Automated Visual Monitoring of Nesting Seabirds

Patrick Dickinson Chunmei Qing Shaun Lawson
University of Lincoln, UK
pdickinson,cqing,slawson@lincoln.ac.uk

Robin Freeman
Microsoft Research, Cambridge
v-robfr@microsoft.com

Seabird populations are considered an important and accessible indicator of the health of marine environments: variations have been linked with climate change and pollution [9], as well as changes in fish stock levels. However, manual monitoring of large populations is labour-intensive, and so necessarily limited in scope. In this paper we present work currently being conducted as a pilot to develop computer vision as a means of automatically monitoring nesting birds. The long-term objective is to provide ecology researchers with behavioural data on a scale not currently available.

We begin by describing the context and objectives for our work, which are centred around on-going manual monitoring of a specific population of Common Guillemots on Skomer Island, West Wales (UK). This project was initiated in February 2010, and as such still in an early stage of development: from this perspective we describe our forthcoming data collection programme, anticipated technical challenges, and initial development of video and image processing techniques.

1. Context

Skomer Island is located in West Wales (UK), covers an area of approximately $3Km^2$ and is home to internationally significant populations of seabirds including Atlantic Puffins, Manx Shearwaters, Guillemots, and Razorbills. It is designated as a "site of special scientific interest" and the surrounding waters are protected as a marine nature reserve. Skomer is administered by the Wildlife Trust of South and West Wales.

A long-term study of Skomer‘s Guillemot population is being conducted by the Evolutionary Biology and Behavioural Ecology Research Group at the University of Sheffield (UK). This study is focussed on the breeding activity and success of a particular population which nest on a cliff face known as "The Amos" (see figure 1). The position of this cliff face is particularly convenient in that it is viewable from an opposing cliff position with minimal disturbance. Currently, field researchers use manual methods to gather data, including population size estimates, samples of chick survival rates, and leg-ringing. However, it is not feasible to manually gather more detailed data about the daily activity of birds; for example, how long individuals spend foraging for food.

The Centre for Computational Ecology and Environmental Sciences (CEES) at Microsoft Research UK is active in developing technological methods of eliciting data concerning the behaviour of birds on Skomer, included the use of GPS trackers to monitor the feeding behaviour of Manx Shearwaters [5]. The project described in this paper is intended to augment the work of CEES by developing computer vision as a means of provisioning detailed population-level data about nesting birds, in a non-intrusive manner.

1.1 Objectives

Ecologists studying Guillemots on The Amos are able to gather characteristic global data about each breeding cycle using manual methods. For example, field workers can determine whether a particular breeding pair successfully fledge a chick. In addition, Guillemots return to the same nest position each year: positions of individual nests can be mapped accurately, and the performance of a pair logged over many seasons. What is unknown is their more detailed hour-by-hour patterns of behaviour. Of particular interest is the amount of time spent attending their nests, as this could be a key factor in chick survival (for many possible reasons, for example see [1]).

The objective of this project is to pilot the development of vision processing algorithms capable of extract-

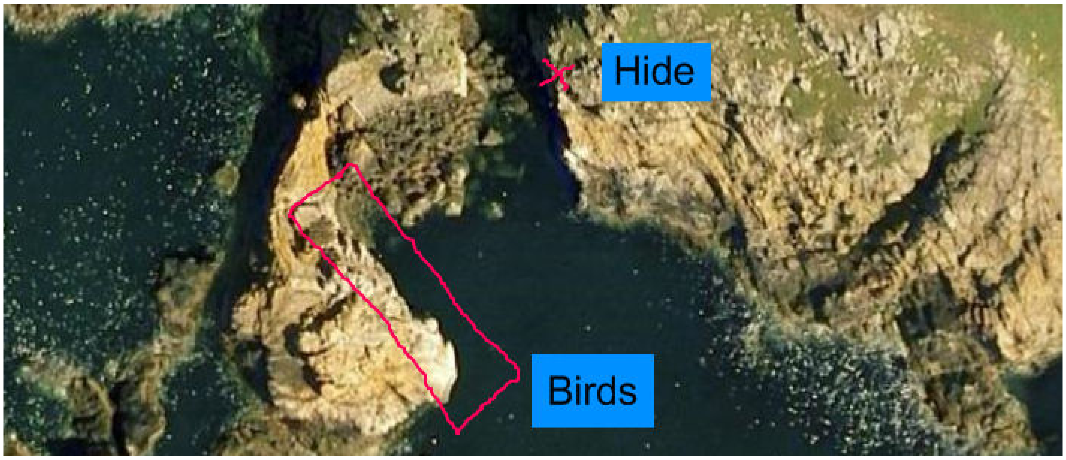

Figure 1. "The Amos" Skomer Island, and hide position.

