2015 IEEE International Conference on Robotics and Automation (ICRA)
Washington State Convention Center
Seattle, Washington, May 26-30, 2015

Gaussian Reconstruction of Swarm Behavior from Partial Data

Glenn Wagner, Howie Choset

***Abstract*— Swarms consist of large numbers of individual agents that generally maintain no fixed relative positions, which makes describing the behavior of the swarm as a whole difficult. Furthermore, the high number of agents leads to frequent occlusions that prevent observations of the entire swarm. In this paper, we represent the behavior of swarms using velocity fields, yielding a description which is invariant to the number of agents in a swarm, and the position, orientation, and scale of the swarm. The velocity field representation allows the behavior of swarms to be modeled as a Gaussian distribution. We demonstrate that this Gaussian model can be used to reconstruct the behavior of the swarm as a whole from partial observations.**

I. Introduction

Biological swarms are decentralized systems consisting of tens to billions of individual agents [16] that can generate complex behavior using purely local communication. Understanding the behavior of swarms is important for understanding and managing the ecology of economically important species such as herring [13] and bees, and may provide insight into the design and control of systems containing very large numbers of robots. However, such dense aggregations of organisms are difficult to observe, as occlusions by other agents or obstacles are common and often block large portions of the swarm from view. Occlusions also make tracking individual organisms for long periods of time difficult.

In this paper, we demonstrate that a simple model of global swarm behavior can be leveraged to reconstruct and classify swarm behavior when only a small portion of the swarm can be observed at any given time. The state of a swarm is represented by the velocity field along which the swarm agents flow. The behavior of the swarm is then modeled by a Gaussian distribution over velocity fields. Using the Gaussian model, the velocity field of the swarm as a whole can be inferred from an observation of a subset of the agents. The qualitative label describing what class of behavior the swarm is executing can be computed from the reconstructed velocity field. We validate our approaches on fish tracking data[20].

II. Prior Work

Reconstructing the global behavior of a swarm from limited observations requires a representation that can describe spatial correlations of behavior across the swarm. The biological community has developed a number of concise descriptions of the overall behavior of swarms [1, 15, 20]. However, these descriptions are hand-coded by the researcher to capture specific behaviors, and thus will only describe "obvious" behaviors. Furthermore, the descriptors are not *generative*, *i.e.* the behavior of the swarm cannot be reconstructed, even approximately, from a set of descriptors, nor can the behavior of specific agents be inferred from the descriptors. As such, these descriptors are not an appropriate model for inferring the global behavior of a swarm from partial observations.

Brown and Goodrich [3] considered the problem of classifying the state of a swarm as one of several qualitative behavioral labels based on observations of a small subset of agents. The classification is made on the basis of the linear and angular momenta of the observed agents, and the number of neighbors with which each agent interacts. Calculating the number of neighbors requires knowledge of how the agents in the swarm communicate and the positions of all the agents in the swarm. As a result, this approach works best as a method of reducing the computational cost of classifying swarm behavior, rather than as a method of classifying the behavior of a partially observed swarm.

The machine learning community has developed many approaches for estimating missing data [6, 8, 9]. Principle Component Analysis [14] and Gaussian inference [17] can leverage training data to improve the reconstruction of novel observations with missing data.

III. Overview

The objective of this paper is to accurately reconstruct and classify the behavior of swarms from observations of a limited subset of the swarm's agents. Thus, a representation of the swarm is required that is independent of the number of agents. Continuum models, which describe a swarm in terms of bulk properties such as density and momentum flux [12, 18, 19], are well suited for this purpose. In this paper, the state of a swarm is represented by a velocity field which describes the bulk motion of the swarm. The overall behavior of the swarm, *i.e.* set of states assumed by the swarm, is modeled as a Gaussian distribution over velocity fields. The Gaussian model contains information on how the velocity field at one point in the swarm is correlated with the velocity field at all other points. Thus, the Gaussian model can be used to infer the value of the velocity field at every point in the swarm from the velocities of a limited number of agents. A label describing the qualitative class of the behavior the swarm is performing can be computed from the reconstructed velocity field. The results presented in this work show that the Gaussian reconstruction produces better reconstructed velocity fields and better behavior classifications from limited

Glenn Wagner is a graduate student at the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 15213 gswagner@andrew.cmu.edu

Howie Choset is a professor at the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 15213 choset@cs.cmu.edu

data than baseline approaches.

