A Rule-Based Method for Automated Footprint Localization and Classification of Small Species

Haokun Geng[1], Cheryl Mills[2], Radu Nicolescu[1], and Reinhard Klette[1]

[1] Department of Computer Science, University of Auckland, Private Bag 92019
Auckland, New Zealand
hgen001@aucklanduni.ac.nz, {r.nicolescu, r.klette}@auckland.ac.nz

[2] Department of Biosciences, Cornwall Campus, University of Exeter
Cornwall, TR10 9EZ, United Kingdom
cm271@exeter.ac.uk

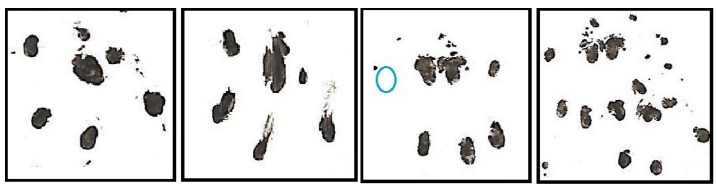
Abstract. In environmental surveillance, ecology experts use a standard tracking tunnel system to acquire tracks or footprints of small animals, so that they can easily measure the presence of any selected animals or detect threatened species based on the manual analysis of gathered tracks. Distinguishing morphologically similar species through analysing footprints requires a great amount of efforts on observation, even experienced wildlife experts can not accomplish this task with highly reliable results. In recent years, image processing technology has become a model example for applying computer science technology to many other study areas or industries, in order to improve accuracy, productivity, and reliability. In this paper, we demonstrate a model/rule-based method for automated footprint identification which includes localization and classification of small species. With appropriate developments or modifications, this method has certain potential for automated identification of any species.

Keywords: environmental surveillance, geometric model, rule-based method, image processing, pattern analysis, automated footprint identification, track analysis.

1 Introduction

Computer-based system has been a common technique of humankind to perform activities that have to be repeated numerous times [3], identifying small species from their footprints is one of such activities. Currently ecological experts need to spend a lot of effort and time on identifying footprints from in inked tracking card highly regarded the experts' knowledge and experiences, and the manual identification analysis often requires to be repeated over and over again on many tracking cards. Therefore, we would like to using pattern recognition technology to provide an automated method to replace or assist the manual identification analysis. The demand for such systems that can process the automated identification of species from their scanned footprint images is most likely to increase in the future [6]. It becomes essential to have such working application, that can be properly incorporated into the current system, handles the monotonous jobs, and outputs accurate and reliable analysis result.

However, the presentations of footprints are massively varied, the "puzzle" is that the images of a footprint may have very different appearances (as shown in Fig. 1). Besides normal footprints, the set of undesirable image data include "sliding footprints", "missing toe footprints", and "overlapped footprints". Before any further analysis can be carried out by the automated recognition algorithm, those varied representations of the footprints need to be transformed into digitalised geometric models. Correctly handling the transformation process is certainly a difficult task.



(a) (b) (c) (d)

Fig. 1. Samples of mice's footprints in different situations. (a) Normal front footprint. (b) Sliding footprint. (c) Missing toe footprint. (d) Overlapped footprints.

In this paper we firstly give the standard geometric models of a common mice's footprints. Based on the geometric models, we describe a rule-based track recognition algorithm that performs automated footprint localization and classification. It follows three major steps: (1) track acquisition: the current standard procedure for collecting tracking cards; (2) geometric model isolation: isolate normal footprints from the massive tracks on a tracking card; (3) algorithm implementation: defines a set of rules to perform footprints localization and classification.

2 Track Acquisition

The *Tracking Tunnel System* is a widely used standard procedure for collecting tracks of small animals to gain an index of the density of target small species in New Zealand [1]. It is a cost-effective method to collect tracks of small species over large areas [8]. Providing reasonable analysis and reliable results on the estimate of species' presence plays an important role in ecological research when ecologists decide to study rare species or assess community composition for environmental surveillance or pest control [5].

Traditionally, tracks or footprints are collected by this tracking tunnel system, and the identification of tracks and footprints is handled manually by experienced wildlife experts [2]. The basic principle of animal tracking is firstly to recognise single footprints from a number of unknown footprints, and then to identify the species based on the analysis of its footprints [8].

The tracking tunnel system is considered the first step when ecologists would like to monitor or study on a selected species. The collected tracks or footprints need to be analysed manually by human experts. In the identification procedure, distinguishing them among many morphologically similar species through analysing their footprints is extremely difficult, and one single tracking card is also possible to contain footprints from different species [8]. Our method aims to ultimately implement an automated recognition process to assist experts in the current identification procedure.

