Some Approaches to Finding Birds in Video Imagery

Paul E. Allen
Charles E. Thorpe

CMU-RI-TR-91-34

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

December 1991

This research is sponsored by DARPA, DOD, monitored by the US Army Engineer Topographic Laboratories under contract DACA 76-89-C-0014.

List of Figures

Abstract

Manually censusing and monitoring bird colonies can be a time consuming, frustrating, and sometimes error prone process. Automation of censusing would efficiently provide consistent and easily collected data which would aid in the monitoring of bird colonies. This paper analyzes some typical scenes from bird colonies and describes our experiments of applying well-known computer vision techniques to those scenes. After a short description of the template matching, convolution, and motion analysis methods we used, we analyze four different classes of scenes using those techniques. They are as follows:

- Kittiwake colony scenes are analyzed by convolution with a filter especially constructed to find light circles with dark edges. Results of that analysis show an improvement over directly thresholding such images.
- Murre colony scenes prove to be highly resistant to analysis because of the coloration and crowding behavior of murres. However, we do show that it is possible to identify murres with profile presentations in such scenes by using template matching.
- Motion analysis techniques are applied to find auklets loitering on the surface of their talus slope nesting areas. Though the birds are practically invisible in these scenes, we show it is possible to find areas of high movement indicating the presence of birds.
- Thresholding and convolution are used to recover counts of birds from scenes of flying auklets. Ideas about tracking individual birds through such scenes are also explored.

Introduction

Manually censusing and monitoring bird colonies can be a time consuming, frustrating, and sometimes error prone process. Automation of censusing would efficiently provide consistent and easily collected data which would aid in the monitoring of bird colonies. The purpose of this paper is to analyze some typical scenes from bird colonies and describe our experiments of applying well-known computer vision techniques to those scenes.

There has been a limited amount of work applying computer vision techniques to censusing and monitoring birds. Bajzak and Piatt [2], and Gilmer et. al. [6] rely on thresholding and some post-thresholding analysis to estimate goose populations from digitized photographs. Anthony [1] uses similar methods to estimate goose populations from video imagery. Strong et. al. [12] describes using multispectral data to inventory geese and Best [4] analyzed the feasibility of using infrared measurements to census geese. Sidle [11] reports on an airborne video system used to monitor the availability of nesting habitat for cranes while Meisner [8, 9] describes a general purpose aerial videography system.

The remainder of this paper begins with a quick review of some of the techniques we employed in our analysis of images from this domain. Following a short section on the source and preparation of the video data we used, the major portion of this paper is devoted to a discussion of four different classes of scenes and analysis of examples from each. The first three classes are typical views of kittiwake, murre, and auklet colonies. The final scene class we analyze consists of views of a flock of flying auklets. After those four sections we summarize our observations about our analysis of the scene classes.

Analysis Techniques

The tools we applied to this domain are common image processing techniques, some of which may not have been applied to the domain before. They include thresholding, template matching, convolution, and optical flow or motion calculation. We will give a brief explanation of some of these techniques, but a full description of them is beyond the scope of this paper. See Ballard and Brown [3] for more detail.

Thresholding. Throughout our analysis we will be using thresholding a great deal but have not given the actual threshold values used. This is because they are generally picked in an ad hoc manner (e.g. "this value seems to work") and depend on the scaling of the image values. We have tried to use thresholding qualitatively instead of quantitatively so that particular threshold values are not important in themselves.

Template Matching. Template matching is a method for computing the "goodness" of a match between a patch of an image and a template. The result of a match calculation at a point in the image can be considered to be a distance measure. If $F(r, c)$ is the image and $T(i, j)$ is a template with $2x+1$ columns and $2y+1$ rows then the sum of differences squared, $d^2(r, c)$, between the image at (r, c) and the template is their match value and is computed as follows:

$$d^2(r, c) = \sum_{i=-y}^{y} \sum_{j=-x}^{x} [F(r+i, c+j) - T(i, j)]^2$$

A small match value at a point corresponds to a good match of the image patch centered at that point with the template; large values indicate a poor match. Depending on the analysis to be performed, the template to be matched can be part of the image itself, part of another image, calculated mathematically, or hand crafted. In our use of template matching, templates come from the images being analyzed.

In both template matching and the convolution technique discussed below, analysis of the edges of images is impossible since the computations depend on having pixel values in a neighborhood surrounding the pixel for which a result is being calculated. Thus there is a narrow border around images that contribute in the calculation of match and convolution for other points but do not have such values calculated for themselves.

Convolution. Convolution is similar to template matching in that a response between the image and a filter is calculated at each point in the image (except for pixels near the edge). If $F(r, c)$ is the image and $T(i, j)$ is the filter with $2x+1$ columns and $2y+1$ rows then the result or response, $F'(r, c)$, of the convolution between the two at (r, c) is computed as follows:

$$F'(r, c) = \sum_{i=-y}^{y} \sum_{j=-x}^{x} F(r-i, c-j)\ T(i, j)$$

Unlike template matching, a large response at a point corresponds to a good match of the filter and the image at that point; a small response indicates a poor match. The filter used for a convolution should have the property that $\sum_i \sum_j T(i, j) = 0$ so that it does not respond to featureless areas of constant intensity.

We used two filters for the convolutions in our analysis. The first is constructed to give a large response at points in the image that are dark circles surrounded by a light edge. The second filter is based on the first, but is specialized to respond at points with a light edge only on two opposite sides of the circle. The filters are shown in figure 1.

