An Extreme Value Theory Model of Cross-Modal Sensory Information Integration in Modulation of Vertebrate Visual System Functions

Sreya Banerjee[1], Walter J. Scheirer[1], Lei Li[2,*]

1 Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, Indiana, USA.
2 Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA

*** li.78@nd.edu**

Abstract

We propose a computational model of vision that describes the relationship between olfactory input and visual sensitivity of zebrafish based on the principles of the statistical extreme value theory. This relationship is mediated through centrifugal input which makes possible the physical connections between the neural retina and olfactory bulb. Motivation for using extreme value theory stems from physiological evidence suggesting that extremes and not the mean of the cell responses direct cellular activity in the vertebrate brain. We argue that the visual system, as measured by retinal ganglion cell responses in spikes/sec, follows an extreme value process for sensory integration and the increase in visual sensitivity from the olfactory input can be better modeled using extreme value distributions. As zebrafish maintains high evolutionary proximity to mammals, our model can be extended to other vertebrates as well.

Introduction

The brain perceives the external world through an integration of stimuli received from different sensory modalities like vision, olfaction, and audition via the centrifugal pathway. A recent study taking inspiration from Cajal's original work on brain mapping [17] describes the current knowledge of the centrifugal olfactory and visual pathways in mammalian species as being incomplete. While, for instance, the signaling pathways mediating brain feedback in human olfaction have been characterized, the origins and effects of signals to visual system functions remain to be examined. In this work, we seek to understand the modulation of the circuits between sensory modalities. A crucial observation, yielding from our own work, points to how due to olfacto-visual sensory integration, measures of visual performance or behavior in response to multi-sensory input are enhanced, when a stimulus in one modality is ambiguous or undetermined. In fact, in all vertebrate species (e.g., teleost, reptiles, birds, rodents, primates) examined thus far, the retina receives brain feedback through the centrifugal visual pathways [16, 18, 32]. Depending on the species under consideration, the centrifugal pathways may originate from different parts of brain, such as the pre-tectal cortex, isthmo-optic nucleus, thalamus, or olfactory bulb.

In zebrafish (*Danio rerio*), the olfacto-retinal centrifugal (ORC) pathway originates from terminalis neurons (TNs) in the olfactory bulb (OB) and terminates in retina.

TNs (Fig. 1A) synthesize gonadotropin-releasing hormone (GnRH) as a major neurotransmitter. In retina, TN fibers synapse with dopaminergic interplexiform cells (DA-IPCs), retinal ganglion cells (RGCs), and possibly other retinal cell types. Insights from relatively recent research [21, 27] has shown that the function of the ORC pathway is regulated by the olfactory input. TN input alters GnRH signaling transduction and decreases dopamine release in the retina, thereby increasing outer retinal sensitivity and inner retinal activity (e.g., firing of ganglion cells). Together, the olfactory input amplifies behavioral visual sensitivity [30].

Zebrafish maintain high evolutionary proximity to mammals, and their retinas share great similarities to humans (e.g., structure, cellular organization, neural circuitry and signaling transmission) [26, 47]. While much progress has been made to understand the anatomy of cross-modal circuitry in zebrafish, our knowledge of the underlying regulatory mechanism and physiological roles of centrifugal input to the retina is still in its nascent stage. Interestingly, Huang et al. [21] demonstrate how the visual sensitivity in zebrafish is increased in the presence of olfactory signals whereas disrupting the ORC pathway impairs visual function. An important observation found in that work reveals the importance of olfactory signals for vision. According to Huang et al., under normal conditions the minimum threshold light intensity to invoke a retinal ganglion cell response (measured in spikes/sec) in a dark-adapted zebrafish embryo may decrease 1-2 logs unit after olfactory stimulation. This demonstrates the sometimes dramatic impact of olfactory signals on vision.

Such a sudden gain in visual sensitivity through olfactory stimulation is an intriguing target for a computational model. We argue that visual sensitivity follows the statistical Extreme Value Theory (EVT). The mean visual sensitivity does not clearly explain the increased sensitivity due to olfactory signals since that scenario is able to sense a stimulus that is an extreme aberration from the norm, i.e., retinal ganglion cell responses without any olfactory stimulation. EVT lays solid groundwork for modeling as it is independent of the underlying distribution of data (cell responses) and is only applicable to the tails of the distribution (the extremes) such that samples which have the least or no possible probability, are distinguished, providing greater discrimination while requiring few statistical assumptions.

At a deeper level, one can ask the following question: is there a theoretical justification for using EVT for neural modeling? Our key insight is that the characterization of the firing behavior of a neuron as repeated integration/thresholding, within a circuit suggests positive answers to these questions. Neurons are generally modeled as an electro-chemical process integrating input (ions) and eventually crossing a threshold whereby they "fire" and release ions. We posit that this inherently leads to an EVT-based model because the distribution of samples that exceed a threshold T likely yields an extreme value distribution (EVD). If all neurons use a fixed threshold T, the inputs to subsequent neurons in the circuit must follow an EVD, with each neuron integrating data from such a distribution and thresholding it. Thus, EVT can provide a plausible consistent multi-layer neuron model.

Beyond the merits of cultivating a better understanding of the operation of cross-modal sensory information integration in vertebrates, there is the possibility that an accurate computational model for this phenomenon could translate into a general algorithm for pattern recognition tasks in computer science. A direct application of this method lies in the development of novel information fusion algorithms that leverage inputs from multiple sensory modalities, i.e, vision and audition [33]. Another practical application is the invention of innovative sensors capable of detecting changes in the environment and then re-configuring on the fly to change operational parameters and power consumption requirements. Currently, sensors are typically designed to sense a single type of physical property such as temperature, pressure or radiation, motion or

proximity. But with a biologically-consistent model they could be remodelled to use multiple observations from the environment for more agile operation. The work presented in this article is in this spirit of leveraging biological observations to forward engineer algorithms that can operate in a general context.

In the following sections, we provide a detailed explanation of our work. Section 2 describes the single unit cell recording procedure from which our analysis is derived and the definition of EVT from which the proposed model is based. Section 3 goes on to describe the exact specification of that model. Section 4 describes our experiments and Section 5 presents the corresponding results. Finally Section 6 concludes by putting this research into a larger biological and computational context.

