Automated detection, tracking, and counting of gray whales

Kevin Sullivan[*a], Mark Fennell[a], Wayne Perryman[b], and David Weller[b]

[a]Toyon Research Coporation, 6800 Cortona Drive, Goleta, CA 93117, [b]National Oceanic and Atmospheric Administration (NOAA) Southwest Fisheries Science Center, Marine Mammal and Turtle Division, San Diego, CA

ABSTRACT

Gray whales in the eastern North Pacific migrate annually between summer feeding areas in the Arctic to wintering areas off Baja California, Mexico. The abundance of this whale population has been documented by shore-based counts in central California conducted by human observers searching for and recording whale sightings during the southbound migration. Here, we describe a new semi-automated system for conducting gray whale counts, and compare such to the human observer based system. This new system consists of infrared cameras which continuously monitor a fixed field of view of the ocean, automated detection software for detecting whale blows, whale-blow verification software, and counting software which estimates the number of whales that have passed by the observation station. This technology is currently being considered to support naval, oil and gas, and merchant marine operations involving marine mammals.

Keywords: Long wave Infrared (LWIR), automatic detection, gray whales, particle filtering

1. INTRODUCTION

Gray whales (*Eschrichtius robustus*) in the eastern North Pacific were hunted to near extinction by the end of the 19th century [1]. Since that time, and coinciding with international conservation measures, this population of whales has recovered to numbers of about 20,000 as of 2011 [2]. The recovery of this whale population has been documented by shore-based counts conducted by NOAA Fisheries at or near Granite Canyon, central California, since the late 1960s. These counts are conducted by human observers searching for and recording whale sightings during the early winter (December-February) when whales are migrating south from summer feeding areas in the Arctic to wintering areas off Baja California, Mexico. Observations are limited to daylight conditions and, in part, the ability to track and record multiple whales simultaneously migrating past the observation station. Perryman et al, [3] used long wave infrared cameras to record video of whales and manually processed the video to estimate the number of whales passing by at night. Automated detection of whales using infrared cameras has been used by Zitterbart et al, [4] for a ship-based application and was reported by Santhaseelan et al [5], and Santhaseelan and Asari [6], for a shore-based application. Here, we describe a new system that not only detects whales, but also estimates the number of whales that are migrating past the NOAA Granite Canyon field site. A block diagram of the system is shown below in Figure 1.

*ksullivan@toyon.com; phone 1805 968-6787

Thermosense: Thermal Infrared Applications XLII, edited by Beate Oswald-Tranta, Joseph N. Zalameda, Proc. of SPIE Vol. 11409, 1140906 · © 2020 SPIE
CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2567187

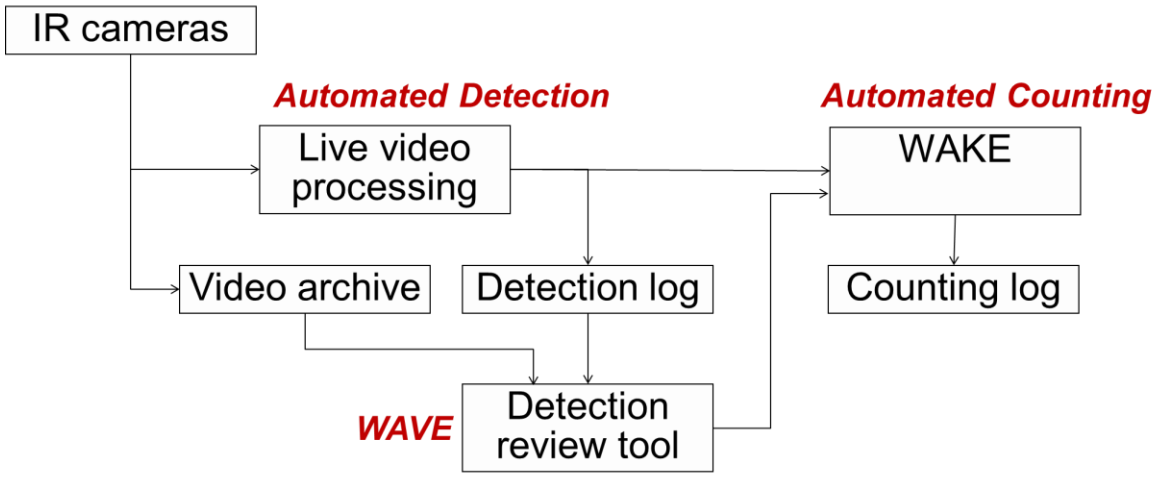

Figure 1: Block Diagram of Whale Detection and Counting System

There are three main components of our system, namely, an automated detection module, a detection review tool, and an automated whale counting module. The automated detection module processes video either from live camera feeds or from stored videos and outputs the location of putative whale respiratory exhalations (from here on called "blows"). This module is further described in Section 2 of this paper. The detection log output by the detector contains a list of times and locations of whale blows along with false alarms from the whales themselves (we only want one detection per blow per whale), other marine mammals in the region, small vessels passing by, splashing birds, and ocean clutter caused by whitecaps and other debris. To eliminate false alarms in a rapid manner, we have created a detection review tool, called WAVE (Whale Analytics Verification Engine), which allows a human to quickly determine blows versus false alarms. WAVE uses recordings of video that were made on a computer's hard drive using VLC video recording software and overlays icons and text. The third major component of our system, and the primary subject of this paper, is the automated counting system. We refer to this module as WAKE. The counting software takes a reviewed detection log and estimates the number of whales that are passing by the location of the cameras. This is done using the times and locations of the whale blows in combination with information on swim speed, surface-dive timing and blow interval behavior of gray whales. The counting software is capable of handling unfiltered detection data with a corresponding false alarm rate, but the manual removal of many false alarms is expected to improve accuracy. The logic behind the counting module is described in Section 3 below.

The system described in Figure 1 was deployed at the NOAA Fisheries Granite Canyon field site between 2014-2016. In January of 2014, NOAA and Toyon performed a comparison between the semi-automated system and a team of human observers that counted the number of whales migrating past the field site during the daytime using handheld 7x50 binoculars, the naked eye, and computer tracking software. The results of this comparison are presented in Section 4. We finish this paper with a brief summary and discussion of additional applications of this technology in Section 5.