3 Geometric Model Isolation

First of all, we isolate normal footprints from the massive tracks on a tracking card. The front foot for a common rat usually has four toes, the hind foot usually has five toes [8]. The toes of the front foot are evenly distributed around the central pad. The hind foot normally has three toes bunched in front of the central pad that can roughly form a straight line. Based on the pervious studies [8] and our experimental set of tracking cards, we isolated normal footprints from tracks. The isolated front footprint model is shown in Fig. 2, and the isolated hind footprint model is shown in Fig. 3.

Analysing the isolated footprints could provide geometric models for both front and hind footprints. The isolated front footprint has a clear geometric structure. The toe prints are marked by blue circles, the central pad and accessorial pads prints are marked by red circles. The central pad is distributed in the middle point of that line segment. The central pad and two accessorial pads clearly form a triangle. Also there are three straight lines all across the central pad, they are 'T1' to 'T4', 'T2' to 'A2', and 'T3' to 'A1' (as shown in Figure 4, left).

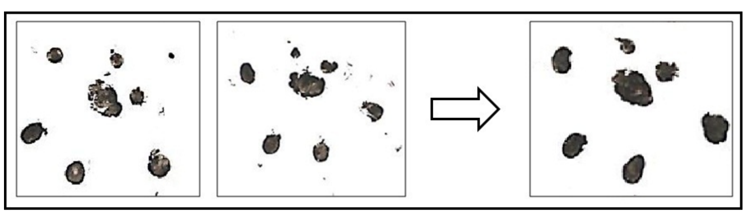


Fig. 2. Isolated front footprint model example.

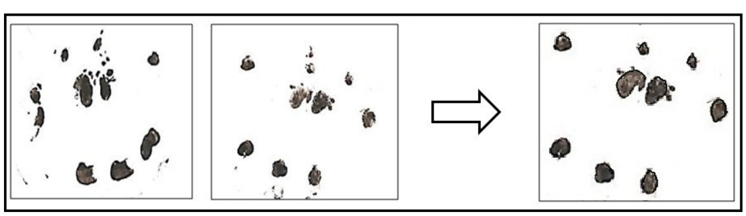


Fig. 3. Isolated hind footprint model example.

The isolated hind footprint has a similar geometric structure to the front footprint. However, by contrast it has three toes in the front that can form a straight line that is parallel with the line formed by the two outer toes. Comparing with the front footprint, it does not only have the three lines we indicated in the analysis of front footprint, but also it has one extra line, which is 'T5' to the middle point of 'A1' and 'A2' (as shown in Figure 4, right).

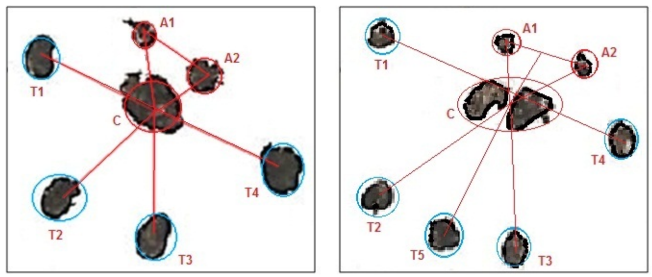


Fig. 4. *Left:* Geometric model for mice front footprint. *Right:* Geometric mode for mice hind footprint.

Since normally every two nearby toes have certain angles in between, a statistic analysis was used to find out the angles between every two nodes of the footprint samples. The corresponding statistic analysis of those angles provide us the following classification rules:

Front or hind footprints classification: for front footprints, the average value for angle $\angle T_2CT_3$ is 46·2° in the range from 43·6° to 48·9°; for hind footprints, the average value for angle $\angle T_2CT_3$ is 56·1° in the range from 53·2° to 59·8°. There is a clear difference between the two ranges.

Left or right footprints classification: if angle $\angle A_1CT_1$ is less than angle $\angle T_4CA_2$, then this is a left footprint; otherwise, this is a right footprint.

4 Algorithm Implementation

In the algorithm implementation, the filtering procedure is given the name 'preprocessing'. Since there are too many interest points detected on the input image. The next challenge would be to filter out insignificant interest points from the image and to define the areas of interest for further analysis. Therefore, we define a set of rules for filtering out the irrelevant interest points:

Rule 1. If an interest point indicates a white blob on a dark background (e.g. a gap between two toes of a footprint), this points should be removed.

Rule 2. If an interest point has a reasonable small radius – less than 6 pixels, then this point should be removed from list.

Rule 3. If an interest point (A) fully contains another interest point (B): if interest point B has a radius greater than or equal to 7/10 of the radius of interest point A, then remove interest point A; otherwise interest point B needs to be removed.

Rule 4. If an interest point (A) partially contains another interest point (B): if the distance between their centres is less than a reasonable length (6 pixels), then these two interest points are recognised as fully containing each other, then apply *Rule 3* to test them; otherwise remove interest point A from list.

