Title: Automated face detection for occurrence and occupancy estimation in chimpanzees

Authors: Anne-Sophie Crunchant[1], Monika Egerer[1], Alexander Loos[2], Tilo Burghardt[3], Klaus Zuberbühler[4, 5, 6], Katherine Corogenes[1], Vera Leinert[1], Lars Kulik[1], Hjalmar S. Kühl[1,7]

1. Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.
2. Fraunhofer Institute for Digital Media Technology IDMT, Ehrenbergstrasse 31, 98693 Ilmenau, Germany
3. Department of Computer Science, University of Bristol, Woodland Road, BS8 1UB Bristol, United Kingdom
4. Department of Comparative Cognition, University of Neuchâtel, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland.
5. School of Psychology and Neuroscience, University of St Andrews, St Andrews KY16 9JP Scotland.
6. Budongo Conservation Field Station, P.O. Box 362, Masindi, Uganda.
7. German Centre for Integrative Biodiversity Research (iDiv) Halle-Leipzig-Jena, Deutscher Platz 6, 04103 Leipzig, Germany.

Short title: Chimpanzee automated face detection

Editorial Correspondence: Anne-Sophie Crunchant, E-mail: as.crunchant@gmail.com

ABSTRACT

Surveying endangered species is necessary to evaluate conservation effectiveness. Camera trapping and biometric computer vision are recent technological advances. They have impacted on the methods applicable to field surveys and these methods have gained significant momentum over the last decade. Yet, most researchers inspect footage manually and few studies have used automated semantic processing of video trap data from the field. The particular aim of this study is to evaluate methods that incorporate automated face detection technology as an aid to estimate site use of two chimpanzee communities based on camera trapping. As a comparative baseline we employ traditional manual inspection of footage. Our analysis focuses specifically on the basic parameter of occurrence where we assess the performance and practical value of chimpanzee face detection software. We found that the semi-automated data processing required only 2-4% of the time compared to the purely manual analysis. This is a non-negligible increase in efficiency that is critical when assessing the feasibility of camera trap occupancy surveys. Our evaluations suggest that our methodology estimates the proportion of sites used relatively reliably. Chimpanzees are mostly detected when they are present and when videos are filmed in high resolution: the highest recall rate was 77%, for a false alarm rate of 2.8% for videos containing only chimpanzee frontal face views. Certainly our study is only a first step for transferring face detection software from the lab into field application. Our results are promising and indicate that the current limitation of detecting chimpanzees in camera trap footage due to lack of suitable face views can be easily overcome on the level of field data collection, i.e. by the combined placement of multiple high resolution cameras facing

reverse directions. This will enable to routinely conduct chimpanzee occupancy surveys based on camera trapping and semi-automated processing of footage.

Keywords: apes; animal biometrics; camera placement; site use; automated image recognition

Research Highlights

Using semi-automated ape face detection technology for processing camera trap footage requires only 2-4% of the time compared to manual analysis and allows to estimate site use by chimpanzees relatively reliably.

INTRODUCTION

Motivation. Biodiversity has declined and continues to decline around the world. This is true of great ape populations, which have dramatically decreased in numbers and distribution over the past three decades [Walsh et al., 2003; Campbell et al., 2008; Greengrass, 2009; Junker et al., 2012; Funwi-Gabga et al., 2014]. In light of multiple drivers of decline (habitat loss [Gates, 1996; Wich et al., 2008; Junker et al., 2012, Wich et al., 2014], hunting [Gates, 1996; Walsh et al., 2003; Kuehl et al., 2009], and infectious diseases [Woodford et al., 2002; Leendertz et al., 2004; Bermejo et al., 2006; Leendertz et al., 2006; Köndgen et al., 2008]), we face the arduous task of conserving and restoring ape populations above critical levels and to secure them as a global community. To do this, it is first necessary to estimate distribution and population sizes accurately in order to allocate conservation efforts to where they are most needed [Kormos & Boesch, 2003; Oates et al., 2007; Plumptre et al., 2010; Morgan et al., 2011; Carlsen et al., 2012 ; IUCN & ICCN, 2012; Maldonado et al., 2012; Dunn et al., 2014; Tweh et al., 2014]. Distribution and density estimates of individuals allow inference on changes in population size. With this information, conservationists can establish and prioritize protected areas and will have a baseline estimate for assessing the effectiveness of their efforts over time [Kormos & Boesch, 2003; Nichols & Williams, 2006; Plumptre & Cox, 2006].

General Approach. To obtain population estimates, monitoring needs to be regular and over a wide range of areas that are inhabited by a species. Long-term monitoring is also important to address various ecological questions, such as the determination of habitat use, resource use, community dynamics and community relationships. Yet, with elusive species, such as apes, direct observations are difficult to obtain without massive habituation efforts, which generates a need for reliable indirect monitoring methods [Kuehl et al., 2008; Head et al., 2013]. An array of indirect monitoring techniques have thus been developed and employed, including line transect nest and dung counts, camera trapping and non-invasive genetic sampling [Plumptre & Reynolds, 1996; Kuehl et al., 2007; Kuehl et al., 2008; Todd et al., 2008; Guschanski et al., 2009; Kouakou et al., 2009; Buckland et al., 2010; Head et al., 2013]. Distribution and abundance can then be inferred using design-based inference, spatial modeling techniques or capture-recapture methods [Buckland et al., 2001; Borchers et al., 2002; Arandjelovic et al., 2010; Head et al., 2013; Murai et al., 2013; Tweh et al., 2014].

Problem Statement. However, while these methods are very useful for conservation research, some of them can nevertheless be labor, time and cost intensive, for they require trained staff, adequate equipment, and regular repetition [Gaston & O'Neill, 2004]. Furthermore, some monitoring methods are vulnerable to human observer biases [Tuyttens et al., 2014]. One exception is camera trapping that is less dependent on human observer skills in the field. However, camera trapping also requires correct identification of individuals to e.g. estimate occupancy or population size [O'Connell et al., 2010] and is ideally only used on demographically closed populations with minimal growth rates and migration [Borchers & Efford, 2008; Head et al., 2013]. Although advantageous to non-invasively observe elusive species and amass large amounts of data [Noss et al., 2012], the technique is, when used conventionally, also labor and time intensive, requiring skilled observers to process the video data.

