A framework for quantifying properties of 3-dimensional movement-based signals

Shaun TD New, Richard A Peters *

ARC Centre of Excellence in Vision Science, Division of Evolution, Ecology & Genetics, Research School of Biology, The Australian National University, Canberra ACT 2601, Australia

Abstract  Understanding how signal properties are optimized for the reliable transmission of information requires accurate description of the signal in time and space. For movement-based signals where movement is restricted to a single plane, measurements from a single viewpoint can be used to consider a range of viewing positions based on simple geometric calculations. However, considerations of signal properties from a range of viewing positions for movements extending into three-dimensions (3D) are more problematic. We present here a new framework that overcomes this limitation, and enables us to quantify the extent to which movement-based signals are view-specific. To illustrate its application, a Jacky lizard tail flick signal was filmed with synchronized cameras and the position of the tail tip digitized for both recordings. Camera alignment enabled the construction of a 3D display action pattern profile. We analyzed the profile directly and used it to create a detailed 3D animation. In the virtual environment, we were able to film the same signal from multiple viewing positions and using a computational motion analysis algorithm (gradient detector model) to measure local image velocity in order to predict view dependent differences in signal properties. This approach will enable consideration of a range of questions concerning movement-based signal design and evolution that were previously out of reach [Current Zoology 56 (3): 327–336, 2010].

Key words  Signal evolution, Sensory Ecology, Movement-based signal, 3D, Display Action Pattern

The design of animal signals has evolved in a complex landscape of selective pressures for efficacy in response to environmental conditions and the signaling context (Endler and Basolo, 1998). Understanding how these signals are optimized, through evolution, is a fundamental tenet of sensory ecology. Progress in this regard is very much linked to the accessibility of technologies that allow us to quantify signals (and noise) in a meaningful way. The sound spectrograph dramatically improved our understanding of the spectral and temporal properties of acoustic signals. Similarly for visual signals, the objective measurement of light is now commonplace in analyses of color patterns. Spectral profiles of the signal and the signaling environment are routinely combined with retinal photoreceptor sensitivities to quantify the structural correlates of conspicuousness from the perspective of receivers (e.g., Vorobyev et al., 1998; LeBas and Marshall, 2000; Macedonia et al., 2002).

Until recently, the inherent difficulty in quantifying motion has restrained our understanding of the factors contributing to movement-based signal designs. Traditional techniques, such as tracking features over time to define display action patterns (DAPs; e.g., Carpenter et al., 1970; Jenssen, 1977; Martins and Lamont, 1998), permit comparisons of movement properties, but only weak inferences can be drawn concerning the relative conspicuousness of signals because comparable measurements of motion noise cannot be made. However, advances in computer-based computation mean that we can now quantify motion signals in a way that informs us of the constraints involved in the perception of relevant, natural motion signals (Zeil and Zanker, 1997; Peters et al. 2002; Elias et al., 2006).

Our interests in movement-based signals have led us to consider how signal properties might change depending on viewing position. Signal directionality has important implications for animal communication strategies (e.g., Yorzinski and Patricelli, 2010) and thus the signal-production structures and/or behaviors that affect directionality may be subject to selection. Some signals, or signal components, have likely evolved to ensure signal properties are robust to changing receiver position, while others may benefit from being more di-
rectangular because of, for example, the possibility of eavesdropping (Witkin, 1977; Larsen and Dabelsteen, 1990; Dantzker et al., 1999; Brumm and Todt, 2003; Dabelsteen, 2005). Indeed, the sensory drive model (Endler and Basolo, 1998) predicts that signals should minimize the probability of detection by unintended receivers; this remains relatively unexplored in the context of movement-based signals.