A. Velocity Field Representation for Swarms

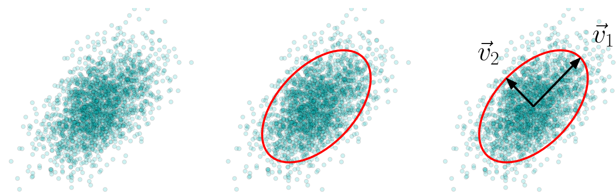
One basic question when analyzing the behavior of swarms is the choice of representation. Much work on the simulation and analysis of swarms has been performed by tracking individual agents within the swarm [4, 5, 10, 21]. Representing individual agents is an excellent approach for exploring how individual agents interact. However, such a representation poses several problems for studying the global behavior of the swarm. Any approach focused on enumerating individuals may become intractable as the number of agents grows. Comparing swarms with different numbers of agents is more difficult when the swarms are represented by the state of the individual agents, and finally the motion of agents within a swarm makes computation of spatial properties difficult. An alternative is to model the swarm as a continuum of infinitesimal agents and define the state of the swarm in terms of aggregate properties such as density and velocity fields [2, 7, 12, 18, 19].

Our interest lies in modeling the global behavior and dynamics of swarms. Therefore, we take the continuum approach and represent the state of swarms as a continuous velocity field. The velocity field of the swarm specifies the velocity field that an agent would have if it were placed at any point within the swarm. The density field is not included in the representation of a swarm for several reasons. Omitting the density field significantly decreases the dimensionality of the representation, and avoids the potential problems with scaling and range associated with performing numerical analysis on data describing more than one physical property. More importantly, the velocity field describes how a swarm changes over time, rather than describing the current state of the swarm, as the density field does.

Finally, we remind the reader of some basic vector field notation. A vector field $\mathcal{X}$ is defined as a section of the tangent bundle TM of some manifold M [11], *i.e.* a function $\mathcal{X} : M \to TM$ which maps each point on the manifold to a vector in its tangent space. The tangent space of a manifold can be thought of as the space of directions in which a point on the manifold can move without leaving the manifold. In a slight abuse of terminology, we define a velocity field as a mapping $\mathcal{V} : \mathcal{W} \to T\mathcal{W}$ from the workspace of a system to a velocity in the workspace.

B. Gaussian Model

We seek a model for swarms that allows predicting the behavior of the swarm as a whole from partial observations. To do this, the model must describe how the velocity field at one point of the swarm depends on the velocity field at other positions in the swarm. We choose to model the behavior of a swarm as a Gaussian distribution over velocity fields, which describes how likely the swarm is to exhibit any given behavior. Critically, the Gaussian model also describes how the velocity field at one point in the swarm is correlated with the velocity field elsewhere in the swarm, allowing an entire velocity field to be inferred from a limited number of observations.



(a) (b) (c)

Fig. 1: Geometric interpretation of Singular Value Decomposition (SVD). **(a)** Samples are drawn from a multivariate Gaussian distribution with $\Sigma = \left[\begin{smallmatrix} .5 & 1 \\ 1 & .5 \end{smallmatrix}\right]$. **(b)** The error ellipsoid is plotted. **(c)** The SVD of the data is computed, then the first $\vec{v}_1$ and second $\vec{v}_2$ right singular vectors are plotted. The first and second right singular vector are equivalent to the semi-major and -minor axes of the error ellipsoid, which are also the principle moments of inertia of the sampled points.