2. AUTOMATED DETECTION OF WHALE BLOWS

We have developed software that can monitor a video stream and automatically detect the presence of multiple simultaneous whale blows. In this section, we describe the algorithms that are implemented in that software and provide some performance metrics. The first step in the processing chain is to segment the incoming video into a sequence of individual frames and operate on each frame of data as it arrives. For best performance, the software needs to finish processing a frame of data before the next video frame arrives. In the case of 30Hz video, the algorithms would have 1/30 of a second to complete this task. In some cases, the algorithms do not finish in time for the arrival of the next frame and this new frame is skipped and the following frame is processed. For the results presented here, we used a COTS desktop computer to process three 640x480 video streams and occasionally skip one frame of data. Performance of the system is little compromised if the number of frames skipped is small, but skipping multiple frames could significantly impact performance.

Whale blows are initially detected using a background model that is created by using a set of prior frames of data. We divide up each frame of video into spatial bins of 100 pixels and create a histogram of the intensity values observed. Given a new intensity sample, we compute a clutter likelihood, l_c, using the histogram. We also assume a pdf, p_w, of whale pixel intensities which is a uniform distribution that is nonzero at the median of the clutter distribution up to the maximum possible intensity value and zero elsewhere. We compute an overall whale likelihood for a single pixel, l_w, as $l_w = p_w/(p_w + l_c)$. We require that 30s of data is collected to create the background model. Thus the system cannot detect whales for 30s after startup. The background model allows us to compute a likelihood that a particular pixel belongs to a whale. We combine the likelihoods, l_w, across multiple pixels by computing their product. The number of pixels used for the likelihood product varies by range. The below table provides detection parameters used for the results presented in this paper.

Table 1: Detection Algorithm Parameters

	Range < 1km	1km < Range < 2km	2km < Range < 3km	3kim < Range < 4km	4km < Range < 5km	Range > 5km
Number of pixels	49	25	9	4	4	4
Minimum Time, s	0.4	0.4	0.4	0.4	0.4	0.4
Maximum Time, s	6	6	6	6	6	6

Once a putative blow is detected, we put it into track [16], and seek to detect it again in subsequent frames of data. We require that all blows persist in the video stream between a minimum and maximum time as specified in Table 1. In order to update a blow track, we search in a spatial window about the predicted location of the blow track. The center point of the search window is computed using the prior location and the prior observed velocity due to the wind at the range of the last detected location. The size of the window is given by +/- the square root of the values for the likelihood product in Table 1.

Every pixel in every frame of data is considered for blow detections allowing for multiple detections to occur in each frame. We combine closely-spaced detections by selecting the single pixel within the search window that has the largest likelihood product.

Blows can have significant motion in pixel coordinates in the presence of wind. For this reason, we estimate the speed of the wind in pixel coordinates using confirmed blows. We compute the wind speed by dividing the difference between the starting point of a blow track and the ending point of the blow track by the difference in time between the two points. We keep track of the speed of the wind in six different range bins because the angular rate of an equivalent cross-range velocity changes significantly with range and because the actual wind speed can change significantly with range. When no confirmed blows are detected, we assume that the wind speed diminishes towards zero at a slow rate.

If large numbers of whales are passing through the area, the background model can become corrupted with pixel intensity values from the whales. This could cause the detector to miss subsequent whales because they look too much like the background. We avoid this problem by ceasing the updating of the background model in the window (size provided in Table 1) about a putative whale detection until after the blow is no longer being observed.

A snapshot of a blow that was automatically detected by the detection software is shown below in Figure 2.

Figure 2: Example Automated Blow Detection

We collected video at Granite Canyon, CA, on January 6 and 7, 2014 using three LWIR cameras manufactured by FLIR, Incorporated. The specific models of cameras used included two F-610 cameras with a ten degree field of view in azimuth and an eight degree field of view in elevation, and one F-606 camera with a six degree field of view in azimuth and a five degree field of view in elevation. These are relatively low cost infrared cameras with the total cost of all cameras being less than $50,000. The cameras were arranged, using markers in the field of view, such that they had adjacent fields of view in azimuth and they were tilted in elevation so that the horizon was about ten pixels down from the top of the image.

The output of the automated detection algorithm is summarized in Table 2 below. The false alarm rates are specified in terms of the number of false alarms per hour per camera. The marine mammal false alarms are detections which are caused by marine mammals present in the scene. Our counting algorithm requires only one detection per blow of the whales, thus any detections on the whale's back or tail are considered false alarms for our purposes. Additionally, any other marine mammals in the area which trigger detections are also considered false alarms even though these detections might be of interest for purposes other than counting the number of migrating gray whales. On January 6, a large group of Risso's dolphins (*Grampus griseus*) entered the field of view of the cameras, splashing around and breaching, and creating a large number of marine mammal false alarms. No such group showed up on January 7, thus lowering the number of marine mammal false alarms substantially. Likewise, the number of false alarms due to the presence of small boats was much lower on January 7 because none were present in the field of view on that day. Large numbers of birds are frequently present at the site and the detection algorithm has proven to do a good job of filtering the vast majority of them out. Clutter false alarms are caused by whitecaps or debris moving in the water and these were higher on January 7 because the winds were stronger in the afternoon.

Table 2: Detector Performance

	Number of single blows detected and verified	Marine mammal false detection rate	Clutter false alarm rate	Bird false alarm rate	Boat false alarm rate
January 6	465	19	0.1	0.4	0.2
January 7	705	1.2	3.4	0.2	0

The detections identified by the automated detection software are reviewed by humans to eliminate the false alarms. The resulting detections output from this process are used to estimate the number of whales present using the automated counting software which infers the number of whales passing by based on the detected blows. The number of blows verified by humans for each day is shown in Table 2.