Quantifying how signal properties vary as a function of viewing angle and distance are relatively simple to predict for movements confined to a single plane of motion. Amplitudes of simple up-down movements characteristic of many lizard displays, for example, can be determined for multiple receiver positions based on simple geometric computations. It is far more difficult to quantify viewing position differences when signals have a complex 3-dimensional (3D) structure. Tail flicking by the Australian Jacky lizard *Amphibolurus muricatus* is one such signal. Tail movements form the introductory part of displays performed by rival males during territorial defense, and serve an alerting function (Peters and Ord, 2003; Peters et al., 2007). Tail positions are not confined to a single plane of movement, but rather move randomly in 3D space around the lizard’s body (see Peters and Allen, 2009). Complicating quantitative analysis further is that the tail is not rigid during displays but constantly changes shape. The signal thus resembles the flicking motion of a whip (see Appendix).

A major constraint in applying traditional DAP analysis and computational algorithms for 3D signals is the use of footage from a single camera, which flattens motion into 2-dimensions (2D). We circumvented this problem here through the use of a stereo camera setup to generate 3D DAP profiles for the complex flicking motion. Although the use of stereo cameras in behavioral biology is increasing (e.g., Fry et al., 2009; Boeddeker et al., 2010), we are not aware of its application to quantifying animal movement-based signals.

Calculating view-specific DAP profiles was achieved using a series of straightforward geometric computations. However, it was clear that the decision of which points along the tail we tracked influenced the resultant profiles. Tracking multiple points along the tail would have improved our sampling of tail movements, but it was never clear how many points would be sufficient given the constant and random shape changes of the tail structure. Indeed the importance of adequate sampling of the tail structure was one of the reasons we have utilized alternative methods for quantifying movement-based signal design (e.g., Peters and Evans, 2003a). The use of computational motion analysis algorithms, for example, provide a richer description of movement-based signal properties by retaining the whole structure and facilitates comparisons with other types of image motion, such as wind-blown plants known to be crucial for determining signal efficacy (Peters et al., 2007). Nevertheless, this approach is still limited to the 2D image sequences obtained from video footage, making it extremely difficult to extrapolate measurements from one view to other viewing positions. We overcome this in the present paper by using the DAP profiles to define the movement of a 3D lizard tail model in animation software, thereby replicating the signal in a virtual world. The animated signal can then be ‘filmed’ from any viewing position, exported as separate image sequences and analyzed using any computational algorithm. After presenting a worked example, we discuss how our approach enables us to consider factors affecting movement-based signals more rigorously or that were previously out of reach.

1 Materials and Methods

1.1 Filming and digitizing the signal

The focal species for our work are Jacky lizards, which are a medium sized (snout-vent length: 100 mm; tail length: 200 mm) agamid lizard native to South Eastern Australia. Filming was undertaken in March 2007 at the Edith & Joy London Foundation, The Australian National University’s coastal campus, Kioloa, NSW, Australia. Signaling behavior was induced in a large outdoor enclosure by presenting a male conspecific, held in a plastic aquarium, to a resident male (see Peters et al. 2007 for details). Signals were filmed simultaneously using three digital video camcorders (Sony DCR-HC21E PAL miniDV handycam) from three locations as indicated in Fig 1. The cameras were calibrated to correct for lens distortion and their alignment in space was later reconstructed using a checkerboard standard and open-source software developed by Bouguet (2005). Footage from each camera was digitized and temporally synchronized using modified open-source software under Linux (dvgrab) and exported as separate AVI files. Movies were then read into Matlab (MathWorks Inc.) where custom-written software (Jan Hemmi, The Australian National University) was used to track the tip of the tail over time. Following the method of Bouguet (2005), stereo triangulation of the data from any two cameras allowed us to compute tail positions in 3D space. As the signaling lizard was free to move within its enclosure, we selected the two
cameras that provided the best view of the signal.

Fig. 1  Schematic illustration of the filming set up for capturing signals
Filming of a checkerboard pattern enabled camera position and alignment in 3D space (see text for details).