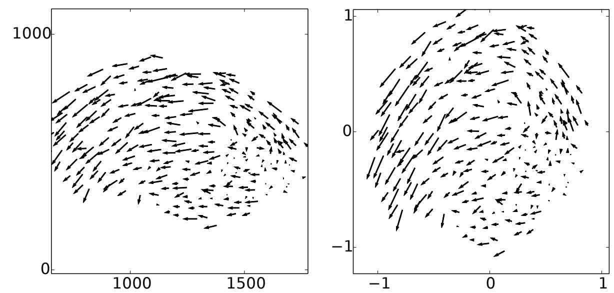
The first step to computing a Gaussian model is to properly register the swarm data. A transformation must be found that maps swarms of similar shape but different size, position, and orientation to a single configuration. Our approach is based on a geometric interpretation of Singular Value Decomposition (SVD) (Figure 1). Let an observation consist of the position and velocity of n agents, which will be referred to as a *frame* of data. Let $\vec{q}_i$ denote the position of the ith agent. The data matrix $\mathbf{Q} = \left[\vec{q}_1, \dots, \vec{q}_n\right]^\intercal$ represents the position of each agent in the swarm as a row. The mean position can be computed as $\bar{q} = \frac{1}{n}\sum_{i=1}^{n} \vec{q}_i$, and used to find the zero-mean data matrix $\bar{\mathbf{Q}} = \mathbf{Q} - \vec{1}\bar{q}^\intercal$. Performing SVD on the zero-mean data matrix, $\bar{\mathbf{Q}} = \mathbf{U\Sigma V}^*$, extracts the matrix of right singular vectors $\mathbf{V}$ which correspond to the principle moments of inertia of the swarm. $\mathbf{\Phi} = \frac{\sqrt{n}}{2}\mathbf{\Sigma}^{-1}\mathbf{V}^*$ describes a linear transform that maps the two-standard deviation ellipse of the positions in $\bar{\mathbf{Q}}$ to the unit circle. A *normalizing transform* can then be defined to map swarm data to a common shape and size (Figure 2):

$$\Phi : \mathcal{W} \times T\mathcal{W} \to \mathcal{W} \times T\mathcal{W} \tag{1}$$
$$\vec{q}, \vec{v} \mapsto \mathbf{\Phi}(\vec{q} - \bar{q}) \times \mathbf{\Phi}\vec{v}$$

Note that the velocity of each agent is scaled by the size of the swarm, effectively expressing velocities in terms of swarm lengths per unit time.

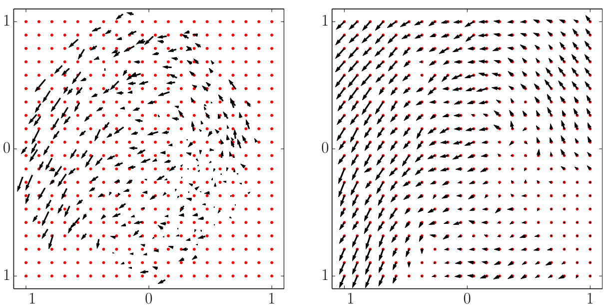
Once the swarm is registered, a velocity field can be generated from the observed agent velocities via knn-interpolation. To be able to compare velocity fields for different frames of data, they must be defined over the same domain. Assuming the spatial distribution of agents is Gaussian, the normalizing transform maps 95% of a swarm to the interior of the unit circle. Therefore, $[-1, 1] \times [-1, 1]$ is an appropriate choice for the domain of the velocity fields.

We model the behavior of the swarm as a Gaussian distribution over velocity fields. The Gaussian model describes the probability of the swarm exhibiting a velocity field.



(a) Original Swarm (b) Transformed Swarm

Fig. 2: Example of the normalizing transform. Both plots have a 1:1 aspect ratio.



(a) Example Swarm (b) Example Velocity Field

Fig. 3: Interpolation of a swarm into a vector field. **(a)** The arrows represent the velocity of swarm agents subject to the normalizing transform. The red dots are the grid points onto which the swam is interpolated to define the discretized vector field. **(b)** Discretized velocity field computed by interpolation.

To compute the Gaussian distribution, the velocity fields must be represented as vectors of finite length. Therefore, the velocity fields must be discretized. Velocity fields are evaluated on a regular, rectangular grid, 20 points on a side, that spans $[-1, 1] \times [-1, 1]$ (Figure 3). The discretized velocity field can now be described by a 3-tensor. The first two indices represent the (x, y) position, respectively, while the third index specifies whether the entry describes the x or y component of the velocity field. The tensor can then be unfolded into a vector, called the *vector representation* of the velocity field.

