Deer Repellent Project Primer

Akhil Vasvani
Aurin Chakravarty
Dan Johnson
Darren Forbes
Grant Kitchen
Justin Pokotylo
Luke Robinson
Seal-Bin Han
Vivek Netrakanti

Deer-Repellent Project Summary and Primer

Abstract - In this paper, we present a summary of the modifications to the current deer-repellent hardware and training algorithm that we would like to modify, but did not have the ability to implement during the semester. We hope that this primer can be used as a way to introduce the next group to this project as well as present some interesting ideas that they can work off of and implement.

I. Introduction

Presently, the most common technique used for gait imaging and recognition employs one or more infrared video cameras to record a target in a monitored area. This technique, however, requires large memory size and fast processing speed to process the data obtained from the camera. Additionally, privacy concerns involved with camera-based systems in public spaces is a barrier to deployment.

Animals such as bats and dolphins have developed a bi-sonar system in which they can navigate dark habitats. The auditory system in bats allows them to detect small changes in frequency of echoes relative to the original pulse sent out. They can then can extract information regarding relative velocity of a moving object.

Inspired by the auditory system in bats, Zhang et al. prototyped an active acoustic sensing system for imaging moving objects, like a walking human or deer. Acoustic surveillance systems offer many advantages over electromagnetic radars: lower cost, easier signal processing, and immunity from electromagnetic interference.

II. Principle of Operation

Waves emanating from an object moving relative to an observer can exhibit frequency dilations. This principle is known as the *Doppler effect*. If the object contains moving parts, each part has its own Doppler shift related to the part's radial velocity component with respect to the receiver. All of the scattered waves are additive, so the resulting modulation is a superposition of the individual components. This is known as the *micro-Doppler effect*.

In the simplest case, the Doppler effect involves a source which emits or scatters a wave, and an observer, which detects the wave emitted by the source. The source and the observer are simple point masses and the emitted wave is a constant tone with wavelength λ_{source} and frequency f_{source}, which are related by

$$\lambda_{\text{source}} = \frac{c_s}{f_{\text{source}}}.$$

c_s is the speed of sound in the medium through which the wave propagates. A dilation occurs when either the source or observer are moving relative to the medium. Doppler modulation contributions from the motion of the source and the motion of the observer are not symmetric.

Consider the case of a stationary source and a moving observer with velocity $v_{observer}$. $v_{observer}$ is only the radial component of the observer's velocity relative to the source. The distance between consecutive wavefronts is no longer λ_{source} because $v_{observer}$ moves the observer closer to or farther from the source. The second wavefront travels a shorter or longer distance before it arrives at the observer. The resulting frequency dilation of the observed signal is

$$f_{\text{observer}} = \frac{c_s + v_{\text{observer}}}{c_s} f_{\text{source}}.$$

This is the dilation due to the Doppler effect for a stationary source and moving observer.

Now consider the case of a moving source and stationary observer. Let v_{source} be the radial component of the source's velocity relative to the observer. When a source is moving, the wavefronts will be crowded together in front of the source and spread out behind it. This occurs because after a wavefront is emitted, the source moves throughout the duration time period before the next wavefront is emitted. The resulting frequency dilation of the observed signal is

$$f_{\text{observer}} = \frac{c_s}{c_s - v_{\text{source}}} f_{\text{source}}.$$

This is the dilation due to the Doppler effect for a moving source and stationary observer.

Monostatic Sonar

Active sonar can use the Doppler effect to measure the velocity of a target. It can record the scattered waveforms and the difference between the transmitted and received frequencies can be measured. The radial velocity component of the target object can then be estimated from this frequency difference.

Regarding the setup, the transmitter and receiver are located at the same coordinates. Assuming a moving target and a stationary sonar, the Doppler phenomenon occurs twice; once as the transmitted signal arrives at the target and once as the target scatters the signal back to the receiver. The dilation of the wave from the first leg involves a stationary source, the sonar transmitter, and a moving observer, the target (Equation 1). The dilation of the wave from the second leg involves a moving source, the target, and a stationary observer, the sonar receiver (Equation 2). The total Doppler shift is found by considering the dilated frequency from the first leg as the source frequency for the second leg. This results in a dilation observed by the monostatic sonar is given by

$$f_{\text{receiver}} = \left(\frac{1 + \frac{v_{\text{target}}}{c_s}}{1 - \frac{v_{\text{target}}}{c_s}}\right) f_{\text{transmitter}}.$$