Animal Biometrics. In response to this problem, animal biometrics has made progress in developing computerized methods for automated detection and individual identification [Gaston & O'Neill, 2004; Kühl & Burghardt, 2013]. Kühl and Burghardt [2013] defined animal biometrics as the utilization of phenotypic characteristics that can identify species and in some scenarios even individuals, by exploiting body morphologies, coat patterns and general appearance, vocalizations or behaviors. Based on phenotypic observations and distinct animal characteristics, biometric software has helped to identify individual elephants from ear nicks [Ardovini et al., 2008], dolphins from dorsal fin shapes [Araabi et al., 2000], zebras from stripe patterns [Lahiri et al., 2011], great white sharks from dorsal fin shape [Hughes & Burghardt, 2015], and great apes from facial characteristics [Ernst & Küblbeck, 2011; Loos & Ernst, 2012; 2013].

Performance Estimation. Assuming perfect ground truth labeling, the performance of automated detection systems can be specified according to a binary classification task. For the task of animal detection, for instance, detections can be categorized into one of four classes: true positives (TP, a manually observed animal is also detected by the software); true negatives (TN, no animal is manually observed nor detected by the software); false negatives (FN, an animal is manually observed, but not detected by the software); false positives (FP, no animal is observed manually but software generates a detection). The performance of the overall detection software can then be characterized by these values. However, performance statistics could also be reported by a combination of recall and false alarm rates; where recall is the proportion of true detections by the software in relation to the total number of detectable events (TP/(TP+FN)) and false alarm rate is the proportion of false detections (FP/(FP+TN)) [Macmillan & Creelman, 2004].

Novelty of Study using Face Detection. Face detection software, as a particular class of animal biometric detection technology, is particularly promising for population assessment, analysis and conservation of great apes with potential for addressing further parameters, as well as population and community ecology questions [Kühl & Burghardt, 2013]. To date, face detection software for animals has been successfully tested under controlled conditions, or was tested based on high-quality image and video datasets which were not gathered by using remote camera devices as in our study [Loos & Ernst, 2012; 2013]. To our knowledge, no studies have successfully used face detection software under completely unconstrained field conditions, and we are not aware of any studies that have directly compared the results of both manual and face detection analyses of camera trap data from the field.

Aims of Study. In this study we evaluate the applicability of previously developed chimpanzee face detection software [Ernst & Küblbeck, 2011] to process field camera trap data. Our primary aim is to assess whether using the software can improve efficiency of the time consuming processing of camera trap footage. More specifically, we are interested in quantifying the amount of time field biologists may save and the expected accuracy of key parameter estimates when using the software compared to purely manual processing. It is not the goal of this study to assess the *performance of the software as an object recognition framework*, this has been already done for high-quality visual footage and the interested reader is referred to [Ernst & Küblbeck, 2011] for a detailed evaluation. Here we focus on quantifying the software's effectiveness for the task of estimating site-specific occurrences of chimpanzees (site occupancy) based on in-frame animal presence/absence [MacKenzie et al., 2002; MacKenzie et al., 2006; Andresen et al., 2014]. We note that this task is fundamentally different compared to evaluating object recognition performance, since neither accurate spatiotemporal localization nor scale information - critical parameters in traditional *performance estimates for object recognition* - retain their importance when focusing on presence/absence information over large time windows only.

Our overall target parameter is site occupancy, i.e. we want to estimate the proportion of an area that is occupied or used by a species during a season [MacKenzie et al., 2002]. This measure is useful in long-term monitoring programs because it can provide data to assess population changes, site colonization and extinction, site use, as well as give insight into multi-species interactions and other ecological parameters [MacKenzie et al., 2002; MacKenzie et al., 2003].

Summary of Objectives. In summary, our objectives are (1) to estimate the performance and efficiency gain when using the face detection software to recognize chimpanzee presence and absence under field conditions, and (2) to estimate site use by two chimpanzee communities from this data. We compare the results of manual processing of camera trap footage with various degrees of automated processing. Though we have chosen to conduct our study on a small scale to test the face detection approach, this approach and software is fit for use at a larger scale where it has the potential to have the greatest benefit and impact of analyzing field data.

DESCRIPTION

Analytical methods

Manual Video Processing. All camera trap videos were first manually screened for the presence of chimpanzees. Detections were also categorized into quality levels of the underlying images (light conditions, chimpanzee distance from camera, visibility time, and face and head positions; Fig. 1). The metadata was recorded together with date, time and GPS location of the capture.

Face Detection System. We used the face detection framework SHORE™ (Sophisticated High-Speed Object Recognition Engine) [Ernst & Küblbeck, 2011; Loos, 2016] developed by the Fraunhofer Institute for Integrated Circuits (IIS) trained to detect chimpanzees (Fig. 2). A software license can be requested from (www.iis.fraunhofer.de). SHORE™ attempts real-time detection and tracking of frontal primate faces in images and videos. Whilst a detailed algorithmic description is published in [Küblbeck & Ernst, 2006; Ernst & Küblbeck, 2011; Loos & Ernst, 2013], here we present a high-level summary of its workings to provide the basic technical context in which the study operates.

General Detection System. SHORE™ [Ernst & Küblbeck, 2011] builds on the key concepts of the well-established object detection framework by Viola and Jones [2001]. SHORE™ utilizes a detection model comprising multiple consecutive classification stages, through which image regions are passing with increasing complexity along an attentional cascade [Viola & Jones, 2001]. In SHORE™, each stage comprises a feature extraction step and a look-up table based classification step, where the classifier is built offline using Real-AdaBoost [Schapire & Singer, 1999]. Real-time capability is achieved by using simple and fast pixel-based features in early stages for a fast and coarse candidate search. Later stages implement slower, but more accurate classifications.

Visual Features. Each stage utilizes one out of three illumination-invariant features: *edge orientation features*, *census features*, or *structure features*. Edge orientation features represent pixel-based gradient directions and are extracted via Sobel operators. In subsequent classification stages more complex census features [Zabih & Woodfill, 1994] are extracted, which encode local brightness changes. In the final classification stages, structure features, which are built out of scaled versions of census features, are extracted on image regions.

System Training. Positive training data, i.e. great ape faces, were used applying slight random variations such as rotation, mirroring, and translation to increase robustness of the classifier to be built. Non-face negative training data was generated by randomly cropping patches from images without great ape faces. Subsequently, further non-face data was gathered by bootstrapping the initial model on images without ape faces.

