An Empirical Model of Goldfish Escaping Strategy

Ziyue Chu[1], Jinxing Yang[1] and Wen-Chi Yang[1,2]

[1] Neuhelium Co. ltd., China
[2] wenchi.yang@neuhelium.com

Abstract

Animals' escaping behavior can be seen as an economic game between predation risk and foraging cost, which is measured by flight distance in the standard deterministic model. However, this fight distance should not be constant even in a stationary environment because it will also be affected by the predator's parameters. In order to better understand animal's escaping behavior, this paper investigates goldfish escaping strategies on the relevant position against a toy submarine in quantification, and further proposed a probabilistic model over the factors that influence the probability to escape. This research has determined goldfish's hyperparameters in their escaping habits based on recording data and found that the flight distance is not only influenced by the risk and cost in previous research but also the approaching direction between goldfish and predator and approaching speed. The proposed probabilistic model shows that flight distance generally increases with the predator's approaching speed and the extent of approaching direction against the vertical plane. The proposed model is finally cross-validated with the true video clip.

Introduction

Most living creatures survive by foraging resources from other species, where the relation between the two is predator and prey. On the other hand, prey could survive by avoiding being predated when they noticed a potential foraging risk nearby. Therefore, different creatures have evolved escaping awareness and various escaping strategies when they face their predators Broom and Ruxton (2005); Hileman and Brodie Jr (1994); Martín et al. (2009); Miyatake et al. (2008); Marchisin and Anderson (1978). To analyze animal's escaping behavior, some mathematical models have been proposed to describe and predict animal's escaping behavior from the perspective of the economic game between predator and prey Ydenberg and Dill (1986); Martín et al. (2009). The economic game models believe that the prey's decision on escaping is determined and balanced by the predation risk and motion maintaining opportunity Ydenberg and Dill (1986); Martín et al. (2009), and a prey will choose to escape when it feels the culminated risk outranges the potential maintaining opportunity rather than the moment when it notices the existence of a predator. The predation risk mainly comes from the fatal death of being caught, while the motion maintaining opportunity comes from the energy consumption of escaping, more resources obtained, and the lower probability of being noticed by the predator in static states compared to the escaping states. Generally, the risk rises and the opportunity shrinks when the predator gets closer to the prey, and the critical point the prey chooses to escape is called the flight distance Ydenberg and Dill (1986). Many factors may influence the flight distance, which can be seen as the hyperparameters that influence the prey's risk and opportunity directly, such as the predator's approaching speed, environment temperature, the prey's group size, and the prey's defense tactics Ydenberg and Dill (1986). These influence factors are analyzed qualitatively in Hurley and Hartline (1974); Rand (1964); Heatwole (1968); Lazarus (1979). In this paper, we want to further explore the detailed influence factors of flight distance for goldfish. Through video clips of the interaction between the goldfish and the toy submarine, we find the goldfish reacts differently when the toy submarine approaches in different directions and at different speeds. Therefore, we want to investigate the two major factors (approaching direction and approaching speed) that influence goldfish escaping specifically with probabilistic statistical data analysis methods.

Experiment settings

The experiment environment is a glass aquarium (0.9 m wide x 0.45 m deep x 0.45 m high) with 15 Celsius water. An oxygen pump and an aquarium filter are placed on the side wall of the aquarium to provide the goldfish with a suitable living environment. A toy submarine is placed in the aquarium and interacts with the goldfish. The goldfish is our research target with a 12cm long streamlined body shape, and the toy submarine can be modeled as a cylinder with 14cm height and 1.25cm radius. The toy submarine is our controller and its movement (sink, float, left and right rotation, forward and back off) is controlled by a 27MHz electromagnetic radiation remote controller. The sensor system in the experiment is mainly computer vision equipment that monitors the goldfish's and toy submarine's movement. The sensor system includes a depth camera (ZED 2) and a computer (Intel(R)

Core(TM) i7-9700K CPU @3.60GHz, NVIDIA GeForce RTX 3080). The depth camera is an integration of two common cameras and is mounted above the aquarium to record videos from horizontal and vertical planes respectively when the goldfish interact with the toy submarine. The recorded videos from two views are then sent to the computer for object detection. The computer uses AI algorithm Yolo V5 to detect the 2d locations of the goldfish and toy submarine from the horizontal and vertical planes respectively. The trained image classification algorithm Yolo V5 can distinguish the glass panes' reflected images from the true ones, and crop the target true toy submarine and goldfish in the video frames with high precision. The computer then combines the 2d locations from the two views and outputs the goldfish and toy submarine's 3d positions. The collected data are stored in the form of video frames $D = \{(F1_t, F2_t)\}$, where $F1_t, F2_t$ refers the front view and the side view respectively at t_{th} frame, and each frame is a 2d picture. The total length of the video is 12 minutes, and the collected data length is $||D|| = 23528$ frames with 34 interaction trials. The total duration of the interaction trials is 3476 frames (equivalent to 1.92 min), and the toy submarine keeps silent (the goldfish don't take the silent submarine as a threat and will not escape) for the remaining time.

