

Complex echo classification by echo-locating bats: a review

Yossi Yovel · Matthias O. Franz · Peter Stilz ·
Hans-Ulrich Schnitzler

Received: 2 February 2010/Revised: 11 August 2010/Accepted: 28 August 2010/Published online: 17 September 2010
© Springer-Verlag 2010

Abstract Echo-locating bats constantly emit ultrasonic pulses and analyze the returning echoes to detect, localize, and classify objects in their surroundings. Echo classification is essential for bats' everyday life; for instance, it enables bats to use acoustical landmarks for navigation and to recognize food sources from other objects. Most of the research of echo based object classification in echo-locating bats was done in the context of simple artificial objects. These objects might represent prey, flower, or fruit and are characterized by simple echoes with a single up to several reflectors. Bats, however, must also be able to use echoes that return from complex structures such as plants or other types of background. Such echoes are characterized by superpositions of many reflections that can only be described using a stochastic statistical approach. Scientists have only lately started to address the issue of complex echo classification by echo-locating bats. Some behavioral evidence showing that bats can classify complex echoes has been accumulated and several hypotheses have been suggested as to how they do so. Here, we present a first review of this data. We raise some hypotheses regarding possible interpretations of the data and point out necessary future directions that should be pursued.

Keywords Echolocation · Classification · Bat · Statistics · Behavior

Introduction

Using echolocation, bats routinely detect, localize, and classify targets in their surroundings (Griffin 1958; Schnitzler et al. 2003). Among these three fundamental tasks, the question how bats use echolocation to classify targets is the least studied and the least understood.

The objects a bat has to classify can be grossly divided into two categories: Foraging bats have to recognize their food, e.g. insects, fruit, etc. while orienting bats have to classify objects in the environment, e.g. water surfaces, meadows, walls, vegetation, etc. The mostly small food targets often differ from the mostly large extended targets in the environment in their complexity. Small food items are mainly comprised of a few reflectors and thus reflect rather simple echoes, whereas extended targets like vegetation produce complex echoes that are usually superpositions of many reflections from the many reflectors a plant contains (i.e. its leaves and branches). Most previous work on classification was done in the context of simple objects. These studies mainly focused on testing and explaining the abilities of bats to discriminate finely controlled spectral and temporal differences. Such studies helped to understand the limits of the system and to reveal its mode of operation, but they generally relied on unnatural deterministic echoes. Complex echoes from natural textures such as plants were either ignored or regarded as disturbing clutter. The recent work that will be described in this review, however, demonstrates how these echoes could contribute to a meaningful acoustic image of the bat's surroundings. The importance of complex echoes to bats is very high. Vegetation can be an

Y. Yovel (✉)
Department of Neurobiology, Weizmann Institute of Science,
Rehovot, Israel
e-mail: yossiyovel@hotmail.com

M. O. Franz
Faculty of Informatics, University of Applied Sciences,
Constance, Germany

Y. Yovel · P. Stilz · H.-U. Schnitzler
Animal Physiology, Institute of Neurobiology,
University of Tuebingen, Auf der Morgenstelle 28,
72076 Tuebingen, Germany

indicator where fruit or insects are typically found (Kalko and Condon 1998; Thies et al. 1998; Schmidt et al. 2000; Levin et al. 2009) and can serve as a landmark to facilitate navigation, roost finding, and habitat selection.

In this review, we will focus on classification of objects that generate complex echoes, mainly vegetation. To date, there is still much that we do not understand regarding complex echo classification. However, in the light of the recent accumulated data, it is worthwhile to try to summarize our knowledge and to point out the issues yet to be addressed.

We will limit our discussion to the classification of stationary objects according to echoes generated by short broadband frequency-modulated (FM), echolocation calls that are emitted by many bat species. Such calls usually have a duration of a few milliseconds and a bandwidth of several tens of kHz (e.g. Pye 1980; Schnitzler et al. 2003). Moreover, in addition to exact temporal information, they deliver plentiful spectral information about the features of the reflecting object; hence, they are advantageous for echo classification.

We will briefly discuss the effects of the bat's motion in relation to the object or the object's motion in relation to the bat (see "Discussion"). The classification of objects that exhibit rapidly oscillating motion is beyond the scope of this summary. This kind of motion, typical for insect wing fluttering for instance, introduces unique acoustic cues that can be detected either by emitting very long calls or by a careful comparison of successive echoes (Schnitzler et al. 1983; Schuller 1984; Sum and Menne 1988; Kober and Schnitzler 1990).

Classification terminology

We will adopt the terminology suggested by Au (1993) who distinguished between discrimination (wherein the subject has to decide whether two echoes come from the same object or from different objects) and classification (wherein the subject has to learn a criterion that differentiates between two classes of objects). A third less common category in which the reference target is not varied at all is termed recognition, which is the least demanding task in terms of generality because it does not require to learn characteristic features of a reference object or class but rather memorize it (e.g. learn that a plant specimen has a large branch sticking out rather than learning the typical spatial distribution of its branches).

Echo creation and acoustic cues

The interaction of a sound wave with the object forms an echo (Skolnik 2001). The impedance of solid objects is

significantly larger than that of air. Therefore, when a sound wave impinges on a solid object, there is almost no absorption. The energy is either reflected or scattered in different directions depending on the size and the geometry of the object. The echo returning from an object reveals its filtering characteristics that convey information regarding its physical properties. The bat has to "inverse" the laws of reflection and propagation to reconstruct the object from its acoustic characteristics, a classical ill-posed inverse problem. The acoustic parameters of an echo contain information about the physical properties of the object from which it was reflected. None of these parameters, however, can be uniquely attributed to a single feature of the object.

Spectral cues

A broadband signal (such as an FM bat call) that is reflected from a sphere with a circumference within the range of the signal's wavelengths will be attenuated at certain frequencies in comparison to others. This attenuation pattern is known as the frequency response of the object and is sometimes referred to as the acoustic color of the object.

The frequency response of an object reveals information on its acoustic cross-section, which is a function of the relation between the object's geometry and the wavelength of the impinging signal (3–10 mm for a typical FM call). Intuitively, the cross-section can be thought of as the effective reflecting area of the object and is therefore related to its size. Echo intensity linearly depends on the cross-section according to the radar equation. For a simple spherical reflector, the cross-section is determined by the ratio between the circumference of the sphere and the wavelength. Its behavior can be classified into three domains, namely (1) the Rayleigh domain: if the circumference is smaller than the wavelength, the cross-section (and the intensity of the echo) decreases rapidly when decreasing the radius or increasing the wavelength, the decrease is inversely proportional to the fourth power of the wavelength; (2) the resonance domain: if the circumference is of the same order of magnitude as the wavelength (up to 10 times larger), the intensity of the echo oscillates depending on the ratio and has peaks that are larger than the projected area of the sphere; (3) the optic domain: if the circumference is more than an order of magnitude larger than the wavelength, the intensity of the echo is equal in all frequencies and depends only on the radius of the sphere.

It is thus clear that if the object is a perfect sphere, its size can be inferred from its frequency response with good precision. Objects with more complex shapes also exhibit these three domains, but the relation between cross-section and the geometry is not simple and also depends on the

impinging angle. Therefore, for these objects the size can only be partially estimated. The intensity of the echo at different frequencies is also highly influenced by the roughness of the surface from which it is reflected. A rough surface will scatter the energy more than a smooth one and will therefore reduce the echo's intensity. The amount of scatter will be determined by the granularity of the rough surface. Grossly speaking, when the grains are smaller than the wavelength the object can be regarded as a single object with many Rayleigh scatterers, while when they are larger than the wavelength it can be regarded as a multiple reflector object.

Objects with two nearby reflectors may convey additional spectral information. The distance between nearby reflectors (within the integration time of the auditory system) is encoded in the spectral domain as notches of energy at frequencies that correspond in wavelength to twice distance between the reflectors and peaks of energy at frequencies that correspond to integer multiples of this distance. A pure spectral analysis of such two wave-front echoes can therefore provide temporal or spatial information. (Simmons et al. 1990a, b; Mogdans et al. 1993).

Temporal cues

Mathematically, the echo received by the bat can be described as a convolution of its emitted signal with the impulse response (IR) of the object. The IR is defined as the echo generated by an object when ensonifying it with a perfect Dirac pulse (infinite bandwidth and infinitesimal duration). The IR is a key term when discussing temporal cues for classification because its statistics shape the temporal statistics of the echo. It is thus sometimes referred to as the object's shape along the axis of range (Saillant et al. 1993) or the acoustical image of the object (Schörnich and Wiegrebe 2008).

Many objects contain discontinuities that act as separate reflectors (reflectors that are spatially spread farther than the shortest wavelength). These objects can be described as three-dimensional (3D) arrays of reflectors. For such objects—in addition to the frequency response of each reflector that reveals its cross-section—the temporal sequence of the echo conveys information regarding its spatial structure. The entire duration of the echo reveals the total depth of the object, while the internal time intervals reveal the distances between reflectors in the direction of the bat, sometimes referred to as the range map of the object. These intervals are extremely important for complex echo classification and most suggested classification models (see below) rely on them. A comparison of the temporal sequence received by the two ears can provide additional information about the positions of the reflectors.