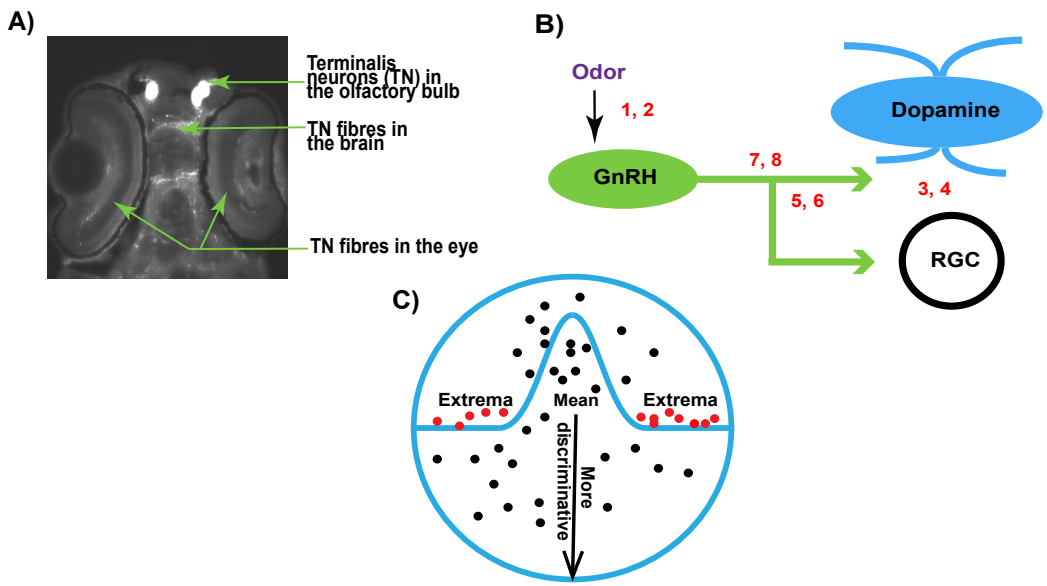

Figure 1. **(A)** A fluorescent image of a zebrafish (dorsal view, anterior is up) showing the terminalis neurons and axons **(B)** Schematic of the experimental setup for RGC recordings in response to olfactory and TN stimulation. The numbers correspond with the following conditions: 1,2 - sham or odor stimulation; 3,4 - activation or inhibition of dopamine receptors; 5,6 - activation or inhibition of GnRH receptors; 7,8 - manipulation of dopamine and GnRH receptors. **(C)** An overview of EVT. Prior work [14, 25, 46] suggests that it may be the extremes (red dots), and not the mean (black dots near the center of the circle), that produce strong responses in the brain.

Materials and Methods

In this section, we explain the methods we use that are crucial for understanding our computational model of cross-modal sensory information integration. This includes the physical experiments that were conducted to collect the source data, as well as the formal elements of EVT.

Single-Unit Recordings and Odor Stimulation

This research builds upon the previous work of Huang et al. [21] (see Fig. 1B). For electrophysiological recordings, zebrafish were anaesthetized with 0.04% 3-amino benzoic acid and immobilized by intraperitoneal injections of $3 - 5$ μl of 0.5 mg ml - 1

gallamine triethiodide dissolved in phosphate-buffered saline (PBS), and then placed on a wet sponge with most of the body covered by a wet paper towel. A slow stream of system water (distilled water with ocean salt added, 3 g gal-1, pH 7.0) was directed into the mouth to keep the fish oxygenized. The eye was slightly pulled out of its socket and held in place by glass rods, thus exposing the optic nerve. Single-unit RGC responses (determined by the spike waveform) were recorded from the optic nerve by using a Tungsten microelectrode (resistance, 5–10 MΩ). Electrical signals were filtered with a band pass filter between 30 and 3000 Hz.

Each fish was dark adapted for 30 min before the first RGC recording was made. The light stimuli (full-field dim white light, generated by a halogen bulb) were directed to the fish eye via a mirror system. The intensity of the unattenuated light beam (logI= 0) measured in front of the fish eye was 670 W cm–2 (Optical Power Meter, UDT Instruments, MD, USA). To determine the threshold, the light intensity was first set at below threshold level (e.g., logI= 6.0) and then increased by 0.5 log-unit steps until the first light-evoked RGC responses were recorded (criteria, 20% above or below the rate of spontaneous firing). This light intensity was noted as the threshold. For each recording, 10 stimuli (600 ms flashes) were delivered at 3 s intervals.

Odors (methionine, 0.5, 2, and 5 mM; total 8 – 10 μl per stimulation) were delivered to the nostril through a glass pipette. The light threshold required to evoke RGC responses was measured before the application of methionine, and was measured again within 10 s following the application of methionine. Thereafter, the measurement was repeated at 1 min intervals for 10 min.

Extreme Value Theory

The extreme value theorem [6] that underpins EVT (Fig. 1C) is very similar to the central limit theorem [23]. Both theorems involve limiting behaviors of distributions of independent and identically distributed random variables as n, the number of random variables, tends to ∞. However while the central limit theorem is concerned with the behavior of entire distributions of random variables, the extreme value theorem only applies to the random variables at the tails of those distributions.

To state this difference precisely, if x_1, x_2, ... , x_n represent the i.i.d. random variables from a distribution, then the central limit theorem describes the limiting behavior of x_1, x_2, ... , x_n while the extreme value theorem describes the limiting behavior of the extremes: $\max(x_1, x_2, \ldots, x_n)$ or $\min(x_1, x_2, \ldots, x_n)$ [6]. It encompasses a number of distributions that apply to extrema.

An extreme value distribution is a limiting model for the maximums and minimums of a dataset. A limiting distribution simply models how large (or small) the data to be modelled will probably get. It is widely used in applications where we are not only interested in estimating the average but also estimating the maximum or minimum [5,15,48,49]. For example, when designing a dam, engineers might not only be interested in the average yearly flood which foretells the amount of water to be stored in the reservoir, but also in the maximum flood, the maximum intensity of earthquakes in the region during the past decade, or maximum strength of concrete to be used in building the dam to avert the possibility of a disaster and for risk management. Castillo et al. [5] list a number of applications where extreme value distribution can be used.

Now that the preliminaries have been covered, we can formally define an extreme value theorem [12]:

Let (s_1, s_2, ... , s_n) be a sequence of independent and identically distributed samples and let $M_n = max(s_1, s_2, \ldots, s_n)$. If a sequence of pairs of real numbers (a_n, b_n)

exists such that each $a_n > 0$ and

$$\lim_{x\to\infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F(x) \tag{1}$$

then if $F(x)$ is a non-degenerate distribution function, it belongs to one of three extreme value distributions: Gumbel, Frechet or Reverse Weibull.

