



**HAL**  
open science

# A Review of Neuromorphic Sound Source Localization and Echolocation-Based Navigation Systems

Eugénie Dalmas, François Danneville, Fouzia Elbahhar, Michael Bocquet,  
Christophe Loyez

► **To cite this version:**

Eugénie Dalmas, François Danneville, Fouzia Elbahhar, Michael Bocquet, Christophe Loyez. A Review of Neuromorphic Sound Source Localization and Echolocation-Based Navigation Systems. *Electronics*, 2024, 13 (24), pp.4858. 10.3390/electronics13244858 . hal-04863043

**HAL Id: hal-04863043**

**<https://hal.science/hal-04863043v1>**

Submitted on 3 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Review

# A Review of Neuromorphic Sound Source Localization and Echolocation-Based Navigation Systems

Eugénie Dalmas <sup>1,\*</sup> , François Danneville <sup>1</sup> , Fouzia Elbahhar <sup>2</sup> , Michael Bocquet <sup>1</sup> and Christophe Loyez <sup>1</sup> 

<sup>1</sup> Institut d'Electronique de Microélectronique et de Nanotechnologies (IEMN), Université de Lille, Centre National de la Recherche Scientifique (CNRS), Université Polytechnique Hauts-de-France, UMR 8520—IEMN, F-59000 Lille, France; francois.danneville@univ-lille.fr (F.D.); michael.bocquet@uphf.fr (M.B.); christophe.loyez@univ-lille.fr (C.L.)

<sup>2</sup> Electronic Waves and Signals Laboratory for Transport (LEOST), Gustave Eiffel University, 59650 Villeneuve-d'Ascq, France; fouzia.boukour@univ-eiffel.fr

\* Correspondence: eugenie.dalmas@univ-lille.fr

**Abstract:** The development of positioning systems has been significantly advanced by a combination of technological innovations, such as improved sensors, signal processing, and computational power, alongside inspiration drawn from biological mechanisms. Although vision is the main means for positioning oneself—or elements relative to oneself—in the environment, other sensory mediums provide additional information, and may even take over when visibility is lacking, such as in the dark or in troubled waters. In particular, the auditory system in mammals greatly contributes to determining the location of sound sources, as well as navigating or identifying objects' texture and shape, when combined with echolocation behavior. Taking further inspiration from the neuronal processing in the brain, neuromorphic computing has been studied in the context of sound source localization and echolocation-based navigation, which aim at better understanding biological processes or reaching state-of-the-art performances in energy efficiency through the use of spike encoding. This paper sets out a review of these neuromorphic sound source localization, sonar- and radar-based navigation systems, from their earliest appearance to the latest published works. Current trends and possible future directions within this scope are discussed.



**Citation:** Dalmas, E.; Danneville, F.; Elbahhar, F.; Bocquet, M.; Loyez, C. A Review of Neuromorphic Sound Source Localization and Echolocation-Based Navigation Systems. *Electronics* **2024**, *13*, 4858. <https://doi.org/10.3390/electronics13244858>

Academic Editor: Toshifumi Moriyama

Received: 21 October 2024  
Revised: 29 November 2024  
Accepted: 4 December 2024  
Published: 10 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** sound source localization; echolocation; neuromorphic computing; spiking neural networks; bioinspiration

## 1. Introduction

From the earliest times, we have always taken inspiration from nature and living beings. The complexity of biological systems has never ceased to be a source of amazement and currently drives many research domains thanks to its efficiency in completing tasks in ideal or challenging conditions.

Bioinspiration can take place at different levels, such as the extraction of data issued from sensors, data encoding, and data processing, but also in a more general approach, in relation to behavioral patterns that an individual may express in a group or in interaction with its environment. Depending on the approach chosen for the creation of a system or subsystem, the design can be completely biomimetic and even bioplausible—as is mostly seen with work derived from neuroscience studies and in the context of biological systems assessments—or, on the contrary, lightly/partially inspired by a biological process or functional aspect. Although originating from an inspiration from biology, it sometimes becomes so negligible or diluted that it is no longer mentioned.

One particular branch of bioinspiration is neuromorphic engineering, which is inspired by the neuronal processing of information in the brain where the computing units of the devices are artificial neurons, connected by artificial synapses acting as memory units that weight the impact of a neuron's output to another neuron, and which use spikes as

a common encoding of all sensor modalities. All sentient beings share the use of action potentials, also called spikes, as a key mechanism for encoding and transmitting neuronal information. The nervous system employs a diverse array of methods to encode and process information, but spikes are crucial and ubiquitous. They are sparse and asynchronous events in continuous time, and hold the information temporally and spatially. Several coding schemes allow the information issued from sensors to be processed by networks of neurons and interpreted at higher levels of abstraction.

Neuromorphic technologies were initially explored in artificial vision and created great enthusiasm that spread to other applications. Spike encoding has recently been the focus of artificial intelligence research for tasks such as image classification [1], speech recognition [2,3], tactile sensing [4,5], or motor control [6], but also object localization [7] and autonomous navigation [8]. Vision is the main means for object detection, but also navigation and mapping, through contour, depth, and motion detection. Several recent surveys have already been published on neuromorphic systems or technologies in the context of positioning and navigation for mobile robots [9,10]; vision-based works, in particular, have been extensively reviewed [11].

Despite being the main contributor, vision is not the only bioinspired means that makes it possible to understand of an environment and the subjects it includes. Hearing also provides non-negligible cues for scene analysis. Noisy subjects are largely present in the environment, sometimes picking up on events that vision failed to identify. Furthermore, sound sources can be localized with the help of paired sound collectors like ears. Some animals even use echolocation with the ultrasonic sounds they emit for hunting or navigating in their environment. This behavior, naturally observed in wildlife, has been an inspiration for techniques employed by sonars and radars. However, the processing in modern sonar or radar-based systems differs from the physiological model of echolocation. It has not been long since neuromorphic technology begun to be used for the signal processing involved in localization.

Neuromorphic positioning and navigation systems relying on the auditory system have not yet been detailed in depth, also because the applications and reported performances are more limited than vision-based systems. In this context, this survey focuses on neuromorphic spiking systems for acoustic source localization and echolocation-based navigation. Its contribution within this scope, which no previous survey has provided, is threefold:

- It provides an exhaustive review of the literature, starting from the earliest works reported;
- Precision and energy efficiency performances are identified and summarized;
- The latest advances are investigated, and the relevance of bioinspired approaches are assessed.

The organization of this paper is as follows: firstly, neuromorphic technology and computing are introduced, focusing especially on the core use of this technology, namely spiking neural networks (SNNs). The fundamental mechanisms of auditory localization are then described, with reference to the mammalian system, and an overview of the underlying neuronal processing is provided. Moreover, artificial bioinspired and/or neuromorphic elements used for modeling of the auditory system are introduced. The neuromorphic sound source localization (SSL) systems are reviewed, categorized by their focus on interaural time and/or level differences that are key binaural cues for localization. Then, an explanation of the echolocation principle is provided, followed by a review of sonar and radar-based neuromorphic navigation systems along with a description of the standard sound and radio ranging technologies. Finally, the overall bioinspiration level, methods, and performances of the reviewed papers are discussed in order also to assess future directions.

## 2. Neuromorphic Technology

Building on principles of neural information processing, neuromorphic technology represents a paradigm shift from traditional computing architectures such as von Neumann

or Harvard [12]. Unlike these conventional systems, which separate memory and processing units and rely on sequential operations, neuromorphic systems integrate memory and asynchronous processing in a highly parallel, event-driven manner.

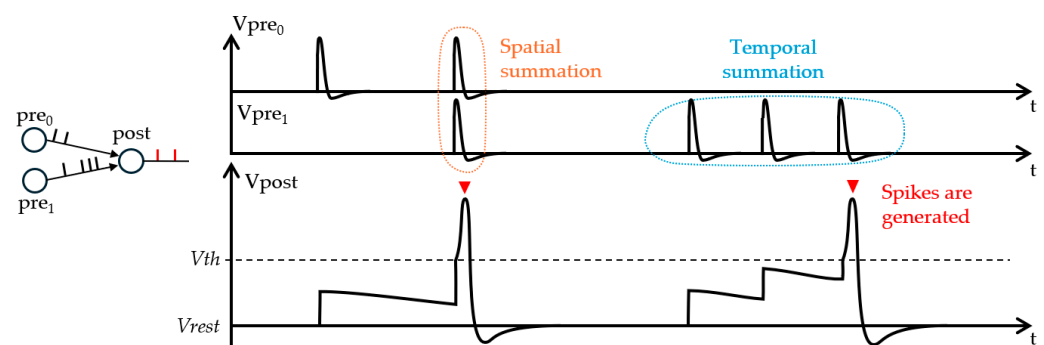
Given the growing interest in event-driven computing, large-scale neuromorphic architectures have been developed for the implementation of SNNs on programmable chips [13]. Commissioned by the Human Brain Project, BrainScaleS [14], and SpiNNaker [15] accelerators were designed for neuroscience simulations. Industries have also developed neuromorphic digital chips, including TrueNorth by IBM [16] and Loihi by Intel [17], with the aim of commercializing the technology. Moreover, ongoing research is being conducted to design analog synaptic and neuron arrays with memristive devices, such as memristor or resistive random-access memory (RRAM)-based arrays [18,19]. Novel photonics components have opened the path for the creation of neuromorphic photonic accelerators, which are still in the early stages of development.

Concurrently, software frameworks were developed that enabled the simulation of complex bioplausible SNNs or completed existing deep learning coding libraries to process spikes [13]. These tools greatly contributed to the creation of new learning rules or network architectures that are either inspired by biology or conventional deep learning.

### Spiking Neural Networks

SNNs revolve around the use of spikes to encode, process, and decode the information, initially introduced for the study of neural systems and the underlying learning mechanisms in the brain. As in biology, populations of neurons are connected by synapses to form neural networks. Unlike conventional artificial neural networks (ANNs) that only implement the concept of neuron weighted integration, SNNs fully reproduce the asynchronous event-based processing of biological neurons, more or less plausible from a biological standpoint depending on the complexity of the artificial neuron model.

Neurons generate a spike when their membrane voltage reaches a certain threshold by temporally and spatially integrating input currents. Inputs may be any stimulus, like the output of a sensor or spikes from other neurons. Figure 1 schematizes the integration of spike inputs into a neuron. The most common neuron model is the leaky integrate-and-fire (LIF) model, which implements the key dynamics of biological neurons with a simple mathematical formalism. Depending on the task and network architectures, different coding schemes have been studied to create suitable representations of the input data in spikes, categorized mainly in rate coding, which could be count or population rate coding, and temporal coding, like time-to-first-spike, burst, or phase coding [20].



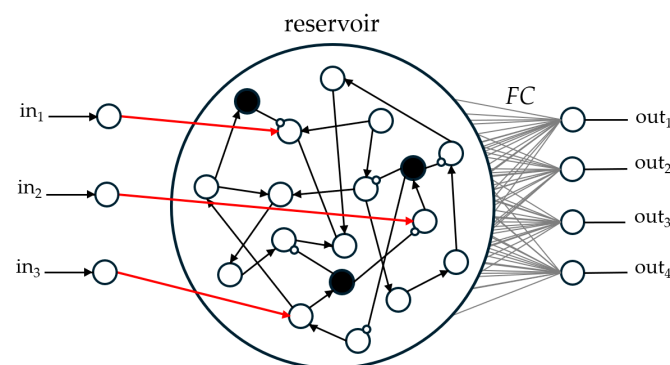
**Figure 1.** Spike generation at a post-synaptic neuron from two pre-synaptic neurons with leaky behavior, that is, a slow decrease over time of the membrane potential  $V_{pre_0}$ ,  $V_{pre_1}$ , or  $V_{post}$ , towards its rest potential  $V_{rest}$ . Upon reaching some threshold potential  $V_{th}$ , the membrane generates a spike resulting from the continuous spatial and temporal integration of pre-synaptic spikes.

In the context of robotics, mobile devices, or more generally of embedded electronics, computational load and resources are limited. Although every task has the objective of reaching the highest accuracies, today the emphasis is on energy efficiency and power

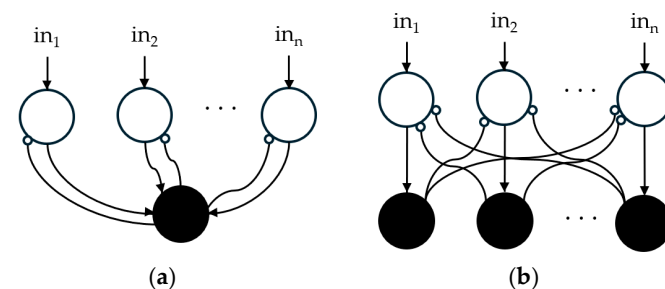
consumption. Extensive research on SSL and echo-based navigation enhanced by deep learning has been conducted, but without strong constraints on power consumption. Conversely, local memory units and the sparsity of event-based processing in spiking systems allow for great energy efficiency.

SNNs have yet to reach the high accuracies observed with ANNs, but the field is still growing. Research on learning schemes for the efficient training of SNNs has been conducted, at first using Hebb’s postulate on synaptic plasticity and based on causality of firing between two neurons, then later by adapting principles used in conventional ANNs [21,22]. The most commonly used schemes are spike timing-dependent plasticity-based learning schemes, a well-studied Hebbian unsupervised rule [23], and backpropagation of gradients by approximation of the spikes (undifferentiability of firing) for supervised training [24].