We now describe how to learn a Gaussian model for swarm behavior from n frames of data, where a frame of data specifies the position and velocity of each agent in a swarm at a single instant of time. A Gaussian distribution is described by the mean velocity field $\vec{\mu}$, and a covariance matrix $\mathbf{\Sigma}$ which describes the spatial correlations of the velocity field. The first step to training the Gaussian model is to compute the normalizing transform for each frame, then apply the transform to project the frame onto $[-1, 1] \times [-1, 1]$. The frame is then interpolated onto the grid, and the vector representation is computed. The mean velocity field can be computed by averaging the vector representations:

$$\vec{\mu} = \frac{1}{n} \sum_{i=1}^{n} \vec{\zeta}_i \tag{2}$$

where $\vec{\zeta}_i$ is the vector representation of the velocity field for the ith frame of data. The covariance matrix is given by

$$\mathbf{\Sigma} = \frac{1}{n} \mathbf{D}\mathbf{D}^T \tag{3}$$

$$\mathbf{D} = \left[\vec{\zeta}_1 - \vec{\mu}, \ldots, \vec{\zeta}_n - \vec{\mu} \right] \tag{4}$$

where $\mathbf{D}$ is the zero-meaned data matrix of vector representations, where each velocity field is represented by a column vector.

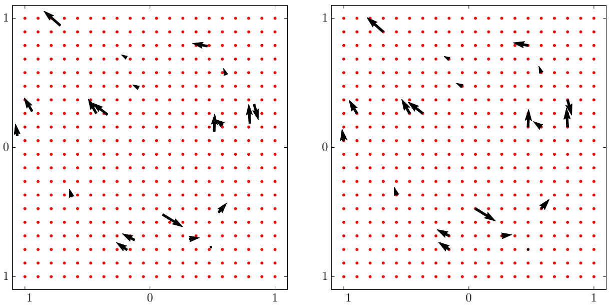
IV. Gaussian Reconstruction from Partial Data

Reconstructing the state of a swarm from partial observations using the Gaussian model described in the previous section is a three step process. First, the Gaussian model is trained on a set of complete observations of a similar swarm, for instance by observing captive fish in a well-instrumented aquarium. In the second step, the partial observation is projected onto the same grid that is used to compute the Gaussian model, thus producing a vector representation of the partial observation. In the final step, Gaussian inference is used to predict any missing values. The end result is a velocity field that describes the behavior of the swarm that produced the partial observation.

The first step, training the Gaussian model, was described in the previous section. In the second step, the normalizing transform is used to project the partial observation onto $[-1, 1] \times [-1, 1]$. Recall that the normalizing transform is a function of the shape of the entire swarm, which is not directly provided by the partial observation. If the agents visible in the partial observation are distributed uniformly across the swarm, then it is likely that the shape of the agents in the partial observation is a good estimate of the shape of the swarm as a whole. In this case, the normalizing transform can be computed directly from the position of the agents in the partial observation. However, if there is a spatial bias in the observation, such as if only one edge of the swarm is visible, more sophisticated approaches for estimating the shape of the swarm as a whole are necessary. In this paper, we assume that the shape and position are given a priori, and thus compute the normalizing transform directly from the complete frame.

Once an observation is transformed, it needs to be projected onto the grid that was originally used to compute the Gaussian model. This operation differs from the interpolation used to compute the velocity field. Instead of computing the best estimate of the velocity field at every grid point, each observed agent assigns its velocity to the nearest grid point, overwriting any previous vector (Figure 4). All other grid points are marked as unobserved. The velocity field on the grid can still be represented as a 3-tensor. However, the entries corresponding to grid points that were not assigned a velocity will not have a defined value (NaN in Python or Matlab). The 3-tensor is unfolded to form the vector representation $\vec{\xi}$ of the partial observation.

Gaussian inference is used to estimate values for the unknown entries of $\vec{\xi}$. The first step is to compute a permutation



(a) Partial Observation (b) Projection onto Grid

Fig. 4: Projection of a partial observation onto a grid. **(a)** The arrows represent the velocity of swarm agents in a partial observation after the normalizing transform. The red dots are the grid points onto which the swam will be projected. **(b)** Vector representation of the partial observation. Note that most grid points do not have a defined velocity.

matrix $\mathbf{P}$ such that

$$\begin{bmatrix} \vec{\xi}_a \\ \vec{\xi}_b \end{bmatrix} = \mathbf{P}\vec{\xi} \tag{5}$$

where $\vec{\xi}_a$ are the entries of $\vec{\xi}$ with undefined values, *i.e.* unobserved locations in the velocity field, and $\vec{\xi}_b$ are the entries with defined values. The permuted Gaussian model is given by

$$\begin{bmatrix} \vec{\mu}_a \\ \vec{\mu}_b \end{bmatrix} = \mathbf{P}\vec{\mu} \tag{6}$$

$$\begin{bmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} \\ \mathbf{\Sigma}_{ba} & \mathbf{\Sigma}_{bb} \end{bmatrix} = \mathbf{P}\mathbf{\Sigma}\mathbf{P}^{-1} \tag{7}$$

