Ecological Informatics 33 (2016) 24–31

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Seabird image identification in natural scenes using Grabcut and combined features

Suxia Xu, Qingyuan Zhu *

[a] Cognitive Science Department, Xiamen University, Xiamen 361005, China
[b] Fujian Key Laboratory of the Brain-like Intelligent Systems (Xiamen University), Xiamen 361005, China

A R T I C L E I N F O

Article history:
Received 23 December 2015
Received in revised form 22 March 2016
Accepted 31 March 2016
Available online 6 April 2016

Keywords:
Seabird identification
Natural scene
Grabcut segmentation
Combined features
Voting

A B S T R A C T

This paper proposes an automated seabird segmentation and identification method that applies to seabird images taken in natural scenes with a non-uniform and complex background. A variety of different bird postures appeared in natural scenes present different features from different points of view, even for the same posture. At first, the Grabcut method is introduced to segment seabird unit from a complicated background. Then, global features, namely the colour, shape and texture characteristics, and local features are integrated to describe the birds regarding various postures. Later, a combined recognition model, which is built using the k-Nearest Neighbor, Logistic Boost and Random Forest models by a voting mechanism that is designed for seabird identification. Finally, the efficiency and effectiveness of the proposed method in recognising seabirds were experimentally demonstrated. The experimental results on 900 field samples (6 seabird species, and 150 samples of each species) achieved a recognition accuracy of 88.1%, which indicates that the proposed method is effective for automated seabird identification.

1. Introduction

There are approximately 321 species of seabirds in the entire world (Harrison, 1987). However, recent research has reported that the monitored portion of the global seabird population has declined overall by 69.7% between 1950 and 2010. In fact, this declining trend could reflect the global seabird population trend (Paleczny et al., 2015). In the newly updated IUCN 2015 Red List, over 40 bird species are now classified as having a high risk of extinction, including many wading shorebirds and other iconic species such as the Helmeted Hornbill, Swift Parrot, Atlantic Puffin, and European Turtle-dove (IUCN, 2015).

Seabirds play an important role in the marine ecosystem, and they constitute a food chain together with the other sea creatures. On the other hand, seabird populations are strongly influenced by threats to marine and coastal ecosystems. The status of seabirds is likely to reflect the underlying state of the coastal and oceanic systems. Seabird productivity and population trends are used as good indicators of long-term and large-scale changes in marine ecosystems (Parsons et al., 2008; Piatt et al., 2007). Seabird conservation can not only protect endangered seabirds but also keep the health and balance of marine ecosystems.

Seabird species identification is the basis of seabird productivity and population trends research. Bird biologists and watchers identify seabird species through monitoring incoming information according to their biological knowledge. However, this task is time consuming and relies heavily on the personal experience of the expert. With the rapid development of image processing technology, there is the potential to apply it to the automated identification of bird species with improved effectiveness and efficiency. Automated recognition and identification of insect and airplane are closed related to bird recognition, as insects and airplanes have wings and variable postures. Many recent studies have been devoted to automated detection and recognition of objects, like airplanes (Du et al., 2013; Zhao and Liu, 2011), and insects (Wang et al., 2012; Wen and Guyer, 2012; Wen et al., 2015). As most automated airplane identification methods are based on infrared images or remote sensing images with low resolution, researches have been devoted to enhancing image quality and extracting shape features, such as geometric characteristics, invariant moments and so on. There are some automated insect identification systems available for commercial or academic deployment. However, most of them do not function well when dealing with the field insect images with high complexity. The study of insect pose estimation has been used to promote field insect recognition, however, the images taken from different viewpoints are not considered in their work (Wen et al., 2015).

Many publications have studied the automated recognition of bird species from acoustics (Anderson et al., 1996, Acevedo et al., 2009, Ventura et al., 2015). There are also some literatures on automated visual recognition of birds based on images (Nadimpalli et al., 2006) and videos (Song and Xu, 2010). Some researchers utilised visual features, such as colour (Marini et al., 2013), shape (Liu et al., 2007), and texture (Burghardt et al., 2004), to automatically identify bird species. However, automated bird species recognition is still a challenge due to the

* Corresponding author at: Department of Mechanical and Electrical Engineering, Xiamen University, Xiamen 361005, China.
E-mail addresses: suxiaxu@xmu.edu.cn (S. Xu), zhuqy@xmu.edu.cn (Q. Zhu).

http://dx.doi.org/10.1016/j.ecoinf.2016.03.005
1574-9541/

difficulties of utilising bird images taken in natural scenes: the images are taken with different lighting and a relatively complicated background (e.g., sky, sea, land, and grass); in natural scenes, birds could show different orientations and postures (e.g., flying in the sky, swimming in the water, standing on the land); the same posture could look completely different when the images are taken from different viewpoints. Recently, a fine-grained recognition method has been applied to bird identification (Göring et al., 2014; Pang et al., 2014). However, most fine-grained bird recognition methods are based on the attribute vocabulary, which requires annotating the main parts of the images manually or by user interaction before identification (Wah and Belongie, 2013).

The main objective of this paper is to develop an automated recognition and identification method that can efficiently identify seabird species based on the bird images taken from natural scenes. A Grabcut method is introduced to segment the seabird unit from the complex background. The global features (the shape, texture and colour) and local features (SIFT) are combined, to conquer the difficulty of performing identification that arises from the variety of postures and orientations. Then, seabirds are identified by combined classifiers based on their integrated features. This method does not require the annotation of bird body parts and attribute vocabulary. An outline of the proposed method is given in Fig. 1.