3. AUTOMATED ESTIMATION OF THE NUMBER OF WHALES

Our system is designed to estimate the number of whales, *N*, that pass by the field site on their migration over a time period *T*. After looking at days of video and live observations, we have seen that the whales tend to travel in closely-spaced groups, often referred to as pods, for periods of time that are typically longer than the time it takes for whales to pass through the region observed by our cameras. We thus compute N by estimating the number of whales in each pod and summing the whales in all pods that pass by during the time interval T. We consider a state vector for the whale tracking problem that is a mixed continuous-discrete state comprising a discrete random variable representing the number of whales in the scene, a continuous random vector representing the locations of the whales in the scene, and a continuous random vector describing the temporal state of the whale in a respiration cycle. Let the state vector at time k be $\boldsymbol{x}_k = [n_k \boldsymbol{r}'_k \; \boldsymbol{b}'_k]$ in which n_k is the number of whales in a pod, $\boldsymbol{r}_k \in R^{4n_k}$ represents the state (2D position and velocity in a local east, north, up (ENU) coordinate system with origin at the location of the cameras) of each whale in the pod at time k, and $\boldsymbol{b}_k \in R^{n_k}$ represents the respiration state of each whale as further discussed in Section 3.1. Note that the size of $\boldsymbol{x_k}$ is variable due to the different number of whales that could be present (thus, realizations of the state can be modeled as a sample from a Poisson Point Process (PPP) [7]). We observe a sequence of measurements $Z^k = (\boldsymbol{z_1}, \boldsymbol{z_2}, \ldots, \boldsymbol{z_k})$ as a pod of whales passes by. The whales are only visible when they surface to blow and only blows are counted as whale detections. Measurements include whale blow detections as well as the absence of a detection as will be discussed further below. Observation of the whale body and tail can trigger detections by the algorithm described in Section 2, but these are treated as false alarms for the purposes of estimating the number of whales that are present. The use of a whale respiration model provides key information for determining how many whales are present, but use of this model requires that only one detection is obtained, only when a whale is blowing. The measurements are related to the state variables through the measurement equation $\boldsymbol{z} = h(\boldsymbol{x_k})$. The state of the whale pod is assumed to change according to the equation $\boldsymbol{x_{k+1}} = f(\boldsymbol{x_k})$, where the function f is a nonlinear time-varying function which is approximated as described in Section 3.1. In order to estimate $\boldsymbol{x_k}$, we compute the density function $p(\boldsymbol{x_k}|Z^k)$ using Bayes' Rule, $p(\boldsymbol{x_k}|Z^k) = \frac{p(z_k|\boldsymbol{x_k})p(\boldsymbol{x_k}|Z^{k-1})}{p(\boldsymbol{z_k}|Z^{k-1})}$.

The counting algorithm addresses a number of challenges. Gray whales surface sporadically and only for brief intervals, and the camera infrared sensors cover a relatively narrow field-of-view. Consequently, we expect to see only a few blows from a pod of whales as they pass by. Their precise swim speed and travel direction are not known, often making it difficult to ascertain immediately whether a sequence of blows is from a single whale, two whales, or possibly more.

The counting algorithm incorporates some of the same techniques used by shore-based human observers facing these challenges. They include, for example, recognizing that a sequence of blows are too closely spaced in time to be from a single whale, or that the distance between successive blows is not consistent with a reasonable swimming speed. Human observers can make these judgments based on experience and by understanding the typical behavior of migrating gray whales. Particularly important is proper characterization of gray whale respiration patterns and normal swimming speeds. The algorithm makes use of these characteristics (as described by Rugh et al. 2008; Perryman et al. 1999) and many other

behavioral statistics by quantifying them as probability density functions (PDF) and performing a Bayesian prediction and updating process.

We update the density function, $p(\mathbf{x}_k|Z^k)$, using a particle filter [8, 9, 10] because of its ability to handle non-linear, non-standard probability density functions for whale motion and behavior, as well as for the sensor measurements of detected blows. The algorithm is implemented in the C++ computer language and runs faster than real time in most situations. Particle filters approximate the PDF of an unknown random quantity by a weighted sum of delta functions. Consider a set of points and associated weights, *w*, in the target state space, $\left\{w_{k-1}^{(i)}, \mathbf{x}_{k-1}^{(i)}\right\}\ i = 1,2,\ldots,n_s$ which represent samples from the updated target state PDF at time k-1. We can write the approximation of the PDF using a weighted sum of delta functions as $p(\mathbf{x}_{k-1} \mid Z^{k-1}) \approx \sum_{i=1}^{n_s} w_{k-1}^{(i)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(i)})$, in which $w_{k-1}^{(i)}$ is the weight associated with the i^{th} point and $\sum_i w_{k-1}^{(i)} = 1$. A filtering algorithm is a process by which the set of points representing $p(\mathbf{x}_{k-1} \mid Z^{k-1})$ is transformed into a set of points, $\left\{w_k^{(i)}, \mathbf{x}_k^{(i)}\right\}$, representing the PDF, $p(\mathbf{x}_k \mid Z^k)$. We are particularly interested in the number of whales in a pod and this is calculated using $E\{n_k\} = \sum_{j=1}^{n_s} w_k^{(j)} n_k^{(j)}$. The summand is simply the weight of the jth particle times the number of whales represented by the jth particle. A key element of the state vector is the schedule of blows for each whale during a surfacing. This is important because the blows are the phenomena being observed by the sensors. Thus, each particle in the filter represents a hypothesis of these state variables:

- Number of whales in the group
- Each whale's position vector on sea surface
- Each whale's velocity vector on sea surface
- A respiration state which indicates the current state of the whale in a repeating respiration cycle which includes a surface interval with multiple blows and a dive with no blows.

The algorithm uses the observed blows to calculate the relative likelihoods of these hypotheses and generates a new set of particles to represent an updated PDF. An important benefit of including a state that represents the respiration behavior is that our algorithm also *updates hypotheses based on the absence of blow detections*. Specifically, all particles that predict a blow during a time interval in which none is seen will be reweighted with a lower likelihood. Those particles may die off during the subsequent resampling step. Note that some of those particles may survive because we need to account for the small chance (15% assumed in this case) a blow is not detected.

We also allow particles to have a value of zero for the number of whales in a group. This represents the hypothesis that there is not actually a group of whales present, thus indicating that an observed blow could have been a false alarm.

Multiple pods of whales may be passing through our sensor coverage at the same time. To account for this fact, we associate each detection with an existing pod, as discussed below, or we create a new pod. The particle filter updates the state likelihoods, $p(\mathbf{x}_k|Z^k)$, separately for each pod using only the detections associated with that pod. We refer to the particles associated with a particular pod as a cloud of particles in the below discussion. Figure 3 shows an overview flow chart of the algorithm's operation and we discuss each step in more detail below.