If the speed of sound is significantly faster than the speed of the target, then the expression can be further simplified by first order Taylor series approximation. This condition is met for most human motions, so the dilation observed by the monostatic sonar is approximated by

$$f_{\text{receiver}} \approx \left(1 + 2\frac{v_{\text{target}}}{c_s}\right) f_{\text{transmitter}}.$$

The Micro-Doppler Effect

The micro-Doppler effect occurs when an object with multiple moving parts scatters a wave. Each part contributes its own Doppler shift related to the object's radial velocity component with respect to the receiver. All of the scattered waves are additive, and the resulting micro-Doppler modulation is a superposition of the individual components. Assuming that there are N moving point masses in a scene where a pure tone wave with frequency f_c is transmitted, the scattered signal seen by the receiver is

$$s_{\text{receiver}}(t) = \sum_{i=1}^{N} A_i(t) \cdot \sin(2\pi f_c t + 2\pi f_i t + \phi_i(t)).$$

Each point mass scatters the wave and modulates the frequency by $f_i = 2\frac{v_i}{c_s} f_c$, which is the Doppler shift in Equation 4. There is also a phase shift $\phi_i(t)$, that is dependant on the range of the point mass. Resolving this phase shift depends on the range of the point mass and the wavelength of the scattered wave. The amplitude of each component depends on the scattering surface and the range of the point scatterer.

There are many challenges in using the Doppler effect to sense and identify actions. By definition, real actions are sequences of motion that evolve in time and three-dimensional space. The micro-Doppler modulations recorded by a single active sonar sensor are only one-dimensional time-series. The observed modulations do not provide much range or spatial information. Over a certain period of time, the frequency modulations provide a histogram of velocities present. Due to the physical limitations of most multi-component objects, like the human body, the realizable set of actions is small.

Another challenge is that many moving objects have symmetry in their motion. For example, a human body may move its right or left arm. A single sonar sensor located at the line of symmetry would have difficulty distinguishing between such actions. One way to overcome this limitation is to use multiple sensors arranged such that no single line of symmetry is common to all sensors.

Sensors and Data Acquisition System

While there are challenges from using the micro-Doppler system, these can be addressed by obtaining an extensive set of multimodal data. This data will be used to train models, especially the use of multiple ultrasound sensors. However, by using this method, the system itself must also have the ability to obtain experimental data from many different types of sensors throughout the process. The ability to synchronize the data stream from all the devices is extremely important in order to make sure the results are accurate and aligned accurately.

One data acquisition technique that takes care of this concern is the acoustic data acquisition system. This system supports multiple distributed sensor modalities and uses a beacon that modulates the frequency in order to make sure data across the stream is aligned properly. The system consists of a set of data acquisition (DACQ) units and a frequency modulated (FM) transmitter that synchronizes the units by broadcasting a time-stamp embedded in the recorded data streams. These stamps are done every 9.6ms and have an accuracy within 6 microseconds. With these DACQ units, one can integrate FM receivers and synchronize all units within the FM transmitter's range, making the differences between data-collecting sensors and devices no longer a concern.

In order to sense movement from animals, we would want to configure ultrasound units for the FM transmitter. Sinusoidal signals are easy to manipulate since they are well-suited for capturing frequency modulations in movement. For this reason, the signal should be transmitted in a sinusoidal carrier, and the devices would record the modulated signals that scatter off the moving objects.