Spiking network architectures also evolved by taking direct inspiration from neurological observations, including several common deep learning topologies such as recurrent and convolutional networks that were initially bioinspired. Stronger plausibility is naturally observed in SNNs that use the asynchronous processing and temporality of spikes. Liquid-state machines (LSM) [25] or winner-take-all (WTA) layers [26] found in the literature, for example, were first studied from a biological perspective and then used in artificial intelligence. LSM, a method of reservoir computing shown in Figure 2, takes advantage of the excitation and inhibition dynamics of synapses by creating random connections in a population of neurons to produce complex spiking patterns, which can be further enhanced by adding delay-lines or connecting populations with different dynamics. Relying on lateral inhibition, WTA is also vastly used to limit spiking activity or for decision making by creating single neuron activations. Two possible architectures are depicted in Figure 3a,b, where inhibition is performed by one global neuron or with the same number of neurons as the layer receiving inhibition.



**Figure 2.** Architecture of LSM reservoir networks, with input and output layers of 3 and 4 neurons, respectively. Neurons chosen as inhibitory are shown as black disks. Excitatory and inhibitory connections are schematized as triangular and white-filled dot arrows, respectively.



**Figure 3.** Architecture of WTA layers: (a) using a global inhibitory neuron; or (b) using a layer of inhibitory neurons.

### 3. Biological Auditory System

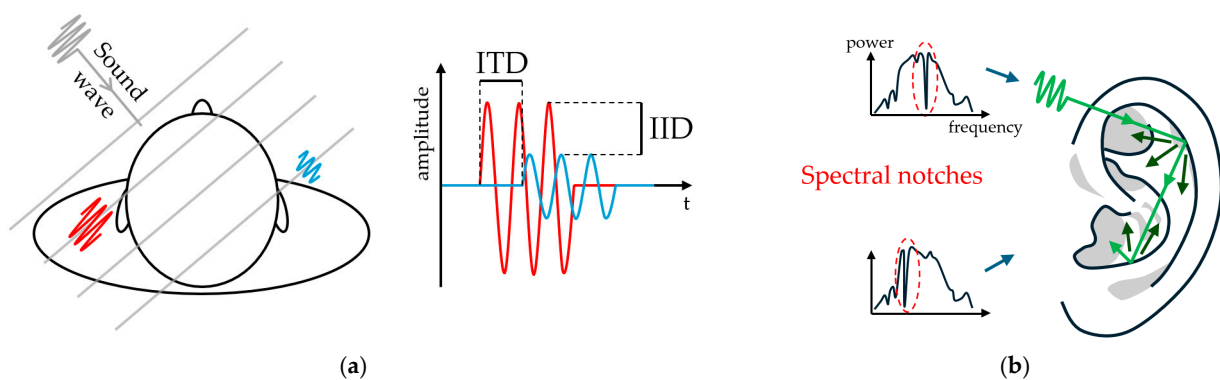
The perception of sounds begins with the intricate working of the ear. In the mammalian auditory system [27,28], the ear is divided into the outer ear (or pinna), the middle ear, and the inner ear. The pinnae and ear canal transform and direct the sound to the eardrum, which translates sound wave pressure into vibrations, while the middle and inner ear provide the necessary processing to send various types of information to the brain.

#### 3.1. Binaural and Monaural Cues

The bilateral arrangement of sound collectors is a common trait of vertebrates, with some variations involving the use of other vibration sensors in insects, arachnids, or marine animals. The use of paired sound collectors is essential for sound localization and general perception of a soundscape [29]. The difference in waveforms received at the two collectors of a binaural pair results in major interaural cues [30]. Among them, the most used are described below.

**Interaural Time Difference (ITD).** Having a binaural pair of sound collectors induces a difference in time of arrival when a sound is perceived, and greatly contributes to the localization of its source. It is maximal when the source is located at either side of the binaural pair, on the interaural axis (frontal plane), and null perpendicular to this axis (sagittal plane), so in front or back equidistant to the two collectors. The ITD perceived differs depending on the interaural distance, or baseline (distance between two sound collectors), which creates a maximum ITD. The smaller the baseline, the higher the frequency at which the ITD is non-ambiguous. In fact, when locating a source from a continuous sound, the ITD is determined by the interaural phase difference thanks to the phase-locked response of the auditory nerve fibers. The ITD can be identified at the onset of sounds knowing the first wavefront, but lateralization becomes ambiguous in continuous sounds when the period of the acoustic waves is smaller than the maximum ITD. In humans, this maximum is around 700  $\mu$ s, leading to a limitation of the frequency at which the ITD is resolved at 1.5 kHz to 1.6 kHz.

**Interaural Level Difference (ILD).** Jointly, the difference in sound magnitudes reliably informs on the sound source location above 3 kHz. At lower frequencies, ILDs are less impacted by head interferences or attenuation from propagation, and become barely noticeable. Magnitude differences can be expressed as a difference in decibel levels (that is, ILD), or directly as a difference in intensity, referred to as interaural intensity difference (IID), which are equivalent to a ratio and a subtraction of sound amplitudes, respectively. ILDs, or IIDs, are particularly reinforced by the pinnae's complex shape and the frequency segmentation by the cochlea, creating monaural spectral cues that play a key role in locating sounds in 3-dimensions. ITD and ILD cues are illustrated in Figure 4a.



**Figure 4.** Binaural sound localization cues. (a) ITD and IID cues are the difference in the time delay and amplitudes of received sounds shown in red and blue between the two ears, respectively. (b) Spectral notches are created by elevation-dependent interferences induced by the shape of the pinna, creating characteristic drops in received power in the spectrum, here circled in red.

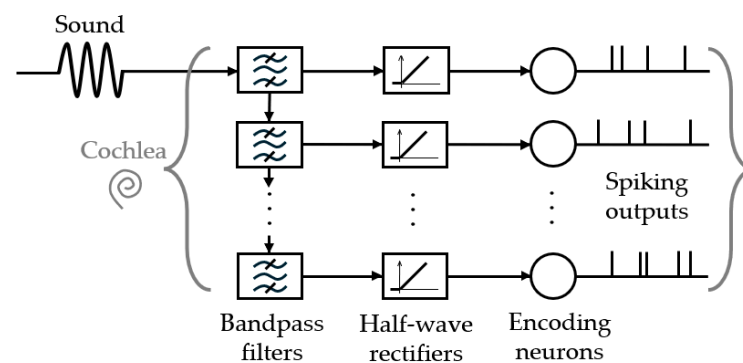
**Spectral notches.** The asymmetric shape of the pinna, relative to all anatomical axis, creates multiple reflections of the direct sound. They interfere constructively or destructively in the ear canal at specific frequencies, resulting in spectral signatures that are location-dependent, as shown in Figure 4b. Spectral notches are sharp drops in spectral gain due to destructive interferences. In particular, when observing the Head-Related Transfer Function (HRTF) of a human subject, a monotonical variation of the spectral content and notches between approximately 5 kHz and 9 kHz allows for the discrimination of sound elevation [31]. Spectral cues created by the pinnae have been artificially reproduced since they greatly contribute to estimating 3-dimensions positions. In applications for robots, it is not unusual to mount microphones in artificial human heads or with artificial pinnae. Not only the pinnae, but also the shape of the head and the torso, contribute to altering the sounds in a location-dependent manner.

### 3.2. Cochlea

Spectral notches can be exploited thanks to the cochlea, an organ part of the inner ear that plays a crucial role in enabling our ability to localize and interpret sounds in our environment by converting sound waves into nerve signals [28].

Composed of a canal in which pressure waves travel, it operates like a cascading filter bank with a band-pass response due to its spiral shape. The varying width and stiffness of the membrane within the canal causes its different regions to resonate at different frequencies, with a spatial arrangement following a tonotopic and logarithmic organization. Hair cells along the membrane move in synchrony with the traveling waves, creating a mechano-electrical transduction from movement to spikes by opening/closing ionic gates. In other words, the cochlea performs a high-resolution frequency segmentation where each spectral component is encoded into spikes for further processing by the brain.

Several cochlea models were described with mathematical formalism, such as Lyon's model [32] or Zilany's model [33], and neuromorphic artificial cochleae were introduced, such as AEREAR(2) [34,35] and NAS [36,37]. A reproduction of the processing in the cochlea is commonly performed following Lyon's model, which can be simplified as a bank of cascading band-pass filters (or parallel like in [38]), half-wave rectification, and suited spike encoding, as schematized in Figure 5.



**Figure 5.** Simplified processing performed by artificial cochleae. Input sounds are amplified to a suitable amplitude, filtered, and half-wave rectified to be encoded by neurons into spikes.

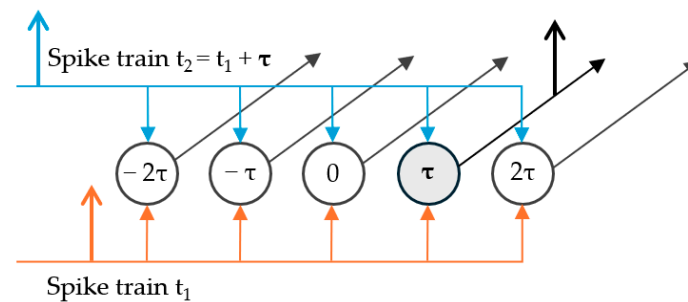
### 3.3. Neuronal Processing

While the cochlea provides the encoding into spikes of the sounds, several neuron populations are responsible for the extraction and integration of the interaural cues for final assessment of the source's position. Although still under investigation, a detailed organization of these populations in the mammalian auditory system has been refined over the years and many bioinspired works are based on this.

The main contributors for the extraction of sound localization cues are located in the superior olivary complex neuron population. It can be decomposed into the medial superior

olive (MSO) and the lateral superior olive (LSO), which both contribute to identifying the direction of arrival of sound sources at different frequency ranges [28].

The MSO extracts ITDs by means of delay-lines and coincidence detections. A simplified model of the MSO was introduced by Jeffress [39], which is represented in Figure 6. Delay-lines propagating spikes issued from the two ears are connected to a population of neurons, such that only one neuron activates for a given ITD. They are considered coincidence detectors because a spike is generated only when they receive simultaneous excitation from both ears. Each coincidence detection neuron is tuned for a specific ITD.



**Figure 6.** Jeffress model of delay-lines and coincidence detection neurons. Only the neuron with the corresponding time difference  $\tau$  generates a spike through temporal summation.

The LSO provides a spiking rate reflecting the IID or ILD by balancing excitation and inhibition between the two ears.

Finally, outputs from the superior olivary complex are integrated into a neuron population called inferior colliculus (IC) for sound localization with other neuronal information, like the response of neurons particularly sensitive to spectral notches or feedback from the auditory cortex.

#### 4. Neuromorphic Sound Source Localization

The localization of objects using the acoustic signals they emit is more commonly referred to as sound source localization. This task of pinpointing the origin of a sound within an environment plays a major role in positioning and navigation systems. Whether for autonomous robots, drones, or assistive technologies, the ability to accurately determine the location of a sound source can significantly enhance situational awareness and decision-making capabilities. Traditional methods for SSL often involve complex signal-processing techniques and extensive computational resources, which can pose challenges in real-time applications and dynamic settings. Neuromorphic computing, inspired by the brain's highly efficient and parallel processing abilities, presents a rather novel approach to address these challenges.

In this section, we review SSL neuromorphic spiking or pulsed systems from their earliest apparitions, in 1989, to the last published works, categorized by the main binaural cues used for the position or sound's direction of arrival estimation, namely ITD and/or ILD.

##### 4.1. ITD Only

Being a prominent cue in binaural hearing, ITD provides precise localization cue in the horizontal plane with low-frequency sounds, making it robust in various listening environments.

With a biomimetic approach, Lazzaro and Mead [40] were the first to build a neuromorphic pulse network in the context of SSL in 1989, although no correspondence between output ITDs and position was studied. The authors built a silicon model of the time-coding pathway of the owl based on observations in [41], which reproduce the Jeffress model. Cochleae of 62 channels implemented on integrated circuits were used [42], and channels were each linked to 170 coincidence detection neurons to constitute an auditory vector of ITDs with direct correspondence to an azimuthal angle. The ITD was then selected by an IC-inspired layer using inhibition and in a WTA fashion.



Then, Horiuchi [43] explored the use of the Jeffress model in a very-large-scale integration (VLSI) circuit in [40] with 15 coincidence neurons per channel and a cochlea processing based on threshold zero-crossing. The output pulses of the ITD vector extractor were combined with eye position units to create a retinotopic auditory output vector. The system was tested with finger snapping sounds at three different angles to show the linear relationship between the output and both azimuth and eye position. The motivation expressed was more about the application of neuromorphic technologies to robotics than the explanation of a biological process.

In 2000, Schauer et al. [38] then further studied the ITD extractor in [40], with slight modifications and added visual support in order to assess its practical application in robotics with speech processing. The model was simulated while considering the constraint of a VLSI design, and tested indoor with speech, pink noise, and moving sound sources. Unlike Lazzaro and Mead [40], who combined the IC layer and WTA inhibition, here a separation in different layers resulted in stronger WTA selection.

Following the description made in [44] of nuclei in the auditory system of rabbits, Voutsas and Adamy [45] designed a neuromorphic SSL system using ITDs. Inspired by the asymmetrical contributions of ipsilateral and contralateral sides, which are excitatory and inhibitory, to the MSO from the cochlea outputs, the authors designed a single delay line coincidence detection model called BiSoLaNN. This model was contrasted with the Jeffress model that uses two delay-lines to extract ITDs. The resulting neural network was tuned with an evolutionary algorithm, and tested with a Darmstadt dummy head in an anechoic chamber. Using frequencies below 1.24 kHz of speech, pure tones, and amplitude modulated sines, an accuracy of 59% ( $\pm 30^\circ$ ) was obtained on the  $\pm 105^\circ$  studied angular range. When reduced to the angular range  $\pm 45^\circ$ , the accuracy increased to 72.5% ( $\pm 15^\circ$ ).