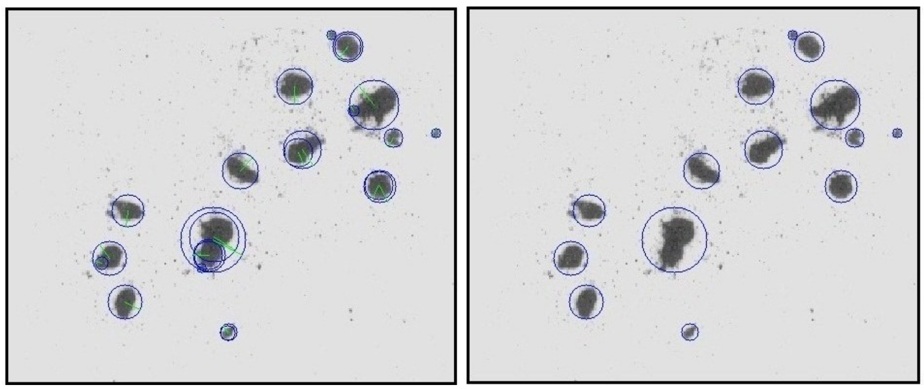


Fig. 5. *Left:* Detected interest points before preprocessing. *Right:* Identified areas of interest after preprocessing.

After the rule-based conditional filtering function is applied, most of the insignificant interest points can be detected and removed. An example of the filtering result is shown in Fig. 5.

The left image shows all the detected interest points on the image, and the insignificant interest points could be the points indicated the same location as other points did, but it has a slightly different radius, or it is partially contained by other interest points with an irrespective small radius.

The right image gives the processing results after we applied the filtering function. Basically all the areas of interest on the image are correctly and completely recognised by our algorithm.

4.1 Identifying Central Pads and Toes

After the areas of interest have been successfully recognised by the algorithm, the next step is to identify which areas of interest are more likely to be a central pad. From analysing the standard models of the front and the hind footprints, it provides us with the following organised truths:

For accurate matching: a central pad normally has the largest area within the six times its radius bounded region, and there should be exactly six (for the front footprint) or seven (for the hind footprint) smaller areas of interest in that particular region.

For loose matching: a central pad normally has a radius larger than the average radius within the six times its radius bounded region, and the number of areas of interest in this particular region should be greater than or equal to four, and less than or equal to ten.

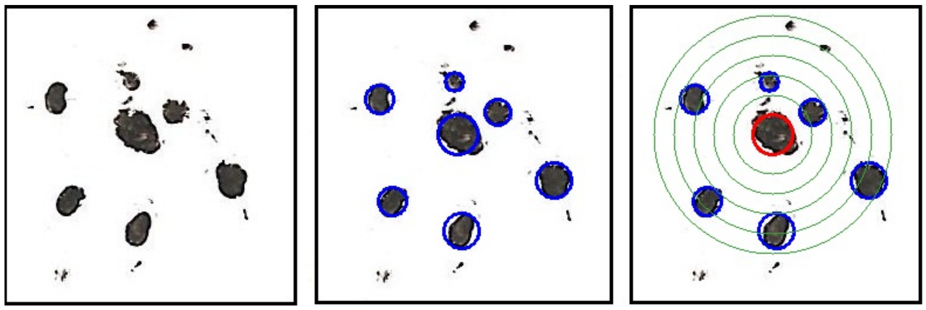


Fig. 6. *Left:* Original image. *Middle:* Preprocessed image. *Right:* Central pad recognition result. Blue circles indicate 'area of interest'; red circle indicate 'recognised central pad'; green circles indicate distance from the centre of the central pad, each gap represents one times the radius of the central pad area.

Figure 6 shows the progress of recognizing a possible central pad on an input image. We use green circles to indicate the distance from the centre of the central pad, which also is the centre of the possible region for a footprint. From the inside to the outside boundary, each gap between every two green circles represents the length of the radius of the central pad area. The outside boundary shows the region of a possible footprint on the image. This region could be valuable when human experts decide to do manual additional analysis of the tracking card image.

4.2 Matching Footprint Models

As the central pad can be recognised, the region of a possible footprint is located with a proper boundary, which is six times the radius of the central pad. The algorithm can then test the interest points within this range whether their distribution matches the pre-defined model. In order to find the best matches footprint, we defined a number of rules for the footprint identification process:

Rule 1. Two accessorial pads should be close to the central pad, generally within the range of three times the radius of the central pad.

Rule 2. Two accessorial pads must have smaller distance to each other than their distance to other areas of interest in this particular region.

Rule 3. Two accessorial pads need to have the smallest radius among the interest points in the relevant region. Considering the ink of the tracking card is not stable, the accessorial pads could possibly leave prints appear bigger than they should be. Thus, it is reasonable to define conditions loosely: the radius of accessorial pads does not need to be the smallest, but only smaller than the average radius.