Biological constraints on the extraction of acoustical cues: the bat's receiver model

Both the spectral and temporal cues that can actually be extracted by bats are constrained by the performance of their auditory system. The exact nature of the bat receiver, however, is currently under serious debate (Schnitzler and Henson 1980; Wiegrebe 2008). A major question is whether the bat's receiver can be compared to a coherent or to a non-coherent matched filter model. Briefly, the matched filter receiver is based on cross-correlating the emitted signal with the received echo. This can be implemented in the time or frequency domain, resulting in a compressed version of the original echo that makes the fine temporal information more visible. The ideal coherent matched filter on the one hand preserves phase information and suggests that bats can access the complete cross-correlation of the emitted and received signals. A semi-coherent matched filter on the other hand does not extract phase information but accesses only the envelope of the cross-correlation function. An ability to rely on the exact IR of an echo for classification thus would imply a coherent matched filter. Consequently, the exact nature of the receiver has critical implications with respect to the type of features that can be extracted for simple object classification (see below). However, because we will concentrate on complex echo classification, the exact fine limits of the bat receiver are of lesser importance in our context. The auditory processing of the bat's inner ear is usually modeled (similarly to most mammalian systems, but with sensitivity to higher frequencies) as a gammatone filter bank (Patterson 1994). The bandwidth of the channels (defining the frequency resolution) depends on their center frequency and can be roughly estimated for FM bats as 1–2% of the center frequency of the channel (von Stebut and Schmidt 2001). The resulting limited spectral resolution is, in our view, not a limiting factor for echo classification, as the conceivable spectral cues do not require such a high spectral resolution. The temporal resolution of the auditory processing is mainly limited by the temporal integration of the inner hair cells (Palmer and Russell 1986). The time integration of the system is usually assessed to be around 100–300 μ s (Sanderson and Simmons 2000; Weissenbacher et al. 2002) although several models have suggested how a better temporal resolution can be achieved despite this (Saillant et al. 1993; Peremans and Hallam 1998; Sanderson et al. 2003; Boonman and Ostwald 2007).

Simple objects versus complex objects

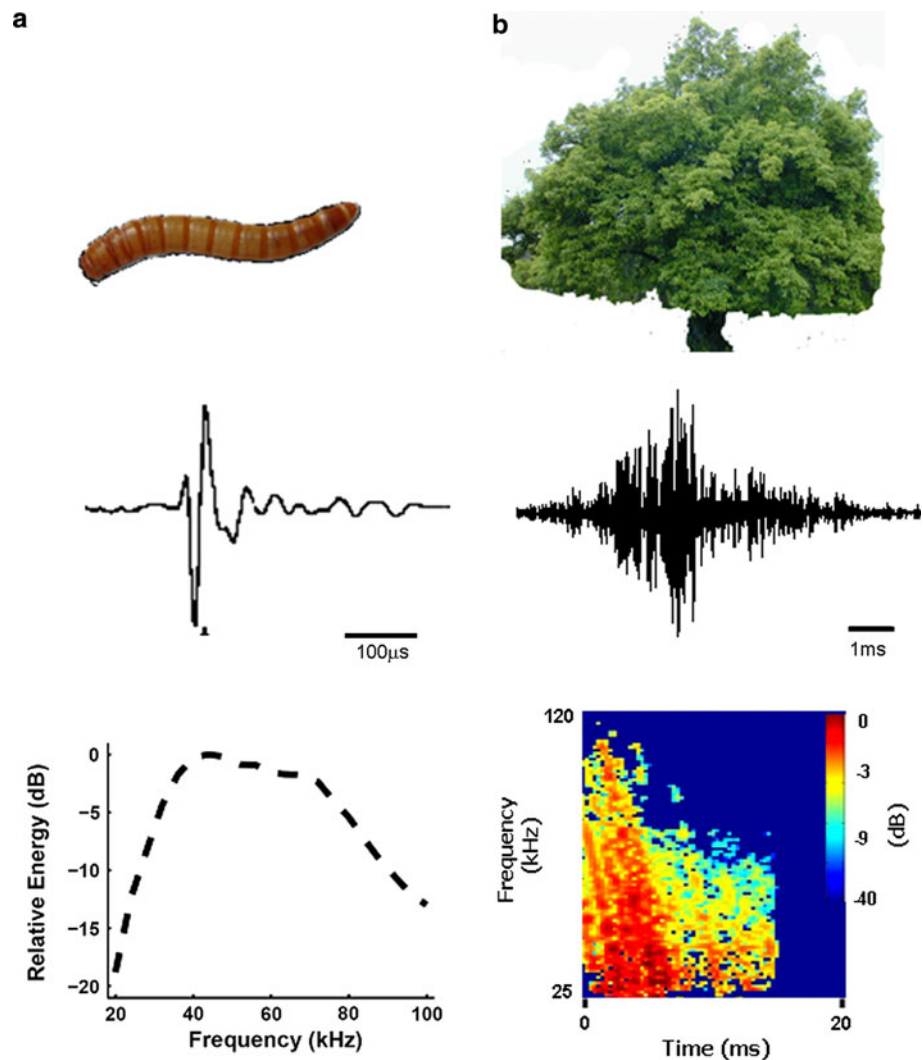
The research done on echo-based object classification can be historically divided into simple versus complex object

classification. There were two main motivations behind this division: From the researchers' point of view on the one side, a simple object can convey exact information that can be extracted using deterministic approaches. As will be elaborated below, such approaches provide insights into the limits and mechanisms of the bat's auditory system. From the bat's point of view on the other side, simple versus complex objects correspond to different categories of echoes a bat has to deal with in its daily life: simple objects represent isolated prey (insect, fruit) while complex ones stand for background (vegetation) or prey in front of background (Fig. 1). A simple division into deterministic versus stochastic echoes that is sometimes suggested, is not adequate since due to the high non-linearity of sound reflection, even the echo of a very simple object cannot be predicted in a one-to-one manner. A physiological definition suggested by some authors refers to objects with more than two reflections returning within a time interval shorter than the resolution of the auditory system ($\sim 1\text{--}50\ \mu\text{s}$) as complex. Auditory models (Saillant et al. 1993; Peremans

and Hallam 1998; Sanderson et al. 2003) are not able to resolve such reflectors suggesting that the exact spatial shape of the object cannot be inferred. However, the temporal vicinity of the reflectors might not be the entire story: even for a deterministic echo consisting of nicely separable glints that can be resolved; a slightly different angle of ensonification might result in a significantly different glint sequence. In order to perceive an object deterministically, the exact positions of its reflectors have to be resolved; for this, it must be scanned from many different angles. In this sense, the echoes of many objects have to be viewed as realizations of random processes, despite their origin in a deterministic reflection process.

Another common approach is to define complexity according to the physical characteristics of the object. On the one hand, a simple object tends to be small and contains up to a handful of reflectors. In many cases, it can be decomposed into simple shapes. An insect for instance, can be thought of as a sphere (head) followed by a cylinder (body) with two or four flapping plates (Ostwald et al.

Fig. 1 Complex versus simple echoes. **a** The echo and magnitude-spectrum of a meal worm, an object typically used in prey catching experiments with bats (re-measured from Simmons and Chen 1989). **b** The echo and spectrogram of an apple tree can be used by the bat to mark a food source, or as a navigation landmark



1998). A complex object on the other hand, can be described as a series of discrete reflectors, distributed along the dimension of the object's range (Altes 1976). Complex objects are usually not small, in many cases, their depth is larger than the duration of the bat's call, and they spread beyond the range covered by the beam of the call. When the latter occurs, the global shape parameters of the object are not available to the bat and the bat actually is faced by what is sometimes terms a texture classification task. The distance from which an object is ensonified will thus influence its perceived complexity. When coming close enough even the most complex object will become simple in the sense that it will only generate a few reflections. The farther the object the more reflections from it will return in temporal vicinity.

It is therefore clear that there is no definite threshold between simple and complex objects. The exact definition of complexity is, however, less important for our purposes. The classification tasks we will discuss in this review are clear cases of complex echoes, i.e. situations in which an object is represented by many reflections that return in temporal vicinity (sometimes shorter than the temporal resolution of the auditory system). This is the case in many behaviorally relevant tasks. A bat flying 1.5 m away from a plant for example, emitting a 70° beam (first zero at 70°) will ensonify the plant with >4 m circular sonar print which will be translated into an echo that is a superposition of dozens of reflections. The behavioral context of the task will usually not allow a deterministic scan of the object from many angles.