Methods

The input of the research is videos of the toy submarine interacting with the goldfish. We first propose a series of assumptions for building the probabilistic model of the goldfish escaping strategy and make explanations of the proposed assumptions point by point. Based on the assumptions, the goldfish's movement is analyzed in each frame of the videos, which aims to extract abnormal movement and quantify features of goldfish escaping behavior. By selecting the goldfish's escaping range and escaping starting point in the interaction videos, this research analyzes potential influence factors of a probabilistic model separately when the goldfish makes an escaping decision and tries to build the probabilistic model that determines whether the goldfish starts to escape.

Assumptions

The probabilistic model that determines whether a goldfish starts to escape is built on the environment assumption of an open water area, where the goldfish interact freely with the toy submarine and won't be affected by the container boundary. Based on this basic assumption, further assumptions on the escaping probabilistic model are listed below:

- The probabilistic model that determines whether a goldfish starts to escape can be separated into a pressure function that drives a goldfish to escape, and a friction function that keeps a goldfish's movement unchanged. A goldfish's determination on whether to escape can be modeled as an economic game between the friction function and the pressure function.
- In the interaction between the goldfish and the toy submarine, the toy submarine can be seen as the predator, and the goldfish's behavior can be seen as the prey's response to predators.
- The probabilistic model is irrelevant to the goldfish's absolute position and is only relevant to the relative position between the goldfish and the toy submarine. Therefore, a dynamic 3d coordinate system can be built on the goldfish as the origin.
- The friction function can be seen as a constant function in this probabilistic model.
- The pressure function in the probabilistic model is mainly determined by flight distance (approaching distance) but is also influenced by the toy submarine's approaching direction and approaching speed.

This simple probabilistic model for escaping triggering is a further improvement of the deterministic economic model in Ydenberg and Dill (1986) that determines when a creature starts to escape facing predators. Previous research on creatures' escaping decisions has established a simple deterministic economic model balancing risk from predators and escaping cost Ydenberg and Dill (1986). Based on the deterministic economic model, we find it's generally complicated to compute a unified flight distance threshold in real scenarios because the simple deterministic model fails in many cases and the animal's behavior decisions seem probabilistic rather than deterministic. Given the economic model, this probabilistic model can be separated into the pressure part that drives a goldfish to escape and the friction part that maintains a goldfish's movement, which corresponds to risk from predators and escaping cost in the economic model Ydenberg and Dill (1986). In the research, the interaction between the goldfish and the toy submarine can be seen as prey escaping away from predation because video analysis of the interaction shows the goldfish always fears and keeps away from an active submarine, and seldom approaches the toy submarine actively. Generally, in an open water area, a goldfish's decision of escaping shouldn't be affected by its exact position in the water, but depends on how the goldfish perceives risk or danger from the predators. The open area assumption reduces the influence of glass pane reflection and the confinement effect of limited water space on the goldfish's behavior. This assumption is made for simplification and can hold in this research as most video data is collected when the interaction occurs in the middle of the water tank. Therefore, the main factor in the probabilistic model should be a relative quantity between the goldfish and the predator (the toy submarine). The friction function in the simple probabilistic model is generally considered to be determined by its inertia, and for simplification, the friction function is set constant in this simple model. The last assumption will be analyzed in the following results.

Goldfish abnormal movement (escaping) detection

In order to build a probabilistic model determining when to start escaping facing predators, the first step is to define what is an escaping behavior in the video analysis. A simple definition for escaping behavior is abnormal behavior or abrupt behavior pattern change in all the video clips. In order to extract escaping behavior in the whole goldfish submarine interaction videos, three escaping behavior characters are used to specify an escaping behavior: escaping time range, abnormal movement patterns, and escaping starting point.

