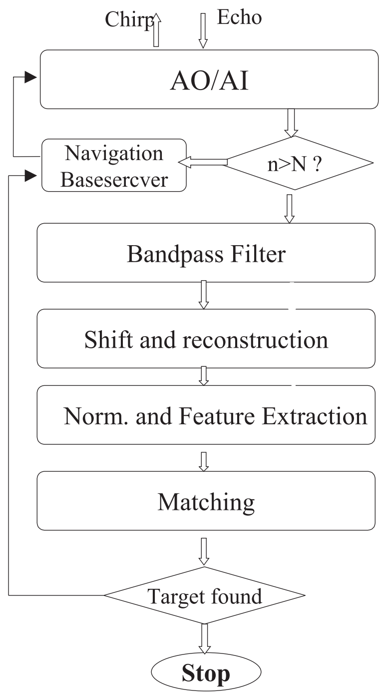


Fig. 4. Process of signal processing

Fig. 5. Three plants: ficus, bamboos and schefflera

## IV. EXPERIMENT RESULTS

Since it is believed that the frequency-modulated (FM) chirps are well suited for target localization [12], a $256us$ duration broadband FM chirp is used here. Natural landmarks in this paper were 3 artificial trees of similar size ($1.7m$ high, see figure5). Because there are normally no geometrical models of the multi-faceted landmarks, a biosonar can not decide precisely its distance to them from unpredictable orientation, we sample deliberately in $\pm 0.15m$ distance range to see how this method can tolerate some distance variance,

TABLE I
FEATURES

| Num | Name | formula | Denotation |
|---|---|---|---|
| 1 | Wave form<br>Distance | $d_{ij} = \sum_{k=1}^{n} (\overline{A}_i(k) - \overline{A}_j(k))$<br>$\overline{A}_i = \sum_{l=1}^{m} A_i(l)/m$ | envelope matching degree |
| 2 | Local maxi. P | $(p^i_{max})$ | local maximums' position set |
| 3 | Sum-area | $Sum_i = \sum_{k=edage1}^{edage2} A_{ik}$ | effective reflection area |
| 4 | Global max ampli. | $X^i_{\max} = \max_i (A_i(j) / \sum_{j=1}^{n} A_i(j))$ | the largest amplitude |
| 5 | Mean amplitude | $X^i_{mean} = \sum_{j=1}^{n} A_i(j)/n$ | average reflection ability |
| 6 | Crest factor | $X^i_c = X^i_{\max} \,/\, X^i_{rms}$ | impulsiveness of echo |
| 7 | Depth | $D_i = (edge(2) - edge(1))$ | scale of target |

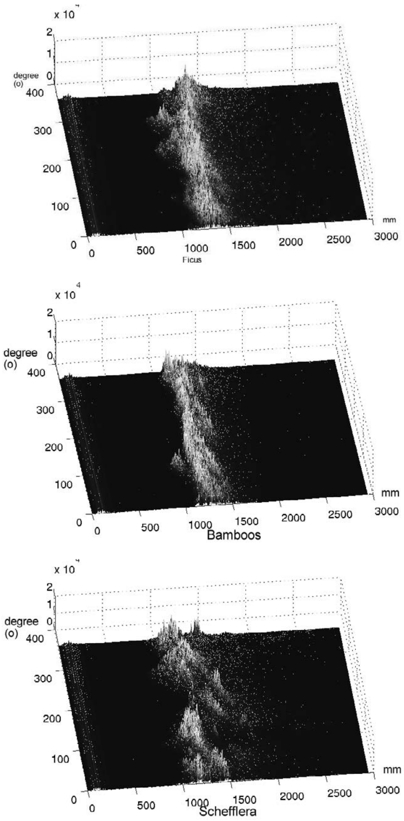


Fig. 6. One round of sequentially sampled echoes of 3 trees

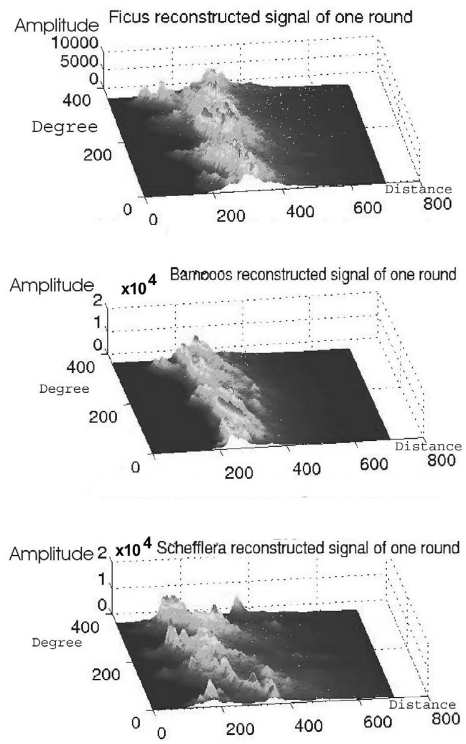


Fig. 7. The reconstructed signal from sampled echoes of 3 trees

which is important for robot's application. The results of one round of sequential sensing of three trees in $1.5 \pm 0.15m$ distance are shown in the figure 6. Their corresponding reconstructed signals are shown on the figure 7, where the direct noises is gotten rid of by cut the beginning part of the signal away .

Firstly, a statistical classification method that uses the features of Distant Interspike Intervals (DII) [8] was performed. It only works with a limited angle range of observation. When sensing from random relative orientations between trees and biosonar head, the classification rates with different orientations of observation are shown in figure 8. The horizontal axis in the figure denotes the observation orientation. From every 10 degree orientation 20 random echoes were sampled.