Kugler et al. [46] completed the field-programmable gate array (FPGA) implementation introduced in [47], which performed simultaneous binaural SSL and classification of six sound sources with SNNs. A filter bank of band-pass filters was used for frequency segmentation, and a piecewise linear approximation implemented the hair cells processing and spike generation. The SSL module extracted the ITD according to the Jeffress model, where artificial neurons are integrate-and-fire (IF) neurons. The final direction of arrival was estimated by the average firing rate of output neurons. On  $\pm 90^\circ$  of angular range in a semi-anechoic chamber, average accuracies of 98.4% ( $\pm 30^\circ$ ) and 61.4% ( $\pm 30^\circ$ ) in azimuth were obtained for frequency-modulated noise and an alarm bell sound, respectively, recorded with a dummy head.

In 2010, Glackin et al. [48] presented an SNN of the MSO from observations in the cat auditory cortex, and studied it using ear canal recordings of a cat [49]. The bioplausible sound locator was proposed as a neuroscience study with supervised spike timing-dependent plasticity learning. This study followed Zilany's cochlea and Jeffress models for ITD extraction with IF neurons. In between is the particular use of a bushy cell neuron layer, which converted the phase-locked burst coding in output of the cochlea to single spikes with LIF neurons and refractory periods. The supervision was performed by excitation of the correct output neuron when the corresponding input was shown. The output layer is in fact the coincidence detector IF neurons for which the delayed excitations from the bushy cells are weighted with the bioinspired learning rule. Finally, the SNN was trained and evaluated with pure tones experimentally measured in a cat, where 91.82% ( $\pm 10^\circ$ ) and 78.64 ( $\pm 5^\circ$ ) azimuth estimation accuracies were obtained.

The same year, Chan et al. [50] also reported a neuromorphic ITD-based binaural SSL system for azimuth estimation mounted on a wheeled robot. The AEREAR cochlea processed incoming sounds recorded by microphones embedded in a spherical head. A WTA SNN using the Jeffress model was enhanced with an additional soft-WTA layer, weakening inhibition to obtain more than one winner, and was evaluated in simulation with an algorithm implementation. By supervising with a gradient descent weight update, root mean square errors of  $6.05^\circ$  and  $4.1^\circ$  for pure tones and noise, respectively, were

reached on the angular range of  $0^\circ$  to  $90^\circ$  and considering a theoretical angular resolution of  $3^\circ$  of the system.

Unlike previous work that used a Jeffress model, Finger and Liu [51] proposed a spike-based algorithm that computed a running histogram of weighted ITDs in output of the AEREAR2 cochlea. The weights corresponded to the interspike interval preceding two paired spikes in order to highlight onsets in continuous localization.

A modified version of the biomimetic SNN incorporating bushy cells in [48] was evaluated by Wall et al. [52] for application to mobile robotics. It was separated into two symmetric SNNs to process the two lateral spaces of a  $\pm 60^\circ$  angular range, trained with supervised spike timing-dependent plasticity. No cochlea was used; instead, sounds were encoded into spike using the Ben's spiker algorithm from dummy head recordings. The system with an angular resolution of  $20^\circ$  resulted in a  $3.8^\circ$  mean absolute error (MAE) for signals with frequencies around 400 Hz.

In 2012, Chan et al. [53] also further studied their SSL system [50] in a reverberant environment and with visual feedback from a transient vision sensor for supervised online learning. After playing pink noise, an azimuth estimation was made, to which the robot turned. Then, flashing light emitting diodes at the sound source gave the true position, and the computed localization error was used to update the WTA network's synaptic weights through gradient descent. The authors reported  $5^\circ$  for pink noise and  $4.4^\circ$  for white noise of root mean square error.

Focusing on hardware implementation, Park et al. [54] developed an ITD extractor in a VLSI system using AEREAR cochleae for spike encoding. Like most works, it used a Jeffress model of coincidence detection and delay-lines connected to a multiplexing neuron layer, and an IC layer. No SSL task was performed although the extractor was introduced in the context of low power-consuming hearing aids.

Until now, neuromorphic SSL systems were designed for binaural hearing and mostly using a biomimetic approach. Working on the assumption that with more sensors better accuracy can be reached, Faraji et al. [55] evaluated with 2 to 8 microphones a WTA network implemented on FPGA. Acoustic signals were encoded into spikes using a voltage comparator with adaptive threshold commanded by the spike count, as a way to limit the spiking activity. It uses the ITD of incoming sounds to estimate the azimuth of its source, positioned 5 m away on the range of  $\pm 90^\circ$ , at different signal-to-noise ratios (SNRs). For an SNR of 10 dB, the localization error went from  $3.49^\circ$  MAE to  $0.99^\circ$  MAE when using 2 and 4 microphones (1 and 6 ITDs). At 0.5 dB SNR, these values dropped to  $5.57^\circ$  MAE and  $1.18^\circ$  MAE.

Among neuromorphic SSL systems, Beck et al. [56] proposed an 8-microphone azimuth estimator taking interesting inspiration from sand scorpions' processing of vibrations. In this work, the biological neuronal processing was extended to acoustic signals. In the end, it still relies, like most works, on time difference but takes advantage of the numerous sensors mounted in a circular array. The system was tested with a continuous 1 kHz sine wave, which resulted in a  $4.05^\circ$  MAE with  $3.01^\circ$  of standard deviation with optimized parameters.

In 2018, Encke and Hemmert [57] reproduced in simulation the auditory system of gerbils for a detailed study of the influences of neuron populations and ITD sensitivity to just noticeable differences in timing across the spectrum (below 1 kHz). The SNN was able to extract small ITDs from speech by linear decoding from firing rates. A comparison of neuron populations' firing rate, called opponent-coding mechanism, was investigated using also an ANN that predicted the ITD.

Without using delay-lines, Luke and McAlpine [58] performed a lateralization task in noisy conditions by directly feeding cochleae outputs from a binaural pair of microphones to an SNN. The network was trained with supervision on frequency-modulated band-pass noise. When tested with added white noise, the lateralization accuracy dropped from 100% to around 87% at 0 dB SNR. Modifying the frequency of the modulated noise in input resulted in accuracy of over 95% at a high SNR but dropped below 60% in high noise at 0 dB SNR.

Schoepe et al. [59] proposed a sound tracking system able to avoid obstacles thanks to visual feedback. The NAS cochlea and AER DVS128 retina chip were used for spike encoding and SpiNNaker for neuromorphic processing. In the overall SNN, a Jeffress-like coincidence detection layer was connected to a WTA subnetwork for ITD-based azimuth estimation. An integration subnetwork then combined auditory inputs with optical information. Spike events from the optical flow provided information about close range obstacles, whereas the SSL subnetwork led the robot to the sound source. The path was then chosen through concurrent excitation and inhibition in a WTA fashion, meaning only one direction could be considered. A simulated 500 Hz sound was swept on the  $\pm 90^\circ$  angular range 2 m away. Performances were reported as a correlation of 89% between the resulting head direction and truth source position.

In 2020, a photonic SNN for azimuth detection was implemented by Pan et al. [60] for the first time, demonstrating feasibility. It consisted of photonic neurons based on excitable vertical-cavity surface-emitting lasers with embedded saturable absorbers that showed similar dynamics to LIF neurons. Using 4 neurons and by analyzing the 2-D map of spiking responses in output, the side from which sounds originated could be identified.

Excellent localization accuracies were reported in Pan et al. [61], who introduced the multi-tone phase coding encoder, inspired by the Jeffress model of coincidence detection for ITD extraction. Two SNNs which were not biomimetic and closer to conventional deep learning architectures, a recurrent SNN, and a convolutional SNN were investigated for processing the temporal cues. The encoder reported better absolute ITD estimation error than naive phase coding methods. An array of 4 microphones allowed both the convolution and the recurrent SNNs to estimate the azimuth of sound sources 1.5 m away with an MAE below  $2^\circ$  on the whole  $360^\circ$  range. Training was supervised with continuous surrogate gradients and backpropagation-through-time for the convolutional and recurrent SNNs, respectively. The recurrent network performed best in all scenarios and in challenging conditions (real-world data). The best MAE reported was  $1.02^\circ$ , which lowered to  $1.76^\circ$  MAE at 15 dB SNR and  $10.75^\circ$  MAE at 0 dB SNR.

In the context of bioplausible SSL systems, Zhong et al. [62] focused on the development of memristor-based oscillation neurons whose spiking activity depended on the ITD of input sounds. A simulated SNN processed the rate-based output of the oscillation neurons for azimuth estimation, and was trained with supervision and backpropagation-through-time. The system obtained 96% ( $\pm 10^\circ$ ) accuracy using generated pulse cycles.

In 2023, Chen et al. [63] demonstrated the advantages of hybrid coding schemes throughout the layers of SNNs for improvements in accuracy. Different combinations of direct, rate, phase, burst, and time-to-first-spike coding were studied in convolutional SNNs for pattern recognition and SSL, on which we will focus. By combining the encoder in [61] at the input layer, burst coding in the hidden layer after ITD extraction, and time-to-first-spike coding in the output layer for efficient decision making, this study reported greater accuracies than [61], which introduced the encoder, and which previously held state-of-the-art SSL precision. The authors reproduced the recurrent SNN from [61], and obtained a  $1.48^\circ$  MAE or 94.3% ( $\pm 5^\circ$ ), tested on the SLoClas dataset. The convolutional SNN with hybrid coding, on the other hand, resulted in a  $0.60^\circ$  MAE or 95.61% ( $\pm 5^\circ$ ) accuracy.

In 2024, Schoepe et al. [37] presented a full hardware SSL neuromorphic system on FPGA, relying on a robotic head-tilting movement for sound source tracking. This study used their previous adaptation of time delay extraction units introduced in [64], inspired by the motion detection in vision, to SSL [65]. Artificial pinnae and NAS cochleae were mounted on a pan-tilt unit, and ITDs were extracted by time delay extraction units. No SNN processed the ITDs; instead, a motor was directly controlled by the extractor spike train outputs to orientate the robotic head toward the source. When evaluating the azimuth estimation after several seconds, MAEs of  $1.92^\circ$  and  $5.5^\circ$  with pure tones and speech were obtained.

Most recently, Dalmás et al. [66] also took inspiration from the time extraction performed in vision, but unlike in [37] they modified the model to give a linear output. From

short impulsive natural sounds filtered only to remove low ambient noise, the authors investigated the potential of their ITD extractor using 3 microphones. Numerical time differences were deduced from the spiking output to estimate spatial coordinates using hyperbolic equations simplified by the microphone array’s configuration. The precision in azimuth reported was 71.3% and 82.8% accuracy on average with a tolerance of 1° and 5°, respectively, for a source at distances between 1 m and 3 m. The proposed system was also studied with sources at a distance of 10 m and in noisy conditions, which resulted in reduced performances.

Table 1 resumes the reviewed works with ITD-only-based SSL neuromorphic systems. The numbers of neurons reported in this table—and the following Tables 2 and 3—are retrieved from the literature or computed from SNN dimensions. Multi-source scenarios are mentioned in several studies, for two and three simultaneous sources. The reported performances in these tables do not take into account these situations. Unless specified, the reported average performance corresponds to the angle precision.

Around 2015, the focus of the research shifted from being centered around biomimetic systems, that would bear resemblance to the first work in this field [40], to investigating lighter bioinspiration with the aim of improving localization accuracy. More than two microphones [55,56,61,63,66] were used, architectures closer to ANN were studied [61,63], and inspiration for the extraction of time differences was taken from vision [37,66]. Satisfying results can be observed, especially in [63], which suggests that adapting the encoding for spike generation enhances the processing and brings out better performances. No attempt was made to estimate the distance of the sound source; however, the latest reviewed work [66] demonstrated the poor estimation precision in distance of an ITD-only SSL system, attributed to the geometric dilution of hyperbolas.

**Table 1.** ITD-only neuromorphic SSL systems.

Reference	Year	# Mics	# Neurons	Learning	SNN Implementation	Angular Resolution	Test Environment	Sound Source	Average Performances
[40]	1989	2	10,540	-	VLSI	-	-	-	-
[42] <sup>1</sup>	1994	2	-	-	VLSI	-	Quiet	Broadband sounds	-
[39] <sup>1</sup>	2000	2	2308	-	Simulation	~2.8°	Quiet Reverberant	Speech, Pink noise	-
[45]	2007	2	25,856	Algorithm	Simulation	30°	Quiet	Pure tones	59% ACC
						90°		SAM signals	90% ACC
[46]	2008	2	210	-	FPGA	30°	Quiet	FM noise	98.5% ACC
								Alarm bell	61.4% ACC
[48]	2010	2	1029	STDP (S)	FPGA	10°	Quiet	Pure tones	91.82% ACC
						5°			78.64% ACC
[50]	2010	2	-	Gradient descent (S)	Simulation	3°	Quiet	Pure tones	6.05° RMSE
								Noise	4.1° RMSE
[51]	2011	2	1024	-	Simulation	~2.6°	Noisy Reverberant	Speech	0.17° MAE
[52]	2011	2	10	STDP (S)	Simulation	20°	Quiet	Pure tones	3.4° MAE
[53] <sup>1</sup>	2012	2	~3000	Gradient descent (S)	Simulation	3°	Reverberant	Pink noise	5° RMSE
								White noise	4.4° RMSE
[54]	2013	2	51,752	-	VLSI	0.9°	Quiet	Narrow band sound pulses	-
[55]	2015	2	>890	State machine	FPGA	0.32°	Quiet	Speech	3.49° MAE
							Noisy		5.57° MAE
		4					Quiet		0.99° MAE
							Noisy		1.18° MAE

Table 1. Cont.