Escaping time range The escaping time range refers to the time period when a goldfish completes an escaping action from predators. In the video analysis, an escaping action is not finished at a time point and it usually takes some time from starting to escape till the end of an escaping. We generally regard the escaping time range as a constant hyperparameter depending on each goldfish when eluding away from predators. This means this research assumes the goldfish generally takes the same time period plus some random noise to finish escaping when it eludes away from the active toy submarine, because a goldfish's escaping behavior is a conditional reflex, and in the economic model, a goldfish makes the decision to escape when it feels the risk outranges the escaping cost, which should be a probabilistic unified movement pattern. The escaping time range is calculated from goldfish submarine interaction input video clips in different scenes statistically. The total input 3d video clips are annotated as $D = \{(F1_t, F2_t)|_j\}$, where $F1_t|_j, F2_t|_j$ refers the front frame and the side frame respectively at t_{th} frame in the j_{th} scene, and each frame is a 2d picture. The goldfish's escaping behaviors are extracted as $D_{ex} = \{(t_{i0}, t_{i1}, s_i)|_j\}, j \in \{1, 2, ..S\}$, where $t_{i0}|_j, t_{i1}|_j$ are the goldfish's i_{th} escaping video clip starting and end frame index in scene j, $s_i|_j$ is the segment index, that represents the number of segments in the i_{th} selected video clip, and S is total scene number. In all the video clips, each escaping period $f_{ij} = \frac{t_{i1}-t_{i0}}{s_i}|_j$ can be then obtained from the extracted data. A single variate Gaussian probability density function $g(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp \frac{-(x-\mu)^2}{2\sigma^2}$ that describes the distribution of escaping time range can be applied to fit the discrete distribution $\{f_{ij}\}$, where the Gaussian parameters σ, μ are estimated from the input data by maximum likelihood estimation (MLE) $\hat{\mu} = \frac{1}{N}\Sigma_{ij} f_{ij}, \hat{\sigma} = \sqrt{\frac{1}{N}\Sigma_{ij}(f_{ij} - \hat{\mu})^2}, N = \Sigma_j i_j$. The estimated $\hat{\mu}$ is the escaping range time that the goldfish takes to complete an escaping action normally, which will also be used as convolution window size in goldfish abnormal movement detection.

Abnormal movement patterns and detection This subsection analyzes whether a goldfish's movement in a frame is abnormal based on specified patterns. A normal movement is defined as swimming freely, smoothly, and continuously in the case of not being interrupted or frightened. In all the designed goldfish submarine interaction scenes, the goldfish tends to flee away from the toy submarine when the submarine approaches. Therefore, all the detected abnormal movements can be seen as the detected escaping behaviors.

High movement speed and abrupt swim direction turning are the two main patterns considered to determine whether the goldfish's movement is abnormal in the specified frame in this research. For the high movement speed pattern, a statistical hyperparameter θ_s is used as the threshold to determine whether the goldfish moves extraordinarily fast in each specified frame, and the continuous-value speed time series $v_t \in \mathcal{R}^3$ is converted to a binary mask time series $mv_t \in \{0, 1\}$ marking whether the goldfish's speed is detected abnormal in each frame according to:

$$v_t \mapsto mv_t = \begin{cases} 1, & ||v_t|| \geq \theta_s \\ 0, & ||v_t|| < \theta_s \end{cases} \tag{1}$$

Similarly, a statistical hyperparameter θ_a is also preset to determine whether the goldfish turns extraordinarily large in each specified frame as the threshold, and the continuous-value swim direction turning angle time series $\psi_t \in (-\pi, \pi]$ is also converted to a similar binary mask time series $ma_t \in \{0, 1\}$ by comparison with the statistical threshold θ_a:

$$\psi_t \mapsto ma_t = \begin{cases} 1, & \cos\psi_t < \cos\theta_a \\ 0, & \cos\psi_t \geq \cos\theta_a \end{cases} \tag{2}$$

Here, a cosine function wraps the swim direction turning angle statistics in the algorithm to avoid the influence of lacking rotation starting axis and the periodicity of the trigonometric functions.

The time series of binary masks $\{m_t\}$ are then convoluted with a window function $g_K(t) = \texttt{step}(t) - \texttt{step}(t - K)$ to best match the continuous escaping patterns under the assumption that a fixed time period is required to finish the action, where the step function is $\texttt{step}(t) = (\texttt{sgn}(t)+1)/2$, and the window size K is the obtained escaping time range in subsection :

$$m_t * g_K(t) = \sum_{\tau} m_t \cdot g_K(\tau - t) \tag{3}$$

Convolution is a popular operation for anomaly detection, and the output of the convolution in equation 3 reflects how the discrete time series in a given period corresponds to the escaping pattern. The larger the convolution is, the more likely this frame slice with length K will be detected as an abnormal movement. The detection of goldfish abnormal movement can be separated into the above two abnormal pattern detection: abnormal swim speed and direction turning. Details of the algorithms that detect whether the goldfish's movement is abnormal in a frame are shown in Algorithm 1 and Algorithm 2.