What can we learn from simple object classification on complex echo classification? Conclusions and open questions

After half a century of studying echo-based classification of simple objects many issues are still open. Behavioral studies have revealed substantial quantitative evidence for the abilities of bats to classify echoes (Griffin et al. 1965; Webster and Brazier 1965; Konstantinov and Akhmarova 1968; Bradbury 1970; Simmons and Vernon 1971; von Helversen and von Helversen 2003; Weissenbacher and Wiegrebe 2003; von Helversen 2004; Schörnich and Wiegrebe 2008 and many others). These studies showed that bats discriminate simple object characteristics such as size, texture, duration, and shape (for a review see Moss and Schnitzler 1995).

Despite this vast amount of behavioral research, we have only little understanding of the acoustic cues used by the bats for classification. One main difficulty in understanding the relevant acoustic cues is that several cues vary simultaneously when any of the physical characteristics

mentioned above is changed. Experiments with virtual played back echoes partially solve this problem, by enabling artificial control on acoustical cues. Echo intensity is probably the most established and examined acoustic cue that was found to be useful for classification. Bats are known to be able to distinguish between intensities that differ by 1–3 dB (Schnitzler and Henson 1980). Simmons and Vernon (1971) for instance found that the absolute intensity of the echoes could mediate classification of equal-shaped objects with different sizes and equal-sized objects with different shapes. Temporal cues were also examined. Simmons and Chen (1989) for instance suggested that intervals between intensity peaks of the echo's IR could explain the bats' ability to classify mealworms from disks. To this end, they assumed that bats integrate information from more than a single echo which was also assumed by other researchers (e.g. Moss and Surlykke 2001), but still not solidly proven. One main caveat in many of the studies is the lack of quantification: only a handful of them attempted to quantify the acoustic cues that might mediate classification. Most authors only described these cues visually and did not try to measure or model the classification performance they afford. This was partially due to the limited computation power and should be easier to do today.

Spectral cues were mainly suggested to facilitate bats' observed hyperacuity in discriminating two wave-front objects, namely, objects with two closely spaced reflecting surfaces (Simmons et al. 1974; Habersetzer and Vogler 1983; Schmidt 1992; Mogdans et al. 1993). These studies suggested that the cancellation of frequencies corresponding to the distance between two nearby surfaces (spectral notches) explain the bats' ability to discriminate between objects with internal distances that differ by less than 1 mm. It was therefore commonly proposed that bats perceive the range of nearby reflectors (closer than the integration time of the auditory system) according to their spectral information content (Simmons et al. 1974). Several researchers presented evidence supporting this hypothesis (Schmidt 1988, 1992; Simmons et al. 1990a, b; Saillant et al. 1993). Others demonstrated how spectral notch patterns can represent the size of hollow hemispherical objects (Simon et al. 2006), however, since the temporal and spectral domains contain identical information, behavioral experiments will not be sufficient to settle the debate which domain is used by the bats, probably until the bats' receiver will be fully understood (see above). Frequency response based classification was tested for the case of classification of mealworms from disks (Griffin 1967) but it was found mostly useless for this task.

Several authors developed biologically inspired models that explain the simple object classification behavior observed in bats (Saillant et al. 1993; Peremans and Hallam

1998; Sanderson et al. 2003). These models typically offer a mathematical framework imitating the auditory processing of the inner ear and the corresponding generation of neural activity (some include additional processing before or after these main parts). Trying to explain simple echo reconstruction (mainly two-wave-front object discrimination), these models focused on the exact ranging of nearby targets. They are thus mainly relevant for classification of simple objects, but they were later generalized to explain complex echo classification as will be described below.

Classification of complex objects

Behavioral evidence for complex object classification

In one of the few behavioral experiments dealing with the classification of complex echoes from natural targets, Denzinger and Schnitzler (unpublished data) trained Natterer's bats (*Myotis nattereri*) to discriminate conifers from broad-leaved trees in complete darkness. During the training phase, the bats were always confronted with the same pair of artificial plastic plants (one conifer and one wide-leaved plant). They learned to associate the wide-leaved plant with food by hanging mealworms on it and correspondingly spent most of their time flying around it. In the test phase, when the bats were presented with new pairs of conifers and wide-leaved plants, they continued to significantly spend more time circling the wide-leaved plants. These results revealed that bats could classify conifers from deciduous plants. The ability of the bats to generalize the knowledge they acquired during training implied that they learned some acoustic features of the plants rather than the specific echoes of the two training stimuli.

Statistical acoustic features

The echo trains returning from any of the plants described above contained many dozens of reflections that represent—due to the large size of the object—only a limited part of it. Similar echo trains are perceived by commuting bats that must classify extended targets like vegetation to make navigation decisions. We assume that a plant is perceived as a spatial arrangement of reflectors with certain statistics. There is therefore probably no point in discussing acoustic cues that represent the global features of the object such as its shape or size if we look at single echoes. To get anything near the global features of a plant, a bat would have to scan the object with several signals from many angles. The only global feature that might be relevant in some cases is the object's depth as will be further discussed below. Most of the cues used for classification of simple

objects (e.g. intensity, temporal statistics) were also tried as possible cues for complex echo classification. The precise deterministic manner in which they were used to describe simple objects, such as the exact range profile of the echo needed to be replaced by distributions of these parameters. Interference patterns (notches) that were widely explored are probably useless when concerning highly complex echoes because of the wide variety of distances between reflectors that do not result in any systematic patterns. In fact the debate whether the bats' receiver preserves phase information with access to the fine temporal structure of the echo ($< \mu\text{s}$) is insignificant because complex echoes are characterized by very complex phase (Yovel et al. 2008). Only cues that take into account the statistics of the echoes are reasonable for classifying complex objects. Spectral cues providing information about reflector sizes as well as simple statistics; such as the distribution of reflector intensities are probably insufficient in most cases due to the large intraspecies variability (individual plants contain reflectors in a large variety of sizes and distances). Therefore, a major question regarding the actual possibility to classify natural complex echoes is whether reliable cues can be extracted from their stochastic temporal representations. We will divide the various models suggested for complex echoes classification into three categories according to the temporal resolution they require (low, medium, or high).

First statistical models (Medium temporal resolution models)

One possibility would be to use statistics that are more sophisticated. In one of the first attempts to investigate complex plant echo statistics Müller and Kuc (2000) compared the echoes of two potted plants (both with ca. 100 leaves) representing species with different spatial structures: (1) The weeping fig (*Ficus benjamina*) that we shall simply term fig has flat, oval leaves arranged in sparse foliage. (2) The spreading yew (*Taxus media*) which we will call yew has small, needlelike leaves, which are much more densely packed compared to the fig. The authors used a biomimetic sonar system to emit Gabor-wavelet-like impulses with spectral energy at frequencies between 30 and 100 kHz, which approximated the IR of the plant from a certain angle and position. On the one hand, the use of the IR requires the bat to have the controversial coherent matched-filter receiver, but on the other hand, most parameters that were extracted from the IRs did not rely on its fine temporal resolution and suggest that the raw echoes also contain this information. The authors scanned each plant from its two opposite sides and from two different bearing angles ($\sim 0^\circ$, $\sim 45^\circ$) for each side. For each of these four main positions, they used a b-scan protocol

that emitted a signal every 1 mm along a cross-range interval of 30 cm. This procedure resulted in several hundred angles for each of the four main positions. To analyze the echo statistics, they measured several parameters and compared their distribution among the echoes of each species. A main analysis relied on fitting the echoes time-series (after some preprocessing) to an α -stable distribution that was found to fit better than a Gaussian one. Two of the distribution parameters were found to differ between species: the characteristic exponent of the distribution reflecting the heaviness of the tails and thus representing the impulsiveness of the plant's structure and the dispersion that reflects the width of the distribution. The authors based their measure of classification on the following seven parameters: (a) The impulsiveness and the dispersion calculated from the fitted α -stable distribution. (b) The crest factor: the ratio between the IRs maximum squared amplitude and its total energy. (c) The total energy of the echo. (d) The echo's maximum amplitude. (e) The variance of range estimates: the standard deviation of the distance of the first peak in the echo when varying a threshold. This parameter is related to the distribution of the reflectors. (f) The regression coefficient: the correlation between time and the decreasing gain, calculated as the time until a logarithmic fit of the echo-envelope attenuates by a certain amount. This parameter is related to the propagation characteristics of sound within the plant.

The authors tested classification of the species using a single or a pair of these parameters. They estimated the probability density functions of the different parameters and estimated the Bayesian errors of classification. The results showed a rather high classification performance with error rates of 1–30%. The most suitable parameters for classification were the regression coefficient and the crest factor. Classification was possible based on just a single echo, and was possible both when the sound incidence was normal or oblique to the plant.

An interesting point made by the authors is the existence of species-specific cues that might be very informative but non-reliable. In the case of fig for instance, the leaves are rather uniformly oriented towards the outside, resulting in strong coherent reflections. When these reflections are observed, they exclusively represent fig (within the fig-yew data set) but they might be absent and are thus not solely sufficient for classification.