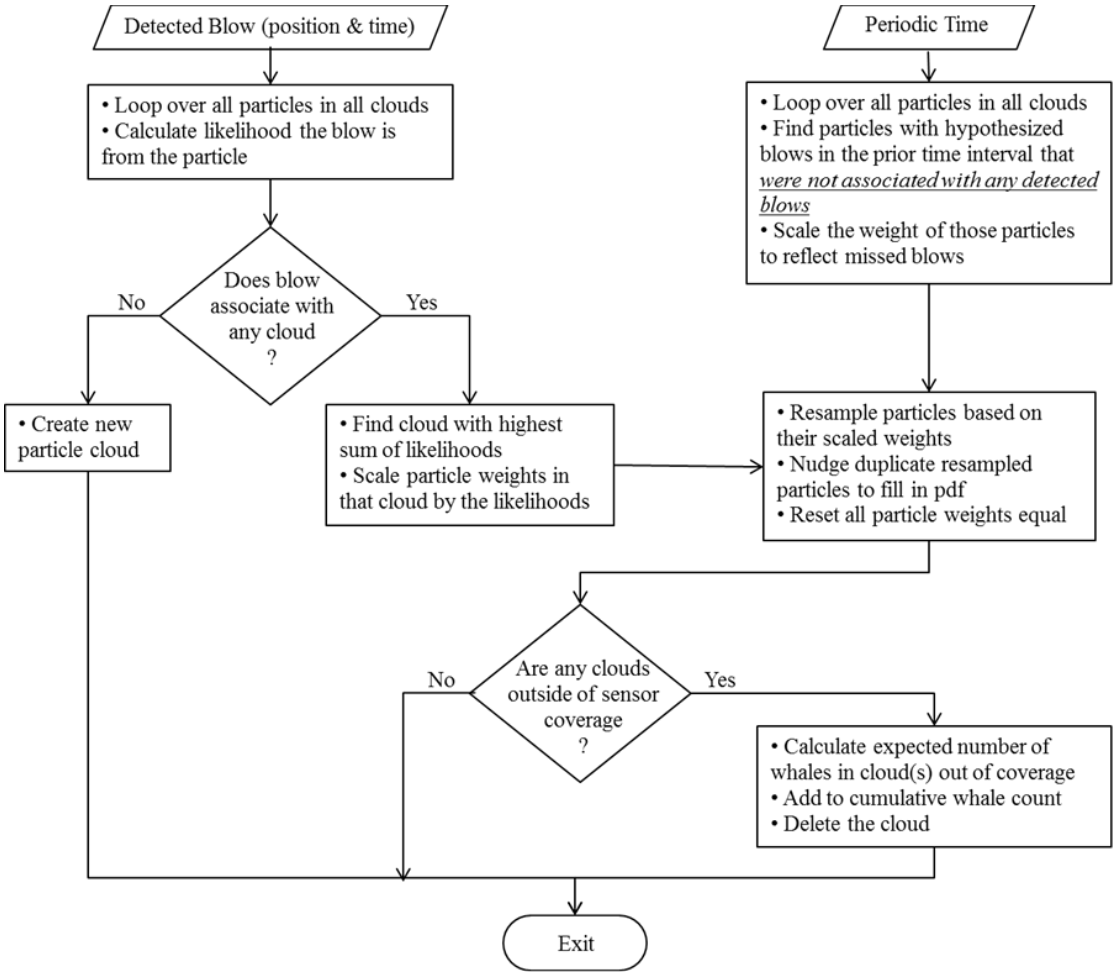

Figure 3: Block Diagram of Whale Counting Logic

Two distinct input paths are shown: updating the Particle Filter with a blow detection, and updating the filter at periodic time intervals. We begin on the left.

A single detected blow is described by a position and time. The first process loops over each particle in every cloud to calculate the likelihood that the hypothesized whale group represented by the particle could have produced the detected blow. When comparing a hypothesized blow to a measured blow, the likelihood is computed by comparing the time and location of the measured blow to the time and location of the hypothesized blow. If the times or locations do not agree within specified bounds (+/- 5 time criteria, +/- 40 m cross-track, and +/- 200m along-track), then the likelihood that the given hypothesized whale state is true is given by the false alarm rate output by the detector or human verification process. If the times and locations agree within the specified bounds, then the spatial likelihood is calculated using the proximity of the detected blow position to that of the particle's hypothesized blow, and accounting for the expected angular uncertainty of the sensor measurement.

We compute a likelihood, *L*, for a particle by first finding the hypothesized blow that best matches the location and timing of the observed blow using

$$i^*, j^* = \arg \max_{i\,=\,1, n_k, j\,=\,1, m_k} L(x_k, z_k, i, j)$$

$$L(x_k, z_k, i, j) =$$

$$\begin{cases} P_d * e^{-0.5\left(\frac{\Delta Az_i}{\sigma_{az}}\right)^2} e^{-0.5\left(\frac{\Delta El_i}{\sigma_{El}}\right)^2} + (1 - P_d) * P_{fa}, & |t_o - t_{ij}| < T_t \; and \; \|p_o - p_{ij}\| < T_p \\ P_{fa}, & |t_o - t_{ij}| \geq T_t \; or \; \|p_o - p_{ij}\| > T_p \end{cases}$$

where

$p_o \in R^3$ = location of the observed blow

$p_{ij} \in R^3$ = location of the j[th] hypothesized blow for the i[th] whale

m_k = the number of whale blows predicted for whale i during the current respiration cycle

$\Delta Az_i = \tan^{-1}\left(\frac{p_{ij(2)}}{p_{ij(1)}}\right) - Az_o$

$\Delta El_i = \tan^{-1}\left(\frac{p_{ij(3)}}{\sqrt{p_{ij(1)}^2 + p_{ij(2)}^2}}\right) - El_o$

Az_o = the observed azimuth of the detected blow

El_o = the observed elevation angle of the detected blow, computed using the difference in pixel coordinates between the horizon and the bottom of the group of pixels which defines a blow

t_o= time of detected whale blow observation

t_{ij}= time of j[th] hypothesized blow for the i[th] whale

σ_{az} = the standard deviation of an assumed Gaussian error in the measurement of the azimuth angle

σ_{el} = the standard deviation of an assumed Gaussian error in the measurement of the elevation angle

P_d = the probability of detection of a whale blow given that there is a blow in the field of view of a camera

P_{fa} = probability that a reported whale blow is in fact a false alarm

T_t = a threshold on the degree to which the projected time of a blow needs to match the observed time of a blow

T_p = a threshold on the degree to which the projected location of a blow needs to match the observed location of a blow

We do not have ground truth indicating the precise number of whales that are present, so we must estimate the values for P_d and P_{fa} using human observations of IR video. We estimated P_d to be 85% and P_{fa} after human verification to be 0.0001.