The maximum likelihood estimator for the unobserved entries of the velocity field [17] is given by

$$\vec{\xi}_a = \vec{\mu}_a + \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\left(\vec{\xi}_b - \vec{\mu}_b\right). \tag{8}$$

Once an estimate for the unobserved positions of the velocity field has been computed, the permutation can be reversed to recover the vector representation of the reconstructed velocity field

$$\vec{\zeta} = \mathbf{P}^{-1}\begin{bmatrix} \vec{\xi}_a \\ \vec{\xi}_b \end{bmatrix} \tag{9}$$

V. Validation

Two approaches are used to measure how well the reconstructed velocity field describes the true state of the swarm; the normalized reconstruction error and semantic classification of swarm behavior.

The normalized reconstruction error measures how well the velocity field computed from a partial observation matches the true velocity field of the swarm. Let $\vec{\zeta}_{partial}$ be the vector representation of the reconstructed velocity field and $\vec{\zeta}$ be the vector representation of the true velocity field. The normalized reconstruction error is

$$err_{\text{recon}}(\vec{\zeta}_{partial}) = \frac{\left|\vec{\zeta}_{partial} - \vec{\zeta}\right|}{\left|\vec{\zeta}\right|}. \tag{10}$$

where $|.|$ is the Euclidean norm. Dividing the error by the norm of the true velocity field is intended to prevent frames where the swarm moves quickly from being given excessive weight when averaged with errors from frames where the swarm moves more slowly.

The normalized reconstruction error measures how close the reconstructed velocity field is to the true swarm velocity field. However, it does not directly measure whether the reconstructed velocity field induces behavior that is qualitatively similar to that induced by the true velocity field. To more directly validate the qualitative behaviors induced by reconstructed velocity fields, we apply a behavior classification technique introduced by Tunstrøm et al. [20]. The classification of behavior is based on two values, a polarization order parameter and a rotation order parameter. The polarization order parameter is analogous to the norm of the total linear momentum of the swarm, and the rotation order parameter is analogous to the norm of the total angular momentum of the swarm about its center of mass. There are four labels: polarized, disordered[1], milling and transition. A swarm with a high polar order parameter and low rotation order parameter is labeled as polarized, which corresponds to the entire swarm translating. A high rotation order parameter and low polarization order parameter leads to milling, which describes states where the swarm rotates about its center. If both order parameters are low, then than the behavior is labeled as disordered, which describes states where there is no coherent motion in the swarm. Otherwise, the behavior is labeled as transition. In [20], the order parameters are computed as a function of the position and velocity of the individual agents. To classify the behavior induced by a discretized velocity field, we simply treat each grid point (Figure 3) as if it were a separate agent for the purposes of calculating the order parameter.

VI. Results

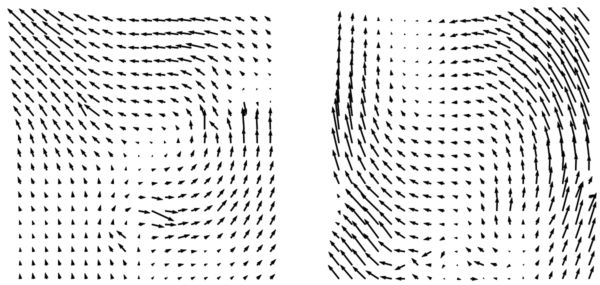
Gaussian reconstruction was tested on three biological data sets provided by Ian Couzin's group [20]. The data sets were collected by filming a school of 300 Golden Shiners, a type of fish, in a broad, shallow pool. The shallow pool means that the swarm was effectively two dimensional, although fish could pass above or below one another. The resulting video was passed through a tracking algorithm to calculate the position and velocity of each fish in the school, which is the form of the data we received. Each data set contains 100,000 frames, corresponding to collecting data for 56 minutes at 30 frames per second. One data set, the same used for algorithm development, was used to train the Gaussian model, with the other two data sets reserved for testing.

