

Marker-Based Tracking of Subsurface Locomotion in Noisy Environments

Miguel M. Serrano, Sarah S. Sharpe, Daniel I. Goldman, Patricio A. Vela
Georgia Institute of Technology
{mserrano6, pvela}@gatech.edu

Abstract

This work presents an automated, marker-based tracking system for monitoring the subsurface locomotion strategies of burrowing animals, in particular small reptiles. High speed x-ray imaging was used to visualize movement of a subject within a controlled granular environment. Given the images, the system returns the trajectories of markers placed on the subject's body and limbs. Unlike existing methods that require unique correspondences between markers in subsequent images, the proposed method is designed to handle false positive marker detections arising from the noise inherent to x-ray imaging. A Bayesian approach to trajectory estimation using linear models with Gaussian noise leads to a combined Kuhn-Munkres algorithm and Kalman filtering formulation. This is the first step towards a robust, marker-based tracking system for articulated subjects undergoing locomotion in noisy environments.

1. Introduction

Recent efforts studying reptiles like the sand-diving and swimming sandfish [2, 9] seek to understand how the interaction between the animal and its environment produce subsurface locomotion. Identifying the interaction mechanisms for modeling purposes requires extensive investigation [10]. Accurate model estimation is contingent on the ability to extract the motion trajectories of the animal's body while fully submerged in the media. The difficulty in studying the subsurface locomotion of these animals arises from the opaque granular media, which renders traditional visible light imaging acquisition techniques useless.

Subsurface imaging is aided by the use of high speed x-ray technology [3, 6]. Through the strategic placement of high density markers on the subject of interest, trajectories of certain body landmarks can be extracted. To be of practical utility however, automated methods for trajectory extraction are necessary. Exam-

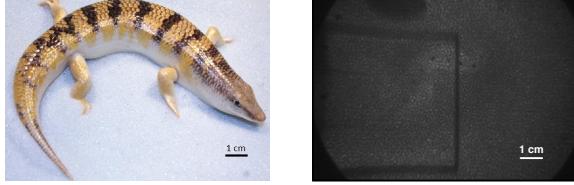
ples of automated marker-based tracking algorithms for x-ray imaging include [1, 4]. The X-Ray Reconstruction of moving Morphology (XROMM) algorithm has been successfully used for motion analysis [1, 5]. It is a template-based approach that requires one-to-one correspondences between detected markers across frames. The technique in [4] uses active appearance models to learn a low-dimensional representation of the subject's shape. Such representations require *a priori* training and a kinematic model that truly has a low-dimensional parametric representation. In the present study, imaging noise associated with the x-ray process produces a one-to-many marker correspondence problem, and the reptile motion has many degrees of freedom, meaning that these approaches are not suitable for the present needs.

This paper presents an automated computer vision system for robustly tracking landmarks placed on an articulated, burrowing animal. The system uses blob detection to identify candidate markers, and assumes a 2D motion model for the markers. A recursive Bayesian estimation strategy for tracking leads to a combined Kuhn-Munkres algorithm and Kalman filtering method for estimating the marker trajectories. The combined approach results in a correspondence cost function that incorporates the filtered states into the costs.

2 Image Acquisition

Animals moved in a container filled with approximately spherical glass bead (diameter, $d \approx 0.3\text{mm}$ or 3mm , and density, $\rho = 2.5 \frac{\text{g}}{\text{cm}^3}$), which was placed between the x-ray source and image intensifier to produce top view (or dorsal) images of animal burial. The media was approximately 10 cm deep. Further information on the methods and techniques can be found in [9].

X-ray recordings of the animals observed entry and movement through the granular medium. Figure 1 shows an extracted image with a sandfish submerged in coarse media. The contrast between the animal's body and the environment is too low to admit tracking of the body. Thus lead markers (1 mm in diameter) are



(a) Mask boundary

(b) Detections

Figure 1. Masking the marker detections for shovel-nosed snake.

placed along the spine of the animal (and also on the legs if appropriate) to enhance the contrast and facilitate tracking. As seen in Figure 1, these markers stand out as dark regions in contrast to the surrounding environment. Under ideal circumstances the markers maintain sufficient contrast for tracking of subsurface locomotion throughout the image acquisition. However, in less than ideal circumstances there are important impediments to tracking. Digital radiography has quantum noise and sensor noise. Further, the coarseness of the media impacts the imaging contrast and texture, while the barrier generates occlusions. In these cases, markers can disappear and some image regions can have marker-like visual characteristics.

