1 Introduction

1.1 Fish Tracking by a Markov Process

This work simplifies the tracking problem by discretizing the plane of all possible fish coordinates into N discrete locations, known as *nodes*. These nodes may be chosen by any process, such as dividing the region into a grid or clustering historical data to calculate N centroids. While the nodes used in this work were stationary, this is not a requirement. The set of nodes will be denoted by V, and the physical location of node i at time t will be represented by the function $V_i(t)$.

The path of a robot is confined to certain *edges* between nodes. An edge is an ordered pair (i, j), with $i, j \in V$, and for each edge a physical path $e(i, j)$ exists that the robot is to follow when moving from node i to node j. There need not exist an edge between any pair of nodes, as in some cases, it may not make sense for the robot to move between such nodes, particularly if finding an obstacle-free path between the nodes would be difficult. If E is the set of all edges, then the directed graph (V, E) is the *state graph*.

Through some analysis of previous fish data, a time-inhomogeneous transition matrix $T(t)$ is generated. $T(t)$ specifies a left stochastic matrix for which the $j, i-$th entry is the probability of the robot transitioning from node i to node j at time t. Since many species of fish show periodic patterns of movement, analysis of previous fish locations may be leveraged to generate a $T(t)$ that predicts future fish locations.

1.2 Transition Matrix Construction

The procedure for creating the transition matrix $T(t)$ roughly falls under two steps. First, previous data is analyzed to create a *transition model* for the fish. This transition model is another time-inhomogeneous transition matrix $T_{\mathrm{model}}(t)$, for which the j, i-th entry is the probability of a *fish* transitioning from node i to node j at time t. This model attempts to find a Markov process that predicts how fish will traverse the state graph.

Second, the transition model is transformed into the robot transition matrix. Since the fish may traverse edges that do not exist for the robot (moored boats or other obstacles may exist), this robot transition matrix must assign a probability of zero to such edges. This work explores how to generate such a robot transition matrix through two approaches. Under the *periodic control* approach, Fourier analysis of fish population densities of each node is used to generate a desired robot population density $\mathbf{p}_t$, and $T(t)$ is selected so that the difference in $\mathbf{p}_{t+1}$ and $T(t)\mathbf{p}_t$ is minimized. Under the *feedback linearization control* approach, a closed-loop feedback control is used to calculate $T(t)$ so as to converge on the desired robot population densities.

2 Transition Model

We will represent the fish's location by Cartesian coordinates in the plane of the surface of the water, with the origin set to some latitude / longitude coordinate O. Latitude / longitude coordinates of the fish measured during the track are mapped to this plane. Let $r[t]$ denote the x-coordinate (distance east of O) and y-coordinate (distance north of O) of the fish at time $t > 0$.

2.1 Interpolation and Weighting

Since the fish's location cannot be determined at regular intervals, the times for which $r[t]$ is defined also fail to occur at regular intervals. Such an uneven sampling is inconvenient for modeling the fish's location. Interpolation is used to generate an approximation $R[t]$ of the fish's location, defined for any $t = kT_{\mathrm{sample}}$,

where k is a nonnegative integer. For convenience, we may use the notation $R[k] = R[kT_{\text{sample}}]$ to index the fish's location by integer k.

2.2   Modeling Individuals vs. Population

After interpolation, there are two possible paths for further analysis. Both paths have the primary goal of a transition model that will place robots as close to the target fish as possible. The *individual* approach creates a transition model in which the probability of the transition $i \rightarrow j$ at time t approximates the probability of a real fish making the transition $i \rightarrow j$ at time t, based on analysis of the data $R[k]$ and $R[k+1]$. However, the *population* approach creates a transition model in which the density of robots at node i and time t approximates the density of fish at node i and time t, based on analysis of the same data. The individual approach makes individual robots behave as individual fish, while the population approach makes the population of robots behave as the population of fish.

2.3   Individual Model

Once the fish's location $R[k]$ is discretized to the N nodes of the graph, we may represent the fish's movement between each time step as $d_i[k]$, defined recursively as follows:

$$
d_i[k] = \begin{cases} R[k+1] & R[k] = i \\ d_i[k-1] & R[k] \neq i \text{ and } k > 0 \\ i & R[k] \neq i \text{ and } k = 0 \end{cases}
$$

Informally, $d_i[k]$ represents the node that the fish will occupy at time step $k+1$ if it's location at time step k is node i. Of course, we only observe the fish's transition from one source node per time step. If $R[k] = i$ and $R[k+1] = j$, then we may easily choose $d_i[k] = j$; however, we are left to guess where the fish *would* have gone from some other node $x \neq i$ to choose the value of $d_x[k]$. The above definition makes the guess that, if $R[n+1] = y$ for the largest $n < k$ at which $R[n] = x$, then $d_x[k] = y$. An alternative definition of $d_i[k]$ may make the guess that $d_i[k] = R[k+1]$ for any i, though such a definition may overemphasize the current location of the fish, at the expense of losing information about the fish's path.

The next step is to convert $d_i[k]$ into a function $T[k]$, which yields the $N \times N$ transition matrix for time step k. This summer, we used a "windowed transition counting model" to extract transition probabilities from $d_i[k]$. Under this model, we define $T[k]_{i,j}$ as follows:

$$
T[k]_{i,j} = \sum_{m}^{L} f(m|k, \sigma/T_{\text{sample}})\langle d_j[m] = i \rangle,
$$

where L is the largest value such that $d_i[L]$ is defined, $f(x|\mu, \sigma)$ is a discrete normal distribution (defined for integers 0 to $L-1$ with a total sum of 1), and $\langle P \rangle$ is 1 if the proposition P is true and 0 otherwise. Note that, by this definition, every column in $T[k]$ has a sum of 1, and every element of $T[k]$ is in the range $[0, 1]$.

2.4   Population Models

Let $\vec{R}[k] = (x_k, y_k)$ be a vector denoting the fish's location at time step k, and let $\vec{V_i} = (x_i, y_i)$ denote the location of node i. From these quantities we may define an *attenuation vector* $\mathbf{a}[k]$ as follows:

$$
\mathbf{a}_i[k] = e^{-||\vec{R}[k] - \vec{V_i}[k]||^2/(2\sigma_{\text{atten}}^2)}
$$

where σ_{atten} is a free parameter. Informally, $\mathbf{a}[k]$ describes the weight of a particular node in the fish's location. If the fish's location at some k precisely equals the location of node i, then $\mathbf{a}_i[k] = 1$; but as

the fish moves away from i, then $\mathbf{a}_i[k] \to 0$. The parameter σ_{atten} sets the "attenuation" of the a node's assigned weight as the fish moves away from it.

We then define a density vector $\mathbf{p}[k]$ as a weighted sum of the attenuation vector:

$$\mathbf{p}[k] = \sum_m^L f(m|k, \sigma/T_{\text{sample}})\mathbf{a}[m].$$

Ideally, we would like to construct a transition matrix $T[k]$ such that $\mathbf{p}[k+1] = T[k]\mathbf{p}[k]$. In some cases, however, such a construction may not be possible. Instead, we find a $T[k]$ that minimizes $||\mathbf{p}[k+1] - T[k]\mathbf{p}[k]||^2$, subject to the constraint that $T[k]$ is a stochastic matrix.

3 Robot Transition Matrix