In any given frame, position and velocity are known for approximately 250 of 300 fish due to occlusions and tracking errors. While not perfect, the data is good enough to be treated as a complete observation of the swarm for the purposes of this paper. To generate partially observed data sets, the test data sets were artificially subsampled in one

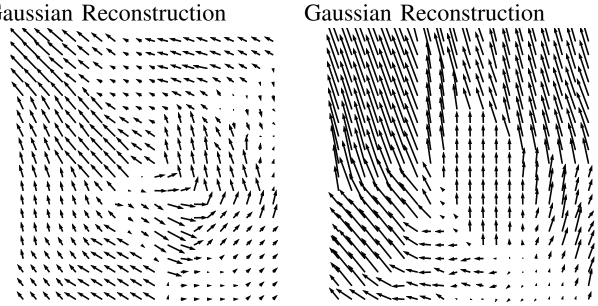
[1] In [20], the disordered label is replaced by the label "swarm"

(a) Uniform Subsampled (b) Visibility Subsampled

Fig. 5: Example of subsampled data



(a) Uniform Subsampled Gaussian Reconstruction (b) Visibility Subsampled Gaussian Reconstruction



(c) Uniform Subsampled Interpolated Reconstruction (d) Visibility Subsampled Interpolated Reconstruction

Fig. 6: Example reconstructed velocity fields from the subsampled data in Fig 5. The Gaussian reconstruction is noticeably better for the visibility subsampled data **(b)** than the interpolated reconstruction **(d)**, because the Gaussian model allows the behavior of fish on the "far" side of the swarm to be predicted.

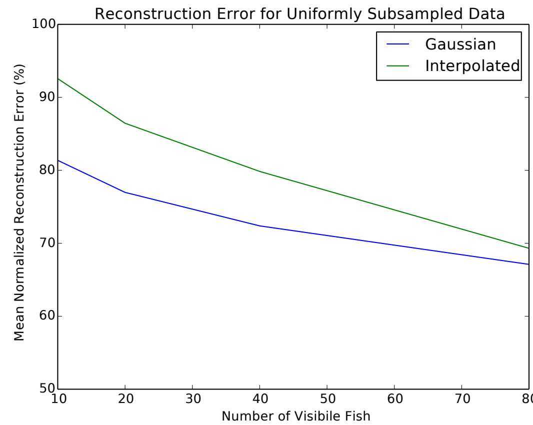


Fig. 7: Mean normalized reconstruction error for uniform subsampled data

	Reconstruction Error (%)
Gaussian	75
Interpolated	95

TABLE I: Mean normalized reconstruction error for visibility subsampled data.

of two manners; *uniformly subsampled* and *visibility subsampled*. In uniformly subsampled data, a specified number of fish are randomly selected as being visible, with the visible fish in each frame chosen independently (Figure 5a). We test subsamplings that kept 10, 20, 40, and 80 fish. Uniform subsampling is intended to model occlusions due to foreground objects or random sensor noise and tracking error. The fish marked as visible in visibility subsampled data are those that would not be occluded by another fish from a point of view of an observer infinitely far from, and in the plane of, the swarm (Figure 5b). Each fish is modeled as a circle 9 millimeters in radius. A fish is considered occluded if the center of the fish is behind another fish.

We evaluate the performance of the Gaussian reconstruction using two criteria (Section V); how well the velocity field reconstructed from partial data matches the velocity field computed from the complete swarm as measured by normalized reconstruction error, and whether the behavior label computed for the reconstructed velocity field matches the behavior label for the full data. As a point of comparison, we compute the velocity field at each point on a grid by running k-NN interpolation directly on the partial data, which we refer to as the interpolated reconstruction. Velocity fields reconstructed using Gaussian and interpolated reconstruction are given in Fig 6. For baseline behavior classification, we compute both the label for the interpolated reconstruction of the velocity field, referred to as the interpolated classification, and compute a behavior label directly from the partial data, without first computing a velocity field. The later approach is referred to as the subsampled classification.

We start by comparing how well the velocity fields reconstructed from partial data match the velocity fields computed for the full data. For uniformly subsampled data (Figure 7), Gaussian reconstruction produces lower normalized reconstruction error than the interpolated reconstruction. As the number of visible fish increases, the benefit of Gaussian reconstruction over interpolated reconstruction decreases, as is to be expected. Gaussian reconstruction has a greater performance advantage over interpolated reconstruction when run on visibility subsampled data. In visibility subsampled data, only the bottom of the swarm is visible. Gaussian reconstruction has an internal model of swarm behavior which can be used to predict the behavior of the far side of the swarm, while interpolated reconstruction assumes that the behavior of the far side of the swarm is similar to that of the near side.