3 Marker Tracking

The proposed tracking solution is decomposed into two phases, the detection phase and the association phase. By design the association phase also incorporates filtering of the estimated marker trajectories.

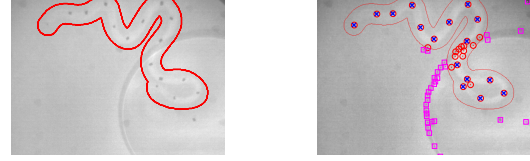
3.1 Marker Detection

The lead markers on the subject have a roughly circular appearance in the x-ray images. These markers are detected using a blob detection algorithm [8] relying on high responses from a circular filter convolved with the x-ray image I . The circular filter is the Laplacian of Gaussian (LoG) filter,

$$r(x, y) = \Delta G_\sigma * I. \quad (1)$$

where Δ is the Laplacian operator and G_σ is the Gaussian kernel with standard deviation σ (also known as the scale of the filter). The highest response r for a marker is produced at the scale best matching the marker radius.

Thus, given an initial frame with manually tagged markers, unique blob detection kernels are designed to detect the markers of the subject, leading to a filter bank of LoG filters. The output of the convolutions are thresholded and the centroid of every disjoint detection region is classified as a candidate marker. As mentioned



(a) Mask boundary

(b) Detections

Figure 2. Masking the marker detections for shovel-nosed snake.

earlier, this approach also leads to false positives in the detection process, and can also lead to missed detections when the contrast between the surrounding media and the marker drops significantly. To reduce the quantity of false positives, a target mask region is utilized. Given the current estimate of the subject’s markers, a mask is generated by considering only those image coordinates that are within a specified distance to any one marker. Figure 2(a) shows the image with the mask region, while Figure 2(b) shows all marker detections. The magenta square detections lie outside of the mask and are rejected, while the red circular detections lying within the region are accepted for further processing. The blue crosses are the true marker locations.

3.2 Correspondence Matching

To successfully generate the frame-to-frame marker associations, a correspondence matching algorithm is needed. Following a Bayesian approach [?], the correspondence can be found by maximizing the probability associated with all possible candidate correspondences given the past detections and the current detection. Assuming a Markov process, the probability optimization is reducible to a recursive estimation procedure. Let the n marker states at time k be denoted by $x_i(k)$ for $i \in \mathcal{S} := \{1, \dots, n\}$ which form the set $\mathcal{X}(k)$, and likewise the m detected candidate marker positions be denoted by $y_j(k)$ for $j \in \mathcal{D} := \{1, \dots, m\}$ which form the set $\mathcal{Y}(k)$. Assuming the markers are independently identically distributed, the probability to maximize is

$$\arg \max_{\alpha(k, \cdot), \mathcal{X}(k)} \prod_{i=1}^n P(y_{\alpha(k, i)}(k) | x_i(k)) P(x_i(k) | x_i(k-1)), \quad (2)$$

where $\alpha(k, \cdot) : \mathcal{S} \rightarrow \mathcal{D}$ is the marker to detected marker correspondence function at time k . The optimization is over the correspondence function and the marker states.

Kalman Filter Model. For simplicity, first consider the case of an individual single marker, labeled by i , with the correct correspondence. Assume that the dynamics of each marker are linear time-invariant with

Gaussian uncertainty, plus that the (correspondence) measurement has Gaussian uncertainty,

$$\begin{aligned} x_i(k) &= Ax_i(k-1) + w_i(k), \\ y_i(k) &= Cx_i(k) + v_i(k), \end{aligned} \quad (3)$$

Both the state equation and the observation equation are affected by white stationary noise defined by $w_i \sim \mathcal{N}(0, Q)$ and $v_i \sim \mathcal{N}(0, R)$, respectively. The Bayesian procedure of equation (2) leads to a Kalman filter [?] for the marker trajectory for these given state equations. The Kalman filter first involves a prediction step,

$$\begin{aligned} x_i^-(k) &= Ax_i(k-1) \\ P_i^-(k) &= AP_i(k-1)A^T + Q \end{aligned} \quad (4)$$

where $x_i^-(k)$ and $P_i^-(k)$ are the predicted state and covariance obtained by propogating the previous estimate through the uncertain dynamics of (3). Then, given current measurement $y_i(k)$, there is a correction step,