1.2  Quantifying the signal
We selected one representative sequence and analyzed tail flick structure in two ways. First, we considered the visual angle subtended by the flicking tail as measured at its tip. Secondly, we used tail position data to reconstruct the signal in animation software and exported image sequences for subsequent analysis. Here we used a gradient detector model to compute local image velocity before averaging across space to generate speed-time profiles.

Our goal was to see how changing viewing position affected the perceived signal, which required us to quantify signal properties from multiple locations. To determine viewing directions, we used the centre of the observed data as a reference point and systematically varied the azimuth, elevation and distance around this point (Fig. 2a). Setting the location of the ‘intruder’ to an azimuth of 0°, azimuth values were spaced at 9° increments around the horizontal plane of the reference point (for a total of 40 azimuth directions). Elevation was the angle subtended between the reference point, the horizontal projection of the reference point and a second point in space (Fig. 2a). Fifteen points above and below the height of the signaler were used (starting at ±4° and continuing at increments of ±6°). The factorial combination of azimuth and elevation resulted in 1200 viewing directions for a given distance (Fig. 2b). We varied viewing distances to simulate actual distances between 0.5 and 50 m. Each viewing position was converted to XYZ coordinates based on the combination of azimuth, elevation and distance relative to a reference point (Fig. 2a inset). Henceforth, we distinguish between viewing direction (azimuth and elevation) and viewing position (azimuth, elevation and distance).

(a)

Fig. 2  (a) Multiple viewing positions were simulated by varying the azimuth, elevation and distance from a reference point. (b) A total of 1200 unique viewing directions were used for a given viewing distance, which define a sphere around the reference point
Inset: Greater viewing distances were computed using the same set of azimuth and elevation angles.

1.2.1  Visual angle
We present data for a tail flick sequence not used in the final analysis in Fig. 3 to illustrate our procedure for calculating visual angles. The sequence in this illustration depicts the 3D DAP profile measured at the position of the tail tip during a flicking sequence lasting almost 10 s (246 frames, PAL at 25 fps). Rather than determining the angular extent of motion over the entire sequence (as in Peters and Allen 2009), we divided the sequence into several bouts of movement interspersed with periods of no movement (see Peters et al. 2007 for further discussion on this aspect of signal structure). For each viewing position, we computed the maximum vis-
visual angle generated in each bout, and selected the largest value across all bouts. In Fig. 3, a single bout of movement is denoted by red circles, with the largest visual angle for a viewer located at \( O \) given by \( \angle AOG \). This process was repeated for all 1200 viewing directions at 0.5, 1 and 2 m and then every 2 m thereafter up to a distance of 50 m. All analyses were undertaken using custom-made programs written in Matlab (Mathworks Inc.).

**Fig. 3** Lizard tail flicking occurs in bouts of intermittent motion
In this example, data points are plotted for a complete sequence and those from a single bout are colored red. For a given viewing position, \( O \), we calculated the maximum angle subtended during the bout to be between the data points \( A \) and \( G \) (\( \angle AOG \)).

### 1.2.2 Local image velocity
The XYZ tail position coordinates were used to accurately recreate the movement of the flicking tail using 3D animation software. To generate the 3D animated replica of the original signal we used LightWave 3D v9.2 (NewTek Inc.) animation software and techniques described in detail elsewhere (Peters and Evans, 2003b, 2007). Briefly, we created a wireframe model of a lizard tail based on morphological measurements and overlaid a texture that simulated the patterning of Jacky lizard skin. The model also comprised bones as invisible elements that controlled the configuration of the tail model. Bones are connected in series and can have variable range of movement and stiffness. The last bone in the series was linked to an invisible target that when moved caused the tail model to move with it. We imported the XYZ coordinate data generated during the digitization process to determine the position of the invisible target. LightWave’s powerful inverse kinematics algorithms determined the shape of the tail in each frame. Importantly, the invisible target position in each frame is also the position of the tail tip, while the precise configuration of the rest of the tail is modeled by the animation software.