We next compare the ability of different approaches to apply the correct semantic label to partial observations. Recall that the classification scheme of Tunstrøm et al. [20] has four labels; polar, milling, disordered, and transition.

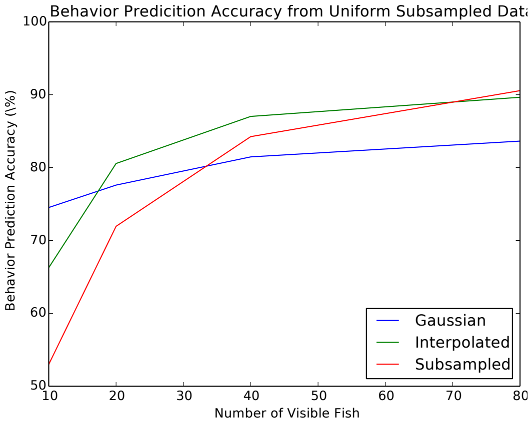


Fig. 8: Behavior classification accuracy for behaviors reconstructed from uniformly subsampled partial data

Predicted	True Behavior Polar	Milling	Disordered
Polar	5245	994	1479
Milling	848	83919	5869
Disordered	127	315	2041
Transition	4196	10507	10456
accuracy (%)	50	88	10

(a) Behaviors from Gaussian classification

Predicted	True Behavior Polar	Milling	Disordered
Polar	6515	1491	2633
Milling	17	21588	460
Disordered	119	64	3744
Transition	3765	72592	13008
accuracy (%)	63	23	19

(b) Behaviors from interpolated classification

Predicted	True Behavior Polar	Milling	Disordered
Polar	5438	1433	586
Milling	1	330	42
Disordered	202	66	9379
Transition	4775	93906	9838
accuracy (%)	52	0	47

(c) Behaviors from subsampled classification

TABLE II: Confusion matrices for classifying swarm behavior from visibility subsampled data. Each entry gives the number of times a frame of data with the true behavior specified by the column was classified as the behavior specified by the row. The last row gives the overall accuracy.

There are a significant number of data frames that are labeled transition that a human would classify as one of the three other labels. However, the frames assigned one of the other three labels clearly belong to the labeled class. Therefore, when evaluating the performance of reconstruction approaches only frames whose label is not transition are considered.

Classification based on Gaussian reconstruction, *i.e.* Gaussian classification, provides better classification accuracy than interpolated or subsampled classification on uniformly subsampled data when only ten fish are visible (Figure 8). As the number of visible fish increases and the importance of estimating missing data decreases, Gaussian classification is outperformed by the alternate approaches because the Gaussian model introduces some bias. On the visibility subsampled data, Gaussian classification is the only approach that can reliably identify milling behavior (Table II). This is to be expected, as only one side of the swarm is visible, and only Gaussian classification has a model that allows it to predict the behavior of the far side of the swarm. All three approaches have similar success rates at classifying polar behavior, but only subsampled classification can reliably identify disordered behavior. This is also not surprising, because Gaussian classification and interpolated classification both perform a velocity field reconstruction step, that will tend to smooth out disordered states.

VII. Conclusions

In this paper, we present a method for representing the global behavior of a swarm using velocity fields. A Gaussian model can then be learned for swarm velocity fields, and used to infer the behavior of the entire swarm from partial observations. We validate Gaussian reconstruction on biological data, comparing the quality of the reconstructed velocity fields to a simple interpolation approach. The performance of Gaussian classification at predicting behavior labels from partial data is compared to both computing a label for an interpolated velocity field and computing a behavior labeled directly from the visible agents. Gaussian classification provides better performance than the alternatives in each case, as long as the number of observed agents is sufficiently small.

The largest limitation to Gaussian reconstruction is that an estimate of the size and center of the swarm must be provided *a priori*. In our future work, we will develop techniques to track and predict the position and shape of the swarm from partial data to facilitate utilization of Gaussian reconstruction.