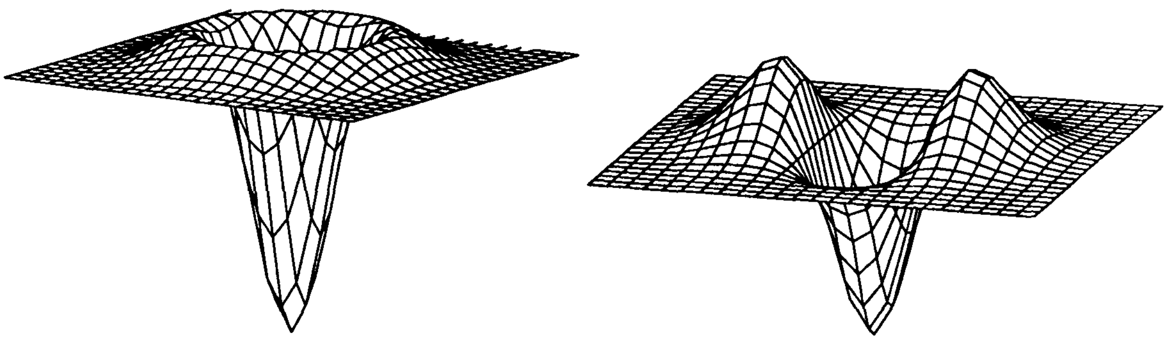

Figure 1: Filters constructed for use in convolution (left: h_1, right: h_2).

The basis for the first filter is the two dimensional Gaussian function $f(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$. Computing the Laplacian of the Gaussian function gives the first filter $h_1(x, y)$:

$$\begin{aligned} h_1(x, y) &= \nabla f(x, y) \\ &= \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \\ &= -2\frac{e^{-\frac{x^2+y^2}{2\sigma^2}}}{\sigma^2} + \frac{y^2 e^{-\frac{x^2+y^2}{2\sigma^2}}}{\sigma^4} + \frac{x^2 e^{-\frac{x^2+y^2}{2\sigma^2}}}{\sigma^4} \end{aligned}$$

The second filter, $h_2(x, y)$ is created by using the one dimensional Gaussian $g(x) = e^{-\frac{x^2}{2\sigma^2}}$ in one direction and the second derivative of a one dimensional Gaussian in the other direction:

$$h_2(x, y) = g(x) \bullet \nabla g(y)$$

$$= e^{-\frac{x^2}{2\sigma^2}} \left(- \frac{e^{-\frac{y^2}{2\sigma^2}}}{\sigma^2} + \frac{y^2 e^{-\frac{y^2}{2\sigma^2}}}{\sigma^4} \right)$$

In using these filters, σ is set to be equal to the expected radius, in pixels, of the inner dark circle to which the filter is to respond. Since the filters approach zero around the edges we found it sufficient to compute the filter to about 4σ pixels from the center. Note that to create a filter that responds to the opposite situation of a light circle with dark edges one need only negate the formulas above.

It should also be noted that the h_1 filter can be approximated by computing the difference of two appropriately sized Gaussians in each of the two dimensions. The primary advantage of this approximation is that the computation of the convolution can be decoupled into horizontal and vertical directions and then recombined to get the response values. This results in a big savings of computation time over calculating the convolution with the original form of the filter.

Optical Flow. The fundamental idea of optical flow or motion analysis is to compare two sequential images of the same scene and compute the speed at which the intensity of the pixels in the scene is changing. The presence of motion at a point in the scene is indicated by a high velocity or speed at that point. We will give only our actual implementation of these methods; see Ballard and Brown [3] for a full treatment of optical flow.

The basic operation we perform is to compute a velocity image, $V(r, c)$, of the scene to be analyzed. If $F_1(r, c)$, $F_2(r, c)$, ... is a time ordered sequence of images taken at regular intervals then they may be considered as a single function, $F(r, c, t)$, which gives the intensity of the image at row r and column c at time t. The velocity image at time t is computed as follows:

$$V(r, c) = \sqrt{[V_r(r, c)]^2 + [V_c(r, c)]^2}$$

where $V_r(r, c)$ and $V_c(r, c)$ are the components of velocity at time t in the row and column directions respectively. They are computed as follows:

$$V_r(r, c) = \frac{dr}{dt} = \frac{\frac{dF}{dt}}{\frac{dF}{dr}} = \frac{\frac{F(r, c, t) - F(r, c, t-1)}{1}}{\frac{F(r, c, t) - F(r-1, c, t)}{1}} = \frac{F(r, c, t) - F(r, c, t-1)}{F(r, c, t) - F(r-1, c, t)}$$

$$V_c(r, c) = \frac{dc}{dt} = \frac{\frac{dF}{dt}}{\frac{dF}{dc}} = \frac{\frac{F(r, c, t) - F(r, c, t-1)}{1}}{\frac{F(r, c, t) - F(r, c-1, t)}{1}} = \frac{F(r, c, t) - F(r, c, t-1)}{F(r, c, t) - F(r, c-1, t)}$$

This gives a measure of motion from an image pair which we will use in the analysis of scenes of an auklet colony.

Video Image Source and Preparation

Video Data Source. The video data we used was gathered at seabird colonies on Little Diomede Island, Alaska in the summer of 1991. With one exception, it consists of scenes typical in sampling and monitoring seabirds by manual methods. Originally recorded on an 8mm camcorder, we received the data after it was transferred to VHS format. Though only some of our analysis required it, we transferred the data once more to 3/4 inch Umatic video tape since our Umatic equipment allows sequential frame-by-frame digitization. Undoubtedly, these transfers have resulted in a loss of quality from the original.

Resolution. One of the biggest challenges in using video imagery to census or otherwise monitor birds is the relatively low resolution of standard video cameras. For effective censusing it is desirable to capture a fairly large number of birds in the field of view of the camera, but tradeoffs usually must be made to ensure that individual birds are recorded by enough pixels to facilitate recognizing the presence and, possibly, the species of a bird. (But see Strong et. al. [12] for a description of their multispectral system with a resolution less than the size of a bird.) The resolution of our data varied with the focal length of the lens and the distance to the foreground of the scene.

Color. Though the original video data we have is in color, all the analysis described here has been done with black and white images derived from the color images. This is done by averaging the intensities for the red, green, and blue bands at each pixel in the image. We believe that this does not result in the loss of much significant data since neither the birds nor the background in the scenes are very colorful. To the human eye, the most significant color in the video data we have is the orange of the bills of crested auklets, but the low resolution of video imagery negates any significance of that color in scenes containing that species.