$$\begin{aligned} x_i(k) &= x_i^-(k) + K(y_i(k) - Cx_i^-(k)) \\ P_i(k) &= (1 - KC)P_i^-(k) \end{aligned} \quad (5)$$

where the Kalman gain K is defined as:

$$K = P_i^-(k)C^T(CP_i^-(k)C^T + R)^{-1}. \quad (6)$$

Given that the distributions are all Gaussian, the Kalman filter process maximizes a probability, as per (2), that is Gaussian. In this case maximization of the probability is equivalent to minimization of the negative log of the probability, as given by the cost function

$$c_i(k) = \|y_i(k) - Cx_i^-(k)\|_{W(k,i)}^2, \quad (7)$$

where $W(k, i) = (K^T (P_i^-(k))^{-1} K + (1 - CK)^T R^{-1} (1 - CK))^{-1}$, and $\|x\|_M^2 = (x^T M^{-1} x)$.

Filtered Marker Correspondences. Suppose that all marker states $i \in \{1, \dots, n\}$ are to be considered as well as all currently detected markers $j \in \{1, \dots, m\}$. The Bayesian filtering problem of equation (2) takes a more complex form, since n states should be estimated from m possible associations. Optimization now requires finding both the associations and the filtered states. Using (7), a candidate correspondence between marker i and detected marker j has the cost

$$c_{i,j} = \|y_j(k) - Cx_i^-(k)\|_{W(k,i)}^2. \quad (8)$$

The association function α must reflect the set of one-to-one correspondences maximizing (2), or equivalently minimizing the sum of (8) over all $i \in \mathcal{S}$ where

$j = \alpha(i)$, e.g., $\sum_{i \in \mathcal{S}} c_{i, \alpha(i)}$. The Kuhn-Munkres algorithm provides an efficient approach to identifying the optimal solution to the association problem [11]. It solves the dual problem of ensuring optimal assignment of the maximum number of correspondence matches with the minimum effective net cost for the assignments made. The algorithm structures the problem using a cost matrix $C := [c_{i,j}]$, where $c_{i,j}$ is the cost of assigning marker $i \in \mathcal{S}$ to candidate marker $j \in \mathcal{D}$. The Kuhn-Munkres algorithm is an established solution to the assignment problem given a feasible cost matrix.

4 Experiments

The algorithm was run on three different reptiles, the sandfish (Fig. 3), the shovel-nosed snake (Fig. 4), and the ocellated skink (Fig. 5). Nevertheless, the setup presented can be used for the study of any small, articulated burrowing animals. Due to space considerations, one experimental results per animal will be given, with the selection being a more challenging sequence for each animal. Given that the closest methods [1, 4] are not suited to track these sequences, comparison is made with an approach that first generates correspondences then filters the results in order to show the importance of coupling the correspondence and filtering more tightly. Each sequence was manually initialized on an initial frame where every marker was visible, and a second-order, constant velocity model was used for the Kalman filter. Ground truth was obtained by manually annotating 20 evenly spaced frames per sequence.

The tracking results, Figs. 3-5, refer to the proposed method as Filtered Association (FA) and the comparison method as Associate-Then-Filter (ATF). Subfigure (a) shows a sample x-ray image and marker positions for each animal (Figs. 3-5) and subfigure (b) shows the

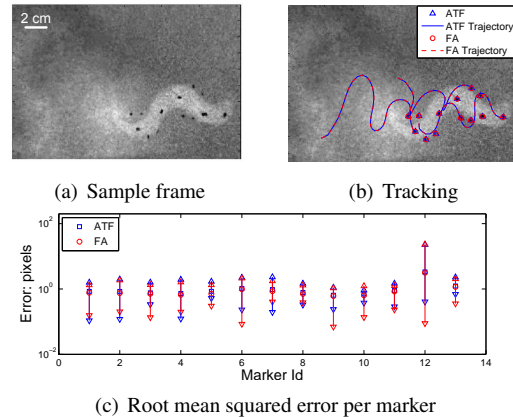


Figure 3. Sandfish tracking.