Algorithm 1 Extremely fast swim speed detection

Input: Goldfish 3d position $\mathbf{x}_t = \{(x, y, z)|_t\}$
Data preprocessing: Goldfish swim speed vector at frame t: $\mathbf{v}_t = \mathbf{x}_t - \mathbf{x}_{t-1}$
Mask: Generate a binary mask time series indicating whether the speed at frame t is extremely fast according to equation 1: $v_t \mapsto mv_t$
Mask convolution: Convolute the discrete binary mask series with a K-length window function $g_K(t)$ according to equation 3
Output: Frame indexes set $t_v = \{t|mv_t * g_K(t) \geq \theta_{os}\}$

Algorithm 2 Abrupt swim direction turning detection

Input: Goldfish 3d position $\mathbf{x}_t = \{(x, y, z)|_t\}$
Data preprocessing: Goldfish swim speed vector at frame t: $\mathbf{v}_t = \mathbf{x}_t - \mathbf{x}_{t-1}$; swim direction turning angle at frame t: $\psi_t = \arccos \frac{\mathbf{v}_t \cdot \mathbf{v}_{t-1}}{||\mathbf{v}_t|| \cdot ||\mathbf{v}_{t-1}||}$
Mask: Generate a binary mask time series indicating whether the goldfish turns an extremely large angle at frame t according to equation 2: $\psi_t \mapsto ma_t$
Mask convolution: Convolute the discrete binary mask series with a K-length window function $g_K(t)$ according to equation 3
Output: Frame indexes set $t_a = \{t|ma_t * g_K(t) \geq \theta_{oa}\}$

Escaping starting point Determining the escaping starting point is essential in building the probability function of escaping because the probability function describes what outer conditions may trigger the goldfish to escape. The detected frames with potential escaping behaviors t_v (Algorithm 1) and t_a (Algorithm 2) generated in subsection is the input in this subsection. All the detected potential escaping frames can be generally categorized into starting frames and frames in the escaping process under the assumption that it usually takes a goldfish some fixed time range K to complete an escaping behavior in subsection . In the cases of potential multiple triggers in a period, escaping frames caused by the former trigger will be mixed up with the ones caused by the latter trigger within K frames, and thus it is generally hard to distinguish which stimulus has triggered the escaping frames. Therefore, this assumption also indicates that a period of goldfish escaping frames can be composed of multiple continuous K time ranges triggered by multiple stimuli. For each escaping frame, combine all the frames into a subgroup when the distance between this and the former escaping frame is shorter than K. The escaping starting points can then be extracted as the indexes of the first frames in each subgroup. Details of the combination are shown in Algorithm 3.

Probability function of escaping

This research aims to further improve the deterministic economic model $f(d)$ that describes the goldfish's escaping de-

Algorithm 3 Subgroup combination and selection of escaping starting point

Input: Detected escaping frames from Algorithm 1 and Algorithm 2: t_v, t_a
Do: Sort all the detected frames with ascending order: $L = \texttt{sort}(\texttt{list}(t_a \cup t_v))$
Initialize: Append the first detected index to the escaping starting point list: $S = [L_0]$, empty the subgroup combination list $G = [\cdot]$
for $i = 1$ **to** $\texttt{len}(L)$ **do**
  if $L_i - L_{i-1} > K$ **then**
    Append:$S \leftarrow S + L_i$
  end if
end for
for $j = 0$ **to** $\texttt{len}(S)$ **do**
  Initialize: empty subgroup list $g = [\cdot]$
  Define: $Range_j = [S_j, S_{j+1})$
  Update: $g \leftarrow L \cap Range_j$
  Append: $G \leftarrow G + g$
end for
Output: First indexes of all the subgroups: the escaping starting point list S & all combined subgroups G.

cision with a probabilistic model $p(d; \boldsymbol{\theta})$. The probabilistic model is a multi-variable conditional probability that describes how different configurations $\boldsymbol{\theta}$ may affect the flight distance d, and thus affect the goldfish's escaping decision. The numerical solution of multi-variable probability $p(d; \boldsymbol{\theta})$ is generally hard to compute or estimate because the configuration space $\boldsymbol{\theta} \in \mathrm{R}^d$ increases exponentially with configuration dimension d. One way to analyze the multi-variable probability $p(d; \boldsymbol{\theta})$ is to consider the influence of different configuration parameters separately and break the multi-variable probability $p(d; \boldsymbol{\theta})$ into a series of two-variable probability $p(d; \theta_i), i \in \{1, 2, .., d\}$. This can be verified by Bayes Theorem with parameter independence assumptions shown in equation 4 (a 3-variable (with 2 parameters) conditional probability simplification).

$$
\begin{aligned}
p(d|\theta_1, \theta_2) &= \frac{p(d, \theta_1, \theta_2)}{p(\theta_1, \theta_2)} = \frac{p(d, \theta_1) \cdot p(\theta_2|d, \theta_1)}{p(\theta_1) \cdot p(\theta_2|\theta_1)} \\
&\approx \frac{p(d, \theta_1) \cdot p(\theta_2|d)}{p(\theta_1) \cdot p(\theta_2)} = \frac{p(\theta_1) \cdot p(d|\theta_1) \cdot p(\theta_2) \cdot p(d|\theta_2)}{p(d) \cdot p(\theta_1) \cdot p(\theta_2)} \\
&= \frac{p(d|\theta_1) \cdot p(d|\theta_2)}{p(d)}
\end{aligned}
\tag{4}
$$