In contrast to the Gumbel or Frechet distributions which are used for unbounded data, the Weibull distribution applies to data that are bounded from below and when the shape (k) and scale (λ) parameters are positive (the Reverse Weibull is simply the opposite of the Weibull's non-degenerate distribution function). Moreover, the Weibull is used for modeling minima. In order to use it for modeling data that fall in the upper tail of distribution, a minor adjustment needs to be made by flipping data based on the maximum value and then applying Weibull distribution. The probability distribution function of the two- parameter Weibull distribution is given as:

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}\frac{x}{\lambda}^{(k-1)} e^{(\frac{x}{\lambda})^k}, & x \geq 1 \\ 0, & x < 0 \end{cases} \tag{2}$$

Note that there are other types of extreme value theorems one can make use of, such as the Pickands-Balkema-de Haam Theorem [36]. We limit ourselves to the theorem in Eq. 1 in this work for the modeling of explicit tail data, but we will invoke the Pareto distribution, which is derived from the Pickands-Balkema-de Haam Theorem, in the modeling of the overall distribution. This is described below in the next section.

A Model for Cross-Modal Sensory Information Integration

Now that the relevant background has been introduced, we formally define our computational model for cross modal sensory information integration (Fig.2). It is motivated by the following hypothesis: *The tuning curves for RGC responses with and without olfactory signals are different. The extreme values lying near the tails of the distributions underlying those curves contribute to the determination of the visual sensitivity of zebrafish and should not be discarded as outliers.*

Our intuition behind this hypothesis is that the characterization of the firing behavior of a neuron as repeated integration/thresholding within a circuit resembles an EVT distribution. Neurons are generally modeled as an electro-chemical process integrating input (ions) and eventually crossing a threshold whereby they "fire" and release ions. This phenomenon inherently leads to an EVT-based model because the distribution of samples that exceed a threshold T likely yields an EVD. If all neurons use a fixed threshold T, the inputs to subsequent neurons in the circuit must follow an EVD, with each neuron integrating data from such a distribution and thresholding it.

The single unit recordings that we used for our experiments can be regarded as samples from a large population. One way to infer more about the population statistics is to extrapolate from the available samples by fitting distributions to them and sampling additional data. However, fitting data to a known distribution can be difficult because of limited sample sizes, leaving one to make a "best guess" based on prior information about the behavior of large sample statistics. As an alternative, we use Bayesian Information Criterion (BIC) or Schwarz criterion [41] to select the best distribution among a finite set of candidate models.

Based on the BIC method, the distribution that fit accurately to the overall RGC response data with olfactory stimulation was the Gaussian distribution. For the overall RGC data without olfactory stimulation (or control responses), the generalized Pareto

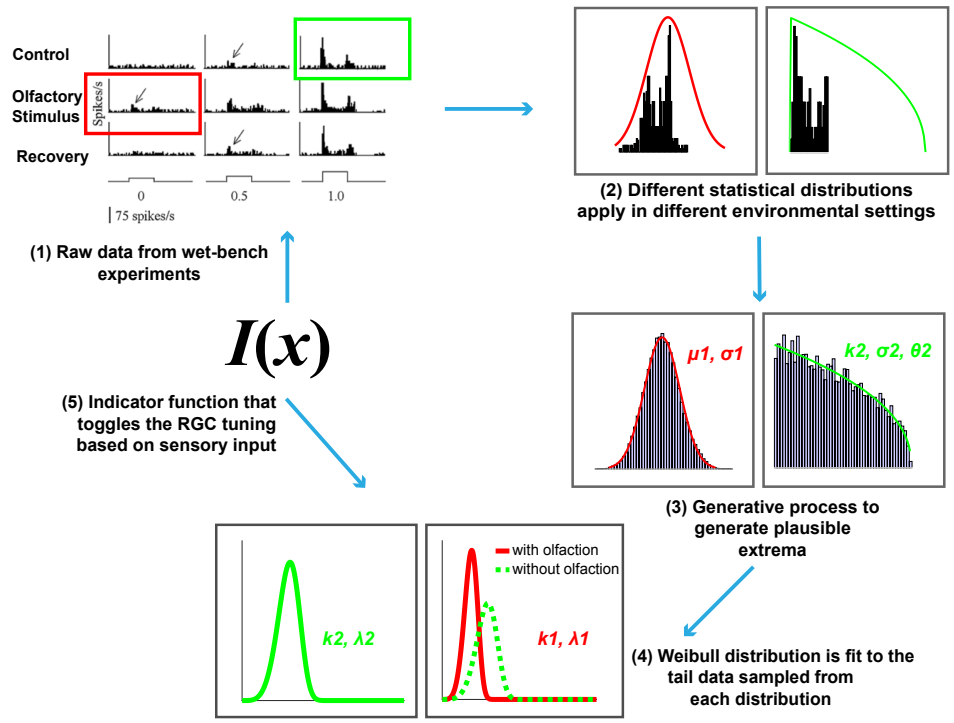

Figure 2. An overview of the proposed computational model of cross-modal sensory information integration. The red curve represents the RGC responses when an olfactory stimulus is present and the green curve represents the responses when there is no olfactory stimulus present. The first step represents the data collection effort from the wet-bench experiments. The second and third steps include fitting distributions to the data collected in order to draw samples for further processing. The fourth step represents fitting a Weibull distribution in order to model the underlying difference between visual sensitivity when olfactory input (i.e., additional sensory information) is present as opposed to when it is not. The final step is identifying an indicator function, *I(x)* that toggles between the two distributions based on the sensory input received.

distribution was the best fit. Interestingly, this distribution is considered to be in the EVT family. The population statistics (mean μ and variance σ^2) of the RGC response data with olfactory signal is found as the unbiased estimates of the distribution parameters and are given by the following equations:

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} \tag{3}$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2 \text{ for all i} \in (1, 2, 3, ..., n) \tag{4}$$

The probability density function for the generalized Pareto distribution with shape parameter k, scale parameter σ and threshold parameter θ is given by the following equation:

$$y = f(x \mid k, \sigma, \theta) = \left(\frac{1}{\sigma}\right) \left\{1 + k\frac{(x-\theta)}{\sigma}\right\}^{-1-\frac{1}{k}} \tag{5}$$

We used maximum likelihood to estimate the parameters k and σ from the two-parameter generalized Pareto distribution by fitting RGC responses without olfaction[1]

Having access to a model of the entire population facilitates generative sampling, which in turn allows for better tail modeling, and support for heightened visual sensitivity under certain conditions. Such generative processes in the brain may be responsible for a number of different phenomena, as they facilitate generalization in learning from limited sampling [37]. We use random sampling and the Metropolis-Hastings algorithm, a Markov chain Monte Carlo sampling method [19] to generate in total $100,000$ simulated RGC responses with and without olfaction respectively. The maximum (or the minimum) RGC response values within these samples follow an EVD. For our analysis, we concentrate only on the maximum RGC responses from the distributions described above because the lowest possible RGC response can be 0 spikes per second, indicating no response. Since the RGC responses (both with and without olfactory signals) can be assumed to be i.i.d samples from continuous distributions that are bounded from below, the Weibull distribution is the correct choice for modelling them. We expect the Weibull cumulative distribution curves (CDFs) for RGC responses with and without olfaction to be widely separated and the threshold RGC response value for an olfactory signal to shift sensitivity leftward (see Fig. 3 for an example), indicating that the cells are now more sensitive. This effect, replicated within the model, would confirm in a more rigorous sense that the presence of olfactory signals increases the fish's sensitivity towards its surrounding and almost endows it with night vision that would be otherwise impossible in absence of those signals.