The thresholds T_t and T_p were selected such that nearly all hypothesized blows which could match with an observed blow would be included in the likelihood calculation. We thus selected 5s for T_t because most blows last a shorter time than this. T_p was set such that the inclusion zone boundary is three times the assumed standard deviation of the angular error. The angular errors were assumed to be 1/100 of the azimuth field of view of a camera and 1/200 of the elevation field of view.

We determine if an observed blow can be associated with any clouds of particles, by checking to see if one or more particles have hypothesized blows within the temporal and spatial windows. If one or more clouds of particles meet this criteria, we find the one with the highest sum of likelihoods where the likelihood of each particle is given by $L(x, z, i^*, j^*)$ and proceed by assuming the blow was generated by the group that the cloud represents. Next, each particle in that cloud has its weight scaled by the likelihood value associated with each particle.

Now that each particle's weight represents the likelihood of the detected blow, the entire cloud can be resampled to update the PDF it represents. The resampling process randomly selects and copies particles proportional to their relative weights in the cloud. Heavily weighted particles are selected more often; lighter weight particles are selected rarely or perhaps not

at all. This process can sometimes create and copy particular particles many times, which can lead to a degenerative PDF with many copies of only a small number of unique particles. To prevent that situation, we perform regularization [10] to fill in the state space nearby the selected particle. Finally, all particles are reset to have uniform weighting such that, individually, they represent equal-likelihood hypotheses. The cloud represents an updated and filled-in PDF of the state variables.

The next step checks each cloud to see if the whale group has exited the sea-surface polygon of sensor coverage. This test checks the whale-position estimates of each particle for polygon inclusion, and then compares the number outside to a user-input threshold. The threshold used for the results presented here was that 19,800 particles out of 20,000 needed to be outside in order to decide that a pod has left coverage. If the whale pod is found to have left the sensor coverage, then we calculate the expected number of whales in the PDF and add it to the cumulative total. Summary data about that whale pod is output to a file and then the cloud is deleted.

The right-hand path through the flowchart in Figure 3 occurs at specific periodic time intervals, typically every five seconds. Five seconds was chosen because it covers the longest time in which most whale blows will persist in the field of view of a camera. In this case, we loop over every particle of every cloud and look for blows that were hypothesized to occur during the prior interval but did not get associated with any detections. The likelihood of missing these blows, *L*, is simply one minus the detection probability of the sensor, raised to the number of missed blows. The weight of these particles is scaled by the calculated likelihood. All particles that did not hypothesize any blows, or those with blows that were associated with detections, are left with their same weight giving

$$L(x) = \begin{cases} (1 - Pd)^N, & N > 0 \\ 1, & N \leq 0 \end{cases}$$

where N is the number of blows that were scheduled to occur in the last time interval, but did not get associated with any observed blows.

All clouds are resampled as described above to generate updated PDFs that reflect the absence of detected blows. And as described previously, the clouds are checked to see if they have left sensor coverage, and if so, the cumulative whale count is updated.

The Particle Filter runs as described in Figure 3 until an entire log of blow detections has been processed and all clouds have left sensor coverage.

4. MIGRATING GRAY WHALE DYNAMICS MODEL

We have created a model for the motion and respiration of a migrating gray whale. The motion model assumes a nearly constant velocity [16], in the horizontal plane. The respiration model assumes that the whales follow a breathing cycle where they have a surface interval that includes multiple blows typically separated by 10 – 40 seconds, over a period of 0.5 – 2 minutes, followed by a dive which typically lasts 2 - 5 minutes. These statistics have been quantified by a number of studies looking at whale behavior [3, 11, 12, 13, 14, 15]. This breathing cycle repeats and the total length of the cycle can vary from one whale to the next and from one breathing cycle to the next. We can characterize the respiration state of a whale using a parameter, b_k, which indicates the location in the breathing cycle as a fraction of the duration of the current breathing cycle.

If no particle clouds are associated with a blow detection, which includes startup and anytime no clouds exist, then a new cloud is created using our system dynamics model that characterizes whale behavior. The initial states of each particle (number of whales, positions and velocities, and blow schedules) are generated by random draws from user-input PDFs and PMFs for the relevant parameters. Figure 4 shows the multi-step process.