The tail flick signal was ‘recorded’ within the animation environment from multiple viewing positions and image motion quantified for each viewpoint. We selected a subset of viewing directions from the full set described above, limiting analysis to 10 azimuths (starting at 9° and then every 37° until 341°). We also restricted our analysis to viewing distances of 1, 5, 10, 20 and 50 m. The recording process generated 150 versions of the same signal against a homogenous background, which we exported as JPEG image sequences (Fig. 4).

Local image velocity was quantified using a custom written Matlab program (Peters et al., 2002). Briefly, a gradient detector model calculated the velocity field in image sequences based on temporal and spatial derivatives of filtered versions of image intensity. We focused only on the angular speed of movement (velocity magnitudes; see middle panel of Fig. 4), extracting the maximum value across space for each frame. These values were then converted from pixels per second to degrees per second based on the filming distance to define a maximum angular speed by time profile (henceforth, speed-time profile) for each sequence (right panel of Fig. 4). We compared speed-time profiles using cross-correlation, selecting the correlation value corresponding to an offset of zero (i.e., sequences were temporally aligned).

**Fig. 4** Animations were exported as image sequences and the image motion computed (a). The maximum angular speed in each frame was extracted (b) and converted to angular speed based on the geometry of viewing position, thereby defining maximum angular speed profiles (c)
2 Results

The selected tail flicking sequence was 24 s in duration (698 frames) and featured eight main flicking bouts. The position of the tail tip was tracked over time in 2D space separately for recordings from cameras 2 and 3 (see Fig. 1), and subsequently transformed to 3D space using stereo triangulation as described above. These data are shown in Fig. 5a, with the simulated ‘intruder’ located at the position of the X-axis label. To illustrate how viewing position changes the apparent shape of the distribution we calculated the 3D convex hull (Fig 5b), and rotated the viewing direction (Fig. 5c,d). Changing the azimuth (at the same elevation) in a counter-clockwise direction we initially see a narrowing of the apparent spread of tail tip positions (Fig. 5c); however, changing the viewing direction a little more revealed a wide distribution of points (Fig. 5d).

2.1 Visual angle

For each bout of continuous movement, we computed the maximum visual angle from the 32,400 viewing positions (1200 viewing directions × 27 viewing distances) and selected the maximum value from any given bout. Visual angles generated by the flicking tail vary with receiver position to some degree, as illustrated in Fig 6 for a viewing distance of 1 m. For this signal, the best viewing positions reflect what could be predicted from the convex hull (see Fig. 5d). Interestingly the location of the intended receiver did not coincide with these optimal positions (Fig. 6 bottom). To illustrate the effect of changing viewing distance we present surface plots in Fig. 7 relating azimuth position, viewing distance and visual angle for three viewing elevations: above (Fig. 7a), at the same height (Fig. 7b) and beneath the signaling lizard (Fig. 7c). Differences as a function of azimuth direction are clear at close distances, particularly at the height of the signaler where visual angles differed by up to 20°. The profiles flatten out quickly beyond 1 m, however, reducing differences to less than 1° by 8 m (see inset in Fig. 7b).

2.2 Local image velocity

Speed-time profiles were calculated for each of the 150 animation sequences. Our motion analysis algorithm was unable to detect movement for most sequences recorded at a viewing distance of 50 m. We concentrate, therefore, on viewing distances of 1, 5, 10 and 20 m. At close viewing distances, speed-time profiles were quite similar, as illustrated in Fig. 8 for viewing positions at the same elevation as the signaler; the same pattern was observed for other elevations (not

---

**Fig. 5** (a) Position of the tail tip for the entire sequence. The apparent distribution of points depends heavily on the position of the viewer, as illustrated by the changing shape of the 3D convex hull that encloses the tail tip data (b-d)
Cross correlating profiles between azimuths and within elevations showed that changing viewing direction did not alter signal structure (Fig. 9). There were also only minor differences between profiles at 1 m when comparing across all azimuth and elevation directions (Fig. 9). Azimuth direction profiles remained relatively similar at 5 and 10 m, but were dissimilar at 20 m (Fig. 9).