Reference	Year	# Mics	# Neurons	Learning	SNN Implementation	Angular Resolution	Test Environment	Sound Source	Average Performances
[56] <sup>2</sup>	2016	8	>10	-	Simulation	-	Quiet	Pure tones	4.05° MAE ±3.01° SD
[57]	2018	2	-	-	Simulation	-	-	Speech	-
[58]	2019	2	855	Surrogate gradients (S)	Simulation	90°	Noisy	FM noise	87% ACC
[59] <sup>1,2</sup>	2019	2	1122	-	FPGA	-	Quiet	FM sound	89% correlation
[60]	2020	2	4	-	Photonic	-	Quiet	Pulses	-
[61]	2021	4	1981	Surrogate gradients (S)	Simulation	1°	Quiet	Speech, Noise	1.02° MAE
							Low noise		1.07° MAE
							High noise		10.75° MAE
[62]	2022	2	>36	Backpropagation (S)	Simulation	15°	Quiet	Pulses	96% ACC
[63]	2023	4	4261	Backpropagation (S)	Simulation	2.5°	Noisy	Speech, Broadband, sounds	0.6° MAE 95.61% ACC
[37] <sup>1,2</sup>	2023	2	326	-	FPGA	5°	Quiet	Pure tones,	1.92° MAE
							Speech		5.5° MAE
[66]	2024	3	6	-	Simulation	1°	Quiet	Natural sounds	71.3% ACC
						5°			82.8% ACC

<sup>1</sup> Use vision. <sup>2</sup> Use body movement. ACC—Accuracy; MAE—Mean Absolute Error; RMSE—Root Mean Square Error; SD—Standard Deviation; S—Supervised; FM—Frequency Modulated; STDP—Spike Timing-Dependent Plasticity.

#### 4.2. ILD Only

Similarly to studies conducted on ITD alone, ILD or IID have been the focus of several works, complementing ITD for comprehensive localization across the frequency spectrum.

A bioplausible SNN emulation of the LSO for IID-only sound source localization was studied by Wall et al. [67]. The SNN is composed of an inhibitory neuron population node, an LSO layer, and a receptive field layer fully connected with the output neurons. In the same manner as the previous work [52], two symmetric SNNs individually process the two hemispheres of the range of  $\pm 60^\circ$  with  $10^\circ$  angular resolution. The network is trained with supervised learning ReSuMe (Remote Supervised Method) and tested on HRTF data. An average accuracy of 91.39% ( $\pm 10^\circ$ ), or 62.62% generalized accuracy, is reported on sounds with distinct frequencies between 4.8 kHz and 25.2 kHz.

A software implementation of the mammalian auditory pathways was presented by Feng and Dou [68] using only IF neurons for processing ILD cues. Here, a one-shot learning algorithm set the synaptic weights, and the output neurons in the IC layer with the highest spike rate correspond to the estimated azimuth with  $10^\circ$  angular resolution. An accuracy of 90.56% ( $\pm 10^\circ$ ), or 74.56% generalized, was obtained for white noise inputs with frequencies between 1.22 kHz and 20 kHz. This work was completed to process both ILD and ITD in [69], with the claim that this was more faithful to the physiology organization than previous works.

Escudero et al. [70] implemented on FPGA a rate model of the LSO for IID-based SSL with the NAS cochlea. The azimuth was estimated as a direct correspondence to the output spiking rate, and with good robustness to noise. A “spike hold and fire” block is specifically designed here to reproduce the extraction of the IID in the LSO with spike rates. Evaluated with three pure tones between 1 kHz and 5 kHz, the MAE obtained was  $3.41^\circ$  at 28 dB SNR and  $5.63^\circ$  at  $-5$  dB SNR. A motor was controlled by the output firing rates to perform a tracking of sound sources with head rotation.

A neuromorphic implementation of the LSO on TrueNorth was reported by Oess et al. [71]. Binaural inputs, recorded with a dummy head and pinnae in a sound-attenuated room,

were converted to spectrograms (algorithm) to encode temporal and spectral intensities into spike trains. The resulting spike rates were normalized and processed by an SNN. An additional 19 readout neurons, external to the FPGA implementation, were tuned to a certain ILD value by receiving a weighted sum of the SNN's last layer, and thus led to a resolution of  $10^\circ$  on the  $\pm 90^\circ$  angular range. A mean accuracy of 62.5% ( $\pm 10^\circ$ ) was reported for synthesized and natural sounds among which noisy recordings were present.

With the increase in digital neuromorphic accelerators, Schmid et al. [72] conducted investigations on a generic procedure to map a rate-based SNN to different neuromorphic FPGA hardware. A binaural SSL model inspired by the LSO was taken as a reference, using only ILDs for 2-D localization, and originating from [71]. A three-step mapping was proposed, focusing first on translating connections and neurons with regards to a generic neuron model, then mapping them onto the hardware-specific simplified connection patterns and spike-based neurons, to finally implement the SNN using the hardware platform's framework, and fine-tuning parameters. While the TrueNorth implementation was thoroughly tuned for performance, involving additional pre- and post-processing steps, on the SpiNNaker only the LSO mechanism was studied for evaluation of the generic procedure. In the end, a similar qualitative behavior in both implementations was observed, with reports of  $18^\circ$  and  $29^\circ$  MAE for TrueNorth and SpiNNaker, respectively, in real-world data.

**Table 2.** ILD-only neuromorphic SSL systems.

Reference	Year	# Mics	# Neurons	Learning	SNN Implementation	Angular Resolution	Test Environment	Sound Sources	Average Performances
[67]	2012	2	-	ReSuMe (S)	Simulation	$10^\circ$	Quiet	Pure tones	91.39% ACC
[68]	2016	2	12,500	One-shot (S)	Simulation	$10^\circ$	Quiet	Pure tones, White noise	90.56% ACC
[70]	2018	2	>128	-	FPGA	$15^\circ$	Quiet Noisy	Pure tones	3.41° MAE 5.63° MAE
[71]	2020	2	1043	-	FPGA	$10^\circ$	Noisy	Natural sounds	62.5% ACC
[72]	2023	2	1024	-	FPGA (TrueNorth) FPGA (SpiNNaker)	$10^\circ$	Quiet	Natural sounds	18° MAE 29° MAE

ACC—Accuracy; MAE—Mean Absolute Error; S—Supervised.

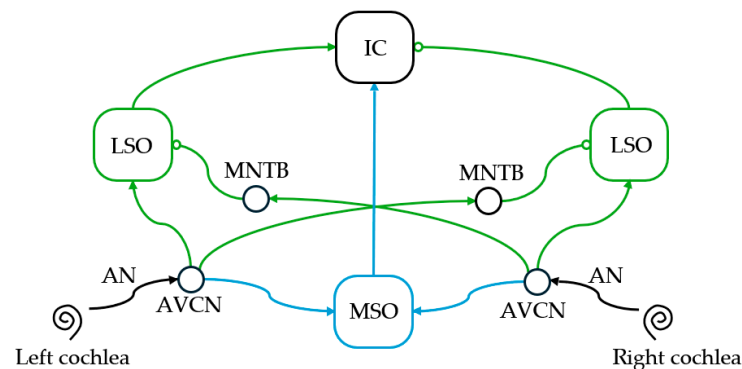
Table 2 resumes the reviewed works with ILD-only-based SSL neuromorphic systems. Relying only on ILD (or IID) does not lead to results as encouraging as systems using ITD. The method for extracting intensity or level differences was biomimetic in all the works, based on the LSO. Some variations in the tuning, implementation, or test sounds resulted in different performances that in fact lead to the conclusion that the ILD cue is not sufficient to estimate the azimuth of a sound source.

#### 4.3. ITD and ILD

Finally, both ITD and ILD were used to fully take advantage of the localization information carried in incoming sound waves.

Unlike previously mentioned works, Liu et al. [35] used both ITD and ILD to estimate the azimuth of white noise and speech in a biomimetic approach. The MSO and LSO were modeled in layers with different delayed connections to the outputs of a gammatone filter bank, and both finally integrated into an IC output layer, as shown in Figure 7. All the layers were matrices encoding in all frequency channels the ITD, ILD, and azimuth estimation, using corresponding neurons. No competition was added to the SNN, such that the authors also tested the SSL system with two simultaneous speakers. The frequency segmentation allowed the discrimination of multiple sources that did not share the same

spectrum. In the end, the SSL system resulted in accuracies of 80% ( $\pm 30^\circ$ ) and 90% ( $\pm 10^\circ$ ) on the angular ranges  $\pm 90^\circ$  and  $\pm 45^\circ$ , respectively, with the use of an artificial head.

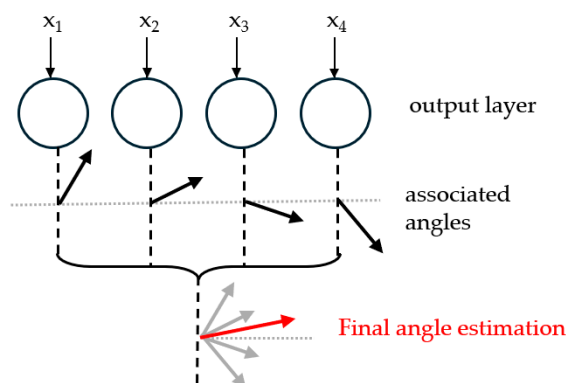


**Figure 7.** Diagram of the mammalian auditory pathway with ITD- and ILD-specific pathways highlighted in blue and green, respectively. Populations of neurons are represented by circles or rounded-edge rectangles. AN—auditory nerve; AVCN—anteroventral cochlear nucleus; MNTB—medial nucleus of the trapezoid body.

The binaural spiking SSL system presented by Goodman and Brette [73] also used ITD and ILD cues with HRTF data and cochlea filtering. The authors of this work used the synchrony patterns induced by location-dependent filtering for angle estimation. They also introduced a filtering of cochlea spikes by successive feed-forward neuron layers, before identifying the azimuth and elevation from overlapping neuron assemblies of coincidence detection. An MAE of  $2^\circ$  to  $7^\circ$  in azimuth and of  $7^\circ$  to  $20^\circ$  in elevation were reported for broadband white noise, speech, and musical instrument sounds.

Dávila-Chacón et al. [74] further studied the biomimetic SNN from [35] with a Nao robot. Test sounds were decomposed in 16 frequency channels by a Patterson–Holdsworth filter bank between 200 Hz and 4 kHz. Performances resulted in  $2.5^\circ$  MAE and  $27^\circ$  MAE for white noise and speech, respectively, on the whole angular range of  $\pm 90^\circ$ , but the average errors dropped to  $0^\circ$  for a  $\pm 30^\circ$  angular range.

Gao et al. [75] implemented in a memristor array an analog SNN for binaural SSL processing both ITD and ILD, and supporting in-situ training. Pulses were sent in the memristor array to modify the weights according to a multi-threshold update scheme where the number of applied pulses depends on several threshold values on the weight change. Each neuron output corresponded to a vector, such that the final angle estimated was the combination of all, as depicted in Figure 8. The authors reported a root mean square error of  $5.7^\circ$  for a two-layer memristor SNN in simulation, and  $12.5^\circ$  experimentally with a single layer.



**Figure 8.** Final output direction vector from summation of individual output vectors.

Focused on neuromorphic computation, Xu et al. [76] studied their patterning process for organic electrochemical synaptic transistor array fabrication with an SSL function simulated in a cross-grid array. Synaptic transistors of these arrays show long-term memory effects thanks to conductance modulation when suppression or excitation pulses are applied. The array is trained with supervision in a backpropagation manner using a correction voltage determined by an algorithm. Binaural signals from HRTF recordings were preprocessed by Fourier transform to generate 60 characteristic values, then fed to the feed-forward fully connected neural network. The output layer corresponded to the  $\pm 120^\circ$  angular range with steps of  $40^\circ$  where the final azimuth estimation was a weighted combination of all vectors, similarly to [75]. By the 100th epoch, the mean square error was lower than 0.11 normalized.

In 2023, Li et al. [77] encoded binaural sounds into spikes with a Zilany model cochlea, then processed by a LSM of 800 LIF neurons. The readout layer was connected to an ANN classifier for azimuth estimation. The overall network relied on both ITD and ILD by directly providing the filtered sound to the LSM, and hyperparameters were chosen through Bayesian optimization. Moreover, the authors used soft labeling in the training of the classifier to reach accuracy of 94.29% ( $\pm 15^\circ$ ) and 86.33% ( $\pm 10^\circ$ ) with HRTF recordings.

Another study that used LSM was presented by Roozbehi et al. [78], who introduced a dynamic rescaling of the reservoir's size. Based on the small-world connection technique, the generation of new neurons allows the network to increase its computation capacity. Neurons of the LSM had spatial coordinates that represented the real space in which a sound is recorded. The location of the source was estimated by the coordinates of the neuron with the highest membrane potential. The network then grew depending on the localization error to give a more precise estimation. In the test scenario, two microphones were placed so the clapping or speech sound source moved at a distance lower than the baseline. The system achieved 69.8% accuracy in locating the source or  $3.4^\circ$  MAE in azimuth and 0.38 m MAE in distance.

Table 3 resumes the reviewed works with ITD/ILD-based SSL neuromorphic systems. Little research was performed using both ITD and ILD cues. Starting with similar biomimetic systems around 2010, new approaches were then explored from 2022 onwards in order to successfully integrate both cues. Emerging technologies are also being investigated [75,76], more in phase with the context of low-power consumption.

We have yet to see a full neuromorphic system using both cues with localization performances surpassing the best precision obtained with ITD only [63]. With added information, the task should be easier, unlike what is now observed. Localizing with both cues would likely benefit from propositions of novel ILD extractors or enhanced processing of the standard LSO's output.