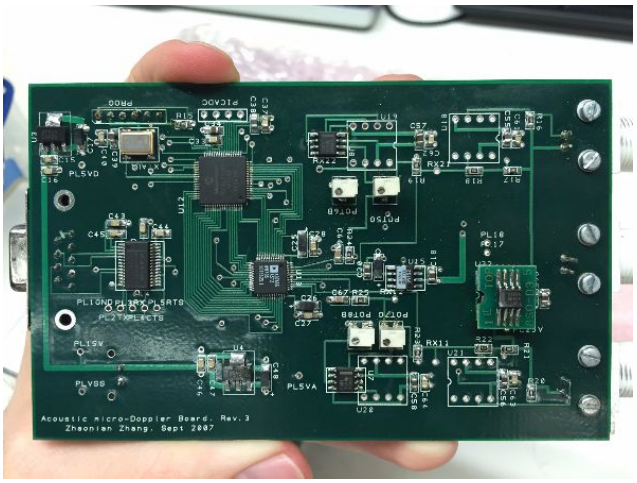
III. Proposed Improvements

III.a Hardware

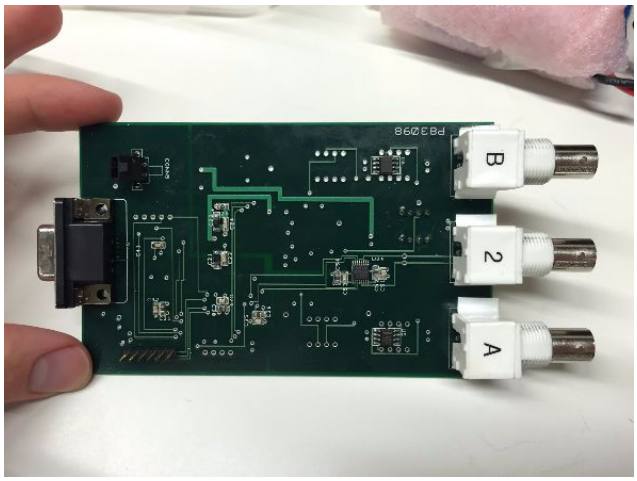
The existing prototype for this project has been created on a ~12.7cm x 7.5cm PCB with all the connections between parts correctly placed. While it is semi-functional, the size of the PCB is slightly too large for practical use in a real-world setting. For this reason, the primary upgrade to this prototype would be shrinking its size so that it can be encased and will not take up much space. It is known that some of the parts on the existing prototype are outdated, as the model was created over a decade ago. Based on Moore's Law, many of these parts have modernized versions that will most likely be smaller, which can be used to replace the current parts. In addition, the processing power of the device as a whole would improve by using parts that are more up to today's standards, which would improve overall functionality. For example, one of the replacements on the board to be made would be the RS232 with a Wireless XBee Wifi shield. This will not only make the device wireless, a highly desirable capability, but will also save space because the chip is noticeably smaller than the the RS232 mount. With continuous improvements like this, the size of the board will be able to shrink down to a practical size.

Besides replacing parts with smaller, similar functioning devices, there are also a few pieces in the circuit that might be obsolete, as there might be parts today that handle the function of these multiple devices in one chip. This would not only optimize the circuit so that one part does the job of multiple, but will also be a valid way to save space on the PCB.

The casing will be 3D-printed and will take up a lot of space that the device needs. However, it is important that this be designed correctly, as it will be made in order to snugly hold the board in place and protect it from external interference such as wind and water. Overall, it would be ideal for the device to be about the same size as the palm of a hand so that it can save space and effectively blend in with stones or plants where it will be used.

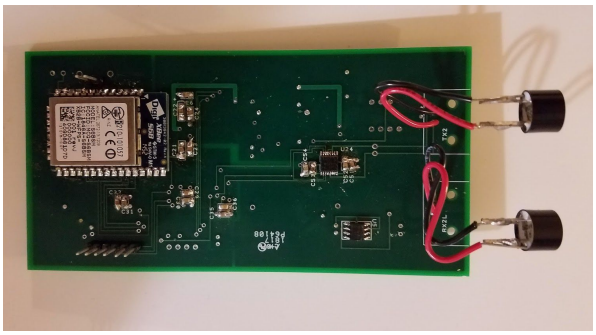


Front side of prototype PCB

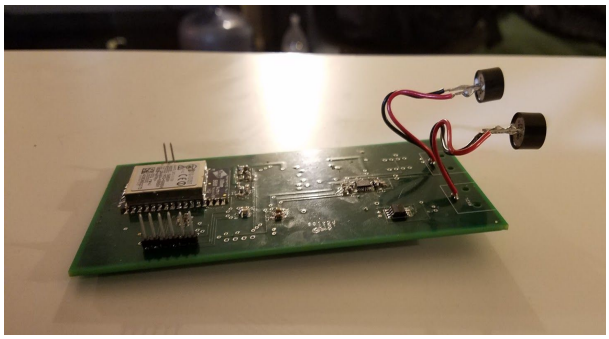


Back side of prototype PCB

The newest iteration of our device has been reduced to fit into a 3D printed enclosed case. The PCB board was reduced to 12.7cm x 6.3cm (16% reduction), and the new box had final dimensions of 15cm x 7cm x 5cm. While the box was only 16% smaller than the previous design, the new box also has the ultrasonic transducers enclosed within the structure, reducing the 2 1.5cm x 5 cm extensions in the original device. In addition, the new board only has one transmitter and receiver. However, this was not a decrease in functionality, as only one receiver was plugged in at a time in the original device. Furthermore, the new design eliminated the RS-232 serial communication tethering, and instead is upgraded to an Xbee Wifi module, which can support data transmission rates up to 72 Mbps with optimal wifi connection.