Shift Correction. During the digitization process the video images sometimes get shifted a few rows and/or columns one way or the other. It is important that this shift be taken into account if the analysis to be done relies on matching sequential frames of the same scene. The cause of this shifting is unknown but probably has its roots in the many transfers the video footage we are using has gone through. The shifting is easily corrected for by using template matching to calculate a match between a small patch from each image. This results in a pair of offsets such that $F_1(r, c) = F_2(r + row\_offset, c + column\_offset)$. This correction works best if the matching template contains a distinct edge or some other high contrast feature. Note that using patches with only horizontal edges or only vertical edges is to be avoided since that will most likely result in a meaningful correction for only rows or columns respectively. Some of the video data had a time and date stamp in the lower right corner of the frame; we used a portion of that as the template when correcting for shift in that data. In other scenes we used the intersection of approximately orthogonal edges for calculating shift correction.

Kittiwakes

Kittiwakes are gulls that breed in colonies and build nests on ledges and outcroppings of cliffs. Their head, shoulders, and underside are white while the rest of their body is gray. The density of birds depends mainly upon nest density which can be fairly high when nests are situated along a ledge. Cliff faces and outcroppings at colonies are often covered with substantial amounts of guano turning naturally dark colored rock white. Figure 2 shows two different kittiwake cliffs at different distances.

There are two fundamental problems to overcome before automating the censusing of kittiwake colonies. The first is to assure that guano-covered rocks and outcrops are not mistaken for the white portion of kittiwakes. The second is to take into account the fact that birds are sometimes so close to one another that they appear as a single bird in a video image. Such close proximity of birds often happens only when the two adults tending a nest are simultaneously present.

As can be seen from figure 3, thresholding by itself is insufficient for counting kittiwakes because of the problem with guano-covered rocks (and, in this case, bright water reflections in the upper left corner of the left image). For our analysis we used convolution with a filter that was the negative of the h_1 filter developed above. This gave a light-on-dark filter with a large response to small, light-colored kittiwake blobs on dark rocks while minimizing the response to larger, light-colored guano-covered rocks. The result of the convolving the images from figure 2 with the filter are shown in figure 4.The σ parameter in the filter is set to 4 pixels for the left image and 6 pixels for the right image.

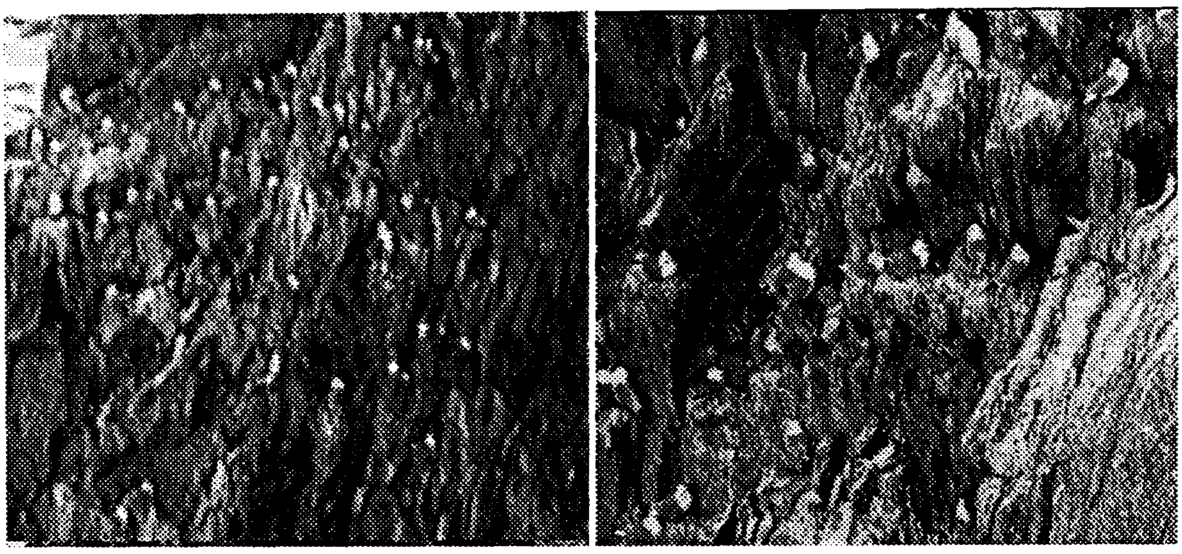

Figure 2: Left: Long distance image of kittiwake nesting cliff.
Right: Closer view of different cliff.

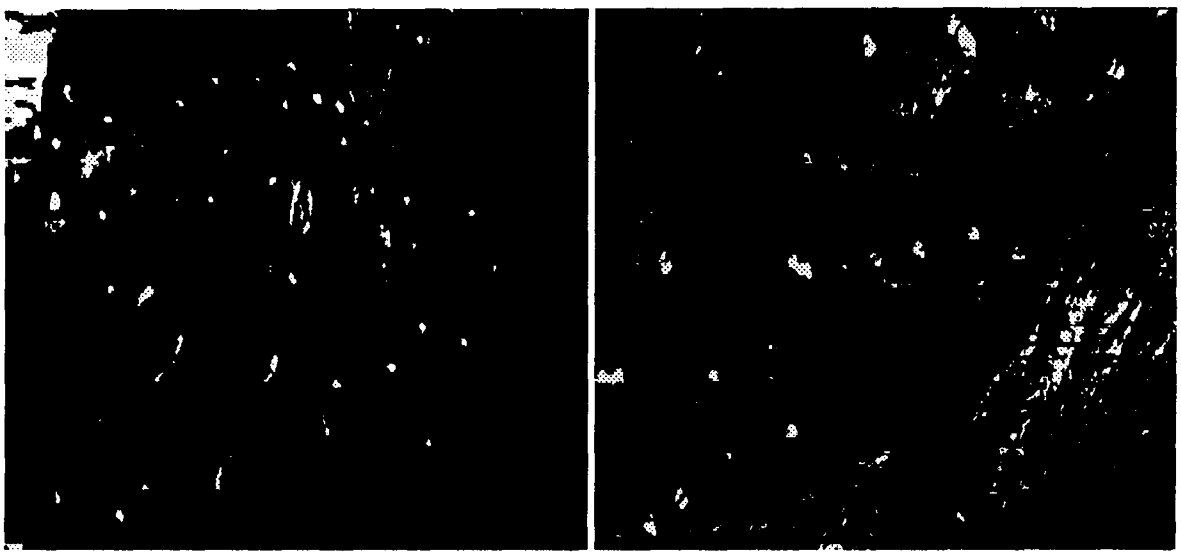

Figure 3: Result of thresholding images in figure 2.