The question of distance estimation is not studied in any works (including the previous Section 4.2), apart from [66] (reviewed in Section 4.1), which reported incorrect distances from ITD only, and the study by Roozbehi et al. [78], which gave spatial coordinates but constrained to within the baseline of the microphone array. Apart from [78], sound sources are always placed at a fixed distance, although ITDs vary with the distance to the microphones, which casts doubts on the interpretation of results of the commonly used LSO.

**Table 3.** ITD/ILD neuromorphic SSL systems.

Reference	Year	# Mics	# Neurons	Learning	SNN Implementation	Angular Resolution	Test Environment	Sound Sources	Average Performances
[35]	2010	2	-	-	Simulation	$10^\circ$	Low noise	White noise, Speech	80% ACC
[73]	2010	2	106	-	Simulation	$15^\circ$	Quiet	White noise, Speech, Instruments	$2^\circ$ – $7^\circ$ MAE (azimuth) $7^\circ$ – $20^\circ$ MAE (elevation)
[74]	2012	2	-	-	Simulation	$15^\circ$	Quiet	White noise, Speech	$2.5^\circ$ MAE $27^\circ$ MAE



Table 3. Cont.

Reference	Year	# Mics	# Neurons	Learning	SNN Implementation	Angular Resolution	Test Environment	Sound Sources	Average Performances
[75]	2022	2	67 261	Backpropagation (S)	Memristor array Simulation	40°	Quiet	-	12.5° RMSE 5.7°
[76]	2023	2	-	Backpropagation (S)	OESTs array	40°	Quiet	-	0.11 normalized MAE
[77]	2024	2	1516	Bayesian optimization	Simulation	10°	Quiet	Speech	86.33% ACC
[78]	2024	2	100	Modified STDP (S)	Simulation	-	Noisy	Natural sounds	69.8% ACC 3.4° MAE 0.38 m MAE <sup>1</sup>

ACC—Accuracy; MAE—Mean Absolute Error; S—Supervised; OESTs—Organic Electrochemical Synaptic Transistors; RMSE—Root Mean Square Error. <sup>1</sup> Distance estimation average performance.

## 5. Echolocation for Navigation and Ranging

Identifying objects in space using the acoustic sound they produce works well with noisy sources but is not possible with quiet and motionless subjects. While sound source localization can be described as a passive process that analyzes ambient sounds, echolocation involves the active emission of sound by the organism or system.

Extending these principles to technological applications, sonar (sound navigation and ranging) and radar (radio detection and ranging) are systems designed to perform echolocation-like tasks. Sonar uses ultrasound waves on land or underwater to detect objects and map the surroundings, functioning similarly to how bats or dolphins navigate and hunt. Radar, on the other hand, employs radio waves to detect the range, speed, and other characteristics of objects in the air or on the ground, akin to echolocation but utilizing electromagnetic waves instead of sound. Both sonar and radar exemplify how human technology can mimic natural echolocation to achieve tasks in various environments.

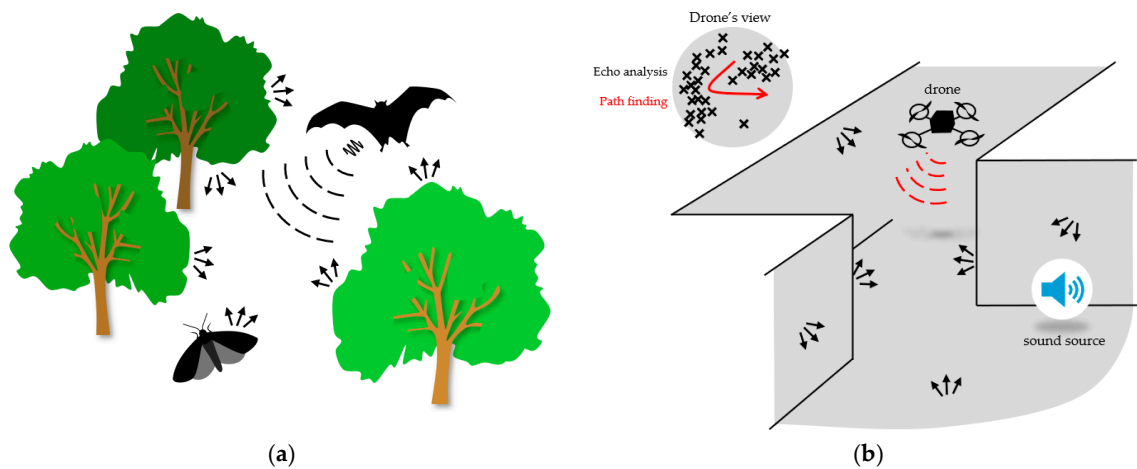
### 5.1. Echolocation Principle

Some animals use echolocation to hunt and navigate in their environment over a few meters to a few kilometers as a better means for object detection and environment mapping than vision under low visibility conditions. Among the animals that use echolocation, bats and cetaceans are the most commonly studied echolocating mammals, but studies have found evidence of this behavioral use in birds [79] and tree mice [80], although it is less sophisticated. Humans with complete or partial blindness have also developed the ability to differentiate objects' location, material, or shape by using acoustic echolocation [81].

Clicks or chirps, which are broadband sounds or frequency-modulated ultrasonic vocalizations, are produced by the individual and broadcast in any non-obstructed direction. The waves propagate in the surrounding space and are reflected as echoes upon reaching an obstacle. Depending on the material, the texture, and the location of the object, the frequency composition of the echoes changes due to wave absorption or dispersion, and the time difference perceived between an echo and the emitted echolocating sound varies with the distance. Figure 9a illustrates the biological echolocation behavior for navigation and hunting, adapted to modern navigation systems in Figure 9b, which typically are sonar- or radar-based. Jointly with continuous echolocation and the memorization of the perceived environment, the mapping of the surroundings increases in spatial accuracy with an increasing call rate, comparable to the framing rate of a video.

Echolocation and sound localization rely on shared auditory pathways to process different types of sound information, but there are also specialized adaptations for each function. Echolocators use IID and monaural spectral cues like frequency shifts of the perceived echoes to infer the azimuth of the detected object, and for bats in particular, spectral notches created by their pinnae help with decoding the elevation. Several models have been proposed to describe the neuronal processing which enables the individual to

understand the delays between calls and echoes [82]. In bats echolocating with frequency-modulated biosonar signals for example, specific detection neurons tuned for particular delays spike when inhibitory rebounds from the call and its echo coincide.



**Figure 9.** Echolocation principle. (a) Illustration of the navigation in-flight of a bat with echolocation. Echoes, shown by black arrows, returning with different elevations and spectral content inform on the surrounding environment mapping and texture. The same behavior is used during hunting for locating prey. (b) Application of sonar or radar to drones for navigation in environments where vision cannot be used. Combined with SSL, full tracking of sound sources is possible. Crosses in the drone's view are identified obstacles from echoes; a path within the empty space can then be determined.

## 5.2. Sonar

Sonar is a technology fundamentally rooted in the principles of echolocation. Inspired by this natural behavior, sonar systems were developed to perform similar tasks, particularly underwater, where visibility is limited. These systems have evolved from simple echo-based distance measurements to complex signal-processing techniques. There are two main types of sonar: active and passive. Active sonar involves sending out sound waves and listening for their return, while passive sonar listens for sounds emitted by objects themselves. In this section, only active sonars that use neuromorphic processing in pulsed or spiking systems will be reviewed.

The majority of the reviewed works on neuromorphic echolocation with spiking systems were conducted by or with the same author, Horiuchi [83–88], whose team proposed several models of the bat midbrain while focusing on ITD, ILD processing, and mapping memory essentially through a bioplausible approach.

Cheely and Horiuchi [83] presented a neuromorphic VSLI implementation mimicking the neuronal processing in the bat for the temporal analysis of echoes using post-inhibitory rebounds. Here, 13 neurons with tuned delay-lines were used to detect echo-pulse delays ranging from 1 to 30 ms with a theoretical average power consumption of 550  $\mu$ W.

Shi and Horiuchi [84] studied the processing of ILD for azimuthal echolocation with a neuromorphic VSLI that implements the LSO, dorsal nucleus of the lateral lemniscus, and IC neuron populations of the bat. The emitted 40 kHz ultrasonic signals were encoded into spike trains by logarithmic envelope extraction of echo-pulses and sent to an LSO layer where the first spike latency of the input was used to encode the ILD. Single spikes were thus provided to the dorsal nucleus and IC layers for azimuth estimation, where spike responses were observed for a target placed 90 cm away and moved on a  $\pm 70^\circ$  range with  $5^\circ$  steps.

Circuits in [83,84] were further studied in [85], which focused on enhancing the system with regards to data representation, parameters, and hardware optimization.

Horiuchi [86] presented a spiking implementation of an algorithm for obstacle avoidance in navigation connected to the binaural sonar head introduced in [84] and to a WTA

network for steering selection in output. Winning directions were selected by calculating the echo direction based on ILD and generating direction-specific inhibition with the sonar head, resulting in delayed spikes in those directions.

Following previous works on LSO neuromorphic implementations for echolocation, Wen and Horiuchi [87] described the development of square- and cube-root compression of echo envelopes for direction of arrival detection in VLSI hardware, and focused on single neuron LSO. A circuit performing amplification, compression, and filtering led to neuron responses sensitive to amplitude ratios. This work observed the output activity for different distances demonstrating the independence of the amplitude scaling with a resulting working range of 0.9 m to 2.1 m, enabled by the proposed power-law compression. As a result, average errors evaluated on an angular range of  $\pm 30^\circ$  were below  $4.2^\circ$  but still increased with the distance.

Isbell and Horiuchi [88] constructed echo view cells, inspired by place cells relying on odometry and sensory signals, which enabled the recognition of previously memorized echo patterns for efficient navigation and mapping. Sonar echoes were recorded between  $\pm 5^\circ$  in  $1^\circ$  steps at different positions in a room, and were processed by a multi-layer SNN with delay-lines, a variant of reservoir computing. Trained with supervised randomized learning, it obtained 93.5% accuracy in predicting from which known locations a sonar pattern originated.

Beside the echolocation systems mentioned above, other researchers have investigated spike encoding and neuromorphic implementations in sonar-based navigation.

The biological processing of echoes in the bat was also taken as inspiration by Kuc [89] to process more efficiently reverberation artifacts from conventional sonar in a pulse system. Further studied in [90], the system was used for estimation with a moving sonar of passing range for navigation. The author employed delays and multiresolution coincidence detection for the object localization and classification of surface roughness from strong echo artifacts.

Fontaine and Peremans [91] argued for the plausibility of known neural architectures of the frequency-modulating bats' midbrain and investigated a first spike latency coding for HRTF binaural and monaural cues extraction in comparison with spike rate coding. As was revealed in [84], ILD impacts the latency of generated spikes in the LSO, which thus also encodes the spectral peak and notches. In this study, a detailed simulation of the biological system was performed, and a machine learning regression algorithm was used for extracting azimuth and elevation from IC spiking patterns, such that in azimuth/elevation, an MAE of  $8.9^\circ/6.5^\circ$  and  $4.05^\circ/6.5^\circ$  was observed from monaural and binaural experiments.

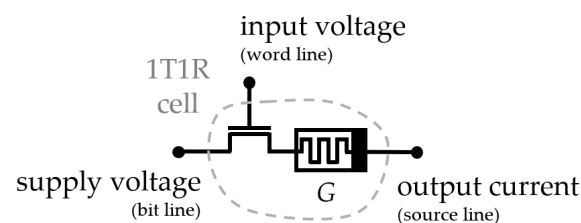
Tandon et al. [92] reproduced the system in [84] based on echo intensity for simultaneous obstacle avoidance and source tracking similar to an SSL task with a spike rate encoding of the azimuth to control an artificial head. The authors reported 68% of successful sources reaching accuracy and poor collision avoidance under noisy conditions with the highest operation frequency. Without noise, the best root mean square error was approximately 10% in head aim for target tracking in azimuth.

Reinforcement learning was investigated by Amaravati et al. [93] using a biomimetic approach in a digital accelerator. Ultrasonic echoes were converted to pulses and then processed digitally in a fully connected network with a WTA output layer. The network was trained by online learning using backpropagation while the system was mounted on a moving robot for obstacle avoidance. This study focused on optimizing energy efficiency, which was measured at an average of 1.25 pJ/multiply-accumulate (MAC) operation by the chip alone, or 1.5 nJ and 670 pJ energy consumption at training and inference, for a supply voltage of 0.8 V.

An artificial head was designed by Khyam et al. [94] with mobile ears and nose inspired by the Horseshoe bat, relying on combined constant frequency and frequency modulation biosonar signals, and was used to record numerous echoes in environments with various foliage types. Recordings were then filtered and encoded into spikes from

envelopes, modeling the auditory system of the bat. The sonar was then used in [95] in the input of a convolutional ANN for passageway detection in foliage.

At the limit between a sonar mapping and the SSL systems described in Section 4, Moro et al. [96] performed object detection with an ultrasonic emitter. Two receivers made it possible to work with an ITD that depended on the azimuth of the object detected. The ITD extractor was biomimetic, reproducing the Jeffress model, and the focus of the paper was its implementation in a RRAM-based neuromorphic circuit platform using parallel one-resistor-one-transistor structures for which one cell is represented in Figure 10. Characterization of the system led to a theoretical angular resolution of  $4^\circ$  and  $10^\circ$  precision for objects located 50 cm away from the receivers, when considered for preprocessing in an SSL task, for a total power consumption of 81.6 nW by the system.



**Figure 10.** Single one-transistor one-resistor (1T1R) cell for RRAM-based array. The resistive element schematized is a memristor whose conductance is akin to a synaptic weight.