The authors emphasized two key points regarding the statistics of foliage echoes: The profound stochastic non-Gaussian nature of echoes and their non-repeatability, or as they put it, “the same echo from a foliage is never seen twice.” They therefore concluded that invariant features suitable for classification might be found only in the realm of statistical expectations. This work set the first footstep towards a better understanding of the statistics of natural

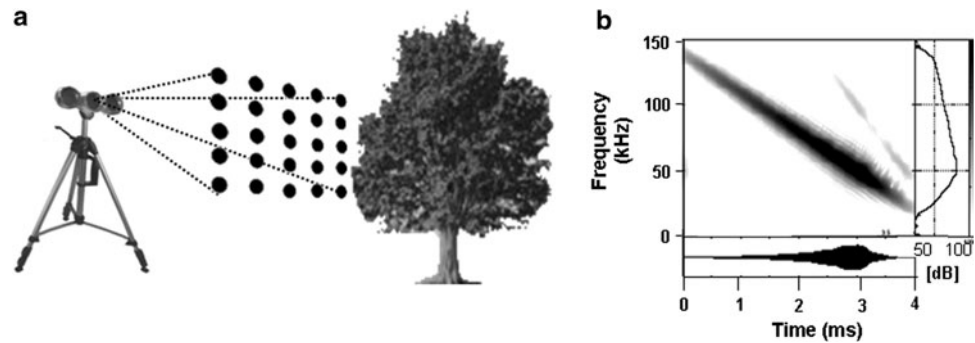
echoes as perceived by echo-locating bats. It provided evidence that complex echo classification might not be as difficult as previously thought. Its main limitation was, however, the restricted sample of plants that represent only a limited part of the variation bats encounter outdoors. The sample was restricted in two senses: the examined plants were small (perhaps representative of a natural plant branch), and only a single specimen was ensounded for each species. With complex spherical-symmetric objects such as plants, it is legitimate to assume that echoes from different rotational angles represent new specimens (McKerrow and Harper 2001). Still, a sample of only one specimen is probably not sufficient to model the natural variation bats encounter.

Another drawback of this study lies in the fact that it examined two plant species that differ greatly in structure: wide leaved versus conifer. The authors were aware of this point and concluded that due to the high similarity they observed in the parameters of these extreme cases “The distance between them in the studied feature space leaves little room for reliable discrimination of intermediate foliage types.” They therefore suggested using methods “to increase the resolving power of foliage classification” such as integration of information from several echoes. Yovel et al. (2008) indeed demonstrated that classification of these two extreme types of foliage is rather easy, but they were also able to suggest cues for classification of other cases of more similar foliage using only a single echo.

First large data-set (Medium temporal resolution models)

Stilz (2004) extended the work of Müller and Kuc (2000). He used a biomimetic sonar to emit 4 ms bat-like FM down-sweeps (140–25 kHz) and recorded the echoes of five plant species, representing a variety of the common species in the environment of the bats of central Europe (Fig. 2). Fifty specimens of each species were ensounded in situ from 25 angles each, creating the largest natural-echo database available today. The five species ensounded differ in the global and local physical structure and include Apple tree (*Malus sylvestris*), Norway spruce tree (*Picea abies*), Common beech tree (*Fagus sylvatica*), Blackthorn bush (*Prunus spinosa*) and Corn field (*Zea mays*). Stilz examined the ability to classify species based on a series of parameters extracted from the echoes or their IRs: the maximum amplitude, the mean amplitude, the crest factor, the amount of peaks above two different thresholds, and the fourth moment (see below). Using discriminate function analysis he was able to achieve high classification accuracy (above 80%), especially when using more than one parameter. The classifiers he suggested rely on the

Fig. 2 Acquisition of complex echoes. **a** A biomimetic sonar head was mounted on a tripod and taken to the field to record the echoes of five plant species in situ from 25 angles each. **b** The emitted signal was a 4 ms bat-like FM linear down-sweeps (140–25 kHz)



extraction of simple acoustical parameters from the echoes that do not require any fine temporal or spectral resolution. The simplicity of Stilz's approach implies the existence of clear species-specific cues in complex plant echoes. Stilz also addressed the question of integration information from several echoes, and in some cases, he found it advantageous to use the standard deviation of a parameter as measured from several angles.

Classification of virtual complex echoes

The group of Wiegand introduced phantom echoes to study classification of complex echoes by bats. This enabled them to generate controlled artificial complex echoes that vary according to a desired statistical parameter. Using this experimental design they showed that bats of the species *Phyllostomus discolor* can learn to discriminate echoes with IRs that have different levels of roughness (Grunwald et al. 2004). Roughness, defined as the base-10 logarithm of the fourth moment, indicates the impulsiveness of a signal and the non-Gaussianity of its distribution. It has been shown to be potentially advantageous as a parameter for discrimination of plant echoes (Stilz 2004). In the behavioral experiments, bats were first trained to discriminate between two echoes that largely differed in their fourth moment (a smooth class and a rough class). In the test phase, new echoes were presented with roughness levels as the training ones or with intermediate levels of roughness. The bats classified echoes with lower roughness as smooth, which suggested an ability to generalize from the training phase. The IRs used to create the echoes were random and extremely complex, each comprising up to 4,000 reflectors. These IRs were objectless in the sense that they were not created by a real object or a model of a real object.

Neural inspired model (High temporal resolution models)

Müller (2003) suggested a neural inspired approach to classify the echoes of four deciduous plant species:

sycamore (*Platanus hybrida*), linden (*Tilia cordata*), field maple (*Acer campestre*) and hornbeam (*Carpinus betulus*). He ensounded a single specimen of each species with a 3 ms FM (120–25 kHz) bat-like linear down-sweep from a variety of positions. Next, an algorithm based on the processing of the auditory system, translated the echo into a sequence of interspike intervals. The model included a fourth-order gammatone filter bank with a quality Q (the ratio between the center frequency and the filter bandwidth) defined at 3 dB, a half-wave rectification and finally, low-pass filtering with a first-order recursive low-pass filter ('leaky integrator') with different time constants τ . This signal was normalized such that waveform shape and not energy (influenced by target range and size) will play a role. Spikes were then determined from the normalized signal according to a vector of 1,024 thresholds. Analysis was based on the interspike intervals between the times the signal crosses the different thresholds. The authors only considered interspike intervals longer than 30 μ s (i.e. long interspike intervals) thus assuring that the spikes were created by separate reflectors and that they are feasible despite the time resolution physiological limits. Classification was finally based on three parameters calculated from each sequence of interspike intervals: The average duration of long intervals, the average time of occurrence of long interspike intervals and the number of long interspike intervals. Note that these measures are related to the roughness measure suggested by several authors, but they do not require the fourth moment estimation. The 3D joint probability density functions of these parameters in each species were sufficiently different to classify individual echoes with a success rate of more than 90%. Interestingly, in this work four deciduous plants were used thus exhibiting a much more difficult problem than the deciduous versus conifer one. The limited sample size was the main caveat of this study because only a single specimen (artificially constructed from several branches) was used for each species. This forced the author to rely on potentially risky procedures such as using the same specimen to model the species statistics and to test classification performance.

Other neural models

Several more recent works suggested biologically inspired models that could explain complex echo perception (Matsuo et al. 2004; Boonman and Ostwald 2007; Wiegrebe 2008). These models concentrated on the attempt to measure exact time intervals between reflectors and came up with very high temporal resolution in the order of at least 50 μ s.

The exact time-intervals approach could explain some findings, but it is usually limited when many reflections return with very small-time intervals as is the case for very complex objects such as plants. A simple simulation of a plant with evenly distributed leaves with a typical distance of 20 cm between them (like in a hedge) and a bat positioned 1.5 m from the first row of leaves, reveals that dozens of reflections will return within the integration time. This will be even more severe if the bat is farther, as is often the case in real life, because the beam impinging on the object will become wider. In addition, although they assume that perception of complex echoes is based on a highly exact temporal representation reflecting distances, none of these models cope with the angular (azimuth and elevation) position of the reflectors, which is necessary in order to reconstruct the actual 3D shape of the object. The ongoing debate whether bats create 3D spatial representation of simple objects seems much less relevant in the case of complex objects. Whereas for an object with a few reflectors, a bat might be able to localize all of them in space based on binaural information, this is probably undoable for highly complex echoes. The main obstacle preventing this is the difficulty to identify the two corresponding reflections returning from a single reflector in the two audio streams received by the ears. This is the case because many reflections return practically at the same time. The 3D reconstruction problem is thus ill-posed, meaning that each echo could have been generated by more than one arrangement of reflectors.

From a functional point of view, the benefit of reconstructing an exact 3D representation of the object is not clear for complex echoes. In the context of our discussion, these objects are used as landmarks (for food, roost, navigation, etc.). The position, orientation, and shape of individual reflectors in foliage are not immediately relevant for these behavioral goals. In contradiction to tasks like landing on an object or catching it, the tasks above will require further feature extraction (in order to classify the object) even if a perfect 3D representation of it is calculated.

