GAUSSIAN PROCESS DYNAMIC MODELING OF BAT FLAPPING FLIGHT

Matthew Bender[1]*, Xu Yang*[2]*, Hui Chen*[2]*, Andrew Kurdila* [1]*, and Rolf Müller*[1,3]

[1]Department of Mechanical Engineering, Virginia Tech, Blacksburg, Virginia, USA
[2]School of Information Science and Engineering, Shandong University, Jinan, China
[3]Shandong University – Virginia Tech International Laboratory, Jinan, China

ABSTRACT

The flapping flight of bats can serve as an inspiration for flapping-wing air vehicles. Obtaining an understanding of bat flight requires detailed, occlusion-free kinematics data that can only be collected using large numbers of cameras. Here, we have explored the use of low-cost cameras with low frame rates that result in nonlinear, large-baseline motions in image space. To create a better model for predicting the motion of features under these circumstances, we have applied Gaussian Process Dynamic Modeling (GPDM) to manually digitized flight data in order to learn a lower dimensional manifold near which the motion evolves. The primary contribution of this work is the first nonlinear dimensionality reduction for the representation of bat flight.

Keywords: Gaussian Process Dynamic Models, Motion Capture, Dimensionality Reduction, Manifold Estimation

1. INTRODUCTION

Research in flapping flight has gained new attention in recent years [1, 2, 3]. Bats are particularly agile, stable, and efficient compared to other flapping-wing fliers which makes them excellent models for the design of flapping-wing air vehicles. Their flight characteristics, however, are due to a highly articulated skeleton and deformable wing membranes which are difficult to study via imaging due to frequent incidents of self-occlusion. Such occlusions can be avoided by increasing the number of cameras to cover many different vantage points. In order to be able to accomplish this, the current work has looked at low-cost/low-frame rate cameras. The low frame rates pose a challenge because they result in large-baseline nonlinear motions of points in the image that cannot be captured by image space trackers such as optical flow that have been used in prior work [4, 5]. However, methods for improving estimation and correspondence for such image sets have been presented previously. One such method for solving correspondence is to bootstrap correspondence with the state estimation. In [6], the authors bootstrap optical flow with a Square Root Unscented Kalman Filter to perform tracking. Because the motion of the bat is ultimately unknown, the authors assume constant acceleration of all generalized joint coordinates. The authors show that this method works well for their experiments, however, the videos were recorded at frame rates over 250 frames per second (fps). For our experiments which are conducted with much lower frame rate cameras (120fps), a better motion model is needed.

One method for determining a representation of motion is to learn a model using dimensionality reduction. In [7], principle component analysis (PCA) is used to determine the number of modes required to reconstruct bat flight. The authors report that $\sim$13 modes are required to capture 95% of the motion. This reduction, from 56 skeletal DOF to 13 PCA components, is substantial, but the representation is still linear. Our work aims to learn a low-dimensional nonlinear manifold to represent bat flight for the bat species used in our experiments. The advantage of learning such a model is that fewer degrees of freedom should be able to represent the motion while still representing the motion adequately.

While nonlinear dimensionality reduction of bat flight has not been studied, such analysis is popular in human motion studies. Applications include avatar and image pose synthesis for computer simulations, human robot trajectory planning, and studying musculo-skeletal pathologies. This paper uses the methods presented in [8, 9, 10] which develop Gaussian Process Dynamic Models (GDPM) for studying human motion. GDPM is an extension of Gaussian Process Latent Variable Models (GPLVM, [11]). The GDPM method is designed to preserve temporal continuity in the data. This attribute is attractive when a limited set of training data available. GPLVM can also be used to learn a manifold, but because continuity is not preserved, more training data is required to ensure that there are no gaps in the manifold.

The purpose of this work is to learn a dynamic model of bat flight which can be used as a more accurate motion prior than random walk or PCA models. To the best of our knowledge, this is the first nonlinear dimensionality reduction approach for bat flight. To begin, the kinematic model is described in Section 2. The methods for estimating joint states from images are detailed in [12] and will not be recreated in this paper. Section 3 presents the theory for performing GPDM on motion capture data. The experimental methods and identified model are presented in Section 4. Finally, conclusions and recommendations for future work are presented in Section 5.