The works described in this section take strong inspiration from biology, essentially from bats which live in environments with which we are familiar, such as forests or cities, and to which drones can be compared. Research in neuromorphic sonar echolocation is similar to SSL previously reviewed in their approach and trend. Echolocation also shares the use of the auditory system in the domain of ultrasounds, so naturally ITD, with the Jeffress model in [96], and ILD, with the LSO in [84,87,92], are cues used for the processing of echoes. More specific pathways such as inhibitory rebound in [83] can be found but are far from representing the majority.

### 5.3. Radar

Radar relies on the emission of electromagnetic waves in short pulses or with frequency-modulated continuous waves (FMCW) in the radio frequency domain; these are also used by animal species and called clicks and chirps, respectively. The analysis of echoes enables the localization and tracking of objects, but also makes it possible to determine their velocity from frequency shifts induced by the Doppler effect, or identify objects and map environments by radar imaging. Radars exist in largely varying sizes and energy efficiency depending on the application, ranging from a few milliwatts, like in gesture recognition and short-range sensing, to a few gigawatts, for example in air traffic or astronomical body observation systems.

Although radar-based neuromorphic systems are less explored, recent studies have shown the potential of spiking systems for extracting or processing radar waveforms and images. Of the few works reported in the literature, most engaged in a recognition process of objects or movements with SNNs [97,98], which is loosely related to positioning and navigation. Nevertheless, some studies undertook a localization or mapping task with radar data.

Vogginger et al. [99] proposed theoretical SNN alternatives to radar processing steps for object detection and tracking from range-Doppler and range-angle maps in the output of FMCW radars. Because of the continuous emission and small phase shifts in FMCW signals, relying on binaural cues would be scarcely possible, so the spiking Fourier transform was suggested as the only viable solution so far. The authors proposed two SNNs, one for the identification of object reflections from noise based on constant false alarm rate algorithms, and another for object classification. Conjectures regarding possible adaptations of known

path integration techniques with SNN for simultaneous localization and mapping were discussed but not investigated.

A collision avoidance method was introduced by Van Damme et al. [100] that processed FMCW radar images for robot navigation. Radar images were translated to a range azimuth matrix with conventional object detection algorithms and encoded into spike trains with rate coding for direction choice by a SNN trained with spike timing-dependent plasticity and reinforcement learning. Knowing a target location, the SNN was tested indoors with a mobile robot and six obstacle scenarios, which resulted in an average collision rate of 0.027 for an average success rate of 0.54 of target locations reached (weighted by the traveled path length).

Recently, Safa et al. [101] performed simultaneous localization and mapping using an event-based camera and a FMCW radar, where sensor images were encoded into spike images and processed by an SNN with online continual unsupervised spike timing-dependent plasticity. While a drone explored a storage room, the SNN mapped and learned the environment. The SNN output was then processed by the RatSLAM algorithm, which also relies on odometry data for mapping correction and loop closure detection. Localization and mapping MAE averaged to 0.66 m and 0.31 m, respectively, on all flight sequences, or 0.83 m and 0.46 m, respectively, without camera images.

Jointly with the implementation of the first complete neuromorphic radar hardware architecture NeuroRadar, Zheng et al. [102] evaluated the radar's performances in an indoor tracking case study with a convolutional SNN in simulation. The radar sensor used multiple self-injection locking modules to capture amplitudes and phases of reflections. The received frequency shifts were demodulated for the generation of a baseband carrying the motion information then encoded into spikes by LIF neurons. Backpropagation and ANN-SNN conversion were used for training the convolutional SNN that processed the generated spikes, and a mean square localization error of 1 m was reported for a power consumption of 2.03 mW considering the front-end and signal processing.

To summarize, Table 4 provides key information on the reviewed sonar and radar-based echolocation systems for navigation. Neuromorphic computing only started being associated with radar systems recently in the midst of the growing interest in the technology. Neuromorphic radar-based systems do not rely on bioinspiration; instead, the research conducted explores the potential of this paradigm with radar, whose applications in embedded systems are not few. SNNs are tested here in the context of object detection and mapping, with architectures in feed-forward layers or convolutional to process 2-D data.

**Table 4.** Neuromorphic sonar and radar-based echolocation systems for navigation.

Reference	Year	Inputs	Emission	Learning	SNN Implementation	Task	Experimental Precision Performances
[83]	2003	Sonar	P	-	VLSI	Neurological study	-
[84]	2007	Sonar	P	-	VLSI	Neurological study	-
[90]	2007	Sonar	P	-	Simulation	Mapping	-
[85]	2007	Sonar	P	-	VLSI	Mapping	-
[86]	2009	Sonar	P	-	VLSI	Collision avoidance	-
[91]	2009	Sonar	FMCW	SVM regression (S)	Simulation	Neurological study	4.05° MAE azimuth <sup>1</sup> 6.5° MAE elevation <sup>1</sup>

Table 4. Cont.

Reference	Year	Inputs	Emission	Learning	SNN Implementation	Task	Experimental Precision Performances
[92]	2016	Sonar, Sound	P	Backpropagation (S)	Simulation	Collision avoidance Navigation Tracking	36.6% ACC 68% ACC 10% RMSE
[87]	2018	Sonar	P	-	VLSI	Object localization	<4.2° mean error
[93]	2019	Sonar	P	Backpropagation (RL)	VLSI	Obstacle avoidance	-
[94]	2019	Sonar	CF-FM	-	-	Spike encoding Dataset creation	-
[88]	2020	Sonar, Odometry	P	Randomized learning (S)	Simulation	Mapping Navigation	93.5% ACC
[96]	2022	Sonar	P	-	RRAM array	Object localization	-
[99]	2022	Radar	FMCW	-	Simulation	Object detection Tracking	-
[100]	2023	Radar	FMCW	STDP (RL)	Simulation	Collision avoidance Navigation	2.7% ACC 54% ACC
[101]	2023	Radar, Video	FMCW	STDP (US)	Simulation	Navigation Mapping	0.66 m MAE 0.31 m MAE
[102]	2024	Radar	FMCW	Backpropagation (S) ANN-SNN conversion	VLSI and Simulation	Object localization	1 m MSE

P—Pulses; FMCW—Frequency-Modulated Continuous waves; S—Supervised; RL—Reinforcement learning; CF-FM—Constant Frequency and Frequency Modulation; US—Unsupervised; MAE—Mean Absolute Error; ACC—Accuracy; MSE—Mean Square Error; RMSE—Root MSE. <sup>1</sup> Results for binaural inputs.

## 6. Discussion

### 6.1. Methods and Precision Performances

Neuromorphic systems with spike representation in SSL and echolocation are mostly driven by research on the biological neuronal pathways in auditory systems. Most works have reproduced the mammalian binaural auditory system. The creation of bioplausible models of SSL or echolocation is usually not with the aim of enhancing performances but rather of introducing new possibilities for spike processing. Beyond the commonly reproduced MSO, LSO, and IC, other minor neuron populations have been simulated to refine the understanding of binaural cues and subtle differences related to HRTF data [57,88].

However, being biomimetic often does not benefit localization precision, as was demonstrated in [61,63]. More extensive work has been carried out on SSL tasks compared to sonar and radar-based echolocation systems, which revealed the advantages of exploring bioinspiration with conventional processing or architectures used in ANNs. Reproducing the brain structure and neuronal pathways will not enable the outstanding performances seen in mammals without providing an equivalent computational capability. The number of neurons and synapses is difficult to replicate, and simply playing with this parameter is bound to have its limitations. Goodman and Brette [73] used 106 neurons for a minimum of 2° MAE in azimuth, whereas Pan et al. [61] obtained ~1° MAE with ~2 k neurons, and Chen et al. [63] reported less than 1° MAE with ~5 k neurons. In fact, the use of network architectures and supervised learning methods from conventional deep learning adapted to spike encoding allows systems to reach higher levels of precision [63]. Reinforcement and

unsupervised learning have been used in echolocating systems which are more bioinspired, but supervision represents the majority of the trainings.

As neuromorphic technology and its capabilities are being studied, different architectures and processing techniques naturally appear, changing the focus of research in this field. Over the years, test conditions and system topologies have become more complex and sometimes greatly differed. The reviewed neuromorphic navigation systems used different test scenarios and evaluation metrics. In SSL systems, some focused on pure tones and/or common noise distributions, while others used speech and/or natural sounds for performance evaluation. An increase in precision and angular resolution of the proposed SSL solutions is observed, although mainly in ITD-only systems. The angular resolution of works using ILD remained at  $10^\circ$  and with precisions not yet competing with ITD-only systems. Additionally, SSL [35,67,68,70–78] and sonar-based [84–88,91,92] systems that relied on IID or ILD mostly did not evaluate the evolution of the precision with the distance. Models inspired by or mimicking the LSO for the extraction of intensity-related cues are typically used, and only a few studies [78,87,88,92] have considered varying distances, mostly sonar-based systems as part of a navigation task.

Similar to how mammals process their environment, localization or navigation can be assisted by or combined with multimodal information such as video [37,42,43,53,59,101] or odometry [88]. Furthermore, mobile mount can enhance the systems' perception by incorporating movements, like in [37,56,59] for SSL or with the reviewed echolocators. Using data from several sensors of different modalities would certainly lead to more efficient systems, with increased robustness in real-world applications by having supplementary information like vision and hearing provide. The radar echolocator in [101] reported path integration performances with vision-only, radar-only, and both sensors, where the latter performed better thanks to data fusion. Moreover, bioinspiration can also be transmodal, that is, the adaptation of the biological process to data issued from other sensors, like visual motion [37,59] or sand scorpion vibration detection [56].

Few studies have been conducted to create sonar spiking object localization and/or navigation systems, and even fewer for radar-based systems. The recent first investigations and the development of a neuromorphic radar [102] described previously will likely encourage further research in the context of autonomous vehicles and mobile robots. Concentrated in the past two years, these works already provide great insights into the possible implementation of neuromorphic computing in radar systems.

## 6.2. Hardware and Energy Efficiency

Enabled by the surge of neuromorphic processors, FPGA implementations of SNNs are growing in all fields, and this can be seen in SSL as well. The majority of hardware implementations were made with VLSI circuits in the early studies of neuromorphic SSL, which then shifted to FPGA and, more specifically, with TrueNorth and SpiNNaker accelerators. Moreover, memristive arrays have recently been developed and tested with sonar echolocation [96] and SSL tasks [75,76]. These technologies will certainly continue to be investigated as extensive research is ongoing to create extreme energy-efficient memory units for synaptic weight retention.

Neuromorphic solutions aim to improve the energy efficiency of localization and navigation tasks, especially among deep learning methods, by emulating the neural mechanisms of biological auditory systems while having sufficient precision to be used in real-world applications, for which progress in real-world experiments remains to be achieved.

The asynchronous event processing and sparsity of spike encoding allow for a reduction in the power consumption, which could lead to implementations being more energy-efficient than conventional methods. Table 5 reports the power consumption and energy efficiency of the reviewed hardware implementations. Amaravati et al. [93] reported a power consumption at least two orders of magnitude lower than other sonars with ANNs. In the study by Zheng et al. [102], the power consumption of the NeuroRadar was also 97% lower compared to a standard radar tracking system.

**Table 5.** Power consumption and energy efficiency of neuromorphic SSL and echolocation systems.

Reference	Year	Total Power Consumption	Energy Efficiency
[83]	2003	550 $\mu$ W <sup>1</sup>	-
[70]	2018	58 mW	-
[93]	2019	690 $\mu$ W	1.25 pJ/MAC operation
[75]	2022	0.306 $\mu$ J	-
[62]	2022	-	0.9 nJ/spike
[76]	2023	-	2.03 fJ/synaptic event
[37]	2023	972 mW	-
[102]	2024	2.03 mW	-

<sup>1</sup> Theoretical value.

In fact, little information is available on the energy efficiency of most reviewed papers unless the hardware is the focus of the work. Especially for SNNs, which claim to offer higher energy efficiency, it would be relevant to look at the precision performances against the power consumption or computational cost for a more accurate comparison of the different approaches. The combination of these metrics is an indicator of the system's complexity, as balancing high performance and low-power consumption remains a challenge. Here, some low and ultra-low power consumptions are reported, but superiority in power consumption is not always showcased in SNNs: unless workloads, processing, and learning methods adapted to the temporality of spikes and aligned with neuromorphic architectures are used, overall performances fall behind those of ANNs according to recent studies [103,104]. Digital processing has the advantage of being easier to implement, but lacks compatibility with spike representation, leading to additional computational resources being spent to emulate SNNs. Therefore, unless a power consumption is given, interpretation of the results may be uncertain and unsatisfying. In the future, efforts should be made to evaluate the energy efficiency of neuromorphic implementations as this is the main motivation that drives neuromorphic technology.

Systems using emerging materials and technologies [62,75,76] reported energy-related performances but appeared to provide more of a proof of concept than a possible device implementation, with extensive characterization of the novel components. Further maturation would be required, also regarding variability, stability, and scalability, to consider these technologies for positioning in real-world conditions. They are tested in the context of sound localization and echolocation within laboratories, and cannot insure the same functioning outside laboratory conditions or over multiple instances of the same device yet, unlike FPGA boards, which are more flexible and stable. However, emerging technologies hold the key to ultra-low power consumption, which digital hardware has difficulty reaching, by providing greater compatibility with spike computing with asynchronous and analog processing, among other characteristics.

## 7. Conclusions

Ultimately, then, bioinspiration holds the promise of low power consumption systems with sufficient precision for most applications. The underlying motivation is the reproduction of mechanisms observed in the living in order to replicate their performances in terms of precision, but also the reduction of energy costs, as can be seen in the most recent reviewed works, which reveal this trend. Nevertheless, the evaluation of energy efficiency is yet to be sufficiently assessed considering its relative significance to neuromorphic computing.