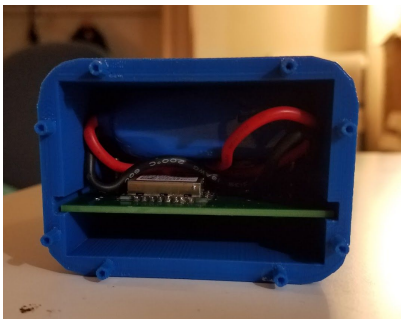


Top of new PCB

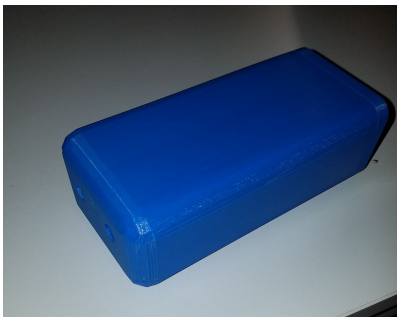


Side of New PCB

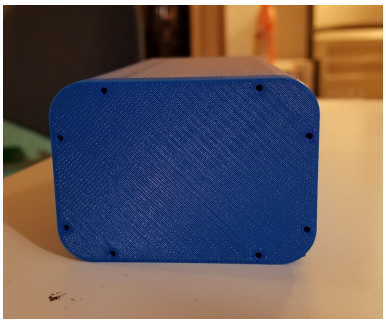
In addition, the new pcb board includes room and ports for internal mounts to be placed to support the extensions of the ultrasonic transducer pair. The box was designed using Creo Parametric, and then 3D printed in ABS. Furthermore, the back of the box includes additional screw holes for future use.



Open Back of New Box



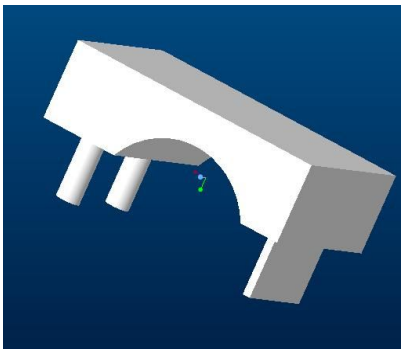
New Box



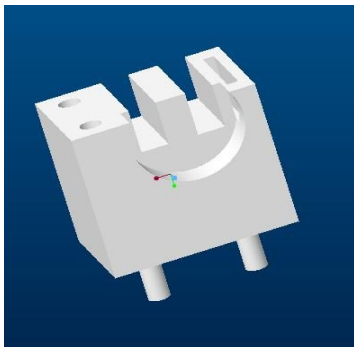
Back of New Box

For the next iteration, we are hoping to move away from the PIC board for processing. In addition, we aim to include a designated battery slot as well as 2 mounts for the ultrasonic transmitter and receiver pair. This will allow the ultrasonic transducers to be supported much like in the original design, except with a 3D printed mount we can eliminate the 5 cm metal supports.

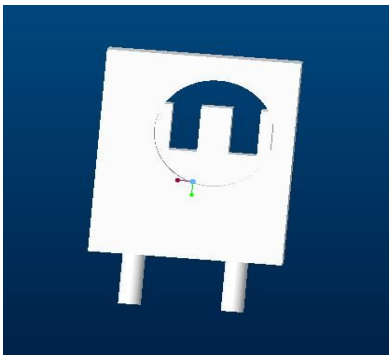
The mounts are designed to attach to the PCB board and support the ultrasonic transducers. This will allow for the board to be able to slide in and out of the box easily and mate with the holes in the front of the box. A rendering of the CAD models for the mounts are shown below.