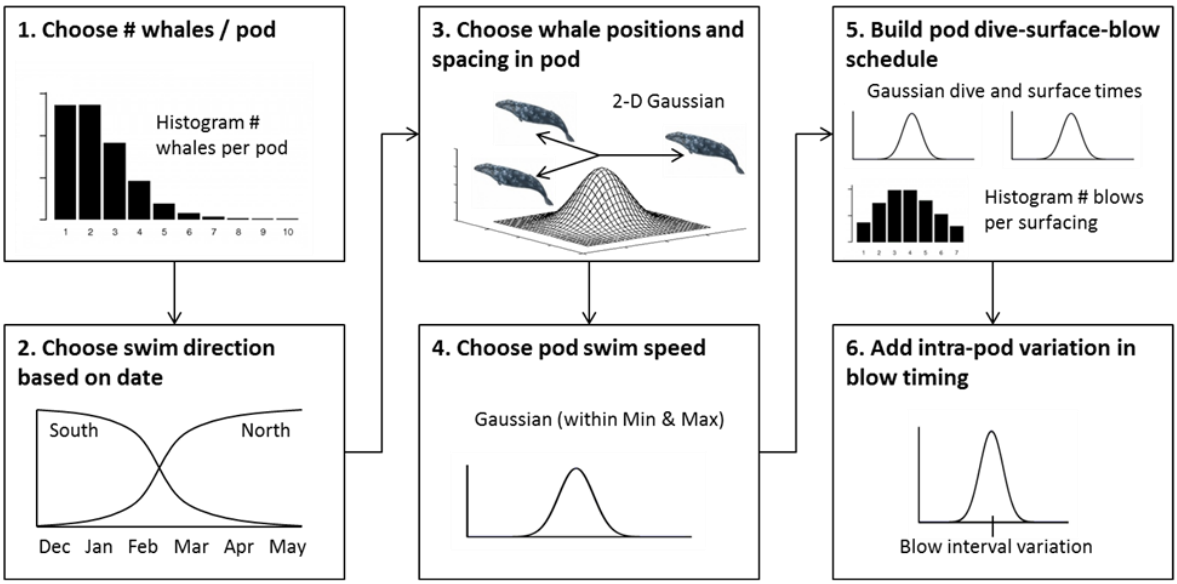

Figure 4: Whale Pod Motion Model

Step 1. Choose the number of whales in the pod with a random draw from a probability mass function specified by an input histogram. The histogram was created after reviewing references [2, 3, 11] and by manually processing previously-collected video. This state-space variable stays the same for the life of this particle.

Step 2. Choose the swim direction of the pod. The whales are either heading south or north depending on the time of the year. The precise direction along a migration axis is perturbed slightly with Guassian noise, drawn from a distribution with three degrees standard deviation.

Step 3. Choose the whale relative positions. Each whale's position in the pod is drawn from a two-dimensional Gaussian density with specified standard deviation in the cross-track (5m standard deviation) and along-track (10m standard deviation) dimensions. We are not aware of any data which supports or denies this assumed spacing distribution. We estimated these values based on manual observation of prior videos and overhead images. The whales are assumed to maintain their relative positions for the duration of time within the sensors' coverage.

Step 4. Choose the swimming speed from a Gaussian PDF truncated by a minimum and maximum swim speed. We used a Gaussian with a mean value of 1.5 m/s, a standard deviation of 0.25 m/s, a minimum value of 0.1 m/s and a maximum value of 2.5 m/s.

Step 5. Build the dive-surface-blow schedule for the pod. First we randomly select the number of blows in a particular surfacing by drawing from an input histogram. Table 2 provides the histogram used for the results presented here. Next we space the pod blows in time by drawing from a truncated Gaussian PDF (mean value of 20s, standard deviation of 10s, minimum value of 10s, and maximum value of 45s). After the final blow of a surfacing event, the whale pod enters a deep dive during which no blows will occur. The dive time is drawn from yet another Gaussian PDF (mean value of 180s, standard deviation of 60s, minimum value of 45s, and maximum value of 600s). After the dive, Step 5 is repeated for the subsequent surfacing event.

Step 6. Add variation to the pod blow timing for individual whale blows. We are unaware of any studies quantifying the correlated nature of respiration rates within a whale pod, but we have observed that the blow timing is correlated. To account for this, we start with the single blow time for the pod selected in Step 5 and randomly select times around the pod blow time for each individual in the pod. The random selections are drawn from a Gaussian PDF centered on the pod blow time with a standard deviation of 5s with a maximum possible perturbation of 45s.

Table 3: Probability Mass Function for the Number of blows by a Pod of Whales in a Surface Interval

Number of blows:	1	2	3	4	5	6	7	8	9	10
Probability:	.15	.25	.25	.15	.10	.05	.02	.01	.01	.01

Figure 5 illustrates the result of creating a particle cloud using the above-described dynamics model. The cloud is comprised of 'N' distinct particles, each representing a unique hypothesis about the respiration events of a whale pod. These hypotheses are generated each time a new whale pod is detected. Note that a particle can have zero whales, representing the hypothesis that a detection is, in fact, a false alarm, and that there is not really a group of whales present.