Research in neuromorphic SSL and sonar-based navigation was driven, first and foremost, by the motivation to validate in simulation neurological processes; then, as the spike toolbox expanded and new computational capability was unlocked, better results were obtained. Methods found in the literature are centered around the main binaural cues, ITD and ILD, typically extracted after spectral segmentation; further research is required for the successful integration of both cues to compete with current solutions in ITD-only systems.



Radar-based navigation has just started to implement neuromorphic computing in its processing, but it has already proven its great potential for embedded applications with improved adaptability to the environment and a frugal power consumption.

The increasing diversity of models and approaches seems to have grown in particular in the past few years. The limit of this growth has yet to be reached, and constant progress can be observed in the design of efficient neuromorphic hardware, SNN architectures, and learning methods, which should bring SNNs closer to achieving overall higher performances in embedded applications.

**Author Contributions:** Conceptualization, E.D. and C.L.; methodology, E.D. and C.L.; investigation, E.D.; writing—original draft preparation, E.D.; writing—review and editing, F.D. and F.E. and M.B. and C.L.; supervision, F.D. and C.L.; project administration, C.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is supported by the French Agence Nationale de la Recherche (ANR) project ULP SMART 3D COCHLEA.

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Niu, L.Y.; Wei, Y.; Liu, W.B.; Long, J.Y.; Xue, T.H. Research progress of spiking neural network in image classification: A review. *Appl. Intell.* **2023**, *53*, 19466–19490. [[CrossRef](#)]
- Wu, J.; Yilmaz, E.; Zhang, M.; Li, H.; Tan, K.C. Deep spiking neural networks for large vocabulary automatic speech recognition. *Front. Neurosci.* **2020**, *14*, 199. [[CrossRef](#)] [[PubMed](#)]
- Yang, C.C.; Chang, T.S. A 71.2- $\mu$ W Speech Recognition Accelerator With Recurrent Spiking Neural Network. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2024**, *71*, 3203–3213. [[CrossRef](#)]
- Gu, F.; Sng, W.; Taunyazov, T.; Soh, H. Tactilesgnet: A spiking graph neural network for event-based tactile object recognition. In Proceedings of the 2020 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October–24 January 2021.
- Kang, P.; Banerjee, S.; Chopp, H.; Katsaggelos, A.; Cossairt, O. Boost event-driven tactile learning with location spiking neurons. *Front. Neurosci.* **2023**, *17*, 1127537. [[CrossRef](#)] [[PubMed](#)]
- Juarez-Lora, A.; Ponce-Ponce, V.H.; Sossa, H.; Rubio-Espino, E. R-STDP spiking neural network architecture for motion control on a changing friction joint robotic arm. *Front. Neurobot.* **2022**, *16*, 904017. [[CrossRef](#)]
- Barchid, S.; Mennesson, J.; Eshraghian, J.; Djéraba, C.; Bennamoun, M. Spiking neural networks for frame-based and event-based single object localization. *Neurocomputing* **2023**, *559*, 126805. [[CrossRef](#)]
- Abubaker, B.A.; Ahmed, S.R.; Guron, A.T.; Fadhil, M.; Algburi, S.; Abdulrahman, B.F. Spiking Neural Network for Enhanced Mobile Robots' Navigation Control. In Proceedings of the 2023 7th International Symposium on Innovative Approaches in Smart Technologies (ISAS), Istanbul, Türkiye, 23–25 November 2023.
- Wang, J.; Lin, S.; Liu, A. Bioinspired Perception and Navigation of Service Robots in Indoor Environments: A Review. *Biomimetics* **2023**, *8*, 350. [[CrossRef](#)]
- Yang, Y.; Bartolozzi, C.; Zhang, H.H.; Nawrocki, R.A. Neuromorphic electronics for robotic perception, navigation and control: A survey. *Eng. Appl. Artif. Intell.* **2023**, *126*, 106838. [[CrossRef](#)]
- Novo, A.; Lobon, F.; Garcia de Marina, H.; Romero, S.; Barranco, F. Neuromorphic Perception and Navigation for Mobile Robots: A Review. *ACM Comput. Surv.* **2024**, *56*, 1–37. [[CrossRef](#)]
- Schuman, C.D.; Kulkarni, S.R.; Parsa, M.; Mitchell, J.P.; Kay, B. Opportunities for neuromorphic computing algorithms and applications. *Nat. Comput. Sci.* **2022**, *2*, 10–19. [[CrossRef](#)]
- Javanshir, A.; Nguyen, T.T.; Mahmud, M.P.; Kouzani, A.Z. Advancements in algorithms and neuromorphic hardware for spiking neural networks. *Neural Comput.* **2022**, *34*, 1289–1328. [[CrossRef](#)] [[PubMed](#)]
- Schemmel, J.; Brüderle, D.; Grübl, A.; Hock, M.; Meier, K.; Millner, S. A wafer-scale neuromorphic hardware system for large-scale neural modeling. In Proceedings of the 2010 IEEE International Symposium on Circuits and Systems (ISCAS), Paris, France, 30 May–2 June 2010.
- Furber, S.B.; Lester, D.R.; Plana, L.A.; Garside, J.D.; Painkras, E.; Temple, S.; Brown, A.D. Overview of the SpiNNaker system architecture. *IEEE Trans. Comput.* **2012**, *62*, 2454–2467. [[CrossRef](#)]
- Akopyan, F.; Sawada, J.; Cassidy, A.; Alvarez-Icaza, R.; Arthur, J.; Merolla, P.; Imam, N.; Nakamura, Y.; Datta, P.; Nam, G.J.; et al. Truenorth: Design and tool flow of a 65 mw 1 million neuron programmable neurosynaptic chip. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2015**, *34*, 1537–1557. [[CrossRef](#)]
- Davies, M.; Wild, A.; Orchard, G.; Sandamirskaya, Y.; Guerra, G.A.F.; Joshi, P.; Plank, P.; Risbud, S.R. Advancing neuromorphic computing with loihi: A survey of results and outlook. *Proc. IEEE* **2021**, *109*, 911–934. [[CrossRef](#)]

18. Gökgöz, B.; Gül, F.; Aydın, T. An overview memristor based hardware accelerators for deep neural network. *Concurr. Comput. Pract. Exp.* **2024**, *36*, e7997. [[CrossRef](#)]
19. Duan, X.; Cao, Z.; Gao, K.; Yan, W.; Sun, S.; Zhou, G.; Wu, Z.; Ren, F.; Sun, B. Memristor-Based Neuromorphic Chips. *Adv. Mater.* **2024**, *36*, 2310704. [[CrossRef](#)]
20. Auge, D.; Hille, J.; Mueller, E.; Knoll, A. A survey of encoding techniques for signal processing in spiking neural networks. *Neural Process. Lett.* **2021**, *53*, 4693–4710. [[CrossRef](#)]
21. Yi, Z.; Lian, J.; Liu, Q.; Zhu, H.; Liang, D.; Liu, J. Learning rules in spiking neural networks: A survey. *Neurocomputing* **2023**, *531*, 163–179. [[CrossRef](#)]
22. Guo, Y.; Huang, X.; Ma, Z. Direct learning-based deep spiking neural networks: A review. *Front. Neurosci.* **2023**, *17*, 1209795. [[CrossRef](#)]
23. Bi, G.Q.; Poo, M.M. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **1998**, *18*, 10464–10472. [[CrossRef](#)]
24. Dampfhofer, M.; Mesquida, T.; Valentian, A.; Anghel, L. Backpropagation-based learning techniques for deep spiking neural networks: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *35*, 11906–11921. [[CrossRef](#)] [[PubMed](#)]
25. Maass, W.; Natschläger, T.; Markram, H. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comput.* **2002**, *14*, 2531–2560. [[CrossRef](#)] [[PubMed](#)]
26. Feldman, J.A.; Ballard, D.H. Connectionist models and their properties. *Cogn. Sci.* **1982**, *6*, 205–254. [[CrossRef](#)]
27. Moore, D.R.; Fuchs, P.A.; Rees, A.; Palmer, A.; Plack, C.J. *The Oxford Handbook of Auditory Science: The Auditory Brain*; Oxford University Press: Oxford, UK, 2010; Volume 2. [[CrossRef](#)]
28. Kunchur, M.N. The human auditory system and audio. *Appl. Acoust.* **2023**, *211*, 109507. [[CrossRef](#)]
29. Carlini, A.; Bordeau, C.; Ambard, M. Auditory localization: A comprehensive practical review. *Front. Psychol.* **2024**, *15*, 1408073. [[CrossRef](#)]
30. Schnupp, J.; Nelken, I.; King, A. Neural Basis of Sound Localization. In *Auditory Neuroscience: Making Sense of Sound*; MIT Press: Cambridge, MA, USA, 2011; pp. 177–221.
31. Zonooz, B.; Arani, E.; Körding, K.P.; Aalbers, P.R.; Celikel, T.; Van Opstal, A.J. Spectral weighting underlies perceived sound elevation. *Sci. Rep.* **2019**, *9*, 1642. [[CrossRef](#)]
32. Lyon, R. A computational model of filtering, detection, and compression in the cochlea. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'82), Paris, France, 3–5 May 1982.
33. Zilany, M.S.; Bruce, I.C. Representation of the vowel /ε/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats. *J. Acoust. Soc. Am.* **2007**, *122*, 402–417. [[CrossRef](#)]
34. Chan, V.; Liu, S.C.; van Schaik, A. AER EAR: A matched silicon cochlea pair with address event representation interface. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2007**, *54*, 48–59. [[CrossRef](#)]
35. Liu, J.; Perez-Gonzalez, D.; Rees, A.; Erwin, H.; Wermter, S. A biologically inspired spiking neural network model of the auditory midbrain for sound source localisation. *Neurocomputing* **2010**, *74*, 129–139. [[CrossRef](#)]
36. Jiménez-Fernández, A.; Cerezuela-Escudero, E.; Miró-Amarante, L.; Domínguez-Morales, M.J.; de Asís Gómez-Rodríguez, F.; Linares-Barranco, A.; Jiménez-Moreno, G. A binaural neuromorphic auditory sensor for FPGA: A spike signal processing approach. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 804–818. [[CrossRef](#)]
37. Schoepe, T.; Gutierrez-Galan, D.; Dominguez-Morales, J.P.; Greateorex, H.; Jimenez-Fernandez, A.; Linares-Barranco, A.; Chicca, E. Closed-loop sound source localization in neuromorphic systems. *Neuromorphic Comput. Eng.* **2023**, *3*, 024009. [[CrossRef](#)]
38. Schauer, C.; Zahn, T.; Paschke, P.; Gross, H.M. Binaural sound localization in an artificial neural network. In Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Turkey, 5–9 June 2000.
39. Ashida, G.; Carr, C.E. Sound localization: Jeffress and beyond. *Curr. Opin. Neurobiol.* **2011**, *21*, 745–751. [[CrossRef](#)] [[PubMed](#)]
40. Lazzaro, J.; Mead, C.A. A silicon model of auditory localization. *Neural Comput.* **1989**, *1*, 47–57. [[CrossRef](#)]
41. Knudsen, E.I.; Konishi, M. A neural map of auditory space in the owl. *Science* **1978**, *200*, 795–797. [[CrossRef](#)]
42. Lyon, R.F.; Mead, C. An analog electronic cochlea. *IEEE Trans. Acoust. Speech Signal Process.* **1988**, *36*, 1119–1134. [[CrossRef](#)]
43. Horiuchi, T. An auditory localization and coordinate transform chip. *Adv. Neural Inf. Process. Syst.* **1994**, *7*, 787–794.
44. Yin, T.C. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In *Integrative Functions in the Mammalian Auditory Pathway*; Oertel, D., Fay, R.R., Popper, A.N., Eds.; Springer: New York, NY, USA, 2002; Volume 15, pp. 99–159.
45. Voutsas, K.; Adamy, J. A biologically inspired spiking neural network for sound source lateralization. *IEEE Trans. Neural Netw.* **2007**, *18*, 1785–1799. [[CrossRef](#)]
46. Kugler, M.; Iwasa, K.; Benso, V.A.P.; Kuroyanagi, S.; Iwata, A. A complete hardware implementation of an integrated sound localization and classification system based on spiking neural networks. In Proceedings of the Neural Information Processing: 14th International Conference (ICONIP 2007), Kitakyushu, Japan, 13–16 November 2007.
47. Iwasa, K.; Kugler, M.; Kuroyanagi, S.; Iwata, A. A sound localization and recognition system using pulsed neural networks on FPGA. In Proceedings of the 2007 International Joint Conference on Neural Networks, Orlando, FL, USA, 12–17 August 2007.
48. Glackin, B.; Wall, J.A.; McGinnity, T.M.; Maguire, L.P.; McDaid, L.J. A spiking neural network model of the medial superior olive using spike timing dependent plasticity for sound localization. *Front. Comput. Neurosci.* **2010**, *4*, 18. [[CrossRef](#)]
49. Tollin, D.; Koka, K. Postnatal development of sound pressure transformations by the head and pinnae of the cat: Monaural characteristics. *J. Acoust. Soc. Am.* **2009**, *125*, 980. [[CrossRef](#)]