2. KINEMATIC MODEL

The bat skeleton can be modeled as a multibody system having the connectivity of a topological tree: each limb of the tree forms a kinematic chain with rigid links as in [12, 13]. The degrees of freedom within this skeleton are modeled using the Denavit-Hartenberg (DH) convention. Diagrams of each DOF and complete DH tables are included in the aforementioned references, so they will not be repeated here. While the model used in this paper assumes that the system is constructed from a collection of kinematic chains, there are two major differences between the model used in this paper and the model used previ-

ously. First, the body displacements and rotations are modeled using the Denavit-Hartenberg convention instead of using x-y-z translations and yaw-pitch-roll angles. The second difference is that the body is not assumed to be a single rigid body, but two rigid bodies connected by a single rotational degree of freedom at the base of the tail. Thus, the shoulders of the bat can can move laterally with respect to one another. We have found through numerous system identifications that, adding this DOF reduces the noise in the base body position and orientation. These modifications of the DH table are straightforward and are left to the reader.

3. GAUSSIAN PROCESS DYNAMIC MODELS

This study uses a framework for generating a GDPM from experimental data which was originally developed for tracking human walking. The model is learned using previously developed code [14], but the main principles are explained in this section. Given a sequence of data $\mathbf{Y} := \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ ... \ , \boldsymbol{y}_K\}$ which is assumed to be zero mean along each dimension where $\boldsymbol{y}_k \in \mathbb{R}^N \quad \forall k \in [1, 2, ..., K]$, we want to determine the model

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}, \mathbf{A}) + \boldsymbol{n}_{x,k} \tag{1}$$

$$\boldsymbol{y}_k = \boldsymbol{g}(\boldsymbol{x}_k, \mathbf{B}) + \boldsymbol{n}_{y,k} \tag{2}$$

where, $\boldsymbol{x}_k \in \mathbb{R}^D$ is a D-dimensional latent space which supports the dynamics, $\boldsymbol{f}$ is a function such that $\mathbb{R}^D \times \mathbb{R}^{D\times K} \rightarrow \mathbb{R}^D$, $\boldsymbol{g}$ is a mapping such that $\mathbb{R}^D \times \mathbb{R}^{D\times K} \rightarrow \mathbb{R}^N$, and $\mathbf{A}$ and $\mathbf{B}$ are matrices of hyper parameters. We assume that

$$\boldsymbol{f}(\boldsymbol{x}, \mathbf{A}) = \sum_i \boldsymbol{a}_i \phi_i(\boldsymbol{x}), \tag{3}$$

$$\boldsymbol{g}(\boldsymbol{x}, \mathbf{B}) = \sum_j \boldsymbol{b}_j \psi_j(\boldsymbol{x}), \tag{4}$$

where $\mathbf{A} := \begin{bmatrix}\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\end{bmatrix}$ and $\mathbf{B} := \begin{bmatrix}\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots\end{bmatrix}$ are matrices of scaling parameters and $\phi(\boldsymbol{x})$ and $\psi(\boldsymbol{x})$ are scalar output functions such that $\mathbb{R}^D \rightarrow \mathbb{R}$.

Some authors use expectation maximization techniques to fit this regression model to the data, i.e., they solve for $\mathbf{A}$ and $\mathbf{B}$. The difficulty in doing this is that the solution for these parameters can be specific to a specimen or a trial, and the solution is not applicable to similar systems. Instead, authors in [8, 9, 10] formulate the joint probability of the model and the data by marginalizing over the parameters. Therefore, the learned model becomes useful for predicting motions. To perform this optimization, we formulate the joint probability of latent variables $\mathbf{X}$, collected data $\mathbf{Y}$, hyperparameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, and a weighting matrix $\mathbf{W}$ as

$$\begin{aligned} &p(\mathbf{Y}, \mathbf{X}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{W}) \\ &\quad = p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \mathbf{W}) p(\mathbf{X}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\boldsymbol{\beta}) p(\mathbf{W}). \end{aligned} \tag{5}$$

