Vision Based Elephant Recognition for Management and Conservation

Ranga Dabarera*, and Ranga Rodrigo†
Department of Electronic and Telecommunication Engineering,
University of Moratuwa, Moratuwa 10400, Sri Lanka
* Email: dabarera@ent.mrt.ac.lk
† Email:ranga@ent.mrt.ac.lk

***Abstract*—In elephant management and conservation, it is vital to have non-invasive methods to track elephants. Image based recognition is a non-invasive mechanism for elephant tracking, albeit the inefficiency in manual method due to difficulties in handling large amount of data from multiple sources. To mitigate the drawbacks in manual method, we have proposed a computer vision based, automated, elephant recognition mechanism, which mainly relies on appearance based recognition algorithms. We have tested feasibility of the system running on a web based interface, which can facilitate researchers and conservationists all around the world to actively participate in elephant conservation.**

***Keywords*—elephant recognition, non-invasive, appearance based**

I. Introduction

Elephant conservation requires the observation and tracking of individual elephants. Some of the most widely used methods are invasive and probable of causing harm to the elephants. One such method is tracking the individual elephant by having a GPS collar, which needs the elephant to be tranquilized. In this invasive tranquilization process immobilizers like Etorphine ($M99$) are injected to the elephants and apart from the panic and pain, some elephant end up dead in this process, which is not the ultimate goal of elephant conservation. It is claimed that around 2% of elephant deaths occur due to this type of invasive methods, which urges for a non-invasive solutions.

Image based recognition and tracking of elephants is a non-invasive method. Nevertheless, it has not been utilized efficiently due to the drawbacks of existed manual method, even though it is a less expensive and non-invasive solution. In manual image based elephant identification the quality of results are mainly dependent on the person or the small research group involving in that identification task. More precisely in manual identification, person or set of researches have to remember the features of each and every elephant in a particular area. Another disadvantage of this manual mechanism is, its limitation in getting the active participation of all the volunteers and fellow researchers around the globe. Alternative mechanisms are required to alleviate those deficiencies in manual image based elephant identification to get the advantages of non-invasive methods.

In this paper we apply computer vision algorithms to automate the process and develop a mechanism to identify individual elephants using images. In the next section, some efforts of using automated methods for conservation of animals, including elephants, is discussed. To the best of our knowledge, we are the first to use computer vision algorithms for the appearance based recognition of individual elephants using images. Given the frontal face image of an elephant, the proposed method search to identify the individual elephant using vision algorithms and gives the result as, the details of the already identified elephant, or as a new identification. We have tested the initial prototype with a web based interface which gave recognition accuracy around 72%. The accuracy can be escalated further by incorporating other features unique to elephants, which has been discussed under future works. Once developed, the system with acceptable accuracy will facilitate the elephant conservation process a collaborated truly global effort. Researchers and people around the world who are interested in elephant conservation can actively participate in this elephant recognition and habitat monitoring system via online through a web portal.

II. Background and Related Work

Computer aided photograph matching in studies using individual identification has been proposed and tested on Serengeti Cheetahs [1]. In 2007 Goswami,*et al.* [2] came up with an application of Photographic Capture-Recapture (CR) modelling to estimate demographic parameters for male Asian elephants. They have developed methods to individually identify elephants and integrated them into CR sampling designs to get capture histories and from them estimate the abundance of male elephants in southern India. The aforementioned scenarios give evidence for the significance of image based individual identification methods for habitat monitoring and information gathering. Some attempts have been proposed to automate the image based elephant identification using the edge of the elephant ear and facial features [3]. They propose to analyse ears of the elephants using fixed points and contours, where a similar approach of contour analysing has been proposed for Dolphins from Fin Profiles [4] in digital images, which has been enhanced and implemented with the additional supportive features [5]. Most facial recognition techniques proposed for elephant recognition are based on distinctive image features such as relative positions of eyes, top and tip of the tusk and horizontal trunk wrinkles in the face input by the user.

This requires the familiarity with distinctive features of the elephant, which is not robust to matcher inexperience. There are exceptions, for the set of rules that we define, in order to recognize elephants. More precisely, there can be more than one elephant, having the same relative locations of the selected distinctive image features. This leaves the problems in image based elephant identification unsolved.