The continuous two-variable conditional probability $p(d; \theta_i \in \mathrm{R})$ can be further reduced to single-variable conditional probability $p(d; \theta_i \in \Theta_i)$, where Θ_i some finite discrete parameter space, by discretizing the original continuous parameter space for simplification. Fixing and enumerating all possible $\theta_i \in \Theta_i$ will reduce the two-

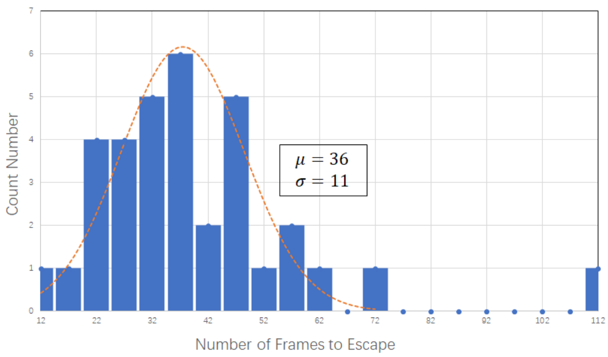

Figure 1: Range K count distribution

variable conditional probability $p(d; \theta_i)$ to single-variable conditional probability $p(d|\theta_i)$, where θ_i is a constant. Model-based parameter estimation can be then applied to fit the single-variable conditional probability $p(d|\theta_i)$ within a list of potential distribution candidates. The single-variable probability distribution fitting for $p(d|\theta_i)$ is optimized by maximizing the expected likelihood of the data:

$$\Omega_i = \arg\max_{\Omega_i} \prod_{j=1}^{N_j} L_k(d_j; \Omega_i), \quad d_j \sim p(d|\theta_i \in \Theta_i) \quad (5)$$

$$p(d|\theta_i \in \Theta_i) \approx \arg\max_{L_k} L_k(d; \Omega_i), \quad L_k(d; \Omega_i) \in \mathcal{L}, \quad (6)$$

where $\mathcal{L}$ is a series of potential distributions.

Results and evaluation

Determining Escaping time range

A Gaussian model with mean 36 and std 11 has been applied to fit the time range that the goldfish takes to complete the escaping behaviors in 5 different interaction scenarios (approaching, catching, front, plane-back, and vertical up-down). Details of the fitted Gaussian model and the data distribution are shown in Figure 1. Therefore, the estimated mean of the escaping time range $\hat{\mu} = K = 36\ Frames$, which means the goldfish generally takes $36\ Frames/30\ FPS = 1.2s$ to complete one escaping behavior.

Escaping behavior extraction

For abnormal speed detection, Figure 2 shows the goldfish's speed distribution in the whole fish-submarine interaction. In most normal cases, the goldfish tends to swim slowly in the aquarium and reacts fiercely when the toy submarine is near. Therefore, a long tail exists in the goldfish's movement speed distribution graph, and above 95% weight is distributed below 28 cm/s. Here, we take the 95% quantile of the total goldfish speed distribution as an empirical statistic threshold $\theta_s = 28cm/s$ for detecting abnormal movement speed in a frame. In this process, 2490 frames are selected as abnormal.

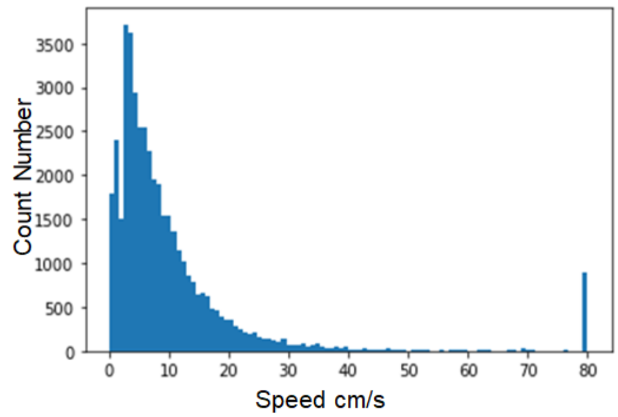

Figure 2: Goldfish abnormal movement speed count distribution

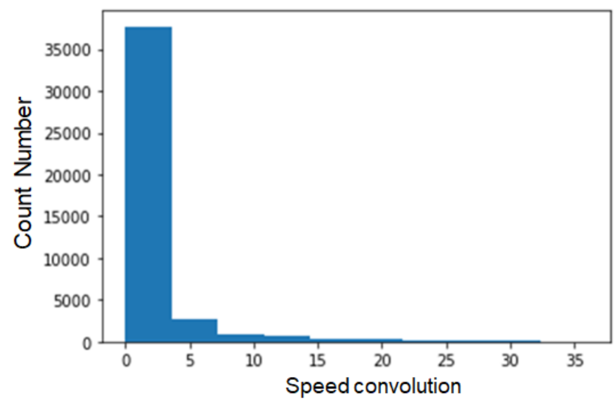

Figure 3: Goldfish abnormal movement speed convolution count distribution

The convolution result of the time series has been counted and transformed into statistic count bins in Figure 3. The vast majority of the convolution weight is distributed below 5 times, while only 5% of the continuous 36-length windows contain more than 8 abnormal speed frames. Similarly, we take the 95% quantile of the total convolution distribution as an empirical statistic threshold $\theta_{os} = 8$ for determining whether the detected frame is associated with a complete escaping behavior under the assumption that the goldfish normally takes K frames to complete an escaping.