The first term on the right hand side is obtained by placing a Gaussian prior on the columns of $\mathbf{B}$ and marginalizing $\boldsymbol{g}$. This eliminates the dependency of the model on the parameters, and only hyper-parameters must be identified. The marginalization can be done in closed form [15] to obtain

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \mathbf{W}) = \frac{|\mathbf{W}|^K}{\sqrt{(2\pi)^{KN}|\mathbf{K}_y|^N}} * \\ \exp\left(-\frac{1}{2}\text{trace}\left(\mathbf{K}_y \mathbf{Y}\mathbf{W}^2\mathbf{Y}^T\right)\right), \end{aligned} \tag{6}$$

where $\mathbf{W} := \text{diag}(\begin{bmatrix}w_1, w_2, ...w_N\end{bmatrix})$ serves to weight $\mathbf{Y}$ so that dimensions which are large in magnitude do not dominate the optimization. Additionally, $\mathbf{K}_y$ is the covariance kernel with hyper-parameters $\boldsymbol{\beta}$. This matrix is assembled as

$$\begin{aligned} (\mathbf{K}_y)_{ij} = \beta_1 \text{exp}\left(-\frac{\beta_2}{2}||\boldsymbol{x}_i - \boldsymbol{x}_j||^2\right) + \\ \beta_3^{-1}\delta(\boldsymbol{x}_i, \boldsymbol{x}_j). \end{aligned} \tag{7}$$

This equation quantifies the "closeness" of two points in the latent space.

The second term on the right hand side of equation 5 describes the dynamics on the latent space. This probability is formulated as

$$p(\mathbf{X}|\boldsymbol{\alpha}) = \int p(\mathbf{X}|\mathbf{A}, \boldsymbol{\alpha}) p(\mathbf{A}|\boldsymbol{\alpha}) d\mathbf{A}. \tag{8}$$

Formulating the probability distribution in this manner removes the dependency on the parameters in $\mathbf{A}$ with the smaller set of hyper-parameters contained in $\boldsymbol{\alpha}$. Assuming a Gaussian prior on $\mathbf{A}$, and assuming the process is Markovian, this probability can also be computed in closed form as

$$\begin{aligned} p(\mathbf{X}|\boldsymbol{\alpha}) = \frac{p(\boldsymbol{x}_1)}{\sqrt{(2\pi)^{(K-1)D}|\mathbf{K}_x|^D}} * \\ \exp\left(-\frac{1}{2}\text{trace}\left(\mathbf{K}_x^{-1}\mathbf{X}_{2:K}\mathbf{X}_{2:K}^T\right)\right). \end{aligned} \tag{9}$$

Note that the weighting matrix is not required here because the latent variables are nondimensional. To define the covariance function we choose the linear plus radial basis function kernel

$$\begin{aligned} (\mathbf{K}_x)_{ij} = \alpha_1 \text{exp}\left(-\frac{\alpha_2}{2}||\boldsymbol{x}_i - \boldsymbol{x}_j||^2\right) + \\ \alpha_3 \boldsymbol{x}_i^T \boldsymbol{x}_j + \alpha_4^{-1}\delta(\boldsymbol{x}_i, \boldsymbol{x}_j). \end{aligned} \tag{10}$$

The only terms that remain in the joint distribution in equation 5 are the priors on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. These serve to constrain the hyper-parameters and the weight matrix $\mathbf{W}$.

This probability distribution can be maximized by minimizing the negative log likelihood of the distribution

$$\begin{aligned} L = \frac{D}{2}\ln||\mathbf{K}_x|| + \frac{N}{2}\ln||\mathbf{K}_y|| - K\ln|\mathbf{W}| + \\ \frac{1}{2}\text{trace}(\mathbf{K}_x^{-1}\mathbf{X}_{2:K}\mathbf{X}_{2:K}^T) + \frac{1}{2}\boldsymbol{x}_1^T\boldsymbol{x}_1 + \\ \frac{1}{2}\text{trace}(\mathbf{K}_y^{-1}\mathbf{Y}\mathbf{W}^2\mathbf{Y}^T) + \\ \sum_j \ln(\beta_j) + \frac{1}{2\kappa^2}\text{trace}(\mathbf{W}^2) + \sum_j \ln\alpha_j. \end{aligned} \tag{11}$$