Thresholding the convolution responses from the right image of figure 4 at two different values gives the images in figure 5. The right image corresponds to a relatively high threshold value while the left image uses a lower threshold. A manual count of the kittiwakes in the original image gives 13 kittiwakes within the range of the convolution filter. (The two leftmost kittiwakes are out of filter range as is the kittiwake closest to the lower left corner.) Using a high threshold gives 10 areas of high response, 9 of which are actually kittiwakes. The remaining high response area is a rock. As shown in the left image of figure 5, lowering the threshold to get the most number

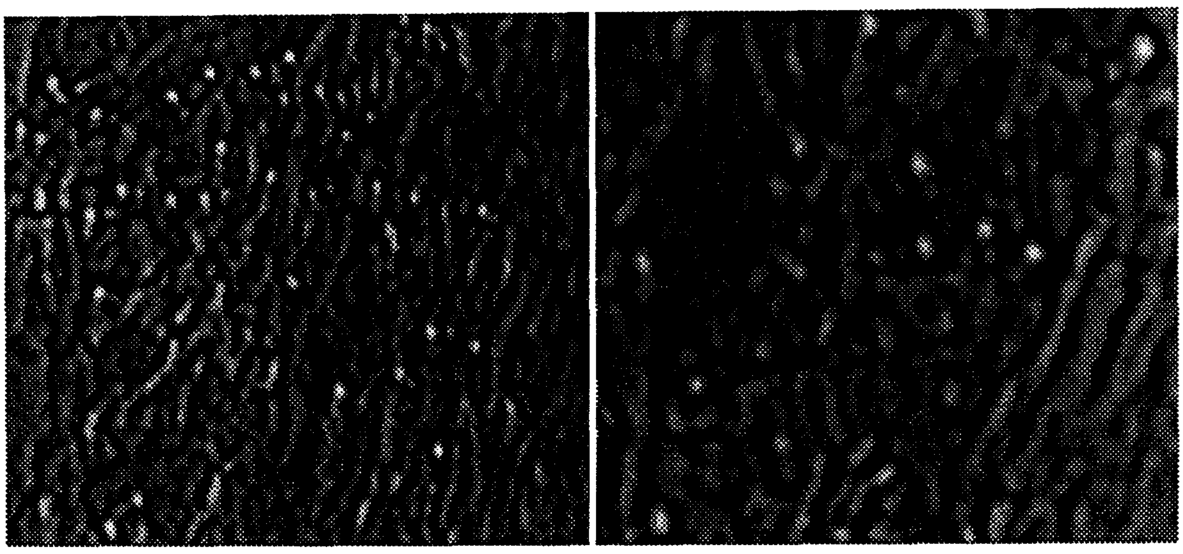

Figure 4: Convolution of images in figure 2 with $-h_1$ filter (left: $\sigma = 4$, right: $\sigma = 6$).

of kittiwakes results in finding only 12 kittiwakes and many other high response areas that are rocks and murres. Note that the filter would never find all 13 kittiwakes in this image because the two kittiwakes just left of center are standing at the same nest and are too close together for the filter to distinguish separately.

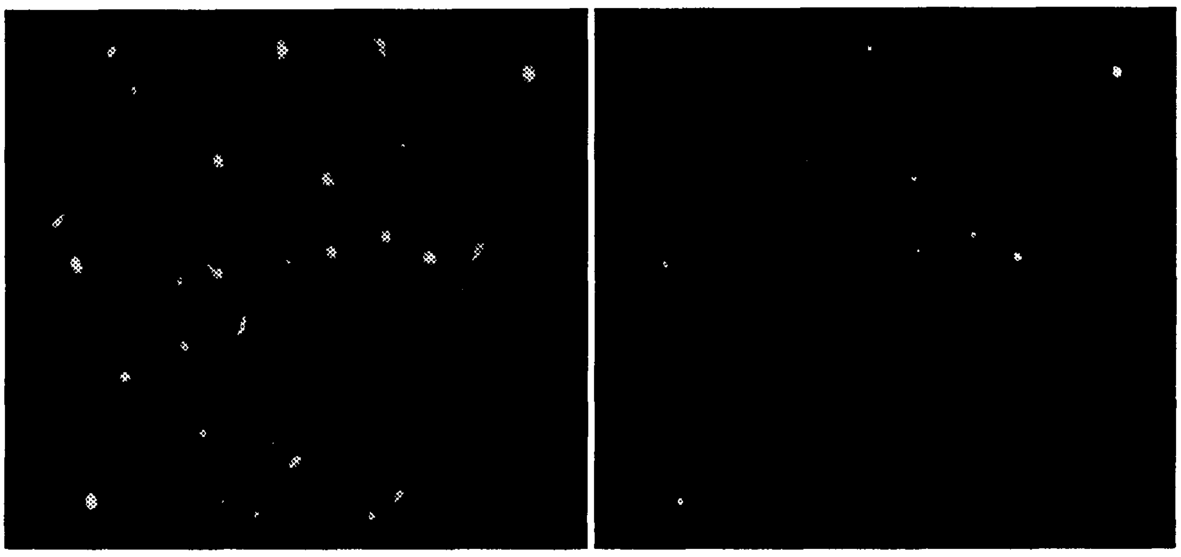

Figure 5: Result of thresholding the right image in figure 4 (right: relatively high threshold, left: lower threshold).