Face Detection. During detection, the gray scaled input image is initially convolved with a 3x3 mean filter kernel to compensate noise. While the detection model is fixed with a size of 24x24 pixels, the mean filtered image is downscaled multiple times using a scaling factor of 1.24 to build an image pyramid. A real-time capable, coarse to fine search is applied by shifting the detection window across every pyramid level to achieve scale invariance. Detections in multiple pyramid levels are subsequently merged to a single detection with mean size and location by applying non-maxima suppression.

Slicing and Face Tracking. As stated earlier, SHORE™ is not only capable of detecting faces in single frames, but also to track them through a scene. Once a face has been detected, a unique identifier is assigned to it. During consecutive frames, the tracking algorithm then tries to maintain the association between ID and face. The subsequent paragraph briefly reviews the tracking algorithm used within SHORE™. For a more detailed explanation the interested reader is referred to [Küblbeck & Ernst, 2006]. As described, the static detector repeatedly searches for faces in all levels of an image pyramid in order to find faces of different sizes. Assuming scale consistency of faces, it is sufficient to scan pyramid levels only a few times per second. Therefore, the image pyramid is partitioned into slices which are processed alternatingly. In practical applications Küblbeck and Ernst [2006] observed a performance improvement by a factor of two to three, depending on the number of faces in the scene. A motion model is then applied to connect the detections of subsequent frames. A linear Kalman filter [Kalman, 1960; Welch & Bishop, 2006] is applied in order to estimate the current state of a tracked face from the detection results. Additionally, the first and second order derivatives are included in the state vector to represent the velocity and the acceleration of a face. Association of object-ID and detected face in consecutive frames is done by using a minimum distance criterion: A detected face in the current frame is associated with the face detected in the previous frame which is closest to the current object position. It was shown in [Küblbeck & Ernst, 2006] that based on the observations of past frames it can be decided if a tracked object actually represents a valid face, which significantly reduces the number of false positive detections while the detection rate is maintained.

Application of Software. We used the face detection software SHORE™ to extract chimpanzee occurrence from all video footage via R (version 3.0.2; R Development Core Team, 2013; https://www.r-project.org) The software was carefully trained by computer vision experts and the detection score was selected based on evaluation on an entirely different dataset. We included videos that did not contain chimpanzees in the analysis. We did not modify the software provided by the Fraunhofer Institute and recognize their contribution to our methodology. The software provides detections of primate faces contained in images and videos. Note that the software *only* detects chimpanzee faces and not whole bodies, its ability to detect chimps in videos is limited to videos where face views are visible. The software then produces a script of codes and coordinates as output for each respective visual image processed. This contained the species detected (chimpanzee or gorilla) and the age class (infant, juvenile, adult) for each individual. Additionally, for each frame where an individual was detected, the output gave the probability of species and the most probable species, the probability of each age class and the most probable age class, as well as positions of the face, eyes and mouth.

Setups and Post-processing. Automated processing can lead to misclassifications, whose impact can bias estimates for species occurrence and site occupancy estimates [MacKenzie et al., 2003; MacKenzie & Royle, 2005; Andresen et al., 2014]. Choosing a suitable annotation procedure and evaluation approach is therefore essential to rate software performance appropriately [Mathias et al., 2014]. To better understand software misclassification, but to also account for the fact that we used software to detect faces and not any body part of chimpanzees, we applied consecutive and increasingly complex test steps after the manual and software processing. In the first step, we rated detections made by the software against all videos manually classified as containing at least one chimpanzee (i.e. the full set of positives). Second, since the software is based only on the detection of near-frontal faces and not bodies, we only considered videos that contained at least one face view of a chimpanzee (i.e. a subset of all positives). Post-processing then took place in the third and fourth steps. In the third step, we aimed at filtering out false positives, i.e. instances where the software responded to an object other than a chimpanzee, such as a swinging branch or a point on a tree (Fig. 2). Since these false detections are usually stationary objects (e.g. leaf or bark), their location estimates are stationary compared to variable whenever chimpanzees move across the scene. We calculated the cumulative distance between the detected face locations in consecutive video frames and removed detections whose cumulative distance was lower than 0.02 (i.e. 2% of the frame width). This threshold was based on the inspection of true and false positive detections with the aim of minimizing the loss of true detections. Lastly, in our fourth step, we only considered video clips where at least one chimpanzee individual's face was in a frontal position (i.e. both eyes facing the camera) and the associated detection was moving over a detectable cumulative distance (i.e. greater than 2% of the video size).

Performance of face detection approach

We tested the performance of the software at three levels: 1) simple presence/absence, 2) sightings vs. time relation to detect chimpanzees manually compared to automatically, and 3) occupancy modeling.

1) Confirming presence/ absence: We determined how often the face detection software correctly recognizes chimpanzee presence and absence (see above). We then applied the four consecutive processing steps and calculated the proportion of each detection category.

2) Detection time: For both the manually and automatically processed video data we derived accumulation curves showing the cumulative number of cameras with which chimpanzee presence was confirmed as a function of time.

3) Occupancy modeling: We interpret the commonly used term 'occupied site' as a 'site used by chimpanzees'. 'Naïve occupancy' is defined as the proportion of sites where a species is present within the surveyed period relative to all surveyed sites. To estimate the number of sites used by chimpanzees at both locations, we used a single-season model. We applied the "occu" function from the "unmarked" package in R [Fiske & Chandler, 2011]. This model estimates two parameters: 1) the probability that a species is present within a site, i.e. probability of occupancy (Ψ), and 2) the probability that a species present is detected within a site, i.e., probability of detection (p). More details about this model can be found in MacKenzie and colleagues [2006]. The model is based on four assumptions that need to be respected to avoid any bias of estimators: 1) sites are closed, meaning that no emigration and no immigration occurs during the study; 2) probability of detection is constant across all sites and surveys or is a function of site-survey covariates; 3) probability of occupancy is constant across sites or is a function of covariates; and 4) detection of species and detection histories at each location are independent of one another [MacKenzie et al., 2002; MacKenzie et al., 2006; Fiske & Chandler, 2011]. We divided the sampling period into sampling occasions (SO) of four days each. We removed one of two sites close by, surveyed during the same time period and separated only by approximately 50 meters and we removed sites with less than five sampling occasions. We also combined close and consecutively surveyed sites to avoid violating independence of detection among sites. We took only the first ten SO per camera into account for several reasons: first, the number of sites with more than ten SO was low and thus the value of detection probability could be biased and have lower precision; second, MacKenzie and colleagues [2002] recommend at least six SO in order to obtain a relatively unbiased occupancy probability; third, we limited the length of the study in order to meet the assumption of site closure; lastly, ten SO represent a total length of 40 days, a length compatible and reasonable with field surveys.

