Pattern Recognition 48 (2015) 1059–1073

Contents lists available at ScienceDirect

Pattern Recognition

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

Sloop: A pattern retrieval engine for individual animal identification

James Duyck [a], Chelsea Finn [a], Andy Hutcheon [b], Pablo Vera [c], Joaquin Salas [c,*], Sai Ravela [a]

[a] Earth Signals and Systems Group, Earth Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA
[b] Department of Conservation, Dunedin, New Zealand
[c] CICATA-IPN, Cerro-Blanco 141, Colinas del Cimatario, Queretaro 76090, Mexico

ARTICLE INFO

Article history:
Received 2 January 2014
Received in revised form
7 June 2014
Accepted 19 July 2014
Available online 1 August 2014

Keywords:
Photo-identification
Animal biometrics
Individual identification
Relevance feedback
Crowdsourcing
Conservation
Scale-cascaded alignment
Local features
Hybrid shape contexts
Gecko
Skink
Whale shark
Salamander

ABSTRACT

Identifying individuals in photographs of animals collected over time is a non-invasive approach for ecological monitoring and conservation. This paper describes the design and use of Sloop, the first image retrieval system for individual animal identification incorporating crowd-sourced relevance feedback. Sloop's iterative retrieval strategy using hierarchical and aggregated matching and relevance feedback consistently improves deformation and correspondence-based approaches for individual identification across several species. Its crowdsourcing strategy is successful in utilizing relevance feedback on a large scale. Sloop is in operational use. The user experience and results are presented here to facilitate the creation of a community-based individual identification system for conservation planning.

1. Introduction

The development of effective conservation strategies for rare and endangered species requires unbiased and precise information on their life history and population ecology [1]. Capture-Mark-Recapture (CMR) studies enable researchers to track individual animals over time to answer questions related to individual growth, survival, dispersal and reproductive strategies. CMR studies typically use techniques in which animals are physically marked or tagged. These methods are intrusive to varying degrees and inefficient to implement in large numbers. Alternative identification techniques that overcome these limitations are needed.

Numerous efforts exist to identify individual animals using photographs. The simplest one to conceive is manual identification, but its high recall comes at exceptional cost. Manual searches are only feasible for small collections; at 10 s per comparison a 10,000-sized catalog will take approximately 15 person-years to analyze. Manual searches typically also employ *ad hoc* strategies using individual markings. This process is difficult to automate and extend and may lead to imprecise quantitative analysis.

Computer-based pattern recognition approaches [1–10] decidedly benefit the identification problem. They scale to large collections, extend to multiple species and are convenient to use. However, despite the advantages of computational speed and advances in automatic pattern recognition, it is not feasible to automatically deliver the high recall needed for correctly tagging many species. Some degree of human involvement arguably benefits the identification process. A problem that emerges is how to determine the optimal level of human involvement. The earliest approach used deformable template matching [2,3] that could be used for multiple species. However, it demands user inputs for every compared pair of photographs; clearly too much.

Motivated by an image retrieval approach to face recognition [11], Ravela and Gamble [8] advanced the interactive retrieval paradigm for animal biometrics. Their retrieval approach characteristically begins with an unlabeled collection of photographs that are segmented, rectified and illumination corrected. Initially automated, these steps were later performed in a semi-automated

---

[*] Corresponding author. Tel.: +52 55 5729 6000x81015.
E-mail addresses: jsalasr@ipn.mx, salas@ieee.org (J. Salas).

http://dx.doi.org/10.1016/j.patcog.2014.07.017
0031-3203/© 2014 Elsevier Ltd. All rights reserved.

fashion to limit errors from propagating to subsequent stages. Typically, generic appearance or geometry features are extracted from regions of interest. The spatial representations of features including vectors, histograms or graphs are constructed and compared, usually using Bayesian inference, to rank images.

The key step then follows. The user verifies a limited number of top ranked retrievals and thus makes the final decision for labeling cohorts[1] with shared identities [8,1,9,10]. Once an initial collection is indexed, over time, this linear approach sequentially assimilates new images by comparing them with existing cohorts or singletons to maintain updated capture histories. The system becomes an interactive search engine for identities because the features are generic, appropriate representations of images are precomputed, and comparisons are fast. By ranking images and presenting top matches, the retrieval approach only requires the user to view pairs that are most likely to be matches. This precision enables human recall that is much more efficient than manual matching. With this approach, a 10,000-size collection may typically require about three person-months to analyze; this is the state of the art.

A retrieval approach contrasts with an automated recognition approach for identifying an animal photograph. The latter is preferred when there are fixed identities with labeled exemplars and, most importantly, when indexing is not the objective. This is not typically true for small animal species where new individuals frequently enter the database and demand persistent indexing and high recall. In these cases, full automation is not yet feasible and will decidedly benefit from a retrieval approach.