Top of US Mount

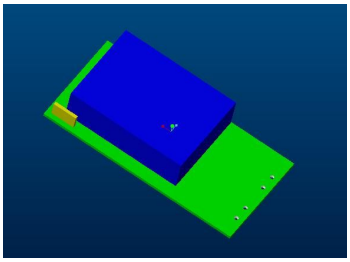


Bottom of US Mount

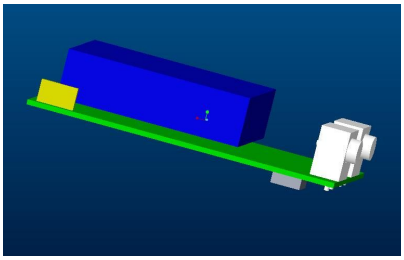


Assembled US Mount

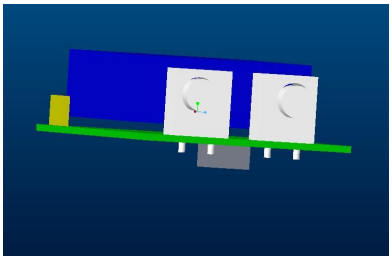
When fully assembled, the PCB board should look as shown below, as rendered in Creo Parametric.



Rendering of New PCB



New PCB Board with Mounts



New PCB Board with Mounts

III.b Training Algorithm

This paper was presented by Gian Delgo Tilpani and Fabio Ramos in
http://www2.informatik.uni-freiburg.de/~tipaldi/sites/default/files/tipaldi09iros.pdf
We are basing our ideas off of this paper and would like to extend their model to our dataset.

Introduction: Murray et al. mentions that the model is trained with an HMM using product of experts. We propose a new method that involves clustering the features of the micro-Doppler signal extracted from the field using a Conditional Random Field (CRF) clustering model. We believe that the Markov assumptions made by training an HMM, namely that given the state at time t, the state at time t+1 is independent of all previous states and the observation y_t is independent to all other states may not be an accurate way to train this model. For example, for a deer's motion, each segmental state may not be just dependent on the last time step of the motion but instead on a variety of factors such as the environment the deer is in. Additionally, the HMM training learns a joint distribution over our parameters, but we

ultimately are interesting in learning a conditional distribution p(Y | X) where Y is the location of the deer and X is the Doppler signal data that we have. For these reasons, we present a CRF Clustering solution to train the model and learn the motion of a deer.

Conditional Random Fields: A CRF is an undirected model that learns the conditional distribution p(X | Z) where X is some hidden state contained in the observed data Z.

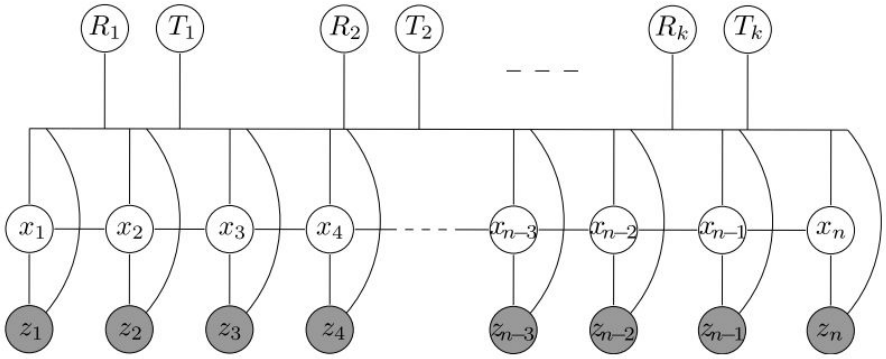


Fig. 2. Graphical representation of the CRF-Clustering model. The hidden states $\mathbf{x}_i$ indicate the cluster number. The observations $\mathbf{z}_i$ correspond to the local features. R_j and T_j indicates rotation and translation of the cluster j.

In order to learn the cliques in the graph, we define a variable C which is the set of cliques over the graph in the CRF. A CRF factorizes the conditional distribution into a product of the clique conditionals, defined as

$$\phi_c(\mathbf{z}, \mathbf{x}_c) = \exp\left(\mathbf{w}_c^T \cdot \mathbf{f}_c(\mathbf{z}, \mathbf{x}_c)\right)$$

In this equation, $w_c{}^T$ represents the weight vectors assigned to the $f_c(z, x_c)$ function, which extracts the features from a set of data. In our case, the features will be the time series signal generated by the deer or other moving object.

The conditional distribution becomes defined as

$$p(\mathbf{x} \mid \mathbf{z}) = \frac{1}{Z(\mathbf{z})} \exp\left\{\sum_{c \in \mathcal{C}} \mathbf{w}_c^T \cdot \mathbf{f}_c(\mathbf{z}, \mathbf{x}_c)\right\}$$

Where Z is a normalizing constant that sums over the hidden states and is a product over all the cliques in the set.