There are many proposed methods for solving this equation: maximum a posteriori (MAP) estimation, Balanced GPDM methods, manually specify hyper-parameters, or two stage MAP. We use Balanced GPDM where the hyper-parameters are initialized and held fixed while latent positions are optimized. Hyper-parameters are then refined while holding latent positions fixed. This process iterates until an error threshold is reached or a maximum number of iterations (100) occurs.

After optimizing hyper-parameters in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, we propagate dynamics in the learned latent space using the distribution

$$\tilde{\boldsymbol{x}}_k \sim \mathcal{N}\left(\mu_X(\tilde{\boldsymbol{x}}_{t-1}); \sigma_X^2(\tilde{\boldsymbol{x}}_{t-1})\mathbf{I}\right), \tag{12}$$

with

$$\mu_X(\boldsymbol{x}) = \mathbf{X}_{out}^T \mathbf{K}_x^{-1} \boldsymbol{k}_X(\boldsymbol{x}), \quad (13)$$

$$\sigma_X^2(\boldsymbol{x}) = k_X(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_X(\boldsymbol{x})^T \mathbf{K}_x^{-1} \boldsymbol{k}_X(\boldsymbol{x}). \quad (14)$$

These dynamics can be used for simulation of trajectories in the latent space and compared to those identified from the experimental data. Once trajectories have been simulated in latent space, they can be projected into feature space to generate simulated motions similar to the training data.

To project latent trajectories back into feature space, we can write

$$\mu_Y(\boldsymbol{x}) = \mathbf{Y}^T \mathbf{K}_y^{-1} \boldsymbol{k}_Y(\boldsymbol{x}), \quad (15)$$

$$\sigma_Y^2(\boldsymbol{x}) = k_Y(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_Y(\boldsymbol{x})^T \mathbf{K}_y^{-1} \boldsymbol{k}_Y(\boldsymbol{x}). \quad (16)$$

Note that the projected trajectories will be zero mean due to the subtraction of the mean applied to the original data.

4. RESULTS

Data collection was performed at the Shandong University - Virginia Tech International Laboratory in Jinan, China. The bat species used in the experiments was Pratt's roundleaf bat (*Hipposideros pratti*). The experimental setup used 21 Hero 3+ Black cameras (GoPro, Inc., San Mateo, CA), and a modified version of a micro-controller-based synchronization system (Genlock Dongle and MewPro, Orangkucing Labs, Tokyo, Japan). The synchronization system was used to trigger all cameras at the same time (precision: $\frac{1}{720}$ of a frame). The frame rate was set at 120fps and the resolution to 720x1280. White markers were placed in a grid pattern across the body and wings of the bat to capture skeletal and membrane motion during flight. Figure 1 shows a typical image captured using the system described above. Straight and level flight is the only regime presented in this paper.

Fig. 1. An example of a single video frame showing a bat in flight. White markers are applied to the wings for point-based reconstruction. Direction of flight is from right to left.

To process the motion capture data, robust methods for identifying feature points have been developed [16]. Correspondence of the points, however, is still established manually. After the points are correlated, sensor fusion is performed using a spatially recursive Kalman filter developed in [13] to estimate the position and pose of the skeleton. A latent space is then identified for all joint coordinates except for the torso displacement. This generates a model which is independent of the specific path of the bat. To perform optimization of equation 11 we initialize the latent variables, $\mathbf{X}$, by using the first D modes of a PCA decomposition of the original data, $\mathbf{Y}$.

The identified manifolds are shown in Figure 2. In the figure, all manifolds use the RBF plus linear kernel from Equation 10 on the dynamics and the RBF kernel in equation 7 for the mapping between latent space and feature space. Subfigures (a), (b), and (c) assume 2, 3, and 4 latent DOFs, respectively. The blue points are the identified latent trajectories, red points are the simulated trajectories, and green points are HMC samples on the simulated trajectories to represent the uncertainty in the manifold. Arrows indicate the direction of motion. Ideally, the manifold should appear to be cyclic and have little variation between cycles.