Detection histories were compiled into a matrix containing four different values: (0) when no detection occurred neither manually nor by the software, i.e. a true negative (TN); (1) when a true positive (TP) detection occurred, meaning that a chimpanzee was detected by the software and confirmed manually; (2) when a false positive (FP) occurred, meaning that a chimpanzee detected by the software was not confirmed manually; and (3) when a false negative (FN) occurred, meaning that a chimpanzee detected manually was not recognized by the software. When no survey was conducted during a SO (e.g. due to camera malfunctioning), we assigned a value of N/A. In the case where several videos with different classifications (i.e. FN, FP, TP) occurred in the same sampling occasion, we prioritized classes as follows: TP>FN>FP>TN. A FN leads to a loss of information and is therefore more important than a FP, easily corrected to a TN when watching the videos. For example, if during a sampling occasion both a video without a chimpanzee but with a detection by the software occurred and a video with a chimpanzee not detected by the software occurred, the sampling occasion was classified as a FN. We ran models for four datasets per site, respectively: the manual dataset including all videos and three other datasets based on the face recognition software output and the fourth processing level (i) one with no manual cleaning, (ii) one, in which false positive were removed and (iii) one, in which the proportional removal of false positive and false negatives was equal.

We developed an assessment study where we “cleaned” false positive and false negative sampling occasions manually by 10% increments; “cleaned” FP SO were transformed into TN SO, and “cleaned” FN SO were transformed into TP SO. We ran 1000 simulations to get occupancy and detection probabilities for each assessment. We used the ‘plogis’ function in order to obtain the occupancy probability (Ψ) at the original scale, with values between 0 and 1. A (0) means that the site is not used by chimpanzees and a (1) means that the site is used by individuals. We calculated the naïve occupancy by taking the number of sites where a chimpanzee was at least once manually detected divided by the total number of sites surveyed.

All analyses and graphs were carried out in R (version 3.0.2; R Development Core Team, 2013; https://www.r-project.org) and map was created in QGIS 2 (version 2.10.1 Pisa; QGIS Development team, 2015; http://www.qgis.org).

EXAMPLE

All field research protocol was in compliance with the EU Commission's legislation for animals used for scientific purposes, and adhered to the legal requirements in both Uganda and Liberia. All data collection at Sapo was performed in accordance with government regulations and approved by the Ministry of Agriculture in Liberia. It adhered to the legal requirements of the Bundesamt für Naturschutz/Federal Agency for Nature Conservation in Germany. Lastly, all field methods and research adhered to the American Society of Primatologists Principles for Ethical Treatment of Non-Human Primates, as well as the ethical guidelines established by the Max Planck Society.

Study sites

The data used in this study were gathered from two research sites with unhabituated chimpanzees as part of the Pan African Programme (http://panafrican.eva.mpg.de/index.php). The first site, the Budongo Conservation Field Station (henceforth Budongo), is located in the Budongo Forest Reserve in Western Uganda and comprises 428 km^2 of continuous forest (Fig. 3). The Budongo Forest is a moist semi-deciduous tropical rain forest situated between 1°37'- 2°03'N and 31°22' - 31°46'E and an average altitude of 1100 m [Eggeling, 1947; Plumptre, 1996]. At the time of data collection the mean monthly rainfall was 125 ± 87 mm and mean minimum and maximum temperatures per day were 16.4 ± 1.3°C and

31.5 ± 2.3°C, respectively (K. Corogenes, unpublished data). The study was conducted in the home range of the unhabituated 'Kamira' community living adjacent to two habituated chimpanzee communities ('Sonso' and 'Waibira'). No information about this specific community has yet been published. The second site is in Sapo National Park in Southwestern Liberia (henceforth Sapo), situated between 5°24' - 5°50'N and 8°24'- 52'W and comprises over 1,800 km^2 of tropical rain forest [Robinson & Peal, 1981]. At the time of data collection mean monthly rainfall was 211 ± 151 mm and mean minimum and maximum temperatures were 21.7 ± 1.5°C and 29.2 ± 3.1°C, respectively (V. Leinert, unpublished data). Around 1,500 chimpanzees are estimated to be in the park [Tweh et al., 2014].

Camera trapping

We installed Bushnell Trophy Cam cameras at both sites, following a standard protocol (http://panafrican.eva.mpg.de/pdf/Pan_African_Field_Protocol.pdf). At Budongo, 18 high-resolution cameras ("HR", Bushnell Trophy Cam 2012 model 119466; 720x1080 resolution) were opportunistically placed in a 2x3 km^2 grid between July 2012 and March 2013 at 24 unique locations. At Sapo, 34 lower-resolution cameras ("LR", Bushnell Trophy Cam 2010 model 119435; 480x620 resolution) were placed at 172 unique locations between January 2011 and May 2012 in a 5x5 km^2 grid. Cameras were attached to trees 1 m above ground at sites where chimpanzee encounters were likely, i.e. feeding spots, natural bridges and trails. Cameras were triggered by movement, which activated a 60 s recording, followed by a minimum 1 sec break before another recording. Cameras were active 24 h a day and checked once a month to change batteries and memory cards.

Results

At Budongo the field sampling effort consisted of 2809 trap days with a mean of 117 trap days per camera location. A total of 6733 HR videos were produced, of which 625 included sightings of chimpanzees (*Pan troglodytes schweinfurthii*) (Table I). The manual analysis found a total of 119 captured frontal face views of chimpanzees, with 110 videos containing at least one frontal face view. In 190 videos, only body parts of chimpanzees were visible. At Sapo, the field sampling effort consisted of 8365 trap days with a mean of 55.4 trap days per location. A total of 8996 LR videos were captured. Of these videos 279 videos contained chimpanzee sightings, with 216 total frontal face views and 148 videos with at least one frontal face view based on the manual analysis (Table I).

Performance of face detection approach

Confirmation of Presence/absence

In general, we found the same trend at both sites, though notably more pronounced for HR videos: as the post-processing level of comparison increased, the number of false detections decreased and true detections increased (Fig. 4). In the second step, after considering only videos containing chimpanzee face views as true detections, we found that TP and FN classifications nearly halved, but as a whole the total number of true detections (TP and TN) remains relatively constant. In the third step, after removing the false detections, we found that true classifications almost doubled and FPs decreased by more than 90% for HR videos and more than 25% for LR data. Finally, after the fourth level of assessment the rate of true detections (TP and TN) was 97% for HR and 98% for LR. For

HR, 25 of 110 videos containing chimpanzees were not recognized as such (i.e. false negatives), while for LR 82 of 148 videos were not recognized. Lastly, the FP rate was at 3% and less than 1% for HR and LR, respectively.

Detection time

We found that a majority of detections (>70%) occur in the first 40 days after camera establishment, when comparing manual and automated detections with all chimpanzee videos (Fig. 5). We also found that after 100 days of sampling, the face recognition software detected chimpanzees on only 50% of the cameras where a chimpanzee was detected manually, because of lack of face views. It is suggestive that chimpanzees walked in different directions and did not show their faces as often and therefore were not detected by the software.

Occupancy modeling

With the method described above, we used a total of 21 sites at Budongo and 100 sites at Sapo. Missing detections in tandem with false detections introduced bias in site occupancy probability estimates when using the LR dataset (Fig. 6B), occupancy probability was correctly estimated for the HR dataset (Fig. 6A). Cleaning only false positives in the case of the LR dataset, does not seem to be accurate. However, balancing the removal of false positives and false negatives seem to be better. When 100% of false positives and 50% of false negatives are cleaned, occupancy estimates are similar to those of the manual dataset and have estimates within the standard error interval of the manual value (Fig. 6).

COMPARISON AND CRITIQUE

Through a combination of manual and face detection approaches to evaluate occurrence, we have found that in its current advanced stage of development, face detection software (“FaceDetect”) is useful and indeed promising for use in the field when looking to determine chimpanzee occurrence. Our key goals that we demonstrated were to show that the software can be successfully used to simply detect presence- absence of chimpanzees in camera trap footage, can be used for site occupancy modeling and most importantly can speed up the process for analyzing field survey data by reducing the required time by up to 96-98%. Currently a critical limitation is that video clips need to contain face views for detection when chimpanzees are present. However, we think that this issue can be easily overcome on the level of field data collection until full body detection software is available. Sets of high resolution cameras can be placed in reverse directions at the same location that is surveyed for chimpanzee occurrence. Such approach should reduce non-detectability of chimpanzees due to lack of face views to an acceptable minimum. In essence combining camera trapping and semi-automated processing of footage will permit to conduct chimpanzee occupancy surveys routinely in an efficient manner.

Evaluation of face detection approach

The face detection software detected videos containing chimpanzee frontal face views with an acceptable low rate of false positives. However, we found that datasets had a large difference from one another: a detection rate of 77% and about 45% at fixed alarm rates of 2.8% and 0.8%, respectively. It is almost certain that this difference is due to camera placements that lead to occlusion of chimpanzee faces, and to differences in video resolution used at both sites. The face recognition software was developed using high quality videos with a resolution of 1280x1024, where visual images were pre-selected and then run through the software for recognition [Ernst & Küblbeck, 2011]. However, videos from camera traps can be of poorer quality due to lower resolution, weather and exposure to the elements. Differences in resolution may thus lead to different analysis of results: HR videos (720x1080, Budongo) had a higher recall rate, while LR videos (480x620, Sapo) had a lower recall rate. Our rate of false alarm of software detections in the last assessment was 2.8% for HR (Budongo) and 0.8% for LR (Sapo) data. This is comparable to similar studies which analyzed high quality images of chimpanzees and gorillas with face detection algorithms [Ernst & Küblbeck, 2011], but is lower than others that have looked at other species such as penguins [e.g. Sherley et al., 2010]. In these studies, as in ours, video quality plays a large role in the ability, accuracy and precision of species detection in data, and we stress the use of quality to improve results.

Time saving is undoubtedly the strongest argument for using face recognition software when comparing manual and automated methods. For example, from the 6733 HR videos (Budongo) we started with, we would only need to check the 285 videos classified as positive detections by the face detection software, and of the 8996 LR videos (Sapo) we started with, we would only need to check the 140 videos classified as positive detections, leaving aside for a moment the condition that chimpanzee presence can only be detected when their faces are visible. This results in a drastic decrease of 95.8% and 98.4% of videos to watch, respectively. When considering that about 3 min/video is needed to manually check for chimpanzee presence (time to open, start and watch the video, and note comments in a sheet), then an estimated 337 h are necessary to derive chimpanzee occurrence for the 6733 HR videos (Budongo). However, in the semi-automated assessment, only 285 videos would need to be reviewed, and thus only about 14.3 h are necessary to obtain occurrence information - a stark difference of 322.7 h.

In our last argument we address the aspect of false negatives and positives. For HR data (Budongo), we found that false negative detections were not a significant issue and relatively little information was lost; only 25 videos containing frontal face views were not detected. LR data (Sapo) had a much higher number of false negatives. Again, non-detections or false negative detections are likely due to poor resolution or occlusion. Additionally, while false positive detections could bias the occurrence analysis when only relying on the face detection software, they can be overcome by manually checking the reduced dataset. Thus we conclude that after post-processing, the face detection software performs well for detection, especially under the necessity that individuals must look directly in the camera and show their faces in order to be detected (see guidelines for field practitioners).

The fact that chimpanzees were detected either relatively quickly by the face detection software in camera trap footage or not at all is not a byproduct of overfitting the detection model, as the software was trained on a completely different dataset. Rather it is more likely that the positioning of cameras differed, which led to a higher or lower chance of recording chimpanzee face views.

Site occupancy modeling

Site occupancy modeling in conjunction with camera trapping can assess the presence of animals. We are aware that cameras were implemented within a small area in the chimpanzee territories and were opportunistically placed. Nevertheless, we know from long-term observations that chimpanzees do not use every part of their territory. We therefore interpret the estimated site occupancy as the used sites. Opportunistic camera placements should not be problematic if we consider only the animal populations within the area we sampled and not the greater region [Bengsen et al., 2011]. Alternatively the opportunistic camera placement we used can be replaced by a completely systematic design of camera placement across larger areas.

Guidelines for field practitioners

To maximize reliability of results, we recommend using high-resolution cameras to maximize the detectability by the face detection software. At least two cameras should be installed facing opposite directions at the site of interest to increase the chance of capturing individual faces. We also suggest that before implementing a study, simulation studies should be carried out to determine the prerequisites for robust estimates [Foster & Harmsen, 2012], minimum sampling effort (i.e., number of cameras), minimum sample area, and minimum sample size (i.e., number of individuals). Furthermore, for large scale studies cameras can be placed systematically, which would help meet the assumptions of occupancy modeling and reduce time to find suitable locations. Together, these aspects will increase result reliability and encourage the use of camera trapping in the field as part of an innovative and effective research approach.

In recent years, despite great strides in technology, many have been cautious of using face detection software to process field data, and have continued to rely arduously on human eye and hand. Yet the arguments for and benefits of using advanced software for data processing are growing and are increasingly hard to ignore. Here, we have demonstrated that the presence and absence of a species within an area can robustly be determined from the face detection software after post-processing video field datasets. We suggest that the time-saving benefits from the software outweigh the false positive detections that may result. Additionally, the long-term goal of this software employment will be to do individual recognition in order to obtain detailed demographic information on communities and populations.

We encourage the use of face detection and recognition software when looking to process large amounts of field data, when on a tight time schedule, and when strapped for skilled or trained human resources. As camera trapping becomes increasingly popular among conservation and community ecologists and researchers, this non-invasive method combined with a semi-automated face detection processing approach shows great potential for population surveys.

ACKNOWLEDGMENTS

First and foremost thank you to the Max Planck Society Innovation Fund, the Krekeler Foundation and the Robert Bosch Foundation for making this research possible. The face detection software was developed as part of the SAISBECO project, a collaboration funded by the Pact for Research and Innovation between the Max Planck Society and the Fraunhofer Society. Many notable thanks to the organizations and people that helped with the field research process at Budongo: the Uganda Wildlife Authority (UWA), the Uganda National Council for Science and Technology (UNCST) and the Ugandan National Forestry Authority (NFA), Geoffrey Muhanguzi, and field assistant Gerald Mayanja. We also thank the Forestry Development Authority Liberia, and to the hardworking field assistants at the Sapo field site: Lester B. Barcley, Samuel K. Freeman, Jefferson B. Kannah, Milton J. Tarnewon, Jerome T. Tarley. We also thank R. Mundry for his statistical advice. Last, thank you to the reviewers for their constructive comments in the editing process. The Pan African Project and Max Planck Institute are in compliance with animal care regulations and the national laws in the countries we work. The authors have no conflict of interest to declare.

REFERENCES

Andresen L, Everatt KT, Somers MJ. 2014. Use of site occupancy models for targeted monitoring of the cheetah. Journal of Zoology 292:212-220.

Araabi BN, Kehtarnavaz N, McKinney T, Hillman G, Würsig B. 2000. A string matching computer-assisted system for dolphin photoidentification. Annals of Biomedical Engineering 28:1269-1279.

Arandjelovic M, Head J, Kuehl H, Boesch C, Robbins MM, Maisels F, Vigilant L. 2010. Effective non-invasive genetic monitoring of multiple wild western gorilla groups. Biological Conservation 143:1780-1791.

Ardovini A, Cinque L, Sangineto E. 2008. Identifying elephant photos by multi-curve matching. Pattern Recognition 41:1867-1877.

Bengsen AJ, Leung, LKP, Lapidge, SJ and Gordon, IJ. 2011. Using a general index approach to analyze camera-trap abundance indices. The Journal of Wildlife Management, 75:1222-1227.

Bermejo M, Rodríguez-Teijeiro JD, Illera G, Barroso A, Vilà C, Walsh PD. 2006. Ebola outbreak killed 5000 gorillas. Science 314:1564-1564.

Borchers DL, Efford MG. 2008. Spatially explicit maximum likelihood methods for capture–recapture studies. Biometrics 64:377-385.

Borchers DL, Buckland ST, Zucchini W. 2002. Estimating animal abundance: closed populations. London: Springer. 314 p.

Buckland ST, Plumptre AJ, Thomas L, Rexstad EA. 2010. Design and analysis of line transect surveys for primates. International Journal of Primatology 31:833-847.

Buckland ST, Anderson DR, Burnham KP, Laake JL Borchers DL, Thomas L. 2001. Introduction to Distance sampling. Oxford: Oxford University Press. 432 p.

Campbell G, Kuehl H, N’Goran Kouamé P, Boesch C. 2008. Alarming decline of West African chimpanzees in Côte d’Ivoire. Current Biology 18:R903–R904.

Carlsen F, Leus, K, Traylor-Holzer, K, McKenna, A. (Editors). 2012. Western Chimpanzee Population and Habitat Viability Assessment for Sierra Leone: Final Report. Copenhagen, Denmark: IUCN/SSC Conservation Breeding Specialist Group – Europe (CBSGEurope). 124 p.

Dunn A, Bergl R, Byler D, Eben-Ebai S, Etiendem DN, Fotso R, Ikfuingei R, Imong I, Jameson C, Macfie L, Morgan B, Nchanji A, Nicholas A, Nkembi L, Omeni F, Oates J Pokempner A, Sawyer S, Williamson EA. 2014. Revised Regional Action Plan for the Conservation of the Cross River Gorilla (Gorilla gorilla diehli): 2014–

2019. New York, NY, USA: IUCN/SSC Primate Specialist Group and Wildlife Conservation Society, 49 p.

Eggeling WJ. 1947. Observations on the ecology of the Budongo rain forest, Uganda. Journal of Ecology 34:20-87.

Ernst A, Küblbeck C. 2011. Fast face detection and species classification of African great apes. In: Advanced Video and Signal-Based Surveillance (AVSS), 8th IEEE International Conference on. IEEE. pp. 279–284.

Fiske I, Chandler R. 2011. Unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software, 43: 1-23.

Foster RJ, Harmsen BJ. 2012. A critique of density estimation from camera-trap data. The Journal of Wildlife Management 76:224-236.

Funwi-Gabga N, Kühl H, Maisels F, Cheyne SM, Wich SA & Williamson EA. 2014 The status of apes across Africa and Asia. In: State of the Apes 2013: Extractive Industries and Ape Conservation. Book authored by the Arcus Foundation, State of the Apes series, Cambridge: Cambridge University Press, pp. 252-277.

Gaston KJ, O'Neill MA. 2004. Automated species identification: why not? Philosophical Transactions of the Royal Society B: Biological Sciences 359:655-667.

Gates JF. 1996. Habitat alteration, hunting and the conservation of folivorous primates in African forests. Australian Journal of Ecology 21:1-9.

Greengrass EJ. 2009. Chimpanzees are close to extinction in Southwest Nigeria. Primate Conservation 24:77-83.

Guschanski K, Vigilant L, McNeilage A, Gray M, Kagoda E, Robbins MM. 2009. Counting elusive animals: Comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. Biological Conservation 142:290-300.

Head JS, Boesch C, Robbins MM, Rabanal LI, Makaga L, Kühl HS. 2013. Effective sociodemographic population assessment of elusive species in ecology and conservation management. Ecology and Evolution 3:2903-2916.

Hughes B, Burghardt T. 2015. Automated identification of individual great white sharks from unrestricted fin imagery. In 26th British Machine Vision Conference (BMVC), pp. 92.1-92.14.

IUCN & ICCN. 2012. Bonobo (Pan paniscus): Conservation Strategy 2012–2022. Gland, Switzerland: IUCN/SSC Primate Specialist Group & Institut Congolais pour la Conservation de la Nature. 65 p.

Junker J, Blake S, Boesch C, Campbell G, du Toit L, Duvall C, Ekobo A, Etoga G, Galat-Luong A, Gamys J, Ganas-Swaray, Gatti S, Ghiurghi A, Granier N, Hart J, Head J, Herbinger I, Hicks TC, Huijbregts B, Imong II, Kuempel N, Lahm S, Lindsell J, Maisels F, McLennan M, Martinez L, Morgan B, Morgan D, Mulindahabi F, Mundry R, N'Goran KP, Normand E, Ntongho A, Okon DT, Petre CA, Plumptre A, Rainey H, Regnaut S, Sanz C, Stokes E, Tondossama A, Tranquilli S, Sunderland-Groves J, Walsh P, Warren Y, Williamson EA, Kuehl HS. 2012. Recent decline in suitable environmental conditions for African great apes. Diversity and Distributions 18:1077-1091.

Kalman R. 1960. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME-Journal of Basic Engineering, 82: 35-45.

Köndgen S, Kuehl H, N'Goran PK, Walsh PD, Schenk S, Ernst N, Biek R, Formenty P, Mätz-Rensing K, Schweiger B, Junglen S, Ellerbrok H, Nitsche A, Briese T, Lipkin WI, Pauli G, Boesch C, Leendertz FH. 2008. Pandemic human viruses cause decline of endangered great apes. Current Biology 18:260-264.

Kormos R, Boesch C. 2003. Regional action plan for the conservation of chimpanzees in West Africa. Washington DC: IUCN/SSC Primate Specialist Group and Conservation International. 24 p.

Kouakou CY, Boesch C, Kuehl H. 2009. Estimating chimpanzee population size with nest counts: validating methods in Taï National Park. American Journal of Primatology 71:447-457.

Küblbeck C, Ernst A. 2006. Face Detection and Tracking in Video Sequences Using the Modified Census Transformation, Image and Vision Computing, 24:564-572.

Kühl HS, Burghardt T. 2013. Animal biometrics: quantifying and detecting phenotypic appearance. Trends in Ecology & Evolution 28:432-441.

Kuehl HS, Maisels F, Williamson L, Ancrenaz M. 2008. Best practice guidelines for surveys and monitoring of great ape populations. Occasional papers of the IUCN Species Survival Commission, Gland, Switzerland: IUCN/SSC Primate Specialist Group. 32 p.

Kuehl HS, Todd A, Boesch C, Walsh PD. 2007. Manipulating dung decay time for efficient large-mammal density estimation: gorillas and dung height. Ecological Applications 17:2403-2414.

Kuehl HS, Nzeingui C, Yeno SLD, Huijbregts B, Boesch C, Walsh PD. 2009. Discriminating between village and commercial hunting of apes. Biological Conservation 142:1500-1506.

Lahiri M, Tantipathananandh C, Warungu R, Rubenstein DI, Berger-Wolf TY. 2011. Biometric animal databases from field photographs: identification of individual zebra in the wild. In: Proceedings of the 1st ACM International Conference on Multimedia Retrieval. ICMR ’11. New York: ACM. pp 6:1–6:8.

Leendertz FH, Pauli G, Maetz-Rensing K, Boardman W, Nunn C, Ellerbrok H, Jensen SA, Junglen S, Boesch C. 2006. Pathogens as drivers of population declines: the importance of systematic monitoring in great apes and other threatened mammals. Biological Conservation 131:325-337.

Leendertz FH, Ellerbrok H, Boesch C, Couacy-Hymann E, Mätz-Rensing K, Hakenbeck R, Bergmann C, Abaza P, Junglen S, Moebius Y, Vigilant L, Formenty P, Pauli G. 2004. Anthrax kills wild chimpanzees in a tropical rainforest. Nature 430:451-452.

Loos, A. 2016. Face Recognition for Great Apes: Identification of Primates in Real-World Environments (Doctoral dissertation). Technical University of Ilmenau, Germany, Department of Electrical Engineering and Information Technology, Institute for Media Technology. 279 p.

Loos A, Ernst A. 2012. Detection and Identification of Chimpanzee Faces in the Wild. In: 2012 IEEE International Symposium on Multimedia (ISM). pp. 116-119.

Loos A, Ernst A. 2013. An automated chimpanzee identification system using face detection and recognition. EURASIP Journal on Image and Video Processing 2013:1-17.

MacKenzie DI, Royle JA. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105-1114.

MacKenzie DI, Nichols JD, Hines JE, Knutson MG, Franklin AB. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200-2207.

MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248-2255.

MacKenzie DI, Nichols JD, Royle JA, Pollock K, Bailey L, Hines J. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. London: Elsevier. 324 p.

Macmillan NA, Creelman CD. 2004. Detection theory: A user's guide. Mahwah, New Jersey: Lawrence Erlbaum Associates Inc. 512 p.

Maldonado O, Aveling C, Cox D, Nixon S, Nishuli R, Merlo D, Pintea L, Williamson EA . 2012. Grauer's Gorillas and Chimpanzees in Eastern Democratic Republic of Congo (Kahuzi-Biega, Maiko, Tayna and Itombwe Landscape): Conservation Action Plan 2012–2022. Gland, Switzerland: IUCN/SSC Primate Specialist Group, Ministry of Environment, Nature Conservation & Tourism, Institut Congolais pour la Conservation de la Nature & the Jane Goodall Institute. 66 p.