We next considered the effect of viewing distance on local image velocity by comparing speed-time profiles for a given azimuth and elevation across distances. Specifically, we compared profiles at a distance of 1 m with those at the same azimuth and elevation at 5, 10 and 20 m. The correlation coefficients for these comparisons are presented in Fig. 10. Angular speed profiles as perceived from 1 and 5 m were similar, irrespective of elevation. Angular speeds seen from 10 m show marginally more differences to 1 m profiles (Fig. 10), whereas we found distinct differences between 1 and 20 m. It is important to note, however, that these correlations reflect temporal structure and do not consider angular speed magnitudes. Two profiles can generate strong cross-correlations, therefore, if the sequence of peaks and troughs are similar regardless of absolute speed differences.

### Discussion

We have presented a new approach for the study of complex movement-based visual signals, demonstrating its application using the Jacky lizard introductory tail-flick. Extending traditional DAP analysis and measurement of local image velocity to retain 3D spatial configurations we were able to quantify differences in signal structure for a range of viewing positions around the signaler. DAP analysis revealed that signal amplitudes vary according to receiver direction (Fig. 6) and degrade with distance (Fig. 7) in line with what would be expected from geometric computations. Interestingly, however, speed-time profiles suggest that at close distances signal structure is consistent for all viewing positions (Fig. 9). Moreover, rather than degrading predictably with viewing distance, changes in averaged local velocity as distance increased were not uniform for all
azimuths (Fig. 10), suggesting that changes in signal properties as a function of viewing distance are not the same from one azimuth direction to another. It will be important to consider the extent to which this prediction is a consequence of the limitations of our computational approach and we recognise that alternative models for image motion analysis may yield contrasting results as parameters in our model, such as spatial filtering, are not based on lizard spatial resolution.

Notwithstanding the need for considering alternative models for motion analysis, we wish to highlight that the DAP analysis and our implementation of gradient detectors produced contrasting predictions about the way signal properties might change with viewing position: one was consistent with geometry and the other posits a less straightforward change. We argue that predictions about changes in signal properties as a function of viewing position based purely on geometry will not do full justice to the detection tasks facing receivers. To illustrate, we consider our estimates for visual angle as a function of distance (Fig. 7) and 0.2°, which is to our knowledge the sole prediction for minimum amplitudes required for signal detection in lizards (Fleishman, 1986). Our results suggest that the visual angle subtended by the tail tip would remain above this threshold as far away as 50 m (Fig. 7b inset). However, the angular size of the tail tip would be approximately 0.002° at this distance and difficult to resolve. Extrapolating from measurements at the tail tip may thus be unreliable. Feature tracking is only one possible means of motion
Fig. 10  Cross correlation coefficients obtained by comparing angular speed profiles as seen from 10 different azimuth positions at 1 m with profiles at the same azimuth position at 5, 10 and 20 m for three elevations: (a) 4°, (b) 32° and (c) 64°.

detection. Salient motion events can also be identified by luminance changes, as modeled here, or second-order computations based on differences in contrast or texture. Creating an animation and modeling the entire signaling structure (tail) does not limit us to an arbitrarily chosen set of points.

Despite the preliminary nature of the results presented, it is intriguing to consider tail flicking as omnidirectional up to 10 m. That the temporal pattern of the signal remains relatively conserved might explain why the lizards do not modify tail flicking for receivers positioned anywhere in this range (Peters and Allen, 2009). Of equal intrigue were the considerable differences between 1 and 20 m for some azimuths (Fig. 10). In the densely vegetated coastal heath that these lizards occupy (see Peters et al., 2008), the only place where a clear line of sight across 20 m is possible would be from an elevated position, where we would expect to find the primary natural predators of adult Jacky lizards, birds (Allen et al., 2009). Our approach makes it possible to consider for the first time whether a compromise between signaling to conspecifics and avoiding detection by predators has shaped the design of movement-based animal signals as has been demonstrably shown in a variety of signals (color: Hemmi et al., 2006; acoustic: Klump et al., 1986; electrical: Stoddard, 1999).