Low-resolution non-stationary spectro-temporal models

Most of the models mentioned up to now (Müller and Kuc 2000; Stilz 2004; Grunwald et al. 2004) used

parameters that were extracted from the entire echo, thus assuming that the statistics of natural echoes are stationary, i.e. they do not change during the duration of the echo. This assumption is not valid in many cases. A typical tree, as was already pointed by Müller and Kuc, does not have the same density of leaves everywhere, it has smaller branches on the outside with scales increasing towards the inside. A model that assumes stationary echo statistics, however, cannot represent this potential source of information.

In a recent study, Yovel et al. (2008) showed that a machine-learning based classification algorithms could learn to classify plant echoes. The main advantages of the classifier suggested in this research (i.e. support vector machines) are that (a) it does not assume stationarity of the input and thus can find temporal cues that might be advantageous for classification (b) it does not assume fine temporal information (actually it requires very low temporal information), (c) it can extract both spectral and temporal cues simultaneously, (d) it provides a decision rule that can be interpreted in the context of the physical structure of the object. In addition, unlike many of the previous models, it did not use a predefined decision rule that relies on the extraction of certain parameters from the data (e.g. maximum amplitude, roughness, etc.). Instead, it used the data itself to derive a classification rule that matches a certain optimization criterion. In the first study, Yovel et al. (2008) showed that such an approach is successful in classifying complex echoes from the same data set used by Stilz (2004). Linear machine learning classifiers were able to classify all five plant species with high (>80%) precision. Using the decision rule that was learned by the classifier, they were able to detect simple spectro-temporal features in the echoes that are suitable for classification. Interestingly, the classifier used temporal cues for classifying the species, suggesting that the temporal information should not be neglected. Some of these features could be explained by characteristic differences in the physical structure of the plant species. The only conifers tested (spruce trees), for instance, could be easily classified according to their longer echoes that had less energy at low frequencies (Fig. 3a–c). Both the longer duration and the decreased energy in low frequencies are a result of their needle-like reflectors that have small cross-sections and thus reflect less energy at low frequencies and enable deeper penetration of an impinging signal compared to wide-leaved plants. Thus, this is the first model to suggest the importance of spectral cues reflecting the distributions of the cross-sections of the object's reflections. The main weakness of this model is that the decision rule might be hard to interpret, especially when non-linear classifiers are used.

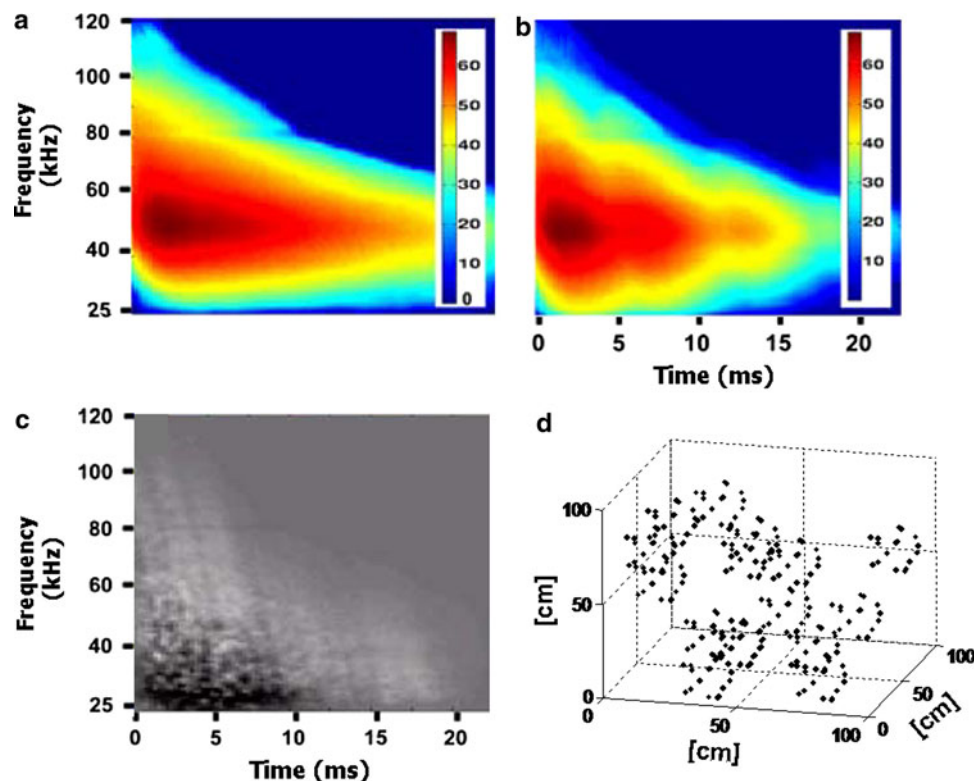


Fig. 3 The machine learning classifier uses a low spectral–temporal representation of the echoes. **a** Average spectrogram of spruce tree echoes. **b** Average spectrogram of echoes of all plants except spruce (apple, black thorn, beech, and cornfield). The *color bars* for both **a** and **b** are in dB. **c** The decision rule learned by the classifier for this classification task (decision echo). *Black* represents negative values, *white* represents positive ones, and *gray* is zero. The classifier was able to detect simple spectro-temporal features in the echoes that are

suitable for classification. In this case the classifier found that spruce echoes are characterized by longer echoes with less energy at low frequencies. [**a–c** were modified from Yovel et al. (2008) with open access permission from PLoS]. **d** An example of a 3D distribution of reflectors created by the model. Notice the clusters of points depicting two levels of branching. [Modified from Yovel et al. (2009) with open access permission from PLoS]

Dependency on angle of acquisition

Bats usually scan plants when on the fly and will thus almost never ensonify a plant from exactly the same position. When classifying complex echoes, bats must therefore develop some invariance to the angle of acquisition. This is a non-trivial task. Even in radially symmetric plants, the echo statistics will be kept from all directions, only when the plant is ensonified in the direction of its vertical symmetry axis (perpendicular to the plants vertical axis, Fig. 4a, b). As soon as the ensonification angle deviates from this direction, the ensonified cross-section of the plant is different from a radial cross-section and thus contains a completely different reflector arrangement.

Müller and Kuc (2000) already addressed the effect of bearing angle on echo statistics by ensonifying their plant from two different aspects (0° and 45°). They found that different plants behave differently in this aspect. The leaves of the fig they used were rather uniformly oriented

towards the outside, thus making strong coherent reflections more likely to be seen at normal sound incidence and resulting in large difference between bearing angles. Yew, however, did not show any directional structure although its statistics also changed with bearing angle, but to a much lesser extent (Fig. 4c).

Interestingly, however, despite the changes, some parameters (such as the impulsiveness, the dispersion and the crest factor) maintained the same difference tendency in both bearing angles, meaning that if the average impulsiveness of one plant species was larger for one angle it was also larger for the other angle but to a different extent. These results might suggest how bats maintain angle independency, but they also suggest that some parameters are beneficial for this while others are not.

The machine learning classifiers (Yovel et al. 2008) showed clear invariance to the viewing angle and were able to detect acoustical features in the echoes that are identical from all impinging aspects.

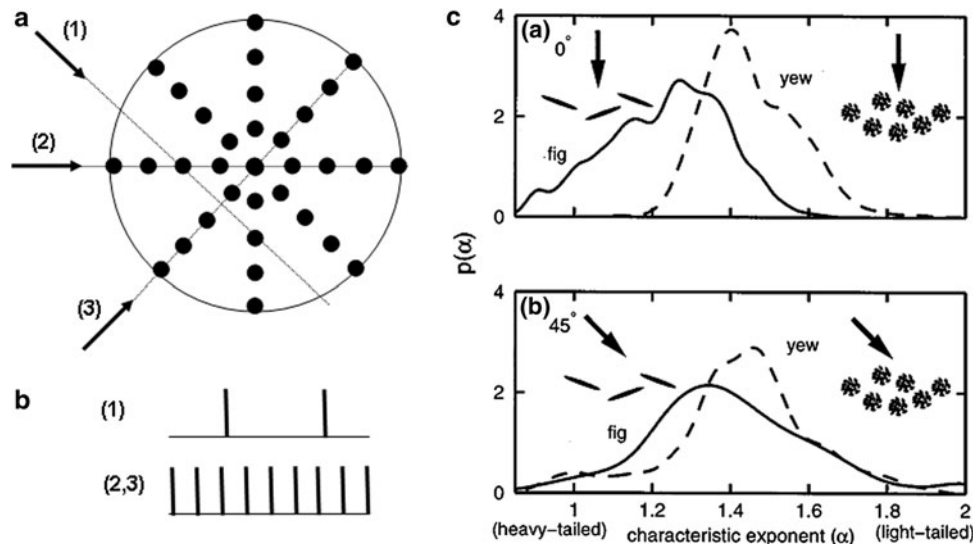


Fig. 4 Invariance to angle of acquisition. **a** Diagram of a circular symmetric plant, black dots represent leaves and arrows depict three ensonification bearing angles. **b** The IRs of the echoes expected to result from the corresponding ensonification in **a** assuming no occlusion and a very narrow beamwidth. The IRs are highly different despite the circular symmetry. **c** The effect of bearing angle on acoustical parameters, Modified from Müller and Kuc (2000). The two

graphs depict the distribution of an acoustical parameter (α , see in text) from two bearing angles (0° and 45°). Fig trees exhibit statistics that strongly depend on angle of acquisition because their reflectors (depicted schematically by black lines representing their cross-section) are arranged in a plane-like orientation depicted schematically by their cross-section. The echoes of Yew trees are not as angular dependent as expected from the omnidirectional orientation of reflectors

Connecting echo statistics with physical plant structure and the generation of natural complex echo statistics

Müller and Kuc (2000) already hypothesized how the differences in some of the echo parameters of the two plant species they examined could result from the differences in the plants' structure. They suggested for instance, that the mirror like planar fig leaves result in higher impulsiveness values than the small facets of the yew needles. They were also the first to argue in favor of echo depth as an acoustic cue that is advantageous for classification. They showed that echoes that originate from inside the foliage of both a conifer and a deciduous plant are visible to the bats, but argued that the underlying mechanisms were different; while in the conifer the small facets of the reflectors enabled signal penetration, in the fig, the sparser distribution of the large planar leaves allowed the sound to travel through the larger voids between the leaves. The authors suggested the correlation between time and echo amplitude as a relevant parameter to represent depth information. They found a logarithmic correlation of this type for yew representing a time-dependent decreasing sound channel gain. For fig, however, no such relationship could be demonstrated.