This process is analogous to the super-additivity phenomenon in the multi-sensory superior colliculus of higher-order organisms like mammals, where the presence of two weak sensory signals from the environment enhances the animal's neural response towards that environment [20]. The RGC threshold value represents an average RGC response for visual sensitivity, which changes throughout an animal's entire life-cycle as it adapts to an ever-changing environment. However, the threshold varies (decreases or increases) in the presence or absence of a sensory stimulus other than visual input. This leads us to the possibility of the existence of some decision making mechanism in the fish's brain that toggles between two different distributions to adjust the tuning of the RGCs based on sensory input. Mathematically, this decision making procedure can be emulated as an indicator function $I(x)$. If θ represents the parameters of an RGC distribution, i.e., the prior information available for RGC responses with or without olfactory signals and x represents a new RGC response due to a stimulus from the environment such that $x \in R^n$ (here $n = 22$, as we have 22 data points from the wet-bench experiments of Huang et al. For further explanation, see Section. 4), then the indicator function $I(x)$ can be represented as:

$$I\left(x \mid \theta\right) = \begin{cases} 1, \text{ if olfactory signal is present} \\ 0, \text{ otherwise} \end{cases} \tag{6}$$

We speculate that the actual neural computation for the overall phenomenon is far more complex and is not restricted to just two modalities. However, given the recordings available for this study, we limit our model to just one particular circuit.

[1]For finding the maximum likelihood estimates of generalized Pareto distribution, we used the matlab function *gpfit* which only returns the estimates of shape, k and scale parameters, σ of a two-parameter generalized Pareto distribution. The function *makedist* was then used to create a probability distribution object from where samples were taken, using the parameters, k and σ estimated via *gpfit*. If θ is not specified in *makedist*, matlab assumes it to be 1.

Choices for an Indicator Function

For the indicator function, we address the following problem: given a set of vectors representing RGC responses in spikes/s with and without olfactory signals, is it possible for an indicator function to identify whether a new RGC response has been triggered after an olfactory signal or not? Our intuition behind using an indicator function is that such a process exists in some capacity in the brain where the presence of one signal enhances the other signal, thereby eliciting responses much different from the situation when the signal is not present.

In essence, this task can be formulated as a binary classification problem with two possible outcomes: presence or absence of olfactory signals. Ideally, any discrimitative supervised learning methods can easily solve the problem. For our analysis, we examine the utility of support vector machines and an artificial neural network which, to some extent, mimics the functions of a biological neuron and is closer to the mechanism that the brain uses to process such signals.

Support Vector Machine

The Support Vector Machines (SVM) is a supervised learning approach that is widely used for classification and regression analysis. Since our data is numeric and high-dimensional, SVM [9] is a natural choice as it has been found to be extremely efficient in high-dimensional spaces for large-scale classification problems. SVMs use a subset of training points in the decision function, which form the "support vectors" that define the decision boundary between classes. As a consequence, it has been found to be memory efficient and has faster execution times if the data is normalized. For analysis, we assumed our data to be linearly separable and used a linear SVM kernel. We normalize our data using min-max normalization.

An SVM model with a set of labelled training data tries to find an optimal hyperplane for classifying new samples based on some constraints. Given a training dataset, $D = (x_i, y_i)$ of size m with $x_i = (x_1, x_2, ..., x_n)$, an n-dimensional feature/attribute vector and label, y_i = -1 or +1, formally the SVM classifier can be defined as a quadratic optimization problem solving the following equation:

$$min \, \|w\|^2 \text{ s.t } y_i(w^T x_i + b) \geq 1 \text{ for all } i \tag{7}$$

where $w = (w_1, w_2, ..., w_n)$ is a weight vector and b is the bias.

An important consideration when designing a SVM model is the parameter C that dictates the trade-off between having a wide margin and correctly classifying training data.

$$min \, \|w\|^2 + C \sum_{1}^{m} \xi_i \;\; \text{s.t} \;\; y_i(w^T x_i + b) \geq (1 - \xi_i), \, \xi_i \geq 0 \text{ for all } i \tag{8}$$

A larger value of C implies a smaller number of mis-classified training samples and is prone to overfitting.

Artificial Neural Network

We also consider a multi layer perceptron (MLP) neural network as the indicator function. Similar to SVM, MLP is a supervised learning algorithm that learns a non-linear mapping from input, $x \in R^n$ where n represents the number of dimensions, to $y \in R^m$ where m can be any number $m < n$, depending on the number of classes in the training dataset. However, unlike SVMs, a simple MLP includes one or more hidden layers consisting of artificial neurons. The hidden layers act as feature detectors and gradually discover the salient features of the training data through

backpropagation [38, 50]. Each neuron includes a non-linear and differential activation function and is connected to every neuron in the previous layer exhibiting a high degree of connectivity between layers. As a result, due to the distributed nature of non-linearities, the learning process is difficult to visualize. However, neural networks are usually assumed to be non-parametric functions, i.e., they can be used as function approximators without having any prior information about the distribution of input or training dataset and hence are well suited to represent the indicator function. If x represents a p-dimensional input vector such that $x = (x_1, x_2, x_3, ..., x_p)$ with $y = (+1, -1)$ as labels and $g : R \mapsto R$ as the activation function, then the equation of a single neuron is given by:

$$y = g\left(b + \sum_{i=1}^{p} w_i x_i\right) \tag{9}$$

where $w = [w_1, w_2, w_3 ... w_p]$ represents the weights learned through backpropagation.

Experiments

One of the basic steps in building a computational model of this nature is to attempt to define the underlying distribution of the data one is trying to explore. We use the data from a study by Huang et al. [21] for our analysis. The data consists of single unit RGC responses measured in spikes/sec before and after olfactory stimulation under varying light intensity (see Fig. 2 from Huang et al.). In terms of raw data organization, it is primarily a histogram with the x-axis representing the visual sensitivity of fish binned into approximately 22 positions representing a timestamp and their corresponding frequency measured in spikes/sec in the y-axis. Under normal condition, the minimum threshold light intensity to invoke a retinal ganglion cell response in a dark-adapted zebrafish embryo is 10^{-5}. However, with olfactory stimulation with methionine, the threshold light intensity decreases to 10^{-6}. We calculated the minimum RGC response threshold to be at 75 spikes/sec. Hence, the data can be separated into two parts: one with olfactory stimulus and the other without it. In total, there were 22 RGC responses with olfactory stimulus and 29 without olfactory stimulation.