Sloop, a distributed interactive system for individual animal identification, emerged through several iterations of the image retrieval methodology [8,1,9,10]. Sloop is neither a complete CMR system nor does it provide statistical analysis, but it is a tool with which users can reconstruct capture or encounter histories. It predominantly uses visual features but ongoing work includes additional textual metadata. Sloop utilizes multiple techniques plus human input, and combines them to deliver high performance. It is deployed operationally and finds mention in popular press [12].

In the community, there are now several retrieval-based systems that, as with Sloop [8,1,9], propose using generic visual (largely local) features with attendant claims of extensibility. Clearly, the freedom that generic approaches provide from *ad hoc* individual markings with linear indexing cost is beneficial and delivers reasonably good performance. However, there are at least two limitations.

First, just as real-world face recognition technologies are highly specialized to satisfy recall demands, high-recall animal biometrics also demands trading invariance of the generic algorithms for selectivity by application. As this paper shows, while generic one-method-for-all approaches are abundant and perform comparably, rarely does any deliver intrinsically high recall at least in the several species considered here. As a consequence, not only are refinements towards higher recall of interest, reusable techniques for tuning multiple arbitrary generic methods, including combining (aggregating) them to deliver a performance better than any one source alone, are also relevant.

Second, as collections grow, even linear human effort is substantial for small teams to undertake. Reductions in mouse-click counts and keystrokes are important but methodology to recycle already-performed work to reduce subsequent human effort is also of great interest. For example, as matches start to form, it should be possible to feedback this relevance information to improve subsequent matching. Neither of these issues receives much attention in the Animal Biometrics community which they must as collections transition to Big Data regimes with richer images, feature varieties and collection sizes.

In this paper, the Sloop system architecture, its workflows, and algorithms are described with application to several species including the jewelled gecko (*Naultinus gemmeus*), grand and Otago skink (*Oligosoma grande* and *Oligosoma otagense*), whale shark (*Rhincodon typus*) and marbled salamander (*Ambystoma opacum*). In each of these cases, *baseline* algorithms using deformable models, patch-based appearance models, local features, and/or shape context models are shown to deliver reasonable performance. Then, new algorithms that uniformly outperform baseline algorithms are presented. This includes hierarchical ranking and retrieval (ordered combination of algorithm results), rank and score aggregation (parallel combination of algorithm results), and relevance feedback including the use of crowdsourcing for scalability. Our experiments suggest that the advantages of these steps compound. For example, indexing a 10,000-size collection can, in principle, be accelerated and completed in a few person-days.

Thus, the central advance of this paper is that whilst fully automated recognition systems are not yet within reach for Animal Biometrics, coupled human-machine systems that deliver high performance are achievable. In these approaches, the algorithms can reduce human effort and the human feedback can improve system performance. Together, they can produce extensible, scalable, and effective large-scale deployments. Within the realm of Animal Biometrics, Sloop appears to be the first such operational system. The result of deploying Sloop on the Grand and Otago Skink Recovery Programme, in Dunedin New Zealand, and the first full year of its use is described. The results are extremely encouraging and may serve as a useful model to integrate biologists, computational vision researchers and citizen scientists in a unified framework.

The remainder of this document is organized as follows. In Section 2, closely related systems are reviewed and compared to Sloop. In Section 3, the Sloop system architecture, workflow and methods are discussed. The application of the methods to individual species is presented in Section 4, operational experience in the Grand and Otago Skink Recovery Programme is described in Section 5, and the paper concludes with a discussion in Section 6.

2. Related work

Identifying individuals among a species population is of increasing interest. Some of the earliest approaches [2,3,13] use 3D deformable matching, which is extensible, but more frequently used are new techniques driven by the need for rapid large-scale matching. Sloop contains a 2D deformation invariant matching algorithm [9], which is only used to improve existing rankings. It is relatively fast and demonstrated to be highly effective on marbled salamanders. Here we show additional improvements by relevance feedback which also facilitates comparisons where one or both sides may be imaged [14].

Within the realm of feature-based methods, there are two main classes of recognition methods; specialized methods [6,15,16] for individual species, and generic methods [8,1,9,17–19,10]. The specialized techniques are of limited interest. For example, a generic correspondence-based approach in Sloop [10] can be adapted for whale sharks with markedly improved performance over an earlier specialized approach [6].

Within the realm of generic methods, Ravela and Gamble [8], motivated by face recognition [11], proposed using multi-scale differential feature histograms, and later using randomized multiscale-PCA [1] on marbled salamanders. SIFT features [20] are popular generic features. For example, Yang and Ravela [9]

---

[1] A cohort is a set of images with the same identity; a singleton is a cohort of size one.