50. Chan, V.Y.S.; Jin, C.T.; Schaik, A.V. Adaptive sound localization with a silicon cochlea pair. *Front. Neurosci.* **2010**, *4*, 196. [[CrossRef](#)]
51. Finger, H.; Liu, S.C. Estimating the location of a sound source with a spike-timing localization algorithm. In Proceedings of the 2011 IEEE International Symposium of Circuits and Systems (ISCAS), Rio de Janeiro, Brazil, 15–18 May 2011.
52. Wall, J.A.; McGinnity, T.M.; Maguire, L.P. A comparison of sound localisation techniques using cross-correlation and spiking neural networks for mobile robotics. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011.
53. Chan, V.Y.S.; Jin, C.T.; van Schaik, A. Neuromorphic audio–visual sensor fusion on a sound-localizing robot. *Front. Neurosci.* **2012**, *6*, 21. [[CrossRef](#)] [[PubMed](#)]
54. Park, P.K.; Ryu, H.; Lee, J.H.; Shin, C.W.; Lee, K.B.; Woo, J.; Kim, J.S.; Kang, B.C.; Liu, S.C.; Delbruck, T. Fast neuromorphic sound localization for binaural hearing aids. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013.
55. Faraji, M.M.; Shouraki, S.B.; Iranmehr, E. Spiking neural network for sound localization using microphone array. In Proceedings of the 2015 23rd Iranian Conference on Electrical Engineering, Tehran, Iran, 10–14 May 2015.
56. Beck, C.; Garreau, G.; Georgiou, J. Sound source localization through 8 MEMS microphones array using a sand-scorpion-inspired spiking neural network. *Front. Neurosci.* **2016**, *10*, 479. [[CrossRef](#)] [[PubMed](#)]
57. Encke, J.; Hemmert, W. Extraction of inter-aural time differences using a spiking neuron network model of the medial superior olive. *Front. Neurosci.* **2018**, *12*, 140. [[CrossRef](#)] [[PubMed](#)]
58. Luke, R.; McAlpine, D. A spiking neural network approach to auditory source lateralisation. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019.
59. Schoepe, T.; Gutierrez-Galan, D.; Dominguez-Morales, J.P.; Jimenez-Fernandez, A.; Linares-Barranco, A.; Chicca, E. Neuromorphic sensory integration for combining sound source localization and collision avoidance. In Proceedings of the 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS), Nara, Japan, 17–19 October 2019.
60. Song, Z.W.; Xiang, S.Y.; Ren, Z.X.; Wang, S.H.; Wen, A.J.; Hao, Y. Photonic spiking neural network based on excitable VCSELs-SA for sound azimuth detection. *Opt. Express* **2020**, *28*, 1561–1573. [[CrossRef](#)]
61. Pan, Z.; Zhang, M.; Wu, J.; Wang, J.; Li, H. Multi-tone phase coding of interaural time difference for sound source localization with spiking neural networks. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2021**, *29*, 2656–2670. [[CrossRef](#)]
62. Zhong, S.; Zhang, Y.; Zheng, H.; Yu, F.; Zhao, R. Spike-Based Spatiotemporal Processing Enabled by Oscillation Neuron for Energy-Efficient Artificial Sensory Systems. *Adv. Intell. Syst.* **2022**, *4*, 2200076. [[CrossRef](#)]
63. Chen, X.; Yang, Q.; Wu, J.; Li, H.; Tan, K.C. A hybrid neural coding approach for pattern recognition with spiking neural networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *46*, 3064–3078. [[CrossRef](#)]
64. Milde, M.B.; Bertrand, O.J.; Ramachandran, H.; Egelhaaf, M.; Chicca, E. Spiking elementary motion detector in neuromorphic systems. *Neural Comput.* **2018**, *30*, 2384–2417. [[CrossRef](#)]
65. Gutierrez-Galan, D.; Schoepe, T.; Dominguez-Morales, J.P.; Jimenez-Fernandez, A.; Chicca, E.; Linares-Barranco, A. An event-based digital time difference encoder model implementation for neuromorphic systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 1959–1973. [[CrossRef](#)]
66. Dalmas, E.; Danneville, F.; Bocquet, M.; Loyez, C. Neuromorphic Coincidence Detector for Interaural Time Difference Encoding and Sound DOA Estimation. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–11. [[CrossRef](#)]
67. Wall, J.A.; McDaid, L.J.; Maguire, L.P.; McGinnity, T.M. Spiking neural network model of sound localization using the interaural intensity difference. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 574–586. [[CrossRef](#)] [[PubMed](#)]
68. Feng, X.; Dou, W. A Biologically Plausible Spiking Model of Human Auditory Pathways for Sound Localization. In Proceedings of the 2016 5th International Conference on Measurement, Instrumentation and Automation (ICMIA 2016), Shenzhen, China, 17–18 September 2016.
69. Xiao, F.; Weibei, D. A biologically plausible spiking model for interaural level difference processing auditory pathway in human brain. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016.
70. Escudero, E.C.; Peña, F.P.; Vicente, R.P.; Jimenez-Fernandez, A.; Moreno, G.J.; Morgado-Estevéz, A. Real-time neuro-inspired sound source localization and tracking architecture applied to a robotic platform. *Neurocomputing* **2018**, *283*, 129–139. [[CrossRef](#)]
71. Oess, T.; Löhner, M.; Jarvers, C.; Schmid, D.; Neumann, H. A bio-inspired model of sound source localization on neuromorphic hardware. In Proceedings of the 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS), Genova, Italy, 31 August–2 September 2020.
72. Schmid, D.; Oess, T.; Neumann, H. Listen to the Brain—Auditory Sound Source Localization in Neuromorphic Computing Architectures. *Sensors* **2023**, *23*, 4451. [[CrossRef](#)] [[PubMed](#)]
73. Goodman, D.F.; Brette, R. Spike-timing-based computation in sound localization. *PLoS Comput. Biol.* **2010**, *6*, e1000993. [[CrossRef](#)]
74. Dávila-Chacón, J.; Heinrich, S.; Liu, J.; Wermter, S. Biomimetic binaural sound source localisation with ego-noise cancellation. In Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2012: 22nd International Conference on Artificial Neural Networks, Lausanne, Switzerland, 11–14 September 2012.
75. Gao, B.; Zhou, Y.; Zhang, Q.; Zhang, S.; Yao, P.; Xi, Y.; Liu, Q.; Zhao, M.; Zhang, W.; Liu, Z.; et al. Memristor-based analogue computing for brain-inspired sound localization with in situ training. *Nat. Commun.* **2022**, *13*, 2026. [[CrossRef](#)]

76. Xu, Y.; Deng, Z.; Jin, C.; Liu, W.; Shi, X.; Chang, J.; Yu, H.; Liu, B.; Sun, J.; Yang, J. An organic electrochemical synaptic transistor array for neuromorphic computation of sound localization. *Appl. Phys. Lett.* **2023**, *123*, 133701. [[CrossRef](#)]
77. Li, Y.; Zhao, J.; Xiao, X.; Chen, R.; Wang, L. Brain-Inspired Binaural Sound Source Localization Method Based on Liquid State Machine. In Proceedings of the International Conference on Neural Information Processing (ICONIP 2023), Changsha, China, 20–23 November 2023.
78. Roozbehi, Z.; Narayanan, A.; Mohaghegh, M.; Saeedinia, S.A. Dynamic-Structured Reservoir Spiking Neural Network in Sound Localization. *IEEE Access* **2024**, *12*, 24596–24608. [[CrossRef](#)]
79. Brinkløv, S.; Fenton, M.B.; Ratcliffe, J.M. Echolocation in oilbirds and swiftlets. *Front. Physiol.* **2013**, *4*, 123. [[CrossRef](#)]
80. He, K.; Liu, Q.; Xu, D.M.; Qi, F.Y.; Bai, J.; He, S.W.; Chen, P.; Zhou, X.; Cai, W.Z.; Chen, Z.Z.; et al. Echolocation in soft-furred tree mice. *Science* **2021**, *372*, eaay1513. [[CrossRef](#)]
81. Thaler, L.; Goodale, M.A. Echolocation in humans: An overview. *Wiley Interdisciplinary Reviews. Cogn. Sci.* **2016**, *7*, 382–393.
82. Beetz, M.J.; Hechavarría, J.C. Neural processing of naturalistic echolocation signals in bats. *Front. Neural Circuits* **2022**, *16*, 899370. [[CrossRef](#)] [[PubMed](#)]
83. Cheely, M.; Horiuchi, T. A VLSI model of range-tuned neurons in the bat echolocation system. In Proceedings of the 2003 IEEE International Symposium on Circuits and Systems (ISCAS), Bangkok, Thailand, 25–28 May 2003.
84. Shi, R.Z.; Horiuchi, T.K. A neuromorphic VLSI model of bat interaural level difference processing for azimuthal echolocation. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2007**, *54*, 74–88. [[CrossRef](#)]
85. Horiuchi, T.K.; Cheely, M. A systems view of a neuromorphic VLSI echolocation system. In Proceedings of the 2007 IEEE International Symposium on Circuits and Systems (ISCAS), New Orleans, LA, USA, 27–30 May 2007.
86. Horiuchi, T.K. A spike-latency model for sonar-based navigation in obstacle fields. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2009**, *56*, 2393–2401. [[CrossRef](#)]
87. Wen, C.; Horiuchi, T.K. Power-Law Compression Expands the Dynamic Range of a Neuromorphic Echolocation System. In Proceedings of the 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), Cleveland, OH, USA, 17–19 October 2018.
88. Isbell, J.D.; Horiuchi, T.K. Echo View Cells From Bio-Inspired Sonar. *Front. Neurobot.* **2020**, *14*, 567991. [[CrossRef](#)]
89. Kuc, R. Biomimetic sonar and neuromorphic processing eliminate reverberation artifacts. *IEEE Sens. J.* **2007**, *7*, 361–369. [[CrossRef](#)]
90. Kuc, R. Neuromorphic processing of moving sonar data for estimating passing range. *IEEE Sens. J.* **2007**, *7*, 851–859. [[CrossRef](#)]
91. Fontaine, B.; Peremans, H. Bat echolocation processing using first-spike latency coding. *Neural Netw.* **2009**, *22*, 1372–1382. [[CrossRef](#)]
92. Tandon, P.; Malviya, Y.H.; Rajendran, B. Efficient and robust spiking neural circuit for navigation inspired by echolocating bats. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 938–946.
93. Amaravati, A.; Nasir, S.B.; Ting, J.; Yoon, I.; Raychowdhury, A. A 55-nm, 1.0–0.4 V, 1.25-pJ/MAC time-domain mixed-signal neuromorphic accelerator with stochastic synapses for reinforcement learning in autonomous mobile robots. *IEEE J. Solid-State Circuits* **2018**, *54*, 75–87. [[CrossRef](#)]
94. Khyam, M.O.; Alexandre, D.; Bhardwaj, A.; Wang, R.; Müller, R. Neuromorphic computing for autonomous mobility in natural environments. In Proceedings of the 7th Annual Neuro-inspired Computational Elements Workshop, Albany, NY, USA, 26–28 March 2019.
95. Wang, R.; Liu, Y.; Müller, R. Detection of passageways in natural foliage using biomimetic sonar. *Bioinspir. Biomim.* **2022**, *17*, 056009. [[CrossRef](#)]
96. Moro, F.; Hardy, E.; Fain, B.; Dalgaty, T.; Cléménçon, P.; De Prà, A.; Esmanhotto, E.; Castellani, N.; Blard, F.; Gardien, F.; et al. Neuromorphic object localization using resistive memories and ultrasonic transducers. *Nat. Commun.* **2022**, *13*, 3506. [[CrossRef](#)] [[PubMed](#)]
97. Huang, J.; Vogginger, B.; Gerhards, P.; Kreutz, F.; Kelber, F.; Scholz, D. Real-time radar gesture classification with spiking neural network on SpiNNaker 2 prototype. In Proceedings of the 2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Incheon, Republic of Korea, 13–15 June 2022.
98. Henderson, A.; Harbour, S.; Yakopcic, C.; Taha, T.; Brown, D.; Tieman, J.; Hall, G. Spiking Neural Networks for LPI Radar Waveform Recognition with Neuromorphic Computing. In Proceedings of the 2023 IEEE Radar Conference (RadarConf23), San Antonio, TX, USA, 1–5 May 2023.
99. Vogginger, B.; Kreutz, F.; López-Randulfe, J.; Liu, C.; Dietrich, R.; Gonzalez, H.A.; Scholz, D.; Reeb, N.; Auge, D.; Hille, J.; et al. Automotive radar processing with spiking neural networks: Concepts and challenges. *Front. Neurosci.* **2022**, *16*, 851774. [[CrossRef](#)] [[PubMed](#)]
100. Van Damme, L.; Durodié, Y.; Deckers, L.; Tsang, J.; Latré, S. Collision Avoidance Navigation with Radar and Spiking Reinforcement Learning. In Proceedings of the 2023 IEEE International Radar Conference (RADAR), Sydney, Australia, 6–10 November 2023.
101. Safa, A.; Verbelen, T.; Ocket, I.; Bourdoux, A.; Sahli, H.; Catthoor, F.; Gielen, G. Fusing event-based camera and radar for slam using spiking neural networks with continual stdp learning. In Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 29 May–2 June 2023.
102. Zheng, K.; Qian, K.; Woodford, T.; Zhang, X. NeuroRadar: A Neuromorphic Radar Sensor for Low-Power IoT Systems. In Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems, Istanbul, Türkiye, 12–17 November 2023.

103. Deng, L.; Wu, Y.; Hu, X.; Liang, L.; Ding, Y.; Li, G.; Zhao, G.; Li, P.; Xie, Y. Rethinking the performance comparison between SNNS and ANNS. *Neural Netw.* **2020**, *121*, 294–307. [[CrossRef](#)] [[PubMed](#)]
104. Ottati, F.; Gao, C.; Chen, Q.; Brignone, G.; Casu, M.R.; Eshraghian, J.K.; Lavagno, L. To spike or not to spike: A digital hardware perspective on deep learning acceleration. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2023**, *13*, 1015–1025. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.