Experiment 1. One of our initial experiments was to check whether the raw data we collected from the experiments confirms our hypothesis that the EVT can be applied to build an accurate model. We posit that since the RGC responses with olfactory stimulation represent extreme aberration from the baseline and are non-negative integers, the Weibull distribution is a perfect candidate for modeling our data. But how differently does our data fit with the Weibull distribution versus a central tendency model like the Gaussian distribution? We explore this in our first experiment by comparing the CDFs of the Weibull and Gaussian distributions against our data.

Tests on normality and synthetic data generation. Using the data collected from experiments as a basis, we simulated an expansive data space by fitting distributions over the original data. The goal was to generate as much data as possible for statistical inference. However, in order to fit distributions to generate more samples from the existing data, we need to make some assumptions about the underlying distribution of the existing data. Initially, we assumed that the distribution of RGC responses in zebrafish through its entire lifecycle is Gaussian. To confirm our assumption, we performed two tests of normality: a Kolmogorov Smirnov test [31] and a Lilliefors test [7, 28, 29]. Based on the results of the normality tests, we fit Gaussian distribution to our data with olfactory stimulus. However, the distribution of data without olfactory stimulus was found out to be non-Gaussian. As described above, we used Bayesian Information criteria to find the best fit, and the distribution for the data was found out

to be Generalized Pareto. Afterwards, we generated $100,000$ non-negative samples of RGC responses from the respective distributions for further analysis.

Experiment 2. Our second experiment was conducted to verify whether or not the samples generated after fitting distributions to data with olfactory stimulus and data without olfactory stimulus conform to the our hypothesis of EVT and how olfactory signals affect visual sensitivity in fishes throughout its' entire lifecycle. In order to verify this, we fit a Weibull distribution to the top n RGC responses to understand how the the curves vary when olfactory input is present as opposed to when it is not. The value n was selected via empirical observation. Sampling methods used were: random sampling and sampling using Markov Chain Monte Carlo method. Since EVDs like the Weibull only apply to samples at the tails of distributions, it is independent of the underlying distribution of the data as a whole. Hence, irrespective of the overall data distribution and sampling process, the results of Experiment 2 for Weibull distributions for the top n responses should ideally be similar to Experiment 1. We expect the Weibull cumulative distributions for data with and without olfaction to be widely separated, with the curve for data with olfaction to shift leftward, giving some probability to RGC responses below the threshold which is practically improbable under normal situations (without olfaction).

Experiment 3. Additionally, we wanted to corroborate whether we can define a deterministic indicator function such that given some RGC response it is possible for the function to identify if an olfactory stimulus is present or not. In essence, this task becomes a binary classification problem where the presence of olfactory signals can be labelled as 1 and the absence of it as 0. As described above, we use a linear SVM and a multi-layer perceptron as our binary classifier. We assume that each RGC response can be represented as a vector of 22 points. This is because we could only retrieve a minimum of 22 points from wet-bench experiments of Huang et al. for data with olfactory stimulus. We use the $100,000$ samples we collected for each dataset (with olfactory stimulus and without olfactory stimulus) dividing it into 80-percent training and 20-percent testing partitions.

In summary, the entire modeling effort is encapsulated in the following steps (also reflected in Fig. 2):

1. **Data collection and representation.** This step consists of collecting and representing data based on the wet-bench experiments for control (without any stimulation) and experimental zebrafish (with olfactory stimulation) as a histogram and collecting the statistics for further analysis.

2. **Experiment 1.** This step consists of an experiment to evaluate our hypothesis on EVT with the raw data collected in step 1. We fit Gaussian distributions (to the entire data with and without olfaction individually) and Weibull distributions (to top-n RGC responses from the two datasets). The value, n was selected via empirical observation.

3. **Tests on normality and synthetic data generation** Assuming that the distribution of RGC responses in a zebrafish through its entire lifecycle is normal, testing the underlying distribution of data for normality and subsequently fitting distributions to generate one hundred thousand synthetic samples. The data with olfactory stimulus follows Gaussian distribution, whereas the underlying distribution for data without olfactory signal is generalized Pareto.

4. **Experiment 2.** Similar to Experiment 1 but instead uses one hundred thousand generated samples and only Weibull distributions on the top- n samples generated to examine how olfactory signals influence visual sensitivity as reflected by the cumulative distribution curves for the two dataset. The value, n was selected through empirical observation.

5. **Experiment 3.** Identifying an indicator function $I(x)$ that can help to distinguish when an olfactory stimulus is present and when it is not. This function can be a deterministic binary classifier like a linear SVM or a simple artificial neural network like a multilayer perceptron.

Results

Experiment 1. Fig. 3 depicts the result of Experiment 1 which was conducted to examine the difference between central tendency modeling and EVT modeling. The data for this experiment was what was directly collected from the wet-bench experiments for both control (without olfaction) and experimental (with olfaction) zebrafish.

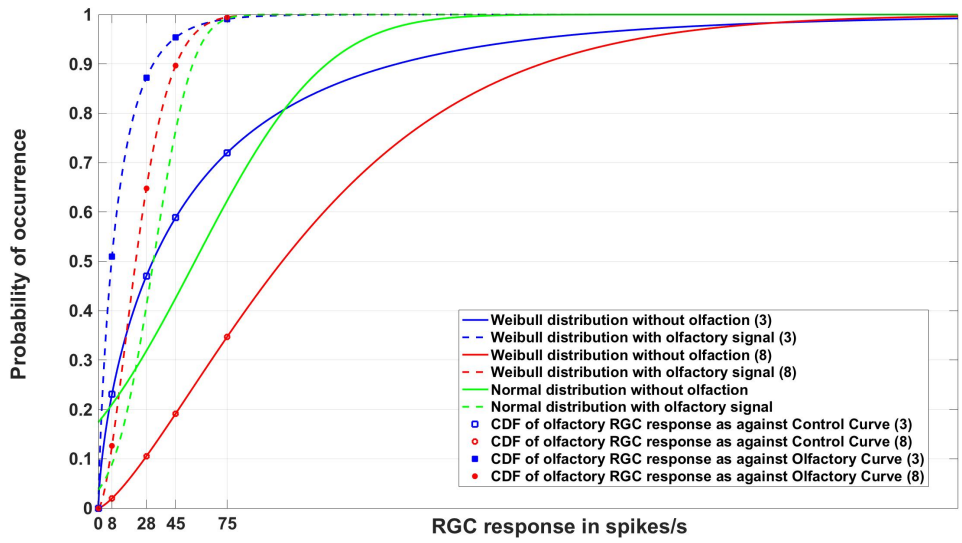

Figure 3. Experiment 1. Cumulative Distribution Functions for zebrafish with / without olfactory stimulation at light intensity $10^{-}5$ **&** $10^{-}6$ **respectively.**. The curves depict the difference between central tendency modeling (green) and EVT (red and blue). As can be seen, tuning becomes more sensitive when the Weibull distribution is used. Number of maximal RGC responses taken are 3 and 8 (as shown in parentheses). Best viewed in color.