Goldfish seldom make abrupt direction turning frequently in a continuous time period. A large angle turning, such as an obtuse angle between two continuous frames, can be seen as an abrupt direction turning. Therefore, the threshold for large-angle turning between two continuous frames can be set as a right angle $\theta_a = 90°$. Based on the convolution result for abrupt direction turning detection in the K-length sliding window, 98% of the total abrupt turnings are distributed below 22 frames, while the maximum of the abrupt direction turning is 28 frames in the 36-length sliding window. Therefore, one statistical convolution threshold for the abrupt direction turning in the K-length sliding window can be set to 98% quantile of the abrupt angle turning convolution distribution, $\theta_{oa} = 22$. In this process, 349 frames are selected as abnormal. Exceeding the convolution threshold θ_{oa} means many abrupt direction turnings occur in the K-length sliding window, and it is likely that this K-length sliding window contains the start of an escaping behavior. To sum up the two aspects, total 2839 frames are classified as abnormal movement starting points, and the precision for the anomaly detection is 81.68%.

Escaping probability function establishment

The traditional flight distance model describes when a creature decides to flee from its predators based on the approaching distance from the perspective of an economic game balancing risk and potential opportunity Ydenberg and Dill (1986). The traditional economic model is a deterministic model and predicts that a creature will choose to flee if the approaching distance is smaller than a distance threshold: $p(flee|d \le d_0) = 1$. However, this traditional deterministic model is not sufficient enough to describe the goldfish's fleeing decision, as it only considers the predator's approaching distance, and a unified approaching distance threshold d_0 is usually hard to obtain on many occasions. Meanwhile, the approaching distance threshold tends to be probabilistic depending on the predator's approaching direction and approaching speed $p(d; \theta_1, \theta_2)$, where θ_1 is the first influence factor (the predator's approaching direction) and θ_2 the second influence factor (the predator's approaching speed). This multi-variate probabilistic model can be further decomposed into a series of single-variate probabilistic models $p(d|\theta_i), i \in \{1, 2\}$ under the independence assumption according to equation 4. Each single-variate probabilistic model $p(d|\theta_i), i \in \{1, 2\}$ is then fitted from a series of the parameterized probabilistic models $\mathcal{L}$[1] [2] with maximized data likelihood according to equation 5.

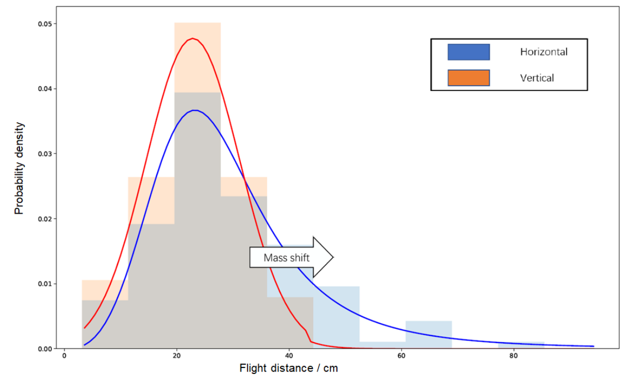

Figure 4: Flight distance distribution when toy submarine approaches horizontally (blue) and vertically (red)

For the first independent influence factor, the toy submarine approaching direction θ_1, we generally discretize the continuous variable θ_1 into binary choices $\theta_1 \in \{horizental, vertical\}$. We first separate all the extracted escaping starting points $data = \{frame_i\}$ in subsection into $data = data_h + data_v$ according to the toy submarine's approaching direction, where $data_h = \{idx|\mathbf{d}^2_{idx}[2] \le \mathbf{d}^2_{idx}[0] + \mathbf{d}^2_{idx}[1], \mathbf{d} = \mathbf{x}_{submarine} - \mathbf{x}_{goldfish}\}$ represents the extracted escaping starting data when the toy submarine approaches horizontally, and $data_v = \{idx|\mathbf{d}^2_{idx}[2] > \mathbf{d}^2_{idx}[0] + \mathbf{d}^2_{idx}[1]\}$ for vertical approaches. For each horizontal and vertical escaping starting point, its approaching distance at the frame is recorded and then fitted from the potential distribution family list $\mathcal{L}$ separately. Figure 4 shows the results of the fitted probabilistic model for horizontal and vertical data respectively. In the figure, a Fisk distribution was fitted for the cases when the submarine approaches horizontally with p-value 1.00 and an NCT distribution for the vertical cases with p-value 0.910.