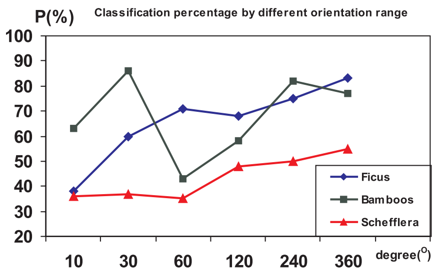


Fig. 8. Classification rate with different orientations by the statistical method using DII features

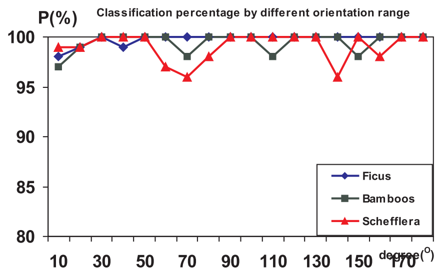


Fig. 10. Sequential classification rate under different orientations with signal reconstruction

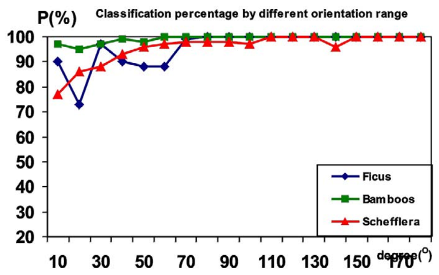


Fig. 9. classification rate with sequential sensing under different orientations without signal reconstruction

Secondly, the proposed sequential sensing method was tested. The signal reconstruction block in the classification process (see figure 4) is at first omitted and 2D Fast Fourier Transform (FFT) is used to pave the sequential echoes smooth before extracting 7 features. The classification rates under sequential observing angles are shown in figure 9, which is calculated from 420 sequential tests. The horizontal axis denotes the orientation of sequential observation.

Then, through another 420 sequential tests with signal reconstruction , the classification rates are shown in figure 10. The improvements in SNR and accuracy are obvious. The method can tolerate the adopt distance variations. The whole process of a 30 degree decision by our present system needs 7.6 seconds(5+2.6), where the scan movement of robot(robin in figure 1) takes about 5 seconds. They can still be shortened by optimizing the software and changing the robot respectively.

## V. Conclusion and future work

This paper has presented a method for natural landmark classification with biosonar. Test results indicate that obvious improvements in classification accuracy can be achieved by the use of this sequential sensing method. They suggest that a mobile robot can achieve the ability to classify natural landmarks like trees only based on sonar.

The research tried to take advantage of perceived properties of bats' prey identification and landmark identification mechanisms and strategies, without the claim to be a precise model of bats echolocation. It can not separate multi-targets from limited orientation ranges at present. Further work will include studying this correspondence problem, quantifying the distance compensation model, realizing different sensing strategies, and utilizing advanced computation algorithms to enhance classification performance.

## Acknowledgment

The authors would like to thank Dr. Rolf Mueller for designing the sonar head with which the experiments were performed. The first author would like to acknowledge the financial support by the German Academic Exchange Service (DAAD) during his PhD scholarship at the University of Tuebingen.

## References

[1] Akbarally and Kleeman. A sonar sensor for accurate 3d target localisation and clasification. In *Proceedings ICTA, Nagoya*, pages 3003–308, 1995.

[2] Dvon Helversen. Object recogition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine. *J Comp Physiol*, 189:327–336, 2003.

[3] Kimoto and Yuta. Sonar based outdoor navigation. In *Proceedings ICAR*, pages 239–244, 1993.

[4] Roman Kuc. Biomimetic sonar recognizes objects using biaural informations. *J. Acoust. Soc. Am*, 102:327–336, Aug. 1997.

[5] Leonard and Durran-Whyte. Mobile robot localisation by tracking geometric beacons. In *IEEE Transactions on Robotics and Automation*, pages 376–382, 1991.

[6] P.J. McKerrow. Robot perception with ultrasonic sensors using data fusion. In *Proceedings IEEE S,Man and Cybernetics Conference*, pages 1380–1385, 1995.

[7] Phillip John Mckrrow and Neil Lindsay Harper. Recognizing leafy plants with in-air sonar. *Sensor review*, 19:202–206, 1999.

[8] Rolf Mller. A computational theory for the classification of natural biosonar targets based on a spike code. *Comput Neural Syst*, 14:595–612, 2003.

[9] Cynthia F. Moss. Auditory scene analysis by echolocation in bats. *J. Acoust. Soc. Amer.*, 110(4):2207–2226, Oct. 2001.

[10] Mueller R. Carmena J. Hallam Peremans, H. A biomimetic platform to study perception in cf-bats. In *Sensor Fusion and Decentralized Control in Robotics Systems III*, volume 4196, Nov. 2000.

[11] PA RACEY R ARLETTAZ, G JONES. Effect of acoustic clutter on prey detection by bats. *Nature*, 414(6865):742–5, Dec. 2001.

[12] H.-U. Schnitzler and E.K.V. Kalko. How echolocating bats search and find food. In *Phylogeny, Morphology, Echolocation and Conservation Biology*, pages 183–196, 1998.

[13] Peremans H. Walker, V.A. and Hallam. One tone, two ears, three dimensions: A robotic investigation of pinnae movements used by rhinolophid and hippoisiderid bats. *J. Acoust. Soc. Amer.*, 104(1):569–579, Aug. 1998.