As can be seen in the figure, with olfactory stimulation the visual sensitivity in zebrafish shifts leftward, making the RGC responses below the normal threshold of 75 spikes/s probable, as indicated by the physiological experiments of Huang et al. [21]. Moreover, if we look closely, the Weibull distribution is a better fit to the data because the RGC responses with olfactory stimulation represent the outliers or extreme responses as opposed to RGC responses without any stimulation. If we fix our attention at the threshold RGC response at 75 spikes/s, the Weibull curves provide a better explanation for getting an RGC response below 75 spikes/sec for olfactory stimulation in comparison to the normal distribution which almost makes those values improbable. In other words, the tuning becomes more sensitive if we use the Weibull distribution. We plotted the curves by varying n of the top-n RGC responses through empirical observation. The tuning becomes more sensitive as n becomes smaller.

Tests on normality and synthetic data generation. Using the data collected from experimentation as a statistical basis, we simulated an expansive data space by fitting

distribution over the original data. Our assumption of a Gaussian distribution for the data with olfactory stimulus was found out to be true via the normality tests mentioned above. However, the results of the normality test for data without olfactory stimulus was negative and using Bayesian information criteria, we found the data distribution to be generalized Pareto. We collected 100,000 non-negative samples or RGC responses for fitting a Weibull distribution to the top n RGC responses in order to understand how the curves vary when olfactory input is present (i.e., when the overall distribution is Gaussian) as opposed to when it is not (i.e., when the overall distribution is Pareto). The sampling methods used are: random sampling and sampling using the Markov Chain Monte Carlo method.

Experiment 2. Figs. 4 and 5 represent both methods, respectively. We get similar results for both sampling procedures. An important observation to note here is that tuning is always more sensitive when olfactory stimulus is present. The cumulative distribution curves for data with and without olfactory stimulation respectively are widely separated and the width of separation increases as we increase n. This reflects how the visual sensitivity threshold changes throughout the fish's lifecycle as it is exposed to an ever-changing environment and acquires new RGC responses as a result of its interaction with the environment.

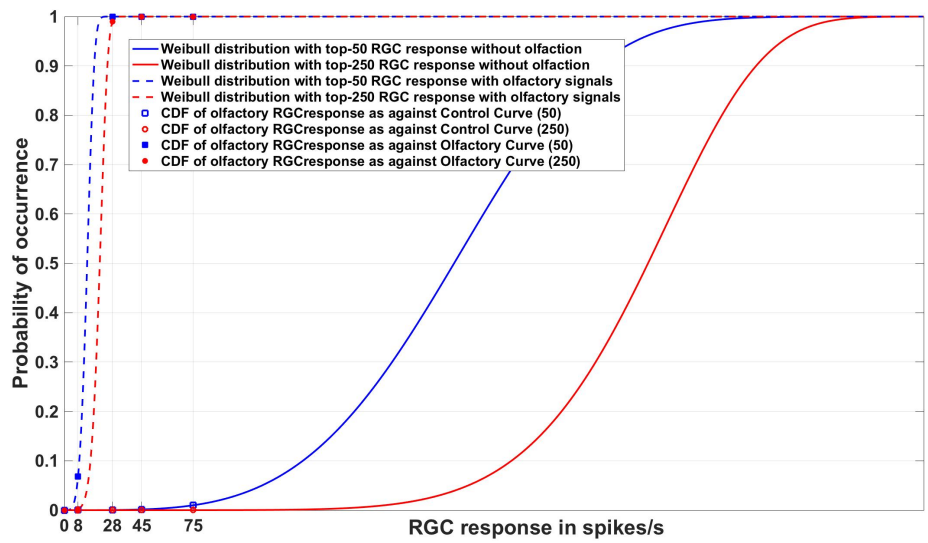

Figure 4. Experiment 2. Cumulative Distribution Functions for zebrafish with and without olfactory stimulation at light intensity $10^{-}5$ **and** $10^{-}6$ **respectively with simulated points through random sampling.**. As can be seen, tuning is most sensitive when an olfactory stimulus is involved. Best viewed in color.

Experiment 3. With respect to testing a possible indicator function, $I(x)$, we use a linear binary SVM classifier and train it using $80,000$ training samples and $20,000$ testing samples. We assume each training sample has 22-dimensions, i.e., 22 RGC responses constitute a single sample. With random sampling, we get a testing accuracy of 95.5 (+/- 0.002) percent but with MCMC sampling, it decreases to 94 (+/- 0.0018) percent. With a multi-layer perceptron model, the accuracy drops to 95.25 (+/- 0.007) percent using the same training-testing split through MCMC sampling. The success of this experiment highlights the importance of the effect of olfactory signals on visual

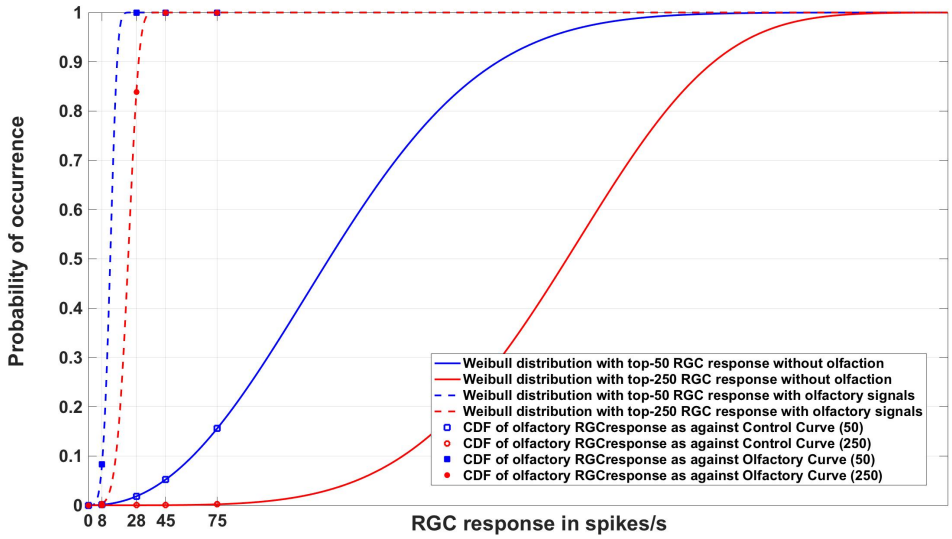

Figure 5. Experiment 2. Cumulative Distribution Functions for zebrafish with and without olfactory stimulation at light intensity $10^{-}5$ and $10^{-}6$ respectively with simulated points through MCMC.. The result is very similar to random sampling. Here also, the tuning is more sensitive when an olfactory stimulus is involved. Best viewed in color.

sensitivity. It provides an explanatory mechanism for to achieve different the RGC responses when an olfactory stimulus is present as opposed to when it is not as pointed out by Huang et al. [21].