ing this type of data from video and/or still images of the nesting site. It is also intended to provide a prototype processing tool which can be used by ecologists to select and extract relevant information. For example, a researcher may wish to determine the attendance at a specific nest from a corpus of video data, and correlate this with manual observations.

The contribution of our work derives partly from the area of application. Whilst some previous work has been conducted into the use of computer vision to monitor birds [2, 6, 7], the extraction of population-level data from cliff nesting sites represents some unique challenges. We are currently developing new processing algorithms to address these challenges, which are outlined in Section 4.

2. Data Collection

A set of previously captured data was made available at the beginning of the project. This comprises a series of high-resolution images captured during a previous season, using a programmable digital still camera. The camera was located at the hide position shown in figure 1. Images were captured approximately every 60 seconds over a period of days: an example image is shown in figure 2. An expanded section is shown in 3, which shows the level of detail in these images. Some video sequences of Guillemots captured at a different site (Lundy Island, UK) were also provided.

This pre-existing data is being used as the basis for initial development work. However, extensive and coordinated data sets will be collected at The Amos during June 2010. Firstly, time-lapse image sequences will be captured using multiple digital still image cameras, using varying image resolutions and frequencies. Secondly, a high-definition camcorder will be used to capture video data over a period of approximately one week. Reference images of the site have been captured prior to the start of the breeding season (when no birds are present). We intend to make these data sets available publicly, through the project website, in due course.

Figure 2. Image of The Amos captured using a 7MP still camera.

3. Technical Challenges

The provision of behavioural data is dependent on segmentation, both in single images and video. This is a key technical challenge of this project. Previous works have utilised standard background differencing techniques [6, 7]; however, such algorithms are susceptible to false-positive detections caused by dynamic

Figure 3. Detail from figure 2.

background processes, and camera movement. The elimination of these types of errors is one initial focus. Other anticipated challenges include poor image quality caused by adverse weather, restrictions on camera placement, and high population density in some areas.

4. Approaches

At the time of writing the project is still in its initial phase. However, work has thus far focussed on two separate methods for segmenting individual Guillemots. This two-pronged approach is motivated by the availability of two quite different modes of data; namely still images and video.

4.1 Segmenting from Video

Segmentation to support automated visual surveillance systems has been well-studied; however, most existing techniques have been developed for monitoring urban and indoor environments. The use of such techniques in natural scenes presents technical challenges which are not yet well-solved (see section 3). Our initial approach to this problem develops from our previous work which has shown that region-based rather than pixel-based models can be effective at eliminating errors caused by dynamic backgrounds. We are currently experimenting with a probabilistic region-based background model represented by an adaptive mixture of Gaussians (MoG) in 5-dimensional space (3 colour co-ordinates and 2D image coordinates). A single component of the model represents a set of background pixels with similar colour and spatial characteristics, and is described using the usual Gaussian form:

$$p(\mathbf{x}_t|\theta_{(j,t)}) = \frac{e^{-\frac{1}{2}(\mathbf{x}_t-\mu_{(j,t)})(\mathbf{\Sigma}_{(j,t)})^{-1}(\mathbf{x}_t-\mu_{(j,t)})}}{\sqrt{(2\pi)^5|\mathbf{\Sigma}_{(j,t)}|}} \quad (1)$$

where $\mathbf{x}_t$ is a pixel observation at time t, and the parameters $\theta_{(j,t)} = \left\{\omega_{(j,t)}, \mu_{(j,t)}, \mathbf{\Sigma}_{(j,t)}\right\}$ describe the weight, mean, and covariance matrix of the j^{th} model component at time t. Training is effected using a recursive component splitting procedure, and refined using the well known Expectation Maximisation algorithm for MoGs. A visualisation of an example background model is shown in figure 4.

Figure 4. Visualisation of a region based background model.