A manual count of kittiwakes in the left image of figure 2 gives about 38 birds. That is a rough figure because the quality of the image is extremely poor and even when watching the video it is difficult to tell which small white blobs are birds and which are guano. (At this level of quality and resolution ground truth counts would have to be made at the site of video data collection.) The left image of figure 6 is the result of thresholding the left image in figure 4. It shows that most of the kittiwake-looking portions of the original image are found as well as some guano-covered rocks. Some positively identified kittiwakes are missing from the threshold image while some positively identified rocks are included. There are many spots of high response for which we are unable to determine the true identity. Notice that large patches of guano from the original image are not mistaken for kittiwakes. Only smaller patches or fingers of larger guano patches that are about kittiwake size give a high response and are ambiguous.

After noticing the vertical striping in the response of the h_1 filter on the left image of figure 2 we applied the negative of the h_2 filter to the original image to try to minimize the striping. The threshold image of the convolution response with the $-h_2$ filter is shown on the right of figure 6. That result is similar to the result with the $-h_1$ filter but responds to different kinds of non-kittiwake features.

We believe that this analysis has shown that, though far from perfect, the application of convolution using these types of filters represents progress over directly thresholding such images. Future work could be directed at using the high response areas to define areas of interest in images where detailed motion analysis could be performed. It also seems the analysis of these types of scenes could benefit from using data in other spectral bands where birds have different intensity than guano.

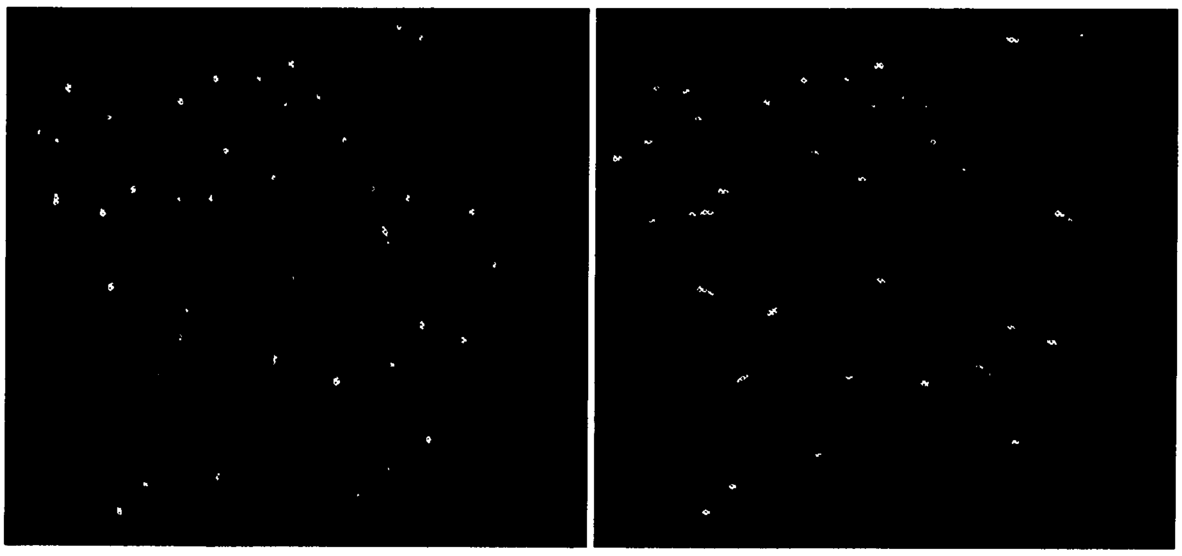

Figure 6: Left: Result of thresholding left image in figure 4.
Right: Result of thresholding convolution response of $-h_2$ filter with left image of figure 2.

Murres

Murres also breed in colonies on cliffs but do not build nests; eggs are laid directly on ledges and outcroppings. These birds are often packed close together so that they touch and crowd each other. The body of murres are black except for their white undersides. Actually, two species of murres are usually found in the same colonies, but differentiation can be made only by fine distinctions visible to the human eye but not apparent in low resolution video imagery. The left image in figure 7 shows a portion of a murre cliff in a presentation typical for censusing while the right image of that figure gives a closeup of the birds.

Figure 7: Left: Typical census presentation of murres with template outlined. Right: Closeup of murres.

Challenges to the automated censusing of murre species are their severe crowding behavior, dark coloration, and the inconsistent presentation of their highly visible white underside. Visibility of the underside may range from a full frontal view to the more common full rear view where the underside is not visible at all.

These challenges make analysis of murre images extremely difficult. We had limited success in applying template matching to murre images to pick out individuals with a profile presentation. Such an individual is outlined in the left image of figure 7. The result of using that outlined area as a template and computing the match of the template across the image is shown in the left image of figure 8. Thresholding that result image gives the image on the right of figure 8 showing three small dark blobs where good matches occur. The rightmost blob corresponds to the outlined bird used as the template. The other two blobs correspond to two other murres with a similar profile posture that have successfully been identified. Though not completely encouraging, this result does show that it is possible to do limited analysis of murre images when the birds cooperate by giving a profile view and exposing their high contrast sides.

Auklets

Auklets breed in colonies situated on talus slopes and nest in the cavities and interstitial spaces created by the talus. They are counted during daily activity periods when they loiter on the surface of the talus. Often more than one species is present, but the species are usually easily distinguished by humans through gross differences in coloration and size. Though those distinctions are apparent even at relatively low resolution videography it is beyond the capabilities of current computer analysis to differentiate the species at the even lower resolutions necessary to make counting efficient. All trials that we describe here consider auklets without regard to species.

Though the birds are indistinguishable from the talus, the left image in figure 9 shows a typical scene of auklets loitering on the surface of their colony. Even in a still video image shown on a display, viewers are hard put to find the auklets in the scene. While watching the video being played, however, it is much easier to find the auklets because the eye picks out their movement against the background. This observation led us to apply motion analysis techniques to this video sequence.