The fitted horizontal probabilistic model $p(d|\theta_1 = horizontal)$ holds a long tail compared with the vertical model $p(d|\theta_1 = vertical)$. This reveals that there still exists some probability for the goldfish starting to flee in the large distance range for the horizontal model, while the probability for the goldfish starting to flee when the submarine approaches vertically from a large distance is very small. The vertical model concentrates all the approaching distance d probability in the short-distance area and distributes in a relatively small range ($d \in (5, 35)$), while the probability mass in the short-distance area for the horizontal model is relatively smaller than the vertical model because the horizontal shares some probability mass in the large distance area (the long tail shown in Figure 4). Besides, the peak and the mean of the vertical probabilistic model are both left-skewed compared with the horizontal model, and the peak in the vertical model

[1]The set of potential distribution family $\mathcal{L}$ = [normal, beta, exponential, gamma, generalized extreme value, logistic, lognormal, triangular, uniform, fatigue-life, generalized gamma, generalized norm, double Weibull, double gamma, right-skewed Gumbel, power normal, Rayleigh, Weibull maximum, Weibull minimum, Laplace, alpha, generalized exponential, Bradford, beta prime, Burr (Type III), Fisk, generalized Pareto, hyperbolic secant, half-normal, half-logistic, inverse Gaussian, inverted gamma, Levy, log-Laplace, log gamma, Maxwell, Mielke Beta-Kappa, non-central chi-square, non-central F distribution, non-central Student's t, Nakagami, Pareto, Lomax (Pareto of the second kind), power log-normal, power-function, Rice, semicircular, inverted Weibull, folded normal, folded Cauchy, cosine, exponential power, exponentiated Weibull, Wald, wrapped Cauchy, truncated exponential, truncated normal, Student's t, R-distributed (symmetric beta)].

[2]Details of these parameterized distribution families can be found on webpage Jones et al. (01) from Scipy 1.9.3 Virtanen et al. (2020).

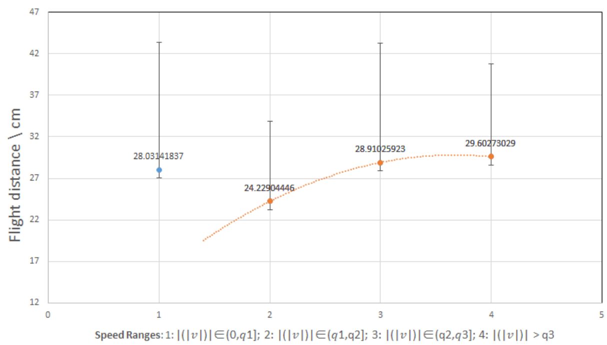

Figure 5: Flight distance distribution under different approaching speed ranges

is higher than the horizontal model, which also reflects the probability mass concentrates more in the small-distance area and around the peak for the vertical model, while the probabilistic model for the horizontal data is relatively flatter than the vertical model. These differences reveal that the goldfish reacts differently with regard to the predator's (the toy submarine) approaching directions. Results in Figure 4 suggest the goldfish is more sensitive to the predator's horizontal approaches and tries to flee even when the predator is still at a relatively far distance. Therefore, the flight distance in the traditional economic model Ydenberg and Dill (1986) is also affected by the approaching direction, and the horizontal approach will increase the flight distance compared to the vertical approach. This finding can also be verified in the recorded video frames.

For the second independent influence factor, the toy submarine approaching speed θ_2, we generally discretize the continuous variable θ_2 into 4 different speed ranges $\theta_2 \in (0, q_0] \cup (q_0, q_1] \cup (q_1, q_2] \cup (q_2, +\infty)$, where q_0, q_1, q_2 are the 1/4,1/2,3/4-quantile of the whole speed distribution in the extracted escaping frames. Similarly, we first separate all the extracted escaping starting points $data = \{frame_i\}$ in subsection into $data = data_1 + data_2 + data_3 + data_4$ according to the toy submarine's approaching speed, where $data_i = \{idx| \; ||v_{idx}|| \in (q_{i-2}, q_{i-1}]\}, i = 1, 2, 3, 4$ represents the extracted escaping starting data whose approaching speed are in the i-th speed range. All the 3 quantiles for the approaching speeds are $q_0 = 5.04, q_1 = 8.55, q_2 = 12.94$. For the extracted escaping frames in each speed range, its approaching distance d at the frame is recorded and then fitted from the potential distribution family list $\mathcal{L}$ separately. Subfigures in Figure 5 show the results of the fitted probabilistic model of different approaching speeds respectively.