Rule 4. Two accessorial pads and the central pad can form a triangle at the back of the footprint. Each angle inside the triangle should be only smaller than or equal to 90°, and the sum of the three angles is exactly 180°.

Rule 5. A line segment can be drawn between every two toes. The longest line segment, which is the line between the left and the right outer toes, must cross the area of the central pad.

Rule 6. If the region with a recognised central pad can not completely match all the rules, it should be considered and marked as a 'possible region of a footprint' on the result image for human experts to review.

Rule 7. If the region with a recognised central pad has more than ten areas of interest within its considerable range, which is six times the radius of the central pad, the algorithm should identify this region as an 'unpredictable region', and it will refuse to do any further analysis.

5 Conclusion

The experimental result (as shown in Table 1) indicates that the algorithm has fairly high success rate for sensitive footprint identification and loose-condition matches for images with clear prints and clean background. The accuracy for dim background and foreground images is reasonably lower than the results for images from Group One. In addition, the accuracy for tracking cards with tracks from unexpected species (e.g. insects) is surprising good; the reason might be that the track recognition algorithm has a filtering function that filters out all the tracks with very small regions, which just matches the appearances of insect tracks.

Table 1. The experimental results for the algorithm accuracy evaluation.

	Percentage of Accuracy		
Classification	Group 1 (72 cards)	Group 2 (42 cards)	Group 3 (22 cards)
Sensitive matches	77·8 %	61·9 %	68·2 %
Loose matches (True)	85·7 %	68·3 %	80·7 %
Loose matches (False)	16·3 %	31·7 %	19·3 %
Did not detect print	1·4 %	9·5 %	9·1 %
Unidentified print	3·8 %	4·8 %	4·5 %

Notes: "Sensitive matches" indicates rate of best matched footprints. The correctly and incorrectly detected possible footprints are assigned as "loose matches (true)" and "loses matches (false)". "Did not detect print" records no footprints detected for a card. If the majority of footprints for a card is not detected, then it is recorded as "unidentified".

Comparing with some pervious studies [4, 6–9] in this research field, we come up with two new ideas for this algorithm:

Rule-based identification. A footprint could be identified being either "fully matched" or "partially matched", which depends on the degree of matching the pre-defined rules. Due to the sparse amount of information provided by the detected interest points, rule-based identification process could be a key to the shortage of information. Moreover, rule-based identification could allow developers to add a new rule or modify the existing rules. This provides a great extensibility to this algorithm.

Geometric models. Footprint geometric models could provide precise mathematical relationships among nodes of the standard footprint for any target species. In practical implementations, numbers, equations and formulas are always considered useful information for footprint identification.

The experimental results provide positive feedback on the accuracy of this algorithm; if the image cards have clear prints and clean background, 85·7 % of them can be detected as "loosely matched" the pre-defined rules in this algorithm. The identification of individual species comes next in the project.

References

1. Blackwell, G. L., Potter, M. A., and McLennan, J. A.: Rodent density indices from tracking tunnels, snap-traps and Fenn traps: do they tell the same story? New Zealand Journal of Ecology 26, 43 – 51 (2002)
2. Brown, K., Moller, H., Innes, J., and Alterio, N.: Calibration of tunnel tracking rates to estimate relative abundances of ship rats and mice in a New Zealand forest. New Zealand Journal of Ecology 20, 271–275 (1996)
3. Gaston, K. J., and O'Neill, M. A.: Automated species identification: why not? Philosophical Transactions of the Royal Society B 359, 655 – 667 (2004)
4. Lowdon, I. M. R., Seaber, A. V., and Urbanlak, J. R.: An improved method of recording rat tracks for measurement of the sciatic functional index of de Medinaceli. Journal of Neuroscience Methods 24, 279–281 (1988)
5. MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A.: Estimating site occupancy rates when detection probabilities are less than one. Ecology 83, 2248 – 2255 (2002)
6. Mayo, M., and Watson, A. T.: Automatic species identification of live moths. Knowledge-Based Systems 20, 195 – 202 (2007)
7. Medinaceli, L. D., Freed, W. J., and Wyatt, R. J.: An index of the functional condition of rat sciatic nerve based on measurements made from walking tracks. Experimental Neurology 77, 634 – 643 (1982)
8. Russell, J. C., Hasler, N., Klette, R., and Rosenhahn, B.: Automatic track recognition of footprints for identifying cryptic species. Journal of Ecology 90, 2007 – 2013 (2009)
9. Watts, C. H., Thornburrow, D., Green, C. J., and Agnew, W. R.: Tracking tunnels: a novel method for detecting a threatened New Zealand giant Weta. New Zealand Journal of Ecology 32, 92–97 (2008)