The rest of the paper is organised as follows. Section 2 presents the detailed implementation of image pre-processing and segmentation. Then, a combination of global and local features for feature extraction is introduced in Section 3. In Section 4, three recognition models are chosen as basic models to build combined models for seabird identification. Experimental results are given in Section 5 to demonstrate the feasibility and performance of the proposed method. Finally, a brief conclusion is given in Section 6.

2. Preprocessing and image segmentation

2.1. Preprocessing

In this paper, preprocessing is primarily used to normalise the enhanced images and remove the noise. For automated bird recognition research, two types of sample images can be used. One type is composed of images acquired with stuffed birds in the laboratory, and the other type comprises the field images taken under natural scenes. Our study focuses on field seabird image recognition. The images are collected from the Internet with permission to perform academic research. Because most of the images are taken with auto-focusing, the bird images are clean but have relatively complicated background (e.g., sky, sea, land, and grass). To enhance the image and remove the noise, a Gaussian filtering method is used to smooth the image. As a linear smoothing filter, the Gaussian filter is suitable for removing Gaussian noise.

2.2. Grabcut-based bird image segmentation

Compared with the images taken from stuffed birds in the laboratory environment, the bird images taken in natural scenes are more diverse: they possess more complicated background, such as sky, sea, land and grass; birds can show different orientations and postures; and the posture varies with any change in the open angle of the wings (as shown in Fig.2). Each column represents one type of bird species with different postures and backgrounds.

Even for the same bird posture, the bird images will present differently if the images are taken from different viewpoints. Consider a flying Northern Fulmar as an example (shown in Fig.3), if observe from an oblique rear view, the back, wings and part of belly of Northern Fulmar are in the sight (Fig. 3(a)); if observe from an oblique front view, details of the bird belly can be found (Fig. 3(b)); if observe from a flank side, the whole right wing opens with a clear texture, and most of the left wing is sheltered by the head and nape (Fig. 3(c)); if observe from a front view, the wing resembles a curve, and the texture feature is difficult to distinguish (Fig. 3(d)); if observe from a high angle view, the whole wings, back and tail with a clear texture are in the sight. (Fig. 3(e)).

The segmentation methods based on histograms or edge detection are effective when the background is uniform; however, these methods cannot extract the exact foreground from a complicated background (Lucchese and Mitra, 2001). To extract the bird unit from a complex background, a Grabcut algorithm is introduced for bird segmentation, which tends to obtain a better segmentation result in a complex environment (Rother et al., 2004). The Grabcut algorithm, an improved algorithm based on the GraphCut algorithm, is an iterative segmentation algorithm based on graph theory. Grabcut is widely used in the extraction of a foreground object in a complex environment. Grabcut creates Gaussian Mixture Models (GMMs) for the background and foreground separately and adopts an iterative procedure that alternates between parameter learning and segmentation estimation to solve the min-cut problem until it converges (Li et al., 2015; Rother et al., 2004). The procedure of the Grabcut algorithm is as follows:

(1) *Colour data modelling*

The colour image is described as an array $z=(z_1,z_2,\ldots,z_n)$ in RGB colour space. The segmentation of the image is expressed as an array of opacity values $\alpha=(\alpha_1,\ldots,\alpha_N),\alpha_n\in\{0,1\}$, with 0 for the background and 1for the foreground. Two GMMs, which consisted of $K=5$ components, were created for the foreground and background. We assigned a unique GMM component from either the background or foreground to each pixel and expressed it with a vector $k=\{k_1,\ldots,k_n,\ldots,k_N\},k_n\in\{0,\ldots\}$. An energy function, the minimum of which corresponds to good segmentation, is defined as follows:

$$E(\underline{\alpha},k,\underline{\theta},z)=U(\underline{\alpha},k,\underline{\theta},z)+V(\underline{\alpha},z). \tag{1}$$

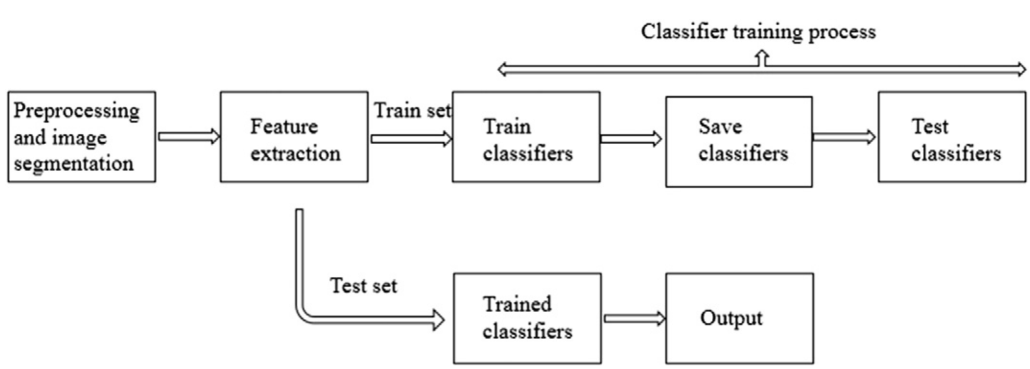

Fig. 1. Outline of the proposed method.