To solve the identification problems, more recent approaches rely on texture and appearance based techniques [6]. Spot the Penguin Project [7] uses modern vision techniques for Penguin observation as a truly non-intrusive behavioural and conservational analysis of population dynamics [7]. Our proposed system carries out individual elephant recognition utilizing initial input images to the system and direct user inputs, to develop a constantly updated habitat monitoring mechanism. Along with the image, the system will keep records of observed location and date, of each individual elephant. As the system evolves, the logged data will help to extract demographic and behavioural information on elephants, which are crucial for elephant habitat protection and to mitigate human elephant conflicts, and by merging with web portal, which will form a user community for people interested in elephant conservation. Our initial method of automated elephant identification stands on the texture based technique of eigenfaces for recognition. The enhanced version will be based on active appearance models.

III. Methodology

Our elephant management and conservation tool consists of two main parts as, computer vision based core for the recognition of elephants, and a web portal as the interface to the user. In the following sections we mainly focus on the vision based elephant recognition algorithm.

When an image of an elephant is uploaded into the system, an image based algorithm searches for and matches the image with an existing annotated elephant database. If it is a previously identified individual, then the system will match the image with information available for that particular elephant. In addition, the database gets updated with the last observed location and the date. If the elephant is previously unidentified, then the system issues a new identity. Utilizing the web portal, researchers and the general public from all around the world will be able to contribute to this process by uploading frontal images of elephants that they have observed. Along with the images, through the web portal, users can give additional information about the elephant, such as relative tail length, tail tuft information, and details of primary and secondary ear folds. This additional information of an elephant can be used with an evidence combining mechanism to recognize individual elephants. Web portal will also include secondary data such as socio-economic and Human Elephant Conflict (HEC) information, that is tagged to geographic locations, so that users can analyse trends and use the system for research and development needs.

A. The image based algorithm

Our initial elephant recognizer was focused on identifying elephants given the full frontal images of faces under relatively uniform illumination conditions. Even thought the traditional manual image based identification uses, locations of distinctive image features such as the eyes, tusks and ears, and measuring the distance between those feature locations; the automated image based methods should not solely replicate the manual counterpart. Automated methods can take the advantage of enhanced computational power and memory using the unique capabilities of computer and vision algorithms for elephant identification. Due to that, our initial implementation is based on Eigenfaces, an observation first made by Kirby and Sirovich [8] for image compression, and popularised by Turk and Pentland [9] for face recognition. An arbitrary face image x can be compressed and reconstructed by starting with a mean image[1] m and adding a small number of scaled signed images u_i;

$$\tilde{x} = m + \sum_{i=1}^{M} a_i u_i \tag{1}$$

Where $\tilde{x}$ is the compressed image of face image x, a_i are coefficients which are unique to a particular image x and M controls the quality of compressed image by deciding the number of coefficients (which is less than or equal to the number of images used for training).Turk and Pentland [9] recognized that the coefficients a_i could themselves be used to construct a fast matching algorithm applying eigenvalue decomposition.

Let the training set of elephant face images be $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_N$ where N is the number of training images. The training images are obtained by reshaping original images as $\Gamma_i = [\ \gamma_{i1},\ \gamma_{i2},\ \ldots,\ \gamma_{i\xi}\ ]$ where γ_{ij} is a pixel and ξ is the dimension of the image, which is equivalent to the number of pixels of that image. The average face can be obtained by $\Psi = \frac{1}{N}\sum_{n=1}^{N}\Gamma_n$ and let the result be $\Psi = [\ \psi_1,\ \psi_2,\ \ldots,\ \psi_\xi\ ]$. To represent the covariance matrix as $C = AA^T$ matrix A is defined as follows;

$$A_{\xi,N} = \begin{pmatrix} \gamma_{11} - \psi_1 & \gamma_{21} - \psi_1 & \cdots & \gamma_{N1} - \psi_1 \\ \gamma_{12} - \psi_2 & \gamma_{22} - \psi_2 & \cdots & \gamma_{N2} - \psi_2 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1\xi} - \psi_\xi & \gamma_{2\xi} - \psi_\xi & \cdots & \gamma_{N\xi} - \psi_\xi \end{pmatrix}$$