Figure 2a shows the 2 DOF manifold. It displays large jumps in the trajectory that are not consistent from cycle to cycle. The manifold in Figure 2b shows a smooth manifold with little uncertainty which is demonstrated by the HMC sample's close proximity to the simulated trajectory. However, the motion is only semi-periodic: the motion is approximately cylindrical, but there are variations in diameter and height. Finally, Figure 2c shows the 4 DOF latent space. This manifold shows substantial uncertainty in the third DOF and the motion in the fourth DOF is not cyclic. Ultimately the only way to evaluate the consistency of the manifold is to project the predictions back into feature space.

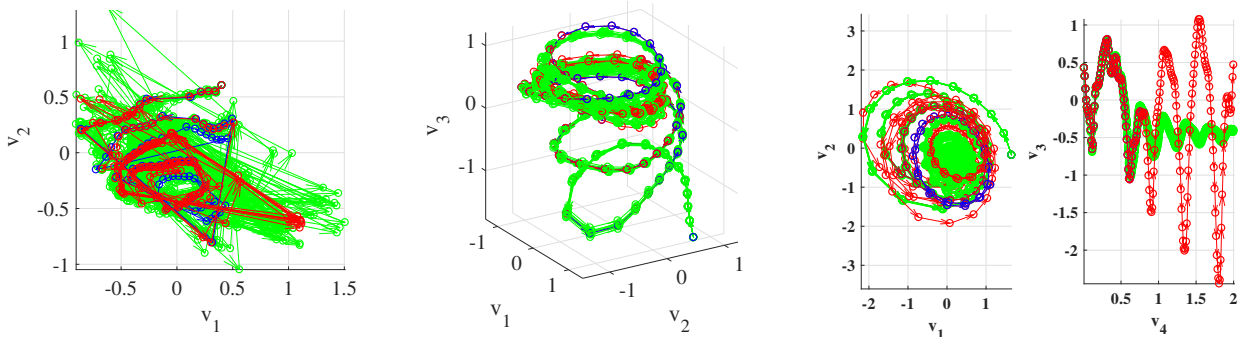
This projection is shown in Figure 3. The predictions (solid line) match the experimental data (dots) very closely. Note that the dark blue cycles for $\theta_1 - \theta_3$ are larger in amplitude than the remaining cyan cycles. This behavior is likely the cause of the semi-periodic latent space trajectories discussed before. Additionally, the extrapolated data (magenta) appears to be periodic repetitions of the cyan portion of the experimental data.

Although the latent coordinates project correctly into feature space, it is still important to ensure that the semi-periodicity of the identified manifold is consistent with the data. If the bat performed each flap cycle exactly the same, there should be no deviation in the latent space. Thus, we posit that the bat is not performing steady, straight and level flight. Upon inspection of body velocity, we can see that the bat is accelerating steadily for the first 0.3s of the experiment and is coasting for the remainder. The change in velocity is illustrated in Figure 4. The large amount of noise is induced by the first order derivative used to calculate velocity. The top of Figure 4 shows the body velocity which is color corresponded to the two manifold views on the bottom. In the blue section, the bat is accelerating. In the cyan section, the bat is coasting. There are two clearly different segments to the manifold.

5. CONCLUSIONS

The work presented here has used Gaussian Process Dynamic Models (GPDM) to learn both a latent space representation for bat flight dynamics and a mapping from latent space back to joint space. This is—to the authors' knowledge—the first nonlinear dimensionality reduction of bat flapping flight. We have successfully identified a model which closely resembles the experimental data provided and produces plausible synthesized motions.

In the future, this model will be used in conjunction with an estimation strategy that respects the geometry of the identified manifold. This will improve the accuracy of predicted



(a) 2 Lat DOF (b) 3 Lat DOF (c) 4 Lat DOF

Fig. 2. Latent Space Visualization. Blue points are the learned latent coordinates, red points are simulated latent coordinates, and green points are fair HMC samples from the distribution. Arrows indicate the direction of motion.