The importance of both echo depth and the amplitude decay rate over time was re-demonstrated in the case of the large natural echo data set (Yovel et al. 2008).

Grunwald et al. (2004) generated artificial echoes from IRs with different fourth moment values (see above). These

echoes can be thought of as representing different foliage statistics. Their main advantage is that they enable to control acoustical parameters and to vary a single cue (roughness) in which they were interested. This model, however, does not contain two types of information shown to be relevant for classification, they do not model spectral filtering resulting from the different cross-sections of the reflectors of different plants (Yovel et al. 2008 demonstrated how conifer species for instance have attenuated energy at low frequencies). Their model also did not capture any of the periodic structure widely observed in plants (Yovel et al. 2009). By using a randomly distributed IRs the authors assume that all distance between reflectors are equally probable, which is not the case for the fractal-like structure that is typical for most plants (i.e. several scales of branches).

Yovel et al. (2009) studied how the physical structure of complex plants contributes to the statistics of their echoes. They showed that the spectra of the echoes can be understood in terms of the sizes (cross-sections) of the reflectors (mainly the leaves) that the plants contain. More interestingly, however, they demonstrated that the distances between reflectors (leaves and branches), which are in many plants species-specific and could thus facilitate classification are encoded in the temporal statistics of the echoes. They suggested that the most informative distances for classification are the large ones because short distances are overrepresented in the echoes: a distance that is

ensonified from a non-perpendicular angle will show up as shorter than it really is in the echo. Large distances are translated into low frequencies (<1 kHz) in the envelope of the echo's IR. Interestingly, these temporal features are amplified by the auditory system filtering. To prove the relation between echo temporal statistics and plant structure, the authors were the first to manipulate real plants (by ripping of leaves).

The authors also attempted to generate artificial echoes that have plant-like echo statistics. Their model first creates a 3D distribution of leaves arranged in several levels, simulating the hierarchical structure that is typical for many plant species. It then sums the reflections returning from these leaves delayed in time according to their distance from the receiver. In contrast to the model suggested by Grunwald et al. (2004), Yovel et al. assumed a 3D object with reflectors that are not randomly distributed, but similar to Grunwald et al. they also did not include cross-section-dependent spectral responses or more complicated interactions between reflectors such as occlusions. Interestingly, however, already this simple model was able to replicate the echo statistics with high precision (Fig. 3d).

Scale invariance is another major problem in object classification. Humans are able to recognize a truck whether it is 5 m high and driving on the highway or 5 cm high and sitting still on a shelf in a toy store. Scale invariance in auditory classification is poorly studied. In humans, it was shown that the auditory system could compensate for the effect of speaker size on perceived speech (Smith et al. 2005). Firzlaff et al. (2007) studied scale invariance in complex echo classification by the fruit eating bats *Phyllostomus discolor*. They trained bats in a 2-AFC paradigm to discriminate between two specific objectless echoes generated from random IRs containing 12 reflectors. In the test phase the bats were confronted with echoes representing scaled (upwards or downwards) versions of the original ones.

A scaled variant of an object differs from the original in both its surface area and its depth. These two changes are translated into an increase or a decrease in the amplitude and a stretch or a compression of the time intervals between the reflectors. The scaled echoes were thus created according to these two principles. The bats' performance in the test phase was significantly above chance level suggesting that to some extent they can classify a complex object they were trained to discriminate independently of its size. The authors showed that a simple model that measures the Euclidean distance between the spectrograms of the scaled echoes and the original ones is not sufficient for classification. The authors also tested neural responses to the scaling of the echoes, the results of which we shall discuss in the next paragraph.

The neural base of classification in the bat brain

Very little is known on the brain processes mediating acoustic scene classification. Most attempts to explain acoustic classification by bats were based on previous findings on the general acoustic processing in the bat brain [see Ulanovsky and Moss (2008) for a review]. It is hard to support any of the above models since for each parameter that was used, corresponding neurons that are sensitive to this parameter have to be found. One of the first attempts to find the neural basis for an observed ability to classify complex objects found evidence for a neural representation of roughness in the inferior colliculus (midbrain) and the auditory cortex (Firzlaff et al. 2006; Borina et al. 2008). In these studies, the researchers found neurons that code roughness in these two brain areas, which might account for the bats' ability to classify echoes according to their roughness. In the first study, the researchers found a significant positive correlation between the activity of roughness sensitive neurons and the activity of these same neurons in response to stimuli with the same power spectra as for the echolocation calls that were sinusoidally amplitude-modulated. Namely, neurons that preferred increased roughness also preferred increased amplitude modulations. Modulation-sensitive units have already been shown in the bats' auditory midbrain (Grothe et al. 2001) and were suggested to be involved in roughness analysis for the processing of stochastic echoes (Grunwald et al. 2004). Interestingly, the authors also found that roughness-prefering units tend to exhibit higher temporal precision, firing at stimulus on-set. This finding could support the neural-inspired model of Müller (2003). In the second study, the authors even demonstrated a clear correlation between the bats behavioral performance when classifying roughness and the results of classification based on neural activity (Fig. 5). Despite this impressive evidence for the importance of roughness, it probably cannot explain the entire classification performance. Yovel et al. (2008) showed that some plant species could not be distinguished according to their roughness. It seems reasonable that a non-stationary measure of roughness (i.e. one that changes along the object's depth) will convey more species-specific information. Modulation-sensitive neurons (Grothe et al. 2001) could also facilitate the use of low frequency information in the echo envelopes for classification. Such information can reflect periodic physical structure of plants that would allow for distinguishing between them even though their echoes have the same roughness values (Yovel et al. 2009).

Another neural study aimed at explaining the neural base of the observed scale invariant classification behavior (Firzlaff et al. 2007). Here, the authors found auditory cortex neurons that are invariant to changes in scale-related features of the echoes representing the surface and the

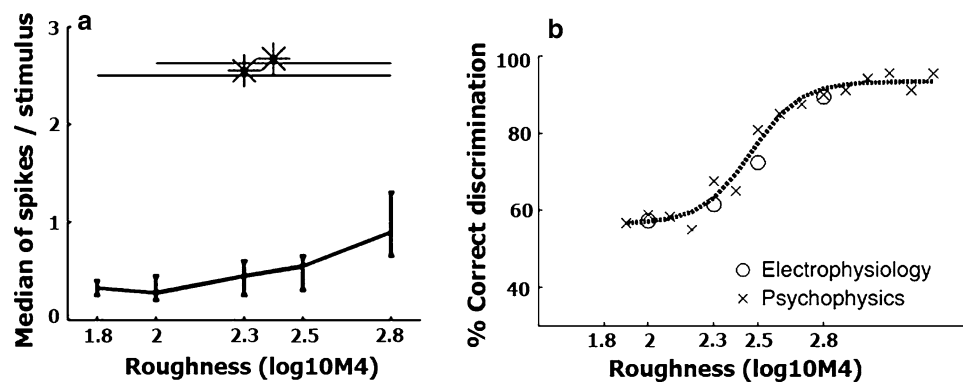


Fig. 5 The neural basis for complex echo classification [Modified from Firzlaff et al. (2006) with open access permission from PLoS]. **a** Neural response strength as a function of IR roughness for four roughness-sensitive units in the auditory cortex of *P. discolor*. **b** Comparison of the psychophysical and neurophysiological roughness discrimination performance. The crosses show the

psychophysical performance with a reference roughness of 1.8; the dotted line is a sigmoid fit to the psychophysical data. The open circles show the performance of an ideal observer (using a receiver-operating-characteristics analysis) basing its decisions on the pooled responses of the 14 roughness-sensitive cortical units found in the study

depth of an object. These neurons maintain a high firing rate independently of increasing or decreasing surface or depth. More interestingly, the authors found that 13% of the examined neurons show an object-size invariant response, thus exhibiting a high firing rate when both scale-related features vary correspondingly.