Discussion

As vertebrates evolved over centuries, sensory organs adapted with the ever-changing environment. In many vertebrate species, at any given time the brain integrates and processes multi-sensory information. In humans, for example, the functions of the olfactory and visual systems are influenced by sensory input from each organ. Most mammals have specialized multimodal neurons in the superior colliculus that are capable of integrating multiple stimuli from the environment and providing a uniform reaction. In lower vertebrates such as fish, however, such advanced mechanisms are absent. In zebrafish, the integration of sensory information from the olfactory system facilitates signaling transduction in the visual pathway. As a consequence, retinal neural activities such as the firing of retinal ganglion cells are increased. While the cellular mechanisms underlying this olfacto-retinal sensory integration have been well characterized, statistical models that describe the phenomenon at the cellular level have not been described. In this paper, we describe a computational model that supports the research into how the visual system integrates information from other sensory modalities.

The idea of building computational models for multisensory input has been explored previously [1, 2, 10]. When it comes to determining the statistical relationship between sensory responses among different sensory organs, the Bayesian model has been a preferred framework. However, almost all of the existing work focuses on higher vertebrates such as mammals. Angelaki et al. [2] attempted to reconcile the difference between the traditional physiological studies on multisensory integration with

computational and psychological studies using Bayesian inference on the visual-vestibular system for the perception of self-motion in macaques. They describe how the multimodal neurons represent probablistic information defined by multiple stimuli and propose that the special neurons accomplish near optimal cue integration through a linear summation of input signals.

With respect to models of simpler animals, Wessnitzer et al. [51] explore multimodal sensory integration for navigation from the physiological perspective of the insect's nervous system. In zebrafish, using a similar linear model [22] the contribution of different types of cone photoreceptor cells to photopic spectral visual sensitivity was determined. This was done by re-modeling the electroretinographic data recorded from the cornea, which include absorbance spectrum of four types of cone photoreceptor cells (cone cells that are sensitive to ultra-violet light, blue light, green light, and red light, respectively) given as the visual pigment template for the appropriate maximum absorption, neural signals obtained from different cone cell types, relative fraction of the individual cone cells across the retina, and linear gains for each cone type [4]. The model incorporated the first-order cellular and biophysical aspects of cone photoreceptor cells and thereby predicts the second-order physiological functions of cone cell-mediated visual sensitivity. Using this model, linear gains that represent the strength of four different types of cone cell-derived neural signals onto four different inferred cone processes in the whole retina can be assessed.

Turning to extreme value theory, the objective of nearly all extant models in computational neuroscience has been to discard the extreme values located at the tails of distributions as noise and concentrate on the mean or average. However, evidence suggests that extremes, and not means, of cell responses direct activity in the brain. For example, the ability of primates, like macaque monkeys, to identify individual faces can be localized to a group of special neurons that fire in response to specific regions of face [14]. An interesting finding that came out of that study was that neurons were tuned to the geometry of extreme facial features. Previous investigations along this line (and a more recent one) concentrated on how the brain fundamentally adapts itself to the statistics of the sensory world, extracting relevant information from sensory inputs by modeling the distribution of inputs that are encountered by the organism [43, 44]. This lead to the advent of "sparse coding" which attempts to explain how neurons encode sensory information using a small number of active neurons at any given point in time [34]. A direct extension of this work suggests that sparse coding is an all-pervasive phenomenon used by all types of sensory neurons in different modalities across different species [35]. EVT builds upon these concepts but is more specialized.

Much prior work related to EVT modeling has focused on various non-biological applications from trend detection in ground-level ozone [45] to quantifying extreme precipitation levels using generalized Pareto distributions [8]. Other applications of EVT include, but are not limited to, finance, telecommunications and the environment [11], and hydrology [24]. Recent work in computer vision and machine learning has extensively used the concept of EVT [3, 13, 39, 40, 42]. For instance, for biometric verification systems, Shi et al. [42] used the General Pareto Distribution to model the genuine and impostor scores and made a significant observation that the tails of each score distribution contain the most relevant information that helps in defining each distribution considered for prediction and the associated decision boundaries, which are often difficult to model.

Our research extends this theory to multi-sensory inputs through a model that demonstrates strong neural fidelity. With a biologically-consistent information fusion algorithm based on retinal circuits in the zebrafish, we believe that we have access to a better general solution to the problem at hand and many other information processing problems of interest. Here, we develop a neural computation model that simulates the

process of multi-organ sensory integration and predicts the consequence of sensory integration in higher-order brain functions. In contrast to Gaussian modeling, we propose that EVT models of the extrema found in the tails of the data can form a powerful basis for cross-modal sensory information integration, facilitating heightened sensitivity in targeted modalities that have been influenced by a stimulus in the environment. This resulted in the development of a computational EVT-based framework for multi-organ sensory integration in the zebrafish that is not only an explanatory model in neuroscience, but also shows promise for applications in machine learning and neuromorphic systems.

Acknowledgments

The research was funded by the Department of Defense (Army Research Laboratory) under the contract, W911NF-18-1-0292.

Author Contributions

WJS and LL initially conceived of the idea. LL was responsible for conducting the wet-bench experiments and preparing the source data. SB designed, analyzed, implemented the model and wrote the paper. WJS supervised the entire modelling effort.

Data Availability Statement

The datasets analyzed for this study and the codes are released for reproducability and can be found in the [Zebrafish_EVT] [`https://github.com/sbanerj2/Zebrafish_EVT`].

Ethics Approval Statement

An ethical review process was not required for our study. All data used in this article come from a previously published study [21]. All experimental procedures in that paper adhered to the NIH guidelines for animals in research.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. T. J. Anastasio, P. E. Patton, and K. Belkacem-Boussaid. Using bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation*, 12(5):1165–1187, 2000.

2. D. E. Angelaki, Y. Gu, and G. C. DeAngelis. Multisensory integration: psychophysics, neurophysiology, and computation. *Current opinion in neurobiology*, 19(4):452–458, 2009.

3. J. B. Broadwater and R. Chellappa. Adaptive threshold estimation via extreme value theory. *IEEE Transactions on signal processing*, 58(2):490–500, 2010.