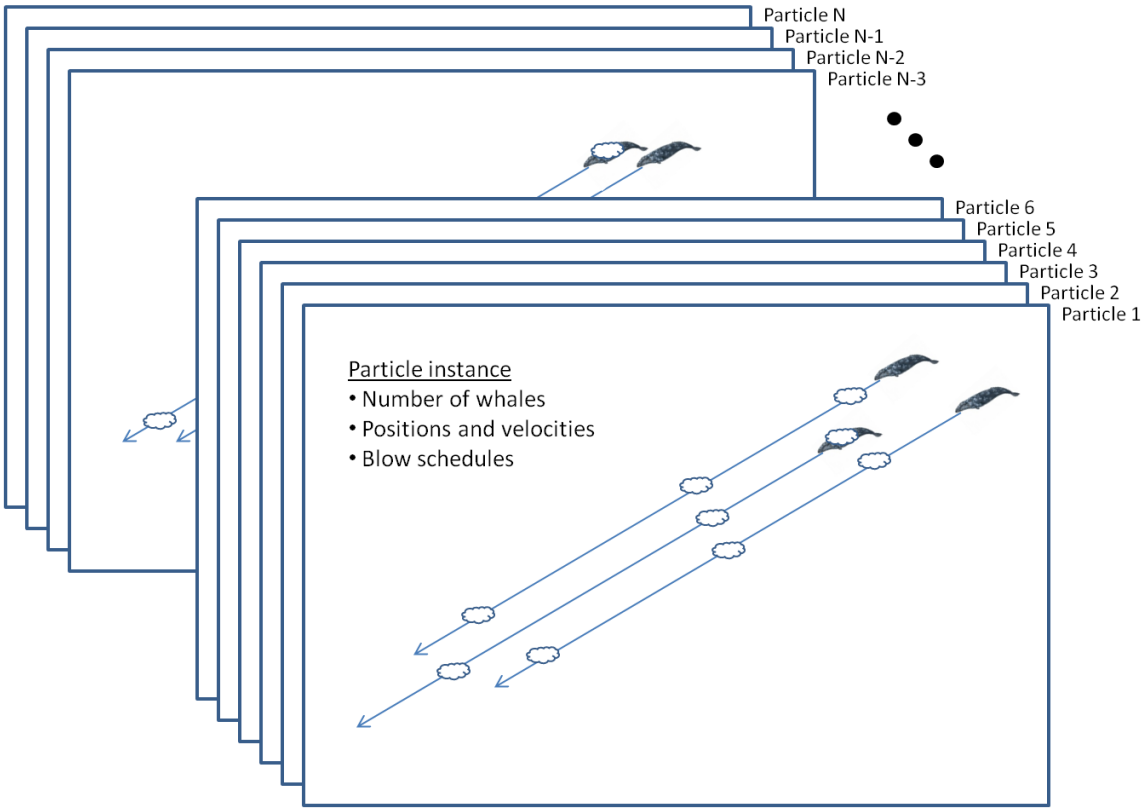

Figure 5: Particle State Representation

5. COMPARISON TO MARINE MAMMAL OBSERVERS

In early January of 2014, we operated our system at the Granite Canyon field site simultaneously with a team of human observers (often referred to as Marine Mammal Observers (MMO)) from NOAA. The cameras were located at 36.437 degrees latitude and -121.92 degrees longitude and the observers were within 100m of the cameras. Visual counts were made by rotating teams of observer pairs using naked eye aided by 7x50 Fujinon FMTRC-SX binoculars with built-in compasses and reticles for distance estimation. The primary observer in the pair kept continual visual watch of the scan area (see Fig. 6) while the secondary observer served as a data recorder but also kept watch and assisted with tracking already identified pods of whales. Each observer had one 90 minute shift as primary observer, followed by a second 90

minute shift as secondary observer and then a 90 minute break. The observers also used custom software (R. Holland, NOAA Fisheries) which aided with data recording and tracking of their observations. At the same time, our three infrared cameras collected and recorded video. A map of the field site along with triangles indicating the viewing directions and approximate range of the cameras and the human observers is shown in Figure 6. We estimate that both the humans and camera-based system have a maximum detection range of about 5km.

The field of view of the human observers and the infrared system were not completely overlapping as shown in the below figure. The humans tend to look further north and the cameras look primarily west. Both systems were attempting to get a total count of whales passing by the site.

Figure 6: Map of Field Site with System Field of View

We recorded detections live at the site during this test and the system is capable of operating in a completely automated mode, but in order to improve counting accuracy, we manually reviewed the detection logs to eliminate many false alarms. The result was a detection log that represented all of the detections found by the infrared cameras along with a small chance of inclusion of false alarms. This detection log was fed into the WAKE tool to get a system count.

Figures 7 and 8 below show a plot of the cumulative count obtained using the IR system and the human observers. There are significant differences in the count obtained by the two systems over short periods of time, but the cumulative count for the day is remarkably consistent. Short-term variations are to be expected due to the fact that the two systems were working with different whale blow observations due to differing fields of view. The humans tended to look north to see the whales as they approached, while the cameras look nearly west so that both southbound and northbound migrations can be equally observed without moving the cameras.

The curves plotted in Figures 7 and 8 for the IR system differ from the MMO curves in that the IR system counts occasionally decrease. This is because the automated counter is constantly estimating the number of whales that pass by and this estimate can go up or down based on the most recent observations. For example, a lone detection might trigger an expectation that a pod of whales is present, but if no further blows are observed, this observation will be negated over time. Counts can also go down for human observers, but humans typically reconcile the count for a group in their heads prior to recording their estimate for the group, so this effect does not show up in the MMO plots.

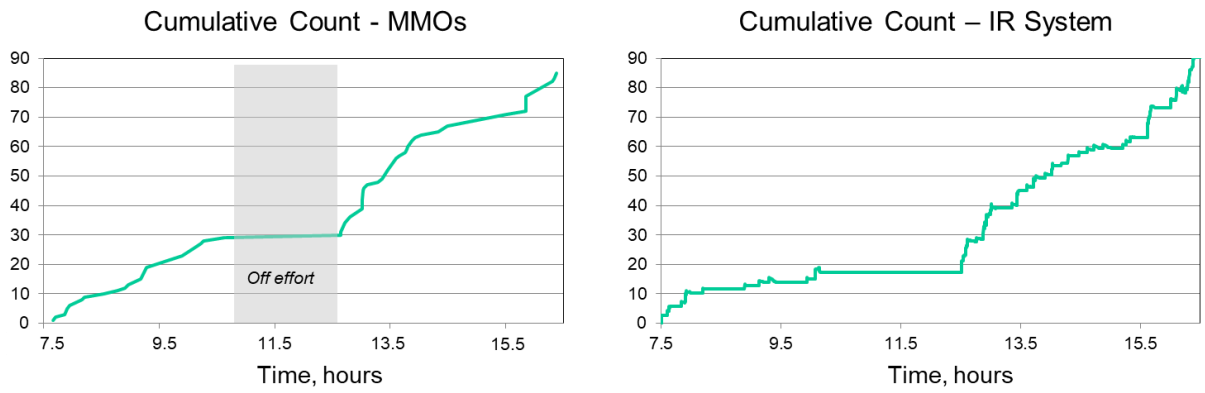

Figure 7: Comparison of Counts – Marine Mammal Observers (MMOs) (left) versus Semi-automated System (right) – January 6, 2014

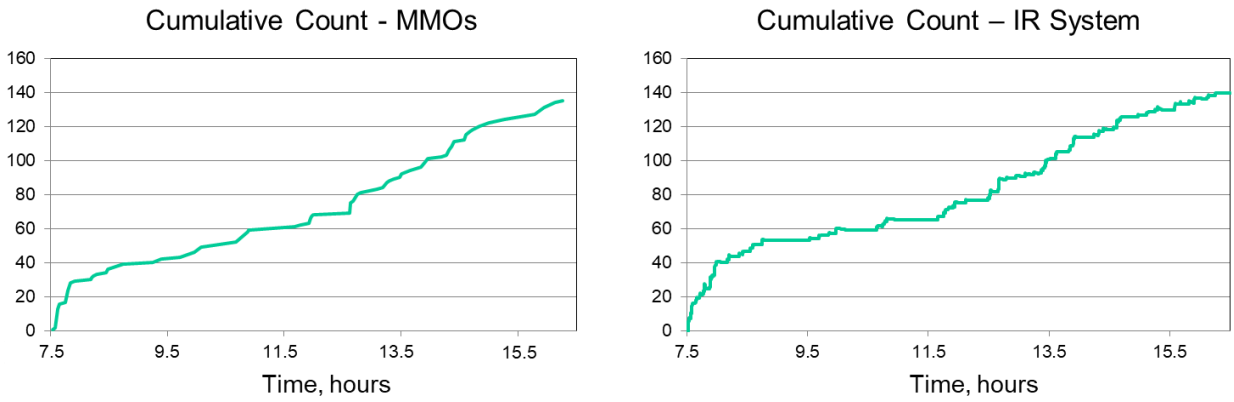

Figure 8: Comparison of Counts – Marine Mammal Observers (MMOs) (left) versus Semi-automated System (right) – January 7, 2014