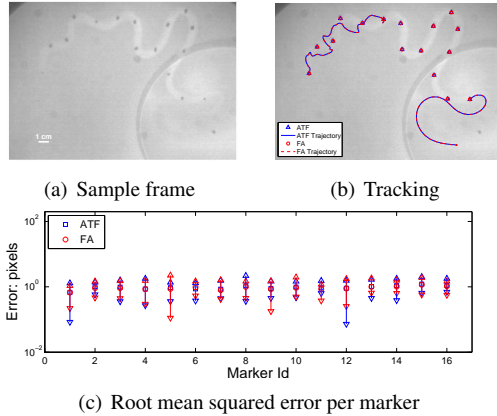


Figure 4. Shovel-nosed snake tracking.

trajectory for select markers (head, tail, and one limb if applicable). Subfigure (c) plots the root mean square error plus the maximum and minimum errors per marker for both FA and ATF. The importance of coupling the overall procedure more tightly can be seen in Figure 5(d), which plots the correspondence correctness for Marker 10 versus time. The top graph of Figure 5(d) shows that the ATF method generates incorrect correspondences that cannot be corrected over time while the (coupled) FA method does not lead to long-term incorrect correspondences. The large error for Marker 12 in Figure 3(c) is due to long-term marker disappearance. Subfigures 3(c)-5(c) show that the proposed approach is accurate to a pixel when the marker is visible.

5 Conclusion

This paper presented an algorithm for automated marker-based tracking of articulated animals in noisy environments. The algorithm admits tracking of burrowing animals imaged using high speed x-ray. A Bayesian-based optimal estimation formulation provides for robust correspondences given false marker detections. The method successfully tracked three articulated reptiles: the sandfish, the ocellated skink, and the shovel-nosed snake.

References

- [1] E. Brainerd, D. Baier, S. Gatesy, T. Hedrick, K. Metzger, S. Gilbert, and J. Crisco. X-ray reconstruction of moving morphology (xromm): precision, accuracy and applications in comparative biomechanics research. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 313(5):262–279, 2010.
- [2] Y. Ding, N. Gravish, C. Li, R. Maladen, N. Mazouchova, S. Sharpe, P. Umbanhowar, and D. Goldman.

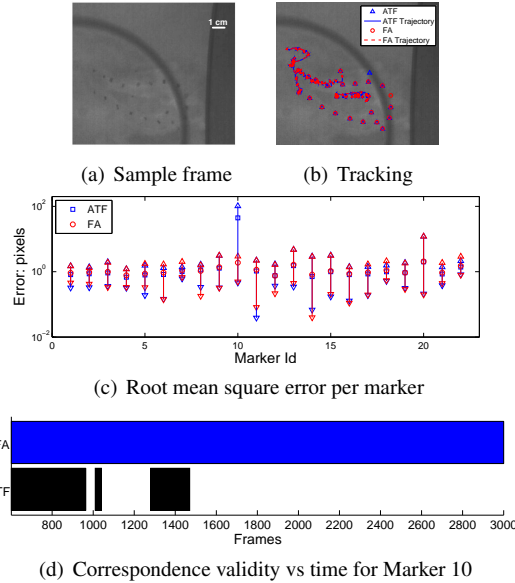


Figure 5. Ocellated skink tracking.

Comparative studies reveal principles of movement on and within granular media.

- [3] S. Gatesy, D. Baier, F. Jenkins, and K. Dial. Scientific rotoscoping: a morphology-based method of 3-d motion analysis and visualization. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 313(5):244–261, 2010.
- [4] D. Haase, J. Nyakatura, and J. Denzler. Multi-view active appearance models for the x-ray based analysis of avian bipedal locomotion. *Pattern Recognition*, pages 11–20, 2011.
- [5] T. Hedrick. Software techniques for two-and three-dimensional kinematic measurements of biological and biomimetic systems. *Bioinspiration & Biomimetics*, 3:034001, 2008.
- [6] T. Hedrick, B. Tobalske, I. Ros, D. Warrick, and A. Biewener. Morphological and kinematic basis of the hummingbird flight stroke: scaling of flight muscle transmission ratio. *Proceedings of the Royal Society B: Biological Sciences*, 2011.
- [7] R. Kalman et al. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.
- [8] T. Lindeberg. Feature detection with automatic scale selection. *International journal of computer vision*, 30(2):79–116, 1998.
- [9] R. Maladen, Y. Ding, C. Li, and D. Goldman. Undulatory swimming in sand: Subsurface locomotion of the sandfish lizard. *Science*, 325(5938):314–318, 2009.
- [10] R. Maladen, Y. Ding, P. Umbanhowar, A. Kamor, and D. Goldman. Mechanical models of sandfish locomotion reveal principles of high performance subsurface sand-swimming. *Journal of The Royal Society Interface*, 8(62):1332–1345, 2011.
- [11] J. Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38, 1957.