Figure 8: Left: Result of template matching between the left image of figure 7 and the outlined template in that image. Right: Result of thresholding left image.

A good count is hard to determine, but we estimate there are at least 20 birds in the scene. From about eleven seconds of video of this scene we selected 23 images spaced one half second apart for the following analysis. Shift correction using a high contrast rock corner was performed to this sequence before analysis. This removes most shifting due to digitization and inadvertent camera movement.

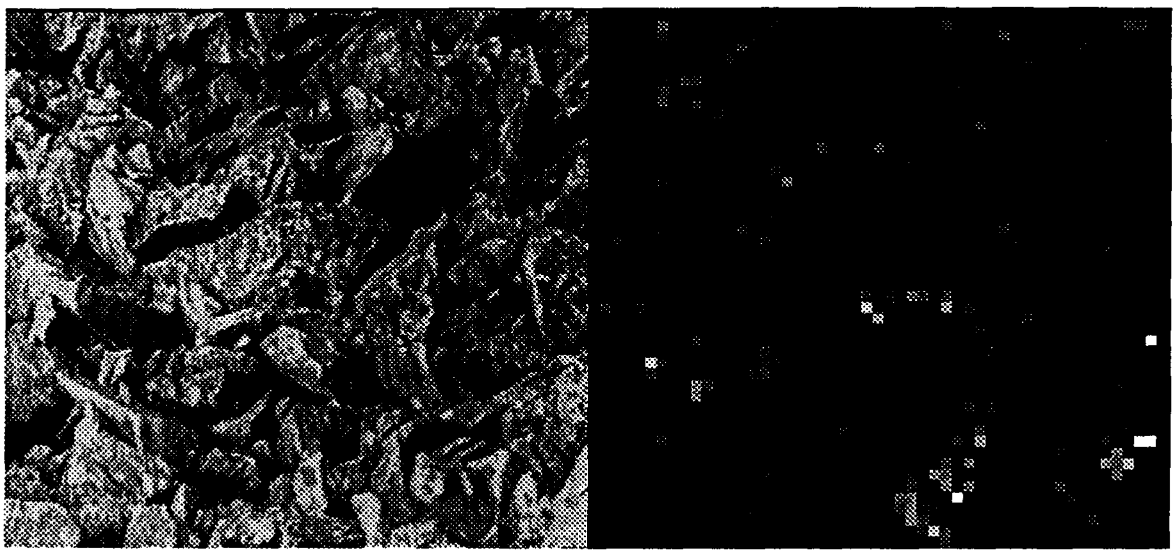

Figure 9: Left: Auklets on talus slope. Right: Result of motion analysis of talus slope sequence.

The output image on the left in figure 9 shows a measure of how much motion occurred over the sequence in each section of the input images. The sections or blocks in the output image correspond to 10×10 pixel regions of the input images. The value of an output block is an average of velocities for points in that block with significant computed motion, calculated as follows. A velocity image was computed for each input image pair and the 16 points from the velocity image with largest values from that pair were selected. The velocity of these high-motion points were used to update the running averages of their corresponding blocks in the output image.

The success of this analysis is difficult to discuss in anything but a general way. The bright blocks in the output image correspond to areas of a great deal of movement in the original sequence. Regions showing many less intense points also correspond to where movement is seen in the original sequence. Black regions in the output image indicate places where no significant movement was found. It is almost impossible to tell if the non-black singleton points in the output image correspond to actual movement or not. It seems that the results of motion analysis would have to be combined with the original sequence in some way so they could be viewed together to see if there is a correspondence between motion observed and motion detected from analysis. Falsely detected movement is possible when critical image features such as sharp edges are between the pixel elements in the digitizer. This can cause an edge to appear in different adjacent pixels in sequential images resulting in the detection of motion when there actually is none.

Flying Auklets

Part of our analysis efforts used some imagery that showed crested auklets flying in a flock fairly close to the camera. As shown on the left of figure 10, the background is of a plain sky while the foreground consists of birds flying away from the camera at an angle very oblique to the direction of view. Though we are unaware of any sampling method that uses this presentation of birds we analyzed it because it looked particularly amenable to automated counting methods. Whether automation of this type could be developed into a sound and meaningful biological sampling method remains to be seen.

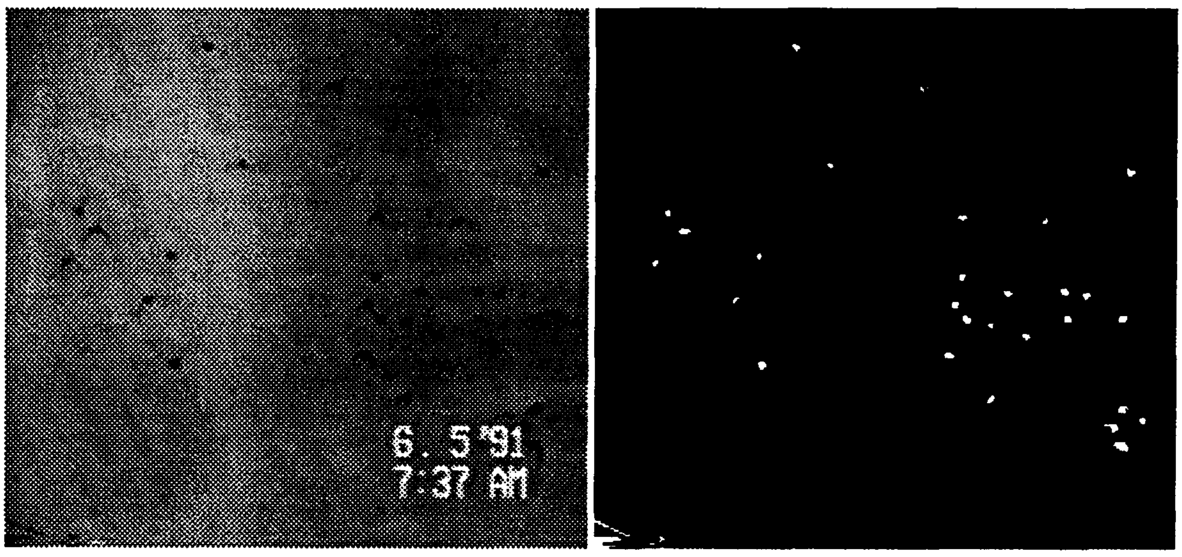

Figure 10: Left: Crested auklets flying. Right: Threshold of left image.