6. SUMMARY

We have created a system for semi-automatically detecting, tracking, and counting gray whales during their migration. This system uses long wave infrared cameras to observe whale blows day and night. The resulting video is processed by a detection module which seeks to report a single detection for each blow made by each animal. False alarms can occur due to multiple detections on a single blow, the presence of other marine mammals, splashing birds, and ocean clutter such as whitecaps. The detections are reviewed by humans to eliminate most false alarms. The reviewed detection log is then fed into counting software which operates a particle filter to estimate the number of whales that are passing by based on the observed blows and any false alarms that make it through the human review process. Challenges in this estimation problem include the fact that the whales are only observed for brief periods (< 10s) out of several minutes of travel time passing by the location of the cameras. Additional complications include the fact that the whales tend to swim in pods where some blows will be missed due to occlusion by other blows, or difficulty resolving closely-spaced blows. In spite of these challenges, the system has proven to produce good results in comparison to human observers during the daytime and has the additional advantage of operating at night when human observers are ineffective and higher blow contrast compared to ocean clutter occurs.

The technology described here can be applied to several other problems involving the need to detect marine mammals. For example, there is currently a need to check for the presence of marine mammals prior to emitting loud acoustic signals to prevent potential damage to the mammals. These loud acoustic signals could arise in both naval and oil and gas operations and techniques for automatically detecting mammals in low visibility conditions (including at night) are of

great interest. Specific applications for the oil and gas industry include seismic surveys which may currently halt operations at night due to the inability of marine mammal observers to see the mammals. For naval operations, any active sonar system has the potential to disturb marine mammals within a range primarily determined by the power and frequency of the transmitted signal. Unmanned systems, such as the Common Unmanned Surface Vessel, are of particular interest in this regard as these systems do not have the ability to host human observers. Another potential application is for the prevention of ship strikes of large whales by fast-moving vessels. This technology could be deployed on ships and used to alert the vessel operator of possible whale locations allowing the operator to potentially take action to reduce speed or otherwise maneuver the vessel to avoid collisions with whales.

Deploying an IR detection system on a ship will require that the motion of the cameras is measured and/or stabilized. Stabilization can be provided and the motion of the camera can be measured using an inertial measurement unit (IMU) attached to the camera. Toyon is currently developing a vessel-based system using this approach.

7. REFERENCES

[1] Reeves, R.R., Smith, T.D., Lund, J.N., Lebo, S.A. and Josephson, E.A. 2010. Nineteenth-century ship-based catches of gray whales, *Eschrichtius robustus*, in the eastern North Pacific. Marine Fisheries Review 72(1):26-64.

[2] Durban, J.W., Weller, D.W., Lang, A.R. and Perryman, W.L. 2015. Estimating gray whale abundance from shore-based counts using a multilevel Bayesian model. Journal of Cetacean Research and Management 15:61-68.

[3] Perryman, W.L., Donahue, M.A., Laake, J.L. and Martin, T.E. (1999). Diel Variation in migration rates of Eastern Pacific gray whales measured with thermal imaging sensors. Marine Mammal Science 15(2):426-445.

[4] Zitterbart, D.P., Kindermann, L., Burkhardt, E. and Boebel, O. 2013. Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. PLOS ONE 8(8): e71217.

[5] Santhaseelan, V., Arigela, S., & Asari, V. K. (2012). Neural Network Based Methodology for Automatic Detection of Whale Blows in Infrared Video. In *Advances in Visual Computing* (pp. 230–240). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33179-4_23.

[6] Santhaseelan, V., & Asari, V. K. (2013, January). Whale blow detection in infrared video using fractal analysis as tool for representing dynamic shape variation. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop* (pp. 520-525). IEEE. doi:10.1109/WACV.2013.6475063.

[7] Streit, R.L., Poisson Point Processes: Imaging, Tracking, and Sensing, Springer, New York, 2010.

[8] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEEE Proceedings-F*, Vol. 140, pp. 107-113, April 1993.

[9] Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T., A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, *IEEE Trans. Signal Processing*, vol. 50, No. 2, 2002.

[10] A. Doucet, N. Gordon and V. Krishnamurty, "Particle filters for state estimation of jump Markov linear systems," *IEEE Transactions on Signal Processing*, Vol. SP-49, pp. 613-624, March 2001.

[11] Mallonee, J.S. (1991). Behaviour of gray whales summering off the Northern California coast, from Patrick's Point to Crescent City. Canadian Journal of Zoology 69:681-690.

[12] Rodriguez de la Gala-Hernandez, S., Heckel, G., Sumich, J.L. (2008). Comparative swimming effort of migrating gray whales (*Eschrichtius robustus*) and calf cost of transport Along Costa Azul, Baja California, Mexico. Canadian Journal of Zoology 86:307-313.

[13] Rugh, D. J., Muto, M.M., Hobbs, R.C. and Lerczak, J.A. (2008). An assessment of shore-based counts of gray whales. Marine Mammal Science 24(4):864-880.

[14] Stelle, L.L., Megill, W.M. and Kinzel, M.R. (2008). Activity budget and diving behavior of gray whales in feeding grounds off coastal British Columbia. Marine Mammal Science 24(3):462-478.

[15] Sumich, J.L. (1983). Swimming velocities, breathing patterns, and estimated costs of locomotion in migrating gray whales, (*Eschrichtius robustus*). Canadian Journal of Zoology 61:647-652.

[16] Blair, W.D., (2011). Design of Nearly Constant Velocity Track Filters for Brief Maneuvers, 14th International Conference on Information Fusion, Chicago, Illinois, USA, July 5-8, 2011, 978-0-9824438-3-5 ©2011 ISIF.