4. D. A. Cameron. Mapping absorbance spectra, cone fractions, and neuronal mechanisms to photopic spectral sensitivity in the zebrafish. *Visual neuroscience*, 19(3):365–372, 2002.

5. E. Castillo, A. S. Hadi, N. Balakrishnan, and J.-M. Sarabia. Extreme value and related models with applications in engineering and science. 2005.

6. S. Coles, J. Bawa, L. Trenner, and P. Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

7. W. J. Conover and W. J. Conover. Practical nonparametric statistics. 1980.

8. D. Cooley, D. Nychka, and P. Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840, 2007.

9. C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

10. J. Driver and T. Noesselt. Multisensory interplay reveals crossmodal influences on 'sensory-specific'brain regions, neural responses, and judgments. *Neuron*, 57(1):11–23, 2008.

11. B. Finkenstadt and H. Rootzén. *Extreme values in finance, telecommunications, and the environment*. CRC Press, 2003.

12. R. A. Fisher and L. H. C. Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press, 1928.

13. V. Fragoso, P. Sen, S. Rodriguez, and M. Turk. Evsac: accelerating hypotheses generation by modeling matching scores with extreme value theory. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2472–2479, 2013.

14. W. A. Freiwald, D. Y. Tsao, and M. S. Livingstone. A face feature space in the macaque temporal lobe. *Nature neuroscience*, 12(9):1187–1196, 2009.

15. J. Galambos. Extreme value theory for applications. In *Extreme Value Theory and Applications*, pages 1–14. Springer, 1994.

16. M. J. Gastiner, R. G. Yusupov, R. D. Glickman, and D. W. Marshak. The effects of histamine on rat and monkey retinal ganglion cells. *Visual neuroscience*, 21(06):935–943, 2004.

17. D. H. Gire, D. Restrepo, T. J. Sejnowski, C. Greer, J. A. De Carlos, and L. Lopez-Mascaraque. Temporal processing in the olfactory system: can we see a smell? *Neuron*, 78(3):416–432, 2013.

18. M. R. Harter and C. J. Aine. Brain mechanisms of visual selective attention. *Varieties of attention*, pages 293–321, 1984.

19. W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

20. N. P. Holmes and C. Spence. Multisensory integration: space, time and superadditivity. *Current Biology*, 15(18):R762–R764, 2005.

21. L. Huang, H. Maaswinkel, and L. Li. Olfactoretinal centrifugal input modulates zebrafish retinal ganglion cell activity: a possible role for dopamine-mediated ca2+ signalling pathways. *The Journal of physiology*, 569(3):939–948, 2005.

22. A. Hughes, S. Saszik, J. Bilotta, P. J. DEMARCO, and W. F. PATTERSON. Cone contributions to the photopic spectral sensitivity of the zebrafish erg. *Visual neuroscience*, 15(6):1029–1037, 1998.

23. E. T. Jaynes. *Probability theory: The logic of science*. Cambridge university press, 2003.

24. R. W. Katz, M. B. Parlange, and P. Naveau. Statistics of extremes in hydrology. *Advances in water resources*, 25(8):1287–1304, 2002.

25. D. A. Leopold, I. V. Bondar, and M. A. Giese. Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, 442(7102):572, 2006.

26. L. Li. Zebrafish mutants: behavioral genetic studies of visual system defects. *Developmental dynamics: an official publication of the American Association of Anatomists*, 221(4):365–372, 2001.

27. L. Li and J. E. Dowling. Disruption of the olfactoretinal centrifugal pathway may relate to the visual system defect in night blindness bmutant zebrafish. *Journal of Neuroscience*, 20(5):1883–1892, 2000.

28. H. W. Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318):399–402, 1967.

29. H. W. Lilliefors. On the kolmogorov-smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325):387–389, 1969.

30. H. Maaswinkel and L. Li. Olfactory input increases visual sensitivity in zebrafish: a possible function for the terminal nerve and dopaminergic interplexiform cells. *Journal of Experimental Biology*, 206(13):2201–2209, 2003.

31. F. J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951.

32. G. Mick, H. Cooper, and M. Magnin. Retinal projection to the olfactory tubercle and basal telencephalon in primates. *Journal of Comparative Neurology*, 327(2):205–219, 1993.

33. A. Nagrani, S. Albanie, and A. Zisserman. Seeing voices and hearing faces: Cross-modal biometric matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8427–8436, 2018.

34. B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision research*, 37(23):3311–3325, 1997.

35. B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–487, 2004.

36. J. Pickands III et al. Statistical inference using extreme order statistics. *the Annals of Statistics*, 3(1):119–131, 1975.

37. R. P. Rao, B. A. Olshausen, M. S. Lewicki, M. I. Jordan, and T. G. Dietterich. *Probabilistic models of the brain: Perception and neural function*. MIT press, 2002.

38. D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533, 1986.

39. W. J. Scheirer, L. P. Jain, and T. E. Boult. Probability models for open set recognition. *IEEE transactions on pattern analysis and machine intelligence*, 36(11):2317–2324, 2014.

40. W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boult. Meta-recognition: The theory and practice of recognition score analysis. *IEEE transactions on pattern analysis and machine intelligence*, 33(8):1689–1695, 2011.

41. G. Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

42. Z. Shi, F. Kiefer, J. Schneider, and V. Govindaraju. Modeling biometric systems using the general pareto distribution (gpd). In *SPIE Defense and Security Symposium*, pages 69440O–69440O. International Society for Optics and Photonics, 2008.

43. E. P. Simoncelli. Vision and the statistics of the visual environment. *Current opinion in neurobiology*, 13(2):144–149, 2003.

44. E. P. Simoncelli and B. A. Olshausen. Natural image statistics and neural representation. *Annual review of neuroscience*, 24(1):1193–1216, 2001.

45. R. L. Smith. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statistical Science*, pages 367–377, 1989.

46. J. W. Tanaka. Caricature recognition in a neural network. *Visual Cognition*, 3(4):305–324, 1996.

47. A. M. Vacaru, G. Unlu, M. Spitzner, M. Mione, E. W. Knapik, and K. C. Sadler. In vivo cell biology in zebrafish–providing insights into vertebrate development and disease. *J Cell Sci*, 127(3):485–495, 2014.

48. W. Weibull. A survey of statistical effects in the field of material failure. *Appl. Mech. Rev*, 5(11):449–451, 1952.

49. W. Weibull et al. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 18(3):293–297, 1951.

50. P. J. Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, 1990.

51. J. Wessnitzer and B. Webb. Multimodal sensory integration in insects—towards insect brain control architectures. *Bioinspiration & Biomimetics*, 1(3):63, 2006.