Future Directions

We regard the approach presented herein to offer a powerful framework for considering the details of movement-based signal designs. Our approach allows precise definition of signal properties and unique flexibility for examining perceptual constraints that have been inaccessible to experimental analysis. In this paper, we have used maximum angular speed as a metric for comparing signal structure, although we do not suggest that this is the cue receivers attend to. The approach provides a detailed representation of the available sensory information, and by considering several different motion parameters it might indeed be possible to identify the attended cue(s). It would also be straightforward to replace the gradient detector model used to analyze image motion herein with an alternative model for motion feature extraction, such as the correlation detector model (e.g., Zeil and Zanker, 1997; Pallus et al., 2010), or other tools for characterizing visual saliency (Gao et al., 2008; Peters in press).

Working within the animation environment also allows us to perform operations that would otherwise not be possible. For example, we might wish to consider the segmentation of thin tails from a distance. It would be straightforward to create tails that do not taper in the natural way, but remain more cylindrical in shape. By comparing the image motion generated by the same sequence of tail positions for normal and cylindrical tail morphologies we are likely to gain insight into how morphology affects the segmentation task. Other aspects of the tail’s morphology can also be manipulated such as tail length and texture patterns. By way of illustration, we considered whether removing the distal two-thirds of the tail, which is responsible for the whip-like motion we associate with tail-flicking, would change angular speed profiles. We generated and analyzed new versions of all animations featuring a partial tail (Fig. 11), which was achieved by overlaying a new texture pattern where part of the tail was the same color as the background (black). Importantly, the movement of the visible portion of the tail was identical in these new clips as the corresponding part of the original full tail versions. Comparing azimuth directions across viewing distances, we found that angular speed profiles for the 1–5 m comparisons showed more variation for the partial tail relative to the full tail, however, not to the extent we
would expect if speed profiles were heavily influenced by the distal part of the tail. Further consideration of the differences between these otherwise identical signals will be the subject of follow-up work. Our strategy also enables us to model the signaling environment with precision. Environmental conditions at the time of signaling affect signal efficacy (Peters et al., 2008; Peters in press), and Jacky lizards adjust the structure of their signal to compensate (Peters et al., 2007). By creating and manipulating virtual microhabitats that closely replicate environmental noise present under natural conditions we can investigate the motion segmentation task in much greater detail.

![Complete tail and Partial tail](image)

**Fig. 11** Representative frames from the original animations (‘complete tail’) and the corresponding frames from partial tail animations (see Discussion)

**Acknowledgements** We thank Martin Stevens for the invitation to submit an article for this special issue, and two anonymous reviewers who provided valuable comments on an earlier version. Thanks also to Robin and Steven Teding van Berkout for onsite support during filming, Jan Hemmi and Jochen Zeil for continuing discussions and valuable comments on early drafts of this manuscript. Housing and experimental procedures were in accordance with guidelines set by The Australian National University’s Animal Experimentation Ethics Committee and the New South Wales National Parks and Wildlife Service. This work was produced with the assistance of the Australian Research Council under the ARC Centres of Excellence Program, and was hosted by the Edith and Joy London Foundation, Kioloa, NSW, Australia. Research was conducted in partial fulfilment of the requirement for a doctoral thesis for STDN at The Australian National University.


**References**


Bougouet JY, 2005. Camera calibration toolbox for MATLAB. MRL-Intel Corp. (See [http://www.vision.caltech.edu/bouguetj/calib_doc](http://www.vision.caltech.edu/bouguetj/calib_doc)).


Peters RA, in press. Movement-based signalling and the physical world: modelling the changing perceptual task for receivers. In: