Virtual plaster cast: digital 3D modelling of lion paws and tracks using close-range photogrammetry

Short title: Digital 3D modelling of lion paws and tracks

Antoine F.J. Marchal[1,2*]

Philippe Lejeune[2]

P.J. Nico de Bruyn[1]

1. Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield 0028, Pretoria, South Africa

2. Management of Forest Resources, Department of Biosystems Engineering, Gembloux Agro-Bio Tech, University of Liège, Passage des Déportés 2, B-5030, Gembloux, Belgium

**Corresponding author:*

marchal.ant@gmail.com

+27 (0)79 548 52 54

Abstract

The ecological monitoring of threatened species is vital for their survival as it provides the baselines for conservation, research and management strategies. Wildlife studies using tracks are controversial mainly due to unreliable recording techniques limited to two-dimensions (2D). We assess close-range photogrammetry as a low-cost, rapid, practical and reliable field technique for the digital three-dimensional (3D) modelling of lion (*Panthera leo*) paws and tracks. Firstly, we tested three reconstruction parameters affecting the 3D model quality. We then compared direct measurements on the paws and tracks versus the same measurements on their digital 3D models. Finally, we assessed the minimum number of photographs required for the 3D reconstruction. Masking, auto-calibration and optimisation provided higher reconstruction quality. Paws masked semi-automatically and tracks masked manually were characterised by a geometric deviation of 0.23 ± 0.18 cm and 0.50 ± 0.33 cm respectively. Unmasked tracks delineated by means of the contour lines had a geometric deviation of -0.06 ± 0.39 cm. The use of a correction factor reduced the geometric deviation to -0.03 ± 0.20 cm (pad-masked paws), -0.04 ± 0.35 cm (pad-masked tracks) and -0.01 ± 0.39 cm (unmasked tracks). Based on the predicted error, the minimum number of photographs required for an accurate reconstruction is seven (paws) or eight (tracks) photographs. This field technique, using only a digital camera and a ruler, takes less than one minute to sample a paw or track. The introduction of the 3D facet provides more realistic replications of paws and tracks that will enable a better understanding of their intrinsic properties and variation due to external factors. This advanced recording technique will permit a refinement of the current methods aiming at identifying species, age, sex and individual from tracks.

Keywords

Tracking; digital 3D model; digital photogrammetry; computer vision; Agisoft PhotoScan; foot; footprint; *Panthera leo*

Introduction

Ecological monitoring provides basic information on status and distribution of animal populations that is crucial for conservation, research and management strategies (Gese, 2001). Using tracks is often considered as a non-invasive, cost- and time-effective way of gaining information on species that are difficult to observe (Gese, 2001; Long et al., 2012). As an integral part of hunting, the earliest human beings have developed the art of tracking that is still used by modern hunter-gatherers such as the San people of Southern Africa (Liebenberg, 1990a). A study in Namibia showed that modern-day San trackers were 96% accurate in interpreting the species, age, sex and individual from tracks for 317 cases (Stander et al., 1997). Track measurements were used to achieve similar levels of identification as that of the San trackers for larger felids such as leopard (*Panthera pardus*) (Stander et al., 1997; Gusset & Burgener, 2005), tiger (*P. tigris*) (Gore et al., 1993; Sharma, Jhala & Sawarkar, 2003; Sharma, Jhala & Sawarkar, 2005), lion (*P. leo*) (Stander et al., 1997) and mountain lion (*Puma concolor*) (Smallwood & Fitzhugh, 1993; Grigione et al. 1999; Jewell, Alibhai & Evans, 2014), and for black (*Diceros bicornis*) and white (*Ceratotherium simum*) rhinoceroses (Jewell, Alibhai & Law, 2001; Alibhai, Jewell & Law, 2008). The most significant example of track use in wildlife studies is the "pugmark census method" that has been implemented for more than three decades to monitor the tiger populations in India (Karanth et al., 2003; Sharma et al., 2005). Designed in 1966, this census involves thousands of rangers that are searching for tracks across India for set periods of time (Choudhury, 1970; Choudhury, 1972). Tracings of the left hind paw's tracks of purportedly nearly all the tigers are then compared for individual identification. This type of census using tracks is highly controversial since the protocol does not take into consideration the variation due to different manipulators and substrates, and the individual identification is highly subjective (Karanth et al., 2003). The pugmark census method and all the above-mentioned track measurement methods are using recording techniques limited to two-dimensions (2D): direct measurement, drawing on acetate sheets or taking photographs. More recently, a Microsoft Kinect depth sensor was used to capture depth images of tracks from captive tigers (Lokare et al., 2014).

The rigorous use of tracks in ecological monitoring requires the variables extracted from them to be sensitive to variation between animals (species, sex, age and individual) and insensitive to external factors (such as substrate and manipulator bias). Three-dimensional (3D) reproduction of an object that is inherently 3D inevitably provides a better representation of reality that will improve the understanding of its intrinsic properties and their variation. Photogrammetry, the "science of measuring in photos" (Linder, 2009), provides a potentially useful tool for such 3D reconstruction. However, any innovative application first requires validation. Here we determine whether close-range photogrammetry can be used as a rapid, practical and reliable field technique for the digital 3D modelling of lion paws and tracks. We first tested the influence of reconstruction parameters on the alignment step, before comparing direct versus digital measurements and finally we assessed the number of photographs required for the 3D reconstruction. This technique was developed with the practical considerations of remote study sites and proximity to potentially dangerous animals in mind. Additionally, digital 3D reconstruction can be computed with a commercially available low-cost non-customised software package that implements both digital photogrammetry and computer vision techniques.

Materials and methods

Study areas

The two study sites, Hluhluwe-iMfolozi Park (HiP, ~900 km^2) and Tembe Elephant Park (TEP, ~300 km^2), are located in the sub-tropical province of KwaZulu-Natal (KZN), eastern South Africa. These two fenced areas are managed by a provincial conservation agency, Ezemvelo KZN Wildlife (EKZNW). Situated in Zululand, HiP is characterised by hilly topography ranging from 40 m to 560 m above sea level with a mean annual rainfall of 650-985 mm. Three major rivers (Hluhluwe, Black iMfolozi and White iMfolozi River) traverse the park. TEP is located in Maputaland along the international border of South Africa with Mozambique and is characterised by sandy plains with ancient littoral dunes and a mean annual rainfall of 700 mm. Dry riverbeds in HiP and sandy roads in TEP provide optimal substrate for tracks. Current lion populations (July 2015) are estimated at ~120

individuals in HiP (Somers et al., in preparation) and ~40 individuals in TEP (Hanekom C., TEP's Ecologist, unpublished data).

Paw and track sampling

Twenty lion paws were opportunistically sampled during nocturnal captures in TEP (Fig. 1a). The captures were part of management activities unrelated to this project (Animal Population Control plan, proposed by Tembe Management Team and accepted by EKZNW Board). Twenty clear lion tracks were sampled in HiP after a direct observation, in front of a camera trap [Cuddeback Attack, Green Bay, Wisconsin, USA] or after identification by means of a tracking book such as Liebenberg (1990b) and Gutteridge & Liebenberg (2013) (Fig. 1b). Both paw and track sampling consisted of (i) directly measuring the length and width of the main pad and toes with the help of a 0-150 mm vernier calliper [Tork Craft, Midrand, South Africa] (Fig. 1b), and (ii) taking photographs to create digital 3D models using close-range photogrammetry. The same manipulator, A.F.J.M., did all the sampling and two different digital single-lens reflex cameras were used: Nikon D7100 (24.1 megapixels) with Nikkor 18-70 mm f/3.5-4.5 and Nikon D80 (10 megapixels) with Nikkor 50 mm f/1.8 for photographing the paws and tracks. The sampling was carried out following the guidelines provided in the photogrammetric package's user manual (Agisoft LLC, 2014a), as well as those described in De Bruyn et al. (2009). The manipulator took 10-15 photographs of the object (i.e. paw or track) with the same focal length from different angles and distances (Fig. 2). During image acquisition, the paw was positioned off the ground on a stand with a clamp making it strictly motionless (Fig. 1a). The photographs have to cover each side of the object to avoid blind spots and they have to overlap with each other (Fig. 2). The object must fill the frame but a feature can be absent in one photograph provided that it appears in others. A ruler, that needs to be visible on at least three photographs, was positioned near the object and remained motionless between photographs to provide a scaling measure (Fig. 1a).

Figure 1 - Paw and track sampling. a) During the paw sampling, the motionless paw is positioned on a stand with a clamp holding the ruler and orientating the paw upward. b) A vernier calliper was used for the direct measurements of paws and tracks.

Figure 2 - Placement of camera stations (blue frames) around the object of interest. The sparse point cloud as well as the camera positions and orientations are the outcomes of the camera alignment step. Note the two markers and the scale bar.

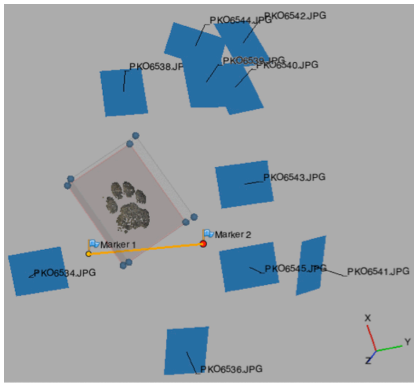
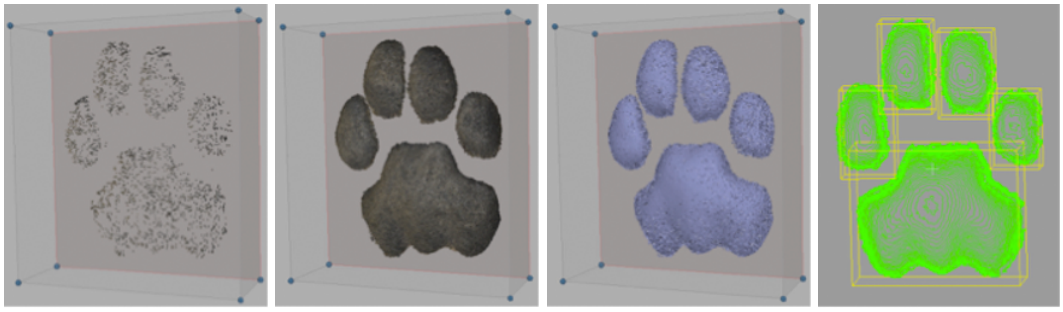


Figure 3 - General workflow in PhotoScan and contour lines. a) Sparse point cloud (2,812 points). b) Dense point cloud (4,572,854 points). c) Polygonal mesh (916,402 faces). d) 0.5 mm contour lines with non axis-orientated bounding boxes computed in CloudCompare.



3D modelling and re-projection error

The 3D modelling was performed with Agisoft PhotoScan Professional Edition version 1.1.4 build 2021 [Agisoft LLC, Saint Petersburg, Russia] (hereafter PS). PS is an image-based 3D modelling solution that can process arbitrary photographs taken in either controlled or uncontrolled conditions and that can reconstruct any visible object from at least two photographs (Verhoeven, 2011; Agisoft LLC, 2014a). PS implements both "Structure-From-Motion" (SFM) and "Dense Multi-View 3D Reconstruction" (DMVR) algorithms (Verhoeven, 2011). The reconstruction of a 3D model comprises three main steps: camera alignment (building sparse point cloud) (Fig. 2; Fig. 3a), building dense point cloud (Fig. 3b) and building polygonal mesh (Fig. 3c). The mesh can then be exported to external programs for further analyses (Fig. 3d). The camera alignment step applies the "bundle adjustment" method to search the feature points (i.e. key points) and match them between photographs (i.e. providing the tie points), find the external orientations (i.e. camera positions and orientations) and estimate the internal orientations (i.e. camera calibration parameters) (Fig. 2; Fig. 3a) (Triggs et al., 2000; Szeliski, 2010; Agisoft LLC, 2014a). The second step applies DMVR algorithms on the aligned image set by operating on the pixel values (Scharstein & Szeliski, 2002; Verhoeven, 2011). The outcome is a dense point cloud (Fig. 3b) that can then be transformed into a polygonal mesh (Fig. 3c). Following the alignment step, PS estimates the "camera error" or "re-projection error" in pixels that can be defined as the "root mean square re-projection error calculated over all feature points detected on the photograph" (Agisoft LLC, 2014a). The re-projection error is basically the distance between a projected point and the measured one (Gargallo, Prados & Sturm, 2007). This error provides crucial information about the quality and accuracy of the alignment step (Verhoeven et al., 2012).

Preferred reconstruction parameters?

There are three reconstruction parameters that can influence the camera alignment and that are tested here: (i) masking, (ii) calibration and (iii) optimisation. For scenario (i), the pictures were either not masked (unmasked) or masked around everything except the main pad and toes (pad-masked). In scenario (ii), the cameras were either automatically calibrated by PS (auto-calibrated) or manually

pre-calibrated (pre-calibrated) in external software. For the third scenario, either we did not apply an optimisation step (non-optimised) or we did (optimised). We selected three paw and three track datasets that contain between 11 to 12 photographs and 10 to 15 photographs respectively. These datasets were representative of our database and complete for the following testing procedures. We manually discarded any blurred photographs and those of lower quality (less than 0.5 units) by using the tool "estimate image quality" in PS. We then aligned the photographs using the highest accuracy (i.e. using original size photographs) and the default settings (Table 1). We positioned two markers (with two projections per marker) by using the "guided marker placement approach" - placing the marker projections on a single aligned photograph and the program automatically projects predictor rays onto the remaining photographs to reduce the chance of misplacing a marker. For each scaled 3D model originating from a specific dataset, we re-launched the alignment step three times for each possible combination of reconstruction parameters (i.e. eight combinations; e.g. combination 1: unmasked/auto-calibrated/non-optimised) and recorded the re-projection error.

Masking is a tool to exclude parts of the photographs, particularly the background, from the processing. The paws were semi-automatically masked in Photoshop Creative Cloud [Adobe, San Jose, California, USA]. After applying the options "sharpen edge" and "auto-contrast" to enhance the edges, we used the "quick selection tool" with an automatic edge refinement of 10-pixel-radius and 50% contrast (see Adobe Photoshop Cloud Creative help file). The "quick selection tool" was not successful for the tracks as the colours and texture were too uniform. Therefore, the tracks were manually masked in PS using the tool "intelligent scissors". For the pre-calibration, we manually estimated the camera calibration parameters using the software Agisoft Lens version 0.4.1 beta build 2021 [Agisoft LLC, Saint Petersburg, Russia] that uses the computer screen as calibration target. The calibration parameters were then imported into PS and used for aligning the photographs through the unfixed calibration mode. Aligning the photographs using image data only (i.e. through the tie points) leads to non-linear deformations originating from calibration errors (Agisoft LLC, 2014a). The optimisation step offers a refined bundle adjustment by adding ground control points to the

Table 1 - Settings used in the camera alignment, optimisation, build dense cloud and build mesh step.

Align photos	
Accuracy	High
Pair pre-selection	Disabled
Key point limit	40,000
Tie point limit	1,000

Optimisation	
Camera accuracy (m)	10
Marker accuracy (m)	0.005
Scale bar accuracy (m)	0.001
Projection accuracy (pix)	0.1
Tie point accuracy (pix)	4

Build dense cloud	
Quality	Ultra high
Depth filtering	Moderate

Build mesh	
Surface type	Arbitrary
Source data	Dense cloud
Face count	High
Interpolation	Disabled

calculations. We used the "scale bar based optimisation" with the default settings (Table 1) by using the two markers as ground control points.

Direct versus digital measurements?

To test for differences between direct measurements on the actual paws and tracks (i.e. length and width of the main pad and the four toes), and the same measurements on their digital 3D models, we reconstructed the 3D polygonal mesh of 20 paws and 20 tracks. The paw and track datasets contain between 12 and 14 photographs and between 10 and 15 photographs respectively. After discarding blurred and lower quality photographs, we launched the camera alignment step using the same settings as above, with two markers and two projections per marker, and using masking, auto-calibration and optimisation. Once the photographs aligned, we built the dense point cloud with the highest accuracy (i.e. using full photograph resolution) and moderate depth filtering (Table 1). Using the dense point cloud as a data source, we then built the mesh with the highest possible details (i.e. highest face count) for arbitrary surface type (i.e. non-topographic photogrammetry) and without automatic interpolation (i.e. only areas corresponding to dense point cloud are reconstructed) (Table 1). We cleaned the meshes by gradually selecting and removing all the patches that did not define the main pad or any of the toes. After automatically closing all the gaps in the meshes, we exported the five shapes. In CloudCompare, we used the tool "Principal Component Analysis (PCA) fit" to create a bounding box that was not axis-orientated (Fig. 3d) (CloudCompare, 2015). This allowed us to automatically extract the lengths and widths of each shape.

To avoid the subjective manual masking of the tracks, we reconstructed the 3D meshes from the same track photographs with the same settings as above but unmasked. After using the tool "PCA fit", we created the contour lines starting at the minimum height (i.e. bottom of the track) with a step of 0.5 mm (Fig. 3d). For each shape, we selected the highest isolated (i.e. non-connected to another shape) contour line as the shape delineation. As with the pad-masked 3D models, the length and width were then automatically extracted.

For the pad-masked paws, pad-masked tracks and unmasked tracks, we calculated the mean geometric deviation as the difference between digital and direct measurements. We estimated a correction factor to adjust the digital measurements using the following equation: Direct = Digital - Digital*Correction Factor ↔ Correction Factor = 1 - Direct/Digital.

Number of photographs required?

To assess the minimum number of required photographs, we selected three paw and three track datasets that all contained more than 10 photographs. Same as for the reconstruction parameters, these datasets were representative of our database and complete for the following testing procedures. We reclassified each dataset into subsets with an increasing number of pad-masked photographs randomly selected with replacement. In each subset, two photographs were always the same as they were used to position the two markers and the scale bar. Thus cancelling the influence of subjective marker positioning on the final 3D models. The random selection of photographs was repeated three times per dataset and per category of number of photographs. We reconstructed, cleaned and measured the mesh of each subset using the same procedure described above. For each category of number of photographs ranging from five to ten, we calculated the predicted error as the percentage of the absolute difference between the corrected digital and direct measurement. The total volume was also recorded.

Data analysis

We processed all our statistical analyses with the program R (R Development Core Team, 2014). We used a Mann-Whitney U test to analyse the difference between the mean re-projection errors for the paw and track datasets taken separately, while a Wilcoxon signed ranks test was used to study the effects of the reconstruction parameters on the same error value. We plotted both the direct versus digital and the direct versus corrected digital measurements for each case (pad-masked paws, pad-masked tracks and unmasked tracks) and we calculated the coefficient of correlation using a Spearman's rank-order correlation test. We plotted the mean predicted error with 95% confidence intervals against the number of photographs for the pad-masked paws and pad-masked tracks. Using a

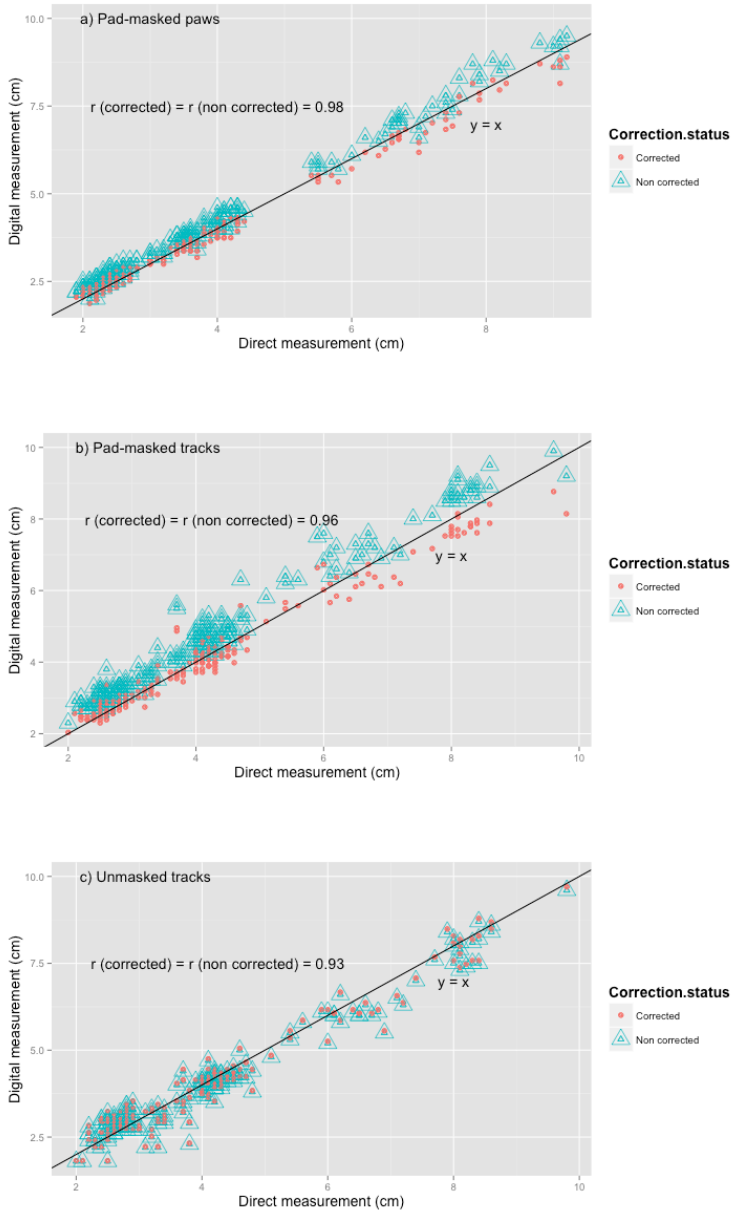
Mann-Whitney U test, we estimated the category of number of photographs in which the asymptote is reached (i.e. when the mean predicted error for that category is not significantly different from that of the category with 10 photographs). The probability values are considered statistically significant at $p \leq 0.05$.

Results

Reconstruction parameters

There is a significant difference (Mann-Whitney U test, $p < 0.001$) between the mean re-projection error for the paws (1.03 ± 0.39 pix) and for the tracks (0.47 ± 0.09 pix). Paws and tracks were therefore considered independently for testing the effects of masking, calibration and optimisation on the 3D model quality. Masking has a significant influence (Wilcoxon signed ranks test, $p < 0.001$) on the alignment of paw photographs. Mean re-projection error is lower for pad-masked (0.68 ± 0.13 pix) than for unmasked (1.39 ± 0.18 pix) paw photographs. The same influence (Wilcoxon signed ranks test, $p < 0.001$) is observed for the alignment of track photographs, with a mean re-projection error that is again lower for pad-masked (0.42 ± 0.09 pix) than for unmasked (0.53 ± 0.02 pix) photographs. Calibration does not have a significant influence on the alignment of paw photographs (Wilcoxon signed ranks test, $p = 0.822$). However, it has a significant influence (Wilcoxon signed ranks test, $p < 0.01$) on the alignment of track photographs, with a mean re-projection error that is lower for auto-calibrated (0.46 ± 0.07 pix) than for pre-calibrated (0.49 ± 0.10 pix) photographs. Optimisation has a significant influence (Wilcoxon signed ranks test, $p < 0.001$) on the alignment of both paw and track photographs. The mean re-projection error of the optimised alignment for paw photographs is lower (1.03 ± 0.39 pix) than in the non-optimised case (1.04 ± 0.39 pix). Similarly, lower re-projection error was observed when optimisation was applied (0.47 ± 0.07 pix) than when it was not (0.48 ± 0.10 pix) for the track photographs.

Figure 4 - Regression of direct versus digital measurements (corrected and non corrected) for a) pad-masked paws, b) pad-masked tracks and c) unmasked tracks. The line represents the true regression line (intercept = 0, slope = 1) and r is the coefficient of correlation using the Spearman's rank-order correlation test.



Direct versus digital measurements

The 3D models of both pad-masked paws and tracks present a positive geometric deviation of 0.23 ± 0.18 cm and 0.50 ± 0.33 cm respectively (Table 2; Fig. 4a; Fig. 4b), while a negative geometric deviation of -0.06 ± 0.39 cm (Table 2; Fig. 4c) is observed for the unmasked tracks. The calculated correction factor is 0.06 ± 0.05 for pad-masked paws, 0.11 ± 0.07 for pad-masked tracks and -0.01 ± 0.12 for unmasked tracks (Table 2). These factors may be used in predictive equations to adjust the overestimation in the case of the pad-masked paws and tracks, and the underestimation in the case of the unmasked tracks (Table 2). The use of the appropriate correction factor reduces the geometric deviation to -0.03 ± 0.20 cm for pad-masked paws, -0.04 ± 0.35 cm for pad-masked tracks and -0.01 ± 0.39 cm for unmasked tracks (Table 2; Fig. 4). The coefficient of correlation, Spearman's r, is 0.98 for pad-masked paws, 0.96 for pad-masked tracks and 0.93 for unmasked tracks, with no difference between non-corrected and corrected (Table 2; Fig. 4).

Number of photographs

For both the pad-masked paws and tracks, the mean predicted error decreases with an increasing number of photographs used to reconstruct the 3D models (Fig. 5). For the paws, an asymptote is reached between six and seven photographs, as the predicted error for seven photographs (5.15 ± 4.05%) is not significantly different from that of 10 photographs (4.22 ± 3.75%) (Mann-Whitney U test, $p = 0.09$) (Fig. 5a). The asymptote is reached for the tracks between seven and eight photographs (Mann-Whitney U test, $p = 0.06$), with a predicted error of 6.00 ± 3.37% for eight photographs and 5.07 ± 3.20% for 10 photographs (Fig. 5b). Other than observing an increasing predicted error when decreasing the amount of photographs, the 3D model volume also shrinks with a decreasing number of photographs. The mean volume for five photographs represents 67.78 ± 5.91% and 84.89 ± 9.31% of the mean volume for 10 photographs for the paws and tracks.

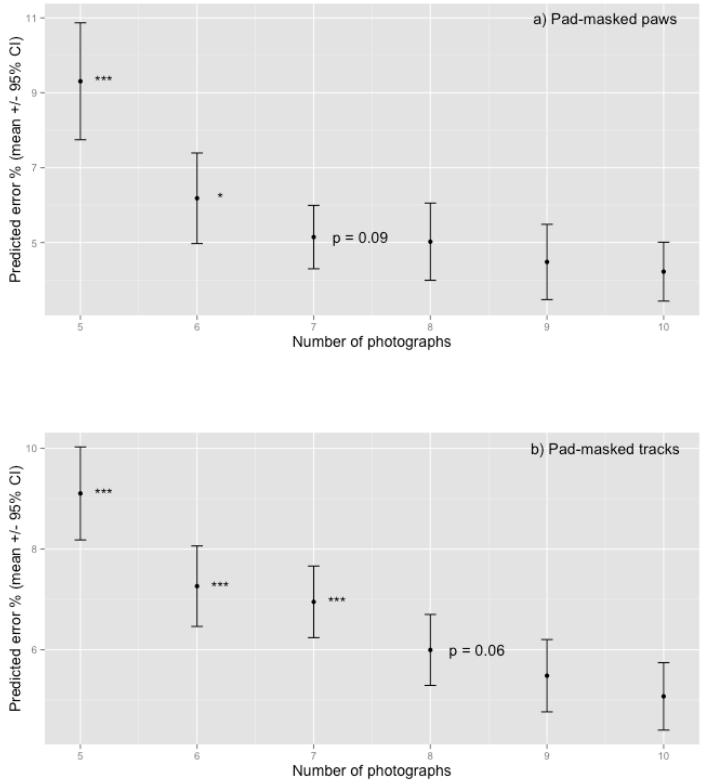
Sampling and processing time considerations

The image acquisition (i.e. photography) of either the paws or tracks took less than 1 minute per object. The manual masking in PS took on average 1.50 minutes per photograph (range: 1.15 to 2.14

Table 2 - Geometric deviation (non-corrected and corrected) and predictive equations to approximate the length and width of the main pad and toes for the pad-masked paw, pad-masked track and unmasked track models. r values are the resultant linear regression fit of direct to digital measurements and direct to corrected digital measurements calculated with Spearman's rank-order correlation test. "Dr" stands for "direct measurement" and "Dg" for "digital measurements".

Model	Geometric deviation (cm)		Equation	N	r	
	Non corrected	Corrected			Non corrected	Corrected
Pad-masked paws	0.23 ± 0.18	-0.03 ± 0.20	Dr = Dg - Dg * (0.06 ± 0.05)	200	0.98	0.98
Pad-masked tracks	0.50 ± 0.33	-0.04 ± 0.35	Dr = Dg - Dg * (0.11 ± 0.07)	200	0.96	0.96
Unmasked tracks	-0.06 ± 0.39	-0.01 ± 0.39	Dr = Dg - Dg * (-0.01 ± 0.12)	194	0.93	0.93

Figure 5 - Mean predicted error (%) and 95% confidence interval (CI) for each category of number of a) paw and b) track photographs. The predicted error is the percentage of the absolute difference between the corrected digital and the direct measurement. An asymptote is reached between 6 and 7 paw photographs, and between 7 and 8 track photographs as the mean predicted error (%) for 7 paw photographs and 8 track photographs is not significantly different from that for 10 photographs.



minutes) and the semi-automatic masking in Photoshop took on average 3.4 minutes (range: 2.42 to 4.20 minutes). For the processing of datasets containing 10 pad-masked photographs, two paw and one track datasets were computed with a laptop Mac Book Pro OSX Yosemite 2.8 GHz Intel Core i7 8GB memory (hereafter MAC), and one paw and two track datasets were computed with a desktop computer Windows 7 Enterprise 3.60 GHz Intel Core i7 16GB memory (hereafter PC). The MAC mean total processing time for paws (tracks) was 53.03 ± 22.86 minutes (5.82 ± 0.32 minutes) with the following breakdown in percentage for the three steps: 2% (5%) photograph alignment, 55% (79%) dense cloud building and 43% (16%) mesh building. Processing with the PC reduced the mean total processing time for paws (tracks) to 11.59 ± 0.94 minutes (1.70 ± 0.30 minutes) with the following breakdown in percentage for the three steps: 3% (8%) photograph alignment, 58% (77%) dense cloud building and 39% (16%) mesh building. Using the same datasets but with only five photographs, the total processing time for paws (tracks) becomes 22.12 ± 1.27 minutes (2.68 ± 0.38 minutes) using the MAC and 9.33 ± 3.98 minutes (0.58 ± 0.11 minutes) using the PC. Five track datasets containing 12 photographs each were processed with the PC in both unmasked and pad-masked condition. The mean total processing time was 46.32 ± 2.86 minutes for unmasked photographs and 2.25 ± 0.64 minutes for pad-masked photographs.

Discussion

In ichnology (i.e. science studying the interaction between organism and substrate), dinosaur tracks have previously been sampled using photogrammetry (Petti et al., 2008; Remondino et al., 2010). To our knowledge, this study represents the first application of close-range photogrammetry to record paws and tracks of extant animals in 3D. This innovative field technique provides an objective and reliable solution to obtain digital 3D models of both paws and tracks. The image acquisition time, less than a minute per paw or track, is ideal for minimising the interaction with immobilized individuals and for working with potentially dangerous species. Furthermore, the necessary equipment for the field data collection is essentially limited to a digital camera and a ruler.

The reconstruction parameters have a significant impact on the alignment step and therefore on the quality of the final 3D models. The comparison of the mean re-projection errors between the different possible combinations showed that masking, auto-calibration and optimisation yielded more accurate 3D reconstruction of both paws and tracks. Other than decreasing the processing time, another advantage of masking is the delineation of the object of interest (paw or track). However, it is important to use a delineation process that is not affected by the manipulator's subjectivity. This was successfully achieved for the paws by means of a semi-automatic masking tool in Photoshop. The tool could easily pick up the interface between the pads and the hair due to a high contrast in colour and texture. This clear contrast is not present in the track photographs and the masking tool showed limited success with the delineation between an imprint and the surface that enables its existence. While subjective manual masking of tracks led to a significant overestimation of the digital measurement (Table 2; Fig. 4b), the semi-automatic segmentation using unmasked tracks and contour lines led to the lowest geometric deviation (Table 2; Fig. 4c). Unmasked tracks present a higher mean re-projection error compared to pad-masked tracks, however this error remains less than that of pad-masked paws (see section Reconstruction parameters). In the case of the pad-masked paws, we believe that the overestimation of the digital measurement might in fact be due to an underestimation of the direct measurement. Since the pads are made of a thick elastic mass of connective tissue (Gutteridge & Liebenberg, 2013), the manipulator tends to compress the calliper on the pads leading to an underestimated measurement. The use of specified correction factors for the measurement estimation from the digital 3D models reduces the geometric deviation by two decimals of a centimetre (Table 2). The accuracy advertised by Agisoft for close-range photogrammetry with PS is 0.1 cm (Agisoft LLC, 2014b). For both paw and track 3D models, we showed that the predicted error increases and the total volume decreases, as the number of photographs used in the reconstruction decreases (Fig. 5). The suggested minimum of seven and eight photographs for paws and tracks respectively, represents a theoretical minimum number of photographs to process the 3D models. From our experience, approximately 7% of the photographs were discarded due to poor quality. Furthermore, it is not only the quantity of photographs that matters but also their position in the 3D space, as they must overlap without any blind spots (Fig. 2). Since more photographs make better

models and to avoid a lack of 2D information, we advise capturing twice as many photographs than the theoretical minimum (i.e. between 14 and 16 photographs). We further recommend masking the paws but not the tracks, and using both the auto-calibration and optimisation functions.

Previous studies using 2D have shown high accuracy (>90%) for objective individual identification from tracks made by black and white rhinoceroses (Jewell et al., 2001; Alibhai et al., 2008), mountain lions (Smallwood & Fitzhugh, 1993; Grigione et al., 1999; Jewell et al., 2014) and tigers (Sharma et al., 2005). Felid tracks were mainly sampled on dusty roads (i.e. producing shallow tracks) as other substrates, such as sand, generated greater variability of the track contour. Unfortunately, optimal dusty roads are not present everywhere. This is particularly the case in our study sites as TEP largely comprises sandy roads while HiP's unpaved roads are often too hard. The above-mentioned studies of wild felids sampled a limited number of individuals (from 3 to 17 individuals). Additionally, the identification accuracy was dependent on the number of tracks per track set (i.e. tracks belonging to the same individuals). Sharma et al. (2005) suggested a minimum of ten tracks per track set. Recording techniques in 2D are affected by the manipulator posture (Smallwood & Fitzhugh 1993) and experience (Karanth et al., 2003) during tracing, while photographs are affected by the time of the day and cloud cover (Grigione et al. 1999). Furthermore, photographs that are not aligned directly over the object can introduce a parallax error. In the same way that 3D has improved facial recognition methods (Chang, Bowyer & Flynn, 2003), we are confident that it will enable a more rigorous, objective and repeatable use of tracks in future studies. By providing more information, 3D replicas of tracks should enable the correct identification of more individuals on a greater variety of substrates with fewer tracks required per individual. Analysing the intrinsic properties of the paws will lead to a better understanding of the tracks they produce. The nature of the sampling technique, which requires several photographs taken from different distances and angles, is expected to be less affected by manipulator bias. This paper shows that working with digital 3D models ought to improve the track segmentation and feature extraction by decreasing the human input. Given the results of our innovative technological approach, we are currently working on improving the technique (e.g. understanding manipulator bias and using different types of cameras) and applying it to identify

individual lions from their paws and tracks. Identifying individuals from their tracks would have major implications in behavioural ecology, conservation biology and wildlife management. Tracking is less invasive than camera trapping, requires less investment and logistics while not being prone to hardware failure and theft.

Acknowledgements

We would like to thank Ezemvelo KZN Wildlife for allowing our research to be conducted in both Hluhluwe-iMfolozi Park and Tembe Elephant Park. We are particularly grateful to Dave Druce and Cathariné Hanekom, the respective park ecologists. Our appreciation is expressed to Birgit Eggers for allowing the paw sampling during the captures, to Jonathan Lisein for providing technical assistance on the photogrammetric package, and to Bronwen Klaas and two anonymous reviewers for commenting on earlier versions of this manuscript. A.F.J. Marchal benefited from a National Research Foundation (NRF) doctoral bursary.

References

Alibhai, S.K., Jewell, Z.C. & Law, P.R. (2008). A footprint technique to identify white rhino Ceratotherium simum at individual and species levels. *Endang. Species Res.* 4, 205–218.

Agisoft LLC (2014a). *Agisoft PhotoScan User Manual, Professional Edition, Version 1.1.* Agisoft LLC, St. Petersburg. URL http://www.agisoft.com/pdf/photoscan-pro_1_1_en.pdf [accessed 1 August 2015]

Agisoft LLC (2014b). *PhotoScan: fully automated professional photogrammetric kit.* Agisoft LLC, St. Petersburg. URL http://airgon.com/Agisoft_PhotoScan.pdf [accessed 1 August 2015]

Chang, K., Bowyer, K. & Flynn, P. (2003). Face recognition using 2D and 3D facial data. *ACM Workshop on Multimodal User Authentication*, 25–32.

Choudhury, S. (1970). Let us count our tigers. *Cheetal* 14, 41–51.

Choudhury, S. (1972). Tiger census in India. *Cheetal* 15, 67–84.

CloudCompare (2015). *CloudCompare (version 2.6.1)*. EDF, R&D, Telecom, ParisTech, Paris. URL http://www.cloudcompare.org/ [accessed 1 August 2015]

De Bruyn, P., Bester, M.N., Carlini, A.R. & Oosthuizen, W.C. (2009). How to weigh an elephant seal with one finger: a simple three-dimensional photogrammetric application. *Aquatic Biol.* 5, 31–39.

Gargallo, P., Prados, E. & Sturm, P. (2007). Minimizing the reprojection error in surface reconstruction from images. *11th IEEE International Conference on Computer Vision*, 1–8.

Gese, E.M. (2001). Monitoring of terrestrial carnivore populations. In *Carnivore Conservation*: 372–396. Gittleman, J.L., Funk, S., Macdonald, D. & Wayne, R.K. (Eds.). Cambridge: Cambridge University Press.

Gore, A., Paranjape, S., Rajgopalan, G., Kharshikar, A., Joshi, N., Watve, M. & Gogate, M. (1993). Tiger census: role of quantification. *Curr. Sci. India* 64, 711–714.

Grigione, M.M., Burman, P., Bleich, V.C. & Pierce, B.M. (1999). Identifying individual mountain lions Felis concolor by their tracks: refinement of an innovative technique. *Biol. Conserv.* 88, 25–32.

Gusset, M. & Burgener, N. (2005). Estimating larger carnivore numbers from track counts and measurements. *Afr. J. Ecol.* 43, 320–324.

Gutteridge, L. & Liebenberg, L. (2013). *Mammals of Southern Africa and their Tracks and Signs*. Sunnyside: Jacana Media.

Jewell, Z.C., Alibhai, S.K., Law, P.R. (2001). Censusing and monitoring black rhino (Diceros bicornis) using an objective spoor (footprint) identification technique. *J. Zool. (Lond.)* 254, 1–16.

Jewell, Z.C., Alibhai, S.K., Evans, J.W. (2014). Monitoring mountain lion using footprints: A robust new technique. *Wild Felid Monitor* 7, 26–27.

Karanth, K.U., Nichols, J.D., Seidenstricker, J., Dinerstein, E., Smith, J.L.D., McDougal, C., Johnsingh, A., Chundawat, R.S. & Thapar, V. (2003). Science deficiency in conservation practice: the monitoring of tiger populations in India. *Anim. Conserv.* 6, 141–146.

Liebenberg, L. (1990a). *The Art of Tracking, the Origin of Science*. Cape Town: David Philip.

Liebenberg, L. (1990b). *A Field Guide to the Animal Tracks of Southern Africa*. Claremont: David Philip.

Linder, W. (2009). *Digital Photogrammetry*. 3rd edn. Berlin: Springer.

Long, R.A., MacKay, P., Ray, J. & Zielinski, W. (2012). *Noninvasive survey methods for carnivores*. Washington: Island Press.

Lokare, N., Ge, Q., Snyder, W., Jewell, Z., Allibhai, S. & Lobaton, E. (2014). Manifold learning approach to curve identification with applications to footprint segmentation. *IEEE Computational Intelligence for Multimedia, Signal and Vision Processing*, 1–8.

Petti, F.M., Avanzini, M., Belvedere, M., De Gasperi, M., Ferretti, P., Girardi, S., Remondino, F. & Tomasoni, R. (2008). Digital 3D modelling of dinosaur footprints by photogrammetry and laser scanning techniques: integrated approach at the Coste dell'Anglone tracksite (Lower Jurassic, Southern Alps, Northern Italy). *Studi Trentini Sci. Nat. Acta Geol.* 83, 303–315.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. URL http://www.R-project.org/ [accessed 1 August 2015]

Remondino, F., Rizzi, A., Girardi, S., Petti, F.M. & Avanzini, M. (2010). 3D Ichnology - recovering digital 3D models of dinosaur footprints. *Photogramm. Rec.* 25, 266–282.

Scharstein, D. & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* 47, 7–42.

Sharma, S., Jhala, Y. & Sawarkar, V.B. (2003). Gender discrimination of tigers by using their pugmarks. *Wildl. Soc. Bull.* 31, 258–264.

Sharma, S., Jhala, Y. & Sawarkar, V.B. (2005). Identification of individual tigers (*Panthera tigris*) from their pugmarks. *J. Zool. (Lond.)* 267, 9–18.

Somers, M.J., Becker, P., Druce, D.J., Graf, J., Gunther, M., Marneweck, D., Trinkle, M., Moleón, M. & Hayward, M.W. (in preparation). Reassembly of the large predator guild into Hluhluwe-iMfolozi

Park. In *Savanna Ecology and Management: Conserving Africa's Mega-Diversity in the Hluhluwe-iMfolozi Park*. Cromsigt, J.P.G.M., Archibald, S. & Owen-Smith, N. (Eds.).

Smallwood, K.S. & Fitzhugh, E.L. (1993). A rigorous technique for identifying individual mountain lions Felis concolor by their tracks. *Biol. Conserv.* 65, 51–59.

Stander, P., Ghau, I., Tsisaba, D., Oma, I. & Ui, I. (1997). Tracking and the interpretation of spoor: a scientifically sound method in ecology. *J. Zool. (Lond.)* 242, 329–341.

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Berlin: Springer.

Triggs, B., McLauchlan, P.F., Hartley, R.I. & Fitzgibbon, A.W. (2000). Bundle adjustment - a modern synthesis. In *Vision Algorithms: Theory and Practice*: 298–372. Triggs, B., Zisserman, A. & Szeliski, R. (Eds.). Berlin: Springer-Verlag.

Verhoeven, G. (2011). Taking computer vision aloft–archaeological three-dimensional reconstructions from aerial photographs with photoscan. *Archaeol. Prospect.* 18, 67–73.

Verhoeven, G., Doneus, M., Briese, C. & Vermeulen, F. (2012). Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *J. Archaeol. Sci.* 39, 2060–2070.