Inference over this model can be done using the well-known belief message passing algorithm, in which all neighbors send information to their neighbors based on their incoming messages and the clique

potentials. A normal CRF would learn the weights that we defined through maximizing the conditional likelihood of the data, but in our case, this is intractable. We must use the pseudo-likelihood of the data as our MLE objective function and it is defined as the sum of local likelihoods $p(x_i \mid MB_i)$ where MB is the Markov Blanket of the node x, defined as the neighbors of a node in the undirected CRF graph.

CRF Clustering: The model attempts to extend the CRF in order to perform clustering.

The model takes use of the rotational, *R*, and translational, *T*, parameters of a motion vector. In the clustering model, the local potentials represent the distance between points and clusters, and pairwise potentials represent constraints within the different clusters.

The local feature takes two inputs, a reference scan *g* and a second scan *s*. The objective of the local feature is to transform parts of *s* according to the clustering assignments, *X*, in such a way that the distance between points *g* and the transformed points in *s* are reduced. This feature is calculated as:

$$\mathbf{f}_{\text{dist}}\left(x_i, s_i, g_i\right) = \left\| T_{x_i} + R_{x_i} s_i - g_i \right\|^2,$$

R and *T* are calculated by minimizing the sum of the distances between the points *s* and the corresponding *g* as such:

$$R_l, T_l = \arg\min_{R,T} \sum_{s_i, g_i \mid x_i = l} \left\| T + R s_i - g_i \right\|^2.$$

The pairwise features attempt to encode information about the clusters themselves by incorporating neighborhood information. For example, if a point i is assigned to cluster k, it is very likely that a neighbor of i would be assigned to cluster k as well. The paper presents three pairwise features:

Neighborhood Feature:

$$\mathbf{f}_{\text{ng}}\left(x_i, x_j\right) = \begin{cases} \lambda_1, & \text{if } x_i = x_j \\ \lambda_2, & \text{if } x_i \neq x_j \end{cases}$$

Weighted Neighbor Feature:

$$\mathbf{f}_{\text{Wng}}\left(x_i, x_j, s_i, s_j\right) = \begin{cases} \lambda_1/\Delta, & \text{if } x_i = x_j \\ \lambda_2/\Delta, & \text{if } x_i \neq x_j \end{cases}$$

Stiffness Feature:

$$\mathbf{f}_{\text{st}} = \left\| (s_i - s_j) - \left[(T_{x_i} + R_{x_i} s_i) - \left(T_{x_j} + R_{x_j} s_j\right) \right] \right\|^2$$

The Neighborhood Feature indicates whether or not two nodes are neighbors, the weighted neighbor feature will assign more weight to a neighbor that is closer in terms of Euclidean distance to node i, and the stiffness feature preserves the distances intra-cluster after transformations are made to preserve the neighborhood properties.

Inference: Now that we have defined all of our parameters, we can define how inference will be done on this model. Regular belief update passing will not work because the values of the observations change within the hidden states. Instead, we have an initial cluster assignment that is randomly assigned. Local features are computed from this initial assignment. Messages are then propagated back and forward to estimate the X values and these new cluster assignments are used to calculate R and T, which give us information about the motion vector. We iterate until convergence.

Number of Clusters: In order to calculate the number of clusters that we ultimately derive, we need to first define p_c which is the probability of the cluster assignment of the CRF. After these probabilities are computed, we can compute the effective number of clusters a point belongs to as

$$N_{eff} = \frac{1}{\sum p_c^2}.$$

For each cluster, we calculate the average number of clusters each point within it belongs too. If that number is 1.7, we merge the cluster with the second best cluster; the cluster that contains most of the point assignments.

Convergence: The experiments run by the authors of the paper converge between 3 and 7 iterations on laser scan data. If we were to use our signal data, it stands to reason that we can reach convergence in a tractable period.

Conclusion: We believe that this CRF approach is a stronger approach to train the model for the deer-repellent system because it does not require a prior distribution and does not model a joint distribution; properties that seem reasonable for this problem. While we implemented and trained a naive CRF clustering model to learn the motion of a deer, as well as other objects, we did not have time to effectively evaluate the model. Going forward, we look to properly test and evaluate the model for future iterations of this project.