Applying eigenvalue decomposition on C

$$\begin{aligned} C &= AA^T \\ &= \sum_{j=1}^{N} \lambda_i u_i u_i^T \end{aligned} \tag{2}$$

Where λ_i are the eigenvalues of C and u_i are the eigenvectors of C. However, as the size of C is $\xi \times \xi$,

[1]Mean of the all training images used to generate the scatter matrix C

finding eigenvalues of C costs more computational power. Due to that fact, in implementation the eigenvalue decomposition is done on scrambled covariance matrix[2] defined as $L = A^TA$ to find eigenvalues and eigenvectors, a fact which has been clearly justified in [9]. As these eigenvectors are corresponding to images of faces, they are commonly referred as eigenfaces. These eigenfaces are ranked according to the order obtained after sorting the corresponding eigenvalues in descending order. Higher the magnitude of the eigenvalue, higher the significance of the corresponding eigenface in image representation. But unlike reconstructing the original face image as in image representation, here our idea is only to identify individual elephants from the frontal face images. For that we can truncate the number of eigenvectors by selecting only the eigenvectors corresponding to higher eigenvalues. We can decide the threshold of higher eigenvalue considering the number of elephant face images we want to analyse and the precision required. By neglecting the additional terms at the latter levels of sorted eigenvalues, it will help the system to perform faster while preserving significant precision.

B. Basic identification process

Before starting the identification process, the system required to be given a known set of elephant face images to build the scatter matrix C of the face images and apply eigenvalue decomposition. After generating the eigenvectors, the corresponding coefficients of each and every existing elephant face image[3] can be calculated by projecting them on to that face space (a hyper-dimensional space where all the elephant face images lie, hyper-coordinate system corresponding to eigenfaces) and the corresponding coefficients a_k are calculated using Eq. (3).

$$a_i = (x - m).u_i \tag{3}$$

When a new image is given, that image is then projected on to the same face space which training images were projected. Corresponding coefficients for the given new image, a can be obtained using Eq. (3). The distances from previously given elephant images to the newly given image can be calculated from the collected coefficients of a_k and a. If the distance of the new image is very high from the elephant face space then we can verify that this image does not properly represent elephant face image[4]. If the distance of the new image is below a predefined threshold, when compared with an existing known elephant image we can make the initial matching which can be further verify using additional user inputs, for instance, observed area, location or distinctive features unique to that particular elephant.

[2]The size of the scrambled covariance matrix, L is $N \times N$ which is much less than the size of C ($\xi \times \xi$). In our implementation the size of L is 38416 whereas the size of C is approximately 1.9×10^{10}

