Binaural Sound Localization Systems Based on Neural Approaches

Nick Rossenbach
June 17, 2016
Introduction

Barn Owl as Biological Example

Neural Audio Processing
   Jeffress model
   Spence & Pearson

Artifical Owl Ruff Localization System

Effect of an Artificial Head to Human Acoustic Perception

Conclusion
Introduction
Motivation:

- sound localization plays an important role for mobile robots
- binaural localization systems are common in nature

Reference: **Biologically Inspired Binaural Sound Source Localization and Tracking for Mobile Robots**, Calmes 2009

- uses barn owl as biological example
- implements system using artificial barn owl ruff
- also uses statistical tracking and visual sensor aids
Barn Owl as Biological Example
Barn Owl

- one of nature's most precise examples of sound localization
- can hunt only by hearing
- special structure of head makes 110-degree hearing possible
- asymmetric ears to distinguish the elevation of sounds
- first research on acoustic hunting was performed by Roger S. Payne in 1971

Tyto Alba by Peter Trimming, Creative Commons 2.0
Neural Audio Processing
Neural Network Basics (Biological)

- **neurons:**
  - create a charge
  - release the charge when triggered/excited
  - stronger impulse - higher frequency of charges

- **synapses:**
  - transfer charges from one neuron to another
  - can increase or reduce the excitation of the target node
  - exhibitory connections:
    - connections increasing the excitation
  - inhibitory connections
    - connections decreasing the excitation
Neural Network Basics (Technical)

- first attempt of mathematical description by McCulloch and Pitts in 1943
- linear combination of weighted inputs  
  ⇒ equivalent of synapses
- apply activation function on the combination  
  ⇒ equivalent of neurons

\[ y = f(w_1 x_1 + w_2 x_2 + ... + w_n x_n) \]

- activation function e.g. sigmoid function

\[ f(x) = \frac{1}{1 + e^{-x}} \]
Jeffress Model

- presented by Lloyd A. Jeffress in 1948
- implemented as delay-line algorithm by Liu et al. in 2000
- a model for the ITD part of the brain
- uses $I$ neurons with delayed inputs from left and right ear for each timestep $n$
- includes delay lines to match phase shifts
- phase shift is computed for each frequency band ($m$) by using fast fourier transformation
- the azimuth spectrum is divided into $I$ parts
Jeffress Model Structure

Dual Line structure (Calmes, 2009)
Jeffress Model Notation

- for each node, the signal is delayed by:

\[ \tau_i = \frac{\text{ITD}_{\text{max}}}{2} \sin \left( \frac{i}{I-1} \pi - \frac{\pi}{2} \right) \]

- to shift a signal in the frequency domain, the complex vector is rotated:

\[ X_{L,n}^{(i)}(m) = X_{L,n}(m) e^{-j2\pi f_m \tau_i} \]

- the azimuth sector is selected by the minimal distance of the complex values:

\[ i_n(m) = \text{arg min}_i [\Delta X_n^{(i)}(m)] \]
Jeffress Model Diagram

3D coincidence map (Calmes, 2009)
• a model for the ILD part of the brain
  (Spence & Pearson, 1989)
• simulates different parts of the barn owl brain
  • NA - frequency filtered signal intensity
    (nucleus angularis)
  • VLVp - sigmoidal shaping of the intensity
    (nucleus ventralis lemnisci lateralis, pars anterior)
  • ICc - peaked response curves determining the ILD sector
    (central nucleus of the inferior colliculus)
• parameters tuned in a way to achieve similar results as the barn owl
Spence & Pearson - Nodes

• each neural node has a predefined activation function
  • equal for every node
  • values determined by research on the barn owl
• voltage $v$ and activity $a$ determined by inputs $g$:

$$v = \frac{g_e \cdot v_e + g_i \cdot v_i + g_l \cdot v_l}{g_e + g_i + g_l}$$

with $e =$ excitatory, $i =$ inhibitory and $l =$ leakage

$$a = \frac{1}{1 + e^{ln(s) \cdot (v - v_t)}}$$

with $s$ determining the steepness of the sigmoidal slope
neural network structure of the implemented Spence & Pearson model
• setting $v_e = 0$, $v_i = -90$, $v_l = -65$ and $g_l = 1$

• achieves similar peak responses as the internal brain structure of the barn owl

• activation function parameters may be randomized

• most active ICc node determines the sound direction
Sound Localization Setup

- combine Dual-Line/Jeffress model with Spence & Pearson model
- select most active nodes from both models
- assign nodes to sectors regarding azimuth and elevation by testing

*ITD/ILD contour lines of simple two-microphone setup (Calmes, 2009)*
Artifical Owl Ruff Localization System
Artificial Owl Ruff

Aim:

- expand the azimuth spectrum above 90 degrees
- make the left ear more sensitive for higher elevated sounds
- make the right ear more sensitive for lower elevated sounds
- achieve frequency distortion with a custom HRTF

*artificial owl ruff setups (Calmes, 2009)*
ITD/ILD contour lines of artificial owl ruff setup

(Calmes, 2009)
Effects of the Artificial Owl Ruff

- achieved to expand the azimuth range above 90 degree
- achieved to focus the ILD part on measuring elevation
- did not achieve to benefit from a custom HRTF...
- ...but:
  - azimuth range further increased
  - ILD sensitivity increased in regards to elevation
- possibly the improvement was too noisy to improve the localization
Effect of an Artificial Head to Human Acoustic Perception
• binaural listening demonstration
Conclusion
Conclusion

• biological inspired neural methods enhance sound localization systems:
  • ITD part: Jeffress model
  • ILD part: Spence & Pearson model

• artificial microphone setups inspired by the barn owl enhance sound localization

• artificial structures have an important effect on acoustic perception
  ⇒ for localization systems as well as humans
thank you for your attention!
Biologically Inspired Binaural Sound Source Localization and Tracking for Mobile Robots, Lauent Calmes PhD thesis at I5 chair of the RWTH, 2009

Biologically Inspired Binaural Sound Localization using Interaural Level Differences, Daniel Peger diploma thesis at I5 chair of the RWTH, 2005