As shown in the right image of figure 10, simple thresholding works quite well and captured all the birds with only a bit of extraneous data in the lower left corner. The result of convolving the h_1 filter ($\sigma = 5$) with the original image from figure 10 is shown in the left image of figure 11. Thresholding that response image gives the image on the right of figure 11 which clearly identifies all the birds and nothing else.

Though it seems that simple thresholding could be made to work with a little postprocessing, convolution provides a better and more robust method of analysis for scenes of this type. Convolution should provide better results in in the face of less than perfect images showing, for example, high contrast clouds in the sky or different lighting conditions. Note that neither of these methods gets the correct result when some birds partially or fully occlude others. One bird partially occluding another will probably result in the occluded bird being considered part of the occluding bird so the two will be counted as one. This problem could be minimized by counting birds in a few images taken several (5-10) frames apart and using the largest or average count from the images.

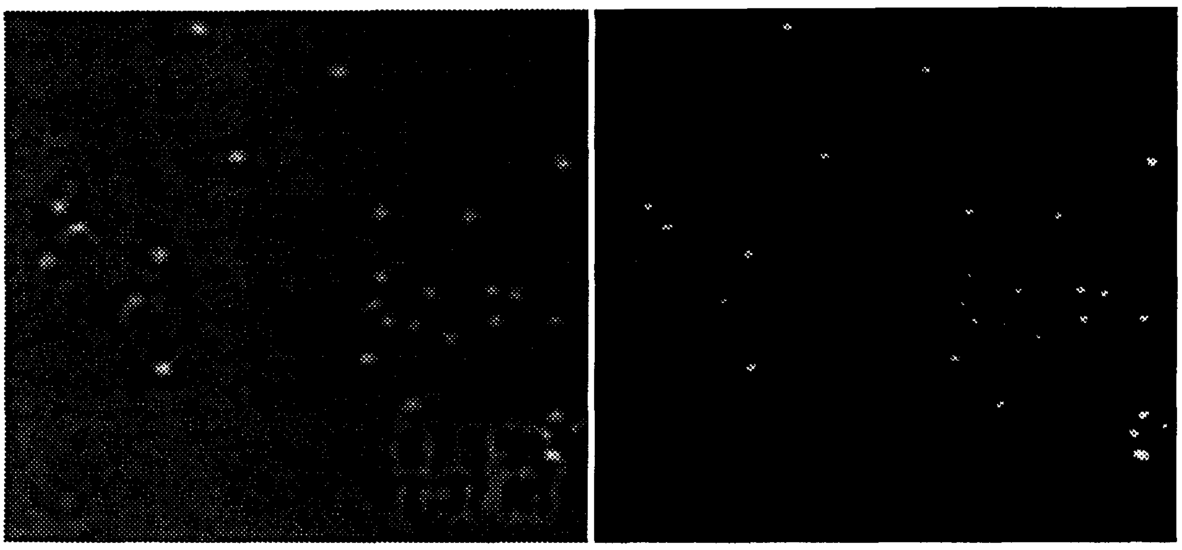

Figure 11: Left: Result of convolving image in figure 10 with h_1 filter and $\sigma = 5$.
Right: Result of thresholding left image.

We also had limited success in tracking individual birds through sequential images of this scene . We digitized 15 sequential frames of the scene and applied a very simple tracking algorithm to the sequence. Digitization shift correction was done using the "AM" portion of the time stamp in the lower right corner of the images. After convolving the h_1 filter over the first image of the sequence, the response was thresholded to find the initial positions of the birds. In subsequent images convolution was used to search the areas around the previous known positions of the birds. A suitably high response in the search area indicated the presence of the bird. Otherwise the search area was increased and the convolution was performed again. This method worked quite well for areas of low bird density but failed spectacularly in high density areas such as the lower right quadrant of the scene. Failure in high density areas was expected because of the simplicity of the algorithm. The method also failed to take into account birds entering and leaving the scene. One simple improvement we made to the algorithm was to calculate a trajectory for each bird and start the search centered in the area of the image where the bird was expected to be instead of where it was found in the previous image. We did not spend too much time on this tracking algorithm because a great deal of work has already been done in the area of tracking multiple moving targets. Reid [10] and David et. al. [5] are just two of many papers available on the topic.

We noted that it is important to perform digitization shift correction on sequences such as this to correct for movement of the image during digitization. Note that in this case the correction was performed on video data not actually in the scene but overlaid on the scene by the camera. In the motion analysis of auklets the shift correction was done using a feature actually in the scene. It is important to realize the difference between the results of these two applications of shift correction. The correction in the motion analysis was being done using a scene feature and thus corrected for digitization shift *and* camera movement at the same time (though in that case camera movement was insignificant because it was mounted on a tripod). In the present analysis, there is significant camera movement because the camera is being hand held. Since we used a feature for shift correction that is not actually in the scene, only digitization shift and not camera movement is corrected.

In the flying auklet sequence the birds, and the camera are both moving with respect to the ground and each other. Thus camera movement is difficult to distinguish from bird movement. The trajectories of some of the birds in the sequence are on the left of figure 12. We can pick a subset of the birds and compute their centroid for each image in the sequence. The movement of such a centroid is shown on the right of figure 12. That graph shows the estimated combined effect of camera and flock movement. If a similar sequence were recorded with a stable camera and the same analysis performed, such a graph would be a good estimate for flock movement alone. In that case, we would hope to see more consistent motion from frame to frame for individual birds. This would make tracking much easier. The availability of new video cameras with built-in "steady-cam" processing that corrects for camera movement might keep even hand held sequences steady.