The model is used to classify pixels as either foreground or background using a multi-class Markov Random Field (MRF). The *maximum a posteriori* field labelling is estimated using Gibbs sampling. This combination of a region-based background model and MRF classifier represents a novel approach to the problem of robust segmentation under dynamic/unstable backgrounds, and is described in detail in [4]. Currently we use a uniform distribution to represent the foreground class; however, we are intending to use trained foreground colour models to refine the segmentation process.

4.2 Segmenting from Still Images

Aside from the method described in section 4.1, we are also developing a feature-based method of locating Guillemots in single images. Our proposed approach is based on a popular face detection algorithm developed by Viola and Jones [8], which learns a set of Haarlike features for known objects of interest from a set of training images. We are currently experimenting with adaptations of these block-based features by incorporating distributions of local intensity gradients [3]. The learning procedure uses Adaptive Boosting (Adaboost) to combine sets of "weak" classifiers into a robust classifier capable of identifying objects in test images, although we are also investigating the use of support vector machines (SVM) with the same types of feature. A small selection of example positive and negative training images are shown in figure 5.

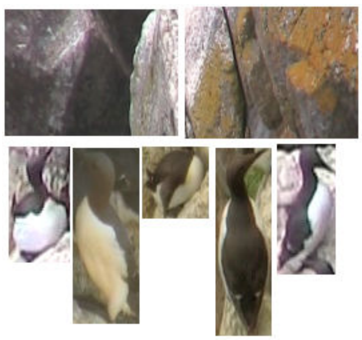

Figure 5. Positive and negative training images for feature-based segmentation.

Our motivation for investigating this method derives from inspection of our initial image set: the Guillemots' black and white plumage presents both strong image gradients, and distinctive and persistent localised shape characteristics which appear relatively insensitive to orientation. Our second rationale is that both algorithms have been shown to be highly efficient in detecting human figures and faces, and facilitate real-time implementation. At the time of writing this method is still in an early phase of development: we expect to be able to quantify performance following data collection during June 2010.

5. Project Updates

Further information on the progress of this project is available at:
http://webpages.lincoln.ac.uk/pdickinson/

6. Acknowldgements

The work presented in this paper is supported by EPSRC grant reference EP/H017143/1.

References

[1] K. Ashbrook, S. Wanless, M. Harris, and K. Hamer. "Hitting the buffers: conspecific aggression undermines benefits of colonial breeding under adverse conditions". *Biology Letters*, 4(6), pp. 630633, 2008.

[2] T. Burghardt, N. Campbell, P. Barham, I. Cuthill and R. Sherley. "A fully automated computer vision system for the biometric identification of african penguins (Spheniscus demersus) on Robben Island". *6th International Penguin Conference*, Tasmania, 2007

[3] N. Dalal and B. Triggs. "Histograms of oriented gradients for human detection". Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 2005.

[4] P. Dickinson, A. Hunter and K. Appiah. "Segmenting video foreground using a multiclass MRF". International Conference on Pattern Recognition, 2010.

[5] T. Guilford, J. Meade, J. Willis, R.A. Phillips, D. Boyle, S. Roberts, M. Collett, R. Freeman, and C.M. Perrins. "Migration and stopover in a small pelagic seabird, the Manx Shearwater Puffinus puffinus: insights from machine learning". *Proceedings Royal Society B*, 276(1660), pp.1215-1223, 2009.

[6] A. Kembhavi, R. Farrell, Y. Lou, D. Jacobs, R. Duraiswami, and L. Davis. "Traking down under: following the Satin Bowerbird". In Proceedings of *IEEE Workshop on Applications of Computer Vision*, Copper Mountain, CO, USA, 2008.

[7] D. Song, N. Qin, Y. Xu, C. Young Kim, D. Luneau, and K. Goldberg. "System and algorithms for an autonomous observatory assisting the search for the Ivory-Billed Woodpecker". In Proceedings of *IEEE Conference on Automation Science and Engineering*, pages 200205, Arlington, VA, 2008.

[8] P. Viola, M. Jones,"Robust real-time face detection". *International Journal of Computer Vision*, 57(2), pp.137-154, 2004.

[9] S. Votier, T. Birkhead, D. Oro, M. Trinder, M. Grantham, J. Clarkand R. McCleery, and B. Hatchwell. "Recruitment and survival of immature seabirds in relation to oil spills and climate variability". *Journal of Animal Ecology*, 77(5), pp. 974983, 2008.