[3]This can be considered as the training phase

[4]This condition hold properly if the number of known elephant images are high

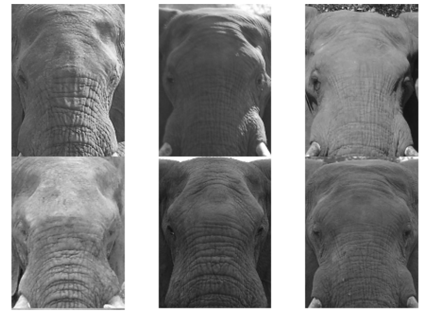

Fig. 1. Warped images obtained after perspective transformations

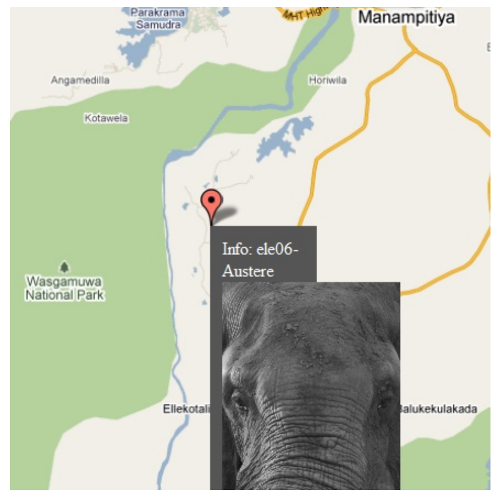

Fig. 2. Identified elephant with location information

IV. Results

Our eigenface based identification mechanism for elephant identification was tested on frontal face images obtained from a database of African elephants *(Loxodonta Africana)* which contains categorized images of 192 African elephants (with couple of images per each elephant). The accuracy of identification using the frontal images, after a perspective transformation stage to warp the face image, was around 72%. This was a significant improvement from the previous implementation where we used direct raw images, which gave an accuracy rate around 50 - 53 %. Figure 1 shows some of the warped images. For warping we do affine transformation taking, the location of the eyes of the elephant in the image, as the initial parameters. The idea behind warping is, taking all the images into a uniform reference level so that the matching becomes easy. Figure 2 shows the trial web based interface designed to integrate the computer vision based elephant identifier and to keep track of individual elephants.

The proposed method has been implemented in C++ with the Intel OpenCV (Open Computer Vision) library. The overall average computation time for training 196 images (with 430×320 resolution) on a PC with a 2.10 GHz Core 2 Duo CPU and $4GB$ memory is $8.895s$. For testing an image with the same resolution, the typical computation time is around $1ms$, without the time for selecting the image and displaying the results.

V. Conclusion

We have proposed a method for non-invasive elephant identification mechanism using computer vision algorithms, for management and conservation of elephants. Proposed method has achieved an accuracy around 72%, only using the frontal face images of elephants. The accuracy can be affected by the low quality of the images, the strange poses of the head positions of the elephant and high illumination variances. The current system can be useful as a semi-automated mechanism for elephant tracking, and can be converted to a fully automated system by further increasing the accuracy.

VI. Future work

To make the elephant recognition system a fully automated system, the recognition algorithm has to be further improved using Active Appearance Models (AAM) [10]. With AAM we can incorporate the shape of the elephant face along with the appearance data as it utilizes a statistical model of the shape and grey-level appearance of the face. Using the location information, the searching domain of the database can be narrowed down, which assists the vision algorithm to give less false positives. Also the system can be assisted with other distinctive features of elephants like relative tail length, ear shape, unique marks etc... and issue the final result of recognized elephant using an evidence combining mechanism.

Acknowledgment

The authors would like to thank Prof. Dileeka Dias and Dr. Ajith Pasqual at the Department of Electronic and Telecommunication Engineering, University of Moratuwa, for their guidance and coordination. The first author would like to thank Sri Lanka Wildlife Conservation Society for giving the financial aid under the Saving Elephants by Helping People Project, partially-funded by Tides Foundation through Google Geo Challenge Grant.

References

[1] M. J. Kelly, "Computer-aided photograph matching in studies using individual identification: An example from serengeti cheetahs," *Journal of Mammalogy*, vol. 2, pp. 440–449, May 2001.

[2] V. R. Goswami, M. D. Madhusudan, and K. U. Karanth, "Application of photographic capture-recapture modelling to estimate demographic parameters for male asian elephants," *Animal Conservation*, pp. 391–399, May 2007.

[3] W. Kreiken and W. van Vliet, "Elephind,amboseli elephant research project (aerp) and the tarangire elephant project (tep)," University of Leiden Institute for Environmental Sciences. Leiden, Tech. Rep., 2005.

[4] J. Stewman, R. Stanley, and M. Allen, "A system to identify dolphins from fin profiles in digital images," *Proceedings of the 8th Annual Florida Artificial Intelligence Research Symposium*, 1995.

[5] S. Hale, "Unsupervised threshold for automatic extraction of dolphin dorsal fin outlines from digital photographs in darwin (digital analysis and recognition of whale images on a network)," Master's thesis, Eckerd College, 2008.

[6] R. Szeliski, *Computer Vision: Algorithms and Applications*, August 2010, vol. draft version, ch. Recognition, pp. 650 – 652.

[7] T. Burghardt and N. Campbell, "Individual animal identification using visual biometrics on deformable coat-patterns," in *International Conference on Computer Vision Systems*, Bielefeld University Applied Computer Science Group. http://biecoll.ub.uni-bielefeld.de: Applied Computer Science Group, Bielefeld University, Germany, 2007.

[8] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990.

[9] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[10] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.