The machine learning based classifier did not rely on any “special” parameters that need to be measured from the echoes and can thus be explained based on our current knowledge of auditory neurons. The classifiers typically relied on low resolution spectro-temporal features for classification (Fig. 2). Neurons in the auditory cortex are tuned according to typical spectro-temporal receptive fields (STRF, Fritz et al. 2003) that resemble these features. An assembly of neurons with the STRFs that codes the decision rule of the machine learning classifier is theoretically sufficient to classify the echo. Moreover, it was shown that neurons can change their STRF in a time scale of seconds according to their input stimulus (Fritz et al. 2003), which could explain how the brain learns to classify new echoes and could facilitate fast learning of landmarks.

Conclusions and open issues

The most important thing that we have learned from research on complex echo classification by echo-locating bats is that complex echoes contain object specific information that can be to some extent used by echo-locating bats. The most common case study, i.e. the analysis of plant species classification revealed that both spectral and temporal cues that are sufficient for classification are available in the echoes. These are available in the high time resolution representations of the echoes (e.g. Grunwald et al. 2004; Wiegube 2008) as well as in the medium

(Müller and Kuc 2000) and low (Stilz 2004; Yovel et al. 2008) representations.

Distance invariance

In vision, the concepts of scale invariance and distance invariance are the same because up to the resolution range, a distant object is simply a smaller version of a closer one. In echolocation, however, this is not the case because the distances between parts of the object can be measured and these do not change when the bat is closer or farther from the object. Firzlaff et al. (2007) studied scale invariance in its visual context by adjusting the amplitude and distance between reflectors. In addition to the amplitude decay, the main influence of changing the distance from a complex object (such as a plant) is the widening of the beam impinging on it which results in turn in the inclusion of more reflection in the echo. This phenomenon clearly gives an advantage to low-resolution models that do not assume measurement of exact time differences between reflectors because these later ones tend to fail when many reflections return at temporal vicinity.

The parametric models suggested above (Müller and Kuc 2000; Stilz 2004) might also suffer from distance invariance problems, because the addition of more reflectors will change the temporal statistics of the echoes and thus change the parameters they extract from them (it will “smear” the echo).

Single versus multiple echoes

Many researchers assume that bats integrate information from several echoes for classifying an object, but this was never proved. When chasing an insect or approaching a fruit, a bat can collect many echoes from different angles that can

enable it to create a 3D like representation of the object. Useful echo integration of complex echoes, however, might be more difficult due to the behavioral context in which it is usually done. When using a tree as a landmark, for instance, a bat typically passes it quickly and might need to settle for only a few echoes from a random variety of aspect angles. The use of a statistical approach should increase the invariance to the angle of acquisition, and indeed, there are both behavioral as well as statistical evidence that a single complex echo is sufficient for classification (Grunwald et al. 2004; Firzlaff et al. 2006; Yovel et al. 2009).

The effects of the bat's motion

Most cases of complex echo classification occur when the bat is moving in relation to the object. The effects of this motion will depend on the relative radial velocity between the bat and the object as well as on the acoustical parameters of its emitted call (frequencies and frequency modulation). These effects were broadly discussed by Boonman et al. (2003) and generally include an increase in the uncertainty regarding the exact position of the reflectors and a shift in their position. Several authors suggested how bats could improve their signal design to compensate for this effect (Boonman et al. 2003; Holderied et al. 2006). Such compensation might function to a lesser extent for a complex echo because the velocity can differ in relation to different parts of the object (this can be significant when the bat is rather close to the object). In any case, such motion discrepancies could only affect the high-resolution models that rely on the exact distances between reflectors.

Binaural information

A major source of information, neglected by all models reported above is the use the binaural information, i.e. the comparison of the echoes received by the two ears. This aspect of echo classification was hardly ever studied (Holderied and von Helversen 2006). To date, the importance of binaural information for complex echo classification is not clear. In order to exploit the binaural disparities in an echo reflected from a simple object, very fine temporal resolution is required. In the case of large multiple-reflector objects, however, the binaural disparities are often quite significant because some reflectors are closer to one ear while others are closer to the other ear. This could supply additional important information for classification.

The behavioral relevance of complex echo classification

In this review, we restricted our discussion to cases in which a bat is flying at a distance of one to several meters

from the object. The echoes returning in such cases are thus a superposition of dozens of reflections and the bat does not scan the object deterministically from many angles. Behaviorally, such cases apply to any task that the bat has to fulfill that does not require an exact interaction with the object, i.e. landing on it or catching it.

All of the work done on complex echo classification to this date addressed the task of plant species classification (or simulated data representing it). Plant species classification might be advantageous for navigation, habitat selection, and food source finding, but there can be other ways in which using complex echo statistics are beneficial for the completion of a goal in the bat's daily life.

Classification of plant species is rather easy for many cases (Stilz 2004; Yovel et al. 2008) due to the major structural differences between species. In some behavioral contexts, a finer within-species classification might be necessary. Field observations suggest that some bats forage on grassland, but will only do so when the grass reaches a certain height. In other cases, the density of a certain bush might be relevant for the bat's decision whether to pursue prey within it. The changing statistics of a hedge might provide navigational information on when to stop following it in route-following behavior. In all of these cases, the bat will have to face the task of discriminating structural differences within a species, or independent of the species.

Recognition of individual objects (or arrangement of several individuals) might also play an important role when returning to the roost or when following a flight path very precisely (e.g. turning at an individual apple tree). The importance of visually memorizing exact complex arrangements of objects was shown to be important for homing in wasps for instance and similar acoustic behavior might be used by bats.

Usage of complex echo statistics could also have some importance when following a simple object (e.g. an insect) with complex echoes (e.g. plants) in the background. A good representation of the background statistics could perhaps assist to highlight the simple objects or to avoid background objects, but such interactions have never been studied or proved. All of these possibilities have not yet been explored and therefore provide much room for future behavioral and theoretical experiments.

Summary and future research

When classifying complex echoes, bats probably rely on some statistical analysis of their acoustic parameters. One of the prerequisite for a deeper understanding of complex echo classification is thus, a better knowledge of the statistics of natural complex echoes. In last years, several successful attempts to create models for complex echo classification showed that this task is possible and probably

easier than previously thought. More sophisticated models that use non-stationary statistics and match the real natural statistics (McDermott et al. 2009) should be developed.

More important, however, are behavioral experiments that test these models and compare them to the bats' behavior. Such studies are extremely sparse today and the few that did compare the behavior to a model (Grunwald et al. 2004; Firzlaff et al. 2007) did so using very simple Euclidean models that might not be sufficient to explain the behavior of the bats. Finally, studies that explain the neural basis of complex echo classification are clearly missing.

Acknowledgments We would like to thank R. Boonman for the fruitful discussion and our two anonymous reviewers for their comments. This work was funded by SFB 550, by the Graduiertenkolleg Neurobiologie, and by grant SCHN 138/27-1 of the German Science Foundation. The work was also supported by the University of Applied Sciences at Konstanz, Germany. This work was also supported by the human resources and mobility activity Marie Curie host fellowships for early stage research training under contract MEST-CT-2004-504321 PERACT by the European Union.