Mathias M, Benenson R, Pedersoli M, Van Gool L. 2014. Face detection without bells and whistles. European Conference on Computer Vision (ECCV). Springer International Publishing. pp. 720-735.

Morgan BJ, Adeleke A, Bassey T, Bergl R, Dunn A, Fotso R, Gadsby E, Gonder K, Greengrass E, Koulagna DK, Mbah G, Nicholas A, Oates J, Omeni F, Saidu Y, Sommer V, Sunderland-Groves J, Tiebou, Williamson E. 2011. Regional Action Plan for the Conservation of the Nigeria-Cameroon Chimpanzee (Pan troglodytes ellioti). IUCN/SSC Primate Specialist Group and Zoological Society of San Diego, CA, USA. 49 p.

Murai M, Ruffler H, Berlemont A, Campbell G, Esono F, Agbor A, Mbomio D, Ebana A, Nze A, Kühl HS. 2013. Priority areas for large mammal conservation in Equatorial Guinea. PLoS one 8:e75024.

Nichols JD, Williams BK. 2006. Monitoring for conservation. Trends in Ecology & Evolution 21:668-673.

Noss AJ, Gardner B, Maffei L, Cuéllar E, Montaño R, Romero-Muñoz A, Sollman R, O'Connell AF. 2012. Comparison of density estimation methods for mammal populations with camera traps in the Kaa-Iya del Gran Chaco landscape: density estimation with camera traps in the Chaco. Animal Conservation 15:527-535.

Oates J, Sunderland-Groves J, Bergl R, Dunn A, Nicholas A, Takang E, Omeni F, Imong I, Fotso R, Nkembi L, Williamson L. 2007. Regional Action Plan for the Conservation of the Cross River Gorilla (Gorilla gorilla diehli). IUCN/SSC Primate Specialist Group and Conservation International, Arlington, VA, USA, 36 p.

O'Connell AF, Nichols, JD, & Karanth, KU. 2010. Camera traps in animal ecology: methods and analyses. New York: Springer, 271 p.

Plumptre AJ. 1996. Changes following 60 years of selective timber harvesting in the Budongo Forest Reserve, Uganda. Forest Ecology and Management, 89:101-113.

Plumptre AJ, Cox D. 2006. Counting primates for conservation: primate surveys in Uganda. Primates 47:65-73.

Plumptre AJ, Reynolds, V. 1996. Censusing chimpanzees in the Budongo forest, Uganda. International Journal of Primatology, 17:85-99.

Plumptre AJ, Rose, R, Nangendo, G, Williamson EA, Didier K, Hart J, Mulindahabi F, Hicks C, Griffin B, Ogawa H, Nixon S, Pintea L, Vosper A, McClennan M, Amsini F, McNeilage A, Makana JR, Kanamori M, Hernandez A, Piel A, Stewart F, Moore J, Zamma K, Nakamura M, Kamenya S, Idani G, Sakamaki T, Yoshikawa M, Greer D, Tranquilli S, Beyers R, Furuichi T, Hashimoto C, Bennett E. 2010. Eastern Chimpanzee (Pan troglodytes schweinfurthii): Status Survey and Conservation Action Plan 2010–2020. Gland, Switzerland: IUCN/SSC Primate Specialist Group 52 p.

Robinson PT, Peal A. 1981. Liberia's wildlife: the time for decision. ZooNooz 54:7-21.

Schapire R, Singer Y. 1999. Improved boosting algorithms using confidence-rated predictions. Machine Learning, 37:297-336.

Sherley R, Burghardt T, Barham P, Campbell N, Cuthill I. 2010. Spotting the difference: towards fully-automated population monitoring of African penguins Spheniscus demersus. Endangered Species Research 11:101-111.

Todd AF, Kühl HS, Cipolletta C, Walsh PD. 2008. Using dung to estimate gorilla density: modeling dung production rate. International Journal of Primatology 29:549-563.

Tuyttens FAM, de Graaf S, Heerkens JLT, Jacobs L, Nalon E, Ott S, Stadig L, Van Laer E, Ampe B. 2014. Observer bias in animal behaviour research: can we believe what we score, if we score what we believe? Animal Behaviour 90:273-280.

Tweh CG, Lormie MM, Kouakou CY, Hillers A, Kühl HS, Junker J. 2014. Conservation status of chimpanzees (*Pan troglodytes verus*) and other large mammals in Liberia: a nationwide survey. Oryx 49:710-718.

Viola P, Jones M. 2001. Rapid object detection using a boosted cascade of simple features. Conference on Computer Vision and Pattern Recognition (CVPR), (Kauai, Hawaii, USA), pp. 511-518.

Walsh PD, Abernethy KA, Bermejo M, Beyers R, De Wachter P, Akou ME, Huijbregts B, Mambounga DI, Toham AK, Kilbourn AM, Lahm SA, Latour S, Maisels F, Mbina C, Mihindou Y, Obiang SN, Effa EN, Starkey MP, Telfer P, Thibault M, Tutin CEG, White LJT, Wilkie DS. 2003. Catastrophic ape decline in western equatorial Africa. Nature 422:611-614.

Welch G, Bishop G. 2006. An introduction to the Kalman Filter. Tech. Rep., Department of Computer Science, University of North Carolina.

Wich SA, Garcia-Ulloa J, Kühl HS, Humle T, Lee JS, & Koh LP. 2014. Will oil palm's homecoming spell doom for Africa's great apes? Current Biology 24:1659-1663.

Wich SA, Meijaard E, Marshall AJ, Husson S, Ancrenaz M, Lacy RC, van Schaik CP, Sugardjito J, Simorangkir T, Traylor-Holzer K, Doughty M, Supriatna J, Dennis R, Gumal M, Knott CD, Singleton I. 2008. Distribution and conservation status of the orang-utan (Pongo spp.) on Borneo and Sumatra: how many remain? Oryx 42:329-339.

Woodford MH, Butynski TM, Karesh WB. 2002. Habituating the great apes: the disease risks. Oryx 36:153-160

Zabih R, Woodfill J. 1994. Non-parametric local transforms for computing visual corespondence. European Conference on Computer Vision (ECCV). Berlin Heidelberg: Springer, pp. 151-158.