It's obvious the probabilistic models for different approaching speed ranges don't belong to the same distribution family. Apart from the smallest approaching speed range, the mean of the probabilistic models in different speed ranges increases with the approaching speed. This difference reveals that the goldfish reacts differently with regard to the predator's (the toy submarine) approaching speed. Results in Figure 5 suggest the goldfish is more sensitive when the predator approaches faster and tries to escape even when the predator is still at a relatively far distance. Therefore, the flight distance in the traditional economic model Ydenberg and Dill (1986) is also affected by the approaching speed, and the flight distance increases with approaching speed in general. This finding can also be verified in the video frames.

Conclusion and future work

This research has built a probabilistic model for flight distance given two observed features, the predator's approaching speed and direction based on the traditional economic model Ydenberg and Dill (1986). The probabilistic model can be a benchmark to study creatures' escaping behavior, and its performance is cross-validated with empirical interaction data. In general, this research has found that the predator's different approaching speeds and directions will influence the goldfish's flight distance when escaping. To be specific, goldfish are generally more sensitive to the horizontal approach than the vertical approach. Although it's not universally true that the flight distance will increase as the approaching speed increase, this result still holds in the large approaching speed range (>q0).

However, this research is far from perfect, and much work can still be done to improve the understanding of animals' escaping behavior under the baseline of this probabilistic model in the future. For instance, more biological understanding behind the probabilistic model still needs to be investigated, such as why goldfish are more sensitive to horizontal approaches than vertical ones in biology. Is it caused by limited eyesight range or different distribution/sensitivity of the goldfish's pressure sensor? How do these factors influence the animals' perception of potential risks and maintain opportunities, which can relate these biological findings to the original economic model? Further research about whether these findings are universal among most species still needs to be performed. Applying this probabilistic model to different species, and checking out whether these finding still holds in that species is one way to confirm the finding's generality. If the findings are not universal, how differently other species may react and what biological reasons may lie behind are the questions to explore. Besides, the probabilistic model itself is an ideal model based on a series of assumptions in subsection , and these assumptions are made to simplify the probabilistic model. The validity of these assumptions still needs to be justified to better reflect the true environment. For example, the third assumption excludes the confinement effect on goldfish behavior in the experiment, and the fourth assumption in subsection supposes the friction function is constant everywhere in the fish tank for simplification. However, this is not universally true, because, for instance, lower water space may contain more food and thus has a larger value for the friction function compared with the upper water space. Detailed friction functions may be designed with empirical data to better reflect the environment in the future.

Acknowledgement

The authors are grateful for all the reviews' precious comments and suggestions for this paper, which can be potential research directions in the future. This work is supported by Shanghai Technology Development and Entrepreneurship Platform for Neuromorphic and AI SoC and in part by the Shanghai Research and Innovation Functional Program under Grant 17DZ2260900.

References

Broom, M. and Ruxton, G. D. (2005). You can run—or you can hide: optimal strategies for cryptic prey against pursuit predators. *Behavioral Ecology*, 16(3):534–540.

Heatwole, H. (1968). Relationship of escape behavior and camouflage in anoline lizards. *Copeia*, pages 109–113.

Hileman, K. S. and Brodie Jr, E. D. (1994). Survival strategies of the salamander desmognathus ochrophaeus: interaction of predator-avoidance and anti-predator mechanisms. *Animal Behaviour*, 47(1):1–6.

Hurley, A. C. and Hartline, P. H. (1974). Escape response in the damselfish chromis cyanea (pisces: Pomacentridae): a quantitative study. *Animal Behaviour*, 22(2):430–437.

Jones, E., Oliphant, T., Peterson, P., et al. (2001–). SciPy: Open source scientific tools for Python.

Lazarus, J. (1979). The early warning function of flocking in birds: an experimental study with captive quelea. *Animal Behaviour*, 27:855–865.

Marchisin, A. and Anderson, J. D. (1978). Strategies employed by frogs and toads (amphibia, anura) to avoid predation by snakes (reptilia, serpentes). *Journal of Herpetology*, pages 151–155.

Martín, J., Luque-Larena, J. J., and Lopez, P. (2009). When to run from an ambush predator: balancing crypsis benefits with costs of fleeing in lizards. *Animal Behaviour*, 78(4):1011–1018.

Miyatake, T., Tabuchi, K., Sasaki, K., Okada, K., Katayama, K., and Moriya, S. (2008). Pleiotropic antipredator strategies, fleeing and feigning death, correlated with dopamine levels in tribolium castaneum. *Animal behaviour*, 75(1):113–121.

Rand, A. S. (1964). Ecological distribution in anoline lizards of puerto rico. *Ecology*, 45(4):745–752.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

Ydenberg, R. C. and Dill, L. M. (1986). The economics of fleeing from predators. In *Advances in the Study of Behavior*, volume 16, pages 229–249. Elsevier.