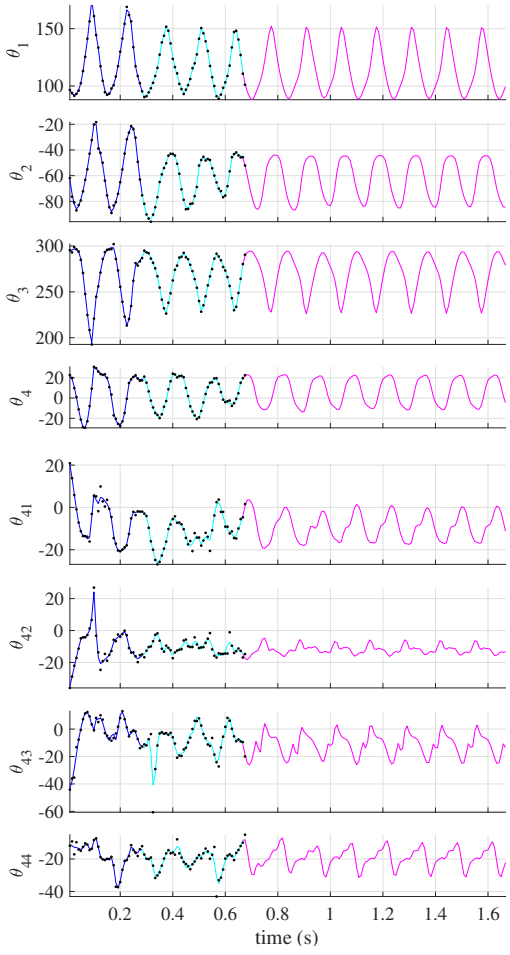


Fig. 3. Feature Space Projection of Latent Trajectories. θ_1 through θ_4 are shoulder and elbow rotations for the right wing. $\theta_{41} - \theta_{44}$ are second phalanx rotations on the right hand.

marker locations substantially and should help improve the accuracy with which feature correspondences can be established. Furthermore, it is very difficult to induce a bat to perform exactly the same maneuver more than once. The authors would like to determine a way to identify small deviations from a typical flap cycle so that a periodic manifold can be identified from semi-periodic data.

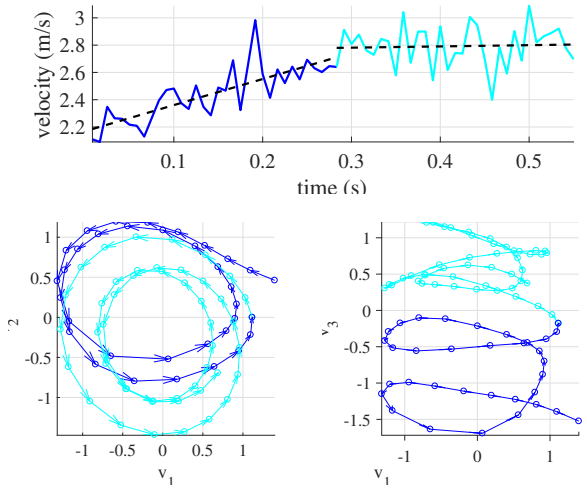


Fig. 4. [Top] Magnitude of Body Velocity. [Bottom] Two views of the manifold. The bat is clearly accelerating (dark blue) for half the test and coasting (cyan) for the remainder.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge the following funding sources for their support: National Natural Science Foundation of China (grant numbers 11374192 and 11574183); Fundamental Research Fund of Shandong University (grant no. 2014QY008); Minister of Education of China Tese grant for faculty exchange; U.S. National Science Foundation (grant no. 1510797); and Virginia Tech Institute for Critical Technology and Applied Science (ICTAS, through support for the BIST Center).

The authors would also like to thank the following individuals for their support in data collection and processing: Li Tian, Xiaozhou Fan, Chenhao Wang, Mengfan Wang, Yuxian Ye, Yuxiang Zhu, Xizhe Ding, Yang Shao, Yiwei Jiang, Kefan Chen, Yanan Zhao, Derek Liu.