The current approach considers a single frame of data. Performance could likely be improved by aggregating multiple observations using filtering technique.

References

[1] Bjørn Erik Axelsen, Tycho Anker-Nilssen, Petter Fossum, Cecilie Kvamme, and Leif Nøttestad. Pretty patterns but a simple strategy: predator-prey interactions between juvenile herring and Atlantic puffins observed with multibeam sonar. *Canadian Journal of Zoology*, 79(9):1586–1596, 2001.

[2] Spring Berman, Vijay Kumar, and Radhika Nagpal. Design of control policies for spatially inhomogeneous robot swarms with application to commercial pollination. In *2011 IEEE International Conference on Robotics and Automation*, pages 378–385. Ieee, May 2011.

[3] Dan Brown and Michael A Goodrich. Limited Bandwidth Recognition of Collective Behaviors in Bio-Inspired Swarms. *13th International Conference on Autonomous Agents and Multiagent Systems*, pages 405–412, 2014.

[4] Andrea Cavagna, Alessio Cimarelli, Irene Giardina, Giorgio Parisi, Raffaele Santagati, Fabio Stefanini, and Raffaele Tavarone. From Empirical Data To Inter-Individual Interactions: Unveiling the Rules of Collective Animal Behavior. *Mathematical Models and Methods in Applied Sciences*, 20(supp01):1491–1510, September 2010.

[5] Iain Couzin, Jens Krause, Richard James, Graeme D. Ruxton, and Nigel R Franks. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, pages 1–11, 2002.

[6] David Gross. Recovering low-rank matrices from few coefficients in any basis. *Information Theory, IEEE Transactions on*, 57(3):1548–1566, March 2011.

[7] Daniel Grünbaum, Karen Chan, Elizabeth Tobin, and Michael T Nishizaki. Non-linear advection-diffusion equations approximate swarming but not schooling populations. *Mathematical Biosciences*, 214(1-2):38–48, 2008.

[8] Qifa Ke and Takeo Kanade. Robust L Norm Factorization in the Presence of Outliers and Missing Data by Alternative Convex Programming. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 739–746. Ieee, 2005.

[9] Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from noisy entries. *The Journal of Machine Learning Research*, 11:2057–2078, 2010.

[10] R Kilgour and TH Scott. Leadership in a herd of dairy cows. In *Proceedings of the New Zealand Society of Animal Production*, pages 36–43, 1959.

[11] Jerrold E Marsden, Tudor Ratiu, and Ralph Abraham. *Manifolds, tensor analysis, and applications*. Springer-Verlag Publishing Company, Inc, third edit edition, 2004.

[12] Jennifer M Miller, Allison Kolpas, Joao Plinio Juchem Neto, and Louis F Rossi. A continuum three-zone model for swarms. *Bulletin of mathematical biology*, 74(3):536–61, March 2012.

[13] L Nøttestad and BE Axelsen. Herring schooling manoeuvres in response to killer whale attacks. *Canadian Journal of Zoology*, 77(10):1540–1546, 1999.

[14] P Oliveira. Interpolation of signals with missing data using PCA. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 828–831, Toulouse, 2006.

[15] Julia K Parrish, Steven V Viscido, and Daniel Grünbaum. Self-Organized Fish Schools: An Examination of Emergent Properties. *The Biological bulletin*, 202(3):296–305, June 2002.

[16] DV Radakov. *Schooling in the ecology of fish*. Halsted Press, New York, New York, USA, 1973.

[17] Matthew Tesch. *Improving Robot Locomotion Through Learning Methods for Expensive Black-Box Systems*. PhD thesis, Carnegie Mellon University, 2014.

[18] John Toner and Yuhai Tu. Long-range order in a two-dimensional dynamical XY model: how birds fly together. *Physical Review Letters*, 75(23):4326–4329, 1995.

[19] Chad M Topaz, Andrea L Bertozzi, and Mark a Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of mathematical biology*, 68(7):1601–23, October 2006.

[20] Kolbjørn Tunstrøm, Yael Katz, Christos C. Ioannou, Cristián Huepe, Matthew J. Lutz, and Iain D. Couzin. Collective States, Multistability and Transitional Behavior in Schooling Fish. *PLoS Computational Biology*, 9(2):e1002915, February 2013.

[21] T Vicsek, A Czirók, E Ben-Jacob, I Cohen, and O Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 1995.