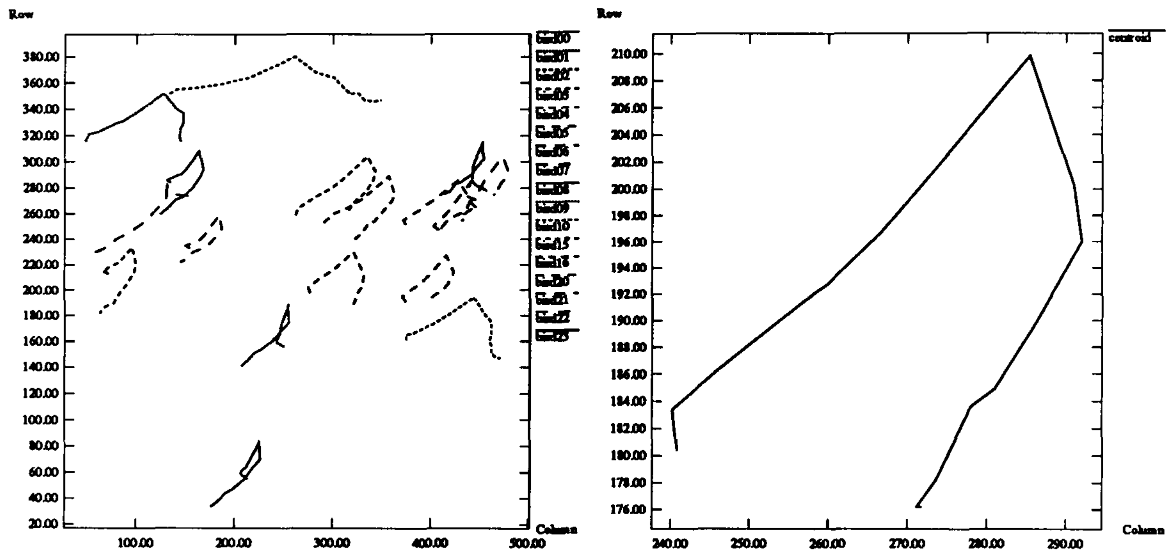

Figure 12: Left: Result of hand tracking several birds from flying auklet sequence.
Right: Centroid of selected birds throughout sequence.

Summary

In an attempt to come closer to automated bird censusing we have applied fairly simple, well-known computer vision techniques to scenes commonly encountered in bird colonies.

- Kittiwake colony scenes were analyzed by convolution with a filter especially constructed to find light circles with dark edges. Results of that analysis showed an improvement over directly thresholding such images.
- Images taken of murre colonies proved to be highly resistant to analysis because of the coloration and crowding behavior of murres. We did, however, show that it is possible to identify murres with profile presentations in such scenes through the use of template matching.

- We explored using motion analysis to find auklets loitering on the surface of their talus slope nesting areas. Given the practical invisibility of these birds in typical talus slope scenes, we feel that our initial results of finding areas of high movement in the scene are promising, but are a very, very long way from any practical application.
- Though recovery of good counts from scenes of flying auklets could be done with thresholding and a little post-threshold processing, we believe convolution with a filter would prove to be more robust in images with different lighting conditions and backgrounds. Ideas about tracking birds through such scenes were also explored. We do not know if the counts recovered from scenes of flying birds have any meaning with respect to biological sampling.

We hope that this work will be of use to other researchers interested in using computer vision techniques in this domain. With the growing availability of higher resolution video cameras and infrared video equipment [7] significant progress should be possible in the near future.

Acknowledgments

We would like to thank Scott Hatch and David Irons of the U.S. Fish and Wildlife Service for providing the video data used in this analysis.

References

[1] Anthony, R. Michael.
Estimating Brant Populations from Video Images.
Technical Report, Alaska Fish and Wildlife Research Center, Anchorage, Alaska 99503, 1991.

[2] Bajzak, Denes and Piatt, John F.
Computer-Aided Procedure for Counting Waterfowl on Aerial Photographs.
Wildlife Society Bulletin 18(2):125-129, 1990.

[3] Ballard, Dana H. and Brown, Christopher M.
Computer Vision.
Prentice-Hall, Edgewood Cliffs, New Jersey, 1982.

[4] Best, R. G.
Infrared Emissivity and Radiant Surface Temperatures of Canada and Snow Geese.
Journal of Wildlife Management 45(4):1026-1029, 1981.

[5] David, Philip; Balakirsky, Stephen; and Hillis, David.
A Real-Time Automatic Target Acquisition System.
Technical Report, Harry Diamond Laboratories, Adelphi, Maryland 20783, 1991.

[6] Gilmer, David S.; Brass, James A.; Strong, Laurence L.; and Card, Don H.
Goose Counts from Aerial Photographs using an Optical Digitizer.
Wildlife Society Bulletin 16(2):204-206, 1988.

[7] Kodak.
High Resolution in a Real-Time IR System.
1990
Advertisement and data sheet.

[8] Meisner, Douglas E.
Fundamentals of Airborne Video Remote Sensing.
Remote Sensing of Environment 19(1):63-79, 1986.

[9] Meisner, Douglas E. and Lindstrom Orville M.
Design and Operation of a Color Infrared Aerial Video System.
Photogrammetric Engineering and Remote Sensing 51(5):555-560, 1985.

[10] Reid, Donald M.
An Algorithm for Tracking Multiple Targets.
IEEE Transactions on Automatic Control AC-24(6):843-854, 1979.

[11] Sidle, John G. and Ziewitz, Jerry W.
Use of Aerial Videography in Wildlife Habitat Studies.
Wildlife Society Bulletin 18:56-62, 1990.

[12] Strong, Laurence L.; Gilmer, David S.; and Brass, James A.
Inventory of Wintering Geese with a Multispectral Scanner.
Journal of Wildlife Management 55(2):250-259, 1991.