7. REFERENCES

[1] S.A. Ansari, R. Zbikowski, and K. Knowles, "Aerodynamic modelling of insect-like flapping flight for micro air vehicles," *Progress in Aerospace Sciences*, vol. 42, no. 2, pp. 129–172, 2006.

[2] W. Shyy, H. Aono, S.K. Chimakurthi, P. Trizila, C.-K. Kang, C.E.S. Cesnik, and H. Liu, "Recent progress in flapping wing aerodynamics and aeroelasticity," *Progress in Aerospace Sciences*, vol. 46, no. 7, pp. 284–327, oct 2010.

[3] Wei Shyy, Hikaru Aono, Chang-kwon Kang, and Hao Liu, *An Introduction to Flapping Wing Aerodynamics*, Cambridge University Press, 2013.

[4] Xiaodong Tian, Jose Iriarte-Diaz, Kevin Middleton, Ricardo Galvao, Emily Israeli, Abigail Roemer, Allyce Sullivan, Arnold Song, Sharon Swartz, and Kenneth Breuer, "Direct measurements of the kinematics and dynamics of bat flight.," *Bioinspiration & biomimetics*, vol. 1, no. 4, pp. S10–S18, 2006.

[5] D K Riskin, J Iriarte-Díaz, K M Middleton, K S Breuer, and S M Swartz, "The effect of body size on the wing movements of pteropodid bats, with insights into thrust and lift production.," *The Journal of Experimental Biology*, vol. 213, pp. 4110–4122, 2010.

[6] Attila J. Bergou, Sharon Swartz, Kenneth Breuer, and Gabriel Taubin, "3D reconstruction of bat flight kinematics from sparse multiple views," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1618–1625, 2011.

[7] Daniel K. Riskin, David J. Willis, Jose Iriarte-Díaz, Tyson L. Hedrick, Mykhaylo Kostandov, Jian Chen, David H. Laidlaw, Kenneth S. Breuer, and Sharon M. Swartz, "Quantifying the complexity of bat wing kinematics," *Journal of Theoretical Biology*, vol. 254, no. 3, pp. 604–615, 2008.

[8] Jack M Wang, David J Fleet, Senior Member, and Aaron Hertzmann, "Gaussian Process Dynamic Models for Human Motion," vol. 30, no. 2, pp. 283–298, 2008.

[9] Jack Wang, David Fleet, and Aaron Hertzmann, "Gaussian process dynamical models," *Advances in Neural Information Processing Systems*, pp. 1441–1448, 2006.

[10] R Urtasun, J D Fleet, and P Fua, "Gaussian Process Dynamical Models for 3D people tracking," , no. June, 2006.

[11] N.D. Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data," *Computer*, vol. 16, no. 5, pp. 329–336, 2004.

[12] Matt J. Bender, Hunter M. McClelland, Andrew Kurdila, and Rolf Müller, "Recursive Bayesian Estimation of Bat Flapping Flight Using Kinematic Trees," *AIAA Modeling and Simulation Technologies Conference*, , no. January, pp. 1–12, 2016.

[13] Matthew J Bender, Hunter G McClelland, Andrew Kurdila, and Rolf Müller, "Spatially Recursive Unscented Kalman Filter for Flapping Flight Motion Estimation," *AIAA Journal of Guidance Navigation and Control*, , no. In Review, 2017.

[14] Jack M Wang, David J Fleet, and Aaron Hertzmann, "Gaussian Process Dynamic Models MatLab Code," 2008.

[15] David J C Mackay, *Information Theory , Inference , and Learning Algorithms*, Cambridge University Press, Cambridge, England, 7.2 edition, 2003.

[16] Yousi Lin, Yang Xu, Hui Chen, Matthew J Bender, A Lynn Abbott, and R Müller, "Optimal Threshold and LoG Based Feature Identification and Tracking of Bat Flapping Flight," in *WACV 2017: IEEE Winter Conference on Applications of Computer Vision*, 2017, pp. 1–9.