References

- Altes RA (1976) Sonar for generalized target description and its similarity to animal echolocation systems. *J Acoust Soc Am* 105:59–97
- Au WWL (1993) *The sonar of Dolphins*. Springer, New York
- Boonman A, Ostwald J (2007) A modeling approach to explain pulse design in bats. *Biol Cybern* 97:159–172
- Boonman AM, Parsons S, Jones G (2003) The influence of flight speed on the ranging performance of bats using frequency modulated echolocation pulses. *J Acoust Soc Am* 113:617–628
- Borina F, Firzlaff U, Schuller G, Wiegrebe L (2008) Representation of echo roughness and its relationship to amplitude-modulation processing in the bat auditory midbrain. *Eur J Neurosci* 27:2724–2732
- Bradbury JW (1970) Target discrimination by the echolocating Bat *Vampyrum Spectrum*. *J Exp Zool* 173:23–46
- Firzlaff U, Schörnich S, Hoffmann S, Schuller G, Wiegrebe L (2006) A neural correlate of stochastic echo imaging. *J Neurosci* 26:785–791
- Firzlaff U, Schuchmann M, Grunwald JE, Schuller G, Wiegrebe L (2007) Object-oriented echo perception and cortical representation in echolocating bats. *PLoS Biol* 5:e100
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223
- Griffin DR (1958) *Listening in the dark*. Yale University Press, New Haven
- Griffin DR (1967) Discriminative echolocation by bats in animal sonar systems biology and bionics. In: Busnel RG (ed) *Laboratoire de Physiologie Acoustique*. Jouyen-Josas, France, pp 273–300
- Griffin DR, Friend JH, Webster FA (1965) Target discrimination by the echolocation of bats. *J Exp Zool* 158:155–168
- Grothe B, Covey E, Casseday JH (2001) Medial superior olive of the big brown bat: neuronal responses to pure tones, amplitude modulations, and pulse trains. *J Neurophysiol* 86:2219–2230
- Grunwald JE, Schornich S, Wiegrebe L (2004) Classification of natural textures in echolocation. *Proc Natl Acad Sci USA* 101:5670–5674
- Habersetzer J, Vogler B (1983) Discrimination of surface-structured targets by the echolocating bat *Myotis myotis* during flight. *J Comp Physiol A* 152:275–282
- Holderied MW, von Helversen O (2006) 'Binaural echo disparity' as a potential indicator of object orientation and cue for object recognition in echolocating nectar-feeding bats. *J Exp Biol* 209:3457–3468
- Holderied MW, Jones G, von Helversen O (2006) Flight and echolocation behaviour of whiskered bats commuting along a hedgerow: range-dependent sonar signal design, Doppler tolerance and evidence for 'acoustic focussing'. *J Exp Biol* 209:1816–1826
- Kalko EK, Condon MA (1998) Olfaction and fruit display: how bats find fruit of flagelliferous cucurbits. *Funct Ecol* 12:364–372
- Kober R, Schnitzler H-U (1990) Information in sonar echoes of fluttering insects available for echolocating bats. *J Acoust Soc Am* 87:882–896
- Konstantinov AI, Akhmarova NI (1968) Discrimination of targets by echolocation in *Myotis oxygnathus*. *J Biol Sci Moscow Univ* 4:22–28
- Levin E, Yom-Tov Y, Barnea A (2009) Frequent summer nuptial flights of ants provide a primary food source for bats. *Naturwissenschaften* 96:477–483
- Matsuo I, Kunugiya K, Yano M (2004) An echolocation model for range discrimination of multiple closely spaced objects: transformation of spectrogram into the reflected intensity distribution. *J Acoust Soc Am* 115:920–928
- McDermott JH, Oxenham AJ, Simoncelli E (2009) Sound texture synthesis via filter statistics. Workshop on applications of signal processing to audio and acoustics. Mohonk, NY
- McKerrow P, Harper N (2001) Plant acoustic density profile model of CTFM. *Ultrasonic sensing IEEE sensors J*, pp 245–255
- Mogdans J, Schnitzler H-U, Ostwald J (1993) Discrimination of two-waveform echoes by the big brown bat, *Eptesicus fuscus*: behavioral experiments and receiver simulations. *J Comp Physiol A* 172:309–323
- Moss CF, Schnitzler H-U (1995) *Behavioral studies of auditory information processing*. Springer, Berlin
- Moss CF, Surlykke A (2001) Auditory scene analysis in echolocating in bats. *J Acoust Soc Am* 110:2207–2226
- Müller R (2003) A computational theory for the classification of natural biosonar targets based on a spike code. *Comput Neural Syst* 14:595–612
- Müller R, Kuc R (2000) Foliage echoes: a probe into the ecological acoustics of bat echolocation. *J Acoust Soc Am* 108:836–845
- Ostwald J, Schnitzler H-U, Schuller G (1998) *Target discrimination and target classification in echolocating bats*. Plenum Press, New York
- Palmer AR, Russell IJ (1986) Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res* 24:1–15
- Patterson RD (1994) The sound of a sinusoid: spectral models. *J Acoust Soc Am* 96:1409–1418
- Peremans H, Hallam J (1998) The spectrogram correlation and transformation receiver, revisited. *J Acoust Soc Am* 104:1101–1110
- Pye JD (1980) Echolocation signals and echoes in air. In: Busnel RG, Fish JF (eds) *Animal sonar systems*. Plenum Press, New York, pp 309–353
- Saillant PA, Simmons JA, Dear SP, McMullen TA (1993) A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: the spectrogram correlation and transformation receiver. *J Acoust Soc Am* 94:2691–2712
- Sanderson MI, Simmons JA (2000) Neural responses to overlapping FM sounds in the inferior colliculus of echolocating bats. *J Neurophysiol* 83(4):1840–1855

- Sanderson MI, Neretti N, Simmons JA (2003) Evaluation of an auditory model for echo delay accuracy in wideband biosonar. *J Acoust Soc Am* 114:1648–1659
- Schmidt S (1988) Evidence for a spectral basis of texture perception in bat sonar. *Nature* 331:617–619
- Schmidt S (1992) Perception of structured phantom targets in the echolocating bat, *Megaderma lyra*. *J Acoust Soc Am* 91:2203–2223
- Schmidt S, Hanke S, Pillat J (2000) The role of echolocation in the hunting of terrestrial prey—new evidence for an underestimated strategy in the gleaning bat *Megaderma lyra*. *J Comp Physiol A* 186:975–988
- Schnitzler H-U, Henson OW (1980) Performance of airborne animal sonar systems: I. Microchiroptera. Plenum Press, New York
- Schnitzler H-U, Menne D, Kober R, Heblich D (1983) The acoustical image of fluttering insects in echolocating bats. In: Huber F, Markl H (eds) *Neuroethology and behavioral physiology: roots and growing points*, pp 235–250
- Schnitzler H-U, Moss CF, Denzinger A (2003) From spatial orientation to food acquisition in echolocating bats. *Trends Ecol Evolut* 18:386–394
- Schörmich S, Wiegrebe L (2008) Phase sensitivity in bat sonar revisited. *J Comp Physiol A* 194:61–67
- Schuller G (1984) Natural ultrasonic echoes from wing beating insects are encoded by collicular neurons in the CF-FM bat, *Rhinolophus ferrumequinum*. *J Comp Physiol* 155:121–128
- Simmons JA, Chen L (1989) The acoustic basis for target discrimination by FM echolocating bats. *J Acoust Soc Am* 86:1333–1350
- Simmons JA, Vernon JA (1971) Echolocation: discrimination of targets by the bat, *Eptesicus fuscus*. *J Exp Zool* 176:315–328
- Simmons JA, Lavender WA, Lavender BA, Doroshov CA, Kiefer SW, Livingston R, Scallet AC, Crowley DE (1974) Target structure and echo spectral discrimination by echolocating bats. *Science* 186:1130–1132
- Simmons JA, Ferragamo M, Moss CF, Stevenson SB, Altes RA (1990a) Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: the shape of target images in echolocation. *J Comp Physiol A* 167:589–616
- Simmons JA, Moss CF, Ferragamo M (1990b) Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus*. *J Comp Physiol A* 166:449–470
- Simon R, Holderied MW, von Helversen O (2006) Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat. *J Exp Biol* 209:3599–3609
- Skolnik MI (2001) *Introduction to RADAR systems*. McGraw-Hill, New York
- Smith DR, Patterson RD, Turner R, Kawahara H, Irino T (2005) The processing and perception of size information in speech sounds. *J Acoust Soc Am* 117:305–318
- Stilz P (2004) *Akustische untersuchungen zur echoortung bei fledermauesen*. Tier Physiologie. University of Tuebingen. <http://www.biosonarlab.uni-tuebingen.de/public/diss-stilz/diss.pdf>
- Sum YW, Menne D (1988) Discrimination of fluttering targets by the FM-bat *Pipistrellus stenopterus*? *J Comp Physiol A* 163:349–354
- Thies W, Kalko EKV, Schnitzler H-U (1998) The roles of echolocation and olfaction in two neotropical fruit-eating bats, *Carollia perspicillata* and *C. castanea*, feeding on piper. *Behav Ecol Sociobiol* 42:397–409
- Ulanovsky N, Moss CF (2008) What the bat's voice tells the bat's brain. *Proc Natl Acad Sci USA* 105:8491–8498
- von Helversen D (2004) Object classification by echolocation in nectar feeding bats: size-independent generalization of shape. *J Comp Physiol A* 190:515–521
- von Helversen D, von Helversen O (2003) Object recognition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 189:327–336
- von Stebut B, Schmidt S (2001) Frequency discrimination threshold at search call frequencies in the echolocating bat *Eptesicus fuscus*. *J Comp Physiol A* 187:287–291
- Webster FA, Brazier OG (1965) *Experimental studies on target detection, evaluation and interception by echolocating bats*. Springfield, VA
- Weissenbacher P, Wiegrebe L (2003) Classification of virtual objects in the echolocating bat *Megaderma lyra*. *Behav Neurosci* 117:833–839
- Weissenbacher P, Wiegrebe L, Kössl M (2002) The effect of preceding sonar emission on temporal integration in the bat *Megaderma lyra*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 188(2):147–155
- Wiegrebe L (2008) An autocorrelation model of bat sonar. *Biol Cybern* 98:587–595
- Yovel Y, Franz MO, Stilz P, Schnitzler HU (2008) Plant classification from bat-like echolocation signals. *PLoS Comput Biol* 4:e1000032
- Yovel Y, Stilz P, Franz MO, Boonman A, Schnitzler H-U (2009) What a plant sounds like: the statistics of vegetation echoes as received by echolocating bats. *PLoS Comput Biol* 5:e1000429